Advanced Lecture on Internet Infrastructure

11. QoS (BW, Delay) Guarantee

Masataka Ohta

mohta@necom830.hpcl.titech.ac.jp

ftp://chacha.hpcl.titech.ac.jp/infra11e.ppt

What is QoS (Quality of Service) Guarantee?

- QoS guarantee
  - mostly, BW, delay and jitter (fluctuation of delay)
- necessary within network, impossible by end
  - end cannot know content of undelivered packet
- BW and delay closely related
  - according to queueing theory, enough BW means small delay
  - narrow BW is OK, if infinite delay is allowed

Networks

- Physical Distribution Networks
  - postal service, parcel services, convenience stores
- Information Communication Networks
  - Publishing Network (Book, News Paper, CD, Movie)
  - Financial Network
  - Phone Network
  - Broadcast Network
  - the Internet

Publishing Network

- Mass Distribution of Same Information
- Delay of the Distribution may be Tolerated
- Protected by Copyright Act
- The First Victim of the Internet
  - Collapsing

Financial Network

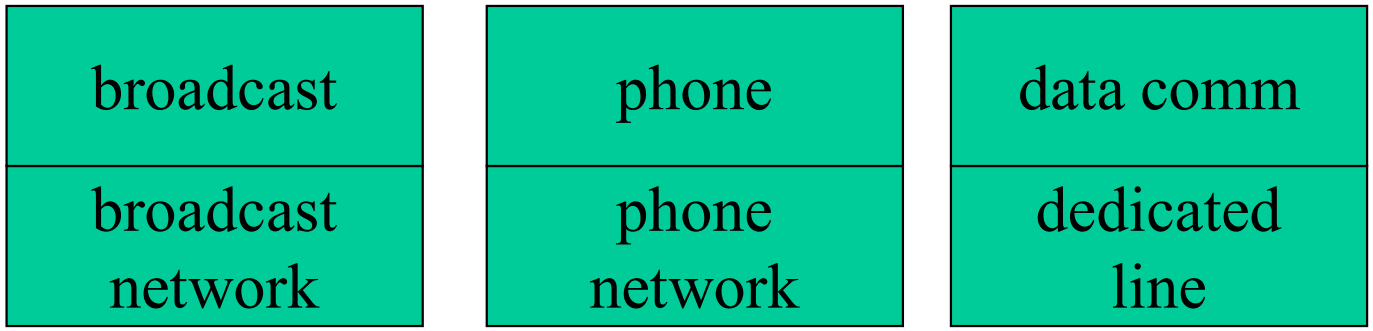
- Manage Transfer of Money
- Partly, Phisical Distribution Network, but, today, mostly ICN
- Security!!!
  - Not that there is no accident
  - Who will pay the loss on accidents

Phone Network

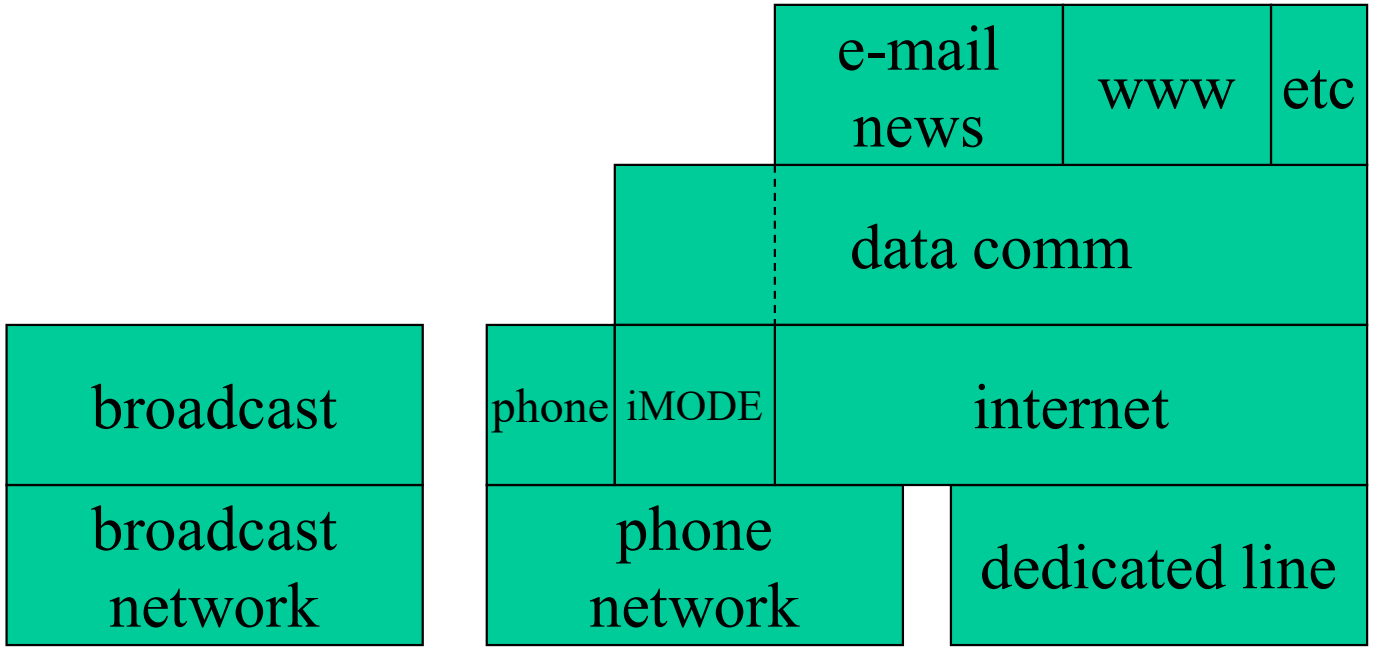
- Network for Realtime Voice Transfer
  - Allocate bandwidth for voice transfer
  - Minimize (guarantee) delay for voice transfer
- Dedicated line service may be Offerred
  - but, primary service is voice transfer
- Slow and Expensive
- Was Protected as National Company
  - Leberated by Telecommunication Business Act

Broadcast Network

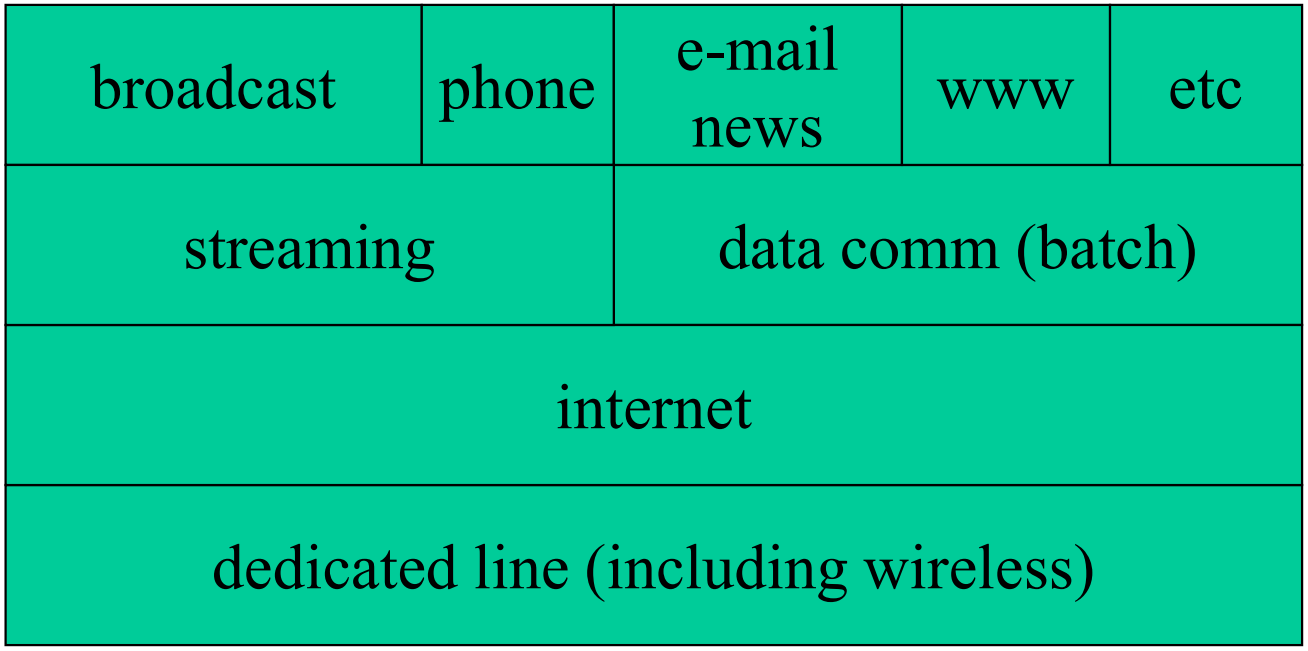
- Network to Transfer Voice/Image to Many in Realtime
  - Allocate bandwidth for the transfer
  - Minimize delay
- Wide Area One to Many Communication over Radio Waves
  - Broadcast/Multicast
- Protected by Broadcast Act



networks before the Internet



networks with the Internet



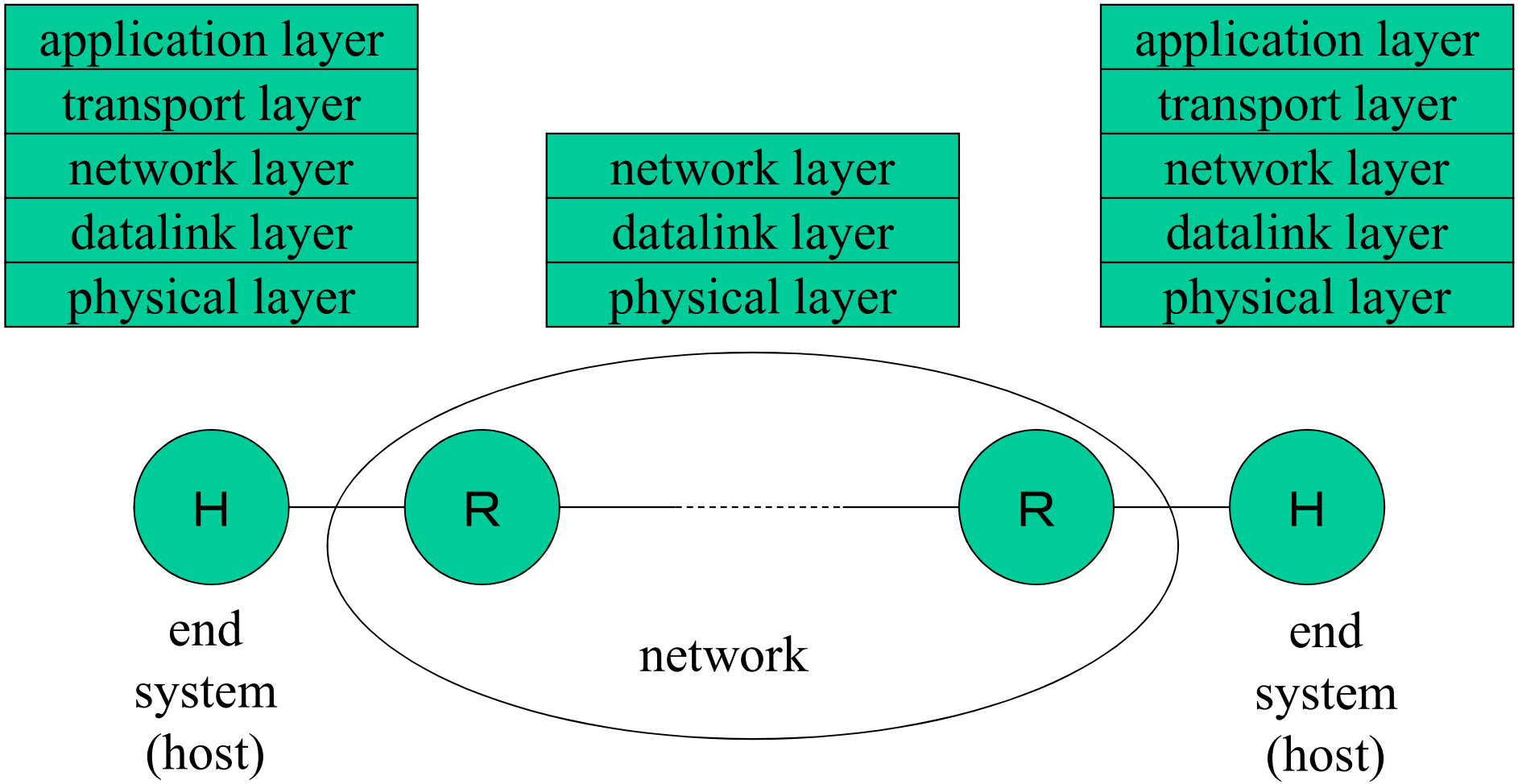
network in the future

Examples of QoS

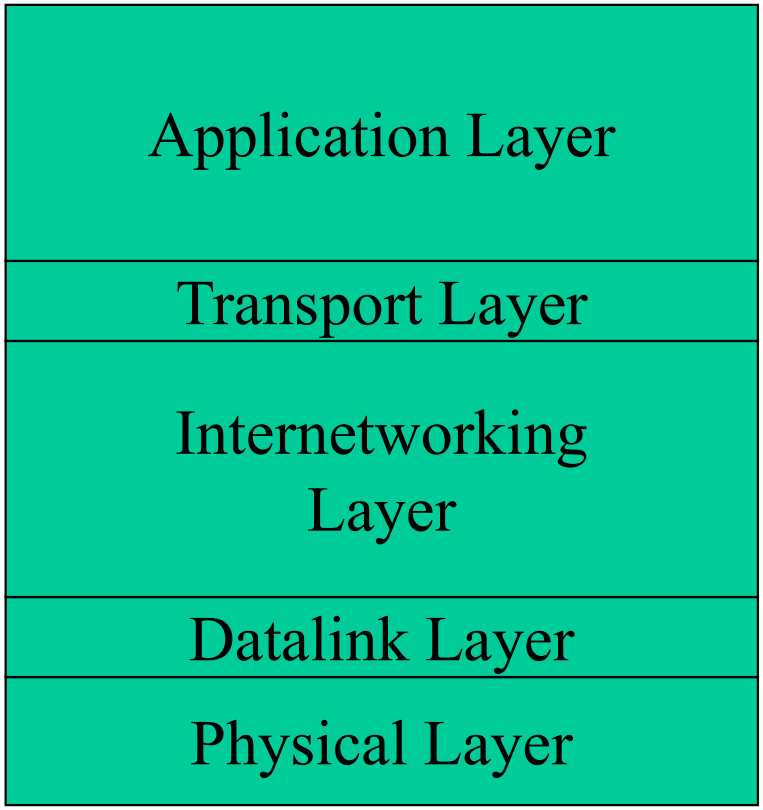
- phone
  - BW: 64kbps, delay < 0.1s
- CD play
  - BW: 1.5Mbps, delay < 1s
- TV broadcast
  - BW: 6Mbps, delay < 1s

Unit of QoS Guarantee

- different QoS requirement for communications between same pair of hosts
  - file transfer
    - best effort
  - voice conversation
    - minimize delay
  - image transmission
    - wide BW
- communications distinguished by port #s

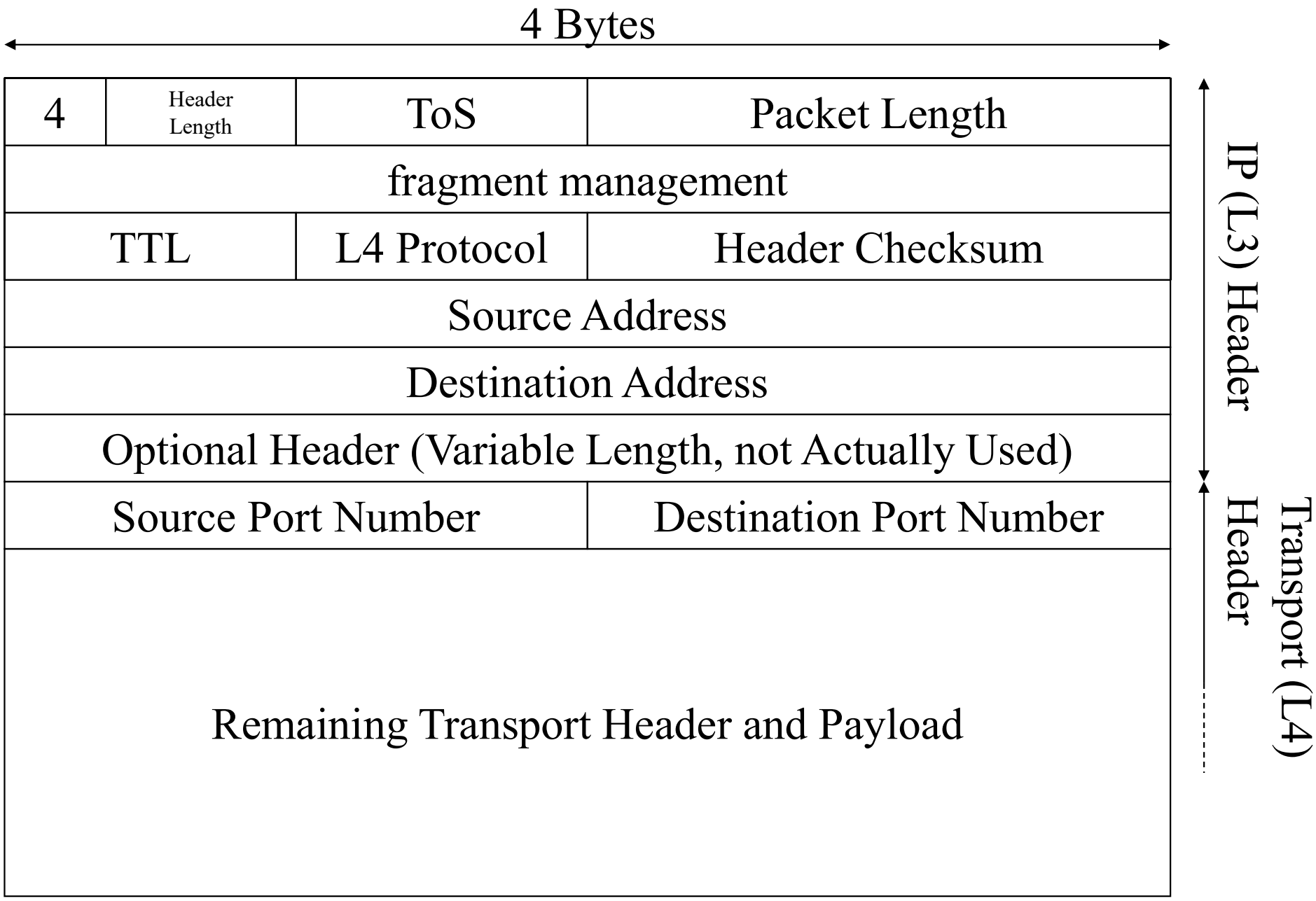


best effort internet

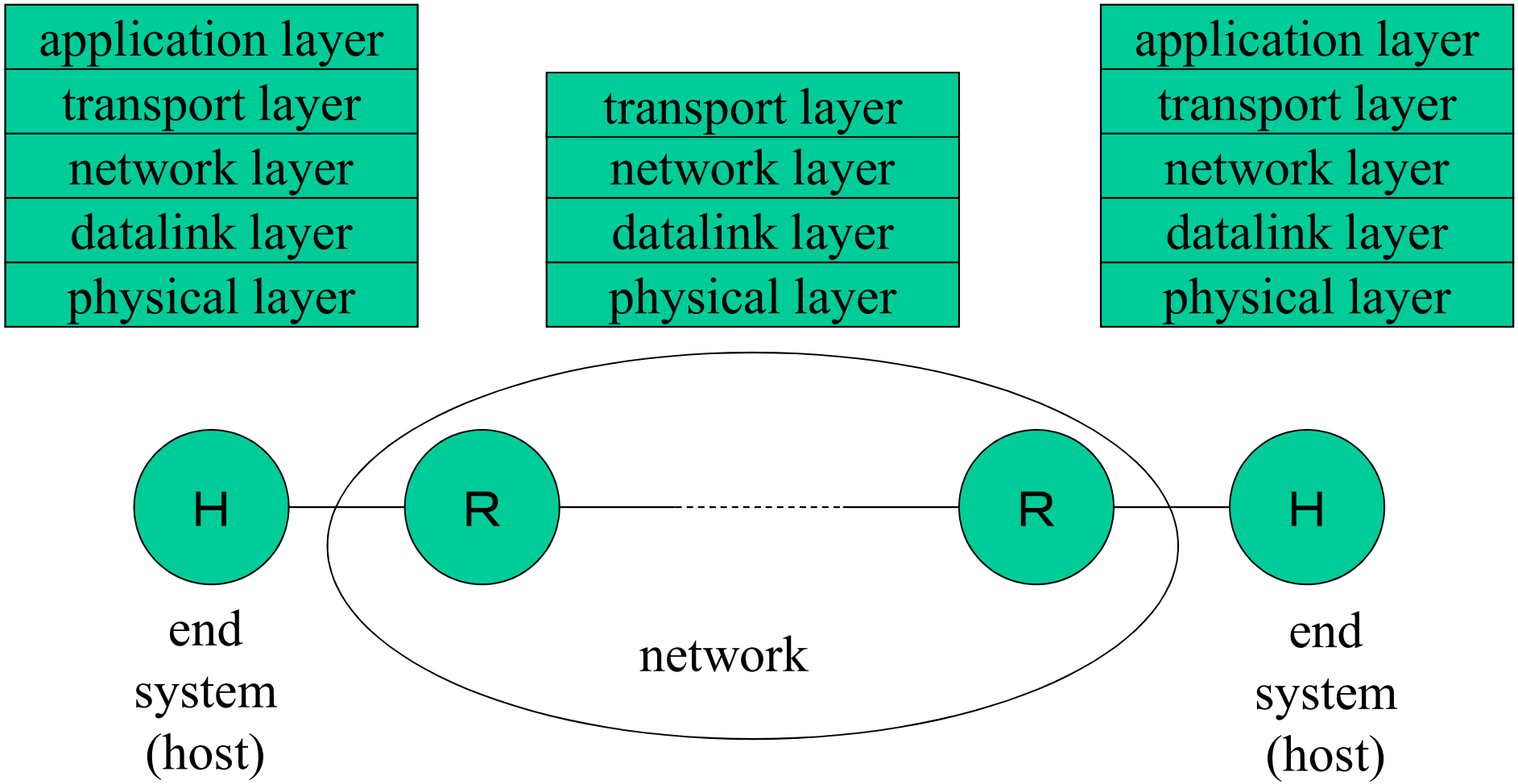


Here is the Essence of the Internet

Layering Structure of the Internet



Format of IPv4 Packets (rfc791)



internet with QoS guarantee

To Guarantee QoS

- jitter absorption
  - absorbed by end using timestamp
- BW and delay guarantee on each router
  - necessary as function of network
  - statistical multiplexing (not 100% guarantee) important
- QoS routing
  - NP complete if two or more additive constraints
  - stabilization of route difficult

Cause of Delay/Jitter

- propagation delay
  - not very large (electric/photonic signals propagate with speed of light)
  - route change causes jitter
- delay by queueing is dominant
  - delay may be suppressed by fine control of queueing
  - having lengthy queue is a bad idea only to make delay longer

Difficulty of QoS Guarantee

- queueing
  - fair queueing is the well known queueing method for QoS guarantee
    - does not scale
    - delay is impractically large (proportional to (packet length)/(communication speed))
- routing
  - route become unstable
  - hierarchical aggregation of route information?

QoS Guarantee by ATM/PNNI

- queueing
  - fair queueing is used
    - lack of scalability ignored
    - delay improved by small packet (cell) size (53B)
- routing
  - instability of route
    - no dynamic rerouting
  - hierarchical aggregation of route information
    - left for implementation, if inaccurate crank back

QoS Guarantee by IETF/RSVP

- queueing
  - fair queueing is used
    - lack of scalability: give up QoS and assure CoS
    - impractically large delay: ignored
- routing: no QoS routing
  - instability of route
    - centralized computation by PCE
  - hierarchical aggregation of route information?

Proper QoS Guarantee

- queueing
  - rely not on fair queueing but on statistical properties
    - scales well
    - guaranteed delay is small ((packet length)/(link speed)
- QoS routing
  - use PATH message to carry knowledge necessary for correct QoS routing
    - route will be stable
    - with hierarchically aggregated route information

BW Control

- have different queues for communications with different QoS requirements
- output packet frequently from queues for wide band communications

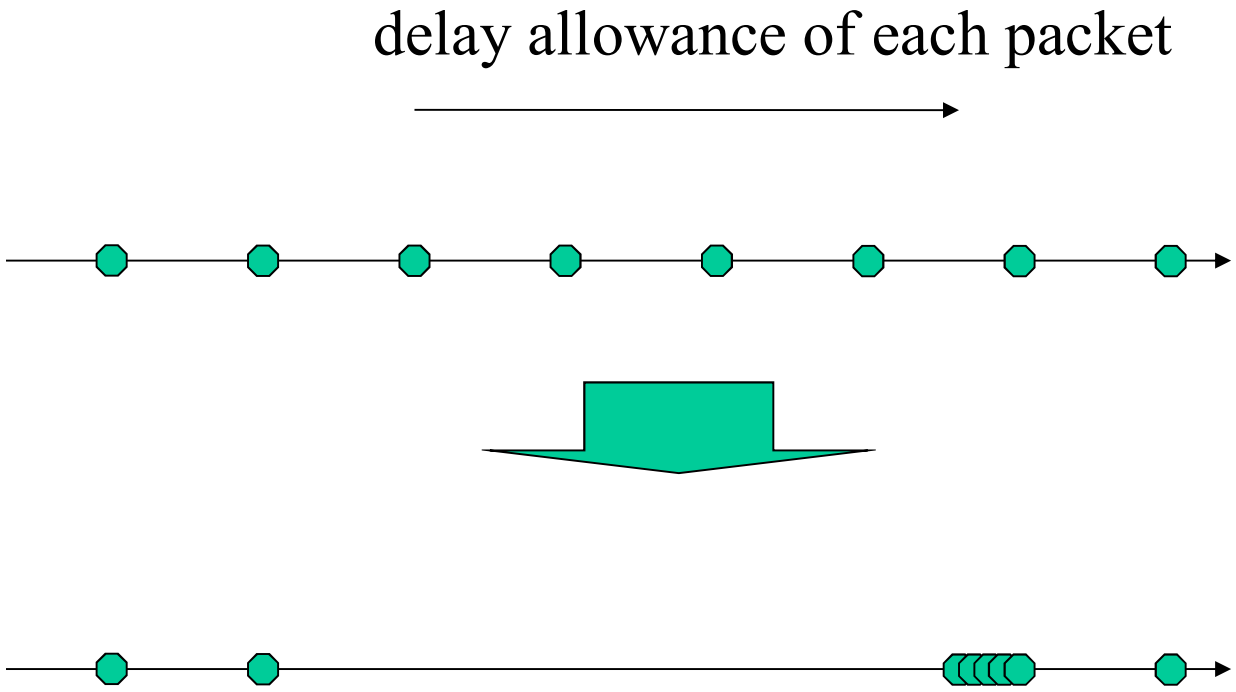
Policing

- frequency of output from a queue depends on BW reserved in advance
- if, in practice, more BW is used
  - queue will overflow and packets (of other communication shareing the queue) will be dropped
- policing is to inspect BW of communication
  - if violated, drop packets (or reduce priority)
  - inspection allowing some jitter necessary
    - based on traffic model (expected jitter behavior)

Traffic Model

- token bucket model
  - base on perfectly periodic packet sequence
    - jitter below certain threshold is allowed
  - algorithm for policing exists
- Poisson model
  - not more bursty than Poisson
  - infinit jitter may occur with low probability
  - policing as statistical test

Allowed Jitter with Token Bucket Model



Controlling Queue

- have multiple queues
- proper queue is selected for each communication (distinguished by address/port)
- choose the next queue for output by proper algorithm
  - fair queueing?
  - PPQ!

Fair Queueing

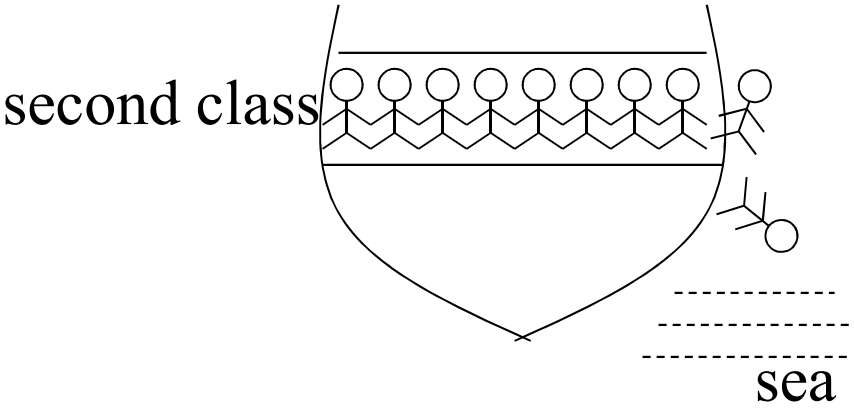
- separate queue for each communication
  - ordering queues by priority involves sort
    - slow, if # of communication is large
      - O(log(# of communication)) or worse
- guaranteed delay is
  - (packet length)/(BW of communication)*(# of hops)
- delay is 100% guaranteed, 100% of link BW may be reserved

Problems of Fair Queueing

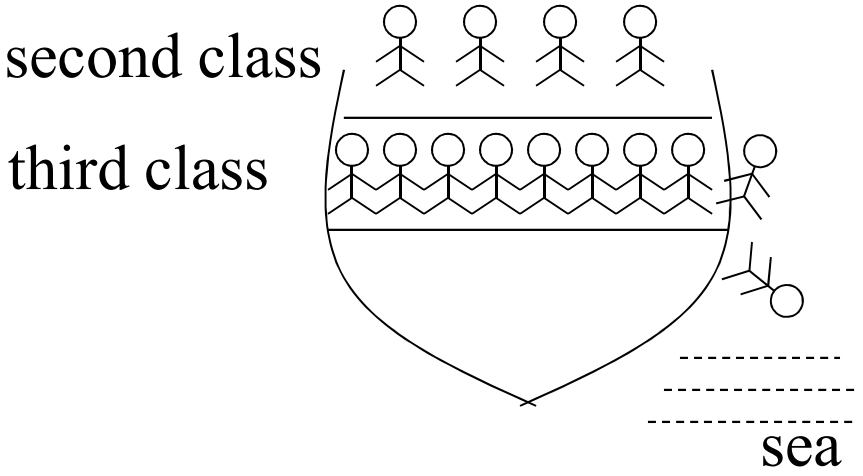
- guaranteed delay is too large
  - 1500B@64kbps means
    - 0.2s delay for each hop
- heavy processing load, not usable at backbone
  - at backbone, link BW is wide, # of QoS communication is large
  - large amount of buffer necessary
- arrival of packet is not 100% guaranteed
  - if arrive, delay is 100% below guarantee

To Use Fair Queueing

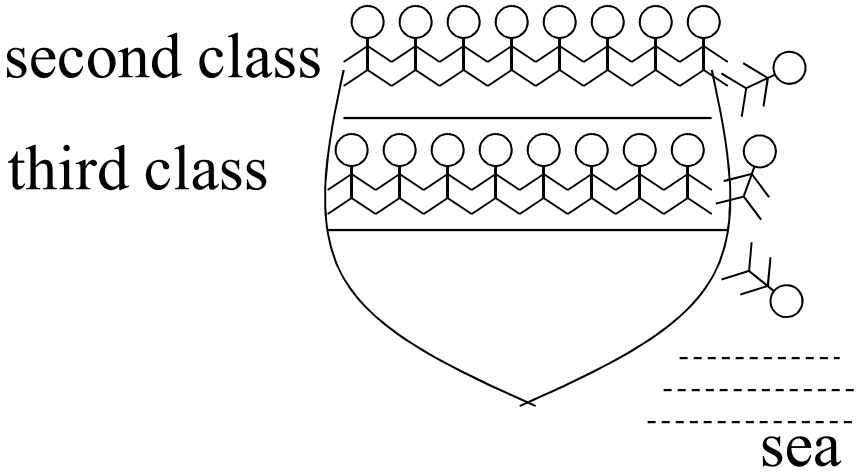
- shorten packet size
  - ATM cell is 53B long (6.6ms@64kbps)
  - maybe good for telephone, not a fundamental solution
- QoS guarantee for a set of communications
  - guarantee on traffic class (DiffServe)
    - CoS (Class of Service) guarantee



a) internet, so far



b) internet with DiffServe



c) internet with DiffServe (upon congestion)

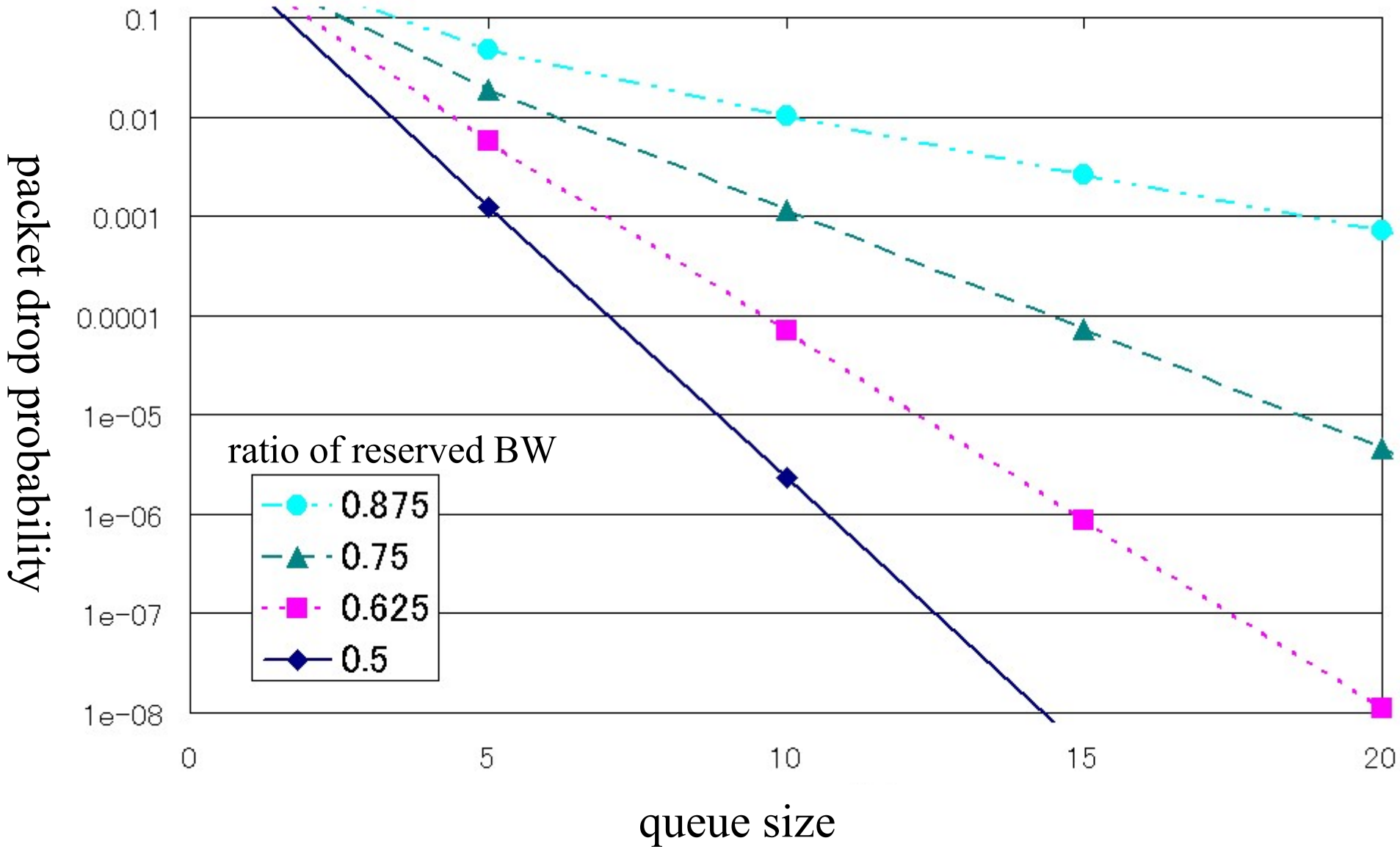
meaning of DiffServe classes

PPQ (Policed Priority Queueing)

- rely on statistical model (Poisson)
  - assume communications are statistically independent
- guarantee delay 99.9999% of time or so
  - (packet length)/(link speed)*20 or so
- only 2 queues (for BE and for QoS)
  - queue for QoS has absolute priority
- each communication is policed individually

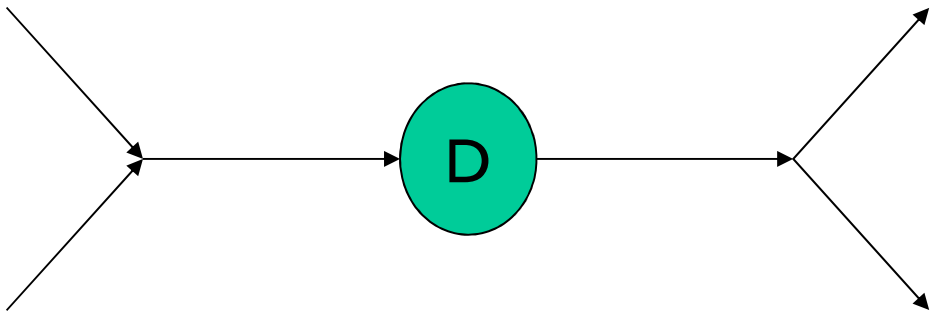
Merit and Demerit of PPQ

- small delay regardless of communication speed
  - 3ms on 100Mbps link with 1500B MTU
- simple and high speed packet processing with small (20 packet?) amount of buffer
- 80% of bandwidth may be reserved
  - rest may be used by BE
- sometimes, packet delays longer than guarantee
  - no different from packet drop by error
- what if, multiple communication synchronizes?
  - actively prevent synchronization


packet drop probability
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
1e-08
0
5
10
15
20
queue size
ratio of reserved BW
0.875
0.75
0.625
0.5

Definition of “no more bursty than Poisson”?

- desired to preserve property
  - independent flows are merged
    - solved
  - after passing a router (G/D/1)
  - independent flows are separated
    - seems to be not a problem



merge pass a router separation

Example Definition of “no more bursty than Poisson”?

- can be defined as “if event rate is λ, variance of # of evens within interval of length L is not more than $L\lambda$”
  - preserved after merge because of independence

Example Definition of “no more bursty than Poisson”?

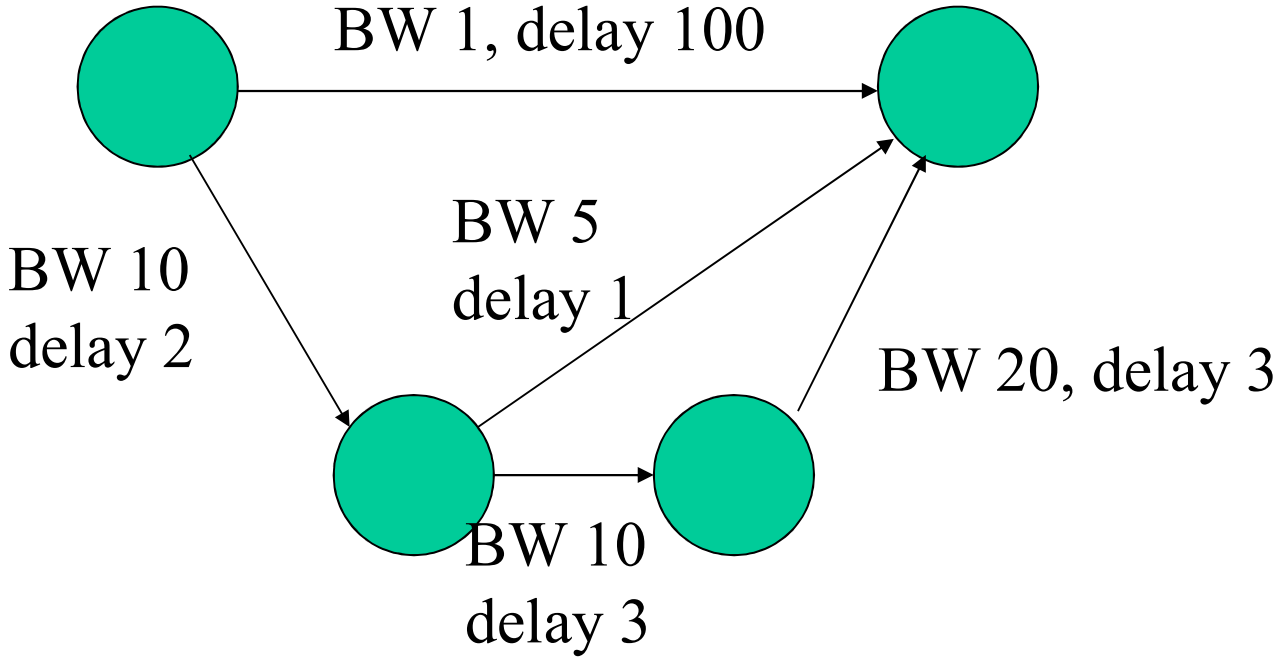
- Poisson process, of course
- perfectly periodic events

RSVP (Resource ReSerVation Protocol, rfc2205)

- multicast capable
  - source send PATH message
  - destinations send RESV messages to oppsite direction of PATH
  - is compatible with any multicast routing protocol
- no QoS routing
- no charge model

QoS Routing

- find shortest path satisfying desired QoS from given link properties

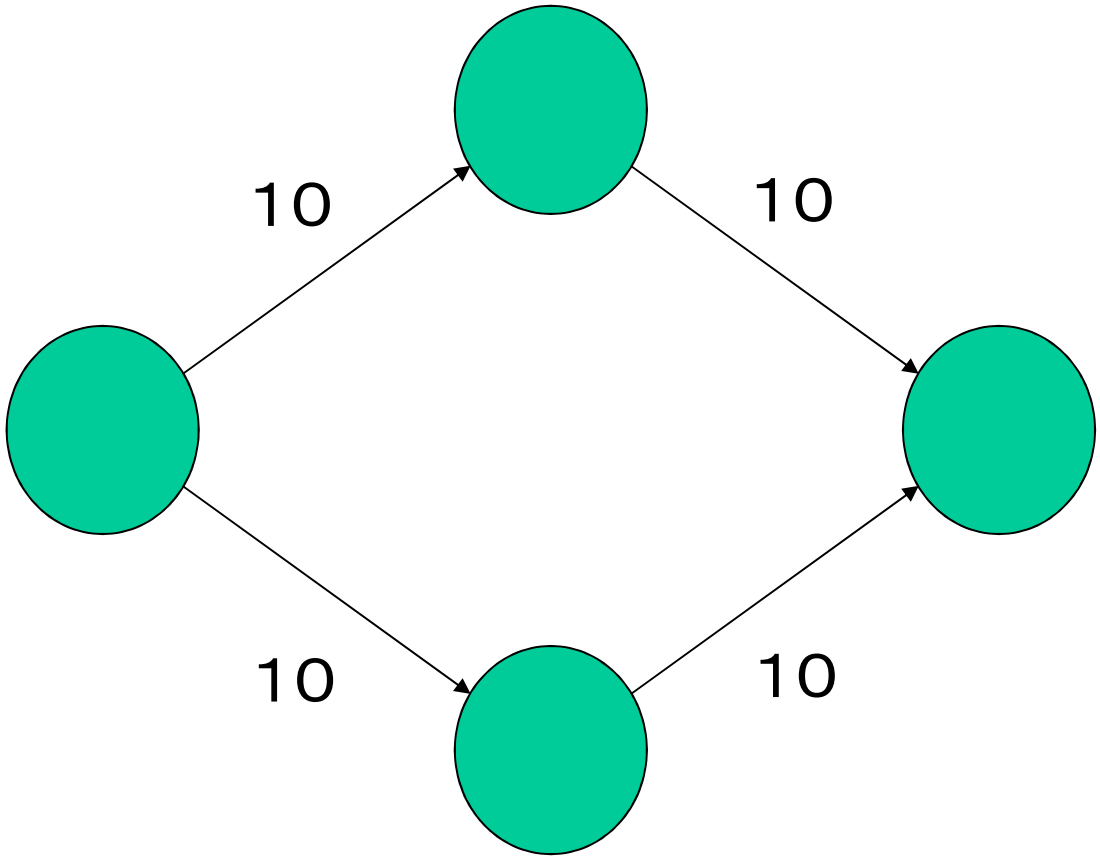


Difficulty of QoS Routing (1) NP Completeness

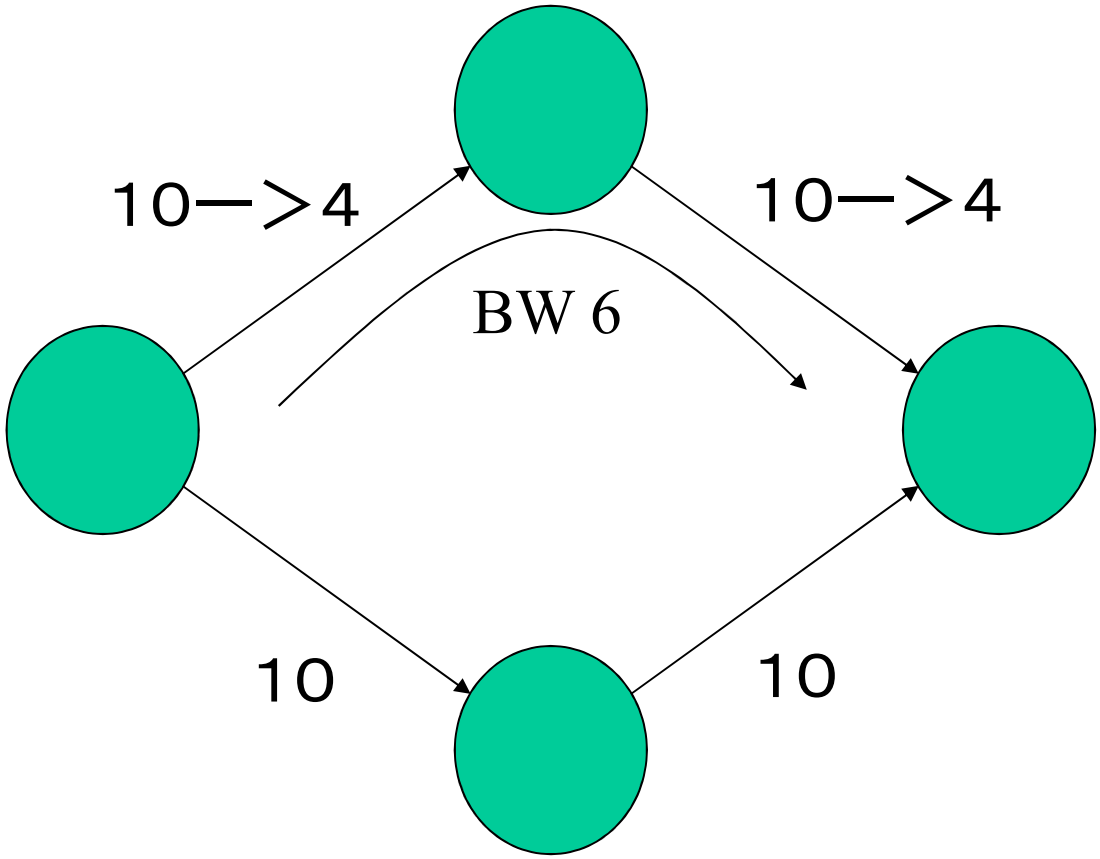
- satisfying BW requirement is trivial
  - ignore links with insufficient BW
- as delay and charge are additive
  - problem with one additive constraint is solved by Dijkstra
  - solving problem with two additive constraints is NP complete
    - practically insolvable

Difficulty of QoS Routing (1) Instability of Route

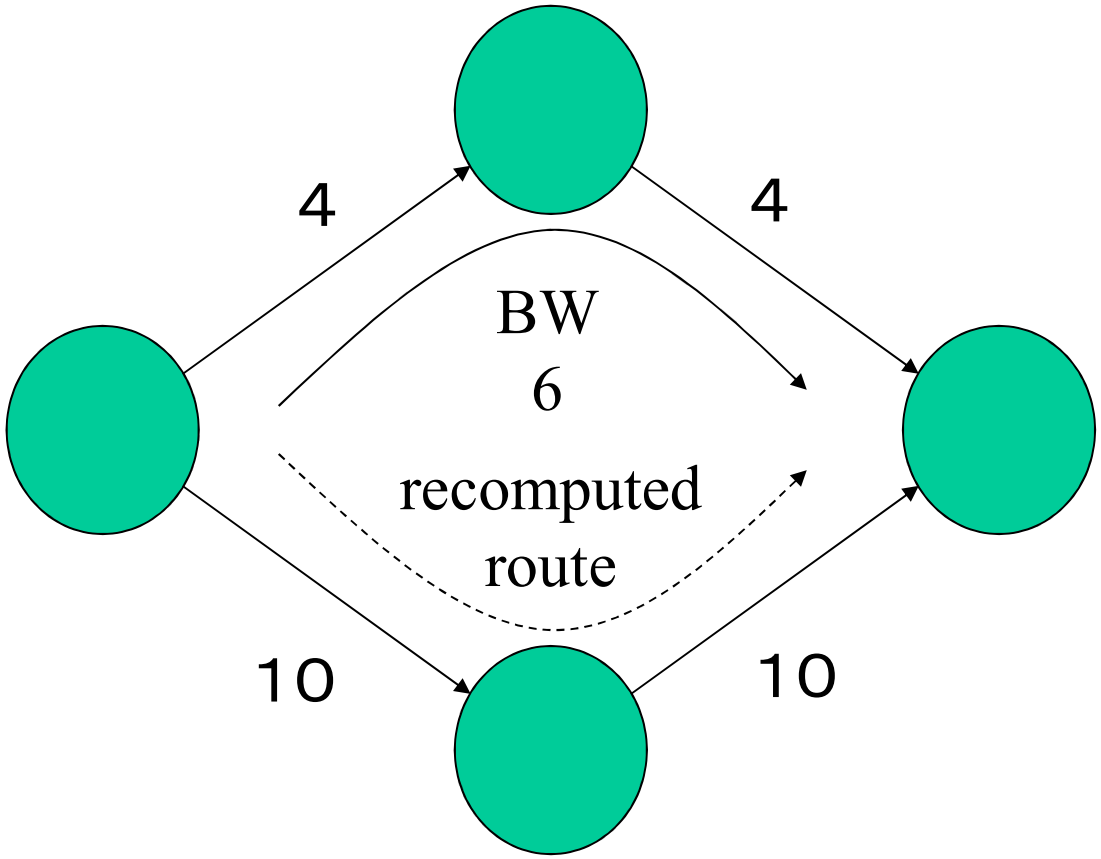
- if some communication consume BW
  - available link BW reduced
- if route is recomputed
  - the communication may not be able to use original route
- is route recomputation impossible?
  - what if, a new member of a multicast group dynamically appears?



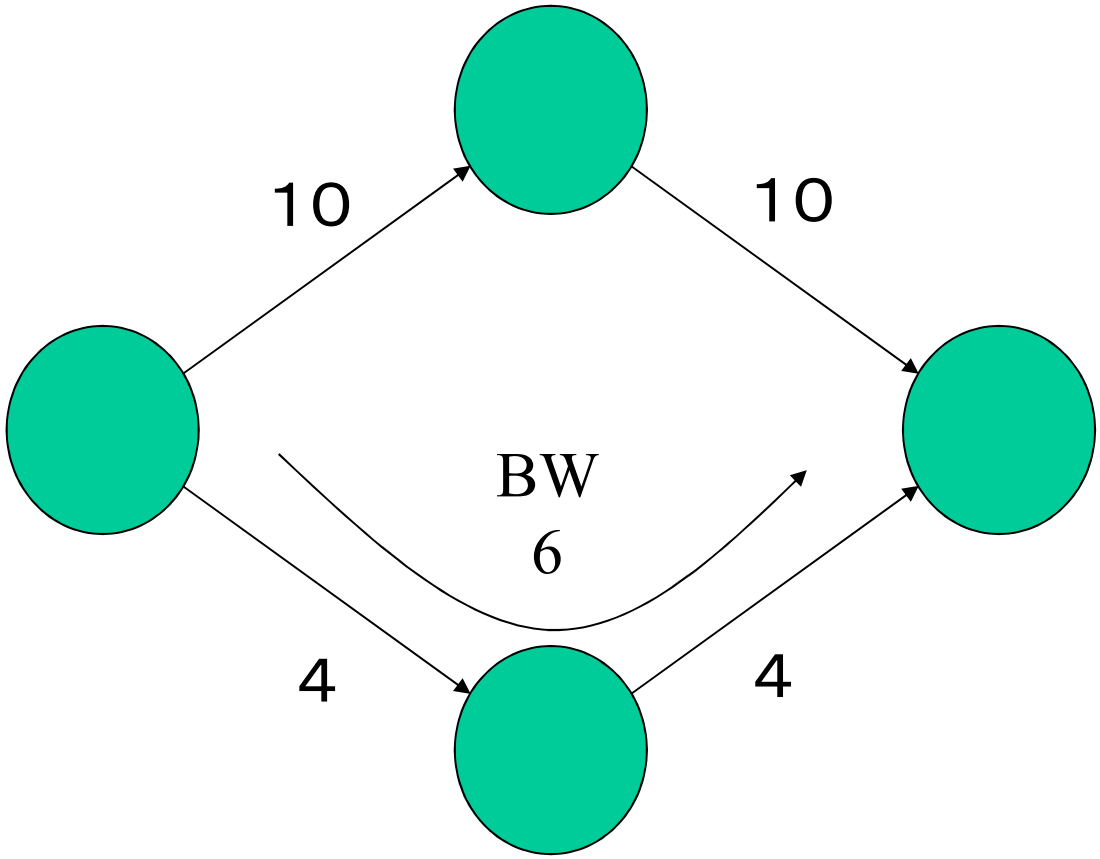
link BW consumed by communication (initial state)



link BW consumed by communication (communication with BW 6)



link BW consumed by communication (recomputation of route)



link BW consumed by communication (route flipping)

To Stabilize Route

- advertise BW of all communications on all links
  - does not scale
- never recompute route
- compute all route by source
  - does not scale for large scale multicast
- PATH QoS Collection
  - RSVP PATH message of communication carry BW actually available to the communication

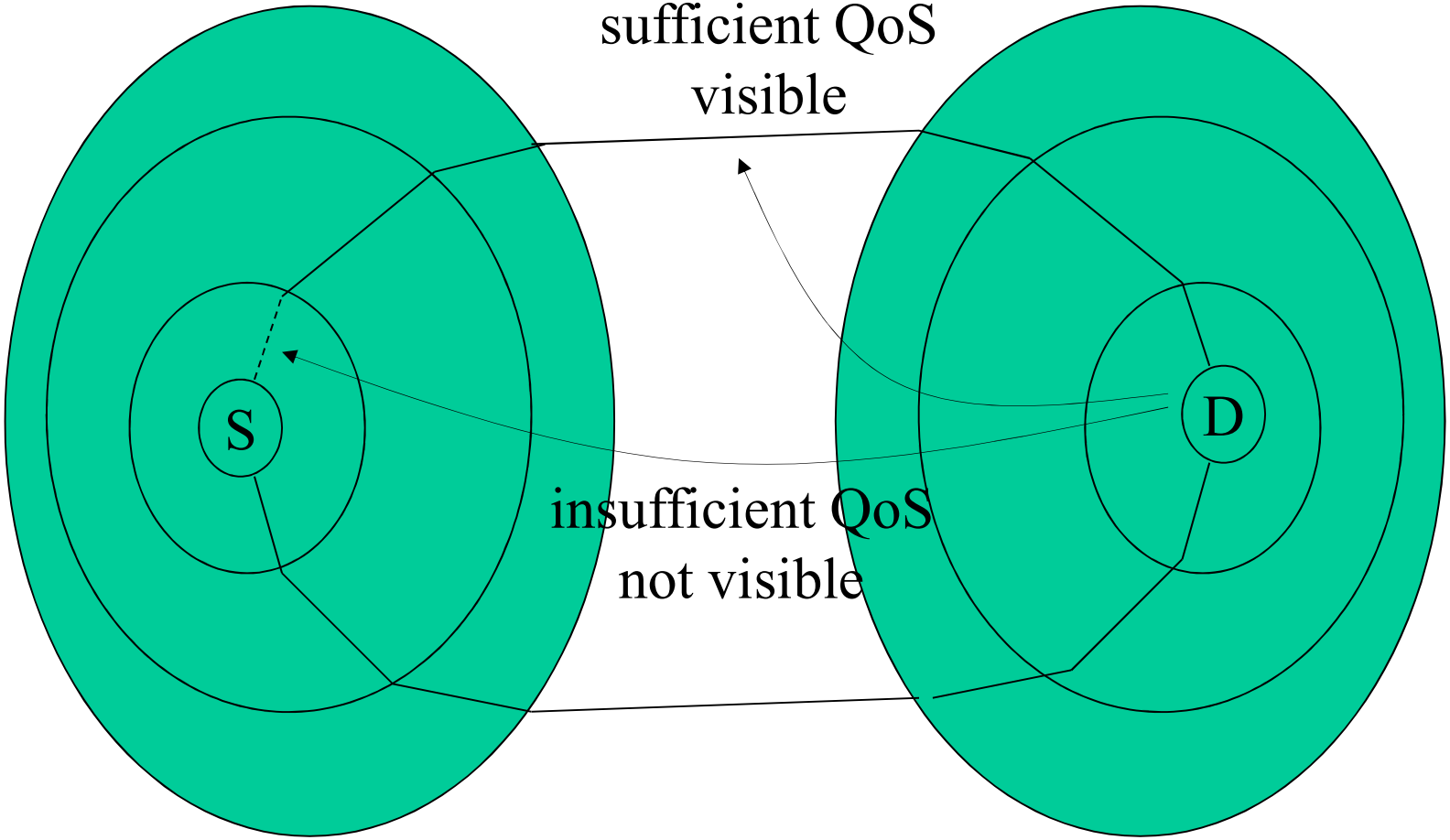
Another Way to Stabilize Route Path Computation Element

- compute route at some central point
  - as routers do not have much computation power (???), let's have a dedicated computer
    - obviously wrong
  - as PCE has all the BW consumption information, there is no instability
    - as ISPs have their own PCEs individually, inter-PCE cooperation is inevitable (no central point)
      - if PCEs can cooperate to prevent instability, all the routers should be able to cooperate to prevent instability

Difficulty of QoS Routing (3) Hierarchy and Crank Back

- hierarchical routing necessary for scaling
  - hierarchical hiding of detailed route information
- to compute route, route information of all the hierarchy necessary
  - detailed route around oneself should be ready
  - detailed route around peer is not
    - computed route at top level may enter area of peer at a point with insufficient internal QoS
- crank back (try to choose other route) is complicated and poorly scale

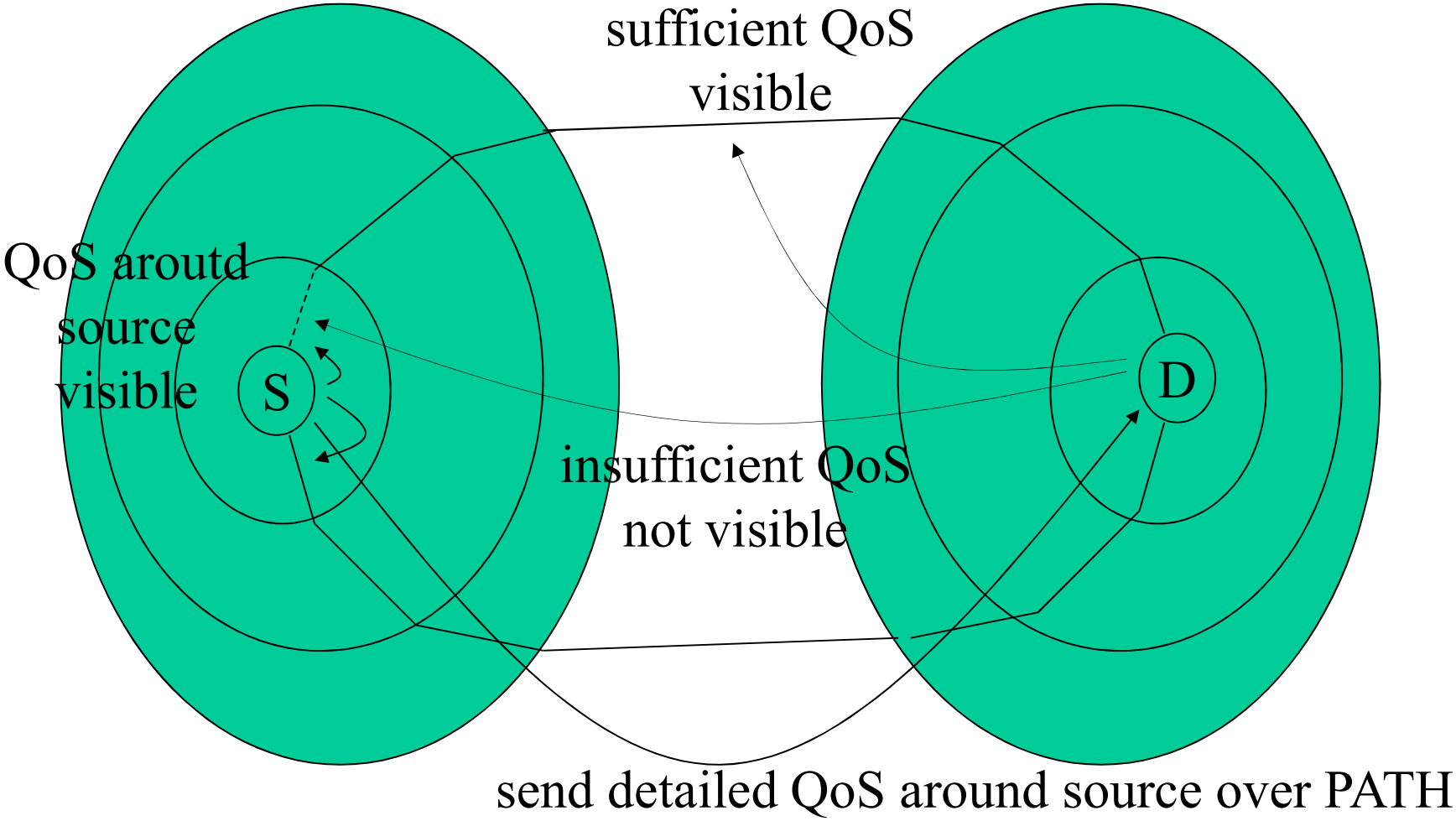
QoS around Source Viewd from Destination with Route Hierarchy



First Aggregated QoS for Hierarchical QoS Routing

- send detailed QoS information around source by PATH messages
  - source send lowest level QoS around itself
  - router between area add higher level QoS
- on each router
  - hierarchical QoS around it is delivered by routing protocol
  - hierarchical QoS around source is delivered by FAQ

QoS around Source Viewd from Destination with Route Hierarchy

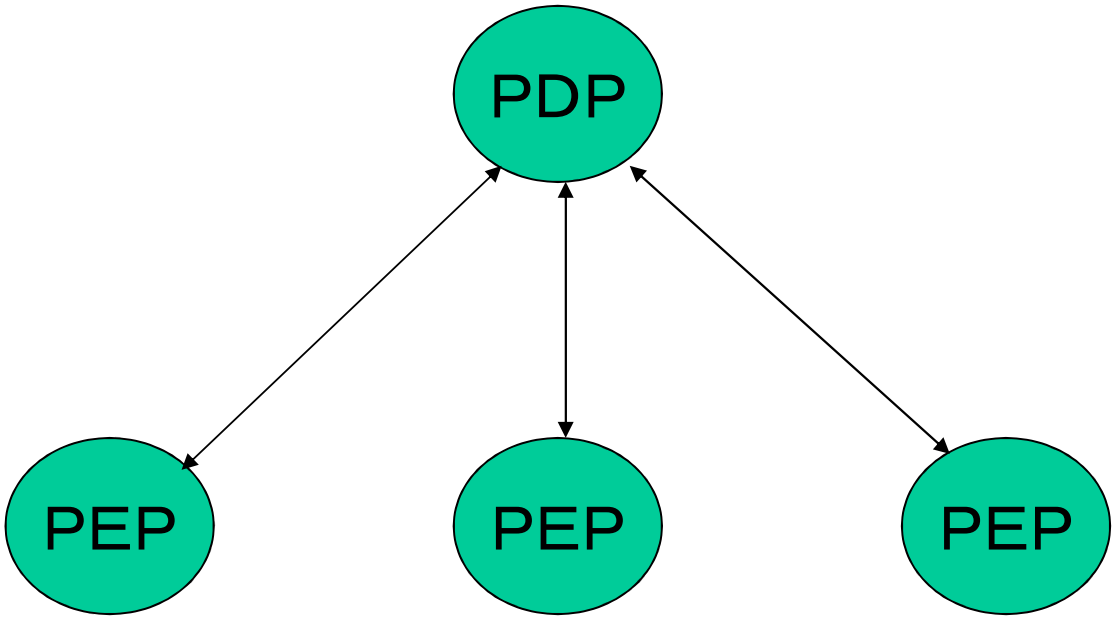


QoS Guarantee and Policy

- QoS guaranteed communication is prioritized
  - not meaningful if all are prioritized
- which communication should be prioritized?
  - judge policy for individual communication?
    - maybe OK for LAN, does not scale in WAN
- let money talk
  - usage based charge based on priority, duration, BW, etc.
  - communication for emergency should have absolute priority

PEP and PDP?

- framework for policy control
- PEP (Policy Enforcement Point)
  - where policy is implemented, routers
- PDP (Policy Decision Point)
  - PCE centrally compute policy
- against the E2E principle
  - load concentration, lack of reliability, not enough information, etc.


PDP
PEP
PEP
PEP

Policy and Charge

- flat rate for best effort communicaiton
  - ISPs throw traffic to other ISPs ASAP
  - policy can be determined by ISPs
- usage base for QoS guaranteed communication
  - ISPs keep traffic as much as possible
  - policy must be determined by users
    - detailed charging information must be offered to users
      - ISPs can’t make charging infomation secret

Multicast and Charging

- policy is determined by users
  - route is determined by users
- different policy between destinations means different route
  - cannot share multicast tree
- policy must be determined by group manager
  - destinations follow or leave the group

SRSVP
(Simple RSVP)

- integrate multicast and QoS routing
  - PIM-SM style multicast
    - simplified without a lot of extensions
  - users may choose route based on weighted sum of delay and chage
    - source routing is also possible
  - link state routing protocol is also developed
    - HQLIP (Hierarchical QoS Link Information Protocol)

Wrap-up

- QoS guratantee is function of network
  - not possible by end
- PPQ with statistical traffic model for queueing
- PQC to stabilize route, FAQ for hierarchy
- integration of multicast and QoS routing essential
- usage based charge necessary for QoS guarantee
  - BGP is useless here