ETSI TS 101 569-1 V1.1.1 (2013-10)

Integrated Broadband Cable Telecommunication Networks (CABLE); Cable Network Transition to IPv6 Part 1: IPv6 Transition Requirements

Reference
DTS/CABLE-00003

Keywords
CABLE, IPCable, IPv6

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://ipr.etsi.org).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Integrated broadband cable telecommunication networks (CABLE).

The present document is part 1 of a multi-part deliverable covering the Cable Network Transition to IPv6, as identified below:

Part 1: "IPv6 Transition Requirements";

Part 2: "Cable IPv6 Transition Test Plan".

DOCSIS® is a Trade Mark registered and owned by CableLabs.

Introduction

The present document accommodates an urgent need in the industry to define requirements that enable seamless transition of Cable Networks to IPv6. Considering the depletion of IPv4 addresses, transition to IPv6 is required in order to enable continued growth of the customer base connected to Cable Networks and ensure service continuity for existing and new customers. High-quality connectivity to all kinds of IP-based services and networks is essential in today's business and private life.

A plethora of transition technologies have been proposed in IETF, other standardization organizations and by manufacturers of IP technology to allow coexistence of IPv4 and IPv6 hosts, access and core networks as well as services. Each of these technology options is specified, implemented and deployed in various forms and stages. The present document takes into account the conclusions drawn in TR 101 569 [i.1]. The report assessed the current status of Cable Equipment standards that are implemented and deployed in the components that comprise integrated broadband cable and television networks with regard to their readiness for IPv6 and analysed the transition technologies in order to evaluate the suitability and applicability in a Cable Network environment.

1 Scope

The present document defines end-to-end system requirements for broadband Cable Networks as well as equipment specifications that enable the transition from IPv4 to IPv6. Since the time-to-market is a factor considering the depletion of IPv4 addresses, the present document accommodates an urgent need in the industry and is based on the technical analysis provided in the companion Technical Report TR 101 569 [i.1].

Abstract

Equipment vendors across the global industry are currently including IPv6 support as an important requirement in the manufacture of their network equipment. However one of the key issue for the industry and in particular network operators is the requirement to support the IPv4 legacy network for the foreseeable next 8 years and significantly longer period for certain networks depending on the scale and investment of their legacy network. With IPv4 address space depletion worldwide, certain network providers particularly in non European markets have recently been managing the depletion pool of IPv4 addresses such as in the United States by pooling IPv4 addresses to increase depletion timelines. Irrespective of this the IPv4 address space continue to be rapidly exhausted requiring the industry to make decisions on the evolution of their IPv4 network investments. Consequently industries are being driven into transition technologies to allow the home and business networks to continue to function on IPv4. Industries are being faced with making key decisions to ensure the continuity of their services given costs and investments in legacy networks with two key decisions, one being a 'chicken and egg' scenario with the Internet and its services and the second being the home and business networks. Major services on the Internet either have switched or are switching over to IPv6 with a dual-stack topology but despite this important transition the full removal of IPv4 is not expected to happen for several years due to several factors but mainly due to legacy home network equipment such as IPTVs, security devices, legacy operating systems and other such equipment that have for 90 % of the in-home environment an approximate life expectancy of up to 8-10 years. Considering the recent development of devices for the home giving IPv6 capability the industry is now at a tipping point where new devices can reach the IPv6 Internet. So legacy is the industries main long term concern, for the short term the concern is functionality and business continuity. Once an MSO has depleted available IPv4 address spaces they have a choice of about 9 main transitional technology methods. The present document specifies in detail the requirements for the main transition technologies, NAT64, 464XLAT, NAT44, DS-Lite and MAP-E.

European cable network providers have in the main opted for CGN solution mainly due to development of DS-Lite comparative to other transition technologies. DS-Lite is currently being seen as the preferred transition technology specification to deploy within European Cable ISP's, whilst certain other regional markets in Asia and South America are opting for DS-Lite or NAT44. Preference to NAT44 is seen by major European Cable ISPs to not provide a viable solution as it stems the tide of IPv6 development and requirement for deployment within provider's networks. An ISP has a key role to help introduce IPv6 by enabling IPv6 within their CORE and Access networks. Technologies such as MAP-E has lower consideration as a transition technology by industry at present although it is technically viable but currently lacking development support by majority equipment suppliers and therefore is yet to be completed to allow full functionality that would lead to non-service deprecating deployment. Current industry focus has been on developing certain transition technologies with the three key leaders within the product ranges and functional possibilities for deployment being; DS-Lite as the primary transition technology option that gives a roadmap for network providers to migrate from a legacy IPv4 network, to transition based with DS-Lite to IPv6. The next potential option is NAT44 with 6rd trailing behind. The latter would require substantial further development to manage numerous failures in transport that would not provide the customer with a seamless transition.

MAP-T, Stateless NAT44, Stateless DS-Lite/Lightweight DS-Lite, Teredo and 4over6 are being developed, or being developed further, in some form or another, although MAP-T may not move much further, but time will tell. These technologies all have something to offer as possibilities for the next stages in transition development, but as an industry within Europe it would be a far improved responsibility as a group to choose a single viable technology or two, dependent on specific physical and logical requirements, and move forward as a single entity developing and improving on those choices.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are necessary for the application of the present document.

[1] ETSI TS 103 161 (all parts): "Access, Terminals, Transmission and Multiplexing (ATTM); Integrated Broadband Cable and Television Networks; IPCablecom 1.5".

[2] ETSI EN 302 878 (all parts): "Access, Terminals, Transmission and Multiplexing (ATTM); Third Generation Transmission Systems for Interactive Cable Television Services - IP Cable Modems".

[3] Recommendation ITU-T G.8013/Y.1731: "OAM functions and mechanisms for Ethernet based networks", July 2011.

[4] IEEE 802.1ag-2007: "IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks - Amendment 5: Connectivity Fault Management".

[5] IEEE 802.3ah-2004: "IEEE Standard for Information Technology - Local and Metropolitan Area Networks - Part 3: CSMA/CD Access Method and Physical Layer Specifications - Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks".

[6] IETF I-D draft-boucadair-pcp-extensions-03: "Some Extensions to Port Control Protocol (PCP)".

NOTE: Available at http://tools.ietf.org/id/draft-boucadair-pcp-extensions-03.txt.

[7] IETF I-D draft-boucadair-pcp-failure-06: "Analysis of Port Control Protocol (PCP) Failure Scenarios".

NOTE: Available at http://tools.ietf.org/id/draft-boucadair-pcp-failure-06.txt.

[8] IETF I-D draft-boucadair-pcp-rtp-rtcp-05: "Reserving N and N+1 Ports with PCP".

NOTE: Available at http://tools.ietf.org/id/draft-boucadair-pcp-rtp-rtcp-05.txt.

[9] IETF I-D draft-donley-behave-deterministic-cgn-05: "Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments".

NOTE: Available at http://tools.ietf.org/id/draft-donley-behave-deterministic-cgn-05.txt.

[10] IETF I-D draft-donley-softwire-dslite-flowlabel-02: "Using the Flow Label with Dual-Stack Lite".

NOTE: Available at http://tools.ietf.org/id/draft-donley-softwire-dslite-flowlabel-02.txt.

[11] IETF I-D draft-dupont-pcp-dslite-05: "The Port Control Protocol in Dual-Stack Lite Environments".

NOTE: Available at http://tools.ietf.org/id/draft-dupont-pcp-dslite-05.txt.

[12] IETF I-D draft-eastlake-trill-lldp-01: "Transparent Interconnection of Lots of Links (TRILL) Support of the Link Layer Discover Protocol (LLDP)".

NOTE: Available at http://tools.ietf.org/id/draft-eastlake-trill-lldp-01.txt.

[13] IETF I-D draft-grant-tacacs-02: "The TACACS+ Protocol Version 1.78".

NOTE: Available at http://tools.ietf.org/id/draft-grant-tacacs-02.txt.

[14] IETF I-D draft-ietf-softwire-dslite-mib-02: "DS-Lite Management Information Base (MIB)".

NOTE: Available at http://tools.ietf.org/id/draft-ietf-softwire-dslite-mib-02.txt.

[15] IETF I-D draft-ietf-softwire-map-06: "Mapping of Address and Port with Encapsulation (MAP)".

NOTE: Available at http://tools.ietf.org/id/draft-ietf-softwire-map-06.txt.

[16] IETF I-D draft-ietf-softwire-map-dhcp-03: "DHCPv6 Options for Mapping of Address and Port".

NOTE: Available at http://tools.ietf.org/id/draft-ietf-softwire-map-dhcp-03.txt.

[17] IETF I-D draft-murakami-softwire-4to6-translation-00: "4via6 Stateless Translation".

NOTE: Available at http://tools.ietf.org/id/draft-murakami-softwire-4v6-translation-00.txt.

[18] IETF RFC 854: "Telnet Protocol Specification", May 1983.

NOTE: Available at http://tools.ietf.org/rfc/rfc854.txt.

[19] IETF RFC 919: "Broadcasting Internet Datagrams", October 1984.

NOTE: Available at http://tools.ietf.org/rfc/rfc919.txt.

[20] IETF RFC 1112: "Host Extensions for IP Multicasting", August 1989.

NOTE: Available at http://tools.ietf.org/rfc/rfc1112.txt.

[21] IETF RFC 1142: "OSI IS-IS Intra-domain Routing Protocol", February 1990.

NOTE: Available at http://www.ietf.org/rfc/rfc1142.txt.

[22] IETF RFC 1191: "Path MTU Discovery", November 1990.

NOTE: Available at http://tools.ietf.org/rfc/rfc1191.txt.

[23] IETF RFC 1195: "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", December 1990.

NOTE: Available at http://tools.ietf.org/rfc/rfc1195.txt.

[24] IETF RFC 1350: "The TFTP Protocol (Revision 2)", July 1992.

NOTE: Available at http://www.ietf.org/rfc/rfc1350.txt.

[25] IETF RFC 1397: "Default Route Advertisement in BGP2 and BGP3 Versions of the Border Gateway Protocol", January 1993.

NOTE: Available at http://tools.ietf.org/rfc/rfc1397.txt.

[26] IETF RFC 1441: "Introduction to version 2 of the Internet-standard Network Management Framework", April 1993.

NOTE: Available at http://tools.ietf.org/rfc/rfc1441.txt.

[27] IETF RFC 1631: "The IP Network Address Translator (NAT)", May 1994.

NOTE: Available at http://tools.ietf.org/rfc/rfc1631.txt.

[28] IETF RFC 1772: "Application of the Border Gateway Protocol in the Internet", March 1995.

NOTE: Available at http://tools.ietf.org/rfc/rfc1772.txt.

[29] IETF RFC 1912: "Common DNS Operational and Configuration Errors", February 1996.

NOTE: Available at http://www.ietf.org/rfc/rfc1912.txt.

[30] IETF RFC 1918: "Address Allocation for Private Internets", February 1996.

NOTE: Available at http://tools.ietf.org/html/rfc1918.txt.

[31] IETF RFC 1997: "BGP Communities Attribute", August 1996.

NOTE: Available at http://tools.ietf.org/rfc/rfc1997.txt.

[32] IETF RFC 1981: "Path MTU Discovery for IP version 6", August 1996.

NOTE: Available at http://tools.ietf.org/rfc/rfc1981.txt.

[33] IETF RFC 2132: "DHCP Options and BOOTP Vendor Extensions", March 1997.

NOTE: Available at http://tools.ietf.org/rfc/rfc2132.txt.

[34] IETF RFC 2328: "OSPF Version 2", April 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2328.txt.

[35] IETF RFC 2370: "The OSPF Opaque LSA Option", July 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2370.txt.

[36] IETF RFC 2385: "Protection of BGP Sessions via the TCP MD5 Signature Option", August 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2385.txt.

[37] IETF RFC 2439: "BGP Route Flap Damping", November 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2439.txt.

[38] IETF RFC 2452: "IP Version 6 Management Information Base for the Transmission Control Protocol", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2452.txt.

[39] IETF RFC 2454: "IP Version 6 Management Information Base for the User Datagram Protocol", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2454.txt.

[40] IETF RFC 2460: "Internet Protocol, Version 6 (IPv6) Specification", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2460.txt.

[41] IETF RFC 2461: "Neighbor Discovery for IP Version 6 (IPv6)", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2461.txt.

[42] IETF RFC 2462: "IPv6 Stateless Address Autoconfiguration", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2462.txt.

[43] IETF RFC 2463: "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2463.txt.

[44] IETF RFC 2464: "Transmission of IPv6 Packets over Ethernet Networks", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2464.txt.

[45] IETF RFC 2465: "Management Information Base for IP Version 6: Textual Conventions and General Group", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2465.txt.

[46] IETF RFC 2466: "Management Information Base for IP Version 6: ICMPv6 Group", December 1998.

NOTE: Available at http://tools.ietf.org/rfc/rfc2466.txt.

[47] IETF RFC 2473: "Generic Packet Tunneling in IPv6 Specification".

NOTE: Available at http://www.ietf.org/rfc/rfc2473.txt.

[48] IETF RFC 2827: "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", May 2000.

NOTE: Available at http://tools.ietf.org/rfc/rfc2827.txt.

[49] IETF RFC 2865: "Remote Authentication Dial In User Service (RADIUS)", June 2000.

NOTE: Available at http://tools.ietf.org/rfc/rfc2865.txt.

[50] IETF RFC 2866: "RADIUS Accounting", June 2000.

NOTE: Available at http://tools.ietf.org/rfc/rfc2866.txt.

[51] IETF RFC 2918: "Route Refresh Capability for BGP-4", September 2000.

NOTE: Available at http://tools.ietf.org/rfc/rfc2918.txt.

[52] IETF RFC 2966: "Domain-wide Prefix Distribution with Two-Level IS-IS", October 2000.

NOTE: Available at http://tools.ietf.org/rfc/rfc2966.txt.

[53] IETF RFC 3037: "LDP Applicability", January 2001.

NOTE: Available at http://tools.ietf.org/rfc/rfc3037.txt.

[54] IETF RFC 3107: "Carrying Label Information in BGP-4", May 2001.

NOTE: Available at http://tools.ietf.org/rfc/rfc3107.txt.

[55] IETF RFC 3123: "A DNS RR Type for Lists of Address Prefixes (APL RR)", June 2001.

NOTE: Available at http://tools.ietf.org/rfc/rfc3123.txt.

[56] IETF RFC 3315: "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", July 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3315.txt.

[57] IETF RFC 3319: "Dynamic Host Configuration Protocol (DHCPv6) Options for Session Initiation Protocol (SIP) Servers", July 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3319.txt.

[58] IETF RFC 3373: "Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies", September 2002.

NOTE: Available at http://tools.ietf.org/rfc/rfc3373.txt.

[59] IETF RFC 3392: "Capabilities Advertisement with BGP-4", November 2002.

NOTE: Available at http://tools.ietf.org/rfc/rfc3392.txt.

[60] IETF RFC 3411: "An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks", December 2002.

NOTE: Available at http://tools.ietf.org/rfc/rfc3411.txt.

[61] IETF RFC 3478: "Graceful Restart Mechanism for Label Distribution Protocol", February 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3478.txt.

[62] IETF RFC 3513: "Internet Protocol Version 6 (IPv6) Addressing Architecture", April 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3513.txt.

[63] IETF RFC 3567: "Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication", July 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3567.txt.

[64] IETF RFC 3623: "Graceful OSPF Restart", November 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3623.txt.

[65] IETF RFC 3630: "Traffic Engineering (TE) Extensions to OSPF Version 2", September 2003.

NOTE: Available at http://tools.ietf.org/rfc/rfc3630.txt.

[66] IETF RFC 3633: "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", December 2003.

NOTE: Available at http://www.ietf.org/rfc/rfc3633.txt.

[67] IETF RFC 3697: "IPv6 Flow Label Specification", March 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3697.txt.

[68] IETF RFC 3719: "Recommendations for Interoperable Networks using Intermediate System to Intermediate System (IS-IS)", February 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3719.txt.

[69] IETF RFC 3784: "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", June 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3784.txt.

[70] IETF RFC 3787: "Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)", May 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3787.txt.

[71] IETF RFC 3847: "Restart Signaling for Intermediate System to Intermediate System (IS-IS)", July 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3847.txt.

[72] IETF RFC 3927: "Dynamic Configuration of IPv4 Link-Local Addresses", May 2005.

NOTE: Available at http://tools.ietf.org/rfc/rfc3927.txt.

[73] IETF RFC 3954: "Cisco Systems NetFlow Services Export Version 9", October 2004.

NOTE: Available at http://tools.ietf.org/rfc/rfc3954.txt.

[74] IETF RFC 3985: "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", March 2005.

NOTE: Available at http://tools.ietf.org/rfc/rfc3985.txt.

[75] IETF RFC 4213: "Basic Transition Mechanisms for IPv6 Hosts and Routers".

NOTE: Available at http://tools.ietf.org/rfc/rfc4213.txt.

[76] IETF RFC 4250: "The Secure Shell (SSH) Protocol Assigned Numbers", January 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4250.txt.

[77] IETF RFC 4271: "A Border Gateway Protocol 4 (BGP-4)", January 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4271.txt.

[78] IETF RFC 4291: "IP Version 6 Addressing Architecture", February 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4291.txt.

[79] IETF RFC 4330: "Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI", January 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4330.txt.

[80] IETF RFC 4360: "BGP Extended Communities Attribute", February 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4360.txt.

[81] IETF RFC 4364: "BGP/MPLS IP Virtual Private Networks (VPNs)", February 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4364.txt.

[82] IETF RFC 4443: "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", March 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4443.txt.

[83] IETF RFC 4447: "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", April 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4447.txt.

[84] IETF RFC 4448: "Encapsulation Methods for Transport of Ethernet over MPLS Networks", April 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4448.txt.

[85] IETF RFC 4456: "BGP Route Reflection - An Alternative to Full Mesh Internal BGP (IBGP)", April 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4456.txt.

[86] IETF RFC 4486: "Subcodes for BGP Cease Notification Message", April 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4486.txt.

[87] IETF RFC 4552: "Authentication/Confidentiality for OSPFv3", June 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4552.txt.

[88] IETF RFC 4577: "OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", June 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4577.txt.

[89] IETF RFC 4659: "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", September 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4659.txt.

[90] IETF RFC 4684: "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", November 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4684.txt.

[91] IETF RFC 4724: "Graceful Restart Mechanism for BGP", January 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4724.txt.

[92] IETF RFC 4760: "Multiprotocol Extensions for BGP-4", January 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4760.txt.

[93] IETF RFC 4762: "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", January 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4762.txt.

[94] IETF RFC 4787: "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", January 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4787.txt.

[95] IETF RFC 4798: "Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)", February 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4798.txt.

[96] IETF RFC 4861: "Neighbor Discovery for IP version 6 (IPv6)", September 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4861.txt.

[97] IETF RFC 4862: "IPv6 Stateless Address Autoconfiguration", September 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4862.txt.

[98] IETF RFC 4893: "BGP Support for Four-octet AS Number Space", May 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc4893.txt.

[99] IETF RFC 5036: "LDP Specification", October 2007.

NOTE: Available at http://tools.ietf.org/rfc/rfc5036.txt.

[100] IETF RFC 5101: "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", January 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5101.txt.

[101] IETF RFC 5102: "Information Model for IP Flow Information Export", January 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5102.txt.

[102] IETF RFC 5120: "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", February 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5120.txt.

[103] IETF RFC 5245: "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", April 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5245.txt.

[104] IETF RFC 5308: "Routing IPv6 with IS-IS", October 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5308.txt.

[105] IETF RFC 5340: "OSPF for IPv6", July 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5340.txt.

[106] IETF RFC 5382: "NAT Behavioral Requirements for TCP", October 2008.

NOTE: Available at http://tools.ietf.org/rfc/rfc5382.txt.

[107] IETF RFC 5443: "LDP IGP Synchronization", March 2009.

NOTE: Available at http://tools.ietf.org/rfc/rfc5443.txt.

[108] IETF RFC 5508: "NAT Behavioral Requirements for ICMP", April 2009.

NOTE: Available at http://tools.ietf.org/rfc/rfc5508.txt.

[109] IETF RFC 5569: "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)", January 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5569.txt.

[110] IETF RFC 5625: "DNS Proxy Implementation Guidelines", August 2009.

NOTE: Available at http://tools.ietf.org/rfc/rfc5625.txt.

[111] IETF RFC 5668: "4-Octet AS Specific BGP Extended Community", October 2009.

NOTE: Available at http://tools.ietf.org/rfc/rfc5668.txt.

[112] IETF RFC 5880: "Bidirectional Forwarding Detection (BFD)", June 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5880.txt.

[113] IETF RFC 5881: "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", June 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5881.txt.

[114] IETF RFC 5883: "Bidirectional Forwarding Detection (BFD) for Multihop Paths", June 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5883.txt.

[115] IETF RFC 5905: "Network Time Protocol Version 4: Protocol and Algorithms Specification", June 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5905.txt.

[116] IETF RFC 5952: "A Recommendation for IPv6 Address Text Representation", August 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5952.txt.

[117] IETF RFC 5969: "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification".

NOTE: Available at http://tools.ietf.org/rfc/rfc5969.txt.

[118] IETF RFC 6036: "Emerging Service Provider Scenarios for IPv6 Deployment", October 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc6036.txt.

[119] IETF RFC 6052: "IPv6 Addressing of IPv6/IPv4 Translators".

NOTE: Available at http://tools.ietf.org/rfc/rfc6052.txt.

[120] IETF RFC 6092: "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", January 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6092.txt.

[121] IETF RFC 6106: "IPv6 Router Advertisement Options for DNS Configuration", November 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc6106.txt.

[122] IETF RFC 6119: "IPv6 Traffic Engineering in IS-IS", February 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6119.txt.

[123] IETF RFC 6144: "Framework for IPv4/IPv6 Translation", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6144.txt.

[124] IETF RFC 6145: "IP/ICMP Translation Algorithm", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6145.txt.

[125] IETF RFC 6146: "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6146.txt.

[126] IETF RFC 6147: "DNS64: DNS Extension for Network Address Translation from IPv6 Clients to IPv4 Servers", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6147.txt.

[127] IETF RFC 6180: "Guidelines for Using IPv6 Transition Mechanisms during IPv6 Deployment", May 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6180.txt.

[128] IETF RFC 6204: "Basic Requirements for IPv6 Customer Edge Routers", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6204.txt.

[129] IETF RFC 6333: "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", August 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6333.txt.

[130] IETF RFC 6334: "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", August 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6334.txt.

[131] IETF RFC 6519: "RADIUS Extensions for Dual-Stack Lite", February 2012.

NOTE: Available at http://tools.ietf.org/rfc/rfc6519.txt.

[132] IETF RFC 6540: "IPv6 Support Required for All IP-Capable Nodes", April 2012.

NOTE: Available at http://tools.ietf.org/rfc/rfc6540.txt.

[133] IETF RFC 6674: "Gateway-Initiated Dual-Stack Lite Deployment", July 2012.

NOTE: Available at http://tools.ietf.org/rfc/rfc6674.txt.

[134] IETF RFC 6877: "464XLAT - Combination of Stateful and Stateless Translation", April 2013.

NOTE: Available at http://tools.ietf.org/rfc/rfc6877.txt.

[135] IETF RFC 6887: "Port Control Protocol (PCP)", April 2013.

NOTE: Available at http://tools.ietf.org/rfc/rfc6887.txt.

[136] IETF RFC 6908: "Deployment Considerations for Dual-Stack Lite", March 2013.

NOTE: Available at http://tools.ietf.org/rfc/rfc6908.txt.

[137] Broadband Forum TR-069 Issue 1 Amendment 4: "CPE WAN Management Protocol".

NOTE: Available at http://www.broadband-forum.org/technical/download/TR-069_Amendment-4.pdf.

[138] CableLabs CM-SP-eRouter-I09-130404: "Data-Over-Cable Service Interface Specifications; IPv4 and IPv6 eRouter Specification", April 2013.

[139] IETF RFC 3056: "Connection of IPv6 Domains via IPv4 Clouds".

[140] IETF draft-ietf-pcp-bas: "Port Control Protocol (PCP)".

NOTE: Available at http://datatracker.ietf.org/doc/draft-ietf-pcp-base/.

[141] Draft-savolainen-heuristic-nat64-discovery: "Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis".

NOTE: Available at http://datatracker.ietf.org/doc/draft-ietf-behave-nat64-discovery-heuristic/?include_text=1

[142] IETF draft-mdt-softwire-mapping-address-and-port-03: "Mapping of Address and Port (MAP)".

NOTE: Available at http://tools.ietf.org/html/draft-mdt-softwire-mapping-address-and-port-03.

[143] IETF draft-mdt-softwire-map-dhcp-option-03: "DHCPv6 Options for Mapping of Address and Port".

NOTE: Available at http://tools.ietf.org/html/draft-mdt-softwire-map-dhcp-option-03.

[144] IETF draft-ietf-softwire-map-02: "Mapping of Address and Port with Encapsulation (MAP)".

NOTE: Available at http://tools.ietf.org/html/draft-ietf-softwire-map-02.

[145] IETF draft MAP DHCPv6 options: "DHCPv6 Options for Mapping of Address and Port".

NOTE: Available at http://tools.ietf.org/html/draft-ietf-softwire-map-dhcp-01.

[146] CM-SP-eDOCSIS-I25-130404: "Data-Over-Cable Service Interface Specifications; eDOCSIS™ Specification".

NOTE: Available at http://www.cablelabs.com/specifications/CM-SP-eDOCSIS-I25-1130404.pdf

[147] Draft-mdt-softwire-map-encapsulation-00: "MAP Encapsulation (MAP-E) - specification".

NOTE: Available at http://tools.ietf.org/html/draft-mdt-softwire-map-encapsulation-00.

[148] Draft-mdt-softwire-map-translation-01: "MAP Translation (MAP-T) - specification".

NOTE: Available at http://tools.ietf.org/html/draft-mdt-softwire-map-translation-01.

[149] Draft-mdt-softwire-mapping-address-and-port-02: "Mapping of Address and Port (MAP)".

NOTE: Available at: http://tools.ietf.org/html/draft-mdt-softwire-mapping-address-and-port-02

[150] Draft-murakami-softwire-4v6-translation-00: "4via6 Stateless Translation".

NOTE: Available at: http://tools.ietf.org/html/draft-murakami-softwire-4v6-translation-00

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI TR 101 569 (V1.1.1): "Access, Terminals, Transmission and Multiplexing (ATTM); Integrated Broadband Cable and Television Networks; Cable Network Transition to IPv6".

[i.2] ETSI TR 102 881 (V1.1.1): "Access, Terminals, Transmission and Multiplexing (ATTM); Cable Network Handbook", June 2010.

[i.3] ETSI ES 201 488 (all parts) (V1.2.2): "Access and Terminals (AT); Data Over Cable Systems".

[i.4] ETSI ES 202 488 (all parts) (V1.1.1): "Access and Terminals (AT); Second Generation Transmission Systems for Interactive Cable Television Services - IP Cable Modems".

[i.5] IETF RFC 4008: "Definitions of Managed Objects for Network Address Translators (NAT)", March 2005.

NOTE: Available at http://tools.ietf.org/rfc/rfc4008.txt.

[i.6] IETF RFC 4087: "IP Tunnel MIB", June 2005.

NOTE: Available at http://tools.ietf.org/rfc/rfc4087.txt.

3 Symbols and abbreviations

3.1 Symbols

For the purposes of the present document, the following symbols apply:

µs	Microsecond
Gbit/s	Gigabit per second
Mbit/s	Megabit per second
MHz	Megahertz
ms	Millisecond

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

3GPP	Third Generation Partnership Project
4rd	IPv4 Residual Deployment
6PE	IPv6 Provider Edge
6rd	IPv6 Rapid Deployment
6VPE	IPv6 VPN Provider Edge
A	(DNS) A Resource Record
A+P	Adress Plus Port
A6	(DNS) A6 Resource Record (historic)
AAAA	(DNS) AAAA Resource Record
ACL	Access Control List
AFTR	Address Family Transition Router
ALG	Application Layer Gateway
ALP	Application Layer Proxy
ALTS	Application Layer Translation Service
APL	Access Prefix List
ARP	Address Resolution Protocol
AS	Autonomous System
ASCII	American Standard Code for Information Interchange
ASIC	Application Specific Integrated Circuit
B2B	Business to Business
B4	(DS-Lite) Basic Bridging BroadBand element
BCP	Best Current Practise
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
BNG	Broadband Network Gateway
BR	Boarder Router
BR/GW	Boarder Relay / Gateway
BW	Bandwidth
CDP	Cisco Discovery Protocol

CE	Customer Edge
CEF	Cisco Express Forwarding
CGN	Carrier-Grade NAT
CGN-GW	Carrier Grade Network Address Translation - Gateway
CIDR	Classless Inter-Domain Routing
CLAT	Customer-side XLAT
CM	Cable Modem
CMTS	Cable Modem Termination System
CoPP	Control Plane Policing
CPE	Customer Premises Equipment
DAD	Duplicate Address Detection
dCEF	Distributed CEF
DCU	Destination Class Usage (accounting method in routers)
DF bit	Do not Fragment flag (in IPv4 header)
DF	Don't Fragment
DHCP	Dynamic Host Configuration Protocol
dIVI	Dual Stateless IPv4/IPv6 Translation
DMZ	DeMilitarized Zone
DNS	Domain Name System
DOCSIS	Data Over Cable Service Interface Specification
DR	Data Retention
DS	Dual Stack
DSCP	Differentiated Services Code Point
DS-Lite	Dual Stack Lite
DUID	DHCP Unique Identifier
eBGP	External BGP
ECMP	Equal Cost MultiPath
ECN	Explicit Congestion Notification
ESM	Enterprise Subscriber Management
FTP	File Transfer Protocol
GPON	Gigabit Passive Optical Network
GRT	Global Routing Table
GUA	Global Unicast Address
GW	GateWay
HA	High Availability
HE	HeadEnd
HFC	Hybrid Fiber-Coax
HSRP	Hot Standby Router Protocol
HTML	HyperText Markup Language
IA_NA	Identity Association for Non-Temporary Addresses (DHCPv6 option)
IANA	Internet Assigned Numbers Authority
iBGP	Internal BGP
ICMP	Internet Control Message Protocol
ID	Identity
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IMIX	Internet Mix
IP	Internet Protocol
IP-FIX	Internet Protocol Flow Information Export
IPFIX	IP Flow Information Export
IPTV	Internet Protocol Television
IPv4	IP version 4
IPv6	IP version 6
IRB	Integrated Routing and Bridging
ISIS	Intermediate System to Intermediate System
IS-IS	Intermediate System To Intermediate System
ISP	Internet Service Provider
ISSU	In-Service Software Upgrade
L2	Layer 2
L2TP	Layer 2 Tunneling Protocol
L3	Layer 3

LAG/LACP	Link Aggregation Group/Link Aggregation Control Protocol
LAN	Local Area Network
LDP	(MPLS) Label Distribution Protocol
LLDP	Link Layer Discovery Protocol
LSN	Large Scale NAT
MAC	Media Access Control
MAP	Mapping of Address and Port
MAP-E	Mapping of Address and Port - Encapsulation Mode
MAP-T	Mapping of Address and Port - Translation Mode
MFIB	Multicast Forwarding Information Base
MIB	Management Information Base
MoCa	Multimedia over Coax Alliance
MP-BGP	MultiProtocol BGB
MPLS	MultiProtocol Label Switching
MSO	Multiple Service Operator
MSS	(TCP) Maximum Segment Size
MT	(for IS-IS)
MTA	Media Terminal Adapter
MTU	Maximum Transmission Unit
MULPI	Media Access Control and Upper Layer Protocol Interface
NAPT	Network Address and Port Translation
NAT	Network Address Translation/Network Address Translator
NCC	Network Coordination Centre
ND	Neighbor Discovery
NDP	Neighbor Discovery Protocol
NPU	Network Processing Unit
NTP	Network Time Protocol
NUD	Neigbor Unreachability Detection
OAM	Operation, Administration and Maintenance
OSI	Open System Interconnection
OSPF	Open Shortest Path First
OSS	Operational Support System
PC	Personal Computer
PCP	Port Control Protocol
PD	Prefix Delegation
PE	Provider Edge
PIM	Protocol Independent Multicasting
PLAT	Provider-side XLAT
PMP	Port Mapping Protocol
PMTU	Path MTU
PMTUD	PMTU Discovery
PPTP	Point-to-Point Tunneling Protocol
QoS	Quality of Service
QPPB	QoS Policy Propagation via Border Gateway
RADIUS	Remote Authentication Dial-In User Service
RARP	Reverse ARP
RDNSS	Recursive DNS Server
RDT	Reliable Data Transfer
RE	Routing Engine
RFC	Request For Comments
RG	Residential Gateway
RG/CPE	Routing Gateway / Customer Premises Equipment
RIP	Routing Information Protocol
RIPE NCC	RIPE Network Coordination Centre
RIPE	Réseaux IP Européens
RP	Route Processor
RR	(DNS) Resource Record
RTSP	Real-Time Streaming Protocol
SCTP	Stream Control Transmission Protocol
SCU	Source Class Usage (accounting method in routers)
SEND	Secure Neighbor Discovery
SI	(DS-Lite) Softwire Initiator

SI-ID	Softwire Initiator Identifier
SIP	Session Initiation Protocol
SLAAC	StateLess Address Auto Configuration
SNMP	Simple Network Management Protocol
SP	Service Provider
SSH	Secure SHell
SYSLOG	Syslog Protocol
TACACS	Terminal Access Controller Access Control System
TCP	Transmission Control Protocol
TFTP	Trivial File Transfer Protocol
TLV	Tagged Length Value
ToD	Time of Day Protocol
TOS	Type Of Service
TSP	Tunnel Setup Protocol
TV	Television
UDP	User Datagram Protocol
UI	User Interface
UPnP	Universal Plug and Play
US	United States
VLAN	Virtual LAN
VoIP	Voice over IP
VPLS	Virtual Protocol Local Area Network Service
VPN	Virtual Private Network
VRF	Virtual Routing and Forwarding
VRRP	Virtual Router Redundancy Protocol
WAN	Wide Area Network
Web-UI	Web User Interface
XLAT	(Address) Translator
XML	eXtensible Markup Language

4 Specification Construct

In the present document, requirements are presented following specific constructs. The constructs are instrumental to define features in a consistent way. The constructs contain various headings that are structuring the feature definition and provide necessary data and references. The same constructs are used for performance, hardware and compatibility requirements within other parts of this multi-part deliverable. Each feature is broken down to describe the 'feature' its 'premise', its function and the relevant 'RFC reference' as illustrated in the example below.

5 Background and Concept of Transition

IPv6 transition is critical to the long-term sustainability of European and global networks in order to ensure business continuity. With more and more services and industries coming to rely on the global Internet as a fundamental platform the need for ubiquitous connectivity of devices and services becomes very urgent. There is no alternative to eventual transition to IPv6 in all parts of the network also for MSOs. Such near-term strategic areas like Mobile Internet and Smart Grids as well as the continued growth in residential and business broadband access services are poised to introduce massive numbers of devices that require network connectivity, which may not easily be provided by the current Internet (IPv4) networks with their depleting address space.

Widespread adoption of IPv6 has been identified as the best way forward to address the exhaustion of the IPv4 address space. Prompt and efficient adoption offers potential for innovation and leadership in advancing the Internet, while delayed adoption of IPv6 would lead to disadvantages for all users and a weaker competitive position of the industry. In the meantime, the IANA Unallocated IPv4 Address Pool was exhausted on 3 February 2011, and the RIPE NCC IPv4 Address Pool is empty as of 14 September 2012 (http://www.ripe.net/internet-coordination/ipv4-exhaustion). The urgency to transition broadband Internet networks to IPv6 is becoming critical.

Device manufacturers, software developers and network operators are adopting IPv6. For Broadband Cable Network providers the ability to continue to connect new customers and to provide new services to customers while maintaining current levels of investment and ensure growth are the main drivers to mitigate the impact from depletion of IPv4 addresses and to adopt IPv6 in the access and customer network. However, the vision of an Internet running IPv6 only will not become a full reality any time soon. For a considerable period of time, significant numbers of devices and services will exist that customers want to use and that were designed to require IPv4 connectivity. Among the more prominent examples for such devices and services are IPTV sets with static firmware, IP-connected refrigerators and other appliances, security devices and devices that cannot be upgraded through software to enable an IPv6 stack. These devices have a longevity of up to eight or ten years and, thus, transitional requirements may be needed at least until 2020.

An immediate replacement of these IPv4 hosts and networks may not be feasible or not desirable for various technical and economic reasons. It is particularly the task of access network operators and broadband service providers to ensure customer choice in terms of IPv6 technology and services, as well as the continuation of IPv4 connectivity. Appropriate transition technologies enable and allow the coexistence of IPv6 and IPv4 in various parts of the end-to-end network. By deploying transition technologies as needed, services may be consumed and customer premises equipment may be used transparently while fostering a smooth transition to the required extended address space provided by IPv6. Simultaneous connectivity among IPv4 and IPv6 hosts and services by employing appropriate transition technologies ensures e.g. the ability to offer IPv4 services even though the address pool has been depleted and the consumer is connected via IPv6.

5.1 Broadband Cable Network Providers

As of today, integrated broadband cable and television networks go into the home of more than 73 million customers in the European Union providing Digital TV, Broadband Internet, and Telephony services. Broadband Internet services provided by Cable Networks utilizing DOCSIS cable modem technology enable about 23,2 million subscribers in Europe (2010) to connect to the Internet with download speeds of currently up to 200 Mbit/s. This figure has grown by at least 12 % annually and is expected to continue to grow. Cable Network operators provide the platform to satisfy fundamental entertainment, communication and information needs to consumers using equipment supplied a complete ecosystem of manufacturers and service providers. Furthermore, the industry is anticipating a transition to delivery of digital television using broadband cable modem technology, which will dramatically increase the number of broadband connected households.

Cable Networks are recognized as one of the key enablers in supporting Europe's Digital Agenda. To continue to meet the demand of accelerating connectivity of digital devices, a standardized approach for the cable eco-system to rapidly transition to IPv6 is required. A failure for an effective standards driven transition would impair the ability to achieve cost effective solutions on a large scale.

5.2 IP Connectivity in Cable Networks

Integrated broadband cable and television networks are built against various international and ETSI standards. Figure 1 depicts the fundamental architecture of a Cable Network as it is currently deployed with a hybrid fibre-coax approach. [i.2] provides a complete overview on Cable Network architectures and services. The IP communication system in Cable Networks is based on the series of ETSI DOCSIS [i.3], [i.4], [2] and PacketCable standards [1]. Since 2006, the current version of the DOCSIS technology (DOCSIS 3.0) has natively supported IPv6. For various technical reasons, many service providers in Europe have not yet implemented IPv6 support for their customers.

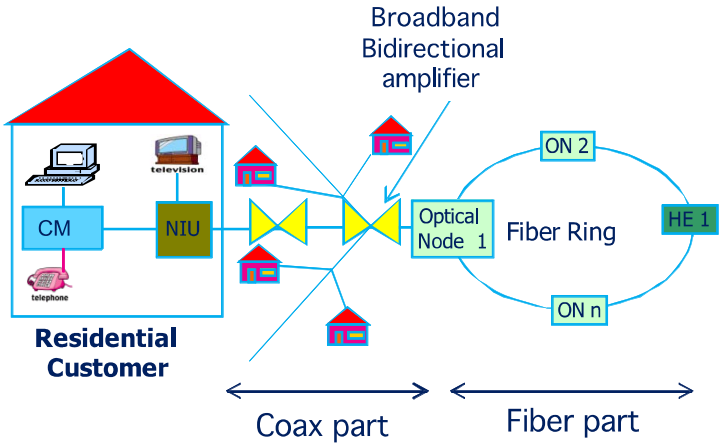


Figure 1: Principle architecture of a hybrid fibre-coax (HFC) Cable Network

In order to achieve the end-to-end connectivity, the Cable Network is interconnected at the Headend (HE) with a backbone network and the Internet. The backbone network may or may not be operated and managed by the same entity as the access network. The home network within the customers' premises is typically installed, configured and operated by the customer.

For the purposes of the analysis in the present document, the end-to-end network is subdivided into various parts, each of which currently supports IPv6 and/or IPv4 with various degrees of probability. The network parts also distinguish themselves by the number of times they occur globally and by the uniformity of their operation and management. Both aspects have an impact on how difficult it is to homogeneously transition to IPv6.

5.2.1 Customer Host

This is the device that the customer uses to consume an IP-based service. It is the final destination and the originating source of IP packets. As Customer Premises Equipment (CPE) it is typically owned and operated by the customer, Cable Network operators are unlikely to be able to support and manage the wide variety of host devices that customers are deploying in their homes.

5.2.2 Home Network

The Home Network extends from the Customer Host to the Access Gateway. It may be constituted by a simple wired or wireless link between the two devices (in which case it becomes irrelevant for this analysis since it is not addressable and does not process protocol messages) or may consist of a complete infrastructure of routers, wireless access points etc. It may connect multiple Customer Hosts to the Access Gateway and it is typically owned, installed and operated by the customer. Cable Network operators may impact the capabilities of the Home Network up to a certain extent by providing information to customers and/or delivery of devices.

5.2.3 Access Gateway

This device is installed at the customer location and constitutes the separation between the Home Network and the Access Network. While the device may be owned by the customer or the Cable Network operator it is typically authorized for usage by the latter. It terminates the Cable Network on the customer side. While the Access Gateway may contain several functions such as a Wi-Fi Access Point or functionality for video decoding, the most fundamental building block for the purposes of the present document is the Cable Modem.

5.2.4 Access Network

The Access Network extends from the Access Gateway to the Headend. It is operated and managed by the Cable Network operator. It typically uses DOCSIS technology to establish IP connectivity and packet transport.

5.2.5 Headend

This device separates the Access Network from the Backbone Network. It terminates the Cable Network on the operator side. For the purposes of the present document it also terminates the portion of the end-to-end IP connection that traverses the operator-managed network. As such it will be the location of devices that are providing transition technology services. However, that does not mean that all Headend functionality is concentrated at a single geographical location.

5.2.6 Internet

This is the Internet cloud that is considered to be the host of IP-based services. It is the final destination and the originating source of IP packets which may be delivered as IPv4 or IPv6. The actual delivery of the service or the path the packet has taken through the Internet cloud is irrelevant for the present document, only the way of addressing is taken into account.

Due to the inherent incompatibility of IPv4 and IPv6 addressing schemes, special measures have to be taken (i.e. transition technologies have to be deployed) in order to ensure connectivity between equivalent network parts supporting different Internet protocols (e.g. IPv4 host talking to IPv6 service across IPv4 home network and IPv6 access network and backbone). The items listed below summarize the types of connectivity that are considered in the present document.

- IPv4 Customer Host via IPv4 Access Network connecting to IPv4 Internet
  This is out-of-scope of the present document as it is provided by native IPv4
- IPv6 Customer Host via IPv4 Access Network connecting to IPv4 Internet
  This is out-of-scope of the present document as it is within customer responsibility
- IPv4 Customer Host via IPv6 Access Network connecting to IPv4 Internet
- IPv6 Customer Host via IPv6 Access Network connecting to IPv4 Internet
- IPv4 Customer Host via IPv6 Access Network connecting to IPv6 Internet
- IPv6 Customer Host via IPv6 Access Network connecting to IPv6 Internet
  This is out-of-scope of the present document as it is provided by native IPv6
- IPv4 Customer Host via IPv4 Access Network connecting to IPv6 Internet
- IPv6 Customer Host via IPv4 Access Network connecting to IPv6 Internet

5.2.6.1 Transition Technology Summary

Internet Protocol version 6 (IPv6) is a technology which is beginning to be adopted by network operators and ISPs across global markets as IPv4 addresses become depleted in all registries. However, despite its maturity as a standard and the gradual momentum for its adoption, the integration of IPv6 within an end-to-end network is being found to be in its early years given the significant investments in legacy network architectures. Technologies that were developed in order to work around the limitations of the available IPv4 address pool such as NAT and CIDR provided stopgaps towards the depletion of IPv4 addresses and, thus, slowed the process of transitioning to a more advanced protocol. All organizations involved in the development and application of Internet protocols taken as a community have failed to come together in a single voice of reasoning. The potential of an IPv6 all-in day where IPv6 should have been adopted as a worldwide agreement allowing all industries to move together has not materialized. Instead, market reality necessitates consideration of transitional functionality and network migration.

The present document - focusing on issues encountered by Cable Network operators when considering IPv6 - studies the potential transition technologies that may be utilized within the MSO end-to-end network to effectively and efficiently enable a transition to IPv6 from IPv4. The evolution in the development of IPv6 has failed to consider interoperability with IPv4 as a key requirement. With hindsight, had this been a consideration during the development of IPv6 then the industry would have been equipped with the means to integrate the IPv6 solution alongside the existing IPv4 deployed network architecture. For example, it may have been possible to use a /96 address range to embed the full IPv4 range in the IPv6 addressing structure. However, since no consideration had been given to interoperability it meant the industry was offered two protocols, IPv4 and IPv6, that are not interoperable with each other.

The industry today is left to manage network topologies using IPv4 and IPv6 since IPv6 had not been designed to be backward compatible with IPv4. Consequently, a network topology may be designed to have end-to-end architectures of both IPv4 and IPv6, unless the industry were to develop a suitable transition technology. The former approach of both protocol versions existing in parallel is not always economic. Therefore, network operators are assessing with some urgency appropriate transition technologies to maintain continuity of their customers' services once the IPv4 addresses are depleted. This presents the industry with challenges to find ways to support network operators with their currently deployed equipment that may not be IPv6 capable by developing suitable transition technologies to allow for a smooth entry into IPv6.

This will result in not just a major change in the basic network infrastructure of the Internet but all of its associative requirements and functions and is poised to have far-reaching effects that one cannot possibly calculate in manpower and administration and that are mainly due to the ubiquity of the Internet today. However, one can place a guess in the direction of some 16 billion US dollars of cost worldwide that would not have to be met without the need to change. In the present document, basic issues of the transition from the current IPv4 networks towards IPv6 are addressed giving a brief overview of how the transition can happen and an introduction to the relevant technical issues in this area along with appropriate solutions.

The primary problem that is being addressed by moving to IPv6 is the lack of IPv4 addresses with the current addressing scheme although this is not the only issue. IPv6 offers 128 bit addresses, which are foreseen to be large enough for future purposes providing some 5 x 10^28 addresses. While changing the address structure where each address is represented by eight sections containing two bytes each typically separated by colons, various other features are being built into the new Internet Protocol to easily enable features like security, QoS, mobility, etc. (e.g. CoPP, Netflow, ECN, ECMP).

NOTE: These features can be used with IPv4 also, but IPv6 is optimized for their application.

The change in the IP address structure impacts the whole networking stack in transit and end nodes (Layers 2 and 3 but not Layer 1 in the OSI layer model). IPv6 does not just disturb the Layer 3 functionality, new protocols such as NDP also change Layer 2 functions to a certain degree with multicast being used instead of broadcast while link-local neighbor tables and destination cache replace the ARP/RARP function. All intelligent protocols that are located within Layers 2 and 3 as well as some protocols associated with Layer 4 has to change to accept the new structure and role IPv6 plays within the stack.

Some systems and applications may not have an upgrade path to IPv6 and, thus, will remain IPv4 clients until they are taken out of service, while others may be expensive or impractical to upgrade in a timely manner. Hence, there is a requirement for co-existence of the two networking technologies and internetworking requirements have emerged.

Since parts of the existing Internet backbone may continue to remain an IPv4 network, the scenario of how IPv6 will be deployed needs to be understood.

Next generation wireless networks are a hot topic in today's world. MSOs are being required to consider integrating mobile services into their own portfolio leveraging technologies and protocols developed in the wireless world. This does actually help to mitigate the cost of potential transition solutions in CPE and other network equipment. Irrespective of whether European or international standards are followed, it is clear that the IPv6 enabled end-user terminal will have to support both voice and data features. While the characteristics of data traffic do not require the device to be a uniquely addressable entity (it is possible to have a scenario of private addresses and usage of an address translation gateway), a voice channel requires the setup of a peer-to-peer connection where the terminal needs to be directly addressable in order to be contacted by external applications. With the arrival of broadband technologies and the always-on Internet, even data applications (for example multi-party gaming) bring in requirements for unique addresses. With the current, rapidly diminishing pool of IPv4 addresses, it would be infeasible to support such requirements. Hence, wireless networks become the most important business case to serve as a driver to migrate towards IPv6.

Even within wired networks, IP has become an essential part of the network particularly for MSOs as well as increasingly for incumbent network operators even for voice services. There is a move away from circuit switched networks towards packet switched IP networks while attempting to retain some of the features of circuit switching (for example end-to-end connectivity) through additional features and applications built on top of IP.

The pressure to migrate to IPv6 is being felt in all parts of the world to various degrees. In some countries such as Japan the scarcity of IPv4 addresses is felt more urgently and, hence, it is taking the lead to move towards IPv6. Countries like the United States, which received huge blocks of IPv4 addresses during the early days of the Internet, are feeling less immediate pressure to move towards IPv6, although that might not be equally applicable to all types of ISPs. As a result, the global Internet will not migrate along a uniform transition timeline, but instead will first transition to IPv6 in regional pockets where the need is greatest.

To briefly cover the deployment status of IPv6, it is widely deployed in test networks especially in Japan and Europe which are being followed by China and Korea. Japanese ISPs have started commercial sales of IPv6 addresses. In the United States, the largest ISPs (including MSOs) have begun deployment of IPv6 to their customers. However, overall deployment is still limited. The availability of early test networks such as the 6bone network (consisting of independent sets of IPv6 networks across the globe, linked together by special tunnelling mechanisms over the IPv4 Internet) provided the industry not only direction to the importance of moving to IPv6 but also provided a test bed for native IPv6 implementations and transition testing.

6 IPv6 Transition Technologies Specifications

Global IPv4 address space is currently projected to be depleted around the middle of 2012 to 2015 for most MSOs. As part of the resulting rollout of IPv6 in the worldwide networking footprint specific measures has to be taken to allow a smooth transition and coexistence between IPv4 and IPv6 capable network devices and services. In [i.1], transitions technologies currently specified or proposed have been subject to a thorough technical analysis focusing on their suitability for deployment in a Cable Network environment and their impact on Cable Network equipment.

The benefits and drawbacks of those transition technologies that are developed to a certain level of maturity were analyzed and prioritized against the specific characteristics of Cable Networks. As a result, five transition technologies as summarized in table 1 have been found to have sufficient possible potential to serve the requirements of MSOs.

In the present document, end-to-end system requirements for broadband Cable Networks as well as equipment specifications that enable the transition from IPv4 to IPv6 are defined for each of the selected transition technologies.

Table 1: Selected Transition Technologies

Technology	Technical Description	Summary Analysis
NAT64	Using NAT64, only native IPv6 connectivity is provided to the end customer. End-to-end IPv6 connectivity is provided natively between end-hosts and Internet services. To provide communication from IPv6 to an IPv4 host or Internet service a NAT64/DNS64 service needs to be deployed in the network.	NAT64 is a transition technology using IPv6 in the home network. The main caveats being: 1 - Applications who do not use DNS will not work across the network. 2 - It requires the complete home network to be IPv6 enabled and it does not support IPv4 only devices in the home network. 3 - It requires an IPv6 native transport from CPE to the NAT64 device. 4 - Extensive ALGs are required due to the fact that the NAT64 device translates IPv6 to IPv4 and vice versa. 5 - It requires PCP to prevent service deprecation on subscribers who previously had IPv4 public addresses and provided Internet access to their home. 6 - Deployment requires NAT64 to be carrier-grade in terms of its performance and features. NAT64 is not widely adopted as a transition technology so far given the above complications mainly mentioned in items 1 and 2.

Technology	Technical Description	Summary Analysis
DS-Lite	DS-Lite is based on a tunnel between the CPE and the AFTR. This tunnel is created through a SI of an IPv6 encapsulation of the IPv4 packet from CPE to AFTR. The IPv4 customer side is based on private addressing and is translated to public addresses at the AFTR. The CPE, therefore, is only required to perform a single encapsulation and the AFTR is performing two, one to encapsulate and decapsulate an IPv4 within an IPv6 packet and the second one to translate private to public IPv4 addresses via NAT.	DS-Lite is an almost non-service deprecating technology with a few caveats: 1 - It requires a new CPE to be delivered to the customer location supporting the features required for SI. 2 - It requires an IPv6 native transport from CPE to AFTR. 3 - Deployment requires DS-Lite to be carrier-grade in terms of performance and features. 4 - It requires PCP to prevent service deprecation on subscribers who previously had IPv4 public addresses and provided Internet access to their home. DS-Lite is a commonly supported in devices connected to Cable Networks due to the functionality potentially built in. DS-Lite does not require to further assign IPv4 private addresses in the network and uses a common form of NAT for IPv4 which is a proven technology.
6rd	When deploying 6rd, end devices are supplied with dual stack addressing. Both IPv4 and IPv6 addresses are provided to the customer. IPv4 connectivity in this model is provided the same way as is done today, using private or public IPv4 addressing. IPv6 connectivity is provided using 6to4 tunneling (RFC 3056 [139]) where the standard 6to4 prefix 2002::/16 is changed by an IPv6 prefix that belongs to the ISP-assigned address space. The IPv6 prefix allocated to the end customer is derived from the IPv4 address assigned to the CM. The v4suffix-length, v6prefix-length, 6rd Border Relay IPv4 (Anycast) Address and 6rd SP Prefix are provided to the CM using DHCP. To deploy 6rd, a 6rd CM needs to be deployed in conjunction with a 6rd BR which provides 6to4 tunneling. More details can be found in RFC 5969 [117].	6rd is a dual stack transition technology providing with some caveats. The main caveats being: 1 - The CM needs to support 6rd and a 6rd border relay needs to be added to the network. 2 - Since the IPv6 address is provided from a IPv4 address, while IPv4 exhaustion is happening, the IPv6 address is subject to NAT with its implications and a second transition to native IPv6 support will be required, which adds costs. 3 - It requires PCP to prevent service deprecation on subscribers who run services from their home. 4 - Deployment requires to be carrier grade feature wise on the 6rd border relay. 6rd is mainly used when the access network cannot transition to IPv6, but given a second transition is required to native IPv6 the technology is not widely adopted.
NAT44	So NAT44 resides at the edge of the network where it interfaces to the public Internet. The NAT has one or more globally unique IPv4 addresses and use a CORE GW and a CPE and as a packet passes from its inside or private interface to its outside or public interface, NAT replaces the packet's private IPv4 address with one of its public IPv4 addresses. The NAT "remembers" which inside device the packet came from by mapping the inside address to the outside address. It requires ALGs and functional support for PCP as it uses NAPT for improved ratios of IP publics to customer utilization.	NAT44 is a legacy technology allowing the use of NAPT to generate much less requirement for IPv4 public addressing. 1 - It requires no new CPEs. 2 - It requires only an additional "NAT" device or blade in the network. 3 - It is an established technology and thus we know what we are getting. 4 - It requires PCP and DeNAT to prevent service deprecation on subscribers who previously had IPv4 public addresses and provided Internet access to their home and to create acceptable levels of DR logging requirements. 5 - It has some failings that it usually enforces the end-to-end into a NAT444 requirement as the home users tend to have their own NAT on their home devices. Nested NAT brings it own problems. 6 - It pushes private address ranges into your network which is not an optimal solution.

Technology	Technical Description	Summary Analysis
464XLAT	The IPv6 address format in 464XLAT is defined in section 2.2 of RFC 6052 [119]. Prefix delegation mechanism such as DHCPv6-PD as described in RFC 3633 [66] are available to assign a dedicated translation prefix to the CLAT. From the delegated DHCPv6 prefix, a /64 is dedicated to source and receive IPv6 packets associated with the stateless translation as described in RFC 6145 [124]. The CPE (CLAT) MAY discover the Pref64::/n of the PLAT via some method such as DHCPv6 option, TR-069, DNS APL RR as described in RFC 3123 [55] or in Internet Draft ietf-behave-nat64-discovery-heuristic.	This technology allows customers to access services natively over IPv6 and through translation over IPv4 and is thus a last resort technology. IPv4 connectivity to IPv4 hosts over home routers (CPEs) and access networks that are provisioned with only IPv6 addresses. Dual-Stack connectivity for hosts connected to IPv6-only access networks. Less need to maintain IPv4 or dual-stack access networks A lightweight solution for providing IPv4 connectivity over IPv6 only access. Single NAT - i.e. no need to have multiple layers of NATs. Multiplexing public IPv4 addresses for large number of customers across a limited number of IPs using port translation. Port forwarding capability on a PLAT using technologies such as: Web-UI, NAT-PMP, UPnP, A+P. It does not require DNS64 as described in RFC 6147 [126] since an IPv4 host may simply send IPv4 packets, including packets to an IPv4 DNS server, which will be translated on the CLAT to IPv6 and back to IPv4 on the PLAT.
MAP-E	Mapping of Address and Port - Encapsulation Mode MAP (MAP-E) is a tunnel technology but unlike CGN it places the processing and state on the CPE. Prior to being standardized, yet still in draft, MAP-E was commonly referred to as IPv4 Residual Deployment (4rd) and is based on that technology. MAP-E enables a service provider to allow IPv4 services to IPv6 (customer) sites to which it provides MSO customer premise equipment (CPE). This approach utilizes stateless IPv4-in-IPv6 encapsulation (i.e. tunneling) to transit IPv6-enabled network infrastructure. The encapsulation could be supported by the CPE, but very few vendors have it on a present product and MAP-E Gateway/Border Relay, which removes the IPv6 encapsulation from IPv4 packets while forwarding them to the Internet. The provider access network can now be on IPv6, while customers see IPv6 and IPv4 service simultaneously.	MAP-E is a viable technology in its basis but lacks vendor support and a full feature set, plus standardization is lacking at present and incomplete. **BR** Native IPv6 WAN support Dual-stack IPv4 / IPv6 LAN support IPv4 to IPv6 header translation in accordance with RFC 6145 [124] NAPT-44 to translate private RFC 1918 [30] addresses to public IPv4 address and port range Provision of IPv6 prefix via DHCPv6 DHCPv6 MAP-T options as defined in draft-mdt-softwire-map-dhcp-option-03 IPv6 to IPv4 header translation Port forwarding mapping for NAT translation Support for UPnP NAT Traversal MSS Clamping for TCP/IPv4 connection negotiation PMTUD support for both IPv4 and IPv6 support Fragmentation / reassembly of IPv6 packets **CPE** IPv6 prefix assigned for IPv6/IPv4 translation (customer-side) IPv4 address for forwarding to / from IPv4 Internet IPv4 to IPv6 header translation in accordance with RFC 6145 [124] IPv6 to IPv4 header translation MSS Clamping for TCP/IPv4 connection negotiation PMTUD support for both IPv4 and IPv6 support Fragmentation / reassembly of IPv6 packets

6.1 CPE Specification

6.1.1 Introduction

The present document describes the requirement for customer premise equipment to support the transition from IPv4-based Internet services to IPv6-based Internet services utilizing DOCSIS equipment.

The requirements to support the variety of transition techniques calls for the implementation of an Embedded Service/Application Functional Entity (eSAFE) device as described by the eDOCSIS specification (eDOCSIS is part of the DOCSIS compendium of standards) as developed by CableLabs. These eSAFE devices is embedded within an eDOCSIS device containing an integrated cable modem that conforms with the DOCSIS specification and is also applicable to EuroDOCSIS devices.

The core features described in the present document will allow for multiple transition techniques to be applied by a cable Multiple Service Operator (MSO) in their transition strategy as IPv4 based address schemes become depleted.

6.1.2 Functional Description

The primary purpose of the device described connects devices within a consumer's home network to a High Speed Internet service provided by a DOCSIS or EuroDOCSIS based Cable Modem Termination System (CMTS). Baseline DOCSIS specifications allow for the connection of multiple local IP clients, each with a unique IPv4 address allocated to the customer premise equipment device.

However, the present document addresses the base requirements for the provision of Internet services to devices in the home through the implementation of an eSAFE device embedded with a cable modem rather than the baseline DOCSIS methodology.

The device MAY have multiple customer facing interfaces connected with options including wireless, fixed Ethernet, or alternate home networking interfaces such as MoCA, HomePlug, G.Hn, etc.

Figure 2, Logical components contained within an eDOCSIS device with the addition of an IPv4 and IPv6-capable eRouter with eSAFE functionality.

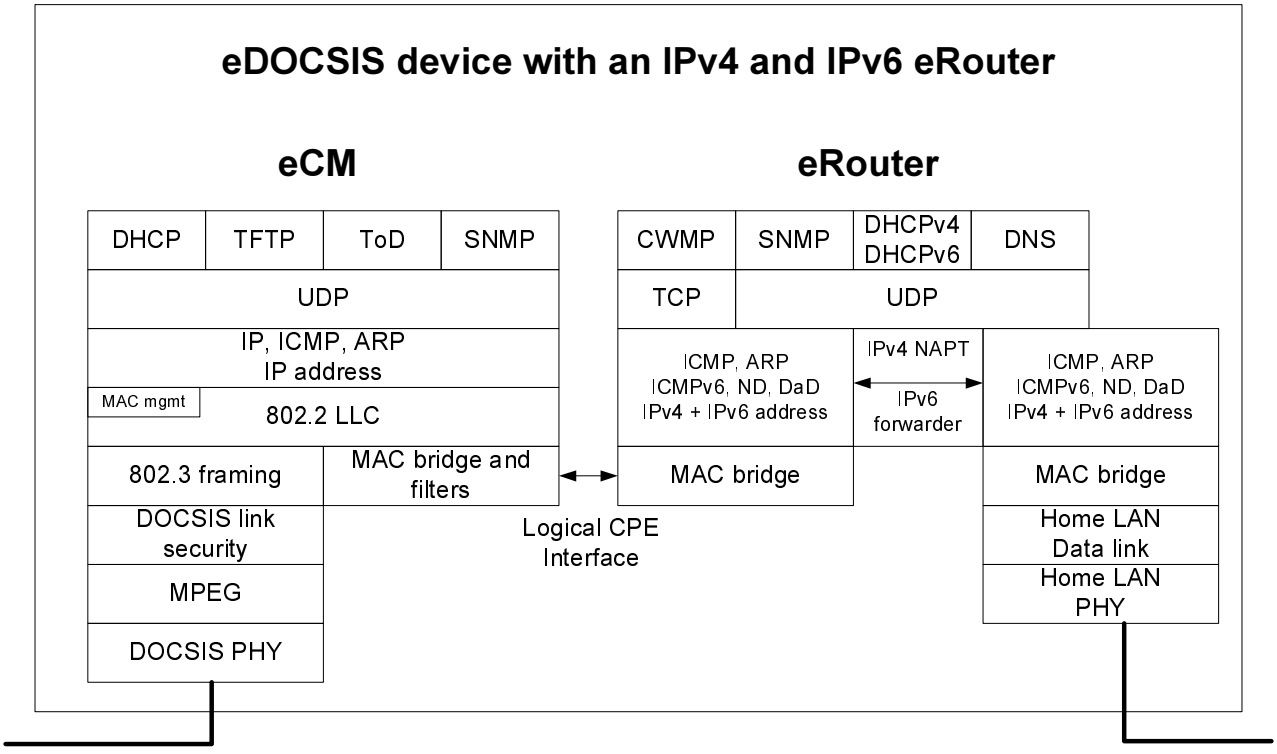


Figure 2: Figure 5-16 from CM-SP-eDOCSIS-I25-130404 [146]

Premise: Device Initialization.

Function: The eSAFE router SHALL obtain its operational parameters during the initial DOCSIS registration process. The correct operational mode is obtained through the use of TLV parameters.

6.1.3 Summary Base Function: Device Initialization

6.1.3.1 Feature:Device Initialization

Premise: In order to operate in the IP mode desired by the cable operator, the eSAFE router needs to receive its operational parameters during registration of the eDOCSIS device's embedded CM.

Function: The CPE device SHALL obtain its operational parameters during initial DOCSIS device embedded CM registration. These parameters are received through the use of TLV statements as defined in "IPv4 and IPv6 eRouter Specification CM-SP-eRouter-I09-130404" as defined by CableLabs [138].

Table 2: Operational Modes (table 6-1 from CM-SP-eRouter-I09-130404) [138]

Mode	IPv4	IPv6
Disabled	No IPv4 provisioning, CM bridges all traffic per [MULPI] spec.	No IPv6 provisioning, CM bridges all traffic per [MULPI] spec.
IPv4 Protocol Enabled	IPv4 Provisioning (section 7) IPv4 data forwarding using NAPT (section 9).	No IPv6 provisioning. No IPv6 data forwarding between Operator Facing Interface and the Customer Facing Interfaces.
IPv6 Protocol Enabled	No IPv4 provisioning. No IPv4 data forwarding between Operator Facing Interface and the Customer Facing Interfaces.	IPv6 Provisioning (section 8) IPv6 data forwarding (section 10).
Dual IP Protocol Enabled	IPv4 Provisioning (section 7) IPv4 data forwarding using NAPT (section 9).	IPv6 Provisioning (section 8) IPv6 data forwarding (section 10).

The CPE device SHALL be capable of operating in any one of these modes in order to be able to implement the range of transition strategies defined.

6.1.4 Disabled Mode

In disabled mode, the device will be operating as a layer-2 bridge for all traffic. No support for transition technologies will be provided in this mode.

In this mode, either a single PC or a third-party wired/wireless home gateway would be required. The support of transition technologies on these devices is outside the scope of the present document.

6.1.5 IPv4 Protocol Enabled Mode

Premise: Support for 6rd Transition strategies.

Functions: The device SHALL provision IPv4 and operate Network Address and Port translation (NAPT) in accordance with [eRouter].

In order to implement 6rd functionality, the CPE device SHALL operate in IPv4 Protocol Enabled Mode only.

MAP and DS-Lite functionality SHALL NOT be enabled when the device is in IPv4 Protocol Enabled Mode.

6.1.6 Dual IP Protocol Enabled Mode

Premise: In order to facilitate a migration to full IPv6 implementation on the CPE device, as well as providing full IPv4 support.

In Dual-stack mode, the cable infrastructure SHALL be able to provide:

- Direct IPv6 connectivity, using an IPv6 GUA

AND

- IPv4 connectivity either via a 'public' IPv4 address, OR
- An 'interior' non-globally routable RFC 1918 [30] IPv4 address where CGN/NAT444 is used.

Functions: The CPE device SHALL implement both IPv4 and IPv6 provisioning in accordance with [eRouter]. IPv6 packets SHALL be forwarded between the WAN and LAN interfaces.

IPv4 packets SHALL be forwarded between the WAN and LAN interfaces using NAPT.

In this mode, any transitional technology (6rd, DS-Lite, MAP) functionality SHALL be disabled.

6.1.7 IPv6 Protocol Enabled Mode

Premise: Support for MAP or DS-Lite Transition strategies.

Functions: In order to implement MAP or DS-Lite transition technologies the CPE device SHALL operate in IPv6 Protocol Enabled mode.

Once the CPE device is configured to operate in IPv6 Protocol Enable Mode (or in Dual IP Protocol Enable mode) it SHALL obtain its IP address for the operator facing network interface using DHCPv6 as defined by RFC 3315 [56]. All other operational parameters required to implement transition strategies SHALL also be obtained via DHCPv6 as indicated by the following logical flow.

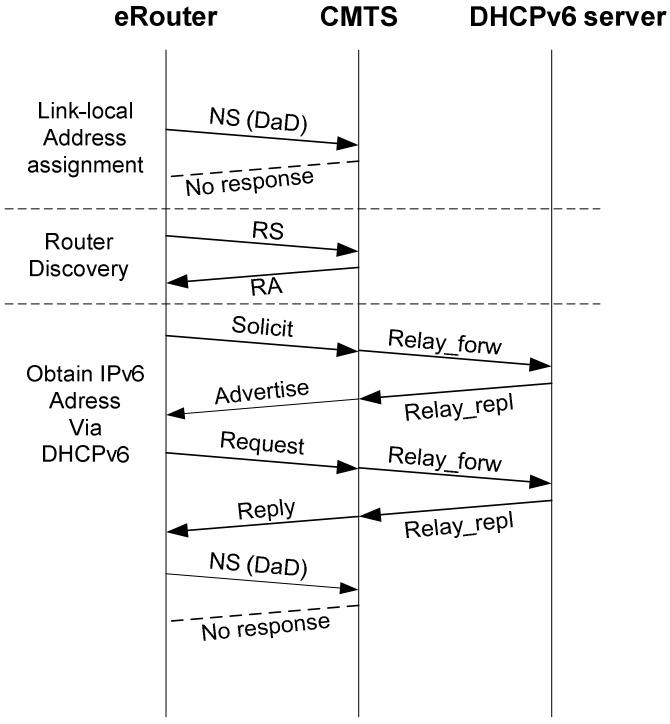


Figure 3: IPv6 Address Assignment Message Flow

6rd functionality SHALL NOT be enabled when the device is in IPv6ProtocolEnabled mode.

6.1.8 Transition Technology Provisioning

Function: In order to implement the correct operational mode the CPE device SHALL receive both the eSAFE operational parameter defined in [138], and the device SHALL also receive the operational parameter required to determine which transition strategy is to be implemented.

These transition strategies as defined by the present document SHALL include:

1) DS-Lite

2) 6rd

3) MAP-E / MAP-T

At this stage the required TLV's have not yet been standardized nor are they in draft. As such the CPE device SHALL receive vendor specific TLV statements to enable the correct operational state.

6.1.9 LAN Functionality

For the LAN interface, the device SHALL comply with the following RFCs:

Table 3: Operational Modes (table 6-1 from CM-SP-eRouter-I09-130404) [138]

RFC No.	RFC Name
RFC 2460 [40]	IPv6 specification
RFC 3315 [56]	Dynamic Host Configuration Protocol for IPv6
RFC 3319 [57]	DHCPv6 Options
RFC 3633 [66]	IPv6 prefix options for DHCPv6
RFC 4291 [78]	IPv6 addressing architecture
RFC 4443 [82]	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification
RFC 4861 [96]	Neighbor Discovery for IP Version 6 (IPv6)
RFC 4862 [97]	IPv6 Stateless Address Autoconfiguration
RFC 6092 [120]	simple security in IPv6 gateways
RFC 6106 [121]	IPv6 route advertisement option for DNS configuration
RFC 6204 [128]	Basic IPv6 Router functionality
RFC 6540 [132]	IPv6 Support Required for All IP-Capable Node

This is in addition to the presently supported IPv4 functions.

6.1.9.1 Address Assignment

For the LAN interface, the device SHALL assign itself a link-local address, and use Duplicate Address Detection (DAD) to ensure uniqueness.

In the case of an IPv6 WAN connection, a prefix will be assigned to the device via DHCPv6. The size of this prefix may be variable between a /48 and a /64.

The cable device SHALL send a DHCPv6 solicit message to the network DHCP server to request a prefix delegation to use for the LAN interfaces. The device SHALL assign a /64 to each LAN segment from the received prefix.

On the LAN side, the device SHALL be able to assign individual host IP addresses via:

- Stateful DHCPv6
- Stateless SLAAC

With the use of SLAAC the device SHALL support one of the following two options for providing DNS server IP addresses to clients:

- MAY support DHCPv6 rapid commit
- Route Advertisement RDNSS

The device SHALL also be able to support prefix sub-delegation via DHCPv6 (DHCP-PD). i.e. if it received a DHCPv6 solicit message for a prefix delegation, it SHALL respond with a suitable sub-prefix.

The device SHALL send periodic multicast router advertisement messages, to advertise itself to the LAN as the default gateway to each LAN segment. These will include the /64 prefix advertisement as assigned.

The device SHALL receive route solicit messages from LAN hosts, and reply accordingly.

6.1.9.2 DNS

The device MAY provide a caching / forwarding DNS server:

- If so, it SHALL advertise its own IP address in DHCPv6 replies and route advertisements.
- If not, then it SHALL advertise the received DNS server IP address(es) in DHCPv6 replies and route advertisements.

Any DNS functionality deployed on the device SHALL support A and AAAA records.

The device SHALL NOT have an Open Recursive Resolver enabled by default.

Note that the device DNS forwarding functionality will vary depending on the transition technology employed. The device SHALL forward DNS queries using the native IP stack employed.

6.1.10 Device Management

The device SHALL support SNMP management on its WAN interface. This is to include the following standards for IPv6.

Table 4: Operational Modes (table 6-1 from CM-SP-eRouter-I09-130404) [138]

RFC No.	RFC Name
RFC 2452 [38]	MIB for TCP/IPv6
RFC 2454 [39]	MIB for UDP/IPv6
RFC 2465 [45]	MIB for IPv6 Textual Conventions and General Group
RFC 2466 [46]	MIB for ICMPv6 Group

Note these are in addition to the presently supported IPv4 MIBs.

6.1.11 Security

In all modes except Disabled, the device SHALL enable a stateful firewall by default. In particular, the device SHOULD provide the home gateway with basic security requirements as specified in [120], and the device SHALL support ingress filtering in accordance with BCP 38 [48].

Devices providing functionality as described in [120] SHOULD implement such functionality according to Appendix I of [138].

The device MAY provide the ability for the user to enable/disable the security features.

6.2 General Requirements for Core Network Elements (AFTR/BR/GN)

6.2.1 Device Management Requirements

6.2.1.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1ab) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: [12**]**.

6.2.1.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [79**]**, [115].

6.2.1.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.2.1.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.2.1.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv6.

RFC Reference: [76**]**.

6.2.1.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18**]**.

6.2.1.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.2.1.8 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26**]**.

6.2.1.9 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.2.1.10 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60**]**.

6.2.1.11 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.2.1.12 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.2.1.13 Feature: Remote Access - TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.2.1.14 Feature: Remote Access - RADIUS, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Radius over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 RADIUS servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [49**]**, [50].

6.2.1.15 Feature: Remote access: Radius, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Radius over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 Radius servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [49**]**, [50].

6.2.2 General Specification Performance Requirements

6.2.2.1 Feature: Node Latency below 0,2 ms

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency should not exceed 0,2 ms.

Function: Protect the network against delay and jitter.

RFC Reference: N/A.

6.2.2.2 Feature: Forwarding performance

The premise, function and related RFC reference relating to this feature is as given below.

Premise: All forwarding and features should be implemented in hardware for both IPv4 and IPv6.

Function: Ensure efficient and performant device capabilities.

RFC Reference: N/A.

6.2.3 General Feature Requirements

6.2.3.1 Feature: Load Balancing ECMP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should be able to load-balance over at least 8 links or interfaces. Hashing should be based on L3 and L4 headers.

Function: Allow for simple adding of capacity in the network.

RFC Reference: N/A.

6.2.3.2 Feature: LAG/LACP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Link aggregation should be supported by the device with a minimum of 8 links. LACP support is optional.

Function: Allow for simple adding of capacity in the network.

RFC Reference: N/A.

6.2.3.3 Feature: Dot1Q interfaces

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support for L3 interfaces on Dot1Q interfaces.

Function: Allow Dot1Q interfaces.

RFC Reference: N/A.

6.2.3.4 Feature: Traffic monitoring

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support for mirroring bi-directional traffic from a physical IF to another physical IF. Support for mirroring bi-directional traffic from a vlan to a physical IF. Support for mirroring bi-directional traffic from a logical IF to a physical IF. Support for mirroring bi-directional traffic from a physical IF to a vlan or logical IF. Support for mirroring bi-directional traffic from a vlan to another vlan or logical IF. Support for mirroring bi-directional traffic from a logical IF to a vlan or logical IF. Filtering of traffic L2-based from the source, freely-configurable. Filtering of traffic L3-based from the source, freely-configurable. Filtering of traffic L4-based from the source, freely-configurable. Support for min. 2 parallel, individually configurable and independent mirroring-sessions. Support for.1q and q-in-q on the source. Support for.1q and q-in-q on the destination. Support for sending mirrored traffic to a remote destination, routed. Support for sending mirrored traffic to a remote destination, MPLS-based.

Function: Allowing network management people to debug the network and discover attacks or anomalies.

RFC Reference: N/A.

6.2.3.5 Feature: Netflow IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Flow-record generation over IPv4 should be supported with a maximum 1:1000 sampoling-rate for both ingress and egress direction. A minimum of 2 collectors should be configurable. Netflow v9 is mandatory. Netflow v10 (IPFIX) is optional.

Function: Protect the network against attacks and export NAT bindings if applicable.

RFC Reference: **[**73**]**, [100], [101].

6.2.3.6 Feature: Netflow IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Flow-record generation over IPv6 should be supported with a maximum 1:1000 sampoling-rate for both ingress and egress direction. A minimum of 2 collectors should be configurable. Netflow v9 is mandatory. Netflow v10 (IPFIX) is optional.

Function: Protect the network against attacks and export NAT bindings if applicable.

RFC Reference: **[**73**]**, [100], [101].

6.2.3.7 Feature: CoPP IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support for rate limiting (configurable down to zero) any traffic destined for the control plane of the device itself on a per protocol (higher layer protocols such as BGP or TFTP) and/or IP source address basis and/or the IP precedence for IPv4 packets. Filtering of IPv4 routing protocol traffic to allow only local link source addresses where appropriate.

Function: Protect the device against attacks and anomalies in the network.

RFC Reference: N/A.

6.2.3.8 Feature: CoPP IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support for rate limiting (configurable down to zero) any traffic destined for the control plane of the device itself on a per protocol (higher layer protocols such as BGP or TFTP) and/or IP source address basis and/or the IP precedence for IPv6 packets. Filtering of IPv6 routing protocol traffic to allow only local link source addresses where appropriate.

Function: Protect the device against attacks and anomalies in the network.

RFC Reference: N/A.

6.2.3.9 Feature: ACL support IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: All filters has to support all IPv4 traffic/prefix filtering. The device has to produce statistics for all filters. Any ACL/filtering feature has to not affect the packet forwarding performance of the interfaces they are applied to. It has to be possible to filter traffic at ingress and egress of all interface types including physical, LAGs and trunked VLAN interfaces, on trunked interfaces the filtering should be per sub-interface, per set of sub-interfaces or on all sub-interfaces in the trunk. Maximum length of filters supported on an interface a minimum of 1 024. Maximum number of filters on the device a minimum of 1 024 filters. Filtering match on higher layer protocols such as TCP, HTTP, etc. Filtering match on source and destination IP addresses or address ranges. Filtering match on presence or absence of the different TCP flags or combination of flags. Filtering match on IP precedence/DSCP/MPLS EXP bits and CoS. Filtering match on all IP protocol types. Filtering match on any wildcard bits combination.

Function: Protect the device against not-allowed traffic.

RFC Reference: N/A.

6.2.3.10 Feature: ACL support IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: All filters has to support all IPv6 traffic/prefix filtering. The device has to produce statistics for all filters. Any ACL/filtering feature has to not affect the packet forwarding performance of the interfaces they are applied to. It has to be possible to filter traffic at ingress and egress of all interface types including physical, LAGs and trunked VLAN interfaces, on trunked interfaces the filtering should be per sub-interface, per set of sub-interfaces or on all sub-interfaces in the trunk. Maximum length of filters supported on an interface a minimum of 1 024. Maximum number of filters on the device a minimum of 1 024 filters. Filtering match on higher layer protocols such as TCP, HTTP, etc. Filtering match on source and destination IP addresses or address ranges. Filtering match on presence or absence of the different TCP flags or combination of flags. Filtering match on IP precedence/DSCP/MPLS EXP bits and CoS. Filtering match on all IP protocol types. Implicit allow of protocol discovery packets, such as neighbor discovery. On an IPv6 interface an implicit deny of router advertisement packets.

Function: Protect the device against not-allowed traffic.

RFC Reference: N/A.

6.2.3.11 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the device can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the device NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.2.3.12 Feature: RP / RE redundancy - Non Stop Forwarding - Non Stop Routing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given RE's within the device has to be an ability within the node. This includes full syncing of all state held by the RE.

RFC Reference: N/A.

6.2.3.13 Feature: Jumbo Frames IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support IPv4 jumbo frames with a minimum size of 9 000 bytes.

Function: Allow for IPv4 jumbo frames to pass the device.

RFC Reference: N/A.

6.2.3.14 Feature: Jumbo Frames IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support IPv6 jumbo frames with a minimum size of 9 000 bytes.

Function: Allow for IPv6 jumbo frames to pass the device.

RFC Reference: N/A.

6.2.3.15 Feature: NDP, NUD, DAD

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor Discovery Protocol should be supported. Router Solicitation - Router Advertisement - Neighbor Solicitation - Neighbor Advertisement - Redirect.

Function: Router discovery: hosts can locate routers residing on attached links. Prefix discovery: hosts can discover address prefixes that are on-link for attached links. Parameter discovery: hosts can find link parameters (e.g. MTU). Address autoconfiguration: stateless configuration of addresses of network interfaces. Address resolution: mapping between IP addresses and link-layer addresses. Next-hop determination: hosts can find next-hop routers for a destination. Neighbor unreachability detection (NUD): determine that a neighbor is no longer reachable on the link. Duplicate address detection (DAD): nodes can check whether an address is already in use. Redirect: router can inform a node about better first-hop routers.

RFC Reference: [96**]**.

6.2.3.16 Feature: Ethernet OAM

The premise, function and related RFC reference relating to this feature is as given below.

Premise: "Link OAM - IEEE 802.3ah [5], clause 57, 802.3-2005, aka ""LFM"". Basic functionality Implementation has to support for the following: OAM Capability Discovery, Link Monitoring, Fault signalling, remote loopback control. Link OAM - support for L3 interfaces. Link OAM - Recovery - Automatic recovery mechanism has to be available. Link OAM - suppress. Service OAM (802.1ag/Y.1731) (aka ""CFM"") support for each configured service. Service OAM messages: Supported message types should be: continuity check, Loopback, Trace route. Service OAM Performance - Performance measurements (delay, delay variation, loss). Service OAM Fault - Fault notification and isolation (AIS/RDI). Service OAM Configuration - configurable domain, level and maintenance (END) point IDs.

Function: Ethernet OAM will contribute to fault management and network stability.

Standards Reference: [3**]**, [5], [4].

6.2.3.17 Feature: QoS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support a basic QoS feature-set. Classification at ingress should be possible on: dot1p bits, DSCP bits, MPLS EXP bits, IP header information, L4 header information. Marking at egress should be possible for: dot1p bits, DSCP bits, MPLS EXP bits. At least 8 queues per interface or subinterface should be supported.

Function: Ensuring SLA's for different types of traffic.

RFC Reference: N/A.

6.2.4 Routing Protocol Requirements

6.2.4.1 Feature: Static Routes, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Configurable, regardless of subnet or mask-length, at least supporting up to 500 per routing-context. Support for annunciations/descriptions/remarks. Configurable route-preference/distance, per route. Support for ECMP, at least 8 concurrent. Configurable route-metric, per route. Configurable to set discarding as next-hop. Support for conditional statics based availability.

Function: Static routing in the network for IPv4.

RFC Reference: N/A.

6.2.4.2 Feature: Static Routes, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Configurable, regardless of subnet or mask-length, at least supporting up to 500 per routing-context. Support for annunciations/descriptions/remarks. Configurable route-preference/distance, per route. Support for ECMP, at least 8 concurrent. Configurable route-metric, per route. Configurable to set discarding as next-hop. Support for conditional statics based availability.

Function: Static routing in the network for IPv6.

RFC Reference: N/A.

6.2.4.3 Feature: OSPFv2

The premise, function and related RFC reference relating to this feature is as given below.

Premise: OSPFv2 should be supported. OSPFV2 Refresh- and Flooding-reduction. OSPFV2 LSA-filtering to/from neighbors. OSPFV2 Redistribution to/from any protocol/static/connected, configurable via policy, support for multiple redistributions in parallel. OSPFV2 Opaque LSA support. OSPFV2 TE-extensions for MPLS and IPFRR. OSPFV2 Passive-mode. OSPFV2 Non-Stop Routing/Forwarding. OSPFV2 Gracefull Restart. OSPFV2 Multiple individual processes/instances supported, including internal redistribution, configurable via policy. OSPFV2 Startup forwarding-delay. OSPFV2 Authentication (per area/link, MD5+clear-text). Minimum 40k routes.

Function: OSPFv2 routing in the network for IPv4.

RFC Reference: [34**]**, [35], [64], [65].

6.2.4.4 Feature: OSPFv3

The premise, function and related RFC reference relating to this feature is as given below.

Premise: OSPFv3 should be supported. OSPFV3 Refresh- and Flooding-reduction. OSPFV3 LSA-filtering to/from neighbors. OSPFV3 Redistribution to/from any protocol/static/connected, configurable via policy, support for multiple redistributions in parallel. OSPFV3 Opaque LSA support. OSPFV3 TE-extensions for MPLS and IPFRR. OSPFV3 Passive-mode. OSPFV3 Non-Stop Routing/Forwarding. OSPFV3 Gracefull Restart. OSPFV3 Multiple individual processes/instances supported, including internal redistribution, configurable via policy. OSPFV3 Startup forwarding-delay. OSPFV3 Authentication (per area/link, MD5+clear-text). Minimum 40k routes.

Function: OSPFv3 routing in the network for IPv6.

RFC Reference: [87**]**, [105].

6.2.4.5 Feature: ISIS IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: ISIS for IPv4 should be supported. ISIS Multi-Topology IPv4 configurable. ISIS Redistribution to/from any protocol/static/connected IPv4, configurable via policy, support for multiple redistributions in parallel. ISIS extensions for TE. ISIS Restart-Signaling incl. 'Overload-bit-on-startup', configurable. ISIS Overload-bit freely configurable. ISIS Non-Stop Routing/Forwarding. ISIS Gracefull Restart. ISIS Authentication (MD5 + Key-Chain + Clear-Text). ISIS Multiple individual processes/instances supported IPv4, including internal redistribution, configurable via policy. ISIS Admin-tags IPv4. ISIS configurable hello-timers/-multipliers, exp. Back-off timers. ISIS configurable metrics with multi-topology per AFI, wide-metrics. ISIS extensions fro MPLS and IPFRR. ISIS passive-mode support (link). Minimum 40k routes.

Function: ISIS routing in the network for IPv4.

RFC Reference: [21**]**, [23], [52], [58], [63], [68], [69], [70], [71].

6.2.4.6 Feature: ISIS IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: ISIS for IPv6 should be supported. ISIS Multi-Topology IPv6 configurable. ISIS Redistribution to/from any protocol/static/connected IPv6, configurable via policy, support for multiple redistributions in parallel. ISIS extensions for TE. ISIS Restart-Signaling incl. 'Overload-bit-on-startup', configurable. ISIS Overload-bit freely configurable. ISIS Non-Stop Routing/Forwarding. ISIS Gracefull Restart. ISIS Authentication (MD5 + Key-Chain + Clear-Text). ISIS Multiple individual processes/instances supported IPv4, including internal redistribution, configurable via policy. ISIS Admin-tags IPv6. ISIS configurable hello-timers/-multipliers, exp. Back-off timers. ISIS configurable metrics with multi-topology per AFI, wide-metrics. ISIS extensions fro MPLS and IPFRR. ISIS passive-mode support (link). Minimum 40k routes.

Function: ISIS routing in the network for IPv6.

RFC Reference: [102**]**, [104], [122].

6.2.4.7 Feature: BGP IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: BGP with address-family IPv4 should be supported. BGP Route-Reflection IPv4. BGP 4-Octet AS support IPv4. BGP Route-Refresh Capability. "BGP Communities supporting normal and extended (all 3 types), well-knowns. BGP Router-ID configurable. BGP NSR/NSF support IPv4. Gracefull Restart/shutdown and negotiation. Peer- and/or session-grouping support. BGP Path-selection modifiers. EBGP Multihop-support. Remove-private-as support. Configurable Prefix-limit support on both session- or group-level. Configurable loop-prevention and migration-techniques like e.g. local-as, as-override. Configurable to allow announcement of inactive prefixes. Community-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination. AS-Path-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination, allowing regular-expressions. Configurable redistribution IPv4 from any static/IGP into or from BGP, simultaneous operation of multiple instances, protocols and address-families, incl. policy-based filtering/modification. MD5-Authentication support for both session- or group-level. Supporting up to 500 individual BGP-sessions. BGP Multipathing, configurable path-amount. Routing-Policy support IPv4. Highly configurable utilizing all known BGP-attributes in any form or combination, simultaneous, supporting configurable lists for AS-Path/Communities. Configurable on both session- or group-level. Useable with redistributions. Supporting PMTUD. Configurable MSS-size per session and/or group-level. Support for utilizing IGP-metric and minimum-IGP-metric into BGP-MED. BFD-Support, IPv4. Load-Balancing per-NH. IPv4. Labeled-unicast BGP support on both session- or group-level. Minimum 40 k routes.

Function: BGP routing in the network for IPv4.

RFC Reference: [25**]**, [28], [31], [36], [37], [51], [54], [59], [77], [85], [86], [90], [91], [98], [111].

6.2.4.8 Feature: BGP IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: BGP with address-family IPv6 should be supported. BGP Route-Reflection IPv6. BGP 4-Octet AS support IPv6. BGP Route-Refresh Capability. "BGP Communities supporting normal and extended (all 3 types), well-knowns. BGP Router-ID configurable. BGP NSR/NSF support IPv6. Gracefull Restart/shutdown and negotiation. Peer- and/or session-grouping support. BGP Path-selection modifiers. EBGP Multihop-support. Remove-private-as support. Configurable Prefix-limit support on both session- or group-level. Configurable loop-prevention and migration-techniques like e.g. local-as, as-override. Configurable to allow announcement of inactive prefixes. Community-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination. AS-Path-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination, allowing regular-expressions. Configurable redistribution IPv6 from any static/IGP into or from BGP, simultaneous operation of multiple instances, protocols and address-families, incl. policy-based filtering/modification. MD5-Authentication support for both session- or group-level. Supporting up to 500 individual BGP-sessions. BGP Multipathing, configurable path-amount. Routing-Policy support IPv6. Highly configurable utilizing all known BGP-attributes in any form or combination, simultaneous, supporting configurable lists for AS-Path/Communities. Configurable on both session- or group-level. Useable with redistributions. Supporting PMTUD. Configurable MSS-size per session and/or group-level. Support for utilizing IGP-metric and minimum-IGP-metric into BGP-MED. BFD-Support, IPv6. Load-Balancing per-NH. IPv6. Transporting Ipv6 over an MPLS network using 6PE should be supported. Minimum 40 k routes.

Function: BGP routing in the network for IPv6.

RFC Reference: [25**]**, [28], [31], [36], [37], [51], [54], [59], [77], [80], [85], [86], [90], [91], [95], [98], [111].

6.2.4.9 Feature: MP-BGP IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: MP-BGP with address-family VPN- IPv4 should be supported. BGP Route-Reflection IPv4. BGP 4-Octet AS support IPv4. BGP Route-Refresh Capability. "BGP Communities supporting normal and extended (all 3 types), well-knowns. BGP Router-ID configurable. BGP NSR/NSF support IPv4. Gracefull Restart/shutdown and negotiation. Peer- and/or session-grouping support. BGP Path-selection modifiers. EBGP Multihop-support. Remove-private-as support. Configurable Prefix-limit support on both session- or group-level. Configurable loop-prevention and migration-techniques like e.g. local-as, as-override. Configurable to allow announcement of inactive prefixes. Community-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination. AS-Path-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination, allowing regular-expressions. Configurable redistribution IPv4 from any static/IGP into or from BGP, simultaneous operation of multiple instances, protocols and address-families, incl. policy-based filtering/modification. MD5-Authentication support for both session- or group-level. Supporting up to 500 individual BGP-sessions. BGP Multipathing, configurable path-amount. Routing-Policy support IPv4. Highly configurable utilizing all known BGP-attributes in any form or combination, simultaneous, supporting configurable lists for AS-Path/Communities. Configurable on both session- or group-level. Useable with redistributions. Supporting PMTUD. Configurable MSS-size per session and/or group-level. Support for utilizing IGP-metric and minimum-IGP-metric into BGP-MED. BFD-Support, IPv4. Load-Balancing per-NH. IPv4. Labeled-unicast BGP support on both session- or group-level. Minimum 40 k routes.

Function: Enabling IPv4 MPLS-VPN's.

RFC Reference: [25**]**, [28], [31], [36], [37], [51], [54], [59], [77], [80], [81], [85], [86], [90], [92], [98], [111].

6.2.4.10 Feature: MP-BGP IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: MP-BGP with address-family VPN-IPv6 should be supported. BGP Route-Reflection IPv6. BGP 4-Octet AS support IPv6. BGP Route-Refresh Capability. "BGP Communities supporting normal and extended (all 3 types), well-knowns. BGP Router-ID configurable. BGP NSR/NSF support IPv6. Gracefull Restart/shutdown and negotiation. Peer- and/or session-grouping support. BGP Path-selection modifiers. EBGP Multihop-support. Remove-private-as support. Configurable Prefix-limit support on both session- or group-level. Configurable loop-prevention and migration-techniques like e.g. local-as, as-override. Configurable to allow announcement of inactive prefixes. Community-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination. AS-Path-list(s) support allowing 'logical-AND' and 'logical-OR' operation as well as a combination, allowing regular-expressions. Configurable redistribution IPv6 from any static/IGP into or from BGP, simultaneous operation of multiple instances, protocols and address-families, incl. policy-based filtering/modification. MD5-Authentication support for both session- or group-level. Supporting up to 500 individual BGP-sessions. BGP Multipathing, configurable path-amount. Routing-Policy support IPv6. Highly configurable utilizing all known BGP-attributes in any form or combination, simultaneous, supporting configurable lists for AS-Path/Communities. Configurable on both session- or group-level. Useable with redistributions. Supporting PMTUD. Configurable MSS-size per session and/or group-level. Support for utilizing IGP-metric and minimum-IGP-metric into BGP-MED. BFD-Support, IPv6. Load-Balancing per-NH. IPv6. Transporting Ipv6 over an MPLS network using 6PE should be supported. Minimum 40 k routes.

Function: Enabling IPv6 MPLS-VPN's.

RFC Reference: **[**25**]**, [28], [31], [36], [37], [51], [54], [59], [77], [80], [81], [85], [86], [89], [90], [91], [92], [98], [111].

6.2.4.11 Feature: BFD IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Bi-directional Forwarding Detection, IPv4. Bi-directional Forwarding Detection, Multihop-Paths, IPv4. Authentication-support, configurable, MD5/Text-based. ISIS BFD-Enabled TLV. Support for.1q IF's. Support for static routes. Support for LAG IF's. Support for OSPFv2, Support for OSPFv3, Support for ISIS, Support for VRRP, Support for BGP, Support for multiple protocols sharing the same BFD-session.

Function: Ensuring fast link failure detection.

RFC Reference: **[**112**]**, [113], [114].

6.2.4.12 Feature: BFD IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Bi-directional Forwarding Detection, IPv6. Bi-directional Forwarding Detection, Multihop-Paths, IPv6. Authentication-support, configurable, MD5/Text-based. ISIS BFD-Enabled TLV. Support for.1q IF's. Support for static routes. Support for LAG IF's. Support for OSPFv2, Support for OSPFv3, Support for ISIS Support for VRRP, Support for BGP, Support for multiple protocols sharing the same BFD-session.

Function: Ensuring fast link failure detection.

RFC Reference: **[**112**]**, [113], [114].

6.2.4.13 Feature: Policy Based Routing IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support policy based routing for IPv4 traffic. This includes: setting a next-hop, setting an outgoing interface, indirect redirection.

Function: Enabling policy based routing for IPv4.

RFC Reference: N/A.

6.2.4.14 Feature: Policy Based Routing IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support policy based routing for IPv6 traffic. This includes: setting a next-hop, setting an outgoing interface, indirect redirection.

Function: Enabling policy based routing for IPv6.

RFC Reference: N/A.

6.2.5 MPLS Requirements

6.2.5.1 Feature: MPLS LDP IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device should support LDP for IPv4. Support for ECMP-Paths. Support for Gracefull Restart. Support for Non Stop Routing. Support for Non Stop Forwarding. ISIS/OSPF IGP-Sync support. In/Egress Filtering-support, freely configurable, policy-based (ACL's/PFX-lists). In/Egress Filtering-support, freely configurable, Interface-based. Support for capability-messages/negotiation.

Function: Allow MPLS transport to and from the device.

RFC Reference: [53**], [**61**], [**99**], [**107**]**.

6.2.6 Service Requirements

6.2.6.1 Feature: IP-VPN's IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv4 IP-VPNs should be supported with a minimum of 128 instances. OSPFv2 and BGP should be supported as PE-CE protocol.

Function: Enable the use of IPv4 IP-VPNs.

RFC Reference: [81**]**, [88].

6.2.6.2 Feature: IP-VPN's IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv6 IP-VPN's should be supported with a minimum of 128 instances. OSPFv3 and BGP should be supported as PE-CE protocol.

Function: Enable the use of IPv4 IP-VPN's.

RFC Reference: [81**]**, [88], [89].

6.2.6.3 Feature: Point-to-Point L2VPN

The premise, function and related RFC reference relating to this feature is as given below.

Premise: L2 VPN- LDP based Ethernet Point-2point tunnels - Draft Martini Should include support for: signalling (e.g. PW status), VC Type 4, VC Type 5. Single and double-tagged frames should be supported.

Function: Enable the use of point-to-point L2VPN's.

RFC Reference: [74**]**, [83], [84].

6.2.6.4 Feature: VPLS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: LDP-based VPLS should be supported.

Function: Enable the use of VPLS.

RFC Reference: [93**]**.

6.2.6.5 Feature: 6PE

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support IPv6 transport over IPv4 MPLS.

Function: The tunnelling of IPv6 packets across an IPv4 CORE using 6PE.

RFC Reference: [95**]**.

6.2.6.6 Feature: 6VPE

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Support IPv6 transport over IPv4 MPLS in separate instances for VPNs.

Function: The tunnelling of IPv6 packets across an IPv4 CORE using 6VPE VRFs.

RFC Reference: [89**]**.

6.3 DS-Lite

DS-Lite uses IPv6-only links between the provider and the customer to carry IPv4 privately addressed packets. The DS-Lite home gateway (CPE or B4 element) is provisioned with only an IPv6 address on its WAN interface. At the LAN-side interface, the CPE operates its own DHCP server handing out RFC1918 private IPv4 addresses to home devices. The CPE does not perform NAT; the NAT function is located on a carrier-grade NAT device in the provider's network, which is also a tunnel terminator for the IPv4-in-IPv6 tunnel. This device is called AFTR.

The IPv4 packet from the home device to an external destination is encapsulated in an IPv6 packet by the CPE and handed to the provider network. The packet is decapsulated at the AFTR, and NAT44 is performed to map the host's private IPv4 address to a public IPv4 address. The IPv6 tunnel source address is added to the NAT table, along with an IPv4 source address and port in order to both disambiguate the customer private address and provide the reference for the tunnel endpoint. If a home device needs to access an IPv6 service, packets are transported "as-is" and routed to an Internet server. With DS-Lite technology, the communications between endpoints remain within their address family without requiring protocol family translation.

NAT services allow service providers to conserve IPv4 addresses and maintain IPv4 Internet access while migrating to IPv6. If implemented in components of the Core Network, an optimized software implementation provides scale, improved transaction rates and complete logging and accounting.

6.3.1 DS-Lite Technology Feature Summary

The DS-Lite technology feature is summarized in table 5 detailing for each function the requirement as required or optional with a brief description of each of the named functions.

Table 5: DS-Lite Technology Feature, Function, Requirement and Description

Functional Name	Requirement	Description
RFC 6333 [129]	Required	Compliance with RFC 6333 [129]. (Except the support for fragmented IPV6 packets from the B4 element - section 6.3 of RFC 6333 [129]) except for specific stated alternatives such as fragmentation
RFC 6334 [130]	Required	Compliance with RFC 6334 [130] - DHCPv6 option for DS-Lite
Redundancy	Required	All critical components has to be redundant in such a manner that they can fail-over without impact to customer or management traffic greater than 1 ms.
Shared GW IP	Required	Single AFTR GW address. The AFTR IPv6 address should be able to be shared amongst different NPU's in the AFTR. A hashing mechanism should be in place to hash all upstream packets based on the source IPv6 address
SI-ID Basic	Optional	The softwire ID for each customer SHOULD
Si-ID Extension	Required	Tunnel/customer identifier SHOULD be based on both IPv6 CPE public and IPv4 Client Private addresses. This requirement is required for **PCP Multisession dynamic forwarding**
AFTR Addressing and Virtual Interfaces	Required	AFTR has to be able to assign a single virtual interface with up to 8 DS-Lite GW addresses for any given DS-Lite instance on the node
Anycast	Required	Anycast AFTR gateway addresses are a requirement to allow simplicity of deployment for a single prefix across multiple AFTR's.
AFTR Address withdrawal	Required	The AFTR should have at least five points of AFTR GW address withdrawal occurrence. The list includes: - loss of route out, - loss of all BGP/IGP sessions, - loss of forwarding, - loss of NPU capacity and certain errors in the NAT caching. Any of the failures should be detectable based on configurable timers with 15 seconds being the default setting.
Physical Interface Load Balancing - ECMP	Required	Functional for both egress downstream and egress upstream traffic. Note it SHOULD be IPv4 encapsulated functional, to simplify, the downstream egress DS-Lite traffic should be balanced on a per flow basis based on the IPv4 address and IPv6 address within the IPv6 encap packet. So each flow balanced is based on IPv4 private destination flows and B4 address to prevent out of ordering. Upstream packets should be based on the IPv4 RFC as per usual.
NPU Load Balancing (hashing)	Required	DS-Lite Inside to Outside hashing performed on the Source IPv6 (128 bits) address of the B4 device. DS-Lite Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Tunnel MTU	Required	The Maximum Transmission Unit within the DS-Lite tunnel, has to be configurable. Expected value will be 1 460 bytes.
MSS Clamping	Required	TCP MSS support is mandatory for the AFTR due to the removal of an end-to-end MTU sizing functionality. This will avoid the need for excessive fragmentation
Fragmentation	Required	Fragmentation should be done on the IPv4 packet. Fragmentation should be placed on the ASIC running in the line card and SHOULD be used in all case unless the DF bit is set (see below)
RFC 6333 [129] Fragmentation	Required	Fragmentation SHOULD be done on the IPv6 packet, ONLY as specified in RFC 6333 [129] and if the DF bit is set, otherwise IPv4 fragmentation SHOULD be preferred. Used in conjunction with Stateful ICMP

Functional Name	Requirement	Description
IANA well known address for IPv4-in-IPv6 tunnel	Required	Use of IANA well known address for configuring IPv4-inIPv6 tunnel. By default the AFTR will assume 192.0.0.1. Customer can configure other value if needed. This can be used for v4 ICMP messaging between B4 and AFTR.
NAT - Network Address and Port Mapping - Endpoint Independent Mapping	Required	For two flows for a common inside source IPv4 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation
NAT - Translation Filtering - Endpoint Independent Filtering	Required	A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.
NAT - Paired IP Address Assignment	Required	Translation to External IPv4 address is done in a paired fashion. A given Inside IPv6 B4 address is always translated to the same External IPv4 address.
NAT - Hair-pinning	Required	Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.
NAT - 1:1 IP Mapping	Required	Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).
NAT - Outside-Service-App mapping for inside-VRF	Required	Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.
NAT - Port Limit configuration	Required	A maximum amount of ports can be configured for every IPv6 source B4 address
NAT - Per-Protocol Timeout configuration	Required	Timeout of mappings is critical and removing a NAT mapping at the appropriate time maximizes the shared port resource.
NAT - Dynamic Port Range start configuration	Required	The start port for dynamic port ranges should be configurable to allow for a range of ports for static port mappings
NAT - Software Load Balancing	Required	NAT Inside to Outside hashing performed on the Source IPv6 B4 Address. NAT44 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Shared Resource per Chassis Regardless of the number of NPUs	Required	All DS-Lite resources should be shared under a single AFTR address, shared under multiple AFTR addresses on the same virtual interface, a single shared IPv4 public range per chassis, shared NAT groups on a single AFTR address, shared NPUs, shared routing resources, memory, shared PCP cache and assignments across multiple NAT groups or a single NAT group.
Port Allocation	Required	In order to reduce the volume of data generated by the NAT device (logging creation and deletion data), bulk port allocation can be enabled. When bulk port allocation is enabled and when a subscriber creates the first session, a number of contiguous outside ports are pre-allocated. A bulk allocation message is logged indicating this allocation.
Deterministic NAT	Required	Algorithmically maps a customer's private IPv4 address to a set of public IPv4 address ports, allowing a significant reduction in logging.
Dynamic Port Assignment Oversubscription (based on threshold)	Optional	Creation of a percentage out of the 60 k block allocation per IPv4 public address. For example the first 100 ports is a primary assignment for each and every B4 within a single IP through DeNAT. The B4 then can get as many, max 3 000, ports as needed over-subscribing until a 80 % (configurable) port assignment is reached per IP, and then a restriction is placed with no B4 being able to request any more ports over the 500 ports mark. Sessions however are not dropped. Once the utilization drops to 70 % once again (configurable) the B4s are free then to oversubscribe once again.

Functional Name	Requirement	Description
Source Filtering per AFTR Address into Separate NAT Groups		NAT grouping should be configurable and able to delineate by IPv6 source prefix. So a single AFTR address (anycast) with the group rules and characteristics for multiple groups based on source IPv6 prefix. This allows separation of B2B, B2C, platform and services separation without using a different AFTR address
Port Reservation	Required	To prevent dynamic or static allocation (within deterministic NAT) of the first 5 000 port blocks, for assignment of with well known ports within static routing
PCP General Implementation	Required	Both PCP RFC draft version 12/13 and 29 should be supported and configurable
PCP Static	Required	Support for PCP to allocate static port bindings
PCP Dynamic	Required	Support for PCP to allocate dynamically assigned port and IP bindings
PCP IP and Port Reservation for encrypted / fixed headers	Required	To allow configuration of IPs and ports to PCP's allocation assignment for both static (portal based customer configured static port forwarding) and PCP UPnP and PMP dynamically assigned port and IP ranges for specific applications that are encrypted or fixed headers
PCP (SI-ID Extension Base) Multi-session Dynamic Forwarding	Required	Allows multiple sessions from multiple PCs behind a single B4 to access the same port by assigning as separate IP to the before (IPv4 public). The ability to dynamically allocate a static port forwards for a different outside IP as the subscriber got assigned when the particular port is in use. Reasoning is bittorrent clients with UPnP and PCP to the AFTR will not be able to negotiate the same port for multiple sessions from different PCs and given all clients use the same "fixed" port by default it is better to allocate a different outside IP
PCP Failure	Required	Application of PCP failover or failure response to prevent unnoticed PCP failure
FTP ALG (Active and Passive)	Required	FTP clients are supported with inside (private) address and servers with outside (public) addresses. Passive FTP is provided by the basic NAT function. Active FTP is used with the ALG
RTSP ALG (ALP)	Required	Remote control protocol for streamers (which use RTP/RTCP or RDT). Our implementation considers the server is located "outside" and clients are "inside". RTSP is used in many streamers like QuickTime or RealNetworks. Both "SETUP" and "SETUP response" are analysed. Based on this information, the translation entry will be created in the table.
SIP ALG (ALP)	Required	SIP requires control and signalling on separate port to the data traffic and therefore requires an ALG to function
PPTP ALG	Required	PPTP, used in many VPN setups, requires control and signalling on separate port to the data traffic and therefore requires an ALG to function
TFTP ALG	Optional	TFTP requires control and signalling on separate port to the data traffic and therefore requires an ALG to function
Use of Application SVI's, connecting multiple routing entities (inside/private VRF, outside/public VRF)	Required	NAT should be possible from any routing context (VRF, GRT) to any routing context (VRF, GRT)
Stateful ICMP	Required	Stateful ICMP mappings between inside and outside ICMP identifiers should be supported with 3 retries on any specific destination host upstream and downstream (based on fragmented packets receive on upstream from the CPE in case the CPE itself does not support stateful ICMP) This is prevent flooding of ICMP messages.

Functional Name	Requirement	Description
Thresholds	Required	Configurable thresholds using watermarks should be supported to monitor the resources on the AFTR. Port percentage per IP IP Address percentage per range assigned and per NAT group Resource per NPU Memory utilization Processor utilization per NPU PCP Percentage IP assignments Clustering utilization (the total utilization if one of the pair of the cluster goes down) Interface utilization BW DS-Lite sessions (dynamic thresholds) SI-ID (extension) numbers per NPU and chassis (based on both private IPv4 and public IPv4 used out of the range assigned)
QoS translation	Required	For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.
Chassis NAT Clustering	Required	Clustering of AFTRs to allow for inter-chassis resiliency under certain configurable parameters
DS-Lite Specific ACLs	Required	Deep packet inspect on the internal IPv4 packet (within the tunnel) before it reaches the SI interface, on the incoming physical interface
Both Active-Active or Active Standby Redundant NPUs	Required	Failover of n+1 or +2
per IPv4 user port limiting	Required (with SI-ID extension)	The ability to assign port blocks per private IPv4 address (SI-ID extension)
SNMPv1, V2 and V3 (c)	Required	DS-Lite specific SNMP MIB compatibility and function to cover all aspects of the DS-Lite feature sets, including per/sec flows, users
Logging via Netflow V9/IPFIX	Required	Netflow v9 support for logging of the translation records. Logging of the translation records can be mandated for Data Retention. The Netflow uses binary format and hence requires software to parse and present the translation records. 1:1 flow ratio is required. This is a backup, radius being the main form of logging, in case countries require further logging requirements than block assignment and IP. Note that the IP-FIX or Netflow template SHOULD be both IPv4 and IPv6 and cannot use separate templates per protocol but one specifically based on DS-Lite function.
Logging via Syslog	Required	Syslog support for logging of the translation records as an alternative to Netflow. Syslog uses ASCII format and hence can be read by users. However, the log data volume is higher in Syslog than Netflow.
Destination based Logging	Required	Destination Based Logging will generate one record per session, including the destination port and address. NFv9 templates including destination IP address and destination port will be used.
Base Logging Fields	Required	AFTR logs the following information when a translation entry is created: Inside instance ID Outside instance ID Inside IPv6 Address Inside Port Outside IPv4 Address Outside Port Protocol Start Time Stop Time
Radius Logging	Required	Logging using Radius accounting messages
XML	Optional	Logging using XML files

Functional Name	Requirement	Description
Static port forwarding (up to 6K static forward entries per npu)	Required	Static port forwarding configures a fixed, private (internal) IPv6 address and port that are associated with a particular subscriber while the AFTR allocates a free public IP address and port. Therefore, the inside IPv6 address and port are associated to a free outside IP address and port.
Static port forwarding 1:1 active/standby	Required	Static Port Forwarding mapping will be kept constant with two npu cards in Active/Standby mode
Performance Requirement		Description
24 DS-Lite instances per npu Card	Required	The ability to stack transition instances on top of one another
40m Translations (per npu) - block assignment	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary
Minimum 50 Gbps bi-directional throughput per npu (Inside-to-outside VRF + Outside-to-inside VRF - GRT to GRT) with IMIX traffic	Required	Per half slot throughput requirement
1M+ primary connections per second setup rate	Required	Primary sessions, port block assignments, per NPU
500 k users per NPU	Required	Per Network Processing Unit a minimum of 500 000 users should be serviceable at any single moment in time
Latency	Required	Latency is between 40 and 200 micro seconds (µs)
6 npu Cards per chassis	Required	Min chassis requirement
IRB/SVI support	Required	Integrated Routing and Bridging/Virtual Interfaces (L3 interface for Bridge Domain)
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 256 k NAT users at the same time.
NOTE: NPU is a transitional processing unit, half slot.		

6.3.2 Main DS-Lite RFC References

Historical documents: RFC 2473 [47], RFC 5382 [106], RFC 5508 [108].

Core protocol: RFC 6333 [129].

Deployment considerations: RFC 6908 [136].

Transition mechanism: RFC 6180 [127].

Translation framework: RFC 6144 [123].

IPv6 transition scenarios: RFC 6036 [118].

DHCPv6 and DS-Lite: RFC 6334 [130].

GI-DS-Lite: RFC 6674 [133].

DS-Lite RADIUS: RFC 6519 [131].

Deterministic NAT: [9].

DS-Lite PCP: [11].

PCP Type 1/Type 2 Encapsulation: RFC 6877 [134].

6.3.3 DS-Lite AFTR CORE Device (LSN/CGN) Specification and Requirements

AFTR is an implementation of an IPv4/IPv6 transition protocol based on Dual-Stack Lite, and is the CORE node that all CPE/CEs running DS-Lite and creating softwires for the technology connect to. It is the termination for the IPv6 tunnel and the originating IPv4 NAT device egressing traffic towards the destination on IPv4 after decapsulation.

This clause concerns itself solely with the DS-Lite specific required functionality to allow the technology to function completely and fully. The specification laid out below allows DS-Lite technology to be deployed on the AFTR as a non-service deprecating form comparative to a native private IPv4 delivery to any given customer within a Carrier Topology or potentially otherwise. The present document is however focuses on Cable requirements in functionality but can stretch further if required.

Note that any given AFTR implementation should be complaint with the associative feature for deployment within a Cable ISP network.

6.3.4 Hardware/Software Requirements

6.3.4.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU within a chassis in a full stateful fashion with no loss of packets or latency beyond 1 ms and constant forwarding. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1 ms failover maximum for any given B4 connectivity of session

RFC Reference: N/A.

6.3.4.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The AFTR has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the AFTR can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the DS-Lite NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.3.4.3 Feature: RP/RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or DS-Lite transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within an AFTR has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.3.4.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given AFTR has to be shared giving the appearance of a single node on the network.

Function: The AFTR has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single address per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.3.4.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given AFTR has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given AFTR to see all or part of the traffic coming into the node. This can be determined through a single AFTR address, so AFTR GW assigned or NAT group assigned under multiple GW addresses. Hashing SHOULD be used to determine the load balancing between all processing/hardware forwarding points with that NAT group or deterministic specification.

RFC Reference: N/A.

6.3.5 Performance Requirements

6.3.5.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1 ms for any given function of the AFTR.

Function: Within the remit of traffic requiring DS-Lite function, decapsulation, encapsulation and NAT function, the AFTR has to perform its function from ingress interface to egress interface in a time measured no higher than 200 micro seconds under 50 % load. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions which require less then 1 ms node latency minimum, 200 microseconds preferred.

RFC Reference: N/A.

6.3.5.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.3.5.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the AFTR Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 gig is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.3.5.4 Feature: Min Sessions per AFTR Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis.

RFC Reference: N/A.

6.3.5.5 Feature: Minimum Customer IPv4 Addresses per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Defines the minimum amount of IPv4 addresses able to be configured per chassis.

Function: If we define the IPv4 address space requirement based on session and block allocation this would be 60 ports per subscriber as a block assignment comparative to 100 k customers online allowing the integer of 1 024. Or if we calculate backwards this would be 100 000 customers divided by 60 per IP (although this changes per deterministic NAT requirement) allowing for a max utilization of 1 000 ports per subscriber and thus giving us 1 666-address requirement if we have maximum assignment utilization. So we can round this up to 2 040-address minimum address pool requirement.

RFC Reference: N/A.

6.3.5.6 Feature: Min Customer B4s per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus IPv6 B4 addresses that can be assigned from a single Chassis or connected

Function: Due to the scaling of any given AFTR solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 B4 addresses. This is solely a minimum and the number may change with the introduction of Private v4 and public v6 Si-ID assignment.

RFC Reference: N/A.

6.3.6 Feature Specifications

6.3.6.1 Feature: Tunnel Identifiers/Client-Customer ID

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For general identification and data retention purposes, tunnel identifiers has to be uniquely associated with a single CPE. The identifier has to be the IPv6 address of the CPE.

SI-ID based on both IPv6 CPE public and IPv4 Client Private.

Function: Unique Si-ID Ipv6 B4 addresses has to be assigned to the CPE, if the DHCPv6 originates from the AFTR or a separate DHCPV6 server, although the SI-ID can be determined with a join between the B4 IPv6 address and the IPv4 private address assigned within the customer Local Network. This allows for specific port block assignments if there are more than 5 clients behind a specific B4 CPE allowing for a greater amount of port allocation.

RFC Reference: [129**]**.

From section 4.2 in [129]:

> "A DS-Lite CPE is an IPv6-aware CPE with a B4 interface implemented in the WAN interface. A DS-Lite CPE SHOULD NOT operate a NAT function between an internal interface and a B4 interface, as the NAT function will be performed by the AFTR in the service provider's network. This will avoid accidentally operating in a double-NAT environment.
> However, it SHOULD operate its own DHCP(v4) server handing out [RFC1918] address space (e.g., 192.168.0.0/16) to hosts in the home. It SHOULD advertise itself as the default IPv4 router to those home hosts. It SHOULD also advertise itself as a DS server in the DHCP Option 6 (DNS Server). Additionally, it SHOULD operate a DNS proxy to accept DNS IPv4 requests from home hosts and send them using IPv6 to the service provider DNS servers, as described in Section 5.5.
> Note: If an IPv4 home host decides to use another IPv4 DNS server, the DS-Lite CPE will forward those DNS requests via the B4 interface, the same way it forwards any regular IPv4 packets. However, each DNS request will create a binding in the AFTR. A large number of DNS requests may have a direct impact on the AFTR's NAT table utilization.
> IPv6-capable devices directly reach the IPv6 Internet. Packets simply follow IPv6 routing, they do not go through the tunnel, and they are not subject to any translation. It is expected that most IPv6-capable devices will also be IPv4 capable and will simply be configured with an IPv4 [RFC1918]-style address within the home network and access the IPv4 Internet the same way as the legacy IPv4- only devices within the home. Pure IPv6-only devices (i.e., devices that do not include an IPv4 stack) are outside of the scope of this document."

6.3.6.2 Feature: ICMPv6: Neighbour Discovery and Stateless Auto-Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the discovery of an IPv6 neighbouring device and the ability for it to assign its own address.

Function: ICMPv6 based NDP and stateless auto-configuration has to be an option for the AFTR to allow for connectivity within NDP for on-link nodes.

RFC Reference: [40**]**, [41], [42], [43], [44].

From section 4 in [44]:

> "The Interface Identifier [AARCH] for an Ethernet interface is based on the EUI-64 identifier [EUI64] derived from the interface's built-in 48-bit IEEE 802 address. The EUI-64 is formed as follows (Canonical bit order is assumed throughout.)
> The OUI of the Ethernet address (the first three octets) becomes the company_id of the EUI-64 (the first three octets). The fourth and fifth octets of the EUI are set to the fixed value FFFE hexadecimal.

The last three octets of the Ethernet address become the last three octets of the EUI-64.
The Interface Identifier is then formed from the EUI-64 by complementing the "Universal/Local" (U/L) bit, which is the next-to-lowest order bit of the first octet of the EUI-64. Complementing this bit will generally change a 0 value to a 1, since an interface's built-in address is expected to be from a universally administered address space and hence have a globally unique value. A universally administered IEEE 802 address or an EUI-64 is signified by a 0 in the U/L bit position, while a globally unique IPv6 Interface Identifier is signified by a 1 in the corresponding position. For further discussion on this point, see [AARCH].
For example, the Interface Identifier for an Ethernet interface whose built-in address is, in hexadecimal, 34-56-78-9A-BC-DE would be 36-56-78-FF-FE-9A-BC-DE.
A different MAC address set manually or by software should not be used to derive the Interface Identifier. If such a MAC address has to be used, its global uniqueness property should be reflected in the value of the U/L bit.
An IPv6 address prefix used for stateless autoconfiguration [ACONF] of an Ethernet interface has to have a length of 64 bits."

6.3.6.3 Feature: IPv6 Global Unicast Address Format

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The AFTR has to be able to forward and be configured with an IPv6 Unicast structured address.

Function: In any given IPv6 topology unicast address structures has to be adhered to with no proprietary configurations. All present IPv6 unicast structures mentioned within 3 513 has to be compatible within the vendor deployment to allow for standardization and monitoring.

RFC Reference: [62**]**.

From section 2.5 in [62]:

"IPv6 unicast addresses are aggregable with prefixes of arbitrary bit-length similar to IPv4 addresses under Classless Interdomain Routing.
There are several types of unicast addresses in IPv6, in particular global unicast, site-local unicast, and link-local unicast. There are also some special-purpose subtypes of global unicast, such as IPv6 addresses with embedded IPv4 addresses or encoded NSAP addresses. Additional address types or subtypes can be defined in the future.
IPv6 nodes may have considerable or little knowledge of the internal structure of the IPv6 address, depending on the role the node plays (for instance, host versus router). At a minimum, a node may consider that unicast addresses (including its own) have no internal structure:

```
|                     128 bits                |
+-----------------------------------------------+
|                  node address                 |
+-----------------------------------------------+
```

A slightly sophisticated host (but still rather simple) may additionally be aware of subnet prefix(es) for the link(s) it is attached to, where different addresses may have different values for n:

```
|           n bits                          |  128-n bits    |
+-------------------------------------------+----------------+
```

```
|         subnet prefix                       |  interface ID  |
+---------------------------------------------+----------------+
```

Though a very simple router may have no knowledge of the internal structure of IPv6 unicast addresses, routers will more generally have knowledge of one or more of the hierarchical boundaries for the operation of routing protocols. The known boundaries will differ from router to router, depending on what positions the router holds in the routing hierarchy."

6.3.6.4 Feature: DS-Lite timers

The premise, function and related RFC reference relating to this feature is as given below.

Premise: These has to allow for all below DS-Lite effecting timers to be configured per the RFC.

Function: Listed below are the timers that SHOULD be configured for the AFTR, note that these SHOULD match the CPE where possible. The timers are not fixed as it is deployment dependent the list you below are the suggested values only and the required fields for any particular vendor:

1) icmp-query --- min 1
2) sip --- min 2
3) tcp-established --- hrs 30 min 0
4) tcp-syn --- sec 30
5) no tcp-time-wait --- min 3
6) tcp-transitory --- min 4
7) udp --- min 5
8) udp-initial --- sec 15
9) udp-dns --- sec 15

RFC Reference: No single defined RFC.

6.3.6.5 Feature: Thresholds and Watermarks

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are two types of thresholds/watermarks, reactive based and logging based, both are required.

Function: Listed below are minimum required thresholds for DS-Lite and SHOULD be adhere to validate the utilization on an event basis due to resource allocation that is required for CGN:

1) Watermarks based on Subscriber Thresholds for the node
2) Watermarks based NPU total resources
3) Watermarks based on Port Allocation Thresholds per Subscriber
4) Watermarks based on Port Allocation Thresholds per NPU
5) Watermarks based on IP Allocation Utilization Thresholds
6) Watermarking Event Reporting through SNMP
7) Watermarking Event Reporting through Syslog
8) Watermarking for dynamic port allocation over the top of Deterministic NAT

RFC Reference: N/A.

6.3.6.6 Feature: Softwire Initialization Dynamic Tunnels

The premise, function and related RFC reference relating to this feature is as given below.

Premise: SI Quick drop and pickup approach.

Function: All tunnelling initialization and dropping should be orientated towards a fast rotation on dynamic allocations given to each B4 node.

RFC Reference: N/A.

6.3.6.7 Feature: Port Block Allocation per IP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Port block allocation is configurable to allow for any ratio assignment per IP.

Function: An assignment of ports based on block allocations for any numeric value with automatic assignment of a single IP determined by the black allocation.

RFC Reference: N/A.

6.3.6.8 Feature: Deterministic NAT / Deterministic Dynamic Allocation and Thresholds

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Deterministic NAT policy and configuration has to be adhered to for AFTRs not only for port logging requirements for data retention but also for processing and functional improvement.

Function: DetNAT should be used in the following manner. It should be configurable to allow for any n-Ports to be allocated to the port block and any number of port block assignments to be dynamic or static per B4 address. It has to also hold a dynamic threshold to further the ability for IP to port ratios by allowing the dynamic section of the port allocation for any single or cluster of B4 nodes to use 1-5 block allocations and to allow for singular oversubscription per B4 based on a threshold. So if the block allocation reaches over 60 % (configurable) of the full amount of ports, 60 000 per IP, (5 000 reserved for well-known port allocation and PCP) a restriction is placed on further port allocations. For example until the 60 % threshold is reached the Subscriber B4 is allowed up to 2 000 ports (configurable) but once the port threshold of 60 % is reached this then drops to 500 ports for all users for all new current requests. No old connections are dropped but no new ones can be added. This allows for controlled oversubscription with heavy port users to over allocate while there are resources.

RFC Reference: [9**]**.

From section 2 in [9].

> "The algorithm is not designed to retrieve an internal host among those sharing the same internal IP address (e.g., in a DS-Lite context, only an IPv6 address/prefix can be retrieved using the algorithm while the internal IPv4 address used for the encapsulated IPv4 datagram is lost).
> Several address assignment algorithms are possible. Using predefined algorithms, such as those that follow, simplifies the process of reversing the algorithm when needed. However, the CGN MAY support additional algorithms. Also, the CGN is not required to support all algorithms described below. Subscribers could be restricted to ports from a single IPv4 address, or could be allocated ports across all addresses in a pool, for example. The following algorithms and corresponding values of A are as follow:
> Sequential (e.g. the first block goes to address 1, the second block to address 2, etc.)
> Staggered (e.g. for every n between 0 and ((65536-R)/(C+D))-1 , address 1 receives ports n*C+R, address 2 receives ports (1+n)*C+R, etc.)

Round robin (e.g. the subscriber receives the same port number across a pool of external IP addresses. If the subscriber is to be assigned more ports than there are in the external IP pool, the subscriber receives the next highest port across the IP pool, and so on. Thus, if there are 10 IP addresses in a pool and a subscriber is assigned 1000 ports, the subscriber would receive a range such as ports 2000-2099 across all 10 external IP addresses).
Interlaced horizontally (e.g. each address receives every Cth port spread across a pool of external IP addresses).
Cryptographically random port assignment (Section 2.2 of RFC6431 [RFC6431]). If this algorithm is used, the Service Provider needs to retain the keying material and specific cryptographic function to support reversibility.
Vendor-specific. Other vendor-specific algorithms may also be supported.
The assigned range of ports MAY also be used when translating ICMP requests (when re-writing the Identifier field).
The CGN then reserves ports as follows:

The CGN removes reserved ports (R) from the port candidate list (e.g., 0-1023 for TCP and UDP). At a minimum, the CGN SHOULD remove system ports (RFC6335) [RFC6335] from the port candidate list reserved for deterministic assignment.
The CGN calculates the total compression ratio (C+D), and allocates 1/(C+D) of the available ports to each internal IP address. Specific port allocation is determined by the algorithm (A) configured on the CGN. Any remaining ports are allocated to the dynamic pool. Note: Setting D to 0 disables the dynamic pool. This option eliminates the need for per-subscriber logging at the expense of limiting the number of concurrent connections that 'power users' can initiate.
When a subscriber initiates a connection, the CGN creates a translation mapping between the subscriber's inside local IP address/port and the CGN outside global IP address/port. The CGN has to use one of the ports allocated in step 2 for the translation as long as such ports are available. The CGN SHOULD allocate ports randomly within the port range assigned by the deterministic algorithm. This is to increase subscriber privacy. The CGN has to use the preallocated port range from step 2 for Port Control Protocol (PCP, [I-D.ietf-pcp-base]) reservations as long as such ports are available. While the CGN maintains its mapping table, it need not generate a log entry for translation mappings created in this step.
If D>0, the CGN will have a pool of ports left for dynamic assignment. If a subscriber uses more than the range of ports allocated in step 2 (but fewer than the configured maximum ports M), the CGN assigns a block of ports from the dynamic assignment range for such a connection or for PCP reservations. The CGN has to log dynamically assigned port blocks to facilitate subscriber-to-address mapping. The CGN SHOULD manage dynamic ports as described in [I-D.tsou-behave-natx4-log-reduction].
Configuration of reserved ports (e.g., system ports) is left to operator configuration.
Thus, the CGN will maintain translation-mapping information for all Connections within its internal translation tables; however, it only needs to externally log translations for dynamically-assigned ports."

6.3.6.9 Feature: IP Ranges per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP ranges should be able to be assigned per NAT group or across multiple NAT groups to allow for full scaling and shared IP resource.

Function: For further study.

RFC Reference: N/A.

6.3.6.10 Feature: NAT Grouping resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT Grouping SHOULD be cross NPU based.

Function: NAT groups SHOULD be able to use the same NPUs and IP allocations regardless of their GW addresses. This allows for complete shared resources across multiple NAT groups. The function states that the NPUs can be grouped in any manner for active standby numbers as the user/administrator wishes. Thus NAT group 1 can use NPU 1,2,3&4, with 1,2&3 as active and NAT group 2 can use NPU 1,2,3,4,5&6 with 1,2&3 as standby. This promotes scaling and no restriction if there are particular requirements to share a resource.

RFC Reference: N/A.

6.3.6.11 Feature: Virtual interface per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: An AFTR has to be able to assign a single virtual interface with up to 8 DS-Lite GW addresses for any given DS-Lite instance on the node.

Function: The virtual interfaces has to be able to be assigned to any number, and all, NPUs within the chassis to allow complete shared resources. So a single or multiple AFTRs acting as the single destination interface for all resources and B4 nodes using the AFTR.

RFC Reference: N/A.

6.3.6.12 Feature: NPU to Interface throughput ratio

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To allow balanced Bandwidth assignment across all virtual or a single virtual interface matching or exceeding the physical bandwidth of the node.

Function: To prevent over buffering, dropped packets and general resource issues the AFTR has to be able to consider a physical topology that matches a balanced logical bandwidth ratio between the interfaces and the NPU throughput.

RFC Reference: N/A.

6.3.6.13 Feature: AFTR Address (es)

The premise, function and related RFC reference relating to this feature is as given below.

Premise: AFTR Addressing.

Function: Listed below:

1) Has to Allow one or multiple addresses as GWs for either one or multiple NAT groups
2) IPv4 and IPv6 addressing SHOULD be placed on separate interfaces
3) The AFTR SHOULD be able to act as the DHCPv6 server for the B4 if required
4) SI-IDs should be able to comprise of both; ONLY the IPv6 B4 address or BOTH the IPv4 private and the IPv6 public address

5) The IPv6 Public address assigned to the B4 has to be unique

6) The IPv6 addresses configured for DS-Lite Gateways has to be Anycast compatible

7) The AFTR has to allow multiple NAT group per IP address and thus per Virtual interface end point

RFC Reference: [62**]**, [78], [116].

6.3.6.14 Feature: Anycast Gateway Address

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Anycast usage for an AFTR gateway or DHCPv6 address has to be accepted within the configuration.

Function: One or many Anycast addresses has to be configurable for the AFTR GWs regardless of any other resource and for the DHCPv6 server address. **Note** that using Anycast might cause failover issues if a path in your network changes and flapping could prevent session stability if clustering and cluster forwarding is not enabled.

RFC Reference: [62**]**.

6.3.6.15 Feature: Mulitple Source Prefix Filtering per AFTR Interface / Address

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The premise is that a definition of sourced addresses SHOULD be configured to allow forced routing to a specific GW address on the AFTR**.**

Function: A user/administrator SHOULD be able to place a source IP statement within a NAT group allowing the B4 packets originating from within that subnet to be routed to a specific NAT group or AFTR interface, so multiple. This allows the delineation of different service IPs instead of relying on separation of traffic based on DNS record responses and thus specific rules and resource allocations for that particular sourcing subnet.

RFC Reference: N/A.

6.3.6.16 Feature: AFTR Address Withdrawal

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The AFTR has to be withdrawn under the circumstances below.

Function: Listed below:

1) The AFTR is not reachable within the routing protocol being used or cannot reach its neighbour's or peers.

2) The NAT group goes down.

3) The DS-Lite process dies.

4) NPUs available cannot hold the present capacity.

5) Shutdown or part or all of the service.

6) External Routing to the internet is down for a fixed period of time (demands AFTR clustering to determine reachability from a redundant AFTR).

RFC Reference: N/A.

6.3.6.17 Feature: Chassis DS-Lite Clustering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To Cluster two or more AFTRs for session state failover.

Function: Clustering will be based on index reference tabling allowing for minimal transfer of data requirement between the cluster to swap state caches. If a single AFTR goes down it has to do the following:

- Remove its AFTR address as a reachable entity within the routing protocol being used
- Forward any remaining packets to the secondary AFTR
- Validate NAT cache state between itself and its redundant AFTR
- SI-ID extensions SHOULD be used in clustering customer identification

The cluster has to sync on the following items:

- The TCP and UDP sessions presently being used to for DS-Lite
- The TCP not used for n amount of time configurable
- The present capacity of interface and NPU to allow for structural syncing
- Sync between all clustered AFTRs due to best path attributes for the Anycast/Unicast Addresses used (also dependent on secondary addressing on the CPE)
- PCP and NAT Caches should be synced

RFC Reference: N/A.

6.3.6.18 Feature: Multiple Transition Technology Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT44, NAT64 and DS-Lite has to function on the same platform to allow for uniformity within European markets.

Function: All three technologies should be able to function and share resources dynamically on any given platform.

RFC Reference: N/A.

6.3.6.19 Feature: NPU / chassis Buffering

Premise: Buffering requirements for the AFTR for transit traffic covering queues and buffer timers.

Function: Listed below:

1) Queue sizing per flow
2) Queue sizing per port block
3) Queue sizing per NPU
4) Queue sizing per ingress interface
5) Buffer sizing per flow
6) Buffer sizing per port block
7) Buffer sizing per NPU
8) Buffer sizing per ingress interface
9) Reordering buffers and out-of-order caching

RFC Reference: N/A.

6.3.6.20 Feature: Tunnel MTU Sizing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Tunnel MTU SHOLD be considered as the IPv4 encapsulated within the IPv6 packet and therefore should be set at 1 500, the IPv6 header being 40 bytes, if IPv4 only the tunnel MTU SHOULD be configurable between 1 420 to 1 460 to allow for all possible header combinations. Note that some stacks consider the TCP header (not just IPv4 and IPv6) as part of the tunnel MTU, although very rare. All considerations SHOULD be taken into account. 1 420 should be the lowest tunnel MTU available especially considering MSS clamping options. Notation here is that 1 280 can be used on point-to-point or point-to-multipoint links within any given network so the full path MTU of the DS-Lite tunnel, of IPv6, should be known before deployment and configuration of specific tunnel MTUs. It is recommended that 1 500 is used through your network up to any given AFTR placement.

Function: For further study.

RFC Reference: [129**]** not completely applicable

From section 5.3 in [129]:

> "Using an encapsulation (IPv4-in-IPv6 or anything else) to carry IPv4 traffic over IPv6 will reduce the effective MTU of the datagram. Unfortunately, path MTU discovery [RFC1191] is not a reliable method to deal with this problem.
> A solution to deal with this problem is for the service provider to increase the MTU size of all the links between the B4 element and the AFTR elements by at least 40 bytes to accommodate both the IPv6 encapsulation header and the IPv4 datagram without fragmenting the IPv6 packet."

6.3.6.21 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: MSS clamping, the MTU value for the TCP max segment size, which should be configurable to as low as 1 420 in case any MSS client considers all header within the MSS value.

Function: MSS clamping should be configurable and functional for both upstream and downstream traffic.

RFC Reference: [29**]**.

6.3.6.22 Feature: DS-Lite Fragmentation and Buffering according to RFC 6333

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Fragmentation has to be placed on the ingress interface (virtual or physical) post decapsulation on upstream and pre-encapsulation on downstream on IPv4.

Function: All fragmentation on DS-Lite on the AFTR and CPE has to be on IPv4. This prevents IPv6 fragmentation requirements and overlay code. IPv6 inherently does not have fragmentation built into it and thus can cause major performance issues on an AFTR forwarding plane or even be placed in software. RFC 6333 [129] states that IPv6 fragmentation and reassembly should be used but the present document does not condone that methodology. IPv4 fragmentation is well coded and has little effect in present day on forwarding performance and thus should be used where possible. If the DF bit is set then RFC 6333 [129] IPv6 fragmentation may be used where MSS is ignored on TCP flows and on UDP/ICMP flows where required. This should be used in conjunction with Stateful ICMP (three ICMP t3/c4 messages per flow max) allowing for path MTU function.

Reassembly has to only be used when the AFTR receives noted IPv6 fragmented packets incoming upstream from the CPE and thus requires full buffering/re-ordering to prevent timeouts and drops.

Pre-fragmented packets has to be re-ordered before being sent on and again requires full buffer queues and wait timers exceeding 3ms.

RFC Reference: [129**]**.

6.3.6.23 Feature: Stateful ICMP with PMTU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Stateful ICMP is the act of holding state for ICMP for messages sent/originated from the AFTR to any know destination on a single flow. This can be used downstream as well as upstream.

Function: A maximum of 3 ICMP packets to big messages has to be sent to single destination for a single flow. This prevents flooding and over compensation. This should ONLY be used when the DF bit is set and should be tied into the fragmentation process of DS-Lite packets.

RFC Reference: N/A.

6.3.6.24 Feature: Port Reservation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Reserving a block of public ports for each private IPv4 client reduces the scaling requirements of per-connection port-mapping.

Function: The AFTR has to be capable of bulk/block port reservation.

RFC Reference: [136**]**.

From section 2.3 in [136]:

> "For operators who desire to aggregate the logs, they can configure the AFTR to preallocate a range of ports to each B4. This range of ports will be used in the NAT44 function, and the AFTR will create one log entry for the whole port range. This aggregation can significantly reduce the log size for source-specific logging."

6.3.6.25 Feature: Static Port Forwards

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Allowing the consumer to configure static port forwarding, using a portal, to configure the AFTR.

Function: Using PCP the customer can log on to their "ISP website" and configure instantaneously a port to forward to their in-home service or server by IP assignment. The AFTR will assign that port return path into the home for a specific public IP for a fixed period of time. Note that this crosses the port block assignment restriction of being within a single IP address and thus creates a new DR log. This based mostly on the AFTR port changes, however PCP may be used to inform the CPE to create a double pop on the port allocation, so the AFTR present the public IPv4 IP on the well known port translates it to private IPv4 and public IPv6 based SI-ID extension keep the destination port to the consumer CPE the same, however if there is already an entry stated for that SI-ID then a double pop maybe necessary requiring the CPE to change the port once again and thus requiring PCP exchange. An extended state SHOULD be held for each double pop port exchange within the extended PCP table.

RFC Reference: N/A.

6.3.6.26 Feature: PCP Mode

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The mode of PCP operation should be has to be primarily plain **option 1** below, but encapsulation SHOULD be supported as well (configurable).

Function: **Note that the AFTR SHOULD support both RFC drafts 12/13 as well RFC draft 19 for PCP (configurable).**

Preference form is to have plain mode on a separate address, this allows two topologies and the ability for processing of the PCP packets to be separated without the need for mixed traffic (DS_lite std traffic and PCP traffic) across the tunnel and therefore further read requirements on forwarding that could impact performance. Also the separation of the traffic allows for separation of topologies in the case of a PCP server existing isolated from the AFTR node.

All sourcing in any given topology or mode has to support the IPv6 address as the identifier to allow for NAT cache matches between the PCP tables and the main NAT CACHE for the SI-ID (softwire initiated identifier). This is a requirement for AFTR redundancy clustering, data retention and other such feature sets.

RFC Reference: [11**]**.

From section 2 in [11]:

> "In the plain mode the B4, the customer end-point of the DS-Lite IPv6 tunnel, implements a PCP proxy ([I-D.bpw-pcp-proxy]) function and uses UDP over IPv6 with the AFTR to send PCP requests and receive PCP responses.
>
> The B4 has to source PCP requests with the IPv6 address of its DS-Lite tunnel end-point and has to use a THIRD PARTY option either empty or carrying the IPv4 internal address of the mappings.
>
> In the plain mode the PCP discovery ([I-D.ietf-pcp-base] section 7.1 "General PCP Client: Generating a Request") is changed into:
>
> 1. if a PCP server is configured (e.g., in a configuration file or via DHCPv6), that single configuration source is used as the list of PCP Server(s), **else**;
> 2. use the IPv6 address of the AFTR. To summary: the first rule remains the same with the precision that DHCP is DHCPv6, in the second rule the default router list is replaced by the AFTR."

From Appendix A in [11]:

> "The encapsulation mode deals at the B4 side with PCP traffic as any IPv4 traffic: it is encapsulated to and decapsulated from the AFTR over the DS-Lite IPv4 over IPv6 tunnel.
>
> At the AFTR side things are a bit more complex because the PCP server needs the context, here the source IPv6 address, for both to manage mappings and to send back response. So the AFTR has to tag PCP requests with the source IPv6 address after decapsulation and before forwarding them to the PCP server, and use the same tag to encapsulate PCP responses to correct B4s. (the term "tag" is used to describe the private convention between the AFTR and the PCP server)."

6.3.6.27 Feature: PCP Failure

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function and reactivity during a failure of the session either at Layer 3, 4 and 5 at the client, server, AFTR or CPE.

Function: Listed below:

1) Relay of failure messages from AFTR to CPE receive

2) Relay of failure messages from CPE to AFTR send

3) Lifetime timer configuration for failure; PCP entry timers

4) Monitoring failure of the of the PCP without message

5) Static mappings are not affected (comparative if the CPE holds those static explicit mappings)

6) PCP table content considering failure (last received packet). PCP table matching when/if required

RFC Reference: N/A.

6.3.6.28 Feature: PCP (SI-ID Extension Base) Multi-session Dynamic Forwarding

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP allocation for multiple instances of the same application on separate PCs behind the same B4.

Function: Allows multiple sessions from multiple PCs behind a single B4 to access the same port by assigning as separate IP to the before (IPv4 public). The ability to dynamically allocate a static port forwards for a different outside IP as the subscriber got assigned when the particular port is in use.

Reasoning is bittorrent clients with UPnP and PCP to the AFTR will not be able to negotiate the same port for multiple sessions from different PCs and given all clients use the same "fixed" port by default it is better to allocate a different outside IP.

RFC Reference: N/A.

6.3.7 Monitoring and Management

6.3.7.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.3.7.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.3.7.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.3.7.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.3.7.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.3.7.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18**]**.

6.3.7.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.3.7.8 Feature: SNMP DS-Lite General

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The AFTR should support specific MIBs to allow full visibility of the resource and traffic.

Function: MIBs should be available for the following:

- Port percentage utilization per IP
- IP Address percentage per range assigned and per NAT group
- Resource utilization per NPU
- Memory utilization
- Processor utilization per NPU
- PCP Percentage IP assignments
- Clustering utilization (the total utilization if one of the pair of the cluster goes down)
- Interface utilization BW

- DS-Lite sessions (dynamic thresholds)
- SI-ID (extension) numbers per NPU and chassis (based on both private IPv4 and public IPv4 used out of the range assigned)

RFC Reference: N/A.

6.3.7.9 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26**]**.

6.3.7.10 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.3.7.11 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60].

6.3.7.12 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.3.7.13 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.3.7.14 Feature: Remote Access: TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.3.8 DS-Lite CPE Requirements

Requirements on the host are determined by the available connectivity in the Home Network. The Home Network may be dual stack or IPv4 only or IPv6 only. In the case that IPv4 is used, the Home Network has to support private addressing.

The following requirements apply to the CPE Router:

- WAN interface facing the Access Network:
  - Request IPv6 global address via DHCPv6
  - Request B4 IPv6 prefix via DHCPv6-PD
- LAN interface facing the Home Network:
  - IPv4 private addressing
  - DHCPv4 server for LAN addressing (stateful DHCP service may be included)
  - IPv4 DNS Proxy
  - IPv4 MTU set to 1 460 Byte
  - TCP MSS clamping to 1 420 Byte
- B4 router:
  - Receive AFTR IPv6 address via DHCPv6
  - Secondary AFTR IPv6 address via DHCPv6
  - Use DNS IPv6 server received via DHCPv6
  - Encapsulate LAN IPv4 packets in IPv6 header
  - Decapsulate IPv6 packets for LAN IPv4
  - Fragment encapsulated IPv6 packets
  - Reassemble received IPv6 fragments
  - PCP
  - Implement the well-known B4 IPv4 address

Assuming that the CPE Router is embedded in a Cable Modem, additional requirements apply:

- Should support bridging of IPv6 packets
- May support IPv4 or IPv6 management
- If IPv6 is implemented, NDP and SLAAC have to be supported as well
- If the Cable Modem is implementing DOCSIS 2.0 and IPv6 it has to comply to:
  - DOCSIS 2.0+IPv6 Cable Modem Specification
  - IPv6 management as defined in the OSSIv3.0 Specification

  - Provisioning Mode Override
- If the Cable Modem is implementing DOCSIS 3.0 it has to comply to:
  - MAC and Upper Layer Protocols Interface Specification
  - Upstream Drop Classifiers
  - OSSIv3.0 Specification
  - Provisioning Mode Override
- If the embedded device implements eRouter it has to support:
  - IPv6 provisioning of CPE devices
  - IPv6 address assignment
  - Identity of DHCPv4 client identifier and DHCPv6 DUID]
  - TR-069 via native IPv6 transport to allow remote management of WiFi and router parameters

In addition to the requirements listed above, the CPE Router should support the following features on its interfaces facing the Home Network (wired, wireless):

- SLAAC
- DNS resolver information
- Stateless DHCPv6 (INFORM) server or, potentially, stateful DHCPv6 server (IA_NA) with the option to switch from SLAAC
- IPv4 NAT/NAPT
- IPv4 static NAT
- IPv4 inbound port forwarding
- IPv4 stateful packet firewall (5-tuple filters), enabled by default
- IPv6 stateful packet firewall (5-tuple filters), enabled by default
- Recursive DNS server option
- Requirements for IPv6 customer edge routers

For purposes of a unified user experience some requirements apply to the user interface:

- Web-UI should be accessible on the LAN IP interface
- Login is initially presented in the format of requesting username and password
- Localization of UI for different languages should be supported
- A Status tab should contain subpages on Software, Connection, Security, Diagnostics
- A Router Basic tab should contain subpages on WAN Setup, LAN&DHCP Server, Backup
- A Router Advanced tab should contain subpages on Option, IP Filtering, MAC Filtering, Port Filtering, Forwarding, Port Triggers, DMZ Host
- A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log
- A Parental Control tab should contain subpages on User Setup, Basic Setup, Content Filter, ToD Filter
- A Wireless tab should contain subpages on Radio, Security, Advanced, Access Control

An MTA tab should contain subpages on Status, DHCP, QoS, Provisioning, Event Log.

6.3.8.1 CPE Definition

6.3.8.1.1 Feature: Cable Router B4 Functionality

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to provide tunnelling of IPv4 based traffic from either directly connected or other CPE devices the DOCSIS cable modem has to provide the capability to implement the required functionality to deliver such traffic via an IPv6 tunnel to an AFTR router. This is defined as the DS-Lite Basic Bridging Broadband - B4 element, This functionality is provided by the implementation of a B4 interface on the DOCSIS interface of the cable modem.

Function: The DS-Lite CPE SHOULD NOT operate any NAT functionality between the B4 interface and any local LAN interfaces as this functionality is provided by the AFTR gateway in the network.

6.3.8.1.2 Feature: SLAAC

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv6 Address Auto Configuration, In order to allow for the generation of a unique IP address without reliance on an external server or manual provisioning the B4 device has to support the IPv6 stateless autoconfiguration method as detailed in RFC 4862 [97]. This methodology ensures that a device can generate a unique and routeable IPv6 address based upon locally available information and advertisements from adjacent routers.

Function: The DS-Lite B4 CPE device has to be capable of supporting IPv6 stateless autoconfiguration.

6.3.8.1.3 Feature: DNS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: As the B4 client is only provided with IPv6 connectivity on its WAN interface, it has to perform all DNS resolution over IPv6.

Function: To reduce the amount of DS-Lite traffic, the cable B4 gateway has to be configured as a DNS forwarder for all LAN client requests. Any client DNS requests received via IPv4 will therefore be forwarded as IPv6 to the external DNS servers.

RFC Reference: N/A.

6.3.8.1.4 Feature: IPv6 LAN IP Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are no specific requirements for IPv6 LAN addressing to support DS-Lite. As DS-Lite provides native IPv6 end-to-end connectivity, IPv6 addressing SHOULD be used for all LAN clients that support IPv6. This being defined by the provisioning policy implanted on the service providers provisioning systems.

RFC Reference: N/A.

6.3.8.1.5 Feature: IPv4 LAN IP Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The B4 has to provide 'private' IP addressing from the ranges specified in RFC 1918 [30]. The accepted normal configuration is to use the 192.168.0.0/16 range, with a single /24 subnet.

The B4 has to use one address from the selected IPv4 subnet as its LAN interface address, and this has to be used as the default gateway for LAN IPv4 traffic.

The B4 has to also use its LAN IP address as the DNS server address in DHCP offers.

RFC Reference: [24**]**.

6.3.8.1.6 Feature: Packet Encapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The B4 router has to encapsulate packets in accordance with RFC 6333 [129]. The B4 encapsulation interface has to drop any packets in the following groups:

Function: Listed below:

- Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].
- Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].
- Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].
- Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: N/A.

6.3.8.1.7 Feature: Packet Decapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The B4 router has to decapsulate packets in accordance with the specification defined in RFC 6333 [129].

Function: Where an IPv4 fragment is received, the B4 router has to forward the fragment unaltered to its destination. It is the responsibility of the destination client to reassemble fragments.

Once the IPv4 packet is recovered, it has to forward the packet to the client specified in the source address.

RFC Reference: N/A.

6.3.8.1.8 Feature: MTU and fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Encapsulation of an IPv4 packet with an IPv6 header to transport across the cable network increases the overall packet size by 40 bytes.

The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes. The maximum IPv4 packet size that can be sent between the B4 and AFTR is therefore 1 460 bytes.

Function: As fragmentation in IPv6 is performed by the sending host, any packet received that is in excess of this size has to either:

Be encapsulated by the B4, and fragmented into two packets for transmission to the AFTR.

Be reduced in size to a maximum of 1 460 bytes before being encapsulated.

Function: IPv6 fragmentation at B4

IPv6 packets are fragmented by the sender and reassembled by the receiver. In the case of DS-Lite, the sender and receiver are the B4 and AFTR, before translation to IPv4.

Every outbound packet that is too big is therefore fragmented by the B4, and reassembled by the AFTR.

Inbound packet is fragmented by the AFTR, and reassembled by the B4 before translation to IPv4.

The fragmentation and reassembly of IPv6 packets is processor-intensive for the B4 gateway. Many applications will send packets using the maximum packet size, so, for example during a file transfer, every packet will require fragmenting or reassembly before it can be translated to IPv4. This is likely to give a poor overall performance.

RFC Reference: N/A.

Function: IPv4 MTU reduction

The MTU of the B4 encapsulation interface can be reduced to 1 460 bytes. This will have three effects:

1) Any client that uses Path MTU Discovery (PMTUD) will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

2) Any IPv4 packet received by the B4 that is in excess of 1 460 bytes, with the DF bit clear, will be fragmented by the B4 before encapsulation. The IPv6 packet will therefore fit within the IPv6 MTU of 1 500 bytes.

3) nbAny IPv4 packet received by the B4 that is in excess of 1 460 bytes, with DF set will be dropped by the AFTR, and an ICMP 'packet too big' message returned to the client. The client is expected to process this message and reduce the packet to the size specified in the ICMP reply.

   Whilst the fragmentation of IPv4 packets involves additional processing on the B4, the IPv4 packet is reassembled by the destination host. This is therefore less impact on the AFTR. Similarly, for inbound packets, an IPv4 packet fragmented by the AFTR is reassembled by the receiving client.

   Fragmentation at an IPv4 layer before encapsulation still has an overhead on the DOCSIS transport as each packet sent still doubles the packet throughput requirement.

6.3.8.1.9 Feature: MSS clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header)

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

MSS clamping has to be implemented on the B4 encapsulation interface, to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size).

Function: For further study.

RFC Reference: N/A.

6.3.8.1.10 Feature: DHCPv4 option 26

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv4 option 26 as defined in RFC 2132 [33] is used to specify the IPv4 MTU that each client should use.

Function: DHCPv4 option 26 as defined in [33] is used to specify the IPv4 MTU that each client should use.

The setting of this parameter will ensure that packets leaving the host are a maximum of 1 460 bytes. However this has the following limitations:

1) Not all client IP stacks and DHCP clients will respect the DHCP option to set the interface MTU.

2) Devices that are manually configured (and hence do not use DHCP) would also require manual configuration of the interface MTU.

3) Reducing the interface MTU of clients would also impact all LAN IPv4 traffic as this would also all be reduced to an MTU of 1 420 - 1 460 bytes.

4) It cannot be guaranteed that all client applications will function if the MTU has been reduced.

RFC Reference: N/A.

6.3.8.1.11 Feature: Recommendations

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate DS-Lite encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU has to be set to allow for full IPv6 encapsulated DS-Lite packets to be sent with no fragmentation.

Where the DOCSIS layer only permits an MTU of 1 500 bytes, the following recommendations are required.

1) The device B4 interface has to be configured with an MTU of 1 460 bytes.

2) The device has to fragment IPv4 packets greater than 1 460 bytes where DF is not set.

3) The device has to drop packets greater than 1 460 byte where DF is set and reply with an ICMP "packet too big" message and specify the MTU of 1 460 bytes.

4) The device has to support MSS clamping to 1 420 bytes.

5) The device has to support DHCPv4 option 26 set to 1 460 bytes.

RFC Reference: N/A.

6.3.8.1.12 Feature: Inbound sessions

The premise, function and related RFC reference relating to this feature is as given below.

Premise: With DS-Lite, the public IPv4 address is not configured on the B4. The ability to control inbound sessions is also removed. These are generally:

1) DMZ host functionality: where all inbound packets can be forwarded to a single client.

2) Port forwarding: where individual TCP or UDP ports can be forwarded to a specific client.

3) UPnP NAT Traversal: where a client requests for the NAT to open one or more TCP or UDP ports to itself for inbound packets.

Function: The B4 router has to be able to replicate the existing behaviour, by communicating these requests to the AFTR, for it to implement the relevant port forwarding to the specific client.

Port Control Protocol (PCP) is the proposed mechanism to provide this functionality. It is currently in draft-ietf-pcp-base [140].

RFC Reference: N/A.

6.3.8.1.13 Feature: Dual Stack Lite QoS using the IPv6 Flow label

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Provision of QoS for IPv4 encapsulated packets The implementation of DS-Lite encapsulates IPv4 traffic within an IPv6 tunnel and thus obfuscates the 5-tuple of the IPv4 traffic and therefore ensuring that all such traffic is classified and prioritized with the same information as all other IPv4 traffic encapsulated within the IPv6 tunnel.

This can prove an issue where differentiated QoS is required for traffic such as VoIP.

Function: The B4 router has to set the IPv6 Flow Label to a non-zero value per IPv4 traffic flow in accordance with RFC 3697 [67].

RFC Reference: [10**]**, [67].

6.3.8.1.14 Feature: MIB Support for DS Lite

Premise: In order to support configuration of the tunnel and NAT functions required by DS-Lite the CPE device needs to implement an enhanced MIB that provides the required functionality that is not supported by the implementation of NAT-MIB [i.5] and the tunnel MIB [i.6].

Function: The B4 router has to implement the requirements of DS-Lite Management Information Base (MIB).

RFC Reference: [14**]**.

6.3.8.1.15 Feature: Port Reservation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Reserving a block of public ports for each private IPv4 client reduces the scaling requirements of per-connection port-mapping.

Function: The AFTR has to be capable of bulk/block port reservation.

RFC Reference: [136**]**.

From section 2.3 in [136]:

> "For operators who desire to aggregate the logs, they can configure the NAT64 to preallocate a range of ports to each CPE. This range of ports will be used in the NAT44 function, and the NAT64 will create one log entry for the whole port range. This aggregation can significantly reduce the log size for source-specific logging."

6.3.8.1.16 Feature: PCP Mode

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The mode of PCP operation should be has to be primarily plain **option 1** below, but encapsulation SHOULD be supported as well (configurable). The mode denotes the separation of traffic or not from the AFTR address and the type of encapsulation.

Function: **Note that the AFTR SHOULD support both RFC drafts 12/13 as well RFC draft 29 for PCP (configurable)**

Preference form is to have plain mode on a separate address, this allows two topologies and the ability for processing of the PCP packets to be separated without the need for mixed traffic (DS_lite std traffic and PCP traffic) across the tunnel and therefore further read requirements on forwarding that could impact performance. Also the separation of the traffic allows for separation of topologies in the case of a PCP server existing isolated from the AFTR node.

All sourcing in any given topology or mode has to support the IPv6 address as the identifier to allow for NAT cache matches between the PCP tables and the main NAT CACHE for the SI-ID (softwire initiated identifier). This is a requirement for AFTR redundancy clustering, data retention and other such feature sets.

RFC Reference: [11**]**.

From section 2 in [11]:

> "In the plain mode the B4, the customer end-point of the DS-Lite IPv6 tunnel, implements a PCP proxy ([I-D.bpw-pcp-proxy]) function and uses UDP over IPv6 with the AFTR to send PCP requests and receive PCP responses.
>
> The B4 has to source PCP requests with the IPv6 address of its DS-Lite tunnel end-point and has to use a THIRD PARTY option either empty or carrying the IPv4 internal address of the mappings.
>
> In the plain mode the PCP discovery ([I-D.ietf-pcp-base] section 7.1 "General PCP Client: Generating a Request") is changed into:
>
> 1. if a PCP server is configured (e.g., in a configuration file or via DHCPv6), that single configuration source is used as the list of PCP Server(s), **else**;
> 2. use the IPv6 address of the AFTR. To summary: the first rule remains the same with the precision that DHCP is DHCPv6, in the second rule the default router list is replaced by the AFTR."

From Appendix A in [11]:

> "The encapsulation mode deals at the B4 side with PCP traffic as any IPv4 traffic: it is encapsulated to and decapsulated from the AFTR over the DS-Lite IPv4 over IPv6 tunnel.
>
> At the AFTR side things are a bit more complex because the PCP server needs the context, here the source IPv6 address, for both to manage mappings and to send back response. So the AFTR has to tag PCP requests with the source IPv6 address after decapsulation and before forwarding them to the PCP server, and use the same tag to encapsulate PCP responses to correct B4s. (the term "tag" is used to describe the private convention between the AFTR and the PCP server)."

6.3.8.1.17 Feature: PCP Failure

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function and reactivity during a failure of the session either at Layer 3, 4 and 5 at the client, server, AFTR or CPE.

Function: Listed below:

1) Relay of failure messages from AFTR to CPE receive
2) Relay of failure messages from CPE to AFTR send
3) Lifetime timer configuration for failure; PCP entry timers
4) Monitoring failure of the of the PCP without message
5) Static mappings are not affected (comparative if the CPE holds those static explicit mappings)
6) PCP table content considering failure (last received packet). PCP table matching when/if required

RFC Reference: [7**]**.

6.3.9 DS-Lite Technical Viability

DS-Lite is at present the most viable technology for transition into IPv6 for the MSO industry. Technically and architecturally it has some drawbacks, but it is as close to a non-service deprecating technology as possible and for some deployments it will be. Functionally it should only be used for business continuity for new additions to your network and thus the issue of centralized placement in the network becomes minimal.

The Pros

- DS-Lite is as close as the industry can get to a non-service-deprecating technology at present, with clustering, PCP and DeNAT most of the CGN issues are mitigated or completely disappear.
- Allows for a smooth migration towards an IPv6 only operator infrastructure. This includes IPv6 only access, aggregation and core network as well as an IPv6 only BNG. In case of a cable operator the BNG is implemented on the CMTS. In case of other access technologies (e.g. GPON) the BNG is implemented on a dedicated platform, but this moves us closer to higher requirements for development of IPv6.
- It removes the need for IPv4 address management and as such a DHCPv4 server is not required anymore.
- As a business continuity process for new customers the scaling is acceptable in live deployment.
- Central intelligence makes development much easier and upgrading only a few nodes in the network whereas CPE state technologies incur massive requirements for upgrades and difficulty with functionality over multiple vendors.
- As an AFTR only handles IPv4 traffic, the capacity requirements per customer on the AFTR will likely decrease over time. The reason is that more and more internet traffic will be on IPv6. This might allow for operator growth over time without the need for large investments on the AFTR.

The Cons

- The cost, particularly of the CPE router, is expected to be increased due to the extension of the required functionality. MSOs that are not deploying WiFi solutions and stick to bridging CPE devices will face the issue of having to exchange CPE just for the reason of introducing DS-Lite and thus the cost.
- DS-Lite requires ALGs (ALPs in most cases) which require the AFTR to do some form of intelligent vicissitudes to the transit packets bearing the risk of drop in performance.
- The requirement for Data Retention capacity is rather large but this is mitigated with DeNAT.
- DS-Lite is a hub-and-spoke topology forcing all tunnelled communication from B4 to pass through the AFTR which causes.

6.4 NAT64 Technology Summary

NAT64 is one of the technologies the current report recommends to consider for standardization mainly due to the lack of a need for inter-communication between IPv4 and IPv6. This technology will allow customers to access services natively over IPv6 and through translation over IPv4.

In order to enable connectivity between IPv6 hosts and the Internet, NAT64/DNS64 presents always an IPv6 address to the host independently if communication is to be established with an IPv6 or IPv4 addressable device. Communication to an IPv4 device is enabled by synthesizing the DNS A record into a AAAA record (DNS64) and by IPv6 to IPv4 address translation via a NAT64 device. As such, the technology is dependent on DNS and requires devices in the home to be natively IPv6 capable. IPv4-only devices and non-DNS based applications will not work in this environment.

NAT64 allows a client in the IPv6 domain to initiate communication with a server in the IPv4 domain by translating source IPv6 address and port to IPv4 address and port. It works in conjunction with a modified DNS known as DNS64. NAT64 relies on DNS64 to provide an AAAA record (corresponding to the server in the IPv4 domain) to IPv6-only hosts initiating communication with the IPv4 server. The AAAA record is created from the A record for the IPv4 address. The IPv4 address is mapped to an IPv6 address prepending a well-known IPv6 prefix assigned to the NAT64 gateway. NAT64 manages a pool of public IPv4 addresses and performs a NAPT function by translating IPv6 source address and port to IPv4 source address and port. This is shown in figure 4.

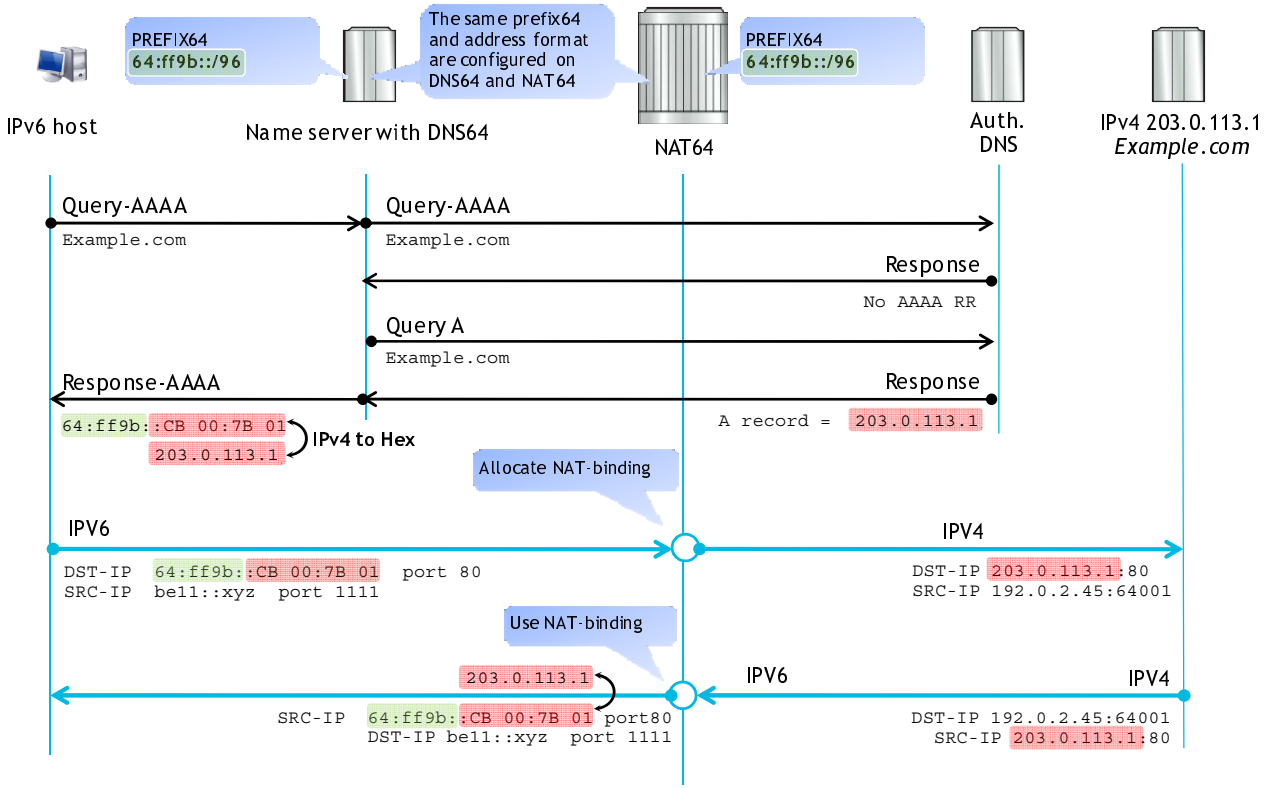


Figure 4: Addressing scheme in NAT64

For TCP and UDP flows, NAT64 maintains mapping between the IPv6 transport address and port and the IPv4 transport address and port and performs header translations. For ICMP, stateful NAT64 needs to maintain mapping between the IPv6 transport address and ICMPv6-identifier and the IPv4 transport address and ICMPv4-identifier.

The NAT64 prefix can be:

- By default the well-known prefix 64:ff9b::/96 (with fixed prefix length of 96 bit).
  This is best practice.
- A network specific prefix.
  The addressing scheme defined for NAT64 [119] allows subnet lengths for the NAT64 prefix to be 32, 40, 48, 56, 64 or 96 bit.

Depending on the prefix length, the IPv6 address with embedded IPv4 address is formatted according to table 6.

Table 6: Embedding IPv4 addresses in IPv6 addresses with different prefix lengths

0-15	16-31	32-47	48-63	64-79	80-96	96-111	112-128	
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix Prefix	0 Prefix	Prefix Prefix	IPv4 IPv4	IPv4 IPv4	/96
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix Prefix	u IPv4	IPv4 IPv4	IPv4 Suffix	Suffix Suffix	/64
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix IPv4	u IPv4	IPv4 IPv4	Suffix Suffix	Suffix Suffix	/56
Prefix Prefix	Prefix Prefix	Prefix Prefix	IPv4 IPv4	u IPv4	IPv4 Suffix	Suffix Suffix	Suffix Suffix	/48
Prefix Prefix	Prefix Prefix	Prefix IPv4	IPv4 IPv4	u IPv4	Suffix Suffix	Suffix Suffix	Suffix Suffix	/40
Prefix Prefix	Prefix Prefix	IPv4 IPv4	IPv4 IPv4	u Suffix	Suffix Suffix	Suffix Suffix	Suffix Suffix	/32

Bits 64 to 71 (u) should always be set to zero even when using a /96 prefix.

The NAT64 translation causes a change in MTU. In addition to the minimum length of 40 Byte for the IPv6 header, 20 Byte length of the IPv4 header have to be taken into account. If after IPv4 to IPv6 translation the IPv6 link MTU is exceeded, it is recommended to fragment the IPv4 packets before they enter the NAT and to set the Max Outside MTU of the NAT accordingly.

Max Outside MTU = IPv6 MTU - 40 Byte IPv6 header - 8 Byte IPv6 fragmentation header + 20 Byte IPv4 header

For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.

6.4.1 NAT64 LSN Technology Feature Summary

The NAT64 technology feature is summarized in table 7 detailing for each function the requirement as required or optional with a brief description of each of the named functions.

Table 7: NAT64 LSN Technology Feature Summary

Functional Name	Requirement	Description
NAT64 - RFC 4787 [94] (UDP)	Required	Compliance with NAT behaviour according to RFC 4787 [94] for UDP.
NAT64 - RFC 5382 [106] (TCP)	Required	Compliance with NAT behaviour according to RFC 5382 [106] for TCP.
RFC 6052 [119]	Required	Compliance with RFC 6052 [119]: IPv6 Addressing of IPv4/IPv6 Translators.
RFC 6145 [124]	Required	Compliance with RFC 6145 [124]: IP/ICMP Translation Algorithm.
RFC 6146 [125]	Required	Compliance with RFC 6146 [125]: Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers.
DNS64	Required	DNS is a primary function of NAT64 with the DNS server function extended to supply the NAT64 LSN GW address with the A record embedded within the last 32 bits of the "false" DNS record sent back to the client.
Redundancy	Required	All critical components has to be redundant in such a manner that they can fail-over without impact to customer or management traffic greater than 1 ms.
Shared Resource	Required	Single NAT64 GW prefix. The NAT64 IPv6 prefix should be able to be shared amongst different NPU's in the NAT64. A hashing mechanism should be in place to hash all upstream packets based on the source IPv6 address.
Si-ID	Optional	Tunnel/customer identifier based on IPv6 CPE address.
NAT64 Addressing and Virtual Interfaces	Required	NAT64 has to be able to assign a single virtual interface with up to 8 NAT64 GW prefixes for any given NAT64 instance on the node.
Anycast	Required	Anycast NAT64 gateway prefixes are a requirement to allow simplicity of deployment for a single prefix across multiple NAT64's.
NAT64 Address withdrawal	Required	The NAT64 should have at least five points of NAT GW prefix withdrawal occurrence. The list includes: - loss of route out, - loss of all BGP/IGP sessions, - loss of forwarding, - loss of NPU capacity and certain errors in the NAT caching. Any of the failures should be detectable based on configurable timers with 15 seconds being the default setting.
NPU Load Balancing (hashing)	Required	NAT64 Inside to Outside hashing performed on the Source IPv6 (128 bits) address of the CPE device. NAT64 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
MSS Clamping	Required	TCP MSS support is mandatory for the NAT64 due to the removal of an end-to-end MTU sizing functionality. This will avoid the need for excessive fragmentation.
Fragmentation	Required	Fragmentation should be done on the IPv4 packet. Fragmentation should be placed on the ASIC running in the line card. Optionally, fragmentation can be done on the IPv6 packet.
NAT - Network Address and Port Mapping - Endpoint Independent Mapping	Required	For two flows for a common inside source IPv6 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation.
NAT - Translation Filtering - Endpoint Independent Filtering	Required	A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.
NAT - Paired IP Address Assignment	Required	Translation to External IPv4 address is done in a paired fashion. A given Inside IPv6 address is always translated to the same External IPv4 address.

Functional Name	Requirement	Description
NAT - Hair-pinning	Required	Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.
NAT - 1:1 IP Mapping	Required	Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).
NAT44 - Outside-Service-App mapping for inside-VRF	Required	Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.
NAT - Port Limit configuration	Required	A maximum amount of ports can be configured for every IPv6 source address.
NAT - Per-Protocol Timeout configuration	Required	Timeout of mappings is critical and removing a NAT mapping at the appropriate time maximizes the shared port resource.
NAT - Dynamic Port Range start configuration	Required	The start port for dynamic port ranges should be configurable to allow for a range of ports for static port mappings.
NAT - Software Load Balancing	Required	NAT Inside to Outside hashing performed on the Source IPv6 CLAT Prefix. NAT44 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Port Allocation	Required	In order to reduce the volume of data generated by the NAT device (logging creation and deletion data), bulk port allocation can be enabled. When bulk port allocation is enabled and when a subscriber creates the first session, a number of contiguous outside ports are pre-allocated. A bulk allocation message is logged indicating this allocation.
Deterministic NAT64	Optional	Algorithmically maps a customer's private IPv6 address to a set of public IPv4 address ports, allowing a significant reduction in logging.
FTP ALG (Active and Passive)	Required	FTP clients are supported with inside (private) address and servers with outside (public) addresses. Passive FTP is provided by the basic NAT function. Active FTP is used with the ALG.
RTSP ALG	Required	Remote control protocol for streamers (which use RTP/RTCP or RDT). Our implementation considers the server is located "outside" and clients are "inside". RTSP is used in many streamers like QuickTime or RealNetworks. Both "SETUP" and "SETUP response" are analysed. Based on this information, the translation entry will be created in the table.
ICMP ALG	Required	ALTS for ICMP to translate from IPv6 to IPv4 and the reverse.
SIP ALG	Required	To allow SIP services to transverse the LSN using an ALP.
TFTP		To allow TFTP services to transverse the LSN using an ALP.
PPTP ALG	Required	To allow PPTP services to transverse the LSN using an ALP for use on some implementations of VPNs.
Use of Application SVI's, connecting multiple routing entities (inside/private VRF, outside/public VRF)	Required	NAT should be possible from any routing context (VRF, GRT) to any routing context (VRF, GRT).
Stateful ICMP	Required	Stateful ICMP mappings between inside and outside ICMP identifiers should be supported.
Thresholds	Required	Configurable thresholds using watermarks should be supported to monitor the resources on the NAT64
QoS translation	Required	For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.
Chassis NAT Clustering	Optional	Clustering of NAT64's to allow for inter-chassis resiliency.

Functional Name	Requirement	Description
Logging via Netflow V9/IPFIX	Required	Netflow v9 support for logging of the translation records. Logging of the translation records can be mandated by for Lawful Intercept. The Netflow uses binary format and hence requires software to parse and present the translation records.
Logging via Syslog	Required	Syslog support for logging of the translation records as an alternative to Netflow. Syslog uses ASCII format and hence can be read by users. However, the log data volume is higher in Syslog than Netflow.
Destination based Logging	Required	Destination Based Logging will generate one record per session, including the destination port and address. NFv9 templates including destination IP address and destination port will be used.
Base Logging Fields	Required	NAT64 logs the following information when a translation entry is created: Inside instance ID Outside instance ID Inside IPv6 Address Inside Port Outside IPv4 Address Outside Port Protocol Start Time Stop Time
Radius Logging	Required	Logging using Radius accounting messages on a per block allocation requirement.
XML l	Optional	Logging using XML files.
Static port forwarding (up to 6K static forward entries per npu)	Required	Static port forwarding configures a fixed, private (internal) IPv6 address and port that are associated with a particular subscriber while the NAT64 allocates a free public IP address and port. Therefore, the inside IPv6 address and port are associated to a free outside IP address and port.
Static port forwarding 1:1 active/standby	Required	Static Port Forwarding mapping will be kept constant with two npu cards in Active/Standby mode.

Performance Requirement		Description
24 NAT64 instances per npu Card	Required	The ability to stack transition instances on top of one another.
40m Translations (per npu) - bloack assignment	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary.
Minimum 50 Gbps bi-directional throughput per npu (Inside-to-outside VRF + Outside-to-inside VRF - GRT to GRT) with IMIX traffic	Required	Per half slot throughput requirement.
1M+ primary connections per second setup rate	Required	Primary sessions, port block assignments, per NPU.
500k users per NPU	Required	Per Network Processing Unit a minimum of 500 000 users should be serviceable at any single moment in time.
Latency	Required	Latency is between 40 and 200 micro seconds (µs).
6 npu Cards per chassis	Required	Min chassis requirement.
IRB/SVI support	Required	Integrated Routing and Bridging / Virtual Interfaces (L3 interface for Bridge Domain).
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 256 k NAT users at the same time.

6.4.2 Main NAT64 RFC References

Historical documents: RFC 4787 [94], RFC 5382 [106].

Core protocol: RFC 6052 [119], RFC 6145 [124], RFC 6146 [125].

Deterministic NAT: [9].

6.4.3 CORE Device (LSN/CGN)

NAT64 is an implementation of an IPv4/IPv6 transition protocol, and is the CORE node that all CPE/CEs running NAT64 LSN (Large Scale NAT) connect to. It is the translation from the IPv6 to IPv4 and the originating IPv4 NAT device egressing traffic towards the destination on IPv4 after translation.

This clause concerns itself solely with the NAT64 LSN specific required functionality to allow the technology to function completely and fully. The specification laid out below allows NAT64 LSN technology to be deployed on the NAT64 as a non-service deprecating form comparative to a native private IPv4 delivery to any given customer within a Carrier Topology or potentially otherwise. The present document is however focuses on Cable requirements in functionality but can stretch further if required.

Note that any given NAT64 implementation should be complaint with the associative feature for deployment within a Cable ISP network.

6.4.4 Hardware / Software Requirements

6.4.4.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU in a full stateful fashion with no loss of packets or latency beyond 1 ms. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1 ms failover maximum for any given CPE connectivity of session

RFC Reference: N/A.

6.4.4.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NAT64 has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the NAT64 can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the NAT64 LSN NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.4.4.3 Feature: RP / RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or NAT64 LSN transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within an NAT64 has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.4.4.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given NAT64 has to be shared giving the appearance of a single node on the network.

Function: The NAT64 has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single address per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.4.4.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given NAT64 has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given NAT64 to see all or part of the traffic coming into the node. This can be determined through a single NAT64 address, so NAT64 GW assigned or NAT group assigned under multiple GW addresses. Hashing SHOULD be used to determine the load balancing between all processing/hardware forwarding points with that NAT group or deterministic specification.

RFC Reference: N/A.

6.4.5 Performance Requirements

6.4.5.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1 ms for any given function of the NAT64.

Function: Within the remit of traffic requiring NAT64 LSN function, translation and NAT function, the NAT64 has to perform its function from ingress interface to egress interface in a time measured no higher than 1 ms. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions.

RFC Reference: N/A.

6.4.5.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.4.5.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the NAT64 Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 gig is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.4.5.4 Feature: Min Sessions per NAT64 Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis.

RFC Reference: N/A.

6.4.5.5 Feature: Minimum Customer IPv4 Addresses per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Defines the minimum amount of IPv4 addresses able to be configured per chassis.

Function: If we define the IPv4 address space requirement based on session and block allocation this would be 60 ports per subscriber as a block assignment comparative to 100k customers online allowing the integer of 1 024. Or if we calculate backwards this would be 100 000 customers divided by 60 per IP (although this changes per deterministic NAT requirement) allowing for a max utilization of 1 000 ports per subscriber and thus giving us 1 666-address requirement if we have maximum assignment utilization. So we can round this up to 2 040-address minimum address pool requirement.

RFC Reference: N/A.

6.4.5.6 Feature: Min Customer CPEs per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus IPv6 CPE addresses that can be assigned from a single Chassis or connected.

Function: Due to the scaling of any given NAT64 solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 CPE addresses. This is solely a minimum and the number may change with the introduction of Private v4 and public v6 Si-ID assignment.

RFC Reference: N/A.

6.4.6 Feature Specifications

6.4.6.1 Feature: Source Ipv6/Client-Customer ID

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For general identification and data retention purposes, tunnel identifiers has to be uniquely associated with a single CPE. The identifier has to be the IPv6 address of the CPE.

Function: Unique Si-ID Ipv6 CPE addresses has to be assigned to the CPE.

RFC Reference: N/A.

6.4.6.2 Feature: ICMPv6: Neighbour Discovery and Stateless Auto-Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the discovery of an IPv6 neighbouring device and the ability for it to assign its own address.

Function: ICMPv6 based NDP and stateless auto-configuration has to be an option for the NAT64 to allow for connectivity within NDP for on-link nodes.

RFC Reference: [40**]**, [41], [42], [43], [44].

6.4.6.3 Feature: IPv6 Global Unicast Address Format

Premise: The NAT64 has to be able to forward and be configured with an IPv6 Unicast structured address.

Function: In any given IPv6 topology unicast address structures has to be adhered to with no proprietary configurations. All present IPv6 unicast structures mentioned within RFC 3513 [62] has to be compatible within the vendor deployment to allow for standardization and monitoring.

RFC Reference: [62**]**.

6.4.6.4 Feature: NAT64 LSN timers

The premise, function and related RFC reference relating to this feature is as given below.

Premise: These has to allow for all below NAT64 LSN effecting timers to be configured per the RFC.

Function: Listed below are the timers that SHOULD be configured for the NAT64, note that these SHOULD match the CPE where possible. The timers are not fixed as it is deployment dependent the list you below are the suggested values only and the required fields for any particular vendor.

i) icmp-query --- min 1

ii) sip --- min 2

iii) tcp-established --- hrs 30 min 0

iv) tcp-syn --- sec 30

v) no tcp-time-wait --- min 3

v) tcp-transitory --- min 4

vii) udp --- min 5

viii) udp-initial --- sec 15

ix) udp-dns --- sec 15

RFC Reference: No single defined RFC.

6.4.6.5 Feature: NAT64 DNS64

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DNS64 implementation to allow for "false" aaaa records with embedded a record to be returned if no aaaa record is found.

Function: NAT64 is based upon DNS initialization and requires a specific piece of DNS code, noted as DNS64, this SHOULD place an IPv4 A record into the last 32 bits of the full 128 bit IPv6 record response to the client with the first 96 bits denoted as the NAT64 LSN "prefix". This is only required if there is no AAAA record (IPv6 address) for the URL requested.

RFC Reference: No single defined RFC.

6.4.6.6 Feature: Thresholds and Watermarks

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are two types of thresholds/watermarks, reactive based and logging based, both are required.

Function: Listed below are minimum required thresholds for NAT64 LSN and SHOULD be adhere to validate the utilization on an event basis due to resource allocation that is required for CGN:

1) Watermarks based on Subscriber Thresholds for the node
2) Watermarks based NPU total resources
3) Watermarks based on Port Allocation Thresholds per Subscriber
4) Watermarks based on Port Allocation Thresholds per NPU
5) Watermarks based on IP Allocation Utilization Thresholds
6) Watermarking Event Reporting through SNMP
7) Watermarking Event Reporting through Syslog
8) Watermarking for dynamic port allocation over the top of Deterministic NAT

RFC Reference: N/A.

6.4.6.7 Feature: Port Block Allocation per IP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Port block allocation is configurable to allow for any ratio assignment per IP.

Function: An assignment of ports based on block allocations for any numeric value with automatic assignment of a single IP determined by the black allocation.

RFC Reference: N/A.

6.4.6.8 Feature: Deterministic NAT / Deterministic Dynamic thresholds

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Deterministic NAT policy and configuration has to be adhered to for NAT64s not only for port logging requirements for data retention but also for processing and functional improvement.

Function: DetNAT should be used in the following manner. It should be configurable to allow for any n-Ports to be allocated to the port block and any number of port block assignments to be dynamic or static per CPE address. It has to also hold a dynamic threshold to further the ability for IP to port ratios by allowing the dynamic section of the port allocation for any single or cluster of CPE nodes to use 1-5 block allocations and to allow for singular oversubscription per CPE based on a threshold. So if the block allocation reaches over 60 % (configurable) of the full amount of ports, 65 000 per IP, a restriction is placed on further port allocations. For example until the 60 % threshold is reached the Subscriber CPE is allowed up to 3 000 ports but once the port threshold of 605 is reached this then drops to 500 ports for all users. No old connections are dropped but no new ones can be added. This allows for controlled oversubscription with heavy port users to over allocate while there are resources.

RFC Reference: [9**]**.

From section 2 in [9]:

> "The algorithm is not designed to retrieve an internal host among those sharing the same internal IP address (e.g., in a NAT64 LSN context, only an IPv6 address/prefix can be retrieved using the algorithm while the internal IPv4 address used for the encapsulated IPv4 datagram is lost).
> Several address assignment algorithms are possible. Using predefined algorithms, such as those that follow, simplifies the process of reversing the algorithm when needed. However, the CGN MAY support additional algorithms. Also, the CGN is not required to support all algorithms described below. Subscribers could be restricted to ports from a single IPv4 address, or could be allocated ports across all addresses in a pool, for example. The following algorithms and corresponding values of A are as follow:
>
> 0. Sequential (e.g. the first block goes to address 1, the second block to address 2, etc.)
> 1. Staggered (e.g. for every n between 0 and ((65536-R)/(C+D))-1 , address 1 receives ports n*C+R, address 2 receives ports (1+n)*C+R, etc.)
> 2. Round robin (e.g. the subscriber receives the same port number across a pool of external IP addresses. If the subscriber is to be assigned more ports than there are in the external IP pool, the subscriber receives the next highest port across the IP pool, and so on. Thus, if there are 10 IP addresses in a pool and a subscriber is assigned 1000 ports, the subscriber would receive a range such as ports 2000-2099 across all 10 external IP addresses).
> 3. Interlaced horizontally (e.g. each address receives every Cth port spread across a pool of external IP addresses).
> 4. Cryptographically random port assignment (Section 2.2 of RFC6431 [RFC6431]). If this algorithm is used, the Service Provider needs to retain the keying material and specific cryptographic function to support reversibility.
> 5. Vendor-specific. Other vendor-specific algorithms may also be supported.
>
> The assigned range of ports MAY also be used when translating ICMP requests (when re-writing the Identifier field).
> The CGN then reserves ports as follows:
>
> 1. The CGN removes reserved ports (R) from the port candidate list (e.g., 0-1023 for TCP and UDP). At a minimum, the CGN SHOULD remove system ports (RFC6335) [RFC6335] from the port candidate list reserved for deterministic assignment.
> 2. The CGN calculates the total compression ratio (C+D), and allocates 1/(C+D) of the available ports to each internal IP address. Specific port allocation is determined by the algorithm (A) configured on the CGN. Any remaining ports are allocated to the dynamic pool. Note: Setting D to 0 disables the dynamic pool. This option
> 3. eliminates the need for per-subscriber logging at the expense of limiting the number of concurrent connections that 'power users' can initiate.

4. When a subscriber initiates a connection, the CGN creates a translation mapping between the subscriber's inside local IP address/port and the CGN outside global IP address/port. The CGN has to use one of the ports allocated in step 2 for the translation as long as such ports are available. The CGN SHOULD allocate ports randomly within the port range assigned by the deterministic algorithm. This is to increase subscriber privacy. The CGN has to use the preallocated port range from step 2 for Port Control Protocol (PCP, [I-D.ietf-pcp-base]) reservations as long as such ports are available. While the CGN maintains its mapping table, it need not generate a log entry for translation mappings created in this step.
5. If D>0, the CGN will have a pool of ports left for dynamic assignment. If a subscriber uses more than the range of ports allocated in step 2 (but fewer than the configured maximum ports M), the CGN assigns a block of ports from the dynamic assignment range for such a connection or for PCP reservations. The CGN
6. has to log dynamically assigned port blocks to facilitate subscriber-to-address mapping. The CGN SHOULD manage dynamic ports as described in [I-D.tsou-behave-natx4-log-reduction].
7. Configuration of reserved ports (e.g., system ports) is left to operator configuration.

Thus, the CGN will maintain translation-mapping information for all Connections within its internal translation tables; however, it only needs to externally log translations for dynamically-assigned ports."

6.4.6.9 Feature: IP Ranges per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP ranges should be able to be assigned per NAT group or across multiple NAT groups to allow for full scaling and shared IP resource.

Function: For further study.

RFC Reference: N/A.

6.4.6.10 Feature: NAT Grouping resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT Grouping SHOULD be cross NPU based.

Function: NAT groups SHOULD be able to use the same NPUs and IP allocations regardless of their GW addresses. This allows for complete shared resources across multiple NAT groups. The function states that the NPUs can be grouped in any manner for active standby numbers as the user/administrator wishes. Thus NAT group 1 can use NPU 1,2,3&4, with 1,2&3 as active and NAT group 2 can use NPU 1,2,3,4,5&6 with 1,2&3 as standby. This promotes scaling and no restriction if there are particular requirements to share a resource.

RFC Reference: N/A.

6.4.6.11 Feature: Virtual interface per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: An NAT64 has to be able to assign a single virtual interface with up to 8 NAT64 LSN GW addresses for any given NAT64 LSN instance on the node.

Function: The virtual interfaces has to be able to be assigned to any number, and all, NPUs within the chassis to allow complete shared resources. So a single or multiple NAT64s acting as the single destination interface for all resources and CPE nodes using the NAT64.

RFC Reference: N/A.

6.4.6.12 Feature: NPU to Interface throughput ratio

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To allow balanced Bandwidth assignment across all virtual or a single virtual interface matching or exceeding the physical bandwidth of the node.

Function: To prevent over buffering, dropped packets and general resource issues the NAT64 has to be able to consider a physical topology that matches a balanced logical bandwidth ratio between the interfaces and the NPU throughput.

RFC Reference: N/A.

6.4.6.13 Feature: NAT64 Address (es)

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT64 Addressing

Function: Listed below:

1) Has to Allow one or multiple addresses as GWs for either one or multiple NAT groups.
2) IPv4 and IPv6 addressing SHOULD be placed on separate interfaces.
3) The NAT64 SHOULD be able to act as the DHPv6 server for the CPE if required.
4) SI-IDs should be able to comprise of ONLY the IPv6 CPE address.
5) The IPv6 Public address assigned to the CPE has to be unique.
6) The IPv6 addresses configured for NAT64 LSN Gateways has to be Anycast compatible.
7) The NAT64 has to allow multiple NAT group per IP address and thus per Virtual interface end point.

RFC Reference: N/A.

6.4.6.14 Feature: Anycast Gateway Address

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Anycast usage for an NAT64 gateway or DHCPv6 address has to be accepted within the configuration.

Function: One or many Anycast addresses has to be configurable for the NAT64 GWs regardless of any other resource and for the DHCPv6 server address. **Note** that using Anycast might cause failover issues if a path in your network changes and flapping could prevent session stability if clustering and cluster forwarding is not enabled.

RFC Reference: N/A.

6.4.6.15 Feature: Multiple Source Prefixes per NAT64 Interface

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The premise is that a definition of sourced addresses SHOULD be configured to allow forced routing to a specific GW address on the NAT64.

Function: A user/administrator SHOULD be able to place a source IP statement within a NAT group allowing the CPE packets originating from within that subnet to be routed to a specific NAT group or NAT64 interface. His allows the delineation of different service IPs instead of relying on separation of traffic based on DNS record responses and thus specific rules and resource allocations for that particular sourcing subnet.

RFC Reference: N/A.

6.4.6.16 Feature: NAT64 Address Withdrawal

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NAT64 has to b withdrawn under the circumstances below.

Function: Listed below:

1) The NAT64 is not reachable within the routing protocol being used or cannot reach its neighbour's or peers
2) The NAT group goes down
3) The NAT64 LSN process dies
4) NPUs available cannot hold the present capacity
5) Shutdown or part or all of the service
6) External Routing to the internet is down for a fixed period of time (demands NAT64 clustering to determine reachability from a redundant NAT64)

RFC Reference: N/A.

6.4.6.17 Feature: Chassis Clustering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To Cluster two or more NAT64s for session state failover.

Function: If a single NAT64 goes down it has to do the following:

- Remove its NAT64 address as a reachable entity within the routing protocol being used
- Forward any remaining packets to the secondary NAT64
- Validate NAT cache state between itself and its redundant NAT64

The clusters has to sync on the following items:

- The TCP and UDP sessions presently being used to for NAT64 LSN
- The TCP not used for n amount of time configurable
- The present capacity of interface and NPU to allow for structural syncing
- Sync between all clustered NAT64s due to best path attributes for the Anycast / Unicast Addresses used (also dependent on secondary addressing on the CPE)

RFC Reference: N/A.

6.4.6.18 Feature: Multiple Transition Technology Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT44, NAT64 and DS-Lite LSN has to function on the same platform to allow for uniformity within European markets.

Function: All three technologies should be able to function and share resources dynamically on any given platform.

RFC Reference: N/A.

6.4.6.19 Feature: NPU / chassis Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Buffering requirements for the NAT64 for transit traffic covering queues and buffer timers.

Function: Listed below:

1) Queue sizing per flow
2) Queue sizing per port block
3) Queue sizing per NPU
4) Queue sizing per ingress interface
5) Buffer sizing per flow
6) Buffer sizing per port block
7) Buffer sizing per NPU
8) Buffer sizing per ingress interface
9) Reordering buffers and out-of-order caching

RFC Reference: N/A.

6.4.6.20 Feature: NAT64 LSN Fragmentation and Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Fragmentation has to be placed on the ingress interface pre-encapsulation on downstream on IPv4. Upstream fragmentation is not required.

Function: All fragmentation on NAT64 LSN on the NAT64 has to be on IPv4 in downstream direction. This prevents IPv6 fragmentation requirements and overlay code. IPv6 inherently does not have fragmentation built into it and thus can cause major performance issues on an NAT64 forwarding plane or even be placed in software.

Reassembly has to only be used when the NAT64 receives noted IPv6 fragmented packets incoming upstream from the CPE and thus requires full buffering/re-ordering to prevent timeouts and drops.

Pre-fragmented packets has to be re-ordered before being sent on and again requires full buffer queues and wait timers exceeding 3 ms.

RFC Reference: N/A.

6.4.6.21 Feature: Stateful ICMP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Stateful ICMP is the act of holding state for ICMP for messages sent/originated from the NAT64 to any know destination on a single flow. This can be used downstream as well as upstream.

Function: A maximum of 3 ICMP packets to big messages has to be sent to single destination for a single flow. This prevents flooding and over compensation. This should ONLY be used when the DF bit is set and should be tied into the fragmentation process of NAT64 LSN packets.

RFC Reference: N/A.

6.4.6.22 Feature: Port Reservation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Reserving a block of public ports for each private IPv4 client reduces the scaling requirements of per-connection port-mapping.

Function: The NAT64 has to be capable of bulk/block port reservation.

RFC Reference: [136**]**.

From section 2.3 in [136]:

> "For operators who desire to aggregate the logs, they can configure the NAT64 to preallocate a range of ports to each CPE. This range of ports will be used in the NAT44 function, and the NAT64 will create one log entry for the whole port range. This aggregation can significantly reduce the log size for source-specific logging."

6.4.7 Monitoring and Management

6.4.7.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.4.7.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.4.7.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.4.7.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.4.7.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv6.

RFC Reference: [76**]**.

6.4.7.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18**]**.

6.4.7.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.4.7.8 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26**]**.

6.4.7.9 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.4.7.10 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60**]**.

6.4.7.11 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.4.7.12 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.4.7.13 Feature: Remote Access - TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.4.8 NAT64 CPE requirements

Requirements on the host are determined by the available connectivity in the Home Network. The Home Network may be dual stack or IPv4 only or IPv6 only. As a transitional functional requirement the CPE should play NO part in the NAT64 deployment only a supporting role for connectivity with only a single exception being handing a separate DNS handoff to the clients from the non-NAT64 home/locations/services.

The following requirements apply to the CPE Router:

- WAN interface facing the Access Network:
  - Request IPv6 address via DHCPv6
- LAN interface facing the Home Network:
  - IPv6 addressing
  - DHCPv6 server or SLAAC for LAN addressing
- CPE router:
  - PCP

In addition to the requirements listed above, the CPE Router should support the following features on its interfaces facing the Home Network (wired, wireless):

- IPv6 stateful packet firewall (5-tuple filters), enabled by default
- Recursive DNS server option

For purposes of a unified user experience some requirements apply to the user interface:

- Web-UI should be accessible on the LAN IP interface
- Login is initially presented in the format of requesting username and password
- Localization of UI for different languages should be supported
- A Status tab should contain subpages on Software, Connection, Security, Diagnostics
- A Router Basic tab should contain subpages on WAN Setup, LAN&DHCP Server, Backup
- A Router Advanced tab should contain subpages on Option, IP Filtering, MAC Filtering, Port Filtering, Forwarding, Port Triggers, DMZ Host
- A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log
- A Parental Control tab should contain subpages on User Setup, Basic Setup, Content Filter, ToD Filter
- A Wireless tab should contain subpages on Radio, Security, Advanced, Access Control

An MTA tab should contain subpages on Status, DHCP, QoS, Provisioning, Event Log.

6.4.8.1 Feature: WAN Interface Address Requirements

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to request an IP address for the interface facing the access network.

Function: The CPE device has to request an IP address via the use of DHCPv6.

RFC Reference: N/A.

6.4.8.2 Feature: Local Area Networking Interfaces

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv6 device addressing.

Function: The CPE device has to implement IPv6 addressing on any locally connected interface regardless of interface type.

RFC Reference: N/A.

6.4.8.3 Feature: IP address provision

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to provide IPv6 addresses to locally connected clients.

Function: The CPE device has to either implement a local DHCPv6 server or allow for the use of SLAAC for LAN addressing requirements.

RFC Reference: N/A.

6.4.8.4 Feature: Integrated IPv6 Statefull Firewall

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a state full firewall function.

Function: The CPE has to be capable of implementing filters based upon a 5-tuple consisting of the following:

1) Source IP address

2) Destination IP address

3) Source Port number

4) Destination Port number

5) Protocol in use

RFC Reference: N/A.

6.4.8.5 Feature: Recursive DNS Server

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a recursive DNS server.

Function: The CPE device has to implement a recursive DNS server which will contact DNS resources to provide an authorative DNS response in request to a resolution request for a given domain.

RFC Reference: N/A.

6.4.8.6 Feature: User Interface Provision

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE has to implement a user accessible web interface on any implemented Local Area Network interface to aid the configuration of the device and to aid in provision of diagnostics information.

Function: The CPE device has to implement a user accessible web interface on port 80.

RFC Reference: N/A.

6.4.8.7 Feature: Access Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement access control to the local web interface.

Function: Access to the web interface has to be controlled by the use of a specified username and password.

RFC Reference: N/A.

6.4.8.8 Feature: Device Localization

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device user interface has to be configurable for Multiple languages.

Function: The language displayed on the user accessible web interface has to be configurable to support display in multiple languages.

RFC Reference: N/A.

6.4.8.9 Feature: CPE Device Status Indication

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE Device has to contain a user interface tab which provides access to the following item.

Function: Listed below:

1) Connection Status

2) Security

3) Diagnostics

4) Software

RFC Reference: N/A.

6.4.8.10 Feature: Router Basic Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Wide Area network Configuration

2) Local Area Network Configuration

3) DHCP Server Configuration

4) Device Configuration Backup

RFC Reference: N/A.

6.4.8.11 Feature: Advanced router control and configuration.

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Advanced Options

2) IP Filtering

3) MAC Filtering

4) Port Filtering

5) Port Forwarding

6) Port Triggers

7) DMZ Host Selection

RFC Reference: N/A.

6.4.8.12 Feature: The CPE device has to implement a Firewall user interface with user accessible configuration control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log.

Function: Listed below:

1) Web Filtering

2) Local Logging

3) Remote Logging

RFC Reference: N/A.

6.4.8.13 Feature: User Interface - Parental Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a specific user accessible tab which provides access to the following functions.

Function: Listed below:

1) User Setup, Basic Setup, Content Filter, ToD Filter
2) Basic Setup
3) Content Filtering
4) Time of Day Access Control

RFC Reference: N/A.

6.4.8.14 Feature: Wireless Status and configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Where the CPE device implements a wireless interface a specific user accessible tab has to be made available.

Function: The Wireless configuration tab has to include the following items:

1) Radio Interface
2) Security
3) Advanced Functionality
4) Access Control

RFC Reference: N/A.

6.4.8.15 Feature: CPE Configuration Options

The premise, function and related RFC reference relating to this feature is as given below.

Premise: **User Interface -**Where the CPE device implements an integrated Multimedia Terminal Adaptor a specific user configurable tab has to be made available.

Function: The MTA tab has to include the following items:

1) Status of the integrated MTA functionality
2) DHCP status, including any additional interfaces configured for the use of the MTA device
3) Quality of Service
4) Device provisioning
5) Event Logging

RFC Reference: N/A.

6.4.9 Technical Viability

There is a number of well-known issues with NAT64 which present topics for further development:

- The main issue with this technology is the fact that it does not operate on IPv4-only devices and that applications need to run via DNS. Applications like P2P do not consult the DNS and, thus, fail to operate since the NAT64 GW address cannot be discovered.

- NAT64 requires ALGs (ALPs to be exact in most cases) which require the NAT64 device to do some form of intelligent vicissitudes to the transit packets bearing the risk of drop in performance and limits in functionality. The risk of service-deprecation is minimized if all functionality is included nativel in the NAT64 device. In addition, any application that communicates an IPv4 address in its downstream payload or a non-local IPv6 address in the upstream payload will not work flawlessly with NAT64.

- With NAT64, all functional intelligence is located in the NAT64 device. Thus, functionality that requires a public address in the local network has to be performed in the NAT64 device itself. An example of this is PCP. In NAT64 is placed on the LSN due to UPnP 1 and 2 requirements.

- To predict scaling requirements comparative to IPv6 utilization can be misleading. This is not specific to NAT64 but occurs similarly in other transition technologies. The issue is that it is largely unknown what portion of the traffic originated by a certain client will be IPv6. Thus, the system has to be scaled for a constant maximum of all clients and designs cannot adapt based on exact traffic predictions.

- NAT64 is a translation technology and, thus, suffers from MTU requirements beyond the norm. This means that MSS clamping, IPv4 fragmentation and IPv6 fragmentation are required in order to avoid ICMP blocking. Fragmentation resends and general PMTU control can have performance effects on customer services. Therefore, there is a major requirement to optimize implementations such that they do not cause a large impact on current RTTs and node latency.

- The requirement for Data Retention capacity is rather large and, thus, making the capacity available may become a major issue for the receiving logging server. Normally, on public addresses there is a single entry or two per week depending on lease times and utilization. Within NAT64, each time an individual subscriber receives a port and IP address assignment, the event has to be logged. This can amount to 200 assignments a day with a start and stop time resulting in 400 logs per user per day. Assuming that an MSO could have 10 % of its customer base online at one time, that would mean a huge amount of logs.

On the other hand NAT64 has some advantages over NAT44 / CGN. NAT64 allows for a smooth migration towards an IPv6 only operator infrastructure. This includes IPv6 only access, aggregation and core network as well as an IPv6 only BNG. In case of a cable operator the BNG is implemented on the CMTS. In case of other access technologies (e.g. GPON) the BNG is implemented on a dedicated platform. NAT64 also removes the need for IPv4 address management and as such a DHCPv4 server is not required anymore.

As a NAT64 only handles IPv4 traffic, the capacity requirements per customer on the NAT64 will likely decrease over time. The reason is that more and more internet traffic will be on IPv6. This might allow for operator growth over time without the need for large investments on the NAT64.

6.5 NAT44

NAT44 shares a single IPv4 address, through port allocation, among a large number of customers. This is done by using CGN, which primarily pulls the address allocation to a more centralized NAT in the service provider network. CGN (sometimes known as Large Scale NAT or LSN) is a highly scalable NAT placed between the customer premises equipment (CPE) and the core of the network that implements.

NAT44 did not get a great deal of attention over the last few years until transition technologies started to be considered once again as the IPv4 depletion lifted above the horizon. It is now seen as a possible implementation method for reserved use of IPv4 public addresses in the MSO space and although has many drawbacks adding NAT444 to home connectivity as an example, it can make the transition to IPv6 easier in some respects.

6.5.1 NAT44 Technical Summary

NAT44 is a technology which translates an IPv4 private address to an IPv4 public address. Private (RFC 1918 [30]) IPv4 addresses are used between the CPE and LSN/CGN-GW. RFC 1918 [30] addresses can be reused in different networks, they are not globally unique, and therefore can cause administrative and scaling issues on an MSO network. There addresses are not publicly routable but SHALL be translated in order to reach the internet.

NAT44 is the basis of DS-Lite in many respects but it uses IPv4 transit instead of an IPv6 tunnel. It requires in a carrier grade implementation similar requirements to DS-Lite including PCP, DeNAT (static port assignments) and NAT Cache Clustering.

NAT has one or more globally unique IPv4 addresses. and as a packet passes from its inside or private interface to its outside or public interface, NAT replaces the packet's private IPv4 address with one of its public IPv4 addresses. The NAT "remembers" which inside device the packet came from by mapping the inside address to the outside address.

NAT44 has numerous methodologies, we discuss only NAT44 under the guise of port range allocation per IP based on the providers CORE network, not on the CPE only, and today is a mixture of the technologies below.

BASIC NAT44

- Source NAT pool with address range
- Translation type is source static

NAPT44

- Source NAT pool with address-range/prefix and port range
- Translation type is source dynamic

DOUBLE NAT44

- Source NAT pool with address range
- Destination pool with address range
- Translation type is source dynamic, destination static

DYNAMIC NAT44

- Source NAT pool with address range
- Translation type is source dynamic

6.5.2 Technical Specifications

Table 8 provides a summary of the technical requirements for the implementation of NAT44.

Table 8: Summary for Implementation of NAT44

Functional Name	Requirement	Description
NAT44 - RFC 4787 [94] (UDP)	Required	Compliance with NAT behaviour according to RFC 4787 [94] for UDP.
NAT44 - RFC 5382 [106] (TCP)	Required	Compliance with NAT behaviour according to RFC 5382 [106] for TCP.
NAT44 - RFC 5508 [108] (ICMP)	Required	Compliance with NAT behaviour according to RFC 5508 [108] for ICMP.
NAT44 - Network Address and Port Mapping - Endpoint Independent Mapping	Required	For two flows for a common inside source IPv4 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation.
NAT44 - Translation Filtering - Endpoint Independent Filtering	Required	A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.
NAT44 - Paired IP Address Assignment	Required	Translation to External IPv4 address is done in a paired fashion. A given Inside address is always translated to the same External IPv4 address.
NAT44 - Port Parity Assignment	Required	Odd internal port is translated to odd external port and even internal port is translated to even external port.
NAT44 - Hair-pinning	Required	Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.
NAT44 - 1:1 IP Mapping	Required	Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).
NAT44 - Outside-Service-App mapping for inside-VRF	Required	Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.
NAT44 - Port Limit configuration	Required	A maximum amount of ports can be configured for every private IPv4 source address.
NAT44 - Per-Protocol Timeout configuration	Required	Timeout of mappings is critical and removing a NAT mapping at the appropriate time maximizes the shared port resource.
NAT44 - Destination-based Timeout configuration	Required	Timeout for sessions on a per destination, per flow, basis.
NAT44 - Dynamic Port Range start configuration	Required	The start port for dynamic port ranges should be configurable to allow for a range of ports for static port mappings.
NAT44 - Software Load Balancing	Required	NAT44 Inside to Outside hashing performed on the Source IPv4 (32 bits) user address. NAT44 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Use of Application SVI's, connecting multiple routing entities (inside/private VRF, outside/public VRF)	Required	NAT should be possible from any routing context (VRF, GRT) to any routing context (VRF, GRT).
Stateful ICMP	Required	Stateful ICMP mappings between inside and outside ICMP identifiers should be supported.
Thresholds	Required	Configurable thresholds using watermarks should be supported to monitor the resources on the NAT44 CGN.
Chassis NAT Clustering	Required	Clustering of NAT44 CGN's to allow for inter-chassis resiliency.
Logging via Netflow V9/IPFIX	Required	Netflow v9 support for logging of the translation records. Logging of the translation records can be mandated by for Lawful Intercept. The Netflow uses binary format and hence requires software to parse and present the translation records.
Logging via Syslog	Required	Syslog support for logging of the translation records as an alternative to Netflow. Syslog uses ASCII format and hence can be read by users. However, the log data volume is higher in Syslog than Netflow.

Functional Name	Requirement	Description
Destination based Logging	Required	Destination Based Logging will generate one record per session, including the destination port and address. NFv9 templates including destination IP address and destination port will be used.
Base Logging Fields	Required	NAT44 CGN logs the following information when a translation entry is created: Inside instance ID Outside instance ID Inside IPv4 Address Inside Port Outside IPv4 Address Outside Port Protocol Start Time Stop Time
Radius Logging	Required	Logging using Radius accounting messages for transition flows
XML l	Optional	Logging using XML files
Port Allocation	Required	In order to reduce the volume of data generated by the NAT device (logging creation and deletion data), bulk port allocation can be enabled. When bulk port allocation is enabled and when a subscriber creates the first session, a number of contiguous outside ports are pre-allocated. A bulk allocation message is logged indicating this allocation.
Deterministic NAT	Required	Deterministic NAT maps every inside private IPv4 address to a deterministic outside IP address and port-range. Mappings are created at configuration time. Dynamic port blocks should be allowed on top of the deterministic port-block.
FTP ALG (Active and Passive)	Required	FTP clients are supported with inside (private) address and servers with outside (public) addresses. Passive FTP is provided by the basic NAT function. Active FTP is used with the ALG
RTSP ALG	Required	Remote control protocol for streamers (which use RTP/RTCP or RDT). Our implementation considers the server is located "outside" and clients are "inside". RTSP is used in many streamers like QuickTime or RealNetworks. Both "SETUP" and "SETUP response" are analysed. Based on this information, the translation entry will be created in the table.
SIP ALG	Required	To allow SIP based application traffic to traverse the NAT
Static port forwarding (up to 6K static forward entries per npu)	Required	Static port forwarding configures a fixed, private (internal) IP address and port that are associated with a particular subscriber while CGv6 allocates a free public IP address and port. Therefore, the inside IP address and port are associated to a free outside IP address and port.
Static port forwarding 1:1 active/standby	Required	Static Port Forwarding mapping will be kept constant with two npu cards in Active/Standby mode.
PCP	Required	Support for PCP to allocate static port bindings.
Performance Requirement		**Description**
24 NAT instances per npu Card	Required	The ability to stack transition instances on top of one another.
40m Translations (per npu) - bloack assignment	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary.
Minimum 50 Gbps bi-directional throughput per npu (Inside-to-outside VRF + Outside-to-inside VRF - GRT to GRT) with IMIX traffic	Required	Per half slot throughput requirement.
1M+ primary connections per second setup rate	Required	Primary sessions, port block assignments, per NPU.
500k users per NPU	Required	Per Network Processing Unit a minimum of 500 000 users should be serviceable at any single moment in time.

Functional Name	Requirement	Description
Latency	Required	Latency is between 40 and 200 micro seconds (µs).
6 npu Cards per chassis	Required	Min chassis requirement.
IRB/SVI support	Required	Integrated Routing and Bridging / Virtual Interfaces (L3 interface for Bridge Domain).
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 256 k NAT users at the same time.

6.5.3 Main NAT44 RFC References

Historical documents: RFC 4787 [94], RFC 5382 [106], RFC 5508 [108].

Deterministic NAT: [9].

PCP: [6], [7], [8], RFC 6877 [134].

6.5.4 CORE Device (LSN/CGN)

NAT44 CGN is an implementation of an IPv4/IPv6 transition protocol, and is the CORE node that all CPE/CEs using private IPv4 addresses connect to via IPv4 routing. It is the originating IPv4 NAT device egressing traffic towards the destination on IPv4.

This section concerns itself solely with the NAT44 CGN specific required functionality to allow the technology to function completely and fully. The specification laid out below allows NAT44 CGN technology to be deployed as a non-service deprecating form comparative to a native private IPv4 delivery to any given customer within a Carrier Topology or potentially otherwise. The present document is however focuses on Cable requirements in functionality but can stretch further if required.

Note that any given NAT44 CGN implementation should be complaint with the associative feature for deployment within a Cable ISP network.

6.5.5 Hardware/Software Requirements

6.5.5.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU in a full stateful fashion with no loss of packets or latency beyond 1 ms. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1ms failover maximum for any given CPE connectivity of session

RFC Reference: N/A.

6.5.5.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NAT44 CGN has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the NAT44 CGN can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the NAT44 CGN NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.5.5.3 Feature: RP / RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or NAT44 CGN transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within a NAT44 CGN has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.5.5.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given NAT44 CGN has to be shared giving the appearance of a single node on the network.

Function: The NAT44 CGN has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single address per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.5.5.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given NAT44 CGN has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given NAT44 CGN to see all or part of the traffic coming into the node. Hashing SHOULD be used to determine the load balancing between all processing/hardware forwarding points with that NAT group or deterministic specification

RFC Reference: N/A.

6.5.5.6 Feature: Routing and MPLS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to route traffic on the inside towards the CPE's and on the outside towards the internet, a NAT44 CGN has to support a rich set of routing and MPLS features.

Function: MP-BGP
BGP community/32 bit AS
MPLS LDP IPv4 natively
ECMP
QoS (IPv4) - classification, priority queuing, etc.
QPPB/SCU/DCU
ACLs/prefix lists/filtering IPv4
CoPP (IPv4/)
IEEE 802.1q
Etherchannel
Ethernet OAM
Policy Based Routing IPv4
IS-IS
Static Routing
OSPFv2
CDP/LLDP
VRRP/HSRP
VLAN mapping/Double Tagging
L3 multicasting/MFIB IPv4
IPv4 forwarding (hardware)
Ethernet technologies
Virtual interfaces IPv4
BFD IPv4
PIM/IGMPv2/v3
CEF/dCEF
Anycast
Route reflection IPv4
Standard IPv4 VPN
DHCP relay IPv4

6.5.6 Performance Requirements

6.5.6.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1ms for any given function of the NAT44 CGN.

Function: Within the remit of traffic requiring NAT function, the NAT44 CGN has to perform its function from ingress interface to egress interface in a time measured no higher than 1ms. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions.

RFC Reference: N/A.

6.5.6.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.5.6.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the NAT44 CGN Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 Gbps is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.5.6.4 Feature: Min Sessions per NAT44 CGN Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis.

Function: Sessions scaling is a major requirement and the session minimum is based on carrier grade requirements.

RFC Reference: N/A.

6.5.6.5 Feature: Minimum Customer IPv4 Addresses per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Defines the minimum amount of IPv4 addresses able to be configured per chassis.

Function: If the IPv4 address space requirement is defined based on session and block allocation this would be 60 ports per subscriber as a block assignment comparative to 100 k customers online allowing the integer of 1 024. Or if we calculate backwards this, would be 100 000 customers divided by 60 per IP (although this changes per deterministic NAT requirement) allowing for a max utilization of 1 000 ports per subscriber and thus giving us 1 666-address requirement if we have maximum assignment utilization. So we can round this up to 2 040-address minimum address pool requirement.

RFC Reference: N/A.

6.5.6.6 Feature: Min Customer CPEs per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus private IPv4 addresses that can be assigned from a single Chassis or connected.

Function: Due to the scaling of any given NAT44 CGN solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 CPE addresses. This is solely a minimum.

RFC Reference: N/A.

6.5.7 Feature Specifications

6.5.7.1 Feature: NAT44 - RFC 4787

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Compliance with NAT behaviour according to RFC 4787 [94] for UDP.

Function: UDP Behavioural Requirements.

RFC Reference: [94**]**.

6.5.7.2 Feature: NAT44 - RFC 5382

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Compliance with NAT behaviour according to RFC 5382 [106] for TCP.

Function: TCP Behavioural Requirements.

RFC Reference: [106**]**.

6.5.7.3 Feature: NAT44 - RFC 5508

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Compliance with NAT behaviour according to RFC 5508 [108] for ICMP.

Function: ICMP Behavioural Compliance.

RFC Reference: [108**]**.

From section 2 in [108].

> "NAT Session - A NAT session is an association between a session as seen in the private realm and a session as seen in the public realm, by virtue of NAT translation. If a session in the private realm were to be represented as (PrivateSrcAddr, PrivateDstAddr, TransportProtocol, PrivateSrcPort, PrivateDstPort) and the same session in the public realm were to be represented as (PublicSrcAddr, PublicDstAddr, TransportProtocol, PublicSrcPort, PublicDstPort), the NAT session would provide the translation glue between the two session representations. NAT sessions in the document are restricted to sessions based on TCP, UDP, and ICMP. In the future, NAT sessions may be extended to be based on other transport protocols such as Stream Control Transmission Protocol (SCTP), UDP-lite, and Datagram Congestion Control Protocol (DCCP).
>
> ICMP Message Classification - Section 3.2.2 of [RFC1122] and Section 4.3.1 of [RFC1812] broadly group ICMP messages into two main categories, namely "ICMP Query" messages and "ICMP Error" messages. All ICMP Error messages listed in RFC 1122 and RFC 1812 contain part of the Internet datagram that elicited the ICMP error. All the ICMP Query messages listed in RFC 1122 and RFC 1812 contain an "Identifier" field, which is referred to in this document as the "Query Identifier". There are however ICMP messages that do not fall into either of these two categories. We refer to them as "Non-QueryError ICMP Messages". All three ICMP message classes are described as follows:
>
> - ICMP Query Messages - ICMP Query messages are characterized by an Identifier field in the ICMP header. The Identifier field used by the ICMP Query messages is also referred to as "Query Identifier" or "Query Id", for short throughout the document. A Query Id is used by Query senders and responders as the equivalent of a TCP/UDP port to identify an ICMP Query session. ICMP Query messages include ICMP messages defined after RFC 1122 or RFC 1812 (for example, Domain Name Request/Reply ICMP messages defined in RFC 1788), as they include request/response pairs and contain an "Identifier" field.

- ICMP Error Messages - ICMP Error messages provide signaling for IP. All ICMP Error messages are characterized by the fact that they embed the original datagram that triggered the ICMP Error message. The original datagram embedded within the ICMP Error payload is also referred to as the "Embedded packet" throughout the document. Unlike ICMP Query messages, ICMP Error messages do not have a Query Id in the ICMP header.

- Non-QueryError ICMP Messages - ICMP messages that do not fall under either of the above two classes are referred to as "Non-QueryError ICMP Messages" throughout the document. For example, Router Discovery ICMP messages [RFC1256] are "request/response" type ICMP messages. However, they are not characterized as ICMP Query messages in this document as they do not have an "Identifier" field within the messages. Likewise, there are other ICMP messages defined in [RFC4065] that do not fall in either of the ICMP Query or ICMP Error message categories, but will be referred to as Non- QueryError ICMP messages.

The reason for categorizing ICMP messages for NAT behavioral properties is that each category has different characteristics used for mapping (i.e., the Query Id and the Embedded datagram), which leaves the Non-QueryError ICMP messages in a separate, distinctive group."

From section 3 in [108].

"This section lists the behavioral requirements for a NAT device when processing ICMP Query packets. The following subsections discuss requirements specific to ICMP Query handling in detail."

From section 3.1 in [108].

"Unless explicitly overridden by local policy, a NAT device SHALL permit ICMP Queries and their associated responses, when the Query is initiated from a private host to the external hosts. ICMP Query mapping by NAT devices is necessary for current ICMP-Query-based applications to work. This entails a NAT device to transparently forward ICMP Query packets initiated from the nodes behind NAT, and the responses to these Query packets in the opposite direction. As specified in [NAT-TRAD], this requires translating the IP header. A NAPT device further translates the ICMP Query Id and the associated checksum in the ICMP header prior to forwarding.

NAT mapping of ICMP Query Identifiers SHOULD be external-host independent. Say, an internal host A sent an ICMP Query out to an external host B using Query Id X. And, say, the NAT assigned this an external mapping of Query Id X' on the NAT's public address. If host A reused the Query Id X to send ICMP Queries to the same or different external host, the NAT device SHOULD reuse the same Query Id mapping (i.e., map the private host's Query Id X to Query Id X' on NAT's public IP address) instead of assigning a different mapping. This is similar to the "endpoint independent mapping" requirement specified in the TCP and UDP requirement documents [BEH-UDP], [BEH-TCP]."

6.5.7.4 Feature: NAT44 - Network Address and Port Mapping - Endpoint Independent Mapping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For two flows for a common inside source IPv4 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation.

Function: This feature aggregates translation flows allowing a full cone deployment with no delination of the same flow type and destination.

RFC Reference: N/A.

6.5.7.5 Feature: NAT44 - Translation Filtering - Endpoint Independent Filtering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.

RFC Reference: N/A.

6.5.7.6 Feature: NAT44 - Paired IP Address Assignment

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Translation to External IPv4 address is done in a paired fashion. A given Inside address is always translated to the same External IPv4 address.

RFC Reference: N/A.

6.5.7.7 Feature: NAT44 - Port Parity Assignment

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Odd internal port is translated to odd external port and even internal port is translated to even external port.

RFC Reference: N/A.

6.5.7.8 Feature: NAT44 - Hair-pinning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.

RFC Reference: N/A.

6.5.7.9 Feature: NAT44 - 1:1 IP Mapping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).

RFC Reference: N/A.

6.5.7.10 Feature: NAT44 - Outside-Service-App mapping for inside-VRF

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.

RFC Reference: N/A.

6.5.7.11 Feature: Private IPv4/Client-Customer ID

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For general identification and data retention purposes, identifiers has to be uniquely associated with a single CPE. The identifier has to be the private IPv4 address of the CPE.

SI-ID based on private IPv4 CPE address.

Function: Unique Si-ID Ipv4 CPE addresses has to be assigned to the CPE.

6.5.7.12 Feature: NAT44 CGN timers

The premise, function and related RFC reference relating to this feature is as given below.

Premise: These has to allow for all below NAT44 CGN effecting timers to be configured per the RFC.

Function: Listed below are the timers that SHOULD be configured for the NAT44 CGN, note that these SHOULD match the CPE where possible. The timers are not fixed as it is deployment dependent the list you below are the suggested values only and the required fields for any particular vendor:

i) icmp-query --- min 1

ii) sip --- min 2

iii) tcp-established --- hrs 30 min 0

iv) tcp-syn --- sec 30

v) no tcp-time-wait --- min 3

vi) tcp-transitory --- min 4

vii) udp --- min 5

viii) udp-initial --- sec 15

ix) udp-dns --- sec 15

RFC Reference: No single defined RFC.

6.5.7.13 Feature: Thresholds and Watermarks

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are two types of thresholds/watermarks, reactive based and logging based, both are required.

Function: Listed below are minimum required thresholds for NAT44 CGN and SHOULD be adhere to validate the utilization on an event basis due to resource allocation that is required for CGN:

1) Watermarks based on Subscriber Thresholds for the node

2) Watermarks based NPU total resources

3) Watermarks based on Port Allocation Thresholds per Subscriber

4) Watermarks based on Port Allocation Thresholds per NPU

5) Watermarks based on IP Allocation Utilization Thresholds

6) Watermarking Event Reporting through SNMP

7) Watermarking Event Reporting through Syslog

8) Watermarking for dynamic port allocation over the top of Deterministic NAT

RFC Reference: N/A.

6.5.7.14 Feature: Port Block Allocation per IP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Port block allocation is configurable to allow for any ratio assignment per IP.

Function: An assignment of ports based on block allocations for any numeric value with automatic assignment of a single IP determined by the black allocation.

RFC Reference: N/A.

6.5.7.15 Feature: Deterministic NAT / Deterministic Dynamic thresholds

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Deterministic NAT policy and configuration has to be adhered to for NAT44 CGNs not only for port logging requirements for data retention but also for processing and functional improvement.

Function: DetNAT should be used in the following manner. It should be configurable to allow for any n-Ports to be allocated to the port block and any number of port block assignments to be dynamic or static per private IPv4 CPE address. It has to also hold a dynamic threshold to further the ability for IP to port ratios by allowing the dynamic section of the port allocation for any single or cluster of CPE nodes to use 1-5 block allocations and to allow for singular oversubscription per CPE based on a threshold. So if the block allocation reaches over 60 % (configurable) of the full amount of ports, 65 000 per IP, a restriction is placed on further port allocations. For example until the 60 % threshold is reached the Subscriber CPE is allowed up to 3 000 ports but once the port threshold of 605 is reached this then drops to 500 ports for all users. No old connections are dropped but no new ones can be added. This allows for controlled oversubscription with heavy port users to over allocate while there are resources.

RFC Reference: [9**]**.

From section 2 in [9]:

> "Several address assignment algorithms are possible. Using predefined algorithms, such as those that follow, simplifies the process of reversing the algorithm when needed. However, the CGN MAY support additional algorithms. Also, the CGN is not required to support all algorithms described below. Subscribers could be restricted to ports from a single IPv4 address, or could be allocated ports across all addresses in a pool, for example. The following algorithms and corresponding values of A are as follow:
>
> 0. Sequential (e.g. the first block goes to address 1, the second block to address 2, etc.)
>
> 1. Staggered (e.g. for every n between 0 and ((65536-R)/(C+D))-1, address 1 receives ports n*C+R, address 2 receives ports (1+n)*C+R, etc.)
>
> 2. Round robin (e.g. the subscriber receives the same port number across a pool of external IP addresses. If the subscriber is to be assigned more ports than there are in the external IP pool, the subscriber receives the next highest port across the IP pool, and so on. Thus, if there are 10 IP addresses in a pool and a subscriber is assigned 1000 ports, the subscriber would receive a range such as ports 2000-2099 across all 10 external IP addresses).
>
> 3. Interlaced horizontally (e.g. each address receives every Cth port spread across a pool of external IP addresses).

4. Cryptographically random port assignment (Section 2.2 of RFC6431 [RFC6431]). If this algorithm is used, the Service Provider needs to retain the keying material and specific cryptographic function to support reversibility.

5. Vendor-specific. Other vendor-specific algorithms may also be supported. The assigned range of ports MAY also be used when translating ICMP requests (when re-writing the Identifier field).

The CGN then reserves ports as follows:

1. The CGN removes reserved ports (R) from the port candidate list (e.g., 0-1023 for TCP and UDP). At a minimum, the CGN SHOULD remove system ports (RFC6335) [RFC6335] from the port candidate list reserved for deterministic assignment.

2. The CGN calculates the total compression ratio (C+D), and allocates 1/(C+D) of the available ports to each internal IP address. Specific port allocation is determined by the algorithm (A) configured on the CGN. Any remaining ports are allocated to the dynamic pool.

Note: Setting D to 0 disables the dynamic pool. This option eliminates the need for per-subscriber logging at the expense of limiting the number of concurrent connections that 'power users' can initiate.

3. When a subscriber initiates a connection, the CGN creates a translation mapping between the subscriber's inside local IP address/port and the CGN outside global IP address/port. The CGN has to use one of the ports allocated in step 2 for the translation as long as such ports are available. The CGN SHOULD allocate ports randomly within the port range assigned by the deterministic algorithm. This is to increase subscriber privacy. The CGN has to use the preallocated port range from step 2 for Port Control Protocol (PCP, [I-D.ietf-pcp-base]) reservations as long as such ports are available. While the CGN maintains its mapping table, it need not generate a log entry for translation mappings created in this step.

4. If D>0, the CGN will have a pool of ports left for dynamic assignment. If a subscriber uses more than the range of ports allocated in step 2 (but fewer than the configured maximum ports M), the CGN assigns a block of ports from the dynamic assignment range for such a connection or for PCP reservations. The CGN has to log dynamically assigned port blocks to facilitate subscriber-to-address mapping. The CGN SHOULD manage dynamic ports as described in [I-D.tsou-behave-natx4-log-reduction].

5. Configuration of reserved ports (e.g., system ports) is left to operator configuration.

Thus, the CGN will maintain translation-mapping information for all Connections within its internal translation tables; however, it only needs to externally log translations for dynamically-assigned ports."

6.5.7.16 Feature: IP Ranges per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP ranges should be able to be assigned per NAT group or across multiple NAT groups to allow for full scaling and shared IP resource.

RFC Reference: N/A.

6.5.7.17 Feature: NAT Grouping resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT Grouping SHOULD be cross NPU based.

Function: NAT groups SHOULD be able to use the same NPUs and IP allocations. This allows for complete shared resources across multiple NAT groups. The function states that the NPUs can be grouped in any manner for active standby numbers as the user/administrator wishes. Thus NAT group 1 can use NPU 1,2,3&4, with 1,2&3 as active and NAT group 2 can use NPU 1,2,3,4,5&6 with 1,2&3 as standby. This promotes scaling and no restriction if there are particular requirements to share a resource.

RFC Reference: N/A.

6.5.7.18 Feature: NPU to Interface throughput ratio

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To allow balanced Bandwidth assignment across all virtual or a single virtual interface matching or exceeding the physical bandwidth of the node to prevent loss of traffic by software configuration.

Function: To prevent over buffering, dropped packets and general resource issues the NAT44 CGN has to be able to consider a physical topology that matches a balanced logical bandwidth ratio between the interfaces and the NPU throughput, however the feature to rate limit and buffer to match ratios on the NAT NPUs used comparative to the physical interface BW.

RFC Reference: N/A.

6.5.7.19 Feature: NAT44 CGN Address (es)

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT44 CGN Addressing.

Function: Listed below:

1) The NAT44 CGN SHOULD be able to act as the DHPv4 server for the CPE if required.
2) The private IPv4 address assigned to the CPE has to be unique.
3) The NAT44 CGN can advertise a default route so all traffic that should be NAT'ed is routed to the NAT44 CGN.
4) For inter-chassis NAT44 CGN redundancy, both NAT44 CGN should keep track of each others availability.

RFC Reference: N/A.

6.5.7.20 Feature: Source IP per NAT44 CGN Interface

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The premise is that a definition of sourced addresses SHOULD be configured to allow forced routing to the NAT44 CGN.

Function: A user/administrator SHOULD be able to place a source IP statement within a NAT group allowing the CPE packets originating from within that subnet to be routed to a specific NAT group or NAT44 CGN interface. His allows the delineation of different service IPs instead of relying on separation of traffic based on DNS record responses and thus specific rules and resource allocations for that particular sourcing subnet.

RFC Reference: N/A.

6.5.7.21 Feature: NAT44 CGN default route Withdrawal

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NAT44 CGN default route has to be withdrawn under the circumstances below.

Function: Listed below:

1) The NAT44 CGN is not reachable within the routing protocol being used or cannot reach its neighbour's or peers.
2) The NAT group goes down.
3) The NAT process dies.
4) NPUs available cannot hold the present capacity.
5) Shutdown or part or all of the service.
6) External Routing to the internet is down for a fixed period of time (demands NAT44 CGN clustering to determine reachability from a redundant NAT44 CGN).

RFC Reference: N/A.

6.5.7.22 Feature: Chassis Clustering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To Cluster two or more NAT44 CGNs for session state failover.

Function: If a single NAT44 CGN goes down it has to do the following:

- Remove its NAT44 CGN address as a reachable entity within the routing protocol being used
- Forward any remaining packets to the secondary NAT44 CGN
- Validate NAT cache state between itself and its redundant NAT44 CGN

The clusters has to sync on the following items:

- The TCP and UDP sessions presently being used to for NAT44
- The TCP not used for n amount of time configurable
- The present capacity of interface and NPU to allow for structural syncing
- Sync between all clustered NAT44 CGNs due to best path attributes

RFC Reference: N/A.

6.5.7.23 Feature: Multiple Transition Technology Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT44, NAT64 and DS-Lite has to function on the same platform to allow for uniformity within European markets.

Function: All three technologies should be able to function and share resources dynamically on any given platform.

RFC Reference: N/A.

6.5.7.24 Feature: NPU/chassis Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Buffering requirements for the NAT44 CGN for transit traffic covering queues and buffer timers.

Function: Listed below:

1) Queue sizing per flow
2) Queue sizing per port block
3) Queue sizing per NPU
4) Queue sizing per ingress interface
5) Buffer sizing per flow
6) Buffer sizing per port block
7) Buffer sizing per NPU
8) Buffer sizing per ingress interface
9) Reordering buffers and out-of-order caching

RFC Reference: N/A.

6.5.7.25 Feature: Stateful ICMP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Stateful ICMP is the act of holding state for ICMP for messages sent/originated from the NAT44 CGN to any know destination on a single flow. This can be used downstream as well as upstream.

Function: A maximum of 3 ICMP packet-too-big messages has to be sent to single destination for a single flow. This prevents flooding and over compensation. This should ONLY be used when the DF bit is set.

RFC Reference: N/A.

6.5.7.26 Feature: Port Reservation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Reserving a block of public ports for each private IPv4 client reduces the scaling requirements of per-connection port-mapping.

Function: The CGN-GW has to be capable of bulk/block port reservation.

RFC Reference: N/A.

6.5.7.27 Feature: Static Port Forwards

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Allowing the consumer to configure static port forwarding, using a portal, to configure the CGN-GW.

Function: Using PCP the customer can log on to their "ISP website" and configure instantaneously a port to forward to their in-home service or server by IP assignment. The CGN-GW will assign that port return path into the home for a specific public IP for a **fixed** period of time. Note that this crosses the port block assignment restriction of being within a single IP address and thus creates a new DR log. This based mostly on the CGN-GW port changes, however PCP may be used to inform the CPE to create a double pop on the port allocation, so the CGN-GW present the public IPv4 IP on the well known port translates it to private IPv4 and public IPv4 based SI-ID to keep the destination port to the consumer CPE the same, however if there is already an entry stated for that SI-ID then a double pop maybe necessary requiring the CPE to change the port once again and thus requiring PCP exchange. An extended state SHOULD be held for each double pop port exchange within the extended PCP table.

RFC Reference: N/A.

6.5.7.28 Feature: PCP Mode

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The mode of PCP operation should be has to be primarily plain **option 1** below, but encapsulation SHOULD be supported as well (configurable). The mode denotes the separation of traffic or not from the AFTR address and the type of encapsulation.

Function: **Note that the CGN-GW SHOULD support both RFC drafts 12/13 as well RFC draft 29 for PCP (configurable).**

Preference form is to have plain mode on a separate address, this allows two topologies and the ability for processing of the PCP packets to be separated without the need for mixed traffic separation of the traffic allows for separation of topologies in the case of a PCP server existing isolated from the CGN-GW node.

All sourcing in any given topology or mode has to support the IPv6 address as the identifier to allow for NAT cache matches between the PCP tables and the main NAT CACHE for the SI-ID (softwire initiated identifier). This is a requirement for CGN-GW redundancy clustering, data retention and other such feature sets.

RFC Reference: [135**]**.

From section 1 in [135]:

> "The Port Control Protocol (PCP) provides a mechanism to control how incoming packets are forwarded by upstream devices such as Network Address Translator IPv6/IPv4 (NAT64), Network Address Translator IPv4/IPv4 (NAT44), and IPv6 and IPv4 firewall devices, and a mechanism to reduce application keepalive traffic. PCP is designed to be implemented in the context of Carrier-Grade NATs (CGNs) and small NATs (e.g., residential NATs), as well as with dual-stack and IPv6-only Customer Premises Equipment (CPE) routers, and all of the currently known transition scenarios towards IPv6-only CPE routers. PCP allows hosts to operate servers for a long time (e.g., a network- attached home security camera) or a short time (e.g., while playing a game or on a phone call) when behind a NAT device, including when behind a CGN operated by their Internet service provider or an IPv6 firewall integrated in their CPE router."

6.5.7.29 Feature: PCP Failure

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function and reactivity during a failure of the session either at Layer 3, 4 and 5 at the client, server, CGN-GW or CPE.

Function: Listed below:

1) Relay of failure messages from CGN-GW to CPE receive
2) Relay of failure messages from CPE to CGN-GW send
3) Lifetime timer configuration for failure; PCP entry timers
4) Monitoring failure of the of the PCP without message
5) Static mappings are not affected (comparative if the CPE holds those static explicit mappings)
6) PCP table content considering failure (last received packet). PCP table matching when/if required

RFC Reference: N/A.

6.5.7.30 Feature: PCP (SI-ID Extension Based) Multi-session Dynamic Forwarding

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP allocation for multiple instances of the same application on separate PCs behind the same CPE.

Function: Allows multiple sessions from multiple PCs behind a single CPE to access the same port by assigning as separate IP to the before (IPv4 public). The ability to dynamically allocate a static port forwards for a different outside IP as the subscriber got assigned when the particular port is in use.

Reasoning is bittorrent clients with UPnP and PCP to the CGN-GW will not be able to negotiate the same port for multiple sessions from different PCs and given all clients use the same "fixed" port by default it is better to allocate a different outside IP.

RFC Reference: N/A.

6.5.7.31 Feature: NAT44 Tunelling

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The creation of a encapsulation tunnel 4in4 for NAT44 for prevent the nested NAT issues common with NAT44 implementations.

Function: Allows L2TP or other tunnelling methodologies to separate the IPv4 private link addressing from CPE to HeadEnd with the transit private addressing structure therefore removing any Nested NAT (NAT444/4444) form end to end.

RFC Reference: N/A.

6.5.8 Monitoring and Management

6.5.8.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.5.8.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: RFC 5905 [115].

6.5.8.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: RFC 5905 [115].

6.5.8.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: RFC 4250 [76].

6.5.8.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: RFC 4250 [76].

6.5.8.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: RFC 854 [18].

6.5.8.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: RFC 854 [18].

6.5.8.8 Feature: SNMP General

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CGN should support specific MIBs to allow full visibility of the resource and traffic.

Function: MIBs should be available for the following:

- Port percentage utilization per IP
- IP Address percentage per range assigned and per NAT group
- Resource utilization per NPU
- Memory utilization
- Processor utilization per NPU
- PCP Percentage IP assignments
- Clustering utilization (the total utilization if one of the pair of the cluster goes down)
- Interface utilization BW
- NAT sessions (dynamic thresholds)
- SI-ID numbers per NPU and chassis (based on both private IPv4 and public IPv4 used out of the range assigned)

RFC Reference: N/A.

6.5.8.9 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: RFC 1441 [26].

6.5.8.10 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: RFC 1441 [26].

6.5.8.11 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: RFC 3411 [60].

6.5.8.12 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: RFC 3411 [60].

6.5.8.13 Feature: remote access: TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: draft-grant-tacacs-02 [13].

6.5.8.14 Feature: remote access: TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: draft-grant-tacacs-02 [13].

6.5.9 CPE Definition

6.5.9.1 Feature: Cable Router B4 Functionality

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to provide tunnelling of IPv4 based traffic from either directly connected or other CPE devices the DOCSIS cable modem has to provide the capability to implement the required functionality to deliver such traffic via an IPv6 tunnel to an CGN-GW router. This is defined as the NAT Basic Bridging Broadband - B4 element, this functionality is provided by the implementation of a B4 interface on the DOCSIS interface of the cable modem.

Function: The NAT44 CPE SHOULD NOT operate any NAT functionality between the B4 interface and any local LAN interfaces as this functionality is provided by the CGN-GW gateway in the network.

6.5.9.2 Feature: SLAAC

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv6 Address Auto Configuration, In order to allow for the generation of a unique IP address without reliance on an external server or manual provisioning the B4 device has to support the IPv6 stateless autoconfiguration method as detailed in RFC 4862 [97]. This methodology ensures that a device can generate a unique and routeable IPv6 address based upon locally available information and advertisements from adjacent routers.

Function: The NAT44 B4 CPE device has to be capable of supporting IPv6 stateless autoconfiguration.

6.5.9.3 Feature: DNS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: As the B4 client is only provided with IPv6 connectivity on its WAN interface, it has to perform all DNS resolution over IPv6.

Function: To reduce the amount of NAT44 traffic, the cable B4 gateway has to be configured as a DNS forwarder for all LAN client requests. Any client DNS requests received via IPv4 will therefore be forwarded as IPv6 to the external DNS servers.

RFC Reference: N/A.

6.5.9.4 Feature: IPv6 LAN IP Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are no specific requirements for IPv6 LAN addressing to support NAT44. As NAT44 provides native IPv6 end-to-end connectivity, IPv6 addressing SHOULD be used for all LAN clients that support IPv6. This being defined by the provisioning policy implanted on the service providers provisioning systems.

RFC Reference: N/A.

6.5.9.5 Feature: IPv4 LAN IP Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The B4 has to provide 'private' IP addressing from the ranges specified in RFC 1918 [30], tools.ietf.org/html/rfc1918. The accepted normal configuration is to use the 192.168.0.0/16 range, with a single /24 subnet.

The B4 has to use one address from the selected IPv4 subnet as its LAN interface address, and this has to be used as the default gateway for LAN IPv4 traffic.

The B4 has to also use its LAN IP address as the DNS server address in DHCP offers.

RFC Reference: RFC 1918 [30].

6.5.9.6 Feature: Packet Encapsulation

Premise: The B4 router has to encapsulate packets in accordance with the specification defined in RFC 6333 [129]. The B4 encapsulation interface has to drop any packets in the following groups.

Function: Listed below:

- Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].
- Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].
- Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].
- Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: N/A.

6.5.9.7 Feature: Packet Decapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The B4 router has to decapsulate packets in accordance with the specification defined in RFC 6333 [129].

Function: Where an IPv4 fragment is received, the B4 router has to forward the fragment unaltered to its destination. It is the responsibility of the destination client to reassemble fragments.

Once the IPv4 packet is recovered, it has to forward the packet to the client specified in the source address.

RFC Reference: N/A.

6.5.9.8 Feature: MTU and fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Fragmentation of an IPv4 and IPv6 packet The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes due to DOCSIS requirements on CMTSes. The maximum IPv4 packet size that can be sent between the CPE and CGN-GW is therefore 1 500 bytes.

Function: As fragmentation in IPv6 is performed by the sending host, any packet received that is in excess of this size is not allowed to fragment but send an ICMP stateful response.

Function: IPv4 MTU reduction.

The MTU of the CPE encapsulation interface can be 1 500 bytes. Any client that uses Path MTU Discovery (PMTUD) (RFC 1191 [22]) will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

6.5.9.9 Feature: MSS clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header).

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

MSS clamping has to be implemented on the CPE interface, to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size).

RFC Reference: N/A.

6.5.9.10 Feature: DHCPv4 MTU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv4 option 26 as defined in RFC 2132 [33], is used to specify the IPv4 MTU that each client should use.

Function: DHCPv4 option 26 as defined in RFC 2132 [33], is used to specify the IPv4 MTU that each client should use.

The setting of this parameter will ensure that packets leaving the host are a maximum of 1 460 bytes. However this has the following limitations:

1) Not all client IP stacks and DHCP clients will respect the DHCP option to set the interface MTU.

2) Devices that are manually configured (and hence do not use DHCP) would also require manual configuration of the interface MTU.

3) Reducing the interface MTU of clients would also impact all LAN IPv4 traffic as this would also all be reduced to an MTU of 1 500 bytes.

4) It cannot be guaranteed that all client applications will function if the MTU has been reduced.

RFC Reference: N/A.

6.5.9.11 Feature: Recommendations

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate NAT44 encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU has to be set to allow for full IPv6 encapsulated NAT44 packets to be sent with no fragmentation.

Where the DOCSIS layer only permits an MTU of 1 500 bytes, the following recommendations are required.

1) The device B4 interface has to be configured with an MTU of 1 460 bytes.
2) The device has to fragment IPv4 packets greater than 1 460 bytes where DF is not set.
3) The device has to drop packets greater than 1 460 byte where DF is set and reply with an ICMP "packet too big" message and specify the MTU of 1 460 bytes.
4) The device has to support MSS clamping to 1 420 bytes.
5) The device has to support DHCPv4 option 26 set to 1 460 bytes.

RFC Reference: N/A.

6.5.9.12 Feature: Inbound sessions

The premise, function and related RFC reference relating to this feature is as given below.

Premise: With NAT44, the public IPv4 address is not configured on the CPE. The ability to control inbound sessions is also removed. These are generally:

1) DMZ host functionality. Where all inbound packets can be forwarded to a single client.
2) Port forwarding. Where individual TCP or UDP ports can be forwarded to a specific client.
3) UPnP NAT Traversal. Where a client requests for the NAT to open one or more TCP or UDP ports to itself for inbound packets.

Function: The B4 router has to be able to replicate the existing behaviour, by communicating these requests to the CGN-GW, for it to implement the relevant port forwarding to the specific client. Port Control Protocol (PCP) is the proposed mechanism to provide this functionality [135].

RFC Reference: N/A.

6.5.9.13 Feature: MIB Support for NAT44

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to support configuration of the tunnel and NAT functions required by NAT44 the CPE device needs to implement an enhanced MIB that provides the required functionality that is not supported by the implementation of NAT-MIB [i.5].

RFC Reference: N/A.

6.5.10 NAT44 Technical Viability

Technically and architecturally NAT44 has some drawbacks, but it is known technology. Functionally it should only be used for business continuity for new additions to your network and thus the issue of centralized placement in the network becomes minimal.

The Pros

- Allows for a smooth migration towards an IPv6 only operator infrastructure. This includes IPv6 only access, aggregation and core network as well as an IPv6 only BNG. In case of a cable operator the BNG is implemented on the CMTS. In case of other access technologies (e.g. GPON) the BNG is implemented on a dedicated platform, but this moves us closer to higher requirements for development of IPv6.
- It As a business continuity process for new customers the scaling is acceptable in live deployment.
- Central intelligence makes development much easier and upgrading only a few nodes in the network whereas CPE state technologies incur massive requirements for upgrades and difficulty with functionality over multiple vendors.
- Cost of the CPE would be the same as most CPEs have NAT44 capability.

The Cons

- NAT44 requires ALGs (ALPs in most cases) which require the AFTR to do some form of intelligent vicissitudes to the transit packets bearing the risk of drop in performance.
- The requirement for Data Retention capacity is rather large but this is mitigated with DeNAT.
- NAT44 is a hub-and-spoke topology forcing all tunnelled communication from CPE to pass through the AFTR which causes.
- No IPv6 network means no development towards a transition to the new protocol.
- Nested NAT breaks many applications.

6.6 464XLAT Technology Summary

464XLAT is one of the IPv6 transition technologies. The main reason for considering 464XLAT for further analysis is its ability to operate without inter-communication between IPv4 and IPv6. This technology will allow customers to access services natively over IPv6 and through translation over IPv4. As IPv6 is not "backward compatible" to IPv4, i.e. the two protocols exist independently of each other without any interaction in most topologies, 464XLAT is introduced to allow for a smooth transition towards IPv6 once no more IPv4 addresses are available and a co-habitation of IPv4 and IPv6 on the same network infrastructure is required. 464XLAT also provides the ability for IPv4-only home devices, applications and OSs to continue to access the Internet with minimal use of IPv4 public addresses from the operator's remaining pool. It also avoids dual or multiple layers of address translation which otherwise introduces its own set of problems.

464XLAT does not require DNS64 since an IPv4 host may simply send IPv4 packets - including packets to an IPv4 DNS server - that are translated on the CLAT to IPv6 and back to IPv4 on the PLAT.

The IPv6 address format in 464XLAT is defined in RFC 6052 [119]. Prefix delegation mechanism such as DHCPv6-PD are available to assign a dedicated translation prefix to the CLAT. From the delegated DHCPv6 prefix, a /64 prefix is dedicated for addressing outgoing and incoming IPv6 packets associated with the stateless translation. The CPE (CLAT) may discover the related prefix of the PLAT via a variety of specified methods such as by means of a DHCPv6 option, using TR-069, DNS APL RR.

In the case that DHCPv6-PD is not available the CLAT does not have a dedicated IPv6 prefix for translation. It, alternatively, performs NAT44 and stateless translation. For IPv4 packets coming from the Home Network and containing a source address from the private IPv4 address space NAT44 is used to forward the packets to the CLAT IPv4 host address. The CLAT will execute a stateless translation such that the IPv4 packets received at the CLAT IPv4 host interface are translated to the CLAT WAN IPv6 address The IPv6 prefix is constructed of the delegated prefix which is completed if needed to form a /64 prefix by adding a subnet ID of 0.

The addressing scheme allows subnet lengths for the XLAT prefix to be 32, 40, 48, 56, 64 or 96 bit. Depending on the prefix length, the IPv6 address with embedded IPv4 address is formatted according to table 9.

Table 9: Embedding IPv4 addresses in IPv6 addresses with different prefix lengths

0-15	16-31	32-47	48-63	64-79	80-96	96-111	112-128	
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix Prefix	0 Prefix	Prefix Prefix	IPv4 IPv4	IPv4 IPv4	/96
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix Prefix	u IPv4	IPv4 IPv4	IPv4 Suffix	Suffix Suffix	/64
Prefix Prefix	Prefix Prefix	Prefix Prefix	Prefix IPv4	u IPv4	IPv4 IPv4	Suffix Suffix	Suffix Suffix	/56
Prefix Prefix	Prefix Prefix	Prefix Prefix	IPv4 IPv4	u IPv4	IPv4 Suffix	Suffix Suffix	Suffix Suffix	/48
Prefix Prefix	Prefix Prefix	Prefix IPv4	IPv4 IPv4	u IPv4	Suffix Suffix	Suffix Suffix	Suffix Suffix	/40
Prefix Prefix	Prefix Prefix	IPv4 IPv4	IPv4 IPv4	u Suffix	Suffix Suffix	Suffix Suffix	Suffix Suffix	/32

Bits 64 to 71 (u) should always be set to zero even when using a /96 prefix.

A 464XLAT translation causes a change in MTU size. In addition to the minimum length of 40 Byte for the IPv6 header, 20 Byte length of the IPv4 header have to be taken into account. If after IPv4 to IPv6 translation the IPv6 link MTU is exceeded, it is recommended to fragment the IPv4 packets before they enter the NAT and to set the Max Outside MTU of the NAT accordingly.

Max Outside MTU = IPv6 MTU - 40 Byte IPv6 header - 8 Byte IPv6 fragmentation header + 20 Byte IPv4 header

For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.

The CPE (CLAT) should implement a DNS proxy. The case of an IPv4-only node behind the CLAT querying an IPv4 DNS server is undesirable since it requires both stateful and stateless translation for each DNS lookup. The CLAT itself should offer a DNS server via DHCP or other means and proxy DNS queries for IPv4 and IPv6 clients. Gateway functions are commonly implemented in CLAT-enabled home routers such that the operation of a DNS proxy is straight forward. The ultimate goal is to simplify traffic flows such that only IPv6 native queries are forwarded across the Access Network. The CLAT should allow for a client to query any DNS server of its choice and bypass the proxy.

6.6.1 PLAT Technology Feature Summary

Table 10 gives a summary of the PLAT technology feature.

Table 10: PLAT Technology Feature Summary

Functional Name	Requirement	Description
RFC 6877 [134]	Required	464XLAT: Combination of Stateful and Stateless Translation.
RFC 6052 [119]	Required	IPv6 Addressing of IPv4/IPv6 Translators.
Redundancy	Required	All critical components has to be redundant in such a manner that they can fail-over without impact to customer or management traffic greater than 1 ms.
Shared Resource	Required	Single PLAT GW Prefix. The PLAT IPv6 prefix should be able to be shared amongst different NPU's in the PLAT. A hashing mechanism should be in place to hash all upstream packets based on the source IPv6 address.
Si-ID	Optional	Tunnel/customer identifier based on both IPv6 CLAT prefix public and IPv4 Client Private addresses embedded in the IPv6 source address.
PLAT Addressing and Virtual Interfaces	Required	PLAT has to be able to assign a single virtual interface with up to 8 PLAT GW prefixes for any given 464XLAT instance on the node.
Anycast	Required	Anycast PLAT gateway prefixes are a requirement to allow simplicity of deployment for a single prefix across multiple PLAT's.
PLAT Address withdrawal	Required	The PLAT should have at least five points of PLAT GW prefix withdrawal occurrence. The list includes: - loss of route out, - loss of all BGP/IGP sessions, - loss of forwarding, - loss of NPU capacity and certain errors in the NAT caching. Any of the failures should be detectable based on configurable timers with 15 seconds being the default setting.
NPU Load Balancing (hashing)	Required	464XLAT Inside to Outside hashing performed on the Source IPv6 (x bits) prefix of the CLAT device. 464XLAT Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
MTU	Required	Due to a larger header size for IPv6 packets, the MTU has to be configurable. Expected value will be 1 480 bytes.
MSS Clamping	Required	TCP MSS support is mandatory for the PLAT due to the removal of an end-to-end MTU sizing functionality. This will avoid the need for excessive fragmentation.
Fragmentation	Required	Fragmentation should be done on the IPv4 packet. Fragmentation should be placed on the ASIC running in the line card.
NAT - Network Address and Port Mapping - Endpoint Independent Mapping	Required	For two flows for a common inside source IPv4 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation.
NAT - Translation Filtering - Endpoint Independent Filtering	Required	A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.
NAT - Paired IP Address Assignment	Required	Translation to External IPv4 address is done in a paired fashion. A given Inside IPv6 CLAT prefix is always translated to the same External IPv4 address.
NAT - Hair-pinning	Required	Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.
NAT - 1:1 IP Mapping	Required	Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).
NAT - Outside-Service-App mapping for inside-VRF	Required	Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.
NAT - Port Limit configuration	Required	A maximum amount of ports can be configured for every IPv6 source CLAT prefix.

Functional Name	Requirement	Description
NAT - Per-Protocol Timeout configuration	Required	Timeout of mappings is critical and removing a NAT mapping at the appropriate time maximizes the shared port resource.
NAT - Dynamic Port Range start configuration	Required	The start port for dynamic port ranges should be configurable to allow for a range of ports for static port mappings.
NAT - Software Load Balancing	Required	NAT Inside to Outside hashing performed on the Source IPv6 CLAT Prefix. NAT44 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Port Allocation	Required	In order to reduce the volume of data generated by the NAT device (logging creation and deletion data), bulk port allocation can be enabled. When bulk port allocation is enabled and when a subscriber creates the first session, a number of contiguous outside ports are pre-allocated. A bulk allocation message is logged indicating this allocation.
FTP ALG (Active and Passive)	Required	FTP clients are supported with inside (private) address and servers with outside (public) addresses. Passive FTP is provided by the basic NAT function. Active FTP is used with the ALG.
RTSP ALG	Required	Remote control protocol for streamers (which use RTP/RTCP or RDT). Our implementation considers the server is located "outside" and clients are "inside". RTSP is used in many streamers like QuickTime or RealNetworks. Both "SETUP" and "SETUP response" are analysed. Based on this information, the translation entry will be created in the table.
SIP ALG	Required	SIP requires control and signalling on separate port to the data traffic and therefore requires an ALG to function.
PPTP ALG	Required	PPTP, used in many VPN setups, requires control and signalling on separate port to the data traffic and therefore requires an ALG to function.
Use of Application SVI's, connecting multiple routing entities (inside/private VRF, outside/public VRF)	Required	NAT should be possible from any routing context (VRF, GRT) to any routing context (VRF, GRT).
Stateful ICMP	Required	Stateful ICMP mappings between inside and outside ICMP identifiers should be supported.
Thresholds	Required	Configurable thresholds using watermarks should be supported to monitor the resources on the PLAT.
QoS translation	Required	For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.
Chassis NAT Clustering	Optional	Clustering of PLAT's to allow for inter-chassis resiliency.
PCP	Required	Support for PCP to allocate static port bindings.
per IPv4 user port limiting	Optional	The ability to assign port blocks per private IPv4 address.
Logging via Netflow V9/IPFIX	Required	Netflow v9 support for logging of the translation records. Logging of the translation records can be mandated by for Lawful Intercept. The Netflow uses binary format and hence requires software to parse and present the translation records.
Logging via Syslog	Required	Syslog support for logging of the translation records as an alternative to Netflow. Syslog uses ASCII format and hence can be read by users. However, the log data volume is higher in Syslog than Netflow.
Destination based Logging	Required	Destination Based Logging will generate one record per session, including the destination port and address. NFv9 templates including destination IP address and destination port will be used.

Functional Name	Requirement	Description
Base Logging Fields	Required	PLAT logs the following information when a translation entry is created: Inside instance ID Outside instance ID Inside IPv6 Address Inside Port Outside IPv4 Address Outside Port Protocol Start Time Stop Time
Radius Logging	Required	Logging using Radius accounting messages.
XML l	Optional	Logging using XML files.
Static port forwarding (up to 6K static forward entries per npu)	Required	Static port forwarding configures a fixed, private (internal) IPv6 address and port that are associated with a particular subscriber while the PLAT allocates a free public IP address and port. Therefore, the inside IPv6 address and port are associated to a free outside IP address and port.
Static port forwarding 1:1 active/standby	Required	Static Port Forwarding mapping will be kept constant with two npu cards in Active/Standby mode.
Performance Requirement		**Description**
64 PLAT instances per npu Card	Required	The ability to stack transition instances on top of one another.
20M+ Translations (per npu)	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary.
Minimum 10 Gbps throughput per npu (Inside-to-outside VRF + Outside-to-inside VRF) with IMIX traffic	Required	Per half slot throughput requirement.
1M+ connections per second setup rate	Required	Primary sessions, port block assignments, per NPU.
Minimum 1M users (private IPv6 CLAT prefixes on the inside)	Required	B4 XLAT Scaling per Chassis for Carrier Grade requirements.
Latency	Required	Latency is between 40 and 200 micro seconds (µs).
Minimum 6 npu Cards per chassis	Required	Min chassis requirement.
IRB support	Required	Integrated Routing and Bridging / Virtual Interfaces (L3 interface for Bridge Domain).
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 256 k NAT users at the same time.

6.6.2 Main 464XLAT RFC References

The PLAT technology feature are described by the following IETF RFCs: RFC 4787 [94], RFC 5382 [106], RFC 5508 [108], RFC 6052 [119], RFC 6877 [134], RFC 6887 [135].

6.6.3 CORE Device (LSN/CGN)

PLAT is an implementation of an IPv4/IPv6 transition protocol based on 464XLAT, and is the CORE node that all CPE/CEs running 464XLAT connect to. It is the termination for the IPv6 prefix and the originating IPv4 NAT device egressing traffic towards the destination on IPv4 after translation.

This section concerns itself solely with the 464XLAT specific required functionality to allow the technology to function completely and fully. The specification laid out below allows 464XLAT technology to be deployed on the PLAT as a non-service deprecating form comparative to a native private IPv4 delivery to any given customer within a Carrier Topology or potentially otherwise. The present document is however focuses on Cable requirements in functionality but can stretch further if required.

Note that any given PLAT implementation should be complaint with the associative feature for deployment within a Cable ISP network.

6.6.4 Hardware Software Requirements

6.6.4.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU in a full stateful fashion with no loss of packets or latency beyond 1ms. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1 ms failover maximum for any given CPE connectivity of session

RFC Reference: N/A.

6.6.4.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The PLAT has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the PLAT can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the PLAT NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.6.4.3 Feature: RP / RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or PLAT transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within a PLAT has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.6.4.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given PLAT has to be shared giving the appearance of a single node on the network.

Function: The PLAT has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single PLAT prefix per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.6.4.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given PLAT has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given PLAT to see all or part of the traffic coming into the node. Hashing SHOULD be used to determine the load balancing between all processing/hardware forwarding points with that NAT group or deterministic specification.

RFC Reference: N/A.

6.6.5 Performance Requirements

6.6.5.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1ms for any given function of the PLAT.

Function: Within the remit of traffic requiring NAT function, the PLAT has to perform its function from ingress interface to egress interface in a time measured no higher than 1ms. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions.

RFC Reference: N/A.

6.6.5.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.6.5.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the PLAT Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 gig is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.6.5.4 Feature: Min Sessions per PLAT Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis.

RFC Reference: N/A.

6.6.5.5 Feature: Minimum Customer IPv4 Addresses per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Defines the minimum amount of IPv4 addresses able to be configured per chassis.

Function: If we define the IPv4 address space requirement based on session and block allocation this would be 60 ports per subscriber as a block assignment comparative to 100 k customers online allowing the integer of 1 024. Or if we calculate backwards this, would be 100 000 customers divided by 60 per IP (although this changes per deterministic NAT requirement) allowing for a max utilization of 1 000 ports per subscriber and thus giving us 1 666-address requirement if we have maximum assignment utilization. So we can round this up to 2 040-address minimum address pool requirement.

RFC Reference: N/A.

6.6.5.6 Feature: Min Customer CLAT's per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus IPv6 CLAT prefixes that can be assigned from a single Chassis or connected.

Function: Due to the scaling of any given PLAT solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 CPE IPv6 prefixes. This is solely a minimum.

RFC Reference: N/A.

6.6.6 Feature Specifications

6.6.6.1 Feature: Customer Identifiers/Client-Customer ID

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For general identification and data retention purposes, customer identifiers has to be uniquely associated with a single CPE. The identifier has to be the IPv6 prefix of the CPE.

SI-ID based on both IPv6 CPE public and IPv4 Client Private.

Function: Unique Si-ID IPv6 CLAT prefixes has to be assigned to the CPE, if the DHCPv6 originates from the PLAT or a separate DHCPV6 server, although the SI-ID can be determined with a join between the CLAT IPv6 prefix and the IPv4 private address assigned within the customer Local Network. This allows for specific port block assignments if there are more than 5 clients behind a specific CLAT CPE allowing for a greater amount of port allocation.

RFC Reference: N/A.

6.6.6.2 Feature: PLAT timers

The premise, function and related RFC reference relating to this feature is as given below.

Premise: These has to allow for all below PLAT effecting timers to be configured per the RFC.

Function: Listed below are the timers that SHOULD be configured for the PLAT, note that these SHOULD match the CPE where possible. The timers are not fixed as it is deployment dependent the list you have below are the suggested values only and the required fields for any particular vendor.

i) icmp-query --- min 1

ii) sip --- min 2

iii) tcp-established --- hrs 30 min 0

iv) tcp-syn --- sec 30

v) no tcp-time-wait --- min 3

vi) tcp-transitory --- min 4

vii) udp --- min 5

viii) udp-initial --- sec 15

ix) udp-dns --- sec 15

RFC Reference: No single defined RFC.

6.6.6.3 Feature: Thresholds and Watermarks

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are two types of thresholds/watermarks, reactive based and logging based, both are required.

Function: Listed below are minimum required thresholds for PLAT and SHOULD be adhere to validate the utilization on an event basis due to resource allocation that is required for CGN:

1) Watermarks based on Subscriber Thresholds for the node

2) Watermarks based NPU total resources

3) Watermarks based on Port Allocation Thresholds per Subscriber

4) Watermarks based on Port Allocation Thresholds per NPU

5) Watermarks based on IP Allocation Utilization Thresholds

6) Watermarking Event Reporting through SNMP

7) Watermarking Event Reporting through Syslog

RFC Reference: N/A.

6.6.6.4 Feature: Port Block Allocation per IP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Port block allocation is configurable to allow for any ratio assignment per IP.

Function: An assignment of ports based on block allocations for any numeric value with automatic assignment of a single IP determined by the black allocation.

RFC Reference: N/A.

6.6.6.5 Feature: IP Ranges per Chassis

Premise: IP ranges should be able to be assigned per NAT group or across multiple NAT groups to allow for full scaling and shared IP resource.

Function: For further study.

RFC Reference: N/A.

6.6.6.6 Feature: NAT Grouping Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT Grouping SHOULD be cross NPU based.

Function: NAT groups SHOULD be able to use the same NPUs and IP allocations. This allows for complete shared resources across multiple NAT groups. The function states that the NPUs can be grouped in any manner for active standby numbers as the user/administrator wishes. Thus NAT group 1 can use NPU 1,2,3&4, with 1,2&3 as active and NAT group 2 can use NPU 1,2,3,4,5&6 with 1,2&3 as standby. This promotes scaling and no restriction if there are particular requirements to share a resource.

RFC Reference: N/A.

6.6.6.7 Feature: Virtual interface per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A PLAT has to be able to assign a single virtual interface with up to 8 464XLAT GW prefixes for any given 464XLAT instance on the node.

Function: The virtual interfaces has to be able to be assigned to any number, and all, NPUs within the chassis to allow complete shared resources. So a single or multiple PLAT's acting as the single destination interface for all resources and CLAT nodes using the PLAT.

RFC Reference: N/A.

6.6.6.8 Feature: NPU to Interface throughput ratio

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To allow balanced Bandwidth assignment across all virtual or a single virtual interface matching or exceeding the physical bandwidth of the node.

Function: To prevent over buffering, dropped packets and general resource issues the PLAT has to be able to consider a physical topology that matches a balanced logical bandwidth ratio between the interfaces and the NPU throughput.

RFC Reference: N/A.

6.6.6.9 Feature: PLAT Prefix (es)

The premise, function and related RFC reference relating to this feature is as given below.

Premise: PLAT Addressing.

Function: Listed below:

1) Has to Allow one or multiple prefixes as GWs for either one or multiple NAT groups
2) IPv4 and IPv6 addressing SHOULD be placed on separate interfaces
3) The PLAT SHOULD be able to act as the DHPv6 server for the CLAT's if required
4) SI-IDs should be able to comprise of both; ONLY the IPv6 CLAT prefix or BOTH the IPv4 private and the IPv6 public prefix
5) The IPv6 Public prefix assigned to the CLAT has to be unique
6) The IPv6 prefixes configured for 464XLAT Gateways has to be Anycast compatible

7) The PLAT has to allow multiple NAT group per IP address and thus per Virtual interface end point

RFC Reference: N/A.

6.6.6.10 Feature: Anycast Gateway Prefix

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Anycast usage for a PLAT gateway or DHCPv6 address has to be accepted within the configuration.

Function: One or many Anycast prefixes has to be configurable for the PLAT GWs regardless of any other resource and for the DHCPv6 server address. **Note** that using Anycast might cause failover issues if a path in your network changes and flapping could prevent session stability if clustering and cluster forwarding is not enabled.

RFC Reference: N/A.

6.6.6.11 Feature: Source IP per PLAT Interface

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The premise is that a definition of sourced addresses SHOULD be configured to allow forced routing to a specific GW prefix on the PLAT.

Function: A user/administrator SHOULD be able to place a source IP statement within a NAT group allowing the CLAT packets originating from within that subnet to be routed to a specific NAT group or PLAT interface. His allows the delineation of different service IPs instead of relying on separation of traffic based on DNS record responses and thus specific rules and resource allocations for that particular sourcing subnet.

RFC Reference: N/A.

6.6.6.12 Feature: PLAT Prefix Withdrawal

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The PLAT Prefix has to withdrawn under the circumstances below.

Function: Listed below:

1) The PLAT is not reachable within the routing protocol being used or cannot reach its neighbour's or peers
2) The NAT group goes down
3) The 464XLAT process dies
4) NPUs available cannot hold the present capacity
5) Shutdown or part or all of the service
6) External Routing to the internet is down for a fixed period of time (demands PLAT clustering to determine reachability from a redundant PLAT)

RFC Reference: N/A.

6.6.6.13 Feature: Chassis Clustering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To Cluster two or more PLATs for session state failover.

Function: If a single PLAT goes down it has to do the following:

- Remove its PLAT prefix as a reachable entity within the routing protocol being used
- Forward any remaining packets to the secondary PLAT
- Validate NAT cache state between itself and its redundant PLAT

The clusters has to sync on the following items:

- The TCP and UDP sessions presently being used to for NAT
- The TCP not used for n amount of time configurable
- The present capacity of interface and NPU to allow for structural syncing
- Sync between all clustered PLATs due to best path attributes

RFC Reference: N/A.

6.6.6.14 Feature: Multiple Transition Technology Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: 464XLAT, NAT44, NAT64 and DS-Lite has to function on the same platform to allow for uniformity within European markets.

Function: All four technologies should be able to function and share resources dynamically on any given platform.

RFC Reference: N/A.

6.6.6.15 Feature: NPU/chassis Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Buffering requirements for the PLAT for transit traffic covering queues and buffer timers.

Function: Listed below:

1) Queue sizing per flow
2) Queue sizing per port block
3) Queue sizing per NPU
4) Queue sizing per ingress interface
5) Buffer sizing per flow
6) Buffer sizing per port block
7) Buffer sizing per NPU
8) Buffer sizing per ingress interface
9) Reordering buffers and out-of-order caching

RFC Reference: N/A.

6.6.6.16 Feature: Stateful ICMP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Stateful ICMP is the act of holding state for ICMP for messages **sent/originated** from the PLAT to any know destination on a single flow. This can be used downstream as well as upstream.

Function: A maximum of 3 ICMP packets too big messages has to be sent to single destination for a single flow. This prevents flooding and over compensation. This should ONLY be used when the DF bit is set.

RFC Reference: N/A.

6.6.6.17 Feature: QoS Translation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For QoS consistency, the Traffic Class of the inbound IPv6 packet should be copied into the DSCP field of the outbound IPv4 packet. The DSCP of the inbound IPv4 packet should be copied into the Traffic Class field of the outbound IPv6 packet.

Function: For further study.

RFC Reference: N/A.

6.6.6.18 Feature: Port Reservation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For further study.

Function: For further study.

RFC Reference: N/A.

6.6.6.19 Feature: Static Port Forwards

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For further study.

Function: For further study.

RFC Reference: N/A.

6.6.6.20 Feature: PCP Failure

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function and reactivity during a failure of the session either at Layer 3, 4 and 5 at the client, server, PLAT or CPE.

Function: Listed below:

1) Relay of failure messages from PLAT to CPE receive
2) Relay of failure messages from CPE to PLAT send
3) Lifetime timer configuration for failure; PCP entry timers
4) Monitoring failure of the of the PCP without message
5) Static mappings are not affected (comparative if the CPE holds those static explicit mappings)
6) PCP table content considering failure (last received packet). PCP table matching when/if required

RFC Reference: N/A.

6.6.6.21 Feature: PCP Fixed port Torrents

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function

Function: Listed below:

1) Relay of failure messages from PLAT to CPE receive

RFC Reference: N/A.

6.6.6.22 Feature: PCP Based Capacity GW Orientation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function.

Function: Listed below:

1) Relay of failure messages from PLAT to CPE receive

RFC Reference: N/A.

6.6.7 Monitoring and Management

6.6.7.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.6.7.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.6.7.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.6.7.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.6.7.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv6.

RFC Reference: [76**]**.

6.6.7.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18**]**.

6.6.7.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.6.7.8 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26].

6.6.7.9 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.6.7.10 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60**]**.

6.6.7.11 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.6.7.12 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.6.7.13 Feature: remote access: TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.6.8 Support System Requirements

6.6.8.1 Feature: DHCPv6 Server

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function.

Function: For further study.

RFC Reference: N/A.

6.6.8.2 Feature: DNS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function.

Function: For further study.

RFC Reference: N/A.

6.6.9 CPE Requirements

In the context of 464XLAT, the customer premise equipment has to operate as a customer-side translator device (CLAT). The primary purpose of the CLAT is to translate private IPv4 from devices on local area network interfaces to IPv6 on its wide area network interface. The CLAT function within the CPE device interoperate with a corresponding provider-side translator (PLAT) which translates from the IPv6 traffic received from the CPE device to the global IPv4 address space.

6.6.9.1 Feature: CLAT functionality

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Customer Premise IPv4 translation to IPv6.

Function: The CPE device has to conform to RFC 6145 [124].

RFC Reference: [124**]**.

6.6.9.2 Feature: Native IPv6 Support

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT CPE has to support the use of native IPv6 addresses on its local area network interfaces.

Function: The CLAT device has to implement a native IPv6 router device and all locally derived IPv6 traffic has to bypass the CLAT functionality for transit directly via the operator IPv6 network.

RFC Reference: N/A.

6.6.9.3 Feature: DHCP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT CPE device needs to provide addressing to client devices on all locally connected network interfaces.

Function: The following DHCP server functionality has to be implemented:

1) DHCPv4 for locally connect IPv4 clients

2) DHCPv6 for locally connected IPv6 devices

RFC Reference: N/A.

6.6.9.4 Feature: DNS

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT CPE device needs to provide DNS translation services to client devices on locally connected network interfaces to minimize excessive translation of DNS queries.

Function: The CLAT CPE device has to implement the following DNS proxy for IPv4 hosts connected to the CPE device locally area networks. The CLAT device.

RFC Reference: [110**]**.

6.6.9.5 Feature: NDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT device when implemented in a gateway device has to provide normal gateway function service.

Function: The CLAT device has to provide Neigbour Discovery Protocol with router advertisement.

RFC Reference: [96**]**.

6.6.9.6 Feature: IPv6 Prefix handing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT device requires awareness of multiple IPv6 prefixes.

Function: The CPE device has to obtain separate IPv6 prefixes for the following:

1) A /64 Prefix for its network side interface.

2) A /64 Prefix for any locally connected network interfaces.

3) A /64 Prefix for the exclusive purpose of sending and receiving packets that have been statelessly translated by its upstream PLAT router.

RFC Reference: N/A.

6.6.9.7 Feature: CLAT discovery of PLAT IPv6 prefix destination

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CLAT device.

Function: The CLAT device has to discover the PLAT-side IPv6 prefix that is used as the destination of the PLAT, it MAY discover the translation prefix via the following means:

1) DHCPv6 option messaging when defined.

2) [Discovery-Heuristic] Savolainen, T., Korhonen, J., and D. Wing, "Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis", Work in Progress, March 2013 [141].

RFC Reference: N/A.

6.6.9.8 Feature: CLAT Intercommunication

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Where peer to peer traffic between CLAT enabled CPE devices on a service provider is required, a means of minimizing excessive translation should be provided.

Function: The CLAT device has to implement Interactive Connectivity establishment (ICE).

RFC Reference: [103**]**.

6.6.9.9 Feature: User Interface Provision

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE has to implement a user accessible web interface on any implemented Local Area Network interface to aid the configuration of the device and to aid in provision of diagnostics information.

Function: The CPE device has to implement a user accessible web interface on port 80.

RFC Reference: N/A.

6.6.9.10 Feature: Access Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement access control to the local web interface.

Function: Access to the web interface has to be controlled by the use of a specified username and password.

RFC Reference: N/A.

6.6.9.11 Feature: Device Localization

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device user interface has to be configurable for Multiple languages.

Function: The language displayed on the user accessible web interface has to be configurable to support display in multiple languages.

RFC Reference: N/A.

6.6.9.12 Feature: CPE Device Status Indication

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE Device has to contain a user interface tab which provides access to the following item.

Function: Listed below:

1) Connection Status
2) Security
3) Diagnostics
4) Software

RFC Reference: N/A.

6.6.9.13 Feature: Router Basic Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Wide Area network Configuration
2) Local Area Network Configuration
3) DHCP Server Configuration
4) Device Configuration Backup

RFC Reference: N/A.

6.6.9.14 Feature: Advanced router control and configuration.

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Advanced Options
2) IP Filtering
3) MAC Filtering

4) Port Filtering

5) Port Forwarding

6) Port Triggers

7) DMZ Host Selection

RFC Reference: N/A.

6.6.9.15 Feature: The CPE device has to implement a Firewall user interface with user accessible configuration control

Premise: A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log.

Function: Listed below:

1) Web Filtering

2) Local Logging

3) Remote Logging

RFC Reference: N/A.

6.6.9.16 Feature: User Interface - Parental Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a specific user accessible tab which provides access to the following functions

Function: Listed below:

1) User Setup

2) Basic Setup

3) Content Filtering

4) Time of Day Access Control

RFC Reference: N/A.

6.6.9.17 Feature: User Interface, Wireless Status and configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Where the CPE device implements a wireless interface a specific user accessible tab has to be made available.

Function: The Wireless configuration tab has to include the following items:

1) Radio Interface

2) Security

3) Advanced Functionality

4) Access Control

RFC Reference: N/A.

Premise: User Interface -Where the CPE device implements an integrated Multimedia Terminal Adaptor a specific user configurable tab has to be made available.

Function: The MTA tab has to include the following items:

1) Status of the integrated MTA functionality
2) DHCP status, including any additional interfaces configured for the use of the MTA device
3) Quality of Service
4) Device provisioning
5) Event Logging

RFC Reference: N/A.

6.6.10 Technical Viability

As with all NAT solutions, 464XLAT has restrictions on inbound connections. 464XLAT also has a restricted payload due to the increased IP header used for IPv6. Fragmentation is more likely to occur.

As the NAT functionality is removed from the CPE, existing features such as UPnP NAT traversal and port forwarding can no longer be used to direct inbound packets on predetermined IPv4 TCP or UDP ports to a specific host.

Application Level Gateways on the CPE cannot be used when 464XLAT is employed. The PLAT has to provide the additional port mapping requirements for each required ALG.

Other issues are mostly well-known and present topics for further development:

- The cost, particularly of the CPE (CLAT), is expected to be increased due to the extension of the required functionality. MSOs that are not deploying WiFi solutions and stick to bridging CPE devices will face the issue of having to exchange CPE just for the reason of introducing 464XLAT.
- 464XLAT requires ALGs (ALPs to be exact in most cases) which require the PLAT to do some form of intelligent vicissitudes to the transit packets bearing the risk of drop in performance and limits in functionality. The risk of service-deprecation is minimized if all functionality is included natively in the PLAT.
- With 464XLAT, all functional intelligence is located in the PLAT. Thus, functionality that requires a public address in the local network has to be performed in the PLAT itself. An example of this is PCP. In 464XLAT, it is placed on the PLAT due to UPnP 1 and 2 requirements.
- To predict scaling requirements comparative to IPv6 utilization can be misleading. This is not specific to 464XLAT but occurs similarly in other transition technologies. The issue is that it is largely unknown what portion of the traffic originated by a certain client will be IPv6. Thus, the system has to be scaled for a constant maximum of all clients and designs cannot adapt based on exact traffic predictions.
- 464XLAT is an address family translation technology and, thus, suffers from MTU requirements beyond the norm. This means that MSS clamping, IPv4 fragmentation and IPv6 fragmentation are required in order to avoid ICMP blocking. Fragmentation resends and general PMTU control can have performance effects on customer services. Therefore, there is a major requirement to optimize implementations such that they do not cause a large impact on current RTTs and node latency.
- The requirement for Data Retention capacity is rather large and, thus, making the capacity available may become a major issue for the receiving logging server. Normally, on public addresses there is a single entry or two per week depending on lease times and utilization. Within 464XLAT, every time an individual subscriber receives a port and IP address assignment the event has to be logged. This can amount to 200 assignments a day with a start and stop timer resulting in 400 logs per user per day. Assuming that an MSO could have 10 % of its customer base online at one time, that would mean a huge amount of logs.
- 464XLAT is a hub-and-spoke topology forcing all communication from CLAT to CLAT to pass the PLAT

On the other hand 464XLAT has some advantages over NAT44 CGN. 464XLAT allows for a smooth migration towards an IPv6 only operator infrastructure. This includes IPv6 only access, aggregation and core network as well as an IPv6 only BNG. In case of a cable operator the BNG is implemented on the CMTS. In case of other access technologies (e.g. GPON) the BNG is implemented on a dedicated platform. 464XLAT also removes the need for IPv4 address management and as such a DHCPv4 server is not required anymore.

As a PLAT only handles IPv4 traffic, the capacity requirements per customer on the PLAT will likely decrease over time. The reason is that more and more internet traffic will be on IPv6. This might allow for operator growth over time without the need for large investments on the PLAT.

6.7 MAP-E

MAP-E is a limited development at the moment with no known products in support of MAP-E at present. There are implementations coming but they are proprietary solutions as the RFC and general standardization is not yet complete, with the MAP-E RFC in its first draft. Note that due to this the specification on MAP-E is not a complete one as the structure is yet to be built.

6.7.1 MAP-E/MAP-T Technology Summary

Mapping of Address and Port - Encapsulation Mode (MAP-E) is a combination of tunnel and translation technologies. Prior to being standardized, MAP-E was commonly referred to as IPv4 Residual Deployment (4rd). MAP-E enables a service provider to rapidly allow IPv4 services to IPv6 (customer) sites to which it provides customer premise equipment (CPE). This approach utilizes stateless IPv4-in-IPv6 encapsulation (i.e. tunneling) to transit IPv6-enabled network infrastructure. The encapsulation SHOULD be supported by the CPE and MAP-E Gateway/Border Relay, which removes the IPv6 encapsulation from IPv4 packets while forwarding them to the Internet. The provider access network can now be on IPv6, while customers see IPv6 and IPv4 service simultaneously.

MAP-E uses the IPv4 / IPv6 address mapping technique "Mapping of Address and Port (MAP)" IETF draft-mdt-softwire-mapping-address-and-port-03 [142], to associate the IPv4 address with an IPv6 address. This allows the automatic creation of tunnels between the Border Router and the CPE Router.

MAP-E also allows the provisioning of an IPv4 address and port range to the CPE Router across the IPv6 network.

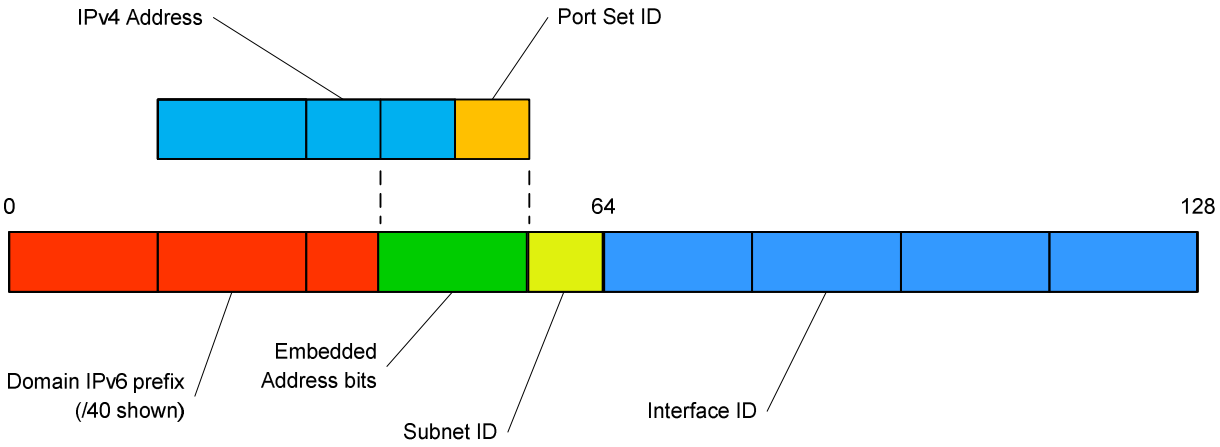


Figure 5: IPv6 address for MAP address and port mapping

Provisioning of the CPE router requires the following attributes:

- IPv6 prefix - Note that the IPv6 prefix is not dedicated for use with MAP-E, and can be part of the customer-assigned IPv6 prefix used for native IPv6 traffic.
- Embedded Address bits - These form the public IP address and port range.
- MAP-E Border Relay IPv6 prefix.
- DHCPv6 options defined in IETF draft-mdt-softwire-map-dhcp-option-03 [143].

The Border Relay has to be configured with a public IPv4 address and the IPv6 prefix for translation, in addition to its own IPv6 prefix.

For each outbound packet, the CPE Router is required to perform NAT-44 translation from the private (RFC 1918 [30]) address to the address provided by the Service Provider, and a port within the range specified. Once translated to a public IPv4 address, the CPE Router then creates the mapped IPv6 packet header in accordance with the stateless IPv4/IPv6 translation mechanism defined in RFC 6145 [124], and using the MAP header mapping rules. The translated packet is then sent across the IPv6 network to the MAP-T Border Relay.

The Border Relay is responsible for translating the received packet back to IPv4, recovering the original IPv4 addresses from the IPv6 header, using the MAP header mapping and RFC 6145 [124] translation technique, to forward on to the destination.

Inbound IPv4 packets are received by a Border Relay. The Border Relay translates the IPv4 header into IPv6, using the same RFC 6145 [124] translation, and MAP header mapping to produce the IPv6 packet, to forward to the CPE router. The CPE router uses the same translation mechanism to retrieve the IPv4 header, and performs regular NAT-44 to pass the packet on to the originating host.

Where an IPv6 server is located between the CPE router and Border Router, this is accessible

As the packet size varies between IPv4 and IPv6, MTU size and fragmentation are handled by the IP/ICMP translation. The use of IPv4 PMTUD will result in ICMP 'packet too big' messages. TCP MSS Clamping is additionally required to ensure the overall IPv4 packet size does not exceed the IPv6 path Maximum Transmission Unit.

The MAP-E Border Relay is responsible for fragmentation and reassembly of any IPv4 packet received prior to the encapsulation to IPv6.

Data Retention/Lawful Intercept systems will be required to capture the mapped IP address and port range for each customer.

6.7.2 Technology Feasibility Synopsis

MAP-E, out of the two technologies, does have technical benefits over centralized CGN solutions such as NAT44 and DS-Lite, but these do not out way the present drawbacks. At present only a few carrier grade vendors are road-mapping MAP into their code and thus it is not an industry wide transition technology. It basically took too long for the delivery to be developed as the replacement for the next stage of technologies such as dIVI and 4rd. With low adoption the development of the technology and lack of interest in adding further complexity and development time into the IPv6 transitional world it is unlikely to present itself in the near future, if at all, as a viable option for most operators to consider for deployment in live to support customers on their journey into IPv6 only experience. As a technology it is considered a possible "nice to have" once developed completely as it has some nice positives but in no means has a future of a non-self-deprecating service unfortunately. Transition from native IPv4 to MAP-E would certainly heavily degrade the customer's experience in comparison to other possible technologies.

Negatives:

- State is held on the CPE - this can be considered both a positive and negative, but the negative aspect of this is this increases the process and memory requirement on the CPE comparative to some other technologies. The other major problem with this is configuration changes and deployment tweaking requires you to update every single CPE using MAP. With all taken into account within the cable industry the modem is our largest cost and complex in development with DOCSIS considering what it is required to do as such a small device, so adding cost due to memory and processing CPE makes the technology potentially less viable and more.
- Performance on the CPE has shown in some cases to be slower using NAT than not.
- Vendor adoption - Vendors in general are not widespread and thus fixed contracts with specific vendors would be required to deliver this into the network. As the vendors are limited in adoption there is little we can do to.
- Standardization - the present state of standardization for MAP-E and MAP-T, although some RFCs are complete, there are a few drafts that yet need to be complete to allow for adoption by vendors.
- Mathematical relationship between IPv4 and Ipv6 addresses creates a sub-optimal solution.
- Fixed port block requirement - this limits your ability to dynamically assign port blocks and thus your scaling.
- Requirement for reassembly on the BR - this forces state on the BR and can cause return path issues removing the benefits of a partially stateless deployment within the CORE.

- No static port bindings which limits service.
- ICMP forwarding issues.
- DR could become an issue dependent on legal requirements.

Positives:

- No centralization like in CGN, however this is mitigated in CGN with the possibility of line card integration into present routers already on your network and using Anycast.
- Almost stateless which allows for less requirements for processing, memory and general resource allocation to the transition processes except for on the CPE.

6.7.3 MAP-E/MAP-T Technology Feature Summary

Table 11 gives a summary of the MAP-E and MAP-T technology feature.

Table 11: MAP-E and MAP-T Technology Feature Summary

Functional Name	Requirement	Description
MAP-E [147]	Required	Main MAP-E, mapping of address and port - encapsulation version, RFC draft.
MAP-T [148]	Optional	Main MAP-T, mapping of address and port - translation version, RFC draft.
MAP-T RFC 6145 [124]	Optional	Compliance with RFC 6145 [124] "IP/ICMP Translation Algorithm".
Compliance with RFC 6052 [119]	Required	Compliance with "IPv6 Addressing of IPv4/IPv6 Translators.
Compliance with [15], [17]	Required	Port and address mapping derivative
Shared Resource	Required	Single BR GW address. The BR IPv6 address should be able to be shared amongst different NPU's in the BR. A hashing mechanism should be in place to hash all upstream packets based on the source IPv6 address.
BR Addressing and Virtual Interfaces	Required	BR SHALL be able to assign a single virtual interface with up to 8 Map-E GW addresses for any given BR instance on the node.
Anycast	Required	Anycast MAP-E gateway prefixes are a requirement to allow simplicity of deployment for a single prefix across multiple BRs.
BR Address withdrawal	Required	The BR should have at least four points of BR GW address withdrawal occurrence. The list includes: - loss of IPv6 route out, - loss of all BGP/IGP sessions, - loss of forwarding, - loss of NPU capacity Any of the failures should be detectable based on configurable timers with 15 seconds being the default setting.
Basic mapping rule	Required	Compliance with: - draft-mdt-softwire-mapping-address-and-port-02.txt [149] - draft-murakami-softwire-4v6-translation-00 [150]
Forwarding mapping rule	Required	Forwarding rules
Tunnel MTU sizing	Required	The MTU physical and protocol requirements for MAP-E tunnels compliance per draft [147].
Fragmentation	Required	The IPv4 and IPv6 fragmentation requirements for MAP-E.
Re-ordering and Buffering out-of-sync packets	Required	The requirement to hold/wait on all ordered packets for each flow to forward.
Customer CPE ID	Required	IPv4 Private and IPv6 address.
Support for 1:1 redundant NPU card	Required	The NPU SHALL not only be able to do full-cone deployment but also 1:1 mapping per NPU.
Configuration of the size of consecutive blocks and sharing ratio	Required	MAP-E requires fixed blocks to be configured on the CPE, this SHALL be configurable through the boot file.
IPv6 ACLs	Required	IPv6 ACLs before the MAP-E processing on the BR.
IPv4 ACLs	Required	The ability for ACLs to be applied before the decapsulation of the IPv6 packet.
UDP-Lite	Required	UDP-lite SHOULD be supported.
TOS rewrite when translating IPv6 to IPv4 and from IPv4 to IPv6	Required	Marking and Remarking SHOULD be supported for all tunnelled traffic.
TCP MSS definition	Required	MSS clamping to enable MSS value on all packets ingress and egress.
Performance Requirement		Description
24 MAP-E instances per BR	Required	The ability to stack transition instances on top of one another.

Functional Name	Requirement	Description
40m Translations (per npu) - black assignment	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary.
Minimum 50 Gbps bi-directional throughput per npu with IMIX traffic	Required	Per half slot throughput requirement.
1M+ primary connections per second setup rate	Required	Primary sessions, port block assignments, per NPU.
500k users per NPU	Required	The BR has far less requirement for processing but it SHOULD be able to handle a minimum of 500 000 users per chassis or NPU.
Latency	Required	Latency is between 40 and 200 micro seconds (µs).
6 npu Cards per chassis	Required	If the BR requires specific NPU service processing units for MAP-E function then 6 NPUs per node SHOULD be the minimum requirement.
IRB/SVI support	Required	Integrated Routing and Bridging/Virtual Interfaces (L3 interface for Bridge Domain)
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 25 6k NAT users at the same time.

6.7.4 CORE (BR/GW)

Void.

6.7.5 Hardware/Software Requirements

6.7.5.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU in a full stateful fashion with no loss of packets or latency beyond 1ms. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1 ms failover maximum for any given B4 connectivity of session

RFC Reference: N/A.

6.7.5.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CGN-GW has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the CGN-GW can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.7.5.3 Feature: RP / RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within an CGN-GW has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.7.5.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given CGN-GW has to be shared giving the appearance of a single node on the network.

Function: The CGN-GW has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single address per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.7.5.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given CGN-GW has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given CGN-GW to see all or part of the traffic coming into the node. This can be determined through a single CGN-GW address, so CGN-GW GW assigned or NAT group assigned under multiple GW addresses. Hashing SHOULD be used to determine the load balancing between all processing/hardware forwarding points with that NAT group or deterministic specification.

RFC Reference: N/A.

6.7.6 Performance Requirements

6.7.6.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1ms for any given function of the CGN-GW.

Function: Within the remit of traffic requiring transition function, decapsulation, encapsulation and NAT function, the CGN-GW has to perform its function from ingress interface to egress interface in a time measured no higher than 1ms. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions.

RFC Reference: N/A.

6.7.6.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.7.6.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the CGN-GW Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 gig is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.7.6.4 Feature: Min Sessions per BR Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis.

RFC Reference: N/A.

6.7.6.5 Feature: Min Customer B4s per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus IPv6 B4 addresses that can be assigned from a single Chassis or connected.

Function: Due to the scaling of any given CGN-GW solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 B4 addresses. This is solely a minimum and the number may change with the introduction of Private v4 and public v6 Si-ID assignment.

RFC Reference: N/A.

6.7.7 MAP-E Feature Specification

6.7.7.1 Feature: MAP-E Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Addressing structure of MAP-E with the embedded IPv4 address.

Function: Address structured based on BMR.

RFC Reference: [15**]**.

From section 7.2 in [15]:

> "The MAP BR has to be configured with the same MAP elements as the MAP CEs operating within the same domain.
> For increased reliability and load balancing, the BR IPv6 address MAY be an anycast address shared across a given MAP domain. As MAP is stateless, any BR may be used at any time. If the BR IPv6 address is anycast the relay has to use this anycast IPv6 address as the source address in packets relayed to CEs.
>
> Since MAP uses provider address space, no specific routes need to be advertised externally for MAP to operate, neither in IPv6 nor IPv4 BGP. However, if anycast is used for the MAP IPv6 relays, the anycast addresses has to be advertised in the service provider's IGP."

From section 6 in [15]:

> "The Interface identifier format of a MAP node is described in Figure 6.

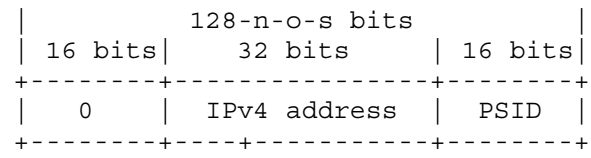


Figure 6: Interface Identifier Format of a MAP node

> In the case of an IPv4 prefix, the IPv4 address field is right-padded with zeroes up to 32 bits. The PSID field is left-padded to create a 16 bit field. For an IPv4 prefix or a complete IPv4 address, the PSID field is zero.
>
> If the End-user IPv6 prefix length is larger than 64, the most significant parts of the interface identifier is overwritten by the prefix."

From section 7 in [15]:

> " For a given MAP domain, the BR and CE has to be configured with the following MAP elements. The configured values for these elements are identical for all CEs and BRs within a given MAP domain.
>
> - The Basic Mapping Rule and optionally the Forwarding Mapping Rules, including the Rule IPv6 prefix, Rule IPv4 prefix, and Length of EA bits
> - The IPv6 address of the MAP BR.
> - Hub and spoke mode or Mesh mode. (If all traffic should be sent to the BR, or if direct CE to CE traffic should be supported)."

6.7.7.2 Feature: IPv6 Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are no specific requirements for IPv6 LAN addressing to support MAP-E. Native IPv6 end-to-end connectivity is required for MAP-E to operate.

Feature: Packet Encapsulation.

Premise: IPv4 encapsulation into IPv6 bearer traffic.

Function: The BR has to encapsulate IPv4 packets with a public IPv4 destination, in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

The IPv4 packet has to be translated using NAPT-44 to a free port within the provisioned Port Set Range.

The IPv6 destination address is created using the MAP Border Relay IPv6 prefix, and the encoded destination IPv4 address.

The IPv6 source address is created using the IPv6 prefix, and the address and port mapping rules.

The router MAP-E interface has to drop any packets in the following groups:

1) Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].

2) Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].

3) Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].

4) Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: [15**]**.

6.7.7.3 Feature: Packet Decapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The BR SHOULD decapsulate embedded IPv4 traffic.

Function: The BR SHOULD decapsulate packets in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

Where an IPv4 fragment is received, the BR has to forward the fragment unaltered to its destination. It is the responsibility of the destination client to reassemble fragments.

Once the IPv4 packet is recovered, it has to forward the packet to the destination specified in the source address.

RFC Reference: N/A.

6.7.7.4 Feature: MTU Size and Fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Encapsulation of an IPv4 packet with an IPv6 header to transport across the cable network increases the overall packet size by 40 bytes.

The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes. The maximum IPv4 packet size that can be sent between the MAP-E CPE Border Router is therefore 1 460 bytes.

Function: Because fragmentation in IPv6 has to be performed by the sending host, any packet received by the BR that is in excess of this size SHOULD either:

1) Be encapsulated by the BR, and fragmented into two packets for transmission to the CPE, or

2) Be reduced in size to a maximum of 1 460 bytes before being encapsulated.

draft-ietf-softwire-map-02 [144] recommends the latter approach is taken.

The MTU of the BR encapsulation interface has to be reduced to 1 460 bytes. This will have three effects:

1) Any destination or source that uses Path MTU Discovery (PMTUD) [22], [32] will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

2) Any IPv4 packet received by the BR that is in excess of 1 460 bytes, with the DF bit clear, SHOULD be fragmented by the BR before encapsulation. The IPv6 packet will therefore fit within the IPv6 MTU of 1 500 bytes.

3) Any IPv4 packet received by the BR that is in excess of 1 460 bytes, with DF set SHOULD be dropped by the BR, and an ICMP 'packet too big' message returned to the client. The client is expected to process this message and reduce the packet to the size specified in the ICMP reply.

RFC Reference: [22**]**, [32].

6.7.7.5 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header).

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

Function: MSS clamping has to be implemented on the BR encapsulation interface to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size). i.e. The device SHOULD support MSS clamping to 1 420 bytes.

RFC Reference: N/A.

6.7.7.6 Feature: Future review of Maximum MTU size

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate MAP-E encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU has to be set to allow for full IPv6 encapsulated MAP-E packets to be sent with no fragmentation.

RFC Reference: N/A.

6.7.7.7 Feature: MAP- E Inbound Session control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: With MAP-E, the ability to control inbound sessions is limited to the port-set range that has been assigned to the CPE. These are generally:

- DMZ host functionality. Where all inbound packets can be forwarded to a single client
- Port forwarding. Where individual TCP or UDP ports can be forwarded to a specific client

- UPnP NAT Traversal. Where a client requests for the NAT to open one or more TCP or UDP ports to itself for inbound packets

Function: The cable BR SHOULD allow inbound TCP or UDP connections within the port-set range that has been provisioned to the CPE. The BR SHOULD NOT allow configuration of ports outside of its assigned port-set range as provisioned by any allowable means such as DHCPV6 options, TR-069 or vendor specific MIB values.

RFC Reference: [16**]**.

6.7.7.8 Feature: Packet Encapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv4 encapsulation into IPv6 bearer traffic

Function: The BR SHOULD encapsulate IPv4 packets with a public IPv4 destination, in accordance with the specification defined in [15].

The IPv4 packet SHOULD be translated using NAPT-44 to a free port within the provisioned Port Set Range.

The IPv6 destination address is created using the MAP Border Relay IPv6 prefix, and the encoded destination IPv4 address.

The IPv6 source address is created using the CPE IPv6 prefix, and the address and port mapping rules.

The BR MAP-E interface SHOULD drop any packets in the following groups:

- Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].
- Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].
- Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].
- Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: [16].

6.7.7.9 Feature: MTU Size and Fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Encapsulation of an IPv4 packet with an IPv6 header to transport across the cable network increases the overall packet size by 40 bytes.

The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes. The maximum IPv4 packet size that can be sent between the MAP-E CPE and Border Router is therefore 1 460 bytes.

Function: Because fragmentation in IPv6 SHOULD be performed by the sending host, any packet received by the cable modem/ eRouter that is in excess of this size SHOULD either:

- Be encapsulated by the CPE, and fragmented into two packets for transmission to the Border Router, or
- Be reduced in size to a maximum of 1 460 bytes before being encapsulated.

draft-ietf-softwire-map-02 [144] recommends the latter approach is taken.

The MTU of the CPE router encapsulation interface SHOULD be reduced to 1 460 bytes. This will have three effects:

1) Any client that uses Path MTU Discovery (PMTUD) [22], [32] will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

2) Any IPv4 packet received by the BR that is in excess of 1 460 bytes, with the DF bit clear, SHOULD be fragmented by the BR before encapsulation. The IPv6 packet will therefore fit within the IPv6 MTU of 1 500 bytes.

3) Any IPv4 packet received by the CPE that is in excess of 1 460 bytes, with DF set SHOULD be dropped by the CPE, and an ICMP 'packet too big' message returned to the client. The client is expected to process this message and reduce the packet to the size specified in the ICMP reply.

RFC Reference: [15**]**, [22], [32].

6.7.7.10 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header).

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

Function: MSS clamping SHOULD be implemented on the BR encapsulation to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size). i.e. The device SHOULD support MSS clamping to 1 420 bytes.

RFC Reference: N/A.

6.7.7.11 Feature: Future review of Maximum MTU size

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate MAP-E encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU SHOULD be set to allow for full IPv6 encapsulated MAP-E packets to be sent with no fragmentation.

RFC Reference: N/A.

6.7.8 Monitoring and Management

6.7.8.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.7.8.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.7.8.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.7.8.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.7.8.5 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18**]**.

6.7.8.6 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.7.8.7 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26**]**.

6.7.8.8 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.7.8.9 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60].

6.7.8.10 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.7.8.11 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.7.8.12 Feature: Remote Access - TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.7.9 MAP-E CPE Specification

6.7.9.1 Feature Device Provisioning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: MAP Provisioning.

Function: The MAP device has to obtain device provisioning via the use of a specified DHCPv6 option parameter. The MAP device MAY obtain this device provisioning information via alternative means.

RFC Reference: [16**]**.

6.7.9.2 Feature: WAN Connectivity

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Whilst there are no direct requirements for the CMTS to be able to support MAP-E, it is required that native IPv6 end-to-end connectivity is provided. The CMTS has to therefore be configured with IPv6 on both the cable and upstream interfaces.

Function: The cable modem / eRouter has to be provided with a GUA /128 for its public WAN interface and at least one /64 prefix delegation for client usage.

RFC Reference: N/A.

6.7.9.3 Feature: Provisioning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To support the provision of MAP-E, some additional attributes has to be provisioned on the CPE:

Function: The following attributes has to be provided, At present, there is a request for a DHCPv6 option for providing these attributes. IETF draft MAP DHCPv6 options [145]. It is recommended that the DHCPv6 option is used. The CPE MAY obtain these additional parameter by alternate means such as SNMP or TR-069 based provisioning:

1) Publically routable IPv4 address
2) Port Set Range
3) Port Set ID
4) Border Relay IPv6 prefix

RFC Reference: [16**]**.

6.7.9.4 Feature: Provisioning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv6 server support.

Function: For the CPE to receive the required provisioning parameters via DHCPv6, the DHCPv6 server has to support the provisioning of the attributes as specified in (IETF draft MAP DHCPv6 options [145].

RFC Reference: [16**]**.

6.7.9.5 Feature: Cable eRouter B4 functionality

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Because the CPE is only provided with IPv6 connectivity on its WAN interface, it has to perform all DNS resolution over IPv6.

Function: To reduce the amount of encapsulated traffic, the cable B4 gateway has to be configured as a DNS forwarder for all LAN client requests. Any client DNS requests received via IPv4 has to therefore be forwarded as IPv6 to the external DNS servers.

RFC Reference: N/A.

6.7.9.6 Feature: LAN Addressing - IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are no specific requirements for IPv6 LAN addressing to support MAP-E. Native IPv6 end-to-end connectivity is required for MAP-E to operate.

Function: IPv6 addressing SHOULD be used for all LAN clients that support IPv6.

RFC Reference: N/A.

6.7.9.7 Feature: LAN Addressing - IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Provision of local LAN addressing for IPv4 client devices.

Function: The CPE Gateway has to provide 'private' IP addressing from the ranges specified in RFC 1918 [30]. The accepted normal configuration is to use the 192.168.0.0/16 range, with a single /24 subnet.

The CPE router has to use one address from the selected IPv4 subnet as its LAN interface address, and this has to be used as the default gateway for LAN IPv4 traffic.

The CPE router has to also use its LAN IPv4 address as the DNS server address in DHCP offers to clients.

RFC Reference: [30**]**.

6.7.9.8 Feature: Packet Encapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv4 encapsulation into IPv6 bearer traffic

Function: The CPE router has to encapsulate IPv4 packets with a public IPv4 destination, in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

The IPv4 packet has to be translated using NAPT-44 to a free port within the provisioned Port Set Range.

The IPv6 destination address is created using the MAP Border Relay IPv6 prefix, and the encoded destination IPv4 address.

The IPv6 source address is created using the CPE IPv6 prefix, and the address and port mapping rules.

The CPE router MAP-E interface has to drop any packets in the following groups:

1) Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].

2) Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].

3) Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].

4) Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: [15**]**.

6.7.9.9 Feature: Packet Decapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The cable modem / e-router has to decapsulate embedded IPv4 traffic.

Function: The CPE router has to decapsulate packets in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

Where an IPv4 fragment is received, the B4 router has to forward the fragment unaltered to its destination. It is the responsibility of the destination client to reassemble fragments.

Once the IPv4 packet is recovered, it has to forward the packet to the client specified in the source address.

RFC Reference: N/A.

6.7.9.10 Feature: MTU Size and Fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Encapsulation of an IPv4 packet with an IPv6 header to transport across the cable network increases the overall packet size by 40 bytes.

The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes. The maximum IPv4 packet size that can be sent between the MAP-E CPE and Border Router is therefore 1 460 bytes.

Function: Because fragmentation in IPv6 has to be performed by the sending host, any packet received by the cable modem/ eRouter that is in excess of this size has to either:

1) Be encapsulated by the CPE, and fragmented into two packets for transmission to the Border Router, or

2) Be reduced in size to a maximum of 1 460 bytes before being encapsulated.

draft-ietf-softwire-map-02 [144] recommends the latter approach is taken.

The MTU of the CPE router encapsulation interface has to be reduced to 1 460 bytes. This will have three effects:

1) Any client that uses Path MTU Discovery (PMTUD) [22], [32] will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

2) Any IPv4 packet received by the CPE that is in excess of 1 460 bytes, with the DF bit clear, has to be fragmented by the CPE before encapsulation. The IPv6 packet will therefore fit within the IPv6 MTU of 1 500 bytes.

3) Any IPv4 packet received by the CPE that is in excess of 1 460 bytes, with DF set has to be dropped by the CPE, and an ICMP 'packet too big' message returned to the client. The client is expected to process this message and reduce the packet to the size specified in the ICMP reply.

RFC Reference: [15**]**, [22], [32].

6.7.9.11 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header).

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

Function: MSS clamping has to be implemented on the CPE encapsulation interface of the cable modem/e-router to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size). i.e. The device has to support MSS clamping to 1 420 bytes.

RFC Reference: N/A.

6.7.9.12 Feature: Client MTU reduction via DHCPv4 option

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv4 option 26 allows the specification of the maximum IPv4 MTU sent by IPv4 client devices on the local area network as defined in RFC 2132 [33].

Function: The cable modem/ e-router MAY implement DHCPv4 option 26 to set the maximum client MTU size, however this approach has the following limitations.

Not all client IP stacks and DHCP clients will respect the DHCP option to set the interface MTU.

Devices that are manually configured (and hence do not use DHCP) would also require manual configuration of the interface MTU.

Reducing the interface MTU of clients would also impact all LAN IPv4 traffic as this would also all be reduced to an MTU of 1 460 bytes.

It cannot be guaranteed that all client applications will function if the MTU has been reduced.

RFC Reference: [33**]**.

6.7.9.13 Feature: Future review of Maximum MTU size

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate MAP-E encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU has to be set to allow for full IPv6 encapsulated MAP-E packets to be sent with no fragmentation.

RFC Reference: N/A.

6.7.9.14 Feature: MAP- E Inbound Session control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: With MAP-E, the ability to control inbound sessions is limited to the port-set range that has been assigned to the CPE. These are generally:

- DMZ host functionality. Where all inbound packets can be forwarded to a single client
- Port forwarding. Where individual TCP or UDP ports can be forwarded to a specific client
- UPnP NAT Traversal. Where a client requests for the NAT to open one or more TCP or UDP ports to itself for inbound packets

Function: The cable modem / e-router SHOULD allow inbound TCP or UDP connections within the port-set range that has been provisioned to the CPE. The cable modem/e-router is not allowed to configurate ports outside of its assigned port-set range as provisioned by any allowable means such as DHCPV6 options, TR-069 or vendor specific MIB values.

RFC Reference: [16**]**.

6.7.9.15 Feature: User Interface Provision

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE has to implement a user accessible web interface on any implemented Local Area Network interface to aid the configuration of the device and to aid in provision of diagnostics information.

Function: The CPE device has to implement a user accessible web interface on port 80.

RFC Reference: N/A.

6.7.9.16 Feature: Access Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement access control to the local web interface.

Function: Access to the web interface has to be controlled by the use of a specified username and password.

RFC Reference: N/A.

6.7.9.17 Feature: Device Localization

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device user interface has to be configurable for Multiple languages.

Function: The language displayed on the user accessible web interface has to be configurable to support display in multiple languages.

RFC Reference: N/A.

6.7.9.18 Feature: CPE Device Status Indication

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE Device has to contain a user interface tab which provides access to the following item.

Function: Listed below:

1) Connection Status
2) Security
3) Diagnostics
4) Software

RFC Reference: N/A.

6.7.9.19 Feature: Router Basic Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Wide Area network Configuration
2) Local Area Network Configuration
3) DHCP Server Configuration
4) Device Configuration Backup

RFC Reference: N/A

6.7.9.20 Feature: Advanced router control and configuration.

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Advanced Options
2) IP Filtering
3) MAC Filtering
4) Port Filtering
5) Port Forwarding
6) Port Triggers
7) DMZ Host Selection

RFC Reference: N/A.

6.7.9.21 Feature: The CPE device has to implement a Firewall user interface with user accessible configuration control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log.

Function: Listed below:

1) Web Filtering
2) Local Logging
3) Remote Logging

RFC Reference: N/A.

6.7.9.22 Feature: User Interface - Parental Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a specific user accessible tab which provides access to the following functions.

Function: Listed below:

1) User Setup
2) Basic Setup
3) Content Filtering
4) Time of Day Access Control

RFC Reference: N/A.

6.7.9.23 Feature: User Interface, Wireless Status and configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Where the CPE device implements a wireless interface a specific user accessible tab has to be made available.

Function: The Wireless configuration tab has to include the following items:

1) Radio Interface
2) Security

3) Advanced Functionality

4) Access Control

RFC Reference: N/A.

Premise: User Interface -Where the CPE device implements an integrated Multimedia Terminal Adaptor a specific user configurable tab has to be made available.

Function: The MTA tab has to include the following items:

1) Status of the integrated MTA functionality

2) DHCP status, including any additional interfaces configured for the use of the MTA device

3) Quality of Service

4) Device provisioning

5) Event Logging

RFC Reference: N/A.

6.7.10 Feature Specification

6.7.10.1 Feature: WAN Connectivity

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Whilst there are no direct requirements for the CMTS to be able to support MAP-E, it is required that native IPv6 end-to-end connectivity is provided. The CMTS has to therefore be configured with IPv6 on both the cable and upstream interfaces.

Function: The cable modem / eRouter has to be provided with a GUA /128 for its public WAN interface and at least one /64 prefix delegation for client usage.

RFC Reference: N/A.

6.7.10.2 Feature: Provisioning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To support the provision of MAP-E, some additional attributes has to be provisioned on the CPE.

Function: The following attributes has to be provided, At present, there is a request for a DHCPv6 option for providing these attributes. (IETF draft MAP DHCPv6 options [145]). It is recommended that the DHCPv6 option is used. The CPE MAY obtain these additional parameter by alternate means such as SNMP or TR-069 based provisioning:

- Publically routable IPv4 address
- Port Set Range
- Port Set ID
- Border Relay IPv6 prefix

RFC Reference: [16**]**.

6.7.10.3 Feature: Provisioning

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv6 server support.

Function: For the CPE to receive the required provisioning parameters via DHCPv6, the DHCPv6 server has to support the provisioning of the attributes as specified in [16].

RFC Reference: [16**]**.

6.7.10.4 Feature: Cable eRouter B4 functionality

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Because the CPE is only provided with IPv6 connectivity on its WAN interface, it has to perform all DNS resolution over IPv6.

Function: To reduce the amount of encapsulated traffic, the cable B4 gateway has to be configured as a DNS forwarder for all LAN client requests. Any client DNS requests received via IPv4 has to therefore be forwarded as IPv6 to the external DNS servers.

RFC Reference: N/A.

6.7.10.5 Feature: LAN Addressing - IpV6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are no specific requirements for IPv6 LAN addressing to support MAP-E. Native IPv6 end-to-end connectivity is required for MAP-E to operate.

Function: IPv6 addressing SHOULD be used for all LAN clients that support IPv6.

RFC Reference: N/A.

6.7.10.6 Feature: LAN Addressing - IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Provision of local LAN addressing for IPv4 client devices.

Function: The CPE Gateway has to provide 'private' IP addressing from the ranges specified in RFC 1918 [30]. The accepted normal configuration is to use the 192.168.0.0/16 range, with a single /24 subnet.

The CPE router has to use one address from the selected IPv4 subnet as its LAN interface address, and this has to be used as the default gateway for LAN IPv4 traffic.

The CPE router has to also use its LAN IPv4 address as the DNS server address in DHCP offers to clients.

RFC Reference: [30**]**.

6.7.10.7 Feature: Packet Encapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv4 encapsulation into IPv6 bearer traffic.

Function: The CPE router has to encapsulate IPv4 packets with a public IPv4 destination, in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

The IPv4 packet has to be translated using NAPT-44 to a free port within the provisioned Port Set Range.

The IPv6 destination address is created using the MAP Border Relay IPv6 prefix, and the encoded destination IPv4 address.

The IPv6 source address is created using the CPE IPv6 prefix, and the address and port mapping rules.

The CPE router MAP-E interface has to drop any packets in the following groups:

- Broadcast IPv4 packets that have a destination address 255.255.255.255 as defined in RFC 919 [19].
- Multicast IPv4 packets that have a destination address 224.0.0.0/5 as defined in RFC 1112 [20].
- Any packets that have a 'private' destination IP Address as defined in RFC 1918 [30].
- Any packets that have a 'IPv4 link-local' destination IP Address as defined in RFC 3927 [72].

RFC Reference: [15**]**.

6.7.10.8 Feature: Packet Decapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The cable modem/e-router has to decapsulate embedded IPv4 traffic.

Function: The CPE router has to decapsulate packets in accordance with the specification defined in draft-ietf-softwire-map-02 [144].

Where an IPv4 fragment is received, the B4 router has to forward the fragment unaltered to its destination. It is the responsibility of the destination client to reassemble fragments.

Once the IPv4 packet is recovered, it has to forward the packet to the client specified in the source address.

RFC Reference: N/A.

6.7.10.9 Feature: MTU Size and Fragmentation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Encapsulation of an IPv4 packet with an IPv6 header to transport across the cable network increases the overall packet size by 40 bytes.

The current DOCSIS2+IPv6 and DOCSIS3 specifications state that the maximum IP packet size that can be carried is 1 500 bytes. The maximum IPv4 packet size that can be sent between the MAP-E CPE and Border Router is therefore 1 460 bytes.

Function: Because fragmentation in IPv6 has to be performed by the sending host, any packet received by the cable modem/ eRouter that is in excess of this size has to either:

- Be encapsulated by the CPE, and fragmented into two packets for transmission to the Border Router, or
- Be reduced in size to a maximum of 1 460 bytes before being encapsulated.

draft-ietf-softwire-map-02 [144] recommends the latter approach is taken.

The MTU of the CPE router encapsulation interface has to be reduced to 1 460 bytes. This will have three effects:

1) Any client that uses Path MTU Discovery (PMTUD) [22], [32] will learn the MTU restriction and reduce any outgoing packets to fit within the MTU.

2) Any IPv4 packet received by the CPE that is in excess of 1 460 bytes, with the DF bit clear, has to be fragmented by the CPE before encapsulation. The IPv6 packet will therefore fit within the IPv6 MTU of 1 500 bytes.

3) Any IPv4 packet received by the CPE that is in excess of 1 460 bytes, with DF set has to be dropped by the CPE, and an ICMP 'packet too big' message returned to the client. The client is expected to process this message and reduce the packet to the size specified in the ICMP reply.

RFC Reference: [15**]**, [22], [32].

6.7.10.10 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Some applications are dependent on the TCP MSS setting to determine the maximum TCP payload to transmit. By default, the TCM MSS size on an Ethernet network is 1 460 bytes (i.e. MTU - 20 byte IP header - 20 byte TCP header).

The MSS is set in the TCP SYN packet of each session. MSS clamping is the process of altering the MSS on outgoing and incoming TCP SYN packets so the client receives the smaller value.

Function: MSS clamping has to be implemented on the CPE encapsulation interface of the cable modem/e-router to reduce the value to the IPv4 MTU - 40 bytes (i.e. TCP/IPv4 header size). i.e. The device has to support MSS clamping to 1 420 bytes.

RFC Reference: N/A.

6.7.10.11 Feature: Client MTU reduction via DHCPv4 option

The premise, function and related RFC reference relating to this feature is as given below.

Premise: DHCPv4 option 26 allows the specification of the maximum IPv4 MTU sent by IPv4 client devices on the local area network as defined in RFC 2132 [33].

Function: The cable modem/ e-router MAY implement DHCPv4 option 26 to set the maximum client MTU size, however this approach has the following limitations.

Not all client IP stacks and DHCP clients will respect the DHCP option to set the interface MTU.

Devices that are manually configured (and hence do not use DHCP) would also require manual configuration of the interface MTU.

Reducing the interface MTU of clients would also impact all LAN IPv4 traffic as this would also all be reduced to an MTU of 1 460 bytes.

It cannot be guaranteed that all client applications will function if the MTU has been reduced.

RFC Reference: [33**]**.

6.7.10.12 Feature: Future review of Maximum MTU size

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Note that the forthcoming DOCSIS3.1 specification is reviewing the DOCSIS MTU size and it is likely to be increased to support packets sufficient in size to accommodate MAP-E encapsulation.

Function: Where the DOCSIS packet size can be increased, the MTU has to be set to allow for full IPv6 encapsulated MAP-E packets to be sent with no fragmentation.

RFC Reference: N/A.

6.7.10.13 Feature: MAP- E Inbound Session control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: With MAP-E, the ability to control inbound sessions is limited to the port-set range that has been assigned to the CPE. These are generally:

- DMZ host functionality. Where all inbound packets can be forwarded to a single client
- Port forwarding. Where individual TCP or UDP ports can be forwarded to a specific client
- UPnP NAT Traversal. Where a client requests for the NAT to open one or more TCP or UDP ports to itself for inbound packets.

Function: The cable modem/e-router SHOULD allow inbound TCP or UDP connections within the port-set range that has been provisioned to the CPE. The cable modem/e-router is not allowed to configurate ports outside of its assigned port-set range as provisioned by any allowable means such as DHCPV6 options, TR-069 or vendor specific MIB values.

RFC Reference: [16**]**.

6.8 6rd

6.8.1 6rd Technology Summary

6rd is defined in RFC 5969 [117] and it is one of the incremental methods for deploying IPv6. IPv6 native traffic is tunneled over IPv4 access network. The IPv4 tunnel end points are the RG/CPE and the Border Router. These two elements perform encapsulation/decapsulation of IPv6-in-IPv4 traffic. BR is placed at IPv6 edge and is addressed via IPv4 (anycast) address for load balancing and resiliency.

6rd is considered stateless which means that no flow states are created in order to forward traffic. In addition, name resolution (DNS) works natively over the IPv6-in-IPv4 tunnel.

From the subscriber management prospective (ESM), there is no IPv6 awareness needed at the BNG (for example DHCPv6, IPv6 ND, etc).

IPv4 traffic is flowing natively between the RG/CPE and the IPv4 network behind the BNG. NAT44 is implemented on the Border Router for IPv4 continuity.

In contrast, IPv6 traffic originating at the subscriber host is encapsulated into IPv4 traffic at the RG/CPE and is carried over IPv4 network to the BR which decapsulates IPv6 traffic and forwards it natively onto the IPv6 network that is directly attached to the BR.

In the opposite direction, the BR encapsulates IPv6 traffic coming from the native IPv6 network into IPv4 tunnel. The tunnel destination is the RG/ CPE which decapsulates IPv4 traffic and forwards native IPv6 traffic to the host.

6.8.2 6rd BR Technology Feature Summary

Table 12 gives a summary of the 6rd BR technology feature.

Table 12: 6rd BR Technology Feature Summary

Functional Name	Requirement	Description
NAT44 - RFC 4787 [94] (UDP)	Required	Compliance with NAT behaviour according to RFC 4787 [94] for UDP.
NAT44 - RFC 5382 [106] (TCP)	Required	Compliance with NAT behaviour according to RFC 5382 [106] for TCP.
NAT44 - RFC 5508 [108] (ICMP)	Required	Compliance with NAT behaviour according to RFC 5508 [108] for ICMP.
Redundancy	Required	All critical components has to be redundant in such a manner that they can fail-over without impact to customer or management traffic greater than 1 ms.
Anycast	Required	Anycast BR gateway addresses are a requirement to allow simplicity of deployment for a single prefix across multiple BR's.
BR Address withdrawal	Required	The BR should have at least four points of BR GW address withdrawal occurrence. The list includes: - loss of IPv6 route out, - loss of all BGP/IGP sessions, - loss of forwarding, - loss of NPU capacity Any of the failures should be detectable based on configurable timers with 15 seconds being the default setting.
Fragmentation	Required	Fragmentation should be done on the IPv4 packet. Fragmentation should be placed on the ASIC running in the line card.
NAT - Network Address and Port Mapping - Endpoint Independent Mapping	Required	For two flows for a common inside source IPv4 address and port, the external address/port translation is independent of the destination IPv4 address and port and when the flows exist simultaneously in the NAT state table they will use the same translation.
NAT - Translation Filtering - Endpoint Independent Filtering	Required	A flow initiated externally can use the existing External/Inside IPv4 address/port mapping and it is independent of the source IPv4 address/port of the senders.
NAT - Paired IP Address Assignment	Required	Translation to External IPv4 address is done in a paired fashion. A given Inside address is always translated to the same External IPv4 address.
NAT - Hair-pinning	Required	Different internal addresses on the same internal interface has to be able reach each other using external address/port translations.
NAT - 1:1 IP Mapping	Required	Ability to configure a one to one type of mapping for particular inside-VRFs: every public IP will be mapped to one and only one private IP (multiple ports are allowed).
NAT44 - Outside-Service-App mapping for inside-VRF	Required	Ability in the inside-vrf to provide the explicit outside serviceapp to be paired.
NAT - Port Limit configuration	Required	A maximum amount of ports can be configured for every IPv6 source B4 address.
NAT - Per-Protocol Timeout configuration	Required	Timeout of mappings is critical and removing a NAT mapping at the appropriate time maximizes the shared port resource.
NAT - Dynamic Port Range start configuration	Required	The start port for dynamic port ranges should be configurable to allow for a range of ports for static port mappings.
NAT - Software Load Balancing	Required	NAT Inside to Outside hashing performed on the Source private IPv4 address. NAT44 Outside to Inside hashing performed on the Destination IPv4 (lower-order 2 bits) address assigned from the pool prefix.
Port Allocation	Required	In order to reduce the volume of data generated by the NAT device (logging creation and deletion data), bulk port allocation can be enabled. When bulk port allocation is enabled and when a subscriber creates the first session, a number of contiguous outside ports are pre-allocated. A bulk allocation message is logged indicating this allocation.

Functional Name	Requirement	Description
Deterministic NAT	Optional	Algorithmically maps a customer's private IPv4 address to a set of public IPv4 address ports, allowing a significant reduction in logging.
FTP ALG (Active and Passive)	Required	FTP clients are supported with inside (private) address and servers with outside (public) addresses. Passive FTP is provided by the basic NAT function. Active FTP is used with the ALG.
RTSP ALG	Required	Remote control protocol for streamers (which use RTP/RTCP or RDT). Our implementation considers the server is located "outside" and clients are "inside". RTSP is used in many streamers like QuickTime or RealNetworks. Both "SETUP" and "SETUP response" are analysed. Based on this information, the translation entry will be created in the table.
SIP ALG	Required	ALP for SIP to function and transverse the NAT.
PPTP ALG	Required	ALP for PPTP to function and transverse the NAT.
Use of Application SVI's, connecting multiple routing entities (inside/private VRF, outside/public VRF)	Required	NAT should be possible from any routing context (VRF, GRT) to any routing context (VRF, GRT).
Stateful ICMP	Required	Stateful ICMP mappings between inside and outside ICMP identifiers should be supported.
Thresholds	Required	Configurable thresholds using watermarks should be supported to monitor the resources on the BR.
Chassis NAT Clustering	Optional	Clustering of BR's to allow for inter-chassis resiliency.
PCP	Required	Support for PCP to allocate static port bindings.
Logging via Netflow V9/IPFIX	Required	Netflow v9 support for logging of the translation records. Logging of the translation records can be mandated by for Lawful Intercept. The Netflow uses binary format and hence requires software to parse and present the translation records.
Logging via Syslog	Required	Syslog support for logging of the translation records as an alternative to Netflow. Syslog uses ASCII format and hence can be read by users. However, the log data volume is higher in Syslog than Netflow.
Destination based Logging	Required	Destination Based Logging will generate one record per session, including the destination port and address. NFv9 templates including destination IP address and destination port will be used.
Radius Logging	Required	Logging using Radius accounting messages.
XML	Optional	Logging using XML files.
Static port forwarding (up to 6K static forward entries per npu)	Required	Static port forwarding configures a fixed, private (internal) IPv6 address and port that are associated with a particular subscriber while the BR allocates a free public IP address and port. Therefore, the inside IPv6 address and port are associated to a free outside IP address and port.
Static port forwarding 1:1 active/standby	Required	Static Port Forwarding mapping will be kept constant with two npu cards in Active/Standby mode.
Performance Characteristic	**Requirement**	**Description**
64 BR instances per npu Card	Required	The ability to stack transition instances on top of one another.
20M+ Translations (per npu)	Required	The requirement to translate sessions within the technology at a certain speed denoted by the sessions themselves. These are all session types not only Primary.
Minimum Gbps throughput per npu (Inside-to-outside VRF + Outside-to-inside VRF) with IMIX traffic	Required	Per half slot throughput requirement.
1M+ connections per second setup rate	Required	Primary sessions, port block assignments, per NPU.
1M users per chassis	Required	Per Chassis a minimum of 1 00 000 users should be serviceable at any single moment in time.
Latency	Required	Latency is between 40 and 200 micro seconds (µs).

Functional Name	Requirement	Description
6 npu Cards per chassis	Required	Min chassis requirement for NPUs.
IRB support	Required	Integrated Routing and Bridging/Virtual Interfaces (L3 interface for Bridge Domain).
Broadband Network Gateway (BNG) support	Optional	32 k BNG sessions and up to 256 k NAT users at the same time.

6.8.3 Main 6rd RFC References

Historical documents: RFC 4787 [94], RFC 5382 [106], RFC 5508 [108], RFC 5569 [109].

Deterministic NAT: [9].

PCP: [6], [7], [8], RFC 6877 [134].

6.8.4 CORE Device (LSN/CGN)

BR is an implementation of an IPv4/IPv6 transition protocol based on 6rd, and is the CORE node that all CPE/CEs running 6rd and creating softwires for the technology connect to. It is the termination for the IPv4 tunnel for native IPv6 traffic and it is the IPv4 NAT device egressing traffic towards the destination on IPv4.

This section concerns itself solely with the 6rd specific required functionality to allow the technology to function completely and fully. The specification laid out below allows 6rd technology to be deployed on the BR as a non-service deprecating form comparative to a native private IPv4 delivery to any given customer within a Carrier Topology or potentially otherwise. The present document is however focuses on Cable requirements in functionality but can stretch further if required.

Note that any given BR implementation should be complaint with the associative feature for deployment within a Cable ISP network.

6.8.5 Hardware / Software Requirements

6.8.5.1 Feature: HA Physical Redundancy NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to perform the ability to switch over on failure of an SI competent NPU in a full stateful fashion with no loss of packets or latency beyond 1ms. If a single NPU fails it has to be grouped so as to allow for the exact functionality to be moved to a new NPU within that NAT group.

Function: Listed below:

1) IP range failover
2) Full stateful NAT session failover
3) No session interruption
4) 1 ms failover maximum for any given CPE connectivity of session

RFC Reference: N/A.

6.8.5.2 Feature: ISSU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The BR has to be fully seamless ISSU compatible allowing process (RE/RP) failover during a hot software upgrade. Due to the nature of the single point of failover on the network the BR can have no requirement during SW upgrading to be rebooted or have any allowance in service failure during the upgrade. ISSU is used as a generic term, not a specific vendor term.

Function: With two or more RE/RPs/Supervisors/Chassis Controllers when adding software to the node in the form of an upgrade the service has to be uninterrupted. This usually occurs through the function of two REs swapping master control of the services on the chassis and then swapping back once the upgrade has occurred. This may require the 6rd NPUs to hold state and function to allow full seamless ISSU.

RFC Reference: N/A.

6.8.5.3 Feature: RP / RE redundancy

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function has to have no traffic effect either management or 6rd transit traffic while swapping processing unit (RE/RP) or on failure.

Function: Full HA redundancy between any two given Res within an BR has to be an ability within the node. This includes full syncing of all NAT caching, forwarding, logging and state held by the RE.

RFC Reference: N/A.

6.8.5.4 Feature: Shared/Split Resources

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The resources within any given BR has to be shared giving the appearance of a single node on the network.

Function: The BR has to have the ability to share or split resources where possible to allow for full function in given topology. This would include a single address per chassis. NPUs has to be able to share a single pool of address space on IPv4 and DHCP pool handoff on IPv6 if the function is required. Memory, processing and disk space should also be a single shared resource where required. In any given configuration the NAT group should specify and determine the resources it requires either by default or configurable.

RFC Reference: N/A.

6.8.5.5 Feature: Traffic Based Load Balanced NPUs

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The NPUs within any given BR has to have the ability to share the traffic and has to be considered a single NPU as per requirement.

Function: The basic function of the above premise is to allow for groups or all of the NPUs within any given BR to see all or part of the traffic coming into the node. This can be determined through a single BR address, so BR GW assigned or NAT group assigned under multiple GW addresses. Hashing SHOULD be used to determine the load balancing between all processing / hardware forwarding points with that NAT group or deterministic specification.

RFC Reference: N/A.

6.8.6 Performance Requirements

6.8.6.1 Feature: Node latency

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Node latency has to not exceed 1 ms for any given function of the BR.

Function: Within the remit of traffic requiring 6rd/NAT function, decapsulation, encapsulation and NAT function, the BR has to perform its function from ingress interface to egress interface in a time measured no higher than 1ms. This includes requirements for fragmentation, reassembly, re-ordering, queuing and ALG proxy functions.

RFC Reference: N/A.

6.8.6.2 Feature: Max throughput per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The throughput of the NPU SHOULD not be less than 10 gig shared.

Function: Due to carrier grade requirements and the ratio of interface requirements on an ISP network any specific NPU SHOULD not have a traffic capacity below 10 gig due to the nature of the capacity on most networks.

RFC Reference: N/A.

6.8.6.3 Feature: Chassis Throughput

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Chassis throughput SHOULD be a minimum of 40 gig to allow for deployment within carrier grade locations.

Function: The full throughput of the BR Chassis needs to be able to meet the requirements for carriers to deploy for a minimum of 100 000 customers in all aspects including forwarding ability, which includes the interface requirements and backplane, to allow for mass connectivity. 40 gig is solely an estimate and this does depend on requirements per ISP.

RFC Reference: N/A.

6.8.6.4 Feature: Min Sessions per BR Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The requirement for Sessions per Chassis should equal a little over the average requirement per customer for 100 000 customers. So the maximum session calculation would be 100 000 * x (average sessions per customer) 100k*60 defines 6 million session availability for a single chassis. As IPv6 is stateless and natively forwarded after decapsulation, it does not consume any session resources.

RFC Reference: N/A.

6.8.6.5 Feature: Minimum Customer IPv4 Addresses per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Defines the minimum amount of IPv4 addresses able to be configured per chassis.

Function: If we define the IPv4 address space requirement based on session and block allocation this would be 60 ports per subscriber as a block assignment comparative to 100 k customers online allowing the integer of 1 024. Or if we calculate backwards this would be 100 000 customers divided by 60 per IP (although this changes per deterministic NAT requirement) allowing for a max utilization of 1 000 ports per subscriber and thus giving us 1 666-address requirement if we have maximum assignment utilization. So we can round this up to 2 040-address minimum address pool requirement.

RFC Reference: N/A.

6.8.6.6 Feature: Min Customer CPEs per NPU

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The function is based on the amount of CPEs and thus IPv4 CPE addresses that can be assigned from a single Chassis or connected.

Function: Due to the scaling of any given BR solution this number should for main be to the min customer number and scaling requirements, so a minimum of 100 000 CPE addresses. This is solely a minimum.

RFC Reference: N/A.

6.8.7 Feature Specifications

6.8.7.1 Feature: Tunnel Identifiers/Client-Customer ID

The premise, function and related RFC reference relating to this feature is as given below.

Premise: For general identification and data retention purposes, tunnel identifiers has to be uniquely associated with a single CPE. The identifier has to be the IPv4 address of the CPE.

Function: Unique IPv4 CPE addresses has to be assigned to the CPE.

RFC Reference: N/A.

6.8.7.2 Feature: IPv6 Global Unicast Address Format

The premise, function and related RFC reference relating to this feature is as given below.

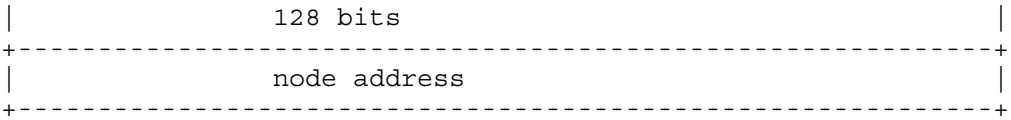
Premise: The BR has to be able to forward and be configured with an IPv6 Unicast structured address.

Function: In any given IPv6 topology unicast address structures has to be adhered to with no proprietary configurations. All present IPv6 unicast structures mentioned within RFC 3513 [62] has to be compatible within the vendor deployment to allow for standardization and monitoring.

RFC Reference: [62**]**.

From section 2.5 in [62]:

> "IPv6 unicast addresses are aggregable with prefixes of arbitrary bit-length similar to IPv4 addresses under Classless Interdomain Routing.
>
> There are several types of unicast addresses in IPv6, in particular global unicast, site-local unicast, and link-local unicast. There are also some special-purpose subtypes of global unicast, such as IPv6 addresses with embedded IPv4 addresses or encoded NSAP addresses. Additional address types or subtypes can be defined in the future.
>
> IPv6 nodes may have considerable or little knowledge of the internal structure of the IPv6 address, depending on the role the node plays (for instance, host versus router). At a minimum, a node may consider that unicast addresses (including its own) have no internal structure:



```
A slightly sophisticated host (but still rather simple) may
additionally be aware of subnet prefix(es) for the link(s) it is
attached to, where different addresses may have different values for
n:

 |               n bits                        |  128-n bits    |
 +---------------------------------------------+----------------+
 |              subnet prefix                  |  interface ID  |
 +---------------------------------------------+----------------+

Though a very simple router may have no knowledge of the internal
structure of IPv6 unicast addresses, routers will more generally
have knowledge of one or more of the hierarchical boundaries for the
operation of routing protocols. The known boundaries will differ
from router to router, depending on what positions the router holds
in the routing hierarchy."
```

6.8.7.3 Feature: 6rd Timers

The premise, function and related RFC reference relating to this feature is as given below.

Premise: These has to allow for all below 6rd effecting timers to be configured per the RFC.

Function: Listed below are the timers that SHOULD be configured for the BR, note that these SHOULD match the CPE where possible. The timers are not fixed as it is deployment dependent the list you below are the suggested values only and the required fields for any particular vendor.

i) icmp-query --- min 1

ii) sip --- min 2

iii) tcp-established --- hrs 30 min 0

iv) tcp-syn --- sec 30

v) no tcp-time-wait --- min 3

vi) tcp-transitory --- min 4

vii) udp --- min 5

viii) udp-initial --- sec 15

ix) udp-dns --- sec 15

RFC Reference: No single defined RFC.

6.8.7.4 Feature: Thresholds and Watermarks

The premise, function and related RFC reference relating to this feature is as given below.

Premise: There are two types of thresholds/watermarks, reactive based and logging based, both are required.

Function: Listed below are minimum required thresholds for 6rd and SHOULD be adhere to validate the utilization on an event basis due to resource allocation that is required for CGN:

1) Watermarks based on Subscriber Thresholds for the node

2) Watermarks based NPU total resources

3) Watermarks based on Port Allocation Thresholds per Subscriber

4) Watermarks based on Port Allocation Thresholds per NPU

5) Watermarks based on IP Allocation Utilization Thresholds

6) Watermarking Event Reporting through SNMP

7) Watermarking Event Reporting through Syslog

8) Watermarking for dynamic port allocation over the top of Deterministic NAT

RFC Reference: N/A.

6.8.7.5 Feature: Softwire Initialization Dynamic Tunnels

The premise, function and related RFC reference relating to this feature is as given below.

Premise: SI Quick drop and pickup approach.

Function: All tunnelling initialization and dropping should be orientated towards a fast rotation on dynamic allocations given to each CPE node.

RFC Reference: N/A.

6.8.7.6 Feature: Port Block Allocation per IP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Port block allocation is configurable to allow for any ratio assignment per IP.

Function: An assignment of ports based on block allocations for any numeric value with automatic assignment of a single IP determined by the black allocation.

RFC Reference: N/A.

6.8.7.7 Feature: Deterministic NAT / Deterministic Dynamic thresholds

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Deterministic NAT policy and configuration has to be adhered to for BRs not only for port logging requirements for data retention but also for processing and functional improvement.

Function: DetNAT should be used in the following manner. It should be configurable to allow for any n-Ports to be allocated to the port block and any number of port block assignments to be dynamic or static per CPE address. It has to also hold a dynamic threshold to further the ability for IP to port ratios by allowing the dynamic section of the port allocation for any single or cluster of CPE nodes to use 1-5 block allocations and to allow for singular oversubscription per CPE based on a threshold. So if the block allocation reaches over 60 % (configurable) of the full amount of ports, 65 000 per IP, a restriction is placed on further port allocations. For example until the 60 % threshold is reached the Subscriber CPE is allowed up to 3 000 ports but once the port threshold of 605 is reached this then drops to 500 ports for all users. No old connections are dropped but no new ones can be added. This allows for controlled oversubscription with heavy port users to over allocate while there are resources.

RFC Reference: [9**]**.

From section 2 in [9]:

> "The algorithm is not designed to retrieve an internal host among those sharing the same internal IP address (e.g., in a 6rd context, only an IPv6 address/prefix can be retrieved using the algorithm while the internal IPv4 address used for the encapsulated IPv4 datagram is lost).

Several address assignment algorithms are possible. Using predefined algorithms, such as those that follow, simplifies the process of reversing the algorithm when needed. However, the CGN MAY support additional algorithms. Also, the CGN is not required to support all algorithms described below. Subscribers could be restricted to ports from a single IPv4 address, or could be allocated ports across all addresses in a pool, for example. The following algorithms and corresponding values of A are as follow:

0. Sequential (e.g. the first block goes to address 1, the second block to address 2, etc.)
1. Staggered (e.g. for every n between 0 and ((65536-R)/(C+D))-1 , address 1 receives ports n*C+R, address 2 receives ports (1+n)*C+R, etc.)
2. Round robin (e.g. the subscriber receives the same port number across a pool of external IP addresses. If the subscriber is to be assigned more ports than there are in the external IP pool, the subscriber receives the next highest port across the IP pool, and so on. Thus, if there are 10 IP addresses in a pool and a subscriber is assigned 1000 ports, the subscriber would receive a range such as ports 2000-2099 across all 10 external IPaddresses).
3. Interlaced horizontally (e.g. each address receives every Cth port spread across a pool of external IP addresses).
4. Cryptographically random port assignment (Section 2.2 of RFC6431 [RFC6431]). If this algorithm is used, the Service Provider needs to retain the keying material and specific cryptographic function to support reversibility.
5. Vendor-specific. Other vendor-specific algorithms may also be supported.

The assigned range of ports MAY also be used when translating ICMP requests (when re-writing the Identifier field).

The CGN then reserves ports as follows:

1. The CGN removes reserved ports (R) from the port candidate list (e.g., 0-1023 for TCP and UDP). At a minimum, the CGN SHOULD remove system ports (RFC6335) [RFC6335] from the port candidate list reserved for deterministic assignment.

2. The CGN calculates the total compression ratio (C+D), and allocates 1/(C+D) of the available ports to each internal IP address. Specific port allocation is determined by the algorithm (A) configured on the CGN. Any remaining ports are allocated to the dynamic pool.

2. Note: Setting D to 0 disables the dynamic pool. This option eliminates the need for per-subscriber logging at the expense of limiting the number of concurrent connections that 'power users' can initiate.

3. When a subscriber initiates a connection, the CGN creates a translation mapping between the subscriber's inside local IP address/port and the CGN outside global IP address/port. The CGN has to use one of the ports allocated in step 2 for the translation as long as such ports are available. The CGN SHOULD allocate ports randomly within the port range assigned by the deterministic algorithm. This is to increase subscriber privacy. The CGN has to use the preallocated port range from step 2 for Port Control Protocol (PCP, [I-D.ietf-pcp-base]) reservations as long as such ports are available. While the CGN maintains its mapping table, it need not generate a log entry for translation mappings created in this step.

```
4. If D>0, the CGN will have a pool of ports left for dynamic
assignment. If a subscriber uses more than the range of ports
allocated in step 2 (but fewer than the configured maximum ports M),
the CGN assigns a block of ports from the dynamic assignment range
for such a connection or for PCP reservations. The CGN has to log
dynamically assigned port blocks to facilitate subscriber-to-address
mapping. The CGN SHOULD manage dynamic ports as described in [I-
D.tsou-behave-natx4-log-reduction].

5. Configuration of reserved ports (e.g., system ports) is left to
operator configuration.

Thus, the CGN will maintain translation-mapping information for all
Connections within its internal translation tables; however, it only
needs to externally log translations for dynamically-assigned
ports."
```

6.8.7.8 Feature: IP Ranges per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IP ranges should be able to be assigned per NAT group or across multiple NAT groups to allow for full scaling and shared IP resource.

Function: For further study.

RFC Reference: N/A.

6.8.7.9 Feature: NAT Grouping resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT Grouping SHOULD be cross NPU based.

Function: NAT groups SHOULD be able to use the same NPUs and IP allocations regardless of their GW addresses. This allows for complete shared resources across multiple NAT groups. The function states that the NPUs can be grouped in any manner for active standby numbers as the user/administrator wishes. Thus NAT group 1 can use NPU 1,2,3&4, with 1,2&3 as active and NAT group 2 can use NPU 1,2,3,4,5&6 with 1,2&3 as standby. This promotes scaling and no restriction if there are particular requirements to share a resource.

RFC Reference: N/A.

6.8.7.10 Feature: Virtual interface per Chassis

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A BR has to be able to assign a single virtual interface with up to 8 6rd GW addresses for any given 6rd instance on the node.

Function: The virtual interfaces has to be able to be assigned to any number, and all, NPUs within the chassis to allow complete shared resources. So a single or multiple BRs acting as the single destination interface for all resources and CPE nodes using the BR.

RFC Reference: N/A.

6.8.7.11 Feature: NPU to Interface throughput ratio

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To allow balanced Bandwidth assignment across all virtual or a single virtual interface matching or exceeding the physical bandwidth of the node.

Function: To prevent over buffering, dropped packets and general resource issues the BR has to be able to consider a physical topology that matches a balanced logical bandwidth ratio between the interfaces and the NPU throughput.

RFC Reference: N/A.

6.8.7.12 Feature: BR Address (es)

The premise, function and related RFC reference relating to this feature is as given below.

Premise: BR Addressing.

Function: Listed below:

1) Has to Allow one or multiple addresses as GWs
2) IPv4 and IPv6 addressing SHOULD be placed on separate interfaces
3) The BR SHOULD be able to act as the DHPv4 server for the CPE if required
4) SI-IDs should be able to comprise ONLY the IPv4 CPE address
5) The IPv4 Public address assigned to the CPE has to be unique
6) The IPv4 addresses configured for 6rd Gateways has to be Anycast compatible
7) The BR has to allow multiple NAT group

RFC Reference: N/A.

6.8.7.13 Feature: Anycast Gateway Address

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Anycast usage for a BR gateway or DHCPv4 address has to be accepted within the configuration.

Function: One or many Anycast addresses has to be configurable for the BR GWs regardless of any other resource and for the DHCPv4 server address.

RFC Reference: N/A.

6.8.7.14 Feature: Source IP per BR Interface

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The premise is that a definition of sourced addresses SHOULD be configured to allow forced routing to a specific GW address on the BR.

Function: A user/administrator SHOULD be able to place a source IP statement within a NAT group allowing the CPE packets originating from within that subnet to be routed to a specific NAT group or BR interface. His allows the delineation of different service IPs instead of relying on separation of traffic based on DNS record responses and thus specific rules and resource allocations for that particular sourcing subnet.

RFC Reference: N/A.

6.8.7.15 Feature: BR Address Withdrawal

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The BR has to be withdrawn under the circumstances below.

Function: Listed below:

1) The BR is not reachable within the routing protocol being used or cannot reach its neighbour's or peers
2) The 6rd process dies
3) NPUs available cannot hold the present capacity
4) Shutdown or part or all of the service
5) External Routing to the internet is down for a fixed period of time (demands BR clustering to determine reachability from a redundant BR)

RFC Reference: N/A.

6.8.7.16 Feature: Chassis Clustering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: To Cluster two or more BRs for session state failover. It is only valid for NAT44 as IPv6 traffic is stateless and natively routed after decapsulation.

Function: If a single BR goes down it has to do the following:

- Remove its BR address as a reachable entity within the routing protocol being used
- Forward any remaining packets to the secondary BR
- Validate NAT cache state between itself and its redundant BR

The clusters has to sync on the following items:

- The TCP and UDP sessions presently being used to for NAT44
- The TCP not used for n amount of time configurable
- The present capacity of interface and NPU to allow for structural syncing
- Sync between all clustered BRs due to best path attributes for the Anycast / Unicast Addresses used (also dependent on secondary addressing on the CPE)

RFC Reference: N/A.

6.8.7.17 Feature: Multiple Transition Technology Resource Sharing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NAT44, NAT64, DS-Lite and 6rd has to function on the same platform to allow for uniformity within European markets.

Function: All four technologies should be able to function and share resources dynamically on any given platform.

RFC Reference: N/A.

6.8.7.18 Feature: NPU / chassis Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Buffering requirements for the BR for transit traffic covering queues and buffer timers.

Function: Listed below:

1) Queue sizing per flow

2) Queue sizing per port block

3) Queue sizing per NPU

4) Queue sizing per ingress interface

5) Buffer sizing per flow

6) Buffer sizing per port block

7) Buffer sizing per NPU

8) Buffer sizing per ingress interface

9) Reordering buffers and out-of-order caching

RFC Reference: N/A.

6.8.7.19 Feature: Tunnel MTU Sizing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The Tunnel MTU SHOULD be considered as the IPv6 encapsulated within the IPv4 packet and therefore should be set at 1 500, if IPv6 only the tunnel MTU has to be configurable to 1 480 to allow for 20 byte header.

Function: For further study.

RFC Reference: N/A.

6.8.7.20 Feature: MSS Clamping

The premise, function and related RFC reference relating to this feature is as given below.

Premise: MSS clamping, the MTU value for the TCP max segment size, which should be configurable to as low as 1 460 in case any MSS client considers all header within the MSS value.

Function: For further study.

RFC Reference: N/A.

6.8.7.21 Feature: 6rd Fragmentation and Buffering

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Fragmentation has to be placed on the ingress interface (virtual or physical) pre decapsulation on upstream and post-encapsulation on downstream on IPv4.

Function: All fragmentation on 6rd on the BR and CPE has to be on IPv4. This prevents IPv6 fragmentation requirements and overlay code. IPv6 inherently does not have fragmentation built into it and thus can cause major performance issues on an BR forwarding plane or even be placed in software.

Reassembly has to only be used when the BR receives noted IPv4 fragmented packets incoming upstream from the CPE.

Pre-fragmented packets has to be re-ordered before being sent on and again requires full buffer queues and wait timers exceeding 3 ms.

RFC Reference: N/A.

6.8.8 Monitoring and Management

6.8.8.1 Feature: LLDP

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Neighbor discovery - LLDP (802.1AB) should be implemented to discover information about adjacent devices, and an operator should have the flexibility to enable/disable this globally or per port.

Function: Allow devices to discover information about adjacent devices.

RFC Reference: N/A.

6.8.8.2 Feature: NTP, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv4 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.8.8.3 Feature: NTP, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: NTP should be supported over IPv6 for configuring the time on the device. Minimum 2 backup peers/sources should be configurable. Authentication should be supported.

Function: Make sure all devices in the network have an equal time configured.

RFC Reference: [115**]**.

6.8.8.4 Feature: SSH, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv4 should be supported.

Function: Accessing the device in a secure way over IPv4.

RFC Reference: [76**]**.

6.8.8.5 Feature: SSH, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SSH2 over IPv6 should be supported.

Function: Accessing the device in a secure way over IPv6.

RFC Reference: [76**]**.

6.8.8.6 Feature: Telnet, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv4 should be supported.

Function: Accessing the device in a simple way over IPv4.

RFC Reference: [18].

6.8.8.7 Feature: Telnet, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using Telnet over IPv6 should be supported.

Function: Accessing the device in a simple way over IPv6.

RFC Reference: [18**]**.

6.8.8.8 Feature: SNMPv2, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv4 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv4.

RFC Reference: [26**]**.

6.8.8.9 Feature: SNMPv2, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv2 over IPv6 should be supported.

Function: Accessing the device using an SNMPv2 tool over IPv6.

RFC Reference: [26**]**.

6.8.8.10 Feature: SNMPv3, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv4 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv4.

RFC Reference: [60**]**.

6.8.8.11 Feature: SNMPv3, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device access using SNMPv3 over IPv6 should be supported.

Function: Accessing the device using an SNMPv3 tool over IPv6.

RFC Reference: [60**]**.

6.8.8.12 Feature: Remote Access - TACACS+, IPv4

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv4 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.8.8.13 Feature: remote access: TACACS+, IPv6

The premise, function and related RFC reference relating to this feature is as given below.

Premise: TACACS+ over IPv6 has to be supported for authentication, authorization and accounting of user access to the device. A minimum of 2 TACACS+ servers should be configurable.

Function: Protecting the devices from non-authorized access and logging the actions performed on the device.

RFC Reference: [13**]**.

6.8.9 6rd CPE requirements

Requirements on the host are determined by the available connectivity in the Home Network. The Home Network may be dual stack or IPv4 only or IPv6 only.

The following requirements apply to the CPE Router:

- WAN interface facing the Access Network
  - Request IPv4 address via DHCPv4
- LAN interface facing the Home Network
  - IPv4 private addressing
  - DHCPv4 server for LAN addressing (stateful DHCP service may be included)
  - IPv4 DNS Proxy
  - IPv6 addressing
  - DHCPv6 server for LAN addressing
  - IPv6 MTU set to 1 480 Byte
  - IPv6 TCP MSS clamping to 1 440 Byte
- CPE router
  - Receive BR IP address via DHCPv4
  - Encapsulate LAN IPv6 packets in IPv4 header
  - Decapsulate IPv4 packets for LAN IPv6
  - Fragment encapsulated IPv4 packets
  - Reassemble received IPv4 fragments
  - PCP over IPv4

In addition to the requirements listed above, the CPE Router should support the following features on its interfaces facing the Home Network (wired, wireless):

- DNS resolver information as defined in [121]
- IPv4 NAT/NAPT
- IPv4 static NAT
- IPv4 inbound port forwarding
- IPv4 stateful packet firewall (5-tuple filters), enabled by default
- IPv6 stateful packet firewall (5-tuple filters), enabled by default

- Recursive DNS server option

For purposes of a unified user experience some requirements apply to the user interface:

- Web-UI should be accessible on the LAN IP interface
- Login is initially presented in the format of requesting username and password
- Localization of UI for different languages should be supported
- A Status tab should contain subpages on Software, Connection, Security, Diagnostics
- A Router Basic tab should contain subpages on WAN Setup, LAN&DHCP Server, Backup
- A Router Advanced tab should contain subpages on Option, IP Filtering, MAC Filtering, Port Filtering, Forwarding, Port Triggers, DMZ Host
- A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log
- A Parental Control tab should contain subpages on User Setup, Basic Setup, Content Filter, ToD Filter
- A Wireless tab should contain subpages on Radio, Security, Advanced, Access Control

An MTA tab should contain subpages on Status, DHCP, QoS, Provisioning, Event Log.

6.8.9.1 Feature: 6rd CPE Base Requirements

The premise, function and related RFC reference relating to this feature is as given below.

Premise: 6rd has been developed as an interim transition mechanism where IPv6 connectivity can be provided using an existing IPv4 access network. As the access network remains as IPv4 there is a requirement to retain services using a public IPv4 address. This places the following requirements on CPE devices configured to act as 6rd routers.

RFC Reference: [117**]**.

6.8.9.2 Feature: Cable Modem management

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to manage the cable modem interface and associated features a management interface has to be provided.

Function: As 6rd implements the transport of IPv6 packets over an IPv4 tunnel, the management of the cable modem has to be performed over IPv4.

RFC Reference: N/A.

6.8.9.3 Feature: 6rd Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The device configuration parameters that are common to all devices within a 6rd domain need to be provisioned to the 6rd embedded cable modem router.

Function: The eRouter / CPE router has to obtain its 6rd configuration parameters via DHCP. The e-router has to include a request for DHCP option OPTION_6RD within its parameter Request list option as defined by RFC 2132 [33]. The e router MAY obtain its 6rd configuration via alternate provisioning methods such as TR-069 [137].

RFC Reference: [33**]**, [117].

6.8.9.4 Feature: 6rd IPv4MaskLen

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Device Configuration.

Function: The 6rd enabled cable modem has to obtain the following configuration parameter.

IPv4MaskLen, the number of high-order bits that are identical across all CE IPv4 addresses within a given 6rd domain.

RFC Reference: N/A.

6.8.9.5 Feature: 6rd Prefix

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The 6rd enabled cable modem has to obtain the following configuration parameter.

Function: 6rdPrefix, the 6rd IPv6 prefix for the given 6rd domain.

RFC Reference: [117**]**.

6.8.9.6 Feature: 6rd PrefixLen

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The 6rd enabled cable modem has to obtain the following configuration parameter.

Function: 6rd prefixLen, the length of the 6rd IPv6 prefix for the 6rd domain that the cable modem router is a member of.

RFC Reference: [117**]**.

6.8.9.7 Feature: 6rd BRIPv4Address

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The 6rd enabled cable modem has to obtain the following configuration parameter.

Function: 6rdBRIPv4Address, the Ipv4 address of the 6rd border relay for the given 6rd domain.

RFC Reference: [117**]**.

6.8.9.8 Feature: 6rd Device Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The 6rd enabled cable modem has to obtain the following configuration parameter.

Function: 6rdPrefix, the IPv6 prefix of the service provider which has to be provided as a 16 Octet IPv6 address. The bits in the prefix after the 6rdPrefixlen number of bits are reserved and has to be initialized to zero by the sender and ignored by the 6rd CPE function within the cable modem/router.

RFC Reference: [117**]**.

6.8.9.9 Feature: IPv6 traffic class marking in IPv4 Encapsulation

The premise, function and related RFC reference relating to this feature is as given below.

Premise: IPv6 in IPv6 encapsulation and manipulation of packet marking is to be performed as specified per RFC 4213 [75].

Function: The IPv6 Traffic Class field has to be mapped to the IPv4 ToS field of tunnelled traffic as the default behaviour. This MAY be disabled or overwritten by further configuration.

RFC Reference: [75**]**.

6.8.9.10 Feature: MTU Size

The premise, function and related RFC reference relating to this feature is as given below.

Premise: As the maximum MTU size within Docsis is well known, the MTU of the 6rd traffic needs to be modified so that the MTU of the DOCSIS transport is not exceeded.

Function: The 6rd tunnel MTU has to be reduced by to 1 480 bytes, If a packet is received on the LAN interface which exceeds this, it has to be dropped and an ICMP 'fragmentation needed' message sent to the source device.

RFC Reference: N/A.

6.8.9.11 Feature: IPv4 Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Provision of IP Addresses to locally connected clients.

Function: The eRouter has to operate as an IPv4 NAT router providing IPv4 addressing to its client devices as per RFC 1918 [30].

RFC Reference: [30**]**.

6.8.9.12 Feature: NAT

The premise, function and related RFC reference relating to this feature is as given below.

Premise: In order to conserve IPv4 address space, the e-router needs to provide NAT functionality from locally connected IPv4 devices to the single IPv4 address provided to the eRouter via DHCP.

Function: The e-router has to implement a NAT function according to RFC 1631 [27].

RFC Reference: [27**]**.

6.8.9.13 Feature: IPv6 LAN Addressing

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The 6rd enabled e-router needs to provide IPv6 addressing to devices requesting IPv6 addresses.

Function: For each connected LAN interface on the eRouter, a IPv6 address with a /64 mask has to be assigned to each interface. LAN IPv6 addresses MAY be provided via DHCP, in which case the e router has to implement a DHCP server. LAN IP v6 addressing MAY be implemented using SLAAC.

6.8.9.14 Feature: IPv6 Statefull Firewall

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a state full firewall function.

Function: The CPE has to be capable of implementing filters based upon a 5-tuple consisting of the following:

1) Source IP address

2) Destination IP address

3) Source Port number

4) Destination Port number

5) Protocol in use

RFC Reference: N/A.

6.8.9.15 Feature: Recursive DNS Server

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a recursive DNS server.

Function: The CPE device has to implement a recursive DNS server which will contact DNS resources to provide an authorative DNS response in request to a resolution request for a given domain.

RFC Reference: N/A.

6.8.9.16 Feature: User Interface Provision

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE has to implement a user accessible web interface on any implemented Local Area Network interface to aid the configuration of the device and to aid in provision of diagnostics information.

Function: The CPE device has to implement a user accessible web interface on port 80.

RFC Reference: N/A.

6.8.9.17 Feature: Access Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement access control to the local web interface.

Function: Access to the web interface has to be controlled by the use of a specified username and password.

RFC Reference: N/A.

6.8.9.18 Feature: Device Localization

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device user interface has to be configurable for Multiple languages.

Function: The language displayed on the user accessible web interface has to be configurable to support display in multiple languages.

RFC Reference: N/A.

6.8.9.19 Feature: CPE Device Status Indication

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE Device has to contain a user interface tab which provides access to the following item

Function: Listed below:

1) Connection Status

2) Security

3) Diagnostics

4) Software

RFC Reference: N/A.

6.8.9.20 Feature: Router Basic Configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Wide Area network Configuration
2) Local Area Network Configuration
3) DHCP Server Configuration
4) Device Configuration Backup

RFC Reference: N/A.

6.8.9.21 Feature: Advanced router control and configuration.

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a user interface tab with access to the following features.

Function: Listed below:

1) Advanced Options
2) IP Filtering
3) MAC Filtering
4) Port Filtering
5) Port Forwarding
6) Port Triggers
7) DMZ Host Selection

RFC Reference: N/A.

6.8.9.22 Feature: The CPE device has to implement a Firewall user interface with user accessible configuration control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: A Firewall tab should contain subpages on Web Filter, Local Log, Remote Log.

Function: Listed below:

1) Web Filtering
2) Local Logging
3) Remote Logging

RFC Reference: N/A.

6.8.9.23 Feature: User Interface - Parental Control

The premise, function and related RFC reference relating to this feature is as given below.

Premise: The CPE device has to implement a specific user accessible tab which provides access to the following functions.

Function: Listed below:

1) User Setup
2) Basic Setup
3) Content Filtering
4) Time of Day Access Control

RFC Reference: N/A.

6.8.9.24 Feature: User Interface, Wireless Status and configuration

The premise, function and related RFC reference relating to this feature is as given below.

Premise: Where the CPE device implements a wireless interface a specific user accessible tab has to be made available.

Function: The Wireless configuration tab has to include the following items:

1) Radio Interface
2) Security
3) Advanced Functionality
4) Access Control

RFC Reference: N/A.

Premise: User Interface -Where the CPE device implements an integrated Multimedia Terminal Adaptor a specific user configurable tab has to be made available.

Function: The MTA tab has to include the following items:

1) Status of the integrated MTA functionality
2) DHCP status, including any additional interfaces configured for the use of the MTA device
3) Quality of Service
4) Device provisioning
5) Event Logging

RFC Reference: N/A.

6.8.10 Technical Viability

As with all NAT solutions, 6rd has restrictions on inbound connections.

As the NAT functionality is removed from the CPE, existing features such as UPnP NAT traversal and port forwarding can no longer be used to direct inbound packets on predetermined IPv4 TCP or UDP ports to a specific host.

Application Level Gateways on the CPE cannot be used when 6rd is employed. The BR has to provide the additional port mapping requirements for each required ALG.

Other issues are mostly well-known and present topics for further development:

- The cost, particularly of the CPE router, is expected to be increased due to the extension of the required functionality. MSOs that are not deploying WiFi solutions and stick to bridging CPE devices will face the issue of having to exchange CPE just for the reason of introducing 6rd.
- 6rd requires ALGs (ALPs to be exact in most cases) which require the BR to do some form of intelligent vicissitudes to the transit packets bearing the risk of drop in performance and limits in functionality. The risk of service-deprecation is minimized if all functionality is included natively in the BR.
- With 6rd, all functional intelligence is located in the BR. Thus, functionality that requires a public address in the local network has to be performed in the BR itself. An example of this is PCP. In 6rd, it is placed on the BR due to UPnP 1 and 2 requirements.
- 6rd is a technology that is not designed to fade out. All IPv6 traffic is passing the BR. As IPv6 traffic is increasing, continuous investments in BR infrastructure will be required. Also to support network operator growth, BR capacity has to be extended.
- The requirement for Data Retention capacity is rather large and, thus, making the capacity available may become a major issue for the receiving logging server. Normally, on public addresses there is a single entry or two per week depending on lease times and utilization. Within 6rd, each time an individual subscriber receives a port and IP address assignment the event has to be logged. This can amount to 200 assignments a day with a start and stop time resulting in 400 logs per user per day. Assuming that an MSO could have 10 % of its customer base online at one time, that would mean a huge amount of logs.
- 6rd is a hub-and-spoke topology forcing all IPv6 communication from CPE to CPE to pass the BR.

The main advantage of 6rd is that operators can run it over their existing IPv4 infrastructure. There is no requirement for IPv6 support on the access, aggregation, core networks and also not on the BNG. As such it is rapidly deployable.

Annex A (informative): IPv6 Survey of Cable Operator sector

IPv6 transition is emerging across the telecom operator sector and ISPs with equipment suppliers developing solutions. The industry is in the current stage of assessment, design and implementation, with experiences and guidance limited.

The survey is used to assess the current directions and considerations from the sectors of industry, Cable Operators, Manufacturers and Industry Forums.

This annex presents the questionnaire targeting the Cable Operator sector of industry.

Standards Research Survey

Technologies and Strategies for Transition to IPv6

Background

IPv6 transition is critical to the long-term sustainability of European and global networks. As more and more services and industries come to rely on the global Internet as a fundamental platform, the need for ubiquitous connectivity of devices and services becomes very urgent. Such near-term strategic areas like Mobile Internet and Smart Grids as well as the continued growth in residential and business broadband access services are poised to introduce massive numbers of devices that require network connectivity, which may not easily be provided by the current Internet (IPv4) networks with its depleting address space.

Widespread adoption of IPv6 has been identified as the best way forward to address the exhaustion of the IPv4 address space. Prompt and efficient adoption offered potential for innovation and leadership in advancing the Internet, and that delayed adoption of IPv6 would lead to disadvantages for all users and a weaker competitive position of the industry. In the meantime, we have seen the exhaustion of the IANA Unallocated IPv4 Address Pool on 3 February 2011, and the exhaustion of the RIPE NCC IPv4 Address Pool is approaching. The urgency to transition broadband Internet networks to IPv6 is becoming critical.

While device manufacturers, software developers and network operators are adopting IPv6, the vision of an Internet running IPv6 only will not become reality any time soon. For a considerable period of time, significant numbers of devices and services will exist that customers want to use and that require IPv4 connectivity. An immediate replacement of these IPv4 hosts and networks may not be feasible or not desirable for various technical and economic reasons. It is particularly the task of access network operators and broadband service providers to ensure customer choice in terms of technology and services. Appropriate transition technologies enable the coexistence of IPv6 and IPv4 in various parts of the end-to-end network allowing services to be consumed and customer premises equipment to be used transparently while fostering a smooth transition to the required extended address space provided by IPv6.

Integrated broadband cable and television networks are recognized as a key enabler in supporting Europe's Digital Agenda. As of 2009, cable networks go into the home of 67 million customers in the European Union, providing Digital TV, Broadband Internet and Telephony services. Broadband Internet provided by Cable Networks utilizing DOCSIS cable modem technology provide services to 20,9 million subscribers in Europe (2009) with access speeds of currently up to 200 Mbit/s. This figure has grown by at least 12 % annually. As such Cable Networks provide the platform to satisfy fundamental entertainment, communication and information needs to consumers. Furthermore, the industry is anticipating a transition to delivery of digital television using broadband cable modem technology, which will dramatically increase the number of broadband connected households. To continue to meet the demand of accelerating connectivity of digital devices, a standardized approach for the cable eco-system to rapidly transition to IPv6 is required. A failure for an effective standards driven transition would impair the ability to achieve cost effective solutions on a large scale.

Subject of Survey

In the framework of its standardization activities related to cable networks, ETSI has set up a Specialist Task Force 440 in order to assess the current status of Cable equipment standards and the approaches for their transition to IPv6. Since the time to market is a factor considering the depletion of IPv4 addresses, this work accommodates an urgent need in the industry for standardization in this area. Resulting from the findings of its assessment, the Specialist Task Force 440 will develop Technical Specifications for transition technologies in cable networks and define requirements on network and device architecture and design. Detailed technical requirements for the components of the transition technology will be specified as well as means for their verification.

In order to complete its assessment of IPv6 transition technologies as well as to define a comprehensive basis for the development of the Technical Specifications, information is solicited from stakeholders and market participants in the format of this survey. The information will be used to identify standardization requirements and enables the Specialist Task Force 440 to take into account a wide set of strategies and approaches.

NOTE: All information gathered by this survey is anonymously present publicly and will never be associated with your company or body.

Please also attach or paste in links and references to support your answers if possible.

Disclaimer

Responses to this survey are considered confidential and therefore individual responses will not be released, shared, or published. Rather survey results will be reported in aggregate data sets. Therefore participants are encouraged to make a copy of their individual responses to enable them to compare their individual responses to industry survey results once published.

In participating in this survey, the participant recognizes that the information provided on this survey will be used in an industry stakeholders' aggregate report and therefore grants the ETSI Specialist Task Force 440 unrestricted use of this information.

Questions

Q.1) What are viable current technologies enabling transition of your networks to IPv6? Why would you recommend specific transition technologies and what is your comparison technically and business strategy wise between the technologies on offer to date?

Q.2) What technologies have you tested and deployed to date? And what is the degree of that deployment?

Q.3) What is the timescale for the final deployment of products or go live date? How many devices do you expect to be deployed on completion and in what timeframe?

Q.4) What are your views regarding the main technical obstacles when migrating to IPv6 in your networks?

Q.5) What are your views regarding the main commercial / financial obstacles when migrating to IPv6 in your networks?

Q.6) How do you assess the maturity of the technologies you are considering and the availability of products?

Q.7) What is your estimated timescale for the removal of your transition technologies?

Q.8) What timescale do you expect to deplete on IPv4 addressing?

Q.9) What feature and requirement sets are required for each transition technology for your network?

Q.10) Do you expect that any service/protocol currently operated will have to be deprecated when the technology is deployed to give a gap analysis between IPv4/IPv6 native and the transition technology? Please provide details on the associated features that are not supported by the technology, including if and why the service deprecation compared to native IPv4 could be accepted and the failings of the technology in general.

Q.11) How will you ensure interoperability of products and services across the chosen transition technology?

Q.12) What future technologies are you considering developing with vendors that are not presently available in code?

Q.13) Will they be used as a replacement or complimentary to the current approach?

Q.14) Why do you prefer the current technologies you have chosen?

Q.15) What product streams will have this available?

Q.16) What capacity and scalability will be available on these platforms?

Q.17) Open Summary. Please add any other information you believe has not been covered by the questions above.

Annex B (informative): IPv6 Survey of Equipment Suppliers Sector

IPv6 transition is emerging across the telecom operator sector and ISPs with equipment suppliers developing solutions. The industry is in the current stage of assessment, design and implementation, with experiences and guidance limited.

The survey is used to assess the current directions and considerations from the sectors of industry, Cable Operators, Manufacturers and Industry Forums.

This annex presents the questionnaire targeting the Equipment Supplier sector of industry.

Standards Research Product Vendor Survey

Technologies and Strategies for Transition to IPv6

Background

IPv6 transition is critical to the long-term sustainability of European and global networks. As more and more services and industries come to rely on the global Internet as a fundamental platform, the need for ubiquitous connectivity of devices and services becomes very urgent. Such near-term strategic areas like Mobile Internet and Smart Grids as well as the continued growth in residential and business broadband access services are poised to introduce massive numbers of devices that require network connectivity, which may not easily be provided by the current Internet (IPv4) networks with its depleting address space.

Widespread adoption of IPv6 has been identified as the best way forward to address the exhaustion of the IPv4 address space. Prompt and efficient adoption offered potential for innovation and leadership in advancing the Internet, and that delayed adoption of IPv6 would lead to disadvantages for all users and a weaker competitive position of the industry. In the meantime, we have seen the exhaustion of the IANA Unallocated IPv4 Address Pool on 3 February 2011, and the exhaustion of the RIPE NCC IPv4 Address Pool is approaching. The urgency to transition broadband Internet networks to IPv6 is becoming critical.

While device manufacturers, software developers and network operators are adopting IPv6, the vision of an Internet running IPv6 only will not become reality any time soon. For a considerable period of time, significant numbers of devices and services will exist that customers want to use and that require IPv4 connectivity. An immediate replacement of these IPv4 hosts and networks may not be feasible or not desirable for various technical and economic reasons. It is particularly the task of access network operators and broadband service providers to ensure customer choice in terms of technology and services. Appropriate transition technologies enable the coexistence of IPv6 and IPv4 in various parts of the end-to-end network allowing services to be consumed and customer premises equipment to be used transparently while fostering a smooth transition to the required extended address space provided by IPv6.

Integrated broadband cable and television networks are recognized as a key enabler in supporting Europe's Digital Agenda. As of 2009, cable networks go into the home of 67 million customers in the European Union, providing Digital TV, Broadband Internet and Telephony services. Broadband Internet provided by Cable Networks utilizing DOCSIS cable modem technology provide services to 20,9 million subscribers in Europe (2009) with access speeds of currently up to 200 Mbit/s. This figure has grown by at least 12 % annually. As such Cable Networks provide the platform to satisfy fundamental entertainment, communication and information needs to consumers. Furthermore, the industry is anticipating a transition to delivery of digital television using broadband cable modem technology, which will dramatically increase the number of broadband connected households. To continue to meet the demand of accelerating connectivity of digital devices, a standardized approach for the cable eco-system to rapidly transition to IPv6 is required. A failure for an effective standards driven transition would impair the ability to achieve cost effective solutions on a large scale.

Subject of Survey

In the framework of its standardization activities related to cable networks, ETSI has set up a Specialist Task Force 440 in order to assess the current status of Cable equipment standards and the approaches for their transition to IPv6. Since the time to market is a factor considering the depletion of IPv4 addresses, this work accommodates an urgent need in the industry for standardization in this area. Resulting from the findings of its assessment, the Specialist Task Force 440 will develop Technical Specifications for transition technologies in cable networks and define requirements on network and device architecture and design. Detailed technical requirements for the components of the transition technology will be specified as well as means for their verification.

In order to complete its assessment of IPv6 transition technologies as well as to define a comprehensive basis for the development of the Technical Specifications, information is solicited from stakeholders and market participants in the format of this survey. The information will be used to identify standardization requirements and enables the Specialist Task Force 440 to take into account a wide set of strategies and approaches.

NOTE: All information gathered by this survey is anonymously present publicly and will never be associated with your company or body.

Please also attach or paste in links and references to support your answers if possible.

Disclaimer

Responses to this survey are considered confidential and therefore individual responses will not be released, shared, or published. Rather survey results will be reported in aggregate data sets. Therefore participants are encouraged to make a copy of their individual responses to enable them to compare their individual responses to industry survey results once published.

In participating in this survey, the participant recognizes that the information provided on this survey will be used in an industry stakeholder's aggregate report and therefore grants the ETSI Specialist Task Force 440 unrestricted use of this information.

Questions

Q.1) What are viable current technologies enabling transition of your networks to IPv6? Why would you recommend specific transition technologies and their comparison?

Q.2) What transition technologies have you tested and placed on the market as a product to date? And what is the degree of that development (what features are missing from your products on each transition technology)?

Q.3) What is the timescale for the final deployment of products or go live date? How many devices do you expect to be deployed and in what timeframe?

Q.4) What are your views regarding the main technical obstacles when migrating to IPv6 in your clients' networks?

Q.5) What are your views regarding the main commercial / financial obstacles when migrating to IPv6 in your clients' networks?

Q.6) How do you assess the maturity of the transition technologies you are considering and the availability of products?

Q.7) What is the timescale for the final removal of your transition technology where IPv4 no longer is considered for products for your customers?

Q.8) What timescale do you expect your MSO clients to deplete on IPv4 addressing?

Q.9) What main feature and requirement sets do you demand for each transition technology for your products?

Q.10) Do you expect that any service/protocol currently operated will have to be deprecated when the technology is deployed to give a gap analysis between IPv4/IPv6 native and the transition technology? Please provide details on the associated features that are not supported by the technology, including if and why the service deprecation compared to native IPv4 could be accepted and the failings of the technology in general.

Q.11) How will you ensure interoperability of products and services across the chosen transition technology?

Q.12) What future technologies are you considering developing?

Q.13) Will they be used as a replacement or complimentary to the current approach?

Q.14) Why do you prefer the current technologies you have chosen?

Q.15) What product streams will have this available?

Q.16) What capacity and scalability will be available on these platforms?

Q.17) Which transition technology, presently available and in development, do you think will provide the least service deprecation compared to native IPv4.

Q.18) Open Summary. Please add any other information you believe has not been covered by the questions above.

Annex C (informative): IPv6 Survey of Industry Forums and Associations Sector

IPv6 transition is emerging across the telecom operator sector and ISPs with equipment suppliers developing solutions. The industry is in the current stage of assessment, design and implementation, with experiences and guidance limited.

The survey is used to assess the current directions and considerations from the sectors of industry, Cable Operators, Manufacturers and Industry Forums.

This annex presents the questionnaire targeting the Industry Forum sector of industry.

Standards Research Survey

Technologies and Strategies for Transition to IPv6

Background

IPv6 transition is critical to the long-term sustainability of European and global networks. As more and more services and industries come to rely on the global Internet as a fundamental platform, the need for ubiquitous connectivity of devices and services becomes very urgent. Such near-term strategic areas like Mobile Internet and Smart Grids as well as the continued growth in residential and business broadband access services are poised to introduce massive numbers of devices that require network connectivity, which may not easily be provided by the current Internet (IPv4) networks with its depleting address space.

Widespread adoption of IPv6 has been identified as the best way forward to address the exhaustion of the IPv4 address space. Prompt and efficient adoption offered potential for innovation and leadership in advancing the Internet, and that delayed adoption of IPv6 would lead to disadvantages for all users and a weaker competitive position of the industry. In the meantime, we have seen the exhaustion of the IANA Unallocated IPv4 Address Pool on 3 February 2011, and the exhaustion of the RIPE NCC IPv4 Address Pool is approaching. The urgency to transition broadband Internet networks to IPv6 is becoming critical.

While device manufacturers, software developers and network operators are adopting IPv6, the vision of an Internet running IPv6 only will not become reality any time soon. For a considerable period of time, significant numbers of devices and services will exist that customers want to use and that require IPv4 connectivity. An immediate replacement of these IPv4 hosts and networks may not be feasible or not desirable for various technical and economic reasons. It is particularly the task of access network operators and broadband service providers to ensure customer choice in terms of technology and services. Appropriate transition technologies enable the coexistence of IPv6 and IPv4 in various parts of the end-to-end network allowing services to be consumed and customer premises equipment to be used transparently while fostering a smooth transition to the required extended address space provided by IPv6.

Integrated broadband cable and television networks are recognized as a key enabler in supporting Europe's Digital Agenda. As of 2009, cable networks go into the home of 67 million customers in the European Union, providing Digital TV, Broadband Internet and Telephony services. Broadband Internet provided by Cable Networks utilizing DOCSIS cable modem technology provide services to 20,9 million subscribers in Europe (2009) with access speeds of currently up to 200 Mbit/s. This figure has grown by at least 12 % annually. As such Cable Networks provide the platform to satisfy fundamental entertainment, communication and information needs to consumers. Furthermore, the industry is anticipating a transition to delivery of digital television using broadband cable modem technology, which will dramatically increase the number of broadband connected households. To continue to meet the demand of accelerating connectivity of digital devices, a standardized approach for the cable eco-system to rapidly transition to IPv6 is required. A failure for an effective standards driven transition would impair the ability to achieve cost effective solutions on a large scale.

Subject of Survey

In the framework of its standardization activities related to cable networks, ETSI has set up a Specialist Task Force 440 in order to assess the current status of Cable equipment standards and the approaches for their transition to IPv6. Since the time to market is a factor considering the depletion of IPv4 addresses, this work accommodates an urgent need in the industry for standardization in this area. Resulting from the findings of its assessment, the Specialist Task Force 440 will develop Technical Specifications for transition technologies in cable networks and define requirements on network and device architecture and design. Detailed technical requirements for the components of the transition technology will be specified as well as means for their verification.

In order to complete its assessment of IPv6 transition technologies as well as to define a comprehensive basis for the development of the Technical Specifications, information is solicited from stakeholders and market participants in the format of this survey. The information will be used to identify standardization requirements and enables the Specialist Task Force 440 to take into account a wide set of strategies and approaches.

NOTE: All information gathered by this survey is anonymously present publicly and will never be associated with your company or body.

Please also attach or paste in links and references to support your answers if possible.

Disclaimer

Responses to this survey are considered confidential and therefore individual responses will not be released, shared, or published. Rather survey results will be reported in aggregate data sets. Therefore participants are encouraged to make a copy of their individual responses to enable them to compare their individual responses to industry survey results once published.

In participating in this survey, the participant recognizes that the information provided on this survey will be used in an industry stakeholders' aggregate report and therefore grants the ETSI Specialist Task Force 440 unrestricted use of this information.

Questions

Q.1) What are viable current transition technologies enabling transition of ISP networks to IPv6? Why would you recommend specific transition technologies and what is your comparison technically and business strategy wise between the technologies on offer to date?

Q.2) What technologies have you developed, tested and considered for deployment in ISPs to date? And what is the degree of that deployment?

Q.3) What is the timescale for the final deployment of products or go live date in general? How many ISPs do you consider will be live with IPv6, and separately, transition technology deployment in their live networks by 2014 in Europe?

Q.4) What are your views regarding the main technical obstacles when migrating to IPv6 in ISP networks?

Q.5) What are your views regarding the main commercial / financial obstacles when migrating to IPv6 in networks in general?

Q.6) How do you assess the maturity of the technologies you are considering and the availability of products?

Q.7) What is your estimated timescale for the removal of all transition technologies worldwide?

Q.8) What timescale do you expect final depletion in European ISPs on IPv4 addressing?

Q.9) What feature and requirement sets are required for each transition technology for ISP networks at present?

Q.10) Do you expect that any service/protocol currently operated will have to be deprecated when the technology is deployed to give a gap analysis between IPv4/IPv6 native and the transition technology? Please provide details on the associated features that are not supported by the transition technologies you consider the most viable, including if and why the service deprecation compared to native IPv4 could be accepted and the failings of the technologies in general.

Q.11) How will you ensure interoperability of products and services across the chosen transition technology?

Q.12.) What future transition technologies are you considering developing that are not presently available in vendor code or fully released?

Q.13) Will they be used as a replacement or complimentary to the current approach?

Q.14) Why in general do you prefer the current transition technologies you have chosen to support and develop in general?

Q.15) What vendor product streams will have this available?

Q.16) What capacity and scalability will be available on these platforms to your knowledge?

Q.17) Open Summary. Please add any other information you believe has not been covered by the questions above.

Annex D (informative): Bibliography

IETF I-D draft-bcx-behave-address-fmt-extension-02: "Extended IPv6 Addressing for Encoding Port Range".

NOTE: Available at http://tools.ietf.org/id/draft-bcx-behave-address-fmt-extension-02.txt.

IETF I-D draft-ietf-v6ops-6204bis-10: "Basic Requirements for IPv6 Customer Edge Routers".

NOTE: Available at http://tools.ietf.org/html/draft-ietf-v6ops-6204bis-10.

IETF RFC 868: "Time Protocol", May 1983.

NOTE: Available at http://www.ietf.org/rfc/rfc868.

IETF RFC 2131: "Dynamic Host Configuration Protocol", March 1997.

NOTE: Available at http://www.ietf.org/rfc/rfc2131.txt.

IETF RFC 2453: "RIP Version 2", November 1998.

NOTE: Available at http://www.ietf.org/rfc/rfc2453.txt.

IETF RFC 3142: "An IPv6-to-IPv4 Transport Relay Translator", June 2001.

NOTE: Available at http://www.ietf.org/rfc/rfc3142.txt.

IETF RFC 4361: "Node-specific Client Identifiers for Dynamic Host Configuration Protocol Version Four (DHCPv4)", February 2006.

NOTE: Available at http://tools.ietf.org/rfc/rfc4361.txt.

IETF RFC 4380: "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", February 2006.

NOTE: Available at http://www.ietf.org/rfc/rfc4380.txt.

IETF RFC 5572: "IPv6 Tunnel Broker with the Tunnel Setup Protocol (TSP)", February 2010.

NOTE: Available at http://tools.ietf.org/rfc/rfc5572.txt.

IETF RFC 6164: "Using 127-Bit IPv6 Prefixes on Inter-Router Links", April 2011.

NOTE: Available at http://tools.ietf.org/rfc/rfc6164.txt.

IETF RFC 6563: "Moving A6 to Historic Status", March 2012.

NOTE: Available at http://tools.ietf.org/rfc/rfc6563.txt.

CableLabs CM-SP-DOCSIS2.0-IPv6-I07-130404: "Data-Over-Cable Service Interface Specifications; DOCSIS 2.0 + IPv6 Cable Modem Specification", April 2013.

CableLabs CM-SP-OSSIv3.0-I21-130404: "Data-Over-Cable Service Interface Specifications DOCSIS 3.0; Operations Support System Interface Specification", April 2013.

IEEE 802.11a: "IEEE Standard for Telecommunications and Information Exchange between Systems - LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: High Speed Physical Layer in the 5 GHz band", 1999.

IEEE 802.11b: "IEEE Standard for Information Technology - Telecommunications and information exchange between systems - Local and Metropolitan networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher Speed Physical Layer (PHY) Extension in the 2.4 GHz band", 1999.

IEEE 802.11g: "IEEE Standard for Information Technology - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Further Higher Data Rate Extension in the 2.4 GHz Band", 2003.

IEEE 802.11n: "IEEE Standard for Information technology - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput", 2009.

History

Document history		
V1.1.1	October 2013	Publication