Title:	Document Version:
Deliverable D2.3 **Operational Interoperability Needs**	2.3

Project Number:	Project Acronym:	Project Title:
IST-2001-37611	6QM	IPv6 QoS Measurement

Contractual Delivery Date:	Actual Delivery Date:	Deliverable Type* - Security**:
30/04/2003	20/06/2003	R – PU

* Type: P - Prototype, R - Report, D - Demonstrator, O - Other
** Security Class: PU- Public, PP – Restricted to other programme participants (including the Commission), RE – Restricted to a group defined by the consortium (including the Commission), CO – Confidential, only for members of the consortium (including the Commission)

Responsible and Editor/Author:	Organization:	Contributing WP:
Emile Stephan	FT	WP2

Authors (organizations):

Miguel Angel Díaz (Consulintel), Jordi Palet (Consulintel), Yann Adam (FT), Vincent Barriac (FT), Jean-Yves Le Saout (FT), Jean-Louis Simon (FT).

Abstract:

This document presents the needs for operational interoperability among different probes, from end to end, across providers, across IPv4 and IPv6 points of measure.

Keywords:

Architecture, Collect, Management, Interoperability, Inter-domain Measurement.

Revision History

The following table describes the main changes done in the document since its creation.

Revision	Date	Description	Author (Organization)
v0.1	05/12/2002	Document creation	Yann Adam (FT)
v0.2	09/12/2002	Update requirement	Emile Stephan(FT)
v0.3	12/12/2002	Update cases studies	Yann Adam (FT) Jean Louis Simon (FT)
v0.4	12/12/2002	Update the dynamic interoperability section and case studies	Vincent Barriac (FT)
v1.0	13/12/2002	Delivery	Emile Stephan (FT)
v1.1	15/01/2003	Update file name	Emile Stephan (FT)
v1.2	27/02/2003	Finalize chapter 4	Yann Adam (FT)
v1.3	28/02/2003	Document review	Emile Stephan (FT)
v1.4	21/03/2003	Update chapter 6 and transfer content of chapter 4 to D.2.6	Vincent Barriac (FT) Jean-Yves Le Saout (FT)
v1.5	21/03/2003	Review	Emile Stephan (FT)
v1.6	26/05/2003	Update	Emile Stephan (FT) Jean Louis Simon (FT) Bernard Jalliffier (FT)
v2.0	06/06/2003	Final version	Emile Stephan (FT) Yann Adam (FT) Bernard Jalliffier (FT)
v2.1	16/06 2003	Final version	Emile Stephan (FT)
v2.2	20/06/2003	Minor corrections	Miguel Angel Díaz (Consulintel)
v2.3	20/06/2003	Final review	Jordi Palet (Consulintel)

Executive Summary

This document addresses the needs for interoperability of measurement solutions across service provider boundaries.

It takes into account the need to measure end-to-end quality of service across administrative areas and over composite networks, and as such examines the exchange of measurement data between asynchronous systems.

It provides recommendations for interoperability for following areas:

- Collecting the information;
- Reporting of the information;
- Consolidation of the information;
- Configuration;
- Test plane.

Table of Contents

Table of Figures

1. INTRODUCTION

Nowadays, the IP networks are widely deployed and more and more services are supported. The convergence of different kind of services on one unique network becomes a reality. In such network, originally best effort based, the needs to offer guaranteed services increase as well as the needs to measure the QoS offered.

IP networks rely on routers that are interconnected by different physical networks. Routing protocols run on routers to assure a good delivery of the information (data packets). The switching capacities of those routers increase regularly and also the bandwidth to interconnect them. DSL technologies deployed at the access of the network allow higher throughputs for the customers and in consequence richer services.

Basically, an IP network is organized in an access network and a core network. To interconnect different IP networks, exchange points are installed.

Service providers are in charge of theirs own networks. In a logical point of view a network is know as an Autonomous System (AS). The needs of measurements are both, inside an AS and between AS.

This document addresses the needs for interoperability of measurement solutions across service provider boundaries. It takes into account the need to measure end-to-end quality of service, and as such examines the exchange of measurement data between asynchronous systems.

2. REQUIREMENTS FOR INTER-DOMAIN INTEROPERABILITY

This section lists a collection of requirements for managing widespread end-to-end measures using the IP performance metrics specified by the IPPM Working Group. It refers to notions introduced and discussed in the IPPM Framework document, RFC 2330. The requirements apply both for active and passive measurement systems.

Sharing points of measure across administrative areas increases both the number of points of measures and the value-add of the results. It increases dramatically the number of feasible end-to-end measures while providing reliable pieces of information for management of SLA, provisioning, troubleshooting, bandwidth management, …

The difficulties in managing a system of end-to-end points of measures are more administrative than technical. The main issues to address are (1) to provide an efficient management of the different steps of the measure, (2) to guaranty the unambiguousness of the measure, (3) to guaranty the trustworthiness of results, (4) to permit efficient troubleshooting and (5) to be flexible enough to accept new measure definitions.

Managing distributed measures requires 3 levels of functionality:

- The test plane timestamps the event sends and receives the packets.
- The control plane manages the network measures. It controls the activation, the deletion and the retrieval of the result of the measure.
- The management plane is in charge of the administrative aspects and of all the issues, which are not covered by the 2 previous planes.

2.1 Conventions

A peer is an identified NMS entity from a trusted administrative area.

The control plane uses a protocol to exchange information, Measurement-Control.

The test plane uses a protocol to perform the measures, Measurement-Packets.

A consolidation measure is an aggregated measure or a statistical measure performed on the result of a network measure or on another consolidation measure.

2.2 Architecture

The control plane may be implemented using different protocols and that the test plane has to be lightweight and easy to implement.

The management plane provides a common interface for a system of points of measures using heterogeneous control protocols while preserving the complexity of the points of measures.

The control plane manages the network measures, and the test plane produces singleton results. The control plane provides the management plane with the primitives to control the network measures.

The management plane directly controls the measure of aggregated and statistical metrics.

2.3 Functional Requirements

- The whole service should provide ability to control, measure, record, and distribute the results of measurements of the metrics defined in the IPPM Working Group. End-to-end measurements is widespread, involve different administrative areas and third parties. In a way to achieve end-to-end measures, peers under the control of the management plan should access the control plane. It controls both the access to the Measurement-Control and to the Measurement-Packets. The granularity of the management distinguishes the control of a measurement instance from the access to the produced results. It provides the peer with a hierarchical grant mechanism to allow the share of the result.
- The specification of the management of the IPPM metrics should be respectful of the Framework for IP Performance Metrics [RFC2330] to provide trusted results.
- The management information exchange should be efficient enough and permit end-to-end troubleshooting.
- The design should guaranty the integration of future IPPM metrics definitions.

2.3.1 Management of the Access to the Points of Measures

Internet flows are widespread. The management plane must permit end-to-end measures across administrative areas, which is a kind of peering management.

2.3.2 Control of the Access

Sharing points of measures across administrative areas squares the number of feasible measures while providing value-added information to the management of high-level services.
The management plane must provide a hierarchical grant mechanism to control the access to the points of measures by peers:

- An administrator of an area allows peers to manage a set of network or consolidation measures on a point of measure.
- A peer allows other peers to perform consolidation measures on the results produced by its measures.
- A peer may be only granted to consult fully identified results.

Sharing network measures results reduces the test plane bandwidth consumption while providing intrinsically consistent results.

2.3.2.1 Naming of the Measure

In a point of measure, a measure instance belongs to a peer. On the network a measure involved several points of measures. The naming of the measure instances must be designed to permit an efficient management.

Currently each point of measure is in charge of the naming of the instance. It results in having different names for one measure instance:

- It is not acceptable, especially for interoperability purpose, because it introduces ambiguity at the beginning of the measure process.
- It does not provide the managers of each point with an unambiguous identifier of a measure instance. Such kinds of confusion increase dramatically the risk of misunderstanding during troubleshooting.

- It is not efficient for troubleshooting because the whole measure set-up duration may increase due to longer negotiation duration exchange with one point of measure while getting an instance name.

The management framework must provide a peer with a naming space where is identified its measure instances:

- It permits the peer to use the same identifier in all the points of measures involved by a measure.
- The results of a measure are numbered in this naming space.
- It provides an efficient share framework to the grant mechanism.

2.3.3 Interoperability Issues

Peering management permits interoperable access to points of measures. Measure must interoperate too.

The configuration of the measure must be unambiguous to permit a high level of interoperability.

2.3.3.1 Time Zone

The time zone parameter may be different at each point of the measure. The timestamp representation should be unique to avoid zone errors during the consolidation.

2.3.3.2 Unit

The comparison of two remote instants needs absolute time. As the measures are typically distributed over Internet there is a need to have an unambiguous and absolute time representation to have more aggregated metrics to be computed automatically.

According to the section "6.1. Metrics" of the Framework for IP Performance Metrics [RFC2330], the timestamp representation should be in UTC.

2.3.3.3 Packets of Type P

The section "13. Packets of Type P" of the Framework for IP Performance Metrics [RFC2330], introduces "the generic notion of a 'packet of type', "because "the value of the metric depends on the type of IP packet(s) used to make the measurement."

The definition part of the measure in the management plane, which represent the source, or the destination of the packet needs to accept any combination of existing Internet protocols suites. Its design should accept the future Internet protocol combination suites.

2.3.3.4 Sampling Law

The section "11.1. Methods of Collecting Samples" of the Framework for IP Performance Metrics [RFC2330] presents the different sampling laws.

The definition part of the measure, which represents the sampling law, needs to accept the different sampling laws discussed in the Framework for IP Performance Metrics [RFC2330].

2.3.4 Trustworthiness of the Results

The management plane must define the result format and the rules to qualify the result.

2.3.4.1 Results Definition

The management plane must specify exhaustively the results of the different measure definitions. It includes the format, the unit and the unusable value transfer rule.

2.3.4.2 Issues Related to Time

The Framework for IP Performance Metrics [RFC2330] considers that the quality of a result relative to time depends on the clock characteristics, which are its accuracy, its skew, its drift and its type of synchronization.

Result values that represent more the variation of the clock than the network behavior are unusable and must be marked as discarded.

Each result of a measure must be validated against the clock variation. A result may not display values more accurate than it was possible at the time of the measure. The result value must be consistent with the accuracy of the clock at the time of the measure.

The instantaneous characteristics of the clock are part of the measure result. So they must be accessible from the management plan. As the result gathering is delayed the history of the instantaneous characteristics of the clock must be available to the management plan.

2.3.4.3 Sampling Law

To avoid network oscillations introduced by network measures the management plane must be able to tune finely the sampling law parameters of a measure configuration.

Meshed measures which use periodic or pseudo random sampling law may introduce network oscillation. The management plane must avoid meshed measures to be synchronized.

2.3.5 End-to-end Troubleshooting

The management plane performance must be compatible with troubleshooting across administrative areas:

- A measure set-up must use a limited number of message exchanges.
- As troubleshooting must work during network congestion, the management plane should support UDP as transport protocol.
- As troubleshooting must work during network congestion a measure set-up may be done using a single message.

2.3.6 Evolution

The Type P must accept the future Internet protocol combination suites.

3. REQUIREMENTS FOR IPV4 AND IPV6 MEASUREMENT INTEROPERABILITY

At network level the following mechanisms are used for interoperability IPv6/IPv4:

- Tunneling IPv6/IPv4: Allows interconnection of IPv6 network through IPv4 network.
- Dual stack: The two IPv4 and IPv6 protocol stacks are implemented on both nodes.
- Translation: IPv6 header is translated into IPv4 header and vice versa.

At application level two mechanisms allow to use IPv4 applications on IPv6 network without modifying them:

- The BIS (Bump In the Stack) technology: By module insertion between the TCP/IP layer and the network board.
- The BIA (Bump In the API) technology: By module insertion that translates from the IPv4 socket API, into the IPv6 socket API and vice versa.

Other mechanisms related to the IPv4/IPv6 migration exist, as for example:

- Definition of an IPv6 address from an IPv4 one.
- DNSv6.

However we will discuss only the NAT-PT process and the Dual stack process respectively.

3.1 NAT-PT

3.1.1 General description

NAT-PT (*Network Address Translation – Protocol Translation*) is defined in [NAT-PT] and is a transition tool used at the border of an IPv6/IPv4 network and provides a bi-directional connector between the IPv6 and IPv4 world.

Figure 3-1: Translation NAT-PT

The main difference between NAT-PT and the others mechanisms (like *Dual Stack IPv6/IPv4*, tunneling IPv6/IPv4, BIA, BIS, DSTM) is that NAT-PT does not need to implement the two protocol stacks IPv6 and IPv4.

The NAT-PT function is a function at level 3 and as such must be deployed on IP routers.

To fulfill this function two operations take place in the process:

- NAT (Network Address Translation): This operation transforms IPv6 addresses into IPv4 ones and vice versa. The address allocation may be static, dynamic or automatic.
- PT (Protocol Translation): This operation takes care of the differences between IPv6 and IPv4 headers. Most of the fields may be translated but not all of them therefore the PT

mechanism is a Best Effort one. As a consequence some important header information may be lost in the process. PT rules are described in [SIIT].

In certain cases NAT-PT may need to access to the packet content because some protocols using IP write addresses inside the packet and those addresses need a NAT treatment. This is done by the ALGs.

An ALG is an algorithm dedicated to an applicative protocol that translates addresses inside the payload protocol packet. [NAT-PT] describes an ALG for DNS and FTP.

The NAT-PT process is state-full which means that it keeps in memory the translation context of ongoing operations. So all the packets belonging to the same flow receive the same treatment. It uses SIIT. SIIT is a state-less process, which means that each and every packet is subject to the same treatment regardless of the context.

3.1.1.1 NAT Process

The translation address mechanism must treat both source and destination address of IP packets.

The destination address is known and is subject to a static translation rule or a DNS resolution in the later case by a call to the DNS ALG algorithm.

The source address may be translated statically or dynamically.

Two types of NAT are described in [NAT-PT].

- The **traditional** NAT-PT. It allows an IPv6 network node to access to an IPv4 network node. The unidirectional sessions are exiting from the IPv6 network. Two flavors exist in the traditional NAT-PT:
  - Basic NAT-PT: A bloc of IPv4 addresses is reserved for the translation of IPv6 addresses.
  - NAPT-PT: Port numbers are translated (UDP and TCP port numbers, ICMP query identifiers). This allows several IPv6 nodes to share the same IPv4 address.
- The **bi-directional** NAT-PT. Sessions may be initialized on both IPv4 and IPv6 side. IPv6 addresses are associated to IPv4 addresses, statically or dynamically, when connections are established in both ways. The naming between IPv6 and IPv4 nodes must be unique. This is achieved through a DNS ALG algorithm usage.

3.1.1.2 PT Process

In order to understand the PT process the IPv4 and IPv6 headers are reminded then the translation mechanism is exposed.

3.1.1.2.1 IPv4 header and IPv6 header

The IPv4 header is defined in [IP] as follows:

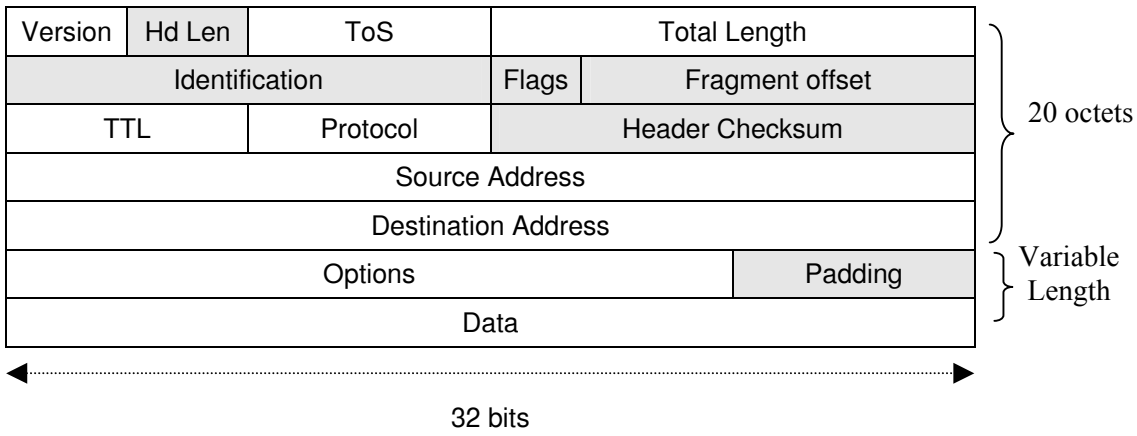

Figure 3-2: IPv4 Header

The IPv4 header has 12 fields (20 bytes total) that can be followed by an optional field followed by a data area that represents the transport level packet.

The above grey fields disappear in IPv6.

The IPv6 header is defined in [IPv6] as follows:

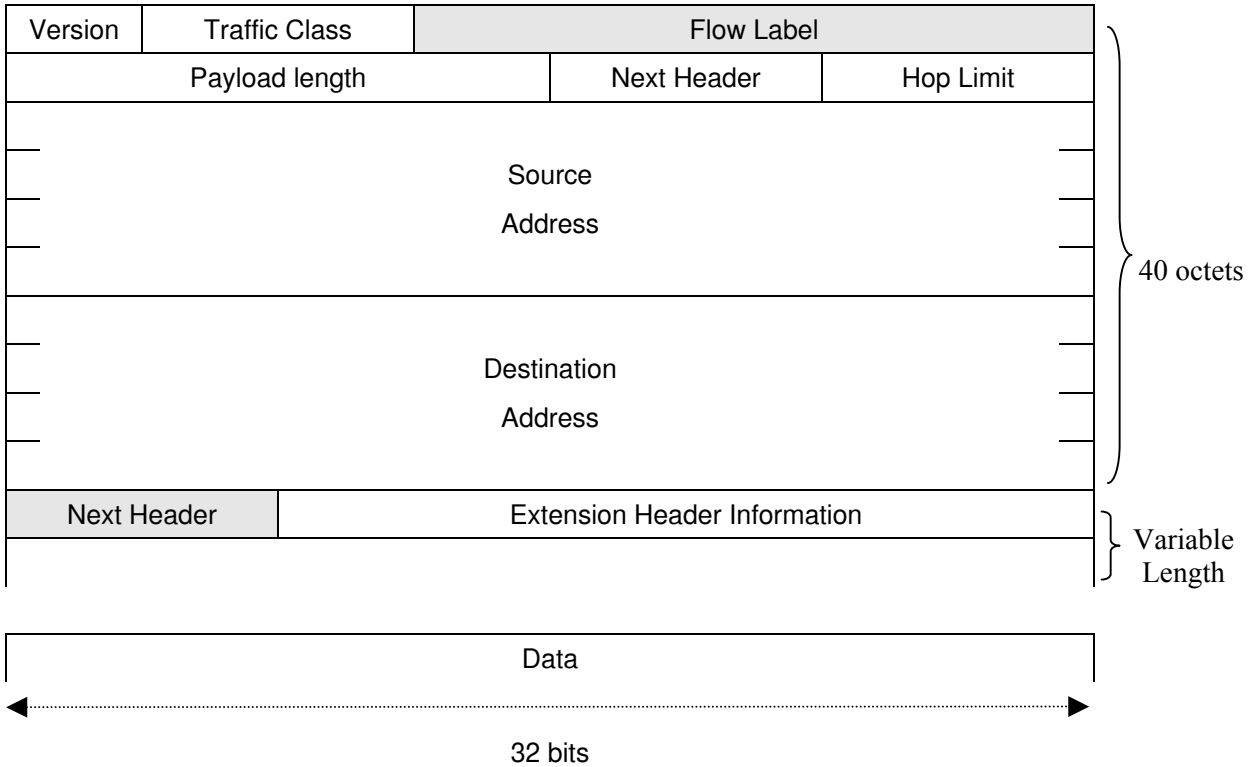

Figure 3-3: IPv6 Header

The IPv6 header has 8 fields (40 bytes total). The IPv4 header fields used for fragmentation are suppressed because this process is the emitter's duty. As well as the checksum field becomes useless, because it becomes mandatory in the IPv6 transport layer (does not need to exist anymore at the level 3). The header and the options are 64 bits aligned.

3.1.1.2.2 Translation from IPv4 fields into IPv6 fields

When the translator receives an IPv4 packet the IPv4 header is translated into IPv6 header then the packet is routed based upon the IPv6 destination address.

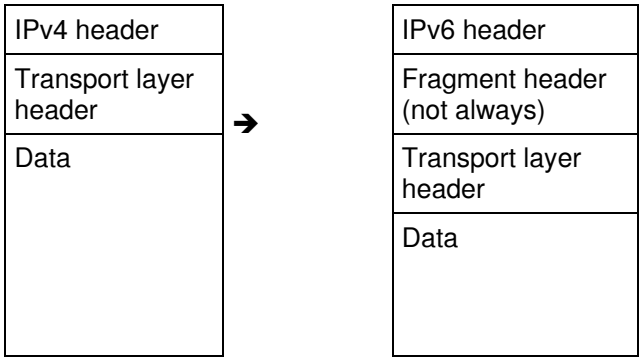

Figure 3-4: Translation Process, IPv4 to IPv6

The translation of the fields is as follow (fragmentation and ICMP excepted):

Field	Value
Version	6
Traffic Class	By default, copy of the IPv4 fields ToS and precedence (the 8 bits are copied)
Flow Label	0
Payload Length	Value of the field IPv4 Total Length less the size of the IPv4 header plus options if options are presents
Next Header	Copy of the field IPv4 Protocol
Hop Limit	Copy of the field IPv4 TTL. As the translator is a router it has to decrement and test its value (= 0 ??)
Source Address	Translated Address (refers to chapter NAT)
Destination Address	Translated Address (refers to chapter NAT)

Figure 3-5: Translation from IPv4 Fields into IPv6 Fields

3.1.1.2.3 Translation from IPv6 fields into IPv4 fields

When the translator receives an IPv6 packet the IPv6 header is translated into IPv4 header then the packet is routed based upon the IPv4 destination address.

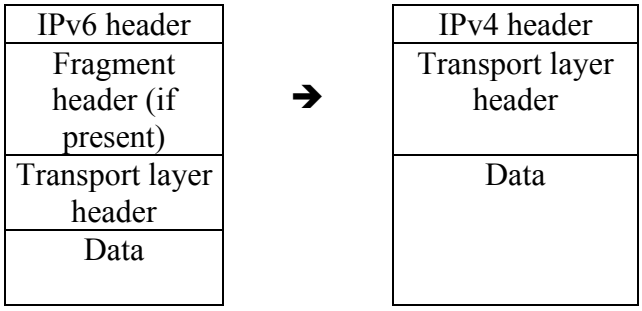

Figure 3-6: Translation Process, IPv6 to IPv4

The translation of the fields is as follow (fragmentation and ICMP excepted):

Field	Value
Version	4
Internet Header Length	5 (no options)
ToS and *precedence*	By default, copy of the field IPv6 Traffic Class (the 8 bits are copied)
Total Length	Value of the field IPv6 Payload Length plus the IPv4 header size
Identification	0
Flags	MF=0 and DF=1
Fragment Offset	0
TTL	Copy of the field IPv6 Hop Limit. As the translator is a router it has to decrement and test its value (= 0 ??)
Protocol	Copy of the field IPv6 Next Header
Header Checksum	Computed when the IPv4 header is created.
Source Address	If the IPv6 source address is an IPv4 translated one (refer to chapter NAT), then the 32 low order bits are copied in the IPv4 address. Otherwise, the source address is set to 0.0.0.0 to avoid the packet destruction.
Destination Address	The translated IPv6 packets have a mapped IPv4 address (refer to chapter NAT). The 32 low order bits are copied in the IPv4 destination address.

Figure 3-7: Translation from the IPv6 Fields into IPv4 Fields

3.1.2 NAT-PT and Active Measurement

3.1.2.1 Risk of Lost of Test Packets in NAT-PT

NAT-PT modifies the application PDU and potentially rejects the packet when the PDU cannot be parsed (for example, if there is no predefined translation for a destination address, the packet

can be drop). So the test packet must look like a PDU of the application metered. Otherwise it may be impossible to know why the packet was lost as well as the reason.

3.1.2.2 Risk of Variation of the Treatment of Test Packets in NAT-PT

If the test packet SDU does not conform to the application metered then the test packet may be processed in a different way and the result will not correspond to the performance of the regular application.

3.1.3 NAT-PT and Passive Measurement

NAT-PT may modify the application SDU (for example when it is scanning a DNS request or answer, NAT-PT has to change the addresses encoded in the IP packet at layer 4) so the hashing result of the same packet on an IPv4 network differs from the hashing result on an IPv6 network. Passive measurement is based on the hash coding of the content of the packet and generally does not take into account the packet header. In the case of NAT-PT hash key can be a problem.

3.2 Dual Stack

3.2.1 General Description

The most obvious method for an IPv6 node to maintain his IPv4 compatibility is to implement both protocol stacks. Such an IPv4/IPv6 node can by design operate in IPv6 native mode as well in IPv4 native mode.

The *dual stack* is an implementation of the TCP/IP protocol suite including the two layers IPv4 and IPv6.

Two implementations exist:

- Doubling the layer 3 only.

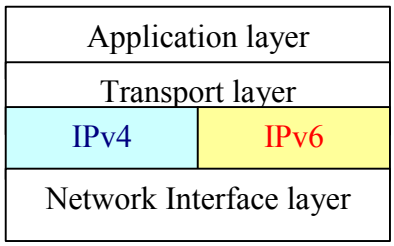

Figure 3-8: Dual Stack with Dual Layer 3

- Doubling the layers 3 and 4 (case in Windows 2003 example).

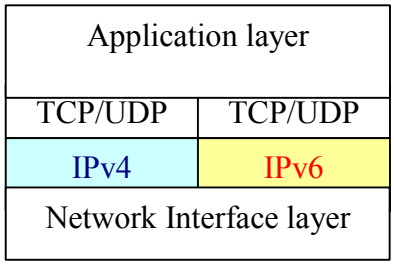

Figure 3-9: Dual Stack with Dual Layer 3 and Dual Layer 4

3.2.2 Dual Stack and Measurement Issues

Dual stack (DS) mechanism is the most appropriate for the IPv4/IPv6 transition process. However, from the measurement point of view, some points must be clearly studied when designing a measurement system:

- At the application level, connections are usually setup with the names of the terminals. In the case of DS, the name resolution is up to the name resolver of the terminal: It has to decide what type of record to send (A or AAAA), and on which transport protocol (IPv4 or IPv6). One recommendation could be to use numeric IP addresses whenever it is possible.
- IPv4 header and IPv6 header don't have the same length (20 bytes for IPv4 and 40 bytes for IPv6 without extensions). There is a consequence for the layer 2: The SDU of an IPv6 packet must be 20 bytes shorter to fit in an Ethernet frame at the maximum size.
- On the measurement path, when routers are DS they may set different metric values for the IGP in IPv4 and the IGP in IPv6. For example, IS-IS can manage two different topologies, one for IPv4 and the other for IPv6, and may configure different metric values for IPv4 and IPv6. The consequence is that the IPv4 and IPv6 paths may be different. This must be taken into account when thinking of the places of probes.

3.3 Multihoming

3.3.1 General Description

Multihoming occurs when you are connected to more than one Internet provider. It usually results in the assigning of multiple IP addresses, and possibly different routes in your own network to reach the different exit points.

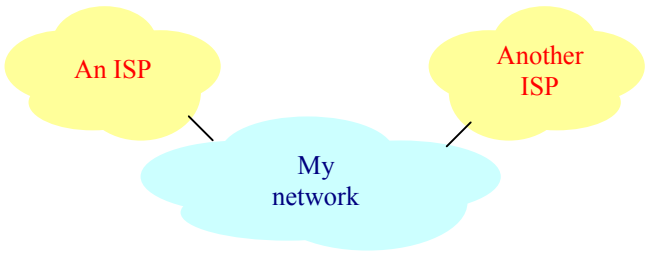

Figure 3-10: Multihoming

3.3.2 Multihoming and Measurement Issues

Multihoming is not exclusive to IPv6, but can bring some issues that must be taken into account:

- The measurement probe may have multiple addresses, especially when it is an IPv6 multihoming because these addresses are usually assigned automatically with the neighbor discovery protocol. The measurement system must be aware of that possibility and make sure to use the correct addresses.
- When one of the ISP service becomes unavailable, the measurement system must be informed in order to take the appropriate actions (adjust metrics, report the event, …).

4. MULTIPOINT MEASUREMENT TEST PLANE INTEROPERABILITY

This section discusses the need of test plane interoperability for measurement of multipoint measure.

As a first step the equipments in charge of the measurement may interoperate easily because they are from the same manufacturer.

However, the reality of networks is no so simple and there is a strong need for test equipments from different manufacturers to get a minimum common comprehension of the measured packets.

4.1 Active Measurement Test Plane Interoperability

ITU-T SG4 is working on a topic with regard to performance measurements of IP networks and services.

Question 4 of Study Group 4 is standardizing test and measurement equipment. The draft Recommendation, O.iptest, focuses on the standardization of a test packet format to perform tests of IP based networks and services.

ITU is standardizing a test packet to perform tests of IP based networks and services in the recommendation O.iptest. It consist in a trailer added at the end of the SDU of a packet, either IPv4 either IPv6. It guaranties interoperability among manufacturers and inter-domain.

The signature carries a measure owner and a measure identifier chosen by the initiator of the measure. The initiator that controls the measure provides the value of this couple of fields. It permits stateless processing of the networks measure in the probes. It facilitates consolidation of the results in the collectors. It simplifies the reporting of the measure to the initiator. The stateless processing of the measure and of the collect is compatible with the exchange of measurement results over composites networks and inter-administrative areas.

The framework of the draft Recommendation is detailed in D2.6.

4.2 Passive Measurement Test Plane Interoperability

To perform snapshot of the behavior of packet, the probes involved in the measure should share the same filtering mechanisms.

5. COLLECTING AND REPORTING INTEROPERABILITY

To gain operational interoperability collecting and reporting interfaces should be standard based.

As developed in D2.7, the IPPM WG is finalizing the IPPM REPORTING MIB and the IPFIX WG has chosen NetflowV9 as the base for standardizing an exporting protocol.

These interfaces may be used in the 6QM project, especially the one of IPFIX, which is dedicated to the exportation of passive statistics and counters.

6. CONFIGURING INTEROPERABILITY

Currently there is not any standard interface for configuring passive measure.

The IPPM REPORTING MIB defines a framework for serializing the IP protocols suites in configuration, which relies on Protocol identifiers standardized by the RMON WG.

This framework is usable for the description of the configuration of passive measure too.

Current protocol identifiers apply only to IPv4. Consequently, in the context of the 6QM project, FT presented the need of standardization of protocol identifiers for IPv6, during the RMON WG session of the 56th IETF. 6QM partners will submit a draft of protocol identifiers for IPv6 for the 57th IETF meeting in Vienna.

That will provide basis interoperability for configuring active and passive measurements.

7. CONSOLIDATION INTEROPERABILITY

The accuracy of the value of the parameters of a consolidation function must be checked before computation and the computation must be cancelled if one parameter's accuracy mismatches:

- This may be due to time sync error. So each probe must be able to indicate its absolute error on time measurement.
- Each probe must use metrics defined in IPPM WG Registry and have to be conform to measurement methods elaborated in IPPM.
- If a probe uses other metric than IPPM ones, it has to give a clear definition and measurement method of these new metrics to ensure a right usage of these measurements.
- If a probe performs an initial aggregation on to its own measurements it has to use standardized technique and it has to report sufficient information to permit a new aggregation on a different level.

  For example:

  - An aggregation of averages needs both each average result and the number of events used to perform this result.
  - An aggregation of temporal averages requires the temporal averages used to belong to a contiguous interval of time.

8. SECURITY CONSIDERATIONS

Security requirement are detailed in the delivery D2.5.

In case of passive test, the hash function cited in the chapter 4.2 has to be distributed in a secure way to avoid malicious action.

Publication of standard hashing key or function is a security hole.

8.1 DoS Attack of the Control Plane

Packet belonging to a dos attack may overflow the control plane.

8.2 DoS Detection by the Test Plane

An attacker may make use of standard hash mechanisms to create packets that will never be triggered by hash key of functions.

9. CASE STUDIES

This section provides a set of case studies coming from authors' experiences.

9.1.1 General Remarks

It is interesting to have a set of small **network tools** on the probes (ping, traceroute, telnet, …) as well as something similar to ethereal to be able to monitor and debug when necessary.

The **front panel of the probes must have a few sets of indicators** (LEDs), which permit to figure out immediately what its state is: IP connectivity, GPS synchronization, measurement activity, connectivity with the manager, …

9.1.2 Synchronization

When GPS is used the distance between the probe and the GPS receiver may be long. The measurement solution must provide a 50 meters cable length.

9.1.3 Reachability

When a probe is deployed in a **private network** (ADSL for instance), it is necessary to activate the NAT/PAT functionality on the access router. **This probe should not encode IP addresses in the packet SDU** because NAT/PAT generally works on the layer 3 only.

When possible, probes must have **routable IP addresses**.

The **lack of public IP address dedicated to measurement tools** is a real problem in an operational network.

The probes must be **802.1Q compatible**.

Communication between probes and the management application must not use private application ports. **Ports 80 and 8080** must be preferred (**security issue, firewall configurations**).

Probes may have 2 different network interfaces, one for the management network, and the other for the measurement network. This is often preferred by network operators, mainly for security reasons (the management network is generally not connected to the public internet, whereas it can be the case for the data collection interface). Nevertheless, it could be interesting that **probes work with only one network interface** for both management and measurement network.

9.1.4 Physical Connectivity

A passive probe really must be non-intrusive and not disturb the link (i.e. not generate additional traffic), excepted obviously when only one interface is used (in this case, a management traffic exists). Generally, this can be easily done by connecting the probe to a hub or a switch and using a port mirroring.

Another important reason why a passive probe must be connected non-intrusively is due to security reasons: if the probe comes down, the traffic mustn't be interrupted. It is highly recommended never to connect a probe in serial on a link.

An active probe must **prevent traffic flooding** coming from its own network interface.

Probes must be protected from external intrusion.

9.1.5 Security and Privacy Protection

Security is a major concern to be kept in mind each time any new equipment (even a measurement probe) is connected to an IP network. This is why, as mentioned above, the separation of the data collection and management planes, the connection of passive probes using port mirroring, or firewalls to protect the probes from external intrusion are commonly used.

But the security of the network is not enough. Access to the information sent and received by the users must be strictly restricted to cases where privacy is not endangered. This is why the content of the sessions or communications (i.e. the payload of packets) must not be recorded and kept in memory by measurement tools. The only exception to this is legal interception, for which the network operators are obliged by the law to provide monitoring points and access to non-encrypted data for legal authorities.

9.1.6 Correlation with End-to-end Quality as Perceived by End-users

The measurement of network QoS by itself is useful for many purposes, but there is one important domain in which it is rarely enough: The estimation of the quality as perceived at the application level, by end users. And in the world of telecommunications, the satisfaction of the customer is the main objective of the network operators (who are often service providers too).

Nevertheless, there are some possibilities to reach a better correlation with users' perception:

- Measurement probes can be soft-based, and thus it is sometimes possible to implement them very close to the end-user's interface, event in the end terminal (in PCs, for instance).
- The coupling of measurement results from different active or passive sources brings very interesting information like call routing, the determination of the sub-path where a problem occurs, and obviously correlation between end-to-end and network metrics.
- Aggregation between different metrics is often needed to understand the applicative QoS, the best example of it is VoIP, where the computation of estimated MOS or R scores based on information like the type of voice coder, packet loss, jitter, round trip delay, etc., is now possible.

Since several types of applications, based on several types of signaling and transport protocols and with several potential classes of service, can be carried on a same link, and thus supervised by the same monitoring probes, these probes must be able to filter those protocols and present the metrics separately for each type of application. Metrics can also be protocol-specific (e.g.: The number of retransmissions of TCP packets, which is not necessary for UDP).

9.1.7 Management Tool

The management tool should have a light main window: A set of indicators (red, orange, green), which permit to figure out immediately what, the state of the system is: GPS synchronization, probes reachability, database state.

The management tool should provide the user with a very simple way of starting a minimum test in order to see if everything is working fine or not.

The management interface must be protected from external intrusion.

Authentication should be used for the management application. Different levels of access rights will be available, so that different users (administrator, regular user, client, legal authorities) can access to different levels of accuracy and information.

The data stored from measurement results can be organized in different ways. But the most efficient one is a session-based presentation: For each session, the information from signaling (originating and destination addresses, date and time, etc.) and transport analysis (packet loss, delays, etc.) is given together and can be easily correlated with end-users' perception. A time-based aggregation of data (e.g.: Mean packet loss rate on a given type of protocol in a given period for a given set of IP addresses) is also necessary. The management tool can also be used to communicate information on the QoS, by sending e-mails, alarms (FTP traps), or by generating measurement reports to be sent to users (internally or externally to the organization) to verify the compliance of an SLA.

10. SUMMARY AND CONCLUSIONS

10.1 Summary

Requirements for interoperability of a measurement system are:

- Measures have to conform to measurement methodology the Framework for IP Performance Metrics [RFC2330] defines.
- The interface between collector and probes is compliant with IPFX WG protocol or with the IPPM Reporting MIB.
- The description of the configuration should use protocol identifiers.
- Each probe must be able to indicate its absolute error on time measurement.
- A probe using other metric than IPPM ones, it has to give a clear definition and measurement method of these new metrics to ensure a right usage of these measurements.
- A probe performing an initial aggregation on its own measurements has to report sufficient information to permit higher level of aggregation.
- Passive measures have to share hashing keys in a way to measure the behavior of the same packets.
- Special care should be taken when configuring measures involving points of measure impacted by NAT-PT or multihoming.
- Watchdog should be implemented to avoid measurement systems to destroy the management network or application.
- Watchdog should be implemented to avoid measurement systems performance to be affected by DoS attacks.

10.2 Conclusions

Measuring end-to-end QoS of IPv6 services deployed across administrative areas and over composite networks requires a high level of interoperability.

Despite the standardization works made by the working groups belonging to standardization entities, for gaining measurement interoperability over IPv6 networks there is a need to initiate the standardization of protocols identifiers for IPv6 and spatial metrics in order to provide passive measurement systems with both standard metrics and interoperability with end-to-end measurements systems.

References

Name	Title	Version	Date
[SIIT]	RFC 2765: Stateless IP/ICMP Translation Algorithm (SIIT)		Feb 2000
[NAT-PT]	RFC 2766: Network Address Translation - Protocol Translation (NAT-PT)		Feb 2000
[RFC2330]	Framework for IP Performance Metrics		May 1998