OptIPuter Networking and Middleware

Joe Bannister
Aaron Falk
Carl Kesselman
Marcus Thiébaux
USC Information Sciences Institute

February 7, 2003

The Need for High-Performance Transport

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- TCP performance suffers in some environments
  - High bandwidth-delay networks
  - Cautious probing of bandwidth
  - Aggressive backoff after presumed packet loss
- Designed for under-engineered networks
  - Stability and farness for all
  - Information comes from the edges of the network
  - Crude congestion signals from the core

Explicit Control Protocol (XCP)

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

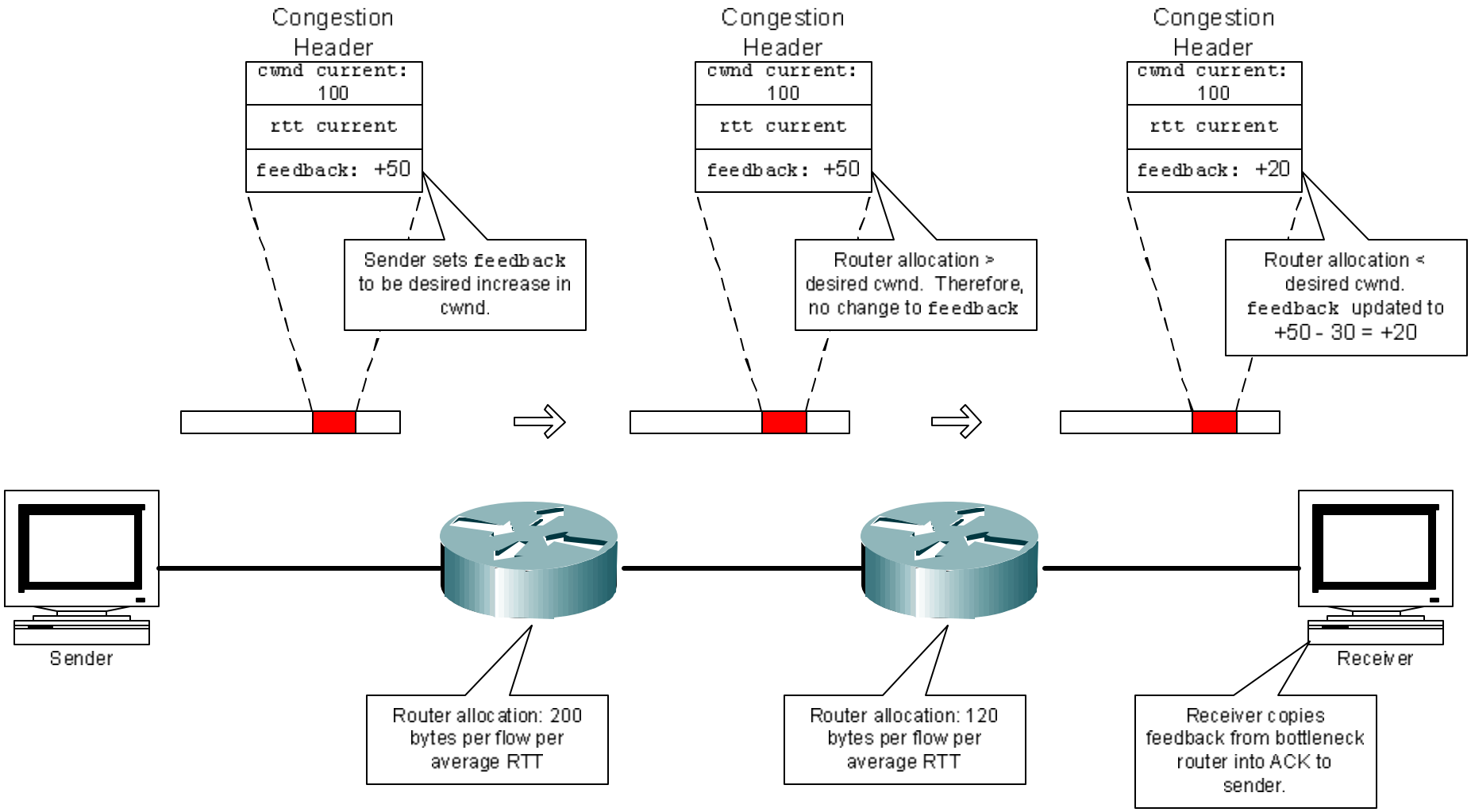
- XCP is a newly proposed congestion-control scheme

  Dina Katabi, Mark Handley, and Charles Rohrs, "Internet Congestion Control for Future High Bandwidth-Delay Product Environments," *Proc. ACM Sigcomm 2002*, Aug. 2002. http://ana.lcs.mit.edu/dina/XCP/

- Decouples fairness control from congestion control
  - Use aggressive window adjustments (MIMD) for congestion control which can grab/release large bandwidth quickly
  - More conservative window adjustments (AIMD) provide fairness control which converges to fair bandwidth allocation
- Bottleneck router explicitly tells sender about bandwidth allocation
  - Applications may request explicit send rate
  - XCP requires modified router behavior but **no per-flow state in routers**
  - State carried in congestion header in packets

Congestion Window Set by Bottleneck Router

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI



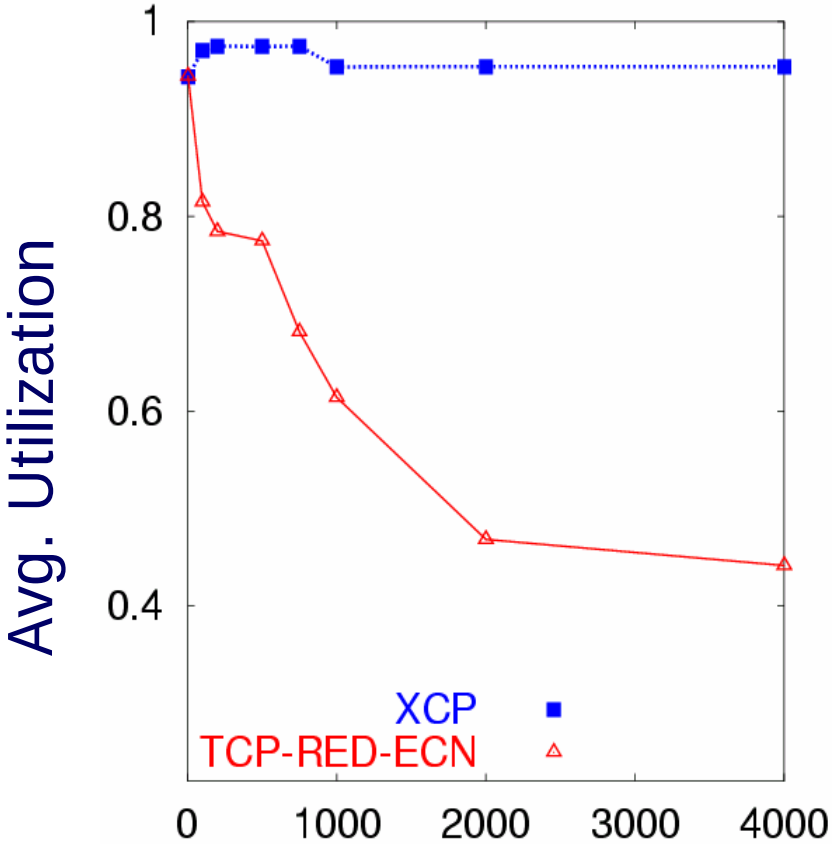
XCP Performance

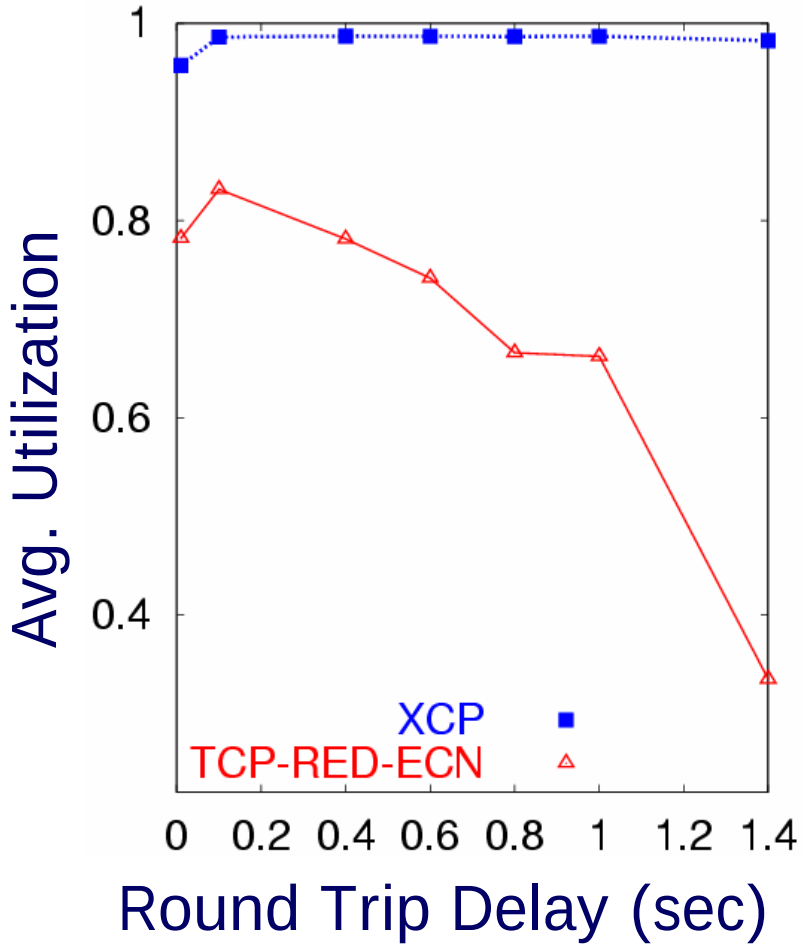
USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- Superior to TCP
  - Almost never drops packets
  - Converges to available bandwidth very quickly, ~1 RTT
  - Fair over large variations in bandwidth and RTT
- Supports existing TCP semantics
  - Replaces only congestion control, reliability unchanged
  - No change to application/network API
- Status
  - Analysis based entirely on ns-2 simulations
  - Software design and coding begun (multiple implementations)
  - Experimentation and interoperability testing planned
  - Researcher and vendor collaborations

Comparison of XCP and TCP

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI





From http://www.ana.lcs.mit.edu/dina/XCP/Sigcomm2002.ppt

Credits and Collaboration

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- Dina Katabi – proposed XCP
- USC ISI
  - Aaron Falk
  - Ted Faber
  - Bob Braden
- MIT LCS
  - John Wroclawski
  - Dina Katabi
- UCB ICIR
  - Sally Floyd
  - Mark Handley (till May 2003)
- UCL
  - Mark Handley (after May 2003)
- PSC
  - Matt Mathis

Progress and Plans

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- Wide ongoing collaboration effort
  - Mailing list
  - Web site
- Design, implementation, and interoperability
  - Protocol design
  - ns-2 simulations
  - Coding
    - Linux/FreeBSD hosts and routers
    - IPX-1200
  - Globus Toolkit integration
    - Application-level performance
  - IETF and standardization
    - A. Falk, Transport Area Directorate, PILC/DCCP WG Chair
  - Vendor adoption
    - Collaboration and prototyping
    - Market plans

ISI Personnel for XCP Development Task

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- Aaron Falk: software development leader, standardization
- Ted Faber: software development, congestion management
- Venkata Pingali (GRA): simulation, performance analysis
- Bob Braden: collaboration liaison, senior advisor
- Joe Bannister: task manager

Why Use XCP?

USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISIDIV7USCISI

- XCP is a serious candidate to replace TCP congestion control in the Internet
- No change to bulk of TCP kernel code or API semantics
- Excellent performance, link utilization, and packet loss across wide ranges of RTT, link bandwidth, flow size
- Capable of fair coexistence with other Internet protocols
- Gradual Internet deployment possible
- But ... requires changes to routers