Project Number:	IST-1999-20841
Project Title:	**SEQUIN**

Deliverable D3.1

Definition of Quality of Service Testbed

Deliverable Type:	PU-Public
Contractual Date:	31 March 2001
Actual Date:	**17 April 2002 (Resubmitted)**
Work Package:	3
Nature of Deliverable:	RE - Report

Authors:	Mauro Campanella	INFN-GARR,
	Massimo Carboni	GARR
	Pierre Chivalier	Renater
	Simon Leinen	SWITCH
	Juergen Rauschenbach	DFN
	Roberto Sabatino	DANTE
	Nicolas Simar	DANTE

Abstract:

This deliverable provides a description of the international testbed that will be used to validate the architectural models and the technology to be used to provide end-to-end QoS services. It provides a description of the functions in hardware required to provide guarantees on the four parameters which characterise QoS, as defined in Deliverable 2.1, together with a description of the testing activity that will be carried out.

Keywords: QoS, GÉANT, IP Premium

1 EXECUTIVE SUMMARY 4

2 QOS PARAMETERS 4

2.1 DELAY 4
2.2 IP PACKET DELAY VARIATION (IPDV) 5
2.3 CAPACITY 5
2.4 PACKET LOSS 5
2.5 MAPPING OF PARAMETERS TO CLASSES OF SERVICE 6
2.6 CONCLUSION 6

3 SERVICES DEFINITION 7

3.1 PREMIUM IP 7
3.1.1 End-to-end Vs Combination of Edge-to-Edge Service. 7
3.1.2 Destination aware Vs destination unaware service. 8
3.2 MULTIDOMAIN EXTENSION 8
3.3 PROVISIONING STRUCTURE 9
3.4 DIFFERENTIATED SERVICES COMPONENTS 9
3.4.1 Traffic Conditioning (Traffic Policing and Traffic Shaping). 9
3.4.2 Packet classification and marking 9
3.4.3 Scheduling 10
3.4.4 Congestion Avoidance 10
3.5 SERVICE ASSURANCE AGREEMENT AND SPECIFICATION 10
3.6 LAN ENVIRONMENT 10
3.7 MONITORING AND ACCOUNTING 10

4 HARDWARE FUNCTIONALITY 11

4.1 INGRESS PACKET HANDLING 11
4.1.1 Classification and Policing 11
4.1.2 Classification and Policing per Destination (optional) 11
4.1.3 Shaping at the Ingress (optional) 11
4.2 EXPEDITED FORWARDING 12
4.3 INTERACTION BETWEEN INGRESS AND EXPEDITED FORWARDING 12
4.4 SHAPING AT THE EGRESS 12
4.5 STUDY OF ROUTER CHARACTERISTICS 12

5 TESTING ACTIVITY 13

5.1.1 Policing: 13
5.1.2 Heterogeneous Diffserv Environment 14
5.1.3 Scheduling 14
5.1.4 Transport Protocol 14
5.1.5 Router Performance 14
5.1.6 Aggregation degree 14
5.1.7 Overprovisioning 14
5.1.8 Shaping 14
5.1.9 Provisioning 15

6 TEST PLAN 15

6.1 INTRODUCTION 15
6.2 BASIC FUNCTIONALITY TESTS 16
6.2.1 Scheduling mechanism 16
6.2.2 Policing and marking 17
6.2.3 Congestion avoidance 19
6.3 SERVICE TEST ON SEVERAL HOPS 20
6.3.1 Premium IP 20
6.3.2 IP+ 21
6.4 TEST ACROSS SEVERAL DOMAINS 21

7 TESTBED INFRASTRUCTURE 21

7.1 PLAGE 23

7.2 GARR-G PILOT 23
7.3 QUASAR 23
7.4 POLISH TESTBED.. 25
7.5 TEST SCENARIO AND TESTBEDS 25

8 OPERATIONAL SERVICE MODEL TESTING ACTIVITY.. 26

9 INTEGRATION OF INTERNATIONAL USER GROUP(S) 26

10 REFERENCES.. 27

11 ACRONYMS.. 28

12 ANNEX A 29

12.1 SEQUIN TESTING ON PRODUCTION NETWORKS.. 29

1 EXECUTIVE SUMMARY

Definition of services, based on users needs, was provided by the SEQUIN Deliverable "D2.1 – Definition of Quality of Service". Four parameters were deduced for these end-to-end QoS services: the one-way delay, the IPDV, the one-way packet loss and the capacity. A brief definition and their mapping on the class-of-services are provided.

The Premium IP model aims to emulate a Virtual Leased Line service. Important aspects of the Premium IP model are discussed such as the end-to-end structure of the model. This end-to-end structure consists of a succession of independent domains. Each domain crossed and the interconnection between domains must implement the Diffserv EF PHB or any other mechanisms providing an equivalent behaviour. The destination aware and destination unaware models and their implications on the admission control inside a domain are discussed. Some hints about the SLA/SLS are provided. The IP+ model is left for further study.

The Diffserv basic building blocks needed for the Premium IP service are listed. Their use and where these features must be applied are explained. The basic building blocks are classification, policing, scheduling mechanism and shaping. A brief survey of the support of these features by Cisco, Juniper and Foundry is provided.

A discussion about the testing activities follows. The discussion provides general information about tests which must be performed. This discussion is followed by the tests plan description. The tests plan consist of three major steps. The first one describes the way the basic building blocks will be tested. These tests have to be carried out on various types of equipment. Concerning the second step, the test of these functionalities on several hops is discussed. The several hops tests is performed in order to get closer to real network behaviour. The last step consists of testing the Premium IP model across several domains and technologies to verify its behavior and the provisioning mechanism.

Finally, the national testbeds, available for the SEQUIN project, are described. The description includes their available hardware and the testbeds interconnection.

2 QOS PARAMETERS

In WP2, a top-down approach to the definition of QoS was taken, in that several international user groups were interviewed on their requirements for QoS. The results of this interview process are summarised in D2.1 In WP3 a complementary, bottom-up, approach has been taken in the form of an analysis exercise to decide what QoS parameters were important. A brief resumé of the QoS parameters and their definitions follow:-.

2.1 DELAY

Delay is defined as the time needed by a packet to be transmitted and fully received by the destination. The total latency can thus be divided into two parts :

- the time needed by the first bit to travel from the source to the destination (as a rule of thumb about 7 microseconds per kilometre). It is a function of
  - the speed of light

  - the number of active and passive equipment crossed along the path and the instantaneous network load (queue length);
- the time needed to transmit all the bits of the frame, which is a function of the transmission speed of the line

The first is defined in [RFC-2330] as "propagation time of a link" i.e. "The time, in seconds, required by a single bit to travel from the output port on one Internet host across a single link to another Internet host."

The one-way delay can span a wide range from almost zero to minutes. The upper boundary is set by the time-to-live field in the IP header. The counter is decreased by one at each hop and the packet is discarded when the counter reaches zero. The range from 50 to 100 milliseconds is considered to be the maximum limit required to sustain normal interactive activity and Voice over IP before service quality degradation can be perceived by a person.

2.2 IP PACKET DELAY VARIATION (IPDV)

Also known as jitter, it is defined, for a pair of packets, as the difference between the One-way-Delay measured for the second packet and the One-way-Delay measured for the first packet of the pair. For a stream of packets it is the difference of the One-way-Delay of a packet and the One-way-Delay of the preceding packet in the stream. For a stream of more than two packets, this is not a single value but a set of values. The IPDV value should be presented as an upper bounded value or a percentile value. The IPDV is mainly a function of queuing in the active equipment crossed along the path.

2.3 CAPACITY

In this context, it is defined as the sustained amount of bits per unit of time that a user is able to transmit from source to destination, independently from other traffic flowing along the same path. The specification of capacity requirements may be detailed providing the following parameters:

- maximum burst size
- peak capacity
- minimum assured capacity or committed access rate
- average value

The definition of a capacity equivalent to a leased line implies that the values of peak, minimum and average capacity are all the same and the burst size is equal to one full MTU packet.

2.4 PACKET LOSS

Packet loss is defined as the percentage of packets sent and discarded by the network. It is a function of instantaneous network load, transmission error rate and failure. Its measurement is detailed in RFC2680.

The IP Performance Metrics working group of IETF is working on how to accurately measure many of the proposed parameters [IPPM-WG].

2.5 Mapping of parameters to Classes of Service

Similar work to that of the IETF IPPM-WG has been done by the ITU, resulting in document [Y-1541]. It was decided to adopt the IETF approach, and define different levels of QoS according to the characterisation of the values of each parameter.

Examples include:

- fixed and guaranteed capacity or minimum guaranteed capacity
- negligible packet loss or bounded packet loss
- bounded or unspecified delay and delay variation

The table below compares a sample list of ranges for each parameter, to the values proposed by ITU-T in the draft recommendation document Y.1541. ITU-T defines classes of service, and values are bound in columns, values are reported between parentheses (IETF does not specify things in strict classes).

Single Values represent the "ideal" or minimum lower bound value for the parameter whilst short, medium and wide range are simply a way to broadly categorise the possible ranges.

	Single value (SV)	Short range (class 0)	Medium (class 1 - interactive)	Wide range (class 2 - non interactive)	(Class 3 - Unspecified)
One-way Delay	Measured value at empty network (baseline)	less than SV + 50 ms (150 ms)	less than SV + 250 ms (400 ms)	less than SV + 10 s (1 s)	(U)
ipdv (a)	Between 0 and the time needed to transmit one full MTU at line speed	25 ms (50 ms)	50 ms (50 ms)	none (1 s)	(U)
Packet loss (Probability)	null	< 10^-4 (10^-3)	< 10^-3 (10^-3)	< 0.1 (10^-3)	(U)
Capacity	(b)	N/A.	N/A.	A minimum of one full MTU size packet per second	(U)

(a) In Y.451 defined as the upper bound on the 1-10^-4 quintile of IPTD minus the minimum IPTD

(b) Values of only a few Kbytes per second are not adequate for advanced services

2.6 Conclusion

Drawing from the results of the interview process undertaken in WP2 (detailed in D2.1) and the ongoing activities of the IETF Diffserv-WG, it has been concluded that the implementation of two services is requested, Premium IP and IP+. These two services respond to the needs of the users as outlined in D2.1 and are capable of delivering the QoS parameters as defined.

3 SERVICES DEFINITION

This chapter defines the type of service we intend to introduce. It shows what type of tests are needed.

3.1 PREMIUM IP

In this section we focus solely on the specification of a premium IP service corresponding to a Virtual lease line service. The detailed specification of IP+ is subject to further study. The detailed specification of Premium IP is outlined in D2.1. In this deliverable we summarise the key implementation issues and assumptions that will be subject to the testing activity in WP5.

3.1.1 End-to-end Vs Combination of Edge-to-Edge Service.

The foreseen service will span multiple administrative domains. In the simplest case the selected traffic will traverse two research networks and the GEANT interconnection backbone (figure 1).

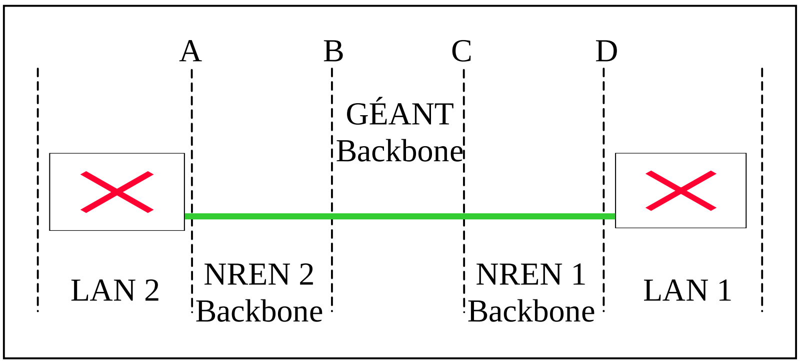


Figure 1 : multiple domains

The goal is to provide an end-to-end service from the computer in LAN 2 to the computer in LAN 1. As the traffic flows between edges of administrative domains (A, B, C, D in Fig. 1) the implementation of the service inside each domain is allowed to vary, providing that the requirements are met in each domain. The end-to-end implementation can thus be split in a set of services and interconnection rules. The implementation on each LAN will be assumed compliant with the requirements.

The interconnection rules and behaviour have to be specified and evaluated on the basis of their impact on the service. The strongest requirement in case of a service spanning multiple domains is that it should exhibit the same performance as if there were a single domain. This is assumed as the goal of the inter-domain rules definition and implementation.

The case of a Local Area Network is a special one. In a LAN the one-way delay is near to null value and, if the local environment is unloaded, possibly thanks to the high link speed, almost all the other QoS values should be near to optimal range.

3.1.2 Destination aware Vs destination unaware service.

Two different type of users' requirement have emerged:

1. a "virtual leased line service" identical in functionality to a point to point link
2. a service in which selected packets should get a preferred treatment, independently on the destination, up to a contractual ingress capacity

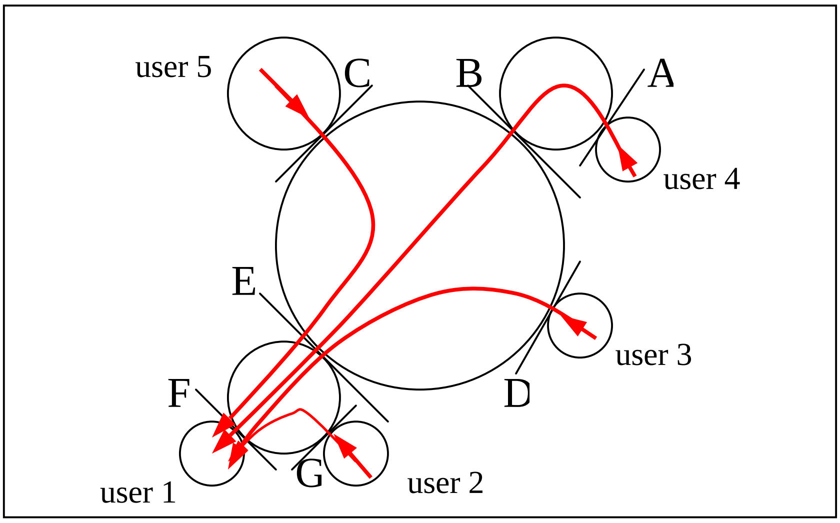


Figure 2

The first service allows a precise dimensioning of the requirements in the network. At each node it is possible to estimate the maximum capacity to be transported. The known routing path allows measurement of the delay and of the delay variation under normal operation. This type of service implies that the admission control is based on the pair (IP-source, IP-destination).

The second request implies that the traffic can be shaped and policed at the ingress according to the contract, but then the traffic paths can be different.

The most important effect is that, at each egress point from the network, the service capacity cannot be known as a priority. This behaviour requires that an asymmetric service capacity be defined at each egress; the egress capacity can, in the worst case, be as large as the sum of all the ingress service capacities. This consideration emphasises the need of estimating capacities and available service capacity at the ingress on a global network scale.

To explain this concept, Fig. 2 displays a case in which users two, three, four and five send traffic at the same time, using the QoS service, towards user one. If the SLA is of the "unknown destination" type, it implies an ingress definition only at the user edges (B, C, D, F and G). The only method to assure that the traffic is handled correctly at the other edges is to dimension the ingress and egress service parameters in such a way that it can sustain a maximum load equivalent to the sum of all the users' services.

3.2 MULTIDOMAIN EXTENSION

The implementation of the IP Premium service on an end-to-end scale in the European environment implies, in general, traffic that crosses multiple domains. The generalisation of Diffserv to multiple domains is not standardised yet.

We will adopt the following requirements to build an IP Premium service end-to-end:

- all domains involved must implement the Diffserv architecture and map the IP Premium traffic to the EF PHB.
- an interface specification is agreed between the various domains to correctly map EF traffic between them. The interface specification may contain mapping between DSCP values, policing rules, capacity assurances and all the parameters needed to ensure a correct propagation of the service.
- The interface should be defined in such a way that when packets cross management boundaries, the packet treatment is compliant to the EF PHB

In this way the two domains are free to have different implementations of the IP Premium service.

For example, one domain may implement a Per Domain Behaviour using a set of dedicated ATM CBR virtual circuits from the edge to its end node and the other may implement the EF PHB using Priority Queuing. The interface specification has to assure that IP Premium Service traffic flows to and from the two domains in a seamless way, without any packet loss and with minimum delay and delay variation. In the case of a known source and destination IP addresses for the IP Premium traffic, an appropriate routing configuration on a router with a POS interface and an ATM interface will implement the interface.

The effectiveness and sufficiency of this approach is subject to experimental validation.

3.3 PROVISIONING STRUCTURE.

In the first implementation, the provisioning structure will not be based on signalling protocols. It will be based instead on manual configuration of the different network elements involved. Policies might be propagated to the Diffserv environment either manually or automatically using, for example, BGP with appropriate extensions.

3.4 DIFFERENTIATED SERVICES COMPONENTS

To implement the Diffserv model, some key building blocks have to be taken into account. Not all of these basic bricks need to be implemented in the same node, or have to be implemented at every hop.

3.4.1 Traffic Conditioning (Traffic Policing and Traffic Shaping).

Traffic conditioning is performed at the edges of a Diffserv domain. Traffic conditioners perform traffic shaping and policing functions to ensure that traffic entering the Diffserv domain conforms to the rules specified by the Service Level Agreement and complies with the service provisioning policy of the domain. Traffic conditioning may range from simple code point re-marking to complex policing and shaping operations.

3.4.2 Packet classification and marking

Packet classification uses a traffic descriptor (for example, the DSCP) to categorise a packet within a specific group in order to define the CoS of that packet. Once a packet has been classified, it can undergo the QoS handling on the network. The packet can then be marked according to the classification result.

3.4.3 Scheduling

Scheduling is achieved through traffic scheduling and traffic queuing. A scheduling mechanism is used to provide guaranteed capacity to the different classes of traffic that are queued in different logical or physical queues.

3.4.4 Congestion Avoidance

Congestion avoidance techniques monitor network traffic loads in an effort to anticipate and avoid congestion at common network bottlenecks. Congestion avoidance is achieved through packet dropping. Among the more commonly used congestion avoidance mechanisms is Weighted Random Early Detection (WRED). With Differentiated Services, it is possible to use the DSCP value when WRED calculates the drop probability of a packet.

3.5 SERVICE ASSURANCE AGREEMENT AND SPECIFICATION

The IP Premium Service will be defined using one or a set of Service Level Agreements (SLA) and corresponding Service Level Specifications (SLS).

The service level agreement will contain at least:

- duration in time, it is suggested that a finite duration be always defined
- the administrative and technical parties involved

The service level specification must detail:

- for the end nodes involved (non-transit ones) the IP address in terms of an officially registered IP address. The IP address can be a single node or a addresses prefix.
- allowed ranges for QoS parameters (capacity, maximum delay, delay variation)
- a common set of rules to classify traffic and policing actions.

Two SLS are needed in principle for each SLA between domain edges. One is for ingress traffic and the other for egress traffic. The two do not need to be identical.

For example, in the first phase of the Service implementation, it is suggested that the egress IP Premium traffic from the core backbone is not policed at the ingress in the NREN or LAN.

3.6 LAN ENVIRONMENT

A LAN will be seen in this context as a separate Diffserv domain and its integration with the other domains will follow the guidelines sketched above. The implementation of the IP Premium Service to a LAN must comply with the same requirements as any other domain that implements the PDB. The particular implementation technique (like 802.1p or VLAN) must be equivalent for the outer domain and an appropriate interface specification must be used.

3.7 MONITORING AND ACCOUNTING

An important element of the implementation of the IP Premium Service is the creation of QoS monitoring and an accounting system. This system is in addition to the system that monitors the normal operation of the network in terms of availability, error rates, routing stability and load, for example. QoS monitoring must ensure that the tools fulfil the SLS specification, as well as provide data for proactive modification of the network in case of performance or load problems.
A series of tools should be put in place to measure the basic QoS parameters.

4 HARDWARE FUNCTIONALITY

The basic router functions needed to implement Premium IP within a domain using Differentiated Services (Diffserv) mechanisms [RFC-2475] will be discussed.

4.1 INGRESS PACKET HANDLING

4.1.1 Classification and Policing

At each interface where Premium IP traffic can enter a different administrative domain, the router must, at least, be capable of recognising Premium IP packets from the DSCP, measuring the arrival rate of those packets and comparing it against a configured traffic profile.

The traffic profile used for ingress policing should be specified as a simple token bucket defined by a rate (in bits per second) and a bucket depth (in octets). For Premium IP, a depth of one MTU should be ideally used.

As a result of comparison against the traffic profile, the router must at least be capable of forwarding packets as-is if they conform to the profile, and dropping them if not. Additionally, it is highly desirable that drops due to non-conformance are counted or flagged so that they can be monitored with network management tools.

4.1.2 Classification and Policing per Destination (optional)

If destination-aware provisioning is used, then policing of Premium IP traffic must also be done according to the destination. This requires that the router be capable of classifying into DSCP sub-aggregates according to the destination address, and performing the metering and policing of these sub-aggregates on separately specified token-bucket traffic profiles.

The sets of destination addresses that define different sub-aggregates could be specified either by using access lists on the addresses themselves, or by using rules which operate on received route announcements and assign different address prefixes to different sub-aggregates (e.g. Cisco QoS Policy Propagation).

Note that there is a potentially high performance impact associated with the mapping from destination address to sub-aggregate. Systems that can perform this mapping as part of the normal destination-based route look-up are at an advantage here, as well as systems that can process such access lists in hardware.

4.1.3 Shaping at the Ingress (optional)

Where the upstream domain (such as an NREN sending Premium IP traffic into GÉANT) is unable to shape its Premium IP traffic according to the profile, the ingress router could help by shaping incoming traffic so that it conforms to the profile when it is sent towards the core. If this feature is not available, Premium IP contracts have to be formulated with the strict shaping requirement on ingress traffic. This makes the service contract simpler to describe, but harder to implement for the upstream domain.

4.2 EXPEDITED FORWARDING

When forwarding Premium IP packets, a router must be able to provide low delay from input to output. Because strict policing at all ingress points ensures that the arrival rate of Premium IP packets is bounded at each hop, Premium IP packets can be forwarded with strict priority with respect to other traffic without incurring the risk of ``starving" other traffic of resources.

Mechanisms that can support timely forwarding of Premium IP packets include:

- Strict Priority Queuing, where Premium IP packets are assigned to a special queue that is worked on with absolute priority over all other queues.
- Weighted Fair Queuing (WFQ [Keshav97]), where Premium IP packets are assigned to a queue which enjoys a very high relative weight with respect to all other queues.
- Weighted Round Robin (WRR), Deficit Round Robin (DRR [Shreedhar95]), or Modified Deficit Round Robin (MDRR, [Sreenivasamurthy]), where all other queues are restricted to a small round-robin time-share so that the Premium IP queue is worked on frequently.

These mechanisms will exhibit varying worst-case delay for Premium IP packets in the presence of other packets.

Typically there will be other implementation specifics that influence forwarding delay, such as the impact of control activity (routing changes, network management access to instrumentation of the forwarding component etc.) and scheduler granularity.

4.3 INTERACTION BETWEEN INGRESS AND EXPEDITED FORWARDING

Ingress classification and policing may be performed on the same router that has to perform expedited forwarding. In this case, it must be ensured that conforming Premium IP packets are not only forwarded with their DSCP intact, but also treated for Expedited Forwarding on the remainder of its path through the router.

4.4 SHAPING AT THE EGRESS

On the egress to a downstream domain (receiving user), it should be possible to shape Premium IP traffic to a token bucket. Ideally, Premium IP traffic from each sending user can be shaped separately.

4.5 STUDY OF ROUTER CHARACTERISTICS

Listed below are some characteristics of the Cisco, Foundry and Juniper routers which can be used to implement an IP Premium service:-

The Cisco 12400 series has the Committed Access Rate (CAR) functionality, which allows the policing of traffic and rewriting of the DSCP field. Shaping is possible on a per interface basis. As a scheduling mechanism, Cisco has implemented Modified Deficit Round Robin (MDRR) which allows the use of two different service modes: Strict-priority mode and alternate-priority mode. The classification into a virtual output queue is based on the precedence field value.

Juniper routers are capable of policing traffic with a token bucket. The policing can be done according to the DSCP field value and the destination address. A Weighted Round Robin mechanism is used as a scheduling mechanism, which allows low delay functionality from input to output for IP Premium packets. The classification into an egress interface output queue is based on the precedence field. Juniper also has a precedence field rewriting functionality.

Foundry routers have CAR functionality in software. They also implement Weighted Fair Queuing (WFQ) and Strict Priority queuing (SP).

The following table summarises the characteristics of each router type:

(1)	Classification	Policing	Shaping	EF support
Cisco 12000 series (2)	CAR in HW and SW	CAR in HW and SW	On a per-interface basis	MDRR, strict priority queue
Cisco 7000 series (2)	CAR in HW and SW, with ACLs	CAR in HW and SW	On a per-interface basis	WRR, CBWFQ
Foundry	CAR in SW	CAR in SW	On a per interface basis	WFQ, SP
Juniper	Done in HW	Token bucket in HW	Different granularities	WRR

(1) information collected on March 2001
(2) the listed functionalities are not available for all the router types within the series and are dependant on the HW release type and engine type

5 TESTING ACTIVITY

The specification and implementation of the IP Premium service builds on the experimental activities carried out by the TF-TANT task force. Different Diffserv related aspects were analysed in the TF-TANT framework , including:

- optimisation of traffic conditioners for TCP traffic
- basic functionality of classification, marking, policing and scheduling
- comparison of different scheduling algorithms to evaluate their suitability to the provisioning of delay and IPDV guarantees, with a variety of traffic patterns and queue configurations
- tuning of WFQ configuration for delay and IPDV minimisation
- tuning of queue sizes
- configuration and performance of queuing systems based on transmission queues
- relationship between the aggregation degree and delay/IPDV/packet loss as a function of the number of aggregation points on the data path

Building on the experience gained from the above-mentioned tests, a complementary set of tests is planned and is outlined below.

5.1.1 Policing:

The optimum token bucket size to be adopted for ingress policing has to be estimated. In principle, for EF-based services, one MTU buffer should suffice. However, many vendors provide a minimum policing granularity which is equal to several MTU packets. Different token bucket configurations have to be tested, possibly in conjunction with real-life applications.

5.1.2 Heterogeneous Diffserv Environment

Quality of Service requires consistent end-to-end performance even in presence of different Diffserv implementations. Interoperability between different solutions adopted by router vendors requires in depth analysis. In addition, the consistency between ATM and non ATM-based Diffserv implementations, and the presence of multiple link speeds in each domain will be the subject of further study.

5.1.3 Scheduling

The functionality of additional scheduling algorithms which were not tested in the TF-TANT framework, such as WRR and MDRR, has to be evaluated.

5.1.4 Transport Protocol

The suitability of the IP Premium service implementation (as defined in this document) to different transport protocols requires better understanding. In particular, we need to verify the applicability to IP Premium aggregates using a mixture of different applications such as real-time audio/video, streaming, Web browsing, …

5.1.5 Router Performance

The impact of Diffserv functional blocks on the router performance requires careful evaluation. The analysis has to cover the entire range of router platforms that will need to handle traffic differentiation in production.

5.1.6 Aggregation degree

In previous test sessions in TF-TANT, UDP traffic generators were used. Future tests will deal with the effect of multiple aggregation points on traffic profiles in presence of production traffic. This gives the possibility to deal with a real mixture of different transport protocols, with real-life packet size distributions and with a significant number of streams multiplexed in a single traffic class. The effect of multiple aggregation/de-aggregation stages on a given data path will be tested.

The following tests will be performed if there is any time left.

5.1.7 Overprovisioning

The design of a Diffserv core network requires full understanding of the behaviour of current high-speed production routers in handling Best Efforts traffic. The existence and location of short term congestion has to be tracked down to evaluate the actual need of scheduling in the core.

5.1.8 Shaping

Shaping is the Diffserv functional block which modifies input traffic according to a target profile. The Diffserv architecture recommends that at domain boundaries the conformance of each traffic class is checked with the service level specification to avoid packet drop due to policing at the ingress of the following domain. Nevertheless, shaping introduces an additional buffering stage which is a potential source of delay and perturbation for the micro-flow. The effect on end-to-end performance of multiple shaping stages, the optimum configuration of a single shaper and the shaping granularity available on production routers requires in depth analysis.

5.1.9 Provisioning

The maximum amount of IP Premium traffic which can be accepted at any router interface in a Diffserv domain has to be identified as a function of the link rate. This is particularly important for the provisioning of the non destination-aware IP Premium service. In this case, the lowest speed-egress port limits the total IP Premium rate that can be accepted from all the ingresses.. This threshold is a function of the queuing algorithms adopted in the domain.

6 TEST PLAN

This section presents the definition of the testing activities, their aims and their requirements.

6.1 INTRODUCTION

The tests must be conducted with some hardware capable of supporting the basic features at line rate. The basic features are described in Section 4 Hardware Functionality. The hardware used should also be representative of what is considered to be the next generation of backbone networks.

The hardware of a next network generation backbone is considered as:-

- gigabits capable routers
- POS STM-16/STM-64 as backbone interfaces
- Access interfaces from POS STM-1 up to STM-16 (scheduling, classification, policing, shaping).

The backbone interfaces should be able to classify packets and to support a scheduling mechanism in hardware. The access interfaces should be able to support the shaping and the policing in hardware in addition to the features supported by the backbone interfaces

The tests should be performed with more than one router in order to represent a more a realistic behaviour by having several hops.

The test plan has several stages. The first stage consists of the test of the basic features requested by the services. This part should provides some expertise on how to configure the features and should bring a better understanding of their effects. These tests will be performed on the type various of equipment available in the testbeds.

The second stage consists of testing the services with several routers on the path. The aim is to observe the behavior of the services and their metrics when the traffic has to cross several hops. The transition low speed – high speed has also to be tested as well as the compatibility with ATM.

The last stage consist of testing the Premium IP model across several domains using various technologies. The aim of this test is to verify the behaviour of the service across several domains and technologies as well as the provisioning mechanisms.

During these tests, the parameters defined for each service must be monitored. The traffic generators should be able to provide information concerning the packets losses, the bandwidth used and the IPDV. The one-way delay can be monitored meaningfully between two synchronised boxes.

6.2 BASIC FUNCTIONALITY TESTS

6.2.1 Scheduling mechanism

The scheduling mechanisms are used to provide guaranteed capacity to the different classes of service. The packets of a class-of-service are classified into a queue dedicated to that class-of-services. Under congestion, Premium IP's one-way delay, IPDV and packet loss values are expected to be better results than those of other class-of-services.

The aim of the first scheduling mechanism's test is to verify the ability of a router to guarantee a fixed capacity to a class-of-service. This mechanism has to be tested on any type of interfaces, as well backbone interfaces than access interfaces. The test is carried out by:

- Sending four flows. Each flow represents an aggregate of traffic, tagged with its own DSCP value.
- Each flow is classified into a different router interface output queue than the other ones.
- A weight is allocated to each queue. Each queue has a different weight.

The sum of the flow capacities should be higher than capacity of the tested interface. The capacity of a flow should be higher than its allocated bandwidth.

The flow bandwidth seen by the destination should be proportional to capacity allocated to the flow's queue.

The second scheduling mechanism's test should provide information about the behavior of the Premium IP and the best-effort one-way delay, IPDV and packet loss. It has to be performed on backbone and access interfaces.

Two flows are sent. The first one is a best-effort flow, its packets are tagged with the dscp value 0. The second one is a Premium IP flow, its packets are tagged with the dscp value 46.

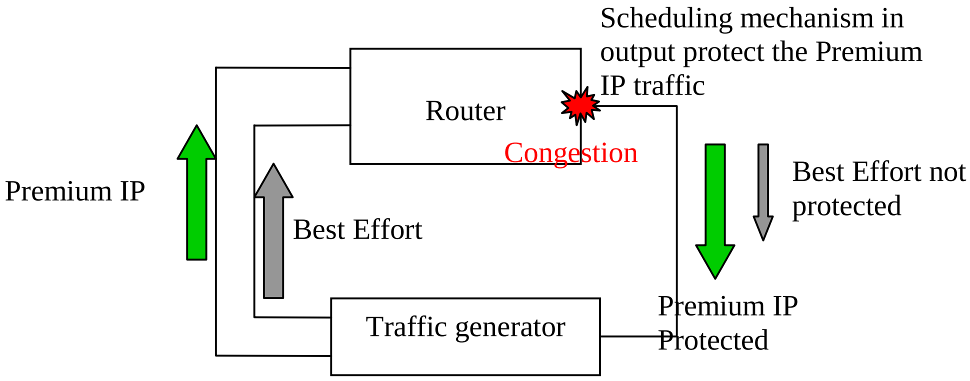

Scheduling mechanism in output protect the Premium IP traffic
Router
Congestion
Premium IP
Best Effort
Best Effort not protected
Traffic generator
Premium IP Protected

Figure 3. Scheduling mechanism protect the Premium IP traffic on the congested interface

Three cases were evaluated:

- The first was a calibration test, a small amount of traffic is sent on an unloaded test infrastructure. The one-way delay, IPDV, packet loss are measured and kept as a reference for the next tests. This corresponds to the best-case scenario where no other traffic is interfering with the calibration flow.
- Second case, the best-effort and Premium IP traffics are classified into the same queue. This is the "worst-case" scenario where Premium IP is treated as best-effort. The best-effort flow throughput varies during the test from 0% to 125% of the tested interface capacity. The Premium IP capacity is equal to 5 % of the interface capacity.
- Third case, the best effort traffic is classified into a queue and premium IP into another one. The strict priority mechanism should be applied to the Premium IP queue. If this mechanism doesn't exist, its behavior has to be emulated with mechanisms such as MDRR, WRR, WFQ or any other scheduler mechanism. The BE flows throughput varies from 0% to 125% of the tested interface capacity. The Premium IP throughput is equal to 5 % of the interface capacity.

The appropriate buffer depth and the weight must be tuned to be in accordance to Premium IP specification.

The bandwidth, packet loss, IPDV and one-way delay are monitored for the both classes of services. They must be compared to the results obtained during the first and second cases tests.

The following results are expected for the third case tests:

- Premium IP:
  - bandwidth protected
  - one-way delay and IPDV closer to best-case scenario than the worst-case one.
  - no packet loss
- Best effort:
  - increase of the delay and the IPDV with the load
  - bandwidth not protected under of congestion

The CPU load of the router, or any other parameters showing a router load variation, has to be monitored during the tests.

6.2.2 Policing and marking

The aim of this test is to verify the ability of a router to police and mark the Premium IP and the IP+ traffic. The policing is mainly performed at the ingress of a network. The re-marking of packet is performed at the ingress or at the egress of a network. These tests must be performed on access interfaces.

Several policing scenarios exist depending on where the policing is applied.

- Policing on the first router encountered on the source-destination path.
- Policing at the ingress of a network, from a Premium IP compliant network, case of destination aware model.
- Policing at the ingress of one network, from a Premium IP compliant network, case of destination unaware model.

The access control rules associated to a policer are the followings:

- The in-profiles packets must be tagged with the Premium IP value used in the network.
- The out-of-profile packets must be dropped.

- The non-conforming or errant packets must be re-tagged as best-effort. A non-conforming packet is a packet using the Premium IP tagging and which is not allowed to do so.

For all the tests, the router CPU load has to be monitored.

Policing on the first router on the source-destination path:

The policing is performed based on the source IP address and destination IP address. The DSCP value can also be used. Several policing instances must be configured on the network ingress interface. Each of them represents the access control rule applied on a flow coming from a user, a group of users or a network.

The test has to be carried out step by step:-

- One conforming flow into one policing instance. The flow capacity has to vary from 50% to 150% of the capacity configured for the policer. Some background traffic has to be added. The aim is to try to find the right bandwidth and burst-size value for configuring the policers and to verify if the policing is performed properly.
- Uses of several conforming Premium IP flows. The aim is to verify the load induced on the router by having several policers configured.
- Usage of various combination of conforming and non-conforming flows. The goal is to verify that the packets are treated as expected.

The kind of tests must be performed for the three-colour marking. Instead of policing, the packets must be tagged with different DSCP values according to the bandwidth used by the flow.

Policing at the ingress of one network, from a Premium IP compliant network, destination aware model:
The policing is performed based on the DSCP value and the network egress point. This is done either by looking at the next AS or by verifying that the destination IP address is included in a destination IP address prefix-list. One policer is configured per network egress point. If the Premium IP tagging values differ between the two networks, the packets must be re-tagged once reaching the ingress point of a network or at the egress point of the upstream network.

Several instances of policers must be configured per ingress interface. The same type of tests than the ones described in section “Policing on the first router on the source-destination path” must be performed.

Policing at the ingress of one network, from a Premium IP compliant network, destination unaware model:

The policing is performed based on the DSCP value. If the Premium IP tagging values differ between the two networks, the packets must be re-tagged once reaching the network or when they are sent to the network. The same type of tests than the previous ones must be conducted.

The total capacity used by the Premium IP flows should be equal to 5% of the circuit capacity.

6.2.3 Congestion avoidance

The aim of the WRED mechanism tests and the uses of several drop profiles is to verify a treatment differentiation for the different class-of-services. A drop profile will be associated to a class-of-service. It will give, in case of congestion, to this class-of-service a different drop probability than the one associated to the other classes-of-services.

The tests must be carried out with TCP and UDP traffic.

The first WRED test consists of finding the right parameters to protect the packets coloured green. This protection is done against the packets coloured yellow. The yellow packets are themselves protected against the red ones. The second test will investigate the distribution of bandwidth amongst several flows.

First test aims to find out the right WRED parameters to configure the colour protection: in presence of congestion, all red (coloured) packets shall be dropped before any yellow. If congestion persists, all yellow packets shall be dropped before any green. The green packets protection is expected.

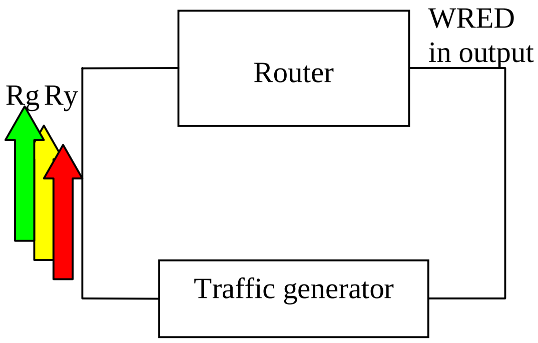
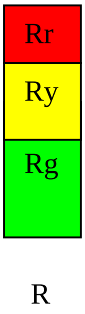


Figure 4. One flow is send with packets tagged with different colours

Packet size distribution will be used to evaluate the parameters.

Packet size	Flow BW proportion
40 bytes	56%
52 bytes	4%
576 bytes	17%
1500 bytes	23%

Red, yellow and green packets in the same queue – queue 0. The traffic generator marks the packets.



R is the sending rate, the interface capacity C = 0.75 R. The green packet rate (Rg) will vary from 0.3R to 0.8R, the yellow packet rate (Ry) from 0.15R to 0.3R. The red packet rate uses the remaining bandwidth. The WRED parameters chosen should have a good behavior with all the red-yellow-green packets distribution. The red packet rate (Rr) will be Rr = R - Rg - Ry.

This test has to be performed on any access and backbone interfaces.

The loss percentage, throughput, goodput and one-way delay must be monitored. These monitored parameters should be presented per colour and per packet size.

Distribution of BW among flow based on marking
For the second WRED test, four sources are used. The sources have the same sending rate. Each sources sends green packets as fixed proportion of the total green rate. The yellow rate of a source is 0.5 of its green rate. The total green rate is increased. The aim is to test the fairness of the distribution of the remaining bandwidth (the red bandwidth) among the sources. The green bandwidth should be “guaranteed”. The yellow one should be proportional to the green one.

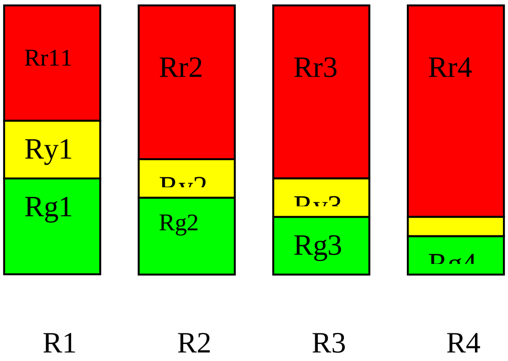


Figure 5. Bandwidth distribution per colour and per source.

4 sources R = R1 + R2 + R3 + R4 (Ri is the source i packet rate).
R1 = R2 = R3 = R4 = 120% of the tested interface capacity
Rg = Rg1 + Rg2 + Rg3 + Rg4 (Rgi is the source i green packet rate) where Rri = Ri – (Rgi + Ryi)

The same WRED parameters than for the previous test are kept.

The following parameters must be monitored: loss percentage, throughput, goodput and delay (as a function of the time). These values should be monitored per colour, source and packet size.

6.3 SERVICE TEST ON SEVERAL HOPS

The second stage consists of deploying a service on several routers defined for the services. The aims are to test the service mechanisms on a backbone topology (at least two hops), on the interconnection between high and low speed circuits, on the interconnections with an ATM network and to observe its impact of several hops on the metric.

6.3.1 Premium IP

Two flows are sent. One corresponding to the Premium IP flow and a second one corresponding to the Best-effort background traffic.

On the interfaces towards the traffic generators, the policing instance must be configured. The classification and scheduling mechanism must be configured on any interfaces involved in the tests. The best effort traffic and the Premium IP traffic are classified in two different queues. The background traffic load varies from zero, for the calibration phase, to 125% of the interface capacity. Background traffic can be injected on various point of the test topology in order to create several congestion points.

The bandwidth, losses, IPDV and one-way delay must be monitored based per class-of-service.

6.3.2 IP+

This test is very similar than the one performed in 6.2.3 section but, in these tests, several routers are crossed. Four flows are sent. Each of them has its packets tagged green/yellow/red according to the proportion previously defined. Another flow, representing the background traffic, is enabled and classified into another queue in order to create additional congestion on the routers.

On the first router towards the traffic generators, the three colour marking should be enabled. The classification and the WRED must be enabled on any involved interfaces. Some best-effort traffic can be injected on various routers in order to create several congestion points.

The bandwidth, delay and losses must be monitored per colour.

6.4 TEST ACROSS SEVERAL DOMAINS

The goal is to test the Premium IP model across several domains using different technologies. This test should give some experience on how provision the Premium IP service across several networks using various technologies. One of the domains should be under congestion.

The test has to be carried out with a flow of 2Mbps. This is typically the Premium IP capacity which could be requested by the end users. This test is close to a real use of the Premium IP service.

A first calibration phase should be conducted. Then the background traffic should be increase in one of the network.

The bandwidth, packet loss, IPDV and one-way delay must be monitored.

7 TESTBED INFRASTRUCTURE

In order to carry out tests to validate the architectural model and the techniques required to implement Premium IP using Differentiated Service mechanisms, a number of local facilities together with a Wide Area Gigabit Network infrastructure are used. Many detailed activities are planned to evaluate effectiveness and configuration challenges of the various classification, policing and queuing techniques required on the routers. These basic tests will use mostly local facilities available by the participants. These tests are all complementary to an extensive set of tests to be performed at an international scale, interconnected by means of the TEN-155 Managed Bandwidth Service based on two national Gigabit testbeds (Plage: http://www.renater.fr/Plage/, GARR-G Pilot).

A third national testbed (QUASAR: http://www.ind.uni-stuttgart.de/Content/Quasar/) will be used to focus on end-user issues and interaction between high-speed and low-speed infrastructures**.** This testbed cannot be connected to Plage and GARR-G Pilot because of the set-up of the national NREN G-WIN. Plage and GARR-G will be used to focus on QoS in a high speed environment whilst QUASAR will be used to try and understand more fully the end-user issues and the effects on moving from a high speed to a low speed infrastructure. Finally, the Polish testbed will be interconnected through the TEN-155 MBS to other testbeds and will be used for long distance and hardware tests. It will also be used to test the applicability of the high speed model to the low speed model. Should other testbeds become available and if they are complementary to the existing ones, their use will be seriously considered.

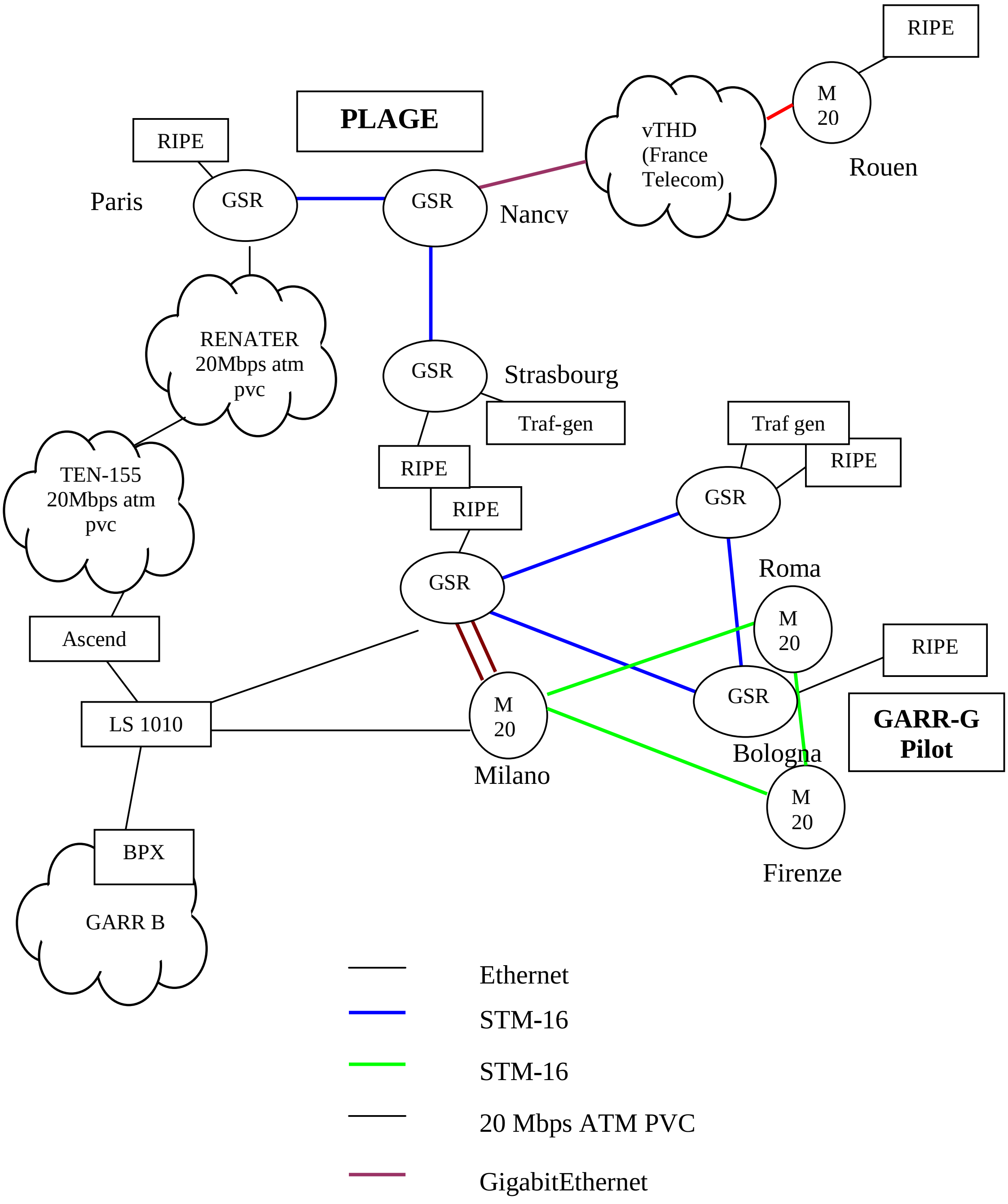


Figure 6 . SEQUIN testbed topology – interconnection of Plage and GARR-G Pilot

The link through TEN-155 will be at a speed of 10 and 20 Mb/s. Even if the link is 100 times lower in speed than the core links, for many delay and ipdv measurements it will be crucial to exploit a international, long, multi-hop link. Most of the effects under investigation increase infact linearly with the number of hops and the long delay implies the crossing of many active and passive equipment, not only router and can validate the feasibility of a complex end-to-end infrastructure on the European scale.

7.1 PLAGE

The Plage testbed is provided by RENATER in a co-operative effort with France Telecom. It comprises a number of 2.5 Gbps links between Paris, Strasbourg and Nancy. Also Rouen is connected to this testbed via France Telecom's vTHD network. The router equipment is composed of Cisco GSR routers (12008) and Juniper M20. Plage is connected to the national RENATER backbone, which allows interconnection to GARR-G Pilot, the other national Gigabit network used by SEQUIN. In terms of traffic generators, Plage makes use of two Smartbit systems from Spirent Technologies. The Smartbits currently available are old models with 100 Mbps interfaces, and efforts are being made to replace them with an up-to-date system with 2.5Gbps interfaces.

For exact one-way delay and jitter measurement it is proposed to use the RIPE TT systems equipped with GPS for time synchronisation which enables transmission of a packet with a time stamp and receipt of the same packet at another RIPE TT system where an accurate measurement can be made. The feasability of this solution is still under study due to operational constraints which have been recently outlined with RIPE. As an alternative to the RIPE TT system, Surveyor systems deployed by the Internet-2 community are being considered. Smartbits are also considered to be suitable for this purpose although their intended use is mainly for the generation of high capacity flows. Plage is currently being rolled-out, with most circuits already in place. Router installations are expected to be complete by end of April 2001.

7.2 GARR-G PILOT

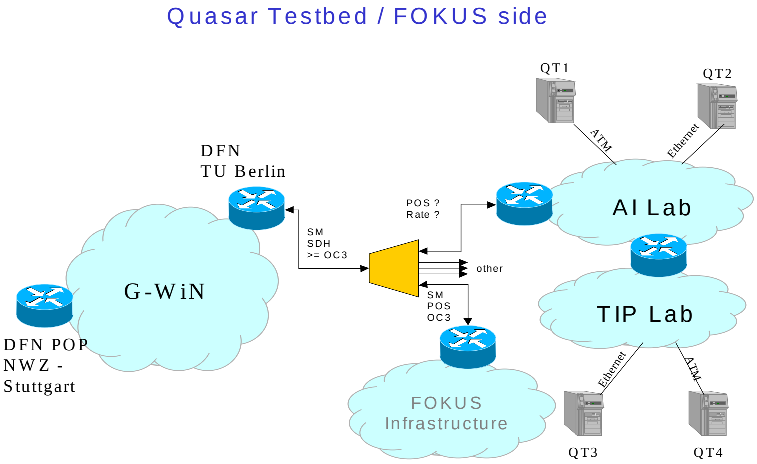
The GARR-G Pilot gigabit testbed connects Milan, Rome and Bologna with 2.5Gbps circuits provided by Telecom Italia. In parallel to this another triangle Milan - Florence – Rome, composed of 2.5 GBps circuits from WIND, is being deployed. The routers used are also Cisco GSR and Juniper M20, as in the case of Plage. As in Plage, Smartbits are being procured as traffic generators and traffic measurement equipment is being evaluated for installation. The GARR-G Pilot is scheduled to be ready for use in June 2001

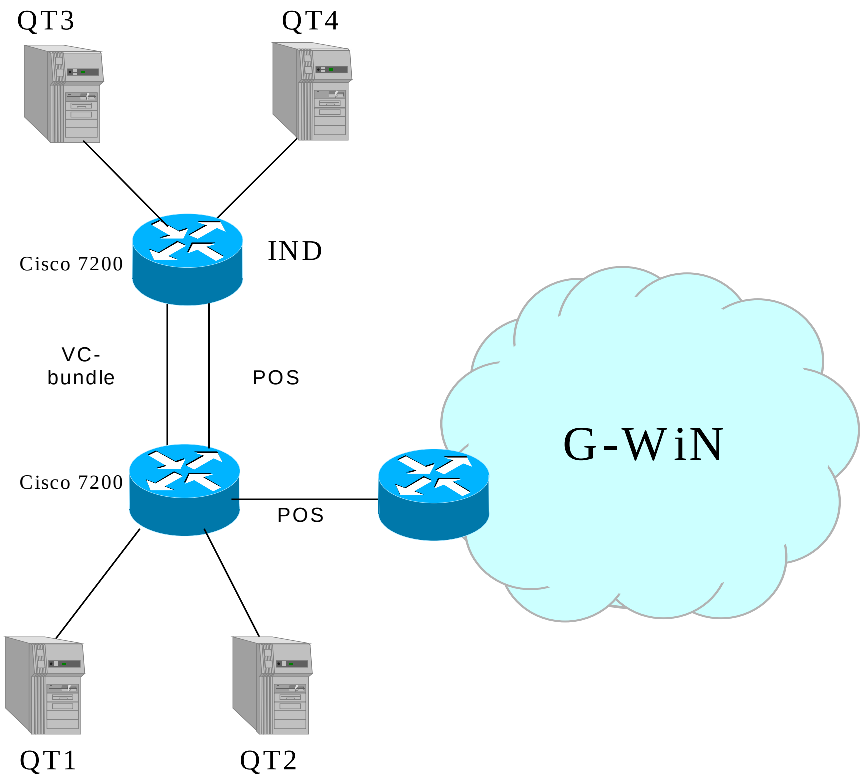
7.3 QUASAR

The Quasar tesbed is provided by DFN in the context of the Quasar (Quality of Service Architecture) project. The main objective of Quasar is the elaboration of a QoS Architecture for the German national research network G-WiN. In the case of communication between two Universities, the National Research Network operator (DFN) interconnects two access networks of two Universities via a backbone network (G-WiN). The QoS provisioning strategy of the campus networks is the responsibility of the Universities and the backbone service provider most probably operates on his own QoS provisioning concept. The Quasar network reflects this case and verifies/validates several QoS provisioning strategies on the three network sections involved in this scenario.

The project results should help to clarify whether a suitable set of service classes can be defined offering a value-added performance profile. The QoS mechanisms should yield improved network efficiency in comparison to a pure over-provisioned network without incurring overbearing complexities and management overheads. The effectiveness of the selected approaches will be demonstrated in a pilot test bed consisting of three major parts. A local testbed at the University of Stuttgart campus based on POS and/or DiffServ over ATM technology. A local testbed at GMD Fokus in Berlin consisting of test terminals in two different labs. The G-WiN core network interconnecting the two sites.

The following diagrams show the detailed network set-up of Quasar:-

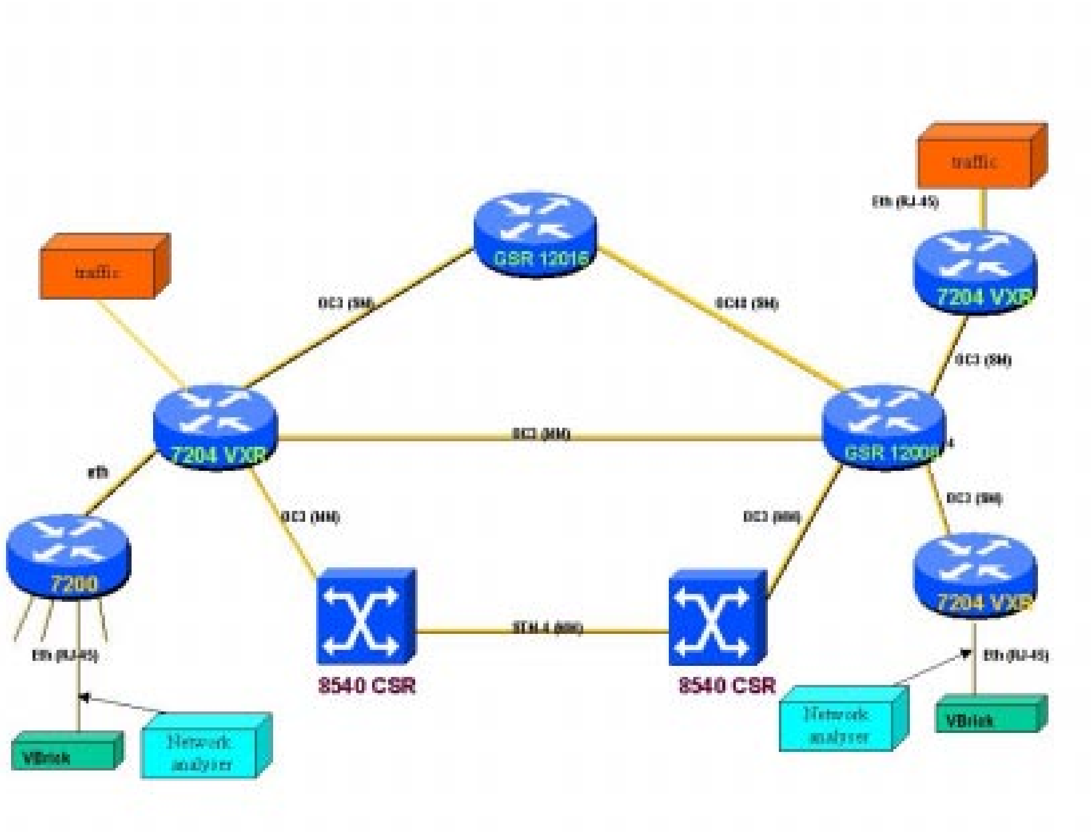




On the Stuttgart side there are 4 Quasar terminals (QT1-QT4). QT1, QT2, QT3 and QT4 are located in one department of the University. The two departments are connected with each other via a Cisco 7200 router and IP/DiffServ over ATM technology. Here, a VC bundle is set-up where IP based QoS parameters can be mapped on ATM layer. In parallel, the two departments are inter-connected via POS technology.

7.4 POLISH TESTBED

The Polish laboratory testbed consist of several Cisco 7200, GSR12000 and Cisco 8540 CSR with interfaces varying from STM-1 to STM-16 (ATM/POS). It also has two different sets of Microsens DWDM (four and eight channels) and two Extreme Networks Black Diamonds with WDM and real traffic generators (IP-MPEG codec). The variety of equipment will allow us to investigate the use of a large set of mechanisms to implement QoS. It will be possible to run long distance test by interconnecting the other testbeds via TEN-155. We will also investigate briefly how to use other technologies to provide QoS. The testbed is available since the end of April.



7.5 TEST SCENARIO AND TESTBEDS

The test beds described this section fulfil the requirements of the test program.

PlaGE, GARR-G Pilot and the Polish testbeds contain one or several gigabit capable routers (Cisco GSR 12000 series). They also contains one or more STM-16 circuits, which represent

the backbone and access circuits. A multi-hop behavior for backbone and access circuits can be emulated on PlaGE and GARR-G Pilot as they have got several gigabit capable routers.
The Polish testbed has some high and low speed circuits where the interaction between the different type of speed can be tested. The Polish and QUASAR testbeds allows the interconnection tests between DiffServ and ATM. As they also contain low speed circuits, the access features can be tested on these testbeds.
Some Smartbit traffic generators will be used on these testbeds. They will allow to load the testbeds as well as provide information on the metrics defined for the services. RIPE TTM boxes, synchronised with GPS antennas, are investigated to allow the one-way delay measurement.
Finally, an 2Mbps ATM PVC will be established between the GARR-G testbed and the PlaGE testbeds. This PVC will cross GARR, TEN-155 and Renater. This will allow to have a better understanding on how to provision the Premium IP service across several domains and technologies. It will provide also provide some information about the behavior of the service across several domains.

8 OPERATIONAL SERVICE MODEL TESTING ACTIVITY

Two significant challenges of the SEQUIN project are the interdomain operation of QoS and the multi-technology environment in terms of the different ways of building IP networks (over ATM, using MPLS, over leased lines, owned optical fibre and lambdas).equipment used. Premium IP is based on the Diffserv model which, by definition, allows independently managed domains to establish their own implementation techniques along with their own interpretation of the IP TOS/PREC bits, or DSCP field. In addition, many router implementations do not yet support the full DSCP specifications. The challenge is, therefore, to perform a mapping of the interpretation of these IP header bits from one domain to another. An SLS model between domains will be developed as part of ongoing activity in WP5 and will be one of the primary focuses of the operational service model testing activity. Another important aspect of the operational service testing activity is the provisioning testing, i.e. the set of activities that need to be performed to satisfy a request for QoS. The tasks to be performed include (but are not limited to):

- static admission control of request (i.e. can the request be accommodated?). This must take into account existing capacity commitments from the aggregate source (NREN) and existing commitments to the destination.
- reconfiguration of policing values on ingress and of shaping values on egress

- constant monitoring of QoS: loss and delay and delay variation must be measured on a per-domain basis and on an end-to-end basis. Each domain is responsible for monitoring these values at a per-hop level, whilst a mechanism should be in place to measure the end-to-end values including capacity

9 INTEGRATION OF INTERNATIONAL USER GROUP(S)

In WP5, the tests outlined in Sections six and seven will be performed initially on test, or artificial, traffic only. In order to validate the techniques, especially the operational service model testing, it is essential to connect a real user group to the testbed. In WP2, several international user groups expressed their willingness to contribute to the testing activity. The DATAGRID user group has connectivity to both Plage and GARR-G and will therefore be involved in this testing activity. Another user group that is being evaluated is the participants of the TF-STREAM task force where evaluation of H.323 over IP is being tested. This is used for tele-teaching, for example. More details of international user groups involved in the testing phase and their connectivity to the testbed will be developed in WP5.

10 References

[Diffserv-WG] http://www.ietf.org/html.charters/Diffserv-charter. html

[EFPHB] An Expedited Forwarding PHB, Bruce Davie, Editor, Anna Charny, Fred Baker, expires Aug 2001,

http://www.ietf.org/internet-drafts/draft-ietf-Diffserv-rfc2598bis-00.txt

[IPPM-WG] http://www.ietf.org/html.charters/ippm-charter.html

[Keshav97] Keshav, S., "An Engineering Approach to Computer Networking", Addison Wesley, January 1997.

[RFC-2475] RFC2475 An Architecture for Differentiated Service. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss. December 1998. (Status: INFORMATIONAL)

[Shreedhar95] Shreedhar, M. and G. Varghese, "Efficient FairQueueing Using Deficit Round Robin", SIGCOMM 1995, pages 231-242

[Sreenivasamurthy] Sreenivasamurthy, D., "Implementation and Evaluation of support for Differentiated Services mapping to ABR service in an Edge/Core Network", Thesis, University of Kansas.

[TF-TANT] http://www.dante.net/tf-tant

[Y-1541] ITU Study Group 13, "Revised draft Recommendation Y. 1541 'Internet protocol communication service - IP Performance and Availability Objectives and Allocations'", November 2000

11 ACRONYMS

ADSL	Asymmetric Digital Subscriber Link
ATM	Asynchronous Transfer Mode
CoS	Class of Service
DSCP	Differentiated Services Code Point
EF PHB	Expedited Forwarding Per Hop Behaviour
IETF	Internet Engineering Task Force
IP	Internet Protocol
ITU	International Telecommunications Unit
IPDV	IP Packet Delay Variation
IPPM	IP Performance Measurement
LAN	Local Area Network
MAC	Medium Access Control
MDRR	Modified Deficit Round Robin
MTU	Maximum Transfer Unit
NREN	National Research and Educational Network
PHB	Per Hop behaviour
PQ	Priority Queuing
QoS	Quality of Service
SLA	Service Level Agreement
SLS	Service Level Specification
SP	Strict Priority
TCP	Transmission Control protocol
TOS	Type of Service
UDP	User Datagram Protocol
WFQ	Weighted Fair Queuing
WRED	Weighted Random Early Detection
WRR	Weighted Round Robin

12 ANNEX A

12.1 SEQUIN TESTING ON PRODU CTION NETWORKS

Considering that a part of the SEQUIN project (WP4) is to deliver an implementation plan on operational networks and considering that some user groups may not be connected to the national testbeds, it may be important to carry out some of the testing on production networks.

To run the tests planned in SEQUIN on TEN-155 or other production networks the following considerations should be taken into account:

- The tests must be performed in laboratory or on the national testbeds first, and preferably at high speed
- The SW revision on the production routers must have the functionality requested. It may take up to 3 weeks (maintenance windows) to make a SW change on production routers, so this does require careful planning ahead. Also, it is preferred that the SW revision has been used in other production environments
- The configuration changes required on the routers must be done in pre-announced “at-risk” windows
- The tests should be run in pre-announced “at-risk” windows
- A scheme to account for major node failure should be in place. On TEN-155 for example this is already in place as follows:

    - NRNs in general have an access PVC landing on the 7k router in the relevant PoP. To protect from failures on the 7k, the NRN has a backup ATM PVC to another node in the network
    - In each PoP where applicable, the 7K is connected to the Juniper M40 router which handles all the core traffic. To account for failure of the M40, the 7k has an ATM connection to another M40 in another PoP.

    NRNs could have something similar, tailored to their particular set-up.

- Appropriate monitoring and measurement schemes should be in place to monitor the behaviour of the router in shipping packets. For example, if the tests are running on a congested node and the scope of the tests is to apply differential treatment to normal and special traffic this should monitored and measured.
- CPU load and memory utilisation, where applicable, should be monitored.

From the TEN-155 perspective, SWITCH, JANET, RENATER and GARR have ATM connections to the Cisco 7507 routers which have many features related to QoS on ATM cards. DFN is connected to the Juniper M40 via POS/STM-4, which has a wide range of QoS features.

In the model presented at the second SEQUIN meeting in Rome in December 2000, TEN-155 would need to police marked packets from the NRNs and perhaps shape them upon delivery to the destination NRN. These functions, on the Juniper M40, are supported on POS interfaces but not on ATM. In some cases TEN-155 may need to apply preferential queuing and scheduling to the marked packets. These functions are in general supported on the TEN-155 core routers except for where there is a core connection using ATM between an M40 and a 7k router.

For the NRNs that have only an ATM interface to TEN-155, the 7507 router is needed to perform the policing and the shaping from/to the NRN.

As the techniques that are being investigated match the Diffserv concept of Per Hop Behaviours, it is possible to make changes on one router at a time, in an incremental manner. This gives a high degree of control of the network changes and makes it very easy to fall back to the previous network configuration in case of unforeseen problems.