Traffic Management on Satellite ATM Networks

Raj Jain

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

Overview

- Trends

- Why Traffic Management?

- Why ATM?

- ATM Traffic Management

- ABR Vs UBR

- 4 Ways to improve ABR over Satellites

- 4 Ways to improve UBR over Satellites

Trends

❑ Communication is more critical than computing

  ❍ Greeting cards contain more computing power than all computers before 1950.

  ❍ Genesis's game has more processing than 1976 Cray supercomputer.

❑ URL is more important than a company's phone number. (54 URLs in first 20 pages of March'97 Good Housekeeping.)

❑ Email is faster than telegrams

Impact on R&D

- Too much growth in one year
  $\Rightarrow$ Can't plan too much into long term
- Long term = 1_2 year or 10_2 years at most
- Products have life span of 1 year, 1 month, …
- Short product development cycles.
  Chrysler reduced new car design time
  from 6 years to 2.
- Distance between research and products has narrowed
  $\Rightarrow$ Collaboration between researchers and developers
  $\Rightarrow$ Academics need to participate in industry consortia

Garden Path to I-Way

❑ Plain Old Telephone System (POTS) = 64 kbps = 3 ft garden path

❑ ISDN = 128 kbps = 6 ft sidewalk

❑ T1 Links to Businesses = 1.544 Mbps = 72 ft = 4 Lane roadway

❑ Cable Modem Service to Homes: = 10 Mbps = 470 ft = 26 Lane Driveway

❑ OC3 = 155 Mbps = 1 Mile wide superhighway

❑ OC48 = 2.4 Gbps = 16 Mile wide superhighway
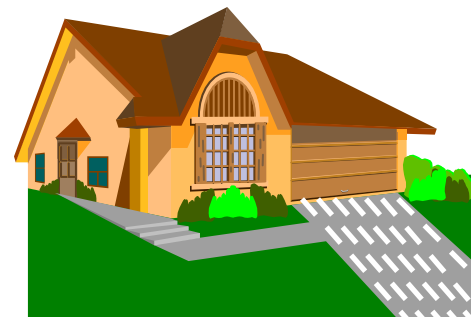
Dime Sale
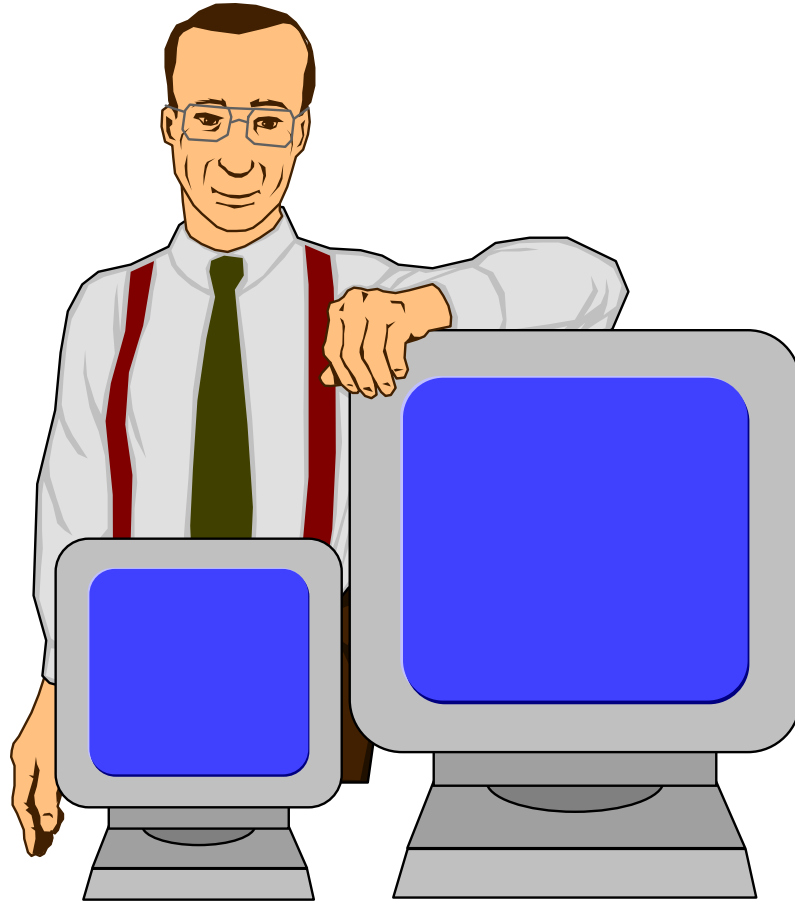
One Megabit memory, One Megabyte disk, One Mbps link, One MIP processor, 10 cents each.....
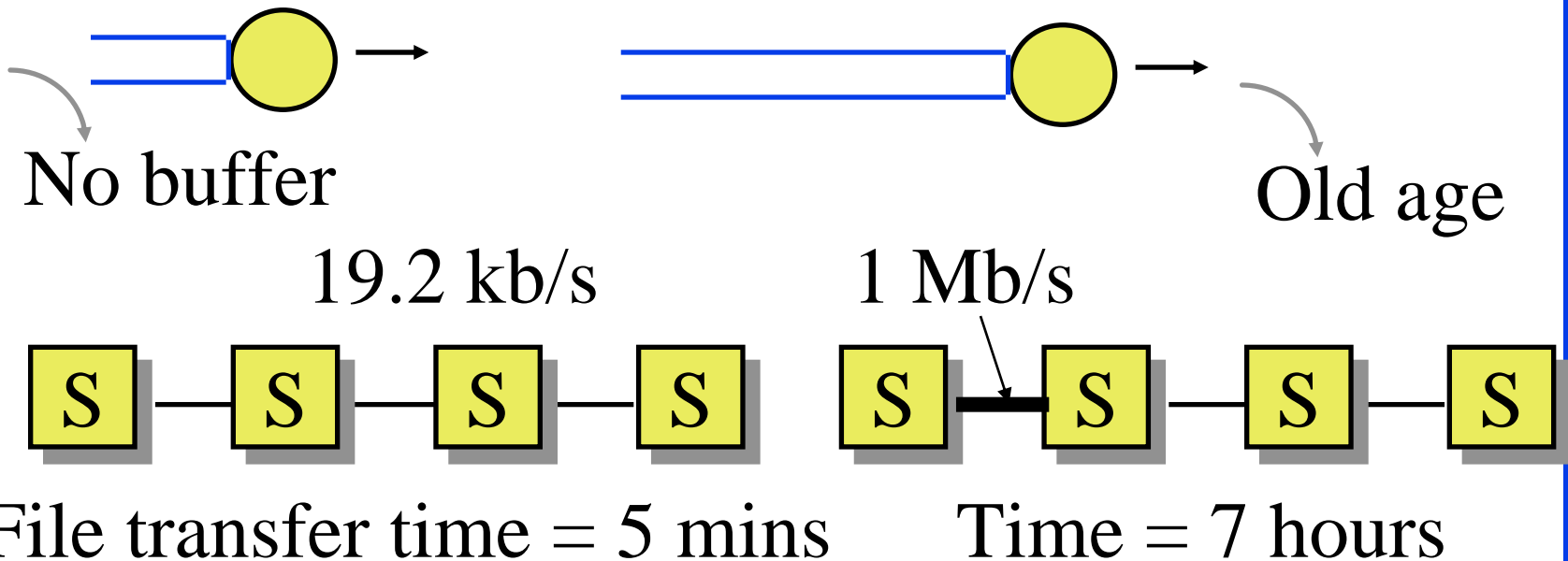
Future

Year

1980

In 1990, the memory will be so cheap that you will not have to worry about paging, swapping, virtual memory, memory hierarchy, and....
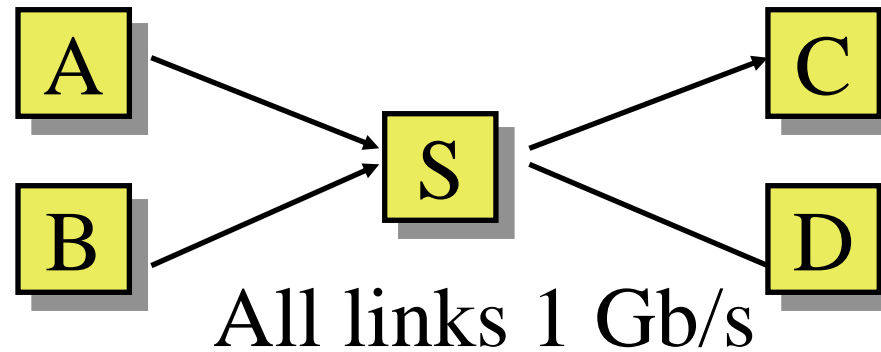
Why Worry About Congestion?

Q: Will the congestion problem be solved when:

❑ Memory becomes cheap (infinite memory)?

❑ Links become cheap (very high speed links)?

❑ Processors become cheap?

A: None of the above.

No buffer

Old age

19.2 kb/s   1 Mb/s

S — S — S — S   S — S — S — S

File transfer time = 5 mins   Time = 7 hours

All links 1 Gb/s

Conclusions:

❑ Congestion is a dynamic problem.
Static solutions are not sufficient

❑ Bandwidth explosion
$\Rightarrow$ More unbalanced networks

❑ Buffer shortage is a symptom not the cause.

ATM

❑ ATM Net = Data Net + Phone Net

❑ Combination of Internet method of communication (packet switching) and phone companies' method (circuit switching)

ATM
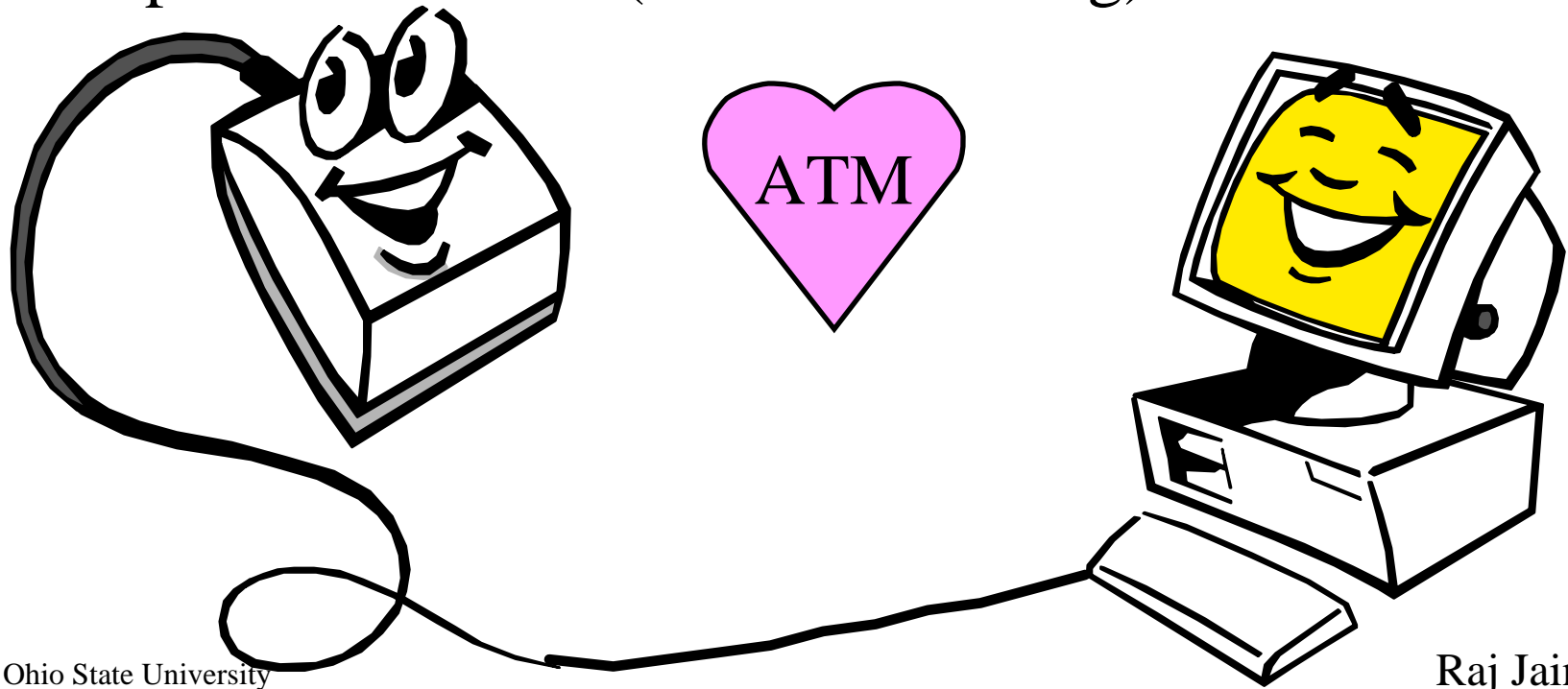
Why ATM?

❑ 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

Old House vs New House
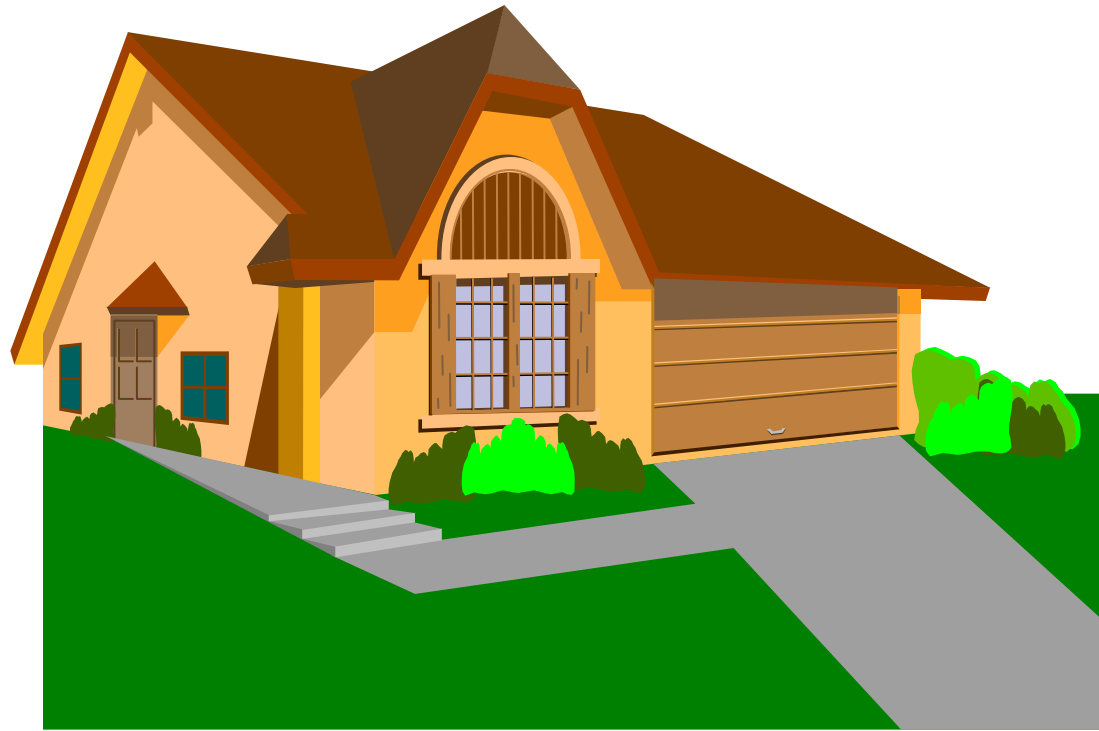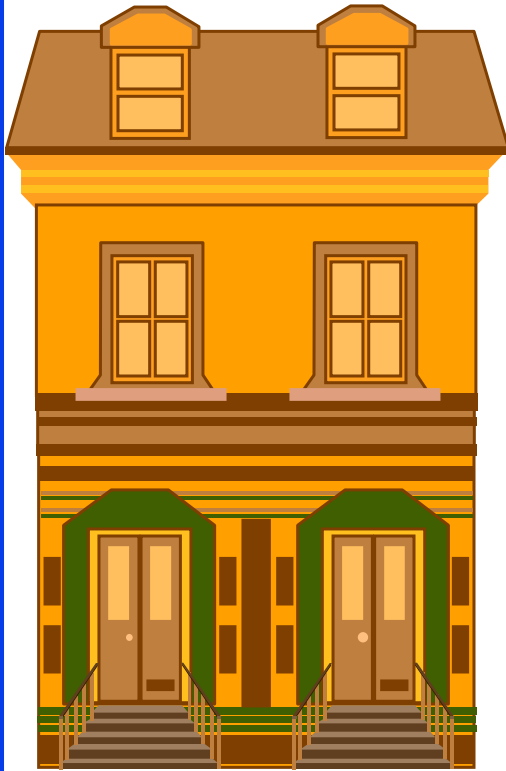
❑ New needs:
Solution 1: Fix the old house (cheaper initially)
Solution 2: Buy a new house (pays off over a long run)

Service Categories

Standby

Guaranteed

Joy Riders

Confirmed

Raj Jain

Service Categories

❑ 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.

Our Goal

❏ Ensure that the new ATM Forum
TM 4.0/5.0 specs are "Satellite-friendly"

❏ There are no parameters or requirement that will
perform badly in a long-delay satellite environment

❏ Users can use paths going through satellite links
without requiring special equipment

❏ Develop optimal solutions for satellite networks

This work is sponsored by
NASA Lewis Research Center.
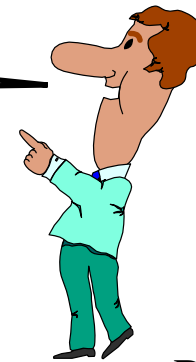
Binary Rate Scheme
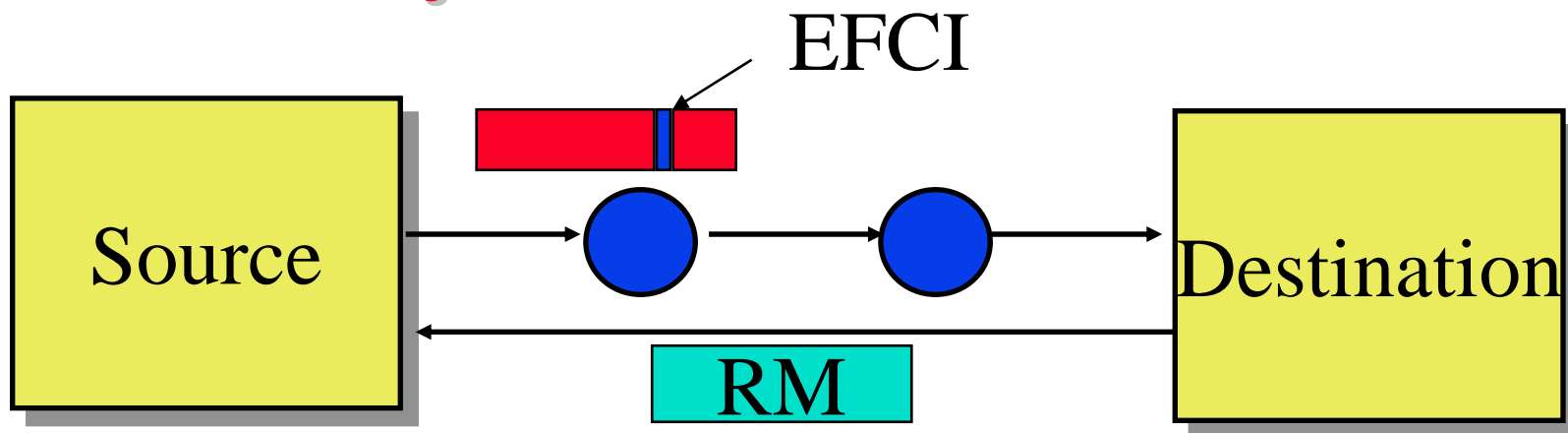
EFCI

- ❑ DECbit scheme in many standards since 1986.
- ❑ Forward explicit congestion notification (FECN) in Frame relay
- ❑ Explicit forward congestion indicator (EFCI) set to 0 at source. Congested switches set EFCI to 1
- ❑ Every nth cell, destination sends an resource management (RM) cell to the source

The Explicit Rate ABR

| Current Cell Rate | Explicit Rate |
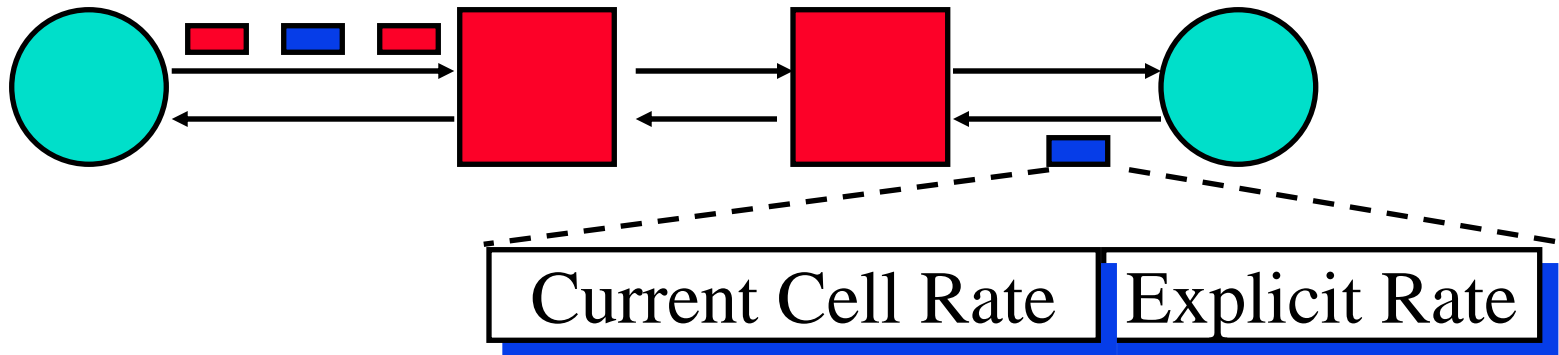
- ❏ Proposed in July 1994
- ❏ Sources send one RM cell every n cells
- ❏ The RM cells contain "Explicit rate"
- ❏ Destination returns the RM cell to the source
- ❏ The switches adjust the rate down
- ❏ Source adjusts to the specified rate
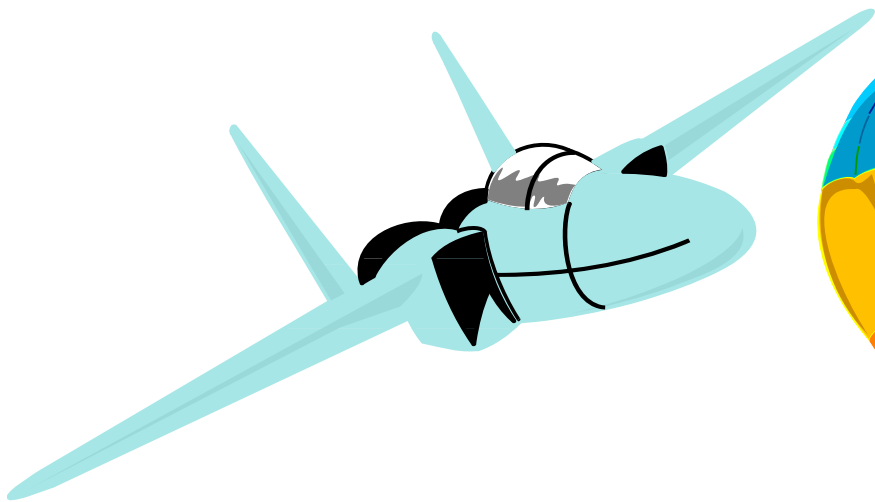
Why Explicit Rate Indication?

❑ Longer-distance networks
$\Rightarrow$ Can't afford too many round-trips
$\Rightarrow$ More information is better

❑ Rate-based control
$\Rightarrow$ Queue length = ΔRate $\times$ ΔTime
$\Rightarrow$ Time is more critical than with windows

ERICA Switch Algorithm

❑ Each manufacturer has its own explicit rate switch algorithm

❑ Explicit Rate Indication for Congestion Avoidance (ERICA) is the most thoroughly analyzed algorithm

❑ Shown to be efficient, fair, fast transient response, able to handle bursty TCP traffic

❑ ERICA+ allows low delay even at 100% utilization and provides stability in the presence of high frequency VBR background traffic

❑ Being implemented by several vendors. Software implementation feasible.

ERICA Scheme: Basic

- ❑ <u>E</u>xplicit <u>R</u>ate <u>I</u>ndication for <u>C</u>ongestion <u>A</u>voidance
- ❑ Set target rate, say, at 95% of link bandwidth
  ABR Capacity = Target Utilization * Link Bandwidth
- ❑ Monitor input rate and number of active VCs
  Overload = ABR Input rate/ABR Capacity
- ❑ This VC's Share = VC's Current Cell Rate/Overload
- ❑ Fair share = Target rate/ Number of Active VCs
- ❑ ER = **_Max(_**Fair share, This VC's share)
- ❑ ER = Min{ER, ABR Capacity}
- ❑ ER in Cell = Min(ER in Cell, ER)

ERICA+: Full Utilization

- Allows operation at any point between the knee and the cliff

- The queue time can be set to any desired value.

- Allows utilization to be 100%

Throughput

Load

Delay

Load

Link Utilization

Time

Queue Length

50
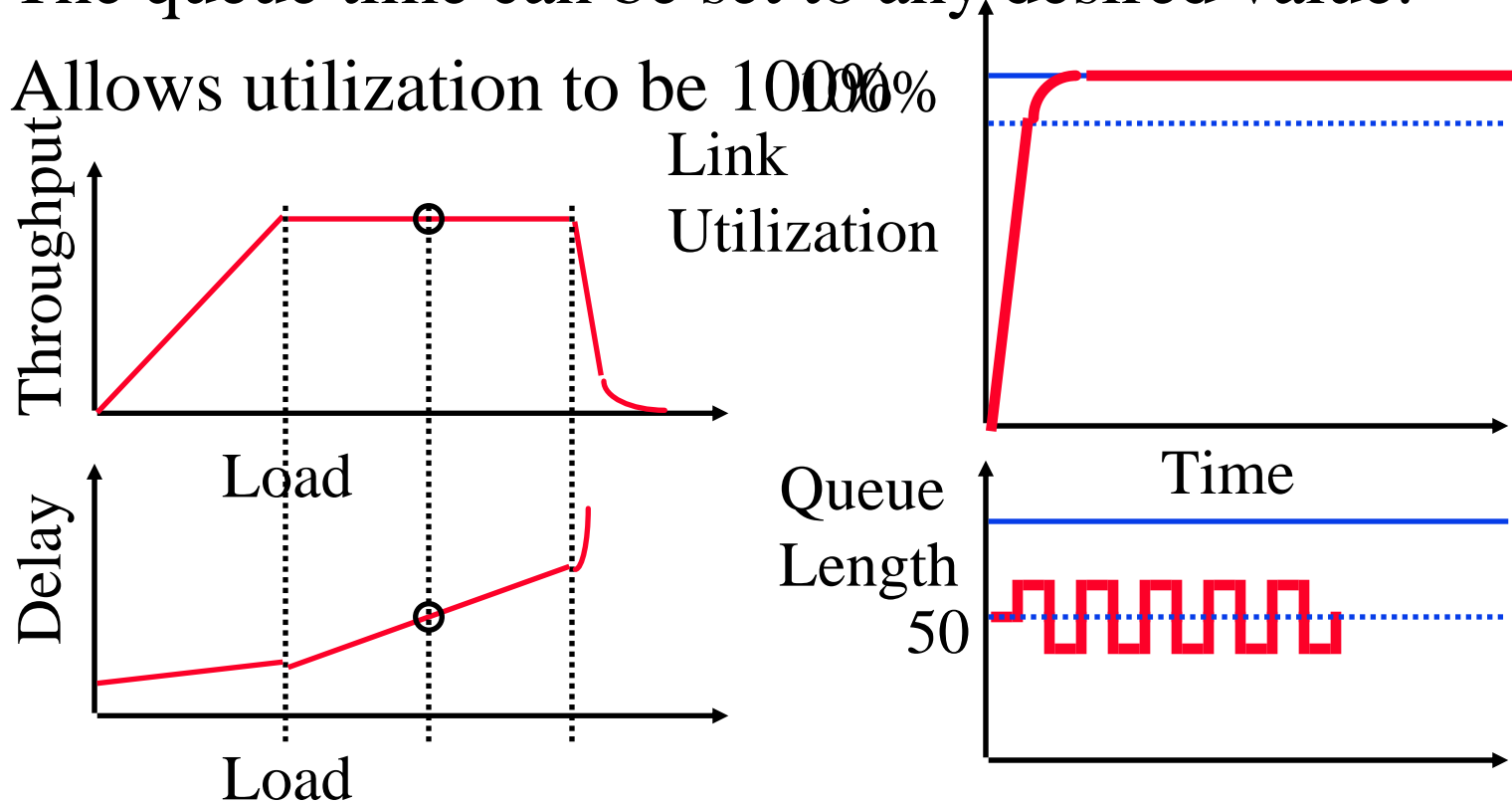
4 Ways to Improve ABR over Satellite

1. Increase the limit on the number of outstanding cells before decreasing $\Rightarrow$ Large TBE

   The size of was increased from 8 bit to 24 bit to accommodate satellite paths.

2. Use larger increase factor
   $\Rightarrow$ RIF=1 $\Rightarrow$ Fast transient Response

3. Implement backward congestion notification (BECN)

4. Implement Virtual Source/Virtual Destination

Effect of TBE

❑ TBE was initially called Crm.

❑ Crm limits the number of cells lost if the link is broken

❑ Source Rule (6):
If you not received feedback from the network after Crm×Nrm cells, reduce your allowed cell rate (ACR):

$$ACR = max\{MCR, ACR - ACR \times CDF\}$$

Effect of CDF

❑ After Crm×Nrm cells:
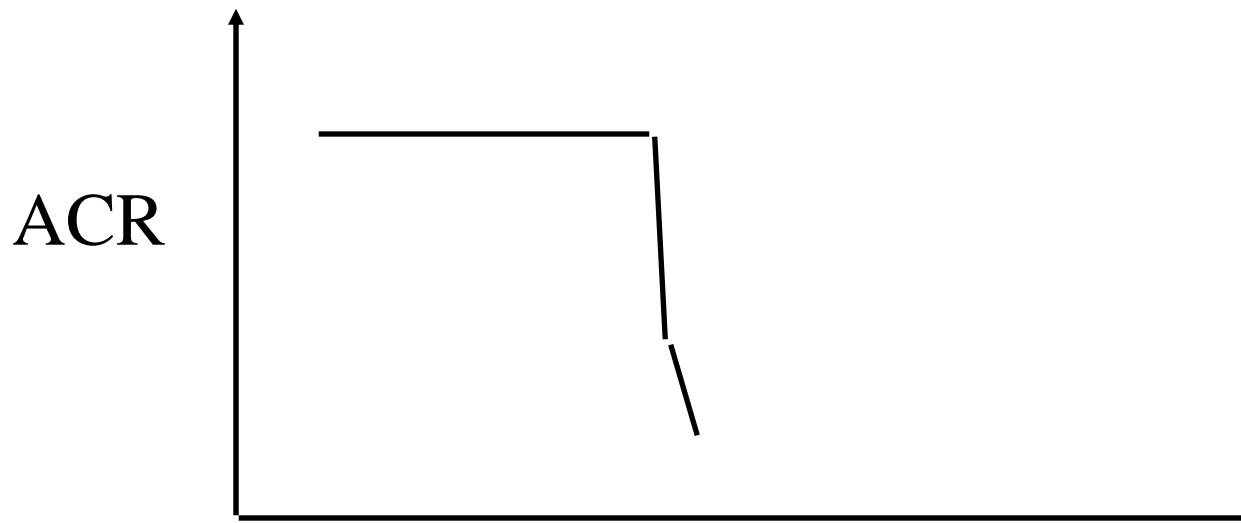
$$ACR = ACR(1 - CDF)$$

❑ After Crm(1 + Nrm) cells:

$$ACR = ACR(1 - CDF)^2$$
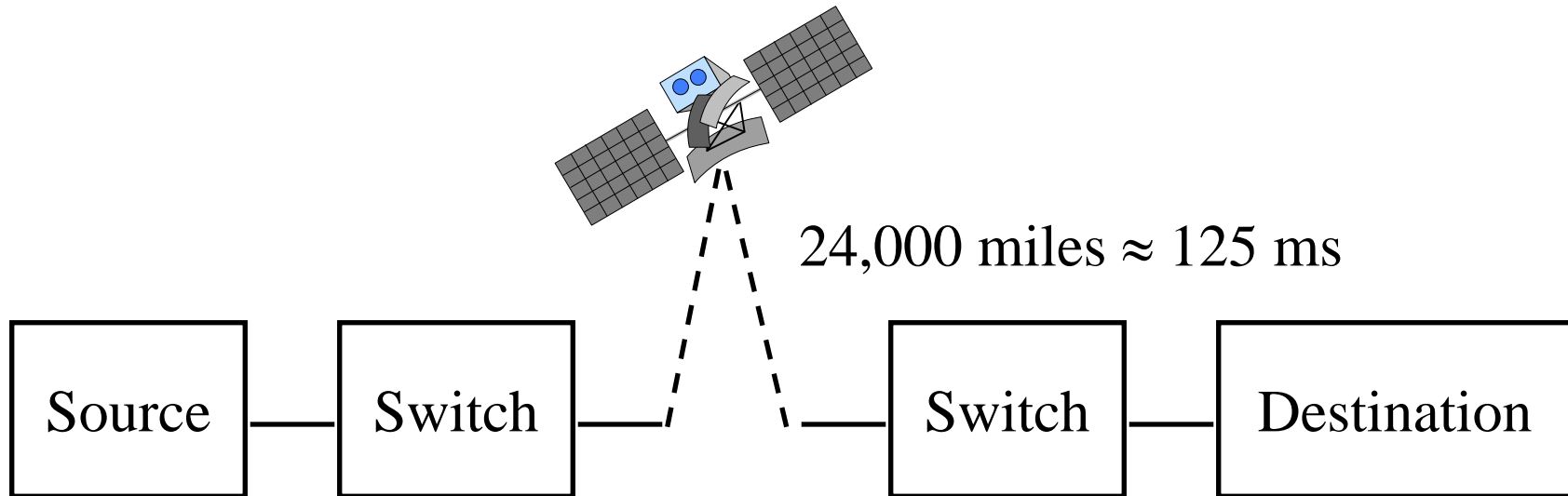
❑ After Crm(k + Nrm) cells:

$$ACR = ACR(1 - CDF)^k$$

Effect of CDF

❏ There is an almost vertical drop after Crm:

ACR

θ The value of CDF has very little effect

θ The source becomes a "Low Rate Source"

θ Initially Crm was a 8-bit quantity with a default of 32

Satellite Links

24,000 miles $\approx$ 125 ms

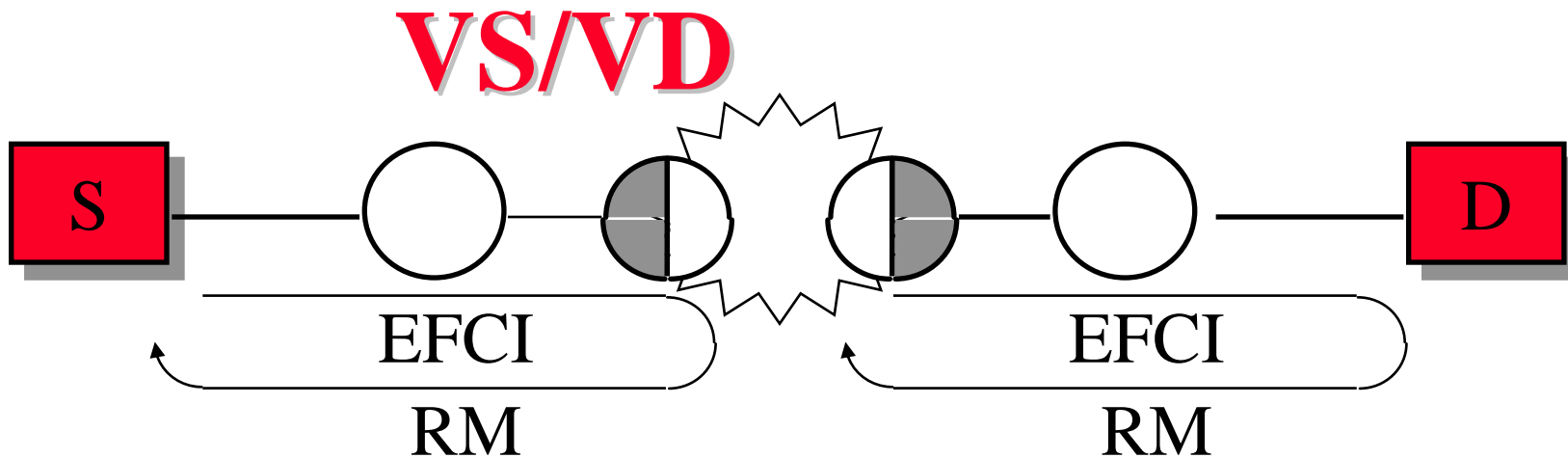| Source	Switch		Switch	Destination

- One-way delay = Up-down = 250 ms
  Round-trip delay = 550 ms (measured)

- Crm = 32
  $\Rightarrow$ Maximum 32×32 = 1024 cells in flight
  before ACR starts coming down

Required Crm

❑ For full throughput

$$\text{Crm} \geq \text{RTTQ}/(\text{Nrm} \times \text{ACR})$$

Where RTTQ = Round Trip Time
including Queueing

❑ For 155 Mbps, $\text{Crm} \geq 6{,}144$

❑ For 622 Mbps, $\text{Crm} \geq 24{,}576$

❑ For two satellite hops: $\text{Crm} \geq 49{,}152$

❑ For n satellite hops: $\text{Crm} \geq 24{,}576n$

$\Rightarrow$ Need 32 bits for Crm

❑ Compromise: Crm is now a 24 bit quantity.

VS/VD
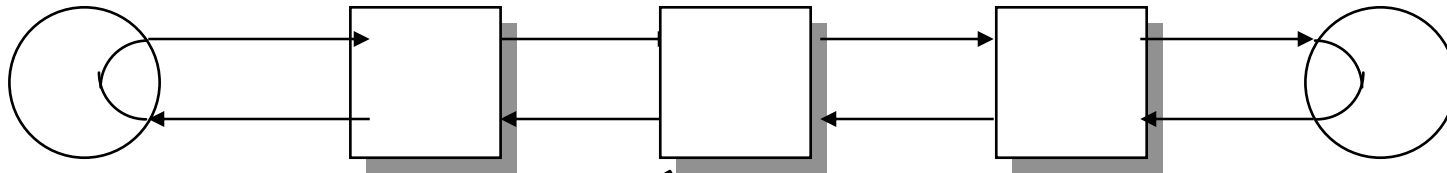

- Virtual source/virtual destinations (VS/VD) follow all notification/control rules
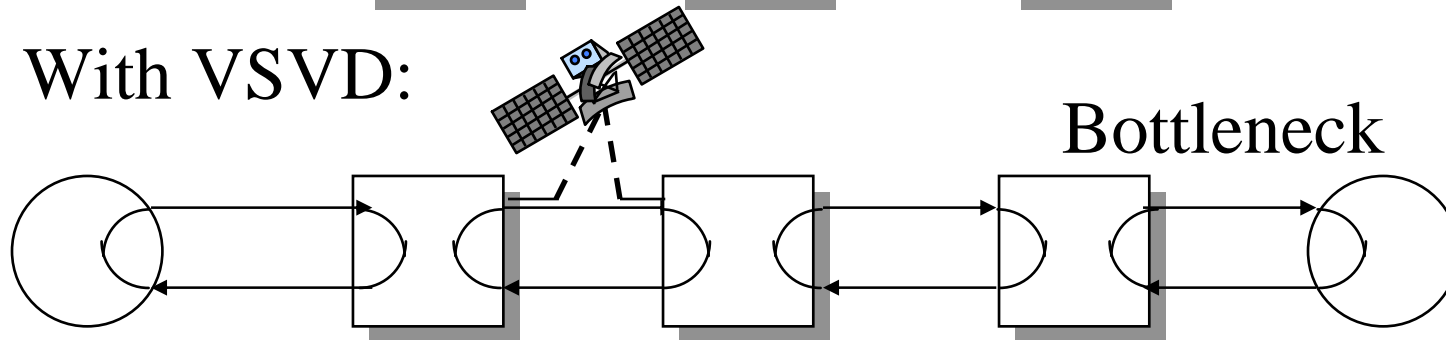
- Can be hop-by-hop
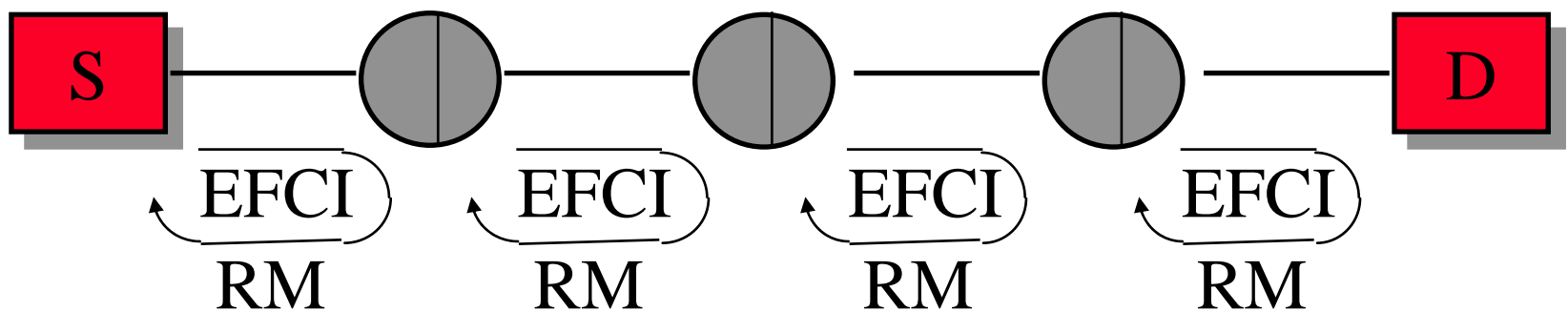
θ Virtual dest/sources maintain per-VC queues.

VS/VD: Results

θ Without VS/VD:

❑ With VSVD:

Bottleneck

Satellite
Link

Workgroup
Switch

θ With VSVD, the buffering is proportional to the delay-bandwidth of the previous loop

⇒ Good for satellite networks

Multicast

RM Cell$_A$

RM Cell$_A$

RM Cell$_B$

RM Cell$_A$

RM Cell$_C$

RM Cell$_{min(B,C)}$
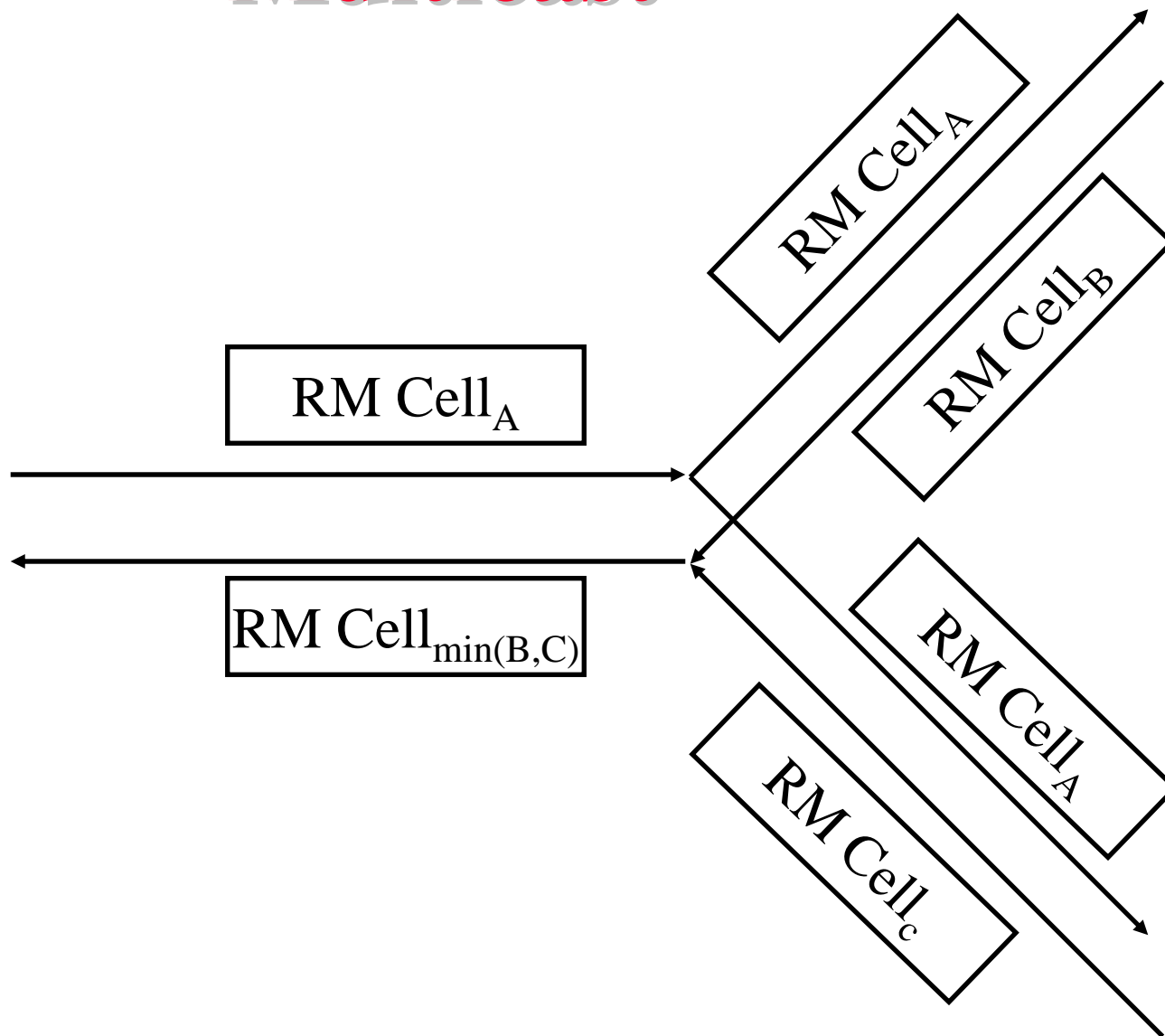
Issues in ABR Multicast TM
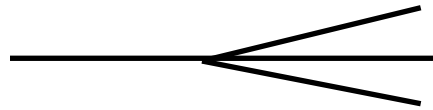
❏ Minimum of ER from branches is sent upstream
Should we wait for all branches?

❏ If you send BRM on every FRM, you may give feedback without receiving any
$\Rightarrow$ Need at least one feedback before sending a BRM. Otherwise, you may give PCR

❏ Not all downstream feedbacks in an upstream feedback $\Rightarrow$ Consolidation noise

❏ Solution: Mulitipoint ERICA

Internet Protocols over ATM

❑ ATM Forum has designed ABR service for data

❑ UBR service provides no feedback or guarantees

❑ Internet Engineering Task Force (IETF) prefers UBR for TCP

ABR or UBR?
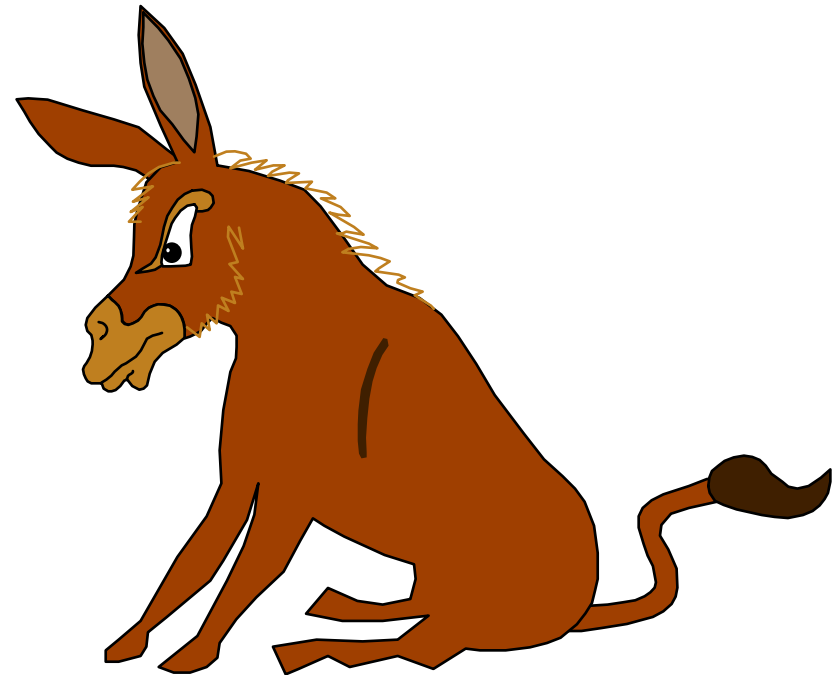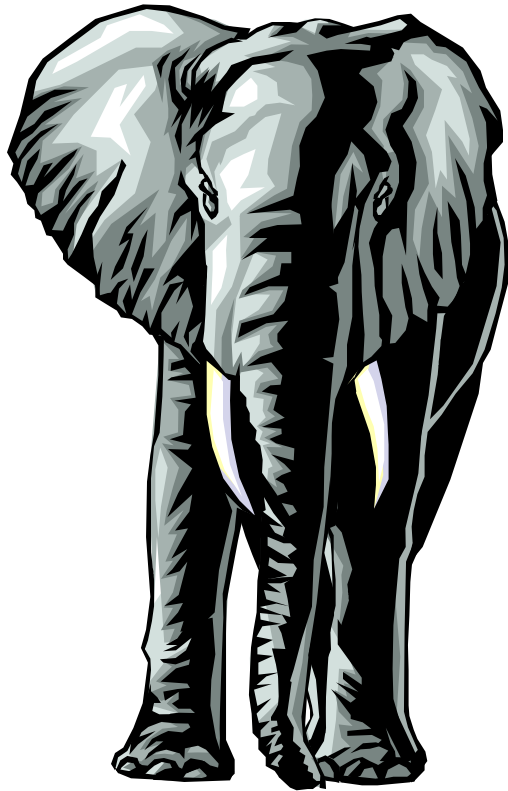
- ❏ Intelligent transport or not?

ABR vs UBR

Source			
Source	Router		

ATM

Dest.
Router	Dest.
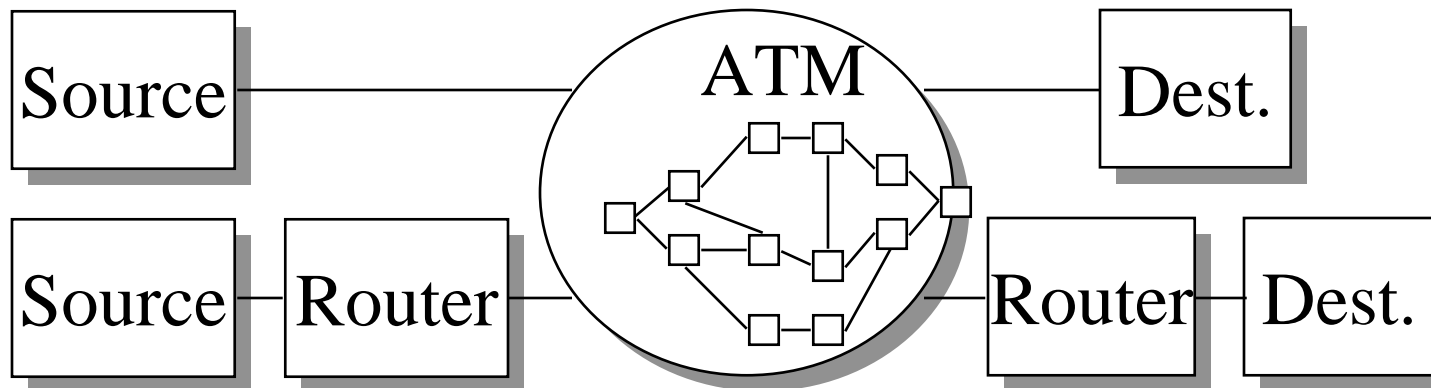
ABR	**UBR**
Queue in the source	Queue in the network
Pushes congestion to edges	No backpressure
Good if end-to-end ATM	Same end-to-end or backbone
Fair	Generally unfair
Good for the provider	Simple for user

Improving Performance of TCP over UBR

	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
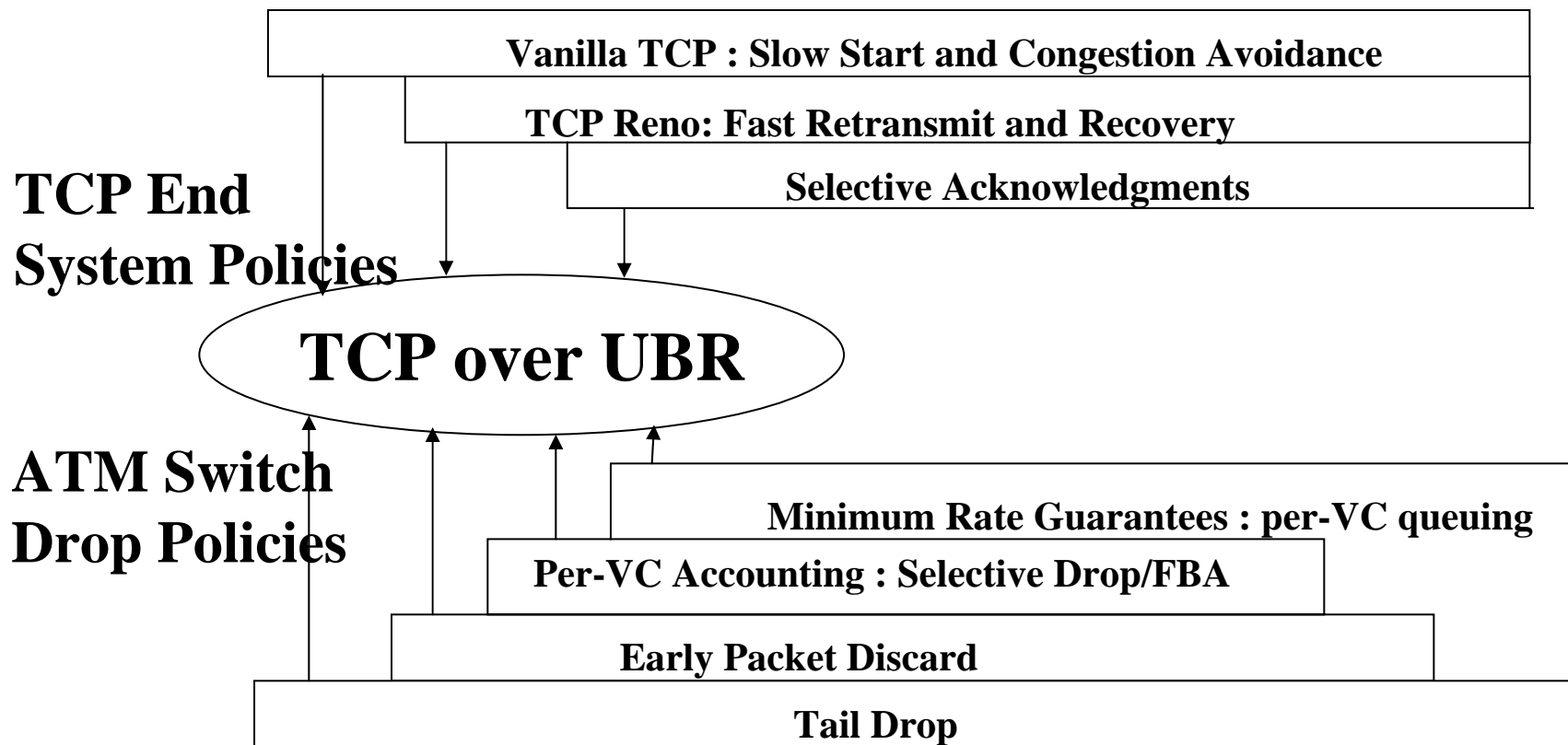
Policies

		End-System Policies			
		No FRR	FRR	New Reno	SACK + New Reno
No EPD					
	Plain EPD				
EPD	Selective Drop				
	Fair Buffer Allocation				

Switch Policies

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.

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.

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+

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.

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.

4 Ways to Improve UBR over Satellites

1. Implement "Selective Acknowledgement" in end-systems

2. Disable "Fast retransmit and recovery" in end-systems

3. Reserve a small fraction of bandwidth for UBR in the switches

4. Fix slow start implementations in end-systems to avoid errors due to integer arithmetic

Summary

❑ Binary feedback too slow for rate control. Especially for satellites. ER switches provide much better performance than EFCI.

❑ ABR service required for long-delay or high-speed networks. UBR+ may be OK for LANs but not for long delay paths.

Summary (Cont)

❑ Implement VS/VD, BECN, RIF=1, TBE=Large to improve ABR over satellites

❑ Implement SACK, Disable FRR, reserve bandwidth for UBR, and correct TCP implementations to improve UBR over satellites.

Our Contributions and Papers

❏ All our contributions and papers are available on-line at
   http://www.cse.ohio-state.edu/~jain/

❏ See Recent Hot Papers for tutorials.