Congestion Control with Explicit Rate Indication

Anna Charny
DEC
Littleton, MA

David Clark
MIT
Cambridge, MA

Raj Jain

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

Disclaimers

- This work was done by Anna Charny for her MS Thesis at MIT (MIT-TR-601, May 1994) supervised by Dave Clark and Raj Jain
- This presentation is not sponsored by Digital Equipment Corporation
- The proposed scheme is a possible variation of the rate based approach
- It is being presented in support of the rate based approach.

- Why bit-scheme in 1984
- Why explicit rate indication in 1994
- The Scheme
- Simulation Results

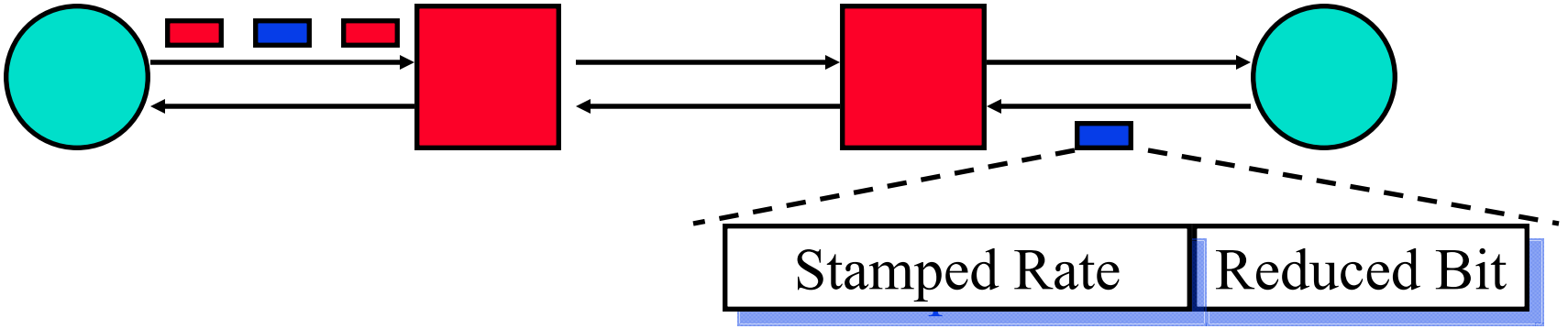
Why Bit Indication?

- Bit $\Rightarrow$ Up or down
- Connectionless networks
  $\Rightarrow$ No knowledge of flows or their demands
- 1984: Big shortage of bits in header
- No new packets
- 1984: No better congestion schemes

Why Explicit Rate Indication?

- Connection oriented networks
  - ⇒ Switches know "who's who"
  - ⇒ More predictability of paths
- Longer-distance networks
  - ⇒ Can't afford too many round-trips
  - ⇒ More information is better
- Rate-based control
  - ⇒ Queue length = ΔRate ´ ΔTime
  - ⇒ Time is more critical than with windows

The Scheme



- Sources send one RM cell every n cells
- The RM cells contain "Stamped (desired) rate" and a "reduced-bit"
- The switches adjust the rate down and sets the reduced bit
- Destination returns the RM cell to the source
- Source adjusts to the specified rate

Source Algorithm

- Always follow the network's specified "stamped rate"
- If reduced bit is set in returned RM Cell
  - Decrease to the rate specified
- If reduced bit is clear in returned RM Cell
  - send a higher rate in "stamped rate" field
  - Increase to the rate returned

Destination Algorithm

- Return all RM cells to the source

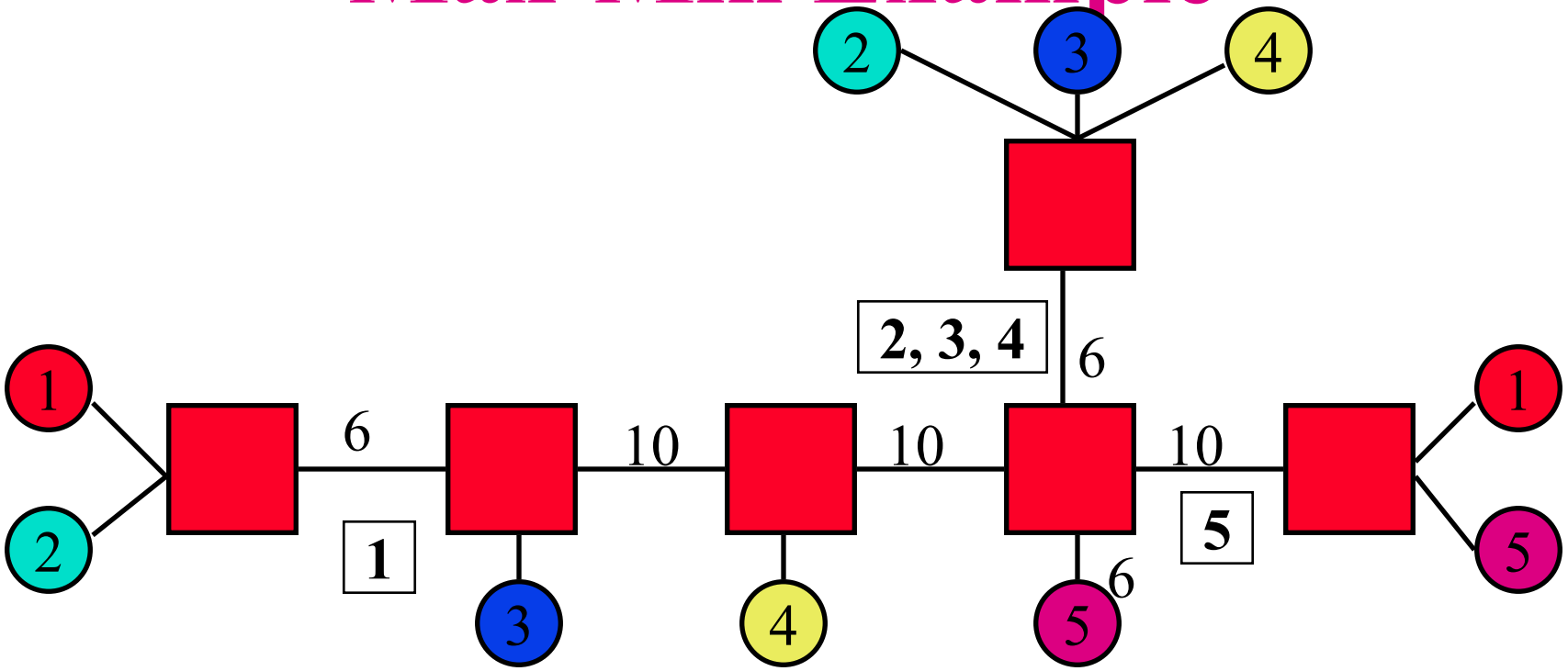
Switch Algorithm

- Optimally allocate available capacity among all VC's
- Optimal =
  - Most money for the provider
  - Most throughput for the link
  - Most power (=Throughput/Delay) for link
  - Max-min Fair allocation

Max-Min Fair Allocation

- At it's bottleneck,
  every VC gets its maximum fair-share.
- Every link is maximally utilized.
- R_{ij} = Rate of ith VC on jth link
- $R_i = R_{ij} = R_i$.
- $\Sigma_j R_{ij} \leq C_j$ = Capacity of the jth link
- At ith VC bottleneck:
  - Let k = # of VC's, $R_i \geq C/k$

Max-Min Example



- Optimal Rates: Flows 2,3,4: 2 each
  Flow 1: 4
  Flow 5: 6

A Sample Switch Algorithm (for Max-Min Optimality)

- Switches compute an "advertised rate"
- RM cells with "stamped ≤ advertised" rate are not touched
- In RM cells with "stamped > advertised" rate, stamped rate is reduced to the advertised rate and reduced bit is set.

Computing Advertised Rate

- Advertised Rate = Capacity/number of VCs
- Underloading VC = Rate < advertised
3. Advertised rate = $\frac{\text{Capacity} - \Sigma \text{ BW of Underloading VCs}}{\text{\# of flows - \# of underloading flows}}$
- If change, go to Step 2

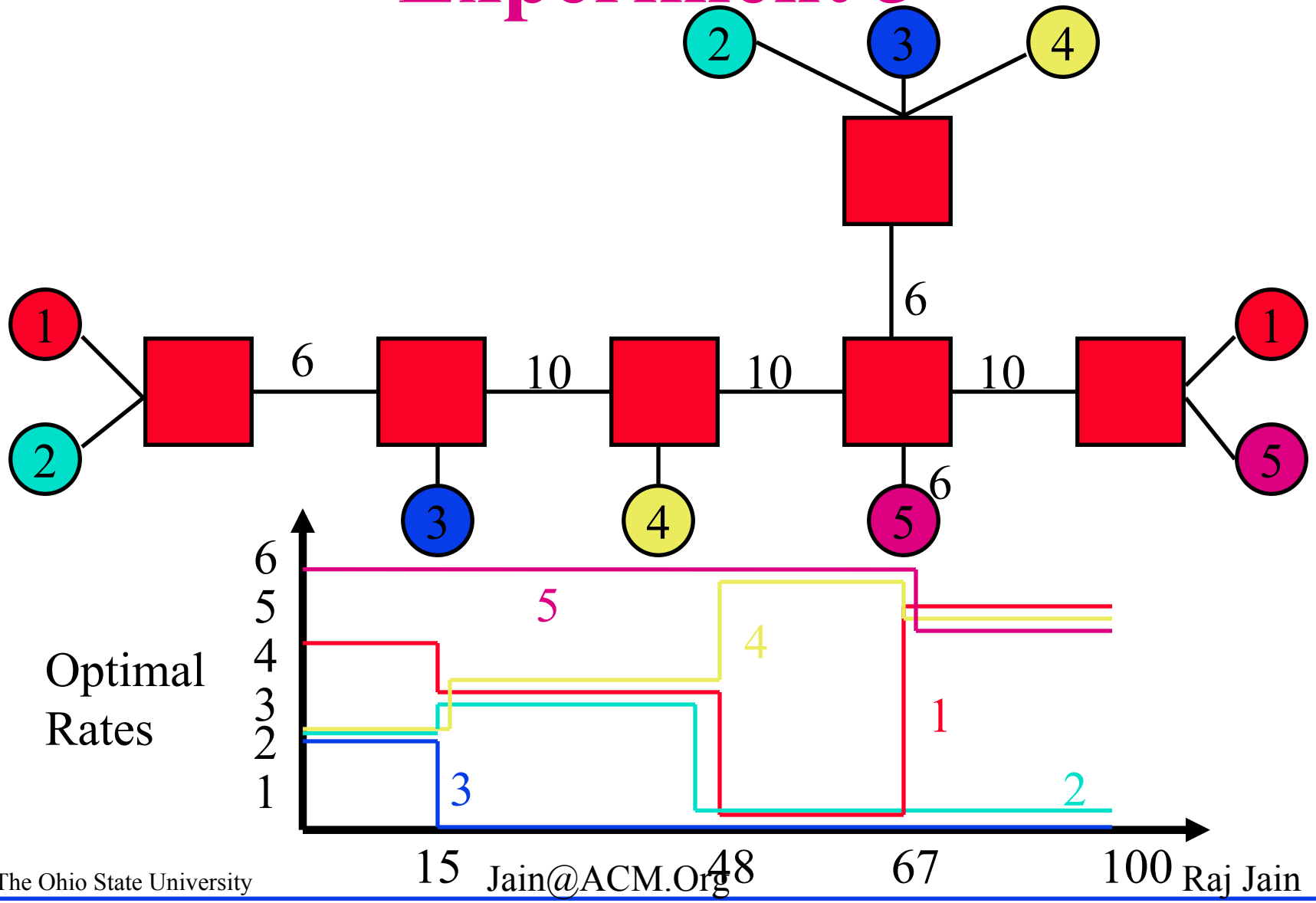
Two iterations are sufficient.

Switches keep a table of stamped rates of all VCs

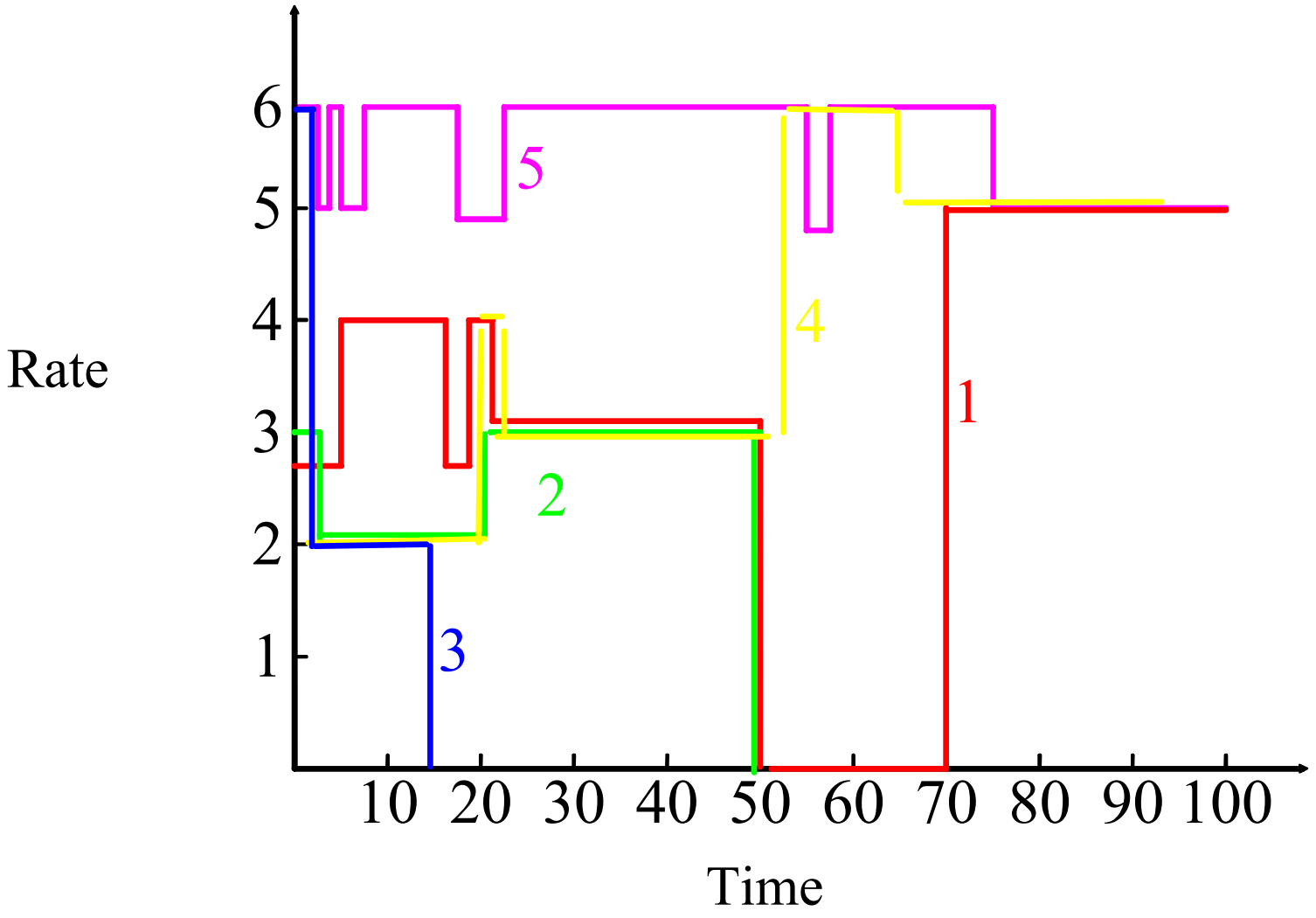
Properties of Scheme

- No guessing of level of overload/underload
- No oscillations
- Convergence within $4k$ round trips where k is the number of bottlenecks
- Initial rate doesn't matter
- Policing is trivial.
  Switches can monitor returning RM cells
- Designed for connection-oriented networks
- Robust to RM cells loss or errors

Experiment 5



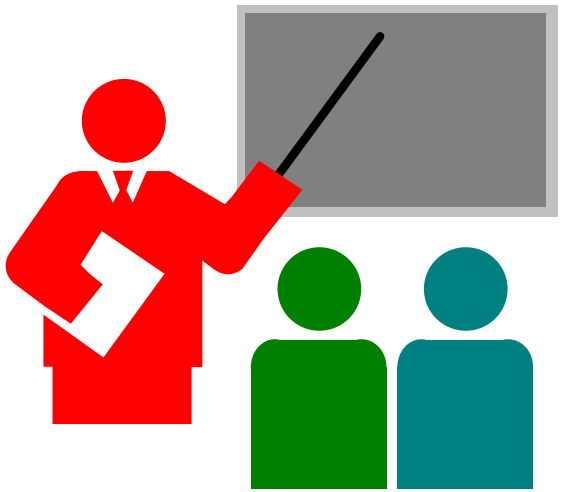
Simulation Results



Round trips for Convergence

Time Session	0	15	48	67
1	4	5		2
2	2	2		
3	2			
4	1	3	2	1
5	7	6	3	1

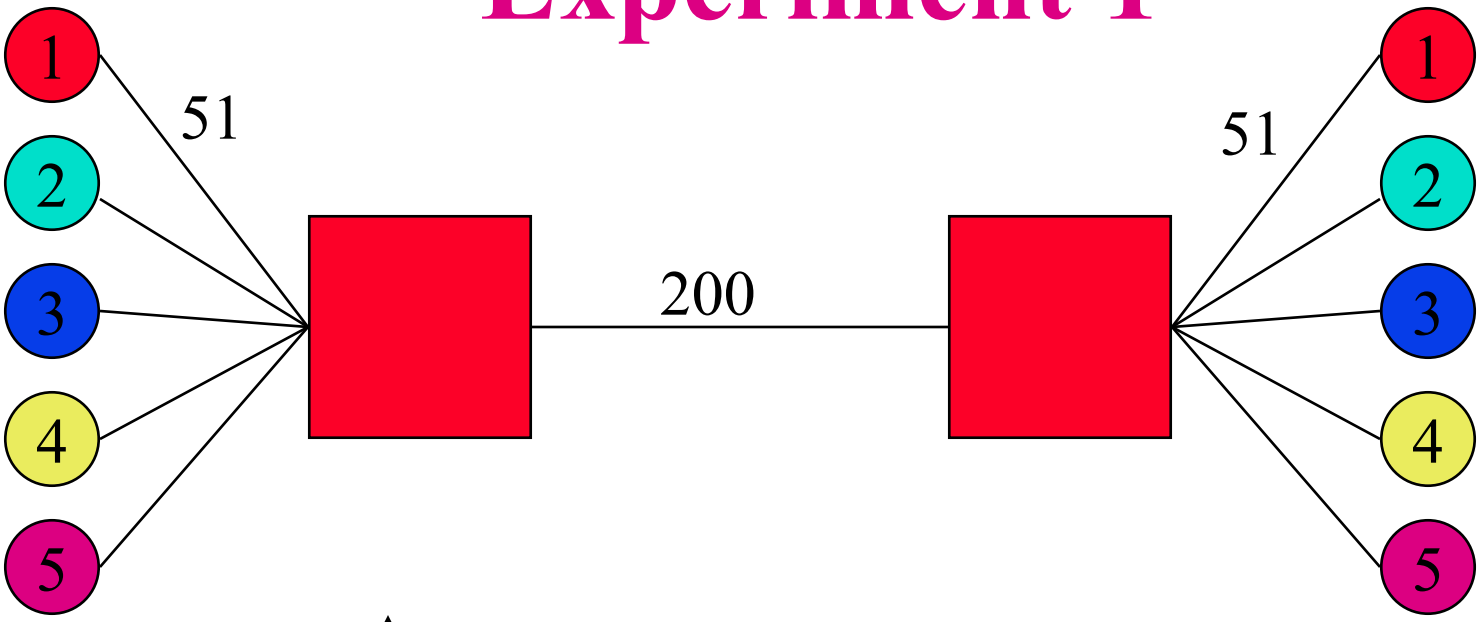
Summary

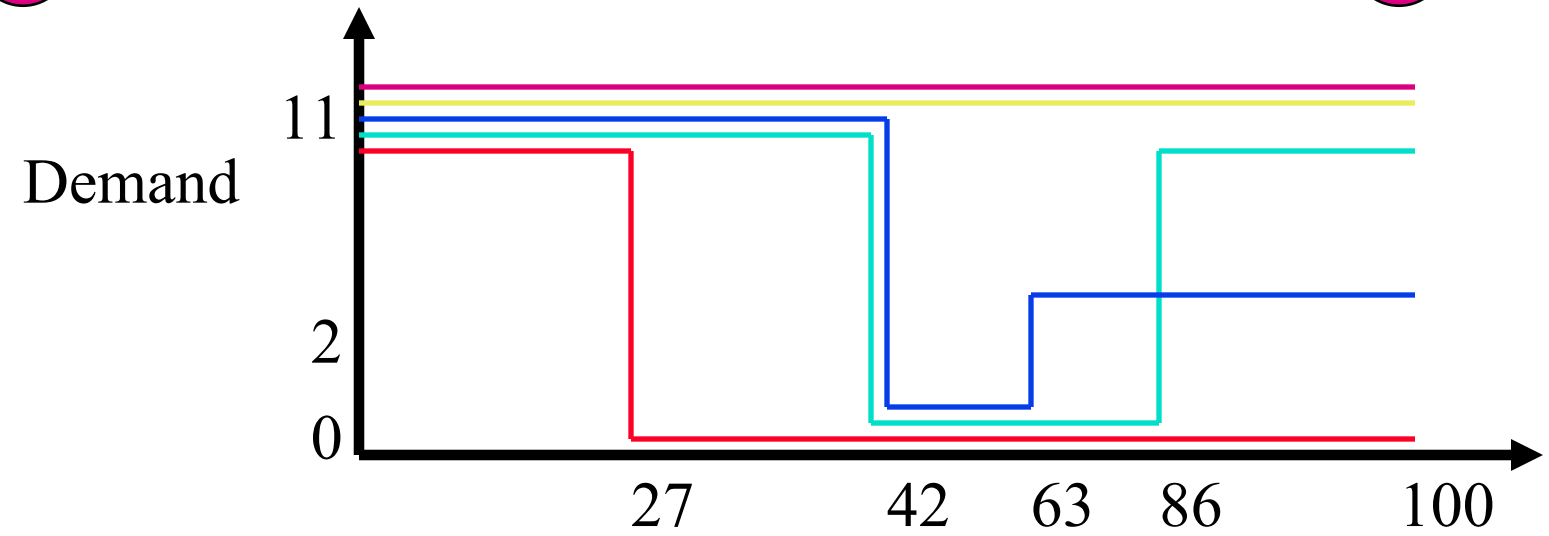


Explicit rate indication

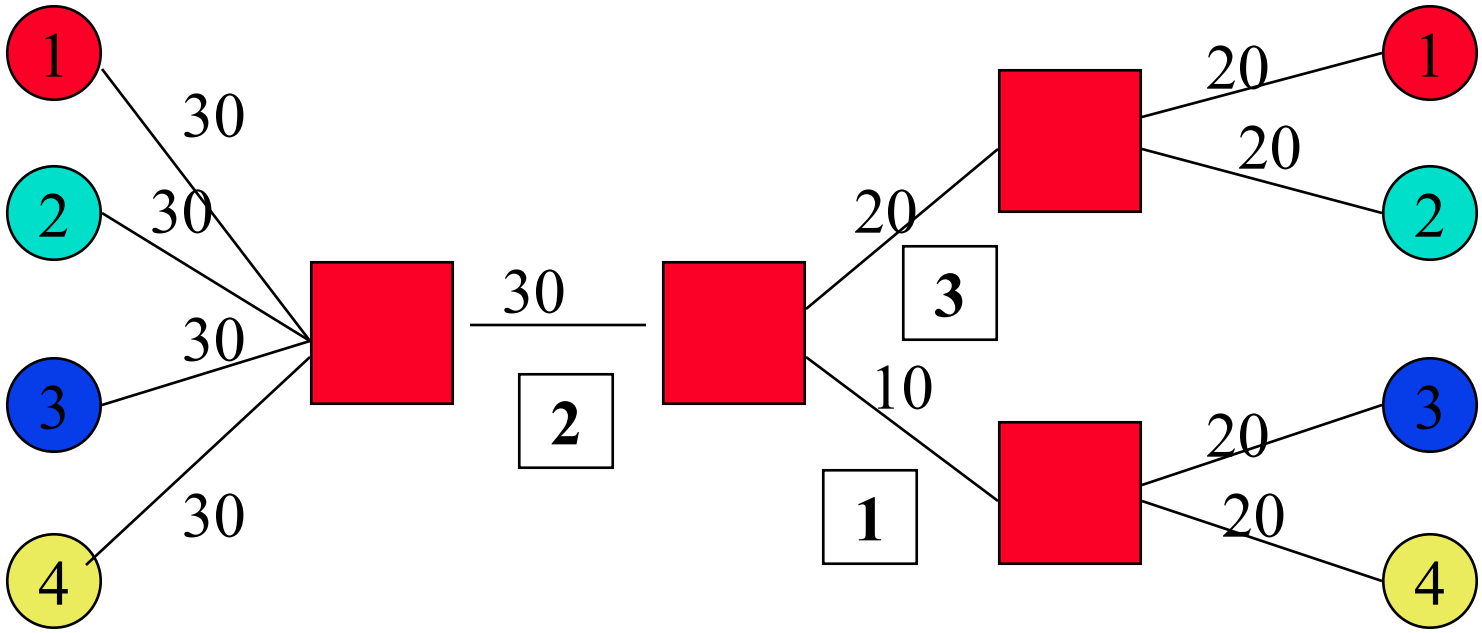
- Provides more information than a single bit
- Converges fast
- Provides a choice of switch optimality criteria
- Easy to police

Experiment 1





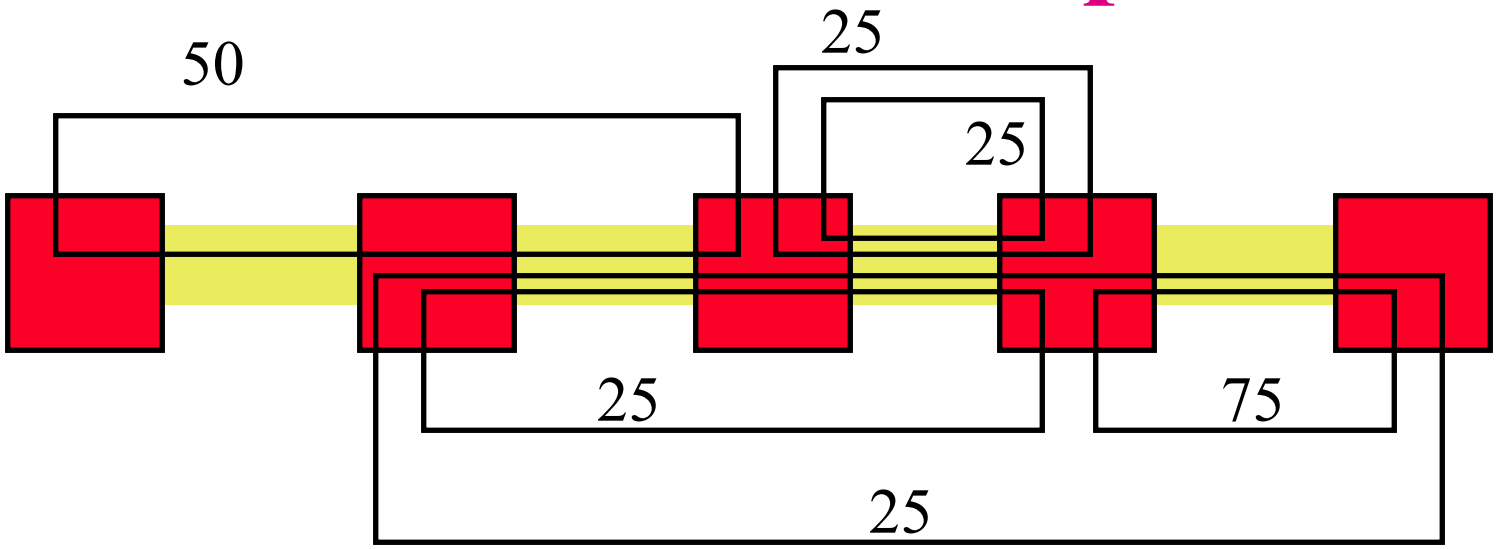
Experiment 3



Optimal Rates for Expt 5

Time Session	0-15	15-48	48-67	67-100
1	4000	3000		5000
2	2000	3000		
3	2000			
4	2000	3000	6000	5000
5	6000	6000	6000	5000

Max-Min Example



- All links at 100 Mbps
- Six VC's at rates 50, 25, 25, 25, 25, 75
- All links are maximally utilized:
  50, 100, 100, 100