An Architecture for Differentiated Services

Van Jacobson

(van@ee.lbl.gov)

Network Research Group
Berkeley National Laboratory
Berkeley, CA 94720

End-to-End Research Group Meeting

MIT LCS

Cambridge, MA

July 8, 1997

There's widespread agreement that the Internet needs some sort of service differentiation (or QoS) to survive. I.e., there must be a way to ensure, via administrative declaration and/or monetary exchange, that particular traffic gets a well-defined level of service.

This facility is fundamental to

- inter-agency or -institution sharing of a common infrastructure
- migration of corporate data traffic from private, leased-line networks to the Internet.

In the long term, economic theory says the Internet cannot survive as a collection of private, competitive entities unless there is service differentiation.

Where Are We Today?

- Most ISPs and their customers want QoS yesterday.
- Major router vendors are starting to deploy the low-level machinery needed for QoS (e.g., Cisco's CBQ and WFQ).
- There's been a decade of research work on Internet QoS.
- Almost every researcher is convinced that QoS is needed "to enable audio & video applications on the Internet".

The Core Problems

QoS isn't about audio and video applications — it's about controlling and administering a scarce resource: network bandwidth.

That means QoS decisions are organizational and economic, not technical.

Traffic flow follows network topology but bandwidth allocation decisions must follow organizational hierarchy.

$\Rightarrow$ Must be separate, organizationally organized, agents to control bandwidth allocation.

The Core Problems (cont.)

Allocation decisions have short- and long-term components:

Long:

- Provisioning
- Budget & budget cycles
- Agency & organization priorities.

Short:

- Particular conversations, tests, experiments, demonstrations, . . .

$\Rightarrow$ Should separate decision of “how much to allocate” from “what conversation(s) are currently using the allocation”.

The Core Problems (cont.)

The QoS mechanism must work at the scale of the Internet (e.g., millions of networks) and at the full range of speeds of the Internet (e.g., Gb/s links).

Scaling implies that

- all state is kept at the edges
- the per-conversation or per-transaction work (e.g., policing) is done at the edges.

$\Rightarrow$ QoS indication must be carried in packet.

Administrative diversity and high speed forwarding both argue for very simple semantics on that indication. E.g., one bit of premium/normal.

The Core Problems (cont.)

Almost all Internet traffic crosses *many* administrative boundaries. End-to-end QoS implies that all those independent units agree to treat the traffic as special. Experience suggests that multilateral agreements rarely work. ISPs are also competitive enterprises. They act in their own best interests and, where necessary, against the interest of their competitors.

$\Rightarrow$ End-to-end QoS must be constructed out of purely bilateral agreements.

QoS must not require extending trust across administrative boundaries.

The Core Problems (cont.)

For economically viable service, customers have to know what they're buying.

- Delivered service *can't* depend on other people's traffic.
- The customer *must* be able to measure conformance.

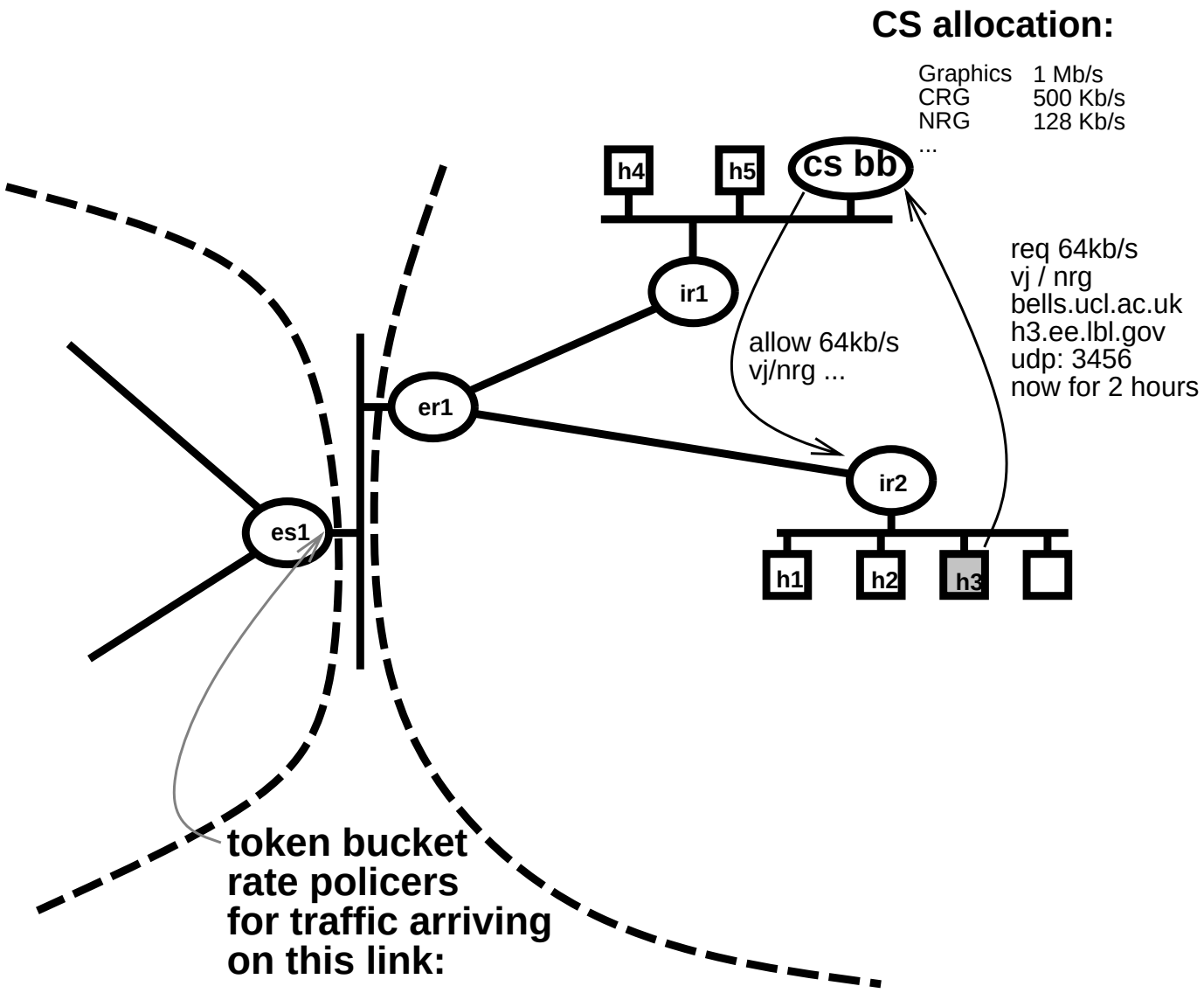
$\Rightarrow$ ISP must not "over sell" premium service.

Some economic issues

- Need to minimize per-transaction cost and setup (take advantage of predictible component of agregate traffic).

- Expect to see advance purchase (e.g., monthly, quarterly) of (predictible) premium bandwidth needs with (higher cost) “spot market” available to handle unanticipated needs.

- At ISP/ISP boundaries, expect agregate predictible component to be handled via advance (e.g., monthly) bilateral agreement (with no settlement).


CS allocation:
Graphics 1 Mb/s
CRG 500 Kb/s
NRG 128 Kb/s
...
h4
h5
cs bb
ir1
req 64kb/s
vj / nrg
bells.ucl.ac.uk
h3.ee.lbl.gov
udp: 3456
now for 2 hours
allow 64kb/s
vj/nrg ...
er1
ir2
es1
h1
h2
h3
token bucket
rate policers
for traffic arriving
on this link:
LBNL 1 Mb/s
NERSC 5 Mb/s
...