GEANT

6/14/02

Click here to start

Author: nicolas

Email: Roberto.Sabatino@dante.org.uk

Table of Contents

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

PPT Slide

H.323 trial

PPT Slide

DANTE

SEQUIN: QoS for GEANT

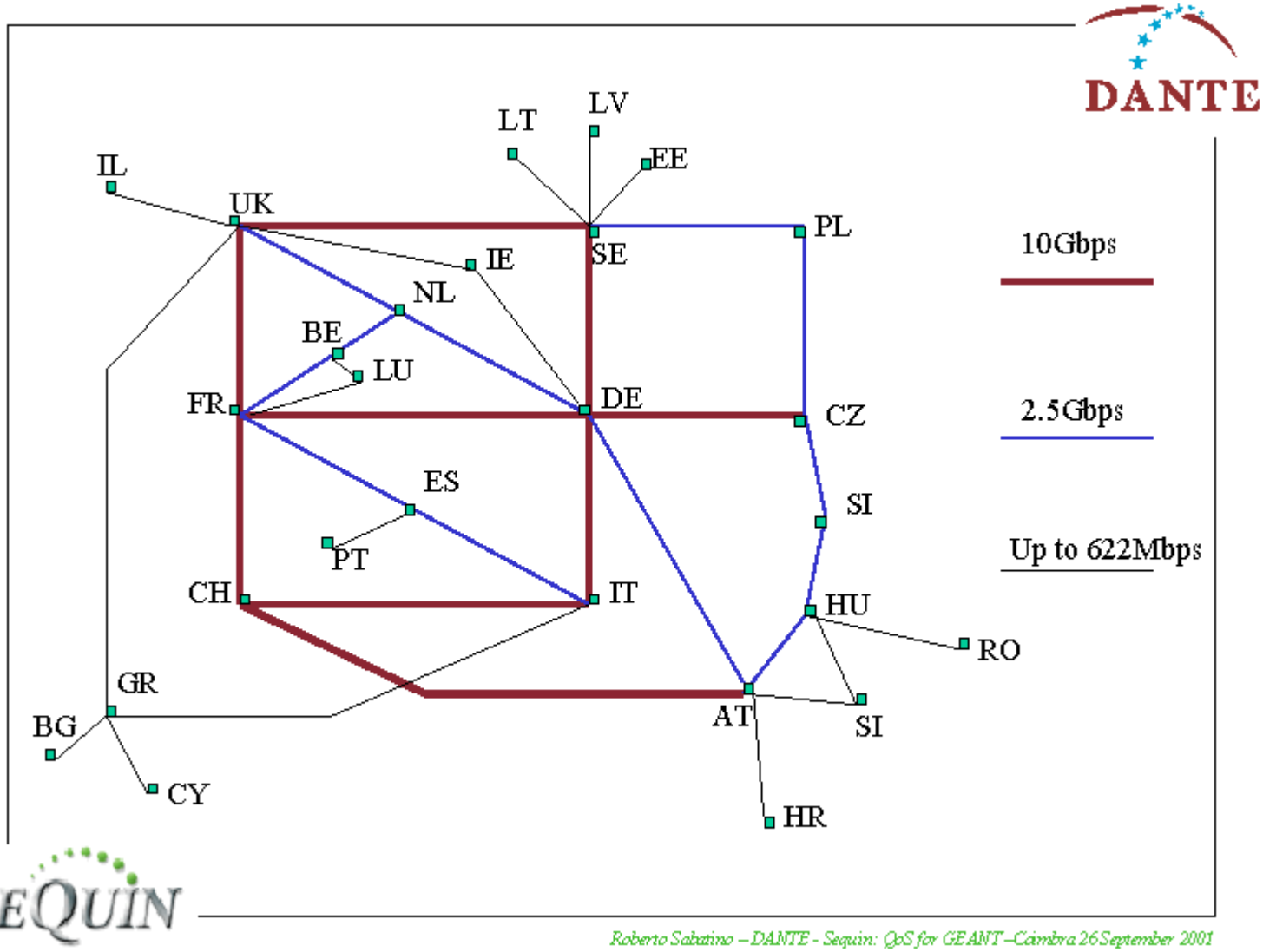
Qofis 2001
Coimbra
26 September 2001

Roberto Sabatino –DANTE
Mauro Campanella - GARR

SEQUIN

GEANT

- A 10 Gbps pan-European network connecting 27 NRENs
- Successor of TEN-155
- First circuits delivered now
- To be fully operational before 30 November 2001


LT
LV
EE
IL
UK
PL
10Gbps
IE
SE
NL
BE
LU
FR
DE
CZ
2.5Gbps
ES
SI
PT
Up to 622Mbps
CH
IT
HU
RO
GR
BG
AT
SI
CY
HR

GEANT services

- Best Effort IP
- Native multicast
- Support for QoS
- L2 VPNs

DANTE

SEQUIN

SErvice QUality across Independently managed Networks

SEQUIN

DANTE

SEQUIN PARTNERS

- DANTE
- DFN
- RENATER
- GARR
- GRNET
- PSNC
- SWITCH
- UKERNA

DURATION

•15 MONTHS from 1 Nov. 2000 + 3 months extension

SEQUIN

Work Packages

- WP1: project management
- WP2: Definition of QoS -> D2.1
- WP3: Definition of Testbed -> D3.1
- WP4: Dissemination and Implementation plan
  - D4.1, D4.2
- WP5: Proof of concept testing
- WP6: Testing in a user environment

Status

- D2.1 and D3.1 complete
- WP4, WP5 and WP6 in progress
  - http://www.dante.net/sequin *on-line*
  - http://www.switch.ch/lan/sequin *work in progress*

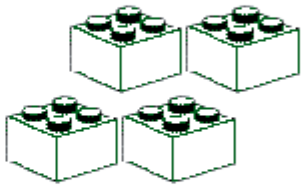
Definition of QoS

- Work of WP2 ->D2.1
- Top-down approach
  - Interview with international user groups on their requirements
- Bottom-up
  - Network performance parameters that may be influenced by configuration of equipment

QoS parameters

From users' requirements and technical considerations :

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



The set is common to IETF and ITU-T.
Naming and definitions are chosen to comply to RFC 2330 (Framework for IP Performance metrics) and follow the ongoing IPPM IETF working group work.

QoS parameters sample value ranges

	Single value (SV)	Short range (class 0)	Medium (class 1 interactive)	Wide range (class 2 non-interactive)
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)
ipdv	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)
Packet loss (Probability)	null	< 10^-4 (10^-3)	< 10^-3 (10^-3)	< 0.1 (10^-3)
Bandwidth (speed 64Kb/s)	Fixed value, greater than time to transmit one full MTU packet	N/A	N/A	a minimum of one full MTU size packet per second

Between parenthesis are ITU-T Y.1541 draft values, Class 3 (unspecified) is not shown

From User Questionnaire

QoS service	One-way-delay	ipdv	packet loss	bandwidth
Best effort	wide	wide	medium	wide
Very good (*Premium IP*)	medium	very small	very small	according to SLA
Prioritised Bandwidth (*IP+*)	medium	medium	medium	according to SLA
Guaranteed bandwidth	medium	medium	very small	single value

	One-way-delay	IPDV	Packet loss	bandwidth
Best effort	Unspecified	Unspecified	< 5%	Unspecified
Premium IP	distance delay + 50 ms	< 25 ms	negligible	according to SLA
IP+	distance delay +100 ms	<25-50 ms	< 2%	according to SLA

Which QoS service

Start with the "very good" service and call it "Premium IP":

- it satisfies all the users' requests
- it is "the best" service possible
- it maps to very high priority scheduling techniques available now
- it is similar to a "virtual wire"

IP Premium goal

Provision QoS for the European research users in the form of an end to end network 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 GÉANT networks

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

Not necessarily available everywhere on day-1, NRENs can join when ready

The implementation and Service Level Agreements have to match the current status of hardware availability and network topology

DANTE

Which QoS framework to use ?

- ✓ • Differentiated Services - RFC2475 -
- • Integrated Services - RFC 1633 -
- ✓ • Overprovisioning

SEQUIN

Premium IP Specification

- ⇒ Differentiated Services Architecture and use the expedited forwarding per hop behavior (EF PHB)
- ⇒ interface definition between domains that behaves as an EF PHB
- ⇒ do not starve best effort traffic (limited percentage of link capacity devoted to Premium IP)
- ⇒ initial provisioning structure: static, no dynamic signaling
- ⇒ IETF IPPM QoS parameters measurement framework
- ⇒ QoS parameters monitoring system is a key element

Basic principles

⇒ minimize number of actions per node
⇒ do not use a signaling protocol
⇒ modular approach that allows different implementation schemes at every hop or domain and allows domains to join the service when ready
⇒ Do not try to solve the most general problem, but rather develop a model that can be implemented in short time using available tools
⇒ ***Keep it simple***

Simplifying the actions for each node

DANTE

In principle, each node might perform an awful lot of tasks:

- admission control and classification

- marking
- policing
- scheduling
- shaping
- congestion control
- QoS rules propagation
- monitoring and accounting

SEQUIN

Admission control

Use the information in the IP header:

- IP source *and* destination
  (prefixes) as near to the source as possible

- the DSCP (or IP precedence equivalent value) along the path

- perform an optional, suggested, admission control based on AS source and destination at inter-domain links (safety measure)

- rules might be based on additional parameters, such as time-of-day

Admission control (continued)

DANTE

The consequences are:

- allowing the computation of total requested Premium IP capacity at each network node in the default case (and for main backup cases too)

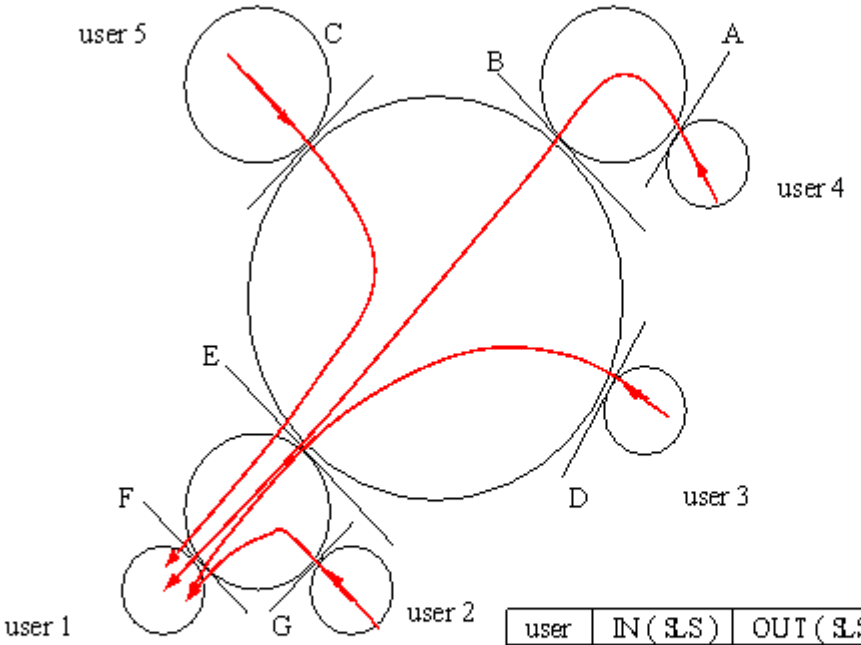
- short access list near users' premise (few users)

- simple control at backbones (IP addresses are not propagated)

- choosing destination aware service (next slide)

SEQUIN

Admission control (continued)



Destination Aware
Vs destination Unaware

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

Examining the tasks for each node

DANTE

In principle, each node might perform an awful lot of tasks:

- admission control and classification — always

- marking
- policing
- scheduling
- shaping
- congestion control
- QoS rules propagation
- monitoring and accounting

SEQUIN

Marking

- Mark each “EF” legal packet at first classification point
- Use the same DSCP value on all domains (Class selector 5 - decimal 40 [RFC 2474] to have interoperability with ToS-only capable hardware) - strongly suggested -
- valid DSCP coupled to invalid IP addresses *may* imply discard to allow easy debugging
- packets with other DSCP values are left untouched

Marking is mandatory at the first classification point, remarking is optional.

Policing

Microflow policing should be done as close as possible to the source according to agreed (through SLA) Premium IP capacity. This step is mandatory

Policing will be done using a token bucket. The depth of the token bucket will be two MTU close to the source and increase to 5 or more along the path if additional policing is required

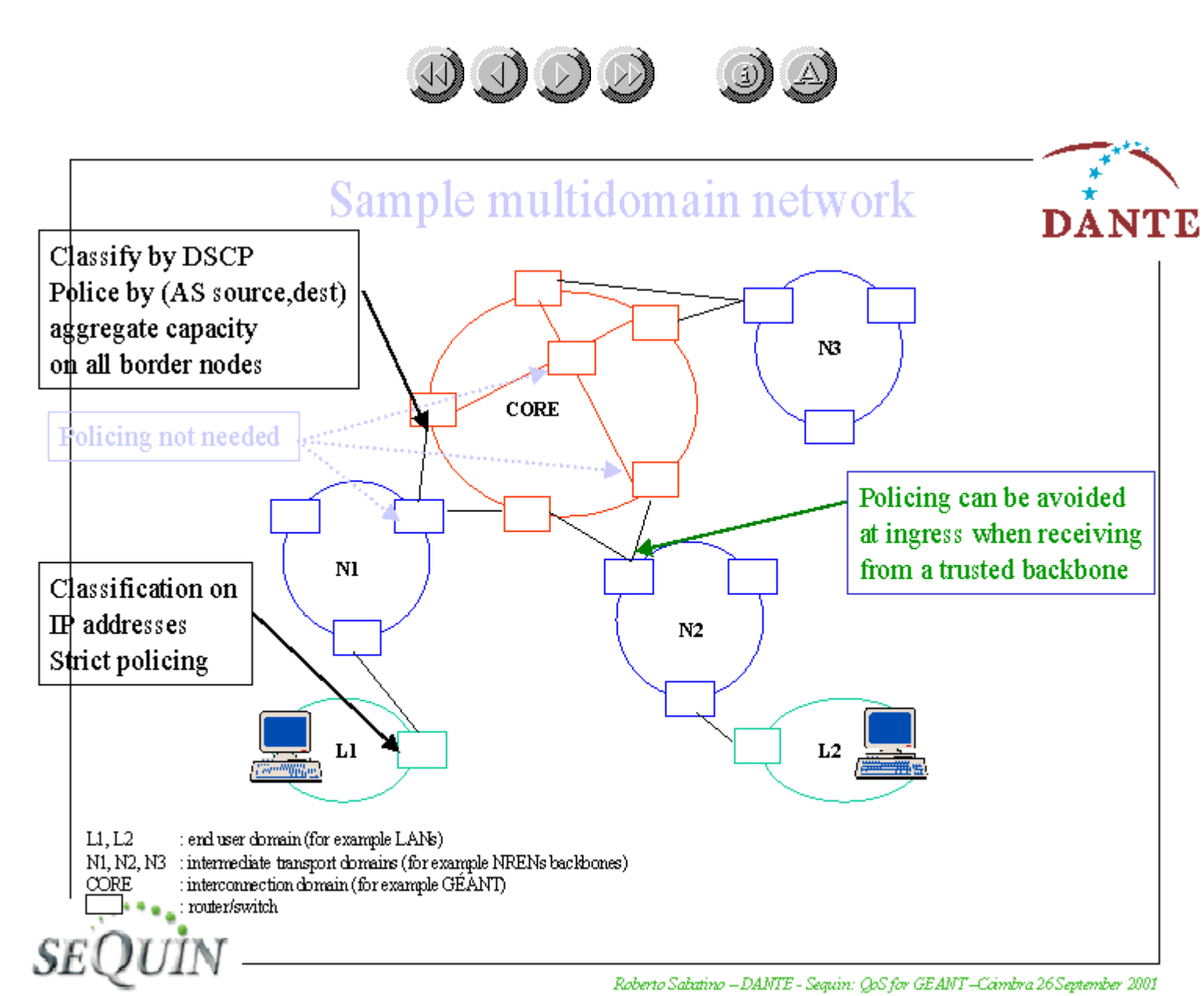
It is suggested to perform only one additional policing stage at the ingress to GÉANT from an NREN, with a larger aggregated capacity value than the sum of the agreements.

"Avoid unwanted packet loss" is the motto.

Policing (continued)

The additional policing stage at the ingress to GÉANT from an NREN serves the purpose of protecting Premium IP traffic from misconfiguration/DoS coming from a single source.

It creates virtual "pipes" for the aggregated Premium flows from each NREN to each other (when needed). The failure of one "pipe" does not influence the others.


Sample multidomain network
DANTE
Classify by DSCP
Police by (AS source,dest)
aggregate capacity
on all border nodes
Policing not needed
CORE
N3
N1
N2
L1
L2
Policing can be avoided
at ingress when receiving
from a trusted backbone
Classification on
IP addresses
Strict policing
L1, L2 : end user domain (for example LANs)
N1, N2, N3 : intermediate transport domains (for example NRENs backbones)
CORE : interconnection domain (for example GÉANT)
: router/switch
SEQUIN
Roberto Sabatino – DANTE - Sequin: QoS for GEANT – Cambra 26 September 2001

Examining the tasks for each node

DANTE

- admission control and classification — always
- marking — Selected locations
- policing — Selected locations
- scheduling — always
- shaping
- congestion control
- QoS rules propagation
- monitoring and accounting

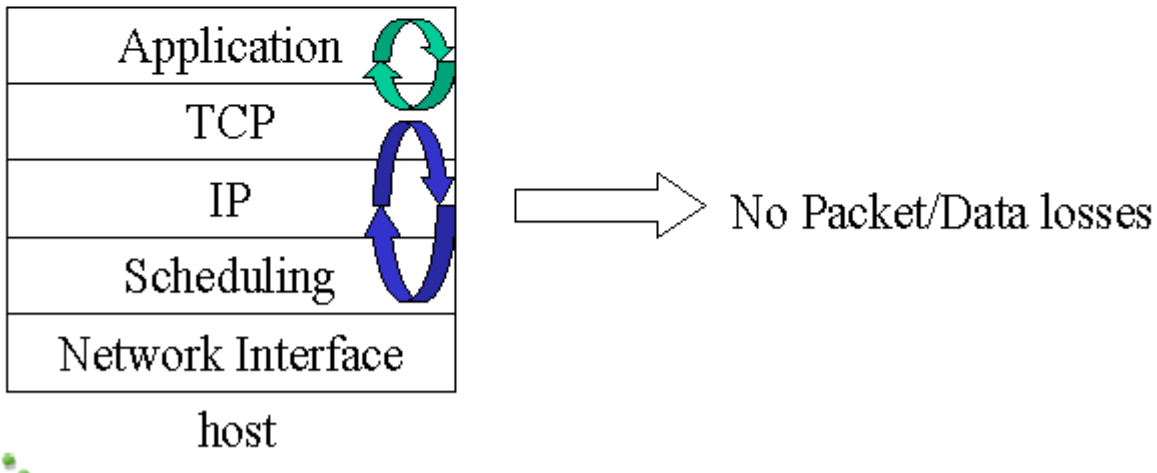
SEQUIN

Shaping

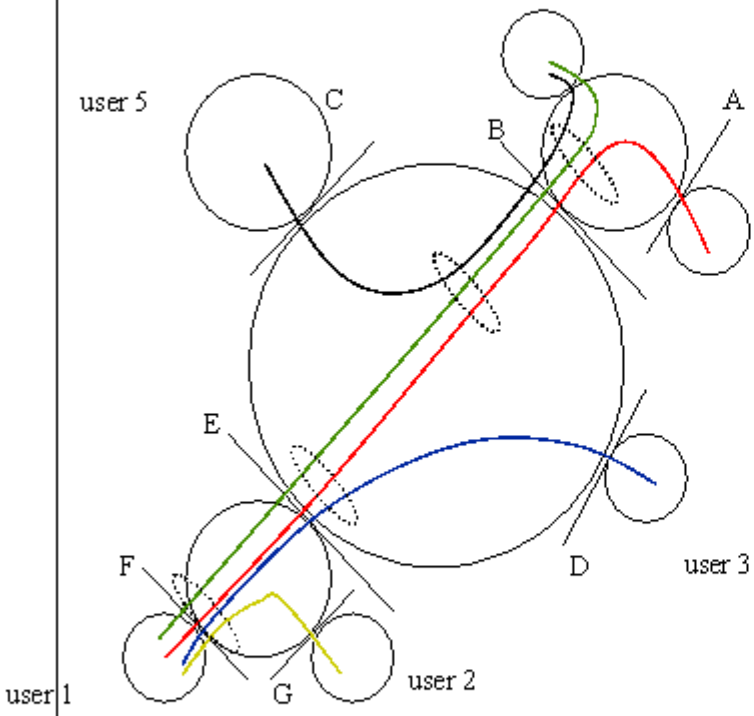
The sending source is required to shape the traffic it produces.

Shaping can be done by the network close to the user

Shaping inside the sending host itself is the preferred way, shaping by the network will in most case lead to packet losses



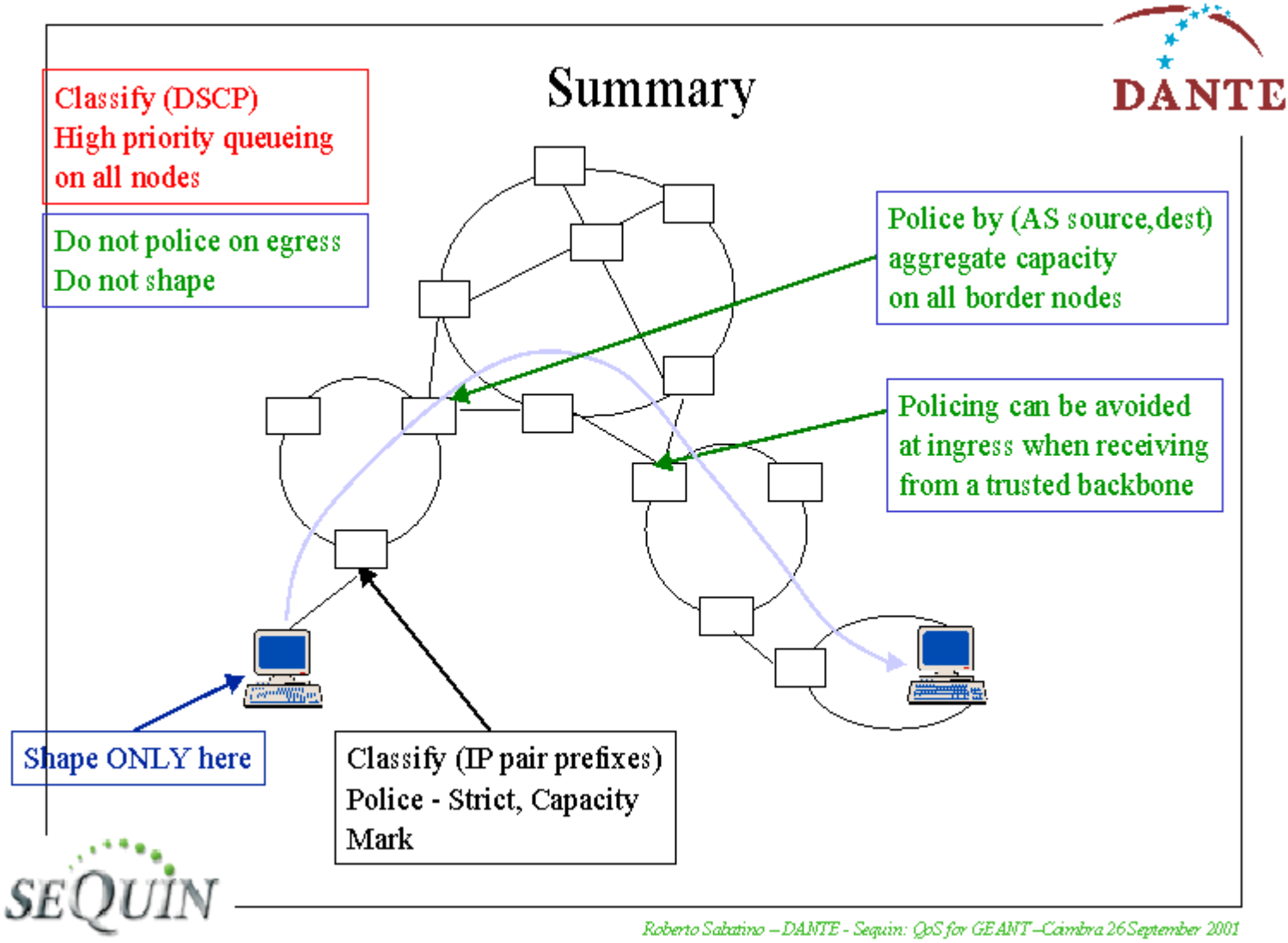
Shaping



Multiple aggregation-separation points and link speed changes.

Examining the tasks for each node

- admission control and classification — always
- marking — Selected locations
- policing — Selected locations
- scheduling — always
- shaping — NO Done by source
- congestion control — not needed
- QoS rules propagation — Selected locations
- monitoring and accounting — Selected locations


Summary
DANTE
Classify (DSCP)
High priority queueing
on all nodes
Do not police on egress
Do not shape
Police by (AS source,dest)
aggregate capacity
on all border nodes
Policing can be avoided
at ingress when receiving
from a trusted backbone
Shape ONLY here
Classify (IP pair prefixes)
Police - Strict, Capacity
Mark
SEQUIN
Roberto Sabatino – DANTE - Sequin: QoS for GEANT – Cambra 26 September 2001

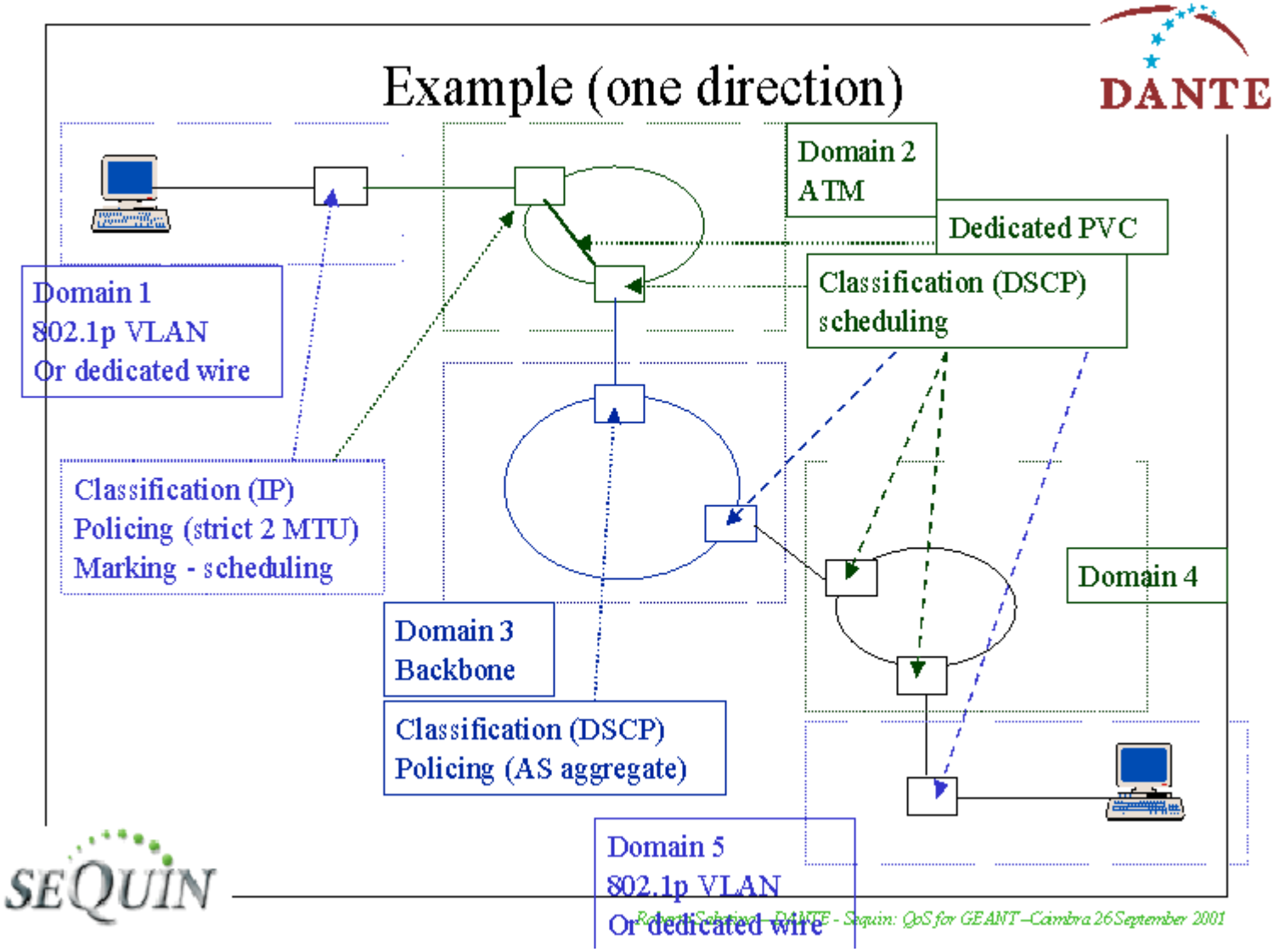
Grey Areas

Exact configuration of buffering and token bucket depth in routers. As a rule of thumb the token bucket depth can be assumed to be 1.2 * (Diffserv active interfaces on router)

Scalability
- the maximum amount of aggregated Premium IP capacity the network can offer
- hardware capabilities

Fast provisioning of the service

Widespread availability and tuning of "last mile" (LANs)


Example (one direction)
DANTE
Domain 2
ATM
Dedicated PVC
Classification (DSCP)
scheduling
Domain 1
802.1p VLAN
Or dedicated wire
Classification (IP)
Policing (strict 2 MTU)
Marking - scheduling
Domain 4
Domain 3
Backbone
Classification (DSCP)
Policing (AS aggregate)
Domain 5
802.1p VLAN
Or dedicated wire
SEQUIN
Robert Sabatino - DANTE - Sequin: QoS for GEANT – Cambra 26 September 2001

H.323 trial

- Selected users from CH, IT, FR
- Using production networks + GARR-G testbed
- Goals
  - verify user perceived quality
  - verify stability of production network
  - verify operational provisioning procedures
  - verify SLS monitoring techniques
- In November 2001

References

Sequin Deliverable D2.1 “Quality of Service definition”
http://www.dante.net/tf-ngn/SEQ-D2.1.pdf

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

GÉANT Deliverable D9.1 - Addendum 1 “Implementation architecture specification for the Premium IP service”:
http://www.dante.net/tf-ngn/D9.1-addendumv2.pdf

GÉANT Deliverable D9.1 - Addendum 2 “Service Level Agreement specification for Premium IP service”:
to be available soon in http://www.dante.net/tf-ngn/

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

Acknowledgements

- NREN Staff
- TF-NGN
- Spirent, who supports our work with 4 smartbits


DANTE
Thank you !
SEQUIN
Roberto S in: QoS for GEANT – Cambra 26 September 2001