QBone and potential impact of IP/optical

Stanislav Shalunov, Internet2

Las Vegas, 2000-01-19

Talk Overview

- What is Internet2?
- What is DiffServ?
- What is QBone?
- How can QBone and IP/optical play together?

Internet2

Internet2, led by over 180 U.S. universities working in partnership with industry and government, is developing and deploying advanced network applications and technologies, accelerating the creation of tomorrow's Internet.

Areas of Internet2 activities

This is a very high-level overview. Color-highlighted areas will be expanded upon.

- Applications
- Middleware
- Advanced networks
- Network engineering

Internet2 Advanced Applications

Advanced network applications allow researchers to collaborate and access information in ways not possible using today's Internet, e.g.:

- Tele-immersion
- Virtual laboratories
- Digital libraries
- Distributed instruction

Middleware

Capabilities such as authentication, authorization and accounting will allow advanced applications to operate seamlessly among many organizations.

- PKI activities: fPKI, PKI-for-NGI
- Shibboleth: web-based authentication
- Medical middleware: issues specific to medical schools and health service
- Directories: education-specific aspects

Advanced networks

National, regional and campus networks provide end-to-end high-performance required by advanced applications.

- Abilene: nation-wide backbone
- vBNS
- Internet2 GigaPOPs
- Internet2 Universities

Network Engineering

Not only faster, but *smarter* networks.

- Multicasting
- IPv6
- Quality of Service

Multicasting and IPv6 operational in native mode in Abilene.

Internet2 Quality of Service

"The Holy Grail of computer networking is to design a network that has the flexibility and low cost of the Internet, yet offers the end-to-end quality-of-service guarantees of the telephone network."
—S. Keshav, in *An Engineering Approach to Computer Networking*, 1997

Internet2 must enable classes of applications that demand a richer set of services than is provided by the current "best-effort" Internet. Such applications need to be able to request and receive assurances from the network of certain end-to-end transmission parameters. These include bandwidth, delay, maximum packet loss rate, and delay jitter.

Focus on the most demanding applications

- Engineering efforts concentrated around *virtual wire* concept
- Once the most demanding applications' needs are met, the rest is easier
- Model application: tele-immersion
- A more modest application: interactive videoconferencing

Tele-immersion

"Tele-Immersion [. . .] will enable users at geographically distributed sites to collaborate in real time in a shared, simulated environment as if they were in the same physical room. This new paradigm for human-computer interaction is the ultimate synthesis of networking and media technologies and, as such, it is the greatest technical challenge for Internet2."
—National Tele-Immersion Initiative

What this application needs: very high bandwidth (hundreds of Mbps to infinity), low latency, very low jitter, almost no loss. Essentially, a very fat pipe all to itself.

We want such applications to work over IP.

Interactive Videoconferencing
High Fidelity Audio

The requirements are similar to that of model application, just not as strict. (Remarkably, bandwidth is much lower.)

- Coast-to-coast RTT is about 70ms
- Comfortable interactivity only gives one about 100ms
- Very hard to get rid of buffering in drivers and hardware

$\Rightarrow$ **No leeway for network jitter.**

Virtual Wire

- Informally: Within an IP network, for two given end-points, everything *looks* as if traffic from the sender to the receiver goes over a dedicated circuit

- Of course, certain jitter is inevitable, but one must be able to compute a worst-case bound for it

- More formally: `draft-ietf-diffserv-pdb-vw-00.txt`

Differentiated Services

[Apologies if this is all-too-familiar to you.]

- Objective: provide scalable, relatively coarse, differentiated treatment within an IP network
- Use a new small field in IP header (Differentiated Services Code-Point–DSCP, 6 bits, taken out of TOS in IPv4 and out of Traffic Class in IPv6) to “paint” packets in a few colors
- Have routers recognize packet color and assign special Per-Hop Behavior (PHB) to each color
- Per-Domain Behavior (PDB) is a concatenation of PHBs
- End-to-end traffic treatment is a concatenation of PDBs

Expedited Forwarding

- EF is the PHB used to create Virtual Wire PDB and "Premium" service

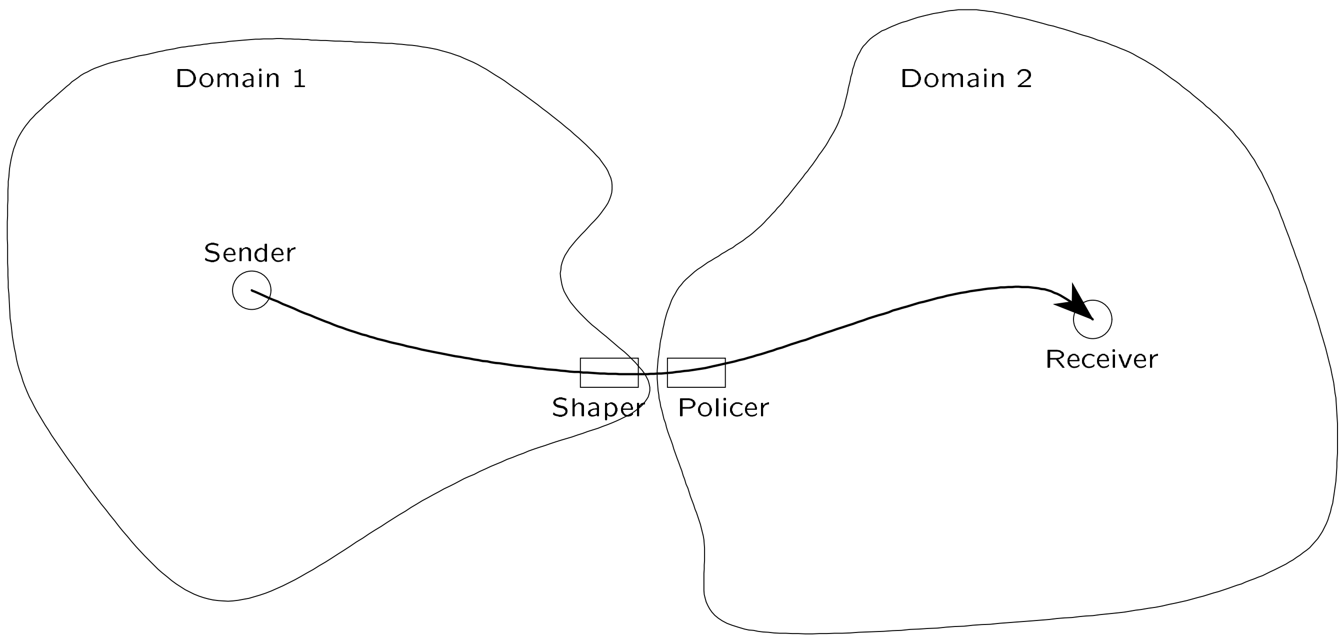
- Precise definition is a subject of controversy within IETF

- Current consensus within IETF DiffServ WG seems to be `draft-charny-ef-definition-01.txt`: *"The intuitive meaning of the packet scale rate guarantee is that as long as there are EF packets in the node, we would like the j-th EF packet of length L(j) to depart no later than L(j)/R seconds after the (j-1)st departed (here R is the configured rate of the aggregate). (L(j)/R is simply the time that it would take to forward the j-th packet at the EF-configured rate R.) Were this always to occur, the EF packets would be forwarded perfectly at the configured rate."*

QBone

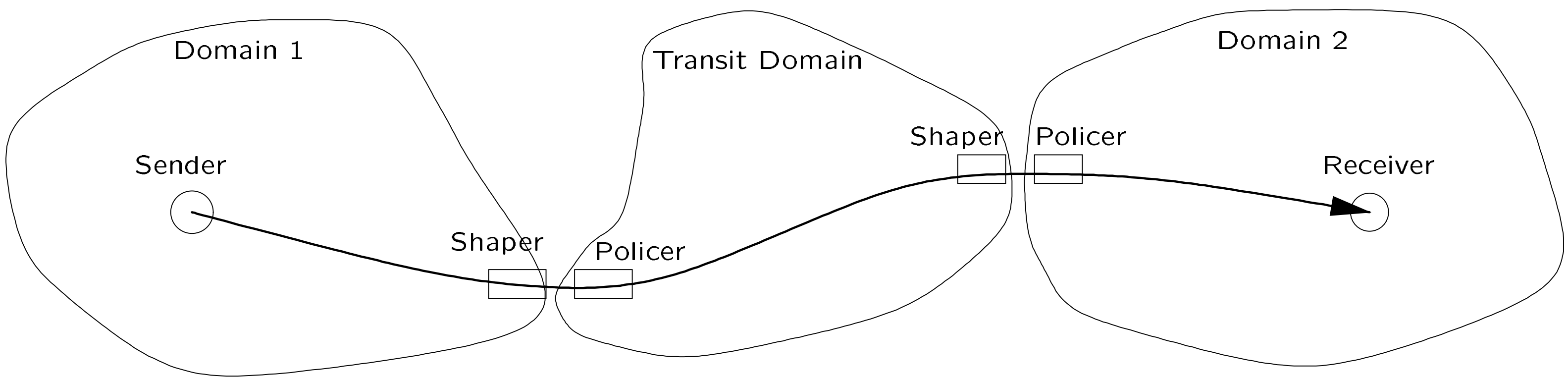
- Internet2 QoS initiative
- Concentrated on prividing “Premium” service based on EF PHB semantics
- Inter-domain interoperability is an important goal
- Scalable because of only considering DiffServ aggregates (not individual flows)
- Measurements are important to provide user assurance and fault isolation functionality

View of two QBone domains



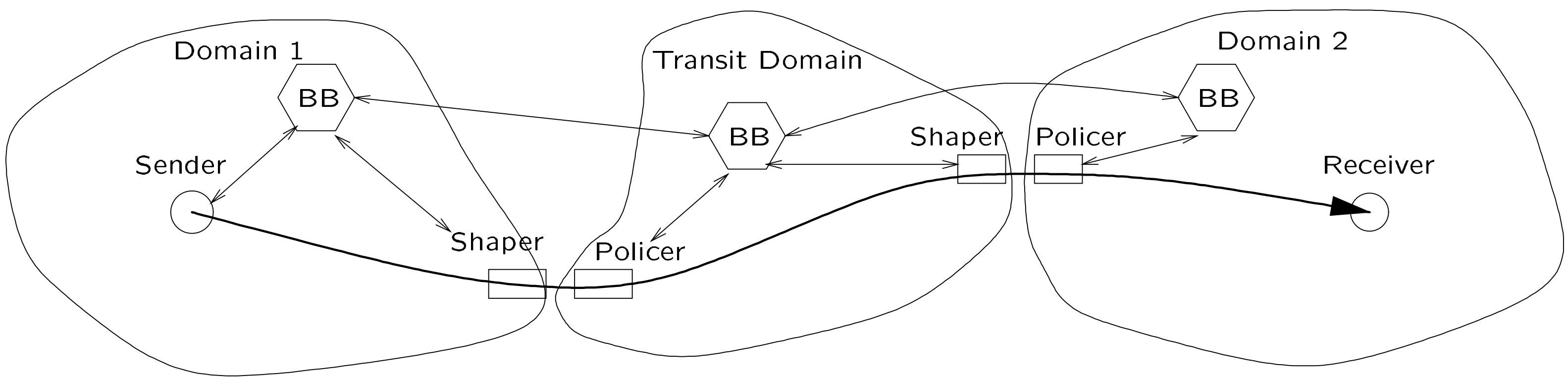
Domain 1 and Domain 2 exchange EF-marked traffic. Domain 2 agrees to accept certain amount of EF-marked traffic at the given peering point. Domain 1 shapes to given profile and Domain 2 polices to enforce the shaping. Non-conforming traffic is dropped.

Transit QBone domain



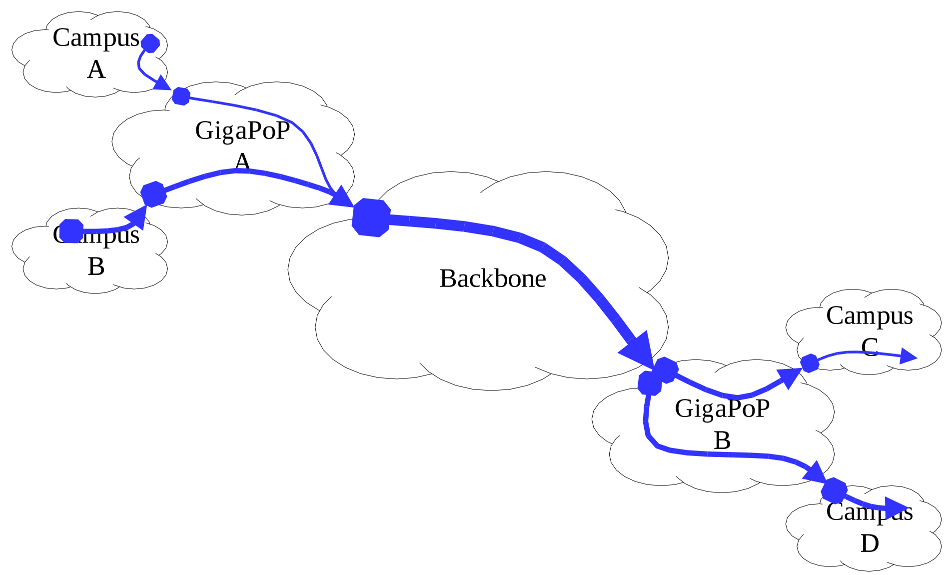
Here, Transit Domain agrees to accept and expeditiously forward certain amount of EF-marked traffic from Domain 1 destined for Domain 2. Domain 2 agrees to accept certain amount (not less than the first) from Transit Domain destined to itself.

Bandwidth Brokers allow dynamic reservations



Here, everything is configured dynamically. Each domain has a Bandwidth Broker (BB). BBs speak in an unspecified way to hosts within same domain, and to BBs of peers using BB protocol. BBs configure border routers.

Example of QBone path aggregation



Premium flows from Campus A to Campus C and from Campus B to Campus D are aggregated within Abilene Backbone. But GigaPOP B has two distinct policers for these flows.

QBone peering and intra-domain treatment

A QBone domain must provide virtual wire-style PDB (so, peering interface is defined), but it's free to implement it in any way it likes internally:

- Large pure IP domains will probably want to use DiffServ inside
- Smaller domains might use IntServ with RSVP signalling
- MPLS (probably in the form of MPLS-TE) might be a supplement
- Some might provision ATM VCs
- Some might use λs
- Where bandwidth is very cheap, overprovisioning is an option

QBone Service Level Agreement

- All agreements are bilateral—just as peering agreements are
- Consider domains A and B, where A peers with domains A_1, $A_2, \ldots A_n$ and B peers with B_1, $B_2, \ldots B_m$. Then, in a static case, a Service Level Agreement (SLA) between A and B may include a number of statements of the following form: "A agrees to accept x_i Mbps from B for A_i" or "B agrees to accept y_j Mbps from A for B_j"
- It may include cost and be dynamic
- May include special provisions ("up to 100 Mbps will always be available to any destination within the U.S. on week's notice"; "reservations over 10 Mbps made less than an hour in advance will be priced twice higher", etc.)

QBone Service Level Specification

- Work exists on formal SLSes: `draft-tequila-diffserv-sls-00.txt`, `draft-salsano-aquila-sls-00.txt`
- But SLS can be quite complex, and we have little experience with even informal specification
- For now, QBone SLSes are informal
- Once BB protocol is finalized and experience gained with using it, we may move forward with formalization
- We feel DiffServ WG meeting at 49th IETF largely shared this point of view

Main Practical Implementation Difficulties

1. EF PHB implementations are far from perfect: may introduce additional latency, jitter

2. Policer implementations (QPPB+CAR-style features) have limitations: token bucket depth, performance

3. Once you enable preferential treatment for a single path, you must have policers on every trust boundary, including every peering point, while some peerings may have older line cards that just don't support this

How optical networking could enhance QBone

- Large transit aggregates of EF-marked traffic within the backbone could simply use a λ

- If number of QoS peers is low, no problems with number of λs

- Provides nearly ideal Virtual Wire PDB implementation without EF PHB (addresses practical difficulty 1)

- *Reduces the need for border policing!* (Addresses practical difficulty 3)

Challenges to using IP/optical to enhance QBone

- Increased complexity of electronic and optical networking co-existence: at a boundary electronic forwarding required, since DSCP defines treatment

- Should Bandwidth Broker be able to automatically provision λs? If not, a full mesh of QoS λs is required, which may not necessarily be an economical way to expend resources; if yes, full automation of provisioning on optical equipment part is required

For more information. . .

Stanislav Shalunov ⟨shalunov@internet2.edu⟩

Internet2: http://www.internet2.edu/

Abilene Premium Service:
http://www.internet2.edu/abilene/qos/

QBone Information: http://qbone.internet2.edu/

QBone Architecture: http://qbone.internet2.edu/arch/

Questions?