ETSI TR 102 155 V1.1.1 (2003-04)

Technical Report

Satellite Earth Stations and Systems (SES);
Broadband Satellite Multimedia;
IP interworking over satellite;
Addressing and routing

Reference

DTR/SES-00072

Keywords

addressing, broadband, internet, interworking,
satellite, IP, multimedia, VPN

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

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, send your comment to:
editor@etsi.org

Copyright Notification

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://webapp.etsi.org/IPR/home.asp).

All published ETSI deliverables shall include information which directs the reader to the above source of information.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Satellite Earth Stations and Systems (SES).

The present document has been generated by ETSI Specialist Task Force STF 214 "Broadband and Satellite integration".

Introduction

In TR 101 984 [3], the general Service and Architectures aspects for BSM (Broadband Satellite Multimedia) systems were introduced. In particular the scenarios defining the position of BSM systems to provide Internet-based services were defined. Mesh and star topologies for both transparent satellites and On-Board Processing (OBP) satellites were described with their generic reference points. IP-over-satellite aspects were introduced in TR 101 985 [4] and functional models were defined for quality of service (QoS), Addressing, Routing and Multicasting.

The present document concerns Addressing and Routing and, based on the above TRs, examines in detail how existing addressing and routing protocols can be used and provides guidelines on modifications or solutions that should be introduced in this area.

The layout is as follows. In clause 4, the reference model used for the study is provided with the relevant definitions. Clause 5 summarizes general requirements for IP Addressing and Routing and outlines the impacts on BMS Systems. In particular, the role of BSMS vis-à-vis Autonomous Systems is detailed. In clause 6, the impact on Addressing and Routing of the position of BSM systems acting either as an access network, a distribution network or a core network is defined.

In clause 7, solutions for routing issues identified in the previous clauses are detailed.

Clause 8 contains general conclusions of the present document and a summary of recommendations for further actions is provided in clause 9.

1 Scope

The present document focuses on addressing and routing in satellite multimedia systems, and specifically on those defined by the term BSMS (Broadband Satellite Multimedia Systems) in TR 101 984 [3] and TR 101 985 [4].

The scope of the present document is to:

- Identify and select use cases and high level architectures applicable to Addressing and Routing of IP packets in broadband multimedia satellite systems.
- Identify satellite-specific requirements for IP Addressing and Routing.
- Identify relevant standardization work in other standards bodies such as IETF, ITU and DVB.
- Recommend topics for standardization.

2 References

For the purposes of this Technical Report (TR), the following references apply.

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

[1] ETSI TR 101 374-1: "Satellite Earth Stations and Systems (SES); Broadband satellite multimedia; Part 1: Survey on standardization objectives".

[2] ETSI TR 101 374-2: "Satellite Earth Stations and Systems (SES); Broadband satellite multimedia; Part 2: Scenario for standardization".

[3] ETSI TR 101 984: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Services and Architectures".

[4] ETSI TR 101 985: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP over Satellite".

[5] ETSI TR 102 156: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP interworking over satellite; Multicasting".

[6] "Diameter Base Protocol (Internet-Draft)".

NOTE: See also the IETF Diameter protocol home page, http://www.diameter.org/.

[7] Global IPv6 Summit: "The IPv6 Implementation Landscape", Peter Loshin, IPv6 Forum, March 14, 2000.

[8] The Economist: "Upgrading The Internet". March 22, 2001.

[9] ISOC MEMBER BRIEFING #6: "The Transition to IPv6", Eric Carmès, January 2002.

[10] "An IP Transport and Routing Architecture for Next generation Satellite Networks", F. Yegenoglu et al. IEEE Network Sept/Oct. 2000.

[11] IEEE Communications Magazine, Vol. 40, no. 6, June 2002 - Special Edition on QoS Routing.

[12] ISO/IEC 3309 (1993): "Information Technology - Telecommunications and information exchange between systems - High level data link control (HDLC) procedures - Frame structure".

[13] ETSI TS 123 107: "Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture (3GPP TS 23.107 version 3.9.0 Release 1999)".

[14] IST BRAHMS project Deliverable 6 (http://docbox.etsi.org/ses/ses/60-WGs/WG_BSM/BRAHMS/Del6_v1.0.zip).

[15] RNAP - A Resource Negotiation and Pricing Protocol, http://www1.cs.columbia.edu/~xinwang/public/projects/protocol.html.

[16] IANA Internet Assigned Number Authority, (http://www.iana.org/).

[17] ETSI EN 300 421: "Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services".

[18] ETSI EN 301 790: "Digital Video Broadcasting (DVB); Interaction channel for satellite distribution systems".

[19] ISO 3166: "Codes for the representation of names of countries and their subdivisions".

[20] TIA Committee TR-34.1: "Satellite ATM Networks: Architectures and Guidelines", May 1998.

[21] IEEE Standard 802.3: "Ethernet LAN".

[22] "Providing IP QoS over GEO Satellite Systems Using MPLS", Ors, T. and Rosenberg, C., International Journal of Satellite Communications, Volume 19, Issue 5, 2001.

[23] IETF RFC 791: "Internet protocol".

[24] IETF RFC 826: "An Ethernet Address Resolution Protocol".

[25] IETF RFC 1517: "Applicability Statement for the Implementation of CIDR".

[26] IETF RFC 1518: "An Architecture for IP Address Allocation with CIDR".

[27] IETF RFC 1519: "CIDR: An Address Assignment and Aggregation Strategy".

[28] IETF RFC 1520: "Exchanging Routing Information Across Provider Boundaries in the CIDR Environment".

[29] IETF RFC 1633: "Integrated Services in the Internet Architecture: an Overview".

[30] IETF RFC 2236: "Internet Group Management Protocol, Version 2".

[31] IETF RFC 2330: "Framework for IP Performance Metrics".

[32] IETF RFC 2332: "NBMA Next Hop Resolution Protocol (NHRP)".

[33] IETF RFC 2333: "NHRP Protocol Applicability Statement".

[34] IETF RFC 2735: "NHRP Support for Virtual Private Networks".

[35] IETF RFC 2990: "Next steps for the IP QoS Architecture".

[36] IETF RFC 3031: "Multi-protocol Label Switching Architecture".

[37] IETF RFC 3077: "A link layer tunnelling mechanism for unidirectional links".

[38] "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)" http://ietf.org/internet-drafts/draft-ietf-dhc-dhcpv6-28.txt.

[39] "Guidelines of Applicability Statements for PPVPNs" http://ietf.org/internet-drafts/draft-ietf-ppvpn-applicability-guidelines-01.txt.

[40] IETF RFC 1166: "Internet numbers".

[41] IETF RFC 950: "Internet Standard Subnetting Procedure".

[42] IETF RFC 1918: "Address Allocation for Private Internets".

[43] IETF RFC 1009: "Requirements for Internet gateways".

[44] IETF RFC 2908: "The Internet Multicast Address Allocation Architecture".

[45] IETF RFC 2730: "Multicast Address Dynamic Client Allocation Protocol (MADCAP)".

[46] IETF RFC 2365: "Administratively Scoped IP Multicast".

[47] IETF RFC 1034: "Domain names - concepts and facilities".

[48] IETF RFC 1035: "Domain names - implementation and specification".

[49] IETF RFC 2131: "Dynamic Host Configuration Protocol".

[50] IETF RFC 951: "Bootstrap Protocol".

[51] IETF RFC 1122: "Requirements for Internet Hosts - Communication Layers".

[52] IETF RFC 3022: "Traditional IP Network Address Translator (Traditional NAT)".

[53] IETF RFC 3135: "Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations".

[54] IETF RFC 1812: "Requirements for IP Version 4 Routers".

[55] IETF RFC 1056: "PCMAIL: A distributed mail system for personal computers".

[56] IETF RFC 2453: "RIP Version 2".

[57] IETF RFC 1583: "OSPF Version 2".

[58] IETF RFC 2328: "OSPF Version 2".

[59] IETF RFC 904: "Exterior Gateway Protocol formal specification".

[60] IETF RFC 1478: "An Architecture for Inter-Domain Policy Routing".

[61] IETF RFC 2386: "A Framework for QoS-based Routing in the Internet".

[62] IETF RFC 2676: "QoS Routing Mechanisms and OSPF Extensions".

[63] IETF RFC 1772: "Application of the Border Gateway Protocol in the Internet".

[64] IETF RFC 2178: "OSPF Version 2".

[65] IETF RFC 2373: "IP Version 6 Addressing Architecture".

[66] IETF RFC 2462: "IPv6 Stateless Address Autoconfiguration".

[67] IETF RFC 1256: "ICMP Router Discovery Messages".

[68] IETF RFC 2080: "RIPng for IPv6".

[69] IETF RFC 2740: "OSPF for IPv6".

[70] IETF RFC 1933: "Transition Mechanisms for IPv6 Hosts and Routers".

[71] IETF RFC 2767: "Dual Stack Hosts using the Bump-In-the-Stack Technique (BIS)".

[72] IETF RFC 2005: "Applicability Statement for IP Mobility Support".

[73] IETF RFC 2002: "IP Mobility Support".

[74] IETF RFC 2543: "SIP: Session Initiation Protocol".

[75] IETF RFC 2784: "Generic Routing Encapsulation (GRE)".

[76] IETF RFC 1661: "The Point-to-Point Protocol (PPP)".

[77] IETF RFC 2516: "A Method for Transmitting PPP Over Ethernet (PPPoE)".

[78] IETF RFC 2364: "PPP Over AAL5".

[79] IETF RFC 1332: "The PPP Internet Protocol Control Protocol (IPCP)".

[80] IETF RFC 1552: "The PPP Internetworking Packet Exchange Control Protocol (IPXCP)".

[81] IETF RFC 2661: "Layer Two Tunneling Protocol L2TP".

[82] IETF RFC 1827: "IP Encapsulating Security Payload (ESP)".

[83] IETF RFC 2547: "BGP/MPLS VPNs".

[84] IETF RFC 903: "Reverse Address Resolution Protocol".

[85] IETF RFC 2748: "The COPS (Common Open Policy Service) Protocol".

[86] IETF RFC 3181: "Signaled Preemption Priority Policy Element".

[87] ITU-T Recommendation Y.1311: "Network Based VPNs - Generic Architecture and Service Requirements".

[88] ITU-T Recommendation Q.2931: "Digital Subscriber Signalling System No. 2 - User-Network Interface (UNI) layer 3 specification for basic call/connection control (Modified by ITU-T Q.2971 (10/1995))".

[89] IETF RFC 1735: "NBMA Address Resolution Protocol (NARP)".

NOTE: Annex A contains a complete list of BGP and IPv6 related protocols.

3 Abbreviations

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

AAA	Authentication, authorization, and accounting
ABR	Area Border Router
ADSL	Asymmetric Digital Subscriber Loop
ALG	Application Level Gateway
AR	Address Resolution
ARP	Address Resolution Protocol
ARPA	Advanced Research Projects Agency
AS	Autonomous System
ASBR	AS Boundary Router
BARS	BSMS Address Resolution Server
BE	Best Effort
BGP	Border Gateway Protocol
BoD	Bandwidth on Demand
BOOTP	BOOTstrap Protocol
BRS	BSMS Route Server
BSM	Broadband Satellite Multimedia
BSMS	Broadband Satellite Multimedia System
BSS	BSMS Signalling Server
CE	Customer Edge
CIDR	Classless Inter Domain Routing
CL	Connectionless
CO	Connection-Oriented
CPE	Customer Premises Equipment
DAMA	Demand Assigned Multiple Access
DHCP	Dynamic Host Configuration Protocol
DiffServ	Differentiated Services
DNS	Domain Name Server
DTCP	Dynamic Tunnel Configuration Protocol
DV	Distance Vector
DVB	Digital Video Broadcasting
DVMRP	Distance Vector Multicast Routing Protocol
EGP	Exterior Gateway Protocol
FCAPS	Fault, Configuration, Accounting, Performance, Security

FIB	Forward Information Base
GEO	Geostationary Earth Orbit
GRE	Generic Routing Encapsulation
HA	Home Agent
HTTP	Hyper-Text Transfer Protocol
IAP	Internet Access Provider
IANA	Internet Assigned Number Authority
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IGP	Interior Gateway Protocol
IP	Internet Protocol
IPv6	Internet Protocol version 6
IPR	Intellectual Property Rights
IS-IS	Intermediate System to Intermediate System
ISO	International Organization for Standardization
ISP	Internet Service Provider
ITU	International Telecommunication Union
ITU-R	ITU Radiocommunication sector
L2TP	Layer 2 Tunnelling Protocol
LAN	Local Area Network
LDP	Label Distribution Protocol
LDS	Label Distribution Server
LER	Label Edge Router
LIR	Local Internet Registries
L-O	Label-Oriented
LS	Link State
LSA	Link State Advertisement
LSP	Label Switched Path
LSR	Label Switched Router
MAC	Medium Access Control
MBONE	Multicast BackbONE
MN	Mobile Node
MPLS	Multi-Protocol Label Switching
NAT	Network Address Translation
NAT-PT	Network Address Translation- Protocol Translation
NBMA	Non-Broadcast Multi-Access
NCC	Network Control Centre
ND	Neighbour Discovery
NHRP	Next Hop Resolution Protocol
OBC	On-Board Controller
OBP	On-Board Processing
OBS	On-Board Switch
OSPF	Open Shortest Path First
PE	Provider Edge
PEP	Performance Enhancing Proxy
PID	Packet IDentifier
PPP	Point-to-Point Protocol
PPPoA	Point-to-Point Protocol over ATM
PPPoE	Point-to-Point Protocol over Ethernet
PPTP	Point-to-Point Tunnelling Protocol
QoS	Quality of Service
RARP	Reverse Address Resolution Protocol
RAS	Remote Access Server
RFC	IETF Request For Comments
RIB	Routing Information Base
RIP	Routing Information Protocol
RIR	Regional Internet Registries
RNAP	Resource Negotiation And Pricing Protocol
RSVP	Resource Reservation Protocol
SD	Satellite Dependent
SDAF	Satellite Dependent Adaptation Functions
SI	Satellite Independent

SIAF	Satellite Independent Adaptation Functions
SIP	Session Initiation Protocol
SME	Small or Medium Enterprise
SLC	Satellite Link Control
SMAC	Satellite Medium Access Control
SOHO	Small Office Home Office
SP	Service Provider (public Internet or private corporate)
SPF	Shortest Path First
ST	Satellite Terminal
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TIA	Telecommunications Industry Association (US)
TTL	Time To Live
UDLR	UniDirectional Link Routing
URL	Universal Resource Locator
VCI	Virtual Circuit Identifier
VLSM	Variable Length Subnet Mask
VP	Virtual Path
VPI	Virtual Path Identifier
VPN	Virtual Private Network
VSAT	Very Small Aperture Terminal (satellite)
w.r.t.	with respect to

4 Overview

4.1 Assumptions and requirements

IP addressing and routing standards from the Internet should be applied to the Broadband Multimedia Satellite System (BSMS) to the maximum extent possible. These requirements include the interworking of the BSMS with the Internet at the IP layer with the aim of making the satellite transparent to the network. The means by which the BSMS and its protocols layers below IP may support addressing, transport and routing will also be discussed.

The requirements for addressing and routing considered in the present document are therefore focussed primarily on the IP layer (assumed to be at the ISO Networking layer - 3), and especially on any requirements that are specific to BSMS as defined as in [3] and [4].

In satellite networks the provision of IP services and interworking is also closely associated with resource management capabilities in order that IP packet transport is carried out efficiently. Compared with terrestrial networks in which bandwidth is readily available for handling bursty traffic, satellites must carefully conserve precious link resources and match IP traffic demand to allocated capacity.

Another important issue is scalability. When the number of hosts or subnetworks interconnected through a satellite access network grows, the complexity of the system may increase non-linearly (e.g. quadratically for mesh networks). Over the coverage of the satellite network, the number of terminals to be interconnected could potentially be of the order of tens of thousands or greater. The scalability of routing protocols implemented in BSMSs is an important factor when aiming to keep the cost and complexity of satellite hardware and capacity utilization low for an increasing number of users. On the other hand if the number of links can be kept low through, for example, star connections then better scalability is achieved.

4.2 Reference model

In the present document, the following reference model, as defined in [4], is used as a general framework. The model defines two components:

- The address resolution function in the Control-plane. This function is used to determine the satellite link address when the address translation is unknown. The results of address translation are stored in the cache for future use.

- The satellite address mapping function in the User-plane. This function maps the IP address to the corresponding satellite link address (e.g. a Satellite MAC address). This function makes use of an address cache which stores the address pairs.

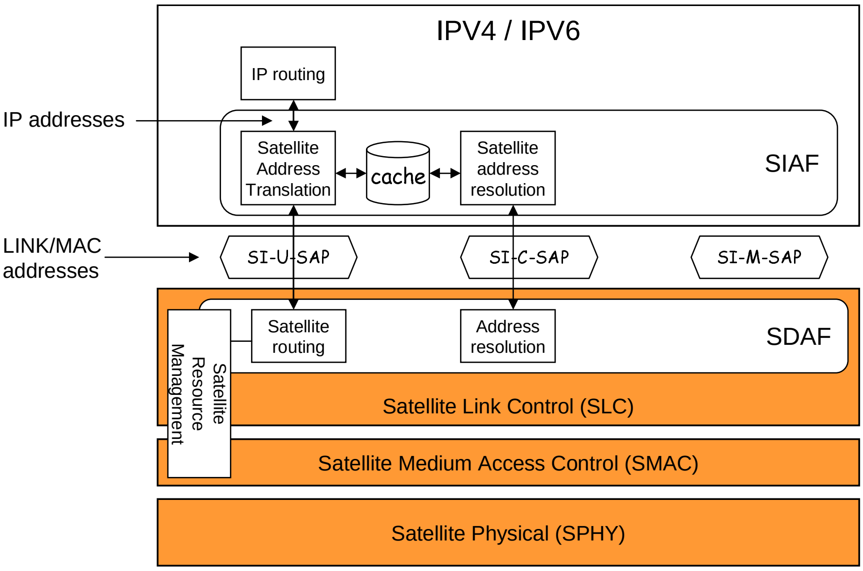

IPV4 / IPV6
IP routing
IP addresses
Satellite Address Translation
cache
Satellite address resolution
SIAF
LINK/MAC addresses
SI-U-SAP
SI-C-SAP
SI-M-SAP
Satellite routing
Address resolution
SDAF
Satellite Resource Management
Satellite Link Control (SLC)
Satellite Medium Access Control (SMAC)
Satellite Physical (SPHY)

Figure 4.2.1: BSMS protocol reference model

5 Background to IP addressing and routing

This clause summarizes general requirements for IP addressing and routing, and outlines the impacts on the BSMS where applicable. A more focussed discussion of the key issues is provided in subsequent clauses.

The majority of the discussion is focussed on IPv4, while IPv6 aspects are summarized in clause 5.4.

5.1 IP addressing

A satellite access network with its wide coverage must envisage multiple usage scenarios shared by many users, both corporate and individual, and by operators or service providers.

Assumption 1	The BSMS is transparent to, and interworks with, the relevant range of addressing schemes (i.e. global, private, IPv4, IPv6 etc.) present in IP flows over the BSMS simultaneously.

This wide range of IP addressing also implies that centralized BSMS routers could have complex routing tables and a low level of route aggregation.

Assumption 2	The BSMS is compatible with IP address assignment protocols.

The allocation of addresses is a major consideration in view of the routing requirements. Addressing in IPv4 is an ever-increasing problem owing to the limited address space of 32 bits (with reserved values) compared with the increasing number of hosts. A number of schemes (e.g. CIDR, subnets, NAT, DHCP - see below) have been introduced to solve this problem by conserving and re-using addresses, which have impacts on Internet protocols and routing.

Allocation of IP addresses to ISPs, particularly to satellite ISPs, is another major issue since it affects their flexibility to change upstream providers easily and to allocate addresses to their own customers.

IPv6 is slowly being introduced, increasing the addressing space to 128 bits, which may resolve many of these problems, but compatibility of IPv6 with and evolution from IPv4 needs to be considered. IPv6 addressing is described in clause 5.2.6.

This clause summarizes the main IPv4 addressing issues and discusses where applicable the impacts on BSMSs.

5.1.1 IPv4 addressing issues

While the initial design of IP addressing enabled the Internet to grow in the last decade, network engineers are constantly challenged to design and implement ever more efficient addressing schemes. The impact of a poorly designed addressing architecture can be catastrophic, particularly in a local context.

IPv4 (the version currently implemented in the Internet) allocates 32 bit addresses to host interfaces.

IP addresses can have three possible uses:

1) The address of an IP network (a group of IP devices sharing common access to a transmission medium - such as all being on the same Ethernet segment). A network number will always have the interface (host) bits of the address space set to 0.

2) The broadcast address of an IP network (the address used to "talk", simultaneously, to all devices in an IP network). Broadcast addresses for a network always have the interface (host) bits of the address space set to 1. Historically the "all 0s" address was also used for a network-wide broadcast; a few systems still use this convention.

3) The address of an interface (such as an Ethernet card or PPP interface on a host, router, print server etc). These addresses can have any value in the host bits except all zero or all 1s.

The expression "unicast" is often used to denote the type of addresses for individual interfaces in IPv4 , following IPv6 terminology (see clause 5.4.2). This differentiates them from multicast addresses which are employed non-uniquely for a group.The Internet currently uses a mixture of two main forms of addressing:

a) Newer CIDR-based (Classless Inter-Domain Routing).

b) Original Class A, B and C addresses, now called Classful addressing.

Almost all new routers support CIDR and the Internet authorities strongly encourage all users to implement the CIDR addressing scheme.

IP addresses are allocated by the Internet Assigned Numbers Authority (IANA) [16] in turn to Regional Internet Registries, then to Local Internet Registries (IRs), and then on to end users. The distribution of IP address space follows the hierarchical scheme described in IETF RFC 1466.

European Internet Registry Policies and Procedures are described in ripe-140.

5.1.1.1 Classful addressing

Five classes of IPv4 address were originally defined (IETF RFC 1166 [40]):

<table>
<tr><td>Class</td><td colspan="8">IP Address bits</td></tr>
<tr><td>A:</td><td>0</td><td colspan="4">Network</td><td colspan="3">Host</td></tr>
<tr><td>B:</td><td>1</td><td>0</td><td colspan="4">Network</td><td colspan="2">Host</td></tr>
<tr><td>C:</td><td>1</td><td>1</td><td>0</td><td colspan="4">Network</td><td>Host</td></tr>
<tr><td>D:</td><td>1</td><td>1</td><td>1</td><td>0</td><td colspan="4">Multicast Address</td></tr>
<tr><td>E:</td><td>1</td><td>1</td><td>1</td><td>1</td><td colspan="4">Addresses reserved for future use.</td></tr>
</table>

In addition to the above table, the all-0s ("this network") and all-1s ("broadcast") host-numbers may not be assigned to individual hosts.

Classful Routers used the first 3 bits of the IP address to determine the type, size and position of the network address.

Class A, B and C networks are now referred to (in CIDR notation - see below) as "/8s", "/16s" and "/24s" respectively according to their network prefix. The /8 address space, for example, occupies 50 % of the total IPv4 address space.

5.1.1.2 Almost all remaining IP network numbers available for allocation today are Class C addresses.Sub-networking

Sub-netting [41] was introduced to utilize the IP address space more efficiently , by taking the "network" address and splitting it up locally for use on several interconnected networks. As far as the world outside the sub-net is concerned, it is still a single IP "network"; sub-networking is a local configuration invisible to the rest of the world.

Subnetting assigns each organization one (or at most a few) network number(s) from the IPv4 address space. The organization is then free to assign distinct subnetwork numbers for its internal networks, and to deploy additional subnets without needing to obtain a new network number.

Subnetting thus adds a third level of hierarchy to the structure of addresses. It divides the classful host-number field into two parts:

1) the subnet-number; and

2) the host-number on that subnet.

The subnet structure of a network is never visible outside of the organization's local network; the route from the Internet to any subnet of a given IP address is the same, no matter on which subnet the destination host is located, since all subnets of a given network number use the same network-prefix but different subnet numbers. The routers within the organization need to differentiate between the individual subnets, but as far as the Internet routers are concerned, all of the subnets in the organization are collected into a single routing table entry. This allows the local administrator to introduce arbitrary complexity into the network without affecting the size of the Internet's routing tables.

5.1.1.2.1 Extended-network-prefix and the Subnet Mask

Internet routers use only the network-prefix of the destination address to route traffic to a subnetted environment. Local routers within the subnetted environment use the "extended-network-prefix" to route traffic between the individual subnets. The extended-network-prefix is composed of the classful network-prefix and the subnet-number.

The extended-network-prefix has traditionally been identified by the "subnet mask". The bits in the subnet mask and in the Internet address have a one-to-one correspondence. The bits of the subnet mask are set to 1 if the system examining the address is to treat these corresponding bits in the IP address as part of the extended-network-prefix i.e. the subnet to which an IP address belongs is found by a bitwise AND operation on the mask and the IP address. The bits in the mask are set to 0 if the system is to treat these bits as part of the host-number.

5.1.1.2.2 Variable Length Subnet Masks (VLSM)

VLSM (Variable Length Subnet Mask) supports more efficient use of an organization's assigned IP address space. One of the major problems with the earlier limitation of use of only a single subnet mask across a given network (in the RIP-1 protocol) was that once the mask was selected, it locked the organization into a fixed number of fixed-sized subnets.

IETF RFC 1009 [43] specifies how a subnet can use more than one subnet mask. When an IP network is assigned more than one subnet mask, it is considered a network with "variable length subnet masks" since the extended-network-prefixes have different lengths. Modern routing protocols, such as RIP-2, OSPF and I-IS-IS, must then be used in the Interior Gateway Protocol (IGP) to enable VLSM by providing the extended-network-prefix length or mask value along with each route advertisement. This permits each subnetwork to be advertised with its corresponding prefix length or mask.

5.1.1.2.3 Private Network Addressing

With the proliferation of IP technology world-wide, including outside the Internet itself, many non-connected enterprises use IP technology and its addressing capabilities solely for local communications (Intranets), without any intention to connect to other enterprises or to the Internet itself. Enterprises themselves also enjoy a number of benefits from their usage of private address space: They gain flexibility in network design by having more address space at their disposal than they could obtain from the globally unique pool. This enables operationally and administratively convenient addressing schemes as well as easier growth paths.

Use of private address space [42] where global uniqueness is not required has the advantage for the Internet at large of conserving the globally unique address space. It also overcomes such problems as the growth of Internet routing tables and the need for local administrators to request another network number before a new network could be installed at their sites.

The IANA has reserved the following three blocks of the IP address space for private networks (IETF RFC 1918 [42]):

1) 10.X.X.X

2) 172.16.X.X to 172.31.X.X

3) 192.168.X.X

The first block is a single class A network number, while the second block is a set of 16 contiguous class B network numbers, and third block is a set of 255 contiguous class C network numbers.

An enterprise that decides to use IP addresses within the above address space can do so without any co-ordination with IANA or an Internet registry. The address space can thus be used by many enterprises. Addresses within this private address space will only be unique within the enterprise and must not be communicated outside the scope of the enterprise.

Any enterprise that needs globally unique address space is required to obtain such addresses from an Internet registry. An enterprise that requests IP addresses for its external connectivity will never be assigned addresses from the blocks defined above.

Hosts within such enterprises may use IP addresses that are unambiguous within their Intranet, but may be ambiguous within the Internet which would prevent access external to the enterprise unless other measures were taken, such as NAT and DHCP.

5.1.1.3 Classless Inter-Domain Routing (CIDR)

Classless Inter Domain Routing (CIDR) (IETF RFCs 1517 [25]/1518 [26]/1519 [27]/1520 [28]) was introduced in order to limit:

1) The growth of routing tables in Internet routers beyond the ability of software (and people) to manage them effectively.

2) The need for allocating new IP network numbers.

The goal of CIDR was to reduce routing entries in the backbone routers, which began to overflow due to the huge number of entries needed for class C networks (up to about 2 million). The problem of scaling in routing mainly relates to Internet backbone routers, since they have to know addresses of all networks on the Internet. After implementing CIDR this number decreased significantly, allowing time for developing long term solutions (especially IPv6). CIDR allows just one routing entry in a router for a whole block of class C networks and defines rules on how to build these blocks.

CIDR replaces classful addressing by the concept of a "network-prefix" to determine the dividing point between the network number and the host number. Routers use the network-prefix, rather than the first 3 bits of the IP address, to determine the network address. CIDR is thus a replication of the private "subnet" concept in the public addressing domain.

Instead of being limited to the classful network identifiers (or "prefixes") of 8, 16 or 24 bits, CIDR currently uses prefixes anywhere from 13 bits to 27 bits. Thus blocks of addresses can be assigned to networks as small as 32 hosts or to those with over 500 000 hosts. This allows for address assignments that fit much more closely to an organization's needs. The prefix is a way of specifying the number of left-most contiguous bits in the network-portion of each routing table entry. For example, a network with 20 bits of network-number and 12-bits of host-number would be advertised with a 20-bit prefix length (a "/20"). All prefixes with a /20 prefix represent the same amount of address space (212 or 4 096 host addresses). Furthermore, a /20 prefix can be assigned to a traditional Class A, Class B, or Class C network number.

5.1.1.3.1 Address allocation strategy

Since the adoption of CIDR, addresses are distributed via a hierarchical set of organizations.

The Internet Assigned Numbers Authority (IANA) allocates large contiguous address blocks to the three Regional Internet Registries (RIRs): RIPE, APNIC and ARIN. They in turn give smaller blocks to two types of Local Internet Registries (LIRs):

1) Provider LIRs (top-level ISPs).

2) Enterprise LIRs (for private networks).

Currently, large blocks of addresses are assigned to the big Provider LIRs, who then re-allocate portions of their address blocks to their customers. These customers, who may be smaller ISPs themselves, may in turn re-allocate portions of their address block to their users and/or customers.

5.1.1.3.2 Hierarchical routing aggregation

The CIDR addressing and allocation scheme enables "route aggregation" in which a single high-level entry can represent many lower-level routes in the global routing tables. Route aggregation brings major advantages for faster forwarding and reduced traffic due to fewer route advertisements between routers.

The Provider LIRs address blocks divided among customers or lower level ISPs allow these lower level networks and hosts to be represented by a single large ISP route entry in the global routing tables. Thus the growth in the number of routing table entries at each level in the network hierarchy is significantly reduced. Currently, global routing tables have approximately 35 000 entries.

The scheme is similar to the telephone network where the network is set up in a hierarchical structure. A high level, backbone network node only looks at the area code information and then routes the call to the specific backbone node responsible for that area code. The receiving node then looks at the phone number prefix and routes the call to its subtending network node responsible for that prefix and so on. The backbone network nodes only need routing table entries for area codes, each representing huge blocks of individual telephone numbers, not for every unique telephone number.

Conclusion 1	In the BSMS, routing aggregation as defined in CIDR, may be limited due to the potentially large number of ISPs networks covered.

A solution to route aggregation for the BSMS could be to use partitioned routing and forwarding tables, between different ISP domains. A centralized route server could also help to solve this problem.

5.1.1.3.3 Difficulties with CIDR

A problem with CIDR may occur when a customer changes ISP but wants to keep his IP addresses: the old ISP still announces the route to the entire block while the new ISP cannot aggregate the old address block as part of its aggregation, so it must inject an exception route into the Internet; there are two routes for that network: the CIDR route and the single route. Possible solutions are to use:

1) The most specific route. This has the disadvantage of needing a new entry in a backbone router, which CIDR would have prevented.

2) NAT: the customer keeps the addresses of the first provider for internal use but uses address translation to translate them into addresses of the new provider when communicating over the Internet.

5.1.1.4 Multicast addressing

The Internet Assigned Numbers Authority (IANA) has assigned the old Class D address space to be used for IPv4 multicast. This means that all IPv4 multicast group addresses will fall in the range of 224.0.0.0 to 239.255.255.255.

Multicast IP addressing does not impose any new requirements in BSMSs compared with terrestrial networks, and the BSMS must be compatible with the addressing adopted for global Internet multicasting.

Conclusion 2	Multicast address resolution or the mapping of IP addresses into satellite layer 2 addresses is specific to the Satellite -Dependent (SD) layers.

Unlike IP "unicast" addresses, IP multicast addresses are not allocated to specific hosts, but instead to services and to groups accessing the services, and a receiving host must identify and "listen" to one or more chosen addresses.

A refinement of Multicast addressing has been obtained by use of a subset of the multicast address space, labelled "administratively scoped" addresses (IETF RFC 2365 [46]), in the domain 239.0.0.0 to 239.255.255.255. This prevents the forwarding of IP multicast packets outside administratively restricted domains. This mechanism is much more efficient than the current use of TTL-scoped addressing (using small TTL values restricts the distribution of multicast packets when large TTL decrements are applied in border routers), using the TTL field in the IP header.

Typical MBONE (Multicast Backbone) usage has been to engineer TTL thresholds that confine traffic to some administratively defined topological region. The basic forwarding rule for interfaces with configured TTL thresholds is that a packet is not forwarded across the interface unless its remaining TTL is greater than the threshold.

Administratively scoped addresses enable multicast technology to be used for communication among small user groups (e.g. videoconferencing) without spreading the associated state information all over the Internet (which would be hard to justify, regarding the fairly limited savings in bandwidth).

For further details refer to [5].

Multicast address allocation is an essential part of using IP multicast. Multicast addresses are an even more limited resource than unicast addresses, and must be allocated dynamically if they are to satisfy expected demand. Though there are in principle 250 million multicast addresses (2^{28}) available in IPv4, these addresses are assigned globally and can become quickly exhausted as multicast usage grows. Multicast addresses must be obtained by the originating server or service provider.

Multicast addresses may be assigned in three ways.

1) by a network administrator (e.g. GLOP - IETF RFC 2770);

2) using a randomly chosen address within a specific range (e.g. Source Demand Routing);

3) using an address leased for a finite period (e.g. MALLOC - see below).

Regarding the latter, the IETF "MALLOC" group is defining protocols which work together to form a global dynamic multicast address allocation mechanism (IETF RFCs 2908 [44] and 2730 [45]).

5.1.1.5 Domain Name System (DNS)

The Domain Name System (DNS) (IETF RFCs 1034 [47] and 1035 [48]) is the method by which a user can find Internet addresses of remote locations starting from mnemonic forms (such as sunc.scit.wlv.ac.uk) by converting them into the equivalent numeric IP address such as 134.220.4.1. To the user and application (e-mail, Web browser, ftp etc.) this translation is a service provided either by the local host or from a remote host via the Internet. The DNS server may communicate with other Internet DNS servers if it cannot translate the address itself.

Proxy DNS servers are often included in service provider or corporate Intranet gateways for higher DNS performance.

5.1.1.5.1 DNS name structure

DNS names are constructed hierarchically, the highest level of the hierarchy being the last component or label of the DNS name. Labels can be up to 63 characters long and are case insensitive. A maximum length of 255 characters is allowed. Labels must start with a letter and can only consist of letters, digits and hyphens.

DNS names can be relative or fully qualified. A fully qualified name includes all the labels and is globally unique. A relative name can be converted by appending the local domain information. For example sunc.scit.wlv.ac.uk is a fully qualified name for the host "sunc" in the domain "scit.wlv.ac.uk".

The final most significant label of a fully qualified name can fall into one of three classes:

1) Advanced Research Projects Agency (ARPA)

   This is a special facility used for reverse translation, i.e. going from IP address to fully qualified domain address. If everything is properly configured a suitably framed query for 1.4.220.134.in-addr.arpa will return sunc.scit.wlv.ac.uk. Details of this will be described later.

2) Three letter codes

   The DNS was originally introduced in the US and the final component of an address was intended to indicate the type of organization hosting the computer. Some of the three letter final labels (edu, gov, mil) are still only used by organizations based in the USA; others can be used anywhere in the world.

   Some of the most common three letter codes are for example:

-	com	Commercial. Now international.
-	edu	Educational.
-	gov	Government.
-	int	International Organization.
-	mil	Military.
-	net	Network related.
-	org	Miscellaneous Organization.

3) Two letter codes

   The final two letter codes indicate the country of origin and are defined in ISO 3166 [19]. The two letter code “us” is used by some sites in the US.

   In some countries there are sub-domains indicating the type of organization such as ac.uk, co.uk in the UK and edu.au and com.au in Australia. Most European countries have not yet adopted this practice.

   For an IP service provider to obtain a domain address it is necessary to identify the administrator of the required domain and then send the administrator the required code and the associated IP address and they will, if they accept the request, include the details in their databases. Conditions for acceptance vary widely between administrators, the administrators for the com and org being, apparently, quite happy to accept anything from anywhere.

5.1.1.5.2 DNS servers and their databases

For any group of computers partaking of the DNS naming scheme there is likely to be a single definitive list of DNS names and associated IP addresses. The group of computers included in this list is called a zone. A zone could be a top-level national domain or a university department. Within a zone, DNS service for subsidiary zones may be delegated along with a subsidiary domain. The computer that maintains the master list for a zone is said to have authority for that zone and will be the primary name server for that zone, there will also be secondaries for that zone.

When any process needs to determine an IP address given a DNS address it calls upon the local host to resolve the address. This can be done in a variety of ways:

- Table lookup. On Unix hosts the table is called /etc/hosts.
- The process communicates with a local name server process. This is commonly called "named" on a Unix system. "named" initially obtains information from /etc/hosts but also maintains a cache of recent requests.
- It sends a message to a remote system that is identified from the information in the file /etc/resolv.conf.
- Finally if a network information system (NIS) is in use DNS service may be one of the facilities provided by the network information system. Most SUN systems work this way although the NIS master will use one or more of the techniques described above to build and maintain the master database.

If a named process cannot resolve an address locally it will call upon higher authority. Ultimately it will attempt to contact the system that is authoritative for the zone in question, however, unless the information is cached or in the hosts local files then it will not know the address of the authoritative server. This problem is resolved by recursive resolution of requests, i.e. any DNS server will pass requests it cannot handle to a higher level server and so on until either the request can be handled (either by sending a message to the identified authoritative host) or until the root of the DNS name space is reached.

There is a small number (e.g. eight) of servers that can serve requests at the root of the DNS name space, all servers should know their IP addresses so that DNS service can be offered even if there are no cached addresses and no local servers indicated by the /etc/resolv.conf file. The root servers will know the IP addresses of the servers for all the national DNS zones and the three letter zones.

Also note that for most users a failure of the DNS service is regarded as a complete network outage, since they are no longer able to use the majority of Internet applications.

5.1.1.5.3 DNS relationship with the BSM

In the case of the BSMS one or more DNS servers may be situated at several points within the network, for example as proxies at gateways of corporate networks, and at the hub stations or satellite gateway stations.

One potential problem is that DNS conflicts with NAT. Domain Names are an issue for hosts which use local DNS servers behind a NAT device. Such servers return site-specific information which may conflict with true Internet names and addresses.

Conclusion 3	Implementing a DNS server in the satellite gateway station is a way of avoiding conflict with NAT.

5.1.2 Unicast address management

Address Management covers allocation of network layer addresses, and mapping to link layer addresses.

5.1.2.1 Dynamic Host Configuration Protocol (DHCP)

The Dynamic Host Configuration Protocol (IETF RFC 2131 [49]) allows hosts to be automatically configured on joining an IP network with reusable IP addresses and other parameters. A client-server model is used with a designated DHCP server located at the ISP for individual users, or on the corporate network. DHCP is based on the earlier Bootstrap Protocol, BOOTP (IETF RFC 951 [50]).

DHCP is the industry standard protocol for dynamic IP assignment, but for individual customers many ISPs allocate IP addresses via PPP (Point-to-Point Protocol) or PPPoE (PPP over Ethernet).

DHCP supports three mechanisms for IP address allocation:

1) Automatic Allocation: DHCP assigns a permanent IP address to a client.

2) Dynamic allocation: DHCP assigns an IP address to a client for a limited period of time (or until the client explicitly relinquishes the address).

3) Manual Allocation: a client's IP address is assigned by the network administrator, and DHCP is used simply to convey the assigned address to the client.

A particular network will use one or more of these mechanisms, depending on the policies of the network administrator.

5.1.2.1.1 DHCP over satellite

All DHCP messages are IP broadcast messages and can be efficiently transported over satellite, particularly if hosts (e.g. residential) are directly connected through the satellite.

However DHCP broadcast messages do not, by default, cross router interfaces. When there is more than one subnet in the network, such as an Intranet, a solution is to put a DHCP server on each segment.In a large organization, placing a DHCP Server on each segment increases cost and administrative effort, and a better option is to use fewer DHCP servers and place these machines in central locations.

Conclusion 4	In a BSMS context, Routers should be configured to pass DHCP/BOOTP messages selectively (BOOTP Relay).

A description of DHCP issues over satellite is given in draft-ietf-udlr-experiments-00.txt.

Forwarding of network-directed broadcasts is permitted but must default to OFF unless specifically allowed. That said, there are other problems with using a network-directed broadcast with DHCP (or BOOTP), namely that a client that does not yet have a subnet mask configured cannot tell the difference between a network-directed broadcast address and a unicast address that happens to have a string of 1s at the tail end. A network-directed broadcast, however, will be sent as a link level broadcast when it arrives at the destination subnet, and according to IETF RFC 1122, clause 3.3.6 [51] should be discarded: A host SHOULD silently discard a datagram that is received via a link-layer broadcast (see clause 2.4) but does not specify an IP multicast or broadcast destination address.

Fortunately, DHCP servers do not in general transmit replies to clients to a broadcast address (see the discussion of the BROADCAST flag in IETF RFC 2131 [49] for exceptions) and when they do it is always to a client on an attached subnet (a BOOTP relay agent to speak to clients on a remote subnet). So there is never any reason for a DHCP server to use a network-directed broadcast in preference to all-1s.

5.1.2.2 Network Address Translation (NAT)

Network Address Translation (IETF RFC 3022 [52]) is the function translating between private IP addresses and global IP addresses.

NAT was invented as a "hack" to circumvent IPv4 address shortage. Meanwhile it has proven to be useful in completely different fields, and is likely to stay with us for much longer, especially considering the progress of IPv6 penetration.

However NATs considerable implementation challenges mean that its use must be carefully considered.

NATs, in common with firewalls, do not just relay packets from one side to another but also control the data flows. They must therefore know as much about every connection as each network device knows about its own connections, i.e. they must keep state information. This requires a significant overhead compared to simply routing packets as in a normal router.

If NAT is used, all packets must go through the NAT-router, i.e. there must not be any alternative routes a packet could take, circumventing the address translation. This should be no problem for private networks, since NAT routers are mostly placed on the borders of internal (leaf) networks.

5.1.2.2.1 Problems with NAT

There are a number of potential problems with the use of NAT; NAT is often regarded as undesirable as it affects the operation of existing applications and security protocols.

Since NAT is application-unaware, applications which include IP addresses embedded within the IP payload (e.g. SIP/SDP, FTP, DNS) each require separate Application Level Gateways (ALG). New IP applications may thus require further ALGs to be developed and deployed. ALGs need to rewrite in-band information in packet payloads that duplicate or rely on the header address or port information.

NAT affects security when using IPSec for the same reason.

NAT also breaks explicit IP fragmentation since only the first fragment of a packet possess information identifying the protocol and the source and destination port used by applications, whilst the remaining fragments are not unique. This makes tracking of multiple simultaneous connections from the same host complex, but it can be achieved by the gateway tracking host/port/fragment IDs, for example.

NAT also violates the end-end integrity check, since it modifies the transport layer checksums within the network itself; the modification risk is however perhaps less than that of a PEP (IETF RFC 3135 [53]).

5.2 IP routing

This clause gives an overview of IP routing requirements. General considerations of the impact on and from satellites are included.

A more in-depth discussion of satellite-related issues for dynamic routing is reserved for clause 6.3.

5.2.1 Routing and forwarding processes

Whilst "routing" is generally used to describe the whole process of IP packet processing, technically a clear distinction is made between the terms "routing" and "forwarding" in IP networking (IETF RFC 1812 [54]).

Routing is more strictly associated with the procedure of determining network layer reachability, identifying the most suitable link on which to send packets. Typically a routing protocol together with local policy is used for this purpose. The result is a Routing Information Base (RIB) which contains the routes for each destination. This information could be used directly to process packets, but is generally compiled to a more readily used format (the Forwarding Information Base, FIB). Routing is a background process which creates and regularly updates a RIB via signalling messages exchanged between neighbouring IP nodes. The algorithm which is followed to calculate the best outgoing interface and next node IP address to reach a remote node can vary according to the Routing Protocol (see below).

Forwarding is the operation of moving packets received on in-bound links to out-bound links. It comprises removing layer 2 protocol information (such as a PPP encapsulation, or Ethernet Frame) of packets received on in-bound links, followed by examination of the layer 3 information in each packet (e.g. addresses, packet type, options). Using a previously constructed table (i.e. the FIB), an appropriate out-bound link is found, and the packets are forwarded to this link where appropriate layer 2 protocol information is added and the packet travels to the next router or a final end host.

The FIB contains associations between destination IP addresses and next hop IP addresses. In addition, an address resolution cache lists link-layer addresses of nodes for next hop IP addresses having an interface on a link. If a link-layer address is not found in this table, an address resolution process is called to create such an entry through a signalling exchange on the link.

5.2.2 IP routing processes

In general there are two main categories of IP routing: Static and Dynamic. Static routing is used wherever possible (particularly in small local networks) due to its simplicity, but dynamic routing is usually needed in complex networks including the Internet.

Static routing involves manual configuration of routes for a list of destinations and the next hops to reach those destinations. It is suitable for small number of destinations or stub networks but, due to its inflexibility, cannot handle node failures or network changes. The "default gateway" address used for access by a terminal on a LAN to destinations not found on the local network is a simple example of static routing. Static routing should be supported in the BSMS.

Dynamic routing involves continuous calculation of routing based on information supplied by other routers in the network, and automatically adapts to network changes.

Since IP is connectionless, each packet is routed to its destination by each node in the IP network according to the address contained within the packet.

Most Internet strategies treat routing of packets with only "best-effort" (without guarantee or QoS), and each router computes one or more next-hops for optimal paths to each destination, based on a set of static metrics (e.g. cost, bandwidth, delay, load, error rate), and packets may be routed independently of others within the stream to or from an end-host. For two-way communications, routing in the forward direction is also performed independently of the return direction (the two paths can be different).

Increasingly more deterministic routing is being introduced known as "Constraint-based" routing with requirements that are based on administration "policy" or are service-oriented (QoS Routing) and may depend on any other intermediate path priorities taken by the network at any time (e.g. congestion or failures).

QoS Routing uses additional routing criteria based on resource availability and requested QoS, where they are available. It is suitable for emerging applications such as IP telephony or video-on-demand which require constant delay and guaranteed bandwidth. This will include QoS provisioning using DiffServ, to supply different QoS treatments to different IP flows, as well as RSVP and IntServ extensions for reserving resources.

In addition "Content-based" routing is emerging, which is suited to Web-caching schemes, and which uses the content of a given information request e.g. and HTTP URL to determine routing. In this way it is intended to achieve load balancing in the network.

The main routing mechanisms are described further below. Whilst these are normally associated with Best Effort (BE) routing, extensions to these algorithms in some cases can also allow QoS routing.

5.2.3 Static routing

An example of the support of static routing by the BSMS is of the routing for an enterprise's regional office network (Extranet or Intranet - see scenario in clause 6.2.2.1), which is described in this clause.

Static routing is used wherever possible (particularly in small local networks) due to its simplicity, and due to the high cost and complexity of dynamic routing.

The regional office subnet is likely to be very simple; the route table of the ST supporting the regional office subnet will likewise be simple. It may contain static route(s) pointing to the elements of the regional office subnet. The NCC will provide the ISP customer with the ability to configure the static route entry. When the ST supporting the regional office subnet receives an IP packet from the satellite link, it will look up a matching route entry which will contain the destination IP subnet, a network mask, the Satellite Next Hop Address, and a cost metric. The ST will then use address resolution to find the link layer address for the Satellite Next Hop.

Figure 5.2.1 shows a telecommuter PC supported by an ST. If the telecommuter frequently interfaces with the regional office, then a practical way of providing routing for the telecommuter ST is static routing. The destination address of this ST matches the regional office subnet and its Satellite NH Address contains the network address of the regional office ST. The NCC should provide the ISP customer with the ability to configure these static route entries which will contain the destination IP subnet, a subnet mask, the Satellite NH Address, and cost metrics. A static Address Resolution (AR) entry is configured for the satellite interface of the telecommuter ST, which resolves the Satellite Next Hop Address of the static route. The telecommuter PC will probably have the ST that supports it as the default router. When packets arrive at the terrestrial interface of the ST, a routing table lookup is done. If the packet is destined for the regional office network, the configured static route is selected.

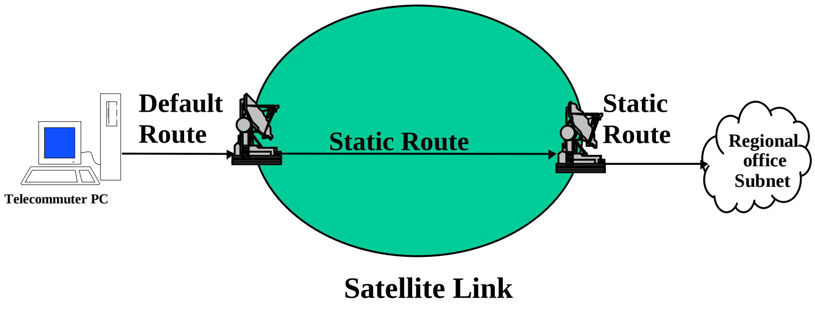


Figure 5.2.1: Static routing scenario

The default route, specified by 0.0.0.0 in the IPv4 network address field of the route entry, specifies the next hop IP layer address when an explicit match is not found in the route table for the IP datagram's destination address. An NSP customer may use this to specify a default router; for example the Access Gateway. Both default and static route table entries will include the Satellite Next Hop Address for the satellite interface.

5.2.4 Dynamic routing

5.2.4.1 Routing algorithms

Routing algorithms calculate optimal (i.e. lowest cost, often the shortest) routes through the network. for destinations, and form the basis for various routing protocols. The current algorithms are:

1) Distance Vector (DV) routers compute the best path from information passed from neighbours by adding distance vectors from router to router. They pass copies of routing tables to neighbour routers. The frequent updates result in slow convergence.

2) Link State (LS) routers each possess a copy of the entire network map and compute best routes from this local map. They rely on updates triggered by events, resulting in faster convergence. They pass link-state routing updates to other routers.

Convergence occurs when all routers have the same routing information. Lack of convergence occurs after there is a change of status of a router or link, causing lost packets and network failure.

5.2.4.2 Routing protocols

Unicast routing protocols are divided into two types:

1) Interior Gateway Protocols (IGP, for routing internal to an AS); and

2) Exterior Gateway Protocols (EGP, for routing external to an AS).

Both IGP and BGP announce to each other the network links under their aegis in order to allow routing decisions to be made. Table 5.2.1 summarizes the current status of routing protocols.

Table 5.2.1: Status of routing protocols

Name	Algorithm type	Running over which protocol	Standard/source	Notes
IGP Protocols				
RIPv1 and v2	DV	UDP	IETF RFC 1056 [55], IETF RFC 2453 [56]	Simple, for small AS, use on broadcast LANs, slow convergence, need full routing table transmissions
OSPFv1 and v2	LS	IP	IETF RFC 1583 [57], IETF RFC 2328 [58]	Newer than RIP, overcomes its limitations. Widely used. Supports CIDR
IS-IS	LS	IP	OSI	OSI version of OSPF
IGRP-EIGRP	DV	IP	CISCO proprietary	Cisco's version of RIP with enhancements
EGP Protocols				
EGP	DV		IETF RFC 904 [59]	Historical status, superseded by BGP
IDPR	LS and DV	TCP	IETF RFC 1478 [60]	Use of TCP requires large windows for BSM
BGPv4	DV and LS	TCP	IETF RFC 1771	Widely deployed, use of TCP requires large windows for BSM. Supports CIDR and route aggregation
IDRP	DV	TCP (optional)	OSI	New development for IPv4 and v6; use of TCP requires large windows for BSM

The most important impact of these protocols on satellites arises from traffic for route advertisements. Since in a wireless environment such as MBMS, the bandwidth is limited advertising routing updates to the rest of the Internet is undesirable, as is the MBMS system "learning" about the rest of the Internet.

Conclusion 5	In a BSMS context, the routing protocols or schemes should incur minimal signalling overhead

The impact of these protocols on the BSMS is discussed further in clause 6.3.

5.2.4.2.1 Use of cost metrics

The network "cost" of links in a routing domain is identified in routing tables and is typically measured in number of "hops" between routers. Routers use this *metric* to determine the least cost path to a destination.

Within a single satellite BSMS mesh network, the cost metric of links between ST routers will be 1, whilst in a star network it will be 2 if the Hub is a router. Routes to destinations outside the BSMS will depend on the total cost of the link, which can be used to choose the nearest gateway stations.

The network operator may also set an artificial metric for a link in order to force traffic along certain paths.

5.2.4.2.2 Interior Gateway Protocols

The growth in networking over the past few years has pushed the previous IGPs such as RIP, which use distance-vector algorithms, past their limits. The primary alternative to distance-vector schemes is a class of protocols known as *Link State, Shortest Path First,* e.g. OSPF.

The important features of these routing protocols are:

- A set of physical networks is divided into a number of areas.
- All routers within an area have an identical database.
- Each router's database describes the complete topology (which routers are connected to which networks) of the routing domain. The topology of an area is represented with a database called a *Link State Database* describing all of the links that each of the routers in the area has.
- Each router uses its database to derive the set of optimum paths to all destinations from which it builds its routing table. The algorithm used to determine the optimum paths is called a *Shortest Path First (SPF)* algorithm.

In general, a link state protocol works as follows. Each router periodically sends out a description of its connections (the state of its links) to its neighbours (routers are neighbours if they are connected to the same network). This description, called a *Link State Advertisement (LSA)*, includes the configured "cost" of the connection.

The LSA is flooded throughout the router's domain. Each router in the domain maintains an identical synchronized copy of a database composed of this link state information. This database describes both the topology of the router's domain and routes to networks outside of the domain such as routes to networks in other Autonomous Systems. Each router runs an algorithm on its topological database resulting in a shortest-path tree. This shortest-path tree contains the shortest path to every router and network the gateway can reach. From the shortest-path tree, the cost to the destination and the next hop to forward a datagram to is used to build the router's routing table.

Link-state protocols, in comparison with distance-vector protocols, send out updates when there is news, and may send out regular updates as a way of ensuring neighbour routers that a connection is still active. More importantly, the information exchanged is the state of a router's links, not the contents of the routing table. This means that link-state algorithms use fewer BSMS network resources than their distance-vector counterparts, particularly when the routing is complex or the Autonomous System is large. They are, however, compute-intensive. In return, users get faster response to network events, faster route convergence, and access to more advanced network services.

5.2.4.2.3 OSPF and the Hello protocol

OSPFs primary means of verifying continuing operation of the network is via its Hello Protocol. Every OSPF speaker sends small hello packets out each of its interfaces every ten seconds. It is through receipt of these packets that OSPF neighbours initially learn of each other's existence. Hello packets are not forwarded or recorded in the OSPF database, but if none are received from a particular neighbour for forty seconds, that neighbour is marked down. LSAs are then generated marking links through a down router as down.

The Hello timer values can be configured, but they must be consistent across all routers on a network segment.

Link state advertisements also age. The originating router re-advertises an LSA after it has remained unchanged for thirty minutes. If an LSA ages to more than an hour, it is flushed from the databases. These timer values are called architectural constants by IETF RFC 1583 [57].

OSPFs various timers interact as follows:

- If a link goes down for twenty seconds, then comes back up, OSPF does not notice.
- If a link flaps constantly, but at least one of every four Hello packets make it across, OSPF does not notice.
- If a link goes down for anywhere from a minute to half an hour, OSPF floods an LSA when it goes down, and another LSA when it comes back up.
- If a link stays down for more than half an hour, LSAs originated by remote routers (that have become unreachable) begin to age out. When the link comes back up, all these LSAs will be re-flooded.
- If a link is down for more than an hour, any LSAs originated by remote routers will have aged out and been flushed. When the link comes back up, it will be if it were brand new.

Conclusion 6	The Hello protocol timers are set not only over the BSM but are shared on all the routers attached to the common network. Hence the operator of a network with an OSPF IGP should set these parameters to ensure the BSMS fully participates in routing.

5.2.4.3 QoS Routing

QoS routing will become important in the future for new applications.

Assumption 3	Although algorithms and protocols for QoS routing are still the subject of research and standardization, the BSMS should ensure interworking will be possible in order to guarantee application QoS.

Under QoS routing (IETF RFC 2386 [61]), a flow requests a specific QoS and is admitted only if the respective QoS can be guaranteed. Paths for flows are selected dynamically, based on matching the demanded QoS for flows with resource availability at network nodes.

Two main categories can be distinguished [11]:

1) Best Path: Global network state information is gathered and routing is based on a global view.

2) Proportional: Local information only is used to select a few candidate paths.

Proportional routing uses only infrequently exchanged information and thus minimal overhead, yet has been shown to achieve good routing performance. This solution would therefore be better suited to satellite implementation.

QoS routing protocols are based on QoS extensions to OSPF routing (IETF RFC 2676 [62]).

5.2.5 Multicast routing

Multicast routing is a primary concern in the terrestrial network where the generally point-to-point nature of the core network, which is ill-suited to multicasting, requires packet replication and carefully choice of routes and tree structure. However in satellites it is important to avoid unnecessary forwarding of multicast traffic to STs that do not request it. For this reason "sparse mode" protocols are preferred, such as PIM-SM. It is also important to minimize the signalling overhead associated with multicast routing, and adaptation of internal satellite protocols may be advantageous.

Within a satellite, physical channels are by nature broadcast within a beam coverage. Multicast routing then concerns selection of routes corresponding to beams or appropriate channels within beams.

If the STs do not contain or attach to local routers but are attached to local IP end hosts only and to a multicast router via a the satellite, then the selection of paths and channels in a satellite over which multicast packets are forwarded is an issue of multicast group membership (via IGMP protocols) rather than of routing.

For further details of multicast routing refer to [5].

5.2.6 Address Resolution

A node (e.g. ST) in the BSMS needs to determine the link layer address corresponding to the "next hop" IP address in order to forward IP datagrams.

In the BSMS network an IP layer must be interfaced to the lower layers (e.g. based on ATM, MPLS, DVB-S/-RCS, "label-based" solutions) by an "interworking" function performing some kind of "address resolution" (AR).

AR is performed only for destination nodes which are indicated as being on a local link by the routing table and for which the sender does not know the corresponding link address. Each node maintains a neighbour AR cache. The AR cache may be filled dynamically using a protocol associated with the router interface; as described below. Static AR entries may also be configured under certain conditions, according to the specific system.

The desirable characteristics of address resolution are:

- Simplicity: avoid numerous static configurations or complex signalling protocol (as in an IP/ATM models).
- Native support for broadcast and multicast, which are expected to be the main support for satellite applications.
- logical partitioning of resources to allow VPNs and sharing of satellite capacity among several Providers.
- Implementation on transparent or regenerative satellites systems.
- Adapted to both star and mesh topologies.

In IPv4, address resolution protocols are defined for existing link layers by, for example:

a) Address Resolution Protocol/Reverse ARP (ARP/RARP, IETF RFC 826 [24]).

b) NBMA ARP (NARP, IETF RFC 1735 [89]).

c) Next Hop Resolution Protocol (NHRP, IETF RFCs 2332 [32], 2333 [33] and 2735 [34]).

These protocols can be applied to satellite links, or satellite-specific protocols can be used (see clause 7.4.2.2).

ARP is intended for Ethernet or any other network that uses a broadcast link layer (e.g. over satellites, at least on the forward link from a Hub Station to STs). Before using IP, the host sends a broadcast ARP request containing the Internet address of the desired destination system. Before using IP, the host sends a broadcast ARP request containing the Internet address of the desired destination system.

NARP is intended is intended for non-broadcast links (e.g. ATM) and for when a conventional address resolution protocol, such as ARP, may not be sufficient to resolve the NBMA (Non-Broadcast, Multi-Access link layer) address of the destination terminal, since it only applies to terminals belonging to the same IP subnetwork, whereas an NBMA network can consist of multiple Logically Independent IP Subnets (LISs - autonomously managed ATM networks).

NHRP is intended to reduce or eliminate the extra router hops required by the LIS model, and can be deployed in a non-interfering manner with existing ARP services.

IPv6 uses Neighbour Discovery (IETF RFC 2461, see clause 5.4.2.2) instead of NHRP etc.

5.2.6.1 AR at satellite interface for customer networks

Because the BSMS is a shared resource where many subnets all use a single communication link, the IP address assigned to the satellite interface of the ST is required to be unique in order to resolve the BSMS MAC address. Satellite-wide uniqueness may be required in some systems because the NCC is used to perform address resolution.

The problem of uniqueness of addressing can be solved in several ways as described in clause 7.

5.3 Relationship of BSMS with Autonomous Systems

The Internet is divided, for network administration purposes, into many separate Autonomous Systems (AS), such as ISPs and corporate networks. Splitting the Internet into ASs allows cohabitation of groups of networks using different routing strategies. An AS is a connected group of one or more IP prefixes run by one or more network operators which has a single and clearly defined routing policy (IETF RFC 1930).

An AS uses one or more interior gateway protocols and common metrics to route packets within the AS.

Depending on how an AS deals with transit traffic, it may be placed into one of the following categories (IETF RFC 1772 [63]):

1) Stub AS: an AS that has only a single connection to another AS, and which only carries local traffic. Such a network requires very simple routing tables which simply direct all packets with destination address outside the AS towards the stub BG.

2) Multi-homed AS: an AS that has connections to more than one other AS, but does not carry transit traffic. Such a topology requires setting up routes via the border gateways, or defining policies to say which border gateway will be used for packets that need to be sent outside the AS.

3) Transit AS: an AS that has connections to more than one other AS, and is designed (under certain policy restrictions) to carry both transit and local traffic.

When IP packets cross an AS boundary they must travel between connected border gateways (BGs) which provide the interfaces between ASs. BGs typically implement the Border Gateway Protocol version 4 (BGP-4) based on CIDR. BGs communicate via routing protocols to exchange routing information. External peering to BGs in adjacent ASs uses External BGP, whilst Internal BGP is used between BGs within the same AS.

Assumption 4	The BSM network will link IP networks belonging either to the same AS, or to different ASs managed by different ISPs. It may also share both of these roles simultaneously. (Furthermore, if satellite on-board routing is considered, the borders of ASs may lie within the BSMS).

Scenarios for the interaction of the BSMS with ASs are shown below, which are applicable to the Use Scenarios described in clause 6.

5.3.1 BSMS as a demarcation zone

The BSMS does not have to participate in the routing protocol exchanges directly if it is used as a "demarcation zone", a mesh of semi-permanent links connecting BGs of ASs. The BSMS can provide layer 2 connections (or even IP/IP tunnels) in this case, as shown in figure 5.3.1, where the BGs are located outside the BSM, at the edge of each AS. For BGP these links then have to be meshed bidirectional links between all interconnected BGs.

Routing protocol traffic between ASs is still carried over the satellite links but is done so transparently.

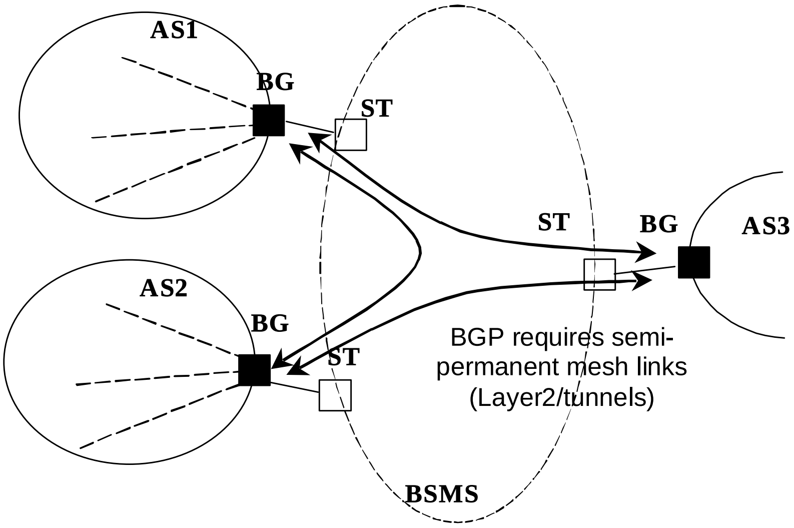


Figure 5.3.1: BSMS as "demarcation zone" for Autonomous Systems

As the number of ASs connected become large and associated STs join and leave the BSMS, management of these types of semi-permanent links becomes inefficient and onerous, unless dynamic connection establishment is used.

5.3.2 BSMS as an Autonomous System

At the other extreme to use as a demarcation zone, when the BSMS is used for interconnection of ASs as a transit network and it implements full IP layer (dynamic) routing, the whole BSMS could be considered as an independent AS with its own BGs, as shown in figure 5.3.2.

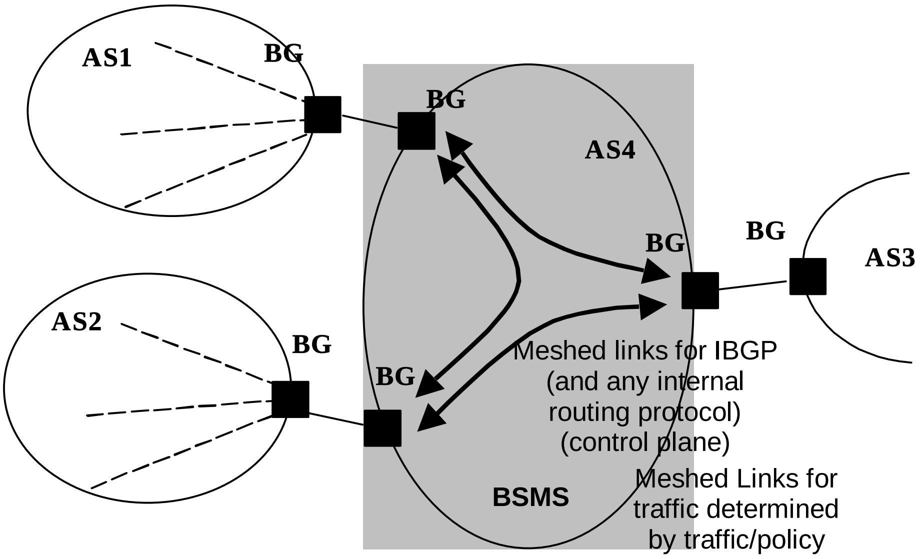

AS1
BG
BG
AS4
BG
BG
AS3
AS2
BG
BG
Meshed links for IBGP (and any internal routing protocol) (control plane)
BSMS
Meshed Links for traffic determined by traffic/policy

Figure 5.3.2: Example of BSMS as an independent AS

The BSMS must then be a repository of routing information about all neighbours reachable through the BSM network, exchanged using the Exterior Border Gateway Protocol (EBGP). All connected ASs must flood the BSMS AS with routing information because it must be aware of all possible destinations. This architectural choice increases the size of the BG routing tables in proportion to the number of networks that the BGP advertises. An important issue to be investigated is routing policy for reducing the size of such routing tables by limiting the number of networks advertised.

Conclusion 7	If the STs are configured as independent EBGP speakers, then they should run the Interior Border Gateway Protocol (IBGP) between them with fully meshed connections to keep routing information updated inside the AS.

This implies more complex STs (combined with routers to perform the BGP peering and internal routing) and BSMS interior routing protocols, but the BSMS can participate in routing policy across the system, and can optimize its interior routing protocols. For example the BSMS can use policy and/or metrics to determine the optimum (lowest cost) ST to reach a destination in the terrestrial network.

5.3.3 BSMS within an Autonomous System

There are several options for interaction of a BSMS within an AS. Within an AS, OSPF (IETF RFC 2178 [64]), for example, allows collections of contiguous networks, hosts and routers to be grouped together into "areas". The topology of an area is invisible from the outside of the area. Conversely, routers internal to a given area know nothing of the detailed topology external to the area. This isolation of knowledge enables the protocol to effect a marked reduction in routing traffic.

Routing in the Autonomous System takes place on two levels, depending on whether the source and destination of a packet reside in the same area (intra-area routing is used) or different areas (inter-area routing is used). In intra-area routing, the packet is routed solely on information obtained within the area; no routing information obtained from outside the area can be used. This protects intra-area routing from the injection of bad routing information.

When the AS is split into OSPF areas, the routers are further divided according to function into the following four overlapping categories:

1) Internal routers

   A router with all directly connected networks belonging to the same area. These routers run a single copy of the basic routing algorithm.

2) Area Border Routers (ABRs)

   A router that attaches to multiple areas. Area border routers run multiple copies of the basic algorithm, one copy for each attached area. Area border routers condense the topological information of their attached areas for distribution to the backbone. The backbone in turn distributes the information to the other areas.

3) Backbone routers

   A router that has an interface to the backbone area. This includes all routers that interface to more than one area (i.e. area border routers). However, backbone routers do not have to be area border routers. Routers with all interfaces connecting to the backbone area are supported.

4) AS Boundary Routers (ASBRs or Border Gateways)

   A router that exchanges routing information with routers belonging to other Autonomous Systems. Such a router advertises AS external routing information throughout the Autonomous System. The paths to each AS boundary router are known by every router in the AS. This classification is completely independent of the previous classifications: AS boundary routers may be internal or area border routers, and may or may not participate in the backbone.

   All routers run same algorithm, in parallel. Each router constructs a tree of shortest paths with itself as root. This shortest-path tree gives the route to each destination in the Autonomous System. Externally derived routing information appears on the tree as leaves.

The BSMS can provide semi-permanent layer 2 links or tunnels within the AS, for example as a backbone, without participating in routing.

When routing is included within the BSMS, it provides links within an Area (intra-area routing) or between Areas (inter-area routing) (or both).

Intra-area routing is determined only by the area's own topology. That is, the packet is routed solely on information obtained within area; no routing information obtained outside the area can be used.

5.3.3.1 Inter-area routing

Inter-area routing is always done via the backbone.

The OSPF backbone is the special OSPF Area 0 (often written as Area 0.0.0.0, since OSPF Area IDs are typically formatted as IP addresses). The OSPF backbone always contains all area border routers. The backbone is responsible for distributing routing information between non-backbone areas. The backbone must be contiguous, and connectivity can be established/maintained through virtual links.

An example for the BSMS is as follows.

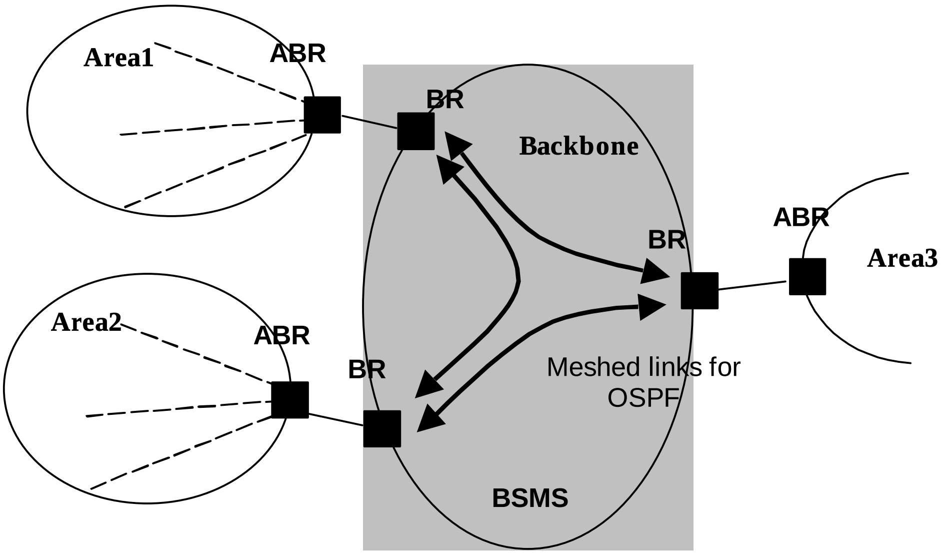


Figure 5.3.3: Example of BSMS as a backbone in inter-area routing

The broadcast capability of satellites is useful in OSPF since neighbouring routers are discovered dynamically on these nets using OSPF's Hello Protocol, which takes advantage of broadcast. The OSPF protocol makes further use of multicast capabilities since Link State Advertisements (LSAs) are flooded throughout the routing domain. The collected link state advertisements of all routers and networks form the protocol's link state database.

5.3.3.2 Intra-area routing

It is the task of the ABR to advertise into Area 1 the distances to all destinations external to the area. AS- external-LSAs are flooded throughout the entire AS, and in particular throughout Area 1. These LSAs are included in Area 1s database. The ABR must also summarize Area 1s topology for distribution to the backbone.

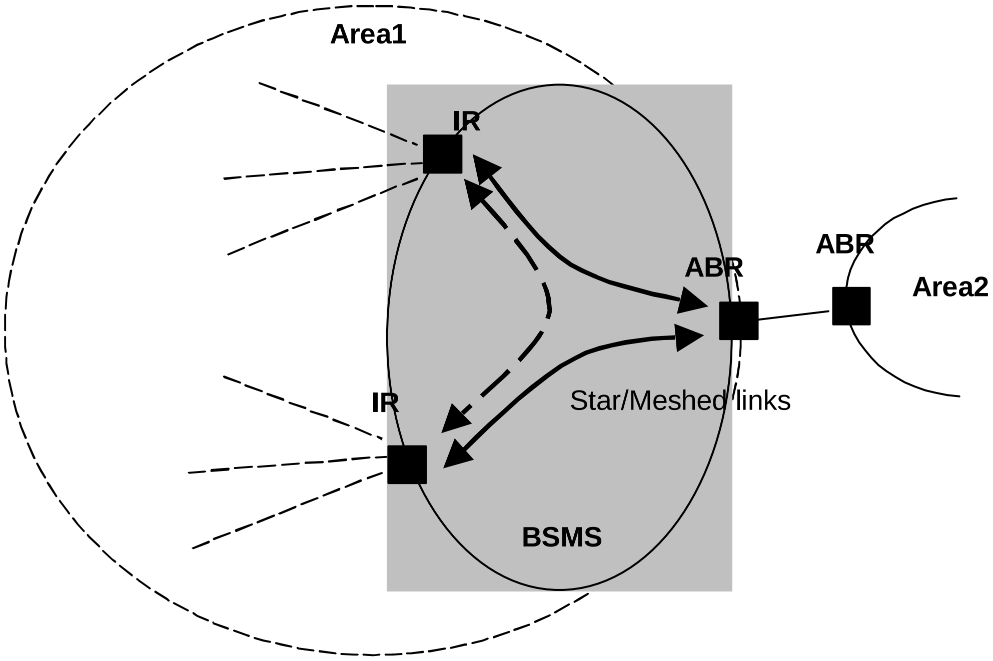


Figure 5.3.4: Intra-area routing example with ABR in the ST

5.4 IPv6 issues

The relevance of IPv6 for satellites is primarily that:

1) Different routing protocols, and particularly Neighbour Discovery, are introduced with which satellites have to be compatible.

2) Transition from IPv4 to IPv6 is complex and introduces a range of network scenarios.

5.4.1 Introduction

The IPv6 specification introduces major modifications to IPv4. Not only is the IP address length extended to 128 bits but also the IP header format and the way header information is processed have been modified. Moving from IPv4 to IPv6 is not straightforward and mechanisms to enable coexistence of and transition between the two versions have to be introduced.

This vastly extended address space is intended to allow one or more global IP addresses to be allocated to host interfaces everywhere, thus simplifying address management and avoiding many associated IPv4 functions.

The set of protocols included and associated with IPv6 are also changed in many ways compared to IPv4, for example those dealing with address resolution and routing, and also mobility.

Whilst the IPv6 protocol itself does not give rise to problems as experiments have shown, and migration from IPv4 could in principle be swift, the continuing need to interwork with IPv4 applications and networks is the main issue.

No general rule can be applied to the IPv4 to IPv6 transition process. In some cases, moving directly to IPv6 will be the answer. For instance IPv6 could be pushed by a political decision to extend the number of IP addresses to sustain the economic growth of a country. Another example is the large-scale deployment of a new IP architecture (such as mobile or home networking) to provide disruptive applications and innovative services.

Disruptive technology [7], [8], [9] is the name given to the new applications that the current incumbents in the networking business have not foreseen, rather than simply upgrades of existing equipment. IPv4 was originally a disruptive technology. Disruptive technologies tend not to be as efficient as mature technologies at the functions for which the latter were designed. However, they create entirely new functions and hence new opportunities.

Some studies foresee that the IPv4 to IPv6 transition period will last till the years 2030-2040, (see [9]).

It should be noted that:

- IPv4-to-IPv6 transition is not always a viable solution. Some disruptive applications will need IPv6 for mass deployment. Deploying transition mechanisms on a large scale can also lead to scalability issues that could heavily limit the IPv6 performance compared to a native solution.

- When IPv4 and IPv6 have to coexist, keeping transition under control is essential to avoid the cost of two parallel Internet infrastructures.

- Transition is not only an issue of addressing or routing. Available and emerging enhanced IPv4 services such as IP QoS, IP security, telephony over IP have to be continuously provided whatever the IP infrastructure might be.

The most important differences from IPv4 are described below, followed by a considerations of the role of the BSMS in the transition form IPv4 to IPv6.

5.4.2 Addressing and address management

IPv6 has a hierarchical addressing structure (IETF RFC 2373 [65]), similarly to IPv4 addresses under CIDR. The address space of IPv6 is intended to simplify some network protocols and to be large enough to obviate private addressing, NAT etc.

IPv6 unicast addresses are aggregatable in order to simplify routing tables, with contiguous bit-wise masks assigned to hierarchical levels as follows:

1) Lowest level: site address.

2) Upper level: "IP Connectivity Services" Provider.

3) Highest level: large Services Providers such as Intercontinental ISPs.

The Connectivity Services Providers address may be structured in different levels, as a provider may sell connectivity while being itself a client of other providers.

The main types of IPv6 address are:

1) Unicast: for communicating to a single interface.

2) Anycast: for one of a set of interfaces (belonging to different nodes).

3) Multicast: for all of a set of interfaces.

There are several forms of unicast address assignment in IPv6, including the global aggregatable global unicast address, the NSAP address, the IPX hierarchical address, the site-local address, the link-local address, and the IPv4-capable host address. Of particular note amongst the latter are "IPv4-compatible IPv6 addresses"; special IPv6 unicast addresses for IPv6 nodes that utilize dynamic tunnelling of IPv6 packets over IPv4. These are assigned the high-order 96-bit prefix 0:0:0:0:0:0 and an IPv4 address in the low-order 32-bits (or "::IP4").

A second type of IPv6 address which holds an embedded IPv4 address, the "IPv4-mapped IPv6 address", is also defined to represent the addresses of IPv4-only nodes (those that do not support IPv6) as IPv6 addresses. This has the format ::ffff:IPv4.

5.4.2.1 Address acquisition

An IPv6-compatible node needs an IP address of global scope if it wants to communicate with the rest of the Internet. This address acquisition procedure is performed at the start of an IP host/terminal session, in case an interface attaches to a new link or if a previous IP address is invalid due to a limited lifetime.

Address acquisition includes creating a link-local address, verifying its uniqueness on the link, and determining what information should be auto-configured (addresses, other information, both).

Addresses can be acquired using either the stateless part of Neighbour Discovery (IETF RFC 2462 [66]), see clause 5.4.2.2, or by stateful mechanisms. ND is used when an operator is not concerned about the exact addresses configured.

The "stateful" Dynamic Host Configuration Protocol (Dynamic Host Configuration Protocol for IPv6 (DHCPv6) [38]) allows for tighter control over configuration parameters for nodes via DHCP servers. It offers the capability of automatic allocation of reusable network addresses and additional configuration flexibility. DHCPv6 can be used separately or concurrently with ND to obtain configuration parameters.

Each IPv6 address has an associated lifetime (possibly infinite) to indicate how long addresses are bound to the interface. On expiration the binding becomes invalid and the address can be assigned to another interface. Interfaces with deprecated IP addresses must use the above procedure to acquire a new IP address.

In summary the design goals of IPv6 address acquisition are:

1) Manual configuration of individual machines should not be needed. Consequently, a mechanism for obtaining unique address for each of the interfaces is needed. This mechanism assumes that an interface can provide a unique (at least within the link) interface identifier.

2) Presence of a stateful server or even a router should not be needed. Plug-and-play communication with other nodes on the link is achieved by using link-local addresses.

3) Even on larger sites a stateful server should not be needed for autoconfiguring site-local or global addresses.

4) Configuration should facilitate graceful renumbering of nodes, for example when changing network providers. This is achieved by leasing of addresses to interfaces and assigning of multiple addresses to the same interface.

5) Administrators need the ability to specify which autoconfiguration system, stateless or stateful, is used.

5.4.2.2 Neighbour Discovery

The Neighbour Discovery functions (IETF RFC 2461) form an important part of the ICMPv6 suite of functions which are intimately involved in IPv6 networking.

It is crucial the BSMS is compatible with the ND protocols if it is to operate in an IPv6 world, since ND plays such a central role in IPv6 in terms of routing, topology discovery, etc.

ND is used by an IPv6 node to collect connectivity data on neighbouring IP nodes at one-hop away. ND performs the functions of a number of separate protocols in IPv4, such as ARP (IETF RFC 826 [24]) and Router Discovery (IETF RFC 1256 [67]). In particular, ND specifies how to perform address resolution in an IPv6 link.

Using ND, IPv6 nodes can:

- detect each other's presence on a link;
- determine each other's link-layer addresses (Address Resolution) and purge out-of-date address resolution entries;
- find routers on a link;
- detect duplicate IP addresses;
- maintain reachability information about the paths to active neighbours and determine the best first hop node in the link for a certain destination node;
- acquire other useful parameters on the link such as subnet prefixes to calculate an IPv6 address dynamically through IPv6 Stateless Address Autoconfiguration (IETF RFC 2462 [66]).

ND performs address resolution similarly to IPv4 nodes with ARP. There are however, some differences such as:

- ND messages are ICMPv6 packets carried inside IP datagrams and are common to any link-layer technology that IPv6 has to operate on, while ARP can vary depending on the link-layer technology in IPv4 links.
- An IPv6 host acquires link-layer addresses of routers on the link through periodic **Router Advertisement** messages, which contain this information, without necessarily exchanging signalling. However, hosts can prompt **Router Advertisement** messages to obtain information on the link quickly.

Similarly to ARP, when an IP datagram has to be sent to a node on the link, but its link-layer address is not available, the sending node multicasts a **Neighbour Solicitation message** that asks the target node to reply with its link-layer address. The target node returns its link-layer address in a unicast **Neighbour Advertisement message.** A single exchange of messages is sufficient for both the initiator and the responder to cache each other's link-layer addresses. Unsolicited Neighbour Advertisement could be used to have neighbouring nodes on a link cache a link-layer address of a new node connecting to the link on its start-up phase. The new node would multicast this message containing its IP address besides its link-layer address, without an explicit request, so as to be immediately reachable by all nodes on the link.

5.4.3 Routing Protocols

For IPv6, adaptations of IPv4 protocols are defined: RIPng (IETF RFC 2080 [68]) or OSPF (IETF RFC 2740 [69]).

5.4.4 Transition from IPv4

The IETF NGTrans Working Group is defining strategies in this field. The main schemes are described below.

5.4.4.1 BSM architectures for IPv4 to IPv6 transition

There are various network topologies in which the satellite link might appear, for example within and between IPv4 and IPv6 subnets and backbones, which affect the choice of transition mechanisms. The BSMS should ideally be able to support many or all possible scenarios for interconnection of IPv6 and IPv4 networks. A diagram of the general interconnection scenario is as follows.

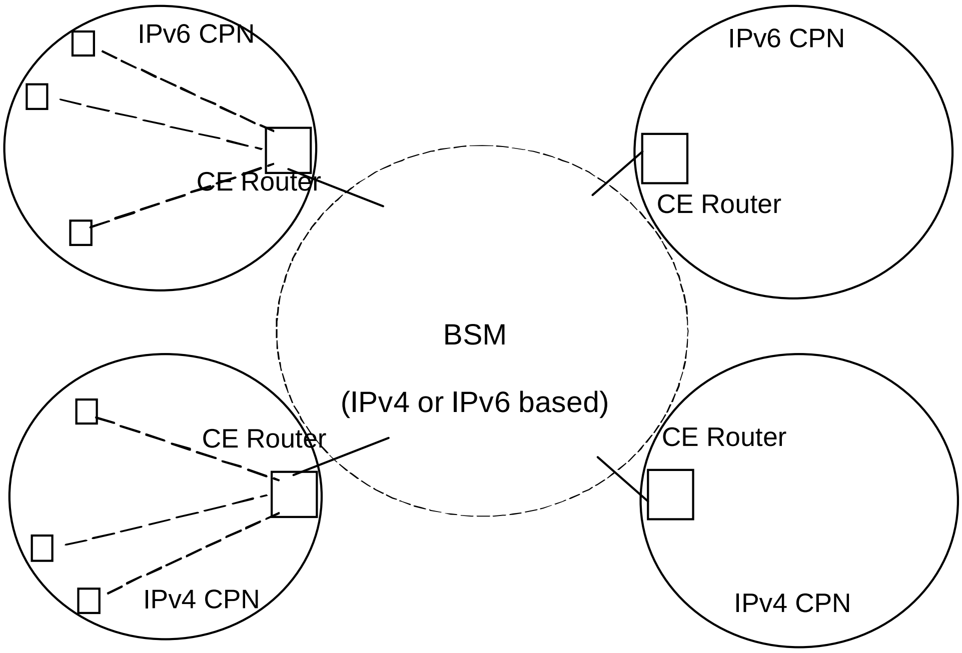


Figure 5.4.1: General IPv4 to IPv6 Interconnection Scenario across the BSM

The Connection of IPv6 Domains via IPv4 Clouds in general has been described in IETF RFC 3056.

New interconnection scenarios include:

- interconnection of IPv6 islands (which is a promising application for satellites when the islands are remote);
- communication and interworking of IPv6 nodes with IPv4 nodes.

The BSMS may be chosen to be either IPv4 or IPv6 based, which has implications on the interworking mechanisms to be included in the BSMS. An IPv4-based BSMS is a natural short-term solution which could be adapted with interworking units at a few specific interfaces to allow for the IPv6 network interconnection scenario. An IPv6-based BSMS is a longer term solution which could be adapted in a similar but inverse way to allow for legacy IPv4 network interconnections.

Assumption 5	The range of situations that need to be considered for BSMS includes all permutations of IPv4 and IPv6 network interconnection through either an IPv4 or IPv6 based BSMS.

Of these, two main aspects may be considered of most immediate importance for strategic and fast adoption of IPv6 (Transition Mechanisms for IPv6 Hosts and Routers - IETF RFC 2893):

1) Interconnection of IPv6 islands through an IPv4 network (BSM and terrestrial). Solutions are generally based on dual stack routers and IPv6 in IPv4 tunnels.

2) Communication and interoperability of IPv6 nodes with IPv4 nodes. Mechanisms rely on dual stack techniques, application level gateways, NAT technology or on temporary allocation of IPv4 address and IPv4 in IPv6 tunnelling.

Both of the above need permanent or temporary allocation of both IPv4 and IPv6 addresses in some network nodes.

Several points should be borne in mind:

- An alternative approach to network level translation or transition is to use dual-stack edge servers at the IPv4/IPv6 border as application level proxies. This could be particularly applicable to the BSMS.

- When IPv4 and IPv6 have to coexist, keeping transition under control is essential to avoid the cost of two parallel Internet infrastructures.

- Transition is not only an issue of addressing or routing. Available and emerging enhanced IPv4 services such as IP QoS, IP security, telephony over IP have to be continuously provided whatever the IP infrastructure might be.

Transition mechanisms proposed by the NGTrans Working Group include engineering tools to build transition strategies.

5.4.4.2 Routing aspects of IPv4 to IPv6 transition

Routing aspects for IPv6 transition scenarios are discussed in IETF RFC 2185, and include:

1) Routing for IPv4 packets (over IPv6).

2) Routing for IPv6 packets (over IPv4):

    - IPv6 packets with IPv6-native addresses;

    - IPv6 packets with IPv4-compatible addresses.

3) Operation of manually configured (static) tunnels.

4) Operation of automatic encapsulation:

    - locating encapsulation;

    - ensuring that routing is consistent with encapsulation.

Three main transition techniques have been defined by the IETF NGTrans working group (IETF RFCs 1933 [70], 2893 and 2767 [71]) as follows:

1) Dual-stack

   This approach requires hosts and routers to implement both IPv4 and IPv6 protocols. At the present time, the dual-stack approach is a popular mechanism for introducing IPv6 in existing IPv4 architectures and will remain widely used in the near future. The drawback is that an IPv4 address must be available for every dual-stack machine.

2) Tunnelling

   Tunnelling (IETF RFCs 2473, 2529 and 3053) enables the interconnection of IP clouds. For instance, separate IPv6 networks can be interconnected through a native IPv4 service by means of a tunnel. IPv6 packets are encapsulated by a border router for transportation across an IPv4 network. Tunnels can be statically or dynamically configured, or implicitly (6to4, 6over4). The TB (Tunnel Broker) approach has been proposed to manage automatically tunnel requests coming from the users and ease the configuration process. ISATAP (Intra- Site Automatic Tunnel Addressing Protocol) is a recent technique to avoid manual tunnel configuration. In later stages of transition, tunnels will also be used to interconnect remaining IPv4 clouds through the IPv6 infrastructure.

3) Protocol Translation

   Translation is necessary when an IPv6-only host has to communicate with an IPv4 host. At least the IP header has to be translated but the translation will be more complex if the application processes IP addresses; in fact such translation inherits most of the problems of IPv4 NATs. ALGs (Application-Level Gateways) are required to translate embedded IP addresses, re-compute checksums, etc. SIIT (Stateless IP/ICMP Translation; IETF RFC 2765) and NAT-PT (Network Address Translation - Protocol Translation; IETF RFC 2766, see clause 5.4.4.3) are the associated translation techniques. A blend of translation and the dual stack model, known as DSTM (Dual Stack Transition Mechanism), allows for the case where insufficient IPv4 addresses are available. Like tunnelling techniques, translation can be implemented in border routers and hosts.

Conclusion 8	For the BSMS, tunnelling is the most inefficient transport mechanism due to the additional encapsulation overhead, and should be avoided on cost grounds in favour of translation.

5.4.4.3 NAT-PT

NAT-PT (IETF RFC 2766) IPv4 to IPv6 protocol translation is essentially a method for communication between IPv6-only and IPv4-only nodes. It resides within an IP router, situated at the boundary between an IPv4 network and an IPv6 network. However NAT-PT is less scalable than other translation methods.

By installing NAT-PT between an IPv6 network and the Internet, all Internet users are given access to the IPv6 network without host modification. Equally, all hosts on the IPv6 network are given access to the Internet with a pool of IPv4 addresses for assignment to IPv6 nodes on a dynamic basis as sessions are initiated across IPv4-IPv6 boundaries.

Its main problem is that, since it does not snoop the payload and is application-unaware, applications which include IP addresses embedded within the IP payload (e.g. SIP/SDP, FTP, DNS) require separate Application Level Gateways (ALG). New IP applications may thus require further ALGs to be developed and deployed to allow an IPv6 node to communicate with a IPv4 node and vice versa.

5.4.4.3.1 Implementation

NAT-PT is relatively simple to deploy as such devices are only necessary at IPv6/IPv4 network boundaries. No client configuration is needed and all NAT-PT translation is totally transparent to the end users.

NAT-PT is an interoperable solution that does not require modifications or extra software (such as dual stacks) to be installed on any of the end user hosts of either IPv4 or IPv6 network. Maintenance is also eased, as any alteration to NAT-PT only needs to be downstreamed to the boundary routers - not to every host that requires contact across an IPv6/IPv4 boundary.

Since NAT-PT is only deployed at network boundaries, administration and maintenance are relatively simple.

The major limitations of NAT-PT are similar to traditional IPv4 NAT devices. In particular end-to-end network layer security is not possible. In addition, translation can only be done on a best effort approach due to the significant differences between the IPv4 and IPv6 headers. Because of these limitations it is always recommended that NAT-PT be used where other mechanisms cannot be used i.e. native IPv6 or IPv6 over IPv4 tunnelling.

5.5 Impacts of mobility on routing requirements

The type of mobility to be considered for a BSMS is limited to "roaming". This includes user, terminal and service mobility between terminal sessions, but avoids the handover and session continuity required for mobility during a session.

This approach fits well with GEO constellations which are the main focus of the present document. Owing to the wide coverage of satellite beams there are rarely changes of routing needed within the satellite even for terminals in motion, compared with terrestrial networks, and the need for users to change frequently their attachment point to alternative gateways is avoided.

Only where non-GEO constellations of satellites with continuously moving coverage are employed will there be a possible need to handle regularly internal BSMS routing changes, and this independently of the mobility of a terminal.

In GEO constellations, roaming can often be handled at the satellite link layer when there is a change of beam access within the same satellite system, without resort to IP layer mobility since the gateway does not need to change.

Only when there is a change of access network between satellites systems or to terrestrial networks is there a change of IP layer attachment. In this case, however, the IP layer procedures required are considered to be adequately described as for terrestrial networks.

The IETF Working Group "IP Routing for Wireless/Mobile Hosts" (mobileip) has addressed these issues (IETF RFCs 2005 [72] and 2002 [73]).

5.5.1 Overview of mobility procedures

In general mobility support can be provided at the network layer, transport layer, or application layer. Network layer approaches are often based on packet forwarding schemes, i.e. a proxy forwards packets destined for the Mobile Node (MN) at the home network to the current location of the MN. This is the case with Mobile IP, which uses a Home Agent (HA) to forward the packets. Transport layer schemes are often based on trying to migrate the ongoing connections when the MN changes its IP address.

The obvious IP level approach for mobility is to use Mobile IP, which has been developed to support mobility through a network solution, and one of the strengths of Mobile IP is that it is transparent to, and serves all, the applications above it. If the mobility solutions were to be implemented at a higher layer, e.g. separately by each application, it might be argued that this would be inefficient. However, for some services, application layer or session mobility (SIP - Session Initiation Protocol) may either partially replace or complement network layer and Mobile IP mobility. For interactive sessions, SIP-based mobility (IETF RFC 2543 [74]) can in principle be used to provide all common forms of mobility, but for terminal mobility, an IPv6-based solution is likely to be preferable, as it applies to all IP-based applications, rather than just real-time applications like Internet telephony and conferencing. In the absence of Home Agents a session level solution implements mobility functions relying on recovery mechanisms at the session layer for fast transport connection reestablishment. No attempts are made at keeping transport level connections alive during network mobility/roaming. The SIP protocol has been chosen by 3GPP as the signalling protocol of choice for Internet multimedia and telephony services, but will not be further described here. There is also a possibility to use SIP over Mobile IP in some applications.

The IETF has standardized IP mobility or "Mobile IP", which provides for transparent mobility, in that it hides the change of IP address when a mobile host changes IP subnets, and users change their point of attachment. The support of mobility management implies modifications to IP protocols. The Mobile IP protocols lack many essential features to solve all the requirements of future all-IP networks. That is the reason why the mobility problem is often divided in two parts: macro-mobility and micro-mobility. Macro-mobility concerns the management of users movements at a large scale, while micro-mobility covers the management of users movement at a local level. The basic model of Mobile IP is not very suitable for handling micro-mobility, i.e. swift handover between subnets. The Mobile IP protocols are beneficial for terminal mobility, (see definition later) excluding micro-mobility.

These procedures include:

- User mobility.
- Terminal mobility.
- Inter-network roaming.
- Mobile IP.
- IPv6 mobility.

The Mobile IPv6 protocols to be described here are enhancements to the standard IPv6 protocol that make it possible for users to be reachable in foreign links.

For the BSM architecture network layer mobility and Mobile IP is the most relevant protocol technology, with the possibility to use the SIP protocols as an alternative in real-time applications. For both protocols only macro-mobility is considered relevant.

Mobile IP allows a Mobile Node (MN) to change its point of attachment to the Internet with minimal service interruption. But Mobile IP in itself does not provide any specific support for mobility across different administrative domains. Sometimes referred to as "triple-A" or just AAA, authentication, authorization, and accounting are fundamental aspects of IP based network management and policy administration. AAA servers will provide the means of administering policy to ensure proper use and management of resources within a mobile and roaming network environment. The principal approach of the IETF in this respect is to integrate authentication during Mobile IP registration with a general Authentication, Authorization and Accounting (AAA) infrastructure based on the IETF Diameter protocol [6].

As Mobility is a wide subject and is outside the main scope of the present document, further discussion is curtailed.

5.6 Unidirectional satellite links

Most current satellite links to user terminals (e.g. those used for digital TV, DVB-S, etc.) are receive-only. For Internet access especially for home use, it is advantageous to use the same or similar low-cost terminals as for broadcasting services. Internet traffic to end-users which consists a majority of flows in the forward link is well-suited to this kind of satellite terminal with the additional of a "thin" terrestrial return link (e.g. telephone modem, ISDN, etc.).

This unidirectional link from the satellite poses problems for the traditional Internet architecture since IP routing protocols have assumed bi-directional links. Multicast routing protocols based on the "reverse shortest path tree" also face such a problem.

For dynamic routing in such links, the method proposed in the UDLR IETF RFC 3077 [37] is based on layer 2 tunnelling which allows emulation of a bi-directional link-layer. GRE (Generic Routing Encapsulation, IETF RFC 2784 [75]) is suggested as the tunnelling mechanism to provide a means for carrying IP, ARP datagrams, and any other layer-3 protocol between nodes. IP packets on the return link are thus encapsulated in the user terminal within another IP packet whose address is instead the address of the bi-directional (terrestrial) IP address of the feeder station, and is reachable via the Internet.

The addresses of the satellite network may be private, whilst the tunnel end points are public.

The DTCP (Dynamic Tunnel Configuration Protocol) is also described to allow receivers to discover the presence of feeds and to maintain a list of operational tunnel end-points in order to forward encapsulated datagrams. Feeds periodically announce their tunnel end-point addresses over the unidirectional link. Receivers listen to these announcements and maintain a list of tunnel end-points.

Note that other types of asymmetrical satellite links can also occur (besides asymmetry in the routing path): see IETF RFC 3449.

6 Use cases/architectures

This clause describes the overall requirements for integration of the BSMS in terms of routing and addressing in IP networks, within the main scenarios of interest.

6.1 Satellite system functions within an IP network

The role of a BSMS in an IP network can take several forms which are described in this clause.

The BSMS may participate in an IP network as a layer 2 subnetwork only, but the main focus of the present document is on IP-aware BSMSs with external IP interfaces in order to offer maximum IP service interworking.

Many types of satellite telecommunication systems of varying complexity have been launched and are being developed. The systems include not only one or more satellites (e.g. in a constellation) but also the ground segment for interfacing to users' and operators' equipment as well as for controlling and managing BSMS operations.

A general functional model of IP interfacing and routing in the BSMS TR 101 984 [3] is as follows.

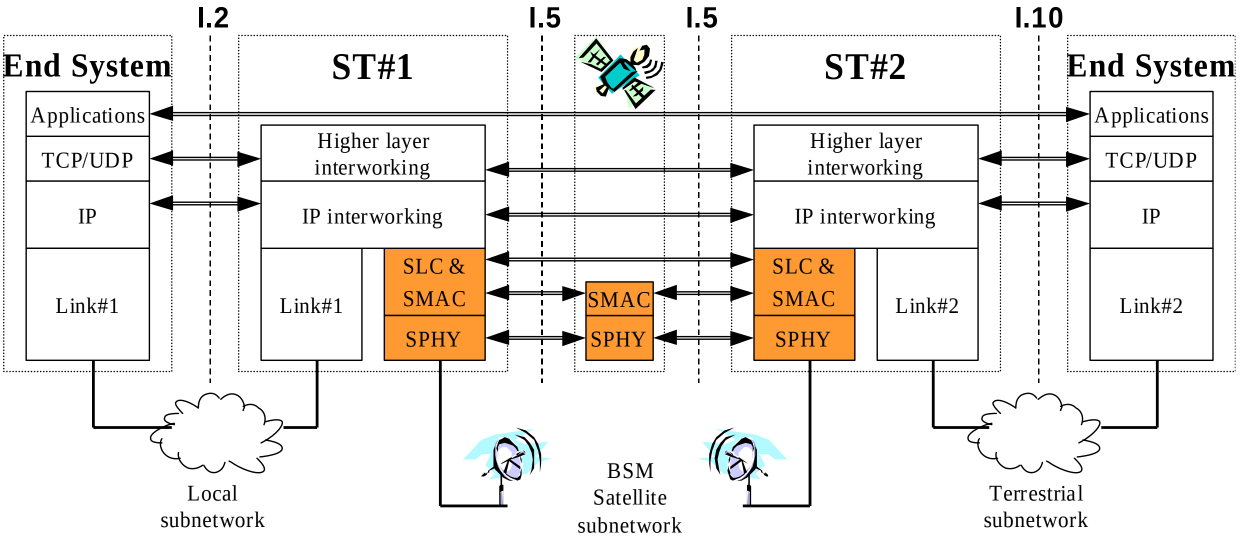


Figure 6.1.1: BSMS General protocol Model

Figure 6.1.1 assumes bi-directional paths through the satellite, increasingly likely to be employed in future systems. Today's systems however are also implemented with the forward (network-to-user) path only carried via satellite and with a terrestrial return path.

Note that the Satellite Terminals (STs) can be divided vertically into two main parts:

1) A satellite interface part with either an access function or gateway function, considered as included within the BSMS.

2) An external (terrestrial) interface part to user equipment or to network operator equipment, and configurable independently of the BSM part.

The IP layers and above (e.g. IP router etc.) in the ST could be considered as belonging to one or the other of these parts, depending on administrative control belonging to either the BSMS operator or the end system user.

The BSMSs main role is to provide IP connectivity between external interfaces. The choice of solution depends on the level of connectivity needed within the system and the complexity of implementation.

6.1.1 Physical connectivity factors

Physical connectivity within the BSMS is often greater than that required at IP level (for example satellite channels are broadcast in nature), and depends on the coverage area. Satellite systems can also provide static or dynamic switching at various layers ranging from layer 1, to layer 3 between ground stations. The IP routing requirements therefore have to be translated in some way, depending on the capabilities of the BSMS and the need for transport efficiency, to the satellite lower layer protocols which determine the internal connectivity.

Satellite systems fall into the following main types which affect their connectivity:

1) GEO systems with national, regional or global coverage. Their mainly stationary antenna beams (e.g. potentially a mix of static or steerable spot beams, regional and global beams) allow stable link configurations to be set up.

2) Non-GEO constellations with quasi-global coverage. Their constant motion requires regular handover of ST links with satellites, associated with greater internal routing complexity.

Within each of these types, the antenna beam configuration plays a major role:

1) Global beam coverage allows easy broadcasting to all STs.

2) Multibeam satellites increase overall capacity, with frequency reuse patterns between beams, for point-to-point links and reduce ST cost. Flexible cross-connection between uplink and downlink traffic is then advantageous, either via time and/or frequency slots (at Intermediate Frequency or Radio Frequency), or via on-board processing of data (e.g. layer 2 switching).

3) Inter-satellite links can provide direct links between satellite coverage areas, avoiding intermediate terrestrial transit and multiple hops.

A further important characteristic of satellites is their capability to process data on-board:

1) Transparent satellites are simpler and independent of layer 1, 2 and 3 protocols; but

2) On-board processing satellites can offer:

   - independence of modulation and coding between uplink and downlink;

   - the potential to switch frames or packets flexibly at layer 2 or even IP;

   - independence of link management (rate adaptation and power control, i.e. to overcome a localized rain fade) and bandwidth control of uplink and downlink.

6.2 Satellite addressing and routing scenarios

6.2.1 Satellite-IP networking scenarios

A BSM network can be used in all parts of the global IP network. It is convenient to divide the global IP network into 3 parts: Core network, Distribution network and Access network as defined in TR 101 984 [3] clause 4.2 and illustrated in figure 6.2.1. The Distribution network is an intermediate IP subnetwork that may be used to connect Access networks to the Core network.

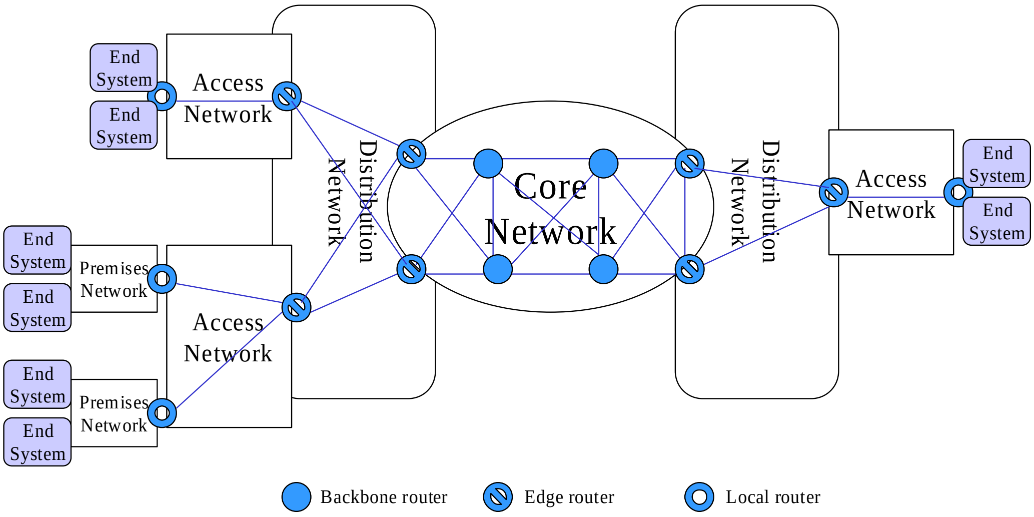


Figure 6.2.1: Core network, distribution network and access network

Each of the networks defined in figure 6.2.1 corresponds to a different domain (i.e. Autonomous System, see clause 5.1.1.4) and these domains are interconnected by edge routers. Different QoS mechanisms may be used within each domain and interworking is through these edge routers. For example, the Core network may use DiffServ, whereas the Distribution network may use IntServ and some Access networks support only Best_Effort.

The types of BSM IP networking scenarios with associated services are summarized in table 6.2.1.

Table 6.2.1: IP networking and service scenarios

	Applicable services	
Access network scenarios	**Point-to-point**	**Multicast**
Corporate intranet	Corporate VSAT network: i.e. site interconnections	Corporate Multicast e.g. Data distribution e.g. Video conferencing
Corporate internet	Internet Access via corporate ISP or via 3rd party ISP	IP multicast RT streaming ISP caching
SME intranet	Small VSAT network	SME multicast
SME internet	Internet Access via 3rd party ISP	IP multicast RT streaming ISP caching
SOHO	Internet Access via ISP Company access via VPN	IP multicast RT streaming ISP caching
Residential	Internet Access via ISP	IP multicast RT streaming ISP caching

Distribution network scenarios	**Point-to-point**	**Multicast**
Content-to-Edge	ISP to Backbone	IP multicast RT streaming Caching at ISP/Edge

Core network scenarios	**Point-to-point**	**Multicast**
ISP interconnect	Trunk interconnect	IP multicast trunking

The above scenarios are described below. The addressing and routing issues related to unicast services in the above access network scenarios are outlined below. Multicast services and scenarios are described in [5].

Assumption 6	Many of the Access, Distribution and Core Network scenarios described below could be implemented simultaneously over the coverage of the BSMS.

Conclusion 9	The BSMS should be able to manage the different and sometimes incompatible addressing and routing requirements for each Access, Distribution and Core Network scenario over its own infrastructure.

6.2.2 Access network scenarios

Access network scenarios include not only access to the Internet but also use of the network as a "backbone" of a Service Provider (SP) to interconnect private networks.

6.2.2.1 Corporate Intranets

"Intranet" is the term used to denote implementation of Internet technologies on a network within a corporate organization, rather than for external connection to the global Internet. An Intranet uses readily available IP technology, such as software and hardware providing Web, e-mail and FTP services, to deliver an organization's information resources to users via the same user-friendly format as the Internet with minimal cost, time and effort.

Components of distributed Intranets include Headquarters, branch offices, Home Offices (SOHOs), and individual roaming Users:

The main features of Intranets are therefore:

- IP-based service provision.
- Private IP addressing (for IPv4) (IETF RFC 1918 [42] and address management (single Autonomous System); no global addressing needed.
- Internal routing protocols (i.e. IGPs) may be used for dynamic routing in large organizations with complex network architectures.
- Manual configuration of Routing Tables (i.e. static routing) is more typical for smaller organizations with simple networks involving few gateways and servers, and where security and authorization is more important.

The Intranet must provide servers for IP configuration and routing for local hosts that would otherwise be done by an ISPs network for Internet access. These include DHCP, DNS, etc. (see clause 5.1).

In IPv6 networks the addressing strategy for corporate networks is intended to be global for simpler overall address management, though still using firewalls.

In addition to a pure Intranet, an Extranet grants controlled access to specific external users, such as customers or trading partners.

6.2.2.1.1 Role of BSMS in Intranets

Use of the BSMS to provide link resources for an Intranet applies to users and subnetworks of the Intranet which are physically remote from other sites. Intranet's imply the provision of Virtual Private Network (VPN) across the BSMS, since the BSMS is considered as a "public" network or more typically as a Service Providers (SPs) "closed" IP backbone. The VPN provides for security, multicast, mobility and QoS support.

The VPN is realised by what is often termed a "VSAT" network, or a private network of small STs over the BSMS.

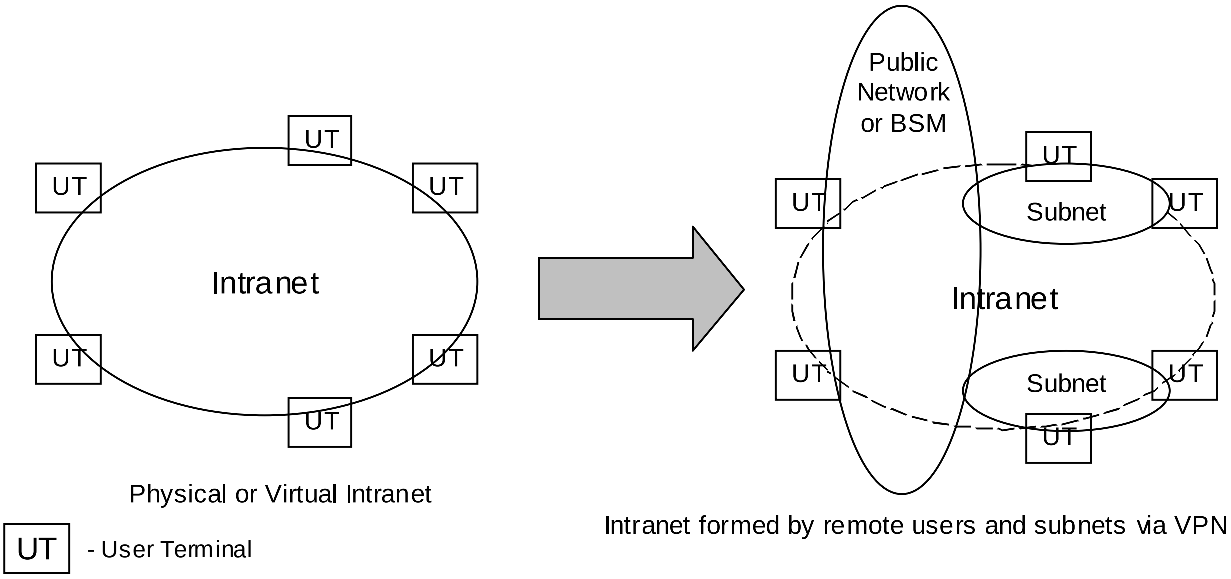


Figure 6.2.2: Intranet Configuration using VPN to connect remote sites

There are many options for VPN architectures. These are described in more detail in clause 6.3.

6.2.2.2 BSMS Intranet IP routing architecture

The configuration of IP nodes in a BSMS used for an Intranet is typically as follows.

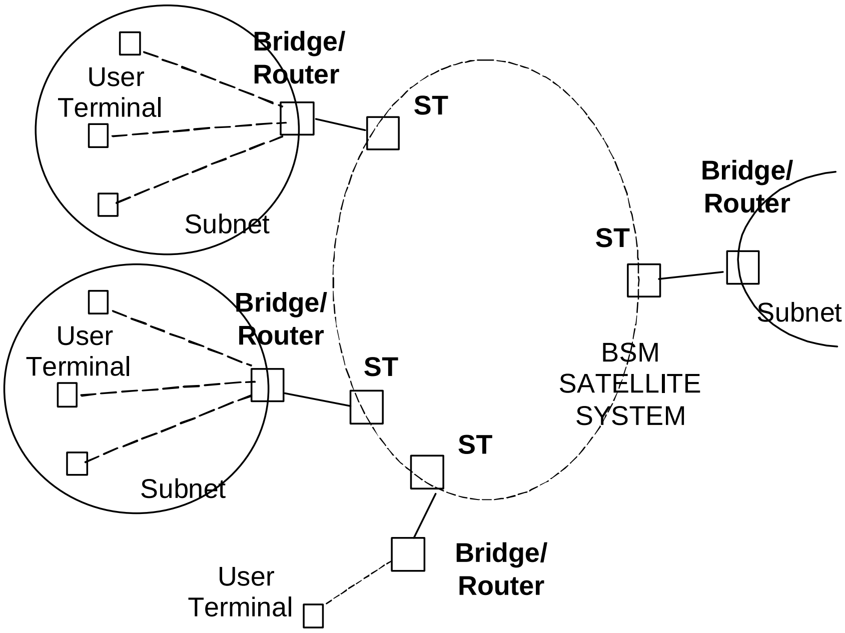


Figure 6.2.3: Configuration of nodes for a BSM-based Intranet

The IP layer connections between sites can be configured in a star configuration or a mesh, or a combination of both, depending on the hierarchy of sites in the Intranet (see clause 6.2.2.1).

The routing and addressing between the subnetworks in this configuration should also take into account their relationships with Autonomous Systems as described in clause 5.3.

The physical layer (of satellite links) may also match the star or full mesh configuration of the IP layer, depending on the traffic and QoS (e.g. delay) requirements between sites, and the system cost. A mesh configuration is suitable where there are many equivalent sites with high speed inter-site traffic. A star network favours an intranet with transactions to and from regional offices centred upon headquarters.

6.2.3 Corporate Internet access

This scenario is an extension of the Intranet/Extranet case covered in clause 6.2.2.2. Main features of this scenario are:

- connection(s) between the corporate network Internet gateway and one or more ISPs via permanent link(s);
- global IP addresses must be allocated to hosts communicating with the Internet e.g. via NAT;
- Proxy server included at the gateway for security (Firewall), NAT, etc.;
- hosts are configured with a "default gateway" address of the Internet gateway/proxy as the next hop when hosts wish to set up a session to an external IP address.

As noted in clause 6.2.2.1, management of addressing in this way relates primarily to current IPv4 practice, based on NAT.

The proxy server acts as an intermediary between a host and the Internet so that the enterprise can ensure security, administrative control, and provide a caching service.

6.2.3.1 Role of BSMS in corporate Internet access

In terms of the role of the BSMS in corporate Internet access, this scenario concerns two cases:

1) The BSMS provides access from remote corporate subnetworks to corporate headquarters which handles Internet access to ISPs.

2) The BSMS provides access from one or more corporate subnetworks to one or more remote third party ISPs.

Case 1) is considered as a similar scenario to an Intranet scenario of clause 6.2.2.1, but in addition one site (the headquarters) manages external Internet connection. The BSMS still forms part of the Intranet and its internal Addressing. The main difference in this case is that global addressing must be used by hosts, or NAT.

Case 2) is illustrated in figure 6.2.4.

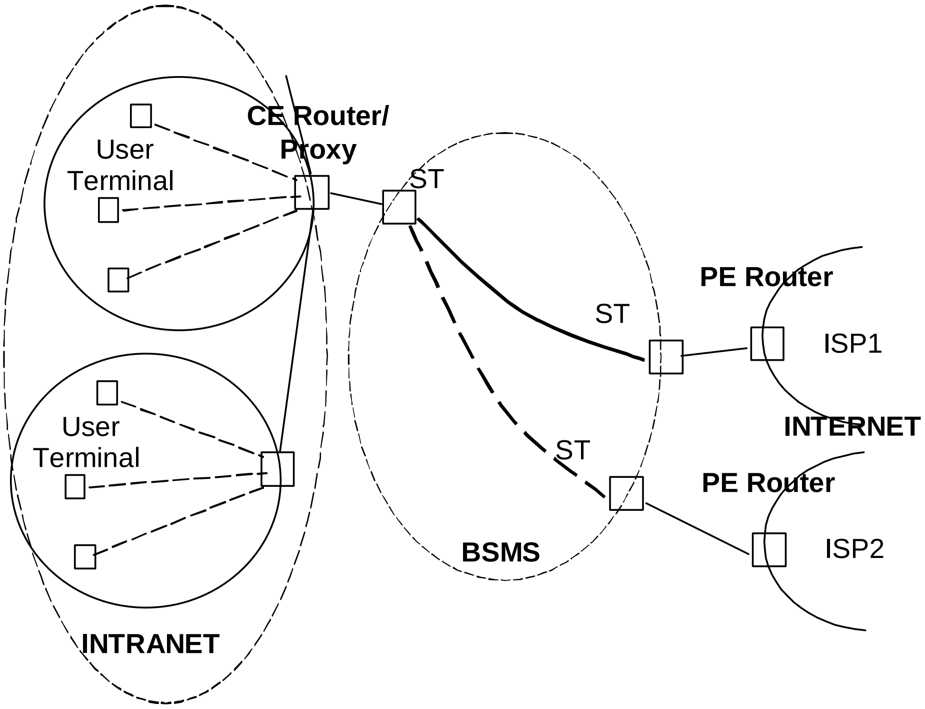


Figure 6.2.4: Configuration of nodes for BSM-based corporate Internet access

6.2.4 SME Intranet and Internet

SME Intranet and Internet Scenarios are considered to be a question of scale, lying between and overlapping with Corporate Networks and SOHO scenarios in terms of complexity of customer's networks. Meshed VSAT networks are the applicable BSMS technology.

SME scenarios are therefore viewed as covered by the descriptions in clauses 6.2.2.1, 6.2.2.2 and 6.2.5.

6.2.5 SOHO

A SOHO is usually associated with a single site based on a small local network of computers, optionally being a remote site of a corporate network to which there is permanent or occasional connection. A SOHO is almost inevitably connected to the outside world via a firewall and concentrator and even a small local (CPE - edge) router. Static internal routing and private local addressing is typically configured. A SOHO may also participate permanently or for short periods in a Corporate network.

Connection to the Internet or to the Corporate network would typically be via permanent (but also dial-up for the smallest examples) access via an ISPs POP.

With satellite access, permanent connection is most likely, which would also allow the SOHO to route to destinations other then the default gateway (i.e. multi-homing). When using permanent Internet access, this scenario is very similar to the Corporate Internet access scenario above, with the exception that the ISP would manage the SOHO's global IP addressing.

SOHO Intranet access is identical to the Corporate Intranet case described in clause 6.2.2.1.

6.2.5.1 SOHO Internet Access

Several scenarios are possible:

- the SOHO is statically attached to only one Service Provider (ISP or Corporate). In this case, the ISP/Corporate HQ is directly connected to the BSMS gateway.
- The SOHO can attach dynamically to several ISPs:
  - The ISPs each have their own BSMS gateways.
  - The ISPs are reached through the same BSMS gateway. In this case the gateway and the ISPs are connected through an IAP (Internet Access Provider). The user is required to choose the ISP at login via a function provided by the IAP. Also the ISPs can outsource some services like Authentication, Address allocation, Accounting, etc to the IAP.

In terms of connections, there are two main methods between the user equipment and the ISP:

- IP over PPP: two methods exists:
  - The CPE Router terminates Users' PPP sessions and routes the IP packets to the ST which in turn routes them to the Gateway and then to the right ISP.
  - The CPE Router does not terminate the Users' PPP session but instead aggregates the PPP (or PPPoE) sessions towards the same ISP over an L2TP tunnel terminated at the ISP Remote Access Server (RAS).
- Native IP: the user IP flows are directly mapped onto the data link layer.

These scenarios are shown in figure 6.2.5.

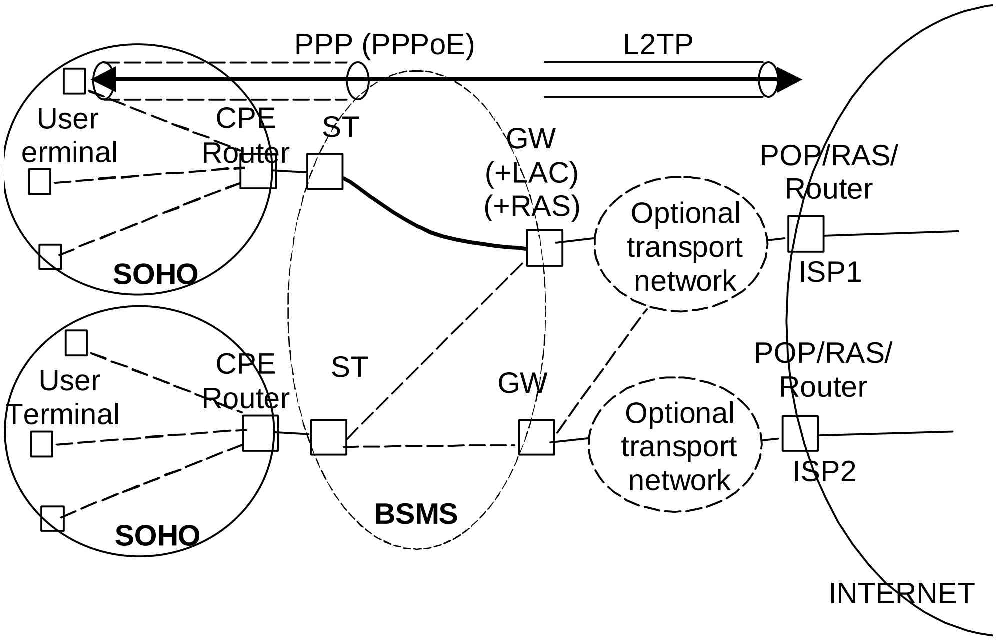
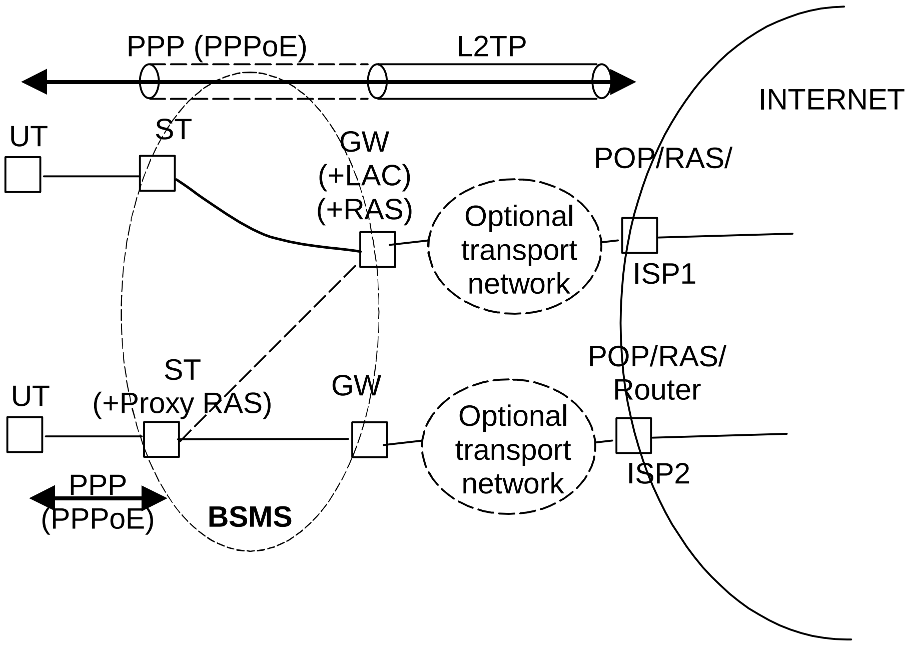


Figure 6.2.5: Configuration of BSMS-based SOHO access to the Internet

6.2.6 Residential

Connection to the Internet by residential users is provided firstly via a connection to one of the POPs of the ISP, or a choice between several, ISPs who then provide access to the Internet. User connection to the ISP is currently typically by "dial-up" terrestrial lines but there is an increasing trend towards "always-on" connections (e.g. via ADSL or cable networks). The connection is often based on a PPP (IETF RFC 1661 [76]) or a PPPoE (IETF RFC 2516 [77]) or PPPoA (IETF RFC 2364 [78]) link to the ISPs Remote Access Server which terminates the PPP protocol and performs all access control functions. PPP is also used as a tunnel protocol (IETF RFCs 1332 [79] and 1552 [80]) to carry any traffic e.g. IP, IPX etc. PPP also includes authentication and security functions. In future PPP may be replaced by native IP procedures using DHCP for address assignment (Dynamic Host Configuration Protocol) and possibly AAA (authentication, authorization, and accounting).



NOTE: Dashed lines in the figure represent optional items (e.g. an ST can connect to alternative Gateways).

Figure 6.2.6: Example of residential access scenario

L2TP (IETF RFC 2661 [81]) provides a means for tunnelling PPP over IP, especially where there is a terrestrial transport network between the gateway and the ISP. For high-speed broadband access to the home, L2TP is also used for aggregation (at the L2TP Access Concentrator) and delivery of PPP connections over packet networks. Proper integration with PPPoE, PPPoA and other tunnelling methods as they "hand-off" to the L2TP portion of the network must be ensured.

L2TP provides:

- An extensible control protocol for dynamic set-up, maintenance, and tear-down of multiple layer 2 tunnels between two logical endpoints.
- An encapsulation method for tunnelling PPP frames between each endpoint. This includes multiplexing of multiple, discrete, PPP streams between each endpoint.

6.2.6.1 Role of BSMS in residential Internet access

For BSMS residential connection services, the ISP may provide dial-up connections, but in a broadband context BSMS users are more likely to be permanently connected. Depending on the service relationship between the BSMS operator and the ISP, the BSMS may have a separate RAS for access control, or the BSMS operator may also be an ISP with a single RAS.

When there is a PPP, PPPoE (or equivalent) tunnel to the ISP no IP routing of connections is involved in this link. A future option would be to provide authentication via native IP (e.g. via a proxy in the ST). This would allow routing to take place in the BSMS to a choice of gateways and ISPs, as shown in figure 6.2.6.

Global IP addresses may be allocated dynamically to a host by the ISP, or may be permanent particularly in the case of permanent connection, depending on the service contract with the ISP.

6.2.7 Distribution and core network scenarios

Compared to access networks where satellite links typically provide several Mbit/s per ST, Distribution and Core network scenarios will typically include fewer physical layer links but of higher, more uniform, bit rates (tens to hundreds of Mbit/s) and relatively fixed, statically configured, connectivity (point-to-point or multicast). Also the size and complexity of equipment such as ground stations and routers involved in such scenarios is also generally different from access networks.

The needs of these scenarios have been traditionally fulfilled by transparent satellites. To traverse even greater parts of the globe than is possible with a single satellite, routing over inter-satellite links is advantageous to avoid the delay due to intermediate ground relay stations.

6.2.7.1 ISP interconnect

ISP interconnect scenarios in both the Distribution and Core networks are intended for nodes located anywhere from the edge to the core of the Internet. Fully meshed bidirectional connections are needed through the BSMS to interconnect all nodes.

In terms of Internet protocols, these scenarios are distinguished by different Internet domains being connected through the BSMS at Border Gateways as shown in figure 6.2.7. The BSMS itself can provide links based either on layer 2-only or layer 3 [4], but in the case of layer 3 the BSMS would need to be an independent AS with its own Border Gateways.

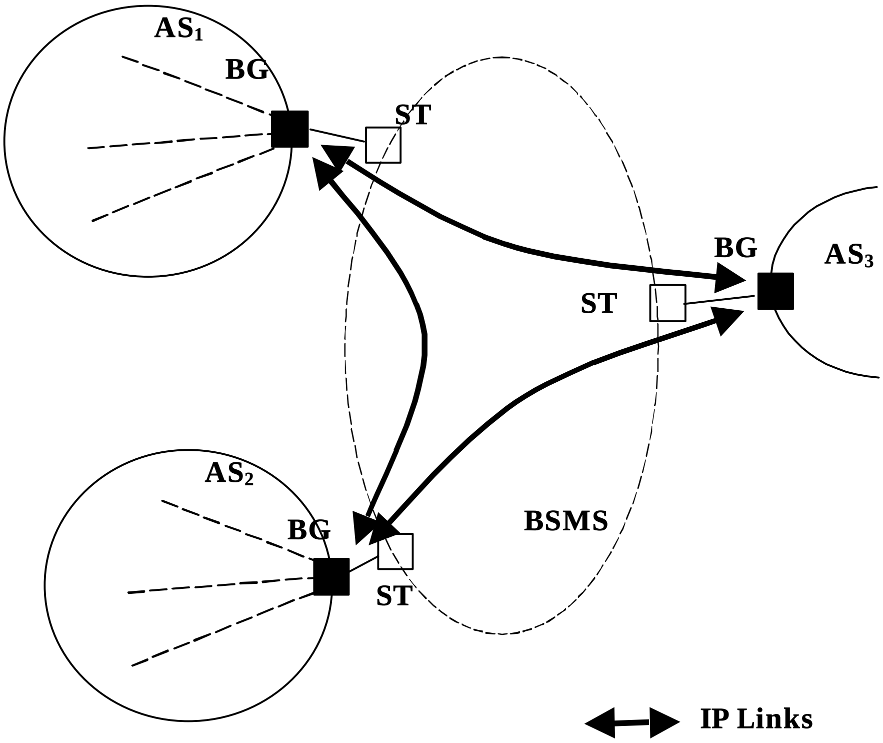


Figure 6.2.7: Core/Distribution scenario and relationship with Autonomous Systems (example with layer 2-only BSMS)

Applicable IP routing and addressing issues are described further in clause 6.3.

6.2.7.2 Content-to-Edge

In the Content-to-Edge scenario for distribution networks, intensively accessed and regularly updated data (e.g. entertainment media, Web caching), is distributed to Edge proxy servers for further localized distribution, to avoid congestion in the core.

A typical service is background updating of caches, rather than interworking with real-time multicast IP services in the Internet. Static routing and/or unidirectional links would be envisaged in this case. A star network from each source would be a suitable configuration.

For more dynamic scenarios involving real-time services, IP multicast services domains would be needed to different Internet Domains via Border Gateways at the BSMS terminals, similarly to the ISP interconnection scenario. The Border Gateways in this case would need to include multicast extensions of routing protocols. Bidirectional links would be preferred to handle dynamic multicast protocols.

Multicasting scenarios are described more fully in [5].

6.3 Virtual Private Networks (VPNs)

The term VPN is defined as the interconnection of remote terminals or subnets through a shared network of another operator in such a way that user services are the same as for a local private network. This implies that the inherent privacy of a private network must be ensured within a VPN.

As public networks or backbones may support several VPNs at the same time, privacy mechanisms are key to provision of VPNs. VPN privacy may encompass many mechanisms such as:

- Dedicated layer 2 links.
- Layer 3 Tunnelling.
- Partitioned routing.
- Data encryption.
- User authentication.

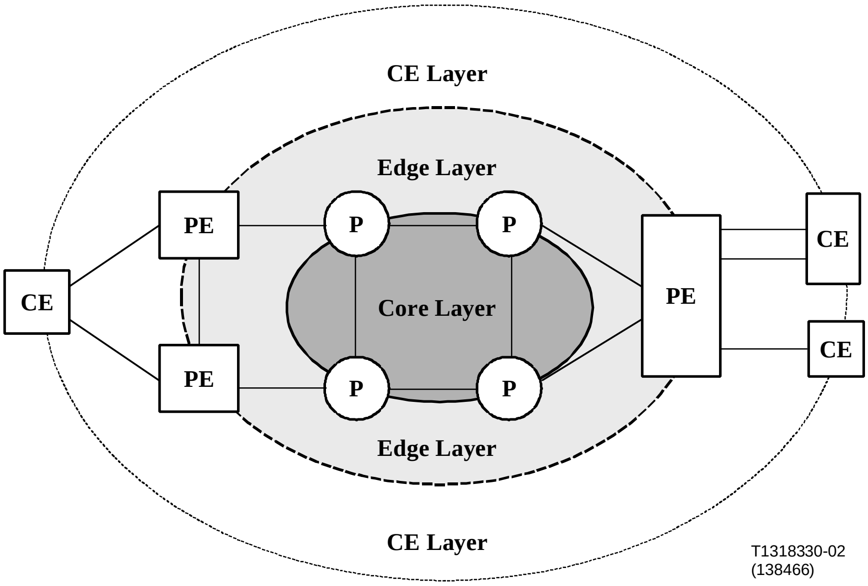
6.3.1 VPN link protocols

The BSMS links forming the VPN of an Intranet do not have to rely on IP layer protocols. As in terrestrial networking. VPN links can be based on different OSI protocol layers from layer 1 to 3.

At layer 1 these links may consist of fixed bandwidth satellite links between all intranet sites, or more flexible TDMA (e.g. DAMA-based) with, in some cases, a configurable matrix of inter-beam links in the satellite.

At link layer more efficient sharing of link capacity can be provided for packet data which may be multiplexed over links to the satellite. Packets may then be forwarded (e.g. in an on-board processing satellite) to the appropriate downlink. Such link layer-based protocols include ATM, MPLS, DVB-RCS.

The main interest in the present document is in IP-based VPNs, in which the SP network provides VPN-specific functions and forwards packets based on layer 3 information. The IETF PPVPN (Provider Provisioned VPN) working group is currently studying these issues (Guidelines of Applicability Statements for PPVPNs [39]). See also ITU-T Recommendation Y.1311 [87].



CE	Customer Edge
PE	Provider Edge
P	Provider

Figure 6.3.1: ITU-T RecommendationY.1311 VPN reference model and designation of network elements

The key motivator for service providers to adopt network-based (or "PE-based") VPN services rather than CE-based services is that they gain considerable capital and operational cost savings. This is done by using one network VPN gateway to provide service for multiple customer sites in an area, and the service provider does have to install or repair CPEs at customers' sites, and has the ability to manage services centrally. Another advantage is the potential offer of additional services, such as classes of service for different priority applications.

For satellite systems these arguments for cost savings over CE-based services may not apply if the ST has to be sited at the customer's premises in each case.

For the BSMS, the choice of VPN approach will affect the functions that need to be included in the satellite network, and notably the ground stations. The implications for the BSMS are outlined below.

For ease of provision of a VPN, by a network provider, it is essential to accommodate addition, deletion, moves and/or changes among sites and members with as little manual intervention as possible. If key VPN network elements can announce their presence to one another through auto-discovery techniques, then the required tunnels can be configured with a minimum of manual intervention. The principle of auto-discovery applies to all types of VPN irrespective of the layer at which the service is offered.

6.3.1.1 Overlapping customer address space

As cited above, Intranets and Extranets often make use of private addressing and this address space is not unique outside of the VPN. When many VPNs share a BSMS, some means of discriminating or hiding private source and destination VPN host addresses in packets delivered to ST routers must be used. A key enabler of VPN provision is therefore the establishment of the "tunnels" which separate the traffic of a given VPN from that of another VPN, and from traffic of the open network across a common infrastructure. These tunnels can use layer 2 addressing, or other techniques as described in clause 6.3.1.2. Similar methods must also apply to IP broadcast packets, and in some cases to IP multicast packets.

6.3.2 IP-based VPN architectures

There are two main approaches to the architecture of IP VPNs, depending on the location and management of the VPN functions:

1) CE-based VPNs, in which the SPs network is used only for transport (e.g. via tunnels) and is not involved in VPN functions which are implemented in the corporate network. The Customer has the task of configuring and maintaining the VPN (e.g. via VPN-enabled routers/firewalls).

2) PE-based (or "network-based") VPNs are a more recent introduction in which the SP offers the VPN functions, together with other IP services such as QoS and firewalls. The burden of VPN management lies with the SP PE router enabling any-to-any connectivity between sites. Security is equivalent to a layer 2 frame relay or ATM VPN service, and IPSec tunnelling/encryption is not typically used.

For CE-based VPNs the IP tunnels between CPEs can use IP/IP, GRE (Generic Routing Encapsulation IETF RFC 2784) or IPSec mechanisms.

PE-based IP VPNs include the following mechanisms:

1) PE-based IPSec.

2) Virtual Router-based.

3) MPLS.

The details of these techniques are described below.

6.3.2.1 CE-based VPNs

1) Tunnelling from the CPN

Tunnel-based IP VPNs allow the encapsulation of an IP packet inside another, between Customer Edge (CE) routers; the lower layer IP packet is routed over an IP network without any particular privacy. Privacy of the upper IP packet is achieved in different ways, for example:

- The user trusts the IP backbone and does not use any encryption mechanism.
- IPSec is used by user terminals to encrypt IP packets.
- PPP PAP/CHAP provides user authentication.
- MPPE: Microsoft Point-to-Point Encryption protocol (IETF RFC 3078).

Tunnelling technologies include GRE, L2TP (layer-2 Tunnelling Protocol, IETF RFCs 2784, 3301, 3355), PPTP (Point-to-point Tunnelling Protocol).

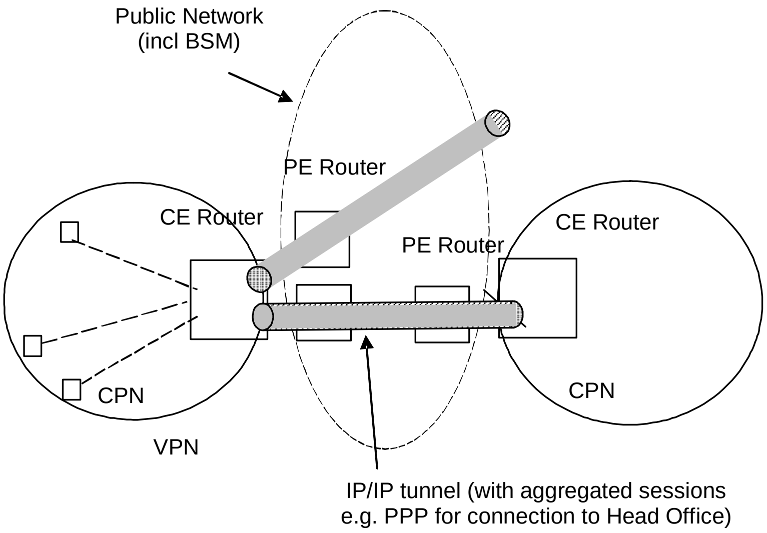

Public Network
(incl BSM)
PE Router
CE Router
PE Router
CE Router
CPN
CPN
VPN
IP/IP tunnel (with aggregated sessions
e.g. PPP for connection to Head Office)

Figure 6.3.2: CE-based VPN using IP/IP tunnelling

2) Use of IPSec between Customer Edge (CE) routers

IPSec-based VPNs rely on IPSec to provide network layer encryption and authentication. IPSec embeds the components for deploying network-wide security across IP networks IETF RFC 1827 [82].

IPSec may work in:

- Tunnel mode: tunnelling between Customer Edge (CE) routers (one encrypted IP layer is encapsulated in another IP packet that is routed in a standard way) together with encryption and key exchange mechanisms.
- Transport mode: adds a specific header between the IP header and the IP Payload. It also uses encryption and key exchange mechanisms.

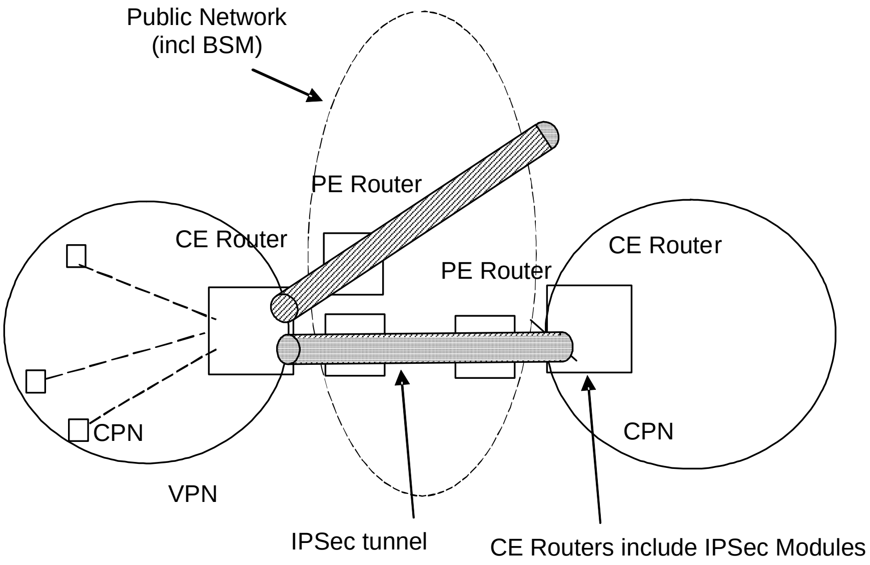


Figure 6.3.3: IPSec-based VPN

6.3.2.2 PE-based VPNs

1) PE-based IPSec

In PE-based IPSec services, the service provider initiates and terminates IPSec tunnelling at network VPN gateways located at edge POPs. IP service switches or edge routers with VPN service capabilities are used as network VPN gateways.

2) Use of Virtual Routers

A Virtual Router is a software "instance" of a router dedicated to each VPN and hosted by a hardware/software platform (typically a Provider Edge (PE) router) that may also support other Virtual Routers. Such a platform is also called a partitioned router or an access router. The Virtual Router's physical and logical partitioning inside an access router provides privacy and enables VPN to use private addressing and overlapping IP addressing planes.

A Virtual Router has the same functions as a standard IP Router: several physical or logical interfaces, a routing table, support of dynamic routing protocols, etc.

The partitioning of a VPNs using VRs also relies on establishing dedicated paths in the backbone between CPNs via tunnelling over IP, or PVCs over ATM, or using MPLS for example, as in the following architecture:

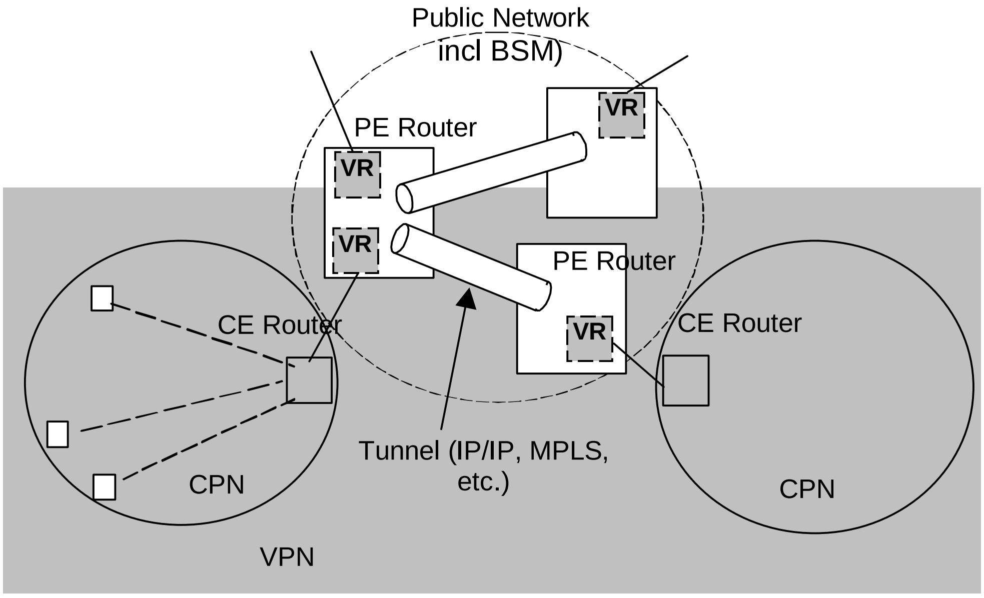


Figure 6.3.4: Virtual-Router based VPN

3) MPLS

MPLS VPNs are based on IETF RFC 2547 [83] which describes the use of MPLS for forwarding packets over the IP backbone and the use of BGP (Border Gateway Protocol) for distributing routes over the backbone.

A "2547" VPN is a private IP network where each remote site has different IP address spaces and has a Customer Edge (CE) router attached to a Provider Edge (PE) router. The route to each of the sites is distributed using the BGP routing protocol. The CE router becomes a peer of the PE router and not a peer to the other CE routers by providing the PE router with route information for the VPN. The PE router stores multiple private routing tables, one for each customer connection (like a Virtual Router).

Inside the public backbone, MPLS Label Switched Paths (LSPs) are created between PE routers by traversing Provider core routers (P routers). This MPLS forwarding role is crucial because the P routers in the core of the VPN provider network do nor need know about the routes connecting the 2547 private network.

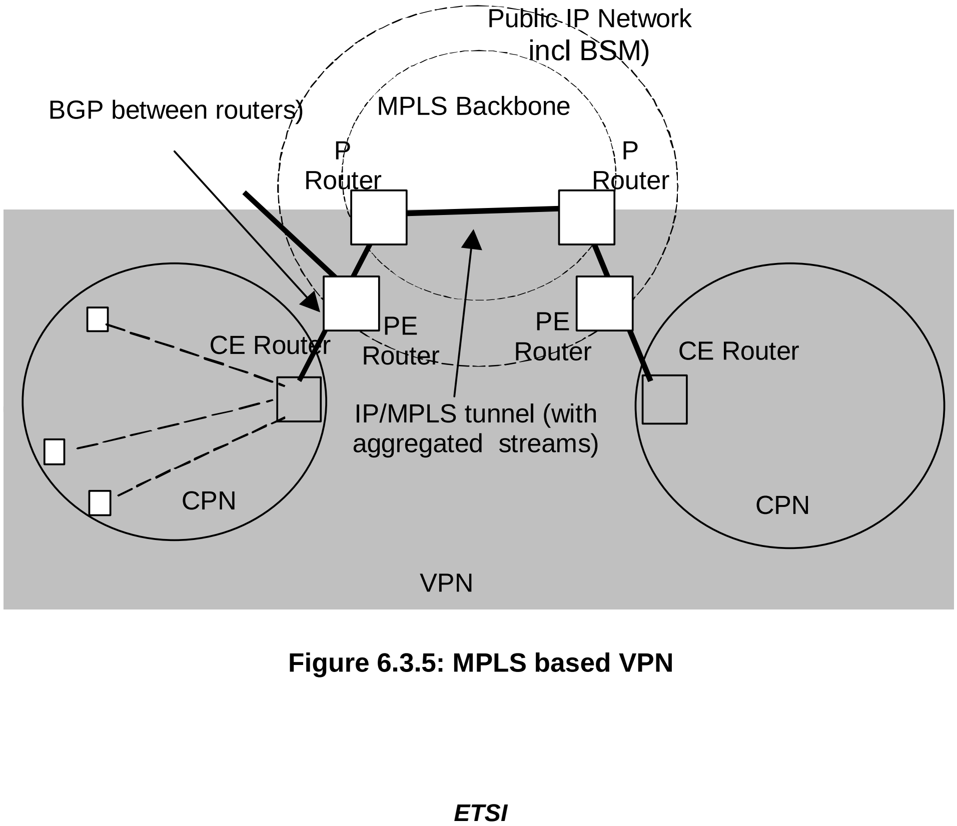


Figure 6.3.5: MPLS based VPN

7 Addressing and routing solutions for BSM systems

7.0 Introduction

This clause focuses on IP addressing routing issues for the BSMS with fixed terminals, focusing on IPv4 and unicast routing. (IPv6 issues are summarized in clause 5.4, multicast in [5]).

Addressing issues concern mainly the means for efficiently allocating IP addresses and resolving layer 2 addresses. A centralized server for these functions is recommended as described in clause 7.3.

IP routing protocols have been designed with terrestrial point-to-point links in mind. The limited bandwidth of satellite implies that internal IP routing schemes be optimized to the specific qualities of satellite links such as broadcast capability, asymmetric links and high delay.

The wide geographic coverage of satellite networks also requires a wide range of interconnection scenarios as outlined in clauses 5.1.1.4 and 6.2 to be accommodated. The coverage also leads to a potentially large number of links between STs, and BSMS IP routing needs to avoid resulting scalability problems.

Clearly, IP routing within the BSMS is only required when at least:

- An IP node in the satellite network has a choice of more than one neighbouring IP node to which to forward the packets (i.e. not simply a "default" gateway).
- The alternative neighbouring nodes are accessible only by different links (e.g. MAC, physical beam paths or channels) rather than via a broadcast channel.

For example, the case of Internet access where all STs connect to the Internet via the same ISP gateway does not require dynamic IP routing in the ST-to-gateway direction, and only static "default gateway" IP addressing. Dial-up access using PPP to a single ISP POP (i.e. from individual users) is also of little interest for routing.

The type of scenarios of most interest for BSMS routing is therefore meshed connection of customers' subnetworks (i.e. corporate networks, SOHOs, ISPs) via permanent (rather than dial-up) links, and connection of these end systems with the Internet. The topology of a BSMS mesh network identified in TR 101 984 [3] is:

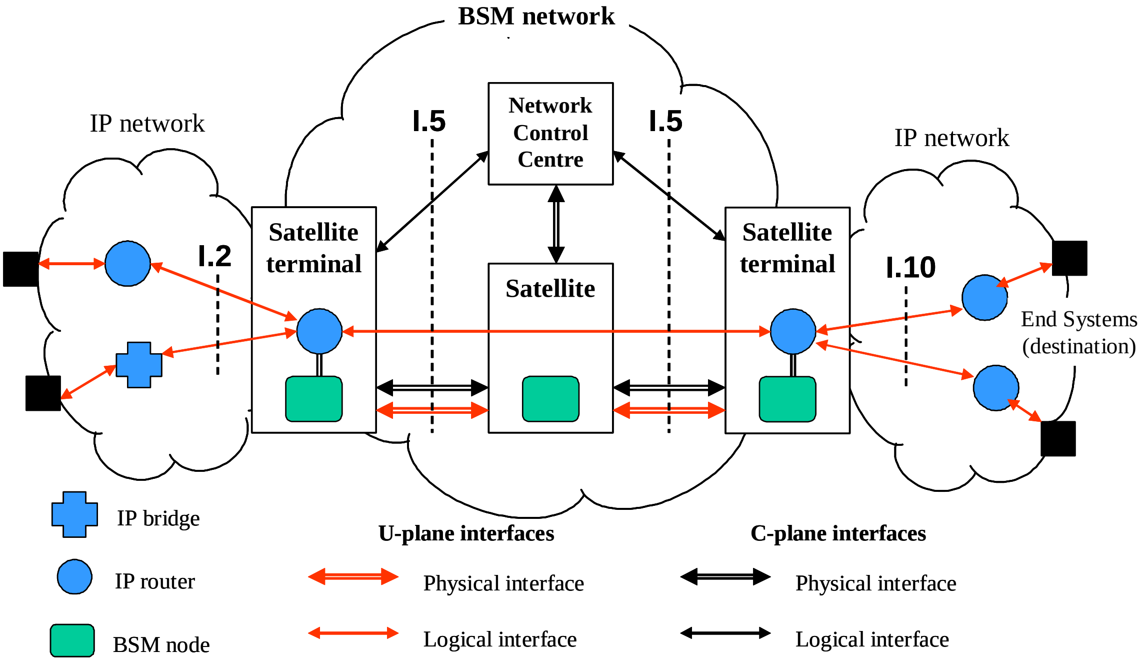


Figure 7.0.1: Example of BSMS mesh topology

The NCC plays an important role in controlling and managing the BSM network. In this role it may be associated with centralized IP addressing and routing functions.

The Control Plane in the above diagram includes the messaging required for addressing and routing protocols. The routing protocols across the BSMS depend on its relationship with the IP domains of the end systems.

For this purpose the main connection scenarios between end systems described in clause 6.2.1 and considered in further detail in this clause are as follows:

Table 7.0.1: Relevant routing scenarios

Connection service scenario	Interconnection w.r.t. Autonomous Systems	Satellite network topology avoiding double hops	Routing protocol across satellite
Internet Access - Single ISP	Single AS	Star	None (Static)
Internet Access - Multihomed	Single AS	Mesh	IGP
Head Office-based Intranet (VPN)	Single AS	Star	IGP (but static default gateway, etc.)
Intranet (VPN), Extranet	Single AS	Mesh	IGP
Independent IP "Island" Interconnect	Multi AS	Mesh	BGP
IP Core Network	Multi AS	Mesh	BGP
IP Edge Network Interconnect	Multi AS	Mesh	BGP
NOTE: The grey shaded lines indicate the scenarios of most interest for BSMS addressing and routing which will be described below.			

Figure 6.2.1 shows the scenarios for interconnection of ASs and the range of routing protocols that a BSMS should be capable of handling across its coverage.

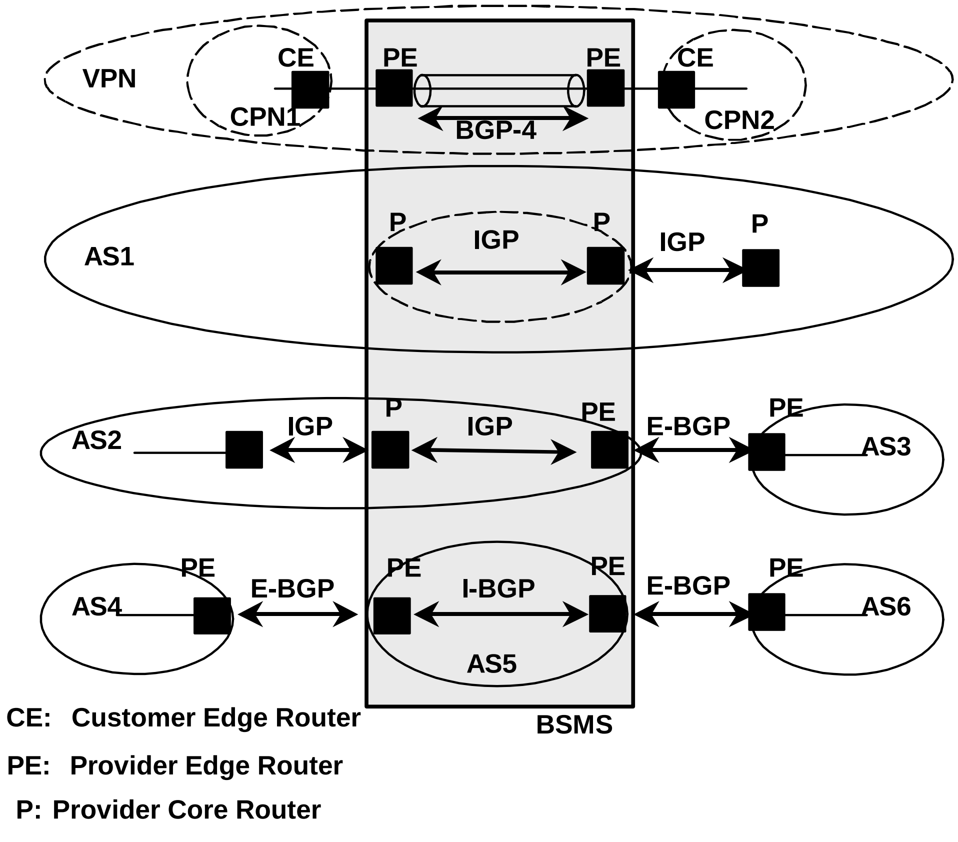


Figure 7.0.2: Routing protocols interfacing to the BSMS

The resulting BSMS routing issues that need to be solved are:

1) How to transport routing protocols with minimum overhead (e.g. build and maintain IP routing tables).

2) What are the most appropriate (internal) protocols to support routing across a BSMS (how to resolve link layer addresses simply).

3) How to handle link layer connections for routing protocols and for IP routes with minimum overhead.

4) Which BSMS architectures offer the most promising solutions for IP Routing?

Solutions to these questions are discussed below.

The issues of routing on the BSMS are summarized in table 7.0.2.

Table 7.0.2

Routing mechanisms	Protocol/ Addressing	BSMS Issue	Comments	Avenue to solution
Peer to peer connectivity - (Dynamic Routing)	IBGP/EGP, IGP (e.g. OSPFv 2)	Overhead, scalability, cost matrix (e.g. OSPF)		Centralized route server
QoS routing	MPLS/other Label-oriented	Label distribution, Overhead, scalability, cost matrix	MPLS network traffic engineering heritage	Special QoS class for signalling
Route/address resolution	NHRP/S-ARP, etc.	Availability of bandwidth Overhead, scalability, cost matrix	No standard	Centralized address and route server

7.1 BSMS configurations

Configurations of mesh-topology BSMSs can be further subdivided as shown in table 7.1.1.

Table 7.1.1

Highest ST protocol layer	Satellite payload type	Satellite payload controller (OBC) - in NCC (control plane)
layer 3	Transparent	Physical layer
	OBP link layer	Link layer
	OBP layer 3 (forwarding) (see note 2)	Layer 3 (routing)
	OBP layer 3 (forwarding and routing) (see note 2)	

NOTE 1: Link layer here means MPLS, ATM, DVB(-RCS), etc. layer 3 implies forwarding in the user plane and routing in the control plane.

NOTE 2: On-board forwarding and/or routing is not considered practicable in the near future, but is included as a technical option.

A summary of the satellite link layer protocols with their IP addressing/mapping relationships is as follows.

Table 7.1.2

IP delivery mechanisms	Address resolution	Link layer connections and signalling	BSMS issue	Comments	Avenue to solution
IP over ATM (or ATM-like)	NHRP	PVCs/SVCs ITU-T Recommendation Q.2931 [88]	Mapping of Uplink-downlink to PVC Segmentation Scalability	ATM over satellite heritage	Centralized signalling NHRP server
IP over Label-based L2:					
1) IP over "Satellite IP-oriented"	S-ARP	Connectionless	Scalability	No standard	S-ARP Server Special QoS class for signalling
2) IP over MPLS	Label distribution with RSVP-TE	CO	Scalability Multicast Refresh rates	MPLS network traffic engineering heritage	Special QoS class for signalling Centralized label server
IP over DVB (-RCS)	MPE	PIDs	Mapping of IP addresses to PIDs Mesh networking Scalability	IP over DVB IETF WG (proposed)	
NOTE: Link layer addressing and IP routing solutions based on the above link layers are discussed in clause 7.4.					

7.2 Solutions for BSMS routing scenarios

Satellites systems providing mesh connectivity (through multiple spot beams) are considered, to allow single-hop links to be set up directly between hosts. Solutions requiring star networking can be considered as a subset of mesh networks. In such a case connections between STs could be made over the BSMS with a star networks and by double hops, first to the default gateway and from there to the destination host.

Mesh systems, and especially the user terminals, can be based either on link layer switching only (e.g. cross-connects, remote bridges, etc.) but layer 3-based STs and satellite systems are the main focus of the present document to ensure IP services and are discussed below.

This clause first discusses solutions to routing over the BSMS. Then solutions to specific link layer-based architectures are described.

7.2.1 Routing within the BSMS

The STs act as routers and should implement the following routing protocols as in figure 7.0.2:

1) E-BGP internally (and optionally externally) if the BSMS is used to set up VPNs dynamically.

2) E-BGP externally if the BSMS is an independent AS or at the edge of an AS.

3) I-BGP internally if the BSMS if the BSMS is an independent AS.

4) IGP internally and externally if the BSMS is part of an AS.

As STs enter or leave the network, their routers need to discover routes and next-hop IP addresses of neighbours by exchanging routing information (by OSPF, BGP, etc.). Routers also participate in route advertisement, since when an area router learns a route from a peer it advertises it to all its peers.

When the number of IP nodes served by the BSMS becomes large, as it may well do over the coverage, scalable IP routing becomes a challenge. Routing protocols between routers in IP nodes or in STs attached to them exchange routing data with all other routers (i.e. full mesh connectivity) over the BSMS coverage, giving rise to a huge volume of routing data, proportional to N^2 (N= no. of routers).

It is important to simplify these routing protocol exchanges over the satellite which risk occupying considerable capacity.

An improved, more scalable, solution to the mesh routing protocol connectivity normally needed between routers is to employ star routing protocol connectivity, by providing peering between attached routers and a central **BSMS Route Server** (e.g. associated with the NCC etc.), instead of the meshed connectivity they would have otherwise. The routing protocol traffic can be reduced significantly in this way [10].

The BSMS Router Server (**BRS**) advertises routes across the BSMS on behalf of simplified "routers" in the STs. The BRS is also used as a "designated router" for OSPF protocols.

The BRS is not a complete router itself since IP user plane traffic between STs should still be meshed (e.g. via OBP switching) to avoid double hops through the satellite, but the BRS implements the routing protocols and algorithms in the control plane of the BSMS. The BRS is forwarded any routing protocol packets which arrive at an ST, rather than the ST processing them internally. The BRS is responsible for updating its routing tables and distributing updates to external neighbours. The STs roles are mainly focussed on the User Plane and on forwarding of packets. The BRS is responsible for sending updates of the forwarding tables of the STs, or sending complete forwarding tables when necessary (e.g. on log-on).

The BRS can also be partitioned amongst different ASs or OSPF Areas within the coverage region.

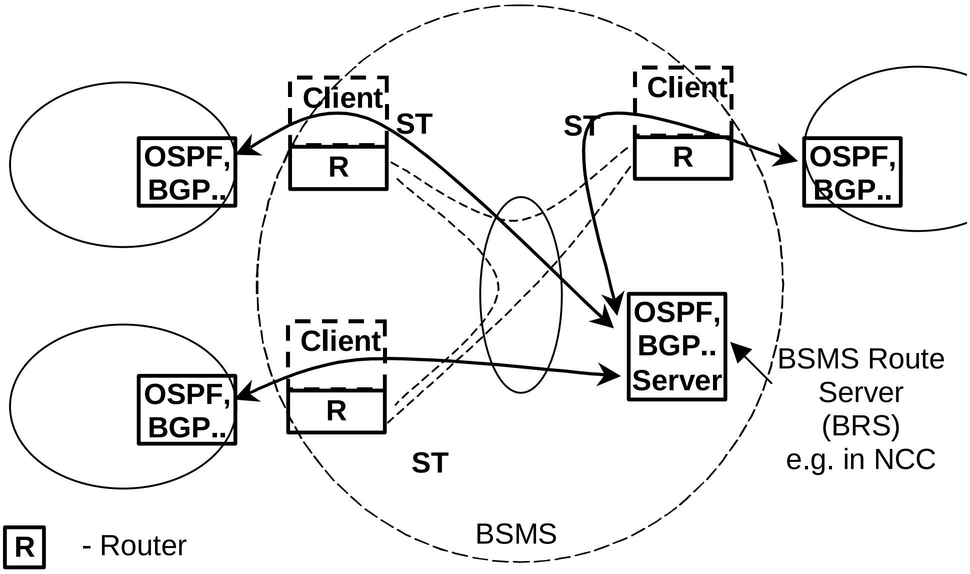


Figure 7.2.1: Proposed BSMS routing architecture

The disadvantage of the architecture is that the BRS is potentially complex and more costly than an ST-only based system because it does not use standard routers.

Figure 7.2.2 shows a reference model for the protocol architecture of the BSMS. It shows the location of the ARP and Routing functions in the protocol stack of STs.

The diagram also shows the SI-SAP interface, which is used to separate the IP Adaptation functions from the satellite-specific functions of the ST. The SI-SAP is used in the interworking mechanisms of IP over the satellite link. An IP network address is assigned to the satellite independent or IP side of the SI-SAP, while a satellite layer 2 address is assigned to the satellite dependent or MAC side of the SI-SAP. Resolution between these two addresses is performed when other STs intend to send IP packets across the satellite link to this ST as the next hop IP router.

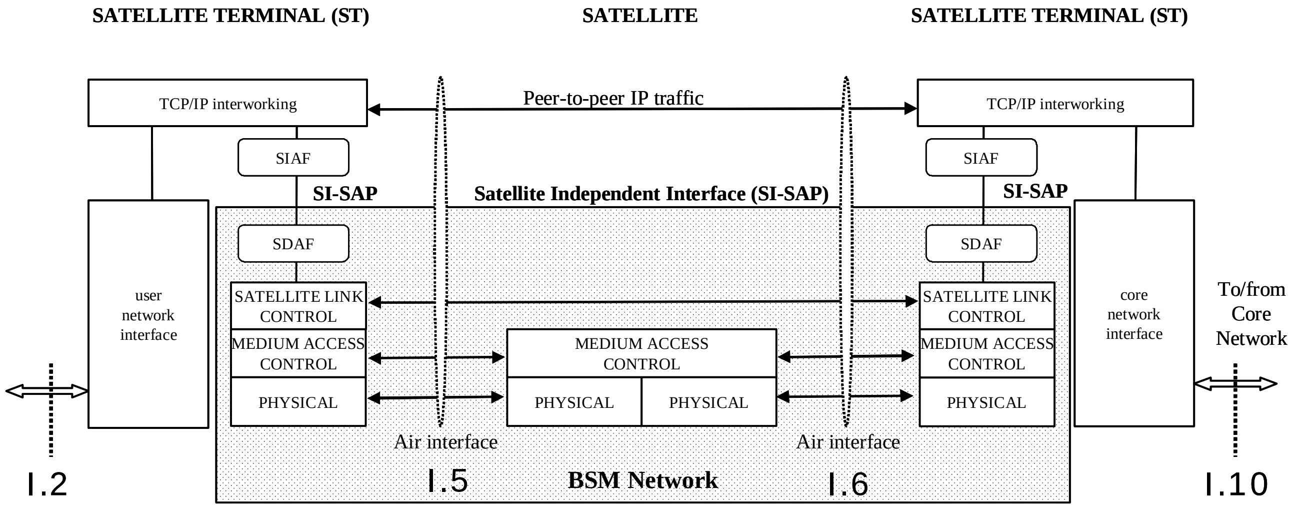


Figure 7.2.2: BSMS Protocol Architecture (OBP satellite)

7.2.2 Dynamic routing with on-board processing

The BSMS must offer a transparent medium for IP packet routing. Figure 7.2.3 shows an example of how the BSMS becomes essentially a dynamic router (using an IGP and either a RIP or OSPF metric) with an on-board processor (combination of an On-Board Controller (OBC) and On-Board Switch (OBS)). When IP packets are presented at the BSMS ingress the ST does not perform the routing decision. Instead it requests for capacity at the OBC and when allocated sends the satellite frames or cells (the S-MTUs - satellite medium transfer units) to the on-board switch with the appropriate addressing. After traversing the switch fabric the S-MTU is placed in a queue that either represents its destination or its type.

The final decision of the appropriate downlink is based on pre-loaded routing information (linking for example a certain destination to a certain beam) as well as instant queue congestion or downlink congestion (in fact the packet could be dropped at that point). After a successful decision the packet is sent to its final destination. If the satellite is transparent then the operations of the OBP are controlled in the NCC and the downlink beam is replaced by a physical port number.

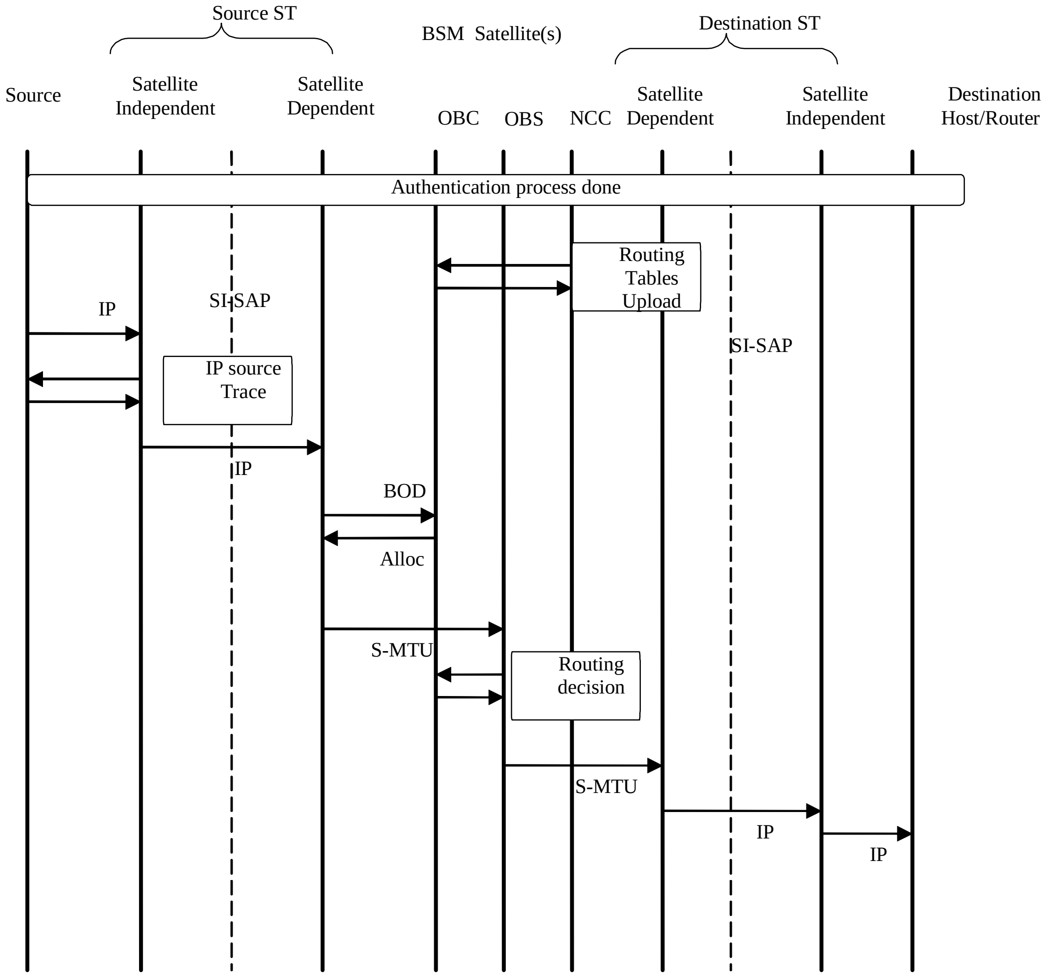


Figure 7.2.3: Dynamic routing with on-board processing

7.2.2.1 BSMS internal routing tables

The BSMS routing table entries will use the familiar format used for most current routing tables and have at the minimum the following entries:

- Input port.
- Destination address.
- Output port (in the BSMS a queue associated with a downlink beam in a OBP or a physical port in a gateway).

For any destination address the output port is determined by the metric used, distance of link state. Also the input port may determine an index as multiple routing tables could be used (this is unlikely in the BSM). Finally specific information in the S-MTU, such as QoS, ECN etc. may also influence its destination. In an MPLS router the input port and output ports are replaced by input labels and output labels. An example for the OBP BSMS is presented in clause 7.2.2.

Input beam	Destination address	Downlink beam
Which uplink?	Based on the ST id or other address within the MTU	Updated by frequent information about the status of the OBS

7.2.2.2 Use of "cost-based" routing in a multi-homed BSMS

In a meshed environment the path to reach a destination may vary, or more than one route may be available. As shown in figure 7.2.4 a single host ST may reach a destination through one or more STs that could be attached to the Internet. In that case the "cost" (see clause 5.2.4.2.1) of going to ST3 or ST4 may change over time depending on the air interface status between the satellite and the STs, the congestion on-board or the congestion in the egress STs. Thus the routing table in ST1 (or the OBP/OBC, depending on the BSMS routing architecture) will evolve over time (there will always be a static default route to the gateway).

Two issues are thus raised:

- How does ST1 (or the OBP/OBC) choose the "best" route to the destination host i.e. how are the costs of the routes evaluated?
- How can the BRS help the determination of the best routes (since the goal of the design is to keep the OBC simple)?

Issues and avenues to solutions in these cases are:

1) for Costs:
   - Level of queue congestion of the main route downlink (if queuing by downlink).
   - Level of queue congestion on the destination queue (if destination based queuing).
   - Link QoS due to weather at destination.
   - "Real" cost in money sense of going to a certain ST.
   - Destination ST on the same administrative domain as the source ST.
   - Destination host on the same administrative domain and the source ST/host.
   - Other policy based decisions.

2) for BRS functions:
   - Computation of routing tables (cost matrix) that involve multidimensional metrics.
   - Semi-static/global routing table computation.
   - Admission control and policy setting.
   - Determination of alternate routes and default routes.
   - Computation of long range statistics that can define main and alternate routes for source-destination pairs.
   - Recording of fault events.
   - Performance and QoS monitoring.
   - Other non real time management/FCAPS functions.

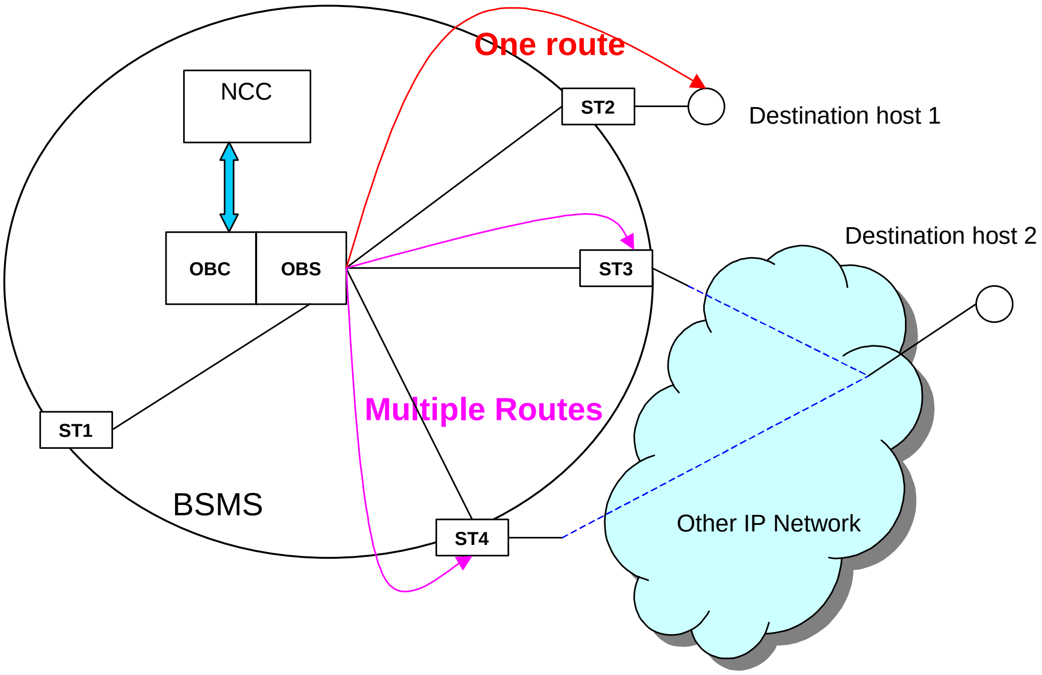


Figure 7.2.4: Example of "cost-based" routing choices with on-board processing

7.2.2.3 Neighbour discovery in OSPF over BSM

In OSPFv2 it is essential that every router on a common network using OSPF enables 2-way communications and keeps its database synchronized. The protocol to discover and maintain neighbours is called the "Hello" protocol. In the Hello protocol OSPF packets of type 1 (Hello) are sent over UDP to all neighbours in a broadcast network and to an adjacent neighbour (discovered via ARP for example) in a non-broadcast environment. Over a BSMS (see clause 5.2.4.2.3) the Hello packet from an upstream router will reach the ST as a UDP packet. This packet is sent as high priority because of its fundamental role in the setting of network topology. If there is no available bandwidth at this point the BOD process in the ST will request and get the appropriate bandwidth and send the packet to its destination(s). There are two issues with this. The Hello packet contains two timing fields:

- HelloInterval - that defines the frequency of the Hello messages – a variable delay over the BSM may disturb this function.
- RouterDeadInterval that defines how often a neighbour has to be heard from to remain active (the time between the last Hello received and the next one).

These are not only set over the BSM but are shared on all the routers attached to the common network. Hence the operator of a network with an OSPF IGP must set these parameters to ensure the BSM fully participates in routing.

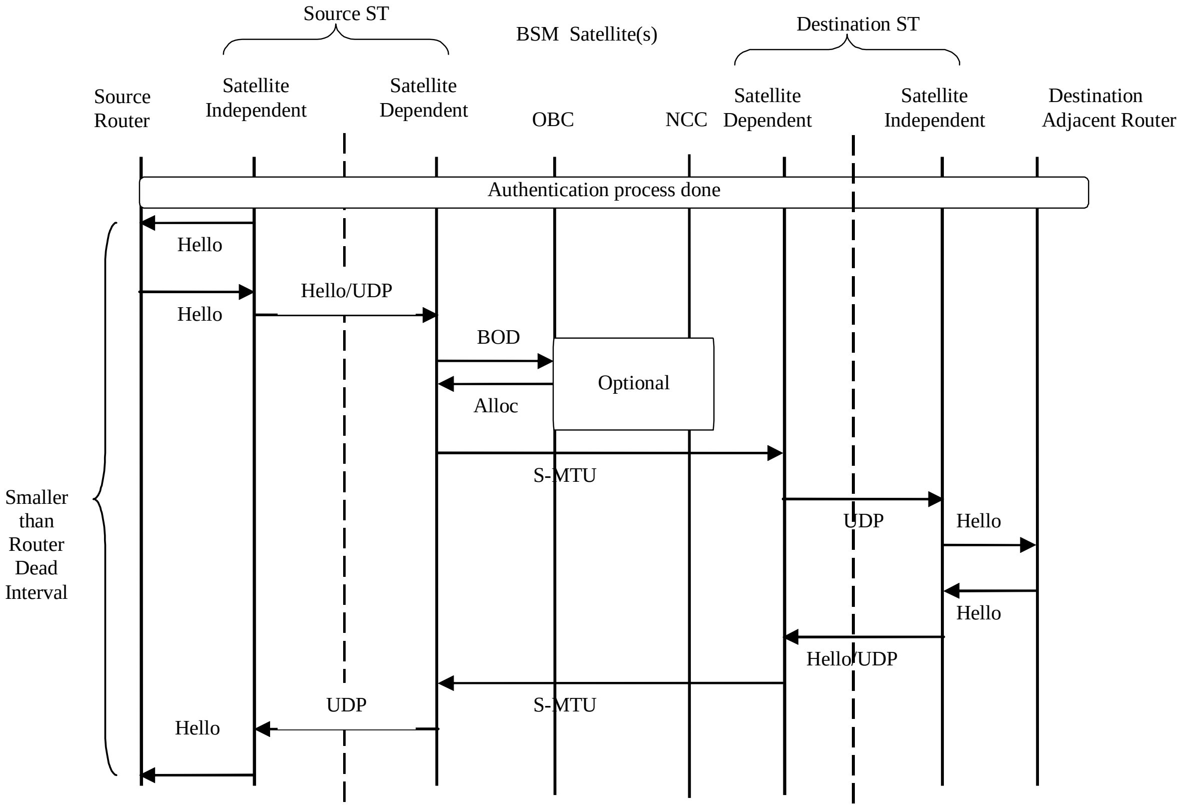


Figure 7.2.5: OSPF neighbour discovery messages

Conclusion 10	Hence to be transparent to the OSPFv2 Hello protocol the BSM should ensure that the functionalities of the protocol are preserved over the BSM.

<table>
<tr><td>Version #</td><td>Type</td><td colspan="2">Packet length</td></tr>
<tr><td colspan="4">Router ID</td></tr>
<tr><td colspan="4">Area ID</td></tr>
<tr><td colspan="2">Checksum</td><td colspan="2">AuType</td></tr>
<tr><td colspan="4">Authentication</td></tr>
<tr><td colspan="4">Authentication</td></tr>
<tr><td colspan="4">Network Mask</td></tr>
<tr><td colspan="2">HelloInterval</td><td>Options</td><td>Rrt Pri</td></tr>
<tr><td colspan="4">RouterDeadROuter</td></tr>
<tr><td colspan="4">Designated Router</td></tr>
<tr><td colspan="4">Backup Designated Router</td></tr>
<tr><td colspan="4">Neighbour</td></tr>
<tr><td colspan="4">…</td></tr>
</table>

Version number: 2

Type: 1 for Hello packets
Packet length: in bytes including headers
Router ID: source router ID
Area ID: identifies the area to which the packet belongs; all OSPF packets are from a single area
Checksum: ensure the validity of the OSPF packet
AuType: the type of authentication used
Authentication: 64 bits of authentication
Network mask: network mask associated with the interface that is sending the Hello message; Hello messages are sent by all interfaces of the router
Options: OSPF options
HelloInterval: the number of seconds between the source's Hello messages
Rtr Pri: router priority
RouterDeadInterval: The number of seconds before declaring a silent router down.
Designated router
Backup designated router: Identified by specific IP addresses
Neighbor: The router ID of each router from whom Hello packets were received recently.(i.e. in the last RouterDeadInterval)

Figure 7.2.6: OSPF Hello packet

7.2.2.4 Proposed messages for route discovery and other signalling

As can be seen in the present document and the other IP-related TRs [3], [4], [5], two of the main problems in any BSMS as regards route discovery and related topics (address resolution, labelling paths in MPLS, RSVP reservation etc.) are the potential limitations in bandwidth at the BSMS ingress terminal and the end-to-end delay. This delay includes not only the physical transmission of the signal through space but also the effects of bandwidth on demand algorithms, requesting bandwidth when it is not there, and other link layer mechanism for those terminals that are not always on. The compound effect of the bandwidth and the delay in allocating it make it possible that signalling messages may time out, delaying the whole network and possibly triggering congestion, or worse that the lack of a response makes the BSMS virtually unavailable at the network layer even if it is available otherwise. Ways of dealing with this problem are highlighted in this clause.

1) Use of a terminal signalling channel

   In those terminals that use an S-Aloha signalling channel, IP signalling and route discovery messages - that are identifiable by their destination addresses - can be piggybacked on the usual terminal signalling for capacity. When a message comes in, it triggers a capacity request and is appended to it. If the request for capacity does not get into a collision then the IP message goes through. The allocated capacity can then either be released or used for other messages. This obviously is better suited when there is little traffic on the BSMS.

2) Dedicated IP low-bandwidth signalling channel

   There is also the possibility of dedicating an IP signalling or minimum bandwidth IP traffic channel to be available permanently. This means that there is one always-on BSMS traffic channel per terminal. While this channel may need very little capacity, as signalling messages are usually short, this is a more suitable solution for the gateway solution and transparent satellite where the signalling messages can be broadcast at least on the forward channel and the small return channel may not use significant satellite capacity (or use another network). However for large mesh-based BSMs even a small always-on channel can be seen as consuming too many of the scarce resources that are needed for "paying" traffic.

3) IP signalling pre-emption class

Another solution, that can be combined in fact with any of the above is the creation of a "pre-emption" class is to allow IP signalling messages (usually short) to be transmitted a soon as they reach the ingress of the BSMS. This is simple: as soon as an "important" message is identified it pre-empts any existing queue and leave in the next available frame (or time slot or transmission opportunity). This way the message is guaranteed to be sent and experiences minimal end to end delay. It does mean however that even highest priority queue will experience a short duration added delay jitter. However if the traffic load at the terminal (or gateway) is fairly high over the duration of the call/session, this jitter will not prove significant.

7.2.2.5 Handling of IPv4 TTL fields

Normally, when a router receives an IP packet, it decrements either the TTL and determines whether the IP packet has reached its packet "lifespan". The originating host initializes the TTL or Hop Limit field; and when the field reaches zero the IP packet is dropped. Whenever a router discards a datagram because its hop count has reached zero, it sends an ICMP time exceeded message back to the IP source address.

Within the BSMS this process could occur for packets received both at the terrestrial interface and satellite link of a ST. However, efficiencies could be obtained if these two steps are combined. When a ST receives an IP packet at its terrestrial interface, it will either successfully resolve the destination IP address of the packet or it will discard the packet. If it finds a route entry which identifies that the next hop IP address is associated with its Ethernet port, then the ST subtracts one from the TTL field and, if zero, discards the IP packet. If the ST finds a route entry, which identifies that the next hop ST is reached via the satellite link, then the ST may subtract 2 from the TTL field and, if zero or negative, discard the IP packet.

When the ST discards the IP packet because of the TTL field is zero or negative, it should send an "ICMP time exceeded message" back to the IP source address.

7.2.2.6 On-board IP routing

A technical solution for the BSMS is to use on-board routing. The advantage is that the BSMS becomes a star network for all IP traffic and protocols, and with minimum transit delay.

The satellite is configured as the default next hop in users' edge routers and address resolution (like NHRP, etc.) and other protocols may be simplified, particularly if the satellite takes over all routing functions. Management of the satellite routing table can be manual, or use IGP, etc. The on-board router views the subnetworks as local networks, avoiding the need for next hop resolution locally. Permanent Virtual Circuits (PVCs) or equivalent links are set up between each edge router and the satellite.

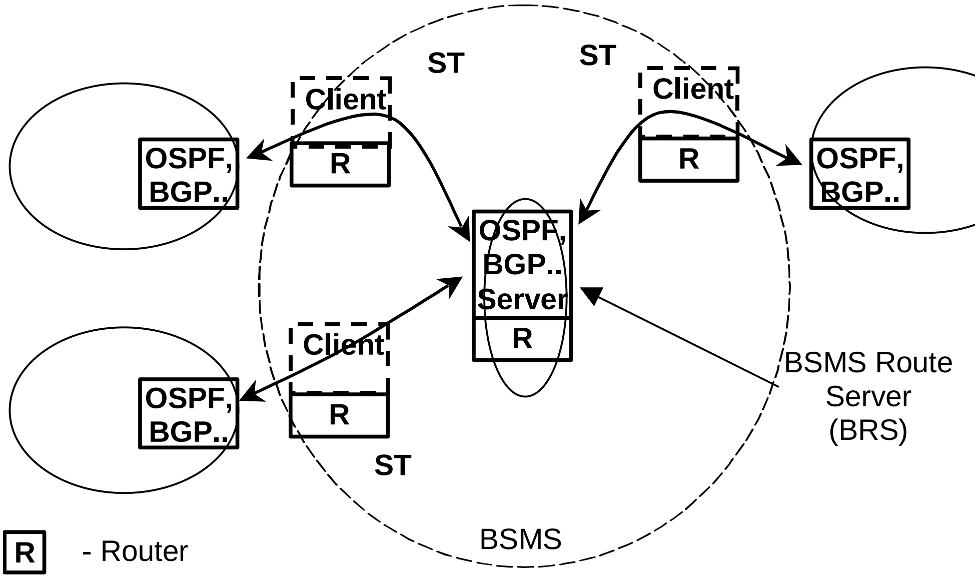


Figure 7.2.7: On-board routing system architecture

An obvious difficulty with this scheme is the on-board router implementation, associated with access and malicious interference security issues. Another problem could be large on-board routing tables due to low aggregation levels.

7.3 Address allocation and resolution

Any BSMS address server functions required for allocating IP addresses and resolving layer 2 addresses should also, like the route server, be centralized (e.g. with the NCC) for the efficiency reasons discussed for routing above. They may or may not be co-located with the route server and there may be more than one route server and address server in the system.

A BSMS Address Resolution Server (BARS) is proposed to encompass any such centralized addressing functions.

Standard mechanisms for address allocation and resolution, DHCP, ARP IETF RFC 826 [24] and RARP IETF RFC 903 [84], see clause 5.2.6, should be handled wherever possible within the BSMS. However BSMS-specific address resolution and allocation protocols could replace the Ethernet framing and transmission of the original Internet ARP protocols. For the IP layer however this should be transparent.

The following clauses describe general procedures. More specific solutions for alternative link layers are discussed in clause 7.4.

7.3.1 IP address uniqueness issues

Because the BSMS is shared between many subnets which may use non-unique private addressing, there is a need to find a way round this problem since a unique ST IP address is required for AR.

For IPv4 networks, an entity called the ***Satellite Next Hop Address*** is proposed to provide the uniqueness required for routing and address resolution across the satellite link. The Satellite Next Hop Address is found in route table entries and is the address that comes from the route lookup process. This same address is also in the AR cache for the satellite interface and is used to resolve to the Satellite Destination MAC Address. Further information regarding the structure of route table entries is in clause 7.2.1.

The STs AR cache for the satellite interface may be composed of the following parameters:

- Satellite Next Hop Address.
- Satellite Destination MAC Address.
- Static AR entry indicator.
- Pending AR entry indicator.
- Stale AR entry indicator.
- AR pass/fail indicator.
- AR entry timer value.

AR cache entries for the satellite link will be configured by the BARS or learned by the ST using Satellite ARP. Configured cache entries will have the "Static AR entry" indicator set and will not be timed out. A likely scenario for using static cache entries is the case when the remote ST is configured with a default route pointing to a hub or a satellite terminal supporting a gateway function. In such a case, a static AR entry ought to be configured to resolve the network address of the Access Gateway.

The AR cache settings for pending, stale, and pass/fail are associated with Satellite ARP discussed in. If an ST is configured to use Satellite ARP and an entry is not found in the AR cache, then the ST participates in the Satellite ARP protocol to resolve the Satellite Next Hop Address to the Satellite Destination MAC address. The ST may insert the Satellite Next Hop Address into the AR cache entry and set the Pending AR entry indicator while it is waiting for a response from the BARS for the Satellite ARP request. Once the Satellite ARP response is returned to the ST, entries learned by way of Satellite ARP are inserted into the AR cache along with the pre-configured cache timeout value. The Pending AR entry indicator is cleared and the AR pass/fail indicator is set according to the value returned in the Satellite ARP Response. When the AR entry times out, the Stale AR entry indicator is set.

There may be conditions when the ST is configured with multiple network layer addresses for the same Satellite Destination MAC address. If the Satellite Next Hop Address cannot be resolved to the Satellite Destination MAC Address, then the ST may drop the IP packet.

7.3.2 Obtaining an IP address by an ST

When a customer's ST, to which one or more hosts are connected, is turned on, it must first go through an authentication/logon procedure, usually under the control of an SP's authentication server. After this, the ST needs to obtain an IP address it will use to receive IP packets.

A DHCP proxy can assign addresses to its attached hosts (STs) or the hosts can request their own addresses. Figure 7.3.1 shows the case where the SPs RAS (or the BARS) runs either a DHCP or RARP proxy and assigns IP addresses to attached STs. The BSM-specific function lies at the Satellite-Dependent interface where the DHCP or RARP command is translated into a BARS command (of the GetIPAdress(Terminal_ID) type) to request an address to the BARS via either static or dynamic assignment.

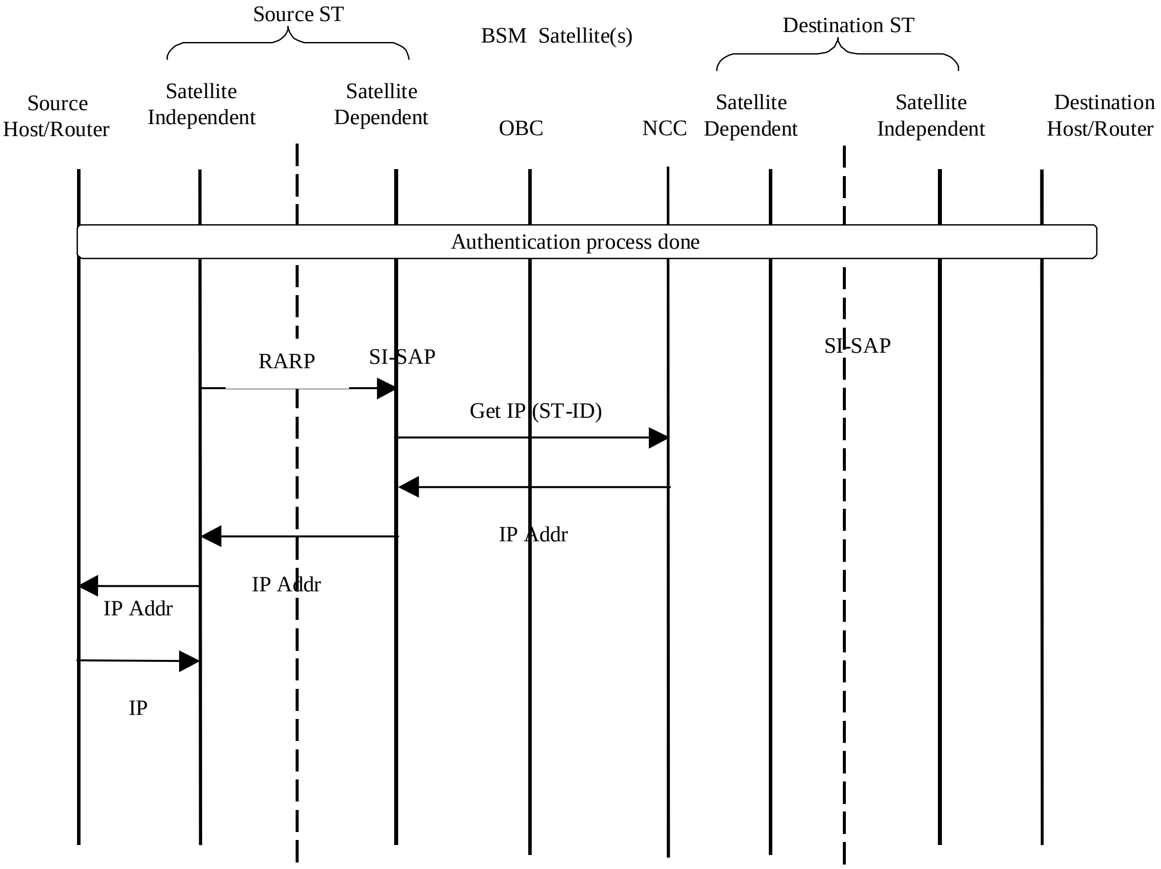


Figure 7.3.1: Obtaining an IP address by an ST

7.3.3 Obtaining a layer 2 address of a destination ST

Layer 2 addressing is the purview of the satellite operator and is specific to the satellite system design. An address resolution protocol (e.g. ARP for IPv4, Neighbour Discovery for IPv6, etc.) provides a means of obtaining and allocating layer 2 addresses of STs.

Before data can be sent by an ST to another, it needs to obtain the destination ST layer 2 address. Figure 7.3.2 shows a simple case where both source and destination are on the same BSMS. The BARS runs an ARP proxy and returns ST-Ids for its attached terminals.

The BSM-specific function here lies at the Satellite-Dependent interface where the ARP command is translated into a BARS command (of the GetSTAdress(IP_address) type) to request an address to the BARS either via a static or via a dynamic assignment. The figure also shows that the BARS may verify the destination terminal is alive before responding to the request.

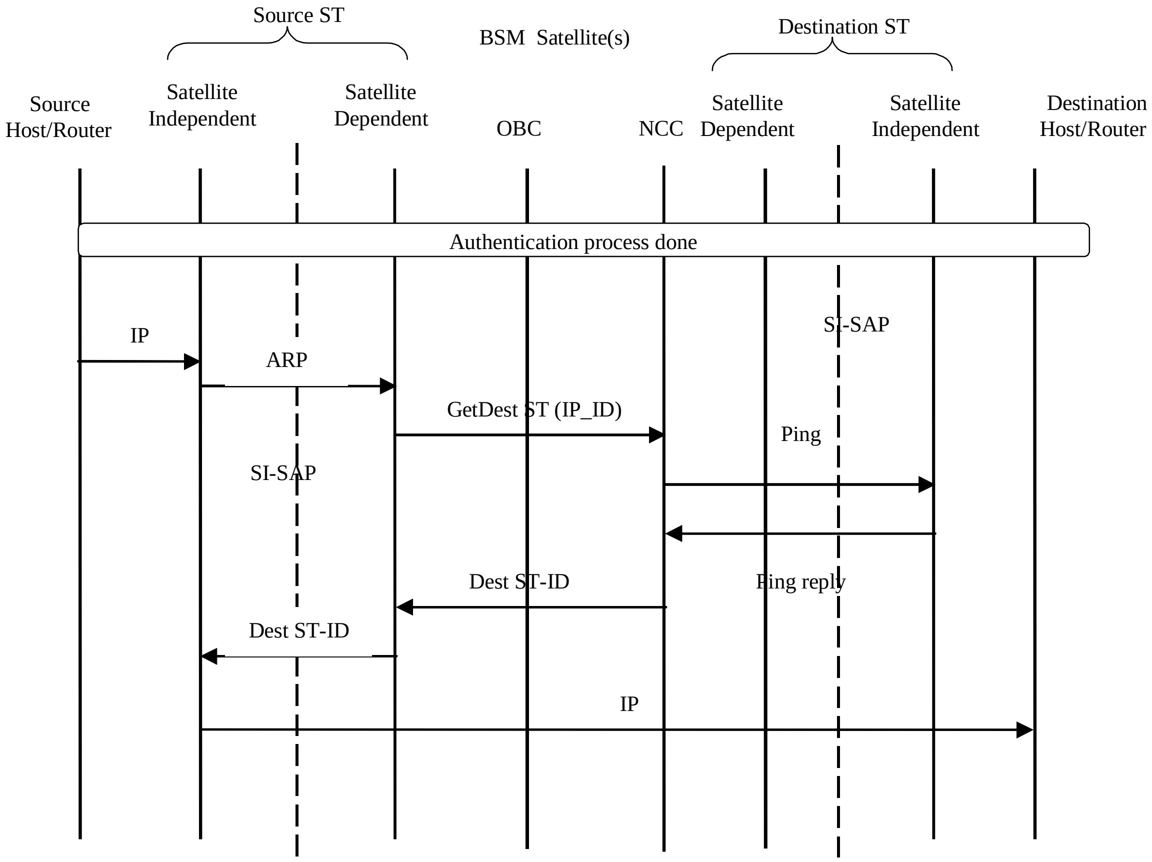


Figure 7.3.2: Obtaining a layer 2 address by an ST

The two messages necessary for this mechanism are Satellite ARP Request and Satellite ARP Response. The Satellite ARP Request has the following format:

- Satellite ARP Version Number.
- Satellite ARP Message Type (Satellite ARP Request).
- Satellite Next Hop Address.
- Requesting access terminal identity.

The Satellite ARP response has the following format:

- Satellite ARP Version Number.
- Satellite ARP Message Type (Satellite ARP Response).
- Satellite Next Hop Address.
- Satellite Destination MAC address.
- Satellite ARP Response Status.
- Satellite ARP requesting access terminal identity.

When the BARS receives a Satellite ARP request, it returns the Satellite ARP Response. The tables containing the address resolution between Satellite Next Hop Address and Satellite Destination MAC Address will be populated as part of the configuration process of a satellite access terminal. A Satellite ARP response status is returned for error conditions or when some other restrictions prevent communication. The BARS returns the Satellite ARP Request ID to the satellite access terminal in the Satellite ARP Response message.

There are some Satellite ARP Response Status values which indicate that the Satellite Next Hop Address was not resolved by the BARS for some reason which is transient in nature. In these cases, the satellite access terminal will not insert the responses in the AR cache, but will retry at an interval which is typically much shorter than the AR cache timeout value specified by the network operator.

7.4 Addressing and routing over specific link layer architectures

Layer 3-oriented BSMS architectures considered are:

1) IP over ATM (or ATM-like).

2) IP over Label-oriented Link/Shim layer protocols (MPLS, etc.).

3) IP over DVB RCS.

4) On-board forwarding/routing.

Solutions for these architectures are described below.

7.4.1 IP over ATM (or ATM-like)

ATM-like link layers are included here since similar protocols have been applied to satellite links where the ATM frame length is adapted to satellites and varies from the standard 53 bytes.

STs need to interconnect via VP/VCs within the BSMS. Because of the ATM point-point connectivity (NBMA), each router has to repeat route advertisements N times (for each of N edge routers) over different ATM paths.

The main questions that arise here are:

1) How to discover the next hop (ATM address resolution)?

2) How to set up corresponding VPC/VCCs?

Because of the ATM point-point connectivity (NBMA) between STs, each router has to repeat route advertisements N times (for each of N edge routers) over different ATM paths.

ATM connections between ST (and at the OBP satellite) should be managed centrally by an ATM signalling protocol server (BSMS Signalling Server - BSS).

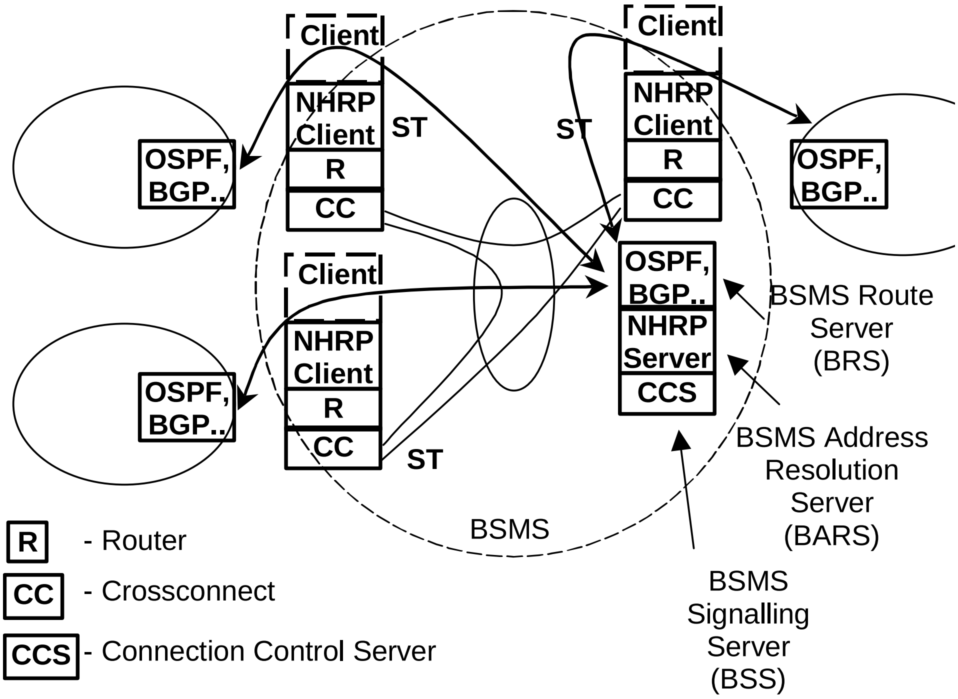


Figure 7.4.1: Layer 3-based IP over ATM architecture

7.4.1.1 Address resolution

The ST routers also need to discover the ATM addresses corresponding to next hop IP addresses, since ATM is a point-to-point connection-oriented protocol. They then need to request set-up of the ATM connection. This resolution process is performed by NHRP for NBMA links. By using a centralized BSMS Address server (the BAS e.g.associated with the NCC) it can manage connectivity within the BSMS as well as to the nearest attached routers.

The NHRP Server can cooperate closely with the ATM connection control server to eliminate additional hops for signalling. The attached routers may belong to different ASs and the BARS can be partitioned into different administrative zones but with cooperation between zones.

The message flows are as follows:

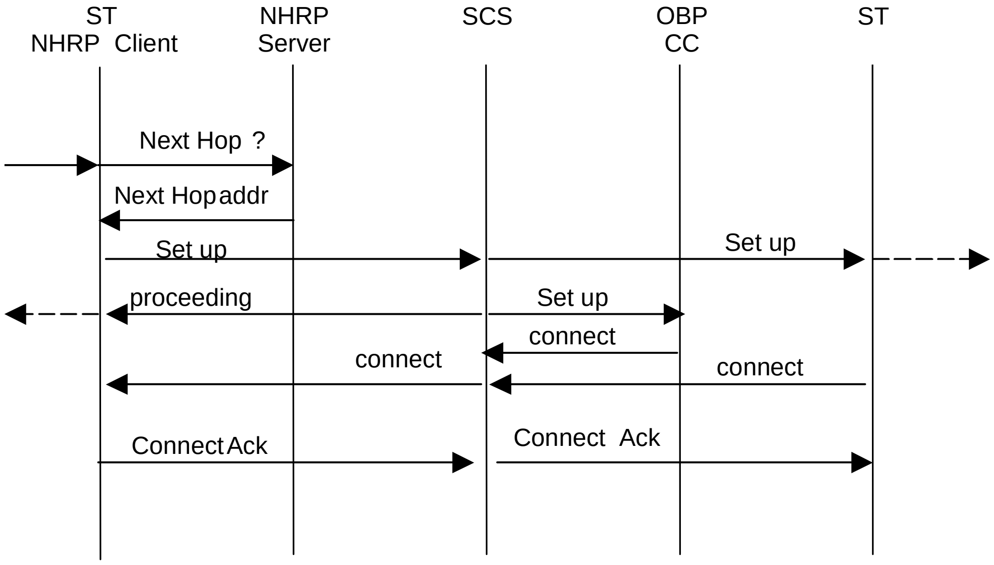


Figure 7.4.2

The disadvantages of the architecture are:

1) Significantly fewer PVCs are needed than a mesh connecting routing protocols but route advertisements continue to be duplicated N-1 times.

2) IP to ATM address resolution is still needed in addition to SVC signalling.

7.4.2 IP over label-based link layer architectures

7.4.2.1 IP over MPLS

MPLS provides connection-oriented links based on IP routing and control protocols. It is generally deployed in an MPLS Domain containing Label Edge Routers, which apply labels to incoming IP packets, and core Label Switching Routers which perform label switching. Hence routing protocols only interact with the LERs. The Label Distribution Protocol is used to assign labels based on routing information and other constraints such as QoS.

The MPLS domain could be restricted to the BSMS or could extend into terrestrial MPLS networks.

An MPLS network uses IP for its control plane. This allows the reuse of all IP technology for auto-configuration, addressing, routing, reliable transport etc. The control plane consists of a private IP network within the MPLS domain carrying a signalling protocol for LSP establishment and possibly other protocols for auto-configuration or routing.

When MPLS is used over ATM, MPLS allows the replacement of ATM signalling protocols (such as ITU-T Recommendation Q.2931 [88] and PNNI for routing) to be replaced by much lighter LDP and IP routing, to set up VCs called Label Switched Paths (LSPs).

The labels and their binding are managed by the LDP, and the LDP should be optimized for the BSMS to minimize signalling overhead. The number of labels in use is an important factor in this overhead due to the potentially large number of STs, QoS parameters etc. Efficient use of labels is therefore another consideration.

A solution to LDP optimization is that all STs should "snoop" the broadcast label resolution responses to other STs and build their forwarding tables. In addition a table of labels could be uploaded at registration of an ST, and updates broadcast periodically to all, taking advantages of satellite multicast.

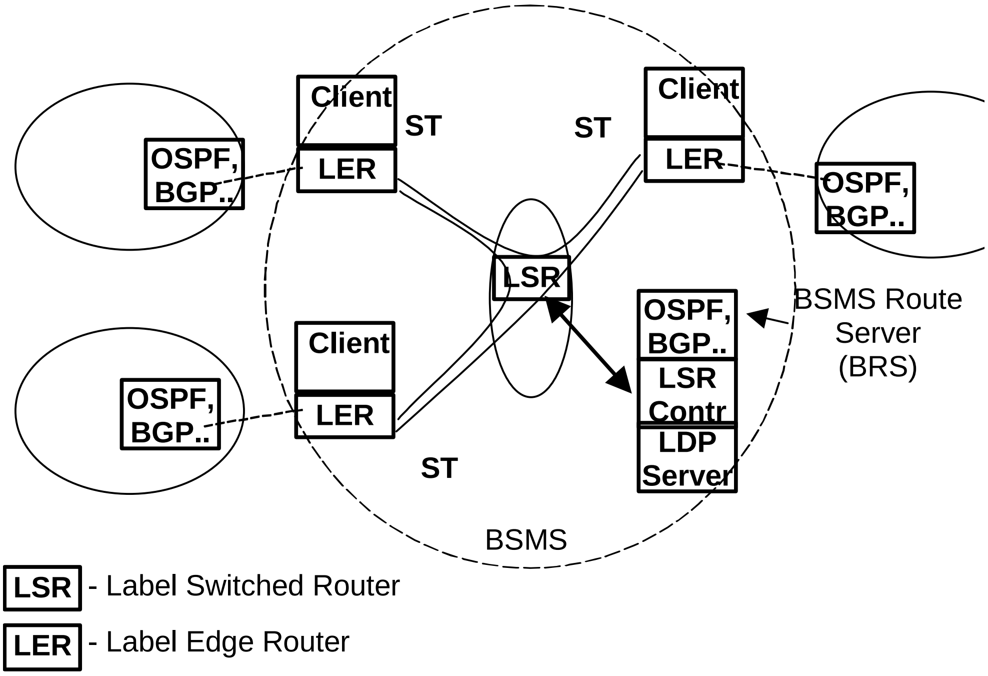


Figure 7.4.3: Layer 3-based IP over MPLS architecture

7.4.2.2 Proposed "label-based" IP-oriented solution

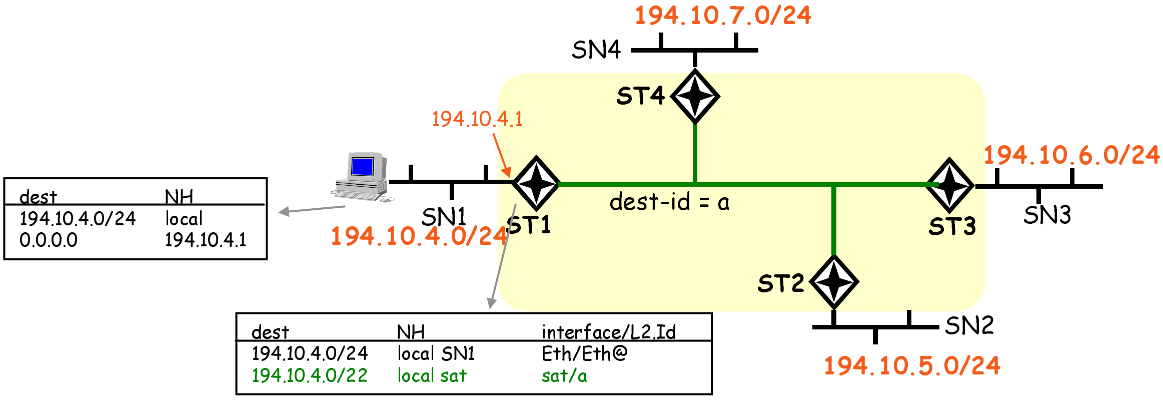
A proposal for the underlying protocol to the IP layer in the BSMS is to employ "label paths" acting as logical broadcast networks: the same label could be used from several STs towards several others.

This is a simpler scheme compared to MPLS, and better adapted to satellites, by taking advantage of satellites' multicast attributes.

Figure 7.4.4 shows an example of a L3 VPN configuration across a BSMS network using the concept of satellite "label paths". Here Network 194.10.4.0/22 consists of four subnets, each of them behind a given ST and in a given spotbeam: SN1 (194.10.4.0/24) to SN4 (194.10.7.0/24). STs satellite interface is configured with the global VPN prefix (194.10.4.0/22) pointing to the satellite local interface (denoted "local sat" NH or Next Hop): any destination in the VPN is considered as local and can be reached directly through the satellite interface through a unique link layer label (denoted "dest-id" or Destination Identifier, which value is "a").

At reception, STs analyse all IP frames emitted with the link layer label "a" and only let in those corresponding to their LAN prefix.

According to the complexity of the topology, these link layer identifiers can be configured manually in the STs or dynamically through a Satellite Address Resolution Protocol (S-ARP).



dest	NH
194.10.4.0/24	local
0.0.0.0	194.10.4.1

dest	NH	interface/L2.Id
194.10.4.0/24	local SN1	Eth/Eth@
194.10.4.0/22	local sat	sat/a

Figure 7.4.4: Layer 3 VPN across BSM

Such solutions could be deployed provided some characteristics are standardized:

- Link layer labelling concepts and label structure.
- Operations for Router STs and for Bridge STs.
- Mapping of labels onto standardized satellite access layer identifiers (PID, VPI/VCI).
- Satellite ARP for dynamic resolution.
- Segmentation and re-assembly functions (when a shared label is used to filter traffic at reception, a function is needed to identify the source of the traffic and therefore perform re-assembly per source).
- QoS support: interaction with IP based (e.g. RSVP) or session based signalling (e.g. SIP).

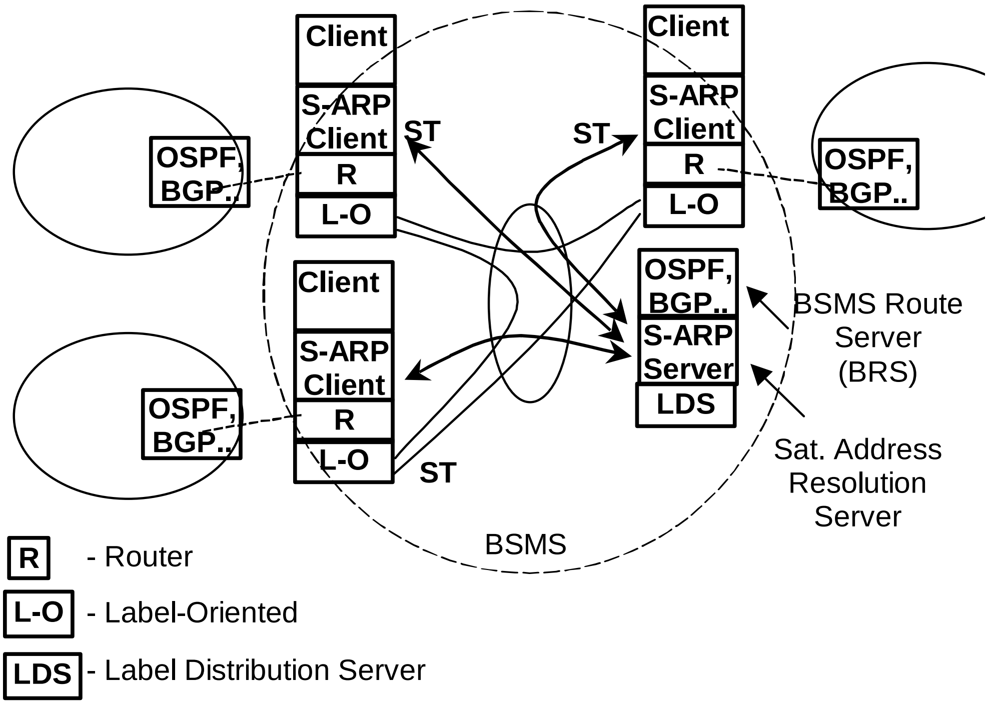


Figure 7.4.5: IP over IP-oriented link layer architecture

7.4.2.2.1 Use of S-ARP

Address resolution is performed after the router interface is determined and makes use of an AR cache, which keeps AR entries for resolving the network address. The AR cache may be populated by a network protocol associated with the router interface; this clause identifies the protocols used for AR. In addition, static AR entries may also be configured under certain conditions. The assumption here is that the details of static configuration protocol is system specific and not part of a generalized satellite ARP protocol. This clause describes a satellite ARP protocol for dynamic address resolution and how the static AR entries might be configured.

A centralized server is recommended for both address resolution and route resolution. Intrinsically, MAC addressing is the purview of the satellite operator and is specific to the BSMS design. Thus the Address Resolution Protocol (ARP) server function is located at the BARS (or in the NCC). The route server function should also be centralized as described above, but it may or may not be collocated with the address server and there may be more than one route server in the system.

7.4.2.2.2 S-ARP description

The ST and the BARS participate in the Satellite ARP protocol. The BARS provides an AR server that contains a database of all Satellite Destination MAC addresses for each satellite in a given network. The Satellite Destination MAC address is unique to a satellite link. Associated with the Satellite Next Hop Address is a Satellite Destination MAC Address.

The ST should query the AR server for the address resolution information for forwarding the packet, and after validating the query, the AR server would return an AR entry. The ST should insert this entry into the AR cache associated with the satellite interface and use it to forward the packet to another ST across the satellite link. The IP packet which triggered the Satellite ARP request (and subsequent IP packets to the same subnet) may be queued.

The two messages necessary for this mechanism are Satellite ARP Request and Satellite ARP Response. The Satellite ARP Request has the following format:

- Satellite ARP Version Number.
- Satellite ARP Message Type (Satellite ARP Request).
- Satellite Next Hop Address.

- Requesting access terminal identity.

The Satellite ARP response has the following format:

- Satellite ARP Version Number.
- Satellite ARP Message Type (Satellite ARP Response).
- Satellite Next Hop Address.
- Satellite Destination MAC address.
- Satellite ARP Response Status.
- Satellite ARP requesting access terminal identity.

When the BARS receives a Satellite ARP request, it returns the Satellite ARP Response. The tables containing the address resolution between Satellite Next Hop Address and Satellite Destination MAC Address will be populated as part of the configuration process of a ST. A Satellite ARP response status is returned for error conditions or when some other restrictions prevent communication. The BARS returns the Satellite ARP Request ID to the ST in the Satellite ARP Response message.

There are some Satellite ARP Response Status values which indicate that the Satellite Next Hop Address was not resolved by the BARS for some reason which is transient in nature. In these cases, the ST will not insert the responses in the AR cache, but will retry at an interval which is typically much shorter than the AR cache timeout value specified by the network operator.

7.4.3 IP over DVB-RCS

DVB-RCS layer 2 solutions can also be foreseen. At present DVB-RCS is considered to have limited scope for mesh networking, though such systems with OBP satellites are being built. The ESA RSAT Group has proposed recommendations for modification of the standard involving regenerative satellites.

Use of DVB-RCS for a mesh network requires a means of allocating PIDs and/or MPE labels to IP different streams. The number of PIDs a typical DVB receiver can decode is limited, and instead filtering of packets at IP level is often necessary. In the same way as above a centralized PID/MPE label server is a suitable solution.

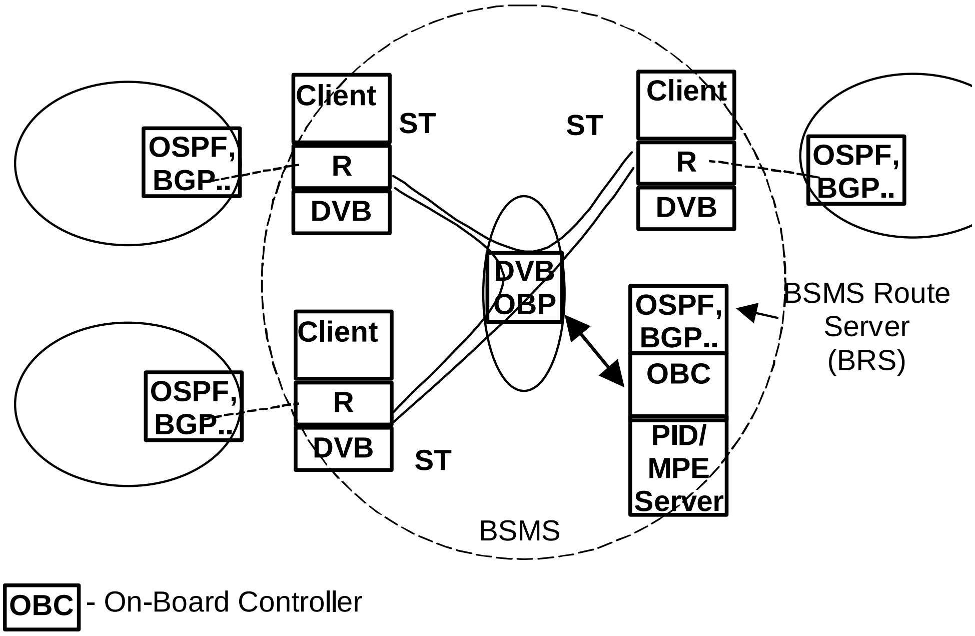


Figure 7.4.6: IP over DVB-RCS Architecture

8 Conclusions

There are many scenarios for Broadband Satellite Multimedia Systems (BSMS) deployment within IP networks. There are also many possible satellite system architectures to provide the services targeted by BSMS. These options lead to a range of functional requirements for IP Routing and Addressing. The most relevant ones have been described and common assumptions or requirements have been identified.

On this basis, conclusions have been reached and technical proposals defined.

Detailed assumptions and conclusions have been indicated within the text of the preceding clauses.

In this clause a summary of these assumptions and requirements is provided together with a summary of the conclusions.

Proposals for further work are covered in clause 9.

8.1 Constraints

IP routing protocols have been designed with terrestrial networks in mind, and mainly for point-to-point links. The **limited bandwidth** in wireless systems in general and in BSMS in particular requires that internal IP routing schemes be optimized to the specific properties of satellite links such as broadcast capability, asymmetric links and relatively high but constant transmission delay.

The wide geographic coverage of satellite networks also requires a wide range of interconnection scenarios to be accommodated. The coverage also leads to a potentially large number of links between Satellite Terminals (ST), and resulting **scalability** problems need to be avoided in BSMS IP routing.

8.2 Assumptions

The main assumptions for IP Routing and Addressing over the BSMS are:

1) BSMS must be able to provide links within and between a wide variety of IP networks, in the access, distribution and core sectors as well as acting as a private IP backbone.

2) BSMS must be transparent to, and interwork with, the relevant range of addressing schemes (i.e. global, private, IPv4, IPv6 etc.) present in IP flows over the BSMS simultaneously.

3) BSMS must be compatible with IP address management protocols.

4) BSMS must be compatible with IP routing protocols.

8.3 Specific conclusions

Taking into account the above assumptions, the conclusions are as follows.

8.3.1 System level

IP routing protocols have been designed with terrestrial point-to-point links in mind. The high cost of satellite capacity implies that internal IP routing schemes be optimized to the specific qualities of satellite links such as broadcast capability, asymmetric links and high delay.

The wide geographic coverage of satellite networks also requires a wide range of interconnection scenarios to be accommodated. The coverage also leads to a potentially large number of links between STs, and resulting scalability problems need to be avoided in BSMS IP routing.

Mesh-connected BSMS networks offer the greatest opportunity and challenge for routing.

The implementation efficiency of a mesh BSMS network depends on the type of lower BSMS protocol layers, i.e. the channel and beam arrangement of the physical layers and the connectivity offered by the link layer. Functions and protocols of alternative BSMS technologies supporting the IP and higher layers are intended to be specific to the lower Satellite-Technology-Dependent layers. The IP layer routing efficiency depends on the connectivity of the underlying mesh network. Nevertheless translation of IP addresses into link layer addresses is closely related to routing and the main link layer options and their address resolution procedures have been described. A general functional architecture for optimizing IP-to-layer 2 address resolution is proposed.

The routing protocols intercepted by the BSMS depend on its relationship with terrestrial Autonomous Systems which it covers.

In particular, in terms of routing, the BSMS should be capable of functions for implementing:

1) Intra-AS routing via IGP, I-BGP, etc.

2) Inter-AS routing via E-BGP.

3) IP-based VPNs, via IPSec, Virtual Router, MPLS, GRE etc., and auto-discovery via BGP-4.

Functions and protocols of alternative BSMS link technologies supporting the IP and higher layers are intended to be specific to the lower Satellite-Technology-Dependent layers.

8.3.2 Space segment

For mesh-connected BSMS networks, OBP switching satellites offer significant advantages in this type of network configuration. The OBP switch is controlled from a centralized NCC on ground.

Star networks and transparent satellites are not excluded from routing scenarios, but their solutions are a subset of mesh network solutions. On-board routing satellites offer more advantages but their technological challenges are considered only solvable in the long term, and if the security issues in clause 7.2.2.6 can be solved.

8.3.3 User segment

The STs should implement the minimum layer 3 functions that ensure compatibility with system requirements, whilst reducing cost and complexity of the equipment.

BSMS-specific optimization of addressing and routing protocols is focussed in the SIAF layer. A set of common primitives at the SI-SAP interface is proposed in clause 9 for optimized addressing and routing.

9 Recommendations for ETSI standards

This clause recommends BSMS IP Addressing and Routing issues that would benefit from standardization.

The BSMS addressing and routing functions addressed by specifications should mainly concern the IP layer and the Satellite technology Independent (SI) protocol layer which serve to provide common solutions and interfaces for interworking with IP networks.

These recommendations are based on the assumptions for BSMS networking scenarios and solutions to addressing and routing described mainly in clause 7.

The recommendations focus on the need for addressing and routing management functions to interface the BSM world to the Internet world (on behalf of STs). This manager (or proxy), based on a server-client model and whose architecture is fairly standard for middleware, uses a specific number of "modules" to implement routing and addressing tasks. It could also be extended to include new modules as routing technology evolves to manage for example quality of service and performance management that are closely related to routing.

In figure 9.1 the structure of the recommended TSs is presented. The rest of this clause details each block of figure 9.1.

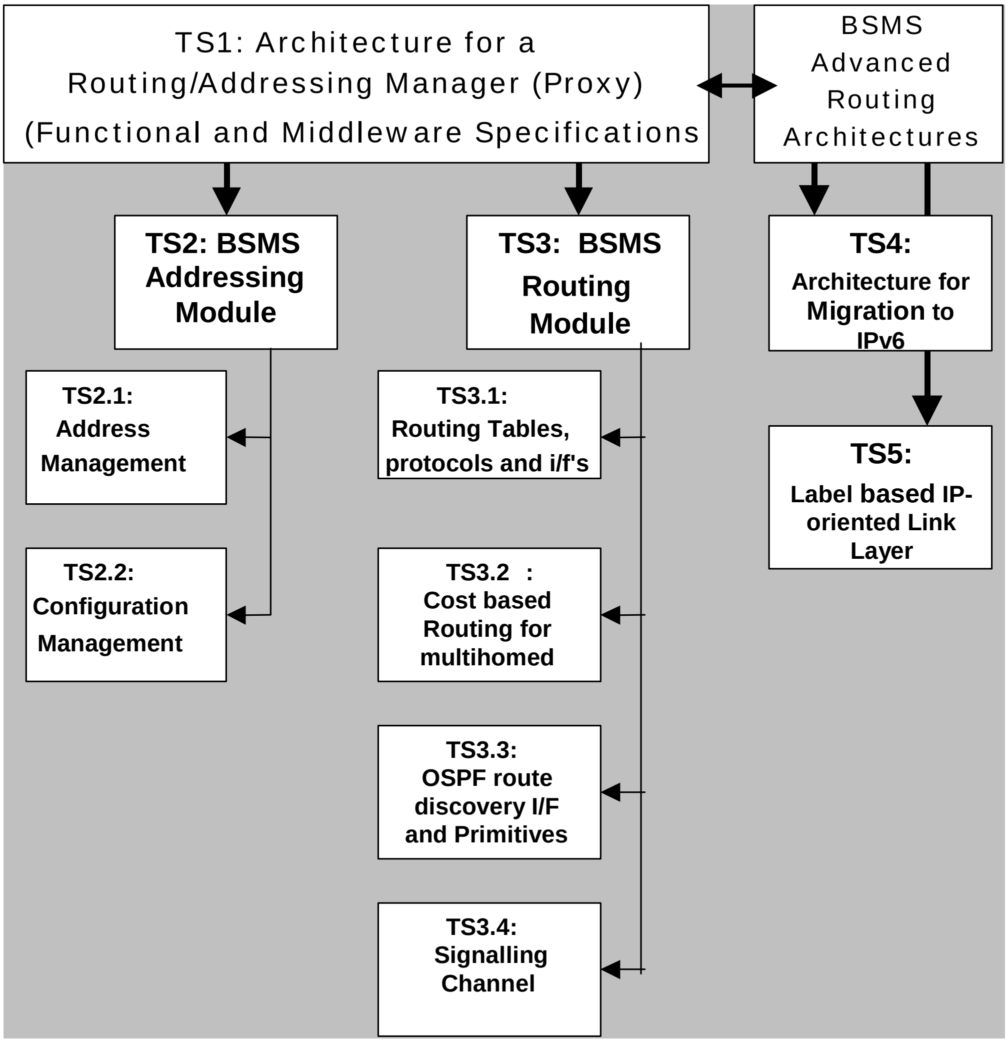

TS1: Architecture for a Routing/Addressing Manager (Proxy) (Functional and Middleware Specifications
BSMS Advanced Routing Architectures
TS2: BSMS Addressing Module
TS3: BSMS Routing Module
TS4: Architecture for Migration to IPv6
TS2.1: Address Management
TS3.1: Routing Tables, protocols and i/f's
TS5: Label based IP-oriented Link Layer
TS2.2: Configuration Management
TS3.2 : Cost based Routing for multihomed
TS3.3: OSPF route discovery I/F and Primitives
TS3.4: Signalling Channel

Figure 9.1: Structure of recommended TSs

9.1 TS1: Routing/addressing management architecture

9.1.1 Aim

An optimized generic BSMS architecture for **routing protocol exchange** has been proposed in clause 7.2. This involves centralized servers, the BRS and BARS (for the Control Plane) which may be associated with the NCC . The architecture covers:

- Routing protocols (BGP, IGP).
- Address resolution (NHRP, LDP, S-ARP, etc.).
- Layer 2 call/connection control.
- Partitioned Route Server for different ASs and OSPF Areas, etc.
- Coordination of roles between Servers (BRS, BARS, NCC, etc.).

The aim of the architecture is to reduce the overhead due to exchange of signalling and routing protocols between STs and end systems, by use of a star network instead of a mesh for this messaging. It also allows grouping of the usually distinct functions related to addressing and routing in the BSM under a single management and configuration entity. This allows this manager to take advantage of functional commonalities between tasks and also to inherit from and interwork with similar functions in other technologies (DSLAM in DSL, head-end in cable networks etc.), software architectures (Open System Gateway Initiative - OSGi - for example) and other ETSI programs (TIPHON, 3GPP). The TS will describe the functional architecture of the server/client model as well as its architecture in middleware.

9.1.2 Requirements

- Based on open standards for addressing and routing protocols, network service control.
- build on BMSS SI-SAP interface primitives and define additional functionality of the SIAF.
- The BSM families should be supported as well as other ETSI standards such as EN 301 790 [18].

9.2 TS2: Addressing modules

TS2 concerns two specific service management aspects within the BSMS proxy architecture, as described below.

9.2.1 TS2.1: Address management SI-SAP interface

9.2.1.1 Aim

Address management includes the allocation and management of BSMS link layer and network layer addresses, and mapping between these addresses (see clauses 7.2.5 and 7.4). The aim is to define a common Address Resolution interface between these layers (at the SI-C-SAP, see clause 4.2) for alternative Link/Shim layers (DHCP, S-ARP, MPLS LDP, DVB-RCS PID/MPE) e.g. based on request from hosts, snooping by hosts and table broadcast by the server.

A robust, secure and link layer independent interface is also needed to support the IP layer address management. This interface should also minimize the need for IP layer signalling and maximize the performance if additional link layer information can be used to aid the IP layer decisions.

NOTE: This specification would address the BSMS-side ST interface; the terrestrial-side ST interface is separately configurable according to the type of user interface.

9.2.1.2 Requirements

Definition of a common address resolution function in the SIAF layer (see figure 4.2.1) of STs and the BARS.

Definition of primitives at the SI-C-SAP interface, for example:

- Provide a static or dynamic address, or explicit assignment procedure.
- Allow upper layers to query the link layer address from the interface.
- Support both point-point and broadcast/multicast link layer interfaces.
- Map IP multicast addresses to link layer addresses.
- Join/leave multicast group.

Also definition of a generic BSMS link layer address format for different link layers for use by the network layer is needed.

9.2.2 TS2.2: Profile management SI-SAP interface

9.2.2.1 Aim

The SI-SAP Interface aims to be generic enough to be applicable to different SD link layers (as indicated in clause 7.4), yet detailed enough to preclude the need at upper layers to use any functionality or information that is specific to a particular SD layer. An SD layer should therefore advertise its capabilities through a profile management function, allowing the higher layers to adjust to the circumstances of the link layer, and to set link layer parameters.

The Profile Management Interface is thus needed for managing the general and default mode of operation, and particularly address management, multicast/broadcast and also general interface functions. It is applicable particularly during ST log-on or session initialization, and will be a quasi-static procedure.

9.2.2.2 Requirements

The SI-SAP configuration primitives should allow retrieval of link layer capabilities and getting/setting modes of operation and operational parameters. These capability parameters include link layer support for and setting of, for example:

- SD medium type (ATM, DVB (-RCS), MPLS, IP-oriented, etc.).
- SD capabilities (SLC/SMAC/PHY functionality) - intended to include all lower layer functions such as encryption, compression, acknowledged delivery.
- Access type: Connection-oriented/connectionless.
- Connection status.
- Network ID/Domain.
- Address auto-configuration support.
- Optimized Address resolution support.
- Multicast/broadcast support.
- Authentication support.
- Idle mode - link layer paging protocol.

9.3 TS3: Routing

TS3 concerns four specific service management aspects within the BSMS proxy architecture, as described below.

9.3.1 TS3.1: BSMS-specific routing tables, protocols and interfaces

9.3.1.1 Aim

A centralized BSMS routing protocol architecture is proposed (see clause 7.2) to optimize routing overhead and coordination across the system. The internal routing protocols present an external interface for IP routing protocols, but internally are based on a star configuration (client-server) in order to minimize routing message exchanges, with centralized server(s) (BRS) and clients in the STs.

The BRS implements routing protocols and issues forwarding tables to STs (and the OBP) which perform IP forwarding only. The STs forward IP routing protocol packets to the BRS. These will rely on specific BSMS routing tables.

9.3.1.2 Requirements

1) Define BSMS routing table entries based on current routing tables.

2) Define functions and messages in the Control Plane of STs (the SIAF layer and above) and of the BRS to perform reliable and secure distribution of ST forwarding tables. General forwarding table updates should be based on broadcast messages and should be issued only when necessary. Ad Hoc ST requests for table download are also needed. (at session set up, etc.).

3) Define generic Static Routing tables based on the above dynamic routing tables, with possibly simplified entries.

9.3.2 TS3.2: "Cost-based" routing in a multi-homed BSMS

9.3.2.1 Aim

An ST may reach a destination through one or more STs attached to the Internet. The "cost" of routing to ST3 or ST4 may change over time depending on the air interface status between the satellite and the STs, the congestion on-board or the congestion in the egress STs. Thus the forwarding table in ST1 (and/or the OBP/OBC, depending on the BSMS routing architecture) may evolve. BSMS functions (see clause 7.2.2.1) are needed to evaluate cost metrics of the routes for choice of the "best" route to the destination host. These metrics can then be included in BSMS routing tables.

A cost metric can be based on RNAP [15] to evaluate OSPF-type link states for fully meshed networks with multiple connectivity. The COPS protocol (Common Open Policy Service – IETF RFC 2748 [85]) could also be considered as a basis for the protocol.

9.3.2.2 Requirements

Define how often the cost metrics need to be refreshed in order to allocate the computation of the routing table between OBC, BRS and ST.

Definition of a BSMS cost-based routing matrix based on conventional IP routing tables and on "Link State" methods. A set of primitives (e.g. at the SI-C-SAP interface) will be defined for these.

ST functions need to be defined for cost metrics of routes depending on:

- Level of queue congestion of the main route downlink (if queuing by downlink).
- Level of queue congestion on the destination queue (if destination based queuing).
- Link QoS due to weather at destination.
- "Real" cost in money sense of going to a certain ST.
- Destination ST on the same administrative domain as the source ST.
- Destination host on the same administrative domain and the source ST/host.
- Other policy based decisions.

BRS functions need to be defined depending on:

- Computation of routing tables (cost matrix) that involve multidimensional metrics.
- Semi-static/global routing table computation.
- Admission control and policy setting.
- Determination of alternate routes and default routes.
- Computation of long range statistics that can define main and alternate routes for source-destination pairs.
- Recording of fault events.

- Performance and QoS monitoring.
- Other non real time management/FCAPS functions.

9.3.3 TS3.3: Route discovery interface and primitives

9.3.3.1 Aim

In OSPF and similar protocols. (i.e. EGP, BGP, etc.) it is essential that every router on a common network enables 2-way communications and keeps its database synchronized (see clause 7.2.2.3).

Over a BSMS the Hello packet from an upstream router will reach the ST as a UDP packet. This packet is sent as high priority because of its fundamental role in the setting of network topology. If there is no available bandwidth at this point the BOD process in the ST will request and get the appropriate bandwidth and send the packet to its destination(s). The main issue concerns timers. The hello timer values can be configured, though they should be consistent across all routers on a network segment.

The Hello protocol timers are set not only over the BSM but are shared on all the routers attached to the common network. Hence the operator of a network with an OSPF IGP should set these parameters to ensure the BSMS fully participates in routing (see clause 5.2.4.2).

Hence to be transparent to the OSPFv2 Hello protocol, for example, the BSM should ensure that the functionalities of the protocol are preserved over the BSMS.

9.3.3.2 Requirements

1) Definition of suitable timer values for the BSMS.
2) Definition of primitives to support route discovery.

9.3.4 TS3.4: Dedicated BSMS signalling channels for routing, etc.

9.3.4.1 Aim

Two of the main problems in a BSMS as regards route discovery and related topics (address resolution, labelling paths in MPLS, RSVP reservation etc.) are the potential limitations in bandwidth at the BSMS ingress terminal and the end-to-end delay (see clause 7.2.2.4). Due to these factors signalling messages may time out, delaying the whole network and possibly triggering congestion, or worse making the BSMS virtually unavailable at the network layer.

Solutions are:

1) Use of a terminal signalling channel.
2) Dedicated IP low-bandwidth signalling channel.
3) IP signalling pre-emption class.

The Pre-emption class of message is the preferred option (which could be based on IETF RFC 3181 [86]).

NOTE: This task could be combined with the relevant QoS work.

9.3.4.2 Requirements

Definition of signalling class function and SI-C-SAP primitives.

The signalling class should allow IP signalling messages (usually short) to be transmitted a soon as they reach the ingress of the BSMS. As soon as an "important" message is identified it pre-empts any existing queue and leave in the next available frame (or time slot or transmission opportunity).

9.3.5 TS4: Strategies for IPv6 service transition from IPv4

This TS is intended to extend and expand on routing and networking architectures addressed in TS1.

9.3.5.1 Aim

The BSMS may be chosen to be either IPv4 or IPv6 based, which has implications on the interworking mechanisms to be included in the BSMS. An IPv4-based BSMS is a natural short-term solution which could be adapted with interworking units at a few specific interfaces to allow for the IPv6 network interconnection scenario. An IPv6-based BSMS is a longer term solution which could be adapted in a similar but inverse way to allow for legacy IPv4 network interconnections (see clause 5.4).

9.3.5.2 Requirements

Definition of BSMS functional architecture to support addressing and routing transition scenarios.

Two main aspects may be considered of most immediate importance for strategic and fast adoption of IPv6 (Transition Mechanisms for IPv6 Hosts and Routers - IETF RFC 2893):

1) Interconnection of IPv6 islands through an IPv4 network (BSM and terrestrial). Solutions are generally based on dual stack routers and IPv6 in IPv4 tunnels.

2) Communication and interoperability of IPv6 nodes with IPv4 nodes. Mechanisms rely on dual stack techniques, application level gateways, NAT technology or on temporary allocation of IPv4 address and IPv4 in IPv6 tunnelling.

Routing aspects for IPv6 transition scenarios include:

1) Routing for IPv4 packets (over IPv6).

2) Routing for IPv6 packets (over IPv4):

   - IPv6 packets with IPv6-native addresses;

   - IPv6 packets with IPv4-compatible addresses.

3) Operation of manually configured (static) tunnels.

4) Operation of automatic encapsulation:

   - Locating encapsulation;

   - Ensuring that routing is consistent with encapsulation.

For the BSMS, tunnelling is the most inefficient transport mechanism due to the additional encapsulation overhead, and should be avoided on cost grounds in favour of translation.

9.3.6 TS5: Label-based IP-oriented link layer

9.3.6.1 Aim

A satellite-oriented solution to the underlying protocol to the IP layer in the BSMS is to employ "label paths" acting as logical broadcast networks. This is a simpler scheme compared to MPLS, and better adapted to satellites by taking advantage of their multicast attributes (see clause 7.4.2.2). Such logical satellite links would be able to support both CO and CL traffic directly: such a "label path" is itself connection-oriented but in a multipoint-to-multipoint configuration STs could use it as connectionless or connection-oriented transport (e.g. by resource reservation) between each other.

9.3.6.2 Requirements

Definition of IP-oriented link layer:

- Labelling concepts and label structure.
- Satellite ARP for dynamic resolution.
- Mapping of labels onto standardized satellite access layer identifiers (PID, VPI/VCI).
- Segmentation and re-assembly functions (when a shared label is used to filter traffic at reception, a function is needed to identify the source of the traffic and therefore perform re-assembly per source).
- QoS support: interaction with IP based (e.g. RSVP) or session based signalling (e.g. SIP).

Annex A: BGP and IPv6 related RFCs

A.1 List of BGP related RFCs

IETF RFC 1265: "BGP Protocol Analysis".

IETF RFC 1266: "Experience with the BGP Protocol".

IETF RFC 1267: "A Border Gateway Protocol 3 (BGP-3)".

IETF RFC 1269: "Definitions of Managed Objects for the Border Gateway Protocol (Version 3)".

IETF RFC 1397: "Default Route Advertisement In BGP2 And BGP3 Versions Of The Border Gateway Protocol".

IETF RFC 1403: "BGP OSPF Interaction".

IETF RFC 1657: "Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2".

IETF RFC 1745: "BGP4/IDRP for IP-OSPF Interaction".

IETF RFC 1771: "A Border Gateway Protocol 4 (BGP-4)".

IETF RFC 1773: "Experience with the BGP-4 protocol".

IETF RFC 1774: "BGP-4 Protocol Analysis".

IETF RFC 1863: "A BGP/IDRP Route Server alternative to a full mesh routing".

IETF RFC 1930: "Guidelines for creation, selection, and registration of an Autonomous System (AS)".

IETF RFC 1965: "Autonomous System Confederations for BGP".

IETF RFC 1966: "BGP Route Reflection An alternative to full mesh IBGP".

IETF RFC 1998: "An Application of the BGP Community Attribute in Multi-home Routing".

IETF RFC 1997: "BGP Communities Attribute".

IETF RFC 2270: "Using a Dedicated AS for Sites Homed to a Single Provider".

IETF RFC 2385: "Protection of BGP Sessions via the TCP MD5 Signature Option".

IETF RFC 2439: "BGP Route Flap Damping".

IETF RFC 2519: "A Framework for Inter-Domain Route Aggregation".

IETF RFC 2545: "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing".

IETF RFC 2796: "BGP Route Reflection An alternative to full mesh IBGP".

IETF RFC 2842: "Capabilities Advertisement with BGP-4".

IETF RFC 2858: "Multiprotocol Extensions for BGP-4".

IETF RFC 2918: "Route Refresh Capability for BGP-4".

IETF RFC 3078: "Microsoft Point-to-Point Encryption protocol".

A.2 List of IPv6 related RFCs

IETF RFC 2185: "Routing Aspects Of IPv6 Transition".

IETF RFC 2460: "Internet Protocol, Version 6 (IPv6) Specification".

IETF RFC 2461: "Neighbour Discovery for IP Version 6 (IPv6)".

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

IETF RFC 2529: "Transmission of IPv6 over IPv4 Domains without Explicit Tunnels".

IETF RFC 2765: "Stateless IP/ICMP Translation Algorithm (SIIT)".

IETF RFC 2766: "Network Address Translation - Protocol Translation (NAT-PT)".

IETF RFC 2893: "Transition Mechanisms for IPv6 Hosts and Routers".

IETF RFC 3053: "IPv6 Tunnel Broker".

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

IETF RFC 3068: "An Anycast Prefix for 6to4 Relay Routers".

History

Document history		
V1.1.1	April 2003	Publication