SRP Store and Forward Performance Comparison with Cut-Through

Donghui Xie

Cisco Systems

May 7, 2001

Agenda

- Store and Forward vs Cut-Through w.r.t delay and jitter performance
- iPT simulation and its results
- DPT-OC3 ring in a similar setup as in iPT simulation
- SRP simulation performance criteria
- SRP performance simulation

  First: DPT-OC3 90% Utilization v.s. iPT 85% Utilization

  Second: DPT-OC3 100% Utilization v.s. iPT 90% Utilization
- Summary
- Appendix: DPT-OC3 vs DPT-OC12 performance

Store and Forward v.s. Cut-Through

- SRP uses a store and forward scheme to transit packets on the ring, which means a packet coming into a transit buffer is first buffer stored in transit buffer and then forwarded to the ring when permitted by SRP-fa. And the transit buffer is usually much larger than dozen of MTUs.
- Cut-Through generally refers to either no transit buffer in the transit path or a transit buffer with 1 MTU size.
- Issue: Vendors who employ Cut-Through in their RPR technology claim Cut-Through scheme gives better end-to-end delay and jitter performance than schemes using Store and Forward.
- There are delay and jitter trade-offs in both schemes. Cut-Through scheme actually pushes the delay and jitter problem onto the ring ingress point and beyond.
- In terms of end-to-end packet delay and jitter performance, Store and Forward may well be the best scheme.

iPT and its Simulation

- iPT technology uses a cut-through scheme for its ring packet insertion and transit forwarding.
- iPT technology is presented at:
  - **www.ieee802.org/rprsg/public/presentations/may2000/rprsg_ipt_overview.pdf**
- Simulation setup:
  - iPT OC-3 network
  - each hop link propagation delay is 10us (2km)
  - node 0 to node 7 sends TCP traffic to destination node 8 in one direction
  - total 8 nodes aggregate traffic on the ring
  - the distance from node 0 to node 8 is 16km
  - node buffer size: 2MB
  - traffic rate from each node 16Mbps in first simulation; in the second simulation, node 0 to 3 22Mbps, node 4 to 7 12Mbps

iPT Simulation Results

- iPT cut-through performance simulation results are presented at:

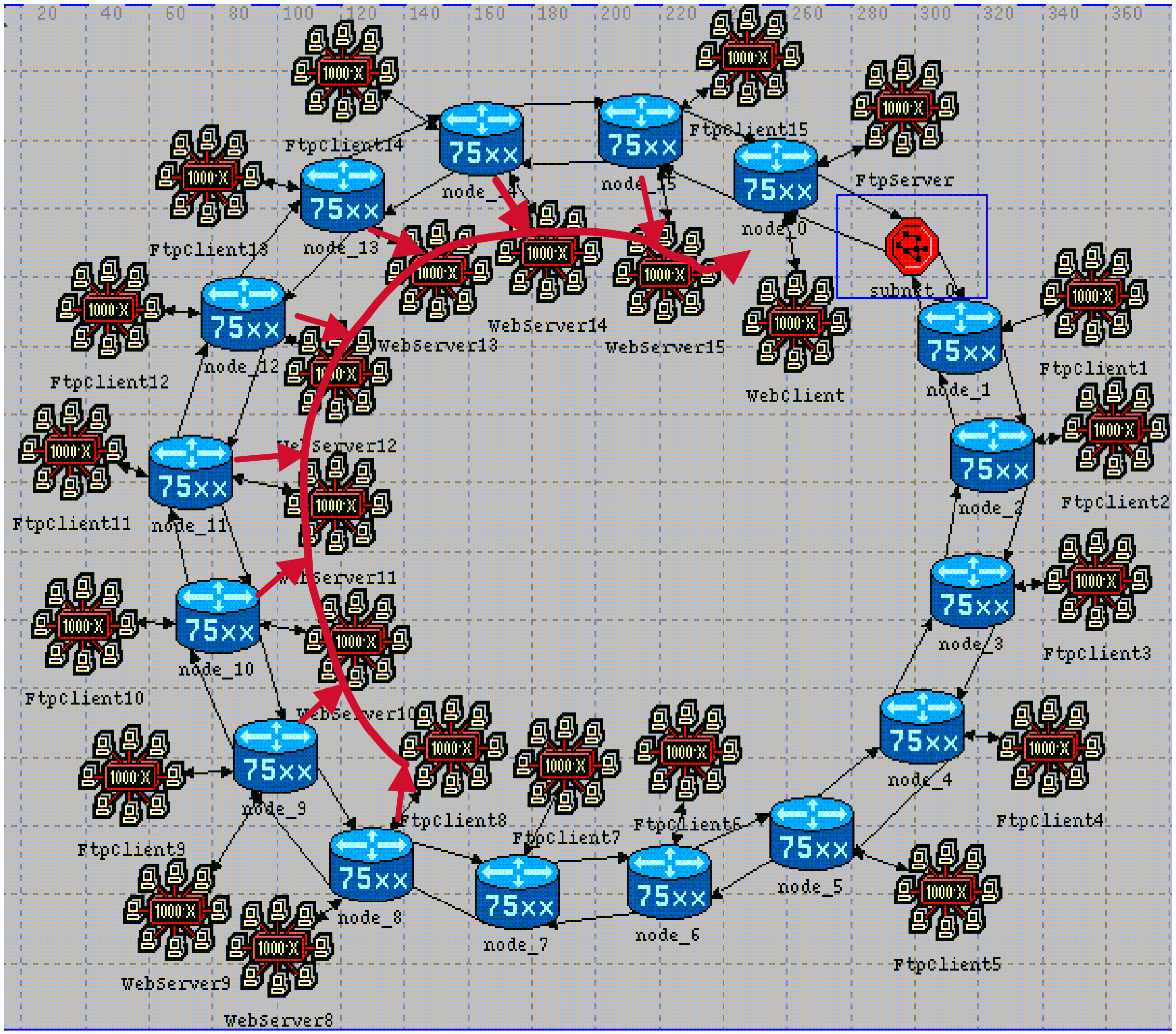
  http://www.ieee802.org/rprsg/public/presentations/may2000/rprsg_ipt_fairness_sim.pdf

- Delay and jitter performance results for iPT cut-through were presented as
  - Head Of Line Delay at page 7, 12 and 19.
  - Ring Access Delay at page _, 13 and 20.
  - Packet End-to-End Delay at page 8, 14 and 21.
- Assume iPT end-to-end delay measurement is from Layer 2 Add Traffic point to ring Drop Traffic point (see iPT node model at page 4 of technology pres).
- We will compare SRP layer 2 end-to-end packet delay with iPT's end-to-end packet delay.

SRP Simulation Performance Criteria

- SRP node traffic input: traffic insertion rate onto the ring from the node
- Ring utilization: the ratio of the throughput on the most congested ring link to the link's bandwidth
- Packet end-to-end delay: the time between a packet being placed into the transmit buffer in source node to the packet being received by destination node and passed onto higher layer

SRP Simulation Setup
DPT-OC3 Ring and Traffic Flows



- Exact ring size as in Nortel's simulation
- Link propagation delay 10us (2km)
- 8 nodes aggregate HTTP and FTP TCP traffic to node_0
- SRP Configuration:
  - LP transit buffer **128Kbytes**
  - LP transmit buffer **512Kbytes**
  - LP Tb low threshold **16Kbytes**
  - LP Tb high threshold **96Kbytes**
  - Max_allow 8000

TCP/HTTP and Ftp Application Configuration

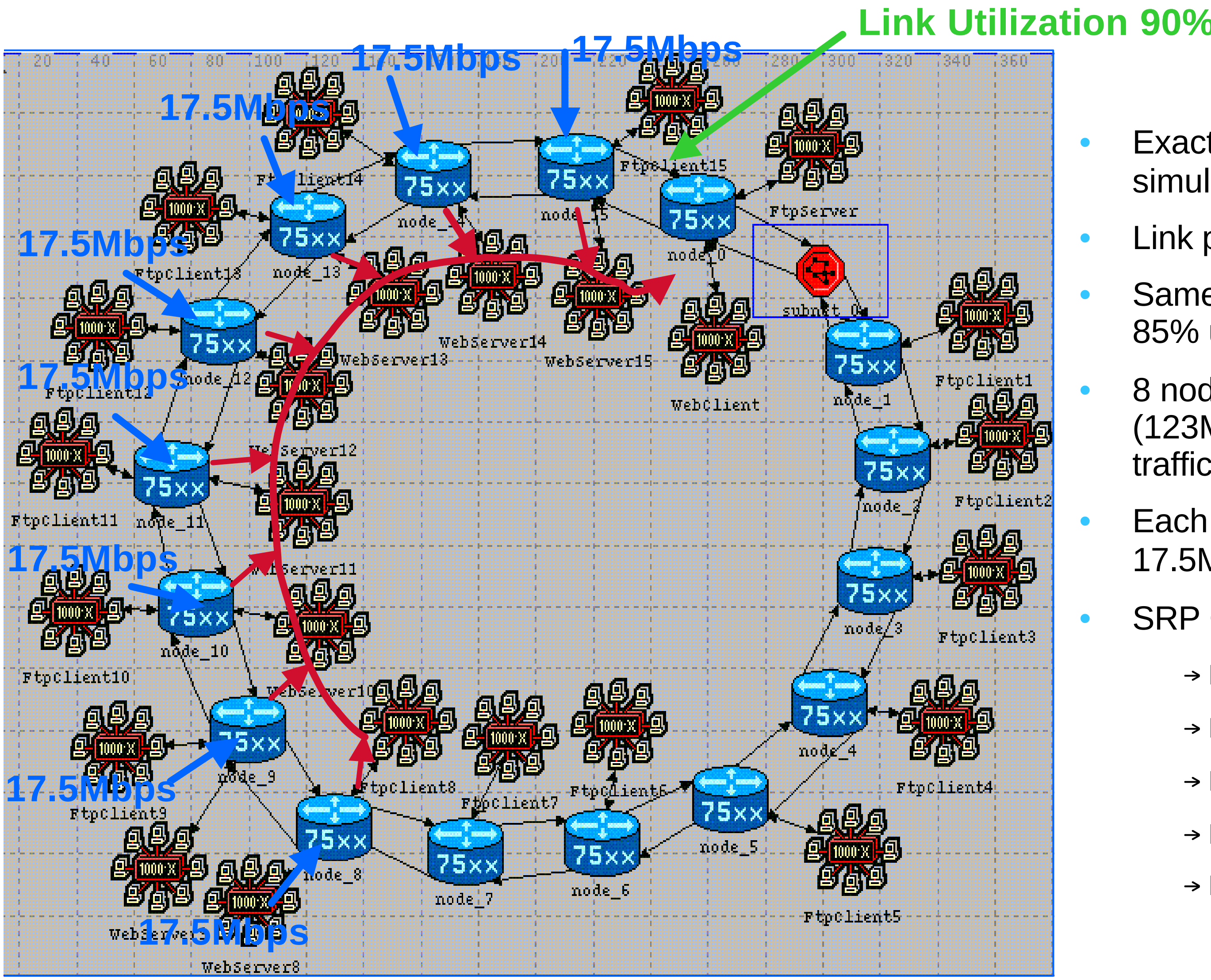
Attribute	Value
Maximum Segment Size (bytes)	Auto-Assigned
Receive Buffer (bytes)	65536
Receive Buffer Usage Threshold (of RCV BUFF)	0.0
Delayed ACK Mechanism	Segment/Clock Based
Maximum ACK Delay (sec)	0.200
Slow-Start Initial Count (MSS)	1
Fast Retransmit	Enabled
Fast Recovery	Disabled
Window Scaling	Disabled
Selective ACK (SACK)	Disabled
Nagle's SWS Avoidance	Disabled
Karn's Algorithm	Enabled
Retransmission Thresholds	Attempts Based
Initial RTO (sec)	1.0
Minimum RTO (sec)	0.5
Maximum RTO (sec)	64
RTT Gain	0.125
Deviation Gain	0.25
RTT Deviation Coefficient	4.0
Timer Granularity (sec)	0.5
Persistence Timeout (sec)	1.0

- HTTP Traffic Configuration
  - HTTP 1.1
  - Exponential page interarrival time
  - Exponential number of objects per page
  - Random object size up to hundred kbytes
  - Exponential number of pages per server
  - 400 ~ 640 simultaneous web users

- FTP Traffic Configuration
  - 20~50 simultaneous users per LAN
  - Exponential ftp request inter-arrival
  - Exponential file size with mean to 100 kbytes

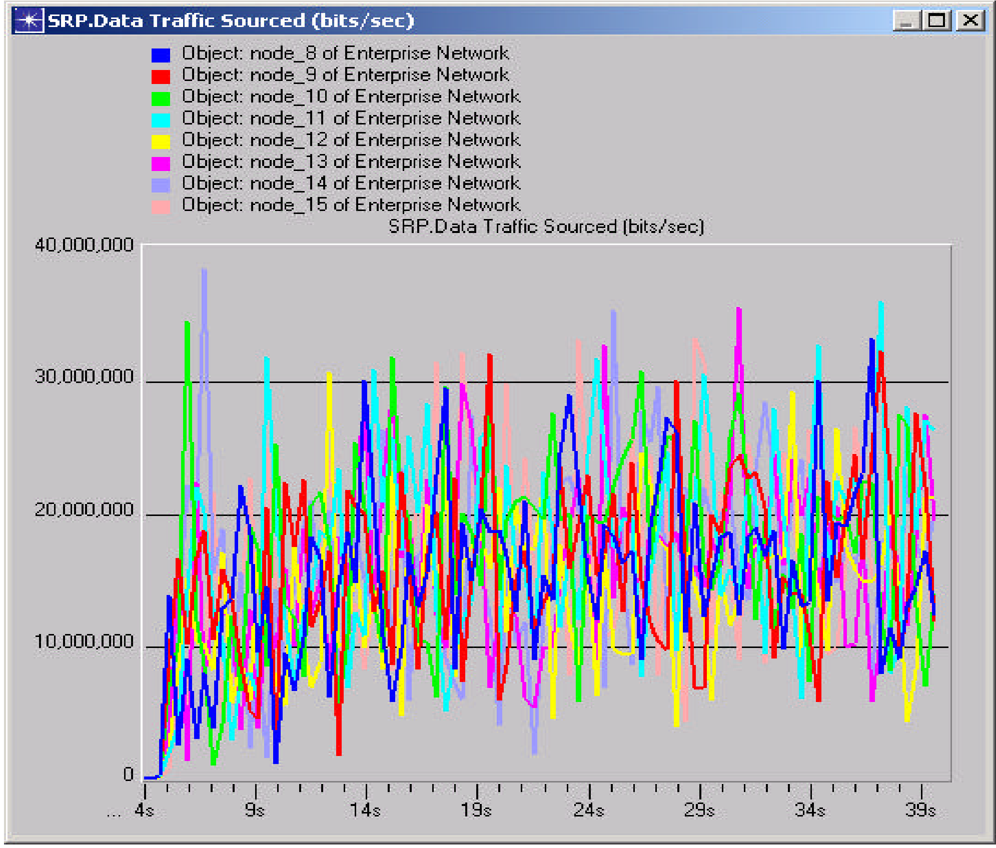
- TCP Configuration
  - TCP Tahoe with fast retransmission
  - No fast recovery
  - Buffer size: 65535 bytes

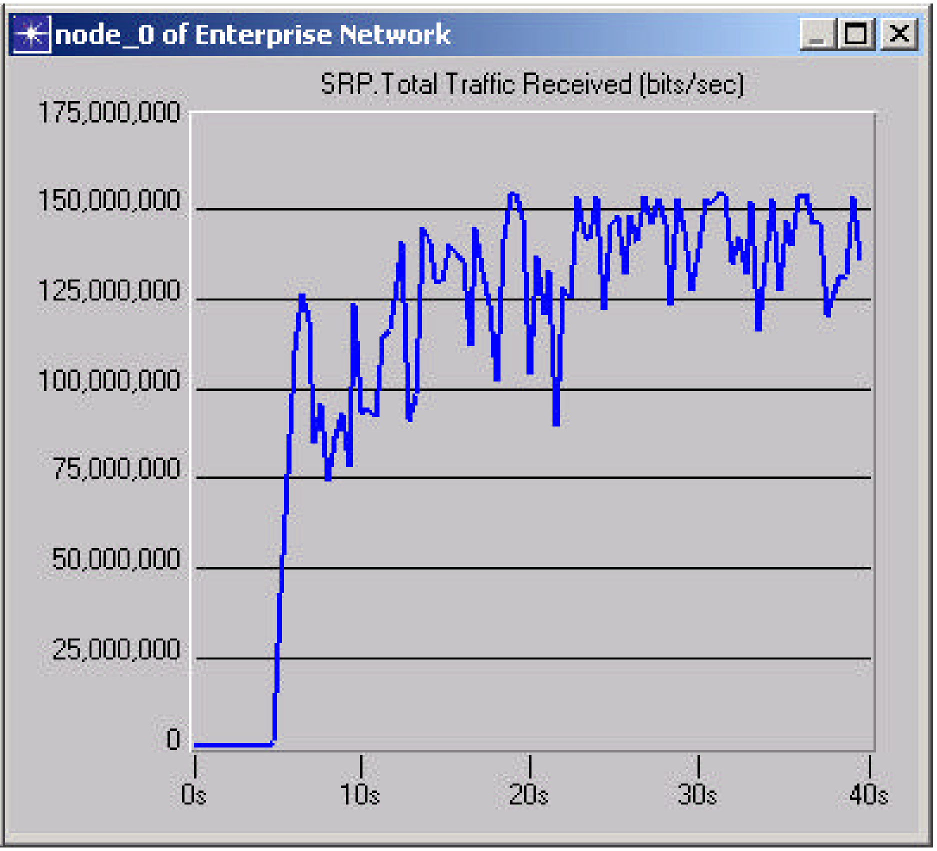
First Simulation: 90% Utilization



- Exact ring size as in iPT's simulation
- Link propagation delay 10us (2km)
- Same traffic patterns as in iPT's 85% util simulation
- 8 nodes aggregate HTTP (123Mbps) and FTP (17Mbps) TCP traffic to node_0
- Each node sources about 17.5Mbps of HTTP and FTP traffic
- SRP Configuration:
  - → LP transit buffer **128Kbytes**
  - → LP transmit buffer **512Kbytes**
  - → LP Tb low threshold **16Kbytes**
  - → LP Tb high threshold **96Kbytes**
  - → Max_allow 8000

90% Utilization
Traffic Input and Total Throughput

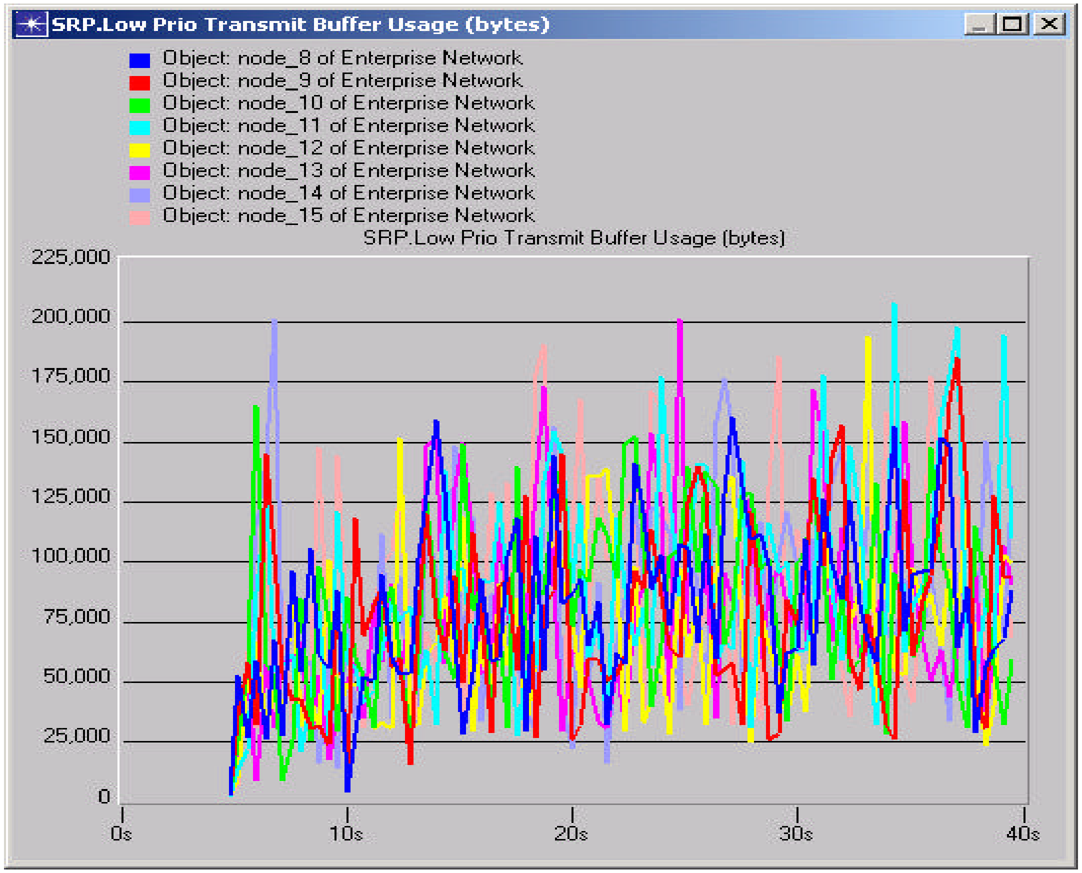




- Node traffic insertion rate onto the ring
- Each traffic source consists of 88% highly bursty HTTP traffic and 12% highly bursty FTP traffic

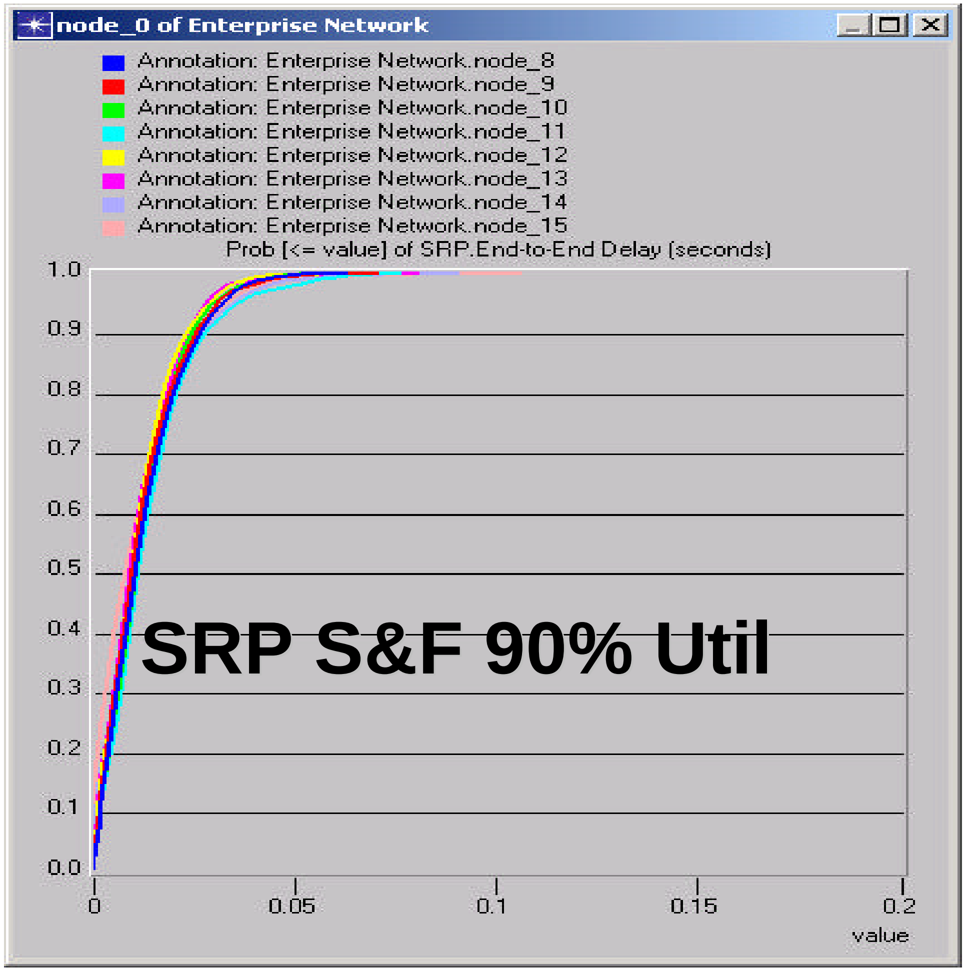
Total traffic throughput on the ring

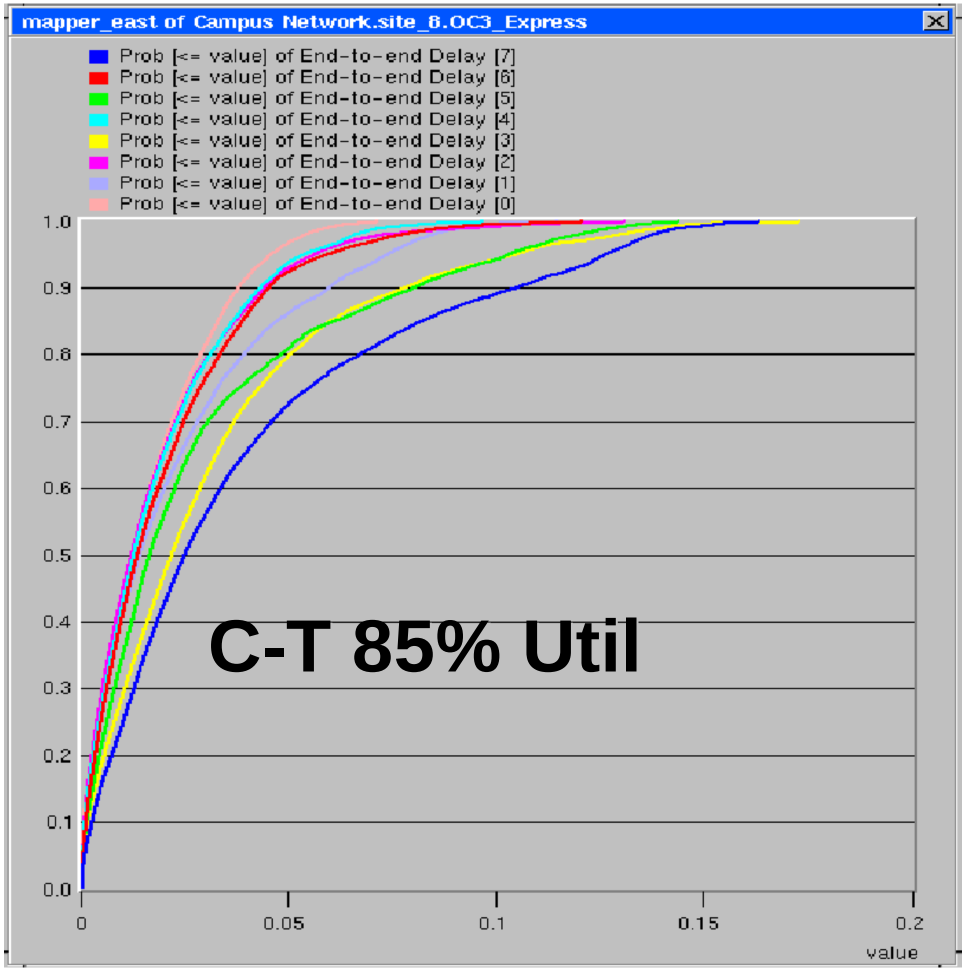
90% Utilization
Transmit Buffer Usage



- Very large variation in buffer usage for all the node due to highly bursty TCP/HTTP and FTP traffic sources
- Highly bursty sources contribute significantly to delay jitters
- No packet loss

90% Utilization
Layer 2 Packet End-to-End Delay

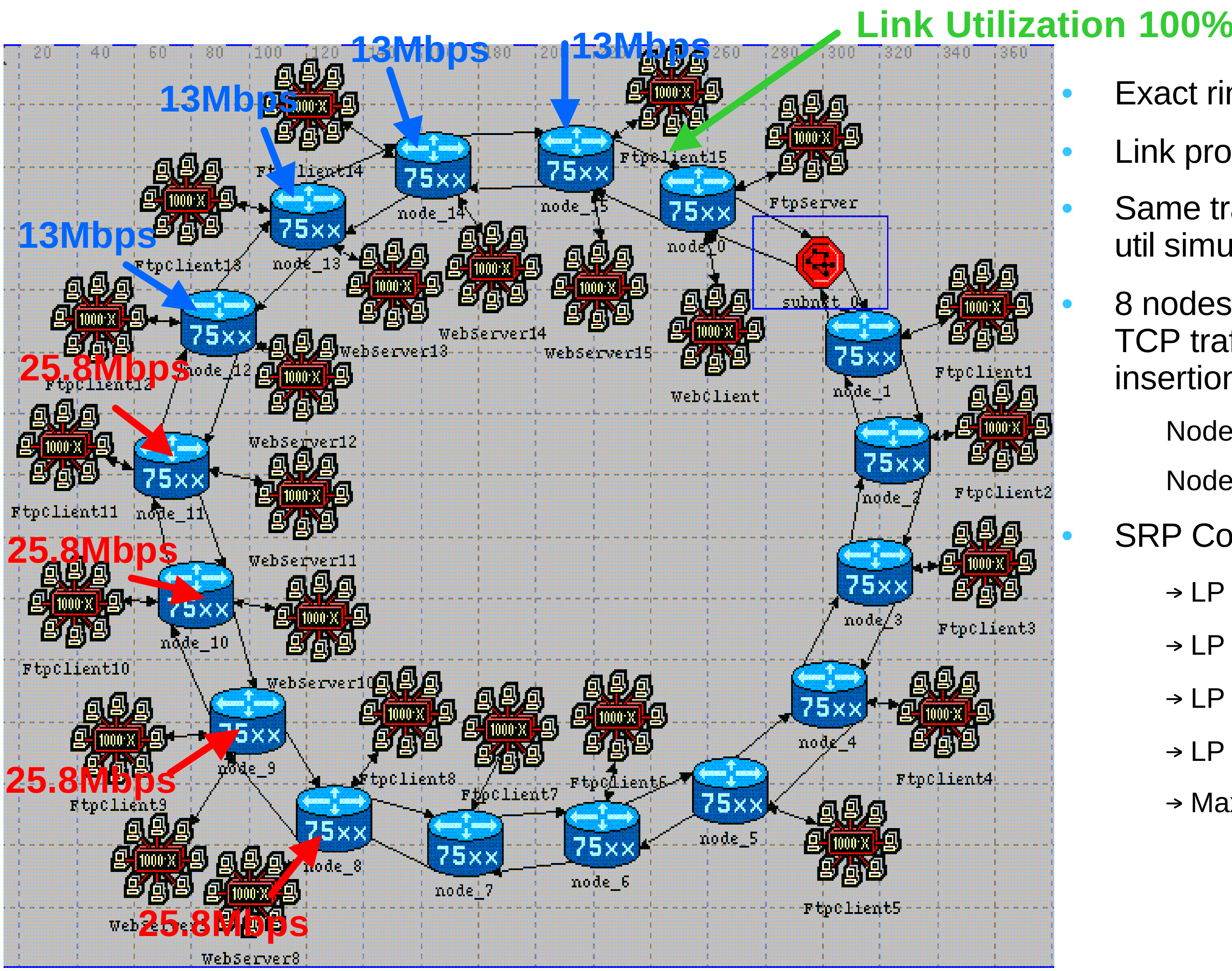




- SRP Ring Packet End-to-End Delay
- Consistent and fair delay and jitters

iPT Ring Packet End-to-End Delay
Screen snapshot from rprsg_ipt_fairness_sim.pdf(p8)

Second Simulation: 100% Utilization

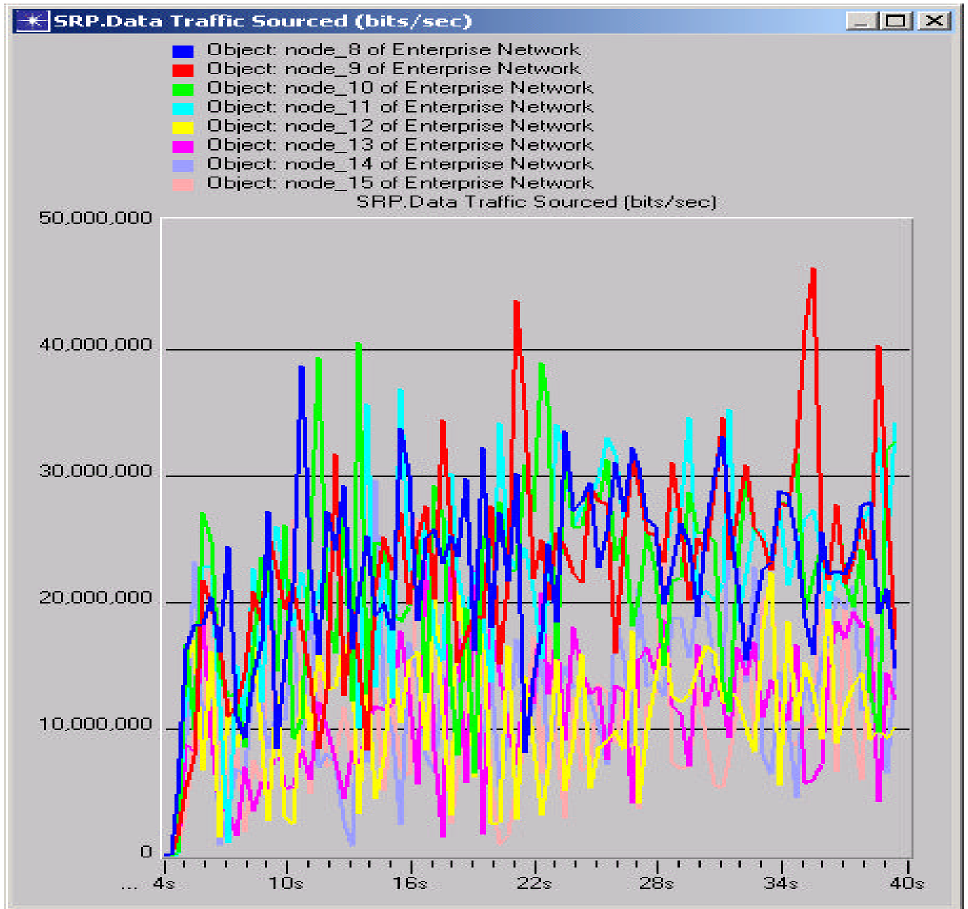


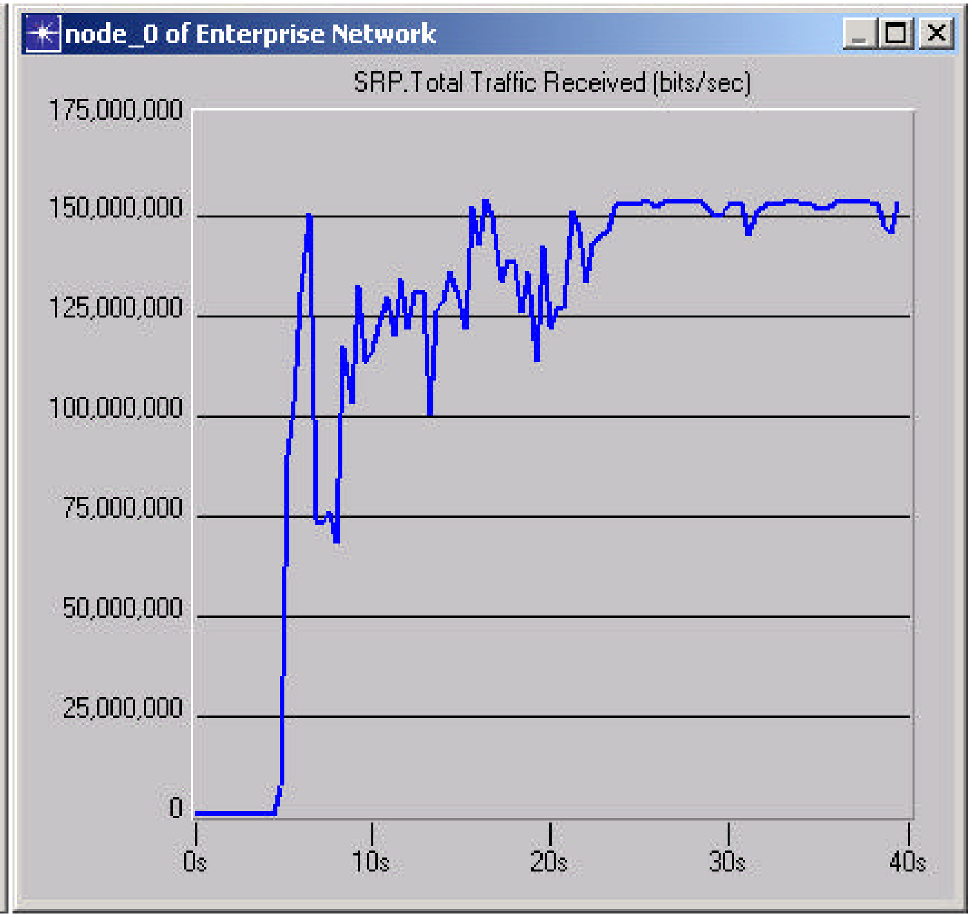
- Exact ring size as in iPT simulation
- Link propagation delay 10us (2km)
- Same traffic patterns as in iPT's 90% util simulation
- 8 nodes aggregate HTTP and FTP TCP traffic to node_0, intended traffic insertion rate:

  Node 8, 9, 10 and 11: 26Mbps

  Node 12, 13, 14 and 15: 13Mbps
- SRP Configuration:
  - → LP transit buffer **128Kbytes**
  - → LP transmit buffer **512Kbytes**
  - → LP Tb low threshold **16Kbytes**
  - → LP Tb high threshold **96Kbytes**
  - → Max_allow 8000

100% Utilization
Traffic Rate on the Ring

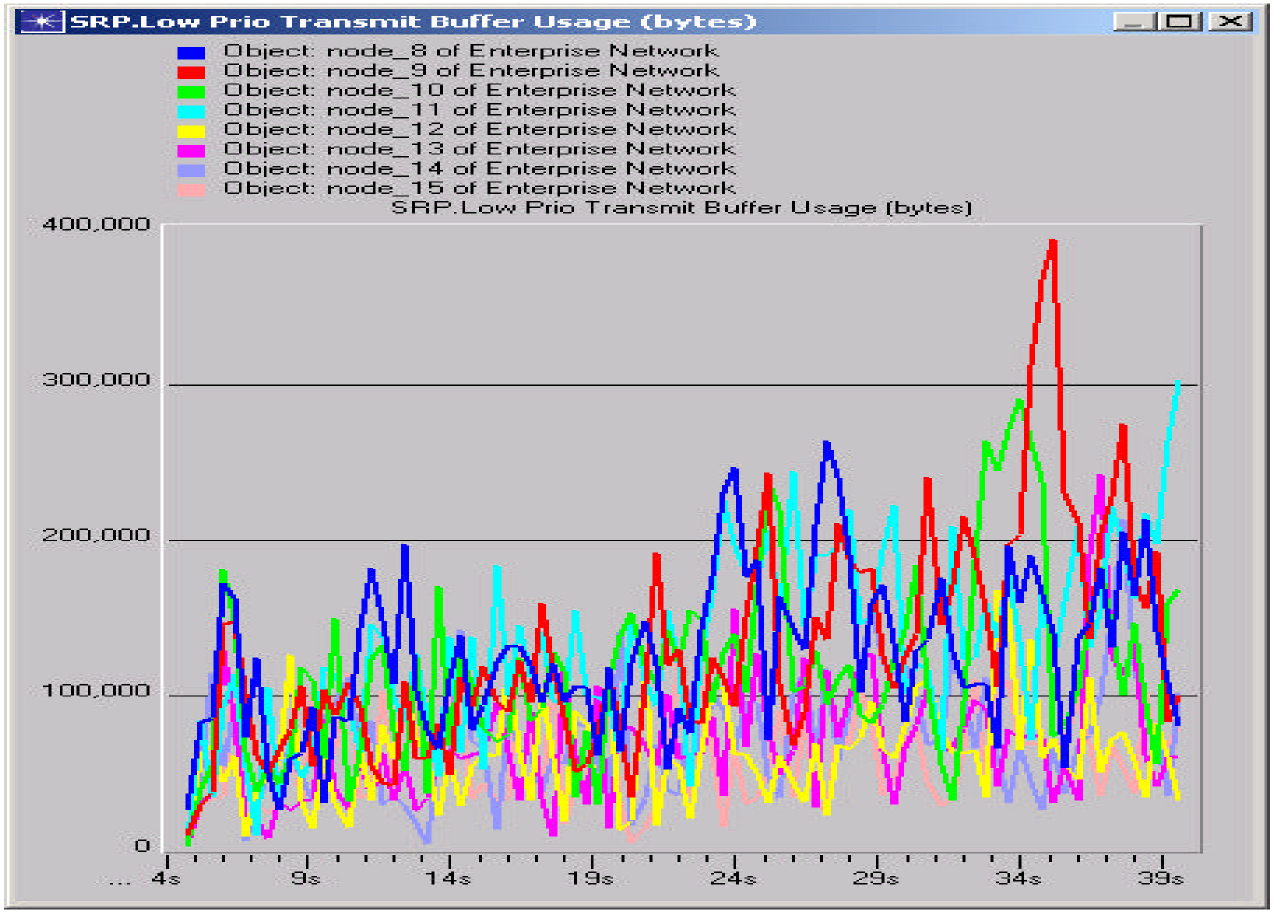




- Node traffic insertion rate onto the ring
- Each traffic source consists of 88% highly bursty HTTP traffic and 12% highly bursty FTP traffic

Total traffic throughput on the ring

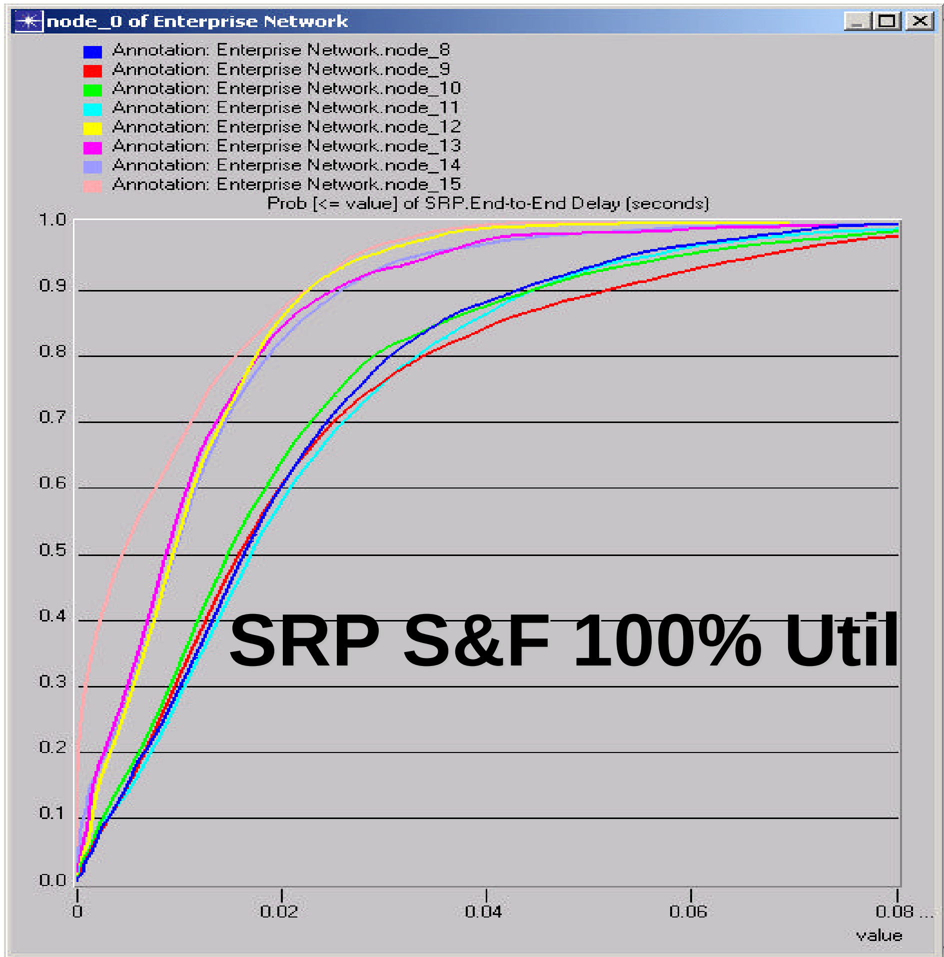
100% Utilization
Transmit Buffer Usage

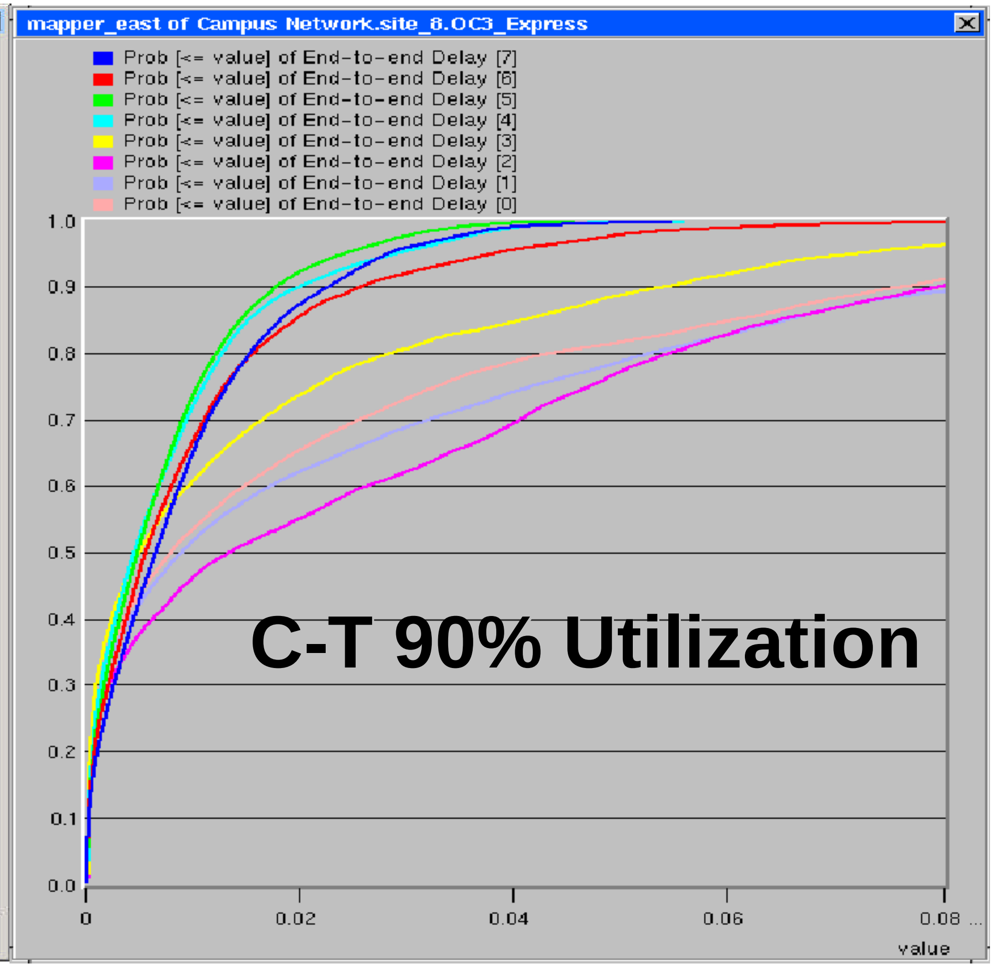


- Very large variation in buffer usage for all the node due to highly bursty TCP/HTTP and FTP traffic sources
- Nodes that insert double amount of traffic have the largest buffer usage and variation.
- No packet loss

Cisco Systems

100% Utilization
Layer 2 Packet End-to-End Delay





- SRP Ring Packet End-to-End Delay
- Consistent and fair delay and jitters
- Only large traffic has large delay and jitters

iPT Ring Packet End-to-End Delay
Screen snapshot from rprsg_ipt_fairness_sim.pdf(p21)

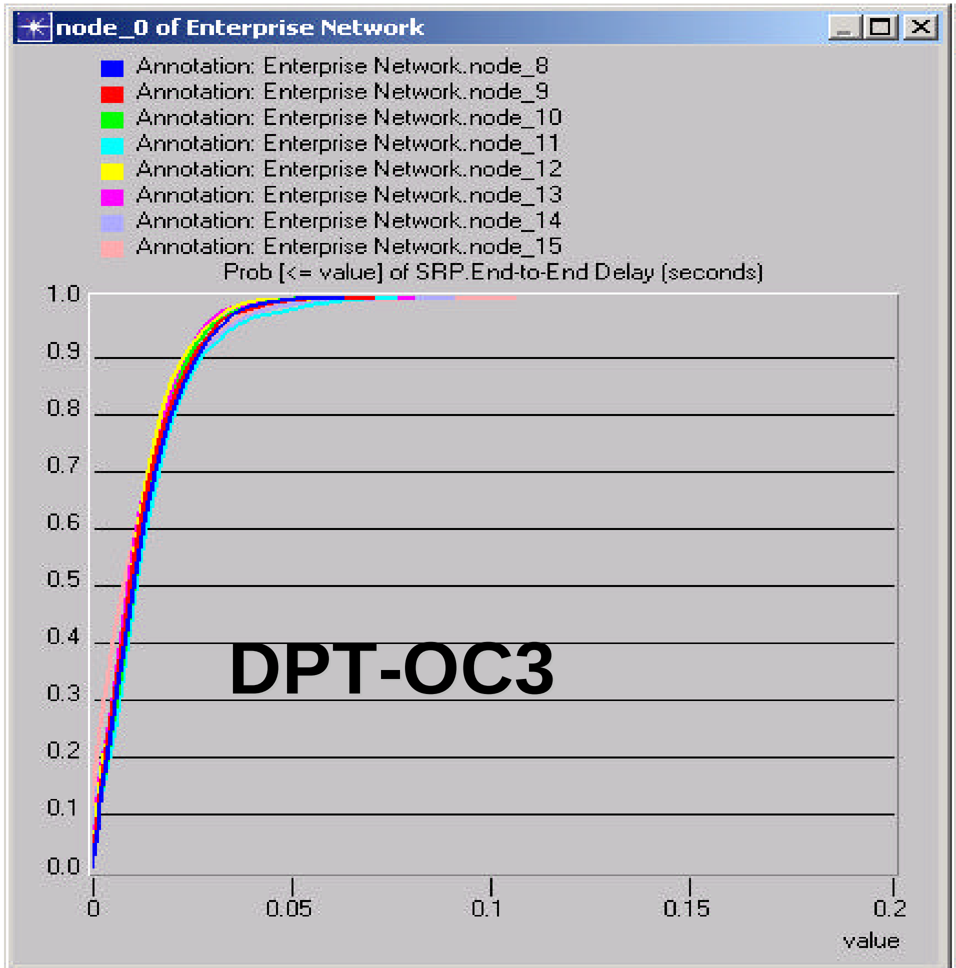
Summary
SRP Store and Forward Performs Better!

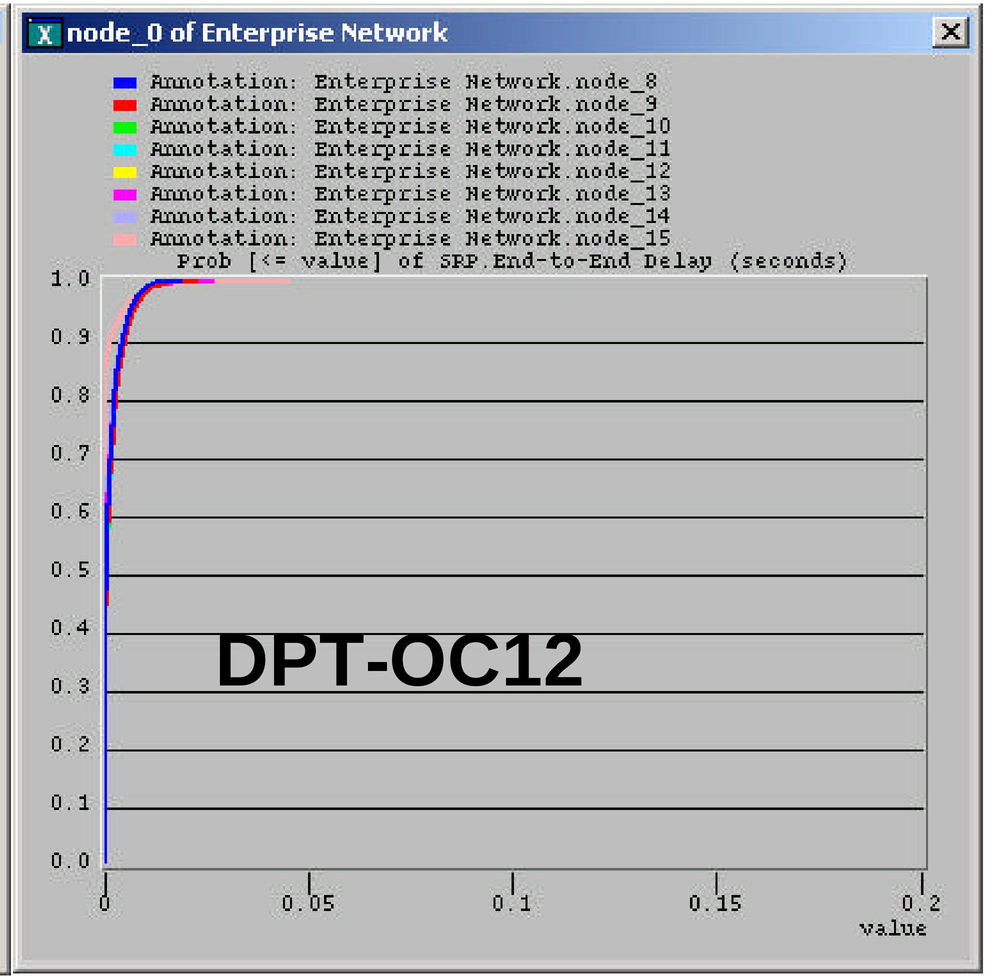
- SRP store and forward technology produced a consistent better packet end-to-end delay and jitter performance as ring utilization increases from 90% to complete 100% subscription.
- As link utilization increases from 85% to 90%, iPT cut-through simulation results indicate much more severe end-to-end delay and jitter performance degradation for many nodes, even though the ring is not oversubscribed.

Appendix:

1. DPT-OC3 vs DPT-OC12 Network Delay Performance with 90% Utilization
2. DPT-OC3 vs DPT-OC12 Network Delay Performance with 100% Utilization

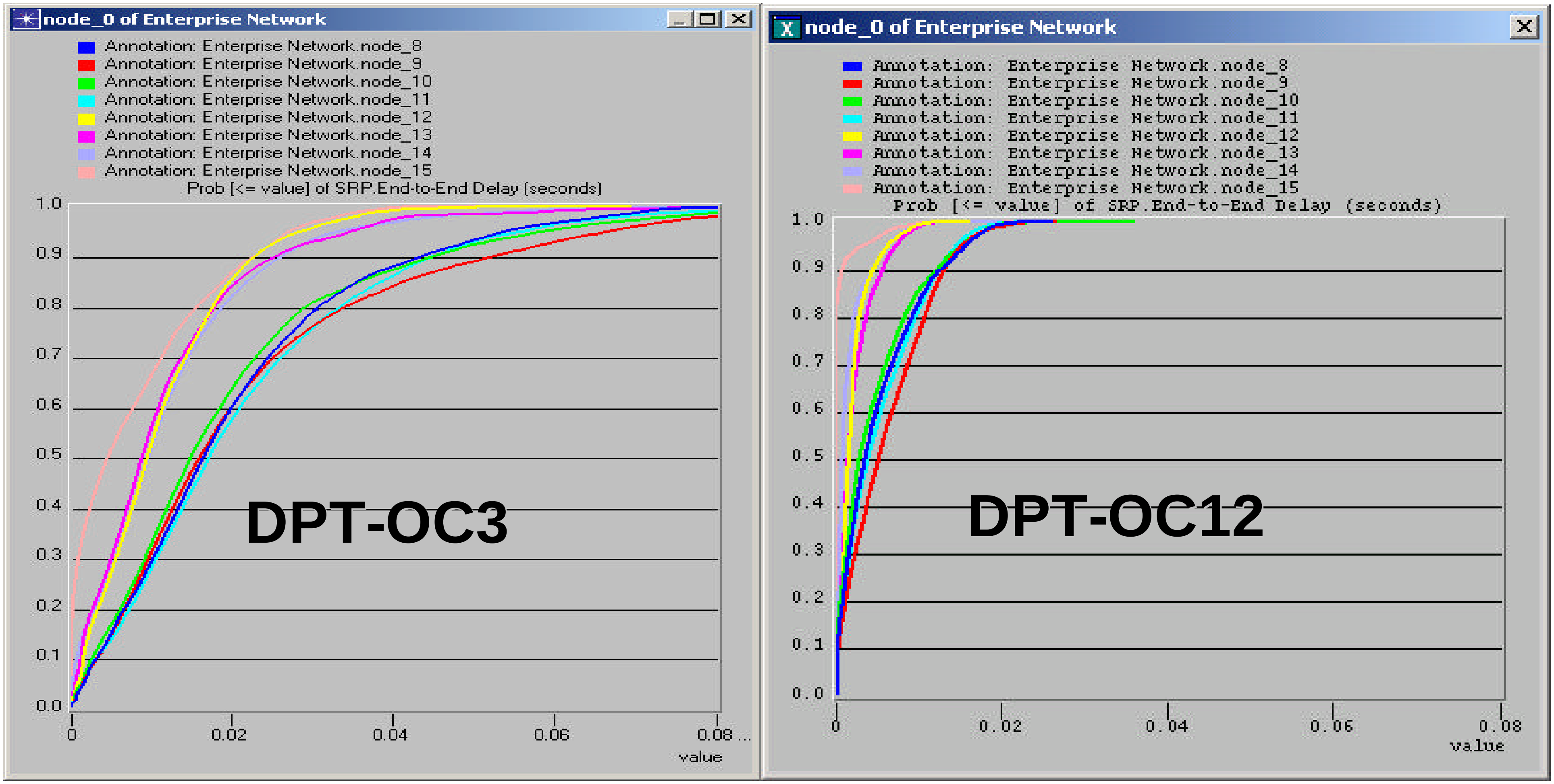
A.1. DPT-OC3 vs DPT-OC12
90% Utilization





As bandwidth increases, packet end-to-end delay and jitter over DPT ring becomes significantly smaller.

A.2. DPT-OC3 vs DPT-OC12 100% Utilization



As link bit rate increases, packet end-to-end delay and jitter over DPT ring become significantly smaller.