



Technical Specification

**IMS Network Testing (INT);
IMS/NGN Performance Benchmark;
Part 1: Core Concepts**

Reference

RTS/INT-0090-1

Keywords

IMS, performance, service

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, please send your comment to one of the following services:

http://portal.etsi.org/chaicor/ETSI_support.asp

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2013.
All rights reserved.

DECT™, PLUGTESTS™, UMTS™ and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and LTE™ are Trade Marks of ETSI registered for the benefit of its Members and
of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Contents

Intellectual Property Rights	4
Foreword.....	4
Introduction	4
1 Scope	7
2 References	7
2.1 Normative references	7
2.2 Informative references.....	7
3 Definitions and abbreviations.....	8
3.1 Definitions	8
3.2 Abbreviations	9
4 Benchmark information model.....	10
4.1 Use-case	11
4.1.1 Call Flow	11
4.1.2 Load Profile	11
4.1.3 Metrics	12
4.1.4 Use-case outcomes.....	12
4.1.5 Scenarios and scenario attempts	12
4.1.6 Design Objective (DO)	12
4.1.7 Scenario	12
4.2 Benchmark test	12
4.2.1 Traffic set.....	12
4.2.2 Background load.....	13
4.2.3 Traffic-time profile	13
4.2.4 Test parameters	13
4.3 Benchmark report	14
5 System Under Test (SUT)	15
5.1 Creation of subscriber data base.....	19
6 Test system.....	19
7 Benchmark metrics examples.....	19
History	21

Intellectual Property Rights

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

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

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee IMS Network Testing (INT).

The present document is part 1 of a multi-part deliverable covering the IMS/NGN Performance Benchmark, as identified below:

- Part 1: "Core Concepts";**
- Part 2: "Subsystem Configurations and Benchmarks";
- Part 3: "Traffic Sets and Traffic Profiles";
- Part 4: "Reference Load network quality parameters".

Introduction

A major change is occurring in the telecommunications domain. Telecom Service Providers (SPs) are quickly evolving their networks from legacy technologies to what might be termed "fourth generation" or "3G Beyond" technologies. The principal features of this technological change are:

- Evolution of "traditional" wireline telecom standards to Voice over IP (VoIP) standards, pre-eminently the cluster of protocols surrounding SIP.
- Evolution of GSM and CDMA networks to 3GPP/3GPP2 standards, such as UMTS.
- Evolution from the UMTS-Technology (3G-Standard) to the [LTE-Advanced](#) standard (4G-Standard) based on the [GSM/EDGE](#) and [UMTS/HSPA](#).
- Introduction of wireless LAN (WLAN) standards, pre-eminently IEEE 802.16 [i.9], for both data and voice communications.
- Fixed-mobile convergence through the various access technologies standardized by 3GPP and TISPAN.
- Potential and actual competition among standards, causing in reaction an attempt at service convergence.

The current convergence point of these trends is a set of technologies termed the IP Multimedia Subsystem (IMS). IMS supports a rich set of services available to end users on either wireless or wired user equipment (UE), provided via a uniform interface by the cooperation between a subscriber's home service provider (SP) and the visited service providers. Services are provided via an "overlay" technique over multiple service provider networks.

Telecom Equipment Manufacturers (TEMs) all along the architectural hierarchy, assuming that IMS represents a growth market, are attempting to develop not merely products for IMS networks, but architectures as well. The request for new architectures represents a view that current processors, network, and server architectures are not sufficient to support wide IMS deployment. Examples of such work include:

- Advanced Telecom Architecture (ATCA), a set of existing and emerging standards for the physical packaging of bladed servers and communications fabrics.
- Fabric standards, such as Infiniband, PCI Express, RapidIO and Gigabit Ethernet (e.g. IEEE 802.1 [i.10] and IEEE 802.3 [i.12] Working Group).
- Technologies for offloading packet processing, such as network processors and other multicore processors.
- Real-time and carrier-grade operating systems, such as the carrier-grade Linux releases of Montavista and Red Hat.
- Middleware architected for high availability, such as defined by the Object Management Group (OMG) and the Service Availability Forum (SAF).
- Application development environments and interfaces, such as defined by the Parlay Group and the Java Community Process (e.g. JAIN).

The number of technological variables is so large that some reasonable ground rules for defining an architecture need to be put in place. SPs require guidance for making decisions among suppliers, and suppliers all along the architectural hierarchy need guidance to develop the right products.

Motivating Example

Figure 1 depicts in a single diagram the major features of the IMS performance benchmark, and will serve as a motivating example. The System Under Test (SUT) is presented with a workload by the Test System. The workload consists of the traffic generated by a large number of individual simulated User Endpoints (UEs), each performing an individual scenario. The collection of scenarios selected for a benchmark test is a traffic set.

The rate at which scenarios from the traffic set are attempted in the benchmark test is governed by a traffic-time profile defined for the benchmark. A traffic-time profile is designed so that the rate of scenario attempts remains constant for sufficient time to collect a statistically significant data set. For example, in the figure, the time between 0 minute and 10 minutes shows the traffic rate ramping up to approximately 120 session attempts per second (SAPS), after which it remains in a steady-state phase for approximately 30 minutes, at an average value of 120 SAPS.

During the test, each scenario attempt (a generalization of "call attempt", accounting for the fact that a scenario may be a registration scenario or an instant messaging scenario) may succeed, may fail, or may succeed but exceed the time threshold determined to be the "design objective" for the scenario. During the steady state phases of a traffic-time profile, the percentage of failures or timeouts, collectively called "inadequately handled scenario attempts", is measured. If it exceeds a certain frequency, which may also be interpreted as a probability of inadequately handled scenario attempt, the SUT has reached its Design Objective Capacity (DOC), the primary comparison metric reported by the benchmark. In figure 1, the DOC is reached at 130 SAPS.

When the traffic-time profile changes to a new average value, the percentage of inadequately handled scenario attempts will in general exhibit a transient spike. The average successful scenario rate and the percentage of inadequately handled scenario attempts is averaged over the steady-state phase excluding the transient.

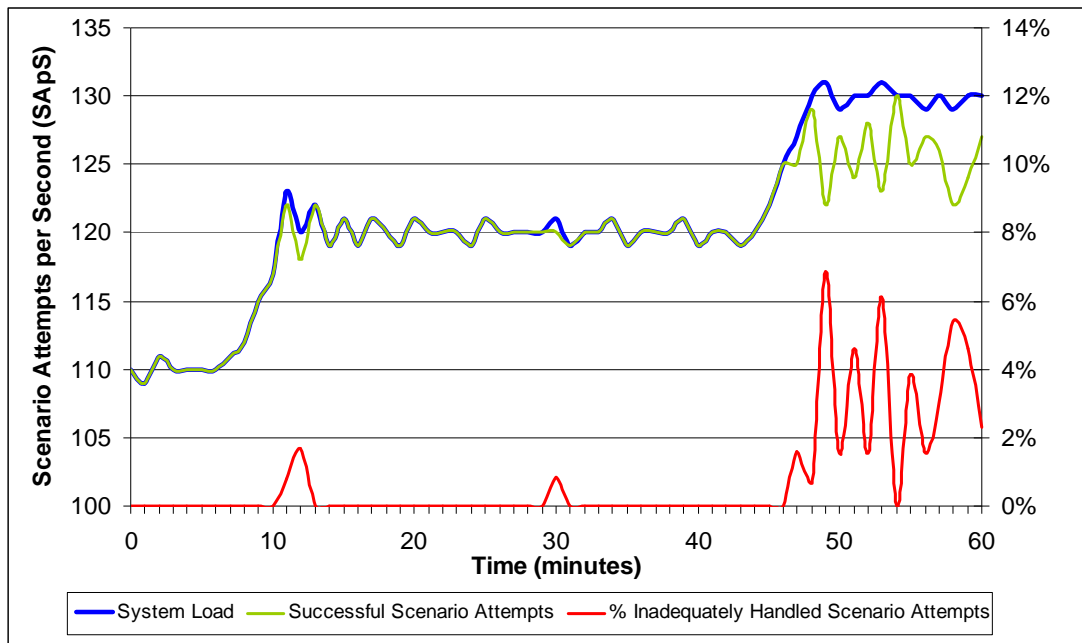


Figure 1: Motivating example of an IMS benchmark

1 Scope

The present document describes the performance benchmark methodology for the IMS based services MMTel, Video Telephony and IMS - PES. The terminology and concepts are described in TR 101 577 [i.11]. The present document is the first part of the multi-part deliverable which consists of four parts.

The present document contains the overall benchmark descriptions, architectures, processes, and information models that are common to all specific benchmarking scenarios.

TS 186 008-2 [i.1] contains the specific benchmarking use-cases and scenarios, along with scenario specific metrics and design objectives. It also defines the SUT configuration parameters. This part also contains any required extensions to the overall descriptions present in the present document, if necessary for the specific scenario.

TS 186 008-3 [i.2] defines an initial benchmark test through the specification of a traffic set, traffic-time profile and benchmark test procedure.

TS 186 008-4 [i.3] defines Reference Load network quality parameters for the use cases defined in TS 186 008-2 [i.1].

2 References

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

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

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

2.1 Normative references

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

- [1] ETSI TS 101 563 (V1.2.1): "Speech and multimedia Transmission Quality (STQ); IMS/PES exchange performance requirements".
- [2] ETSI TS 123 002 (V11.4.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Network architecture (3GPP TS 23.002 version 11.4.0 Release 11)".

2.2 Informative references

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

- [i.1] ETSI TS 186 008-2: "IMS Network Testing (INT); IMS/NGN Performance Benchmark; Part 2: Subsystem Configurations and Benchmarks".
- [i.2] ETSI TS 186 008-3: "IMS Network Testing (INT); IMS/NGN Performance Benchmark; Part 3: Traffic Sets and Traffic Profiles".
- [i.3] ETSI TS 186 008-4: "IMS Network Testing (INT); IMS/NGN Performance Benchmark; Part 4: Reference Load network quality parameters".
- [i.4] Recommendation ITU-T P.862: "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs".

- [i.5] Recommendation ITU-T P.862.1: "Mapping function for transforming P.862 raw result scores to MOS-LQO".
- [i.6] Recommendation ITU-T P.56: "Objective measurement of active speech level".
- [i.7] ETSI TR 121 905: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Vocabulary for 3GPP Specifications (3GPP TR 21.905)".
- [i.8] Recommendation ITU-T P.863 (01/2011): "Perceptual objective listening quality assessment".
- [i.9] IEEE 802.16: "A Technical Overview of the WirelessMAN™ Air Interface for Broadband Wireless Access".
- [i.10] IEEE 802.1, 802.1Q-2011: "IEEE Standard for Local and metropolitan area networks--Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks".
- [i.11] ETSI TR 101 577 (V1.1.1): "Methods for Testing and Specifications (MTS); Performance Testing of Distributed Systems; Concepts and Terminology".
- [i.12] IEEE 802.3: "Ethernet Working Group".

3 Definitions and abbreviations

3.1 Definitions

The IMS benchmarking definitions are described in TR 101 577 [i.11].

For the purposes of the present document, the following terms and definitions apply:

background load: workload applied to an SUT during a benchmark test, for the purpose of consuming SUT resources during a benchmark test and changing the traffic intensity at which the capacity of the SUT is reached

benchmark report: document generated at the conclusion of a test procedure containing the metrics measured during the execution of the test and/or computed from the data collected in the benchmark log

benchmark test: procedure by which a test system interacts with a System Under Test to measure its behaviour and produce a benchmark report

configuration: specification of a subset of IMS/PES architectural elements and metrics for which collection of benchmark tests can be defined

design objective: probabilistic model of delay and failure requirements for SUT, associated with a use-case, specified by threshold values and probabilities for delay and scenario failure

Design Objective Capacity (DOC): largest load an SUT can sustain while not exceeding design objectives defined for a use-case

idle load: load that is not dependent on the traffic or other external activities

maximum capacity: maximum processor load that a processor can handle without rejecting new calls

metric: performance measurement of SUT reported in a benchmark report

parameter: attribute of a SUT, test system, system load, or traffic set whose value is set externally and prior to a benchmark test, and whose value affects the behaviour of the benchmark test

processor load: part of time the processor executes work, normally expressed in percent

NOTE: The processor load consists of Idle load, Traffic load and Usage load.

Reference Call (RC): basic ISUP to ISUP call connected through two MGW in the same MGC domain

SAPS increase amount: increment by which the average SAPS changes between steps of a profile

test parameters: parameters whose values determine the behaviour of a benchmark test

test procedure: specification of the steps to be performed by a benchmark test

test scenario: specific path through a use-case, whose implementation by a test system creates a system load

test system: collection of hardware and software which presents a system load to a system under test and collects data on the system under test's performance, from which metrics can be computed

traffic load: load that results from handling traffic events that are directly related to calls

NOTE: This load varies with the traffic intensity.

traffic-time profile: evolution of the average scenario over a time interval

traffic set: mixture of traffic scenarios

usage load: load that is reserved for the administrations operation and maintenance activities during busy hour

workload: number of reference calls per second (RC/s)

NOTE: It is calculated by multiplying calls per second by its corresponding WLF.

workload factor (WLF): traffic load for different types of calls in relation to the traffic load of the reference call (ISUP call)

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in TR 121 905 [i.7] and the following apply:

AGCF	Access Gateway Control Function
ATCA	Advanced Telecom Architecture
BYE	BYE message
CDMA	Code Division Multiple Access
CN	Core Network
CPU	Central Processor Unit
CS	Circuit Switched
DO	Design Objective
DOC	Design Objective Capacity
ETH	Ethernet
GSM	Global System for Mobile Communications
II NNI	Inter-IMS Network to Network Interface
IM	IPMultimedia
IMS	IP Multimedia Subsystem
IP	Internet Protocol
ISC	S-CSCF – Application Server (ISC) interface
ISDN	Integrated Service Digital Network
ISUP	ISDN User Part
ITU-T	ITU's Telecommunication Standardization Sector
JAIN	Java APIs for Integrated Networks
LAN	Local Area Network
MGW	Media Gateway
MHT	Mean Holding Time
MMTel	Multi-MediaTelephony
NGN	Next Generation Networks
NNI	Network Network Interface
OMG	Object Management Group
PCI	Peripheral Component Interconnect Express
PES	PSTN Emulation Solution
PESQ	PESQ, Perceptual Evaluation of Speech Quality
POLQA	Perceptual Objective Listening Quality Assessment
POTS	Plain old telephone service

RC	Reference Call
SAF	Service Availability Forum
SAPS	Session Attempts Per Second
SIGTRAN	Stream Control Transmission Protocol
SIP	Session Initiation Protocol
SIP-I	SIP-I define a mapping from SIP to ISUP
SP	Service Provider
SUA	SCCP-User Adaptation Layer
SUT	System Under Test
TA	Tones and Announcement
TDM	Time-division multiplexing
TEMs	Telecom Equipment Manufacturers
TR	Technical Report
UDI	Unrestricted Digital Information
UDI/TA	Unrestricted Digital Information with tones and announcements
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
WLAN	Wireless Local Area Network
WLF	Work load factor

4 Benchmark information model

In this clause, "benchmark information model" refers to the structure of the information elements that define the benchmark. This information model is depicted in figure 2.

The information model consists of three primary elements: **use-cases**, which describe the behaviour of an individual user, and which in turn define scenarios; **benchmark tests**, which generate a workload by aggregating the behaviour of individual scenarios in a controlled manner, and collect log files of measurements during the test; and **benchmark test reports**, which report metrics interpreted from the benchmark test log files.

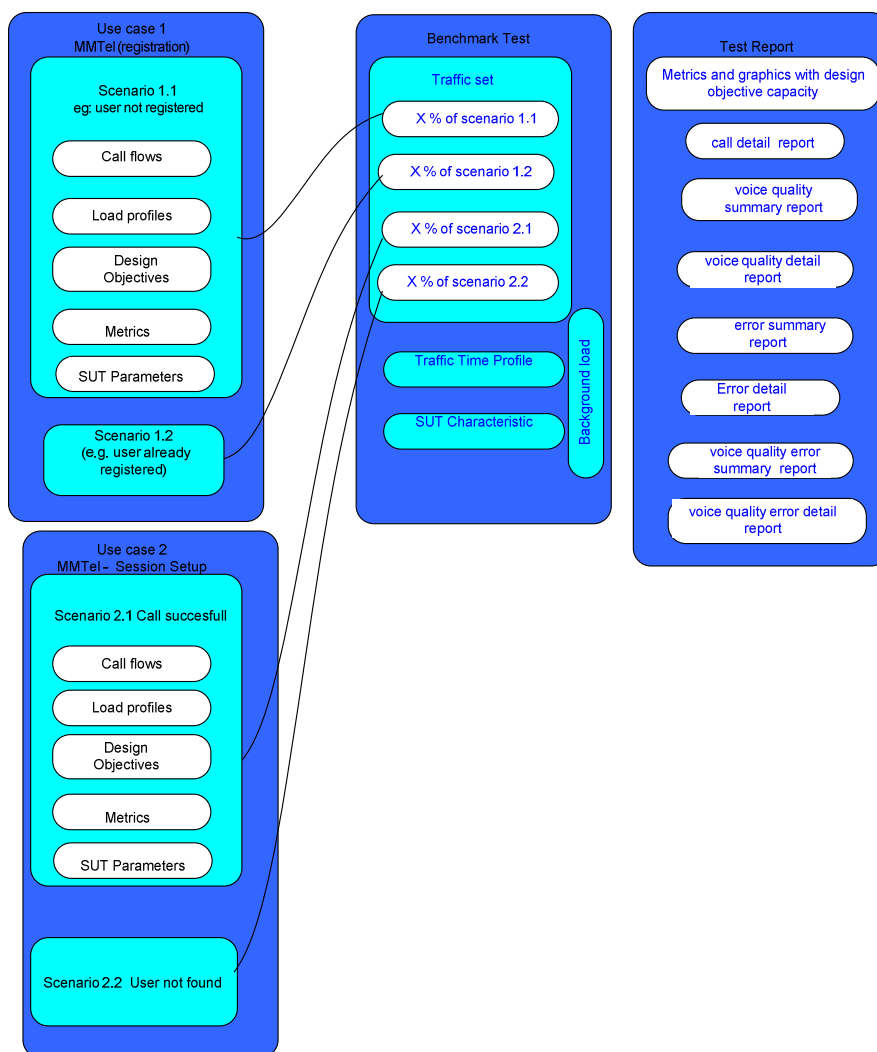


Figure 2: IMS benchmark information model

4.1 Use-case

The top level of the individual behavioural model is the *use-case*. A use-case describes the goal that a user has in interacting with a system, the various actors (e.g. other users, network elements) that participate in the use-case, the basic course of events that are carried out by the user and the SUT, the design objective of the use-case, the possible outcomes that apply to the use-case, and the metrics to be collected. The goal and actors of a use-case are documented in narrative text and diagrams; the other elements are complex information elements, which are described in their respective clauses.

4.1.1 Call Flow

The calls flows define the characteristic message flows, the tones and announcement for a specific interface.

4.1.2 Load Profile

To facilitate the calculation of processing capacity and the appropriate load profile the concept of workload factor has been defined based on the reference call for each combination of traffic case and traffic signalling interface. The reference call (RC) is defined as a basic ISUP to ISUP call connected through two MGW in the same domain.

Based on the workload factors for all different types of calls, the call intensities and the services used, one can express the total traffic load in an equivalent number of reference calls per second.

The dimensioning of any type of network depends on a number of different parameters such as utilization per channel, calls per second, mean holding time, type of accesses being involved, and type of services being requested.

4.1.3 Metrics

The metrics of a use-case describe the measurements collected from the execution of a scenario attempt. Typical metrics include response times and message rates. If a scenario is selected for execution in a benchmark test, its metrics are collected. See clause 7 for more detail.

4.1.4 Use-case outcomes

A use-case outcome is a set of possible outcomes of the scenarios of a use-case. An outcome may be simply "correct", it may reflect an error or failure condition; or it may reflect a correct behaviour that took an excessive amount of time to occur. An instance of a scenario that experiences an error, failure, or timeout outcome is referred to as an inadequately handled scenario attempt.

4.1.5 Scenarios and scenario attempts

A scenario is a trace of a path through a use-case. It is analogous to "call attempt", but applies to all interactions within an IMS/PES network, different Bearer, and application interactions.

A scenario may succeed, fail, or succeed functionally.

The terms "scenario attempt" and "scenario attempts per second" are used in the present document in place of "call attempt" and "call attempts per second" because IMS is a transaction-oriented system with transactions of a variety of types

Traffic sets, and indeed the real world, do not operate according to only one transaction type, so the more generalized term is necessary. It would be incorrect and misleading to attempt to report the capacity of a system in "call attempts per second", "registration attempts per second", etc., for system loads that were other than purely call attempts, registration attempts, etc.

4.1.6 Design Objective (DO)

The Design Objective (DO) describes the maximal acceptable rate handled scenario attempts for a use-case.

4.1.7 Scenario

A scenario describes a single interaction sequence among the actors of a use-case. It is documented by a set of preconditions on its actors (typically specified by parameter values). In case of IMS/PES the scenario is defined as a set of different Bearer Capabilities (e.g. speech, 3,1 kHz audio, UDI, UDI/TA), services (fax, modem) or dial mode.

4.2 Benchmark test

A benchmark by definition measures the behaviour of a population of users. To accomplish this, the behaviours of individual users shall be aggregated into input traffic to the SUT. The input traffic shall be realistic, in the sense that a population of users would perform such actions in the real world, and in the sense that statistical variation in user behaviour is similar to statistical variation that would occur in the real world.

4.2.1 Traffic set

The traffic set is a collection of scenarios which are determined to be likely to co-occur in a real-world scenario. The scenarios do not need to come from the same use-case. Within a traffic set, each scenario has an associated relative occurrence frequency, interpreted as the probability with which it would occur in the course of the test procedure.

4.2.2 Background load

Background load is a workload presented to the SUT in order to consume its resources. It may consist of a stream of traffic presented to the SUT by an external system apart from the test system; or it may be a workload presented to the processing elements, network, or storage subsystem of the SUT.

The purpose of background traffic is to make possible the measurement of a design objective capacity in SUT when the capacity of the test system is insufficient to reach the design objective capacity.

If a benchmark test is published in which background load is used, then the following requirements apply:

- The hardware used to generate the background load shall be fully specified. If the background load is generated by software running directly on the SUT, then the components of the SUT on which the background load is executed shall be fully specified.
- The software used to generate the background load shall be provided in source form, including make files and any other configuration files required to compile the software.

4.2.3 Traffic-time profile

The traffic-time profile is a function describing the average scenario attempt arrival rate as a function of elapsed time during a benchmark test. A traffic-time profile should be chosen in such a manner that, for a given scenario attempt arrival rate, sufficient samples are generated that metrics can be collected with an appropriate confidence bound.

Following Call Profiler Traffic Patterns are used today: Saw Tooth Blast Ramp Steady Call Rate Rolling Blast Poisson Distribution. To get a realistic scenario a combination of at least two scenarios is needed.

4.2.4 Test parameters

The benchmark test parameters are used to control the behaviour of the test script. The data elements required to configure the test system are listed in table 1.

Table 1 is a non-exhaustive list of test parameters defined for the benchmark standard. The list is expected to grow over time, as additional subsystems and system configurations are developed.

Table 1: Test parameters

Parameter	Description
Start time	Amount of time that a system load is presented to a SUT at the start of a test
Stop time	Amount of time that a system load is presented to a SUT at the end of a test
TotalProvisionedSubscribers	The number of simulated subscribers provisioned in the network
PercentSimulatedSubscriber	The average percentage of simulated subscribers
Simulated Maximum simultaneous call legs	The number of simulated Maximum simultaneous call legs
Traffic per subscriber	Traffic per subscriber; default value 0,1 Erlang
PX_PercentRegisteredSubscribers	The average percentage of simulated subscribers that are registered simultaneously.
PX_PercentRoamingSubscribers	The average percentage of simulated subscribers that are roaming (i.e. register in a non-local network).
PX_Simulated Subscriber Registration Parameters	Parameters and distributions of the probabilistic model of simulated subscriber operation
MHT	Mean Holding Time of a call ; default value 110 seconds
Ringing time	Duration between (180 ringing and 200 OK INVITE) Default value 1 s to 5 s
NoS	number of subscribers originating traffic per subscriber
CAPS/BHCA	Call attempts per second/busy hour call attempts
WLF for Call Controller	The workload factor for Call Controller for specific configuration. Default value 1 to 3
WLF for Gateway Controller	The workload factor for Gateway Controller for specific configuration. Default value 1 to 3
WLF for MGW	The workload factor for Media Gateways for specific configuration. Default value 1 to 3
TDM Trunks	Number of TDM trunks
ETH	Number of ETH Connections
Type of call	MMTel to MMTel Video Telephony IMS-PES to IMS-PES MMTel to IMS-PES
Protocol call type and interfaces	SIP-I ISUP SIGTRAN (M2PA; M2UA; M3UA; SUA) SIP II NNI (Ici,Izi) SIP NNI (Mx Interface) Gm Interface
MGCF/MGW/I-BCF/TrMGW performance tests	SIP-I to SIP-I SIP-I to ISUP SIP-I to NNI NNI to NNI SIGTRAN to SIGTRAN (M2PA; M2UA; M3UA; SUA) ISUP (SIGTRAN) to NNI
Transport Interfaces	Voice over LTE (VoLTE) (LTE-Uu, S1-U, S-11, S6a, S11, S5/S8, Rx, Gx, Mw, ISC and Ut interfaces) Ethernet xDSL ISDN POTS (Z)

4.3 Benchmark report

A test report is a document, with accompanying data files, that provides a full description of an execution of a benchmark test on a test system. The SUT and test system, as well as their parameters, are described in sufficient detail that an independent test site can replicate the test. The results of the test include data, represented as charts and data sets, depicting the behaviour of the SUT over the elapsed time of the test; reports of the relevant metrics that are conventionally used to compare benchmark results of differing SUTs; and a full description of other observations and exceptions noted during the test.

5 System Under Test (SUT)

The IMS performance benchmark covers benchmark tests for the MMTel, Video Telephony, and MMTel - IMS PES.

Figure 3 depicts the IMS Reference Architecture. The components of the architecture are the primary building blocks, which are either defined by the IMS standard, or defined by external standards and referenced by IMS. The links between the primary building block represent reference points over which the building blocks communicate with each other.

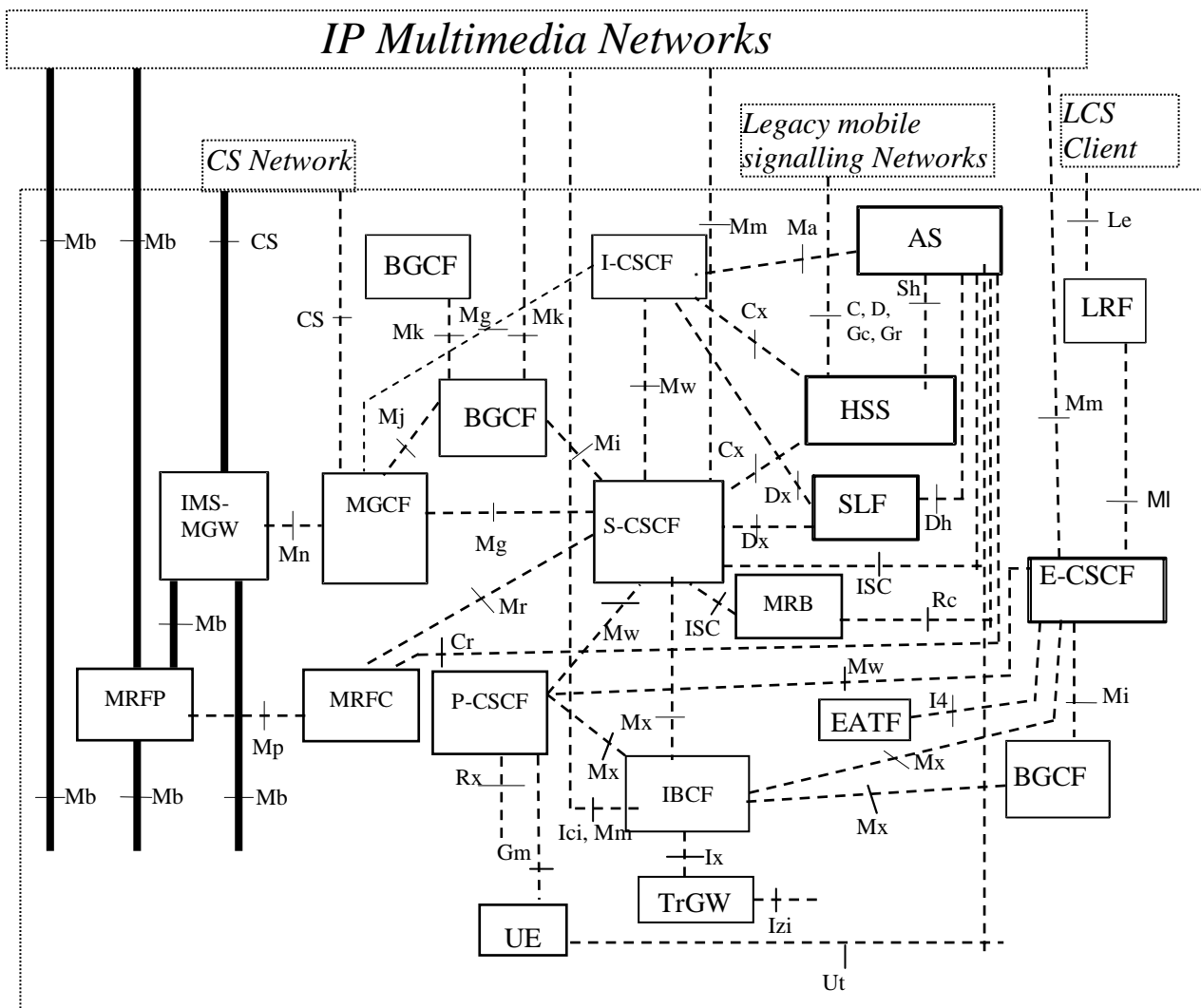
The reference architecture is a logical architecture; no mapping of functional elements to hardware or software component is mandated. And conversely, IMS products as deployed in the real world do not factor neatly into the elements of the reference architecture, which complicates the process of comparing similar products with a benchmark.

The problem can be simplified by observing that there are classes of products which have common subsets of reference architecture elements. For classes defined in this manner, common sets of benchmarks can be defined. The classes defined in this manner are called IMS subsystems, or simply subsystems, and are expected to grow over time, as the IMS marketplace becomes richer.

TS 186 008-2 [i.1] of this multi-part deliverable maintains a list and description of subsystems.

In order to proceed from a subsystem description to a benchmark test, a complete description of all aspects of the subsystem relevant to the performance of the benchmark shall be present; this is referred to as the system configuration, or SUT configuration. This description starts with an enumeration of the elements of the reference architecture, and an enumeration of all reference points that are external to the subsystem (i.e. reference points between elements within the subsystem are "internal"). However, the configuration requires a specification of the hardware elements (e.g. servers, CPUs, network configuration and bandwidth) and software elements (e.g. operating system, database system), because even though the metrics reported by the benchmark tests are measured with respect to such metrics as total round-trip delay of messages and system capacity, the behaviour of IMS systems is still an area of study and a thorough understanding of how CPU and network bandwidth utilization behaves during operation is of interest.

- Figure 4 depicts the Inter-IMS Network to Network Interface between two IM CN subsystem networks.
- Figure 5 depicts the IMS/LTE Basic Configuration.
- Figure 6 depicts the LTE End - to - End Configuration.
- Figure 7 depicts the AGCF/VGW session processing model for IMS PES.



Legend:

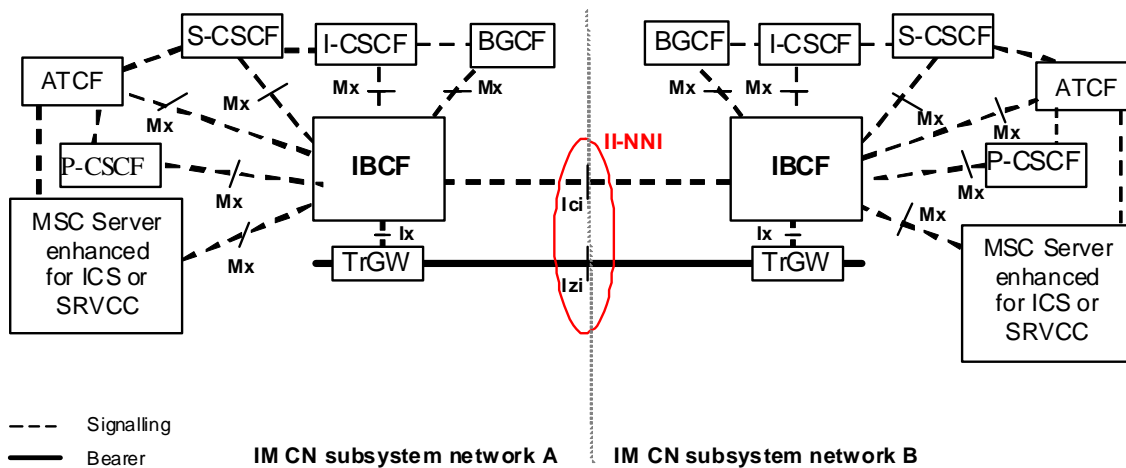
Bold lines: interfaces supporting user traffic;

Dashed lines: interfaces supporting only signalling.

NOTE 1: The reference point CS (Circuit Switched) is not specified in the present document.

NOTE 2: The reference point I5 is not shown in this figure.

Figure 3: Overview of IMS Functional Entities [2]



NOTE: Interface (I-I-NNI) between two IM CN subsystem networks.

Figure 4: Inter-IMS Network to Network Interface between two IM CN subsystem networks

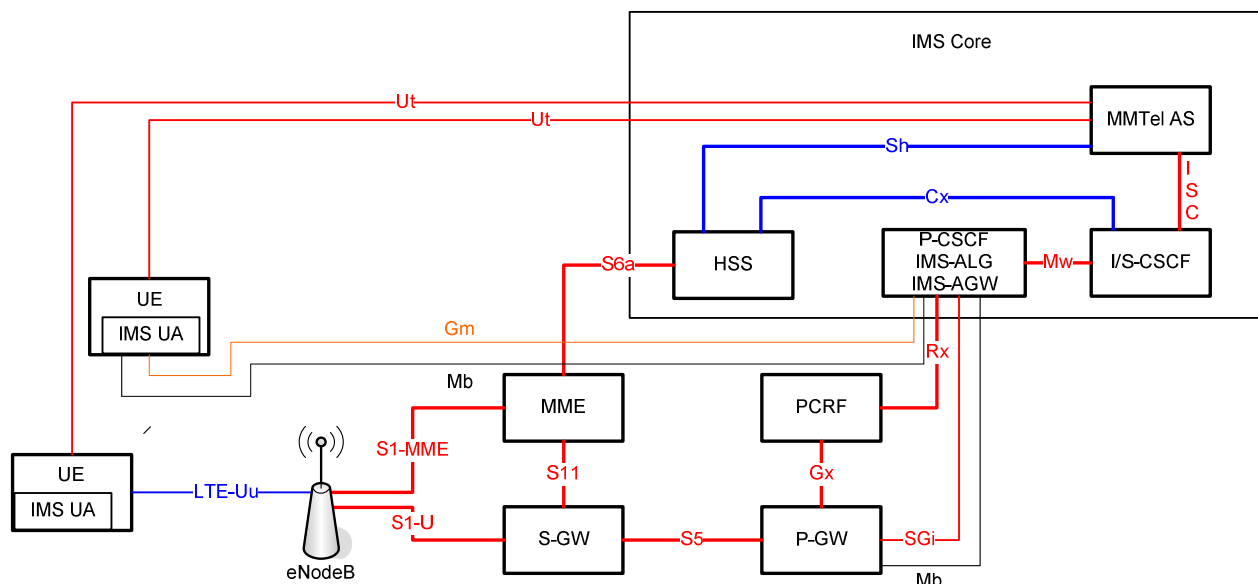


Figure 5: IMS/LTE Basic Configuration

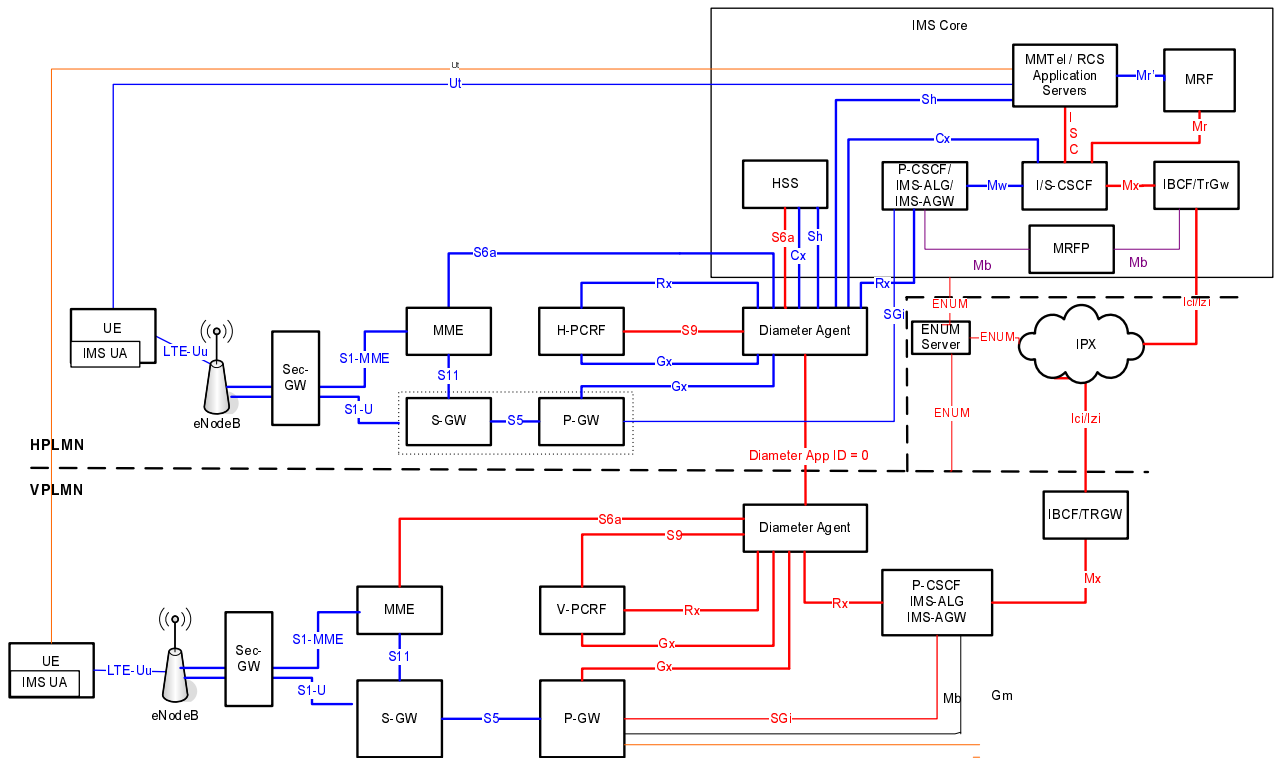


Figure 6: LTE End-to-End Configuration

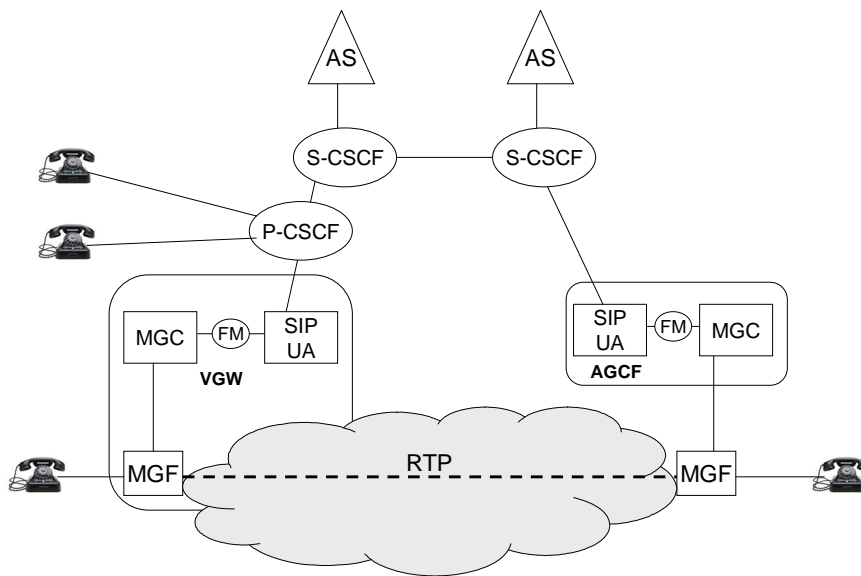


Figure 7: AGCF/VGW session processing models

5.1 Creation of subscriber data base

The subscriber data base is the data-set required to configure the SUT in order to execute a benchmark test. Using the same data, the test system should be able to generate correct traffic.

The present document does not try to specify a complete set, but rather just the subset that will ensure comparable results. All other provisioning information required for correct configuration of a SUT is to be set at the discretion of the SUT provider.

One requirement for a fair benchmark is that the input data is similar for all test-runs. In order for this to happen we have two choices:

- a) Provide data base for the subscriber base. However, because we have to ensure scalability for the benchmark, this solution is not feasible.
- b) Provide rules to generate this data and data generators. Algorithms using random generators will be avoided for data that could possibly influence the results.

6 Test system

The test system is used to generate the appropriate load on the SUT. The present document does not mandate any specific test system to be used, although the details of the test system shall be reported in the benchmark report.

The test system should have two main functions:

- Traffic generation: the test system shall be able to execute use-cases' scenarios following the traffic-time profile. It shall also be able to reproduce the appropriate traffic set (a mix of scenarios with a weight for each of them).
- Network emulation: optionally, network characteristics on the different interfaces should be emulated by the test system.

7 Benchmark metrics examples

The metrics reported by a benchmark test is measured in real time during the execution of the test, or may be computed after completion of the test from event logs collected during the execution. Enclosed is a benchmark matrix based on the TS 101 563 [1].

Table 2: Benchmark metrics examples

Selay parameters	Description
Call request delay	Call request delay is defined as the interval from the instant at which the INVITE message has been received from the SIP subscriber until the 100 Trying from the SBC/P-CSCF is passed back to the subscriber.
Alerting sending	Call request delay is defined as the interval from the instant at which the 180 Ringing is received from the terminating subscriber until the 180 Ringing is passed back to the originating subscriber.
Call set up delay	The time interval starts when the digit collection function determines that the address information received in the INFO or subsequent INVITE message is sufficient for session initiation, and ends when the INVITE message on the Ic or terminating Gm interface has been sent.
Through-connection delay	The through connection delay is defined as the interval from the instant that the 200 OK message is received from the called user at the terminating Gm interface until the through connection is established and available for carrying traffic and the 200 OK message has been sent to the calling user on the originating Gm interface.
Connection release delay	Connection release delay is defined as the interval from the instant when a BYE message is received at the originating or terminating Gm interface until the instant when 200 OK is sent and a corresponding BYE message is sent at the terminating or originating Gm interface respectively.
Speech quality analysis	
Speech Quality	PESQ (Recommendation ITU-T P.862 [i.4]) and Recommendation ITU-T P.862.1 [i.5]. P.863 POLQA "Perceptual objective listening quality assessment" (Recommendation ITU-T P.863 [i.8])
Speech Level - Active Level	Recommendation ITU-T P.56 [i.6].
Speech Level - Peak	
Speech Level - Noise	
Speech Level - Signal to Interval Noise	

History

Document history		
V1.1.1	October 2007	Publication
V1.2.1	November 2012	Publication
V1.2.2	June 2013	Publication