HSTCP-LP: A Protocol for Low-Priority Bulk Data Transfer in High-Speed High-RTT Networks

Aleksandar Kuzmanovic
Edward W. Knightly
Rice University
R. Les Cottrell
SLAC/SCS-Network Monitoring

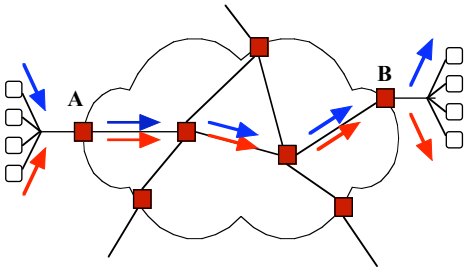
Rice Networks Group
http://www.ece.rice.edu/networks

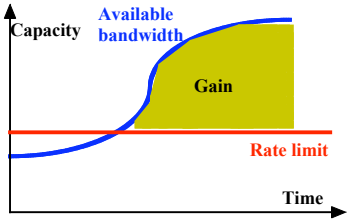
Motivation

- Traditional view of service differentiation:
  - High priority: real-time service
  - Best-effort: everything else
- What's missing?
  - **Low-priority (receiving only excess bandwidth)**
  - **Lower than best-effort!**
    - Non-interactive applications, bulk download
      - »It will last long anyway...
    - Speeds up best-effort service (e.g., web)

Applications for Low Priority Service

- LP vs. fair-share:
  - Bulk downloads
    - **SLAC->CERN**
    - Improve my **other** applications
- LP vs. rate-limiting:
  - P2P file sharing
    - Often rate limited
    - Isolation vs. sharing





Kuzmanovic, Knightly, and Cottrell

RICE

Problem Formulation & Design Objectives

Design end-host-based transmission protocol that emulates the low-priority service

- Low-priority service objectives
  - Utilize the "excess/available" capacity
    - What no other flows are using
  - TCP-transparency (non-intrusiveness)
  - Inter-LP flow fairness (fair-share of the available bandwidth)

Kuzmanovic, Knightly, and Cottrell

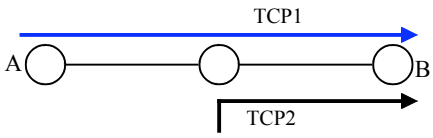
RICE

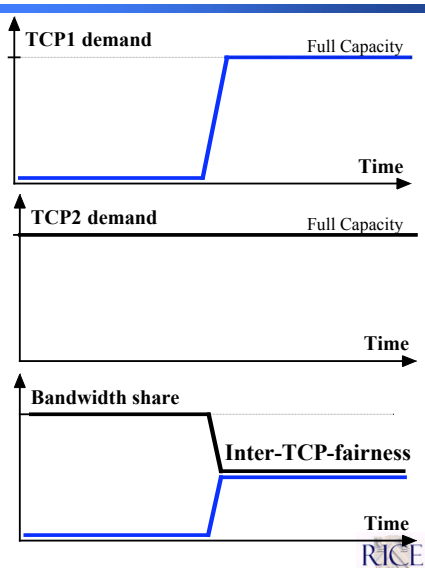
Origins of the Available Bandwidth

- Why is excess bandwidth available when TCP is greedy?
  - Time-of-day effects
    - Night- vs. day-time traffic
  - Short-lived flows
    - Majority traffic is web browsing
  - TCP is imperfect
    - Low-aggregation regimes

RICE

TCP fairness

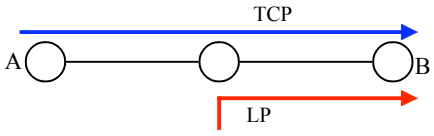


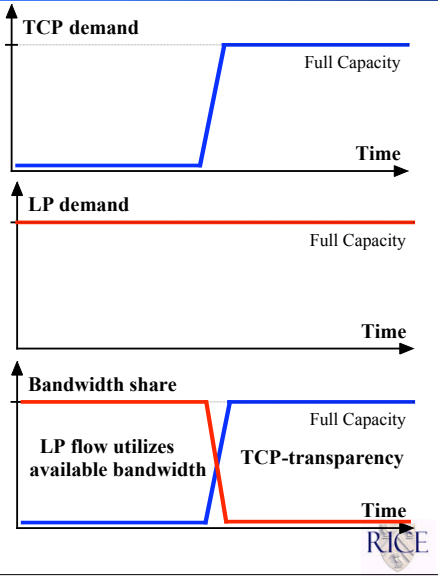


- In presence of TCP cross-traffic:
  - TCP achieves fairness

RICE

Illustration of TCP Transparency

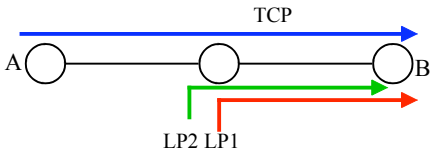


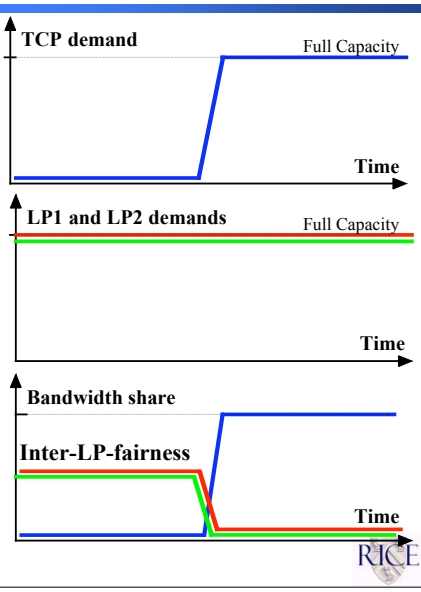


- LP flow utilizes only excess bandwidth
  - Does not reduce the throughput of TCP flows

Kuzmanovic, Knightly, and Cottrell

Fairness Among LP Flows



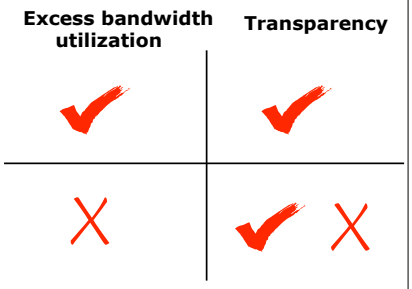


- Inter-LP-fairness is essential for simultaneous file transfers

Kuzmanovic, Knightly, and Cottrell

HSTCP-LP
A Congestion Control Protocol

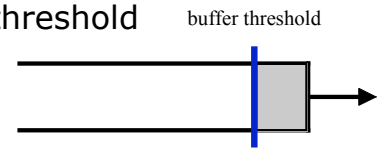
- Problem
  - TCP-LP [Infocom03]
    - AIMD control not suitable for high speeds



- Approach
  - Add HSTCP [Floyd03] high-speed mechanisms
- Challenge
  - Develop large windows and achieve low-prioritization at the same time

Early Congestion Indication

- For transparency, HSTCP-LP must know of congestion before non-LP flows
- Idealized objective: buffer threshold indication

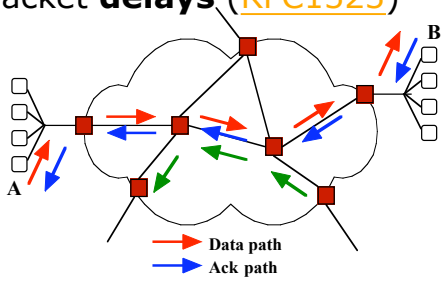


- Endpoint inference: one-way delay threshold

$$sd_i > d_{\min} + (d_{\max} - d_{\min})\delta$$

Reverse Cross-Traffic

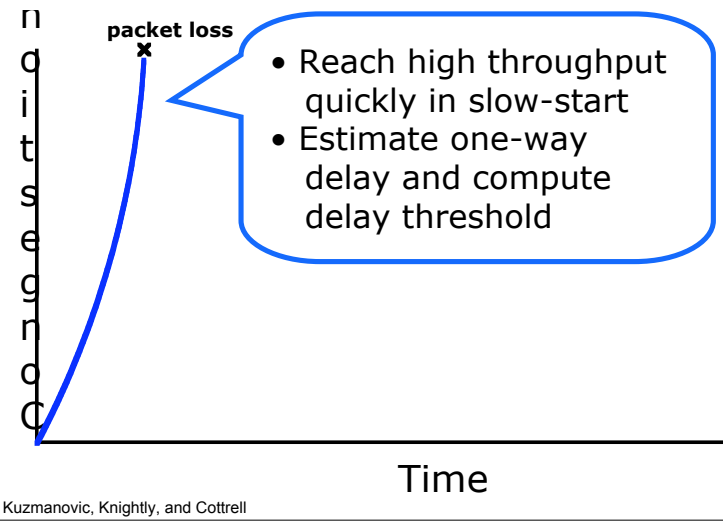
- HSTCP-LP uses **one-way** packet **delays** (RFC1323) for congestion indication
  - Source-destination time stamping
  - Synchronized clocks not needed



- **One-way** delay:
  - Eliminates the impact of **the reverse cross-traffic**
    - HSTCP-LP can loose **Gb/s** of available bandwidth if RTT is used

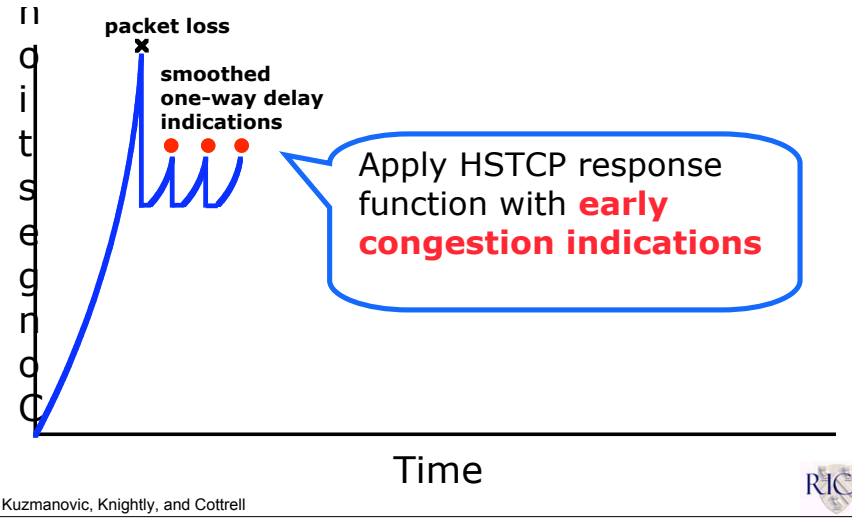
RICE

HSTCP-LP Timeline Illustration



RICE

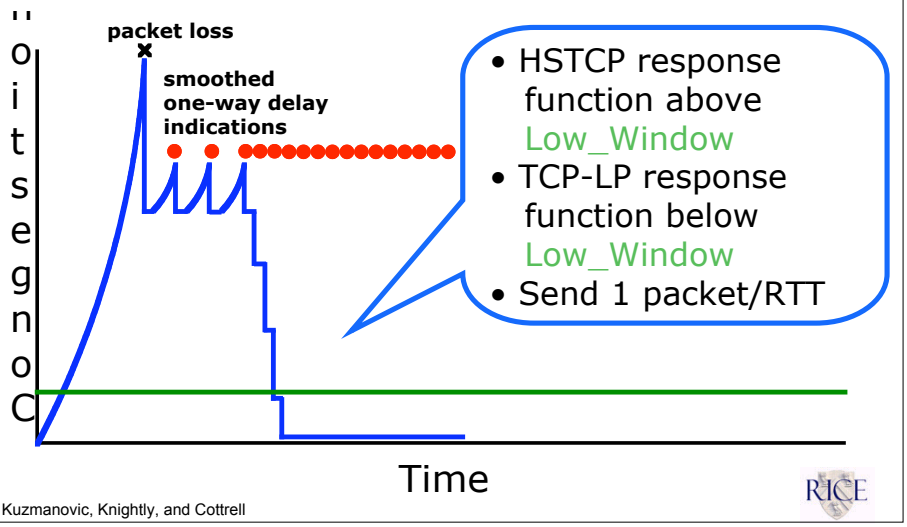
HSTCP-LP Timeline Illustration



Kuzmanovic, Knightly, and Cottrell

RICE

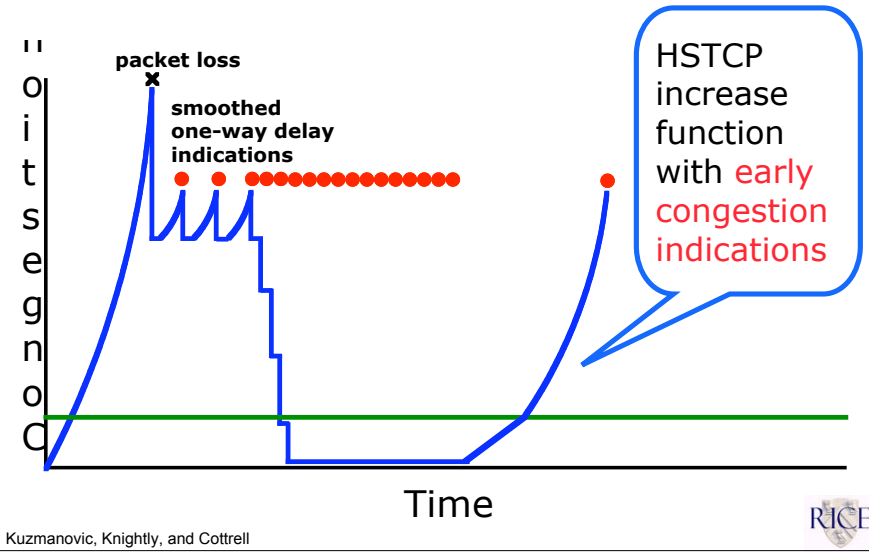
HSTCP-LP Timeline Illustration



Kuzmanovic, Knightly, and Cottrell

RICE

HSTCP-LP Timeline Illustration

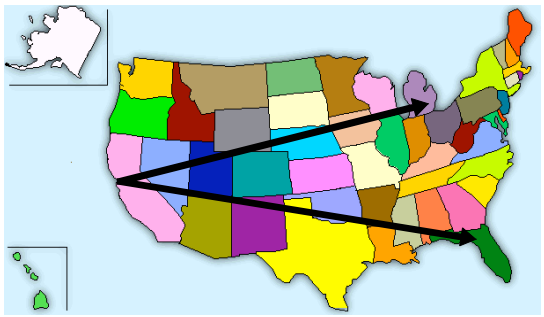


Kuzmanovic, Knightly, and Cottrell

RICE

Implementation and Experiments

- Linux-2.4.22-web100 kernel:
  - HSTCP ⇨ HSTCP-LP
    - Time stamping, modified congestion control...
- Paths:
  - Stanford, CA
    Ann Arbor, MI
  - Stanford, CA
    Gainesville, FL
  - Available bitrate
    ~ 450 Mb/s

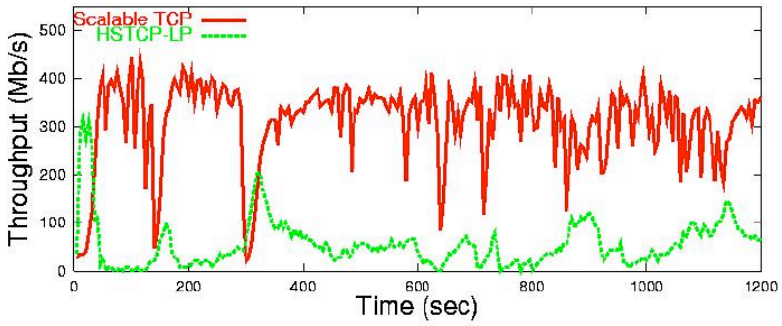


Kuzmanovic, Knightly, and Cottrell

RICE

Low-Priority Background Data Transfer

- HSTCP-LP is highly non-intrusive to other advanced TCP stacks



  - *Strict* prioritization for larger bottleneck queue lengths
  - Lighter prioritization for (max_delay < 50 ms)

Kuzmanovic, Knightly, and Cottrell

RICE

Utilizing Excess Bandwidth

- Extensive experiments on a number of high bandwidth-delay production networks
  - no cross-traffic
  - periodic UDP cross-traffic
  - reverse cross-traffic
  - **http://www.slac.stanford.edu/~hadrien**
- HSTCP-LP utilizes **80-127%** of the available bandwidth when compared to other high-speed stacks
  - 127% when transmission buffer *(tqueuelen)* is large

Kuzmanovic, Knightly, and Cottrell

RICE

Conclusions

- HSTCP-LP: a new protocol for high bandwidth-delay production networks
  - General low priority service (compared to "best-effort")
- Attractive for low-priority bulk downloads: ftp, web updates, P2P
- HSTCP-LP is incrementally deployable
  - Sender side modification of HSTCP without changes to routers
- Implementation and evaluation in the Internet

http://www.ece.rice.edu/networks/TCP-LP

Questions

http://www.ece.rice.edu/networks/TCP-LP