

**96-1173 + 96-1267**  
**Switch Algorithm Testing:**  
**A Case Study with ERICA**

**Raj Jain, Sonia Fahmy, Shivkumar  
Kalyanaraman, Rohit Goyal**

The Ohio State University

<http://www.cis.ohio-state.edu/~jain/>

# Overview

- ❑ Objectives of ATM traffic management
- ❑ The ERICA algorithm
- ❑ Extensions of ERICA
- ❑ Performance evaluation of ERICA and ERICA+

# Objectives of Traffic Mgmt

- ❑ Efficiency and minimal delay
- ❑ Fairness: Max-min allocation and fairness index
- ❑ Good steady state: Minimal oscillations.
- ❑ Fast transient response
- ❑ Adaptation to the presence of multiple traffic classes  $\Rightarrow$  ABR capacity is not fixed
- ❑ Scalability to various speeds, distances, number of switches and number of VCs
- ❑ Need to adapt to high variance in demand and different traffic models

# Source Models

## Increasing complexity:

- ❑ Persistent cell traffic
- ❑ Bursty cell traffic
- ❑ Source bottleneck
- ❑ Persistent TCP sources
- ❑ Bursty TCP sources

# ERICA Scheme: Basic

- ❑ Explicit Rate Indication for Congestion Avoidance
- ❑ 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 Features

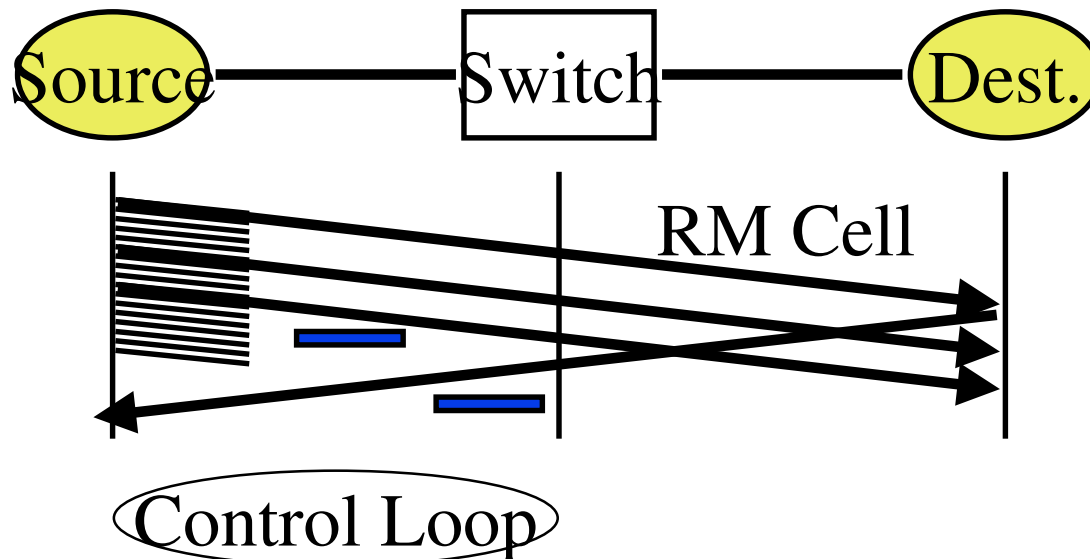
- ❑ Uses measured overload
  - ⇒ If sources use less than allocated capacity, all unused capacity is reallocated to others.
- ❑ Two parameters: Target utilization, Averaging interval
- ❑ Simple Order (1) computation
- ❑ Fast response due to optimistic design
- ❑ Fairness is improved at each step.  
Even under overload.
- ❑ Converges to efficient operation in most cases
- ❑ Max-min fair in most cases

# ERICA Extensions

1. Forward CCR
2. Same feedback in one Interval
3. Fair share first
4. per-VC CCR measurement
5. Time + count based averaging
6. ERICA with VBR
7. Bi-directional Counting
8. Max-min Fairness
9. Averaging of number of sources
10. Boundary cases
11. Averaging of load factor
12. ERICA+ (ERICA with queue control)

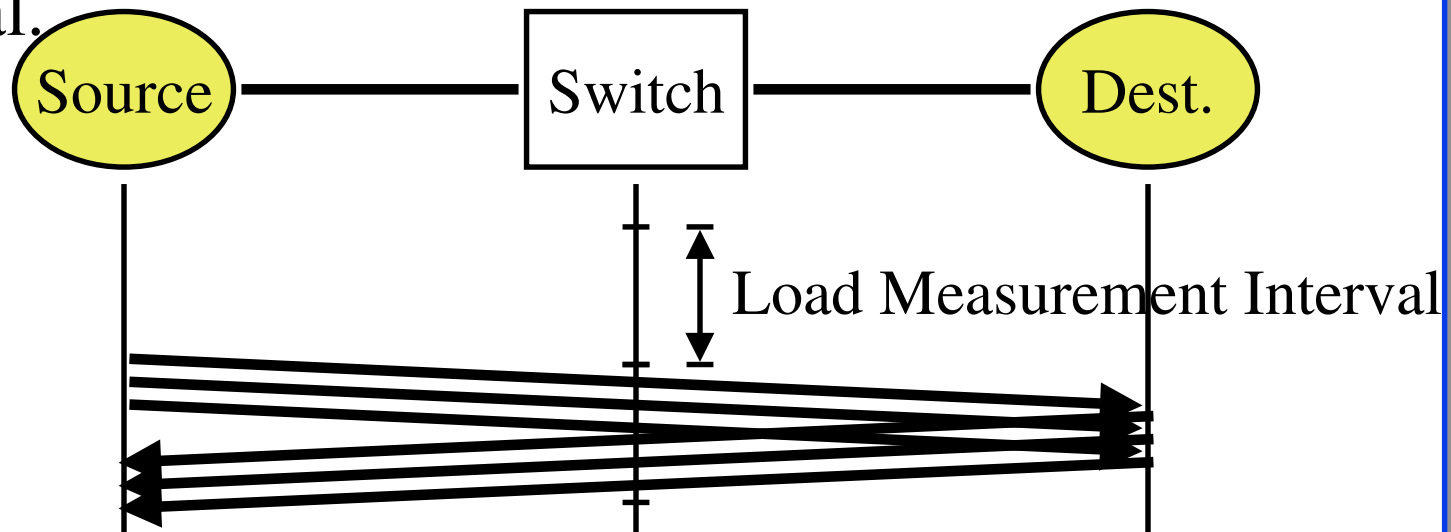
# 1. Innovation: Use forward CCR

- ❑ **Problem:** CCR in backward direction is too old
- ❑ **Solution:** Read CCR in forward RM cells.  
Give feedback in backward RM cells.
- ❑ **Effect:** Shorter control loop for active VCs  
⇒ Faster convergence



## 2. Same Feedback in One Interval

- ❑ **Problem:** Oscillations for high-rate sources
- ❑ **Reason:** Mismatched control and monitoring intervals
  - ❑ Control Interval = Inter-RM cell time
  - ❑ Monitoring Interval = Averaging interval
- ❑ **Solution:** Do not change feedback in one averaging interval.

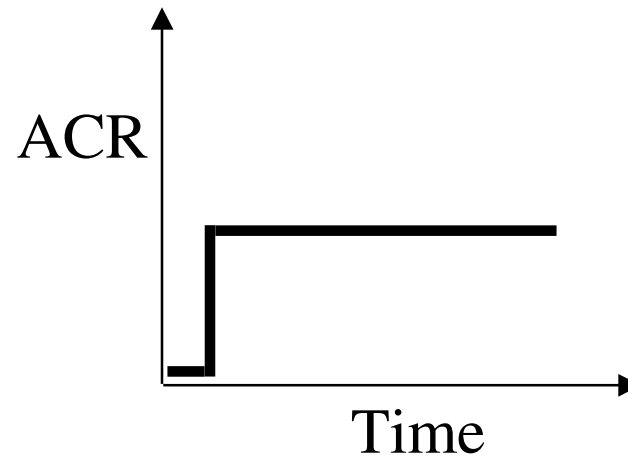
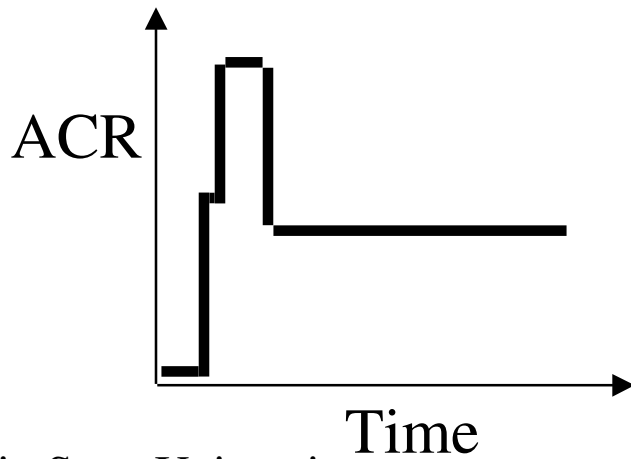


### 3. Innovation: Fair Share First

- ❑ **Problem:** Transient overloads at state changes
- ❑ **Solution:** Source below Fair Share go only up to fair share first.

IF  $CCR < \text{Fair Share}$  and  $ER_{\text{Calculated}} \geq \text{Fair Share}$   
THEN  $ER_{\text{Calculated}} = \text{Fair Share}$

- ❑ **Example:** Two sources  $\{10, 10\}$ ,  $\{50, 10\}$ ,  $\{90, 50\}$ ...



## 4. Per-VC Rate Measurement

- **Problem:** Some VCs are bottlenecked at the source  
CCR does not reflect source rate
- **Solution:**
  - Count number of cells in each VC
  - Source Rate = Number of Cells Seen/Averaging Interval
  - This VC's Share = Source Rate/Overload
- **Advantage:**
  - Also handles sources not using their allocation.  
⇒ Switch based “use it or lose it”

# 5. Time + Count Based Averaging

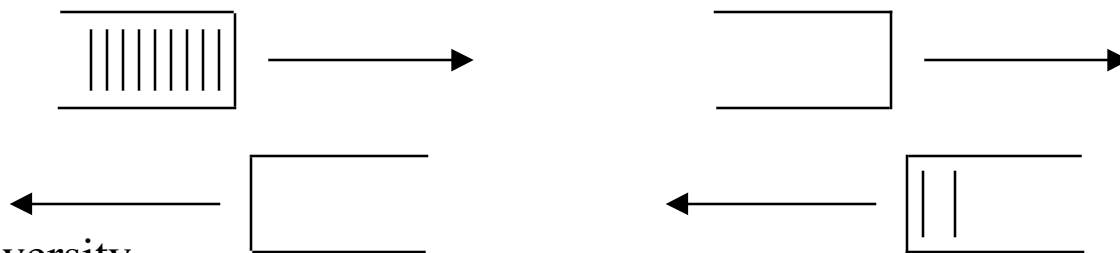
- ❑ **Problem:** Averaging over a fixed interval  
⇒ Sudden overload can cause queue build up
- ❑ **Solution:** Average over  $t$  ms or  $n$  cells whichever happens first.

## 6. Innovation: ERICA with VBR

- Monitor VBR usage
- ABR capacity = Target Rate - VBR input rate
- NOTE: Target utilization applies to total link load  
ABR capacity = Target Utilization  $\times$  Link Rate  
- VBR output rate  
and not  
ABR capacity = Target Utilization  $\times$  (Link Rate  
- VBR output rate)  
 $\Rightarrow$  VBR Output rate  $<$  Target utilization

# Out-Of Phase Effect

- ❑ Bursty load and backward RM (BRM) cells are often out of phase.
- ❑ When there is load in the forward direction, there are no BRMs.
- ❑ By the time the switch sees BRMs, there is no load in the forward direction.
- ❑ The above effect disappears when the bursts become larger than RTT



# 7. Innovation:

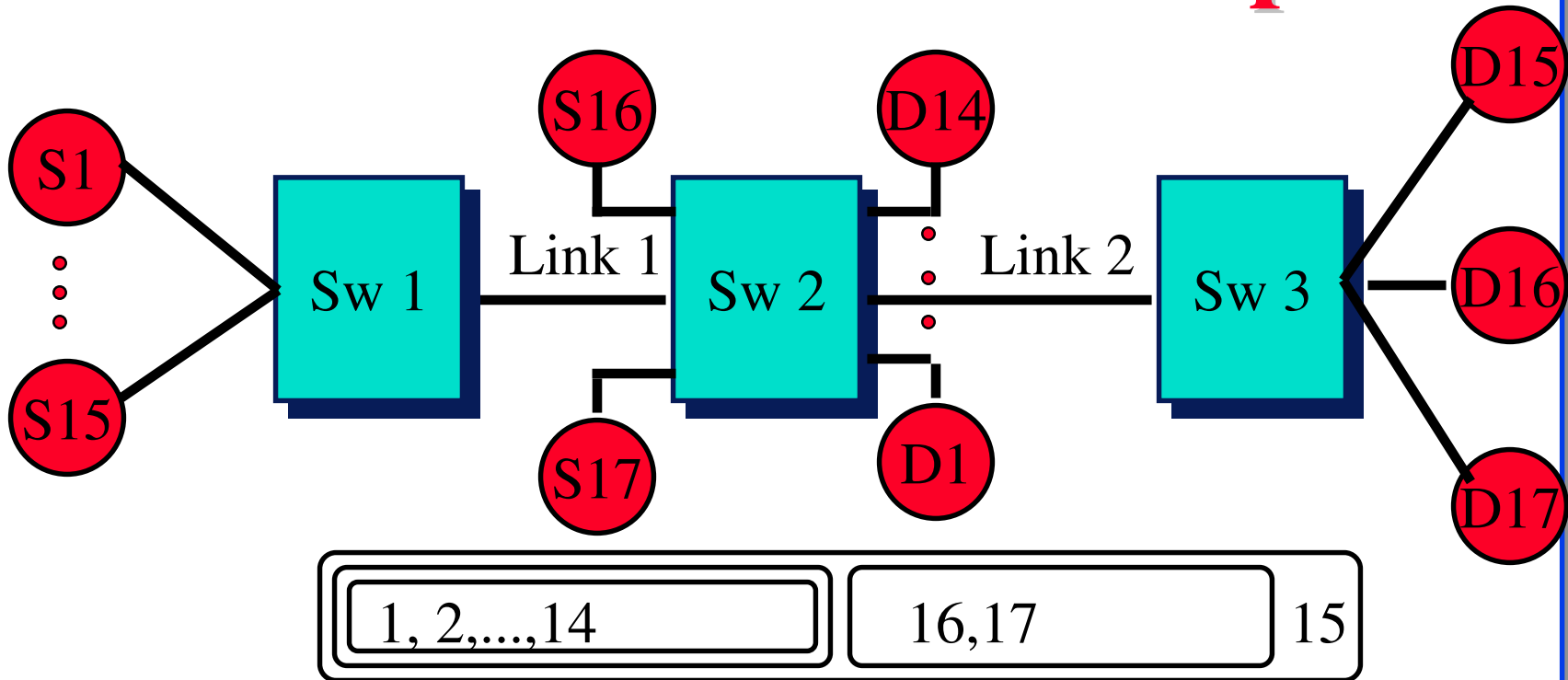
## Bi-directional Counting

- ❑ **Problem:** Data cells or RM cells may not be seen in one direction. Resulting in undercount and overallocation.
- ❑ **Solution:** A VC is active if any of the following holds:
  - ❑ Data cells seen in the forward direction in the last averaging interval
  - ❑ Data cells seen in the forward direction in this averaging interval
  - ❑ BRMs seen in the reverse direction

# Unfairness in ERICA

- ❑  $ER_{\text{Calculated}} = \text{Max}\{\text{Fair Share}, \text{CCR}/\text{overload}\}$
- ❑ ERICA becomes unfair if ALL of the following conditions hold true:
  - ❑ Overload = 1
  - ❑ Some VCs are bottlenecked at other switches
  - ❑ All VCs that are not bottlenecked at other switches have a CCR greater than the fair share
- ❑ Under the above condition, the CCRs do not change at all. The allocation stabilizes. But the stable operating point may not be max-min fair.

# Fairness Problem: Example



- ❑ Max-Min Alloc of 150 Mbps :  $\{10, 10, \dots, 10, 70, 70\}$
- ❑ With  $\{10, 10, \dots, 10, 60, 80\}$ , Link 2 Fair Share = 50, Load = 1,  $\text{Max}\{\text{Fair share, CCR/load}\} = 60$  and 80 for VC16 and VC17.

## 8. Innovation: Fairness Fix

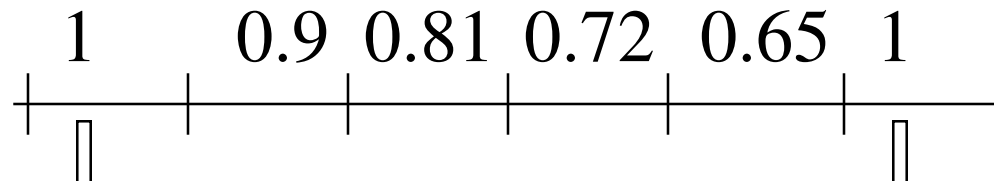
- **Solution:**
- All VCs that are bottlenecked at this switch must get the same allocation = maximum allocation
- Remember maximum ER in the previous interval
- IF overload  $\leq 1+\delta$   
THEN  $ER_{\text{Calculated}} = \text{Max}\{\text{Fair Share, CCR/Overload, Max\_ER}\}$   
ELSE  $ER_{\text{Calculated}} = \text{Max}\{\text{Fair Share, CCR/Overload}\}$

# Fairness Fix (Cont)

- **Example:** On Link 2, Fair Share = 50
  - {10, 10, ..., 10, 60, 80}, Load = 1, ER=10,80,80
  - {10, 10, ..., 10, 80, 80}, Load = 17/15, ER=10, 70.6, 70.6
  - {10, 10, ..., 10, 70.6, 70.6}, Load = 1.008, ER=10, 70.03, 70.03

# 9. Averaging of Number of Sources

- Not all active sources seen in every interval
  - ⇒ Fair share overestimated
  - ⇒ High Allocation
- Solution:
  - Source activity lies between 0 and 1
  - Activity = 1 if the source is seen
  - Activity decays by a factor  $\alpha$ , every interval the source is not seen



# 10. Boundary Cases

ABR Capacity	Input Rate	Over-load	Fair share	CCR/ Overload	Feedback
Zero	Non-zero	$\infty$	Zero	Zero	Zero
Non-zero	Zero	$\infty$	C/N	Zero	C/N
Non-zero	Non-zero	I/C	C/N	CCR*C/I	Max(CCR*C/I, C/N)
Zero	Zero	$\infty$	Zero	Zero	Zero

- If  $N < 1$  then  $N = 1$
- Here,  $I =$  input rate,  $C =$  Capacity,  $N =$  # of Srcs

# 11. Averaging of Load Factor

- ❑ Load Factor = Input Rate / ABR Capacity
- ❑ Load factor is a ratio
  - Both numerator and denominator are variable
    - ⇒ Average numerator and denominator separately
- ❑ Input rate itself is a ratio
  - ⇒ Add number of cells seen and time separately
- ❑ Similarly, for ABR Capacity

# Is Low Queue Length Good?

- ❑ Queue length is close to 1.  
Not good if bandwidth becomes available suddenly  
You can't use BECN to ask sources to increase  
Low rate sources may have long inter-RM cell times
- ❑ Link utilization is 90% or below  
May not be acceptable for high-cost WAN links.
- ❑ Very high queue length is also bad.

## 12. Innovation: ERICA with Queue Control

- ❑ Target utilization is dynamically changed.
- ❑ During steady state: Target utilization = 100%
- ❑ During overload the target may be low, e.g., 80%
- ❑ During underload the target may be high, e.g., 110%
- ❑ Available Bandwidth =  $\text{fn}(\text{Unused bandwidth, Queue length, queue length goal})$
- ❑ Unused bandwidth = Link Rate - VBR output rate
- ❑ Rest is similar to ERICA

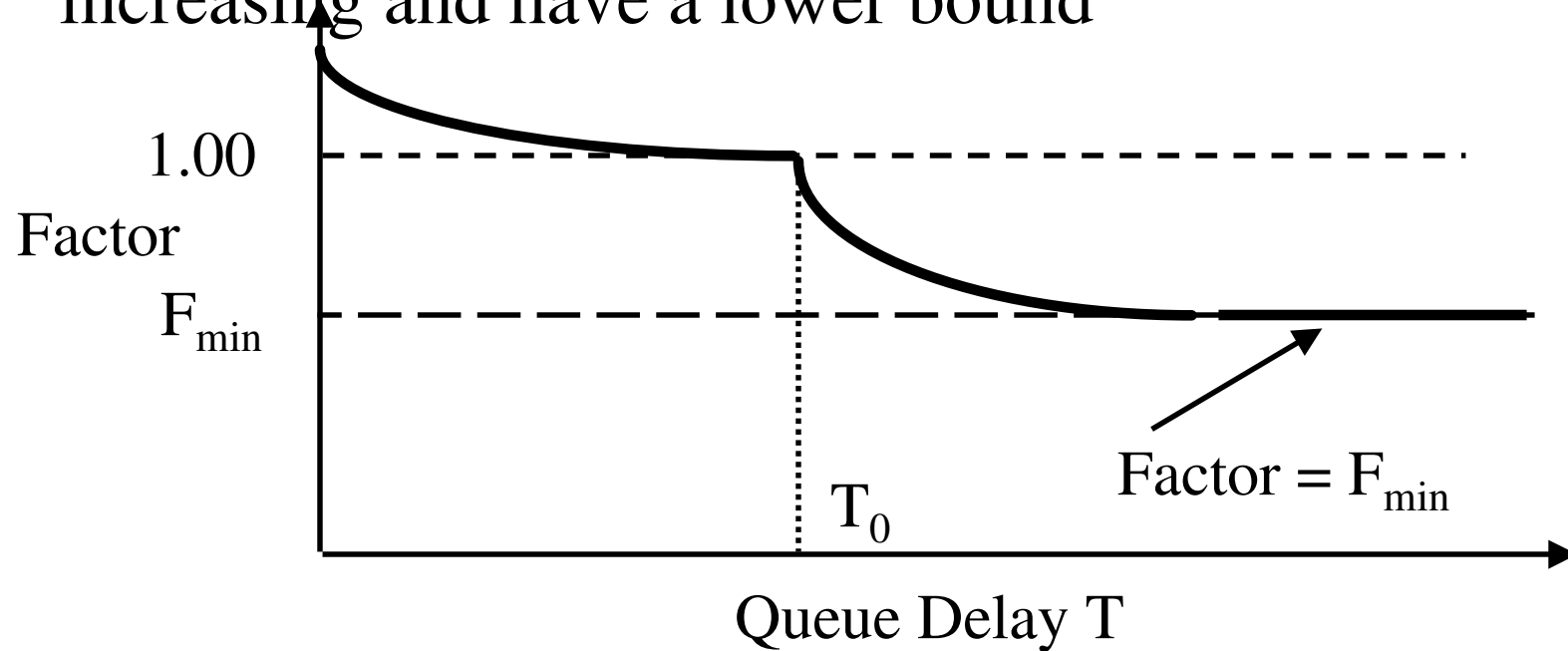
# Innovation:

## Use Queue Delay Threshold

- Since available bandwidth (AB) varies dynamically, a queue of 30 may be too big when AB is 1 Mbps but too little when AB is 100 Mbps.
- Use queue delay instead of queue length  
Queue Delay = Queue length / Available bandwidth
- Available Bandwidth = fn(Unused bandwidth, Queue length, **queue delay goal**)

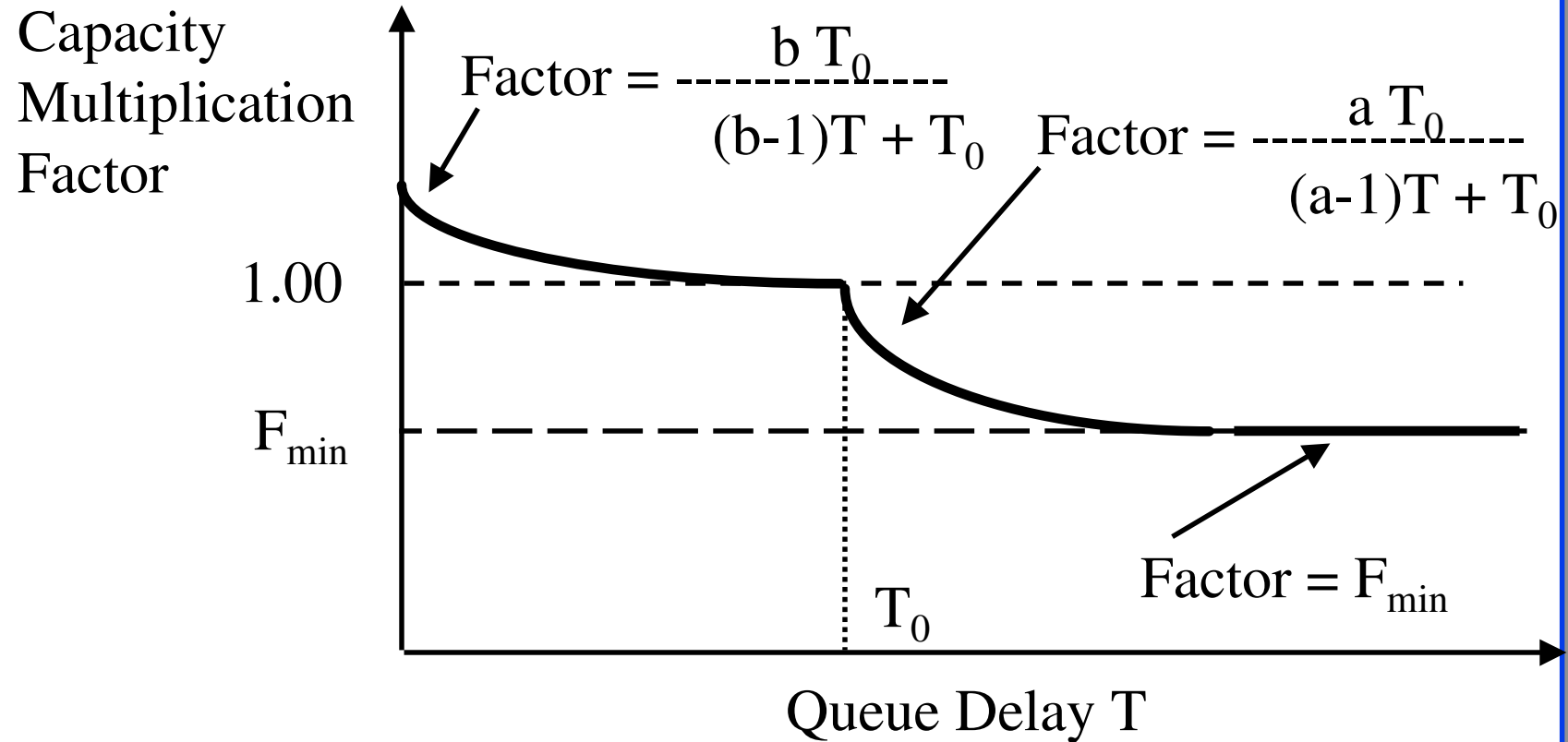
# Innovation: Target Utilization Function

- The function should be monotonically non-increasing and have a lower bound



$$\text{Available Bandwidth} = \text{Unused Bandwidth} \times \text{Factor}$$

# Sample Queue Control Function 1



Parameters:  $\{a, b, T_0, F_{\min}\} = \{1.15, 1.05, 5 \text{ ms}, 0.5\}$

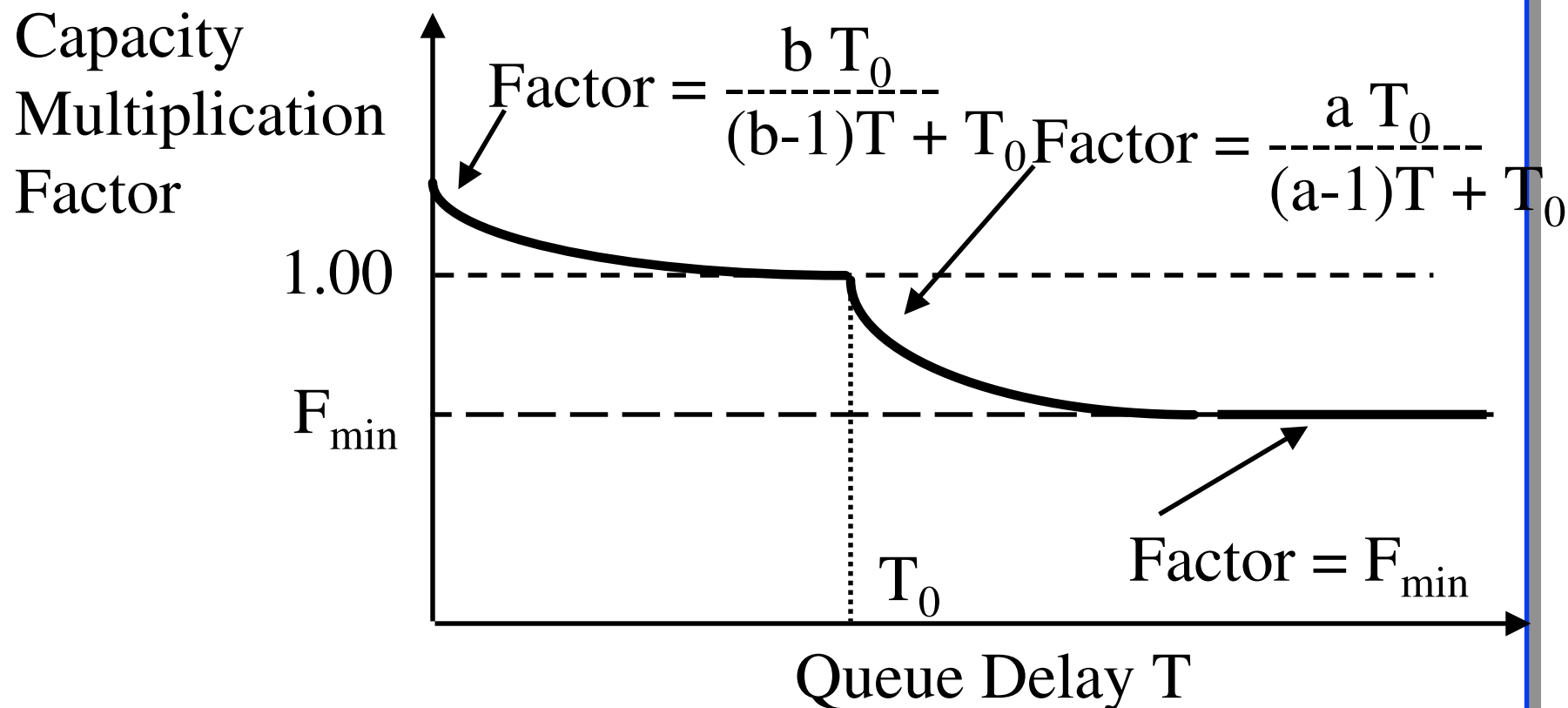
# Advantage of Q-Control

- ❑ Can tolerate errors in measurements:
  - ❑ Number of active sources
  - ❑ VBR load
  - ❑ ABR input rate
- ❑ Allows n-VC TCP operation with buffers  $\approx 1 \times$  RTT
- ❑ 100% Utilization

# Simulation Parameters

- ❑ All links have a bandwidth of 155.52 Mbps
- ❑ All LAN links are 1 Km long and all WAN links are 1000 Km long
- ❑ All VCs are bi-directional
- ❑ RIF = 1
- ❑ TBE = Large  $\Rightarrow$  Disable rule 6
- ❑ Target utilization = 95% (LAN), 90% (WAN)
- ❑ All sources, including VBR are deterministic
- ❑ Averaging interval =  $\text{Min}\{50 \text{ cells}, 1 \text{ ms}\}$  for LANs and  $\text{Min}\{100 \text{ cells}, 1 \text{ ms}\}$  for WANs

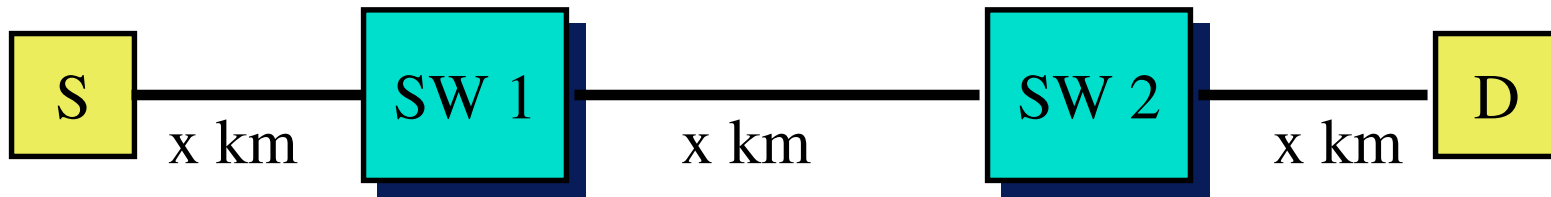
# ERICA+: Parameters



Parameters:  $\{a, b, F_{\min}\} = \{1.15, 1.05, 0.5\}$

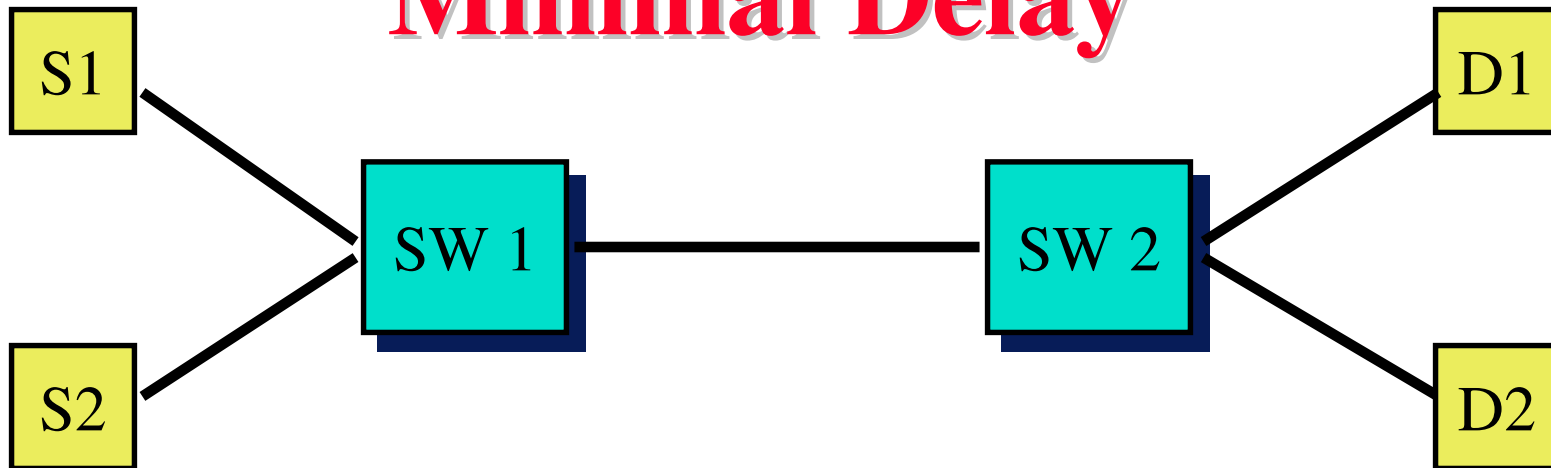
$T_0 = 100 \mu\text{s}$  (LAN),  $500 \mu\text{s}$  (WAN)

# Efficiency Test



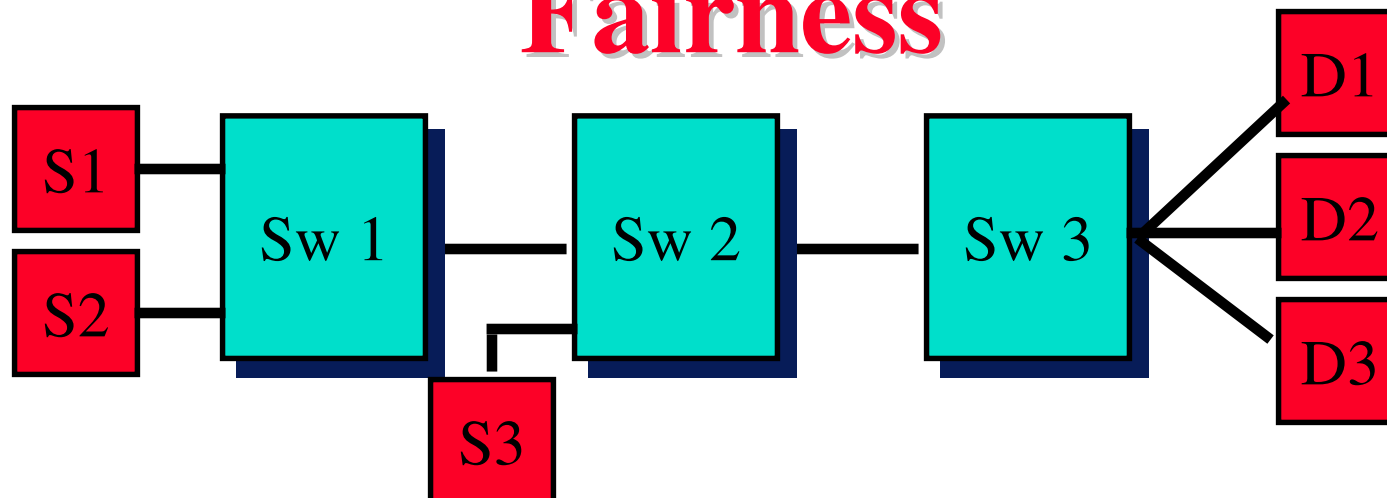
- ❑ Single source configuration:  
Filters out many non-working schemes
- ❑ ERICA achieves efficiency
- ❑ No rate oscillations in the steady state
- ❑ Utilization is at the target (95%)
- ❑ With ERICA+, utilization is 100% with no oscillations and minimal queues

# Minimal Delay



- ❑ Two source configuration
- ❑ For ERICA, convergence is fast, the queue lengths (delays) are small
- ❑ For ERICA+, convergence is fast, the queue length reaches target, no rate oscillations, and 100% link utilization

# Fairness

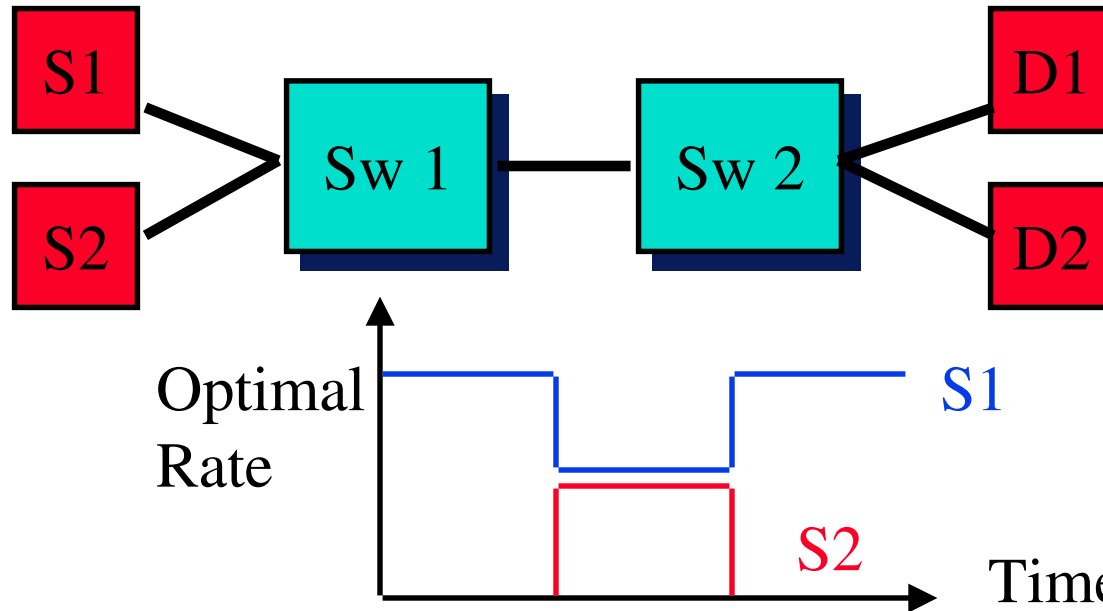


- ❑ Parking lot configuration
- ❑ Max-min allocation =  $1/n$  for all VCs
- ❑ ERICA and ERICA+ allocate the max-min share
- ❑ Parking lot configuration is not sufficient to demonstrate max-min fairness
- ❑ Original ERICA unfair in certain situations, e.g., some VCs bottlenecked at low rates

# Fairness (Cont)

- ❑ Modified ERICA is fair
- ❑ Curves of number of cells received at the destination vs time have the same slope
- ❑ Transient response is slightly worse but the steady state performance is still good

# Transient Performance



- ❑ Modified 2-source configuration,
- ❑ Source 2 is active from 10 ms to 20 ms only
- ❑ Also illustrates the effect of the "fair share first" algorithm
- ❑ ERICA exhibits good transient response

# Adaptation to Variable Capacity

- ❑ VBR source with peak rate of 124.42 Mbps (80%)
- ❑ VBR source is
  - on/off for 1 ms/1 ms (high frequency)
  - on/off for 20 ms/20 ms (low frequency)
- ❑ Fast response to VBR load
- ❑ Utilization drops reflect feedback delay
- ❑ Spikes in the queue lengths also reflect the feedback delay, but the queues are rapidly drained
- ❑ ERICA+ adapts rapidly to changing background
- ❑ Target queue goal is not reached due to the high variance

# Bursty Traffic

- ❑ One persistent + One bursty (request-response) VC
- ❑ Request Size = 16 cells
- ❑ Response Size = 128 (small), 1024 (medium), and 6144 (large) cells
- ❑ Performance of the reverse (response) shown
- ❑ ERICA can adapt to small and medium bursts of data, and the queue lengths are constrained
- ❑ With a target utilization of 90%, not enough capacity to drain large bursts of data from the switch queues before the next burst is received

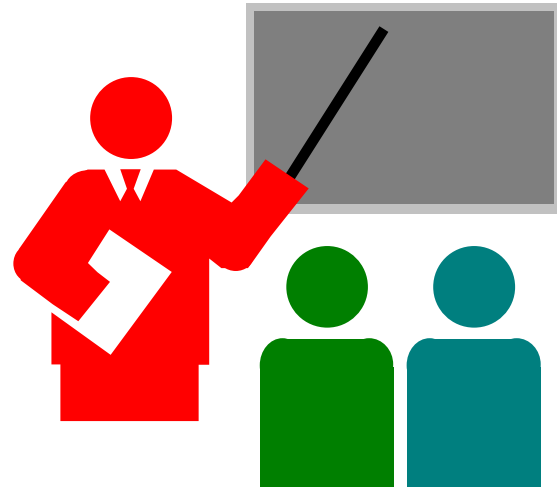
# Bursty Traffic (Cont)

- ❑ Solution 1: Smaller target utilization
- ❑ Solution 2: Bi-directional counting limits the queue sizes for large bursts (out-of-phase effect)
- ❑ Solution 3: Averaging the number of active sources
- ❑ Solution 4: ERICA+ can adapt to bursty traffic better than ERICA
- ❑ With ERICA+ and small burst sizes, the queue delay is below the target
- ❑ Even with large burst sizes, averaging not required for ERICA+

# ACR Retention

- ❑ ACR Retention = Sources cannot use their ACR
- ❑ If they suddenly use ACR  $\Rightarrow$  Overload
- ❑ Larger number of such VCs  $\Rightarrow$  Sudden overload
- ❑ 10 Sources limited to 10 Mbps for first 100 ms only
- ❑ ERICA rapidly detects the overload and gives the appropriate feedback
- ❑ Per-VC CCR measurement option can mitigate the overload situation

# Summary



- ❑ Efficiency and delay requirements
- ❑ Fairness
- ❑ Transient and steady state performance
- ❑ Scalability
- ❑ Adaptation to variable capacity and various source traffic models