Acknowledgements

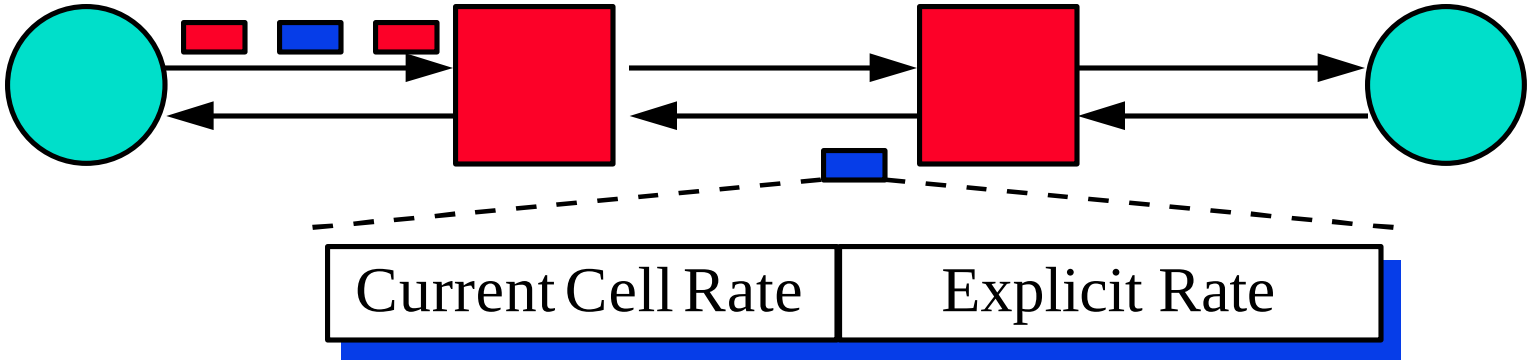
- **Our Team Members**:
  - Shivkumar Kalyanaraman
  - Ram Viswanathan
  - Rohit Goyal
  - Sonia Fahmy
  - Fang Lu
- **Our Sponsors**:
  - Intel
  - Stratacom
  - NASA

- Why worry about congestion in high speed networks?
- When is a network congested: High queue or high input?
- How much bandwidth to allocate each user?
- What is the appropriate goal: Avoidance or Control?
- Can we get full utilization and still have low delay?

Economic Reasons

- Network is a shared resource
  Because it is expensive and needed occasionally
  (Like airplanes, emergency rooms)
- Most costs are fixed.
  Cost for fiber, switches, laying fiber and maintaining them does not depend upon usage
  $\Rightarrow$ Underutilization is expensive
- But overutilization leads to user dissatisfaction.
- Need a way to keep the network maximally utilized

The Explicit Rate Scheme

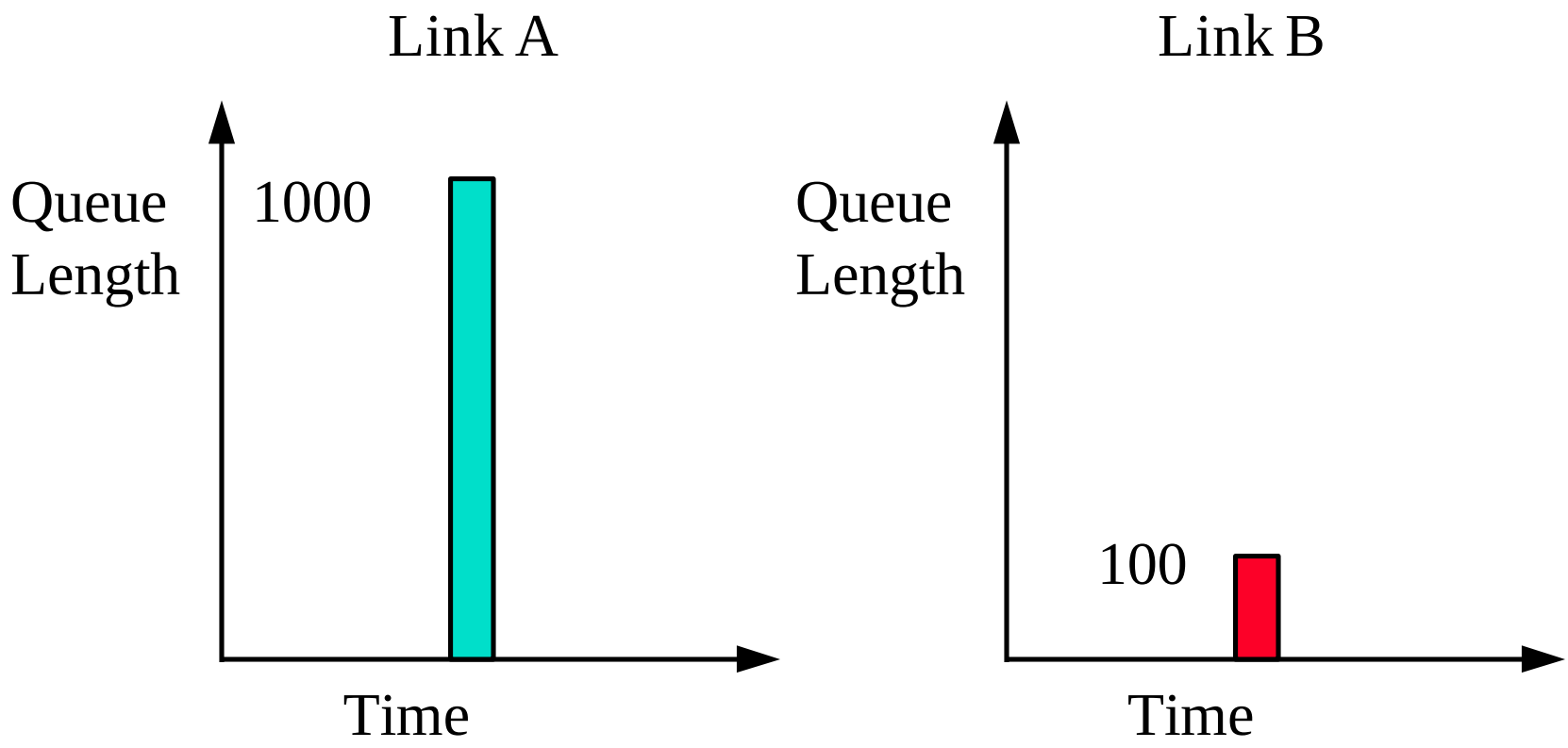


- Every Nrm cells, the sources send a control cell
- The switches measure load over a period
- The destination returns the cell to the source
- The switches specify explicit rate in cell
- The source adjusts the transmission rate

OSU Congestion Principles

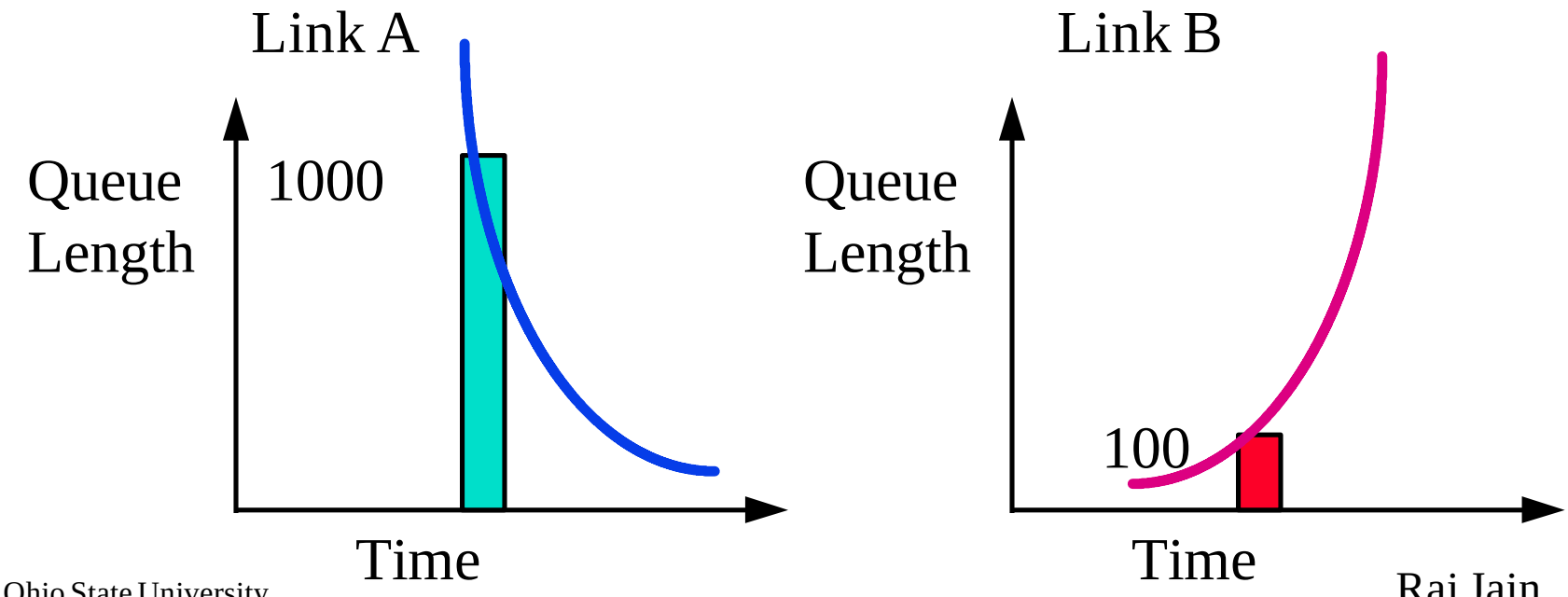
- Input rate (and not queue length) is the load measure
- Transient performance (and not the steady state performance) is more important
- Congestion avoidance (and not congestion control) should be the goal

Which Link is More Overloaded?



Answer: It Depends!

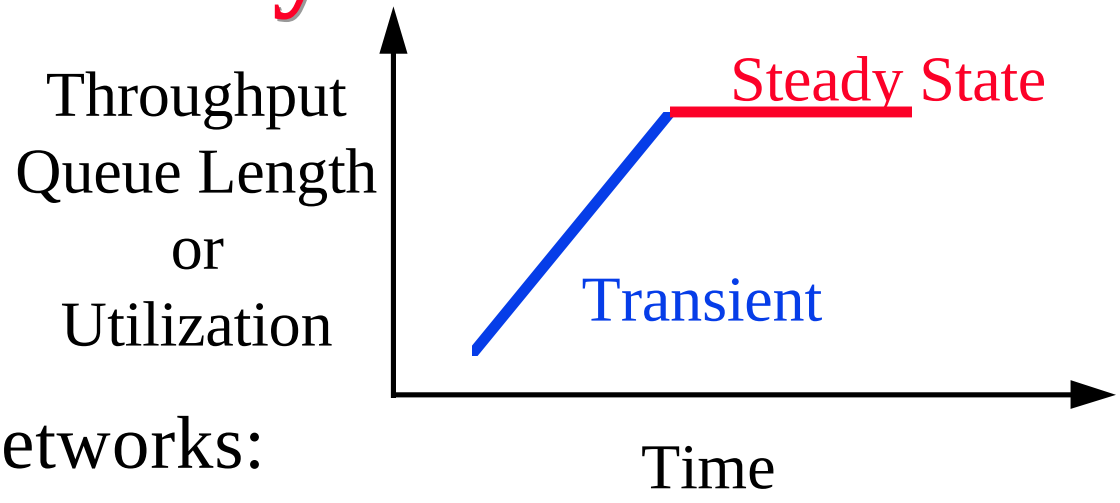
- Link Speed: OC-12 or T1?
- Control: Rate or Window?
  Q = Window, dQ/dt = Rate
- For Rate Control: Monitor Q growth rate



Conclusions

- Instantaneous queue length is not a good indicator of load for a rate controlled system.
  Q(t) = Q(t-1) + Input rate - Service rate
- Using queue length as the load indicator in a rate controlled system leads to unnecessary oscillations.
- Input rate monitoring not only correctly tells whether the system is overloaded, it also tells by what factor.

Why Worry About Transients?

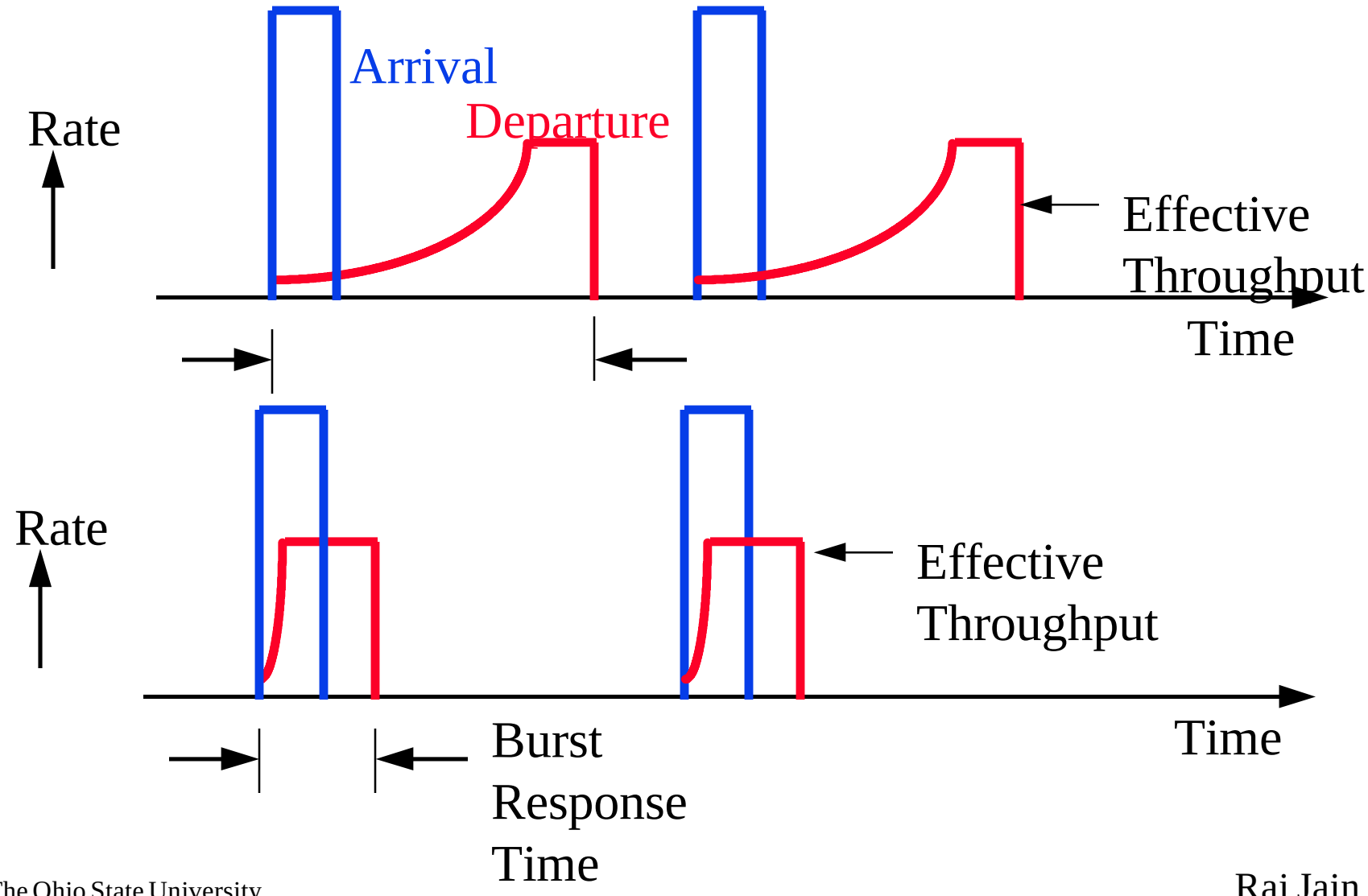


On most networks:

- There are no infinite sources.
- Sources come and go
- VCs may stay but are mostly inactive
- Traffic is highly bursty

⇒ Networks are operating in the transient region, most of the time.

Burst Performance

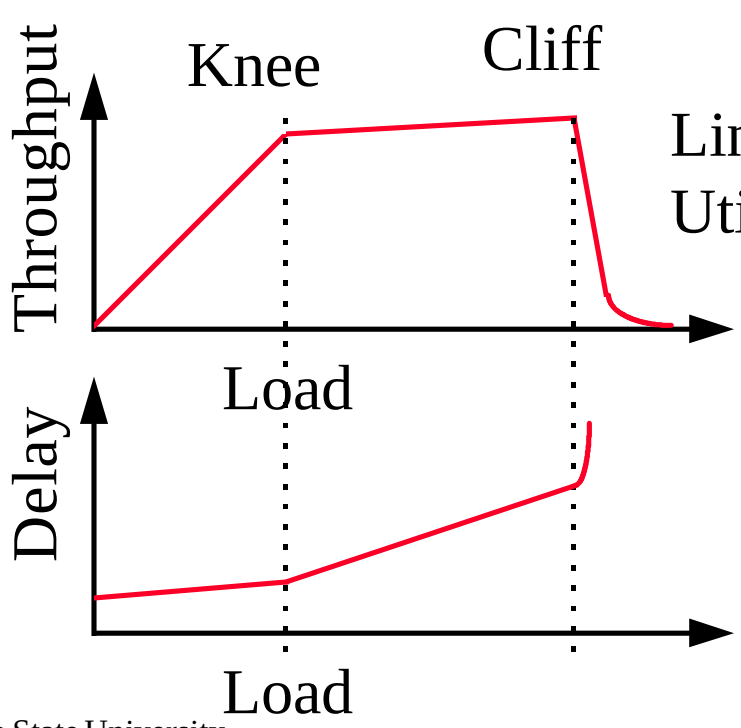


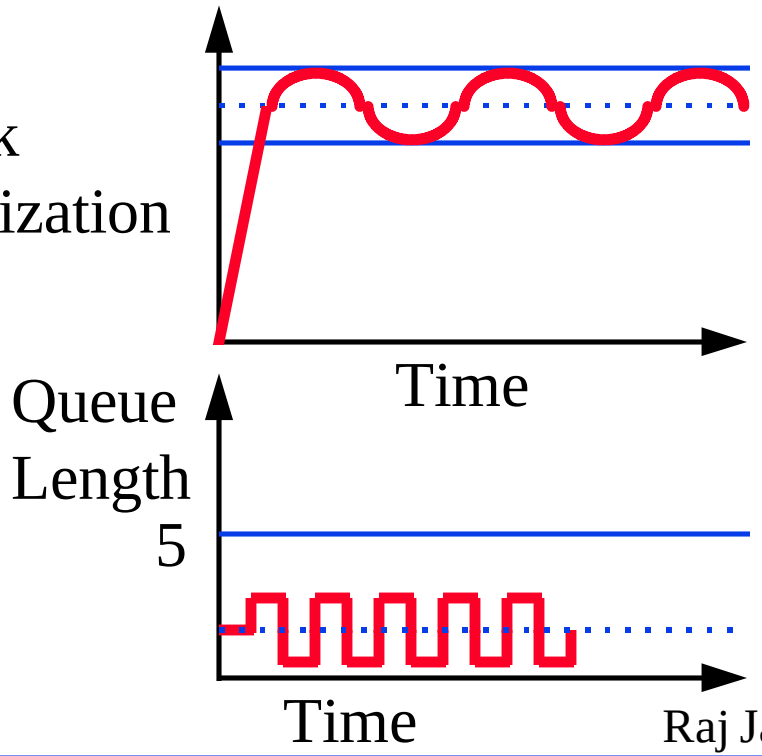
Legacy LANs vs ATM

- Today's LANs have a very fast transient response. Can get to the peak rate within a few microseconds
- On ATM LANs:
  Wait for connection setup and then...
  Everytime, a burst arrives, take several milliseconds to ramp up
- Q: Given 100 Mbps Switched Ethernet and 155 Mbps ATM at the same price, which one would you buy?

Congestion Avoidance

- Congestion Control: Operation at the cliff
- Congestion Avoidance: Operation at the knee
  High throughput, Low delay, Small queues
- Load = Input rate/(Target Utilization*Capacity)





ERICA Switch Algorithm

Explicit Rate Indication for Congestion Avoidance

- Set target rate, say, at 95% of link bandwidth
- Monitor input rate and number of active VCs *k*
  Overload = Input rate/Target rate
- This VC's Share = CCR/Overload
- Fairshare = Target rate/ *k*
- ER = **Max**(Fairshare, This VC's share)
- ER in Cell = Min(ER in Cell, ER)

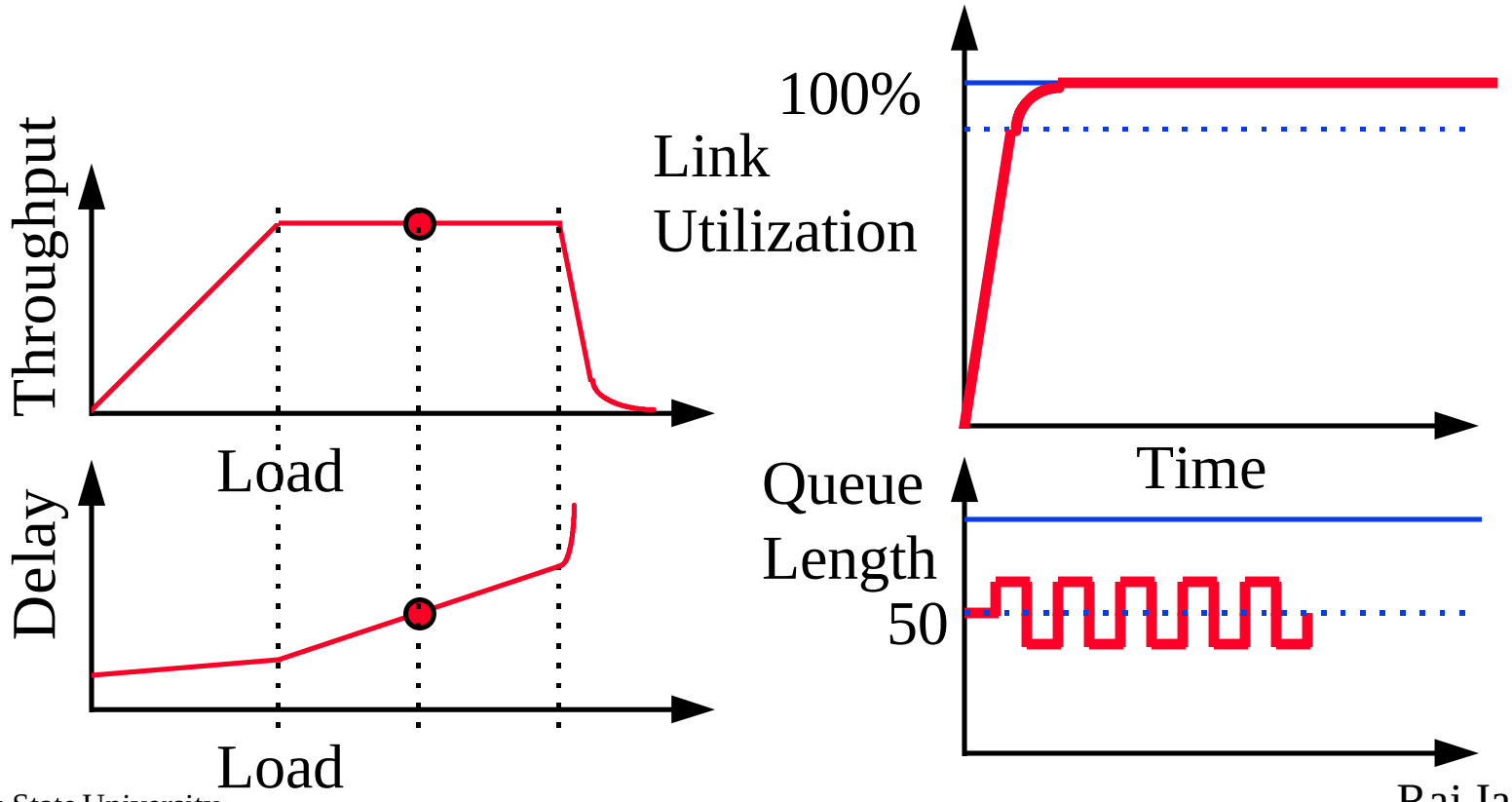
Ref: R. Jain, et al, "A Simple Switch Algorithm,"
AF-TM 95-0179R1, February 1995.

ERICA Features

- Measured overload/load at switch
- Insensitive to source not using their allocated rates
- Small queue lengths during steady state
- Fast response due to optimistic design
- Parameters: Few, insensitive, easy
- Several options: BECN
- Simplified switch algorithm
- Optimized all steps. Eliminated unncessary steps. Eliminated many parameters

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%



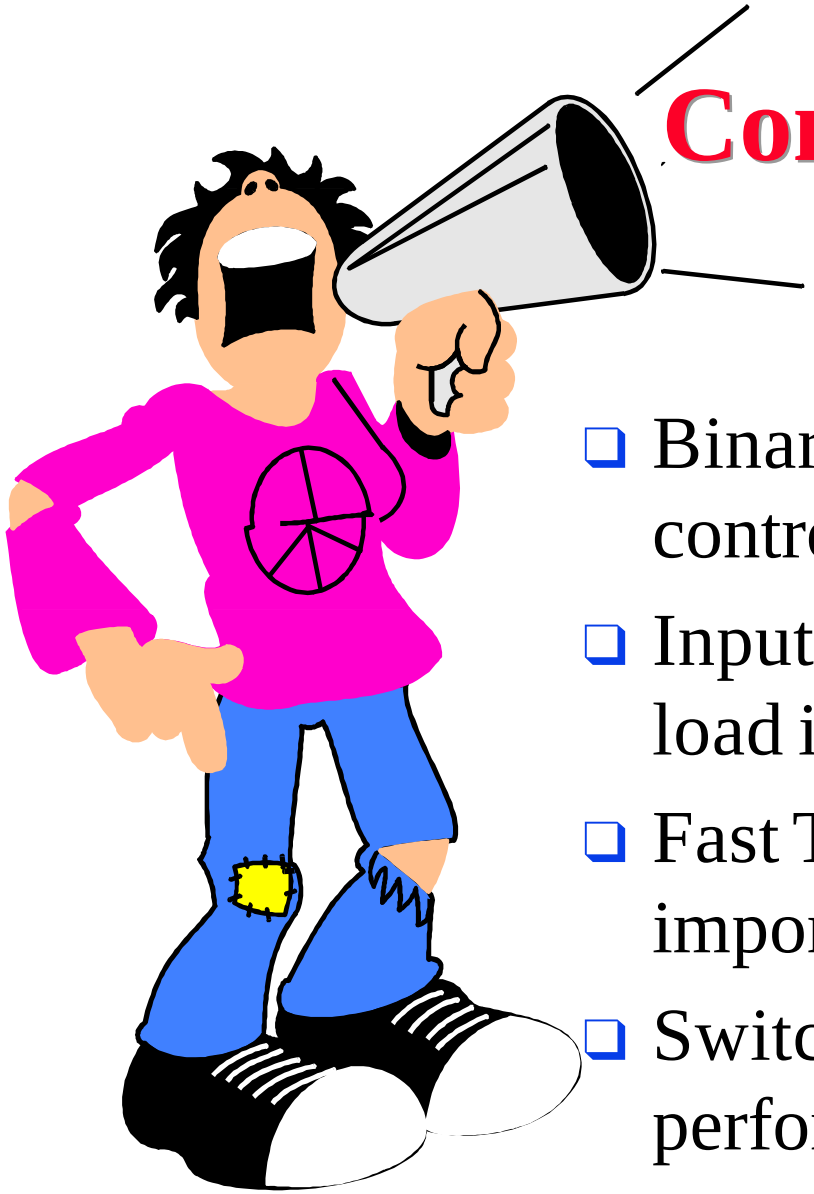
ERICA+: Switch Algorithm

- Target cell rate = Target Utilization × Link Capacity
- Target Utilization
  = fn(Current load, Queue length, Queue drain time goal)
- Rest is similar to ERICA
- Features:
  - Queue length is bounded during overload
  - No queue underflow ⇒ Switches keep ABR cells waiting to be transmitted as soon as the bandwidth becomes available.
  - 100% Utilization even with VBR

Future

- Intermittant sources
- Non-conforming sources
- Optimal Source Strategy
- Out-of-rate cell strategy
- Interoperability of different switch algorithm
- Virtual Source/destination
- Multicast
- Implicit feedback schemes: Heterogeneous Networks

Congestion: Summary



- Binary feedback too slow for rate control
- Input rate (not queue length) is a load indicator for rate
- Fast Transient performance is important
- Switch scheme affects the performance

Our ATM Forum Contributions

All contributions are available on-line at
http://www.cis.ohio-state.edu/~jain/

- R. Jain, S. Kalyanraman, S. Fahmy, and F. Lu, "Parameter Values for Satellite Links,", AF-TM 95-0972, August 1995.
- R. Jain, S. Kalyanraman, S. Fahmy, and F. Lu, "Out-of-Rate RM Cell Issues and Inconsistencies in the End System Behavior and the Pseudocode," AF-TM 95-0973, August, 1995.
- R. Jain, S. Kalyanraman, R. Goyal, "Simulation Results for ERICA Switch Algorithm with VBR+ABR traffic," AF-TM 95-0467, April 1995.

- R. Jain, S. Kalyanraman, R. Goyal,"The Case of Negative ABR Bandwidth: The need for a probe interval parameter," AF-TM 95-0468, April 1995
- R. Jain, S. Kalyanraman, R. Viswanathan, R. Goyal, "A Sample Switch Scheme," AF-TM 95-0178, February 1995
- R. Jain, S. Kalyanraman, R. Viswanathan, R. Goyal, "Simulation Results for the Sample Switch Scheme," AF-TM 95-0179, February 1995
- R. Jain, "Congestion Control in ATM Networks: Recent Advances and a Survey," AF-TM 95-0177, February 1995
- R. Jain, S. Kalyanraman, and R. Viswanathan, "Transient Performance of EPRCA and EPRCA++" AF-TM 94-1173, November 1994.

- R. Jain, S. Kalyanraman, and R. Viswanathan,"Current Default Proposal: Unresolved Issues," AF-TM 94-1175R1, November 1994
- R. Jain, S. Kalyanraman, and R. Viswanathan, "Ordered BECN: Why we need a timestamp or sequence number in the RM cell," AF-TM 94-0987, October 1994.
- R. Jain, S. Kalyanraman, and R. Viswanathan, "Simulation Results: The EPRCA+ Scheme," AF-TM 94-0988, October 1994.
- R. Jain, "Fairness: How to Measure Quantitatively," AF-TM 94-0881, September 1994.
- R. Jain, S. Kalyanraman, and R. Viswanathan, "Rate-based Schemes: Mistakes to Avoid," AF-TM 94-0882, September 1994.

- R. Jain, S. Kalyanraman, and R. Viswanathan, "The OSU Scheme for Congestion Avoidance using Explicit Rate Indication," AF-TM 94-0883, September 1994.

Our Papers

- R. Jain, "Congestion Control in ATM Networks: Recent Advances and A Survey," Invited submission to Computer Networks and ISDN Systems, February 1995.
- A. Charny, D. Clark, and R. Jain, "Congestion Control with Explicit Rate Indication," Proc. ICC'95, June 1995.
- R. Jain, "Myths about Congestion Management in High Speed Networks," Internetworking: Research and Experience, Volume 3, 1992, pp. 101-113.
- R. Jain, "Congestion Control in Computer Networks: Trends and Issues," IEEE Network, May 1990, pp. 24-30.
- K. Ramakrishnan and R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks with Connectionless Network Layer," ACM Transactions on Computers, May 1990.

- R. Jain, "A Delay Based Approach for Congestion Avoidance in Interconnected Heterogeneous Computer Networks," Computer Communications Review, ACM SIGCOMM, October 1989, pp. 56-71.
- D. Chiu and R. Jain, "Analysis of the Increase/Decrease Algorithms for Congestion Avoidance in Computer Networks," Journal of Computer Networks and ISDN, Vol. 17, No. 1, June 1989, pp. 1-14.
- R. Jain, K. Ramakrishnan, D. Chiu, "Congestion Avoidance in Computer Networks with a Connectionless Network Layer," DEC-TR-506, October 1988.

- K.K. Ramakrishnan, D. Chiu, and R. Jain, "Congestion Avoidance in Computer Networks with a Connectionless Network Layer, Part IV: A Selective Binary Feedback Scheme for General Topologies," DEC-TR-510, August 1987, 40 pp.
- R. Jain, "A Timeout Based Congestion Control Scheme for Window Flow-Controlled Networks," IEEE JSAC, Vol. SAC-4, No. 7, October 1986, pp. 1162-1167.
- R. Jain, D. Chiu, and W. Hawe, "A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems," DEC-TR-301, September 1984.