Toward a Reliable Data Transport Architecture for Optical Burst-Switched Networks

Dr. Vinod Vokkarane
Assistant Professor, Computer and Information Science
Co-Director, Advanced Computer Networks Lab
University of Massachusetts Dartmouth, USA

April 22, 2006

IEEE Invited Lecture Presentation

Department of Electrical and Computer Engineering

Indian Institute of Science (IISc), Bangalore, India

Presentation Outline

- Introduction to Optical Transport Paradigms
- Optical Burst Switching
- Reliable Data Transport in OBS
  - Loss Minimization Mechanisms
  - Loss Recovery Mechanisms
- Conclusion and Future Work

Applications Demands

Applications → **Service Requirements** → **Optical Transport Paradigms**

Applications
- Voice Over IP
- Streaming Video
- Grid Computing
- Storage Area Networks
- Multimedia
- Data

Service Requirements
- High Bandwidth
- Dynamic Provisioning
- Reliability
- Low Latency

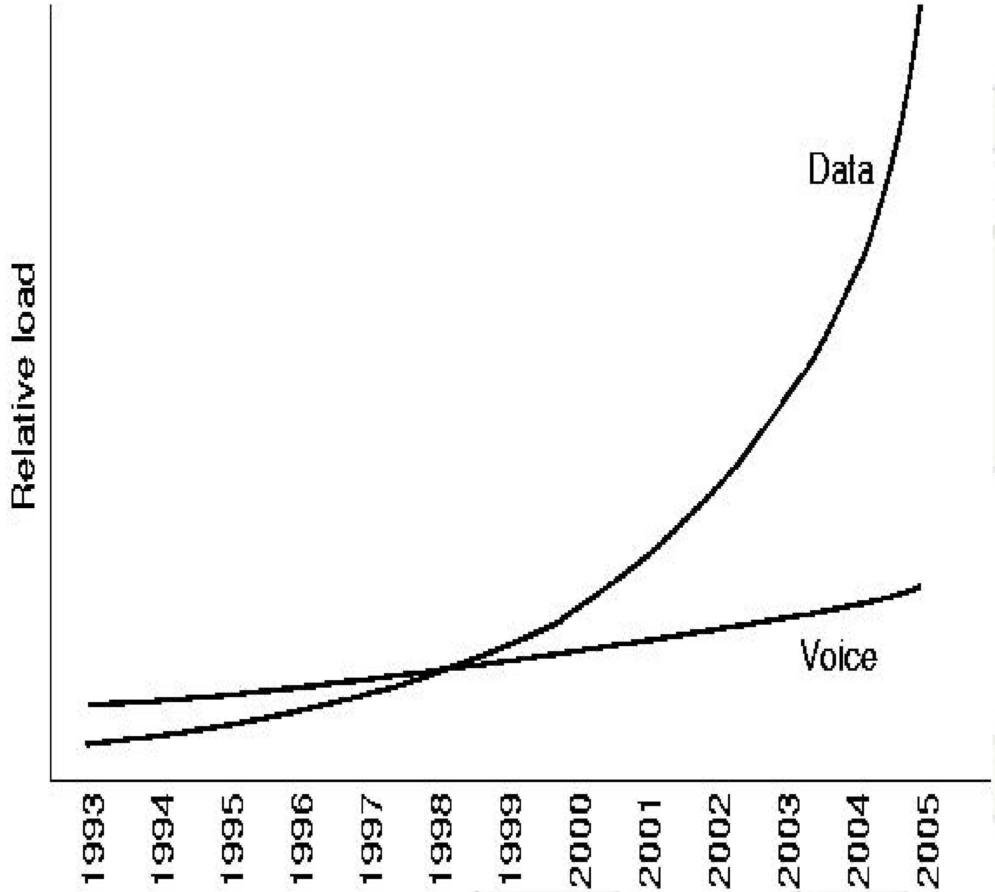
Optical Transport Paradigms
- Optical Circuit Switching
- Optical Packet Switching
- Optical Burst Switching

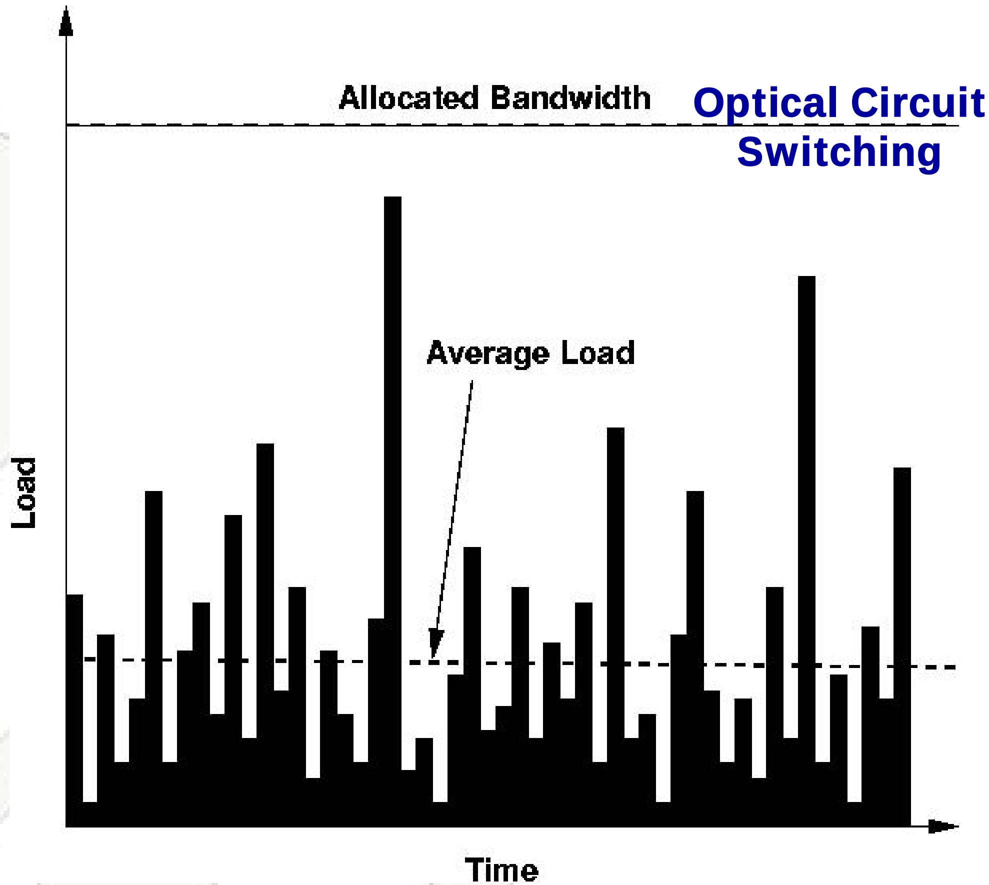
Optical Circuit Switching

- For each request,
  - Set-up a static circuit (lightpath)
  - Transfer data
  - Release connection

- Pros:
  - Suitable for smooth, longer-term, high-bandwidth applications
- Cons:
  - Long circuit set-up latency
  - Inefficient for short-term bursty applications

Optical Circuit Switching (cont.)





- Circuit switched networks optimized for **Voice**
- **Data**: Accounts for majority of total traffic

- Data tends to be **bursty**
- Static bandwidth allocation is **not** efficient

Optical Packet Switching

- A photonic packet contains a header and the payload
- Packet header is processed all-optically at each node and switched to the next hop
- Pros:
  - Statistical multiplexing of data
  - Suitable for bursty traffic
- Cons:
  - Very fast switching speeds (nanoseconds)
  - Synchronization

Optical Burst Switching

- Multiple IP packets assembled into a burst
- An out-of-band control header transmitted ahead of each data burst

- Pros:
  - Statistical multiplexing of data
  - Suitable for bursty traffic
  - Low data-transfer latency
  - Electronic control plane (practically feasible)
  - Optical data plane (high-speed)

Motivation for OBS

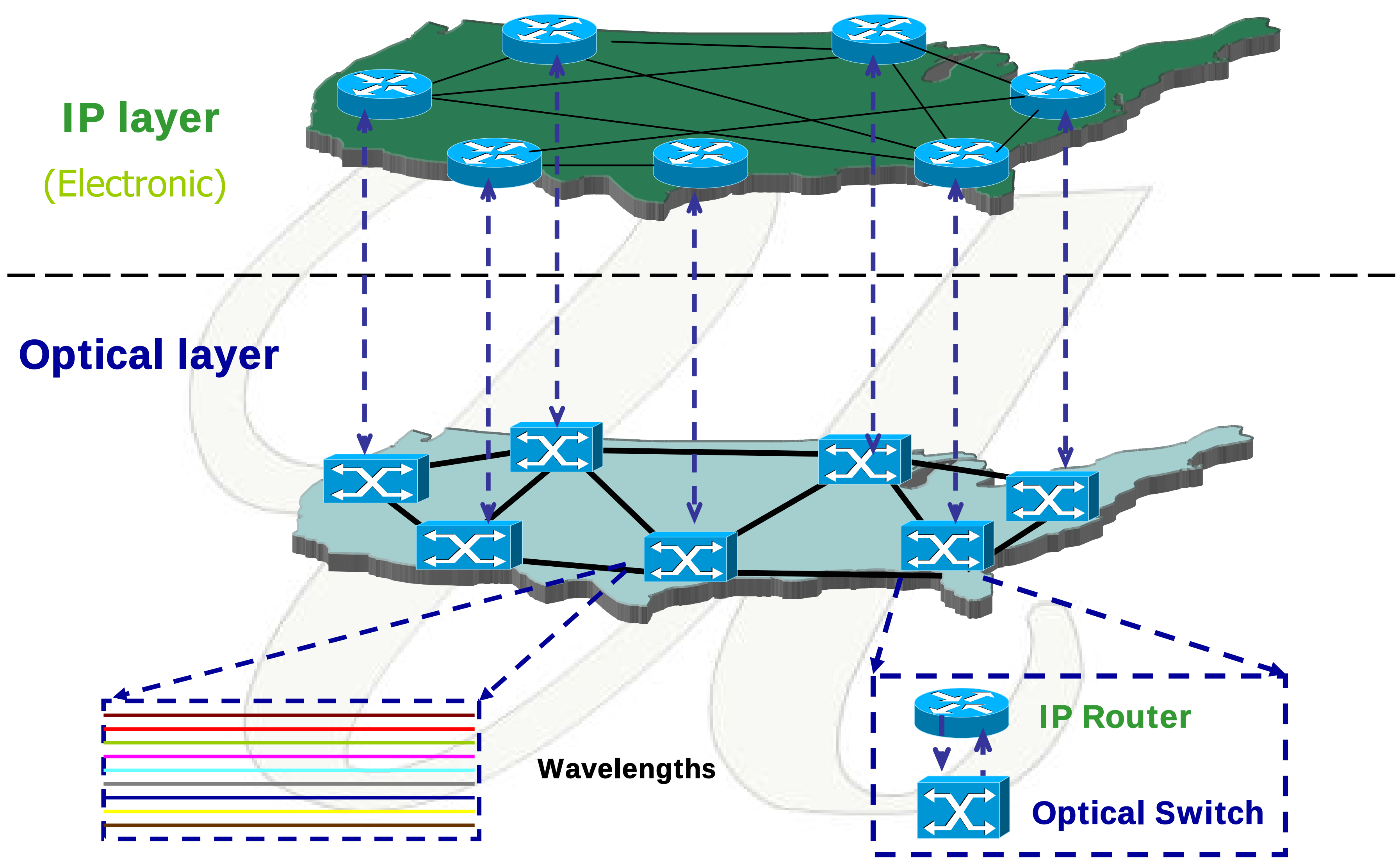
Optical Switching Paradigm	Bandwidth Utilization	Setup Latency	Switching Speed Req.	Proc. / Sync. Overhead	Traffic Adaptively
Optical Circuit Switching	Low	High	Slow	Low	Low
Optical Packet Switching	High	Low	Fast	High	High
Optical Burst Switching	High	Low	Medium	Low	High

OBS combines the best of the two while avoiding their shortcomings

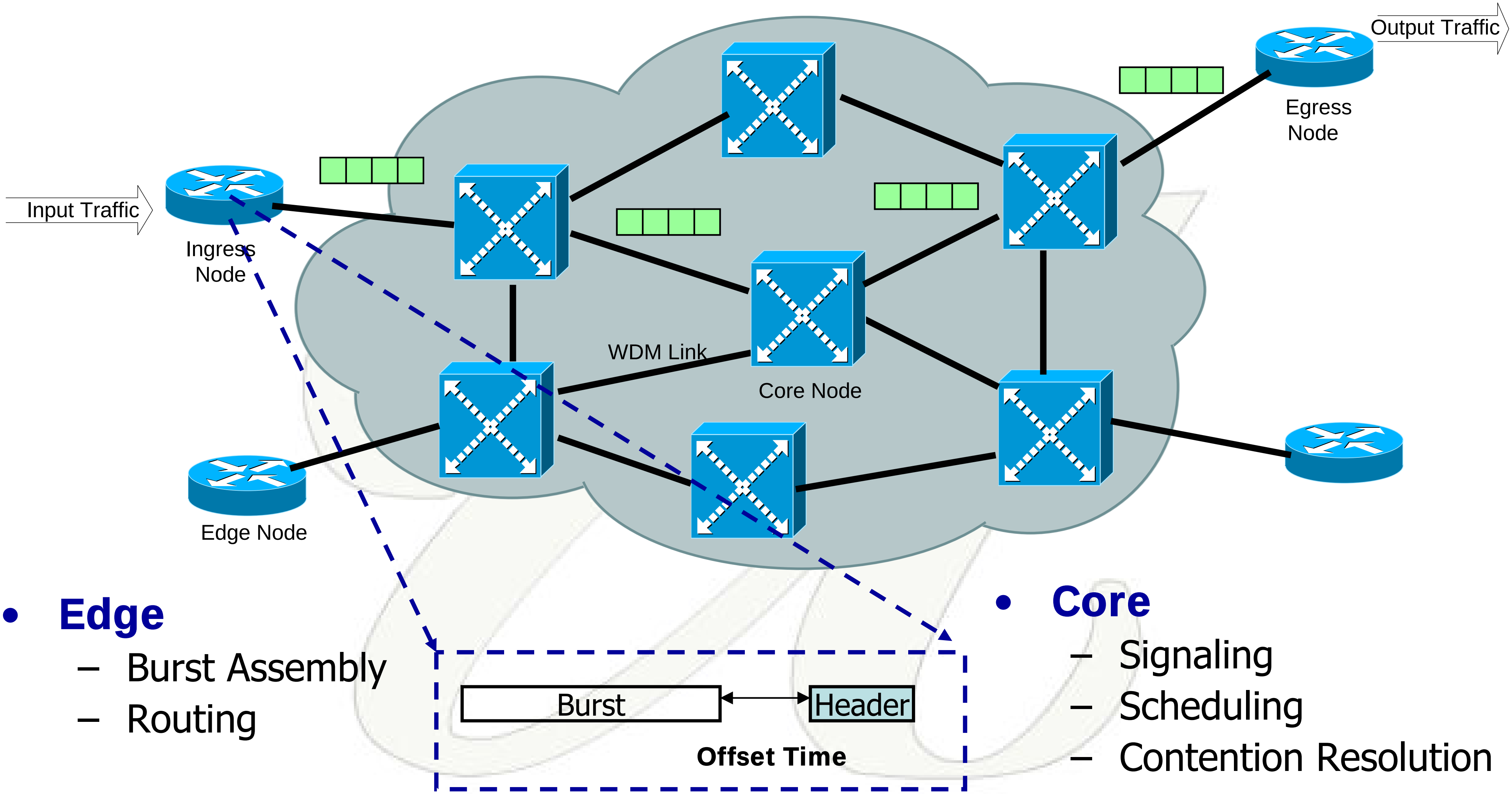
Presentation Outline

- Introduction to Optical Transport Paradigms
- Optical Burst Switching
- Reliable Data Transport in OBS
  - Loss Minimization Mechanisms
  - Loss Recovery Mechanisms
- Conclusion and Future Work

Layered Network Model

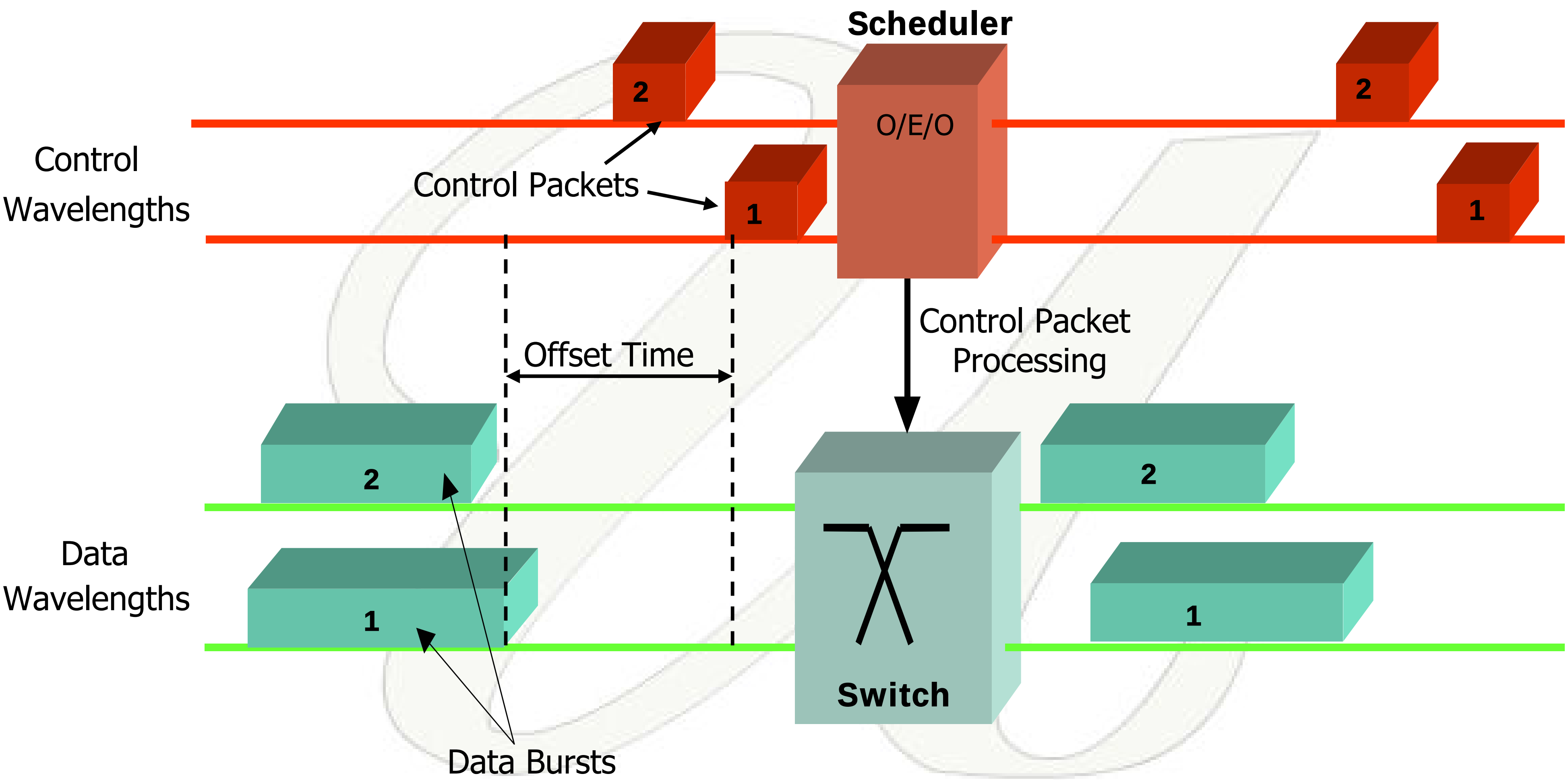


OBS Network Architecture



- **Edge**
  - Burst Assembly
  - Routing
- **Core**
  - Signaling
  - Scheduling
  - Contention Resolution

OBS Node Architecture



Adopted from Qiao

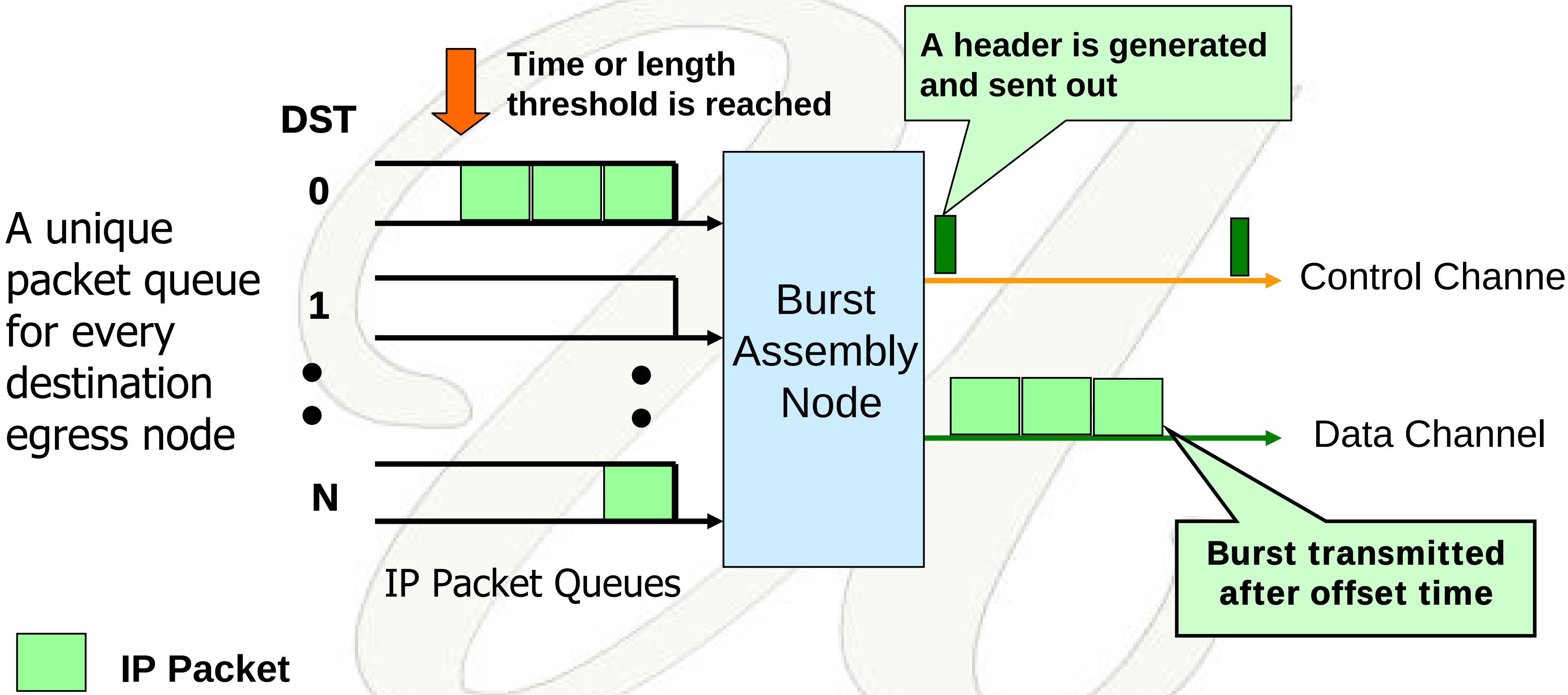
Burst Assembly

- Aggregate multiple (IP) packets going to the same destination into a single burst
- Assembly Mechanisms: Timer-based and Threshold-based

- Timer-based assembly:
  - After a fixed timer interval, all the packets in the queue are framed into a single burst
- Threshold-based assembly:
  - After a fixed length threshold is reached, all the packets in the queue are framed into a single burst.

Burst Assembly

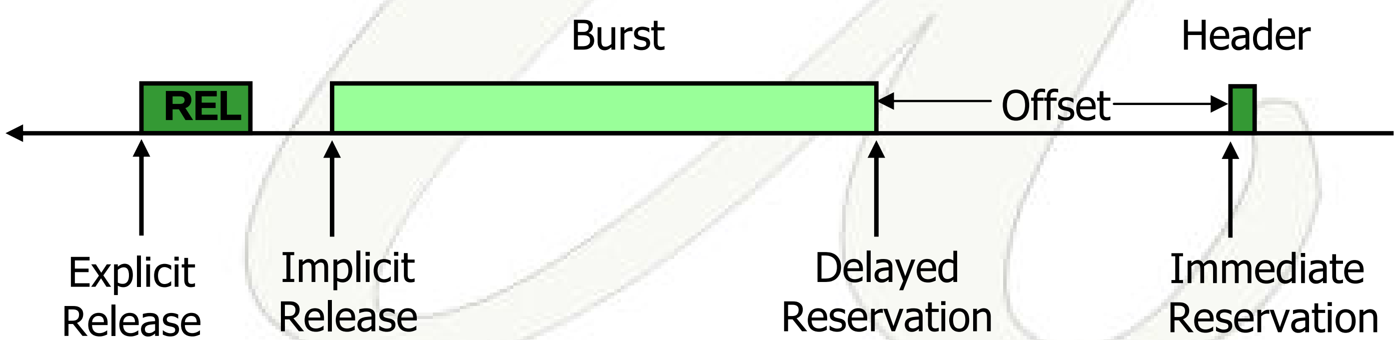
Aggregate multiple (IP) packets going to the same destination into a single burst



Adopted from Qiao

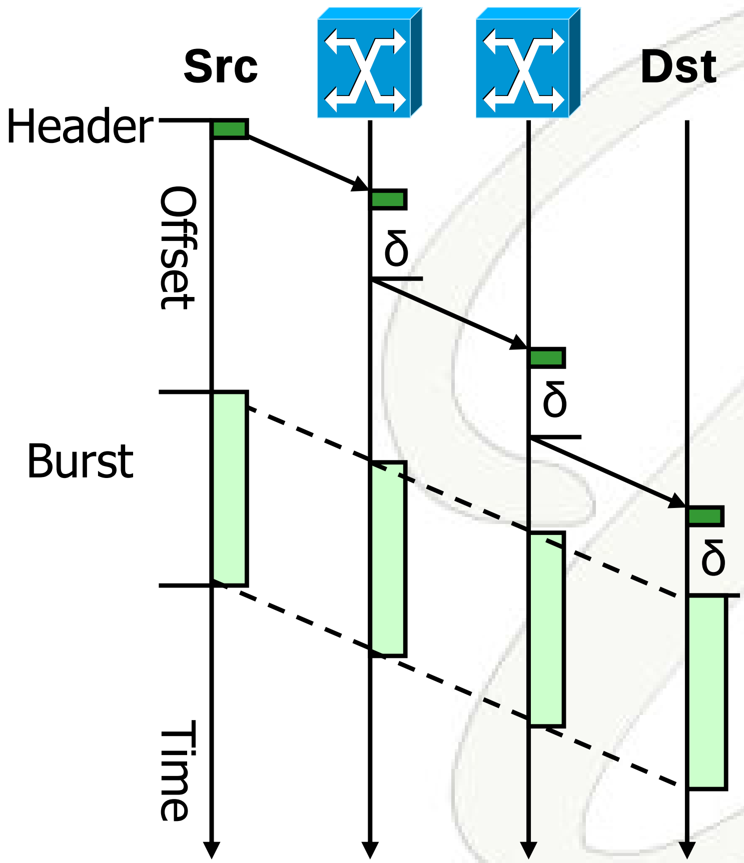
Signaling Technique

- One-way based (un-acknowledged) signaling
- **Reservation Mechanism**: Based on the start of the reservation
  - **Immediate Reservation**: Immediately after the control heater
  - **Delayed Reservation**: At the start of the burst
- **Release Mechanism**: Based on the release of the reservation
  - **Implicit Release**: based on burst length information
  - **Explicit Release**: explicit release control packet used



Tradeoff: Efficiency vs. Simplicity

Just-Enough-Time (JET) Signaling

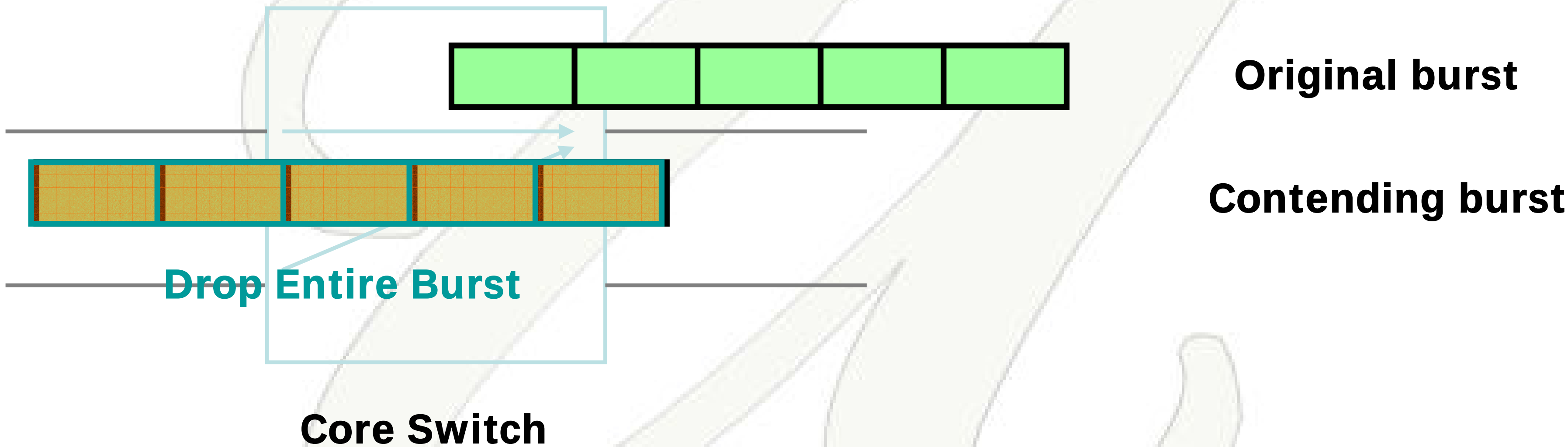


- Delayed Reservation and Implicit Release
- Header contains burst length, offset time, source, destination
- Offset time necessary for processing of header at intermediate nodes **without buffering** the data burst
- **Just-In-Time (JIT)**: Immediate Reservation and Explicit Release

Data Loss in OBS: Burst Contentions

- Contention occurs when more than one burst attempts to go out of same output port (or wavelength) at the same time

- Unique to all-optical networks
  - Traditional networks employ electronic buffering to resolve contentions
  - Lack of optical buffers (cannot store light)



- Drop Policy:
  - One of the bursts will be dropped in its entirety
  - Even though overlap between the bursts may be minimal

TCP over OBS

- Transmission Control Protocol (TCP)
  - Majority Internet applications depend on TCP for reliable data transmission
  - TCP assumes packet loss in the network is due to network congestion
  - TCP congestion avoidance mechanisms will reduce sending rate in the event of a packet loss

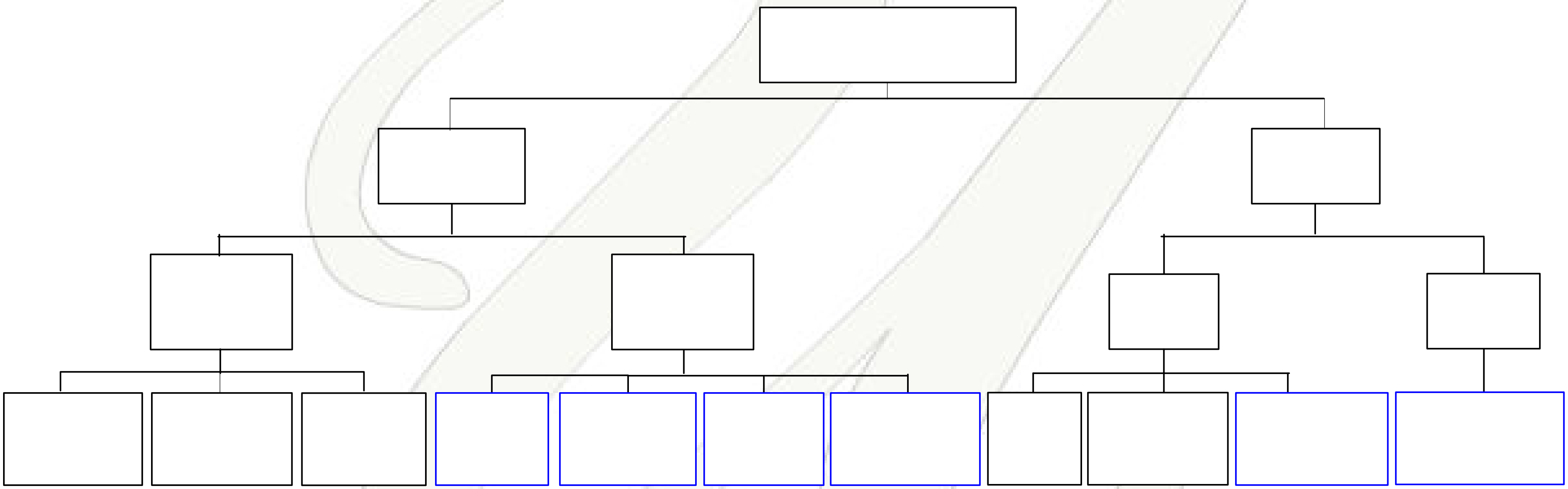
- OBS
  - Random burst loss occurs even when the network is NOT congested

- TCP over OBS
  - TCP falsely reduces sending rate even when the network is NOT congested (False Timeout)
  - Significantly degrade throughput

Presentation Outline

- Introduction to Optical Transport Paradigms
- Optical Burst Switching
- Reliable Data Transport in OBS
  - Loss Minimization Mechanisms
  - Loss Recovery Mechanisms
- Conclusion and Future Work

Toward a Reliable OBS

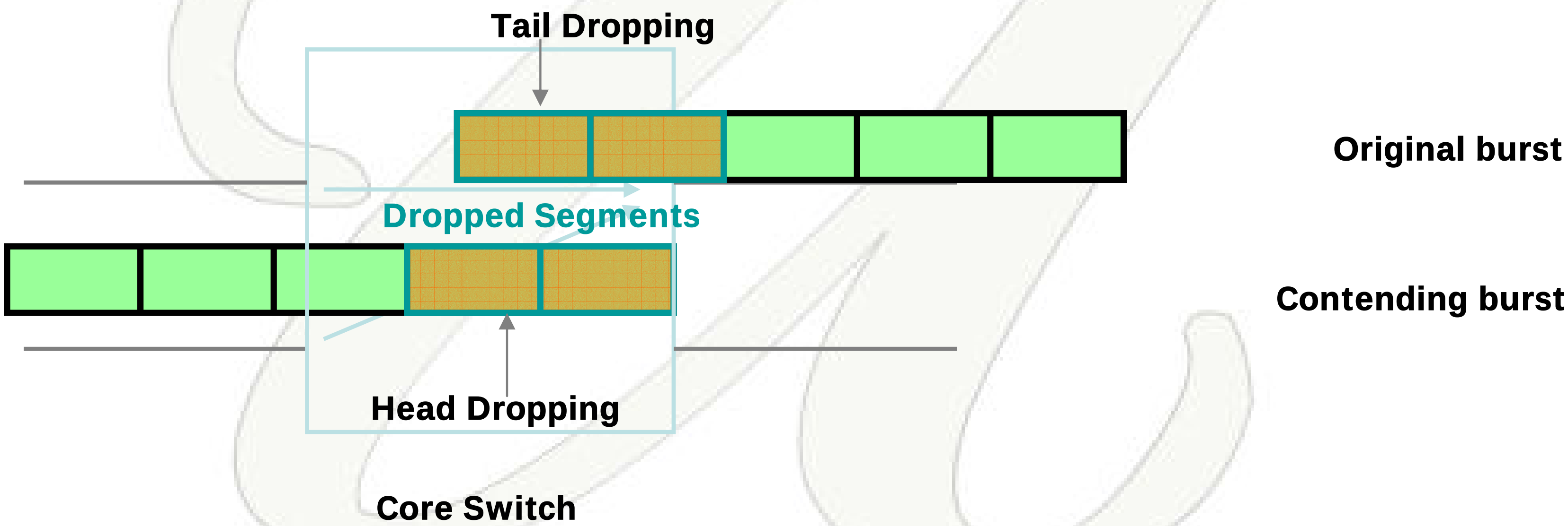


Traditional Contention Resolution

- Optical Buffering (FDLs)
  - Achieved through Fiber Delay Lines
  - Issues: Limited buffer capacity and additional hardware cost
- Wavelength Conversion
  - Converting the wavelength of an incoming channel to another wavelength at the outgoing channel
  - Issues: Additional hardware cost
- Deflection Routing
  - Deflect contending bursts to alternate port
  - Issues: Higher delay and out-of-sequence delivery
  - No additional hardware cost

Burst Segmentation

- When contention occurs, only overlapping segments are dropped
- Two Approaches: Head Dropping and Tail Dropping
- Details: Vokkarane and Jue [IEEE ICC 2002, New York]



Evaluation Criteria

- Evaluation of proposed policies
  - Average end-to-end packet loss probability
  - Average number of hops (delay)
  - TCP Throughput

- Numerical Analysis
  - Analytical modeling
  - Simulation results

Burst Segmentation: Analytical Loss Model

- Burst Arrivals: Poisson Process
- M/G/1/1 Queueing Model

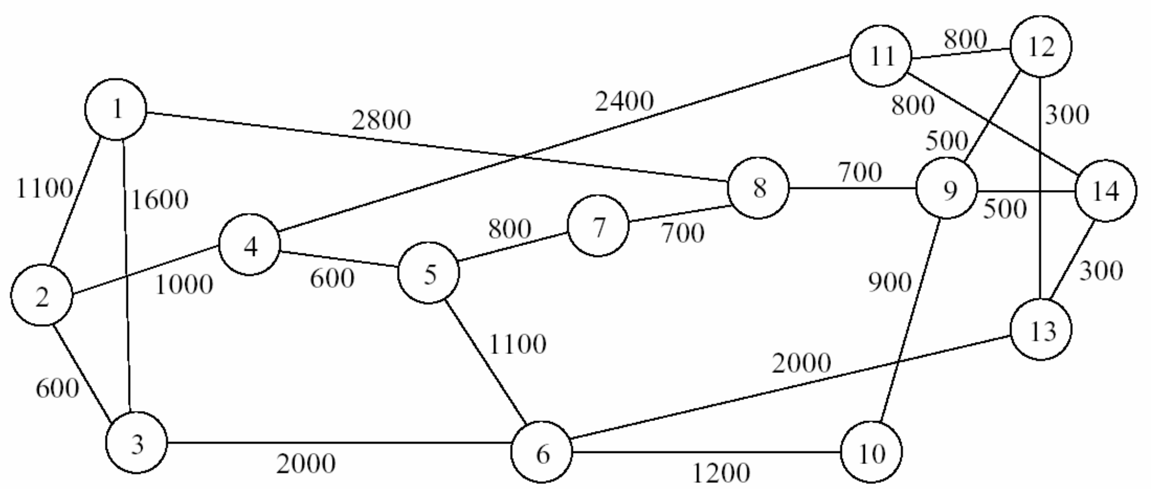
Burst Length Distribution (After *k* hops):

$$G_{l_k^{sd}}(t) = 1 - (1 - G_{l_{k-1}^{sd}}(t))(e^{-\lambda_{l_k^{sd}}})$$
$$= 1 - (1 - G_{l_0^{sd}}(t))e^{-\left(\sum_{i=1}^{k} \lambda_{l_i^{sd}}\right)t}.$$

End-to-End Packet Loss:

$$P_{\text{loss}} = \sum_{s} \sum_{d} \frac{\lambda^{sd}}{\lambda} P_{\text{loss}}^{sd}.$$

Simulation Network

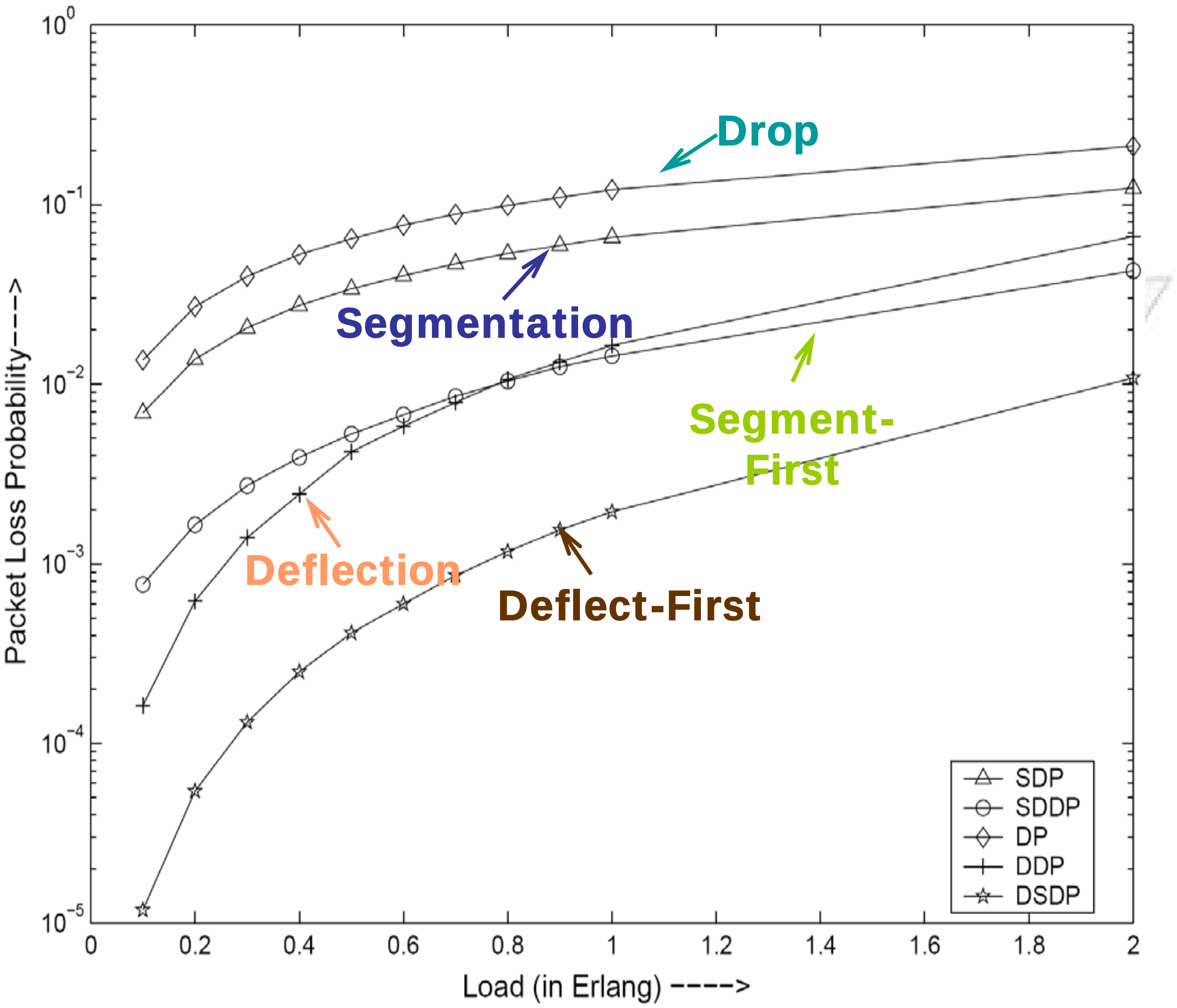


14-node NSFNET

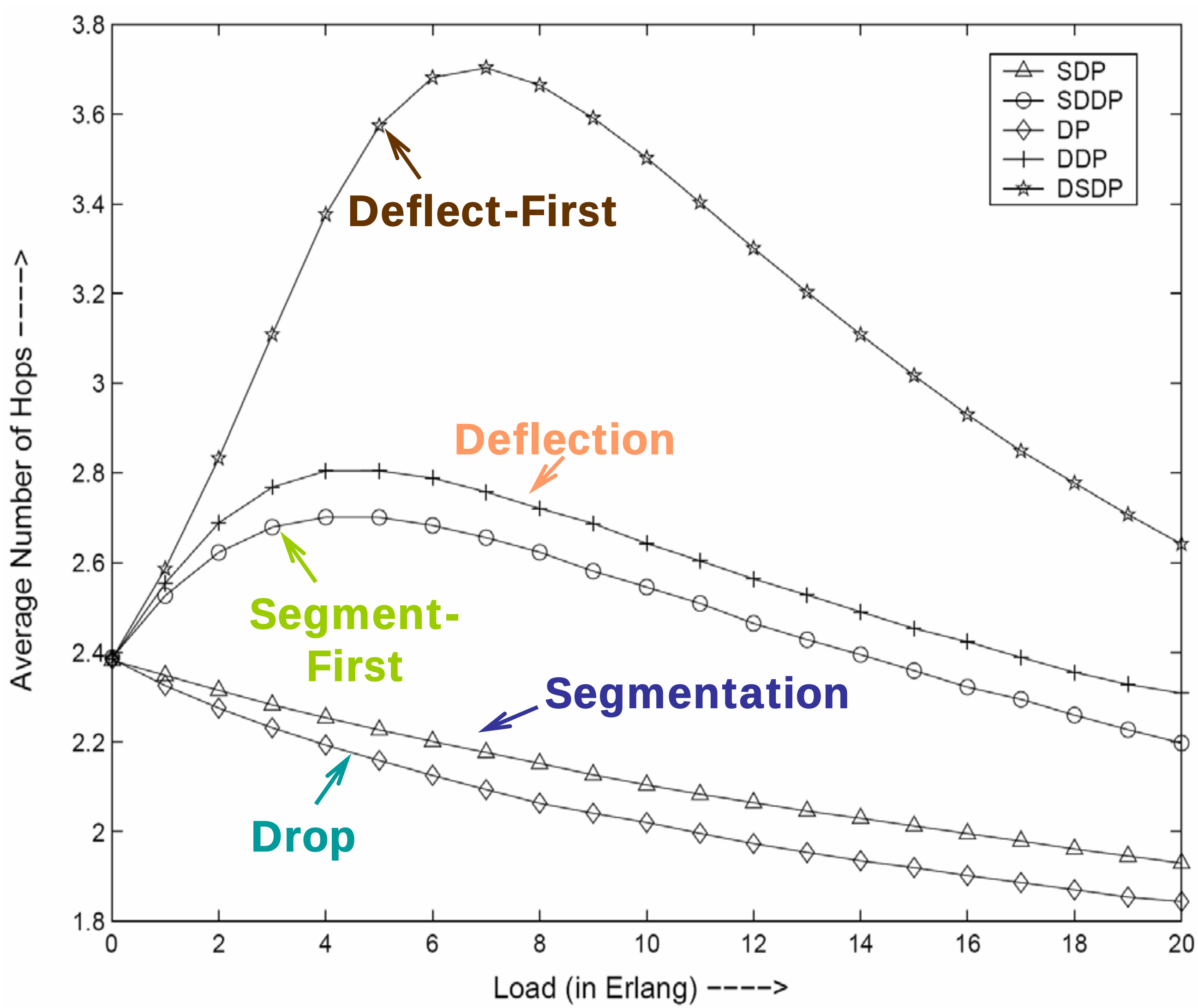
Assumptions

Burst Arrivals	Poisson
Average Burst Length	100 µs (exponentially dist.)
Link Transmission Rate	10 Gb/s
Packet Length	1500 Bytes
Switching Time	10 µs
Optical Buffering	NO
Wavelength Conversion	NO

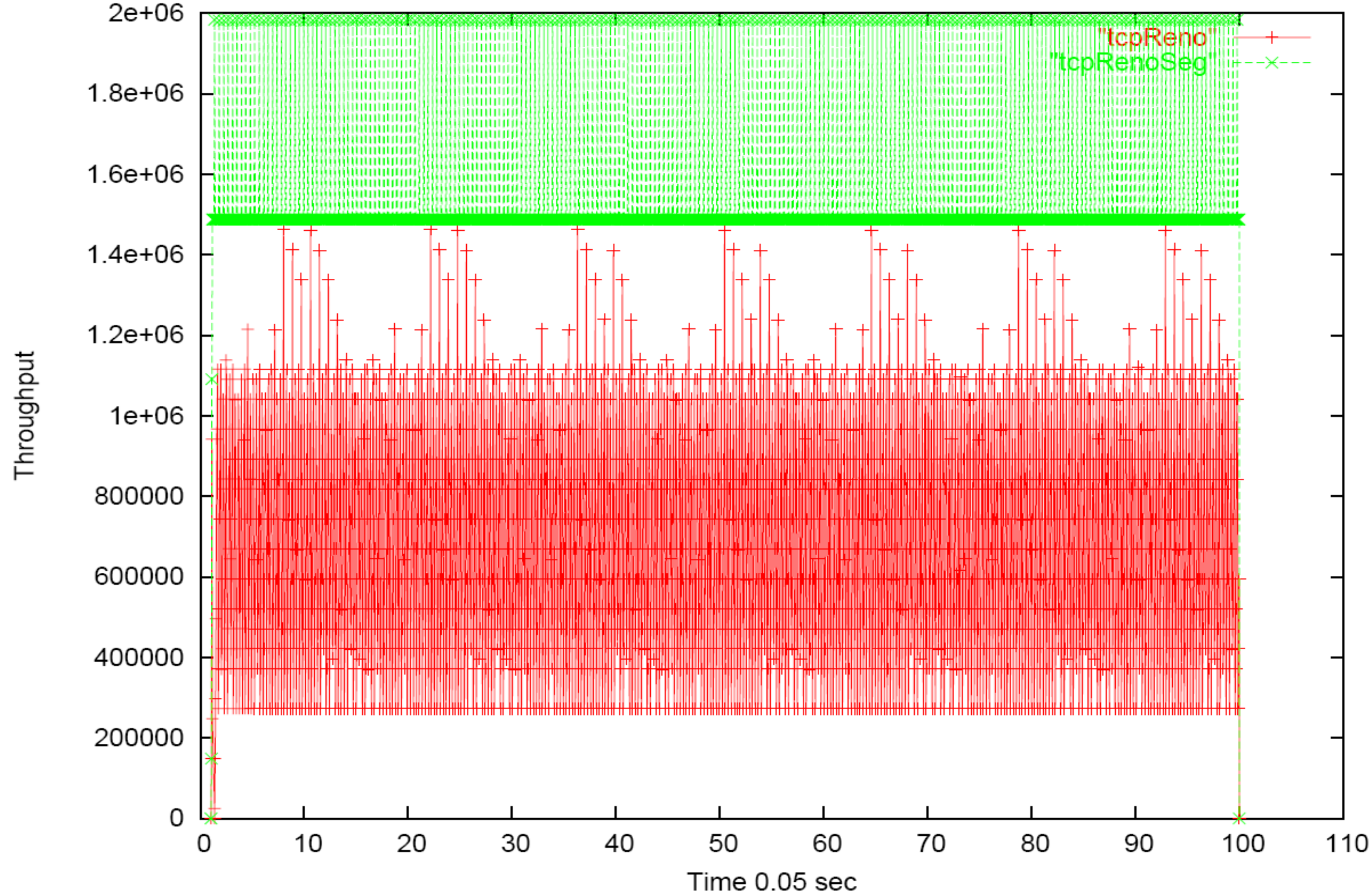
Packet Loss Performance



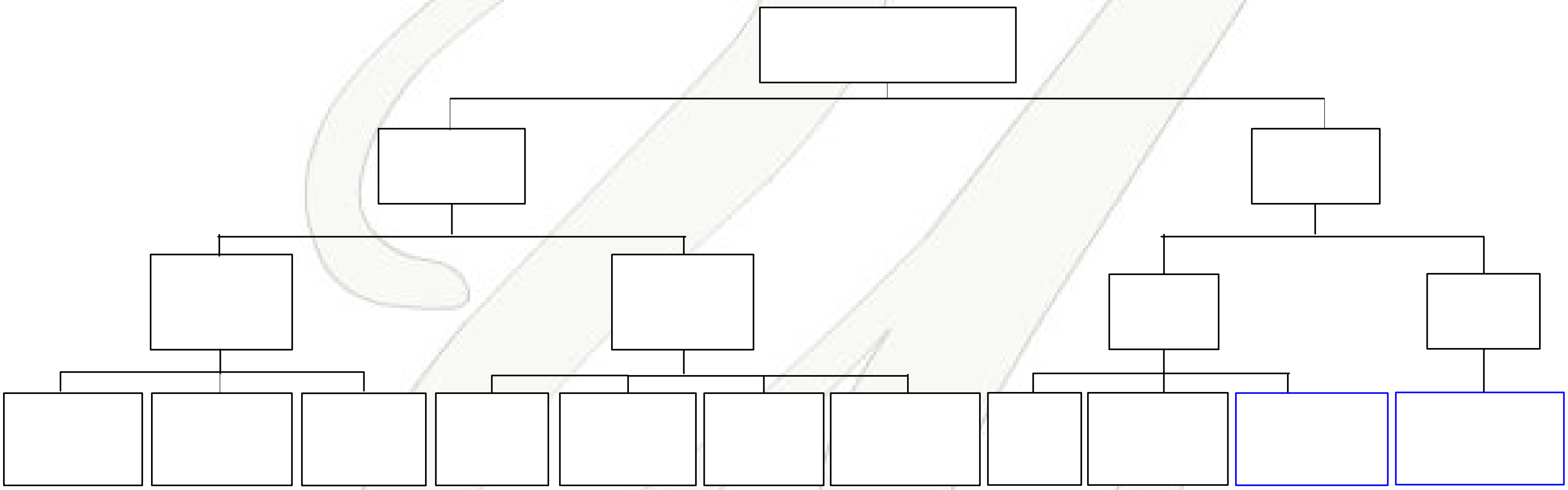
Average Number of Hops (~Delay)



TCP Throughput

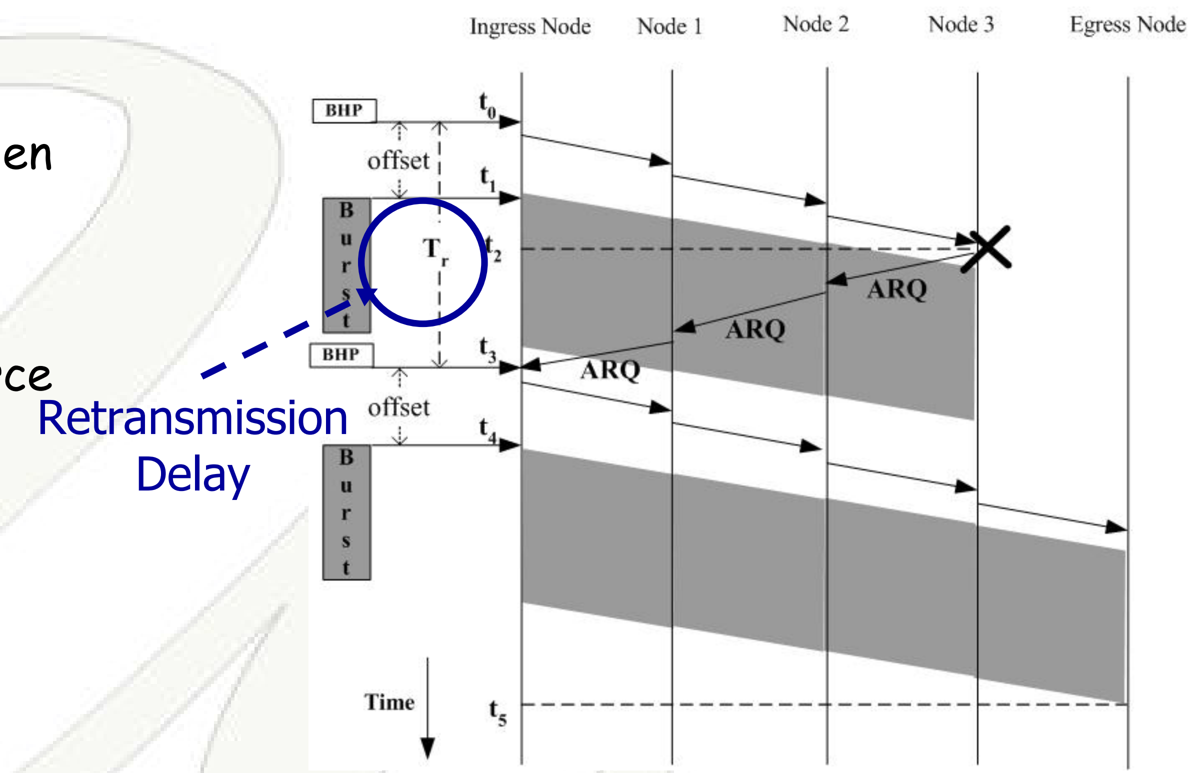


Toward a Reliable OBS



Reactive Loss Recovery: Burst Retransmission

- Objective
  - To recover from burst loss when network is not congested
- Basic idea
  - Retransmit lost bursts at source nodes
  - Stop retransmission when Tr > Delay Constraint (δ)



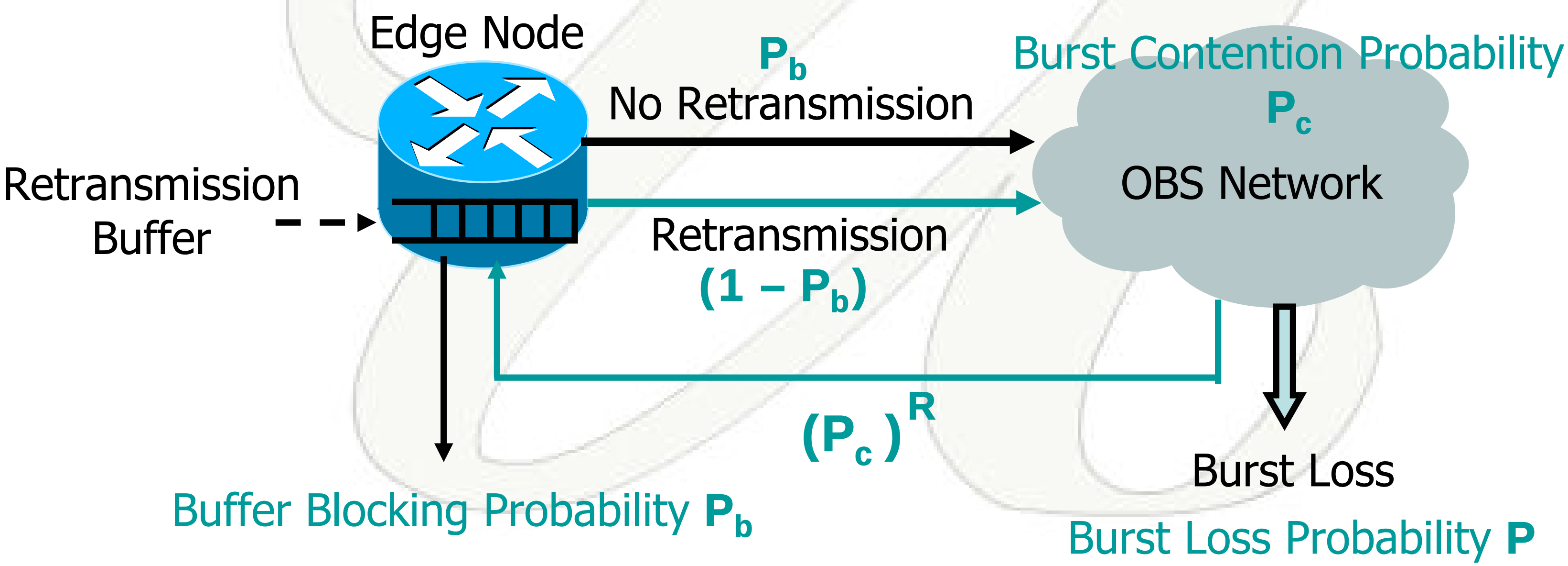
Analysis for Burst Retransmission

Objective

- Analyze the average burst loss probability in the network

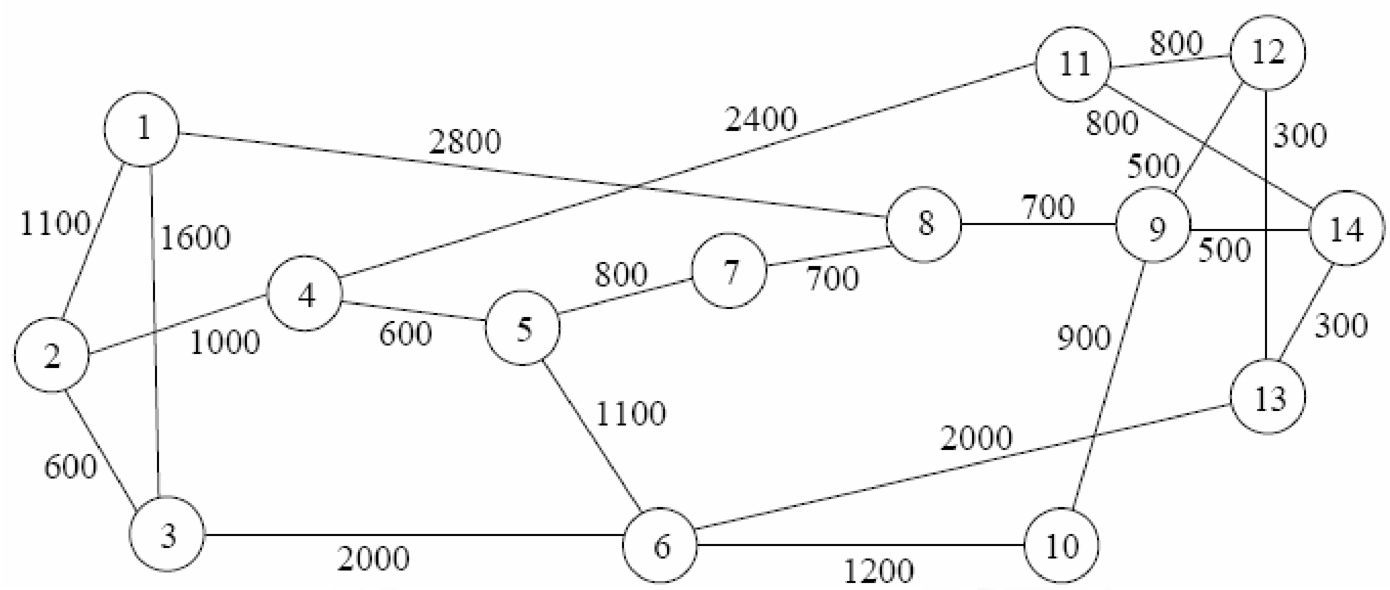
Basic idea

- No retransmission for bursts blocked by retransmission buffers
- $P = P_b P_c + (1 - P_b)(P_c)^{R+1}$

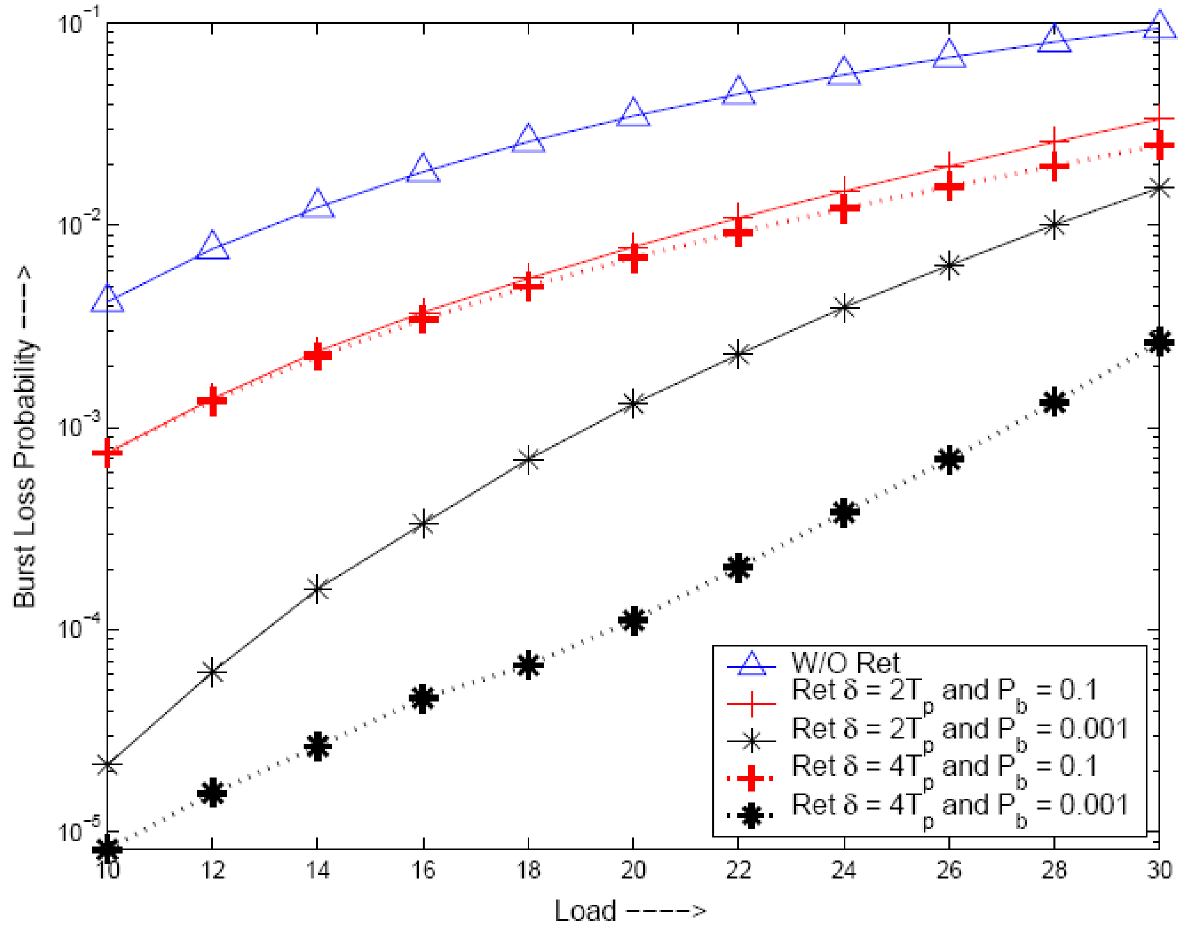


Simulation Assumptions

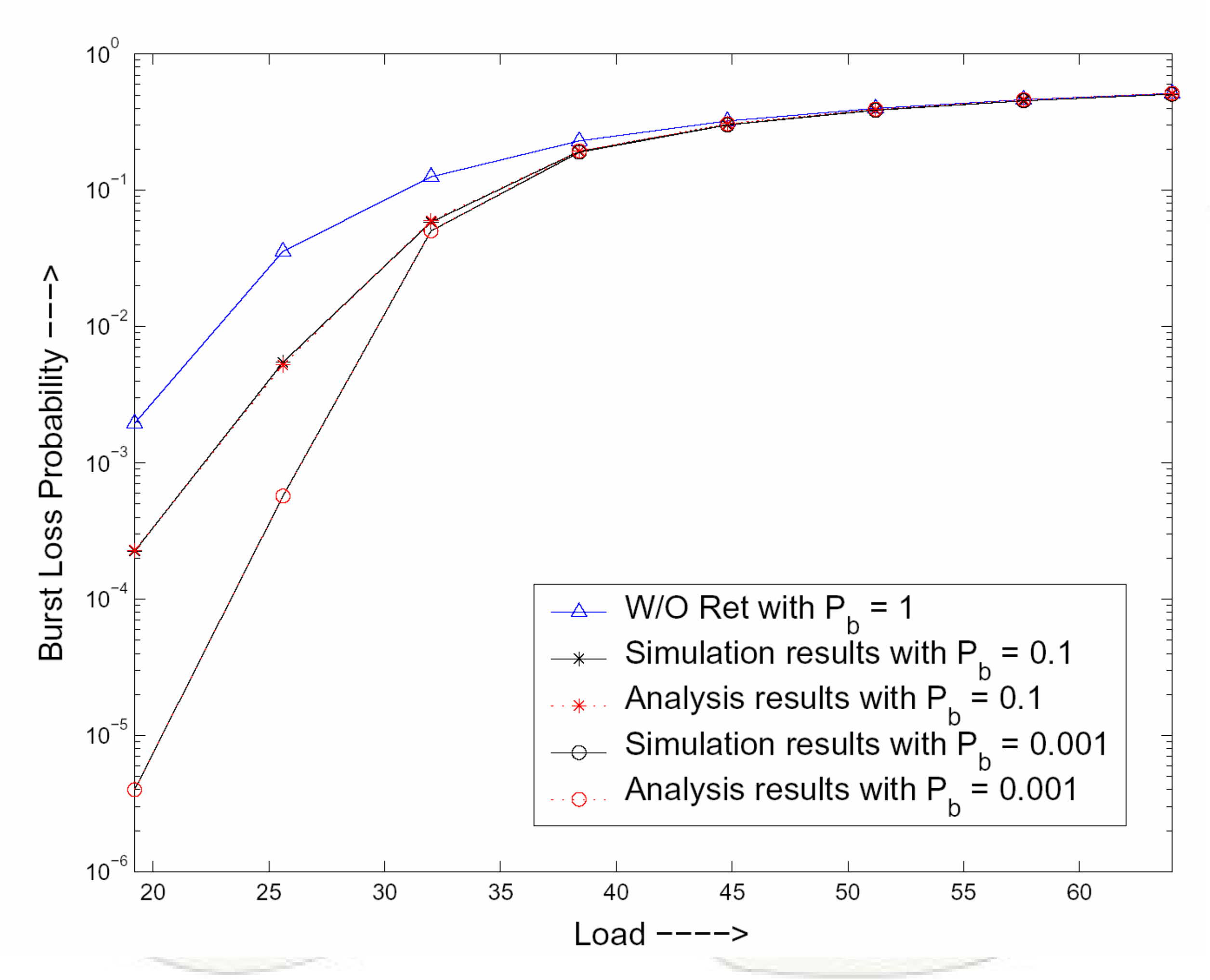
- No. of wavelengths on each link is 4
- Transmission rate on a wavelength is 10 Gb/s
- Burst arrival is Poisson
- Traffic are uniformly distributed
- Average burst length are 100μs



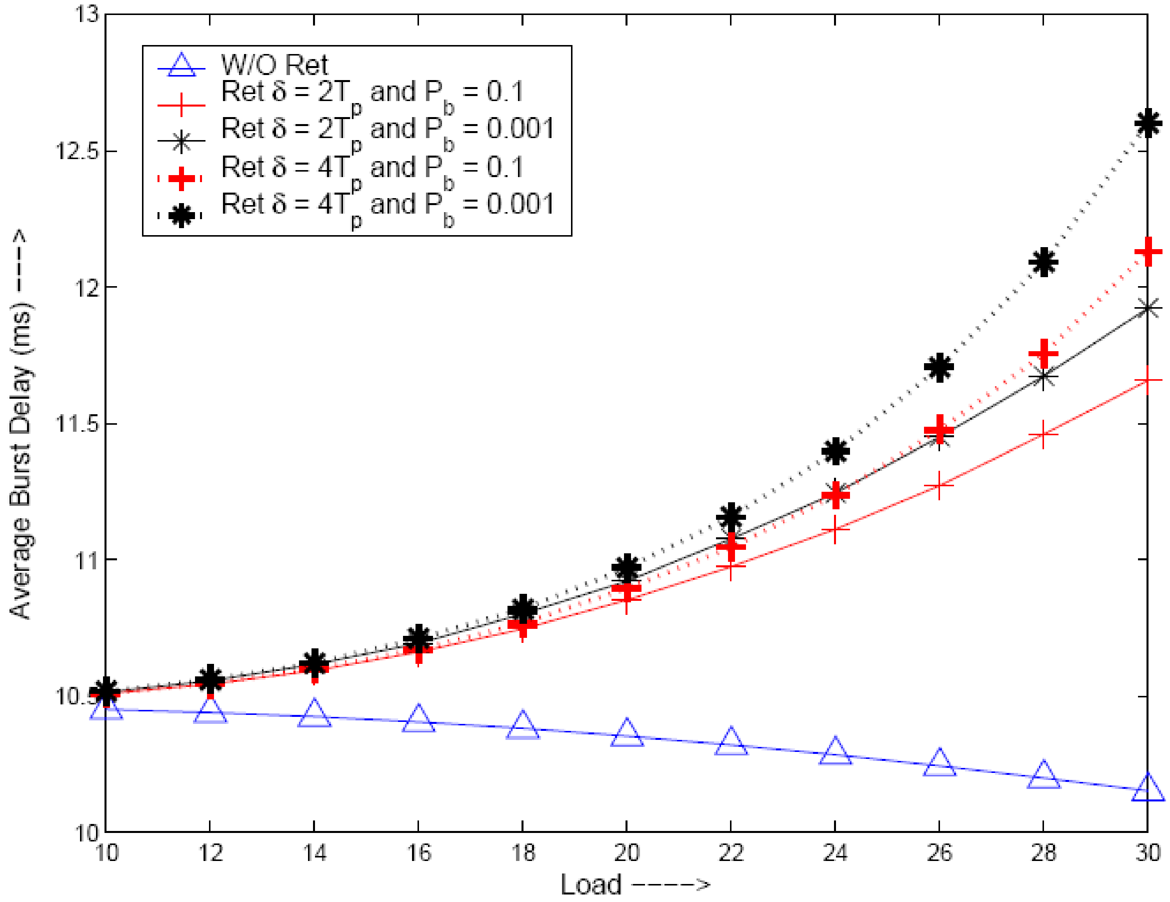
Burst Loss Probability



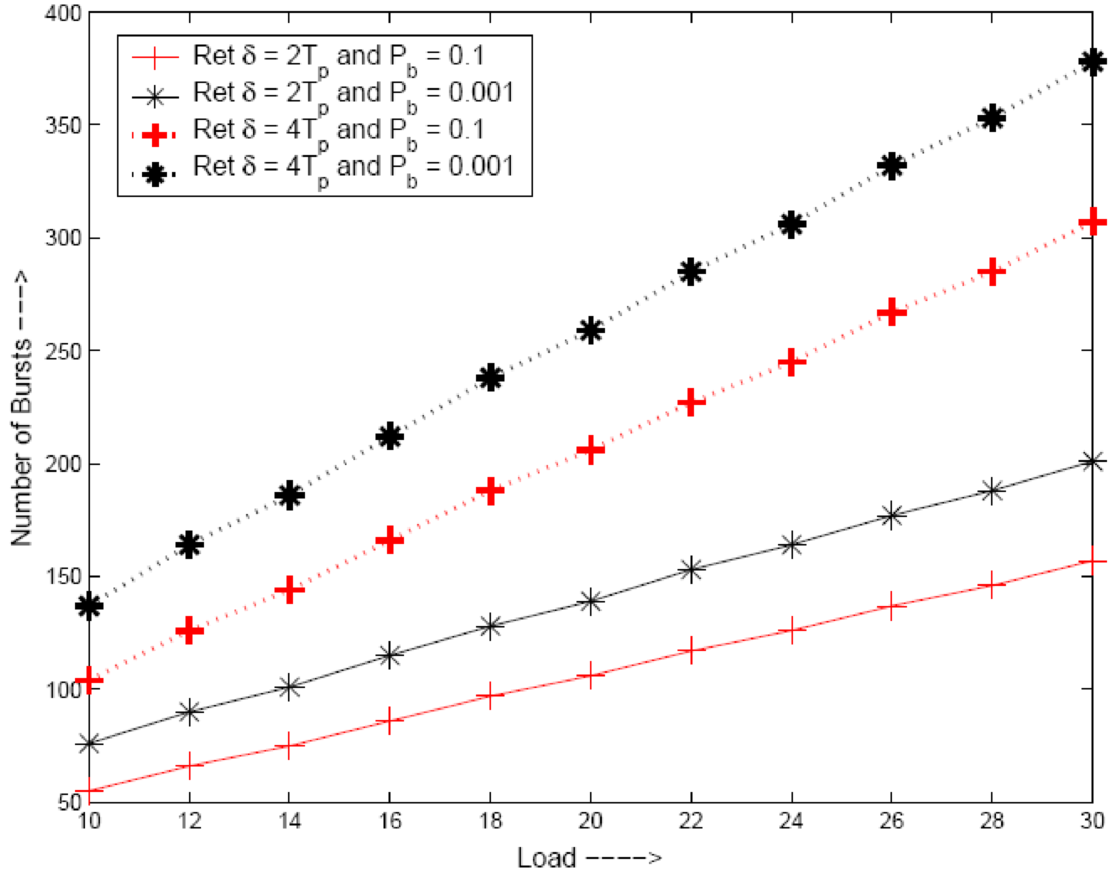
Analysis and Simulation Results for Burst Loss Probability



Average Burst Delay

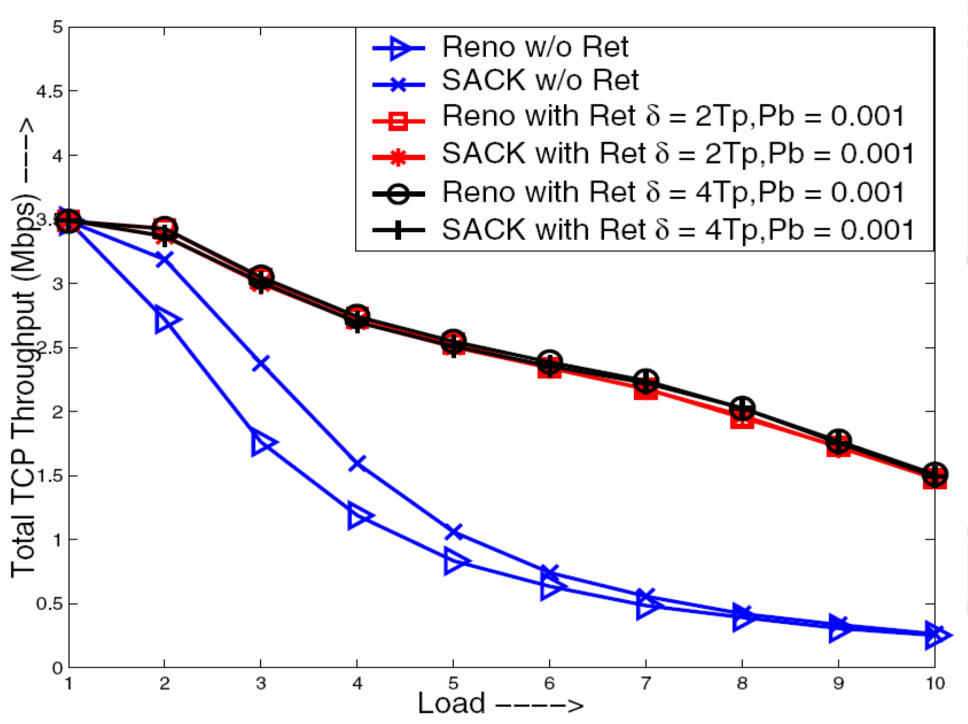


Buffer Capacity at Edge Nodes

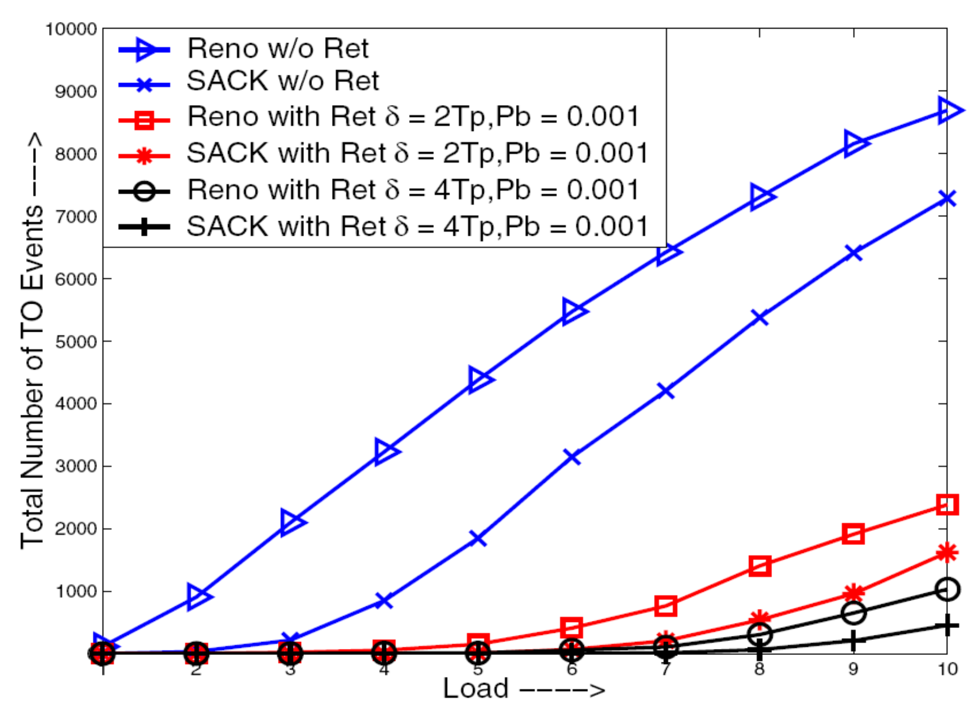


Performance of TCP Versions

Throughput



Num. of Timeouts



Burst Retransmission

- Pros
  - Reduce burst loss probability
  - Correctly indicate network congestion
  - Significantly improve TCP sending rate
- Cons
  - Additional electronic buffers at edge nodes
  - Longer delay for retransmitted bursts
  - Higher burst contention probability

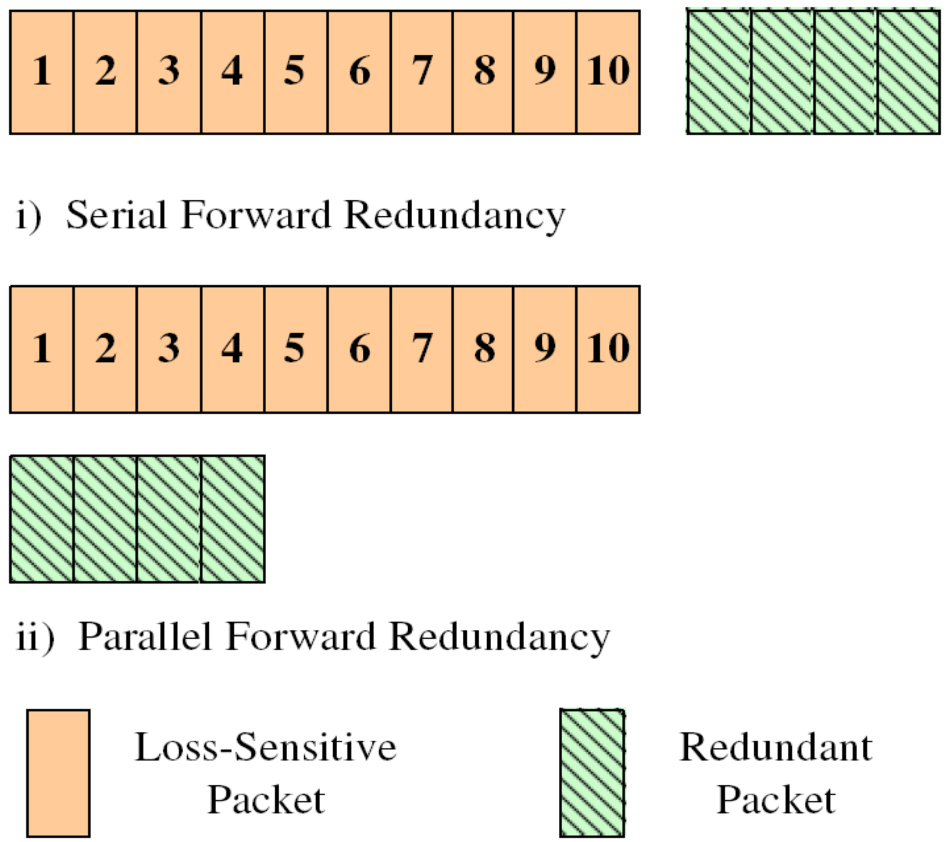
Proactive Loss Recovery: Forward Redundancy (FR)

- Some or all the original packets of a burst are copied and sent in the forward direction from source to destination
- Receiver can recover from selective packet loss in the forward direction

Policies

- Partial (< 100%) or Complete (>= 100%) FR
- Serial or Parallel FR

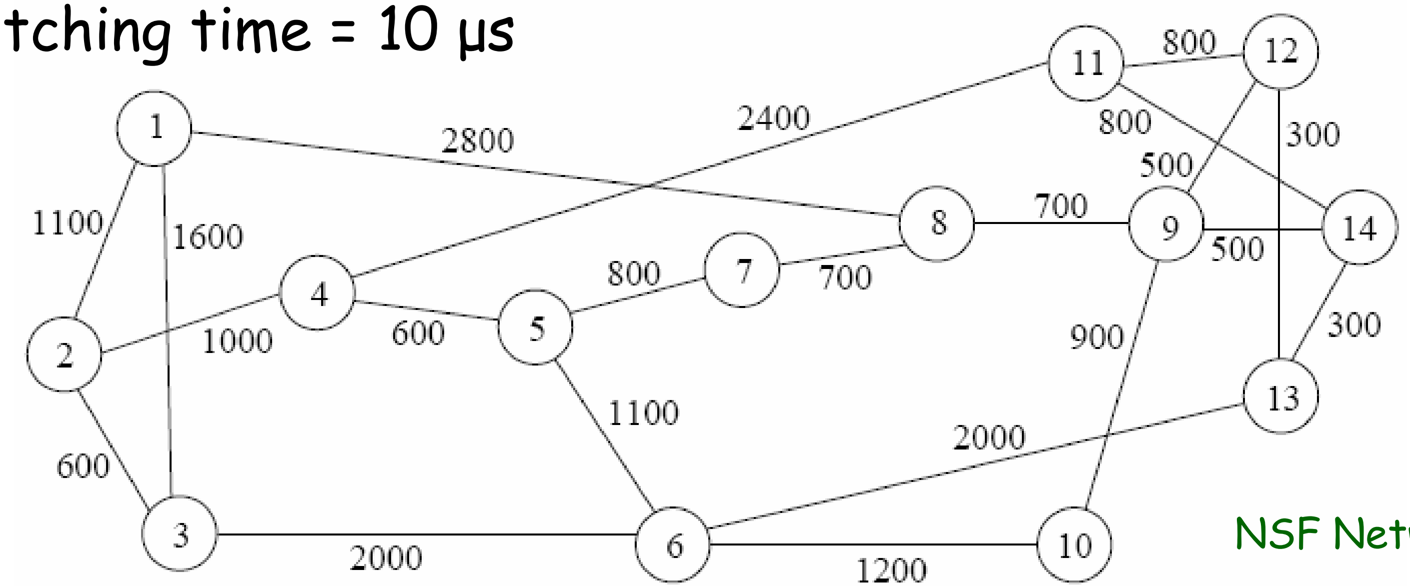
Serial Forward Redundancy (SFR)



SFR: redundant packets are placed at the tail of the original burst

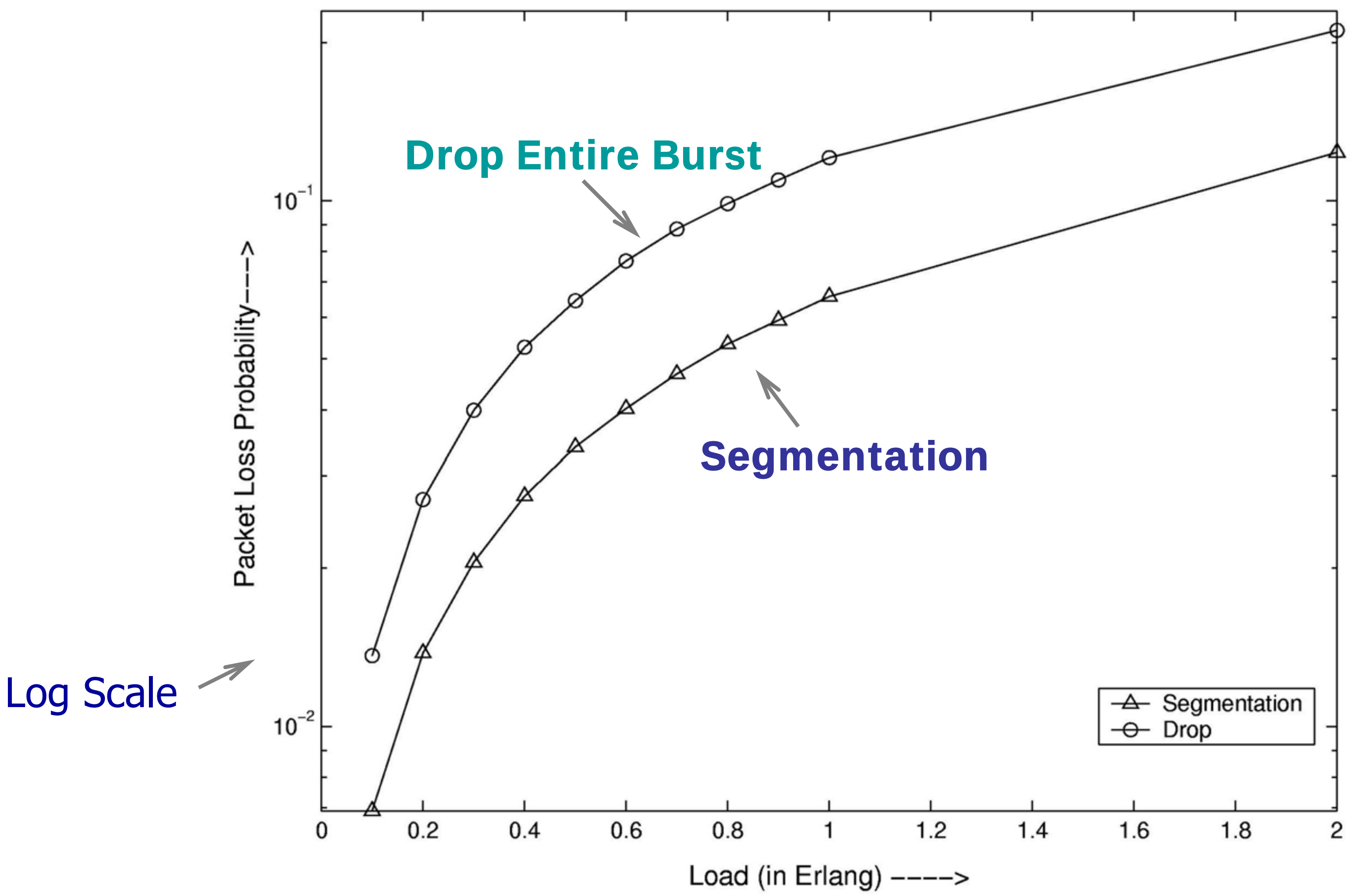
Simulation Assumptions

- No. of wavelengths on each link: 8
- Transmission rate: 10 Gbps per wavelength
- Burst arrivals: Poisson
- Packet size: 1250 byte [10 Kb]
- Fixed burst length: 100 packets [1Mb]
- Traffic: uniformly distributed
- Switching time = 10 μs

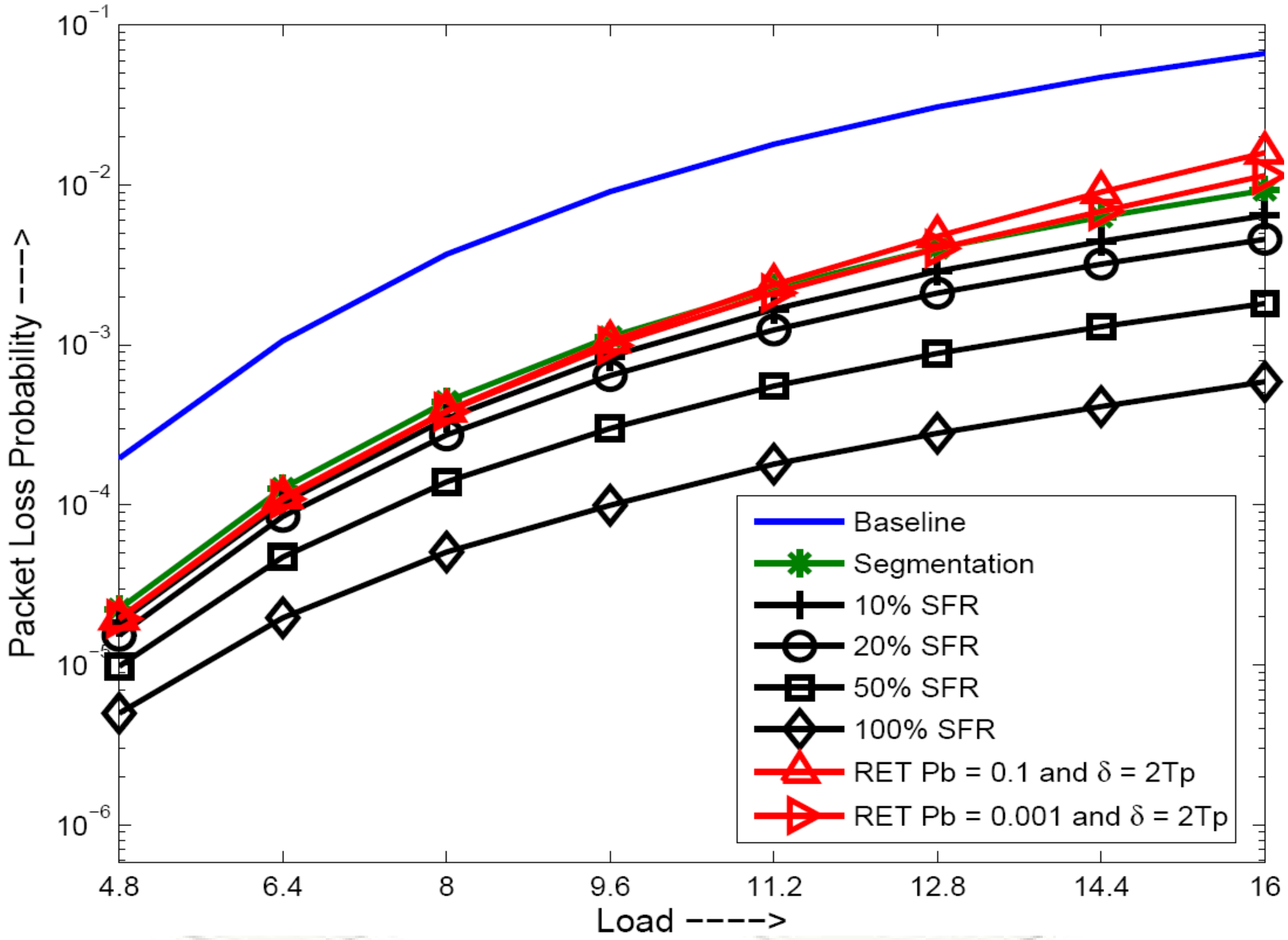


NSF Network

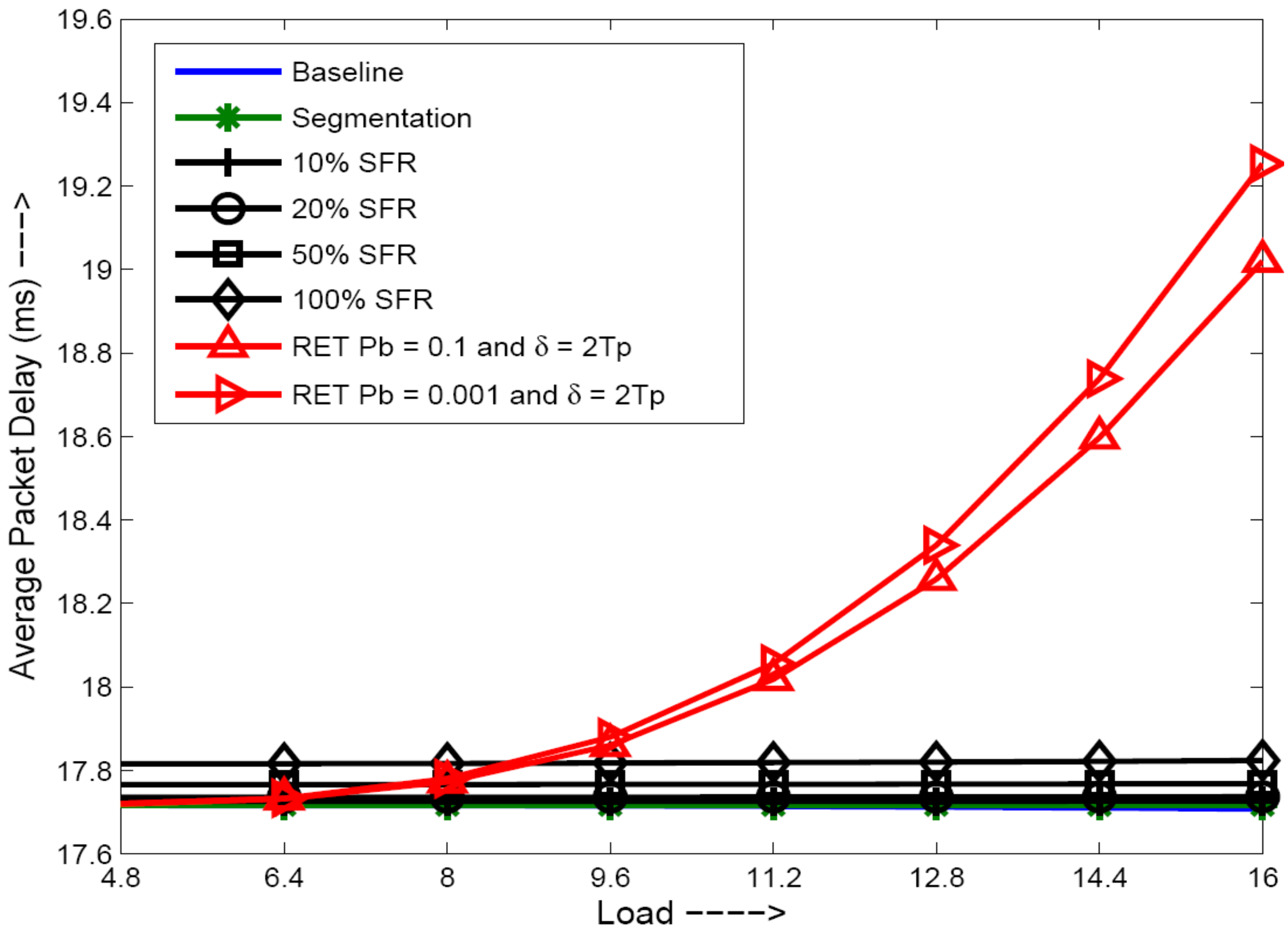
Packet Loss Performance



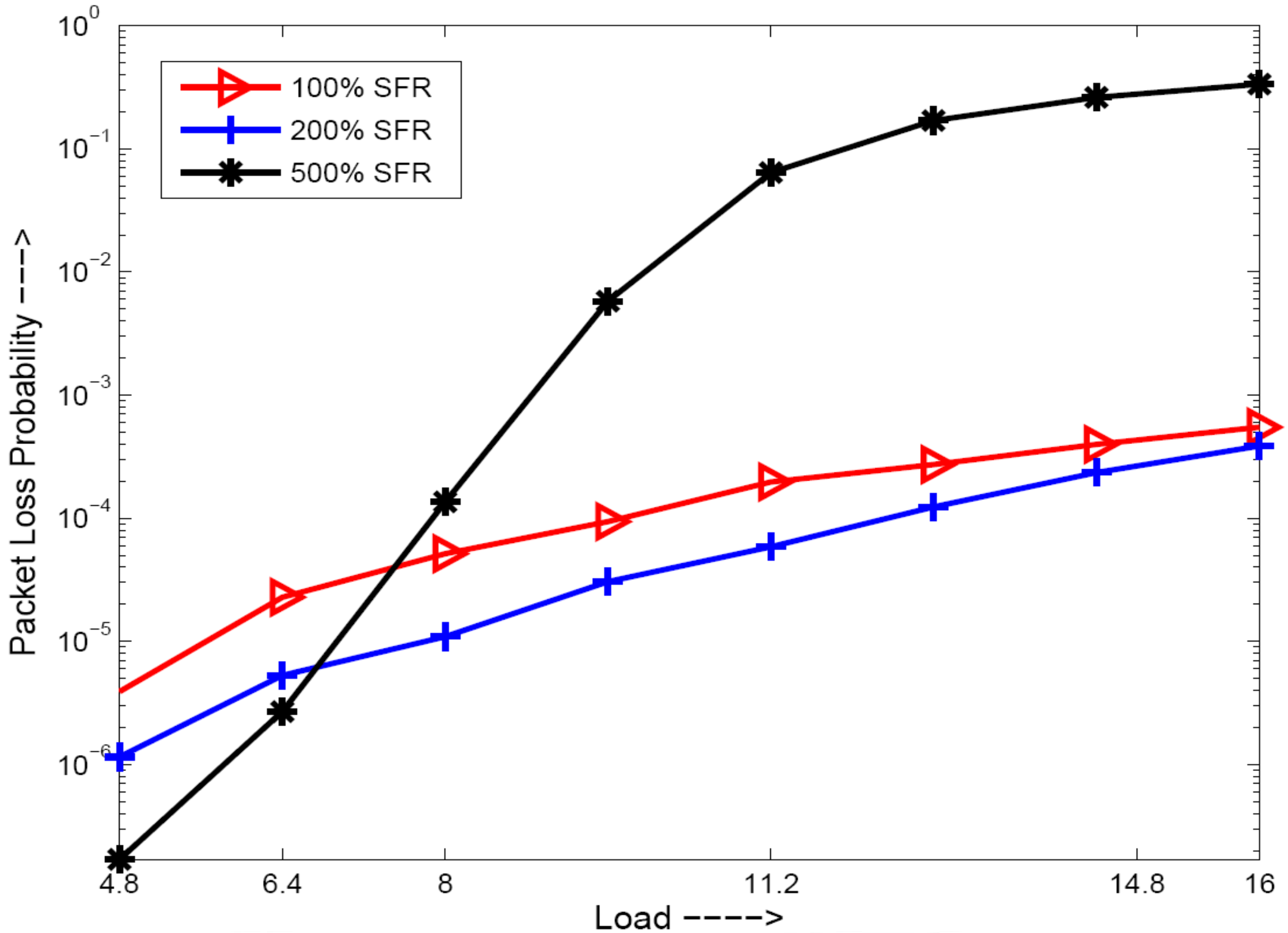
Packet Loss Probability



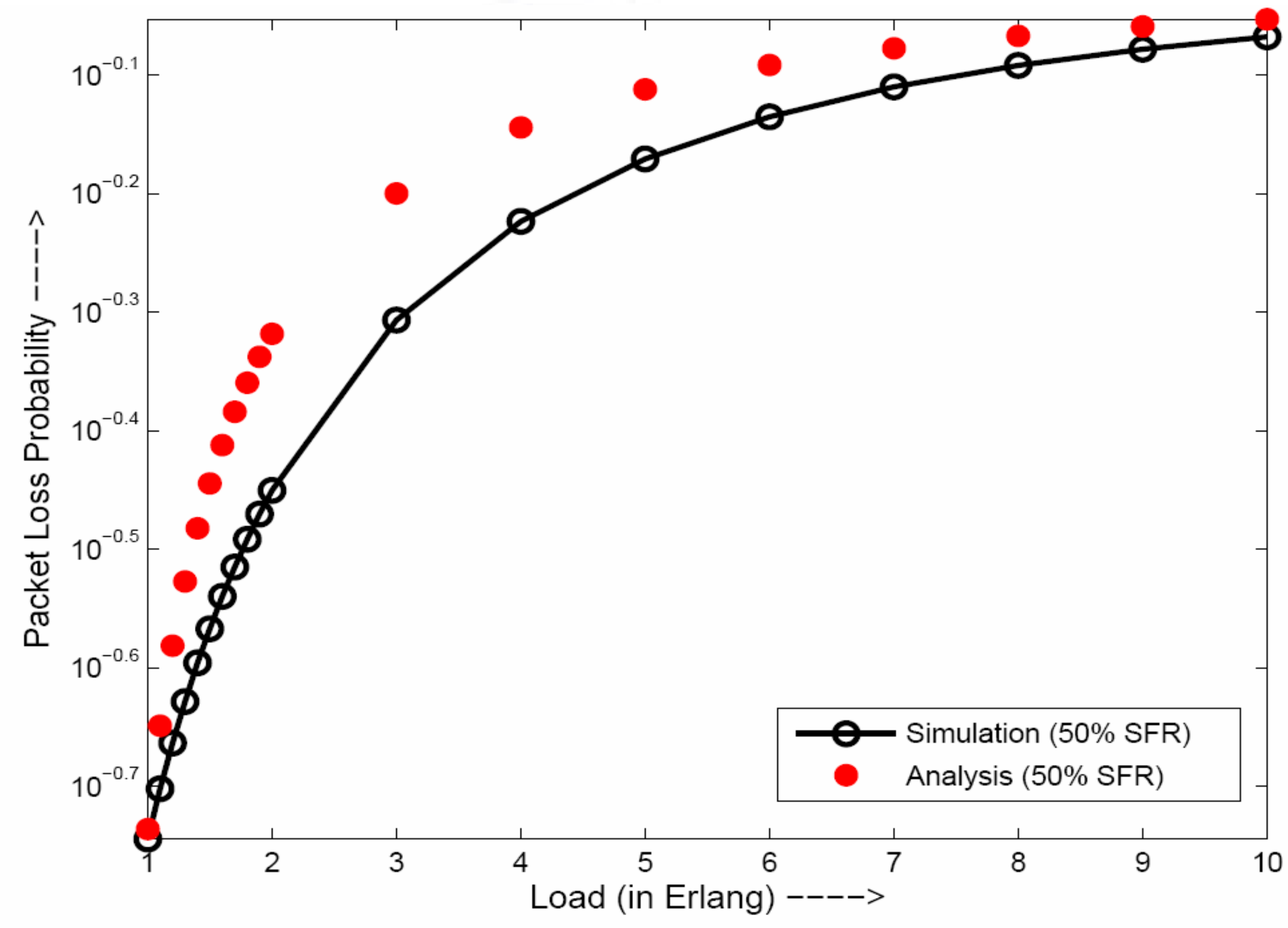
End-to-End Packet Delay



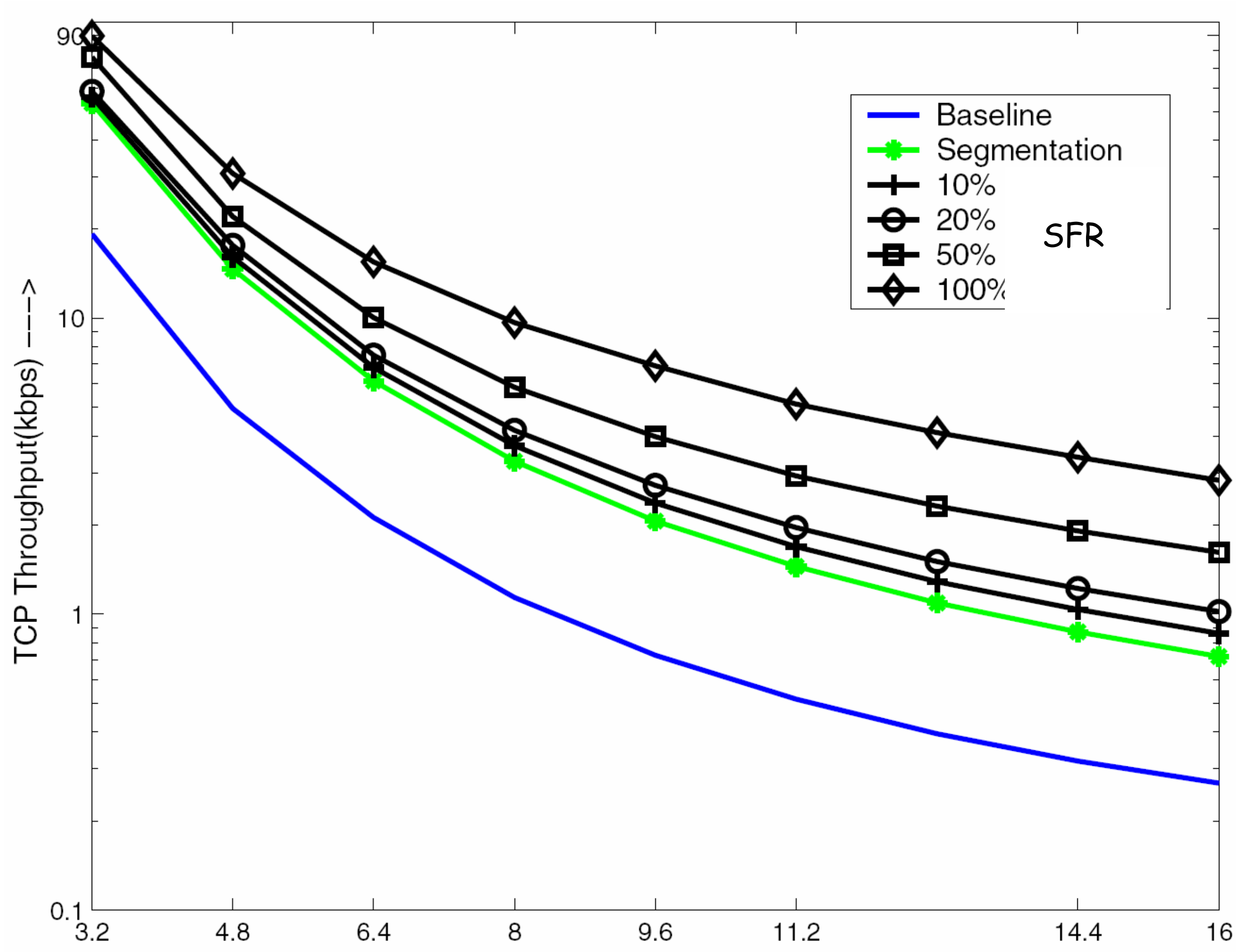
Packet Loss Probability



Analytical Loss Model Results



TCP Throughput



Presentation Outline

- Introduction to Optical Transport Paradigms
- Optical Burst Switching
- Reliable Data Transport in OBS
  - Loss Minimization Mechanisms
  - Loss Recovery Mechanisms
- Conclusion and Future Work

Conclusion

- OBS Network
  - Promising optical core data-transport paradigm
  - Suited for delay-sensitive applications
- Loss Minimization and Loss Recovery Mechanisms
  - Evaluated several new mechanisms
  - Proposed mechanisms significantly improves the reliability
- Future Work
  - Develop Dynamic mechanisms
  - Impact on newer high-speed TCP versions

Thank You

http://www.cis.umassd.edu/~vvokkarane/