Design Issues in Traffic Management for the ATM UBR+ Service for TCP over Satellite Networks

Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

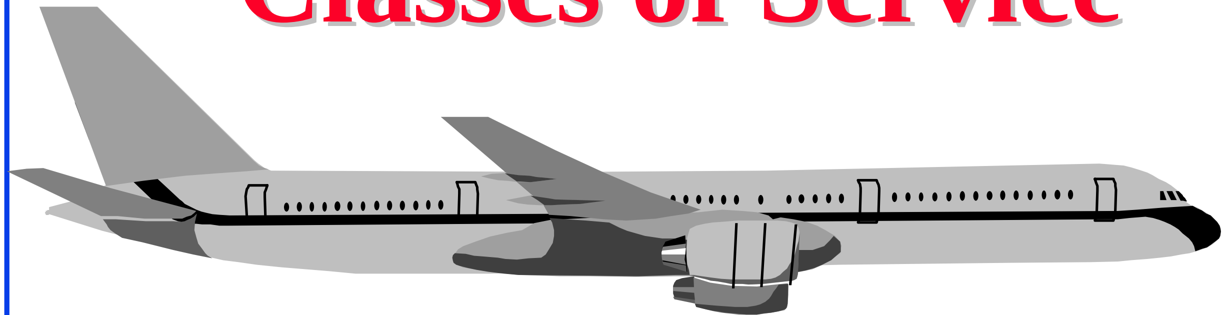
Overview

- Introduction: ABR and TCP Mechanisms
- Statement of Work: TCP over UBR Issues to Study
- Results Todate

Why ATM?

- ATM vs IP: Key Distinctions
  - Traffic Management:
    Explicit Rate vs Loss based
  - Signaling: Coming to IP in the form of RSVP
  - PNNI: QoS based routing
  - Switching: Coming soon to IP
  - Cells: Fixed size or small size is not important

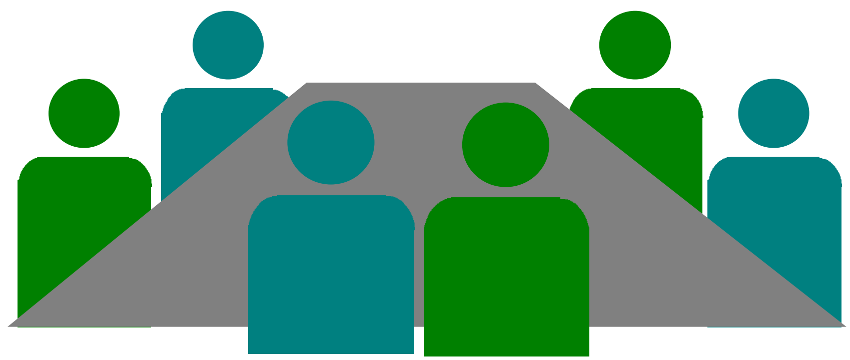
Classes of Service



Standby

Guaranteed

Joy Riders

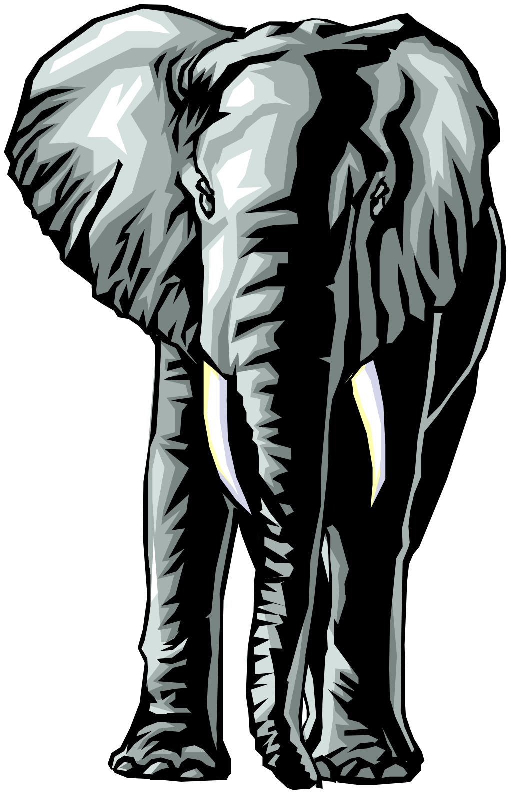


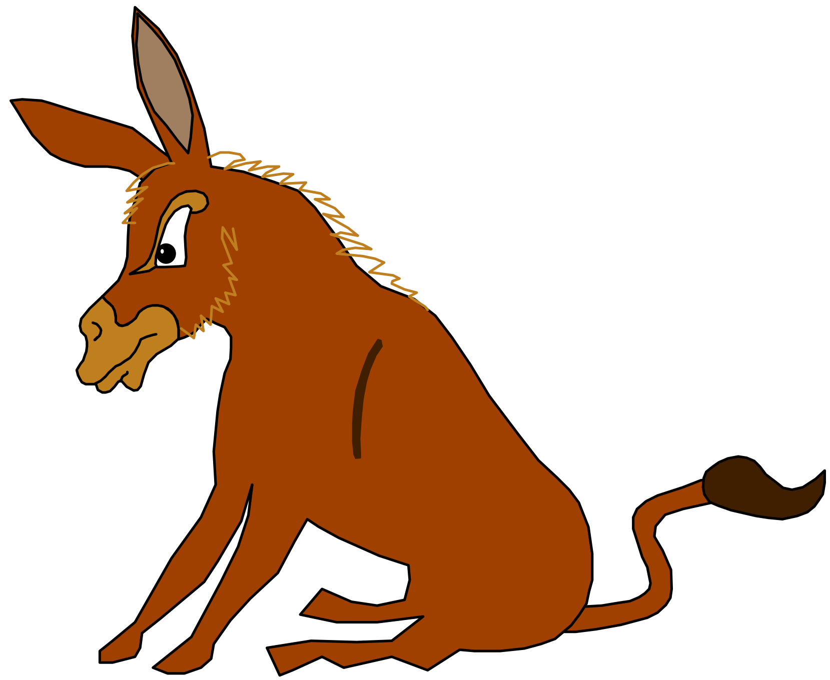
Confirmed

Classes of Service

- ABR (Available bit rate):
  Source follows network feedback.
  Max throughput with minimum loss.
- UBR (Unspecified bit rate):
  User sends whenever it wants. No feedback. No guarantee. Cells may be dropped during congestion.
- CBR (Constant bit rate): User declares required rate. Throughput, delay and delay variation guaranteed.
- VBR (Variable bit rate): Declare avg and max rate.
  - rt-VBR (Real-time): Conferencing.
    Max delay guaranteed.
  - nrt-VBR (non-real time): Stored video.

ABR or UBR?





- Intelligent transport or not?

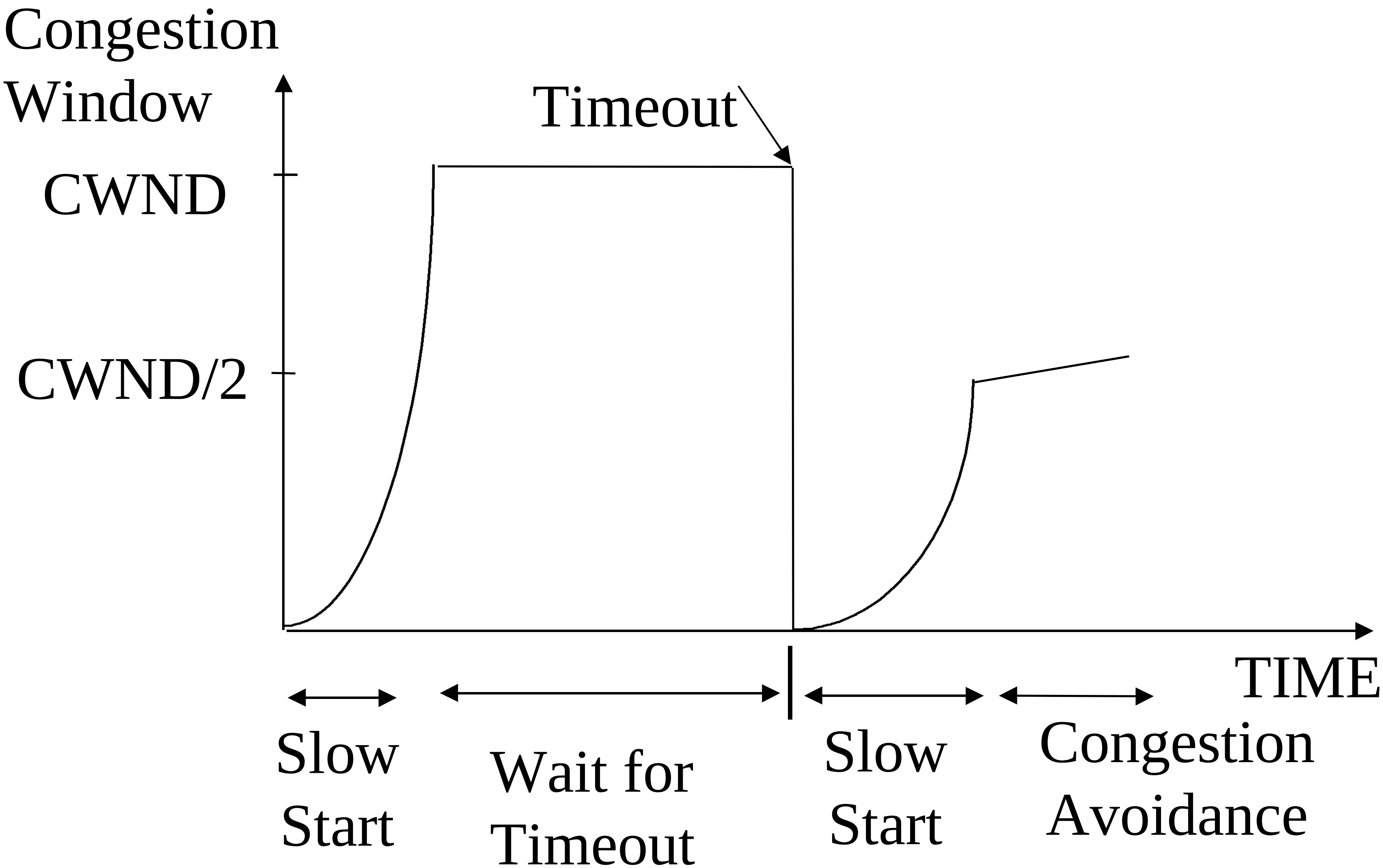
Why UBR?

- Cheapest service category for the user
- Basic UBR is very cheap to implement
- Simple enhancements can vastly improve performance
- Expected to carry the bulk of the best effort TCP/IP traffic.

TCP Congestion Mechanisms

- Slow Start
- Fast retransmit and recovery
- New Reno
- Selective Acknowledgement

Slow Start


Congestion
Window
CWND
CWND/2
Timeout
TIME
Slow
Start
Wait for
Timeout
Slow
Start
Congestion
Avoidance

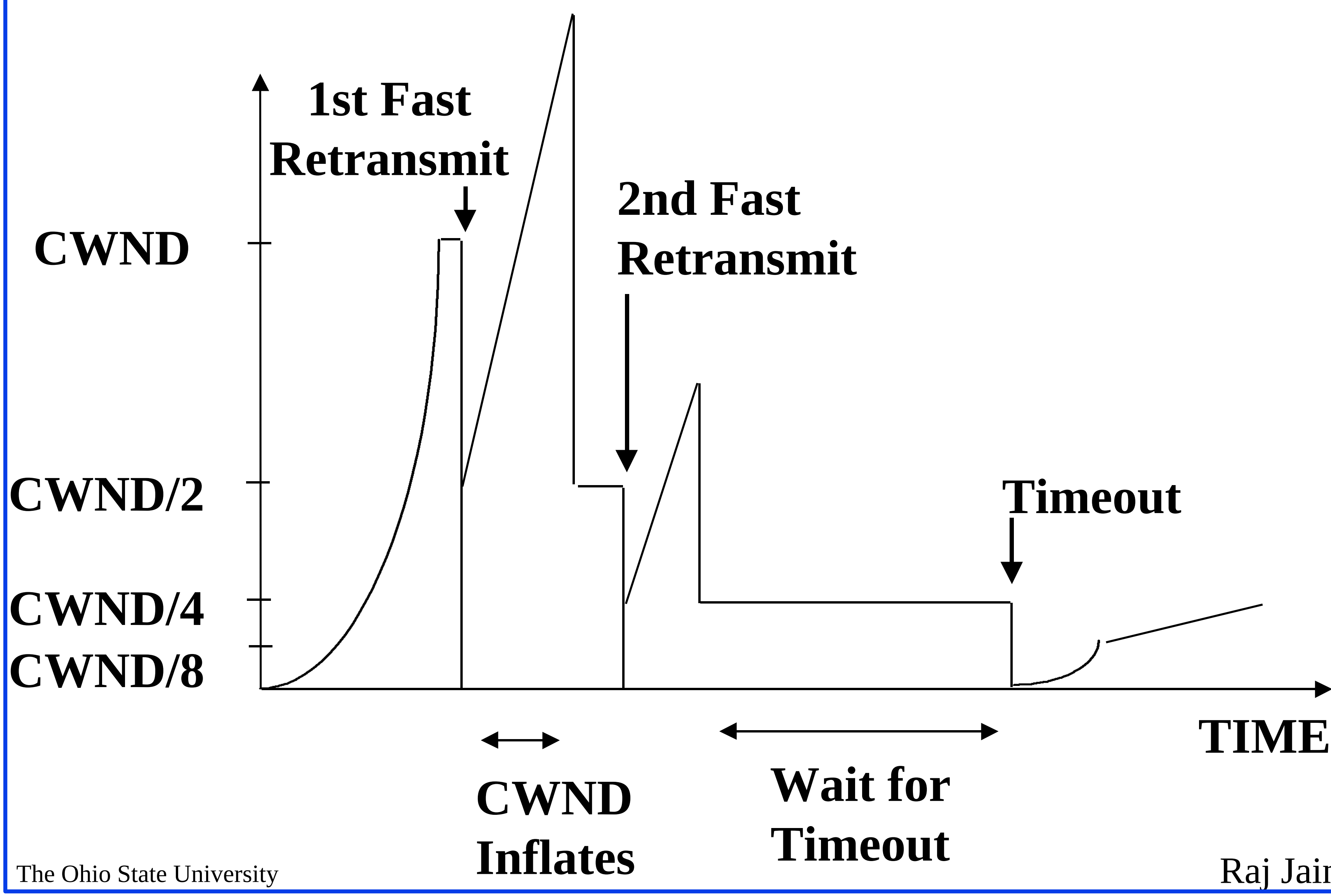
Slow Start (Cont)

- Congestion Window (CWND) and Receiver Window
- Slow Start Threshold
  SSThresh = 0.5 × Congestion Window
- Exponential increase (Slow Start)
- Linear increase (Congestion Avoidance)
- Horizontal line = Timer granularity of 100 to 500 ms

FRR

- Ideas:
  - Don't have to wait for timeout on a loss
  - Don't reduce on single loss due to error
  - Duplicate acks ⇒ Loss
- On three duplicate acks
  - Reduce CWND to 0.5 × CWND + 3 (instead of 1)
  - Set SSThresh to 0.5 × CWND ⇒ Linear increase
- For each subsequent duplicate ack, inflate CWND by 1 and send a packet if permitted
- Problem: Cannot recover from bursty (3+) losses

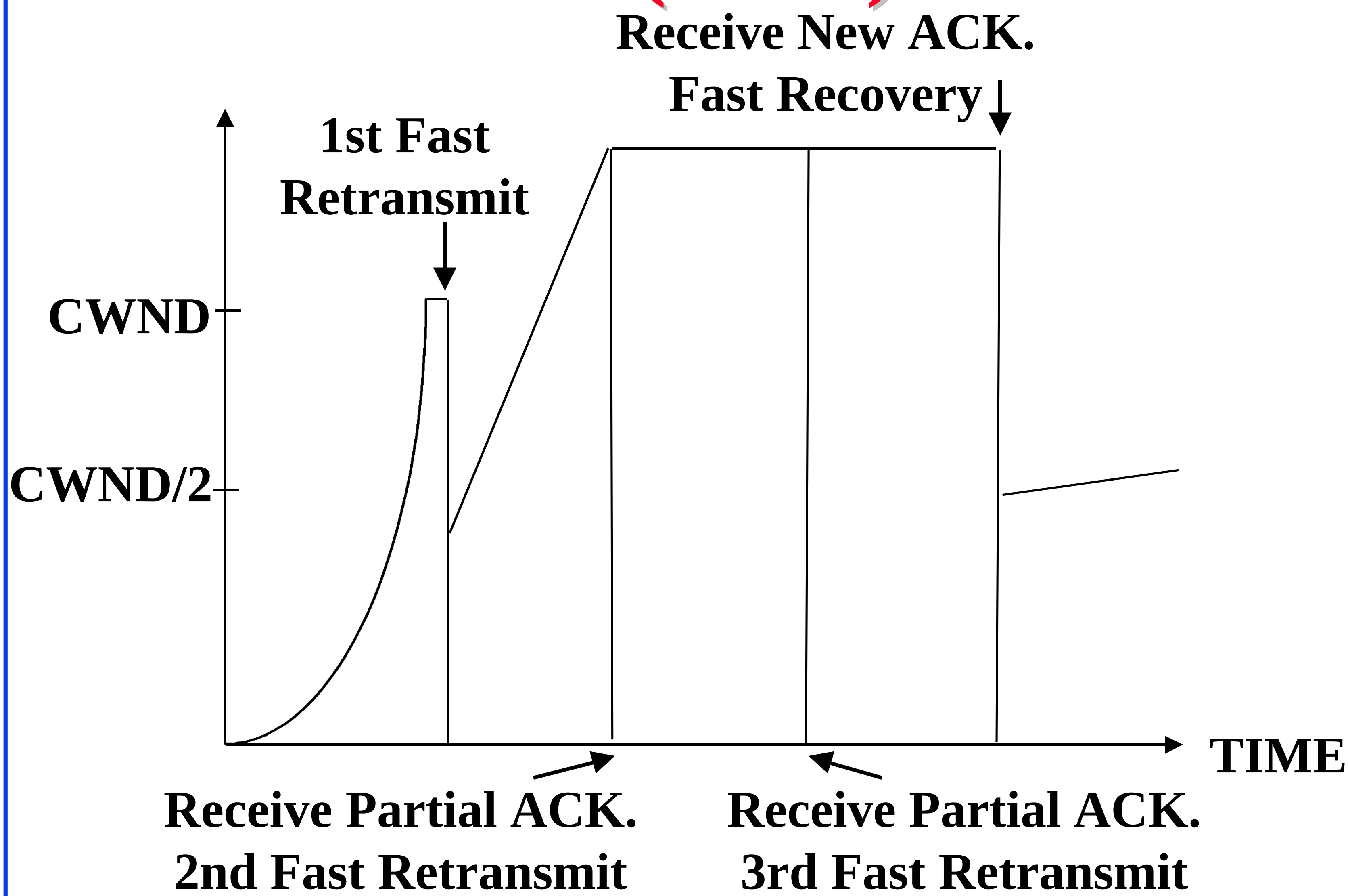
FRR (Cont)


1st Fast Retransmit
2nd Fast Retransmit
CWND
CWND/2
CWND/4
CWND/8
Timeout
TIME
CWND Inflates
Wait for Timeout

New Reno

- Janey Hoe's MS Thesis from MIT
  Published in SIGCOMM'96
- Solution: Determine the end-of a burst loss
  Remember the highest segment sent (RECOVER)
  Ack < RECOVER ⇒ Partial Ack
  Ack ≥ RECOVER ⇒ New Ack
- New Ack ⇒ Linear increase from 0.5 × CWND
- Partial Ack ⇒ Retransmit next packet,
  let window inflate
- Recovers from N losses in N round trips

New Reno (Cont)



Selective Ack

- RFC 2018, October 1996
- Receivers can indicate missing segments
- Example:
  Using Bytes: Ack 500, SACK 1000-1500, 2000-2500
  ⇒ Rcvd segment 1, lost 2, rcvd 3, lost 4, rcvd 5
- On a timeout, ignore all SACK info
- SACK negotiated at connection setup
- Used on all duplicate acks

0-499	500-999	1000-1499	1500-1999	2000-2499
	Lost		Lost	

Goals: Issues

1. Analyze Standard Switch and End-system Policies
2. Design Switch Drop Policies
3. Quantify Buffer Requirements in Switches
4. UBR with VBR Background
5. Performance of Bursty Sources
6. Changes to TCP Congestion Control
7. Optimizing the Performance of SACK TCP

Non-Goals

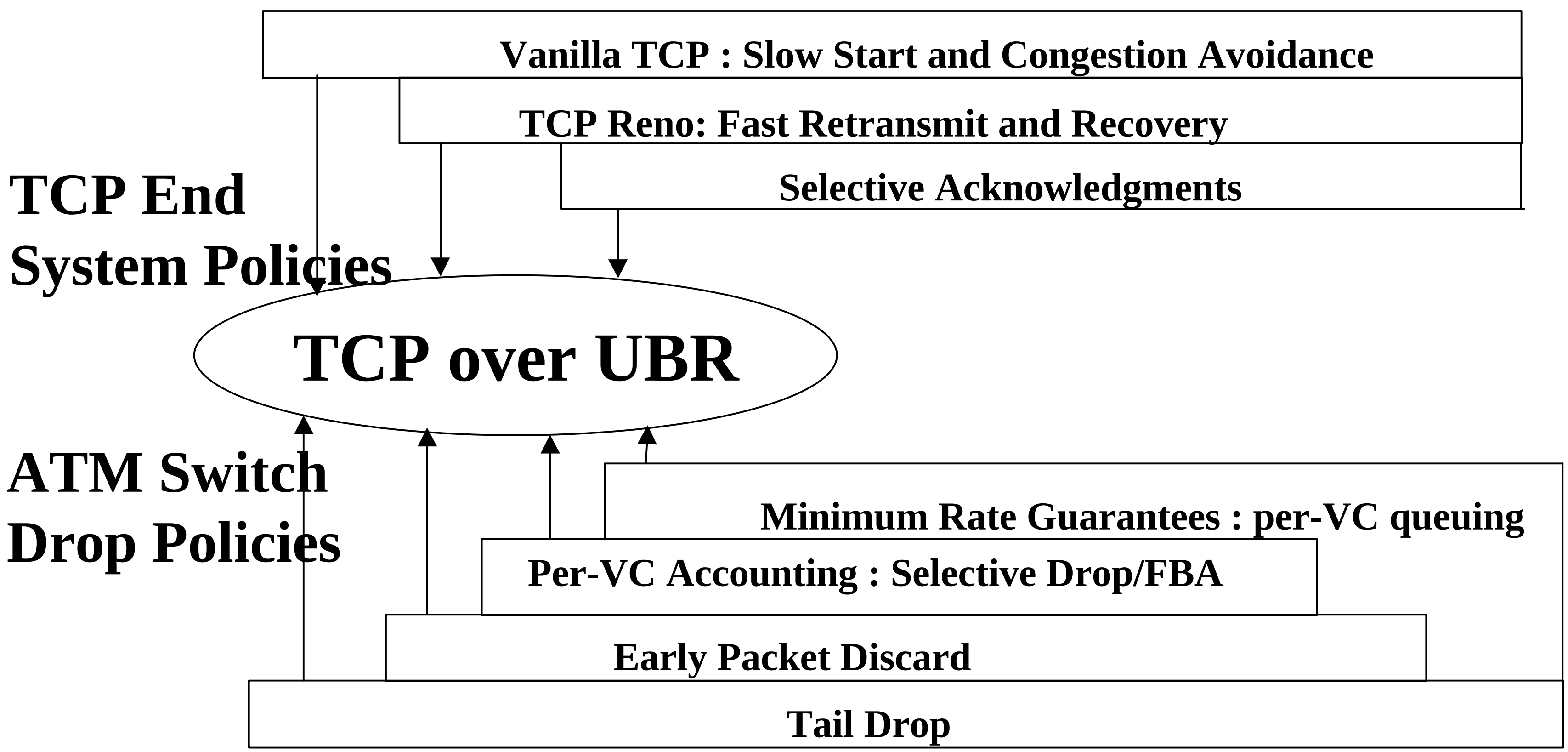
- Does not cover non-UBR issues.
- Does not cover ABR issues.
- Does not include non-TM issues.

1. Policies

End-System Policies

Switch Policies		No FRR	FRR	New Reno	SACK + New Reno
No EPD					
	Plain EPD				
EPD	Selective Drop				
	Fair Buffer Allocation				

Policies


Vanilla TCP : Slow Start and Congestion Avoidance
TCP Reno: Fast Retransmit and Recovery
Selective Acknowledgments
TCP End System Policies
TCP over UBR
ATM Switch Drop Policies
Minimum Rate Guarantees : per-VC queuing
Per-VC Accounting : Selective Drop/FBA
Early Packet Discard
Tail Drop

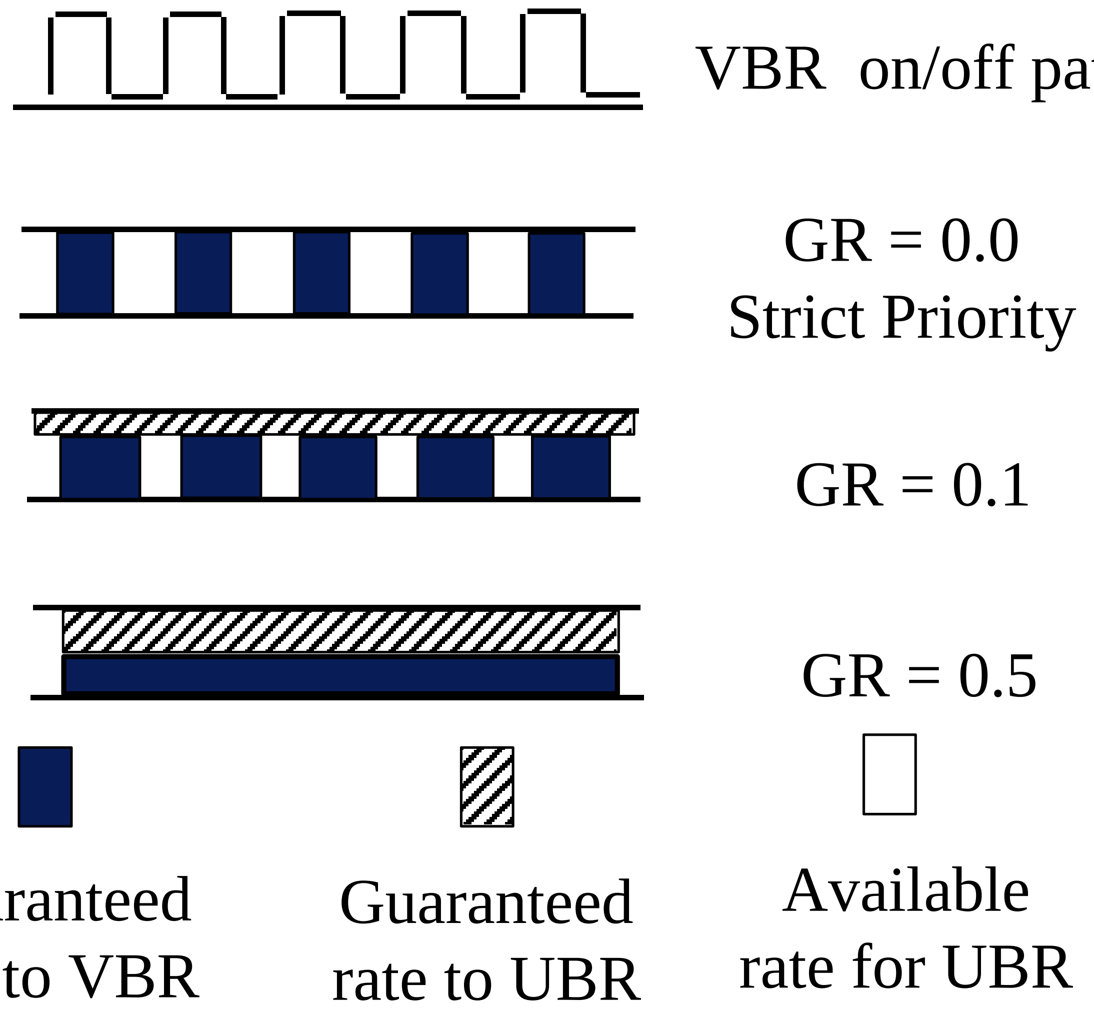
2. Switch Drop Policies

- Selective Drop
- Fair buffer allocation

3. Buffer Requirements

- Assess buffer requirements for TCP over UBR for satellite latencies
- How well can we do with less than 1 RTT buffers?
- How is TCP throughput affected by:
  - Delay-bandwidth product
  - Buffer Size
  - Switch Drop Policies
  - End-System Policies
  - Number of Sources

4. UBR with VBR Background



4a. Guaranteed Frame Rate (GFR)

- UBR with minimum cell rate (MCR)
  $\Rightarrow$ UBR+
- Frame based service
  - Complete frames are accepted or discarded in the switch
  - Traffic shaping is frame based.
    All cells of the frame have CLP =0 or CLP =1
  - All frames below MCR are given CLP =0 service.
    All frames above MCR are given best effort (CLP =1) service.

4b. Guaranteed Rate Service

- Guaranteed Rate (GR): Reserve a small fraction of bandwidth for UBR class.

GR	GFR
per-class reservation	per-VC reservation
per-class scheduling	per-VC accounting/scheduling
No new signaling	Need new signaling
Can be done now	In TM4+

5. Bursty Sources

- Large number of sources
- SPECweb'96 benchmark
- Past Results: ABR is stable.*
- Need to do a similar study for UBR over Satellites.

- ***Ref**: Performance of Bursty World Wide Web (WWW) Sources over ABR, ATM Forum 97-0425, April 1997

6. Problem in TCP Implementations

- Linear Increase in Segments:
  CWND/MSS = CWND/MSS + MSS/CWND
- In Bytes: CWND = CWND + MSS*MSS/CWND
- All computations are done in integer
- If CWND is large, MSS*MSS/CWND is zero and CWND does not change. CWND stays at 512*512 or 256 kB.

Solutions

- **Solution 1**: Increment CWND after N acks (N > 1)
  CWND = CWND + N*MSS*MSS/CWND
- **Solution 2**: Use larger MSS on Satellite links such that MSS*MSS > CWND. MSS ≥ Path MTU.
- **Solution 3**: Use floating point
- **Recommendation**: Use solution 1. It works for all MSSs.
- **To do**: Does this change TCP dynamics and adversely affect performance.

7. Optimize SACK TCP

- SACK helps only if retransmitted packets are not lost.
- Currently TCP retransmits immediately after 3 duplicate acks (Fast retransmit), and then waits RTT/2 for congestion to subside.
- Network may still be congested ⇒ Retransmitted packets lost.
- Proposed Solution: Delay retransmit by RTT/2, I.e., wait RTT/2 first, and then retransmit.

Results

1. Analyze Standard Switch and End-system Policies
2. Design Switch Drop Policies
3. Quantify Buffer Requirements in Switches
4. UBR with VBR Background
   4a. **Guaranteed Frame Rate**
   4b. Guaranteed Rate
5. Performance of Bursty Sources
6. Changes to TCP Congestion Control
7. Optimizing the Performance of SACK TCP

1. Policies: Results

- In LANs, switch improvements (PPD, EPD, SD, FBA) have more impact than end-system improvements (Slow start, FRR, New Reno, SACK). Different variations of increase/decrease have little impact due to small window sizes.
- In satellite networks, end-system improvements have more impact than switch-based improvements
- FRR hurts in satellite networks.
- Fairness depends upon the switch drop policies and not on end-system policies

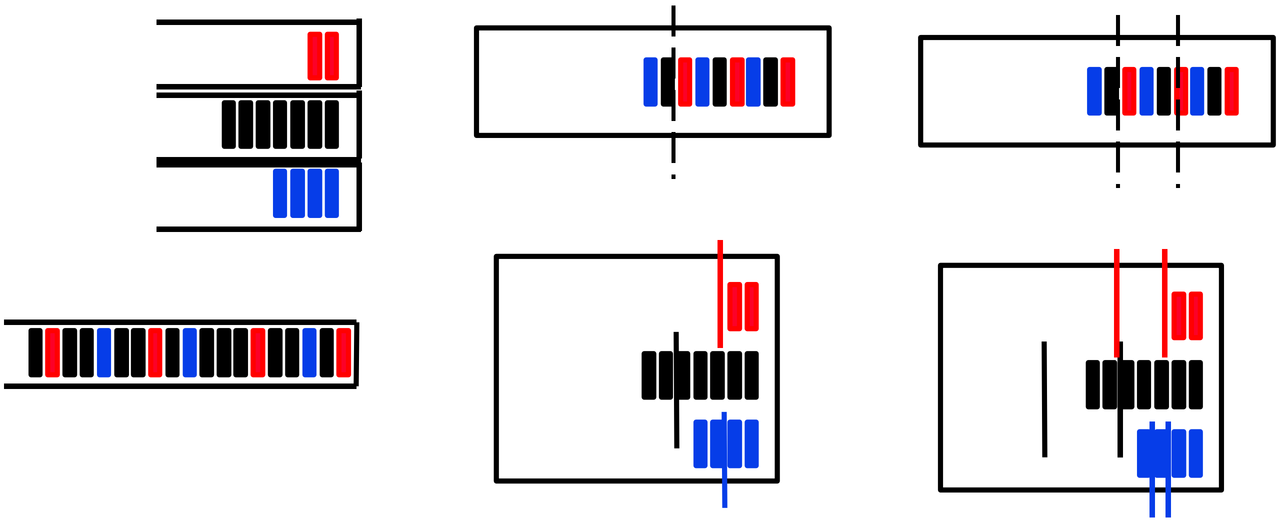
Policies (Continued)

- In Satellite networks:
  - SACK helps significantly
  - Switch-based improvements have relatively less impact than end-system improvements
  - Fairness is not affected by SACK
- In LANs:
  - Previously retransmitted holes may have to be retransmitted on a timeout
    $\Rightarrow$ SACK can hurt under extreme congestion.

3. Buffer Requirements: Results

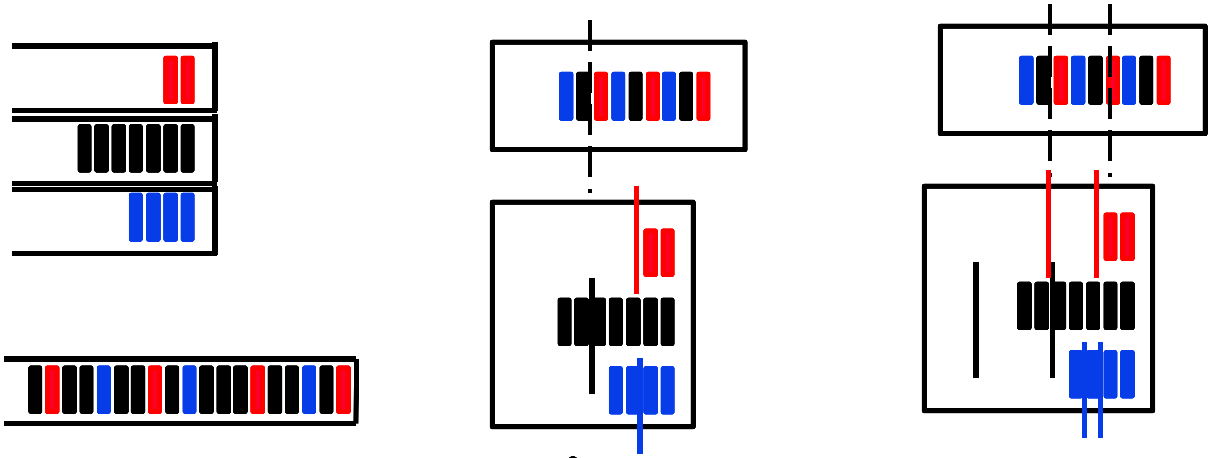
- Very small buffer sizes result in low efficiency
- Moderate buffer sizes (less than 1 RTT)
  - Efficiency increases with increase in buffer size
  - Efficiency asymptotically approaches 100%
- 0.5*RTT buffers provide sufficiently high efficiency (98% or higher) for SACK TCP over UBR even for a large number of TCP sources

4a. GFR Options



Queuing	Per-VC	FIFO
Buffer Management	Per-VC Thresholds	Global Threshold
Tag-sensitive Buffer Mgmt	2 Thresholds	1 Threshold

GFR: Results

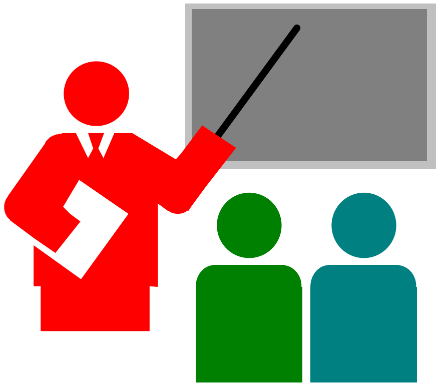


Per-VC Q Fair Excess CLP0

- Per-VC queuing and scheduling is necessary for per-VC MCR. (FIFO ok for TCP w SACK at low loads)
- FBA and proper scheduling is necessary for fair allocation of excess bandwidth
- One global threshold is sufficient for CLP0+1 guarantees Two thresholds are necessary for CLP0 guarantees

4b. Guaranteed Rate: Results

- Guaranteed rate is helpful in WANs.
- For WANs, the effect of reserving 10% bandwidth for UBR is more than that obtained by EPD, SD, or FBA
- For LANs, guaranteed rate is not so helpful. Drop policies are more important.
- For Satellites, end-system policies seem more important.



Summary

- UBR is preferred for TCP/IP over ATM
- TCP policies over UBR may/may not be same as without ATM
- Very comprehensive study of TCP/IP over UBR: existing mechanisms, new mechanisms, parameter selection
- Includes TCP mechanisms, end systems, switches, buffers, traffic patterns, and UBR enhancements.
- Plan to influence the industry

Summary (Cont)

- For satellite networks, end-system policies (SACK) have more impact than switch policies (EPD).
- 0.5*RTT buffers provide sufficiently high efficiency (98% or higher) for SACK TCP over UBR even for a large number of TCP sources
- Reserving a small fraction for UBR helps it a lot in satellite networks

Our Contributions and Papers

- All our contributions and papers are available on-line at http://www.cis.ohio-state.edu/~jain/
- See Recent Hot Papers for tutorials.

Thank You!