Scalable QoS Provision Through Buffer Management

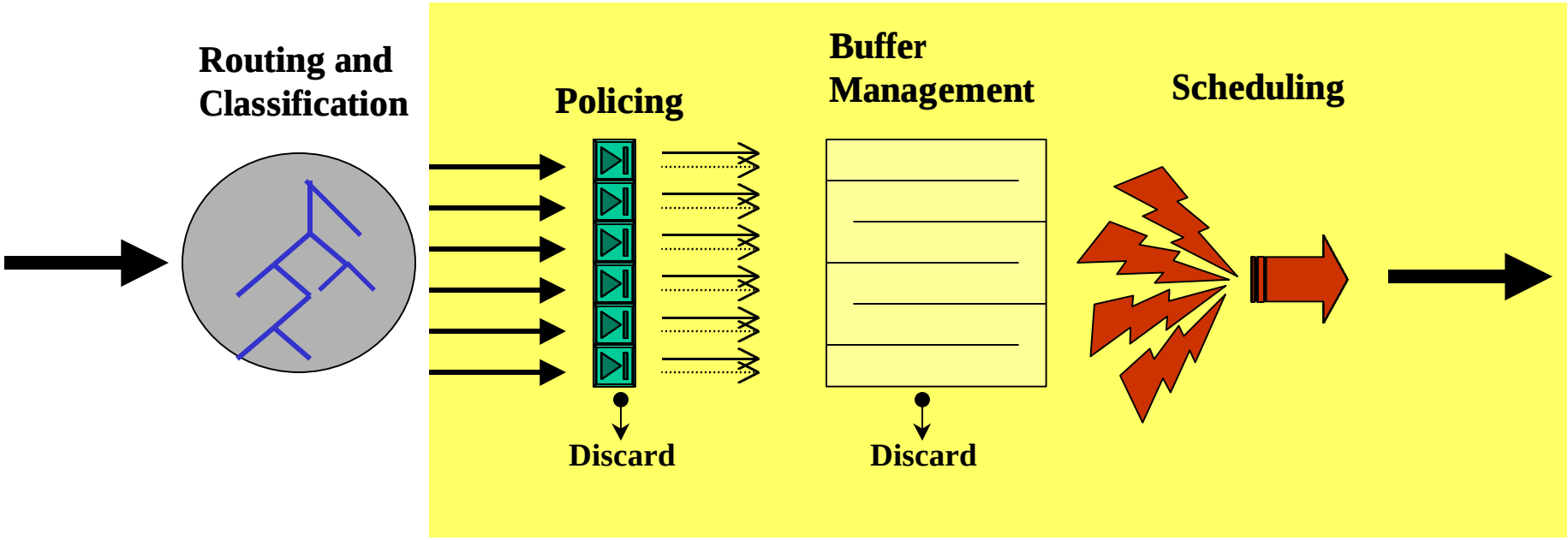
Roch Guerin
Sanjay Kamat
Vinod Peris
Raju Rajan

IBM Thomas J. Watson Research Labs

Outline

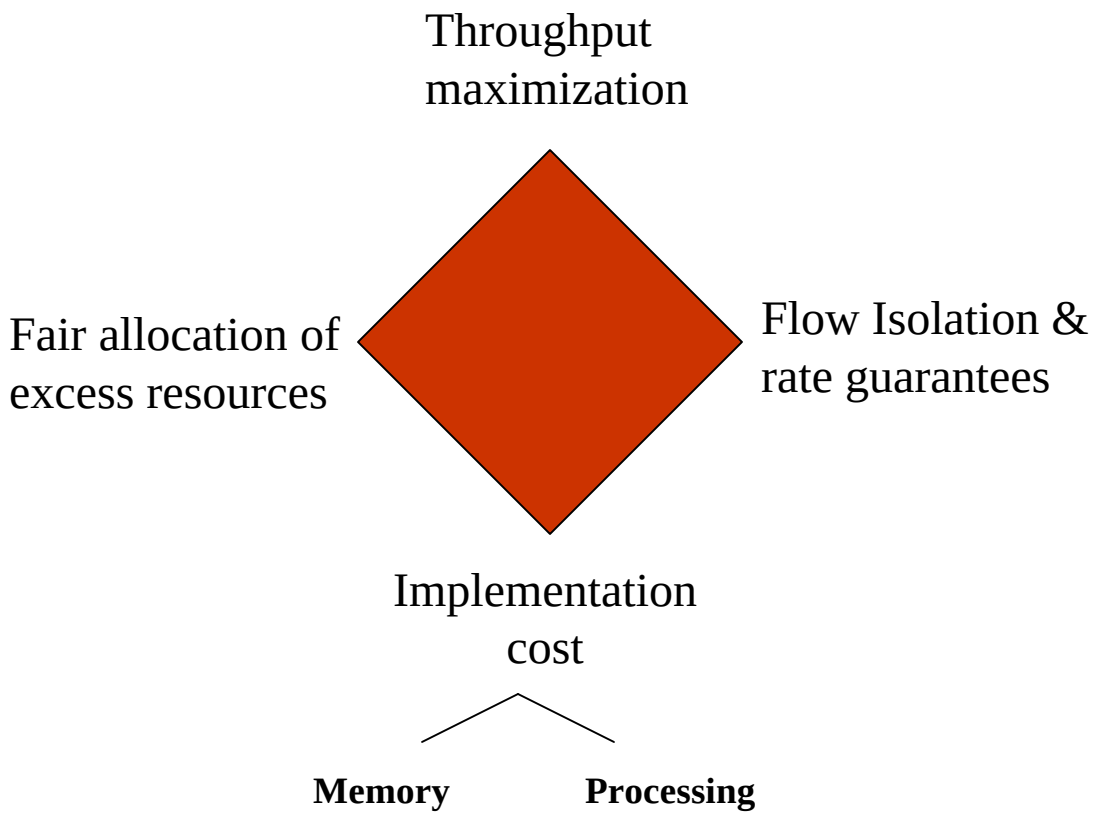
- Differentiated Packet Treatment for QoS
- Performance Objectives
- Design Space
  - Scheduling and Buffer Management Schemes
- Comparing FIFO vs WFQ -- Worst case buffer tradeoffs
- Examining tradeoffs with strict buffer partitioning
- Examining tradeoffs with buffer sharing
- Hybrid Schemes
- Conclusions

Differentiated Packet Treatment



- Flows (Unit of service guarantee -- varying granularity)
- QoS Resources -- Link capacity and Buffer Space
- Scalability
  - Processing time per packet versus number of flows
  - State size versus number of flows

QoS Performance Objectives



QoS Allocation Schemes

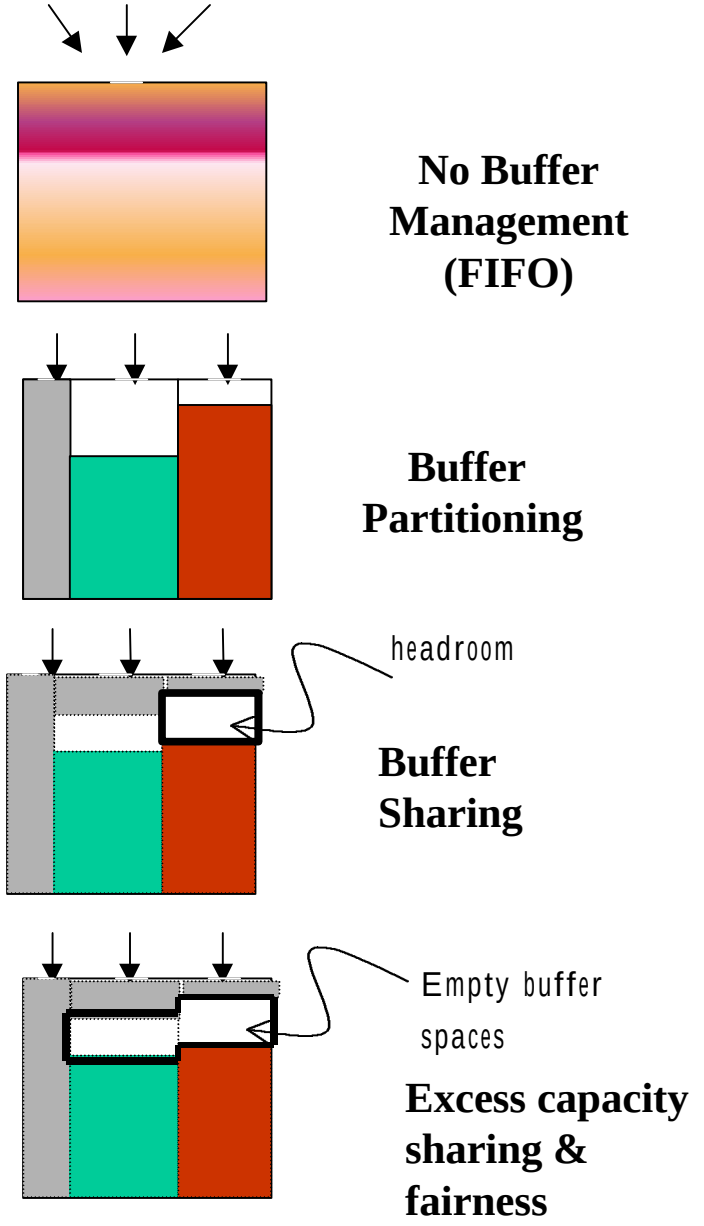
•

Scheduling

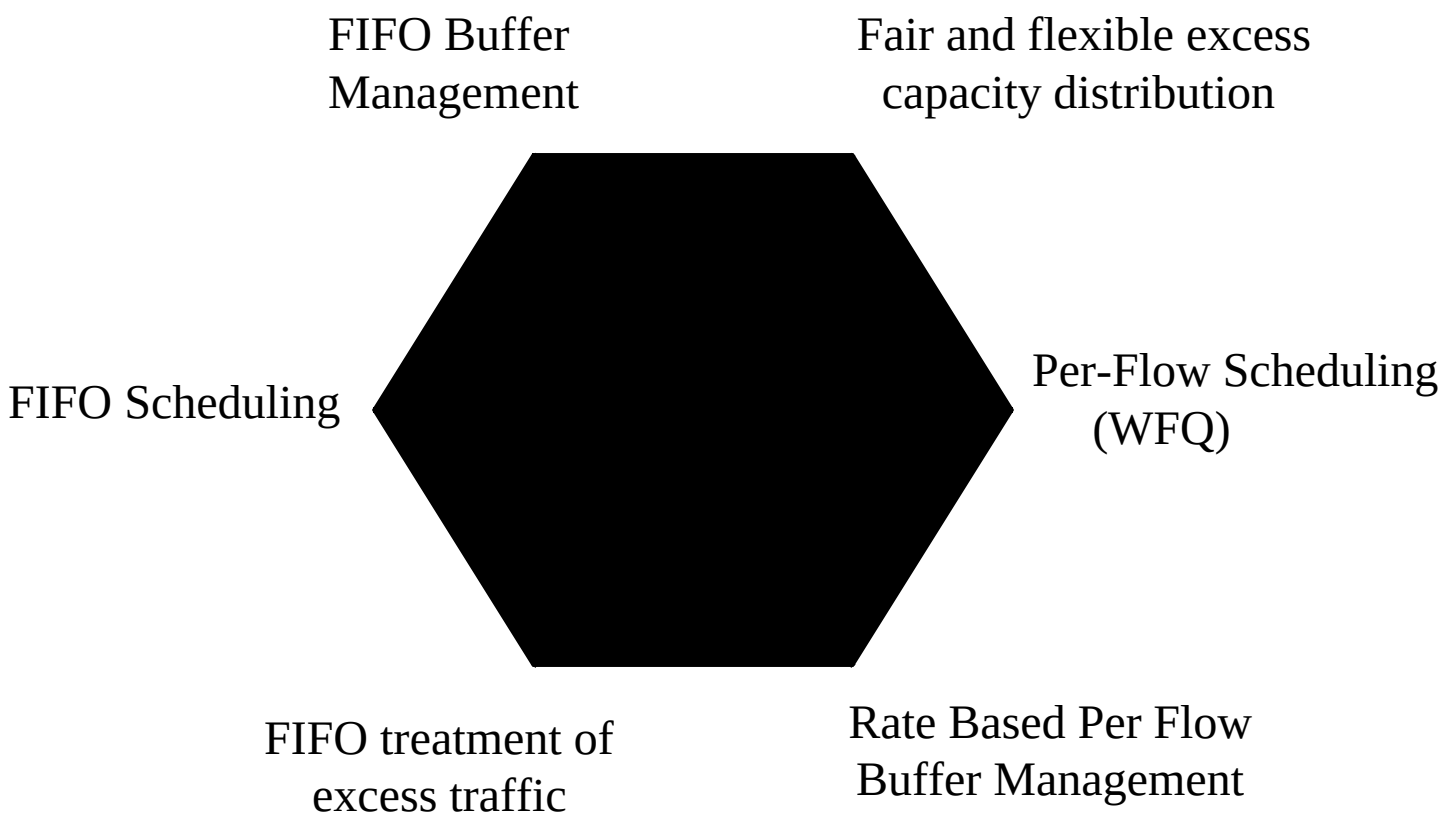
- FIFO
- WFQ

Buffer Management

- No Buffer Management (FIFO)
- Buffer Partitioning
- Buffer Sharing
- Fair Excess Capacity distribution
  (Choudhury and Hahne)

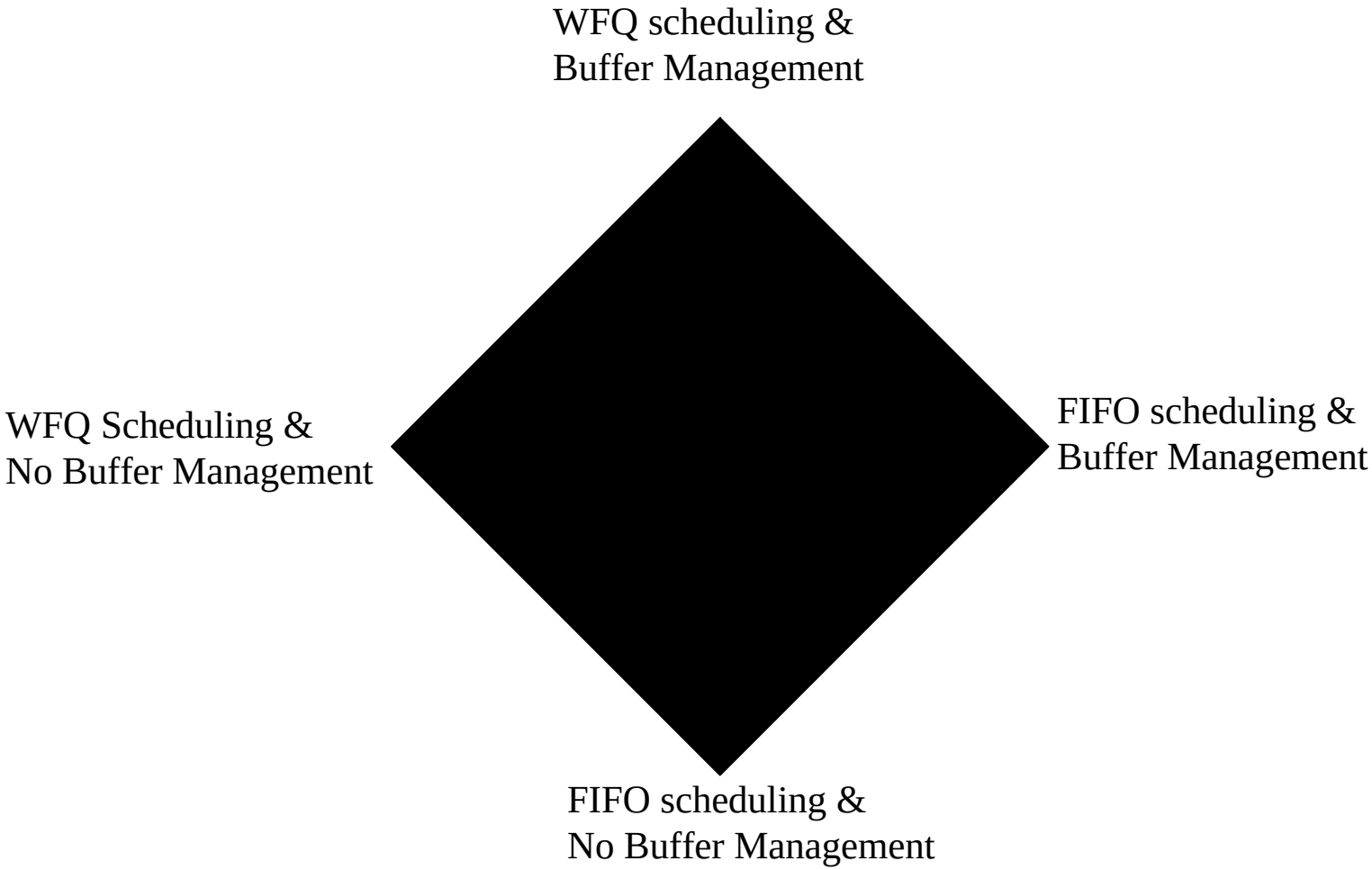


Design Space



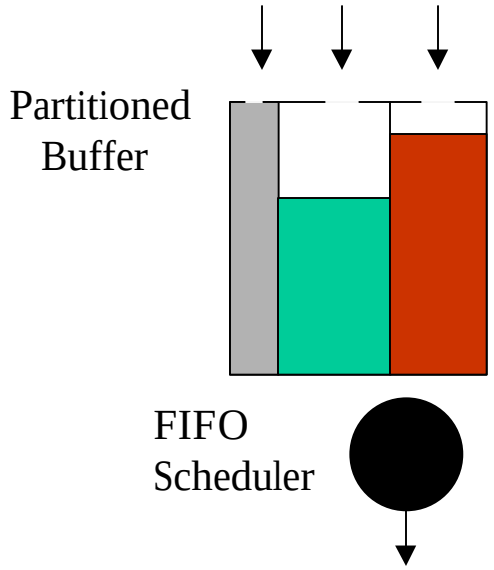
Benchmarks

.


WFQ scheduling &
Buffer Management
WFQ Scheduling &
No Buffer Management
FIFO scheduling &
Buffer Management
FIFO scheduling &
No Buffer Management

Rate Guarantees through buffer management alone

Reservation	**Buffer Partition Size**	**Offered Load**	**Service Achieved**
ρ bits/sec	Bρ/R bits	Conformant	Lossless
		Non-conformant	Losses limited by non-conformance
(σ bits, ρ bits/sec)	σ + Bρ/R bits	Conformant	Lossless
		Non-conformant	Losses limited by non-conformance



Buffer Size B bits
Link Capacity R bits/sec

Buffer Partitioning w. FIFO vs. WFQ

(Worst case comparison)

WFQ Admissibility Checks

$R \geqq \Sigma\rho_i$ Bandwidth availability

$B \geqq \Sigma\sigma_i$ Buffer availability

FIFO Admissibility Checks

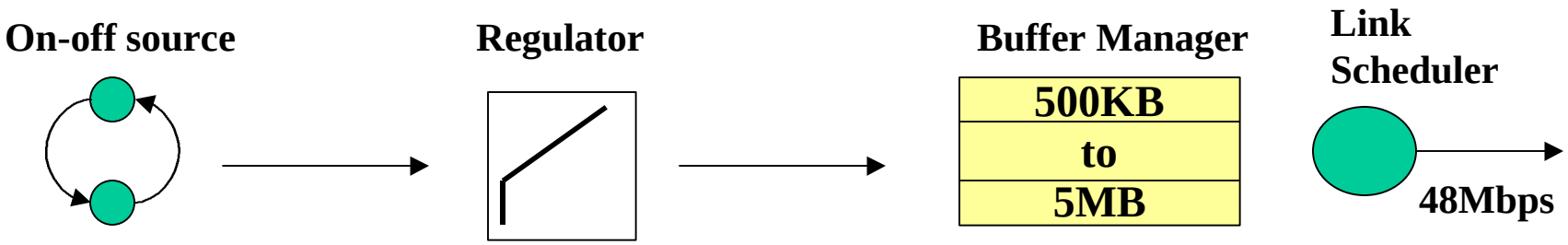
$R \geqq \Sigma\rho_i$ Bandwidth availability

$B \geqq \Sigma\sigma_i + (\Sigma\rho_i)\, B/R$ Buffer availability

Equivalently

$B \geqq \Sigma\sigma_i /(1\text{-utilization})$

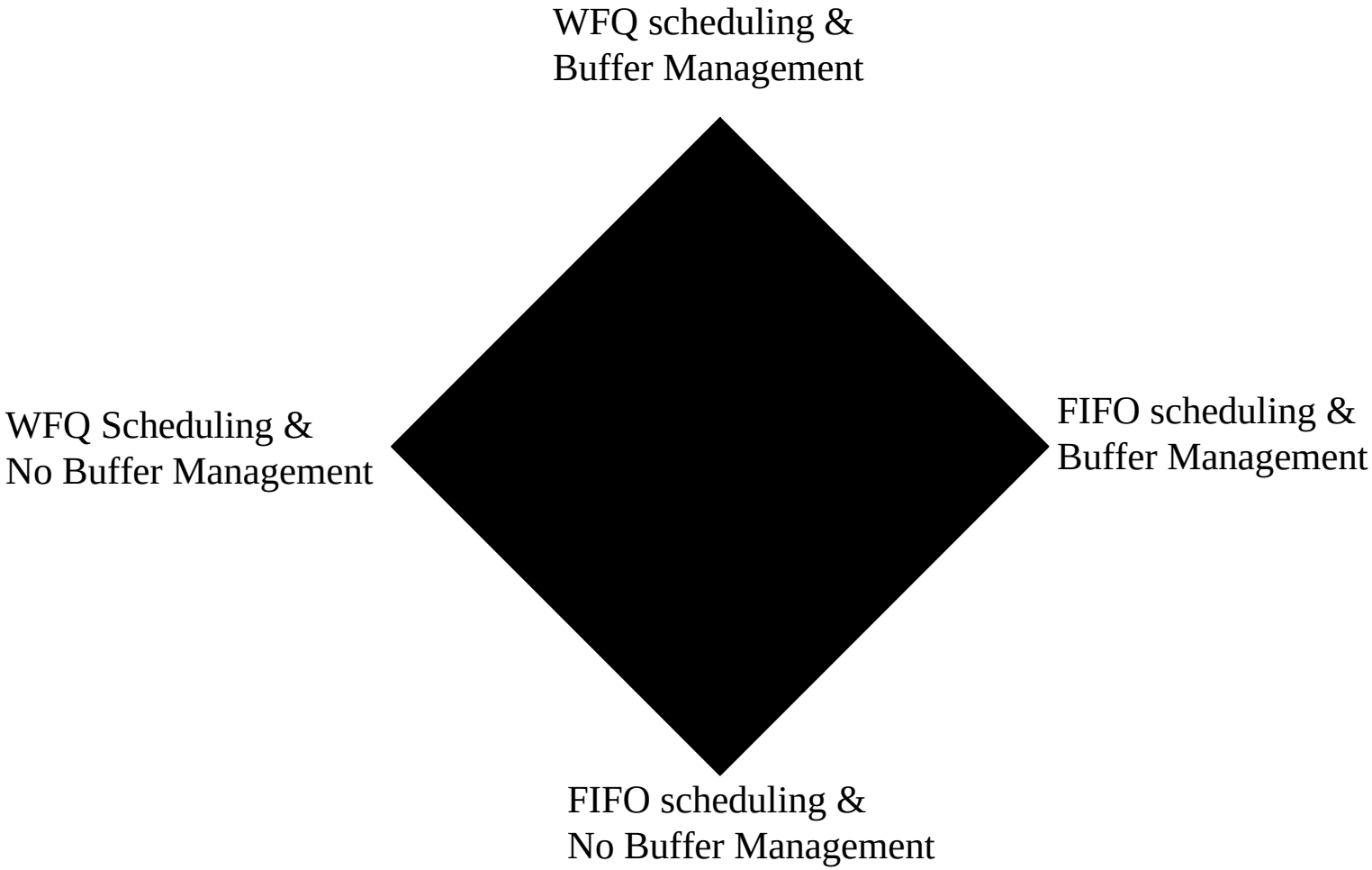
Experimental Setup



Flow Number	Token bucket	Token Rate	Mean Rate	Peak Rate
0	50 kB	2Mbps	2Mbps	16Mbps
1	50 kB	2Mbps	2Mbps	16Mbps
2	50 kB	2Mbps	2Mbps	16Mbps
3	100 kB	8Mbps	8Mbps	40Mbps
4	100 kB	8Mbps	8Mbps	40Mbps
5	100 kB	8Mbps	8Mbps	40Mbps
6	50 kB	0.4Mbps	4Mbps	40Mbps
7	50 kB	0.4Mbps	4Mbps	40Mbps
8	50 kB	2Mbps	16Mbps	40Mbps

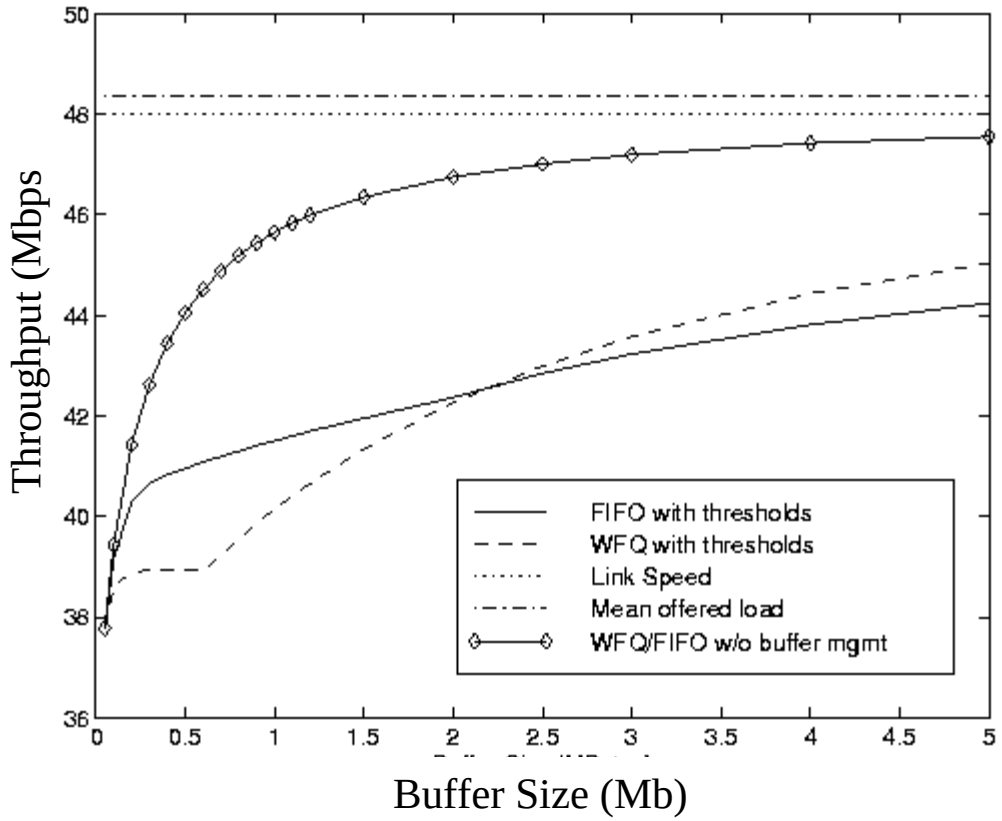
Benchmarks

.


WFQ scheduling &
Buffer Management
WFQ Scheduling &
No Buffer Management
FIFO scheduling &
Buffer Management
FIFO scheduling &
No Buffer Management

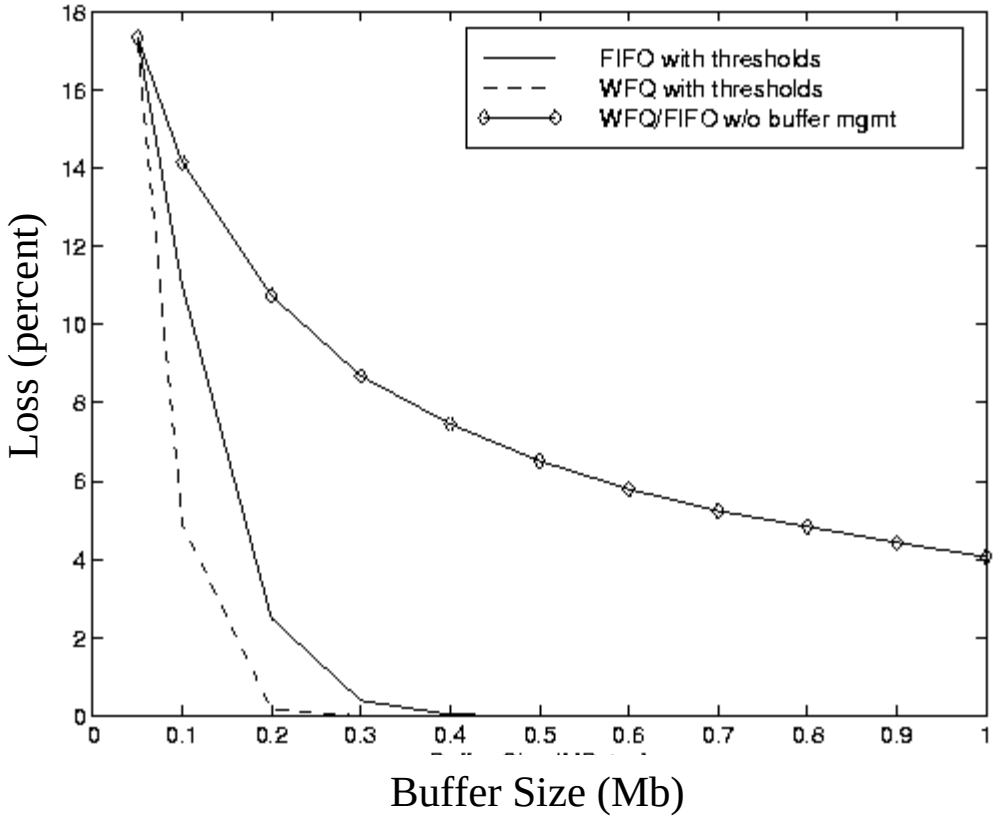
Aggregate Throughput

Buffer Partitioning vs. No Buffer Management

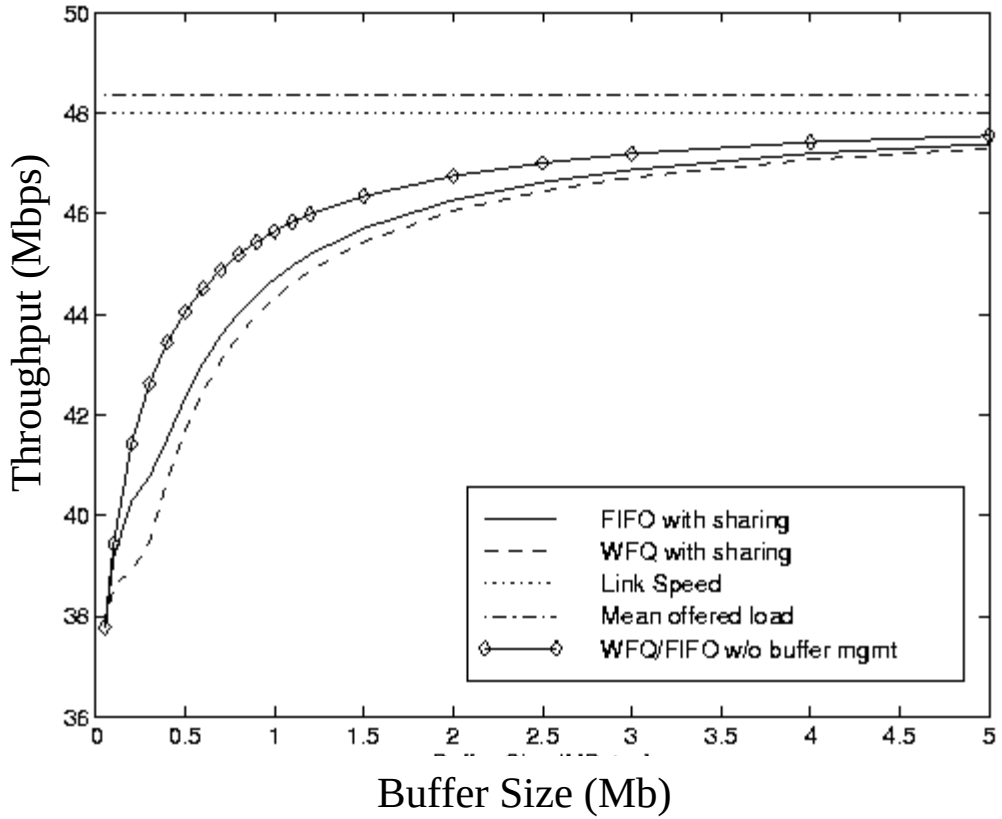


Losses for conformant flows

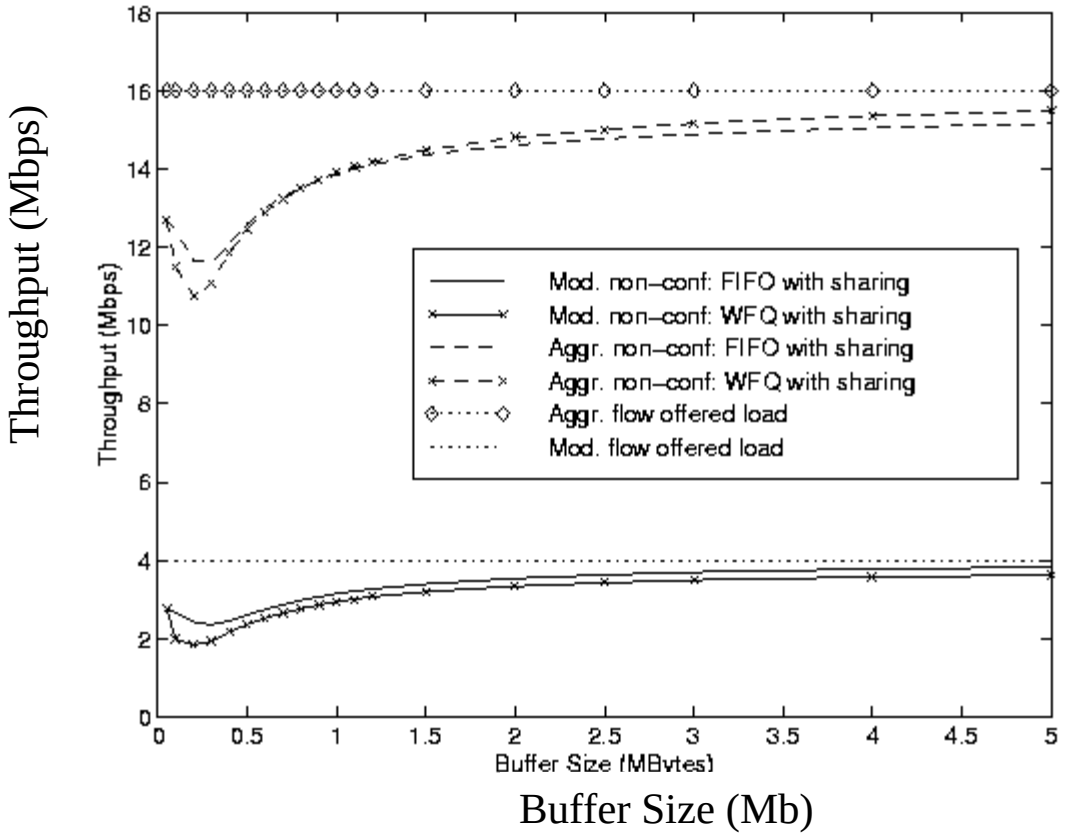
Buffer Partitioning vs. No Buffer Management



Aggregate Throughput
Buffer Sharing vs. No Buffer Management

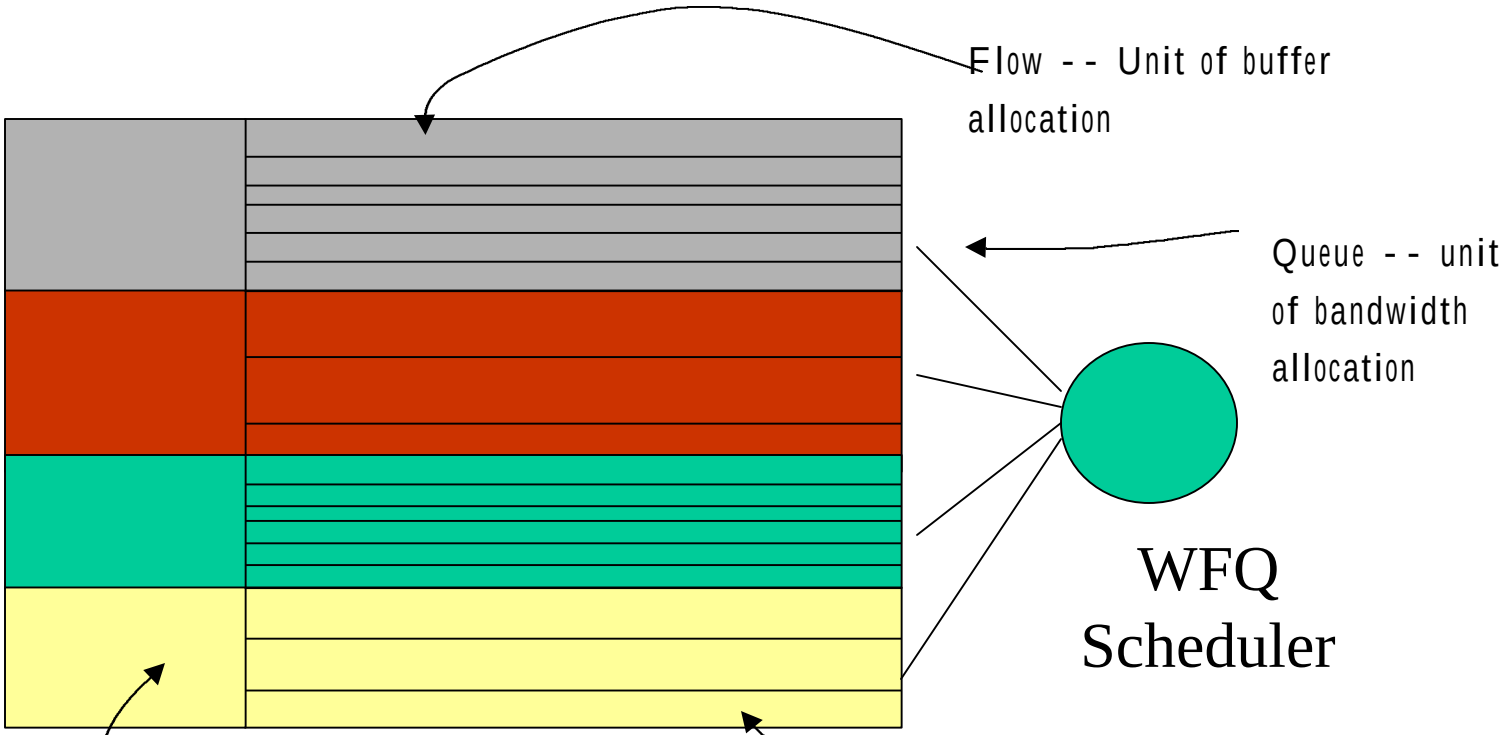


Excess Capacity Sharing



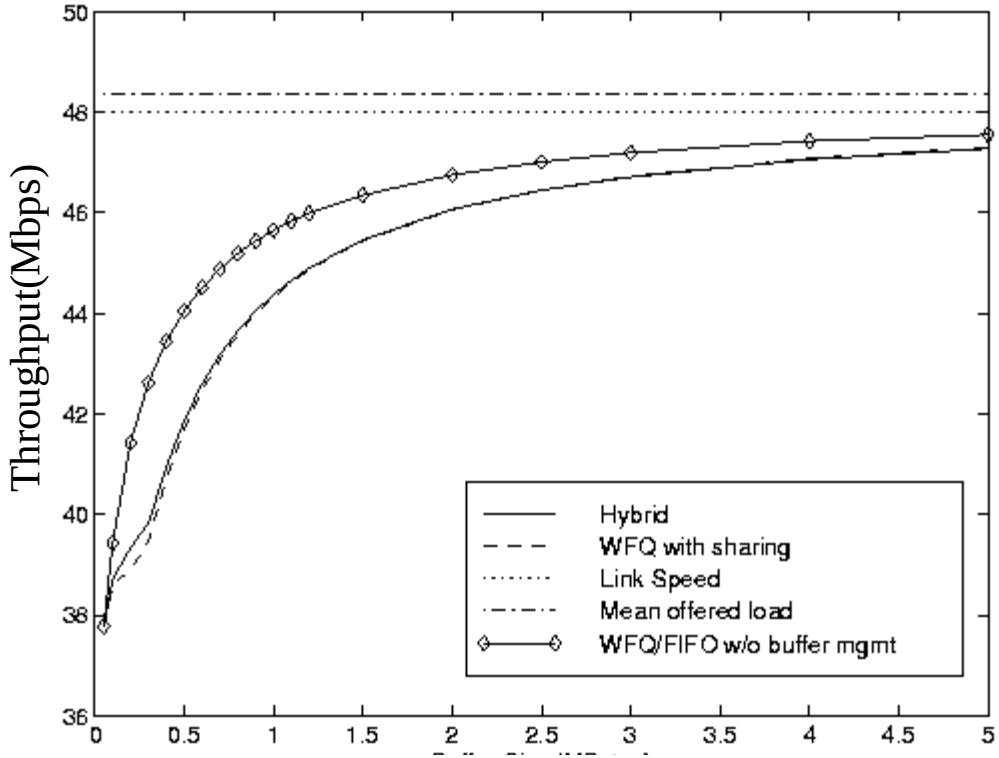
Buffer Size (Mb)

The Hybrid Scheme



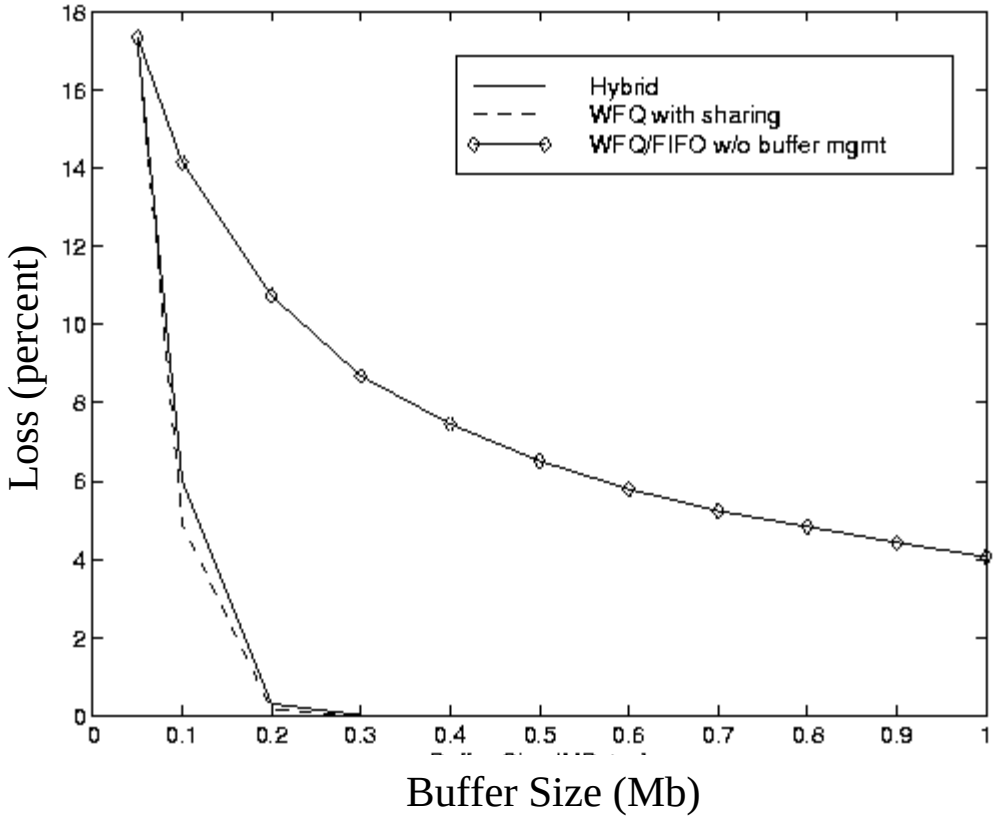
Throughput

Integrated Scheme



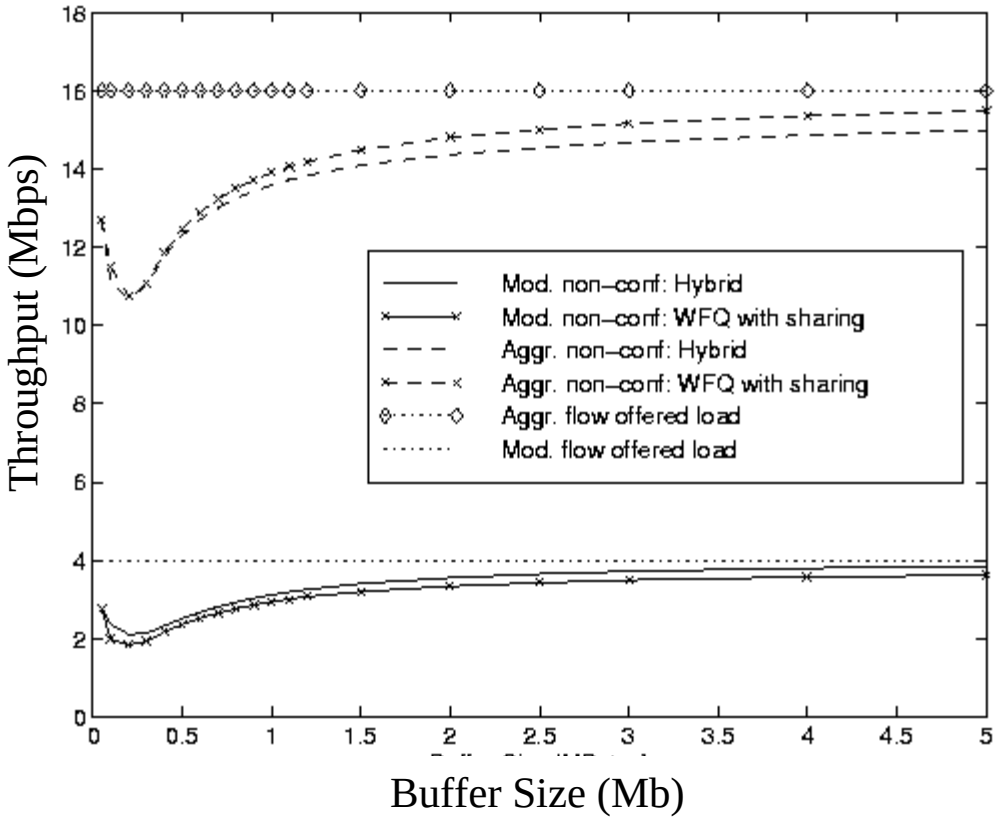
Flow Isolation and Rate Guarantees

Integrated Scheme



Excess Capacity Distribution

Integrated Scheme



Conclusions

The Integrated Scheme

- Scheduling on a fixed number of "Queues" and buffer management on the finer granularity of "Flows"
  - Rules of thumb for flow grouping
- Buffer sharing & flow isolation within a single
  - use "headroom" to limit impact on conformant flows in buffer limited system
- Empty buffers (holes) to regulate the sharing of excess bandwidth
  - flexible notion of fairness

- Flexibility to choose tradeoffs depending on operational environment
- Some quantitative guidelines on the choice of design and setting of parameters