QoS and IP Premium service specification and implementation

Presentation given by Mauro Campanella (GARR) at the TERENA Networking Conference 2001, May 2001

Slide Show

Table of Contents

Author: Mauro Campanella

Email: campanella@mi.infn.it

SEQUIN Home Page: http://www.dante.net/sequin/

Download presentation source

QoS and IP Premium service specification and implementation

Mauro Campanella
INFN-GARR
campanella@mi.infn.it

Research groups

TF-NGN - A joint TERENA and DANTE task force on advanced networking research
http://www.dante.net/tf-ngn

SEQUIN - A RN2 project on QoS on interconnected domains
http://www.dante.net/sequin

TF-NGN Work Groups

Guaranteed Capacity Service Specification and implementation plan	- Hervé Prigent	Crihan-RENATER
Premium IP service specification	- Mauro Campanella	INFN-GARR
Tools for network monitoring / flow measurement	- Simon Leinen	SWITCH
MPLS testing	- Herve' Prigent	Crihan-RENATER
Delay and Jitter sensitive based services	- Tiziana Ferrari	INFN-CNAF
Diffserv AF based services	- Octavio Medina	IRISA
QoS monitoring	- Victor Reijs	HEAnet-SURFnet
Over-provisioned network performance analysis	- Tryfon Chiotis	GRNET
QoS and multicast	- Robert Stoy	DFN
IPv6	- Tim Chown	Univ. of Southampton
MPLS testing	- Hervé Prigent	Crihan-RENATER
Improvement of current multicast service		
User-oriented multicast	- Ladislav Lhotka	CESnet
Multicast developments	- Robert Stoy	DFN
Optical Networking	- Victor Reijs	HEAnet-SURFnet

QoS and IP Premium motivations

- Users' requirement (interviews by Sequin) for services that provide assured capacity and delay and minimum delay variation
- ATM is fading away (no longer any link layer assurances). A replacement is needed for the Managed Bandwidth Service in TEN-155.
- No Overprovisioning over all Europe (yet)
- IP telephony, MPEG2 interactive video, time sensitive applications are here

IP Premium goal

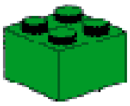
Provision QoS in the network for the European research users in the form of an end to end service offering the equivalent of a leased line.

The service has to be implemented by combining border to border services provided by the NRENs and GEANT networks

The service should be simple, scalable, adapt to network changes easily, based on IP and independent from the transport technology.

QoS parameters

The identified set is :



- ✓ - one-way delay;
- ✓ - one-way packet delay variation;
- ✓ - capacity;
- ✓ - packet loss.

The set matches the IETF and ITU-T ones, naming and definitions will follow RFC 2330 (Framework for IP Performance metrics)

Link layer and routing stability, BER, hardware performance, down time are supposed adequate.

MTU size is supposed to be large enough to avoid fragmentation

IP Premium Specification

⇒ Differentiated Services Architecture

⇒ expedited forwarding per hop behavior (EF PHB) in all domains involved

⇒ interface definition between domains that behaves as an EF PHB

⇒ do not starve best effort traffic

⇒ initial provisioning structure: static, no dynamic signaling

⇒ IETF IPPM QoS parameters measurement framework

⇒ QoS parameters monitoring system is a key element

Implementation

There are still decisions to be taken and open technical issues that can influence each other. The work is in progress.

Caveat (again)

It is assumed that the following ingredients are good enough:
- Link layer : bit error rate (< 10^{-11}), stability, down time
- Silicon : fast (Gb/s), stable, redundant, load-independent performance
- Last mile : minimum level of hardware and capacity (at least 802.1p capable, switched, 10 Mb/s ?)
- MTU size : large enough to avoid fragmentation

The architecture implementation and the SLA have to match reality.

Implementation Decision for the Service Level Agreement

- Admission control rule parameters
- Local Vs global (end to end) agreements
- Asymmetric Service Level Specifications

Admission control rule

In principle might be an arbitrary combination of:

- IP v4 Header contents
  - IP source and destination
  - ToS
- Ports
- Protocol
- time of the day, application type, load….

Just making mandatory or not the list of IP destinations has profound impact on the type of service (destination aware Vs destination unaware or selling Virtual leased lines Vs Aggregate IP Premium Capacity).

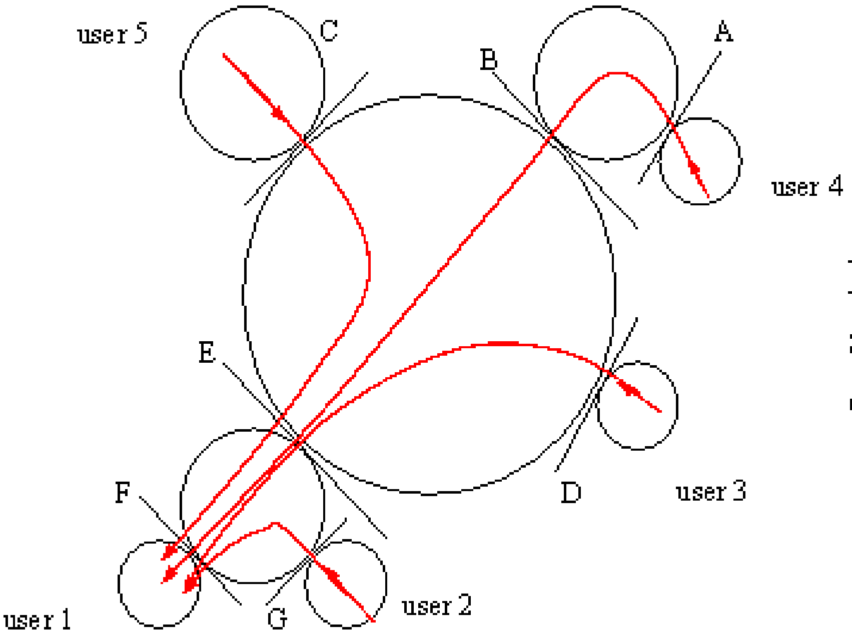
Admission control rule (continued)

Destination aware

- precise dimensioning of resources at each node
- allows known bounds on delay and delay variation

but

- *detailed knowledge of routing*
- *more complex, if sub-aggregates have to be metered separately at each ingress point*
- *sensitive to routing failures*



Destination un-aware and egress bandwidth dimensioning

user	IN (SLS)	OUT (SLS)
1	10	10
2	10	10
3	10	10
4	10	10
5	10	10

Admission control rule (continued)

Destination UN-aware

- simpler configuration of the network elements
- does not need precise knowledge of the network
- weakly sensitive to re-routing

but

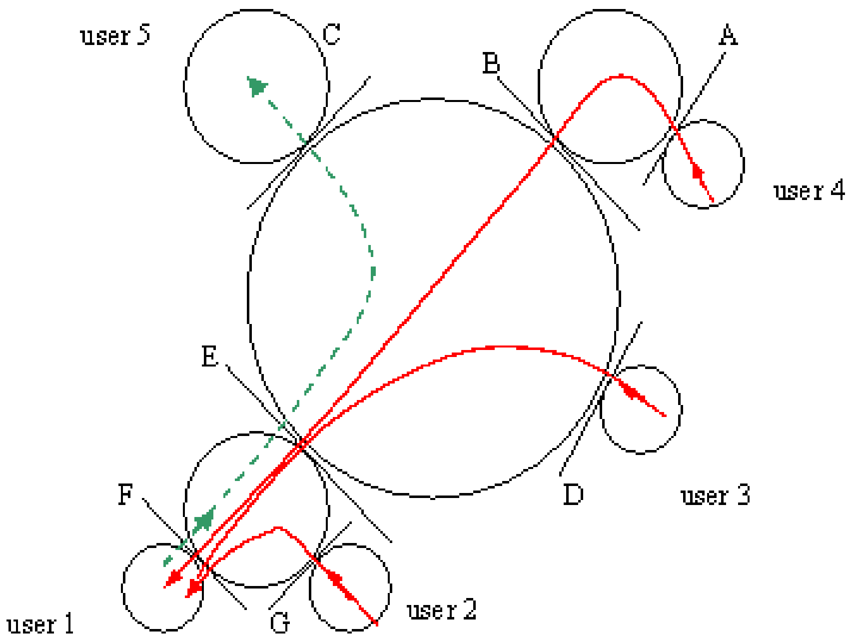
- *allows only extreme bounds on delay and delay variation*
- *implies overprovisioning or absence of policing at the egress*
- *ubiquitous constraint on maximum amount of IP Premium bandwidth configurable on all the links as a function of the lowest speed link*
- *shaping only on aggregates (non per-flow guarantees)*

Asymmetric SLS

There is in principle no reason to avoid asymmetric SLS for ingress and egress on the some boundary, for example for capacity.

If destination un-aware policy is chosen the ingress SLS to a user site has to be left unspecified and can only be assumed to be up to a maximum equal to the sum of all the total egress IP Premium capacity of all the user sources.

Local Vs Global agreements



Suppose user 1 wants to speak IP Premium with user 5 only. Users 2, 3, 4 want to speak with User 1.

If the destination address is known, then it is possible to dimension boundary F, but user 1 will have to discuss with all other users and decide if he accepts to send and receive much more IP Premium traffic then he originally expected.

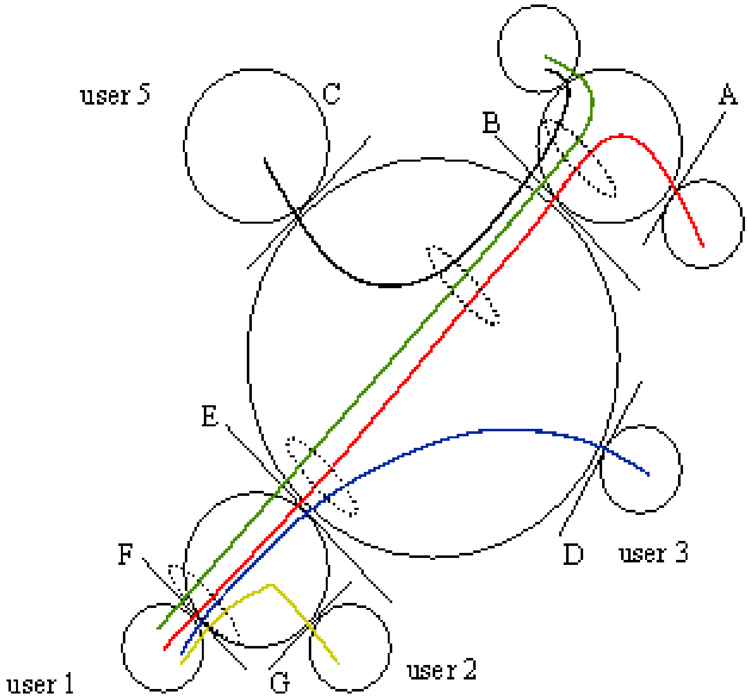
The SLA should be propagated end to end

IP Premium open issues

Technical Issues

- shaping
- aggregation de-aggregation of microflows
- basic (empty) network behavior
- interaction of multiple Diffserv domains
- a LAN as a Diffserv domain
- implementation according to specific hardware and its performance
- tuning (in particular of queuing)
- monitoring architecture
- effects and tuning for protocols other than UDP

Shaping and aggregation



Shaping is required at the source, at least for non elastic protocols...

Along the path there are multiple aggregation -- de-aggregation points and link speed changes.

Study the distortion of shaping and its relation with delay variation.
(switching time for a 1500 Bytes packet is about 5μs at 2.5 Gb/s)

References

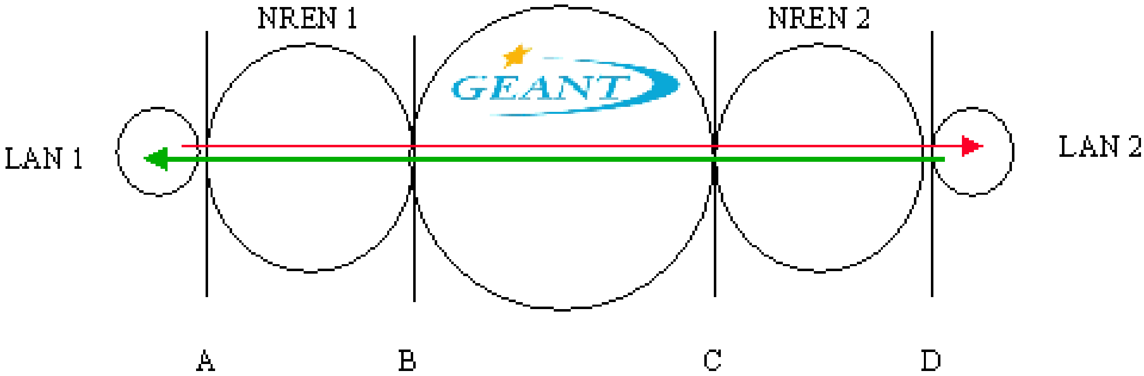
GÉANT Deliverable D9.1 "Specification and implementation plan for a Premium IP service": http://www.dante.net/tf-ngn/GEA-01-032.pdf

Sequin Deliverable D2."QoS definition", to be available at http://www.dante.net/sequin

TF-NGN public documents: http://www.dante.net/tf-ngn/

Thank you

Sample Model



- Classification, marking at A and D only (common value for EF DSCP)
- Strict policing ***ingress*** IP Premium traffic according to IP source and destination at A and D only. Do not police ***egress*** traffic
- Shaping possibly at B
- Priority Queueing or highest weight for EF Traffic
- Switching in the NREN 1,2 and GÉANT core only based on DSCP (ToS)

QoS and IP Premium service specification and implementation

Mauro Campanella

INFN-GARR

campanella@mi.infn.it

Next slide Back to first slide View graphic version

PPT Slide

Research groups

- A joint and

task force on advanced networking research

http://www.dante.net/tf-ngn

- A RN2 project on QoS on interconnected

domains

http://www.dante.net/sequin

Previous slide Next slide Back to first slide View graphic version

PPT Slide

TF-NGN Work Groups

Guaranteed Capacity Service Specification

and implementation plan - Hervé Prigent Crihan-RENATER

Premium IP service specification - Mauro Campanella INFN-GARR

Tools for network monitoring / flow measurement - Simon Leinen SWITCH

MPLS testing - Herve' Prigent Crihan-RENATER

Delay and Jitter sensitive based services - Tiziana Ferrari INFN-CNAF

Diffserv AF based services - Octavio Medina IRISA

QoS monitoring - Victor Reijs HEAnet-SURFnet

Over-provisioned network performance analysis - Tryfon Chiotis GRNET

QoS and multicast - Robert Stoy DFN

IPv6 - Tim Chown Univ. of Southampton

MPLS testing - Hervé Prigent Crihan-RENATER

Improvement of current multicast service

User-oriented multicast - Ladislav Lhotka CESnet

Multicast developments - Robert Stoy DFN

Optical Networking - Victor Reijs HEAnet-SURFnet

PPT Slide

QoS and IP Premium motivations

- Users’ requirement (interviews by Sequin) for

services that provide assured capacity and delay

and minimum delay variation

- ATM is fading away (no longer any link layer

assurances). A replacement is needed for the

Managed Bandwidth Service in TEN-155.

- No Overprovisioning over all Europe (yet)

- IP telephony, MPEG2 interactive video, time

sensitive applications are here

Previous slide Next slide Back to first slide View graphic version

PPT Slide

IP Premium goal

Previous slide Next slide Back to first slide View graphic version

PPT Slide

QoS parameters

The identified set is :

?- one-way delay;

?- one-way packet delay variation;

?- capacity;

?- packet loss.

The set matches the IETF and ITU-T ones, naming

and definitions will follow RFC 2330 (Framework

for IP Performance metrics)

Link layer and routing stability, BER, hardware performance, down time are supposed adequate.

MTU size is supposed to be large enough to avoid fragmentation

Previous slide Next slide Back to first slide View graphic version

PPT Slide

IP Premium Specification

? Differentiated Services Architecture

? expedited forwarding per hop behavior (EF PHB) in

all domains involved

? interface definition between domains that behaves as

an EF PHB

? do not starve best effort traffic

? initial provisioning structure: static, no dynamic

signaling

? IETF IPPM QoS parameters measurement framework

? QoS parameters monitoring system is a key element

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Implementation

There are still decisions to be taken and open technical issues that can influence each other. The work is in progress.

Caveat (again)

It is assumed that the following ingredients are good enough:

- Link layer : bit error rate (<10-11), stability, down time

- Silicon : fast (Gb/s), stable, redundant, load-independent

performance

- Last mile : minimum level of hardware and capacity

(at least 802.1p capable, switched, 10 Mb/s ?)

- MTU size : large enough to avoid fragmentation

The architecture implementation and

the SLA have to match reality.

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Implementation Decision for the Service Level Agreement

- Admission control rule parameters

- Local Vs global (end to end) agreements

- Asymmetric Service Level Specifications

PPT Slide

Admission control rule

In principle might be an arbitrary combination of:

- IP v4 Header contents

- IP source and destination

- ToS

- Ports

- Protocol

- time of the day, application type, load….

Just making mandatory or not the list of IP destinations has profound impact on the type of service (destination aware Vs destination unaware or selling Virtual leased lines Vs Aggregate IP Premium Capacity).

PPT Slide

Admission control rule (continued)

Destination aware

- precise dimensioning of resources at each node

- allows known bounds on delay and delay variation

but

- detailed knowledge of routing

- more complex, if sub-aggregates have to be metered

separately at each ingress point

- sensitive to routing failures

PPT Slide

A

B

C

D

E

F

G

user 1

user 2

user 5

user 4

user 3

Destination un-aware

and egress bandwidth

dimensioning

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Admission control rule (continued)

Destination UN-aware

- simpler configuration of the network elements

- does not need precise knowledge of the network

- weakly sensitive to re-routing

but

- allows only extreme bounds on delay and delay variation

- implies overprovisioning or absence of policing at the egress

- ubiquitous constraint on maximum amount of IP Premium

bandwidth configurable on all the links as a function of the

lowest speed link

- shaping only on aggregates (non per-flow guarantees)

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Asymmetric SLS

There is in principle no reason to avoid asymmetric SLS for ingress and egress on the some boundary, for example for capacity.

If destination un-aware policy is chosen the ingress SLS to a user site has to be left unspecified and can only be assumed to be up to a maximum equal to the sum of all the total egress IP Premium capacity of all the user sources.

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Local Vs Global agreements

Suppose user 1 wants to

speak IP Premium with user

5 only. Users 2, 3, 4 want

to speak with User 1.

If the destination address is

known, then it is possible to

dimension boundary F, but

user 1 will have to discuss with

all other users and decide if

he accepts to send and receive

much more IP Premium traffic

then he originally expected.

The SLA should be propagated end to end

PPT Slide

IP Premium open issues

Technical Issues

- shaping

- aggregation de-aggregation of microflows

- basic (empty) network behavior

- interaction of multiple Diffserv domains

- a LAN as a Diffserv domain

- implementation according to specific hardware

and its performance

- tuning (in particular of queuing)

- monitoring architecture

- effects and tuning for protocols other than UDP

Previous slide Next slide Back to first slide View graphic version

PPT Slide

A

B

C

D

E

F

G

user 1

user 2

user 5

user 3

Shaping and aggregation

Shaping is required at the source, at least for non elastic protocols...

Along the path there are multiple aggregation -- de-aggregation

points and link speed changes.

Study the distortion of shaping and its

relation with delay variation.

(switching time for a 1500 Bytes packet

is about 5ms at 2.5 Gb/s)

PPT Slide

References

GÉANT Deliverable D9.1 “Specification and implementation plan for a Premium IP service”:

http://www.dante.net/tf-ngn/GEA-01-032.pdf

Sequin Deliverable D2.“QoS definition”, to be available at http://www.dante.net/sequin

TF-NGN public documents: http://www.dante.net/tf-ngn/

Previous slide Next slide Back to first slide View graphic version

PPT Slide

Thank you

Previous slide Next slide Back to first slide View graphic version

PPT Slide

A

LAN 1

Sample Model

LAN 2

B

C

D

NREN 1

NREN 2

- Classification, marking at A and D only (common value for EF DSCP)

- Strict policing ingress IP Premium traffic according to IP source and

destination at A and D only. Do not police egress traffic

- Shaping possibly at B

- Priority Queueing or highest weight for EF Traffic

- Switching in the NREN 1,2 and GÉANT core only based on DSCP (ToS)