95-1346R1

ERICA+: Extensions to the ERICA Switch Algorithm

Raj Jain, Shiv Kalyanaraman, Rohit Goyal, Sonia Fahmy, Fang Lu
The Ohio State University

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

- ERICA under VBR
- Scheduling of multiple classes
- ERICA with full utilization
  - Features
  - New Algorithm
  - Simulation Results

Current Switch Algorithm: ERICA

- ERICA = Explicit Rate Indication for Congestion Avoidance
- Monitor:
  Overload = Input rate/Target Rate
  Fair Share= Available rate/# of active VCs
- This VC's Share = CCR/Overload
- ER = Max(Fair Share, This VC's Share)
  ER in Cell = Min(ER in Cell, ER)
- ER in Cell = Min{ER in Cell, Max(Available rate/# of active VCs , CCR/Overload) }
- Use BECN option when appropriate

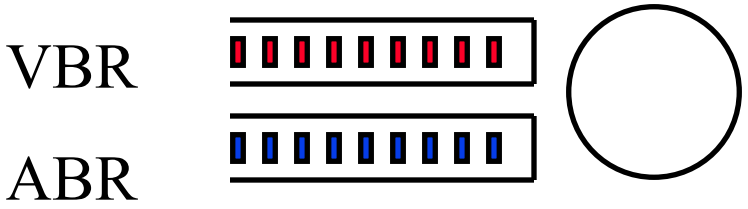
ERICA under VBR

- With VBR, the available bandwidth (AB) changes dynamically.
- Need:
  - An algorithm for setting explicit rates as changes
  - A scheduling algorithm for multiple classes

Innovation 1: Allocation

- Monitor VBR usage
- ABR capacity = Target Rate - VBR input rate
- Overload factor = ABR input rate/ABR capacity
- This VC's share = VC's CCR/overload factor
- Fair share = ABR capacity/Number of active ABR VCs
- ER = Max{Fair share, This VC's share}
- NOTE:
  - ABR capacity = Target Util. × Link Rate - VBR output rate
    and not
    ABR capacity = Target Util.×(Link Rate - VBR output rate)
    ⇒ Target utilization applies to total link load
  - VBR Output rate < Target utilization

Scheduling



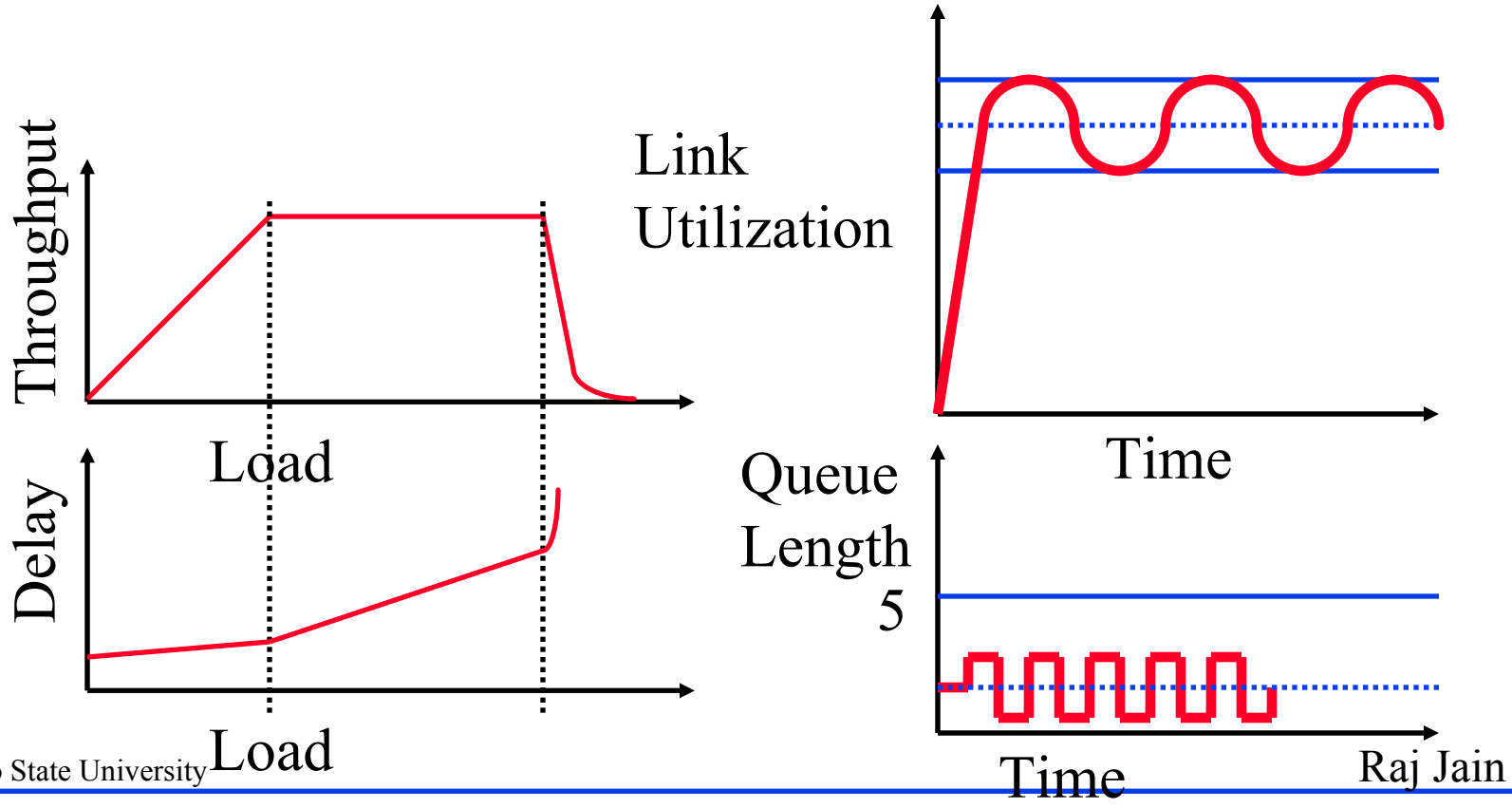
- Allows any desired allocation:
  V_{frac} = Fraction reserved for VBR
  A_{frac} = Fraction reserved for ABR
- Guarantees non-starvation.
  All classes can have a guaranteed minimum.
- No capacity is wasted.
  Capacity not used by one is used by the other class
- Scheduling decision is made per cell or per group of cells
- Keep scores for both (or n) classes
- The class with higher score is serviced next
- Complete pseudo-code in the contribution

Scheduling Variables

- A_{frac} = Minimum Fraction desired for ABR
  V_{frac} = Maximum Fraction desired for VBR
  (A_{frac} ABR cells are transmitted for every V_{frac} VBR cells)
- A_{credit} = Current credit for ABR traffic
  V_{credit} = Current credit for VBR traffic
  (In general, the traffic with higher credit is serviced next.)
- A_{queue} = Number cells in the ABR queue
  V_{queue} = Number cells in the VBR queue
- A_{count} = Number of ABR cells served
  V_{count} = Number of VBR cells served
- The complete pseudo-code is in the original contribution text.

Congestion Avoidance

- High throughput, Low delay
- Small queues
- Bounded oscillations ⇒ Good for Video traffic

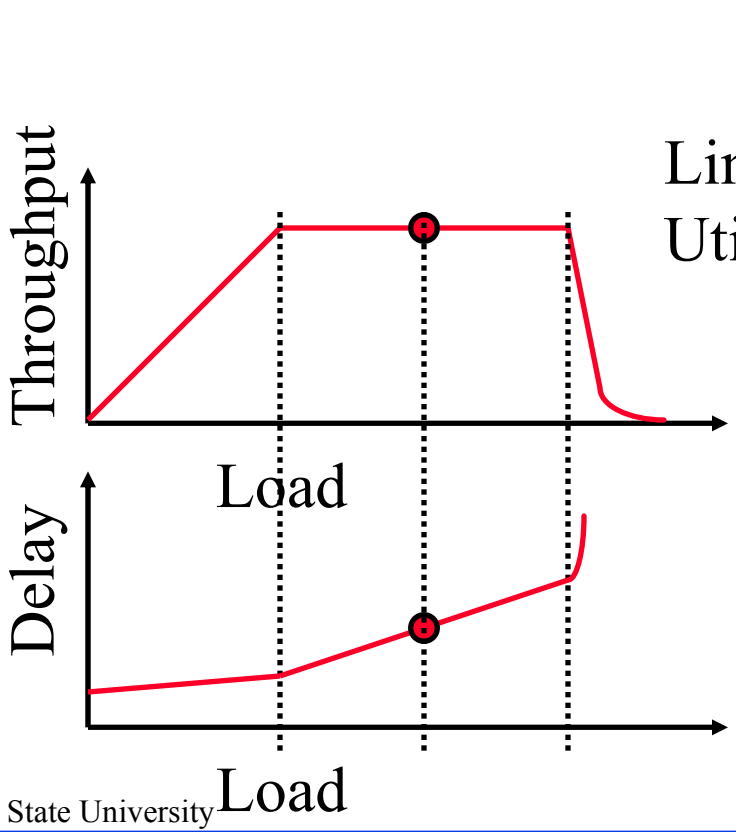


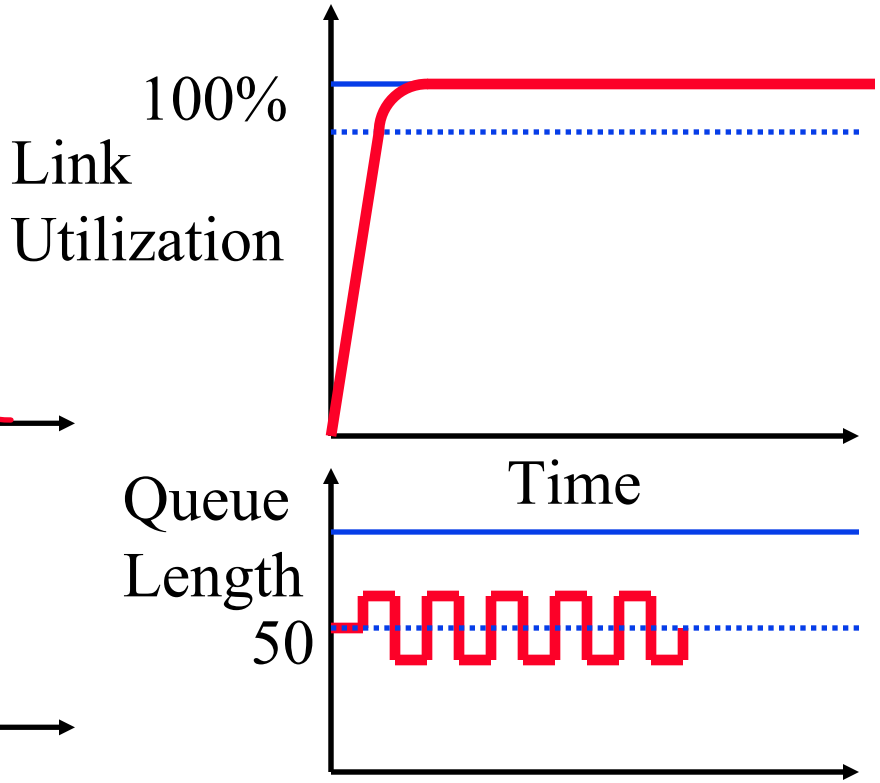
Issues with Current Congestion Avoidance Schemes

- Link utilization is 90% or below
  May not be acceptable for high-cost WAN links.
- 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

Features of the New Scheme

- Allows operation at any point between the knee and the cliff
- The queue length can be set to any desired value.
- Allows utilization to be 100%





Features (Continued)

- Compatible with current ATM Forum TM agreements
- No changes to source operation required
- No changes to destination operation required
- No changes to RM cell format required
- Follows all switch requirements

Innovation 2

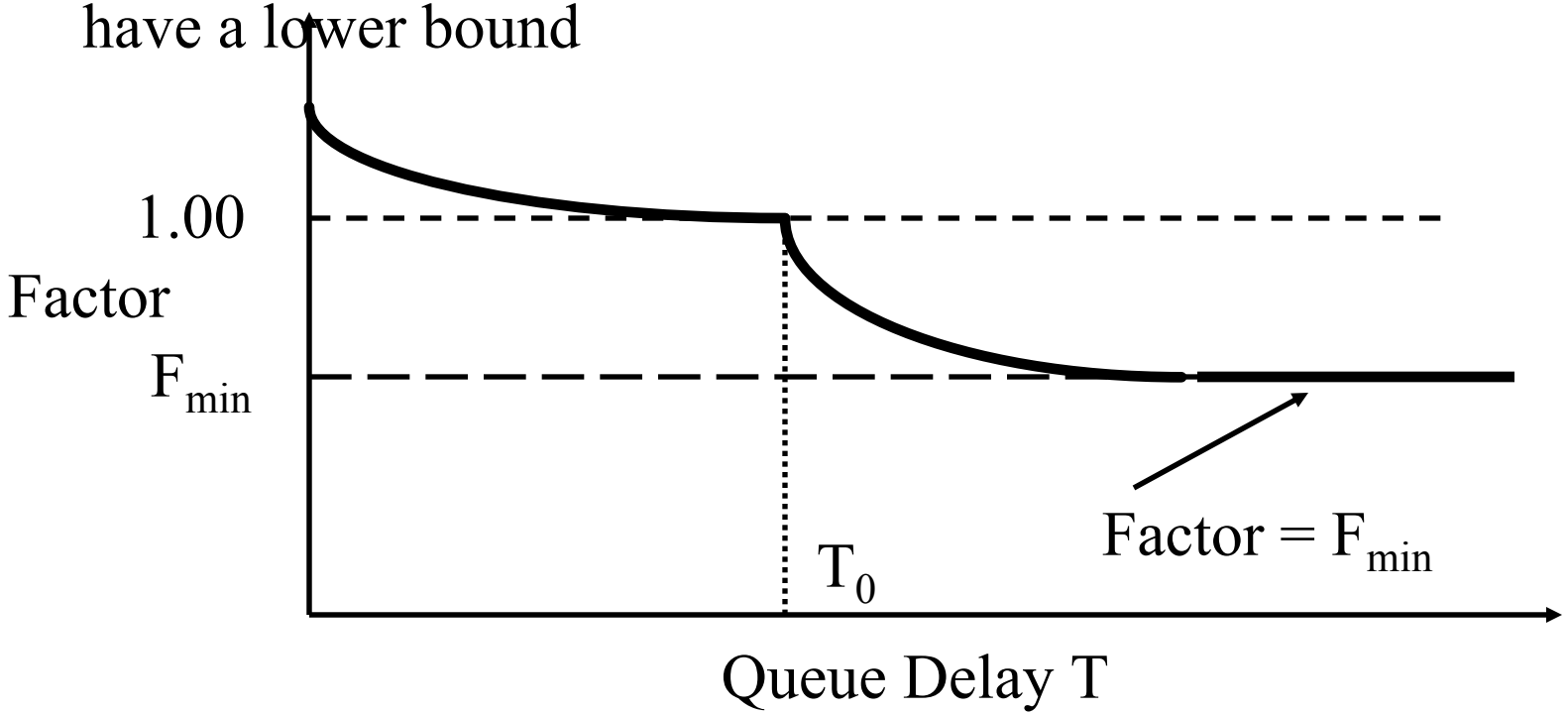
- Target utilization is dynamically changed.
- During steady state: Target utilization = 100%
- During overload the target may be low, e.g., 80%
- During underload the target may be high, e.g., 110%
- Available Bandwidth = fn(Unused bandwidth, Queue length, queue length goal)
- Unused bandwidth = Link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

Innovation 3

- 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

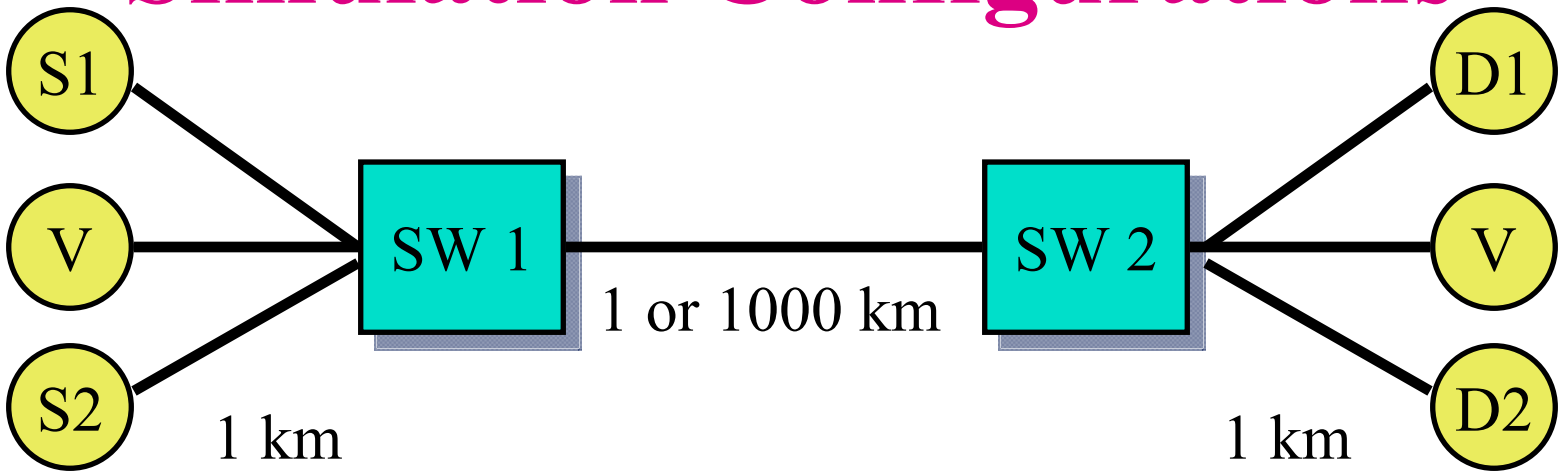
Innovation 4

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



Available Bandwidth = Unused Bandwidth ´ Factor

Simulation Configurations



- All links 155.52 Mbps
- One or Two ABR sources
- With/without VBR background traffic
- LAN or WAN
- The VBR source turns on/off
  - Every 4 ms (LAN) starts at t = 2 ms
  - Every 20 ms (WAN) starts at t = 12 ms

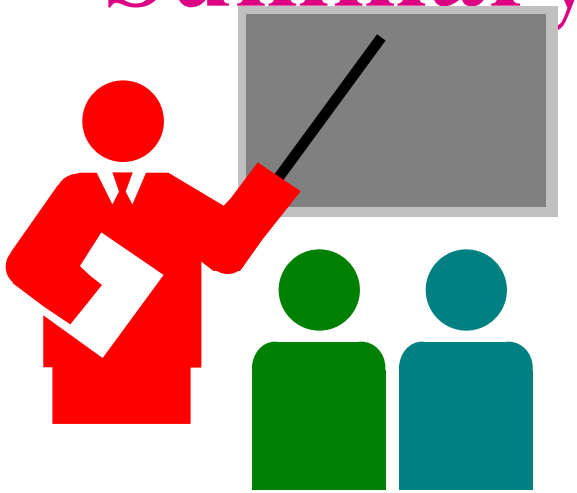
Simulation Parameters

- Source: Standard group #7
  Nrm = 32
  AIRF = 1 ⇒ AIR = PCR/Nrm ⇒ ACR is not limited by AIR
  RDF = 512 cells
  {TDFF, PNI} = {1/8, 0} ⇒ Rule 5 on
  CIF = 512 (LAN), 8192(WAN)
  ⇒ High ICR ⇒ Rule 5 not triggered
  RTT = 10 times the actual propagation delay
  ⇒ High XRM ⇒ Rule 6 not triggered
  XDF = 1/2
- Traffic: Unidirectional
- Switch:
  Averaging interval = min{30 cells, 200 μs}

Simulation Results

- ERICA+ provides 100% utilization
- No overflow or underflow of queues
- ERICA+ converges fast
- Queue length stays at the desired value

Summary



- ERICA with VBR and a general scheduling algorithm
- ERICA+ provides 100% utilization
- Allows operation at any point between the knee and the cliff.
- 100% throughput with any desired queue length possible.
- Provides quick response to transients
- High starts possible in LAN environments