Optical Burst Switching: A Multi-layered Approach to Architecture and Protocol Design

Ph.D. Defense Proposal

Farid Farahmand

03/21/2005

Advisor: Professor Jason P. Jue

Advanced Network Research Lab
The University of Texas at Dallas

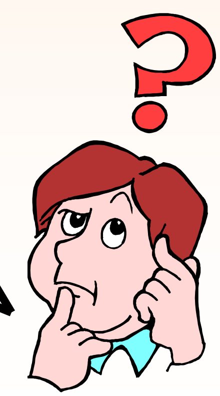
Outline

- Overview of Today's Optical Network
- Future Optical Network Requirements
- Introduction to Optical Burst-Switched Networks
- Research Contributions and Ongoing works
  - Reactive Contention resolution
  - Proactive Contention resolution
  - Data Burst Grooming
- Other Research Contributions
- Selected Publications

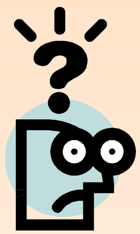
Advances in High Performance Optical Network

UTD

	OADM (point-to-point with no switch)	OADM (with switch)	OXC (Layers approach – no WLC)	OXC with WLC
Topology	P2P WDM	WDM Ring	Mesh/Ring	Mesh
Traffic Type	Static	-	Dynamic	Packet-based / Dynamic
Capacity	-	100 Gb/s – Tb/s	100 Tb/s	Pb/s
Network Types:	Opaque	Opaque	Opaque	Transparent
Switch Type:	Opaque	Opaque	Transparent	Transparent
Generation/ Availability	First	Second	Third	Fourth
	80's	90's	2000-2010	2010+



Internet Growth –
The Myth



> *"Internet traffic is doubling every three months."*

- Good news for carriers such as Level-3!
- The *fluke* of 1995-1996 became the future trend
- We would have needed 600,000 *OC48* links from coast to coast!!

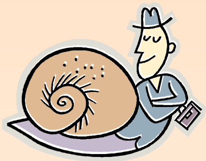
Internet Growth –
The Fact

“Internet traffic is doubling every year.”

- The real growth has been 87 percent per year!

Oooopps!!

Future Network - Basic Trends

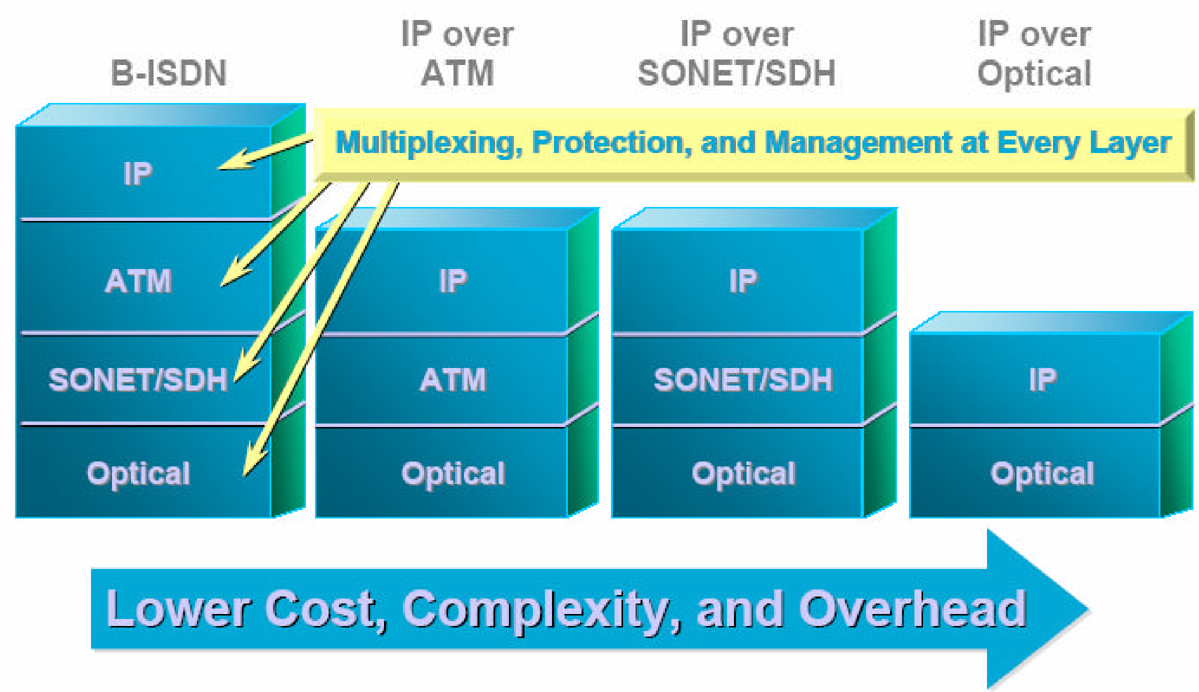


"This is not your father's network."

- **Satisfying on-demand high-BW requests is critical**
  - Fiber-to-curb
    - Digital theater and other "*killer apps*"
  - Many-to-many
    - High quality collaborative work / Game boxes (Xbox)
    - Medical image processing and remote visual steering
- **Offering flexibility**
  - Carefully planning network capacity is impossible
  - No more long network contracts and leases!

Future Network -
Next Generation Characteristics

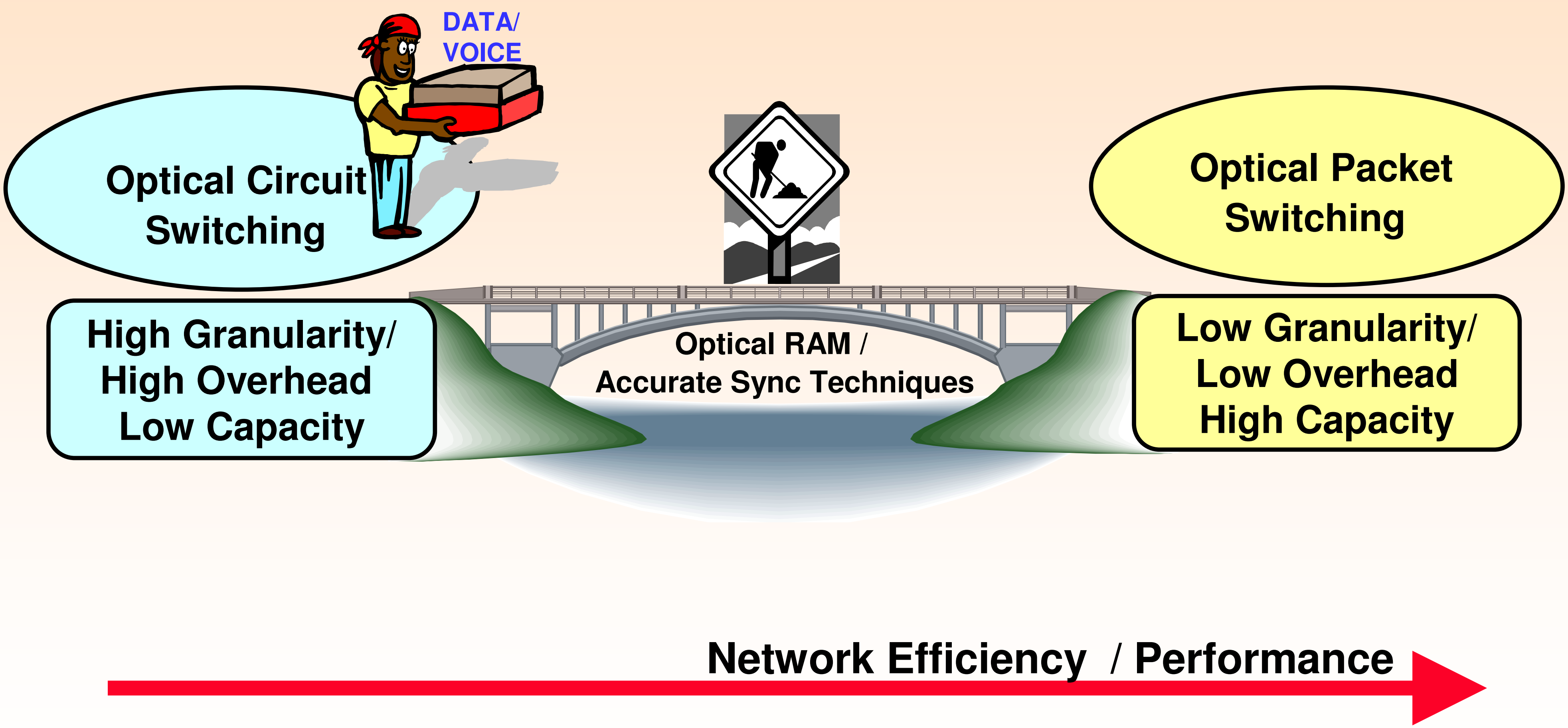
- Simple network hierarchy
  - Moving away from the old IP/ATM/SONET/WDM systems
  - Moving towards IP/WDM



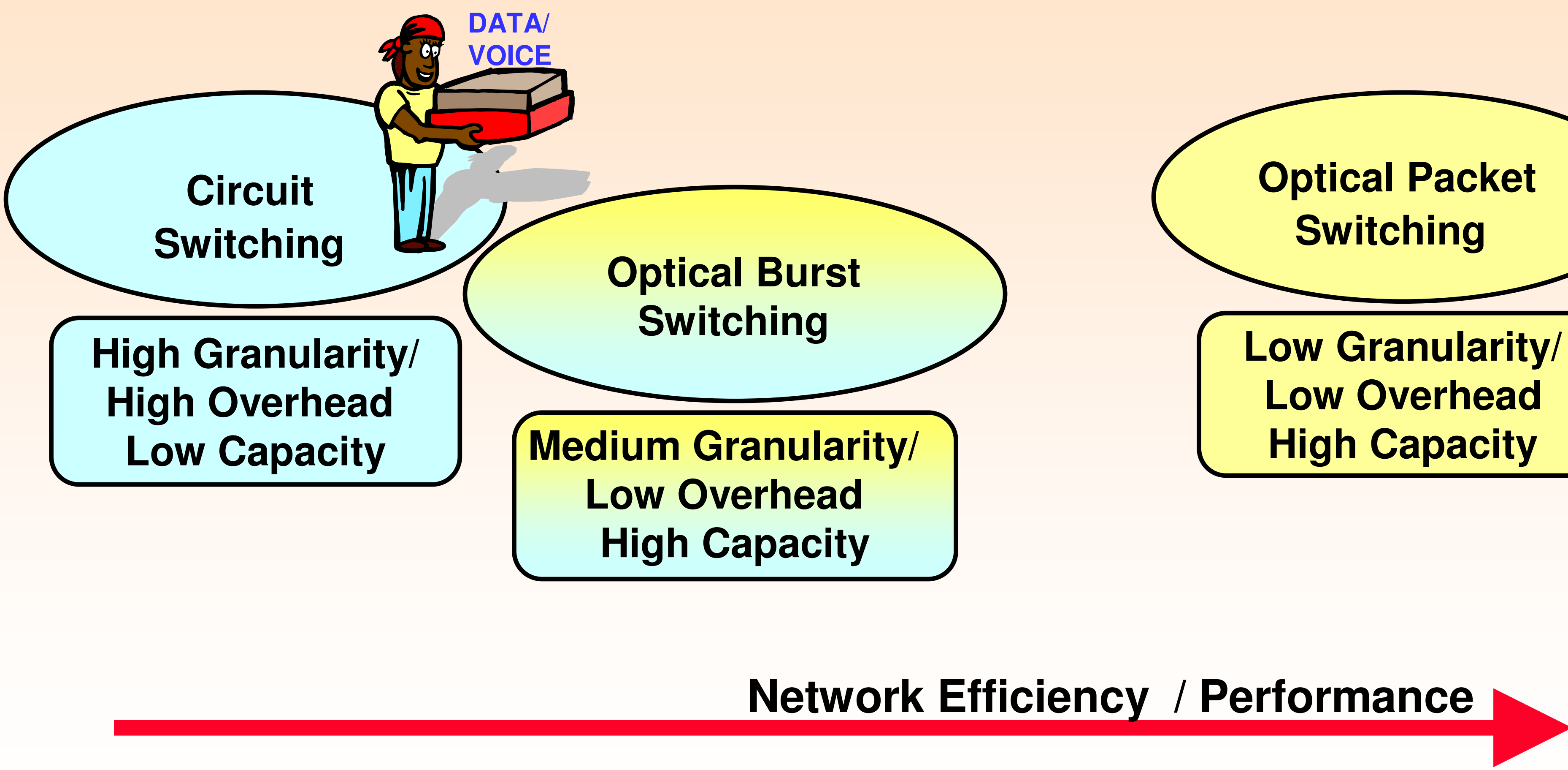
Future Optical Network -
Enabling Technologies

- WDM technology
  - More wavelengths in a single fiber
  - Easy channel access (on-chip DWDM)
- Switching technology
  - Optical circuit switching
  - Optical packet switching
  - Optical burst switching

Packet vs. Circuit Switching



Why Optical Burst Switching?



Burst Switching – A Short History!

UTD

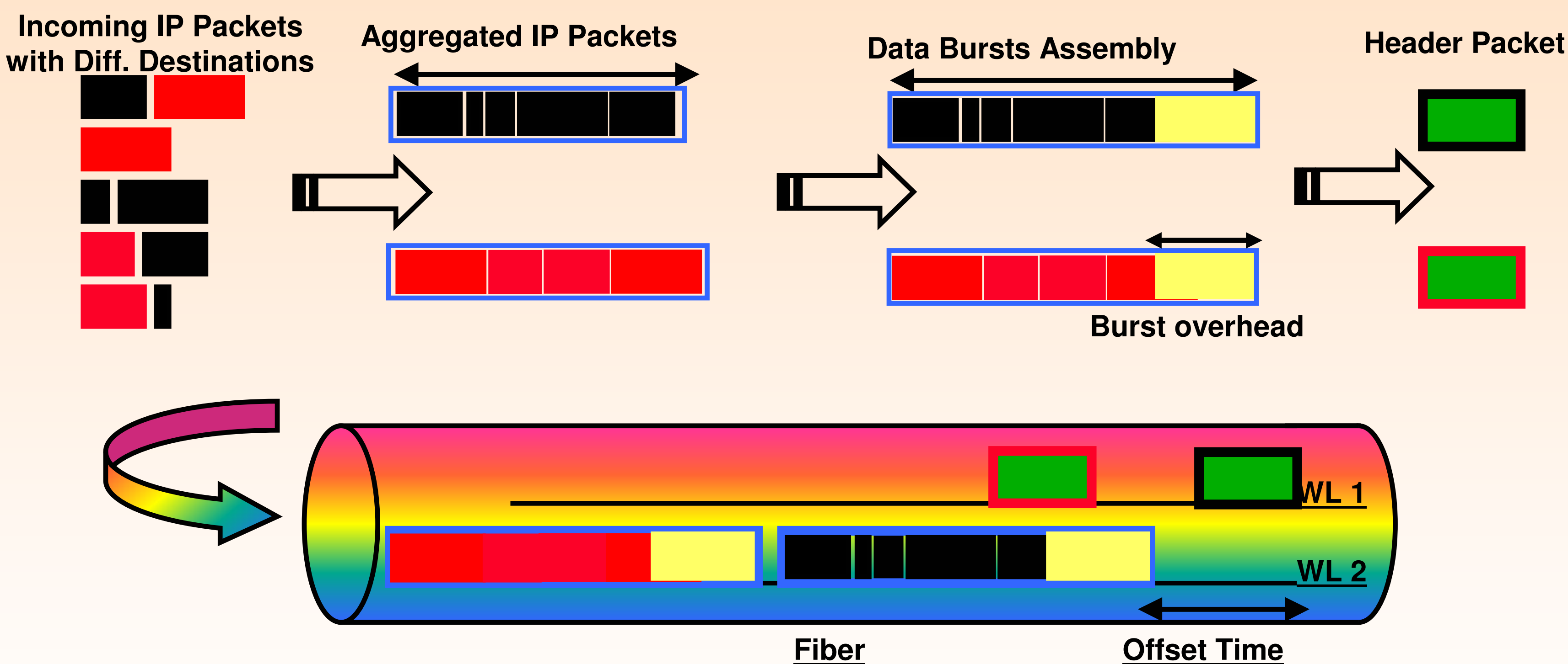
- First formal introduction in early 80's
  - *Burst Switching – an introduction*, by Amstuts
  - An extension to fast packet switching

- The basic concept
  - Handling packets of arbitrary length while employing decentralized shared buffer
- Proposed advantages
  - Providing an integrated switching system
  - Reducing loop length and increasing data transmission rate
  - Providing new services (such as ISDN) at the subscriber instrument

OBS utilizes the burst concept in optical domain!

More Details

Optical Burst Switching – Basic Idea



- Headers are processed electronically
- Data bursts are processed optically

Optical Burst Switching – Potential Advantages

- No fine optical buffering (optical RAM)
  - Fiber delay lines can provide sufficient optical buffering
- Relaxed synchronization requirements
- Electronic processing of header packets
  - Possibly at a different rate!

OBS vs. Others Switching Technologies

Optical Transport Networks	Bandwidth Utilization	Traffic Adaptability	Latency (set-up)	Over head	Optical Buffer Requirements	Data Loss
OCS	Low	Low	High	Low	None	Low
OPS	High	High	Low	High	High	Low
OBS	High	High	Low	Low	Low	High

Traffic Adaptability: such as burtiness

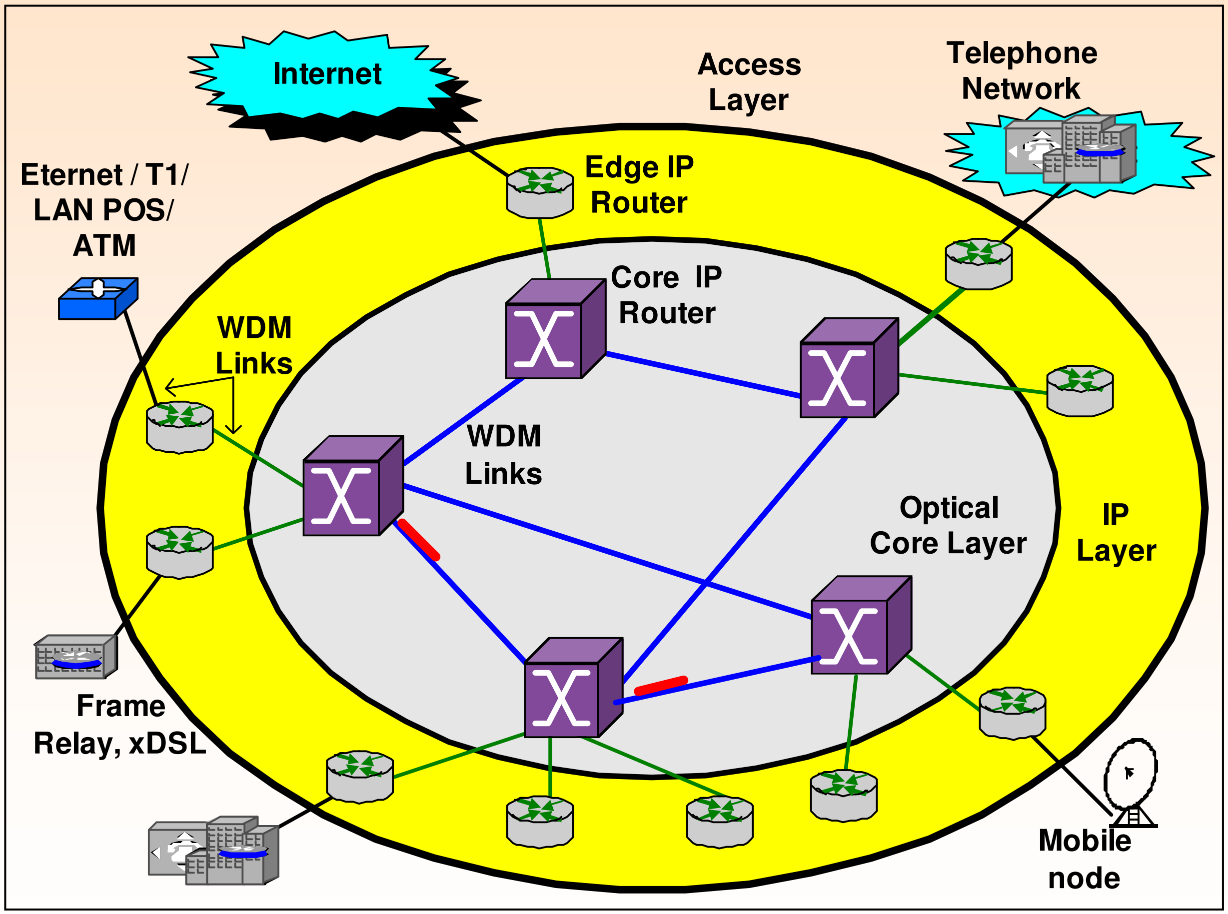
Overhead per unit of data

Source: K. Dolzer

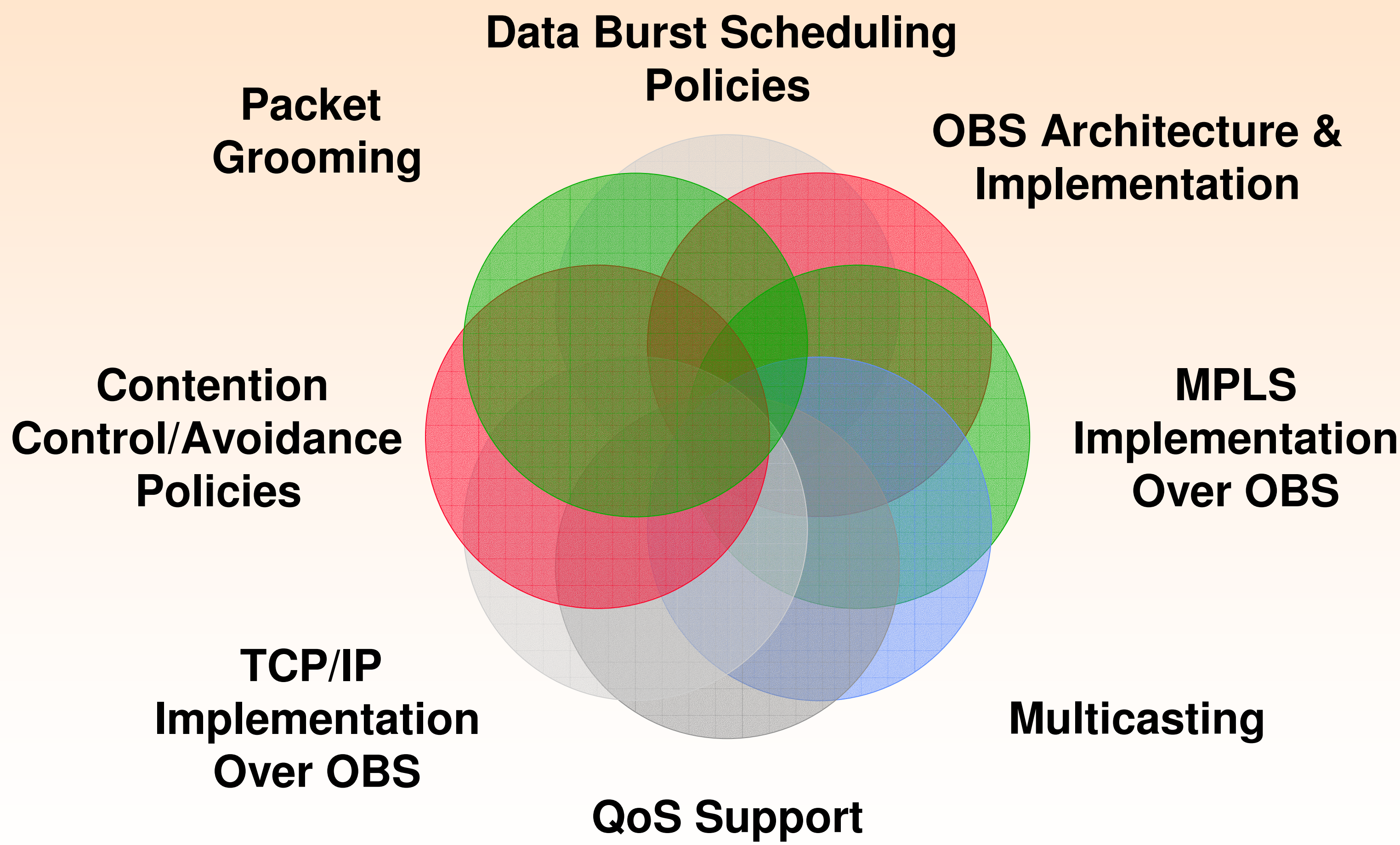
OBS Feasibility and Applications

- Data types handled by OBS
  - Afford random loss
  - Tolerate some delay
- Applications of OBS
  - Grid computing (Grid-over-OBS)
    - The existing peer-to-peer Grid systems is inefficient
    - We assume resource reservation-based Grid is wasteful and inefficient
    - Supporting *Anycasting*
  - Bulk data transfer
  - Distributed data
  - OBS over ring topology

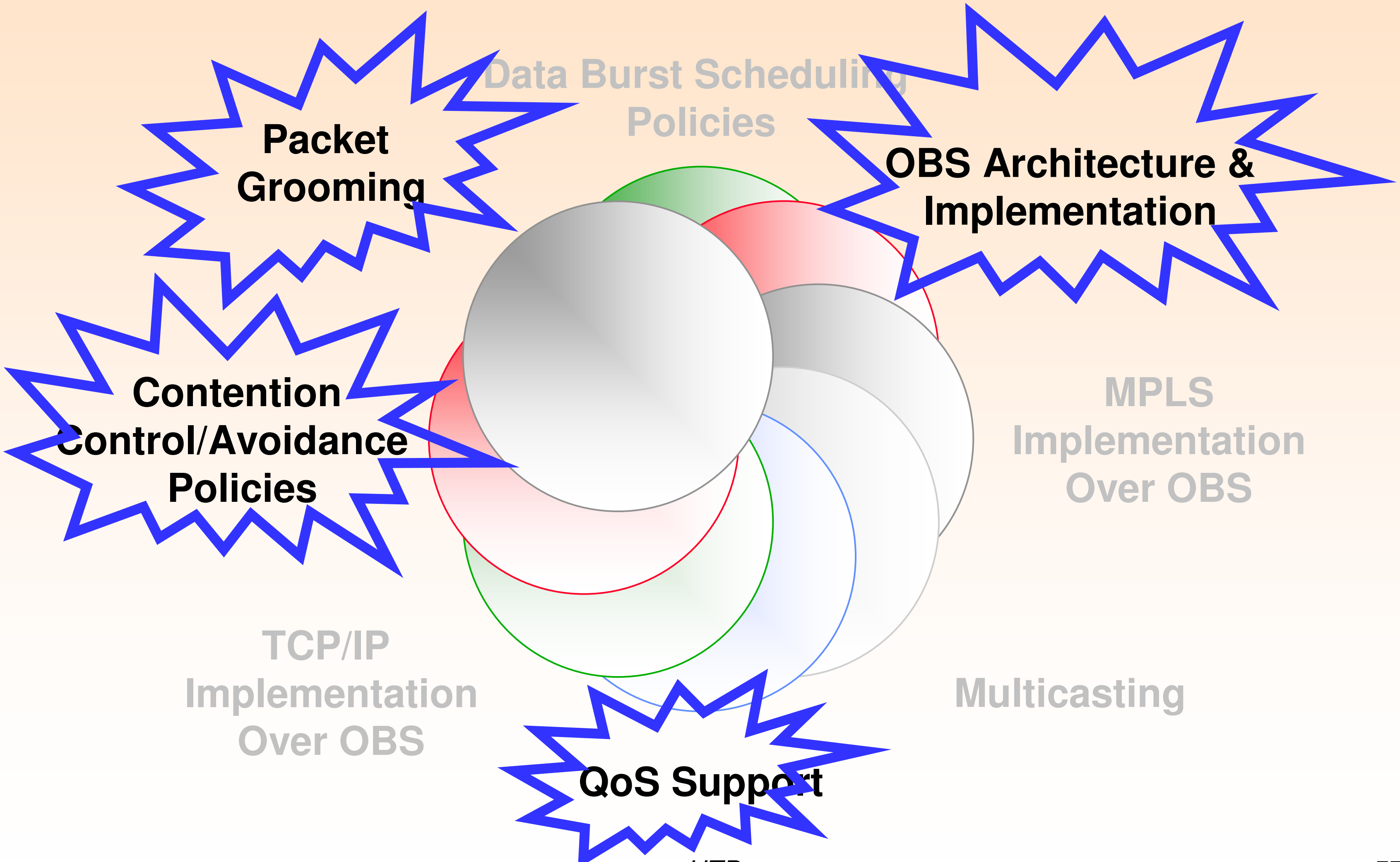
Optical Burst Switching –
Network Components



Optical Burst Switching – Challenges and Issues

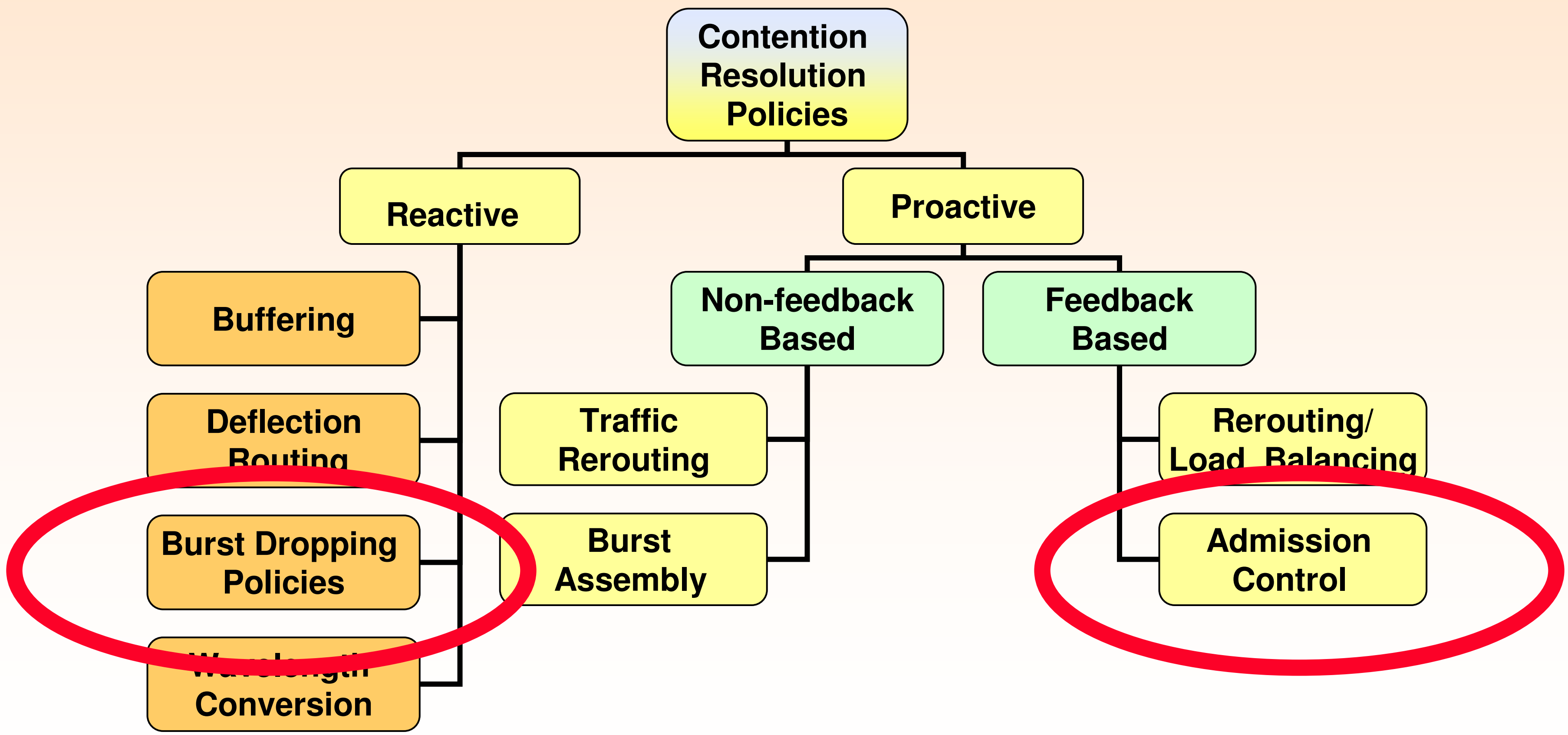


Selected Research Contributions



Contention Policies in OBS

- Contention resolution techniques
  - Resolution of contention between data bursts



Data Burst Dropping Policy

- Existing dropping policies
  - Latest Drop Policy (drop the latest arrival)
  - Segmentation (drop overlapping segments)
- Our proposed dropping policy
  - Called *Look-ahead contention resolution policy*
  - Implementing with any signaling protocol
  - Enabling service differentiation
  - Offering a number of variations

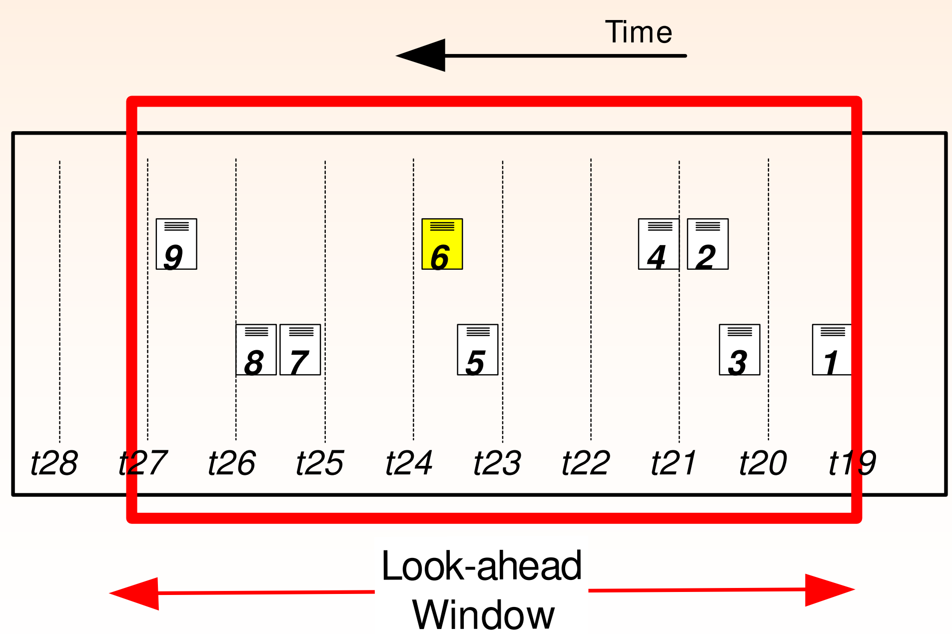
Look-ahead Contention Resolution

Basic Idea

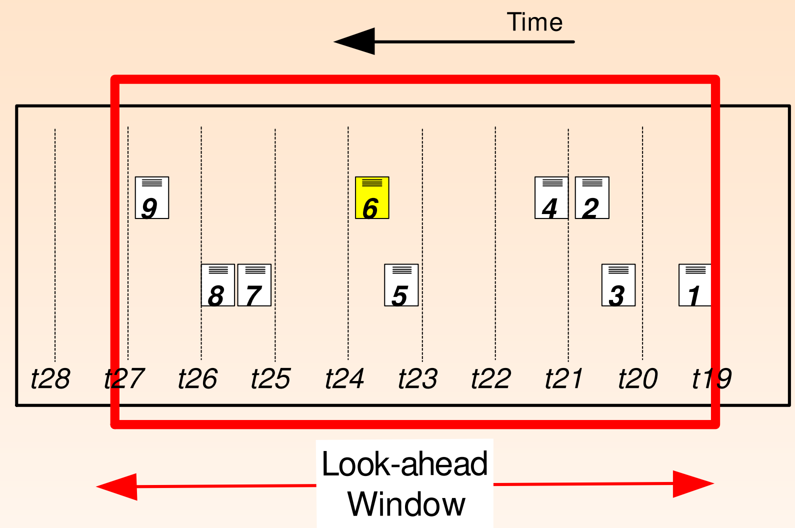
- Takes advantage of separation between data bursts and their headers
- Provides longer view of arriving data bursts
- Offers extended scheduling information

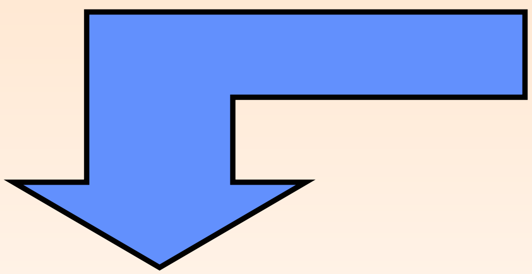
Simple Example
Using Look-ahead Contention Resolution

- Collect all incoming header packets within a look-ahead window W slots long
- W >= 2 x (longest burst size)

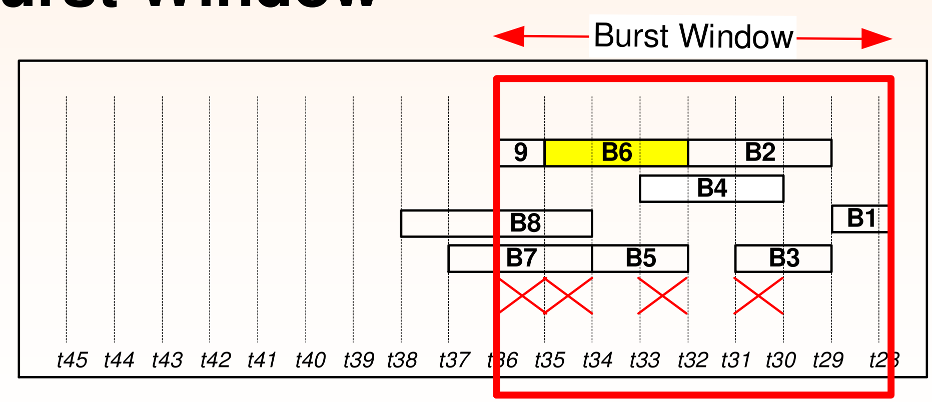


Simple Example
Using Look-ahead Contention Resolution





Burst Window



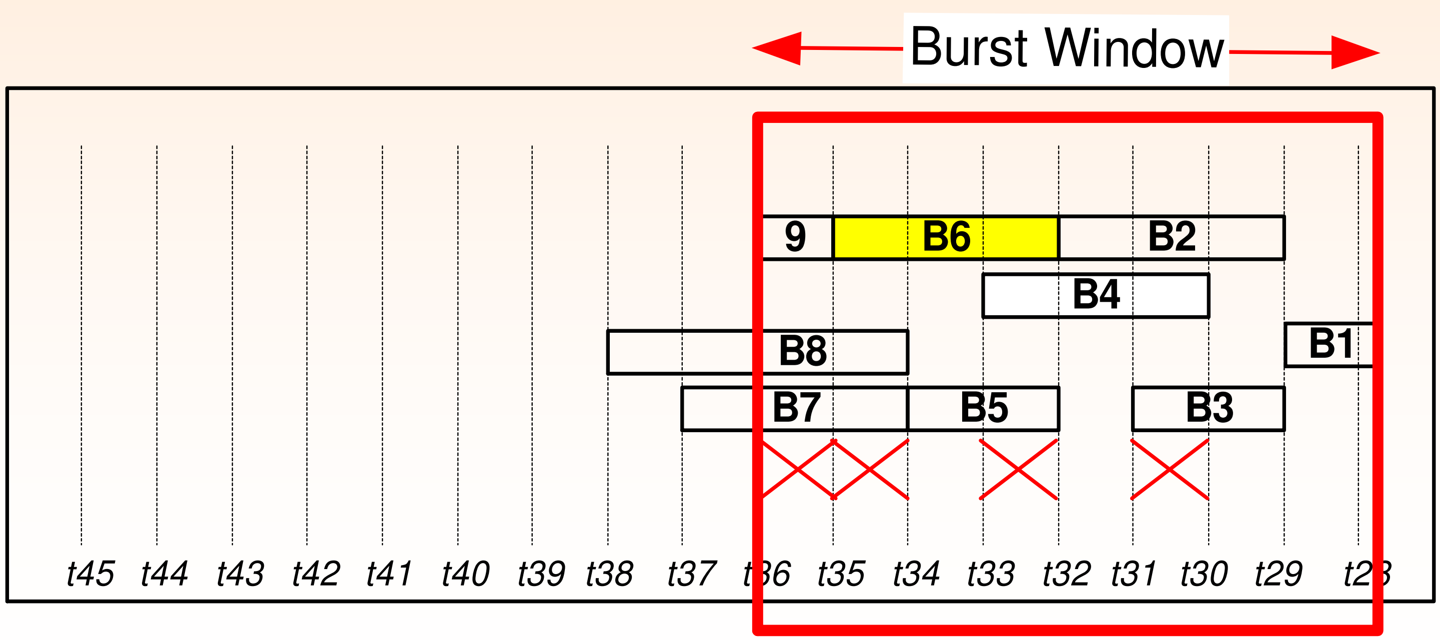
High Priority

Low Priority

Simple Example
Using Look-ahead Contention Resolution

UTD

- Determines contention slots
- Decides which data burst(s) to drop in order to minimize burst loss

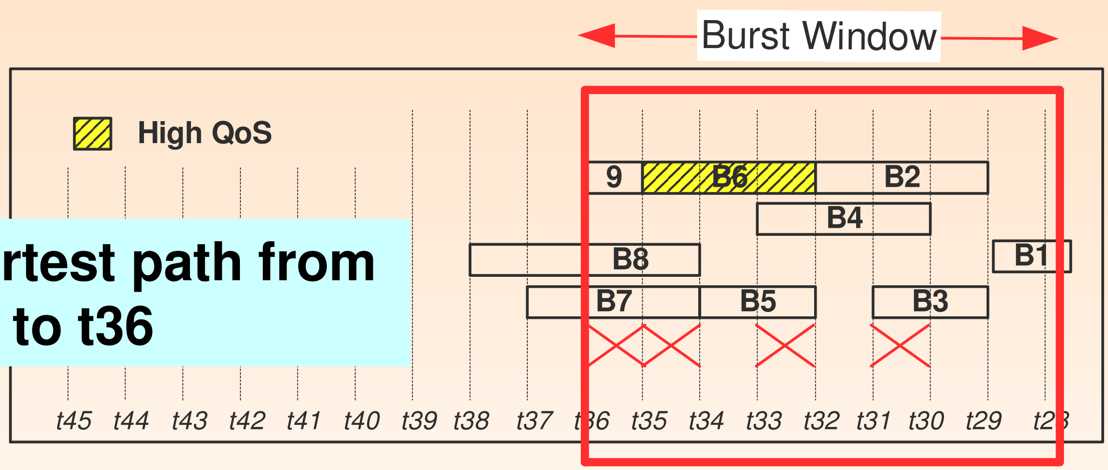


Look-ahead Contention Resolution

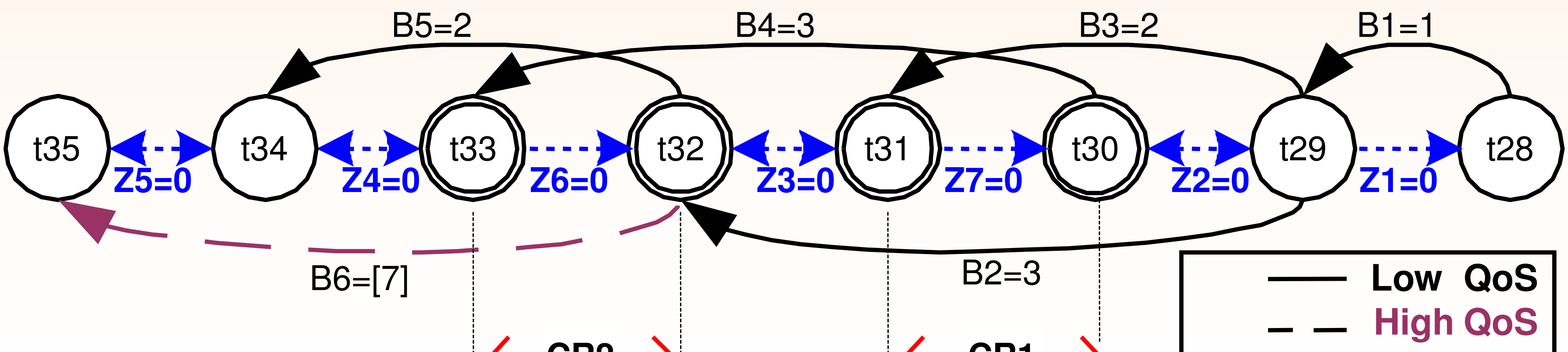
Dropping Section

- Create *auxiliary directed graph*
  - Nodes represent starting and ending times of bursts
  - Arcs represent data bursts
  - Weights indicate data bursts' durations
  - Extra weight dedicated to arcs representing high priority bursts
- Ensure the graph is *connected*
- Apply the *shortest path algorithm* from the start of the FIRST to the END of the last contention slot in the burst window
- Edges on the shortest path represent burst(s) to be dropped

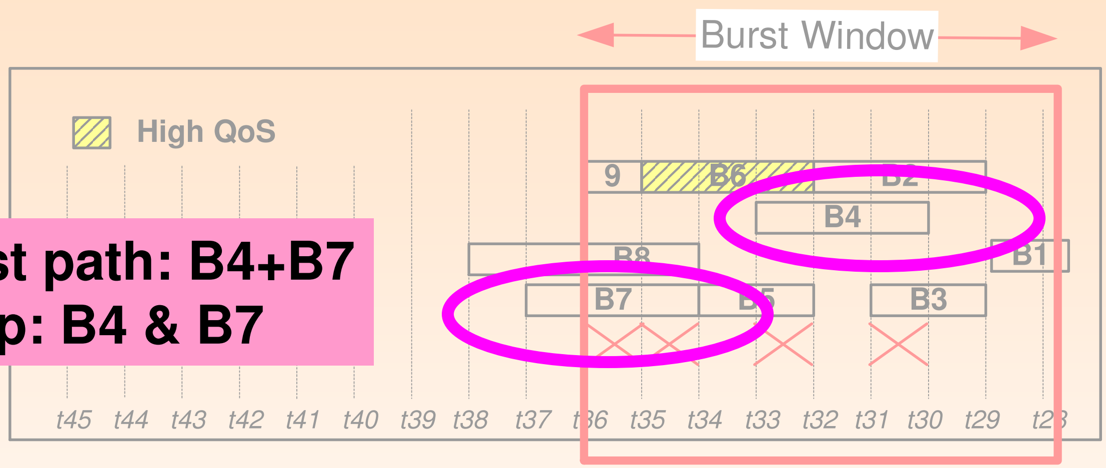
Look-ahead Contention Resolution



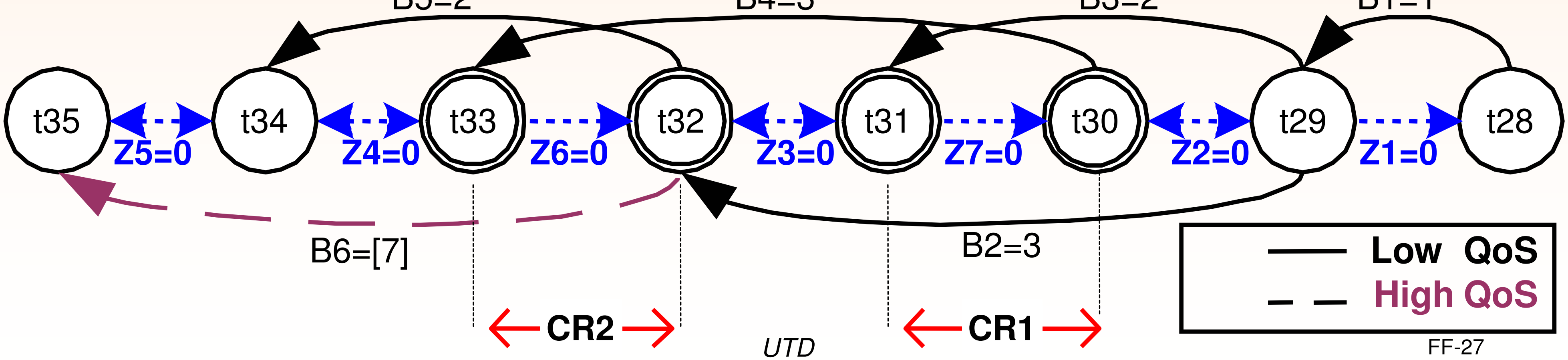
Find the shortest path from t30 to t36



Look-ahead Contention Resolution



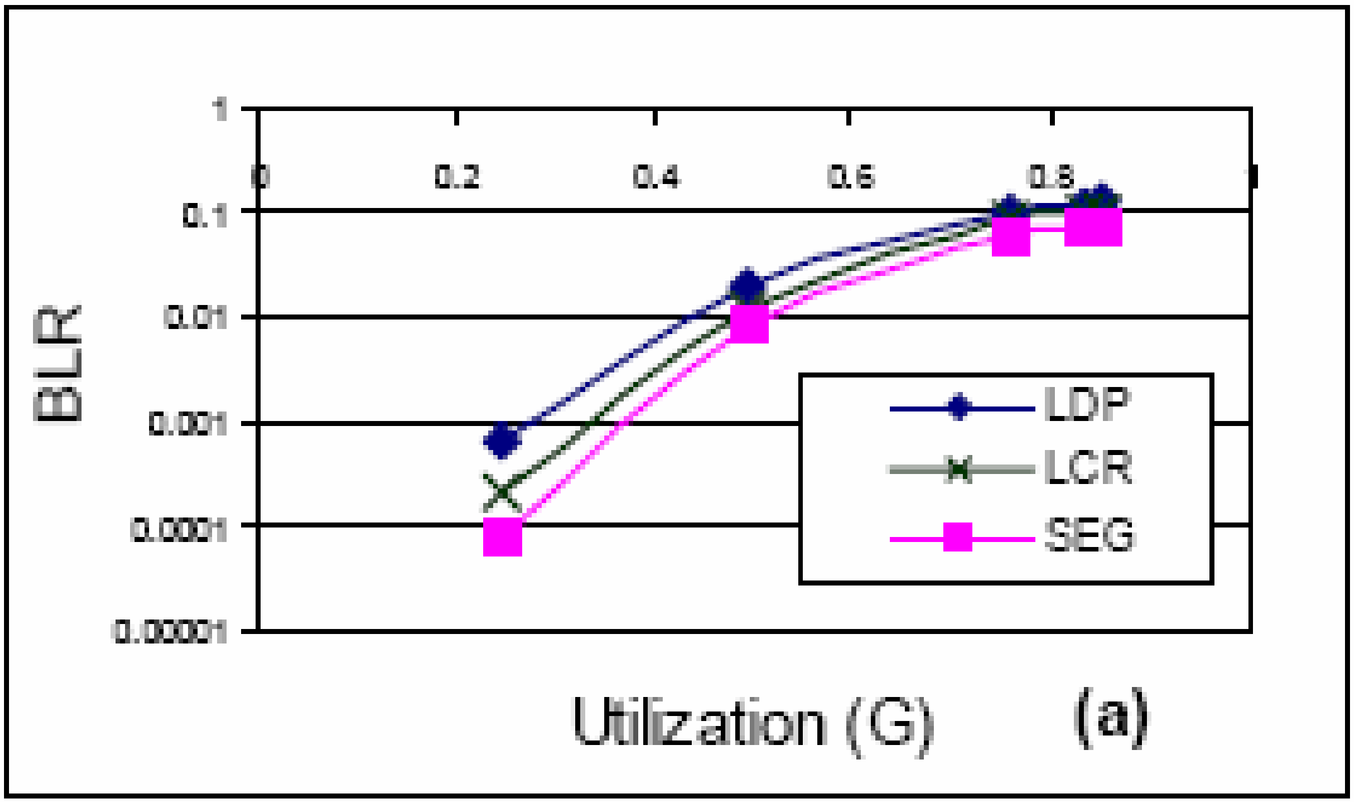
Shortest path: B4+B7
Drop: B4 & B7



Look-ahead Contention Resolution - Variation

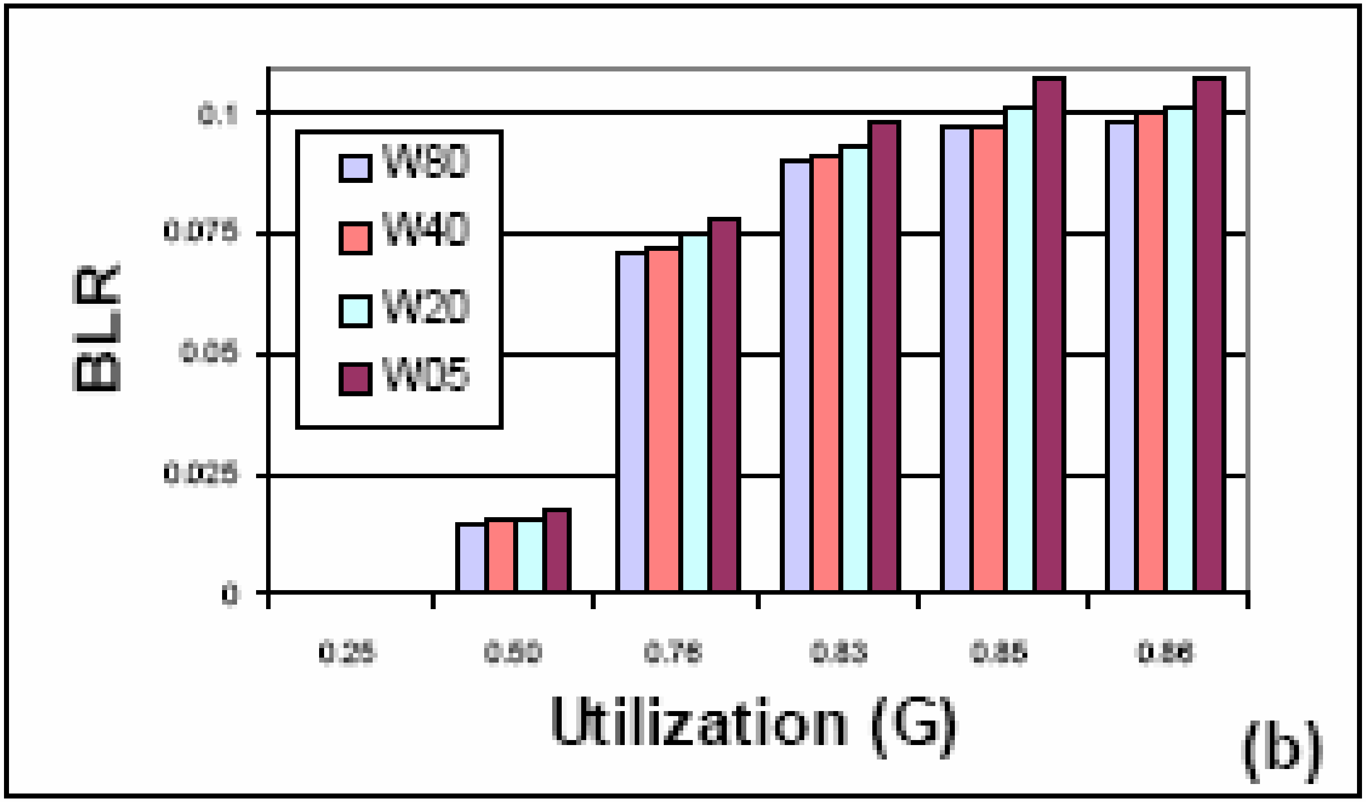
- Called *shortest drop policy*
- Special case of the Look-ahead Contention Resolution
  - W=1 (single slot)
- Header packets processed as soon as they are received
  - Lowering processing delay
- Bursts with highest priority and longest length preempt the rest

Performance Evaluation
Blocking Probability



Window Size = 10 time slots (Max Length)

Performance Evaluation
Window Size impact



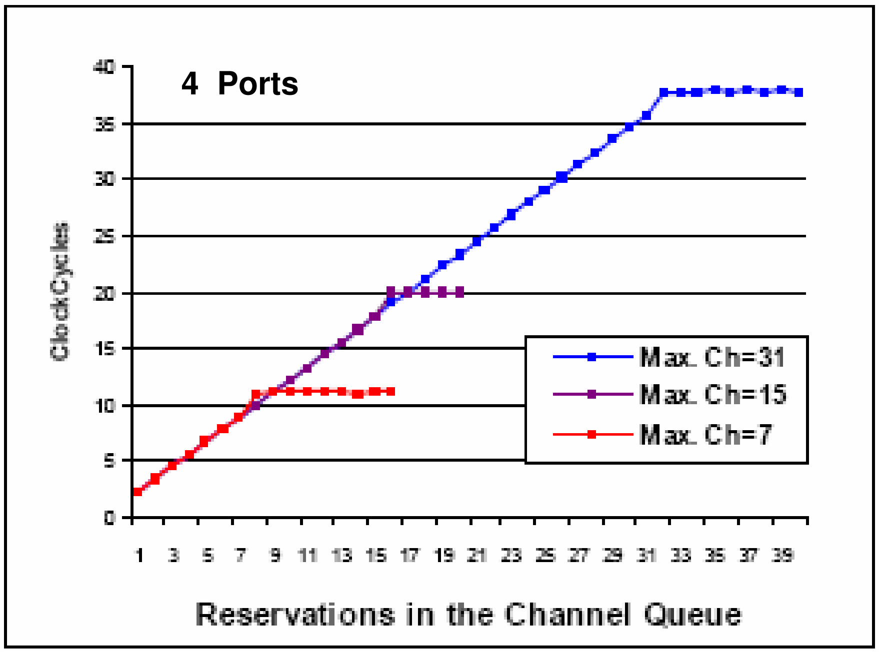
W20 = (2*Max Length)

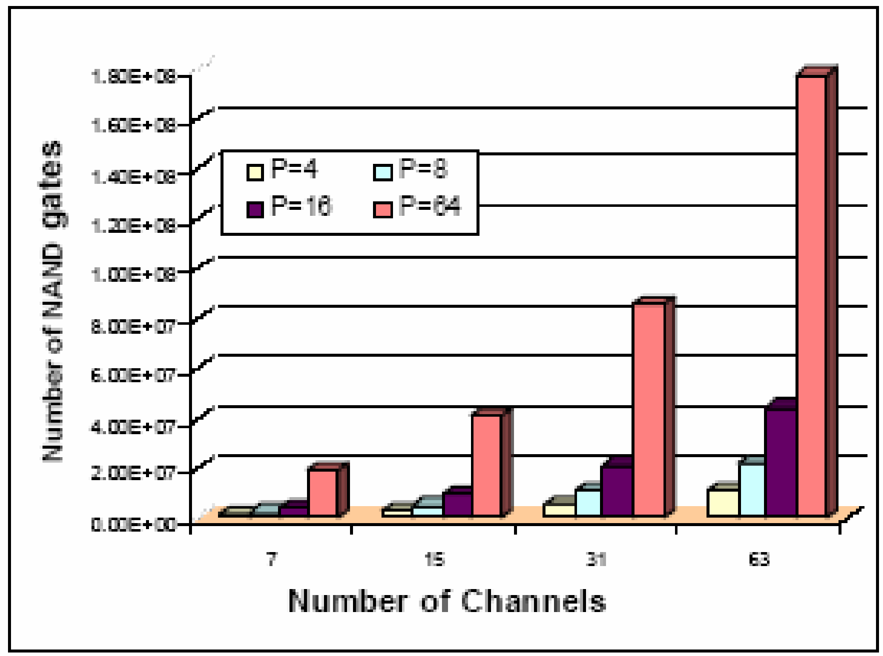
Implementation – Shortest Drop Policy

- Implemented on Altera EP20k400E FPGA
  - 2.5 million gates
  - Maximum clock rate of 840 MHZ
- Designed using VHDL code
- Tested, verified, and synthesized
  - Cadance (NcSim)
  - Quartus II

Objective:
Examine hardware scalability of the algorithm in terms of size and speed

Hardware Performance – Shortest Drop Policy





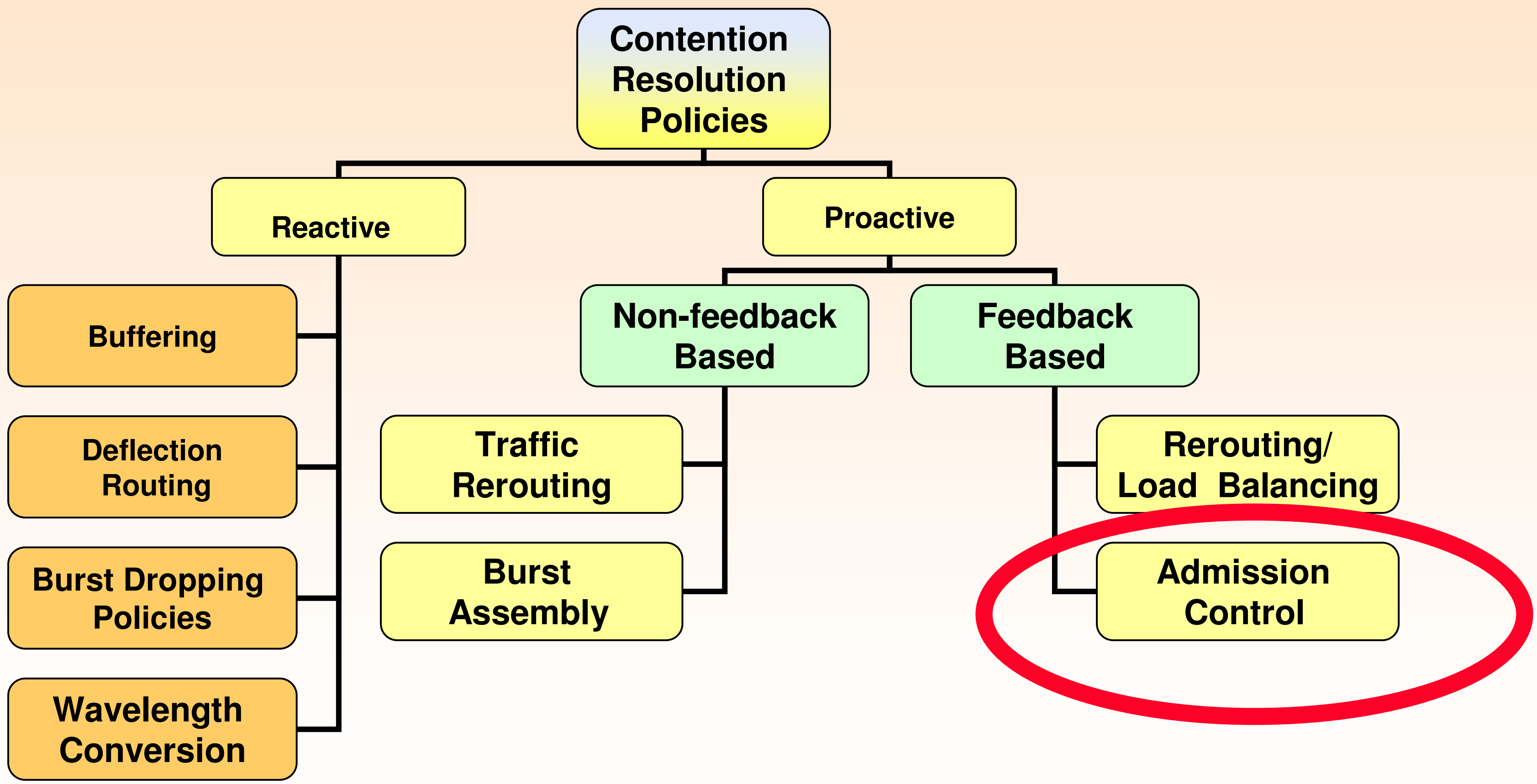
Look-ahead Contention Resolution - Work in Progress

- Developing a more efficient scheduling when ($C>2$)
- Proving that our scheduling approach optimal (locally)
- Developing a loss model for the SDP ($W=1$)

Publications:

1. ***Proceedings, IEEE Globecom 2003,***
2. ***Proceedings, IEEE HPSR 2003,***
3. ***First International Workshop on Optical Burst Switching, 2003***

Contention Policies in OBS

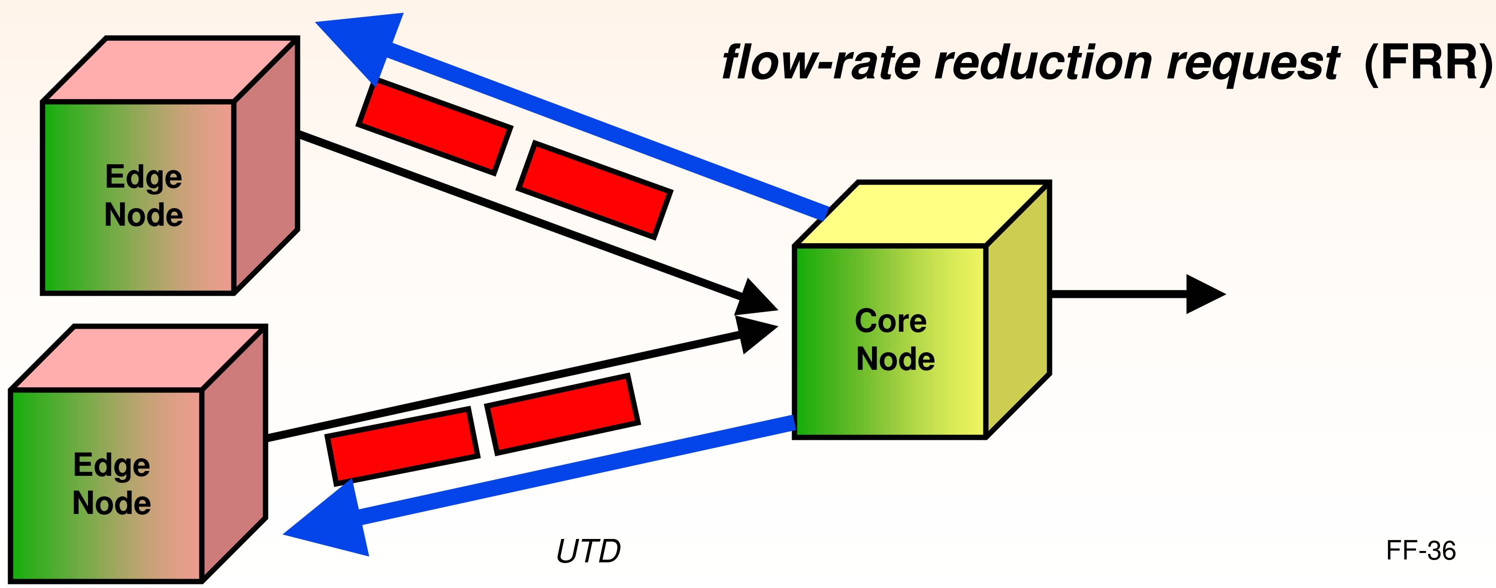


SFC Feedback-based Contention Avoidance – Major differences

- Existing feedback based mechanisms
  - Focus on rerouting or multiple retransmission
- Our proposed feedback based mechanism
  - Flow-rate control
  - Loss-based system rather than a queue occupancy system (no buffers)
  - Feedback signal are sent to the source from the congested nodes (not end nodes)
  - Additive increase / multiplicative decrease

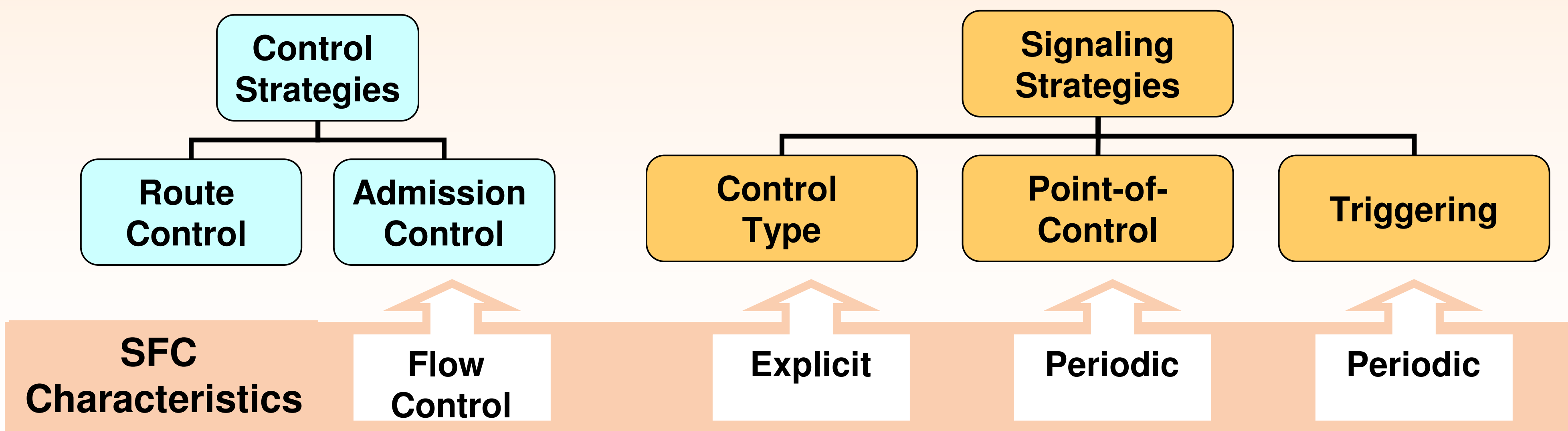
SFC Feedback-based Contention Avoidance – Basic Idea

- The core sends feedback signals to edge nodes
- Feedback signals explicitly request for flow-rate change on a link
- Edge nodes adjust their burst flow rate through admission control



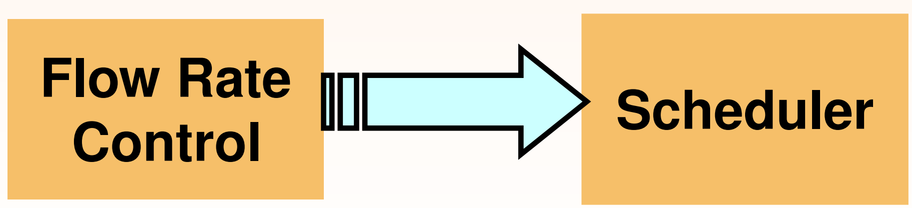
SFC Feedback-based Contention Avoidance – Key Elements

- Control strategies
  - What to do upon receiving the feedback messages
- Signaling strategies
  - How to communicate the current state to other nodes



SFC Feedback-based Contention Avoidance – Edge Node Functionalities

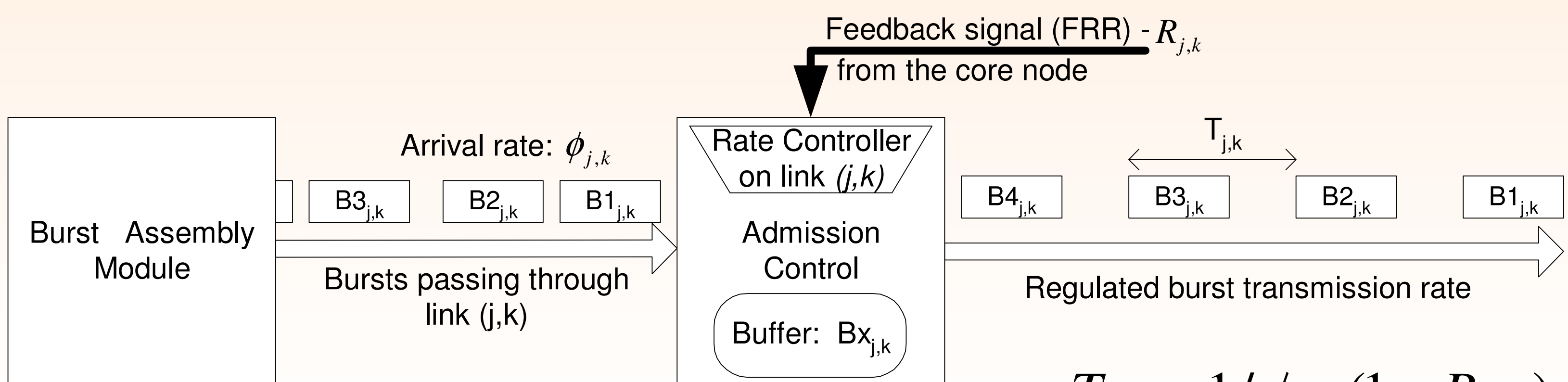
- The source edge node receives all feedback signals ($R_{j,k}$)
- It performs two basic operation
  - Determining the data burst flow rate on the congested link (j,k)
  - Scheduling data bursts



Edge Node's response to the feedback

SFC Feedback-based Contention Avoidance – Flow Control Rate

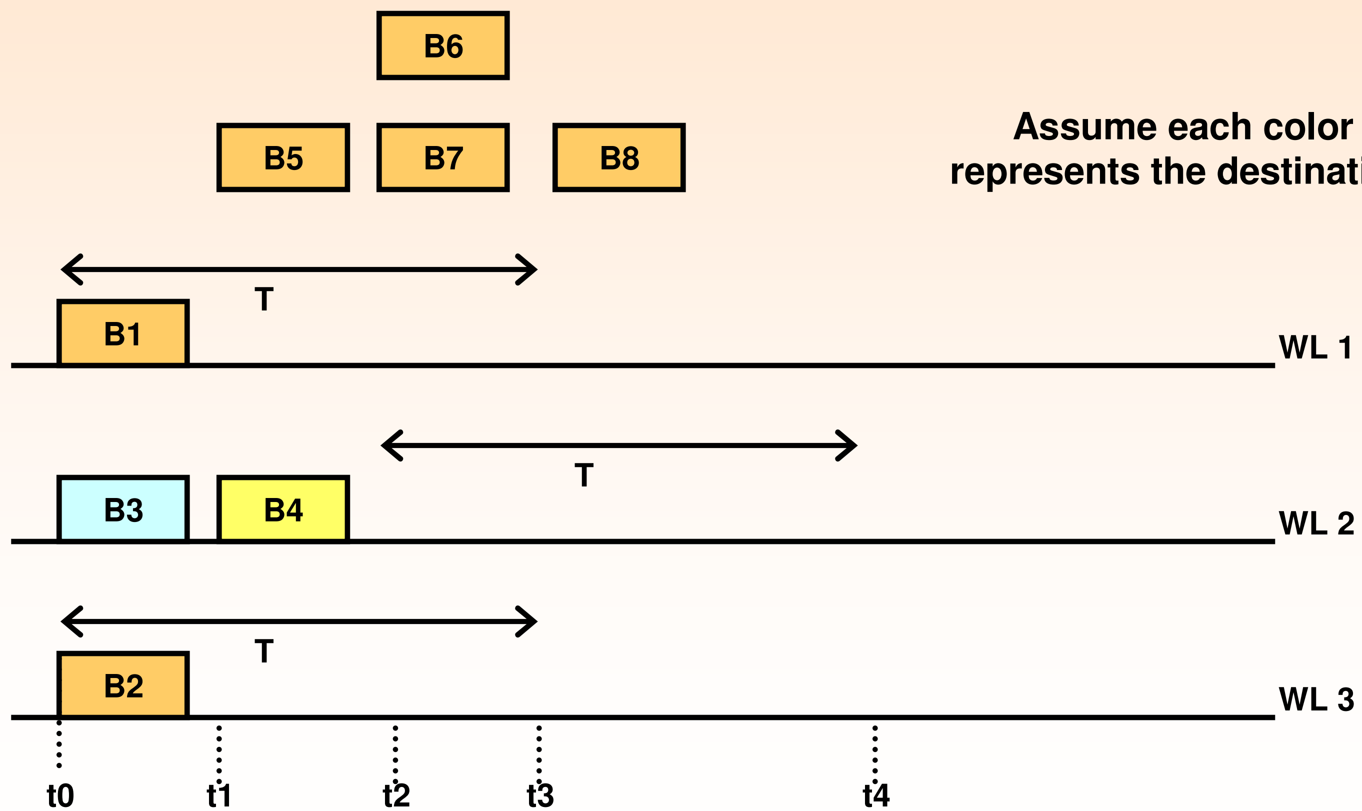
- The core measures the average load no each port
  - Increase the flow rate additively
  - Decrease the flow rate multiplicatively



$$T_{j,k} = 1/\phi_{j,k}(1 - R_{j,k})$$

SFC Feedback-based Contention Avoidance - Performance

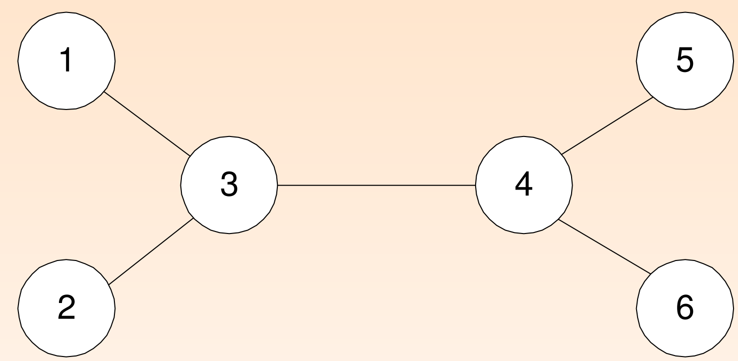
- Data bursts subject to admission control are scheduled on the latest available wavelengths

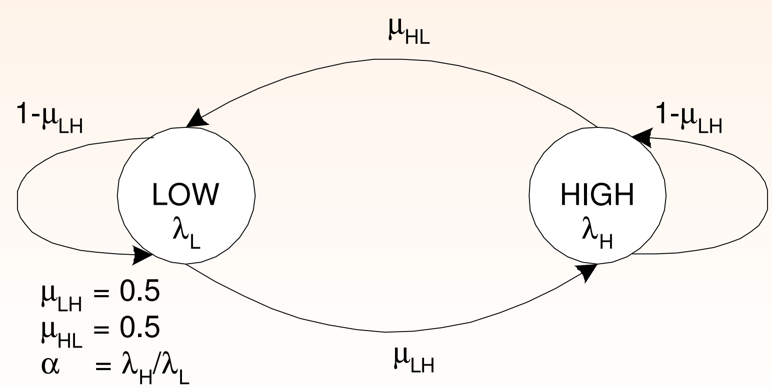


Assume each color represents the destination!

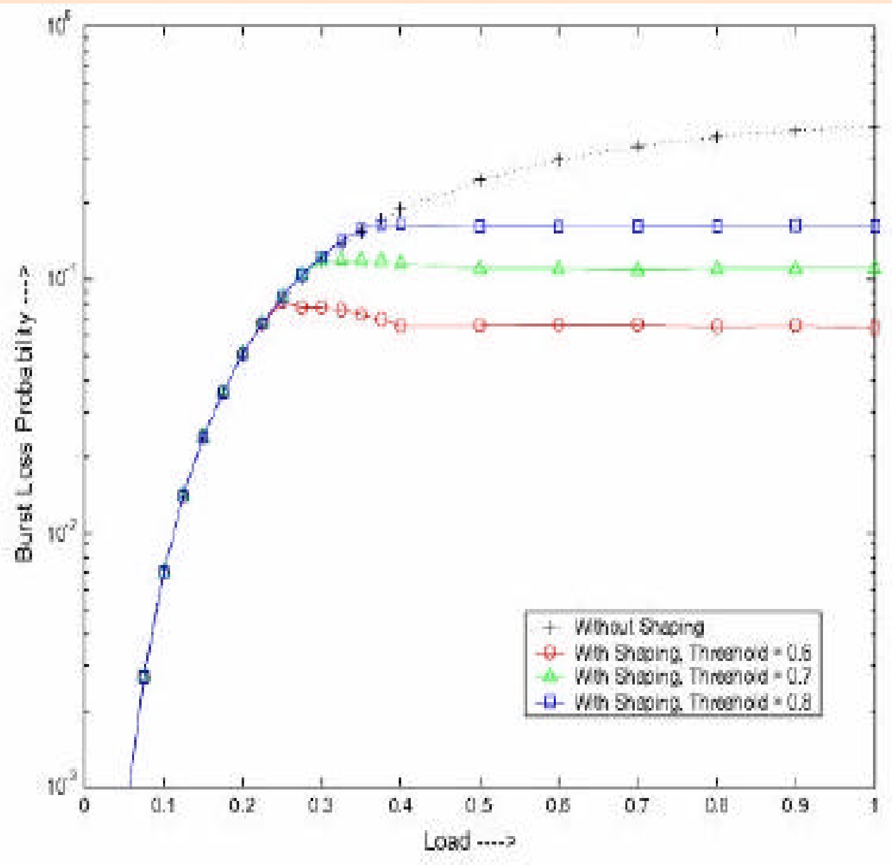
SFC Feedback-based Contention Avoidance - Performance

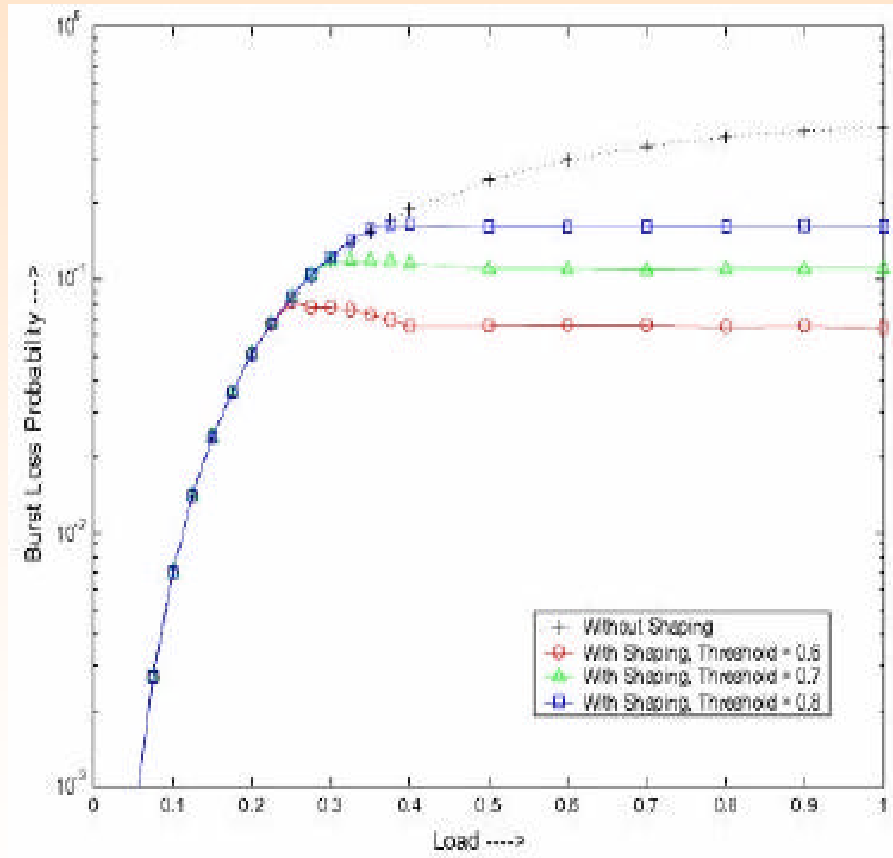
- Traffic models
  - Exponential arrival with fixed size data bursts
  - Markov Modulated Poisson Process with two states
- Significantly reduces the packet blocking probability
- Introduces additional end-to-end packet latency (SDj,k)





SFC Feedback-based Contention Avoidance - Performance





SFC Feedback-based Contention Avoidance - Work in Progress

- Implement the SFC on a larger size networks
- Compare its performance with different averaging mechanisms
- Develop a convergence model

Publications:

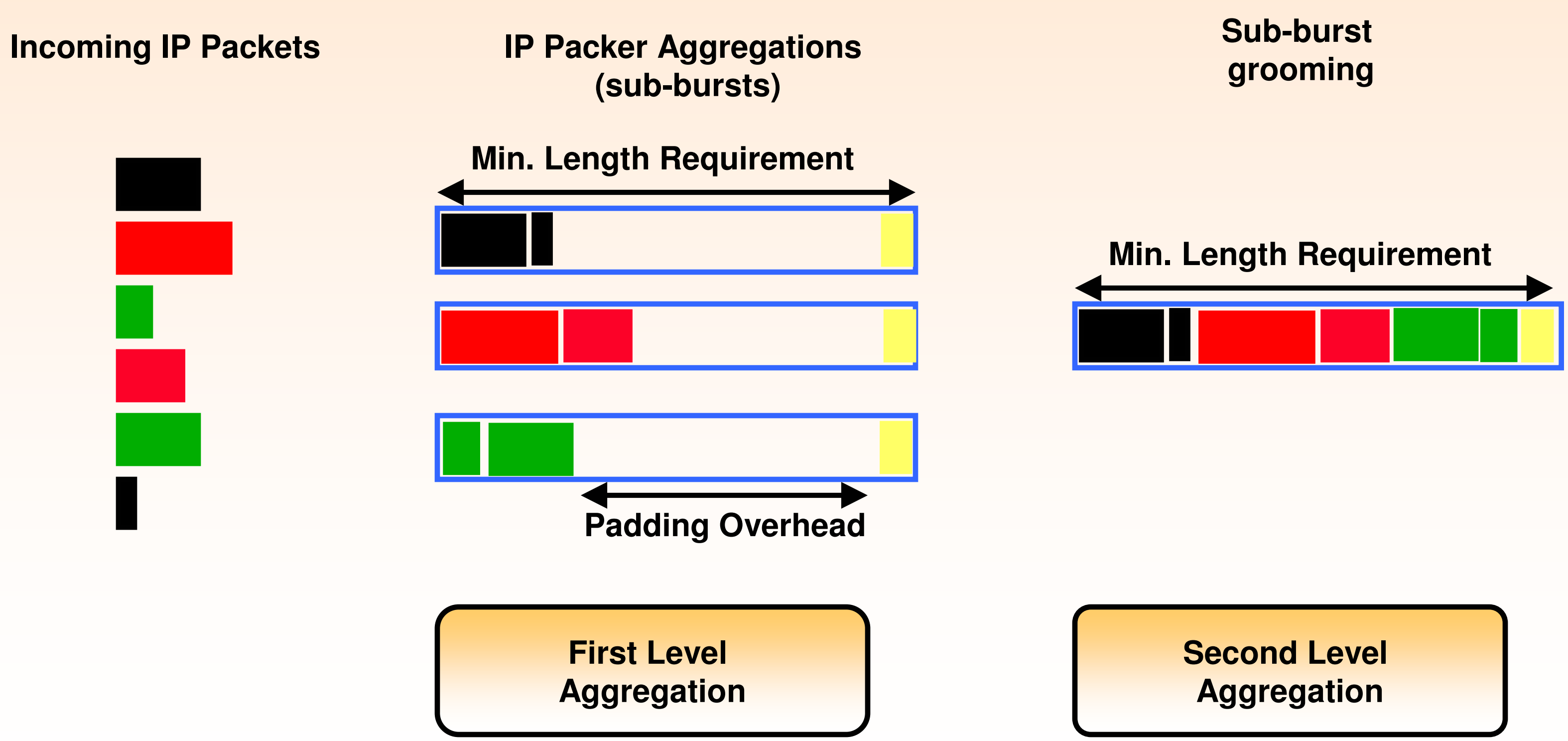
1. *Third International Workshop on Optical Burst Switching, 2003*
2. *Proceedings, IEEE Globecom 2005* (submitted)

Data Burst Grooming

- Data burst grooming
- Motivation
- Algorithms
- Performance

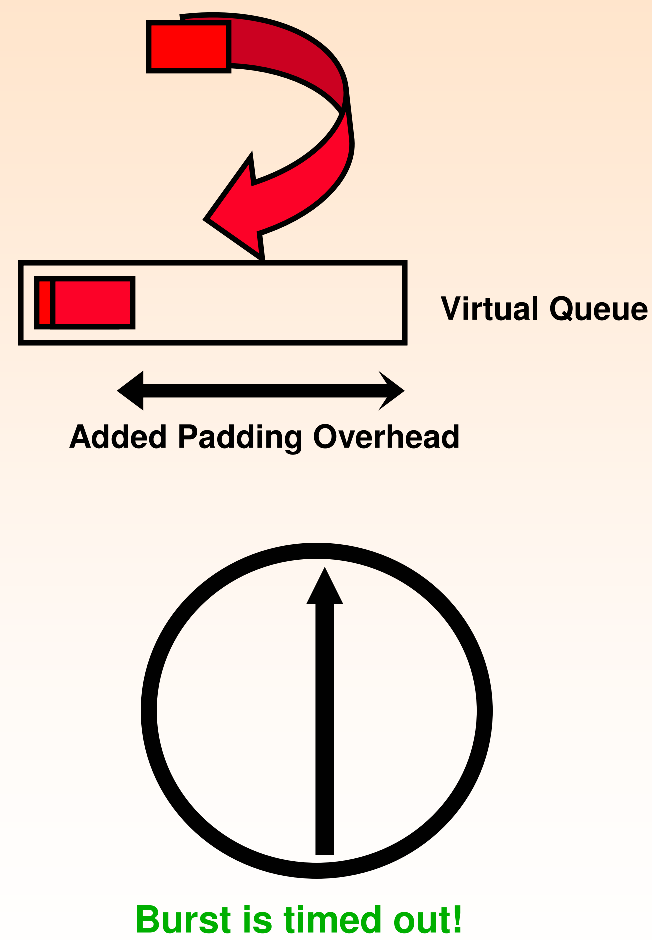
Data Burst Grooming – Basic Idea

Advantage: Reducing overhead, higher throughput, lower packet blocking probability



Data Burst Grooming – Assumptions and Motivation

- Maximum acceptable latency is given
- IP arrival rate is low
- Minimum switching time is fixed
- Bursts must be released after time T
- Each data burst will be much smaller than its *minimum* required length
- Padding overhead must be added
  - Low throughput
  - High packet blocking probability



Data Burst Grooming – Problem and Solutions

PROBLEM

- How to assemble sub-bursts together
  - How many sub-bursts can be groomed in a single burst
  - The total length of the groomed sub-burst
- How to route sub-bursts
  - Shortest path-based routing
  - Allowing alternate routing

SOLUTION

- We Proposed two grooming algorithms
  - No-routing-overhead (NoRO)
  - Minimum-total-padding-overhead (MinTO)

Grooming Algorithms

- No-routing-overhead (NoRO)
  - No routing overhead is allowed
  - Find the largest available sub-bursts
  - Variation: Limit the data-burst length

$$Roh(b_i) = \frac{H_p(S_{b_0}, D_{b_0}) + H_p(D_{b_0}, D_{b_i})}{H_p(S_{b_0}, D_{b_i})},$$

$$TRoh(\mathrm{G}) = \sum_{b_i \in \mathrm{G}, b_i \neq b_0} Roh(b_i).$$

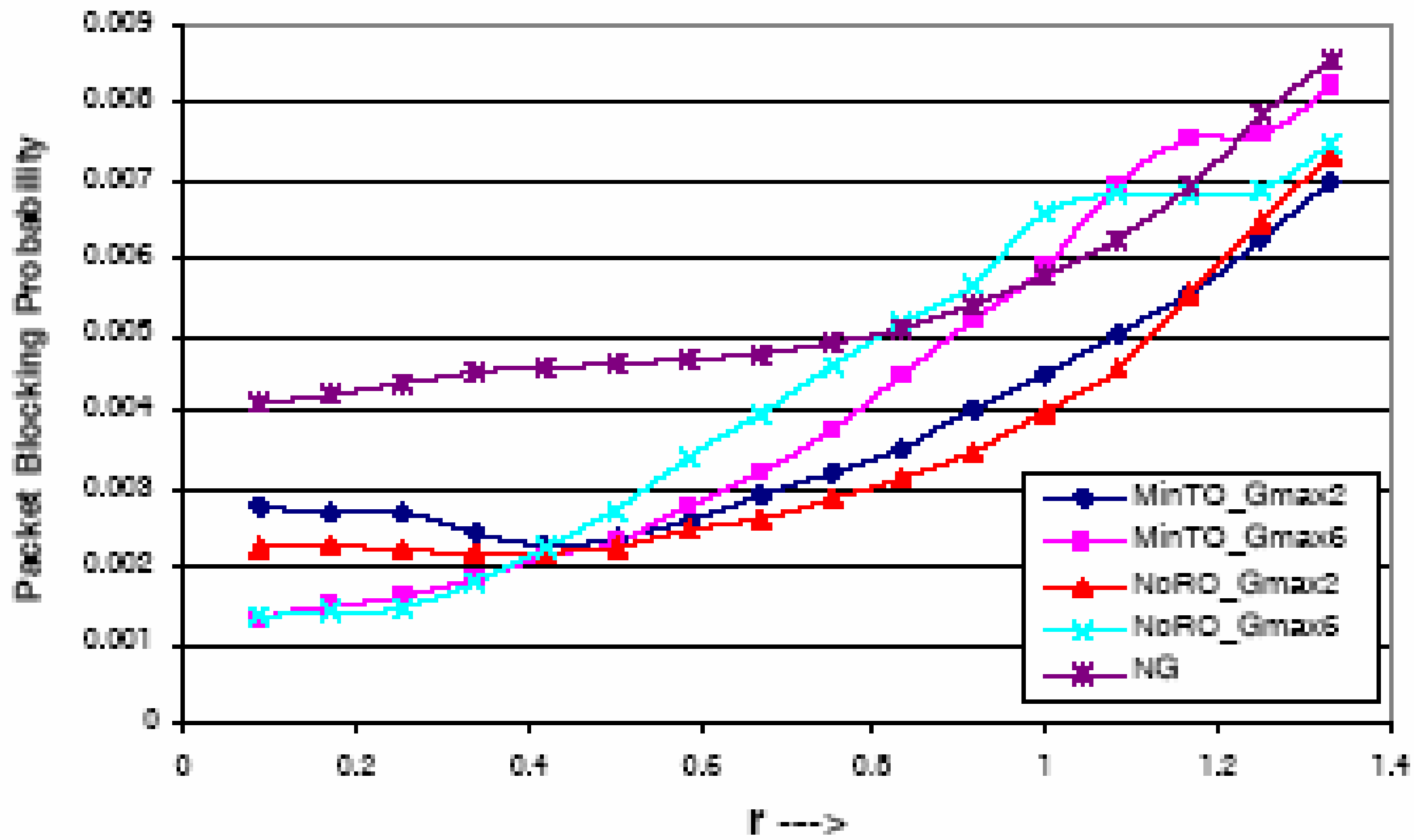
- Minimum-total-padding-overhead (MinTO)
  - Minimize the overall network padding
  - Routing overhead is allowed to compensate for access padding overhead
  - Variation: Limit the routing overhead

$$\begin{aligned} RPoh(b_i) = \{ & max(L^{MIN}, L_G + L_{b_i}) \cdot H_p(S_{b_0}, D_{b_0}) + \\ & \sum_{b_j \in \mathrm{G}}^{b_j \neq b_0} max(L^{MIN}, L_{b_j}) \cdot H_p(D_{b_0}, D_{b_j}) + \\ & max(L^{MIN}, L_{b_i}) \cdot H_p(D_{b_0}, D_{b_i})\} / \\ \{ & \sum_{b_j \in \mathrm{G}} max(L^{MIN}, L_{b_j}) \cdot H_p(S_{b_0}, D_{b_j}) + \\ & max(L^{MIN}, L_{b_i}) \cdot H_p(S_{b_0}, D_{b_i})\}. \end{aligned}$$

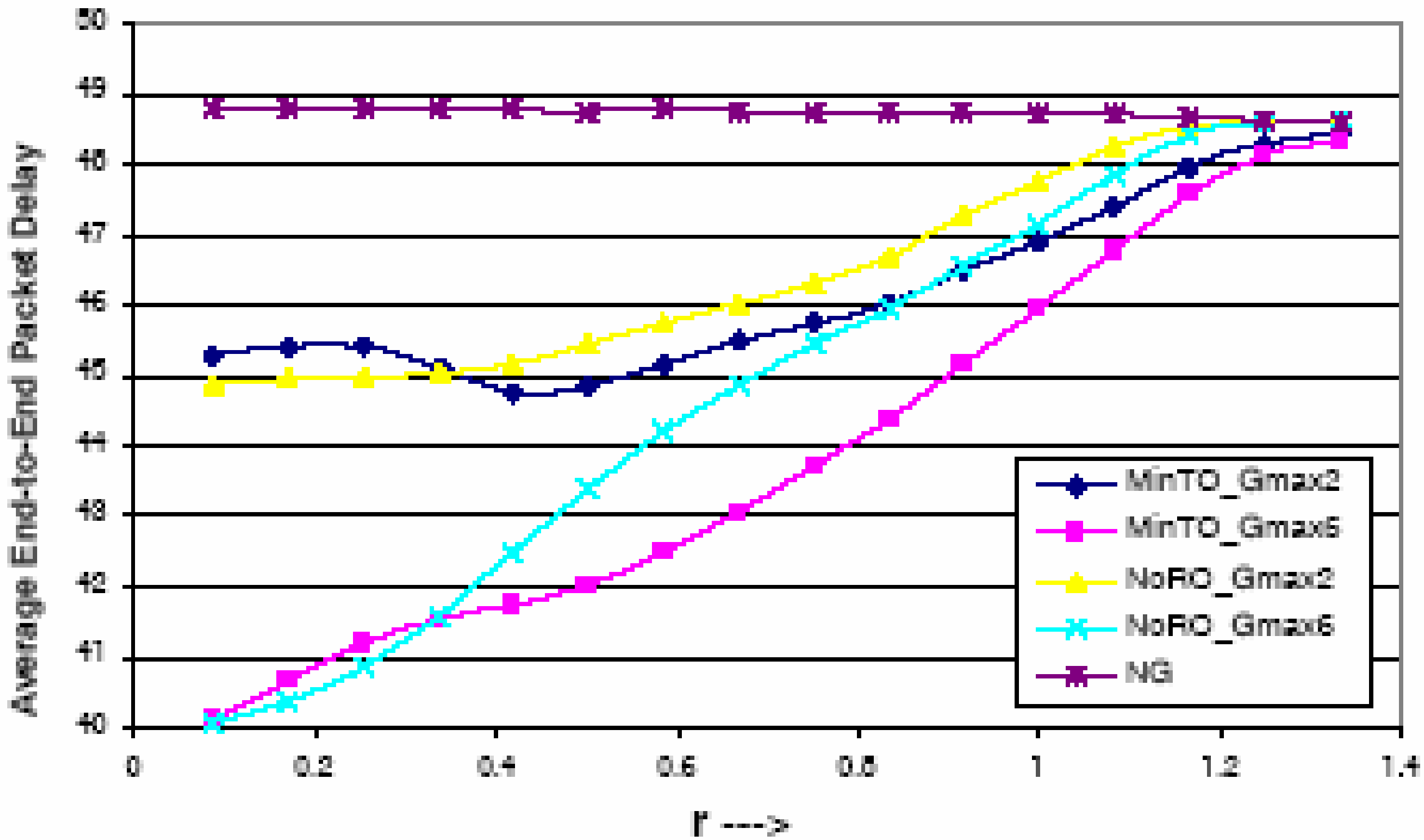
Algorithm Performance

- Very effective at low loads
  - Lower latency
  - Lower packet blocking probability
- Grooming becomes more effective for core nodes with slower switching time
- Grooming becomes more effective if the maximum packet delay is reduced
- Grooming can potentially result in bursty traffic
  - The number of sub-bursts which can be groomed must be limited as the load increases

Algorithm Performance
Blocking Probability



Algorithm Performance
End-to-end delay



Data Burst Grooming – Work in Progress

- THIS WORK IS COMPLETE!
- Other Extensions:
  - Develop a mathematical model to show blocking as a function of grooming size
  - Extend grooming to support QoS
    - Delay priority
    - Lower blocking probability

Publications:

1. ***Journal of Lightwave Technology, December 2004 (submitted)***
2. ***Proceedings, BroadNets 2005 (submitted)***

Other Research Contributions – Traffic grooming,

UTD

- Traffic grooming:
  - Efficient Online Traffic Grooming Algorithms in WDM Mesh Networks with Drop-and-Continue Node Architecture *[Published BroadNets 2004]*
  - An Algorithm for Traffic Grooming in WDM Mesh Networks with Dynamically Changing Light-Trees *[Published IEEE Globecom 2004]*
- Contention Resolution in OBS and GMPLS application:
  - Single-anchored soft bandwidth allocation system with deflection routing for optical burst switching *[Published HPSR 2002]*
- Dual-ring architecture:
  - Near-Optimal Design of WDM Dual-Ring with Dual-Crossconnect Architecture *[Published OptiComm 2002]*

Selected Publications

UTD

- **Submitted Journals**
  - Farid Farahmand, Qiong Zhang, and Jason P. Jue, "Dynamic Traffic Grooming in Optical Burst-Switched Networks," submitted, *IEEE Journal of Lightwave Technology*, December 2005
  - Xiaodong Huang, Farid Farahmand, and Jason P. Jue, "Multicast Traffic Grooming in Wavelength-Routed WDM Mesh Networks Using Dynamically Changing Light-Trees," submitted, *IEEE Journal of Lightwave Technology*, December 2005
- **Conferences**
  - Farid Farahmand, Xiaodong Huang, and Jason P. Jue, "Efficient Online Traffic Grooming Algorithms in WDM Mesh Networks with Drop-and-Continue Node Architecture," *Proceedings, BroadNets 2004*, San Jose, CA, October 2004.
  - Farid Farahmand, Qiong Zhang, and Jason P. Jue, "A Feedback-Based Contention Avoidance Mechanism for Optical Burst Switching Networks," *Proceedings, 3rd International Workshop on Optical Burst Switching*, San Jose, CA, October 2004.
  - Farid Farahmand and Jason Jue, "Supporting QoS with Look-ahead Window Contention Resolution in Optical Burst Switched Networks," *Proceedings, IEEE Globecom 2003*, San Francisco, CA, December 2003.
  - Farid Farahmand and Jason Jue, "Practical Priority Contention Resolution for Slotted Optical Burst Switching Net Works," *Proceedings, First International Workshop on Optical Burst Switching (WOBS 2003)*, co-located with SPIE OptiComm 2003, Dallas, TX, October 2003.
  - Farid Farahmand, Andrea F. Fumagalli, and Marco Tacca, "Near-Optimal Design of WDM Dual-Ring with Dual-Crossconnect Architecture," *Proceedings, SPIE Optical Networking and Communication Conference (OptiComm 2002)*, Boston, Massachusetts, July 2003.
  - Farid Farahmand and Jason Jue, "Look -ahead Window Contention Resolution in Optical Burst Switched Networks," *Proceedings, IEEE High Performance Switching and Routing (HPSR 2003)*, Torino, Italy, June 2003.
  - Xiaodong Huang, Farid Farahmand, and Jason P. Jue, "An Algorithm for Traffic Grooming in WDM Mesh Networks with Dynamically Changing Light-Trees," *Proceedings, IEEE Globecom 2004*, Dallas, TX, November 2004.
  - Timucin Ozugur and Farid Farahmand, "Single-anchored soft bandwidth allocation system with deflection routing for optical burst switching," *Proceedings, IEEE High Performance Switching and Routing (HPSR 2002)*, Kobe, Japan, May 2003.

Outside References for this document

- Semaphore
  http://www.semaphore.com/services/convergence.html
- Andrew Odlyzko has many articles on Internet
  http://www.dtc.umn.edu/~odlyzko/doc/research.html
- Global crossing:
  http://www.gblx.net/xml/network/net_management.xml
- Good search engine
  http://www.firstmonday.org/fm.search

UTD

Thank you!

For more information:

http://www.utdallas.edu/~ffarid