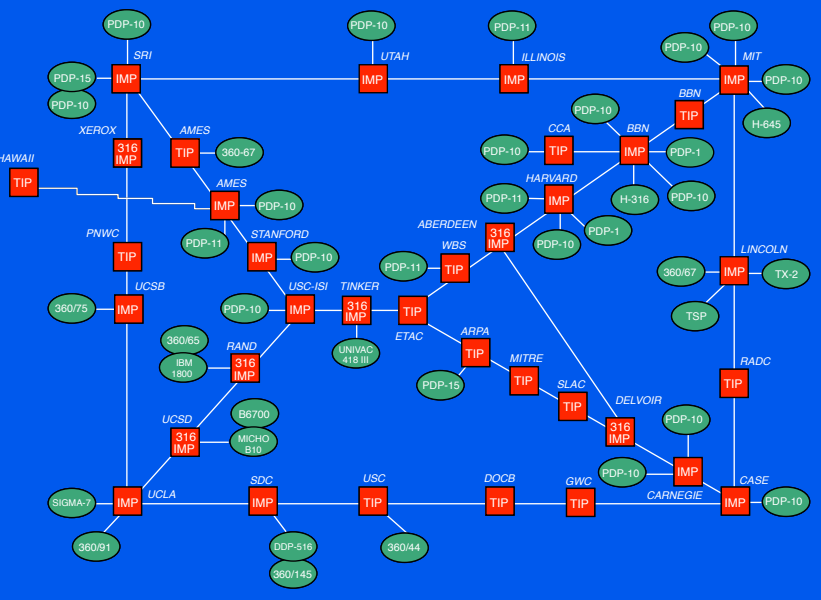
Reality and the Internet of the Future Programs

Scott Bradner
sob@harvard.edu

Historical Picture

- US government has been key to development of the Internet
- basic research
- advance state of the art
- proof of concept
- seed funding
- but total US funding “small”

ARPANET

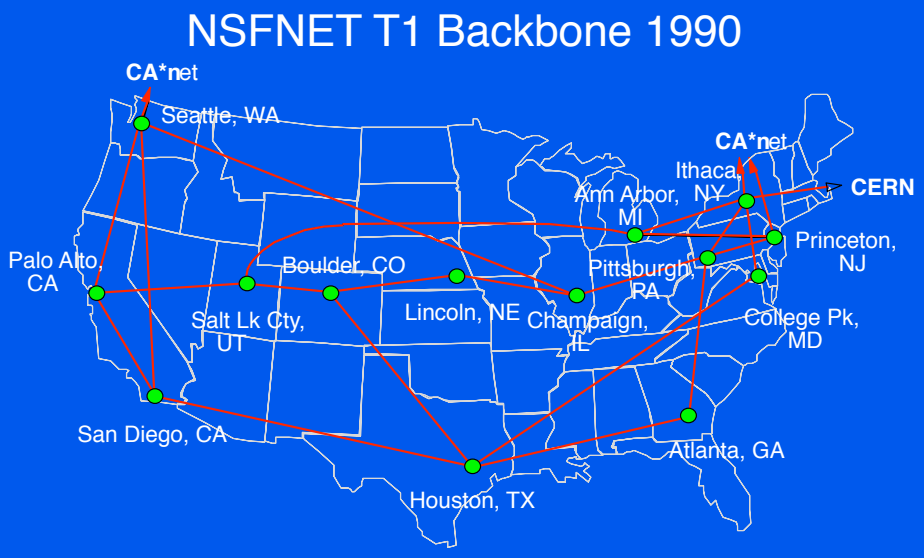


ARPA NETWORK, LOGICAL MAP, JANUARY 1973

ARPANET, contd.

- followed basic datagram decision
  - QoS impact
- routers / routing
  - IMPs, link-state routing
- transport protocols
  - NCP - TCP/IP
- applications
  - FTP, TELNET, SMTP ...
- i.e. everything

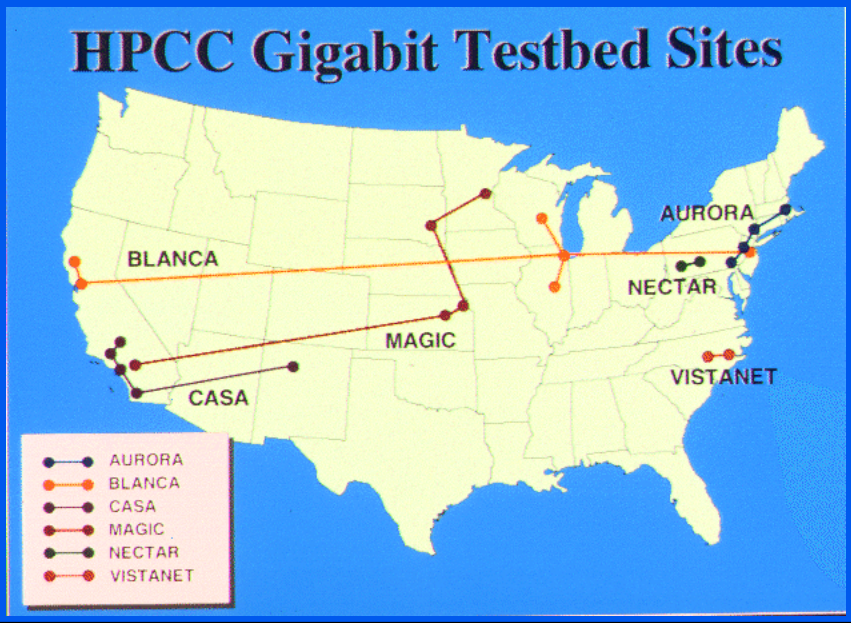
NSFnet



NSFnet, contd.

- TCP/IP only!
- proof of concept for high-speed networks
  - no, the commercial world was not ready
  - in spite of AT&T offer to Congress
- kick start for general use
- AUP forced commercial net development

Gigabit Test Networks



Gigabit Test Networks

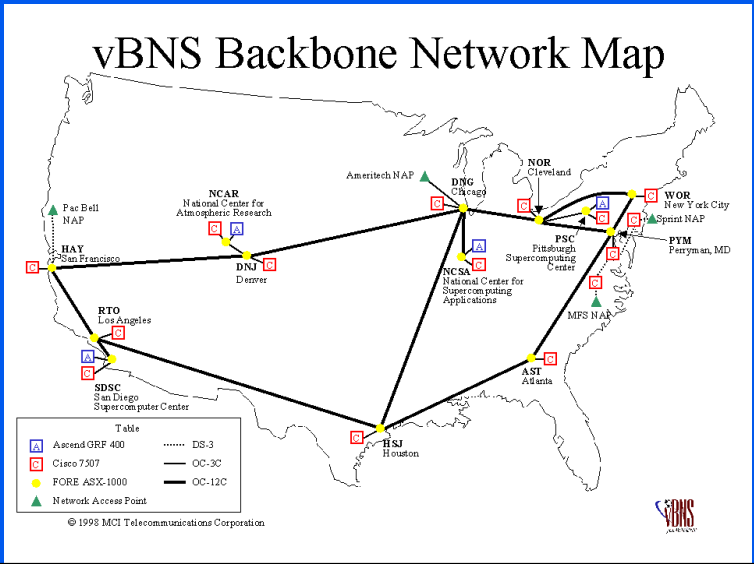
- HPCC - 12 US government agencies
- Government funds long term, “high risk” research (in theory)
- 6 test nets - 24 sites
- ATM @ 155mb & 622 Mb
- SONET @ 2.4 Gb
- look at problems involved with very high speed networking - TCP/IP & native ATM

  ATM is the answer

  (what was your question?)
- aggregated data streams were not permitted

vBNS

- very High-Speed Backbone Network Service



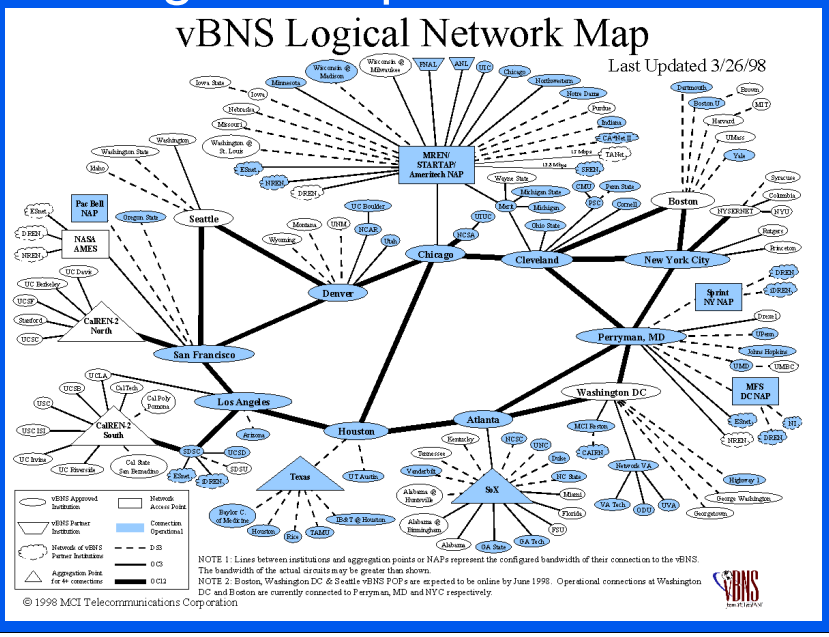
vBNS

- NSF funded, MCI contractor
- “to connect supercomputer centers”
- “platform for developing and testing Broadband Internet Services and equipment for the future”
- increase to gigabit speeds “in1990s”
- TCP/IP network
  reserve-in-advance ATM connections
- AUP = NSF approved sites only

vBNS, Contd.

- now expanding to ~100 sites
  - high-performance connections program
  - currently 4 SC + 42 HPC
- current vBNS grant ends in 1999
  - no currently announced extension or follow on
- maps, stats etc on www.vbns.net

vBNS Logical Map



High Performance Connections

- NSF HPC program
- 2 year grant totaling $350K
  non-renewable (currently)
- NSF also reimburses “user fee”
  up to $40K/year for T3, $130K/year for OC3
- does not require vBNS connection
  vBNS or other approved network
- 92 requests granted so far
  now what if more ask?

International Connections

- NSF helps support international research connections
- must connect to Star TAP
  avoid using vBNS as transit net
- some countries connecting on their own

Next Generation Internet (NGI)

- Clinton administration program
  - started with a campaign speech
- confusion in congress
  - how relate to current networking projects?

NGI, contd.

- research in applications, services and infrastructure
- $100M/yr - 5 year program
  - started 1 October 1997
  - $85M actual for ‘98 - some impounded
- builds on current “very strong agency programs”
- keep US ‘in the lead’
- DARPA, NASA, NIH, NIST, & NSF
  - DoE in future (if congress OKs it)
- 3 sets of goals

NGI Goal 1

- conduct R&D in advanced end-to-end networking technologies
- promote experimentation with next generation networking technologies
  - QoS, security, robustness, network management (including bandwidth sharing), system operations, new routing, security, multicast & mobility protocols, computer operating systems, distributed application environments
- define qualitative metrics for above
- move technologies to commercial net

NGI Goal 2

- establish & operate two testbeds
- a/ >100 sites at 100x current speed ~155Mb
  - built on NSF vBNS NASA NREN, DoD DREN, DoE ESnet (used to include "Internet 2"
  - must be "highly reliable"
- b/ ~10 sites at 1000x current speed (~1Gb)
  - Gb end-to-end
  - built on Advanced Tech Demo net (ATDnet) & DARPA ACTS ATM Internetwork
  - can "break periodically"

NGI Goal 3

- R&D in revolutionary applications
- demonstrate applications that can not be done over "today's Internet"
- e.g.
  - national security response & crisis response,
  - distance education, teleoperation (extreme reliability & guaranteed delay bounds)
- identify a small number of demo apps for each agency + apps from industry and academia

NGI "Fast Facts"

- Internet traffic has been growing 400 percent per year.
- By the year 2000, more than half of the U.S. population is expected to have access to the Internet.
- The Federal government, universities, and businesses are developing medical, environmental, manufacturing, educational, and defense applications that require new high-capacity networks to make them fully functional and widely available.
- The Next Generation Internet (NGI) initiative is a multi-agency Federal research and development (R&D) program to develop, test, and demonstrate advanced networking technologies and applications.
- The NGI initiative, together with investment by academia and industry, is laying the foundation for networks that are more powerful and versatile than the current Internet.

NGI “Fast Facts”, contd.

- The NGI initiative will produce a testbed of Government and university research networks that are 100 to 1,000 times faster than today's Internet.
- NGI-developed applications and technologies will be available to the business sector for incorporation into services for schools, work places, and homes.
- The NGI has virtually unlimited potential to help Americans live better and work smarter.
- Through the NGI initiative, the Government will help create an environment in which advanced networking R&D breakthroughs are fostered.
- The NGI initiative is coordinated by the NGI Implementation Team under the Large Scale Networking Working Group of the Subcommittee on Computing, Information, and Communications (CIC) R&D of the White House National Science and Technology Council's Committee on Technology.

Fed Net Relevance

- the old fed work created the Internet
  - ARPANET & NSFnet
- the not-so old fed work was semi-relevant
  - gigabit test bed, vBNS
- new vBNS work may be a help
- NGI a mixed bag
  - goal 1 (end-to-end technologies) are critical
  - goal 1a is vBNS (basically)
  - goal 1b is redundant with commercial world
  - goal 2 is some demo toys

Internet 2

- higher-ed initiative
- some confusion over goals
- some confusion with NGI

I2 History

- first there were lamentations
  - and then there were more lamentations
- Monterey Futures Group (Mfug)
  - needs (& solutions)
- enter Educom
  - collected Internet I geeks, university pols, ...
  - meetings at FARNET, in Ann Arbor, in Colorado Springs leading up to Oct meeting in Chicago
- 40ish “R1” universities said OK
  - $25K now for organization, “up to” $500K later
- then the prez talked about NGI & I2
  - since then more confusion

I2 Members

Arizona State
Boston
Brown
California Inst of Tech
Carnegie Mellon
Case Western Reserve
ClemsonOhio State
Colorado State
Columbia
Cornell University
Dartmouth College
Duke
Emory
Florida A&M
Florida Atlantic
Florida International
Florida State
George Mason
George Washington
Georgetown
Georgia Inst of Tech
Georgia State
Harvard Univ
Indiana
Iowa State
Johns Hopkins
Kansas State
Lehigh
Massachusetts Inst Tech
Michigan State
Mississippi State
New York
North Carolina State
North Dakota State
Northeastern
Northwestern
Ohio
Univ Houston
Oklahoma State
Old Dominion
Oregon State
Pennsylvania State
Princeton
Purdue
Rensselaer Polytechnic Inst
Rice
Rutgers
Stanford
Syracuse
Texas A&M
Texas Tech
Tulane
Univ Alabama
Alabama Birmingham
Univ Alaska
Univ Arizona
Univ Arkansas
Univ California Berkeley
Univ California Davis
Univ California Los Angeles
Univ Central Florida
Univ Chicago
Univ Cincinnati
Univ Colorado
Univ Delaware
Univ Florida
Univ Georgia
Univ Hawaii
Virginia Commonwealth
Univ Illinois Urbana-Champaign
Univ Iowa
Univ Kansas
Univ Kentucky
Univ Michigan
Univ Minnesota
Univ Maryland
Univ Massachusetts
Univ Missouri
Univ Nebraska
Univ New Hampshire
Univ New Mexico
Univ North Carolina
Univ Notre Dame
Univ Oklahoma
Univ Oregon
Univ Pennsylvania
Univ South Florida
Univ Southern California
Univ Tennessee
Univ Texas
Univ Utah
Univ Vermont
Univ Virginia
Univ Washington
Univ Wisconsin Madison
Univ Wisconsin Milwaukee
Univ Wyoming
Utah State
Vanderbilt

Virginia Tech
Washington State
Yale

So What Is It Not?

What Else Is It Not?

Basic Mission

- pre-competitive technology development environment
- high-speed
- QoS enabled
- support development of next generation applications
- info at www.internet2.edu

I2 Mission, contd.

- demonstrate new research collaboration applications
- demonstrate enhanced delivery of education and other services (virtual proximity)
- facilitate deployment of an affordable QoS supporting communications infrastructure
- promote experimentation with next generation communications technologies
- catalyze government & private sector partnerships

GiGaPoP!?

- part of the given
- definition followed term
- current definition
  - service connection point
  - multiple universities
  - multiple services
    - ISP(s)
    - inter-GP connectivity
    - telephone?

Inter-GigaPop Connections

- "vBNS is a candidate initial connectivity service"
- need QoS hooks
  - whatever that means
- like to have alternatives

Strategic Objectives

- enable advanced applications
  - add functionality to existing apps
  - create new apps
- strengthen the Universities in their research and education mission
- pioneer the introduction of:
  - Quality of Service
  - Advanced Multicast Support
  - IPv6
- establish the gigaPoPs as effective service points

So Why?

- "Quality of Service" control
  - believed to be a key enabler for advanced applications
  - particularly for "real-time" applications
- multicast support
  - one-to-many
  - few-to-few
- IPv6
  - an answer without a question?
  - or a key enabler for growth and for other advanced features?

More on gigaPoPs

- concentrate demand by local universities
  - bottom up not top down GP setup
  - e.g. Harvard/MIT/BU (heard this story before?)
- attract competitive providers
  - multiple ISPs - VC connection to each customers
- diversity of technical and organizational styles

Emerging GigaPoPs

- Alabama, Florida, Georgia, Tennessee
- New England
- Ohio
- DC, Maryland, Virginia
- Westnet states
- Michigan
- Texas
- Southern California
- Metro NYC area
- Chicago region
- Oregon
- Western Pennsylvania
- North Carolina
- Alaska, Washington
- Northern California
- Upstate New York

Diversity of GigaPoPs

- geographic scope
  - campus, metro area, state
- technology
  - TCP/IP, ATM, SONET
- what needs to be the same despite differences?
  - who can talk to who
  - inter-gigaPoP routing policy and design
  - measurement policy, design, and implementation
  - admissions control for QoS
  - inter-NOC trouble tickets
  - security coordination

I2 Issues

- why (in the context of the campus)
- with what money
- production vs. developmental net
- TCP/IP vs ATM
- QoS granularity
- role vs NGI

The Real World

- policy/authentication/settlements
  needed to apply QoS to real world
- confusion in I2 / NGI roles
- NSF / MCI relationship
  NSF pay vBNS user fees
- "scale is the only issue" (Mike O' Dell)
  how can you do scale with 1 dz nodes?

The Real World, contd.

- resource split between NGI goals
  - how important is Fed development of ultra-speed nets vs NGI goal 1 projects?
- will the NGI $ be there for real?
  - or will it be mostly agency $$?
- what is QoS?
  - instance-of-application vs McDonalds?
  - more than one ISP “product”
  - CBR?
  - classes?

Quality of Service, Background

- big call for QoS
- raison d'être for ATM
- push behind RSVP
- confusion over meaning, type and need

Quality of Service, What Is It?

- the ability to define or predict aspects of the performance of systems on a network
- long-time "glass house" requirement
  - SNA is seen as having lots of QoS controls
  - connection-oriented protocol
- one of the original goals for the Internet Protocols
  - "type of service" bits - request different processing for speed, latency & reliability
  - datagram protocol (for robustness)
    - processing "hints" to routers

Why Do You Want QoS?

- cuz the pundits said so?
- better web service?
- migrate SNA applications?
  - memories of control
- want to deploy fake wires? (VPNs)
- want to watch hi definition CNN?
- want to get part of the telephony cash flow?
  - $200B in US in 1997

Where is QoS Needed?

- where there are constrained resources
  - lines
  - interconnect devices
  - servers
- if you have enough resources, QoS controls are generally not required
- “enough resources” hard to define if delay is an issue

QoS Types

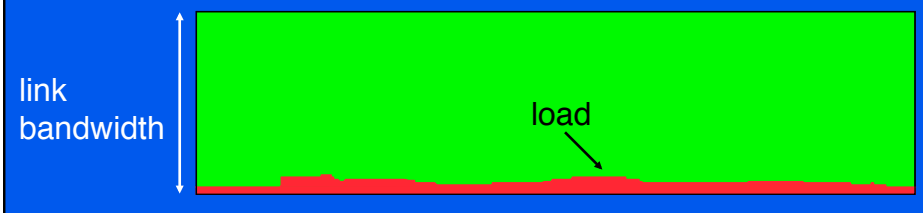
- predictive
  - architect network based on observed loads
  - can also police input loads
- flow based
  - reserve bandwidth for an execution of an application
  - keep track of reservation in each network device in path
- non flow based
  - mark packets to indicate class
  - process differently in network based on marking

Predictive QoS

- QoS in most current datagram networks
- "just" make network "big" enough
- reasonable on a LAN or campus network
- no guarantees
- hard to do for WAN - $$$
  - in spite of 'bandwidth will be free' people
- tends to provide cycles of quality
  - over build for need
  - need catches up and passes capacity
  - over build for new need

Throw Bandwidth at Problem

- with "enough" bandwidth QoS can be easy
  - enough means much more than peaks
  - e.g., gigabit Ethernet for 1 video stream
- still might have to sequence data onto link
  - if bursty traffic



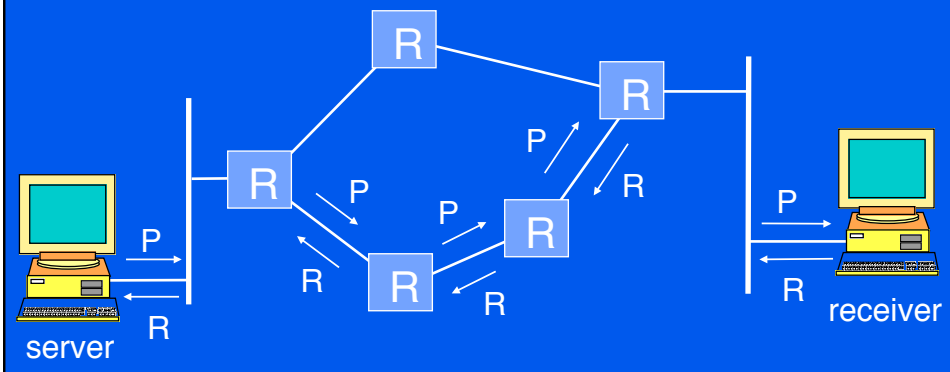
Flow Based QoS

- per flow reservations
- per flow guarantees
- per flow state kept in network elements
- ATM & RSVP QoS are the per flow type
- scaling issues
- authentication issues
- accounting issues

RSVP

- Resource ReServation Protocol (RSVP)
- implementation of INTSRV reservation process (I.e. signaling system)
- can be used to set aside resources for a specific application along a communications path - uses intserve link technology
- can transfer the requests to a new path if rerouted
- simplex (one direction per reservation)
- receiver-oriented
- may make use of QoS-active links - e.g. ATM

RSVP Process



Mixed QoS

- the only sure end-to-end technology is IP
  (for now - extended definition of “now”)
- use IP signaling (like RSVP) to control link-level QoS (like ATM) when present

Flow Based QoS Issues

- scaling issues
- authorization issues
- accounting issues
- advanced reservations *very* hard
- good for long flows (video, audio, large file transfers, VPNs)
  - flow setup cost must be low when averaged over flow length

Flow Lengths in the Internet

from cic nets' Chicago hub

```
IP Flow Switching Cache, 16384 active flows, 0 inactive
  132159644 added, 124468367 replaced, 4892577 timed out, 2782316 invalidated
  statistics cleared 270640 seconds ago

Protocol         Total  Flows   Packets Bytes  Packets Active(Sec) Idle(Sec)
--------         Flows   /Sec     /Flow  /Pkt     /Sec      /Flow     /Flow

TCP-Telnet     5222464   19.2        40    89    785.3       32.9      17.3
TCP-FTP        2087345    7.7         6    87     47.9        7.3      22.7
TCP-FTPD       1275958    4.7        95   390    449.5       21.9      23.6
TCP-WWW       83916123  310.0         9   304   2944.5        5.4      20.9
TCP-SMTP      14106833   52.1         8   173    448.9        6.4      21.6
TCP-X            94849    0.3        81   176     28.6       24.1      17.8
TCP-other     16095661   59.4        38   274   2290.8       20.9      21.5
UDP-TFTP           339    0.0         1   207      0.0        2.3      21.0
UDP-other     5059444   18.6        11   217    208.4        9.4      26.0
ICMP          4201689   15.5         2    83     46.0        5.2      26.8
IGMP            39809    0.1        30   398      4.4       48.2      29.4
IPINIP           9431    0.0      1808   254     63.0      147.1      18.6
GRE             32811    0.1       594   204     72.0       62.1      18.8
IP-other          909    0.0         3   223      0.0        1.2      31.8
Total:      132143665  488.2        15   260   7389.7        0.0       0.0
```

Non Flow Based Qos

- packet headers are “marked” at edge of network
  - precedence bits most common place to mark
- one or more bits used
  - two (priority and best effort) or more levels
- many different mechanisms proposed
  - drop priority
  - queue selector - WFQ on queues
- contract with ISP, contract between ISPs
  - a problem if too much traffic for destination
- creates N predictive Vnets on same Pnet

IETF Diffserv WG

- new diffserv WG formed
- strawman ID published
- components
  - mark bits in TOS byte at network “edge”
  - routers in net use markings to determine packet treatment
  - conditioning marked packets at network boundaries
- deals with flow aggregates
- TOS byte may change in flight

DS Byte

- rename TOS byte to be Differentiated-Services (DS)
- use to designate behaviors
  - not services to "customer"
  - build services from behaviors
- current thinking

PHB per-hop behavior
CU currently unused (must be 0)

PHB

- PHB = 00000 default (best effort)
- PHB = 11100 expedited forwarding

- other patterns to be explored
  - queue selector
  - queue lengths
  - WFQ weights
  - drop algorithm
  - drop preference
  - . . .

PHB, features

- packets in same TCP flow with same PHB must not be reordered
- some PHBs will have in/out bit
  within contract / out of contract indicator

CU

- reserved for future
- could be used for congestion experienced

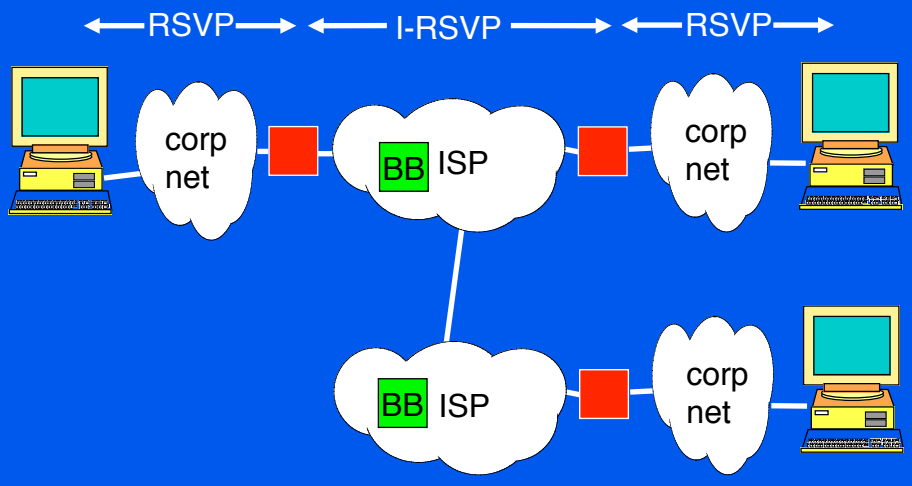
Traffic Conditioners at Edges

- packet classifiers
  - use fields in packet headers to steer processing
- markers
  - set DS byte
- policers
  - monitor traffic & react if profile exceed
    - drop, remark packets
- admission control
  - to accept / reject service request
- shapers
  - modify packet flow to control TCP flows

Signaling

- not part of diffserv work (yet)
- could use RSVP at edge
- could use I-RSVP in core
  - could provide some guarantees
- Bandwidth Broker (BB) proposal
  - can ask for more bandwidth when needed

Diffserv Conceptual Diagram



What's it Good For?

- if ISP is honest then diffserv can be useful
- provide predictable platform for applications
  - e.g. IP telephony
- ISP can sell more than one product
  - might be able to make some money in the business
- "FEDex-like" delivery of electronic goods
- . . .