Robustness to Packet Reordering in High-speed Networks

Sumitha Bhandarkar & A. L. Narasimha Reddy

February 07, 2007

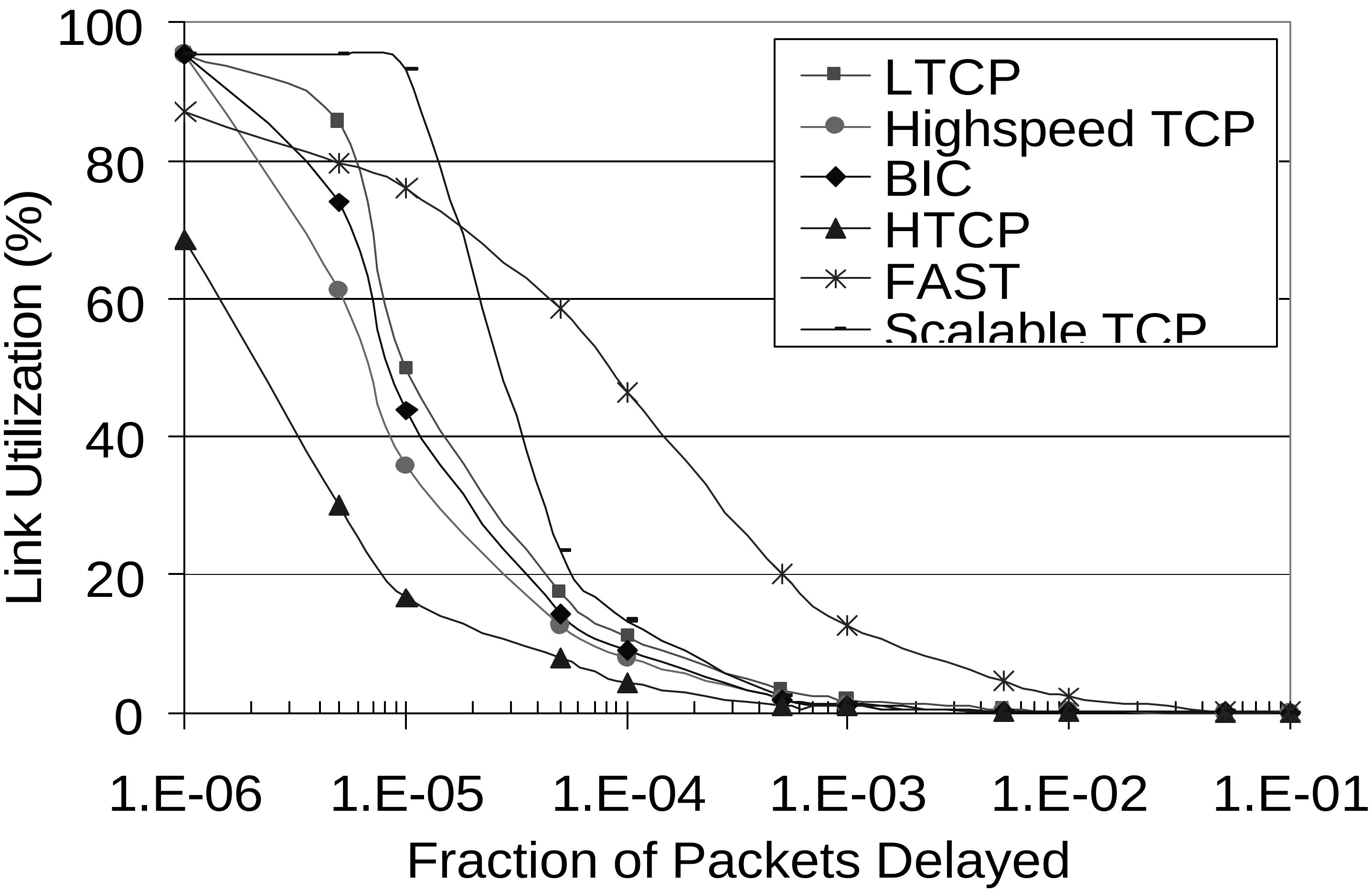
Motivation

- Studies show non-zero amount of packet reordering in the Internet
  - Parallelism in Internet components
    - multiple paths within the switching fabric
    - multi-path routing
    - multi-homing etc.
  - Misconfigured Internet components
  - Reordering may be function of network load
- Most of these causes for reordering are more likely in high speed networks

Motivation

- High speed protocols make probing efficient
  - Allow window to grow very large
- Still use 3-dupack heuristic for identifying loss
  - Extremely susceptible to reordering events
  - Severe degradation of throughput in case of reordering

Packet Reordering in Highspeed Networks



RTT of the flow : 50ms

Model for Packet Delay :

- Uniform distribution for choosing packet
- Normal distribution for delay with mean 25ms and stddev 8ms

Proposed Solution

- Use delayed congestion response with high speed protocols
  - RFC 4653 for TCP
  - How well does this work with high-speed protocols?

TCP-DCR : An Overview

- When dupacks are received,
  - Delay the time to infer congestion by τ
  - Essentially a tradeoff between wrongly inferring congestion and promptness of response to congestion
  - τ chosen to be one RTT to allow maximum time while avoiding an RTO

Highspeed Protocols : An Overview

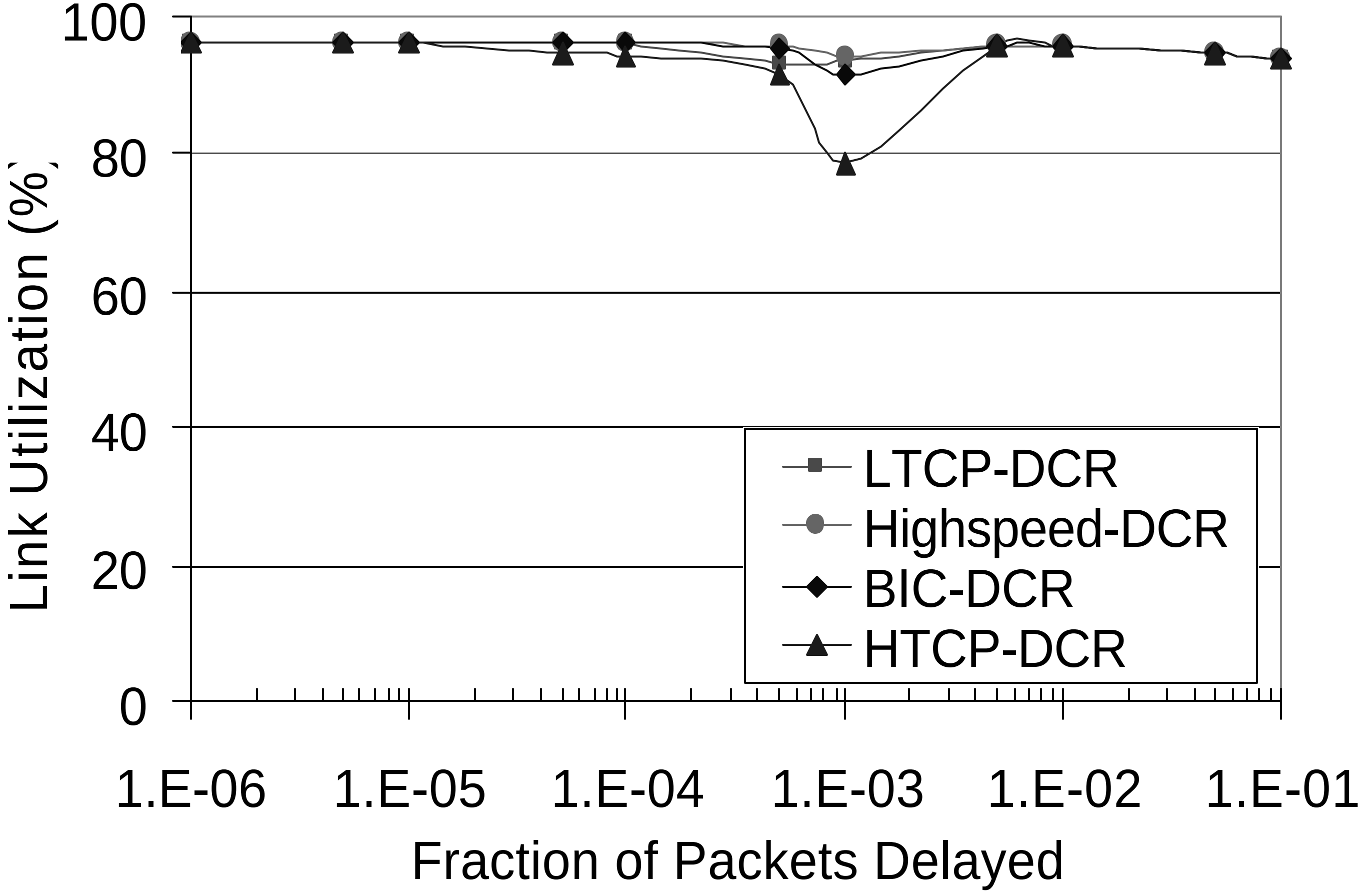
LTCP: Uses the concept of layering to ramp up the bandwidth quickly, while ensuring fairness.

Highspeed TCP: Modifies congestion response parameters based on different response function

BIC TCP: Uses Binary/Additive Increase, Multiplicative Decrease.

H-TCP: Modifies congestion response parameters based on “time since last drop”

Impact of Packet Reordering When DCR is Used



RTT of the flow : 50ms

Model for Packet Delay :

- Uniform distribution for choosing packet
- Normal distribution for delay with mean 25ms and stddev 8ms

Observations

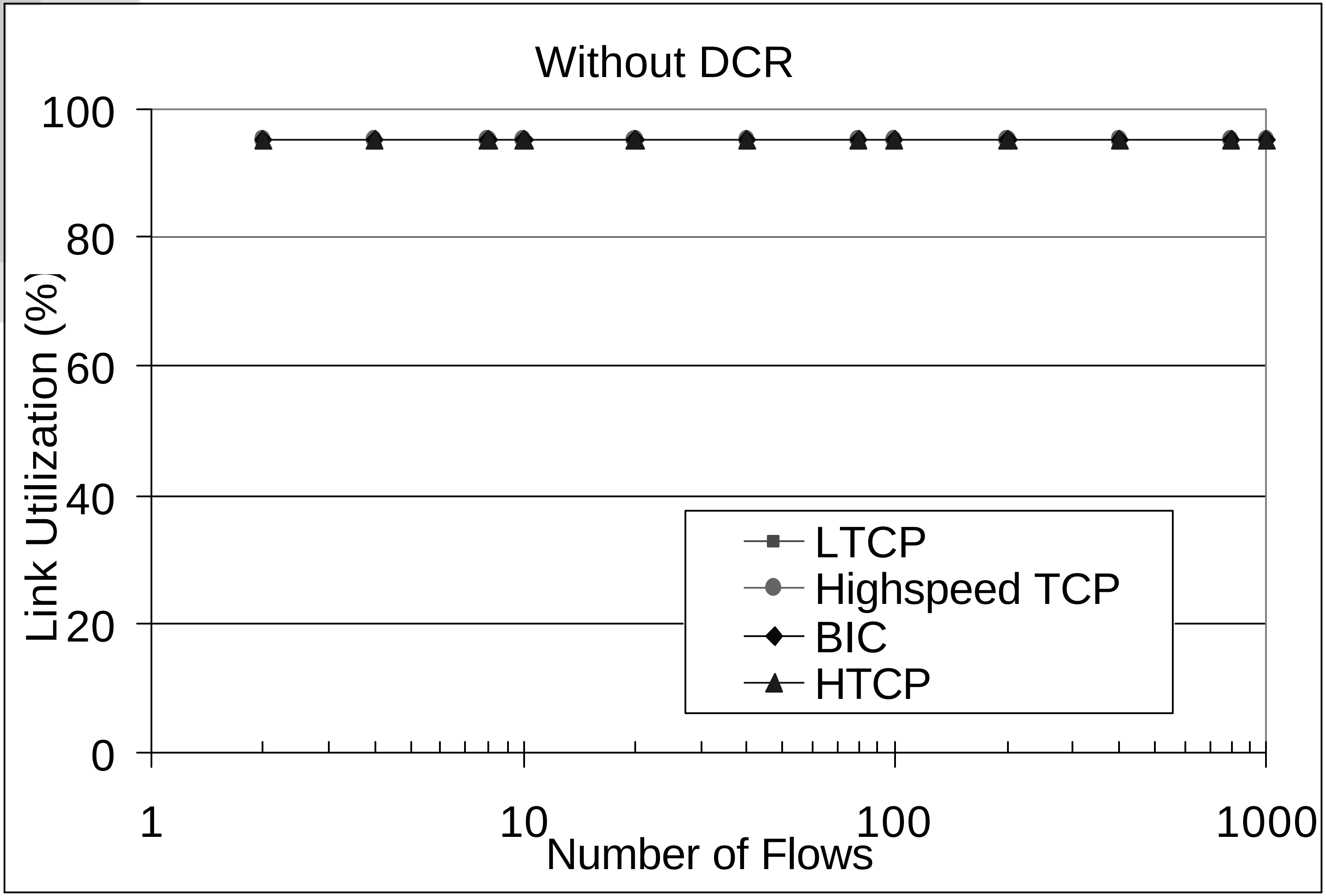
- Using DCR prevents performance degradation
  - Even at high levels of reordering
- DCR delays response by one RTT
  - Reordering of more than 1 RTT can still cause performance degradation

Other issues with DCR

- DCR should not impact other behavior
  - How does congestion impact delaying response?
  - How does delayed response impact drop rates?
  - How does delayed response impact fairness ?

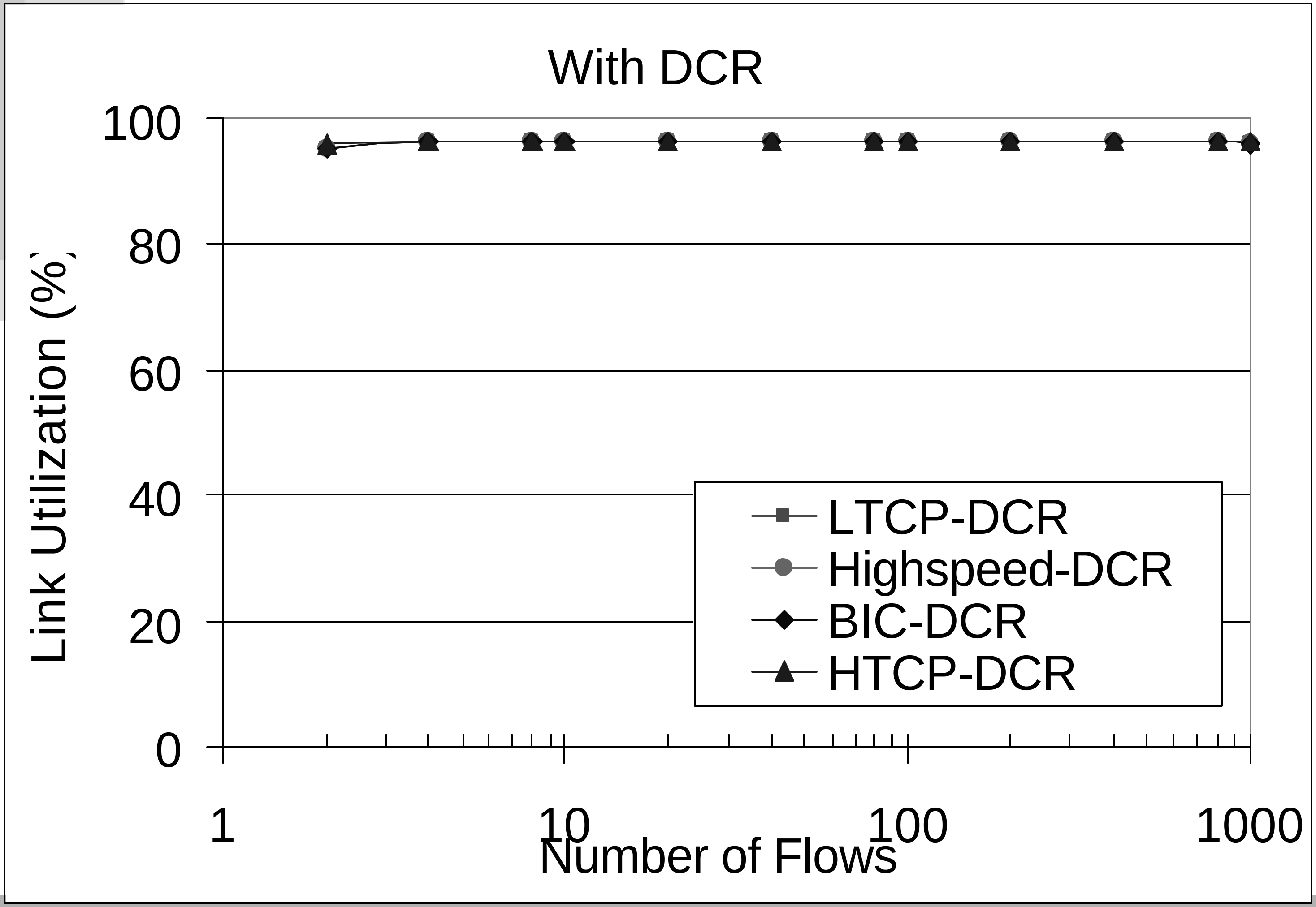
With Congestion Only

Link Utilization



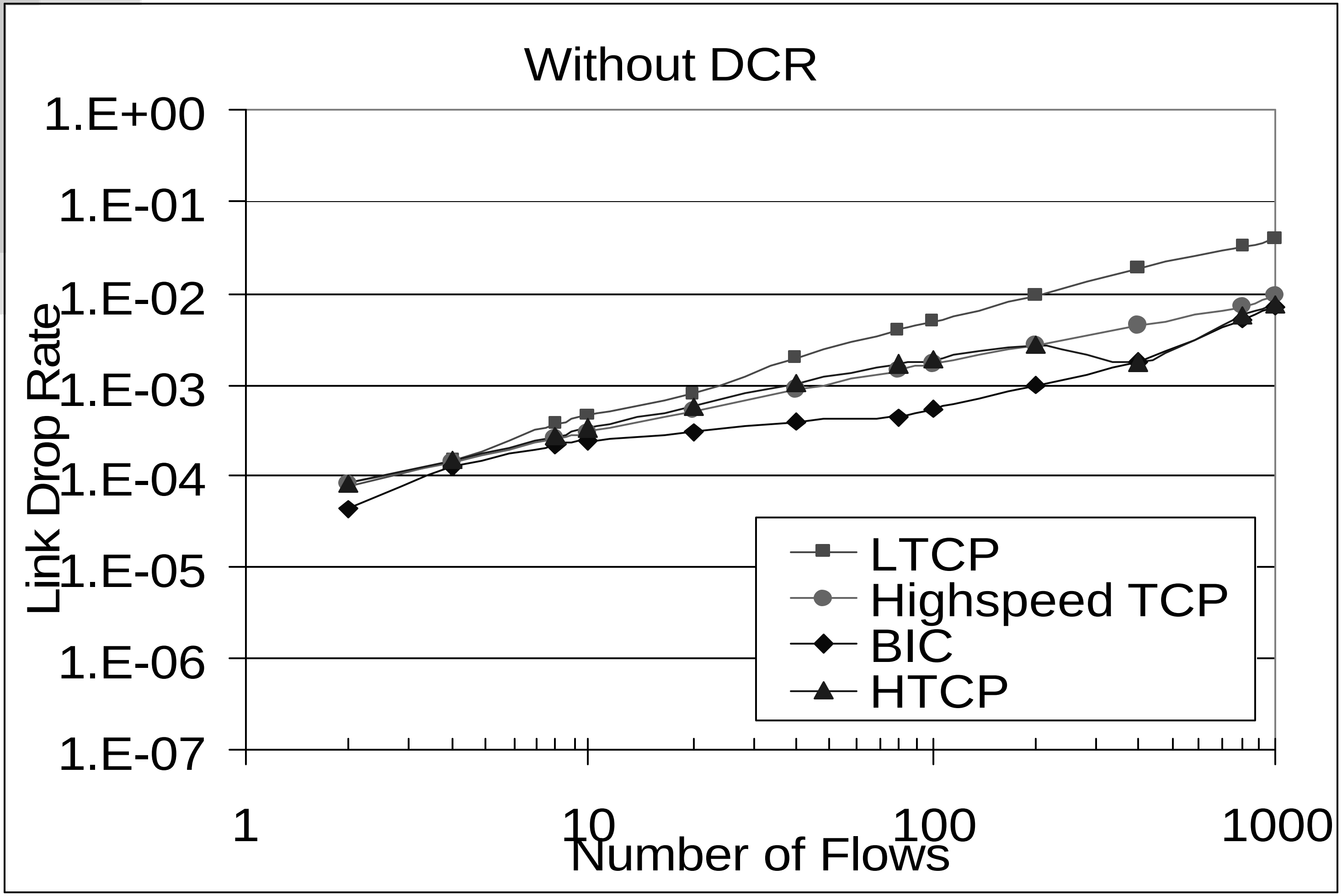
With Congestion Only

Link Utilization



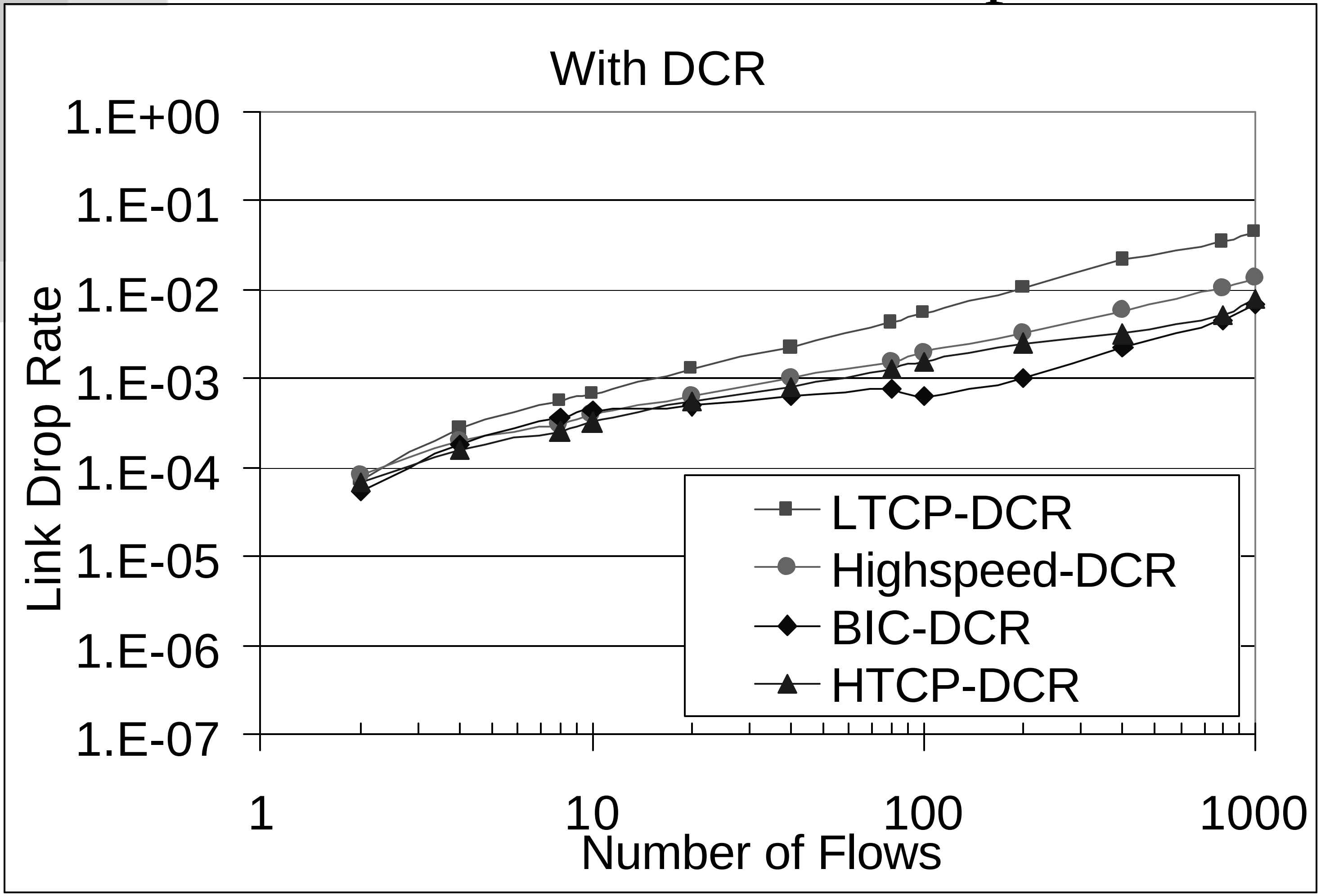
With Congestion Only

Bottleneck Link Buffer Drop Rate



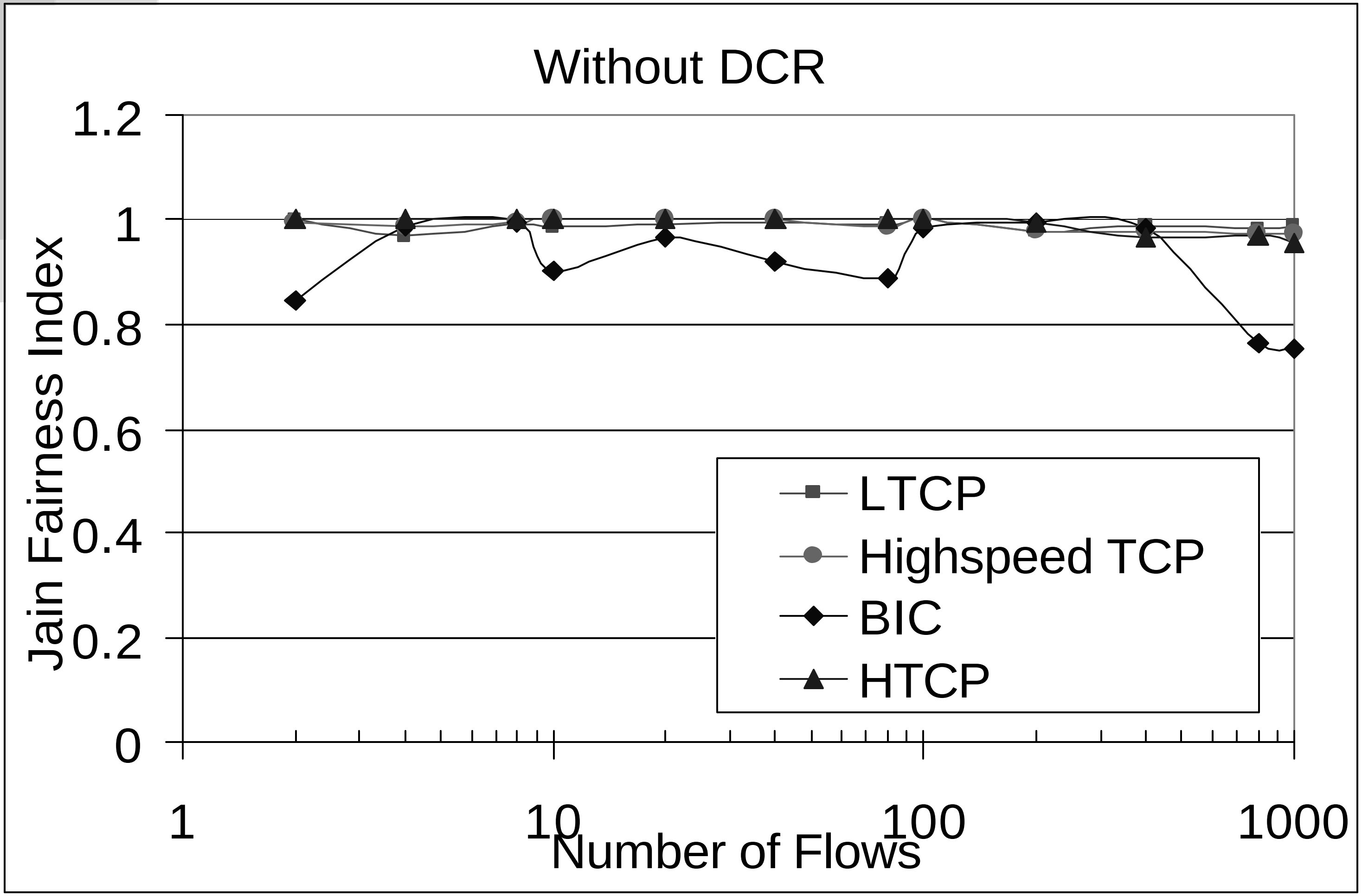
With Congestion Only

Bottleneck Link Buffer Drop Rate



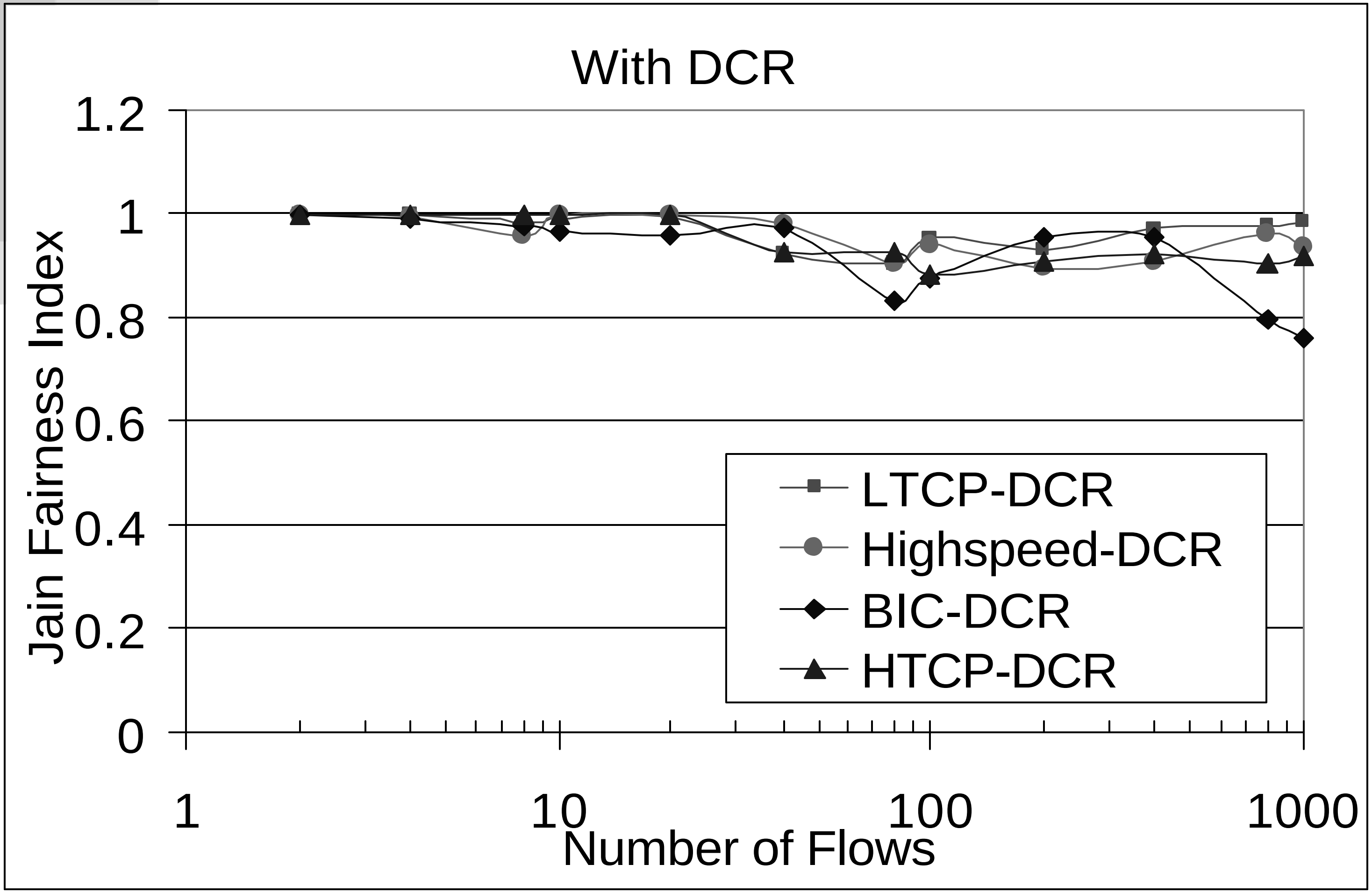
With Congestion Only

Jain Fairness Index



With Congestion Only

Jain Fairness Index

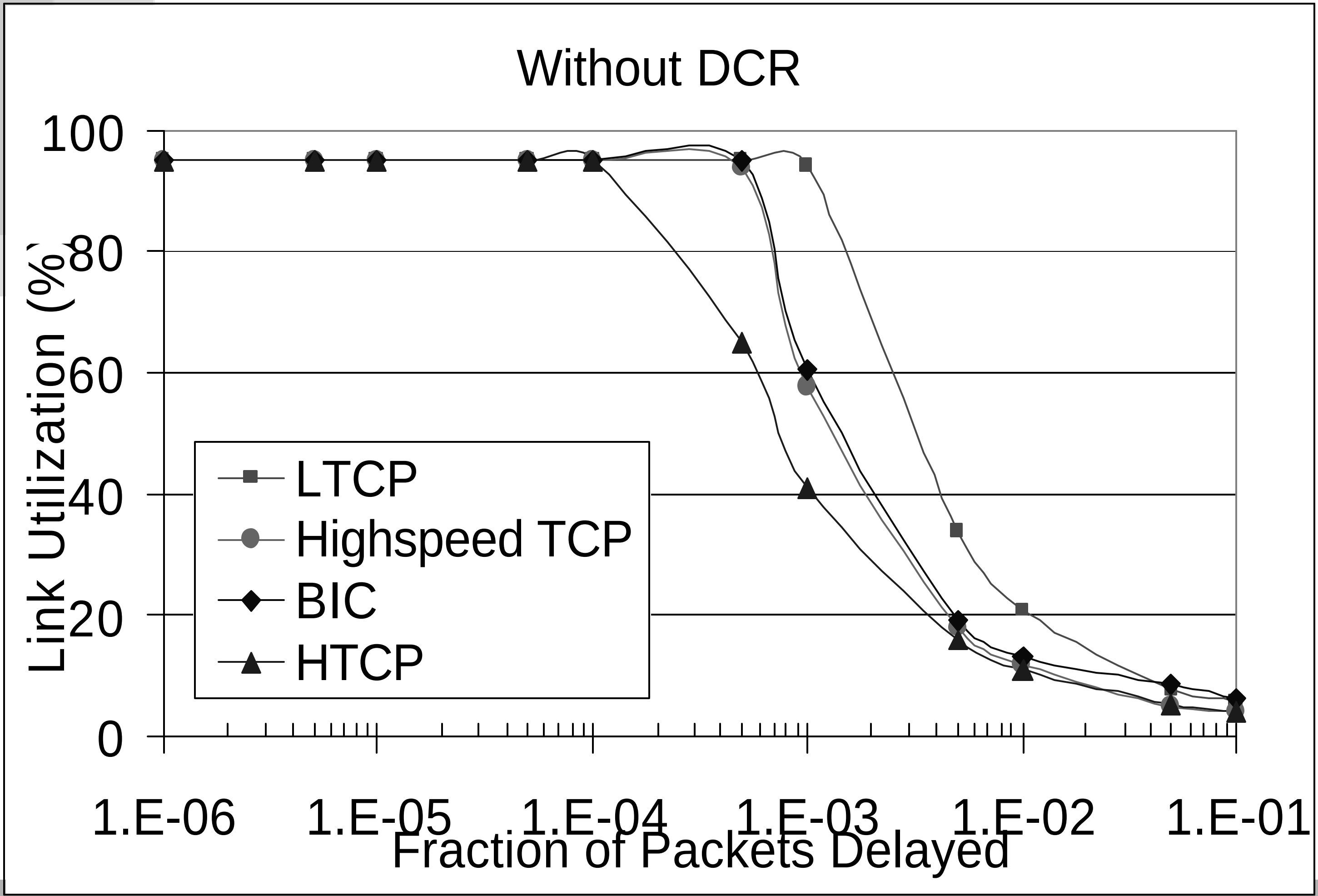


Observations

- When network has only congestion
  - Behavior remains similar with or without TCP-DCR
- Realistic conditions will have both congestion and reordering
  - Next set of results show the impact of both together.

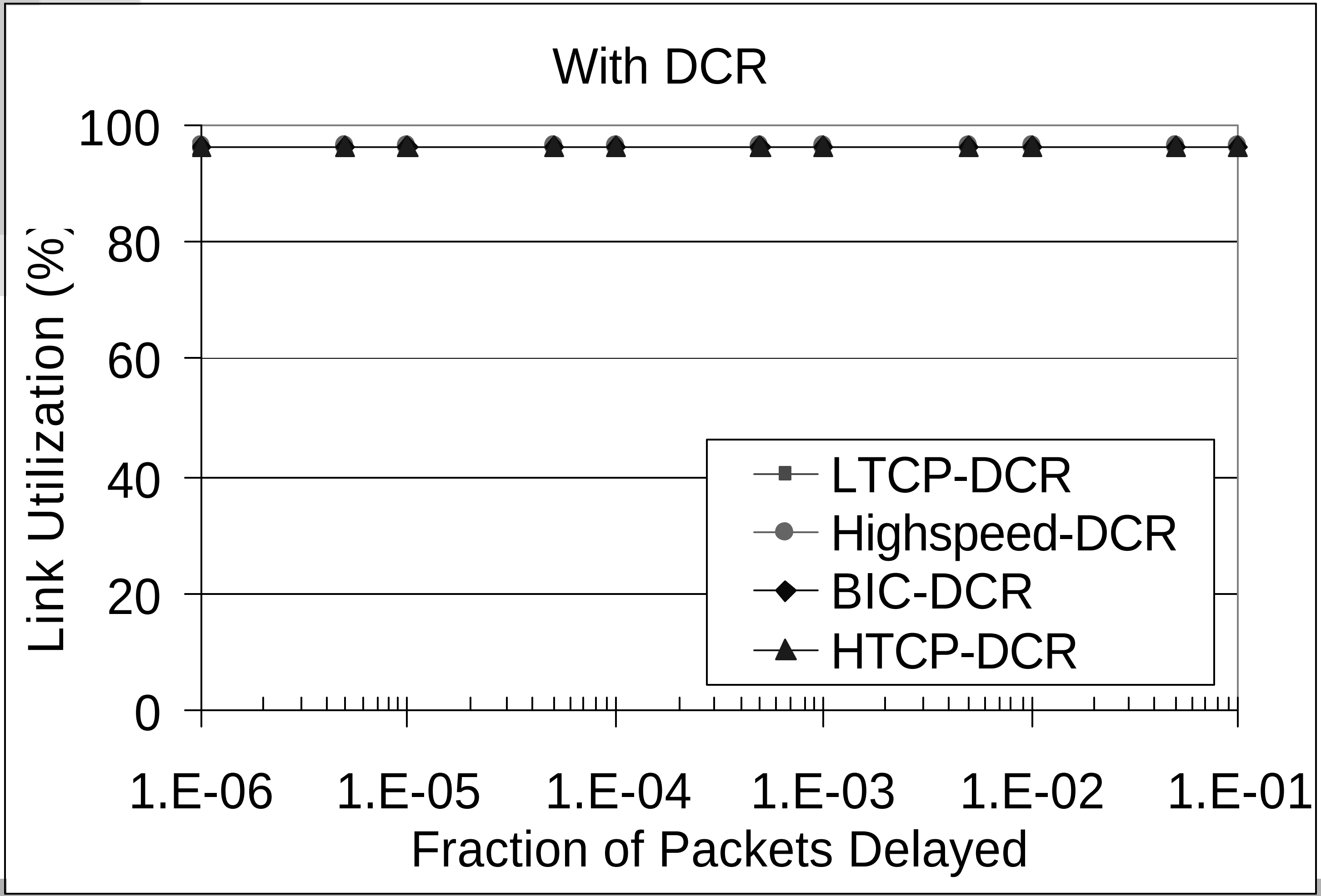
With Congestion & Reordering

Link Utilization



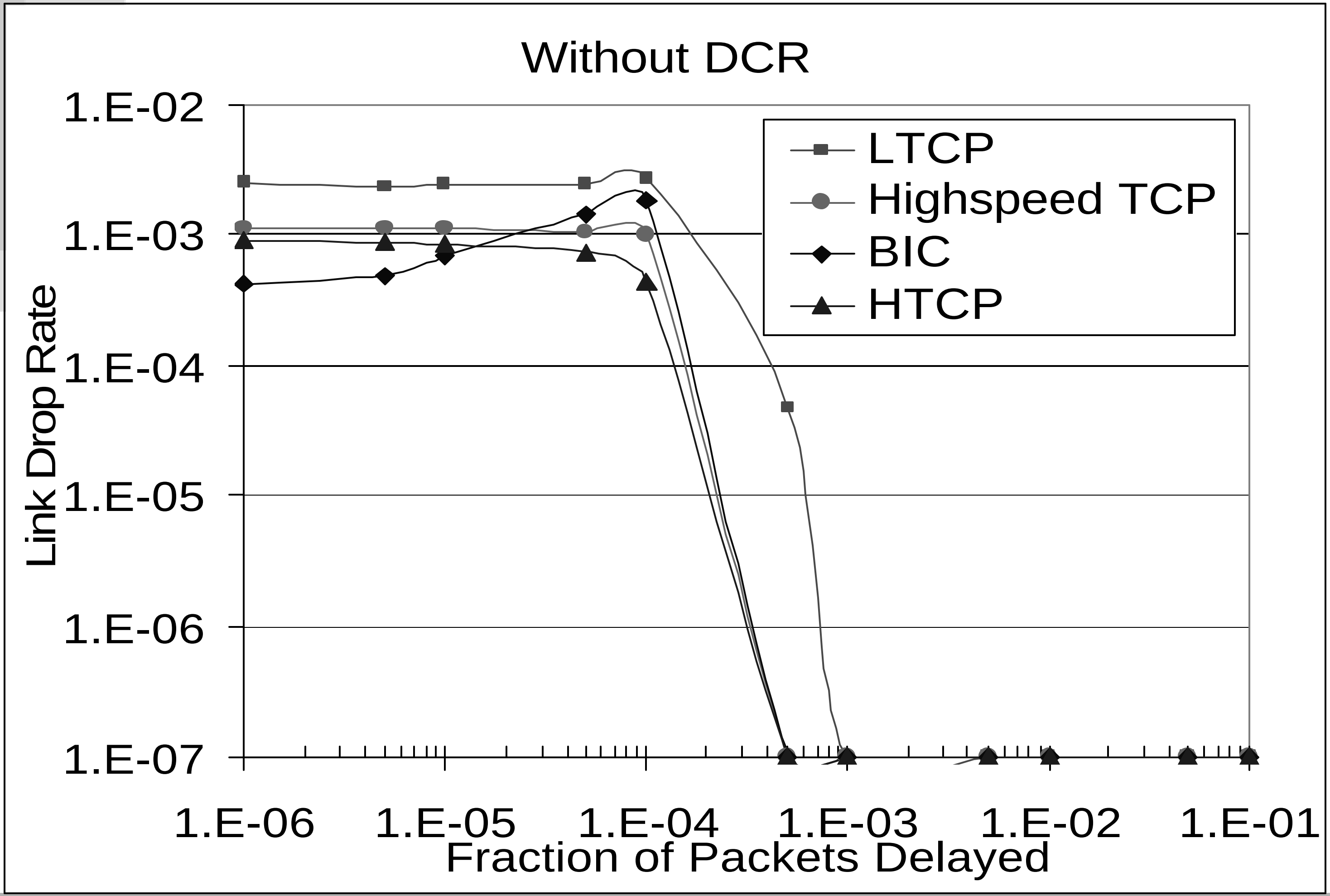
With Congestion & Reordering

Link Utilization



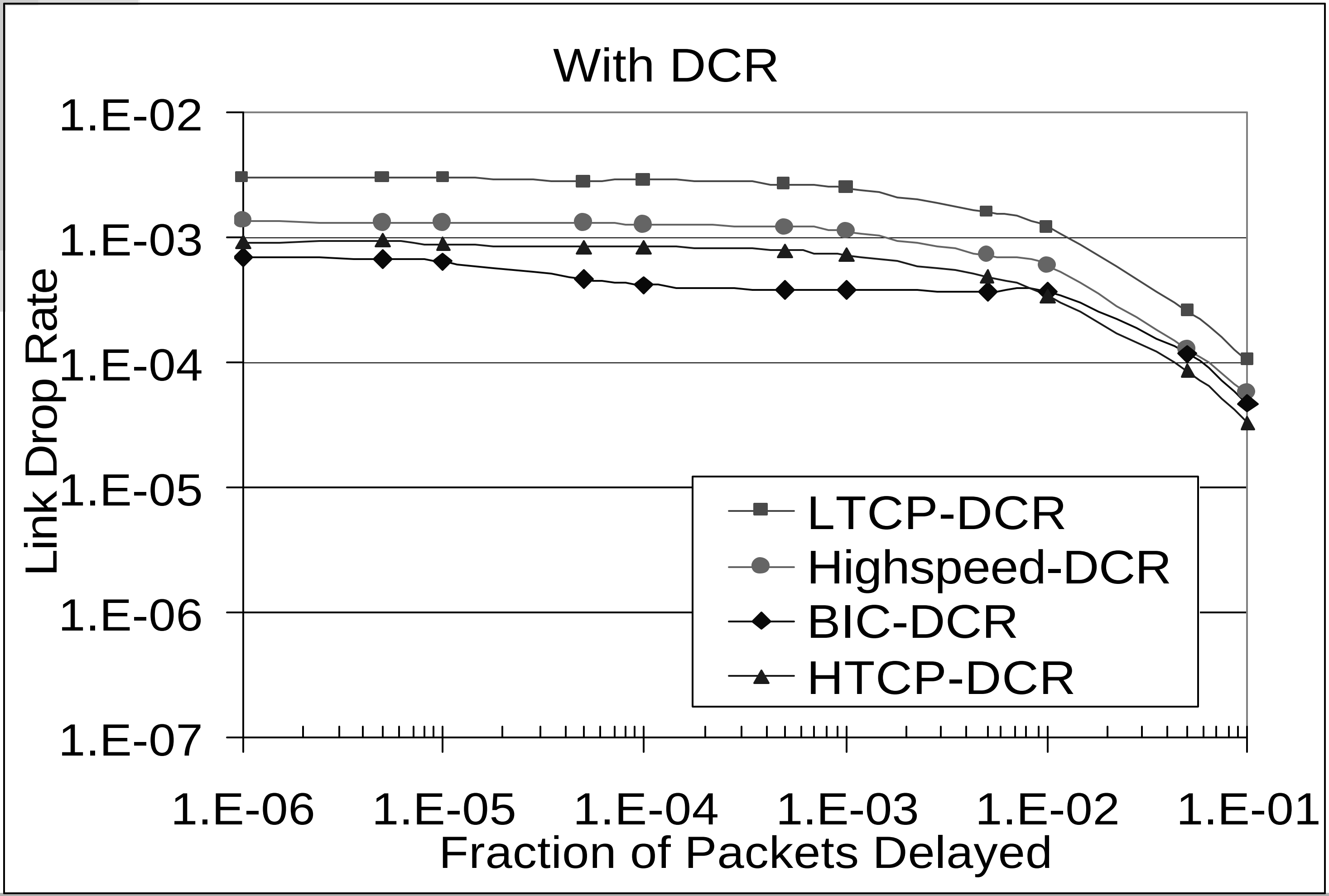
With Congestion & Reordering

Bottleneck Link Buffer Drop Rate



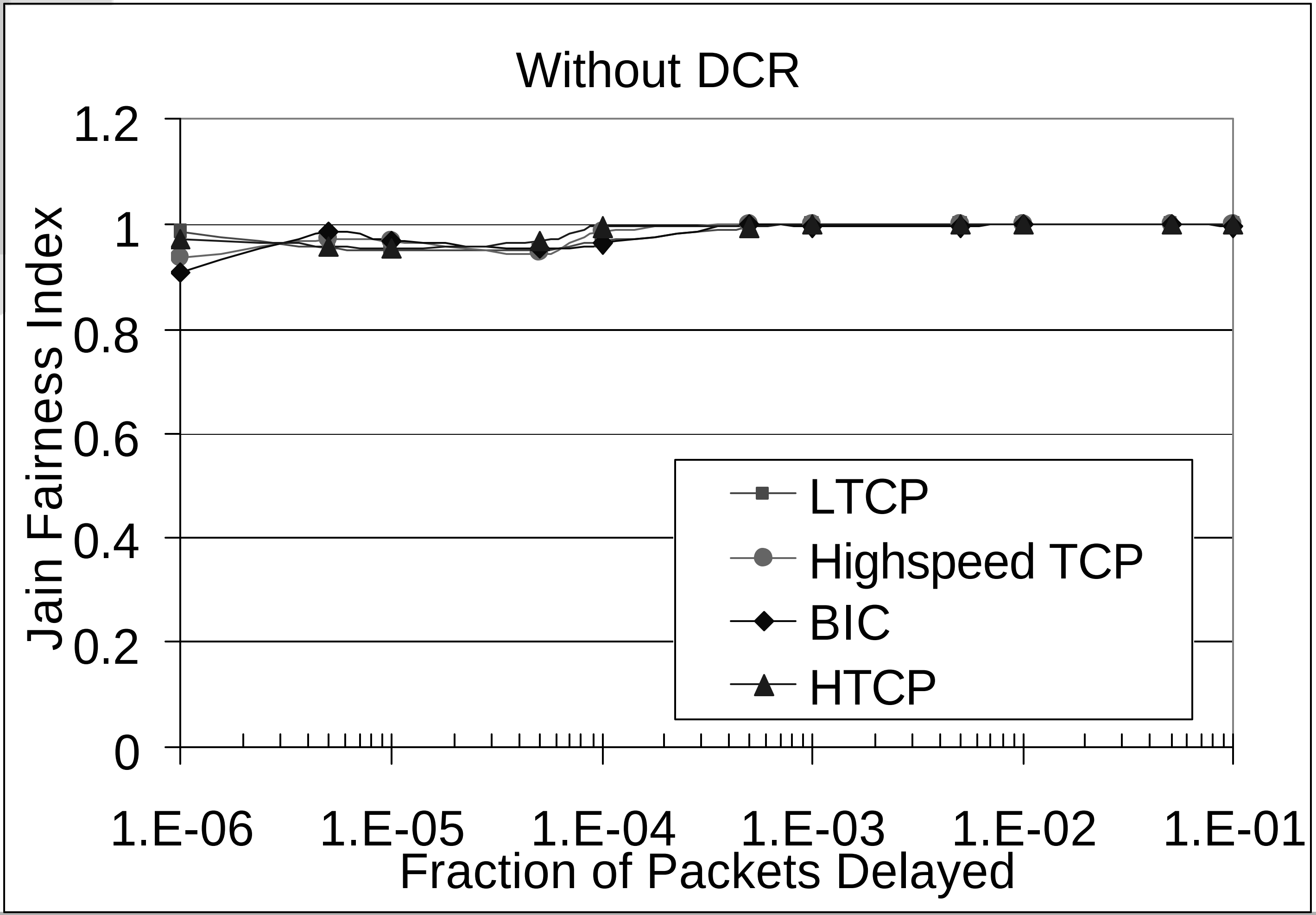
With Congestion & Reordering

Bottleneck Link Buffer Drop Rate



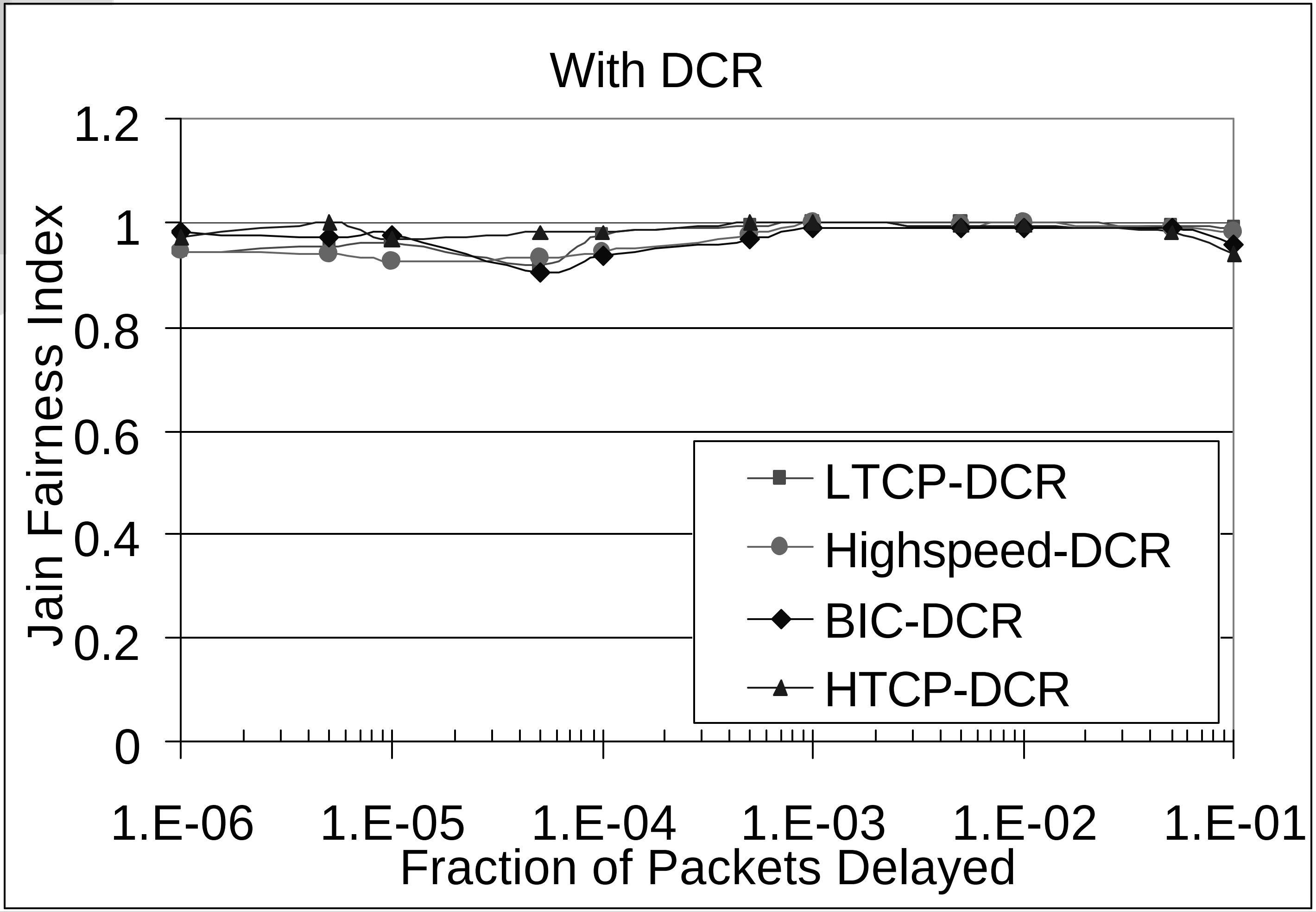
With Congestion & Reordering

Jain Fairness Index



With Congestion & Reordering

Jain Fairness Index

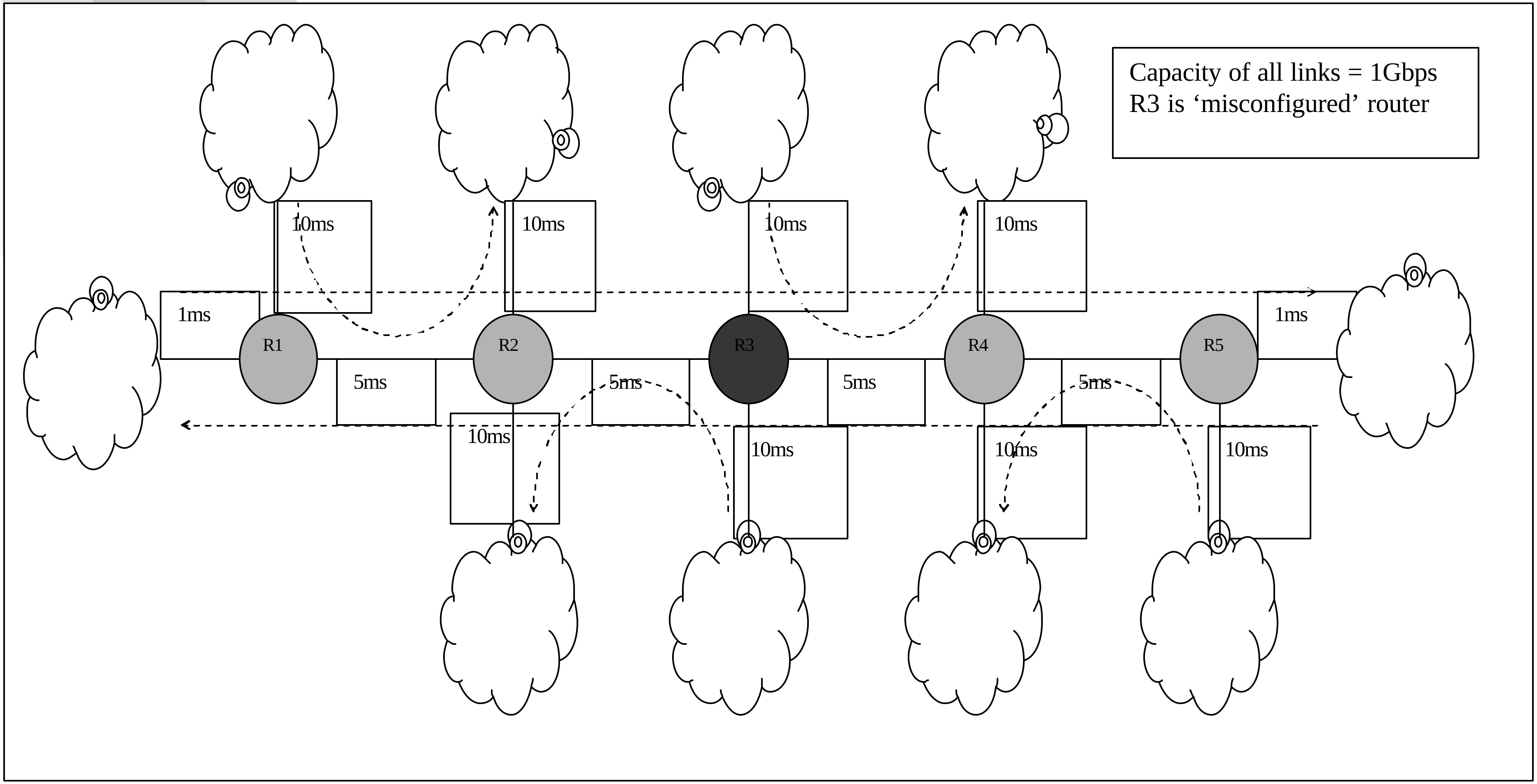


More Complex Scenarios

- Next we investigate more complex scenarios
  - Multiple Bottleneck Links
  - Traffic in both forward and reverse directions
  - Router R3 is 'misconfigured' causing 1% reordering of packets passing through it

More Complex Scenarios

Topology



More Complex Scenarios

Link Utilization

	LTCP	Highspeed	BIC	HTCP
R1-R2	96.87	94.2558	96.4142	94.2525
R3-R2	13.29	8.72681	5.41026	7.9631
R3-R4	11.01	8.83559	9.5864	8.96401
R5-R4	94.65	94.0697	96.5316	93.9085
	LTCP-DCR	Highspeed-DCR	BIC-DCR	HTCP-DCR
R1-R2	96.52	96.74	97.19	94.70
R3-R2	96.65	96.24	95.47	95.74
R3-R4	95.19	97.23	95.37	95.75
R5-R4	95.56	94.40	94.99	96.05

More Complex Scenarios

Bottleneck Link Buffer Drop Rate

	LTCP	Highspeed	BIC	HTCP
R1-R2	1.89E-03	5.79E-04	7.37E-04	5.20E-04
R3-R2	0	0	0	0
R3-R4	0	0	0	0
R5-R4	1.66E-03	6.01E-04	7.23E-04	4.30E-04
	LTCP-DCR	Highspeed-DCR	BIC-DCR	HTCP-DCR
R1-R2	2.57E-03	1.07E-03	8.96E-04	8.52E-04
R3-R2	2.83E-04	1.16E-04	1.06E-04	1.32E-04
R3-R4	2.51E-04	1.06E-04	9.20E-05	1.10E-04
R5-R4	2.41E-03	7.44E-04	6.59E-04	1.15E-03

More Complex Scenarios

Jain Fairness Index

	LTCP	Highspeed	BIC	HTCP
R1-R5	0.983	0.995	0.999	0.997
R1-R2	0.959	0.982	0.967	0.952
R3-R2	0.999	0.999	0.999	0.999
R3-R4	0.999	0.999	0.999	0.999
R5-R4	0.975	0.977	0.967	0.932
R5-R1	0.990	0.996	0.995	0.999
	LTCP-DCR	Highspeed-DCR	BIC-DCR	HTCP-DCR
R1-R5	0.936	0.959	0.904	0.963
R1-R2	0.988	0.972	0.976	0.950
R3-R2	0.996	0.991	0.987	0.994
R3-R4	0.994	0.990	0.989	0.988
R5-R4	0.985	0.978	0.961	0.926
R5-R1	0.961	0.951	0.914	0.971

Conclusion

- Significant improvement in the presence of packer reordering
- Minimal impact in the absence of reordering
- Congestion behavior not significantly modified
  - Similar drop rates
  - Similar Jain Fairness Index
- Safe for widespread deployment