Coordinating Congestion Management and Bandwidth Sharing for Heterogeneous Data Streams

Venkata N. Padmanabhan

Microsoft Research

http://www.research.microsoft.com/~padmanab

NOSSDAV, June 1999

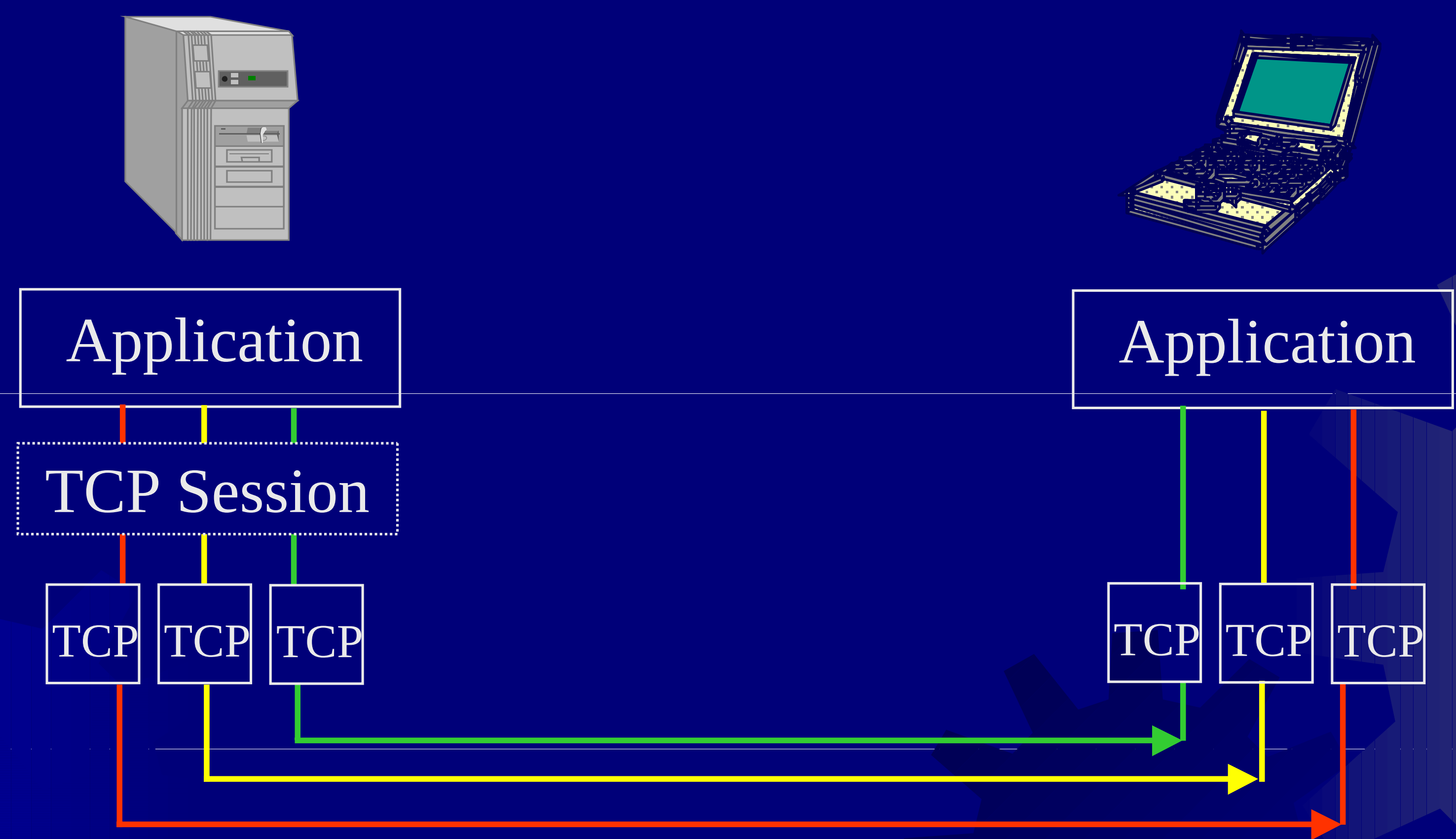
Outline

- Why coordinate?
- Two-host case: TCP Session
- Generalization of TCP Session
- Challenges and Potential Solutions
- Summary

Why Coordinate?

- Decouple congestion behavior from count of application-level data streams
  - *how much* data is what matters, not *which* data
- Be responsive to congestion at a shared bottleneck link
- Apportion bandwidth based on application-level notion of utility

TCP Session [Pad98]

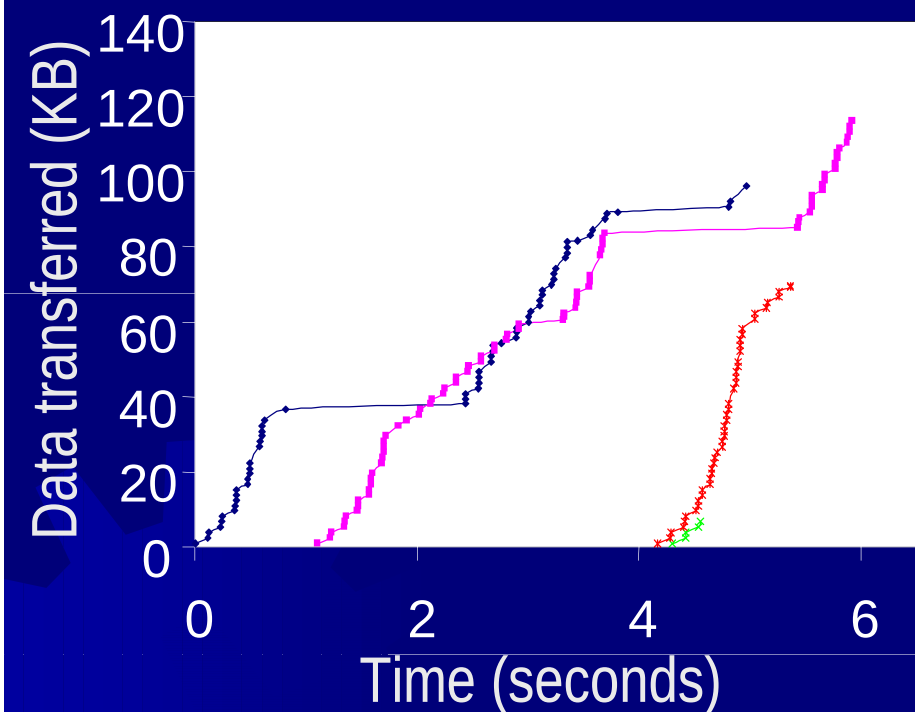


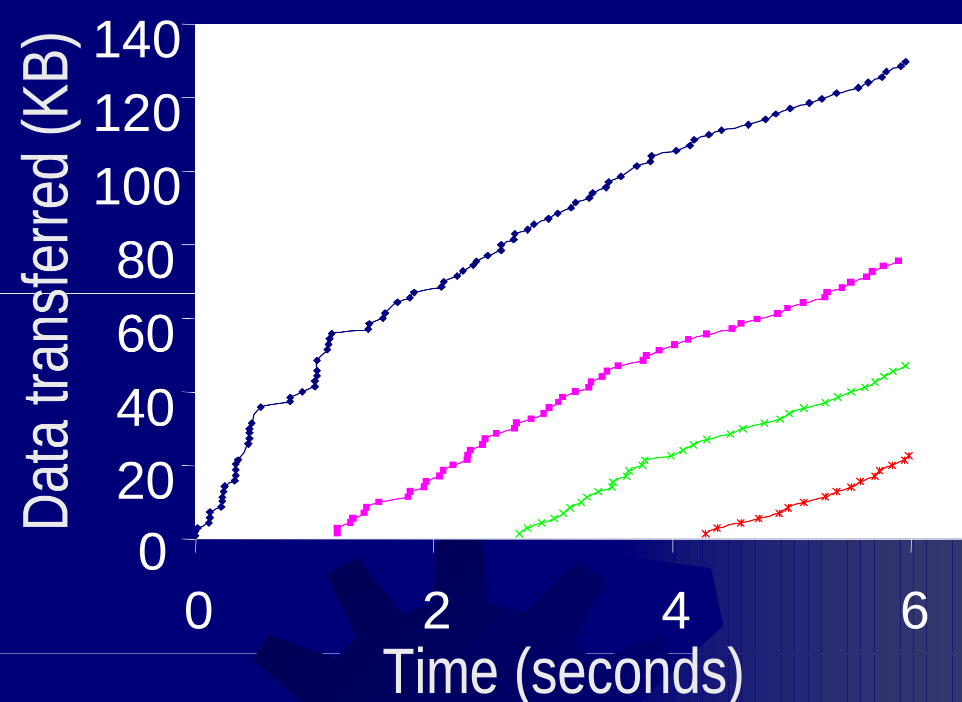
- Coordinate TCP connections between a host pair

TCP Session Components

- Integrated congestion control
  - emulate window dynamics of single connection
- Connection scheduling
  - explicitly apportion bandwidth to connections within a session
- Integrated loss recovery
  - use packet ordering information across connections to infer loss without timeout

Benefits of TCP Session



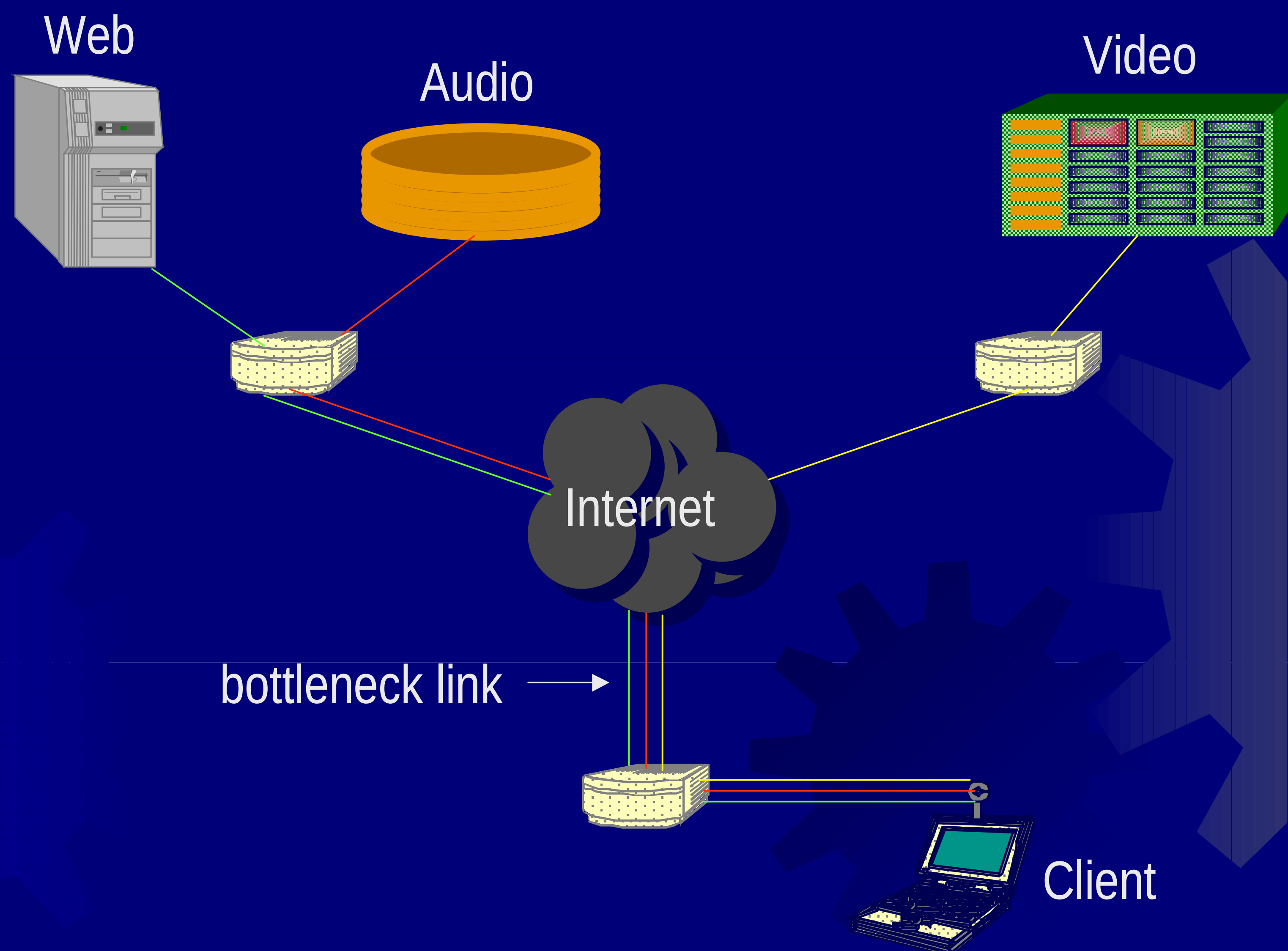


- Coordination ⇒ more predictable performance

Generalization of TCP Session

- Generalization along two axes
  - heterogeneity in data streams
    - TCP (e.g., HTTP) and non-TCP (e.g., RTP/UDP)
  - multiplicity of sources and sinks
    - e.g., server cluster, clients on a LAN
- Need to consider both axes
  - one may imply the other
    - e.g., distinct server nodes for Web & streaming media
  - ⇒ should not ignore either axis (unlike TCP Session, Congestion Manager [RBS99], Detour [SAA+99])

A General Scenario



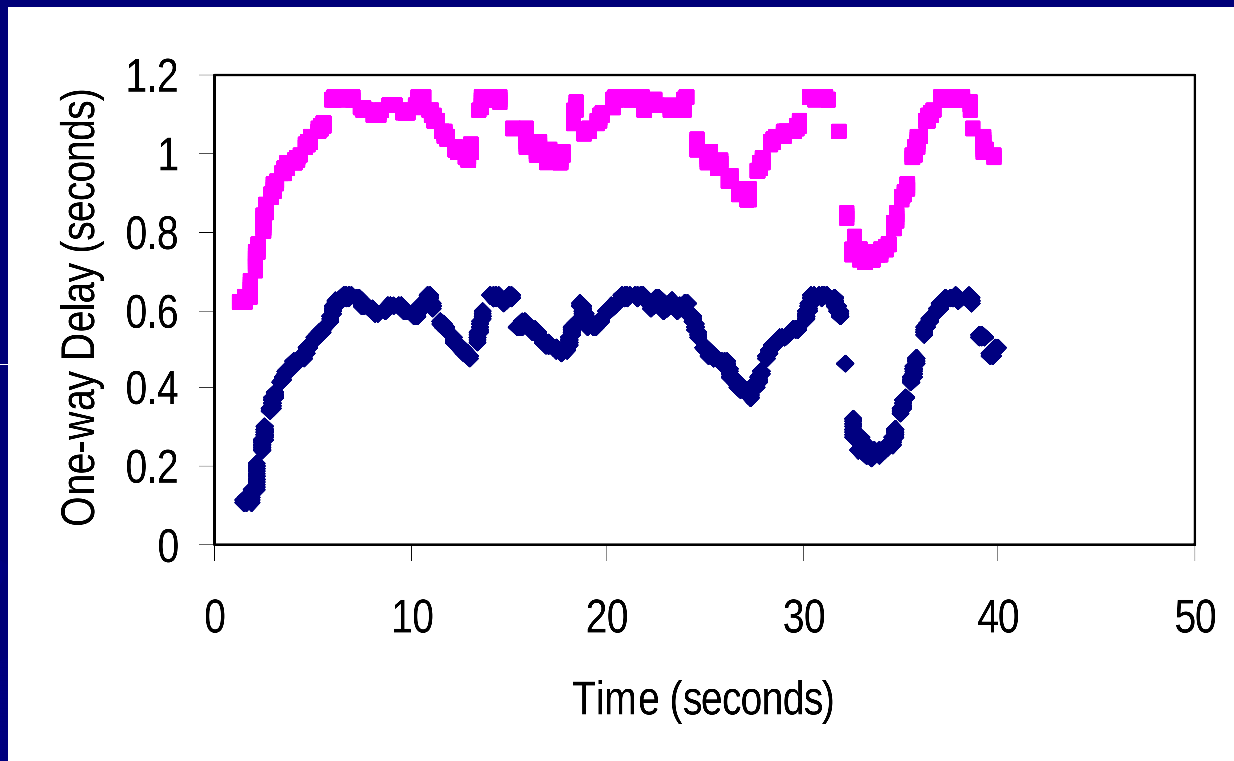
Opportunities for Optimization

- Initialization of congestion state
  - RTP stream as probe for short TCP connection
  - less effective in the presence of diffserv
- Improved loss recovery
  - exploit “heartbeat” provided by RTP stream
  - effective even with priority dropping
- Bandwidth sharing based on utility
  - e.g., progressive Web content delivery [GB99], SCUBA [AMK97]

Challenges & Potential Solutions

- Detecting presence of a shared bottleneck
  - indirect: compare delay/loss patterns
  - direct: ECN augmented with “unique” router ID
- Coordinating distributed senders
  - explicit receiver-driven flow control
  - congestion feedback filtering
- Differing notions of fairness and utility
  - sender versus receiver
  - essentially a matter of policy

Using Delay Dynamics to Detect Presence of Shared Bottleneck

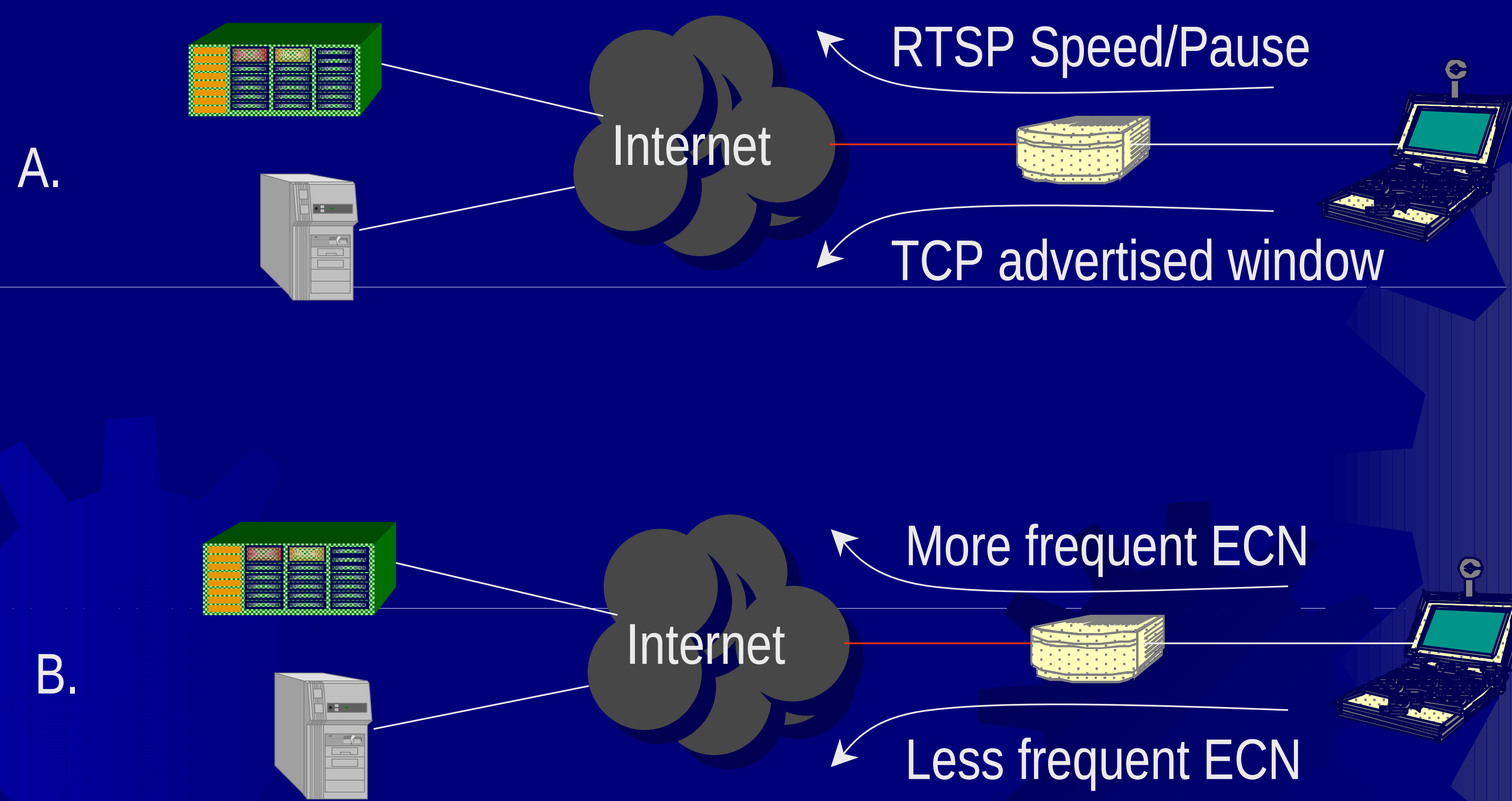


- Resilient to priority dropping
- But not resilient to multiple congested gateways

Challenges & Potential Solutions

- Detecting a shared bottleneck
  - indirect: compare delay/loss patterns
  - direct: ECN augmented with "unique" router ID
- Coordinating distributed senders
  - explicit receiver-driven flow control
  - congestion feedback filtering
- Differing notions of fairness and utility
  - sender versus receiver
  - essentially a matter of policy

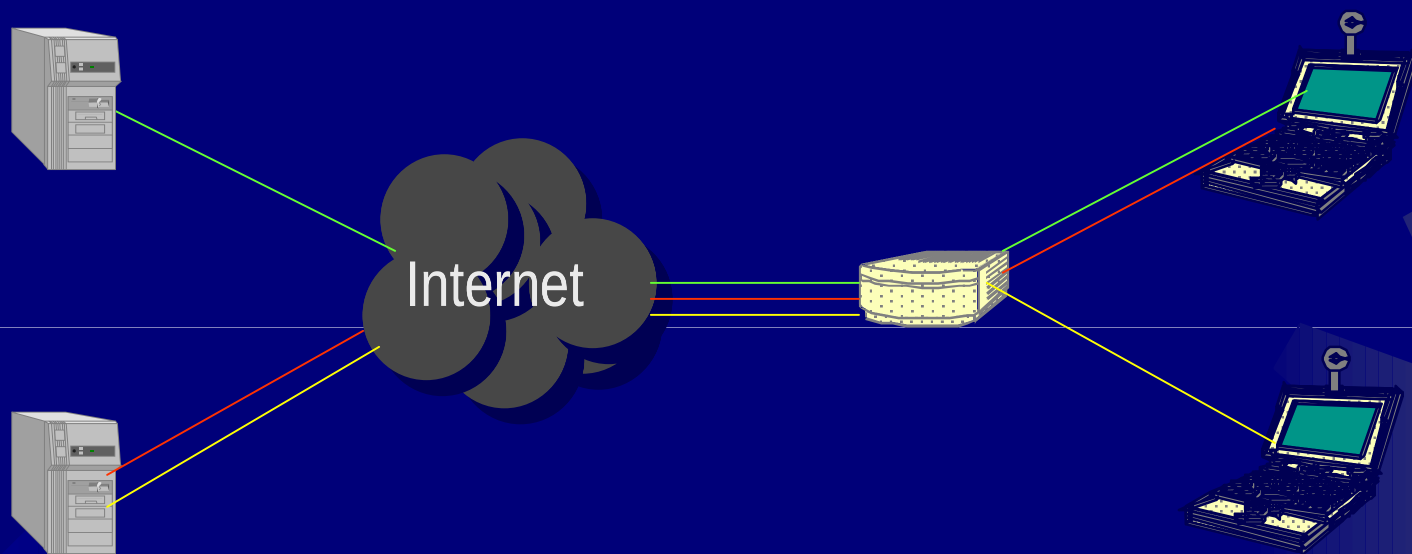
Coordinating Distributed Senders



Challenges & Potential Solutions

- Detecting a shared bottleneck
  - indirect: compare delay/loss patterns
  - direct: ECN augmented with “unique” router ID
- Coordinating distributed senders
  - explicit receiver-driven flow control
  - congestion feedback filtering
- Differing notions of fairness and utility
  - sender versus receiver
  - essentially a matter of policy

Differing Notions of Fairness



- Should the servers split the bottleneck bandwidth equally or should the clients do so?

Multiple Levels of Coordination

- Intra-host
  - fine-grained, packet-level (as in TCP Session)
- Intra-cluster
  - coarse-grained, at the level of congestion epochs
  - coordination bus [McC96] may be used with messages indexed by peer addr/bottleneck link ID
- Inter-cluster
  - more coarse-grained, at the level of user actions
  - gated by the receiver

Summary

- Several potential benefits of coordination between hosts and heterogeneous streams
- Key challenges
  - detecting presence of shared bottleneck
  - mechanism for coordinating distributed senders
  - defining fairness and utility
- Need multi-level coordination scheme with participation of both senders and receivers