Internet2 QBone
—Building a Testbed for Differentiated Services

Authors
Benjamin Teitelbaum (ben@internet2.edu)
Internet2 (UCAID) / Advanced Network & Services
Susan Hares (skh@merit.edu)
Merit Network
Larry Dunn (ldunn@cisco.com)
Cisco Systems
Vishy Narayan (vnarayan@mail.arc.nasa.gov)
NREN/NGI Program, Raytheon/NASA Ames Research Center
Robert Neilson (rneilson@bcit.bc.ca)
British Columbia Institute of Technology
Francis Reichmeyer (franr@iphighway.com)
IPHighway

Benjamin Teitelbaum

is senior engineer with Internet2 and Advanced Network & Services. He chairs the Internet2 Quality of Services Working Group, which is responsible for the QBone initiative- an effort specify and deploy an Architecture for inter domain IP differentiated services. Additionally, Ben leads the QoS engineering group for the high-performance Abilene backbone network. At advanced Network & Services, Ben contributes to the design of the Surveyor IP performance measurement platform and infrastructure.

Outline

- Internet2
- Requirements for Internet2 QoS
- Bandwidth Broker
- options for achieving Resource Allocation
- QBone Architecture
- Security Consideration
- Deployment
- Conclusion

Internet2

- The Internet2 project is a partnership of over 130 U.S. universities, 40 corporations and 30 other organizations.
- One of the primary technical objectives of Internet2 has been to engineer scalable, interoperable, and administrable inter-domain quality of service (QoS) to support an evolving set of new advanced networked applications.

Important advanced application area for Internet2:

- Distance learning
- Remote instrument access and control
- Advanced scientific visualization
- Networked collaboratories

Internet2 QoS

-The study beginning in the fall of 1997 by Internet2 QoS Working Group

Two additional technical requirements:

- Any viable approach must scale, allowing core routers to support thousands of QoS-enabled flows a high forwarding rates.
- Any viable approach must interoperate, making it possible to get well-defined inter domain QoS assurances by concatenating the QoS capabilities of several independently configured and administered network clouds.

Internet2 QBone

-QBone initiative launched in late 1998

Build an open and highly-instrumented testbed for interdomain differentiated services.

Experimental services will be deployed, debugged, analyzed, and refined by networking engineers and researchers working in close collaboration with the users and developers of new advanced networked applications.

-Networks participating in the QBone initiative currently include:

vBNS,	Abilene,
ESNet,	NREN,
CA*Net2,	SURFnet,
TransPac,	MREN,
NYSERNET,	NCNI,
Texas gigaPoP,	

as well as numerous universities and labs.

Internet2 DiffServ

- Differentiated services (DiffServ) approach to QoS gain significant interest in the IETF as a lightweight alternative to the RSVP-integrated services (IntServ) architecture.
- DiffServ design a simple architectural framework for QoS that can provide a variety of scalable, end-to-end services across multiple, separately administered domains, without necessitating complex inter-provider business arrangements or complex behaviors in forwarding equipment.
- At a workshop in May 1998 the working group presented DiffServ as the architecture best suited to meeting the QoS needs of Internet2, and began a dialogue that culminated in rough consensus around the need to build an interdomain testbed to explore and advance DiffServ.

Internet2 DiffServ

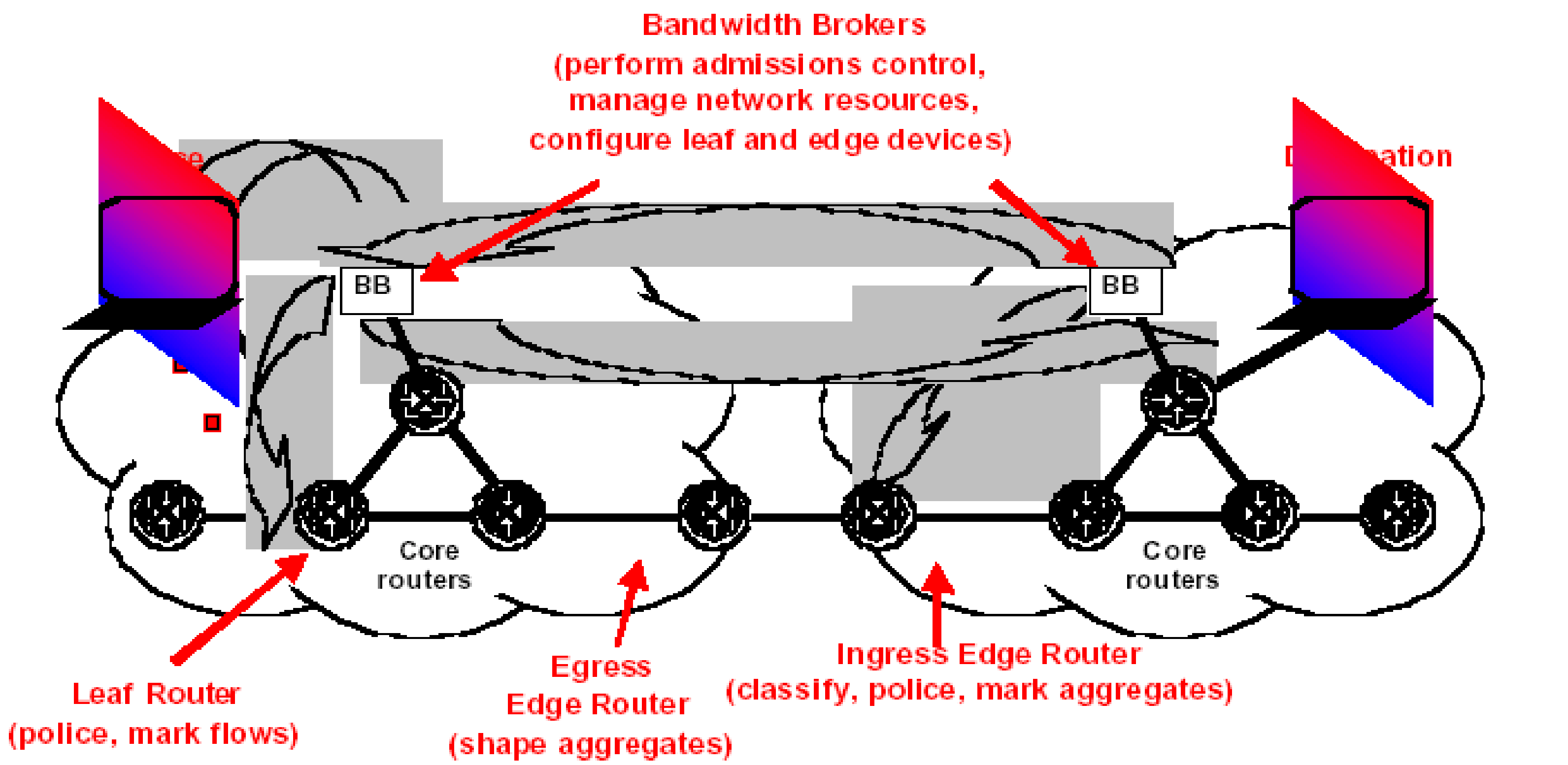


Figure 1. The Differentiated Services Architectural Framework

Internet2 Bandwidth Broker

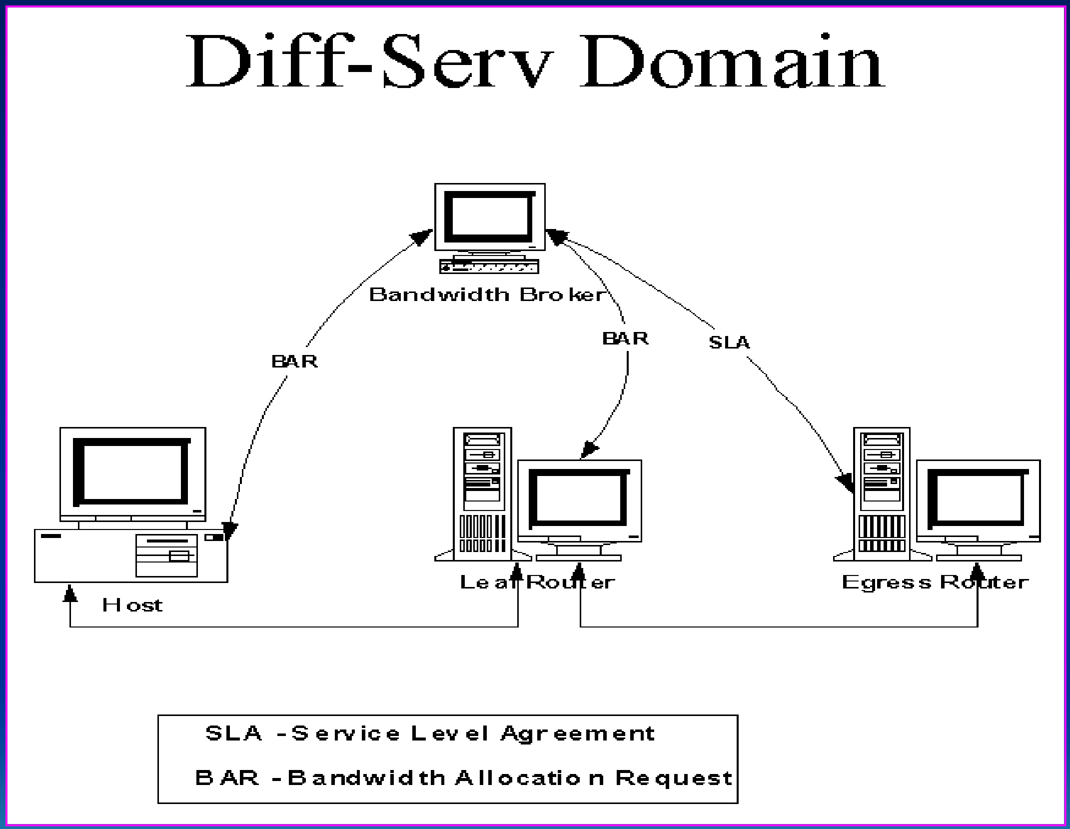
The function of BB:

-To automate the process of SLS negotiation and admission control, and to configure network devices correctly to support the provisioned QoS services, The BB is responsible for ensuring that resources within the DiffServ domain and on links connecting adjacent domains are properly provisioned and not over-subscribed.

-The responsibility of the bandwidth broker including mechanism for signaling QoS requests between hosts and routers, or between DiffServ-enabled domains, and also mechanism for managing the allocation and utilization of DiffServ resources within a domain.

-A bandwidth broker maintains information relating to the SLSes that are defined between a DiffServ domain and its customers. Customers include local users as well as the adjacent networks that provide connectivity to other parts of the Internet. The BB uses this SLS information to configure the routers in the local DiffServ domain, and to make admission control decisions.

Bandwidth Broker



Bandwidth Broker

•Has a database that is used to maintain the Service Level Agreements (SLAs) and the Bandwidth Allocation Requests (BARs)

•Provides a command line interface to add, update and delete the SLAs and the BARs.

•An SLA request contains:
Customer ID, Service type,Rate, Priority, Drop probability, start date and end date.

•A BAR request contains:
SLA ID, Router ID, Source IP, Source Port, Destination IP, Destination Port, protocol, Rate, start date and end date

Bandwidth Broker

SLA Request:

BB maintains the SLAs in the database and deletes them after the period is over.

When a new SLA request is made, the BB compares the request service type, rate and drop probability to the SLAs already in place.

-If no comparable request exists, the BB sends a configuration message to the egress router to setup a new class with the parameters specified. It also assigns a DSCP for the class.

-If a similar request exists, no configuration message are sent.

-An SLA ID is returned to the customer, which will be used in the BAR requests.

Bandwidth Broker

BAR Request:

•BB maintains the list of BAR requests in the database and deletes them when the period is over.

•When the BB receives a BAR request, it identifies the SLA corresponding to the BAR request (using the SLA ID) and determines whether the request can be allowed

-If the request succeeds, the BB sends a configuration message to the leaf router (whose router ID is specified in the BAR) with the DSCP that is assigned for the flow to do the policing and marking. It returns a "SUCCESS" message to the host.

- If the request fails, the BB sends a "FAIL" message to the host.

Bandwidth Broker

BB work procedure

-source must signal its local BB to initiate a service reservation before marked packets from a data source are admitted to a DiffServ domain,.

-If the service reservation is admitted locally, the BB may initiate an end-to-end reservation request along the chain of BBs in the DiffServ networks to be traversed by the data flow.

-When a network-wide admission control decision has been made, the BB will configure the routers in the DiffServ domain to support the requested service profile.

-The bandwidth broker allows separately-administered DiffServ domains to manage their network resources independently, yet still co-operate with other domains to provide dynamically allocated end-to-end QoS.

Bandwidth Broker

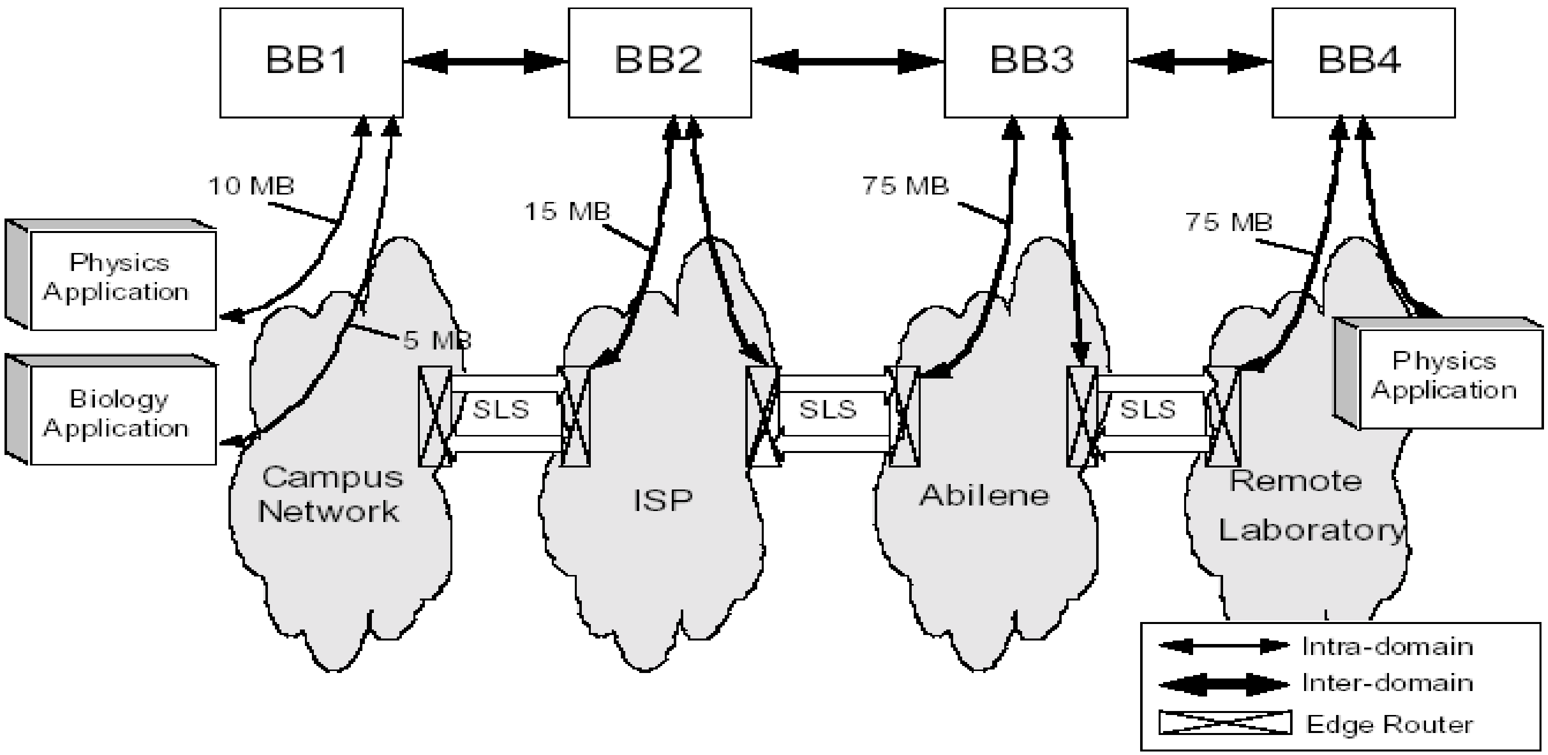


Figure 2. Service Level Agreements on Bandwidth for Academic Institutions

Options for Achieving End-To-End Resource Allocation

Several possible methods for building DiffServ functionality are enumerated below. Each selects a particular balance among:

- which device does packet marking,
- how much signaling information is required,
- the expected frequency of signaling,
- degree to which resource allocation for a flow or aggregate of flows is recognized end-to-end.

Options for Achieving End-To-End Resource Allocation

weak assurances, minimal Administrative and control-plane activity,

stronger assurances ,more administrative and control-plane overhead.

Do nothing

Layer-2 treatment in the campus, static inter-domain bandwidth allocation

Host DS field marking, no signaling

Host DS field marking, no signaling, some flow-recognition near edge

Local signaling, static inter-domain provisioning

Single-ended signaling, with inter-BB communication

Double-ended signaling, inter-BB communication

RSVP

Options for Achieving End-To-End Resource Allocation

Layer2 treatment in campus, static interdomain bandwidth allocation:

Not quite DiffServ

- give packets differentiated treatment via layer-2 marking,
- no explicit DS field marking is done,
- no dynamic signaling is required,
- some local resource allocation exists,
- links between adjacent DiffServ domains are monitored,
- expanded as necessary to give adequate performance.

Options for Achieving End-To-End Resource Allocation

Host DS field marking, no signaling:

Minimalist DiffServ

- A host mark packets with a particular DSCP,the remaining resource provisioning within and between networks is static.
- Layer-3 devices (routers) might be configured in a variety of static ways:
  - from a single behavior aggregate always being given preferential treatment (to the possible exhaustion of bandwidth for best effort traffic);
  - to configuring a proportion of resources (e.g. output bandwidth) at each layer-3 hop for each behavior aggregate or group of behavior aggregates;
  - to more full-blown metering (measurement), policing (distinct handling of outofprofile packets), and output link resource (bandwidth) allocation for each behavior aggregate or group of behavior aggregates.

Options for Achieving End-To-End Resource Allocation

Host DS field marking, no signaling, some flow-recognition near edge:

Extension by adding the feature that some form of flow recognition occurs near the edge.

- Manually configured resource commitments to particular behavior aggregates, flow aggregates, particular flows.
- Once a packet is analyzed and handled (at some level of granularity) at the edge, the packet is subsequently treated only as part of a larger aggregate, as indicated by its DSCP, rather than requiring the host to do it.

Options for Achieving End-To-End Resource Allocation

Local signaling, static inter-domain provisioning:

Dynamically signal for resources

- Only local DiffServ domain knows about the dynamic resource requests.
- A bandwidth broker and policy server might apply administrative policy as to which applications are allowed to generate flows that receive preferential treatment, and dynamically keep track of intra-domain commitments.
- Layer-3 devices might be reconfigured by the BB as new resource commitments.
- The links across DiffServ domain boundaries are still statically provisioned.

Options for Achieving End-To-End Resource Allocation

Single-ended signaling, with inter-BB communication:

Extend by keeping the notion that a host or application might express needs to an intra-domain BB, and adding the notion that BBs in different DiffServ domains communicate with each other.

Options for Achieving End-To-End Resource Allocation

Double-ended signaling, inter-BB communication:

RSVP is used to signal resource requirements in the source DiffServ Domain. Such RSVP messages are tunneled through intermediate DiffServ domains, without elements in Diffserv domain acting on them directly. The RSVP messages, upon arrival at the destination network, are used by destination network to allow more precise destination resource allocation.

QBone Architecture

QBone Premium Service

- Make interdomain, peak-limited bandwidth assurances with virtually no loss, and with virtually no delay or jitter due to queuing effects
- Expedited Forwarding (EF) per-hop forwarding behavior
- QPS reservation {source, dest, route, startTime, stopTime, peakRate, MTU, jitter}

  Token rate = peakRate

  Bucket depth = MTU (Maximum Transmission Unit)

  Low loss, low latency, low jitter

QBone Architecture (cont.)

Measurement Architecture

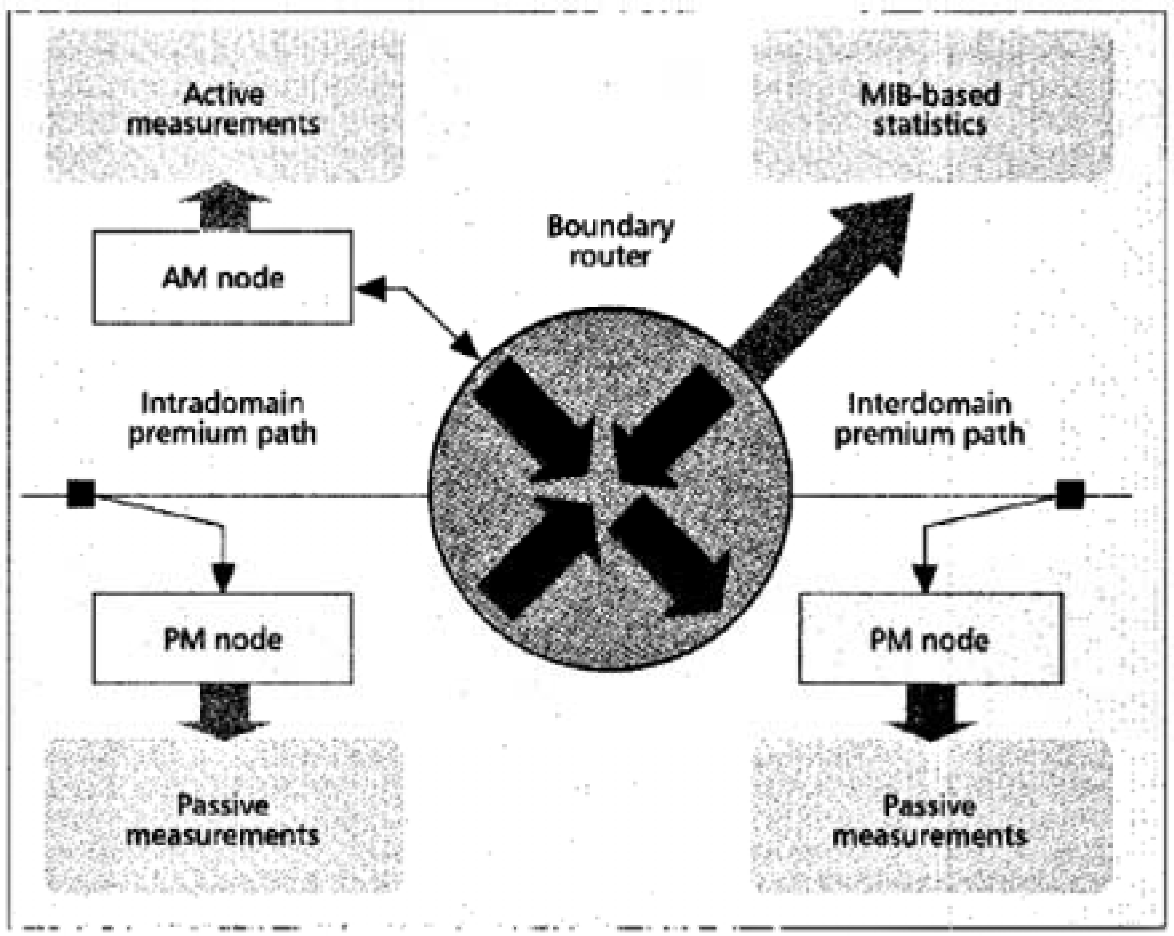
- Each QBone domain must collect and disseminate a basic set of QoS measurements
- Measurement Points and Paths:

  Active measurement

  Passive “sniffing”

  SNMP style polling

QBone Architecture (cont.)



■ Figure 3. *An example QBone measurement node configuration.*

QBone Architecture (cont.)

Measurement Architecture (cont.)

- Required Metrics (3 classes)
  - Active (1st class)
    - IETF IPPM one-way packet loss metrics
    - Instantaneous one-way packet delay variation
    - Periodic traceroutes for each behavior aggregate
  - Passive (2nd class)
    - EF and BE load in packet/sec and bits/sec
  - Polling (3rd class)
    - Link bandwidth in IP bits/sec
    - EF commitment in IP bits/sec
    - EF reservation load in IP bits/sec

QBone Architecture (cont.)

Measurement Architecture (cont.)

- Suggested Metrics
  - EF and BE interface discards
  - One-way packet delay
  - End-to-end burst throughput
- Dissemination Architecture
  - A web site for disseminating and presenting its measurements
  - MRTG-style summary plots
  - Raw measurement data

Security Considerations

- DOS Attacks
  - Altering the DiffServ fields
  - Inject packets with the DiffServ field set to a codepoint to get enhanced service
- Intradomain Reservations
  - Globus environment
  - Once authenticated, use token or encrypted certificate

Security Considerations (cont.)

- Interdomain BB and End-to-End Reservations
  - Peer identity
    - Bilateral peering (trust) relationship
  - Link
    - IPSec
  - Data
    - Resource Allocation Request (RAR)

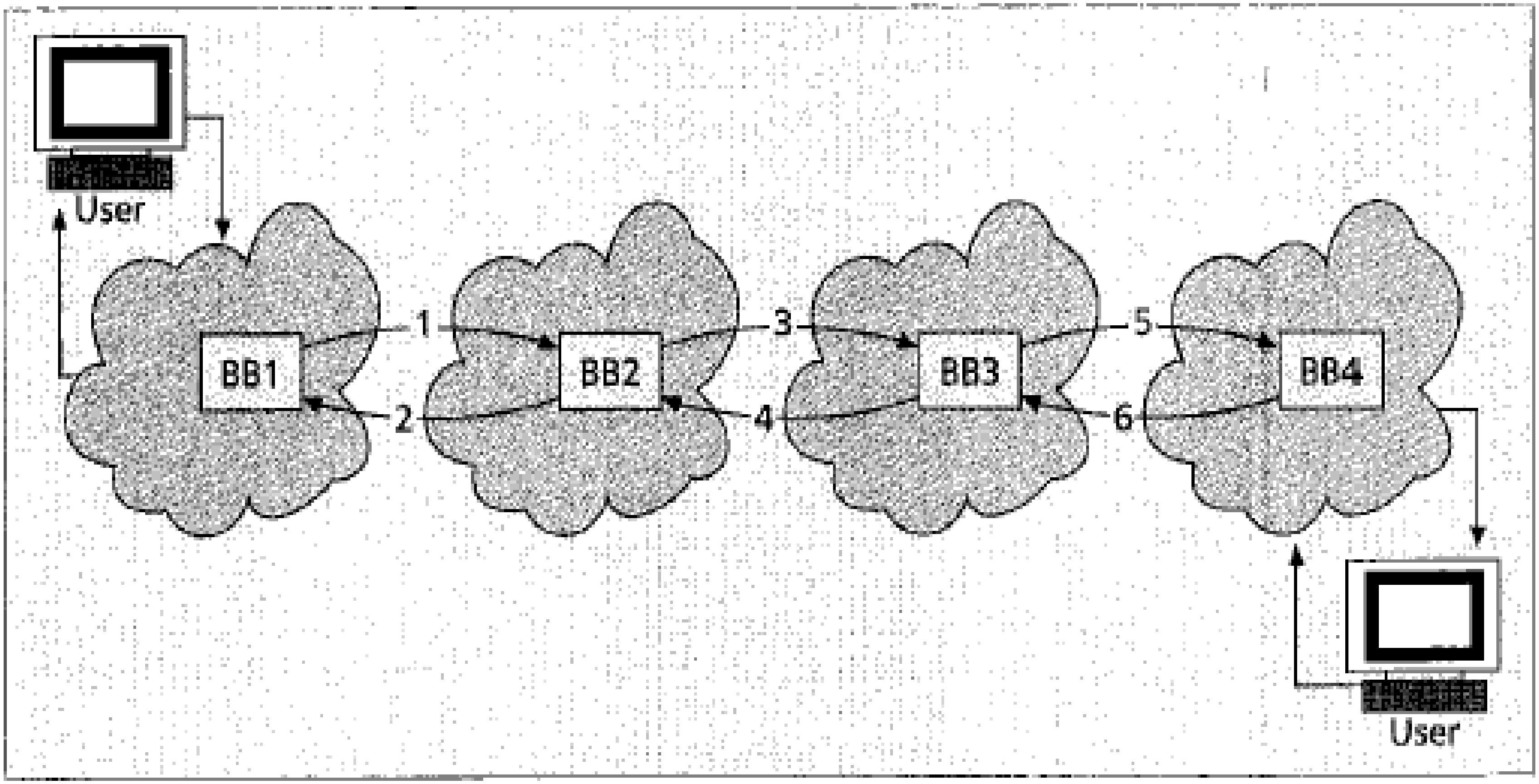
Security Considerations (cont.)

- Interdomain Issues for Ingress Routers
  - Ensure incoming packets are in profile
- Interaction of Non-DiffServ with DiffServ Domains
  - Physical control devices
  - IPSec within DiffServ domain

Deployment Plans and Bandwidth Broker Trials

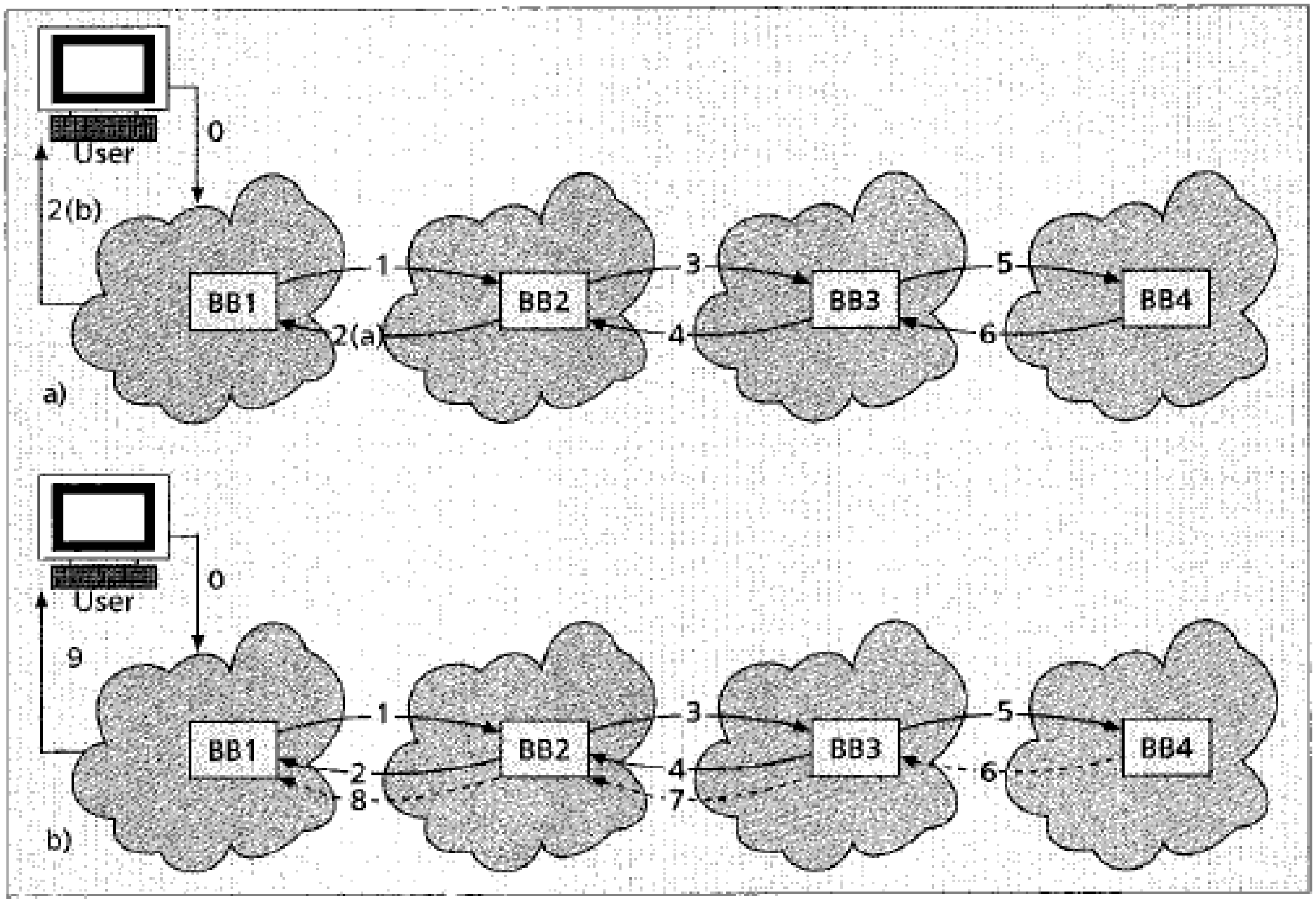
- Initial Deployment
  - Initial Deployment
    - Implement host DS field marking
    - Reservation: long-lived, manual configured
- Bandwidth Broker Deployment
  - Phase 0: Local Admission
    - Single DiffServ domain
  - Phase 1: Informed Admission
    - Query resource on downstream domains

Deployment Plans and Bandwidth Broker Trials (cont.)



■ Figure 4. *An example of signaling in an end-to-end model.*

Deployment Plans and Bandwidth Broker Trials (cont.)



■ Figure 5. *Immediate response examples: a) successful setup; b) failure in domain 4.*

Deployment Plans and Bandwidth Broker Trials (cont.)

- Bandwidth Broker Deployment (cont.)
  - Phase 2: Dynamic Admission
    - Outstanding Issues
      - Appropriate granularity of SLS adjustment
      - The need to couple the BB signaling protocol to interdomain routing
      - The impact of different reservation loads
      - The ability to preempt one reservation fro a higher-priority reservation
      - Techniques to extend DiffServ to multicast flows

Conclusions

- QBone
  - Will be first wide-area test of the evolving DiffServ architecture
  - Will be first experimental deployment of an interdomain DiffServ signaling protocol
  - Aims to start a process that will open the horizon for new advanced networked applications to flourish