Transport Mechanisms for Next Generation Networks and Applications

Chi Zhang
czhang@cs.fiu.edu
School of Computer Science
Florida International University

Chinese American Networking Symposium
CANS 2004

Introduction – Congestion Control

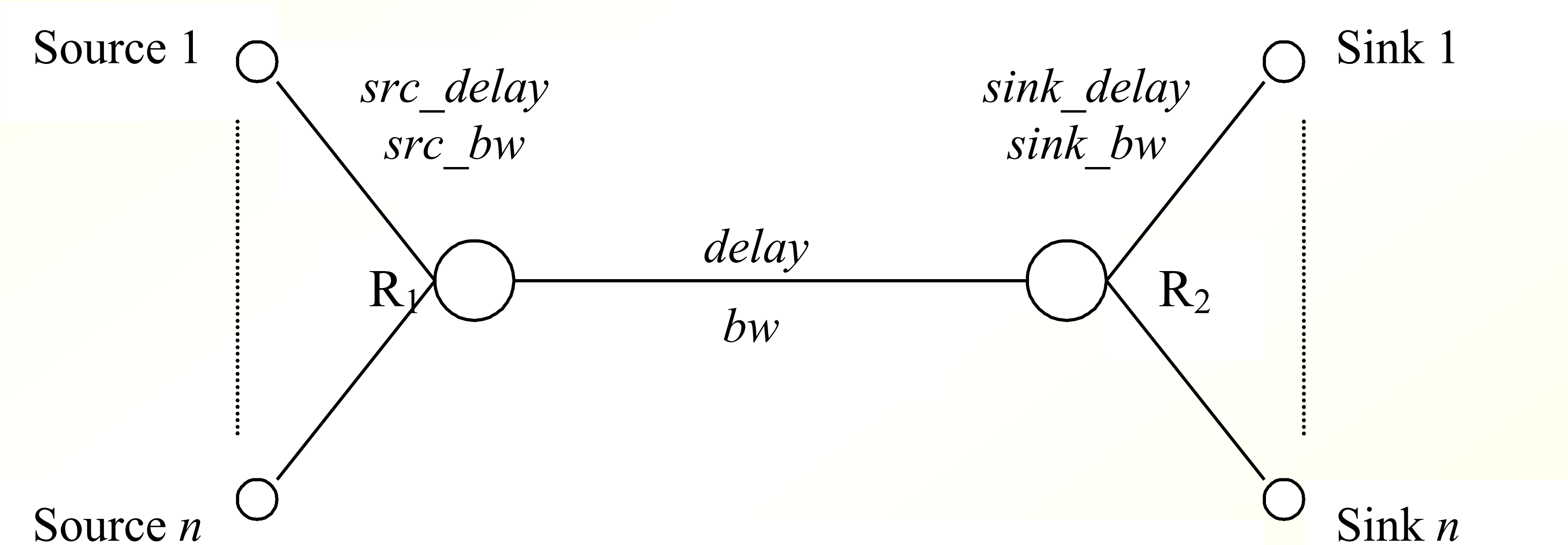
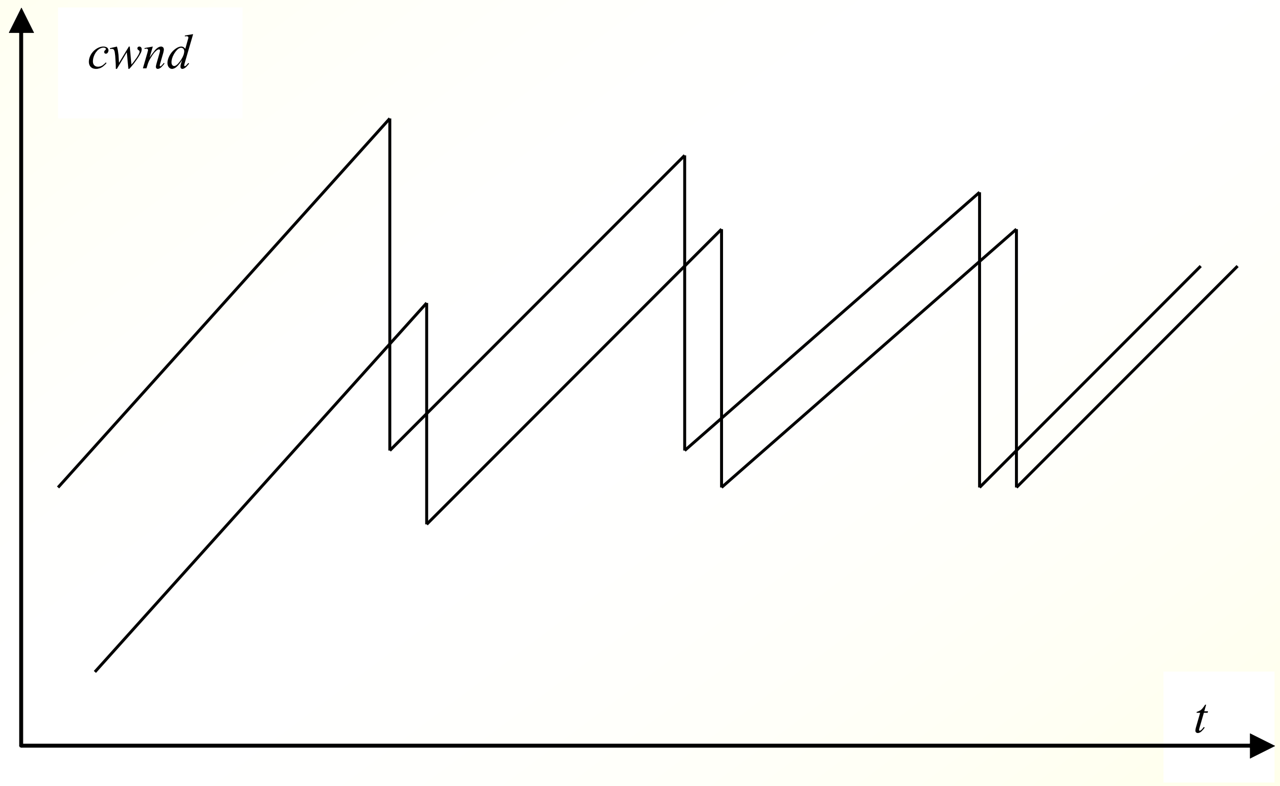


Figure 1. A Simple Network Topology

- Bandwidth Efficiency
- Fairness
- Avoid Congestion Collapse

Introduction - AIMD



- Bandwidth Efficiency
- Fairness
- Avoid Congestion Collapse

Introduction – Problems of TCP

- "Blind" AIMD window adjustment based on packet drops.
  - Packet loss may be caused by random wireless bit errors
  - ACK loss can be judged as packet loss over asymmetric path.
  - Network feedback is received only when congestive drops occur
  - No indication of the level of contention / bandwidth under-utilization
  - Transmission control parameters are static rather than adaptive

Introduction – Problems of TCP

- Appropriate for bulk-data transfer over wired networks
- Problems with the existence of wireless links
  - Unnecessary congestion-oriented response to wireless link errors
- Problems with real time applications
  - MD with a factor of 2 hurts the smoothness.
- Problems with High-Speed Networks for High-Performance Computing
  - Tera- / peta- bytes transfer, 1~100Gbps

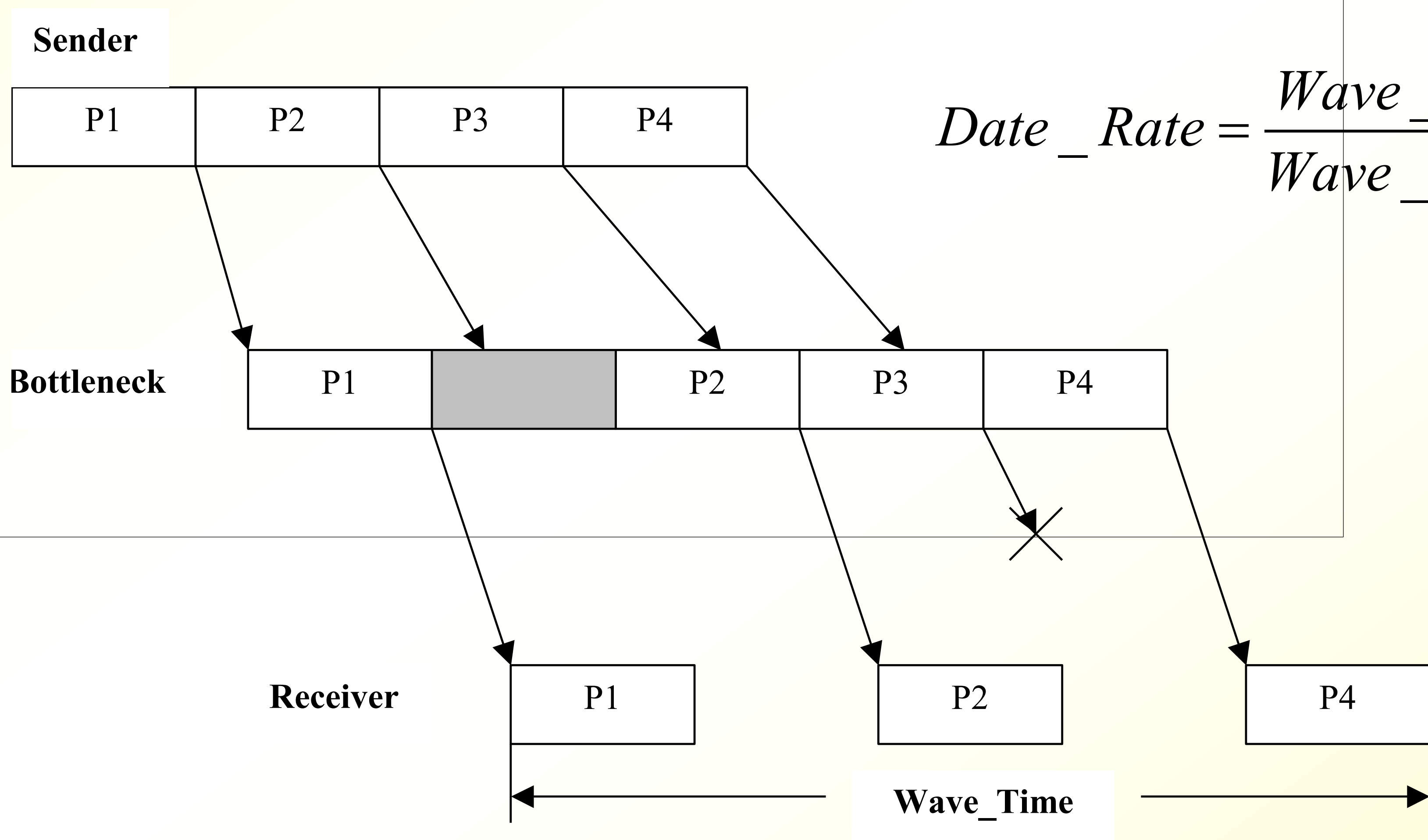
Introduction – Our Solutions

- TCP Real: High-throughput and energy-efficient transport over heterogeneous (wired/wireless) wireless networks
- TCP(α, β, γ, δ): Improve and Stabilize TCP Throughput for Competing Real-time Applications

Novel Congestion Detection of TCP-Real

- Receiver-oriented
  - Solves the Asymmetric-Path Problem
- Measurements-based
  - Wave – data packets sent “side by side”
    - The TCP sender sends packets in waves.
    - Wave-size and wave-sequence number is attached as TCP option.
  - The data-receiving rate of the wave is measured at the receiver and is attached to ACKs sent back to the sender.

Error Detection – Measurement-based



$$Date\_Rate = \frac{Wave\_Size}{Wave\_Time}$$

Distinguish the nature of the errors – Wired or Wireless?

- Data-receiving rate is determined by the interleaving patterns
- The lower the rate, the higher the contention.
- During the period of congestion, the data-receiving rate might decrease significantly, or fluctuate dramatically.
- The data-receiving computation should not be affected if the packet drop is due to transient wireless errors
- **If the recent data-receiving rates do not justify a congestion, the congestion window will not be reduced**

Heterogeneous (Wired & Wireless) Networks

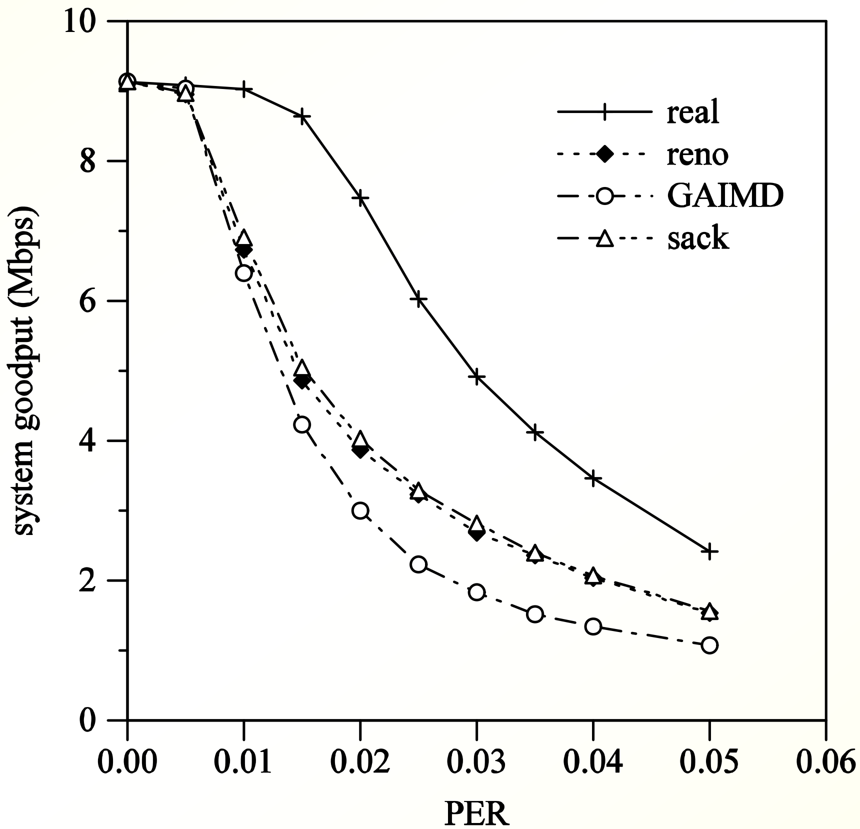


Figure 13. Goodput over Heterogeneous Network (10Mbps bottleneck link, 10 flows)

Heterogeneous Flows

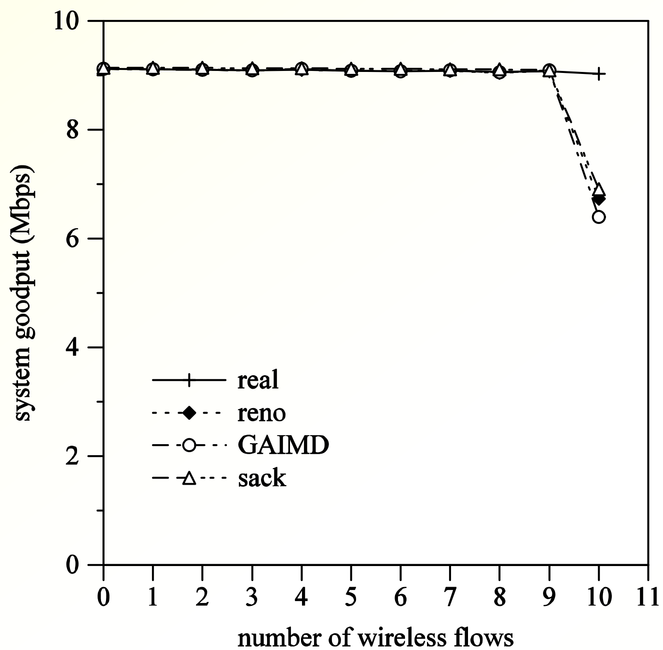


Figure 15. Goodput with Wired and Wireless Flows (10Mbps bottleneck, PER = 0.01, 10 flows)

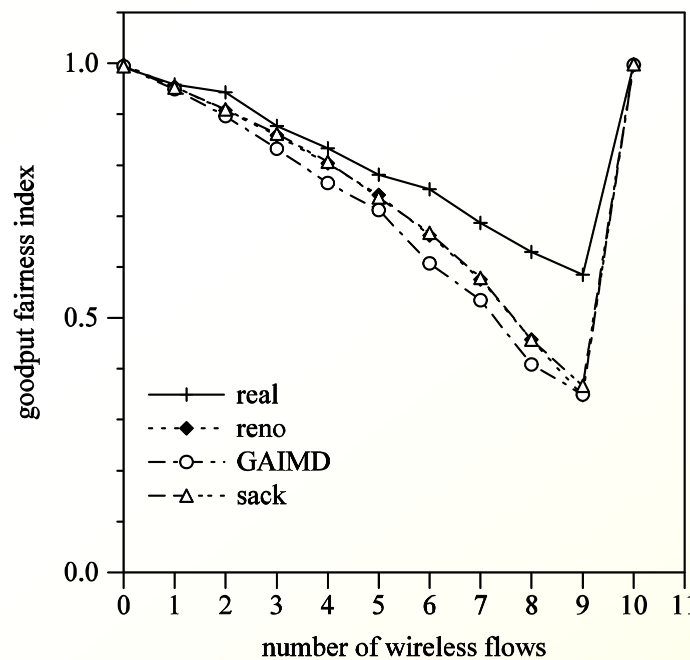


Figure 16. Fairness with Wired and Wireless Flows (10Mbps bottleneck, PER=0.01, 10 flows)

TCP(α,β) Protocols

- parameterize the congestion window increase value α and decrease ratio β
- Increase β to achieve *smooth* window adjustment upon congestion. (β = 0.875)
- At the expense of *responsiveness*: reduce α correspondingly to compete friendly with TCP(1, 0.5), according to a TCP steady-state throughput equation: (α=0.31)

Smoothness Achieved at the Expense of Responsiveness

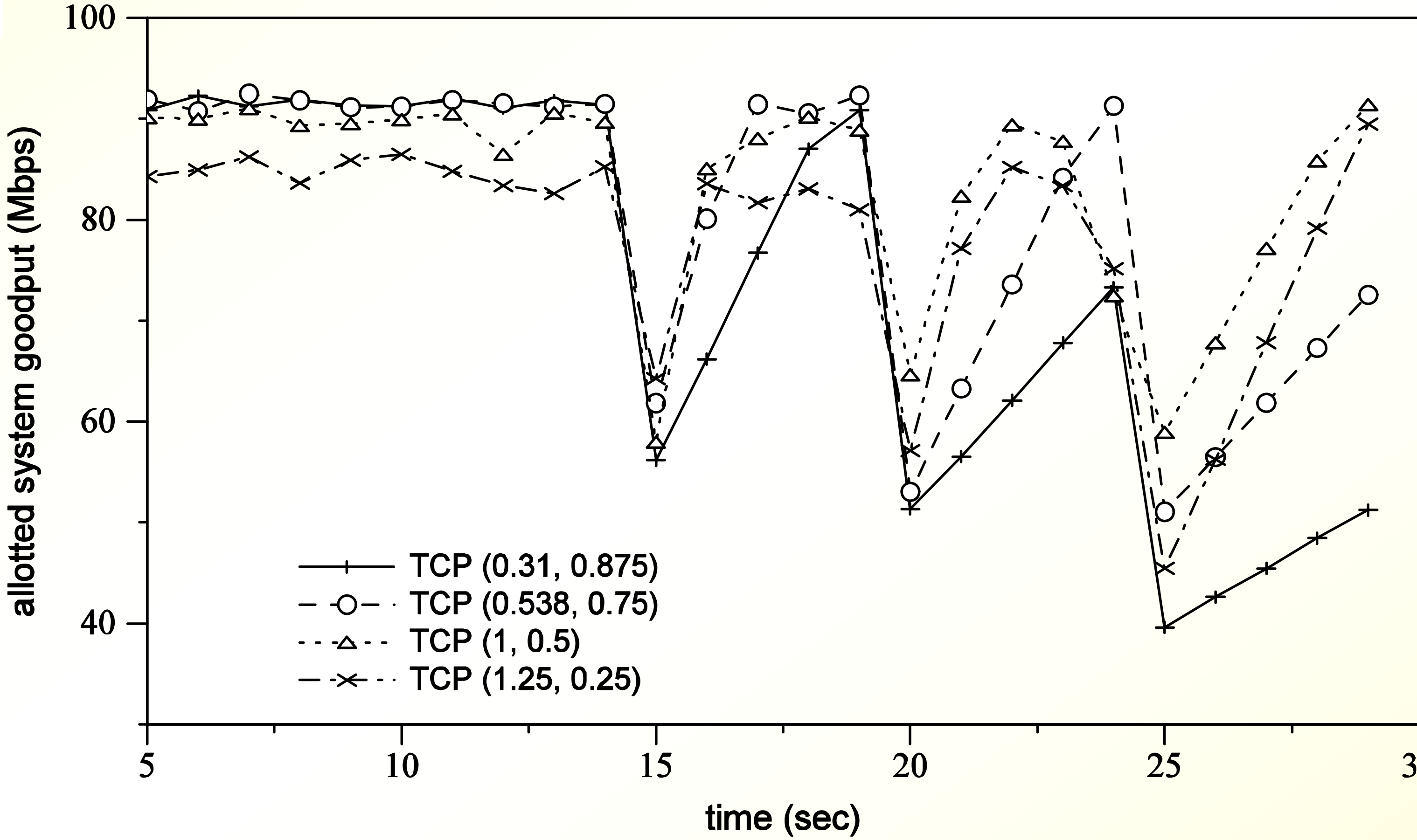


Figure 7. Allotted goodput with decreasing number of flows (100Mbps bottleneck link)

TCP(α, β, γ, δ)

- Measurement-based
  - Based on fine-grained RTT measurements
  - Network feedback is received before congestive drops occur
  - Indicating the level of contention / bandwidth under-utilization
  - Indicating the relative size of buffer
- Adaptive Parameters
- γ: Coordinated Window Adjustments; Congestion Avoidance
- δ: Enhance responsiveness when the capacity is underutilized

Oscillations with Unsynchronized Adjustments

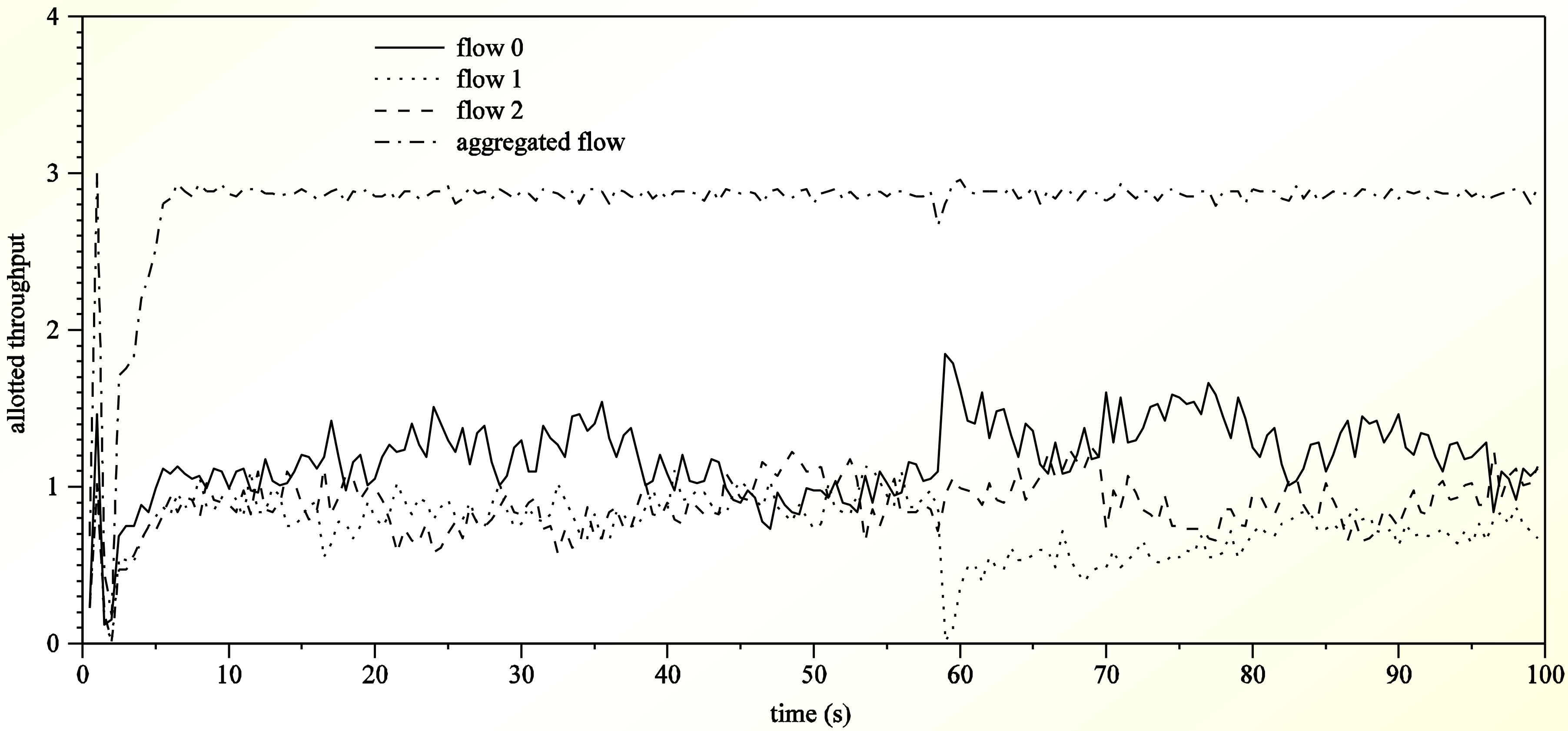


Figure 19. Allotted Throughput with TCP(0.31, 0.875)

Stabilized Throughput and Smoothness

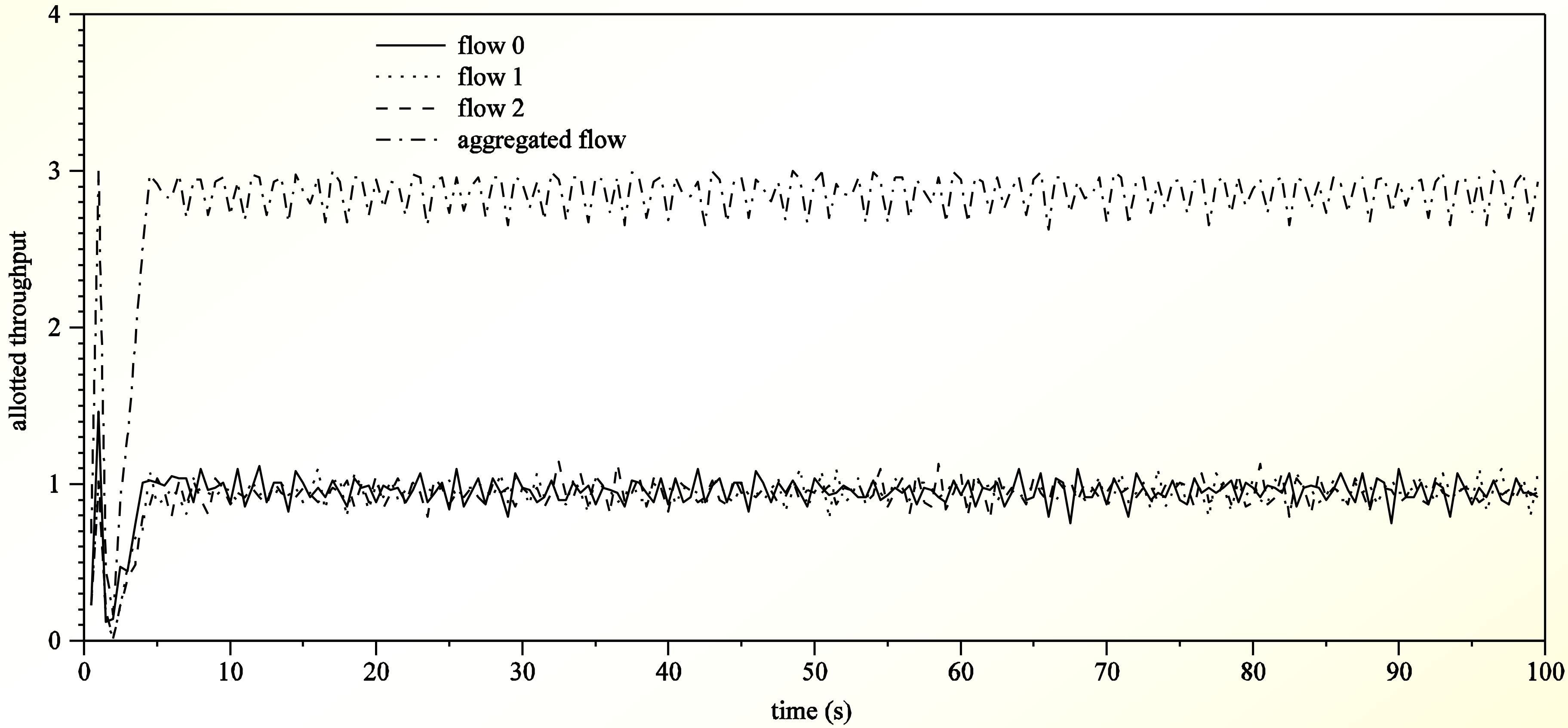


Figure 18. Allotted Throughput with Reno + Gamma

Responsiveness

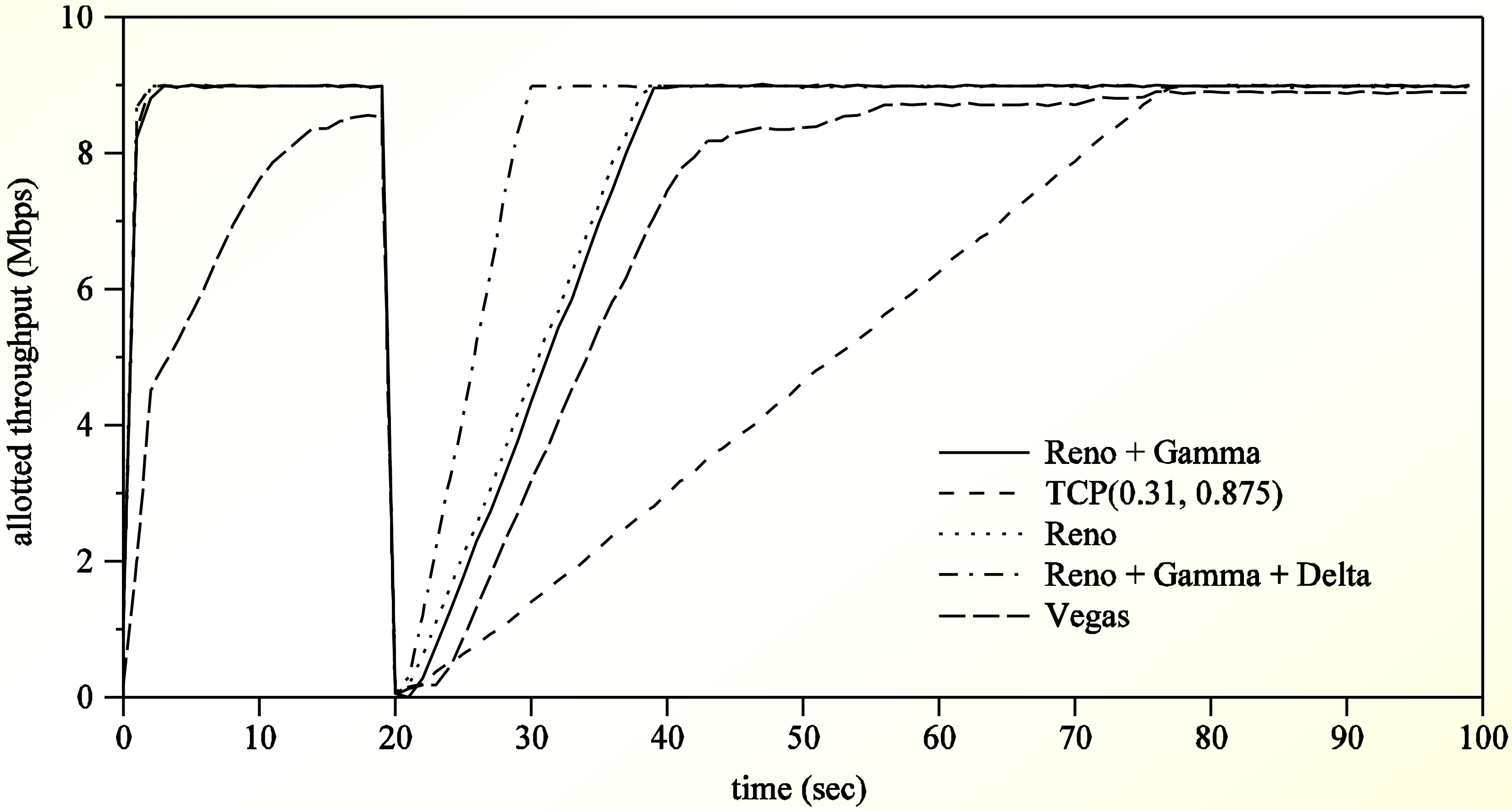


Figure 17. Thoughput with 1.0 Second Handoff

TCP over OC12

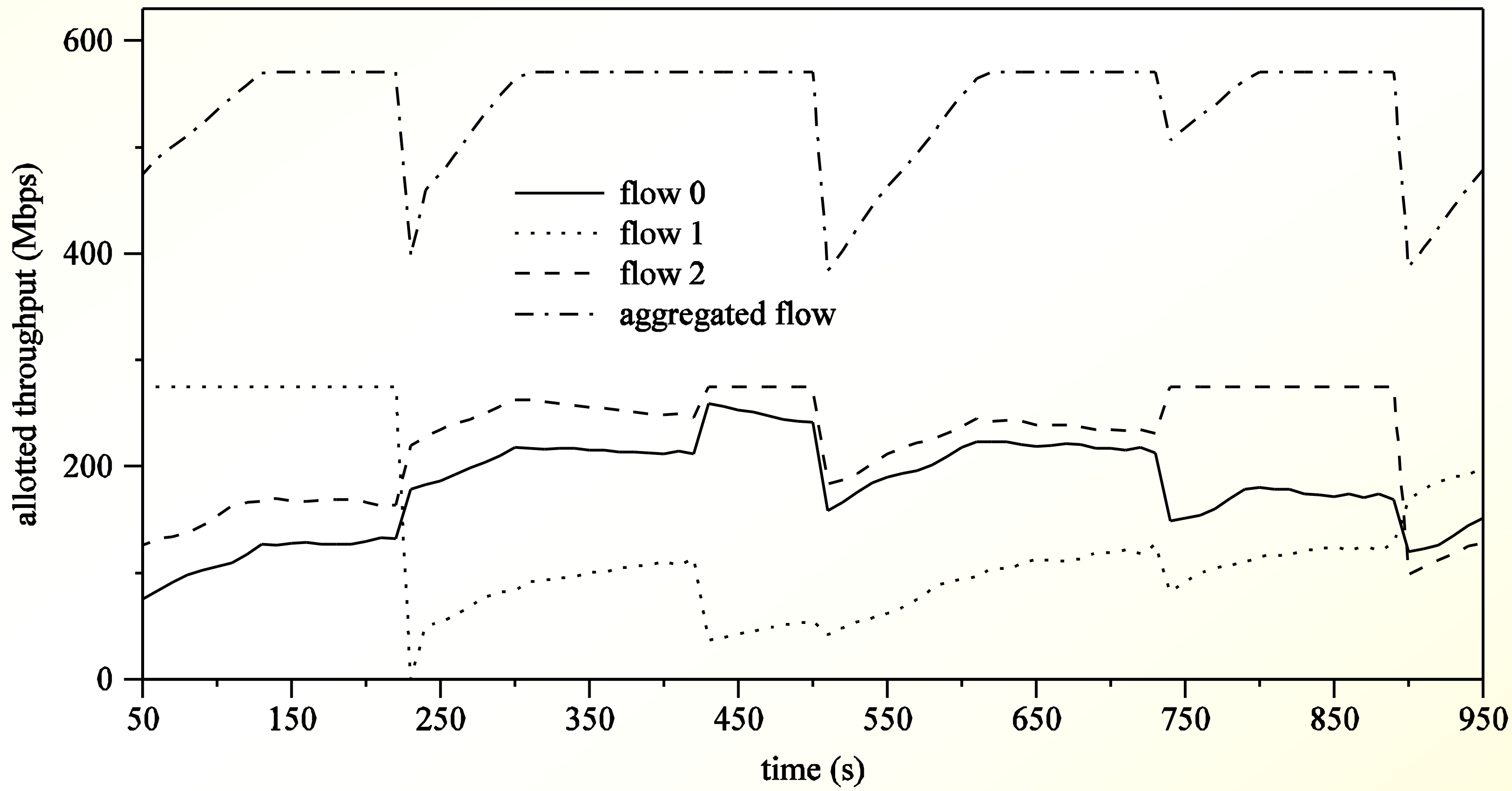


Figure 5. Allotted Throughput with Reno
(3 competing flows over a 622 Mbps bottleneck link)

TCP (α, β, γ, δ) over OC12

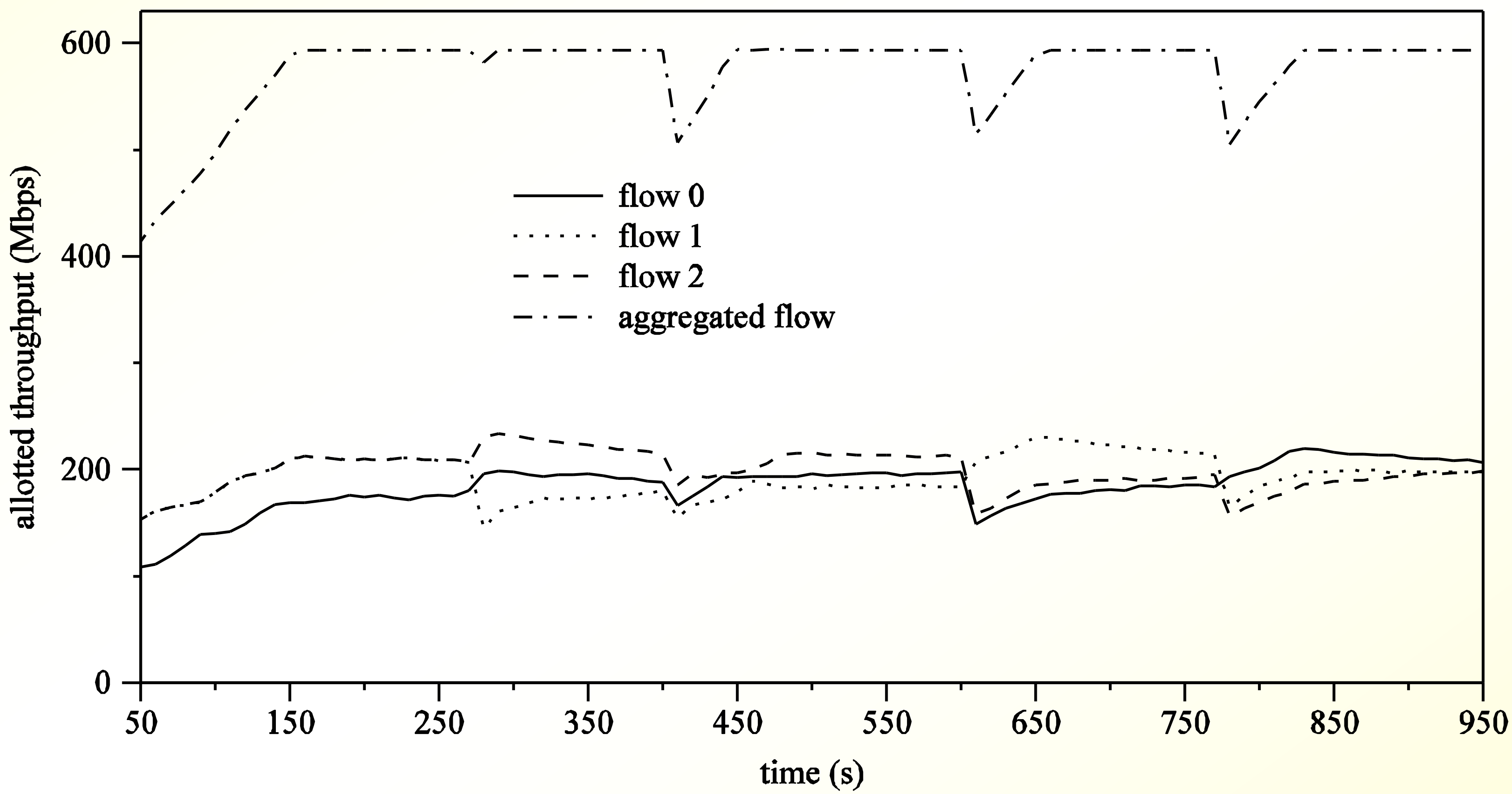


Figure 6. Allotted Throughput with Reno + Gamma
(3 competing flows over a 622 Mbps bottleneck link)

Network Diagnosing

- Network users don’t have to be network experts
- Network Diagnosing is a time-consuming art even for administrators
- Inferring Patterns of Bandwidth Consumption based on Data Mining
  - Netflow, tcpdump, PMA, SNMP, …….
  - The size of monitoring data is large and inaccurate
  - patterns of bandwidth consumption interesting to operators