QoS Experience on European Backbone

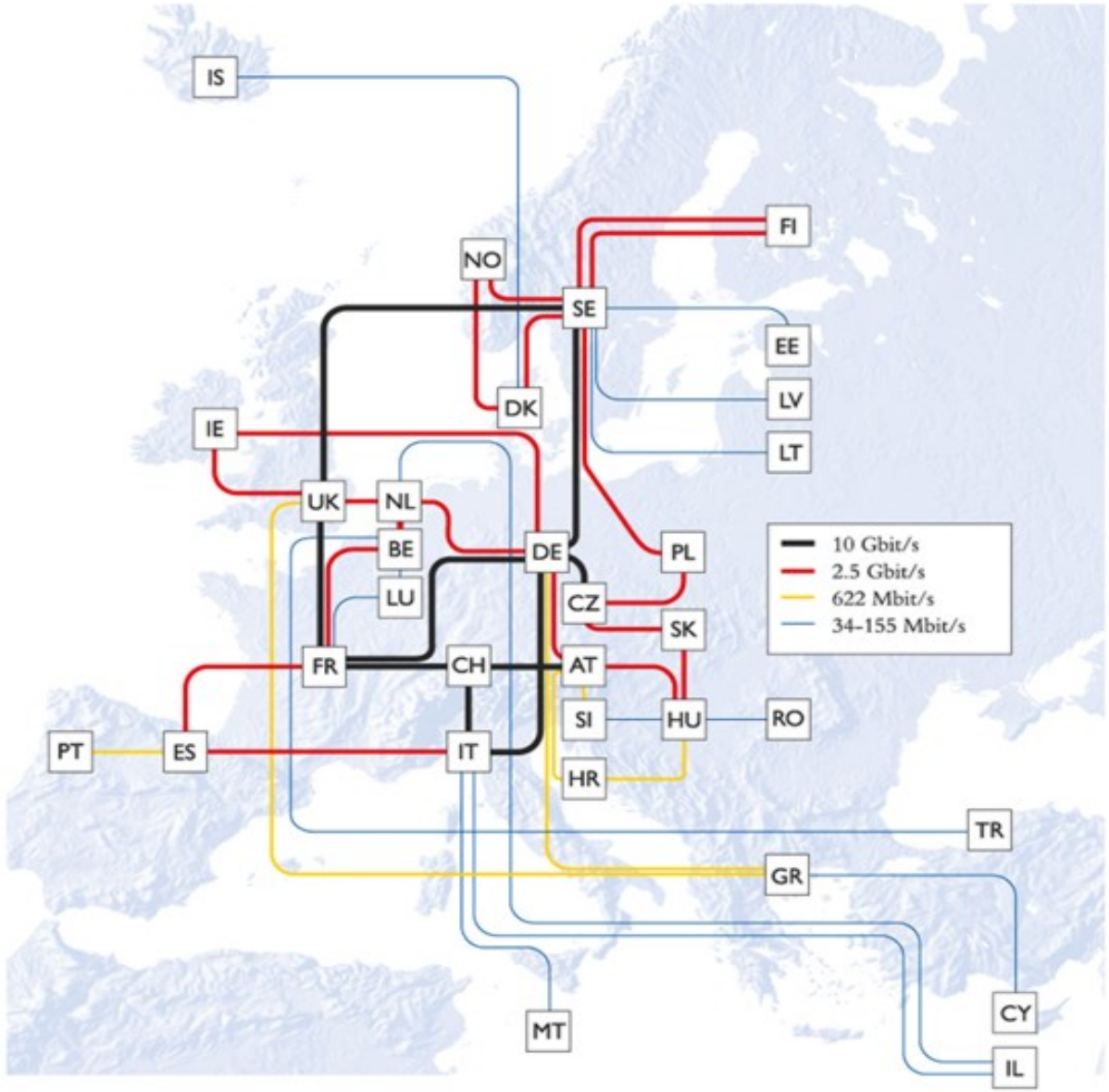
TNC 2003,

Zabgreb (Croatia), 20-05-2003

Nicolas Simar, Network Engineer
DANTE

Agenda

- GÉANT
- Services on GÉANT
- Premium IP
- Less than Best Effort
- Queuing on GÉANT and status
- Next steps

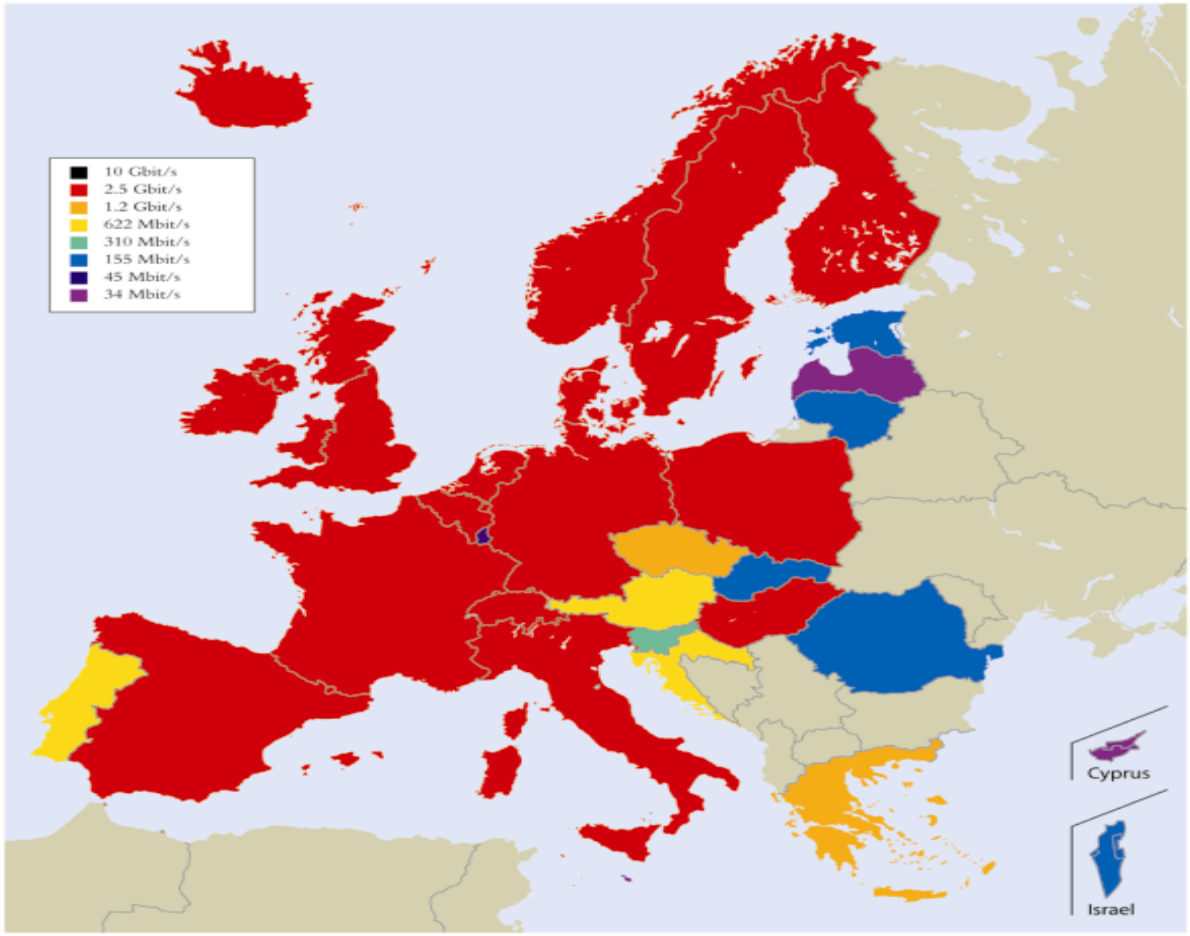


Multi-Gigabit pan-European Research Network

Backbone Topology December 2002

- Connecting 32 European Countries and 28 NRENs
- Backbone capacity in the range of: 34Mb/s-10Gb/s

Backbone Access Speeds-August 2002



GÉANT Services

- **Services available on GÉANT**
  - Best Effort (IPv4)
  - Multicast (IPv4)
  - Premium IP (IPv4)
  - Less than Best Effort (Scavenger) (IPv4).
  - IPv6 (Best Effort)

IP QoS Services on GÉANT

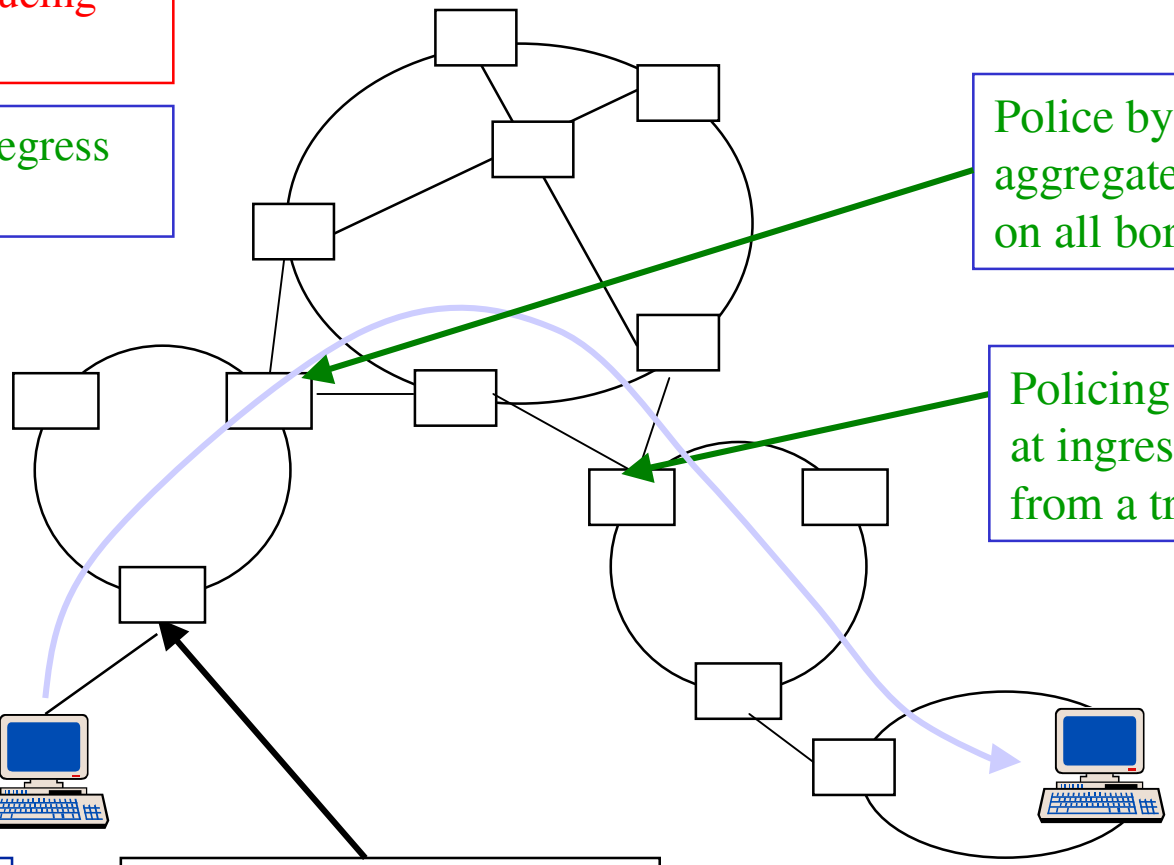
- **Premium IP**
  - Upper-bounded one-way delay
  - Upper-bounded IPDV
  - Negligible packet loss
  - Guaranteed capacity.
- **Less than Best Effort**
  - Class of traffic using the unutilised Best Effort and higher classes of service bandwidth.

Premium IP model

- **End-to-end service across multiple management domains**
  - Using diffserv, ATM CBR or over-provisioning(transition).
  - Packet tagged DSCP 46 (EF - 101110).
  - Destination aware service.
  - Packet with other DSCP (from other service) are left untouched.
  - Premium IP bandwidth limited to 10% of the link capacity.
    - Can cope with 20% in case of circuit failure.

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

Premium IP on GÉANT

- **Protection of authorised Premium IP traffic**
  - Under normal circumstances, the Premium IP traffic is limited at 10% of the circuit capacity.
    - 20% in case of another circuit failure
  - Bullet-proofing of all GÉANT accesses against “un-authorised Premium IP” traffic (tagged DSCP 46) on all the ingress interfaces.
    - If DSCP 46 packet arrives on GÉANT and is part of an un-authorised flow: the packet is not classify in the Premium IP queue and is re-tagged as Best Effort (DSCP 0).
    - if DSCP 46 packet arrives on GÉANT and is part of an authorised flow: check against policer according capacity mentioned in the SLA (in-profile accepted, out-of-profile dropped).

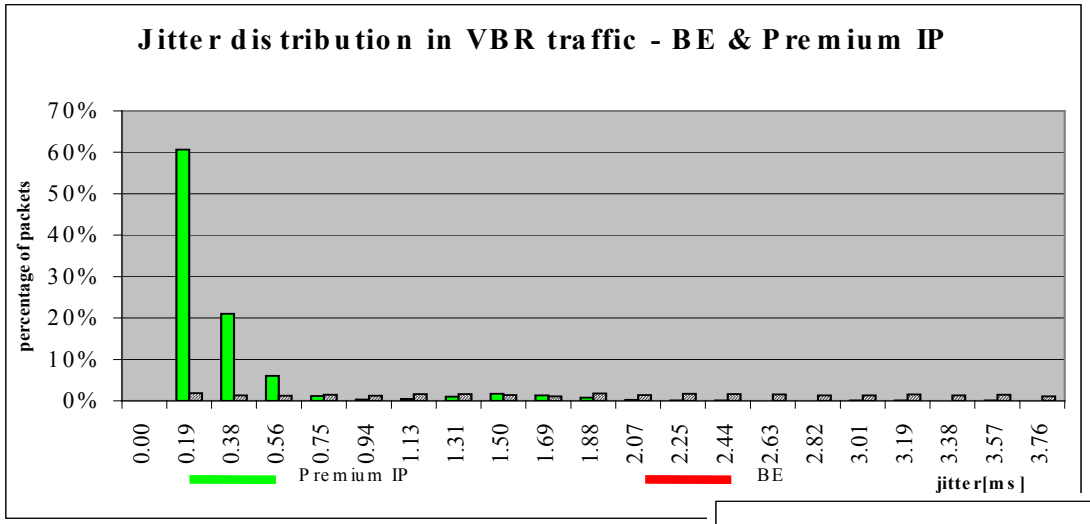
Premium IP on GÉANT

- **Protection of authorised Premium IP traffic** [cont]
  - per next AS rate-limitation (implemented by Juniper for GÉANT)
  - can also do source-destination IP addresses when NREN don't do it.
- **Trust the Premium IP traffic received from a GÉANT backbone interface.**
  - Traffic checked at the GÉANT ingresses.

Premium IP on GÉANT

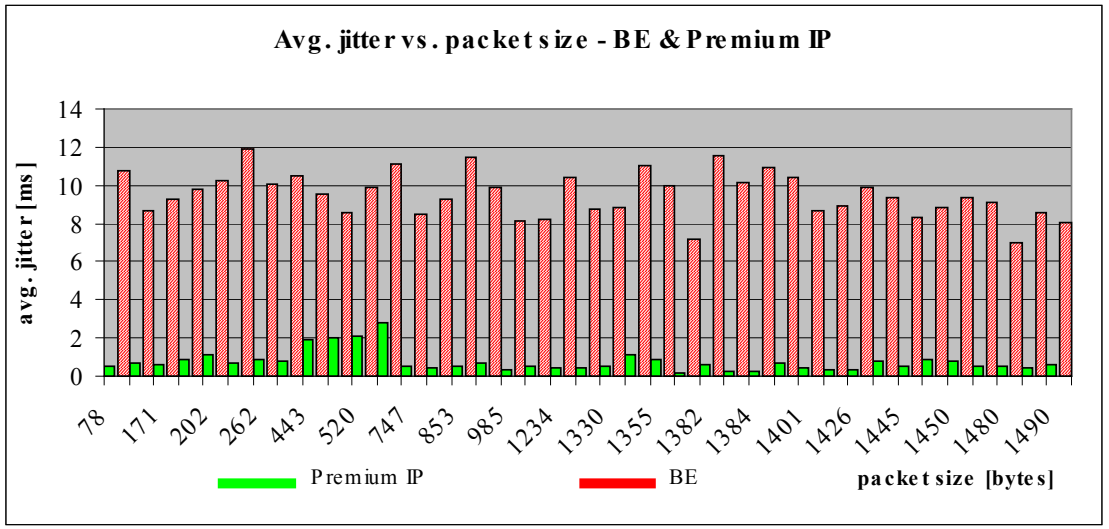
- **Queuing mechanism are configured on the backbone and access interfaces.**
  - 90% of the link capacity is allocated to the Premium IP queue (via WRR).
    - Don't forget that the amount of Premium IP traffic expected in the Premium IP queue is of 10% of the link capacity.
  - 5% for the BE and 5% for the network control.
- **Monitoring of the metric is very important to provide a proper service.**
- **When a end-to-end problem arise, it is very difficult to find out where the problem lies => PERT.**

Test result (end-to-end)



- Traffic sent between NRENs end-sites and crossing GÉANT.
- The packets were sent with a variable bit rate.
- Premium IP provisioned in various way in the domains crossed.

The Premium IP jitter is lower than the BE one and is independent of the packet size.



Less than Best Effort

- **LBE is a class of traffic using the un-utilised Best Effort and higher classes of service bandwidth.**
  - In case of competition for resources, the LBE traffic will de discarded before any Best-Effort or higher classes of traffic.
  - Use the DSCP 8 (001000) - same as Internet2 scavenger service.
- **Congestion on an interface due to LBE.**
  - Should be transparent to the BE or higher classes of services.
  - No BE or higher classes of services packet loss.

Less than Best Effort

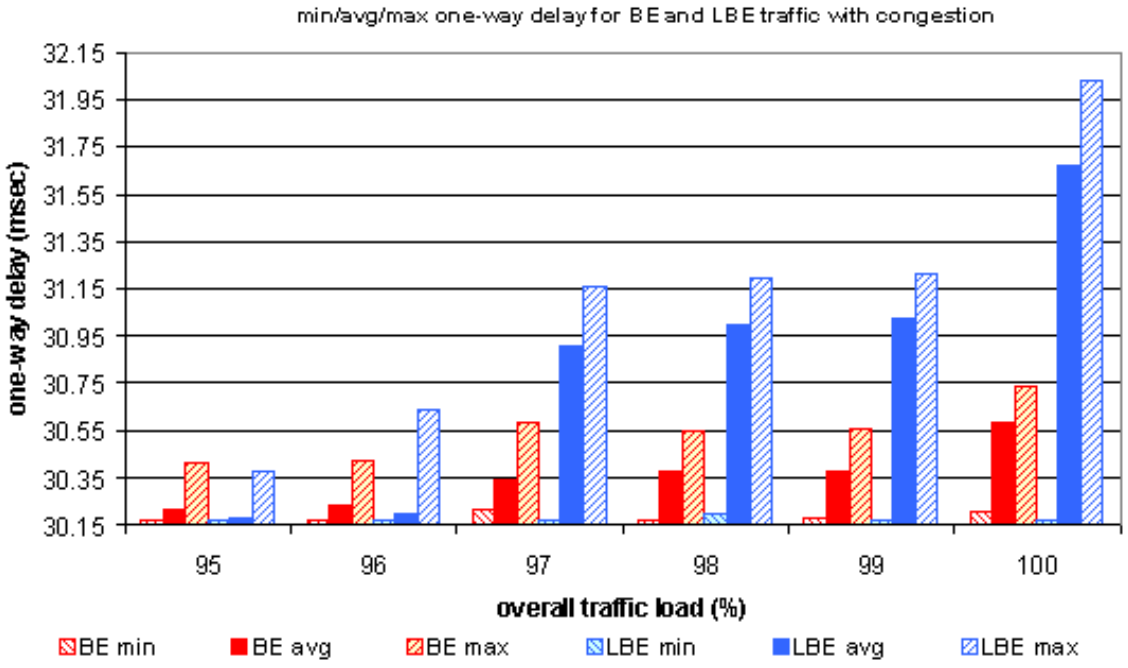
- **No end-to-end guarantees.**
  - No metric needed to quantitatively describe the service.
- **Can be supported on one interface.**
  - Anywhere else, the LBE tagging should be passed transparently.
- **Application scenarios:**
  - mirroring, test traffic, protection of research traffic from student dormitory one.

LBE Queuing technique

- **For algorithm with bandwidth shared assignment, as Weighted Round Robin and Weighted Fair Queuing, a very small bandwidth share is allocated to the LBE queue.**
  - Typically 0% or 1% (depending of scheduling implementation).

Measurement with congestion

- **One-waGdelaG**
  - Increase of LBE maximum one-way delay of 1.5ms.
  - Increase of BE maximum one-way delay of 400µs.



The traffic load (in %) express the amount of traffic generated by the SMB STM-16 interface.

ER2002 Demo - VLBI - dataGRID

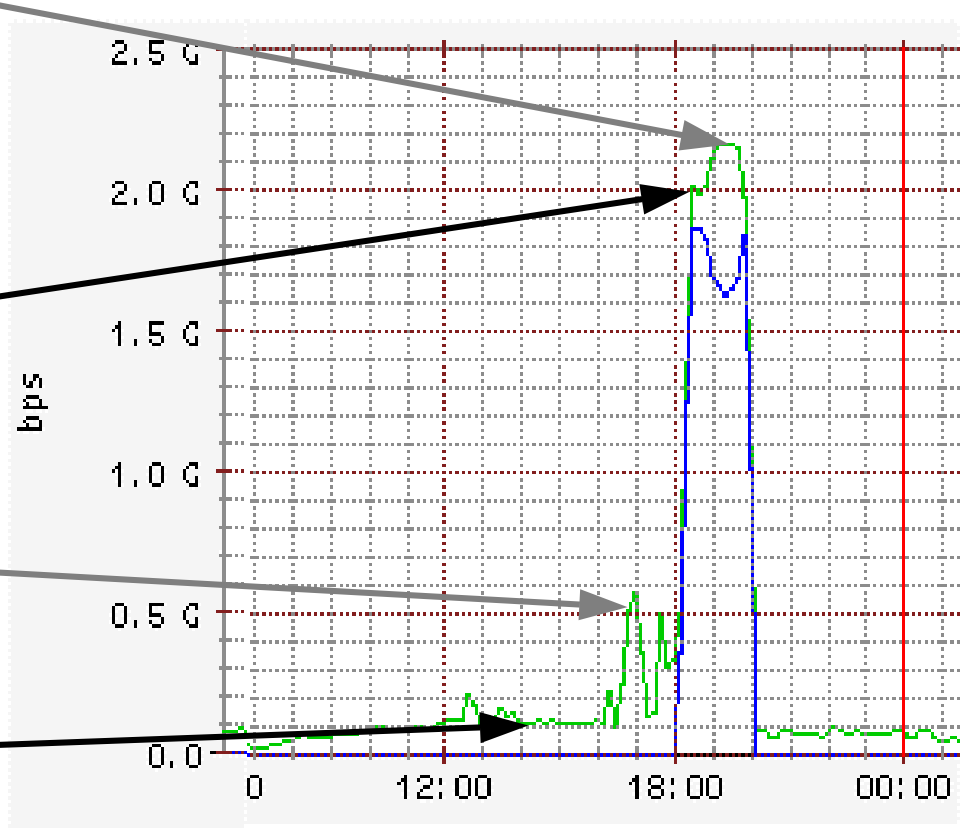
- Normal Traffic +
- Radio Astronomy Data +
- Less Than Best Effort
  (2.0 Gbit/s)

- **Normal Traffic +**
- **Less Than Best Effort**
  (2.0 Gbit/s)

  - Normal Traffic +
  - Radio Astronomy Data
    - 500 Mbit/s

    - **Normal Traffic**

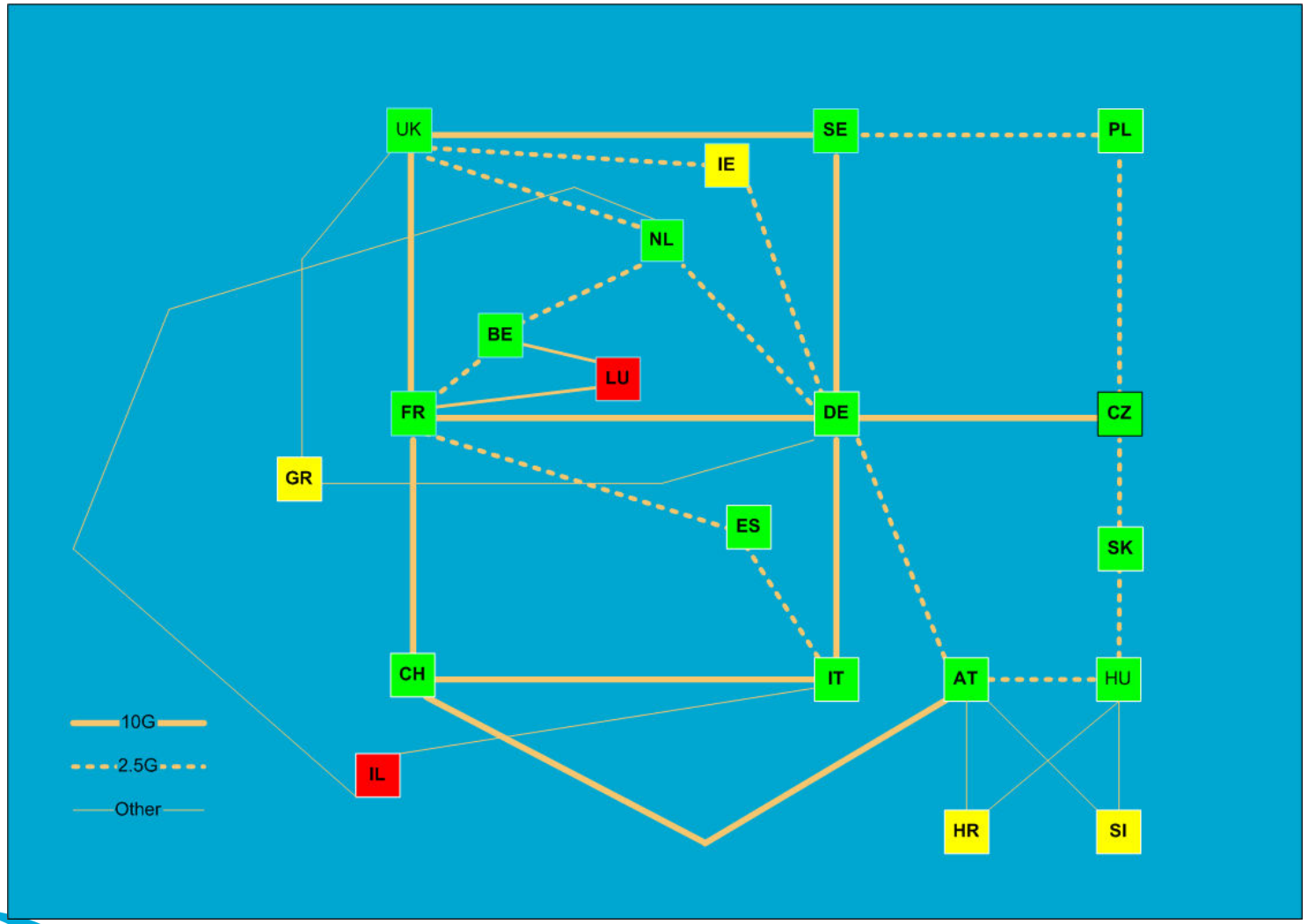
UK-NL STM-16 link utilisation



QoS configuration on GÉANT

- **The configuration has completed on most of the GÉANT routers, allowing Premium IP, BE and LBE to co-exist.**
  - The routers where the three services have been enable are represented as green on the following map.
  - The routers coloured yellow are Juniper routers where "old" FPCs have been re-used from TEN-155. These old FPC's that do not allow for the full functionality of QoS.
    - As such BE is not ideally protected by LBE and the bandwidth is effectively shared. Premium IP only is supported.
  - The red routers are the Cisco 7k's re-used from TEN-155. These routers don't have VIP powerful enough to provide any QoS functionality. Premium IP and LBE not supported.

QoS configuration on GÉANT



Useful tool

- **Feature on the NANOG traceroute to allow the discovery of the DSCP changes along a path.**
  - Very useful as it has allowed to discover most of the configuration problems on GÉANT and between GÉANT and NRENs.

```
[root]# ./traceroute -t 184 193.171.2.1
traceroute to 193.171.2.1 (193.171.2.1), 30 hops max, 40 byte
packets
 1  css7-ATM4-0-0-101-dmsk.man.poznan.pl (150.254.160.62)  1 ms  1
ms  1
ms
 2  150.254.163.118 (150.254.163.118)  2 ms  2 ms  2 ms
 3  z-pozmanu-oc3.poznan-gw.pol34.pl (212.191.127.49)  2 ms  2 ms  2
ms
 4  pol-34.pl1.pl.geant.net (62.40.103.109)  2 ms  2 ms  2 ms
 5  pl.cz1.cz.geant.net (62.40.96.45)  22 ms (TOS=0!)  22 ms  22 ms
 6  cz.de1.de.geant.net (62.40.96.38)  30 ms  30 ms  30 ms
 7  de1-1.de2.de.geant.net (62.40.96.130)  30 ms  30 ms  31 ms
 8  de.at1.at.geant.net (62.40.96.5)  43 ms  43 ms  43 ms
 9  aconet-gw.at1.at.geant.net (62.40.103.2)  43 ms  43 ms  43 ms
10  193.171.2.1 (193.171.2.1)  45 ms *  45 ms
```

Next steps

- **Premium IP reservation tool**
  - Q4 2003, Premium IP requests will be done the NREN via a web-interface.
- **Inter-domain monitoring**
  - Performance monitoring activity

    http://www.dante.net/tf-ngn/perfmonit/
  - Aim at monitoring delay, IPDV, packet loss (to start with) across several domains.
- **PERT**
  - Performance Enhancement Response Team

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

QoS implementation

- Find the correct queuing parameters (for each service)
- "bullet proof" all ingresses for Premium IP
- Once configure, pretty stable (no need to change anything).
- http://www.dante.net/nep/geantqos/

Acknowledgements

- Spirent for a loan of SmartBits 600s (STM-16 interfaces)
- Juniper for the technical support