How Many Wavelengths Do We Really Need?

A Study of Packets Over Wavelengths

Joe Bannister
Joe Touch
USC/ISI

Alan Willner
USC/EE-Systems

Stephen Suryaputra
Nortel Networks

WDM in the Internet

- Improved capacity
  - Link bandwidth (e.g. striping, inverse mux)
  - Connectivity multiplier (diameter reduction)
- Wavelength management
  - Static topology
  - Dynamic topology
    - Globally optimized (e.g. linear programming)
    - Locally optimized (e.g. label switching)

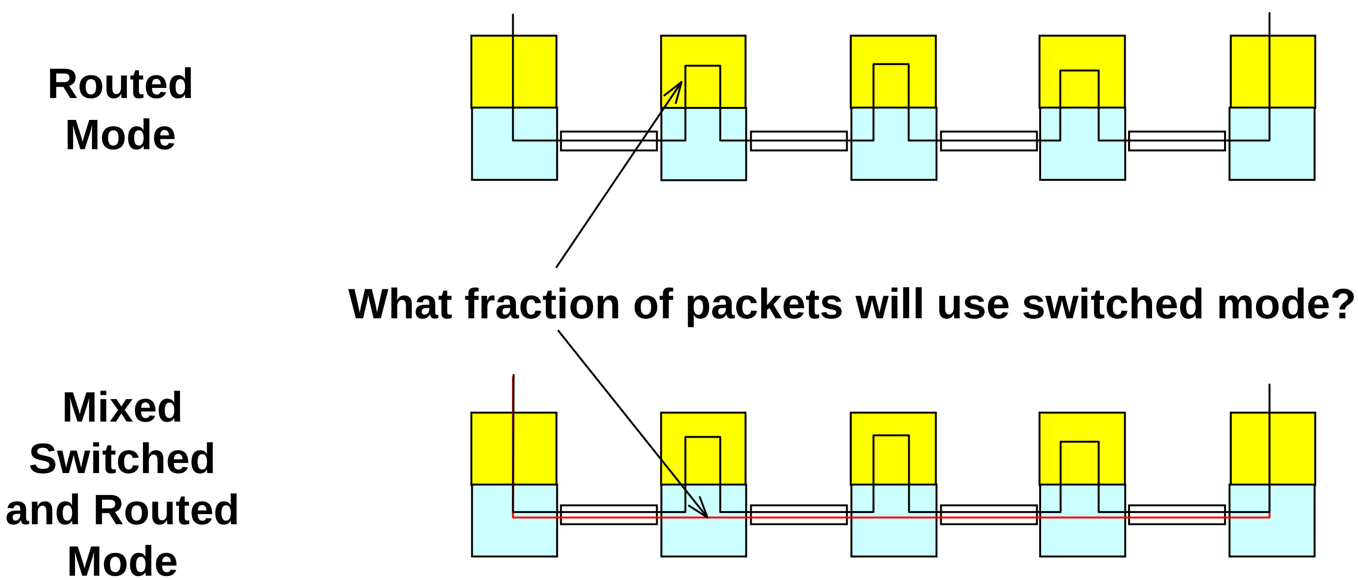
Packets over Wavelengths (POW): Assumptions

- Wide-area switching/routing
- WDM wavelength-selective crossconnect
- Default action is to forward packets
- Detected flows bound to wavelengths in real time
- Common approach

Problem Statement

- What is the expected performance of POW?
  - Wavelength-limited environment vs. channel-rich environment
  - Real traffic and topology
  - Realistic technology assumptions
- What can be done to improve the performance of POW?
  - Optical and protocol technologies

Performance Metric



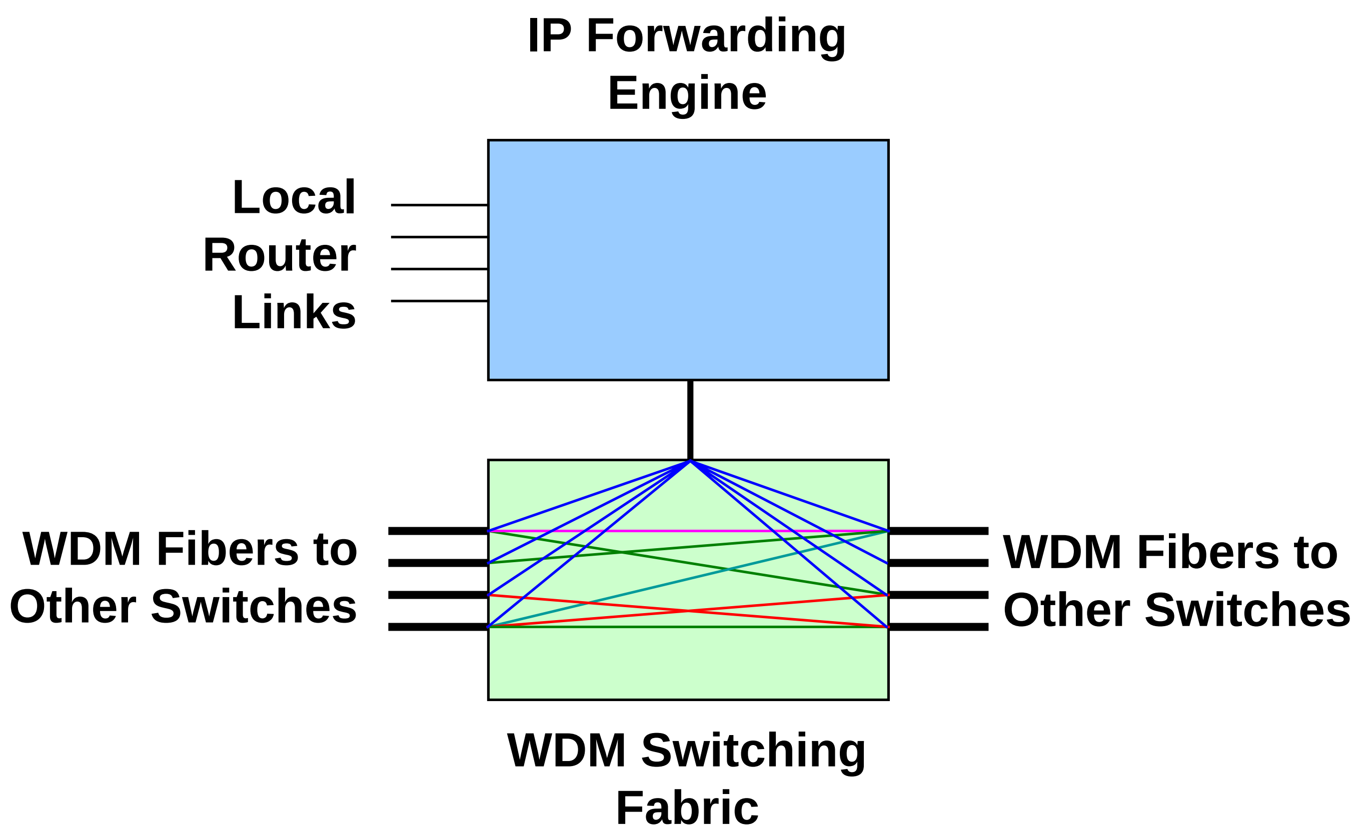
POW Goals

- Scale
  - Number of backbone switches: 10× to 100×
  - Number of active flows: 100× to 1000×
- Performance
  - 80% to 90% of packets switched
  - Gigabit-per-second link speed
  - Higher link utilization
- Robustness
  - No manual configuration
  - Traffic stability after reconfiguration
  - Optical stability after routing changes

POW Assumptions

- State-of-the-practice router
  - 200,000–packets-per-second forwarding rate
  - 50,000-entry routing table
- WDM technology
  - 8 to 64 wavelengths
  - Nonhomogeneity of links

POW Switch–Router Model



Optical Label Switching

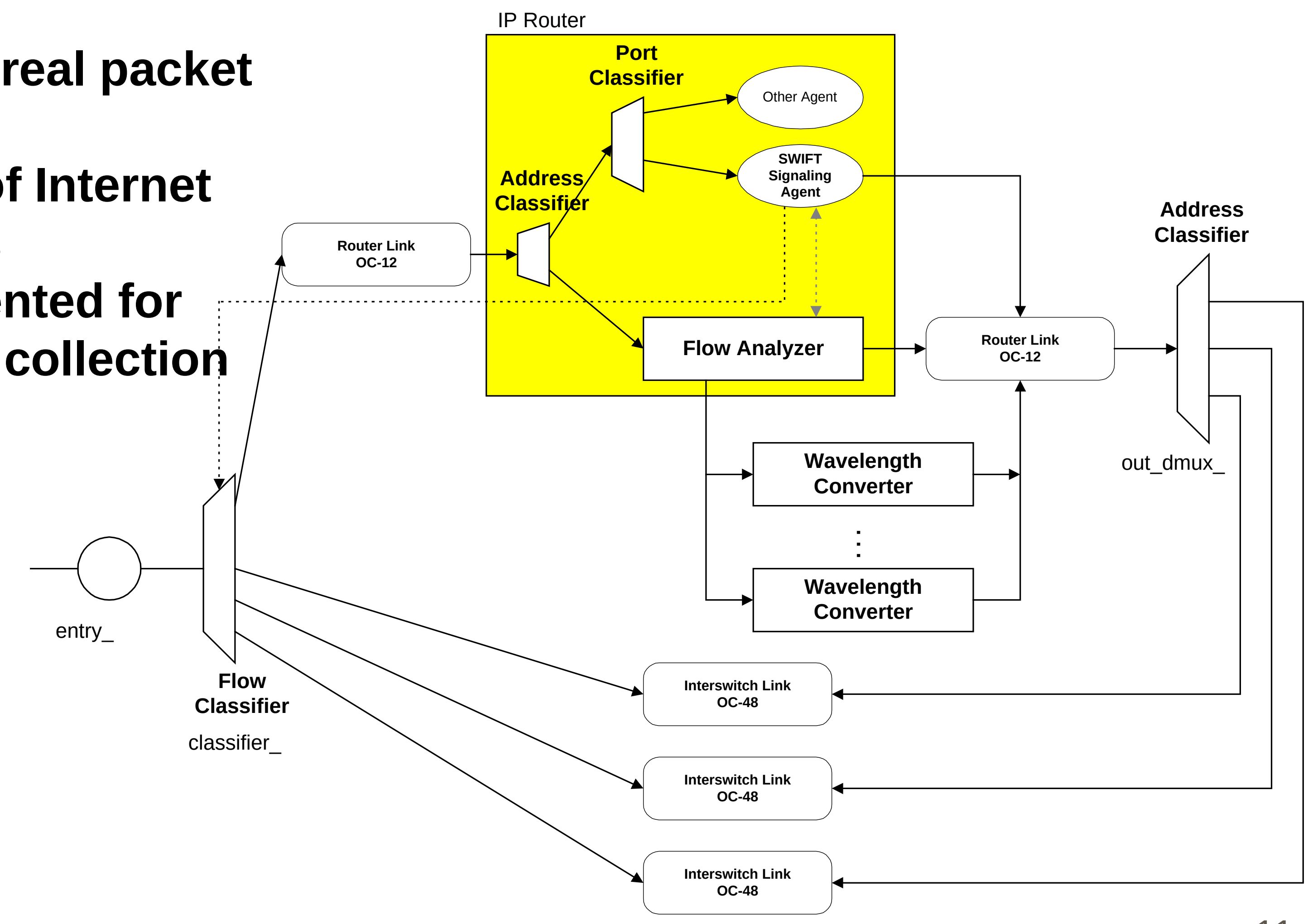
- Like ATM-based label switching or MPLS
  - Far fewer wavelengths than VCs
- Default operation is to route along each hop
- When router detects a *bona fide* flow, it begins a sequence of actions to assign the flow to a free wavelength

Flow Classifier for Label Switching

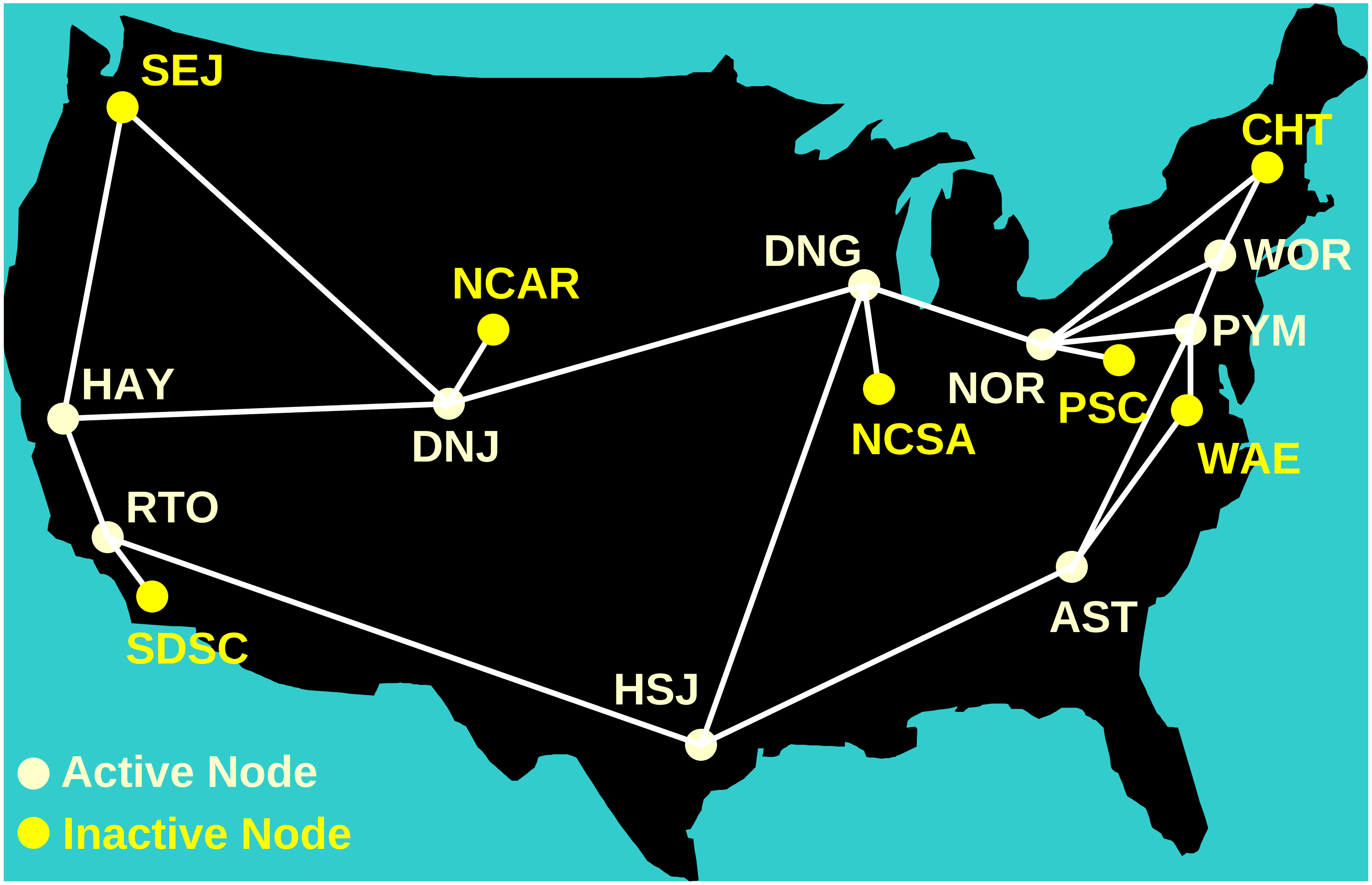
- X/Y flow classifier
- Flow recognition by stream characteristics
  - $\geq X$ packets ($X = 10$)
  - $\leq Y$ seconds ($Y = 20$)
  - Flow is declared switchable
- Flow deletion by stream characteristics
  - $\leq W$ packets ($W = 5$)
  - $\geq Z$ seconds ($Z = 20$)
  - Flow is declared unswitchable

VINT/ns Simulation Model

- Accepts real packet dumps
- Library of Internet protocols
- Instrumented for statistics collection



vBNS Backbone Topology



Traffic Model

- Real traffic traces
  - vBNS traffic matrix
  - LBNL packet traces
- vBNS 9/18/98 hourly statistics
  - All-pairs matrix
  - Low load
- LBNL 1/28/94 tcpdump traces (LBL-PKT-5)
  - 363,000 packets for 1 hour
  - Low throughput

vBNS Traffic Matrix

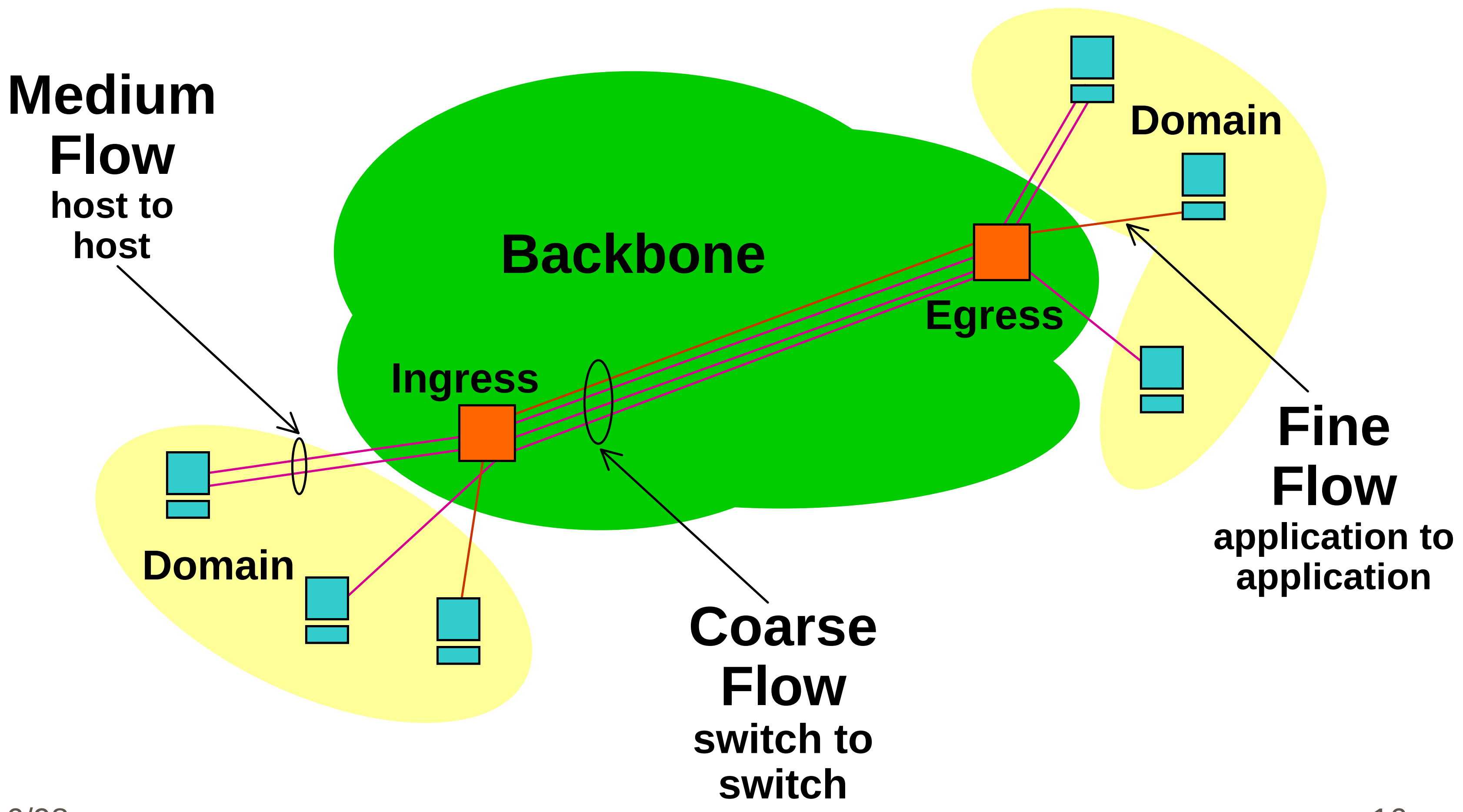
	AST	DNG	DNJ	HAY	HSJ	NOR	PYM	RTO	WOR
AST	0.00000	0.56745	0.03781	0.00515	0.06092	0.09476	0.21693	0.00617	0.01081
DNG	0.08101	0.00000	0.11513	0.00379	0.61243	0.08101	0.08362	0.03870	0.01543
DNJ	0.03311	0.18448	0.00000	0.03595	0.34106	0.01441	0.08419	0.30353	0.00326
HAY	0.15571	0.00263	0.13758	0.00000	0.15903	0.08540	0.35294	0.07668	0.03004
HSJ	0.04519	0.00009	0.78623	0.00666	0.00000	0.03361	0.10832	0.01152	0.00839
NOR	0.01889	0.04205	0.01987	0.61550	0.01504	0.00000	0.11273	0.00149	0.17443
PYM	0.40164	0.00026	0.14268	0.01756	0.12685	0.17532	0.00000	0.04646	0.08924
RTO	0.01402	0.00025	0.89763	0.00399	0.01936	0.00807	0.05224	0.00000	0.00444
WOR	0.01372	0.84503	0.00483	0.00075	0.00740	0.02611	0.09917	0.00298	0.00000

Traffic measured hourly on September 18, 1998, and averaged over the day.

Flow Granularity

- Coarse: aggregated by ingress and egress points
  - <srcdom, destdom>
- Medium: aggregated by IP addresses
  - <srcdomain, srcaddr, destdom, destaddr>
- Fine: no aggregation
  - <srcdom, srcaddr, srcport, destdom, destaddr, destport>

Granularity Illustrated



Routing in POW

- The interior routing protocol must reveal both the next-hop and egress routers
- This capability is provided by modern routing protocols
  - Link-state protocol
  - Distance/path-vector protocol

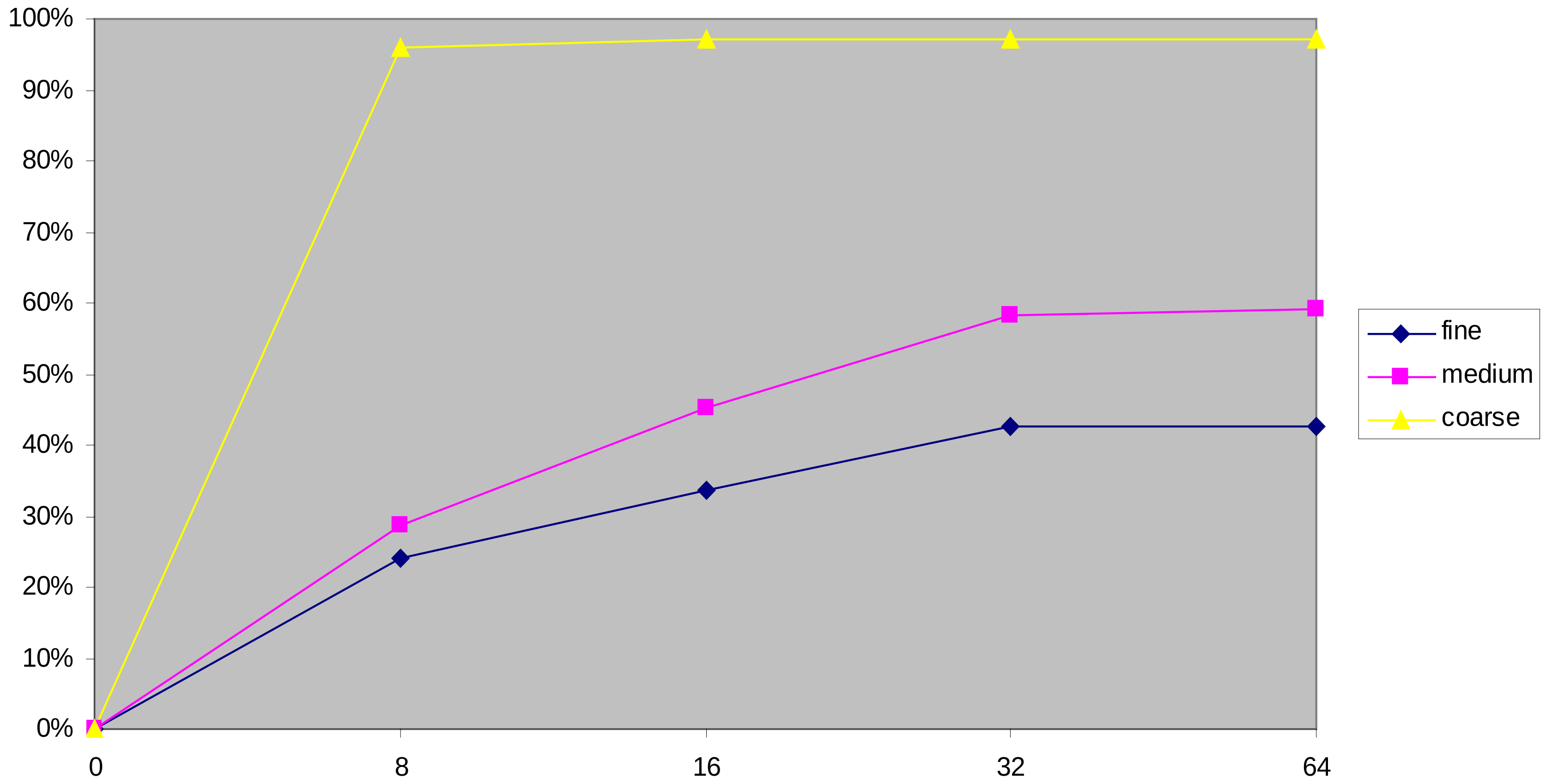
POW Signaling Protocol

- Assume that the interior routing protocol can produce <first_hop, last_hop> information for each destination address
  - IS-IS
  - EIGRP
- Signaling protocol uses first_hop and last_hop information to set up a switched lightpath between these endpoints

Simulation Results

Number of Wavelengths	Flow Granularity	Signaling Overhead	Packets Switched
8	Fine	0.73 %	24.05%
	Medium	0.64 %	28.63 %
	Coarse	0.37 %	96.05 %
16	Fine	1.33 %	33.54 %
	Medium	1.32 %	45.31 %
	Coarse	0.34 %	97.19 %
32	Fine	9.22 %	42.72 %
	Medium	6.09 %	58.39 %
	Coarse	0.34 %	97.16 %
64	Fine	11.91 %	42.61 %
	Medium	7.10 %	59.24 %
	Coarse	0.34 %	97.22 %

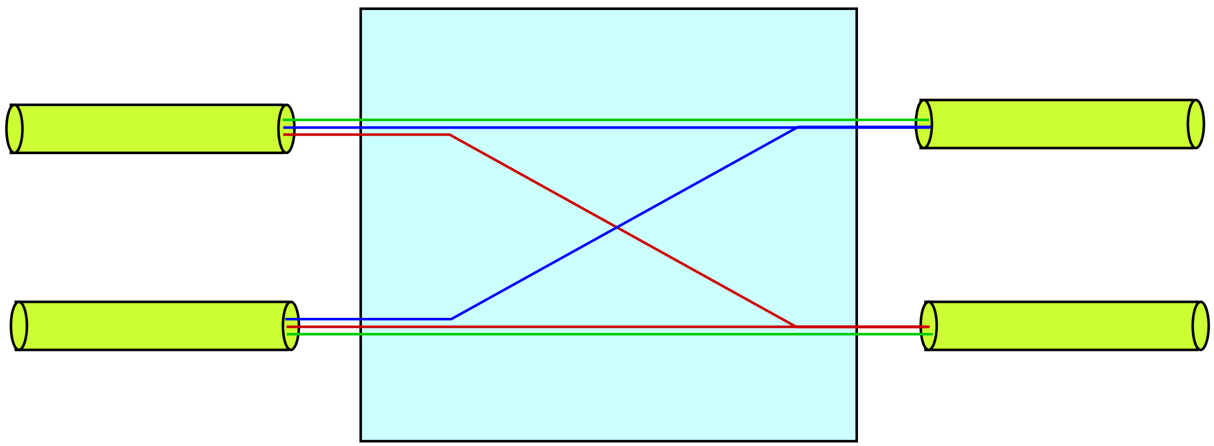
Switching Performance



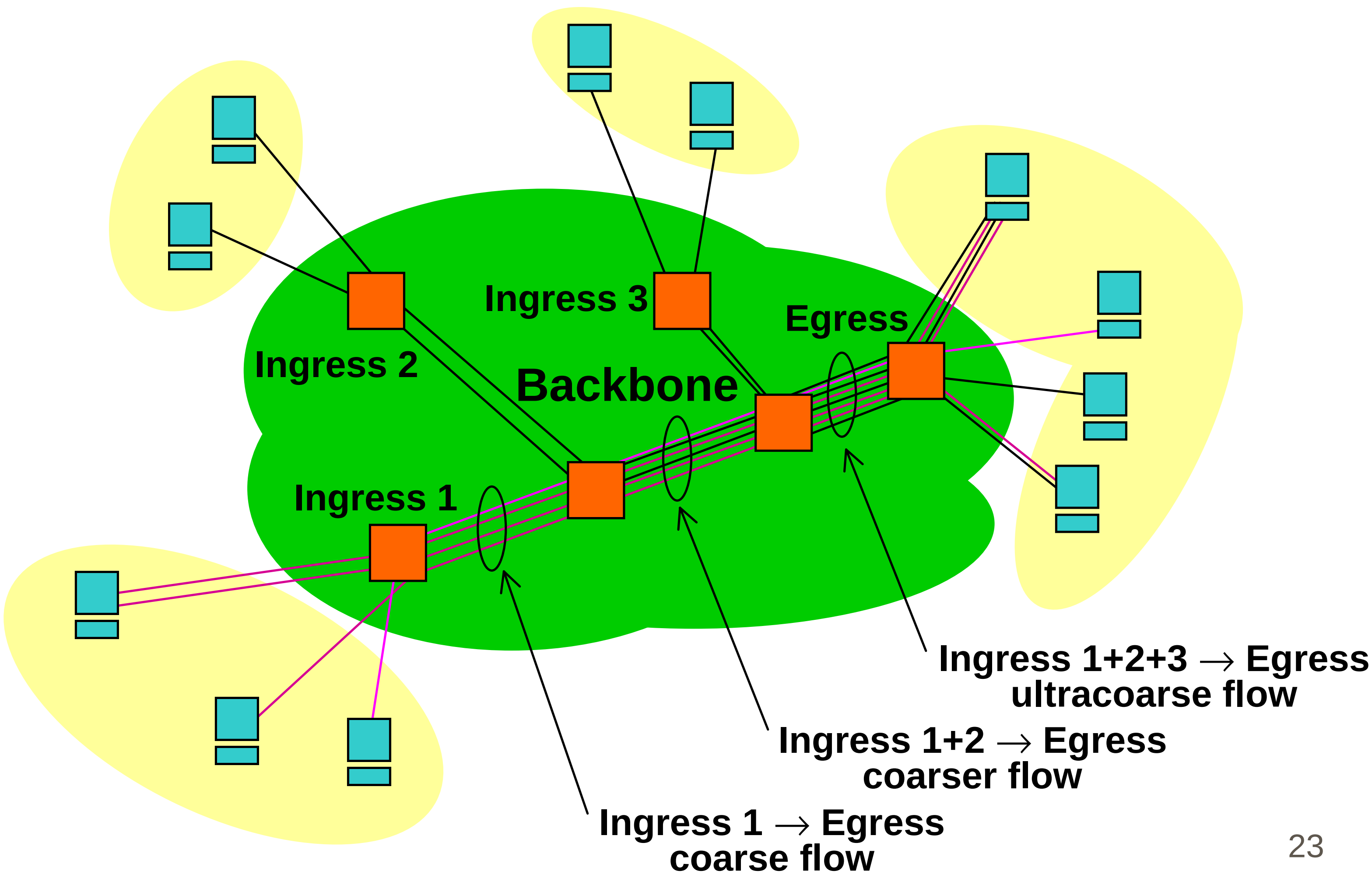
Wavelength Merger

- Further aggregation can be achieved by merging all flows that exit the backbone via the same egress switch
- Requires the use of a WDM switching fabric that can multiplex a wavelength from different input fibers

Wavelength Merger Fabric



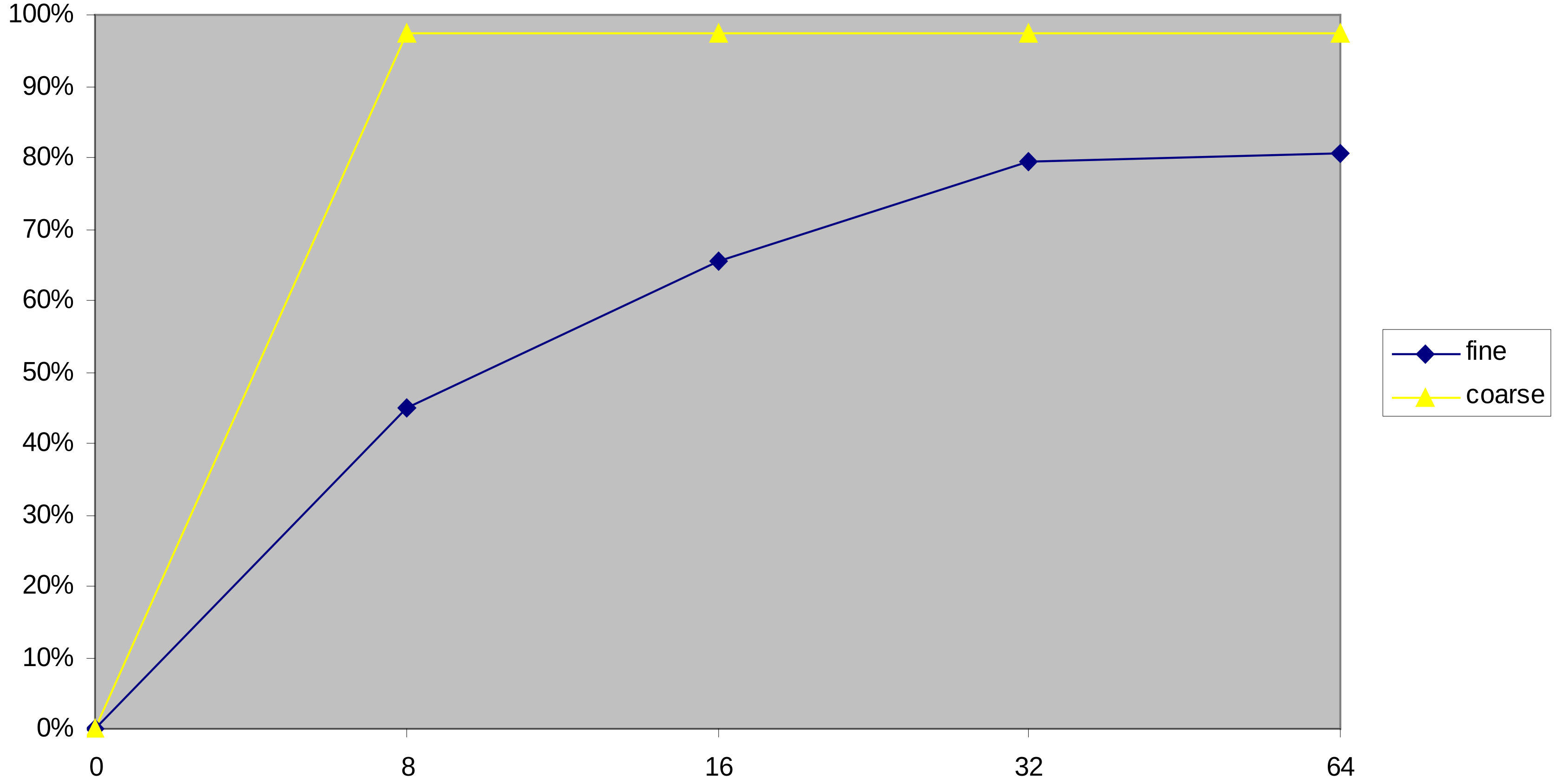
Wavelength Merging



Simulation Results for Merging

Number of Wavelengths	Flow Granularity	Signaling Overhead	Packets Switched
8	Fine	0.37 %	45.03 %
	Coarse	0.18 %	97.33 %
16	Fine	0.97 %	65.39 %
	Coarse	0.19 %	97.28 %
32	Fine	1.73 %	79.43 %
	Coarse	0.18 %	97.36 %
64	Fine	1.80 %	80.67 %
	Coarse	0.19 %	97.26 %

Switching Performance with Merging



Caveat Auditor

- Packet traces were collected in 1994
- Packet traces represent low-throughput traffic
- Technical difficulties prevented the use of more-suitable traces
  - Newer and better traces will be simulated

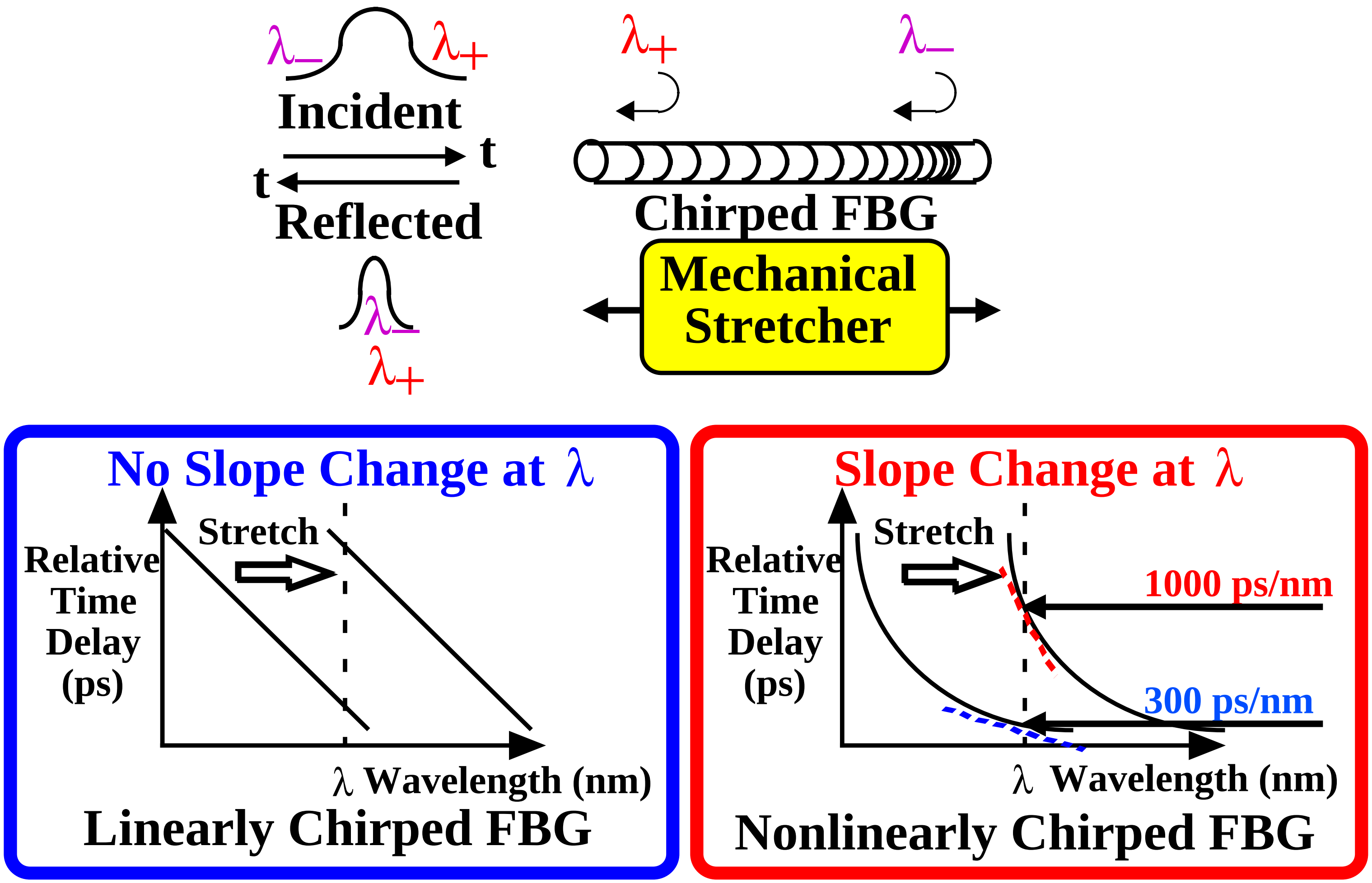
System Hardware Requirements

- Merging of two or more same-wavelength inputs onto one output of the same wavelength
- Configurable from the router in microsecond epochs
- Resolution or avoidance of contention or collisions between data in the merged streams

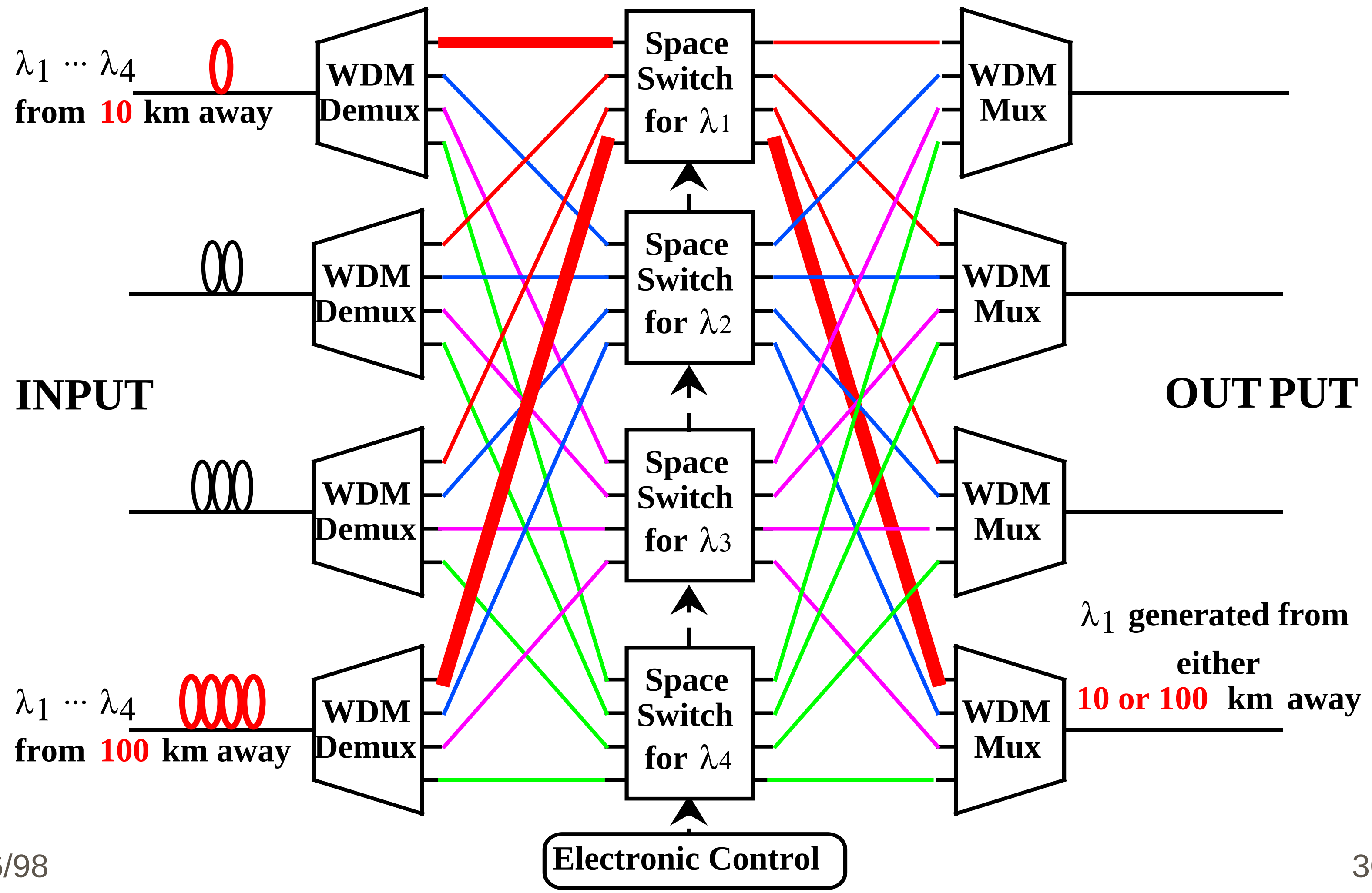
Functions Enabled by Optics

- Tunable dispersion compensation
- Synchronization
- Tunable header recognition for routing
- Time or subcarrier multiplexing of wavelengths
- Baseband and subcarrier signaling

Tunable Nonlinearly Chirped FBG



Dynamic Dispersion Compensation



Dynamic Dispersion Compensation Using Nonlinearly Chirped FBG

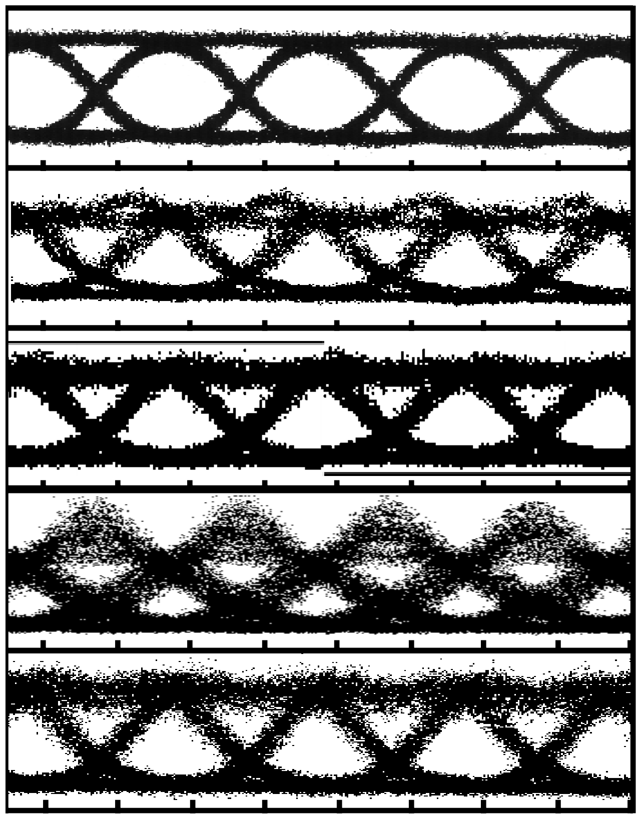
10 Gb/s baseline

50 km no compensation

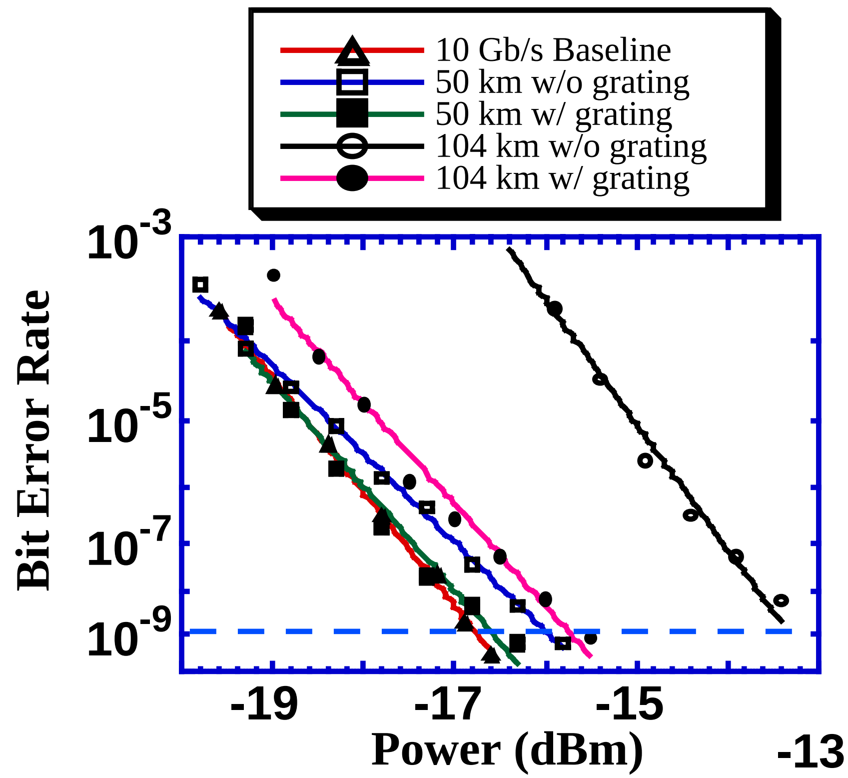
50 km w/ grating compensation

100 km no compensation

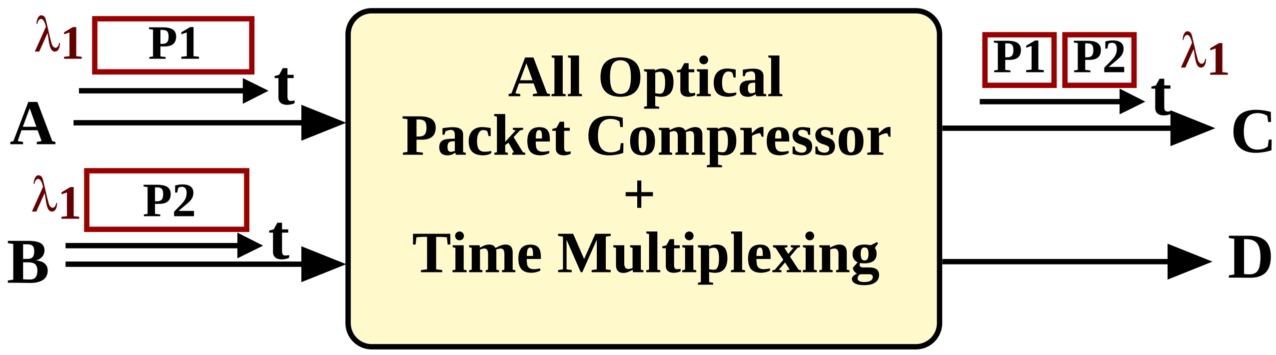
100 km w/ grating compensation



50 ps/div



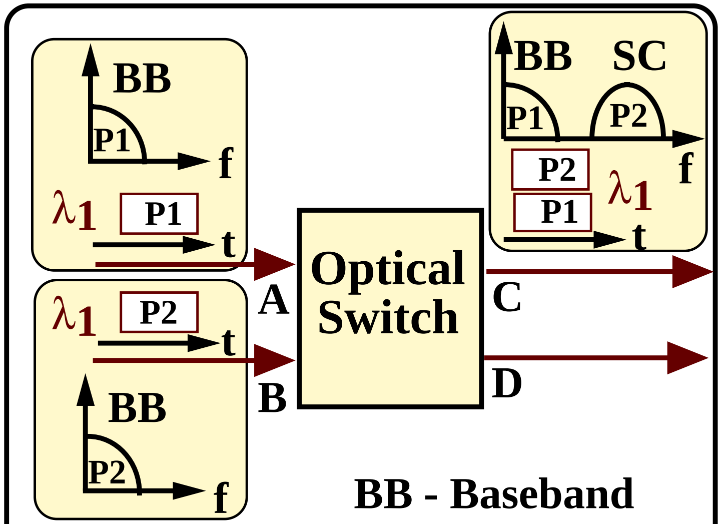
Packet Compression for Contention Resolution



- **Packets can be compressed**
    - **Optical Chirping techniques**
    - **optical buffers**

Multiplexing for Contention Resolution

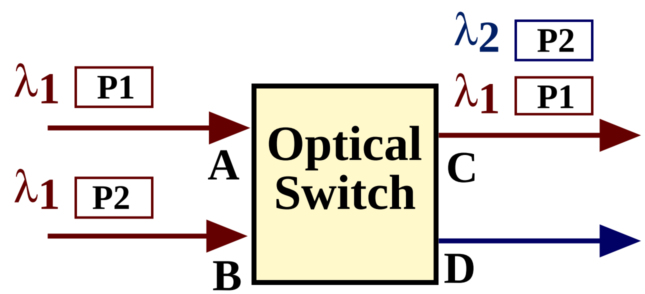
Subcarrier Multiplexing (SCM)



BB - Baseband
SC - Subcarrier

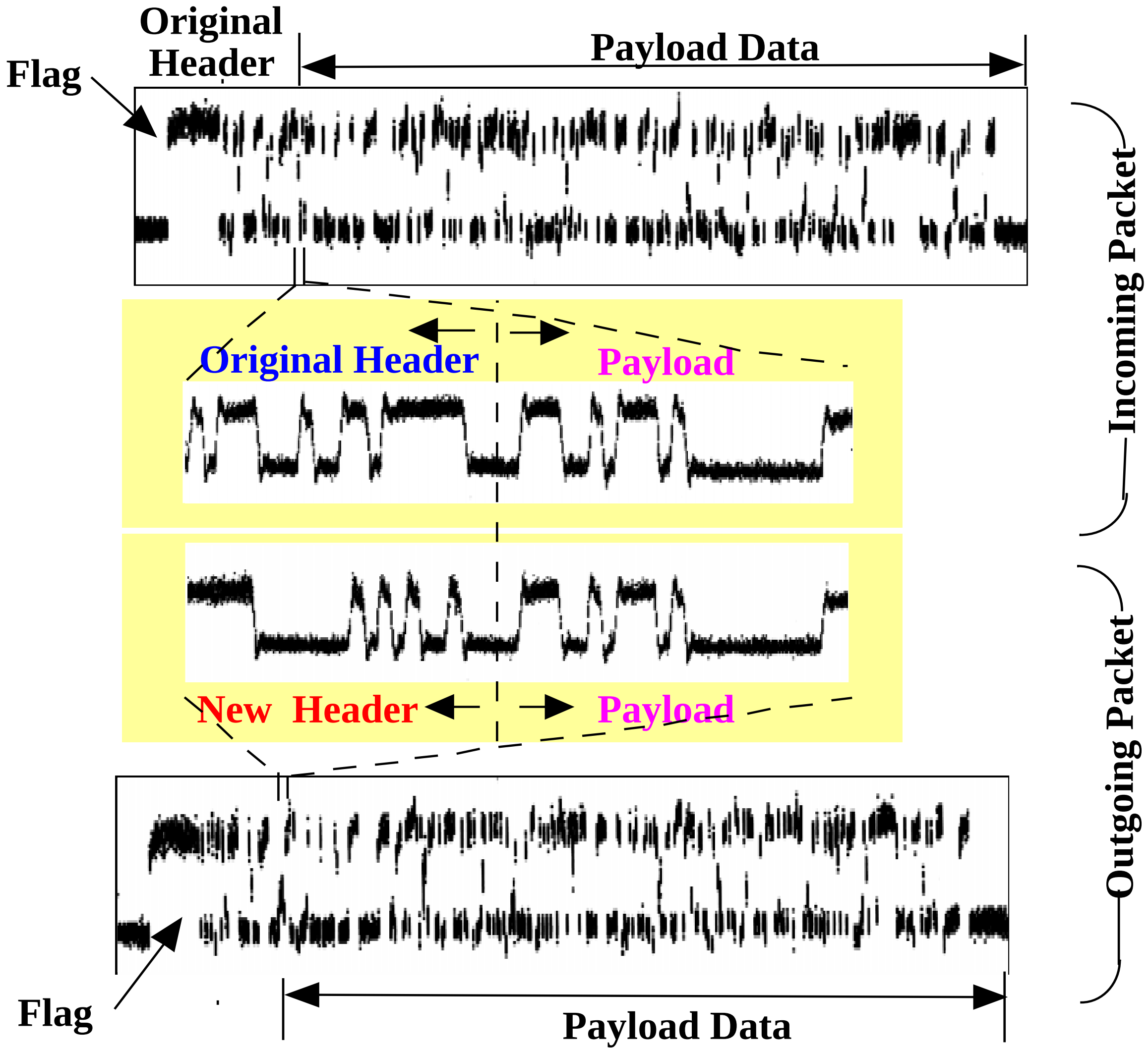
- No extra wavelength required

Wavelength Shifting (WS)

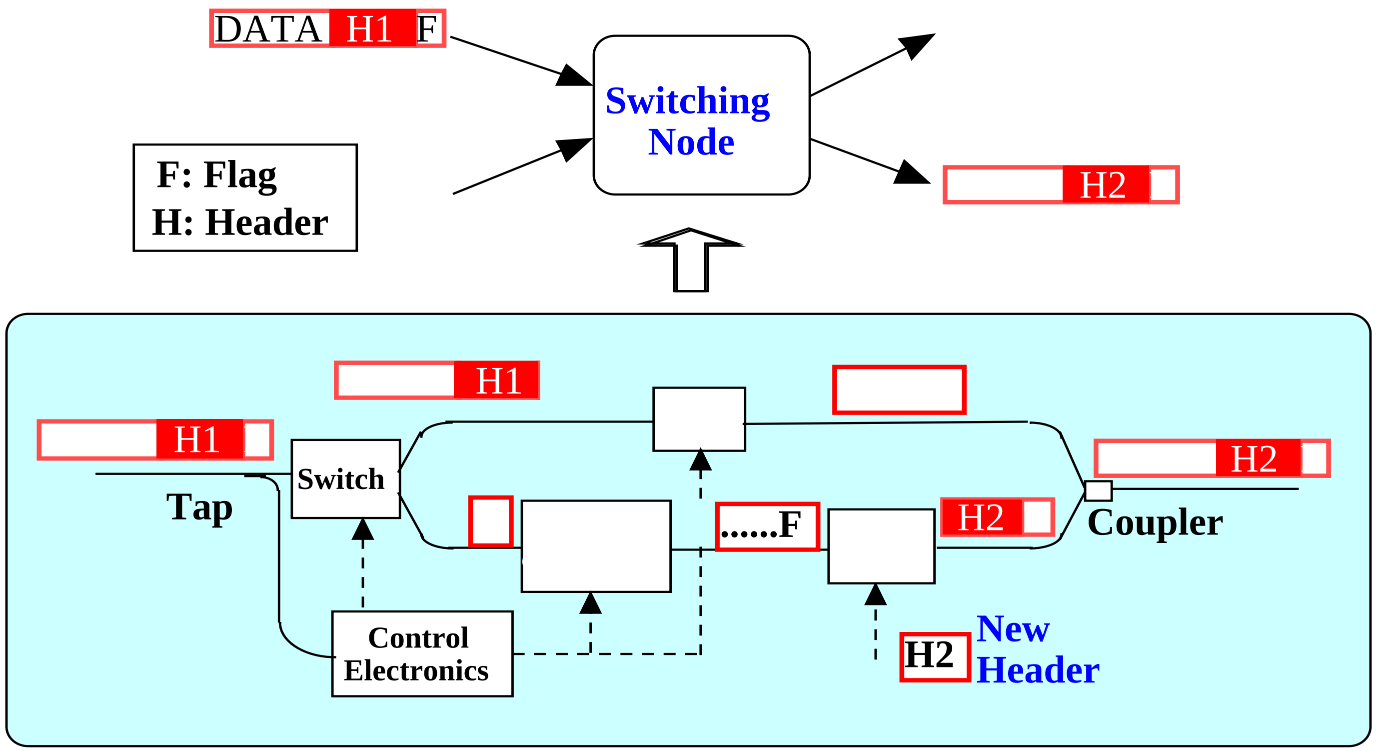


- Uses wavelength domain

Identical-λ Header Replacement



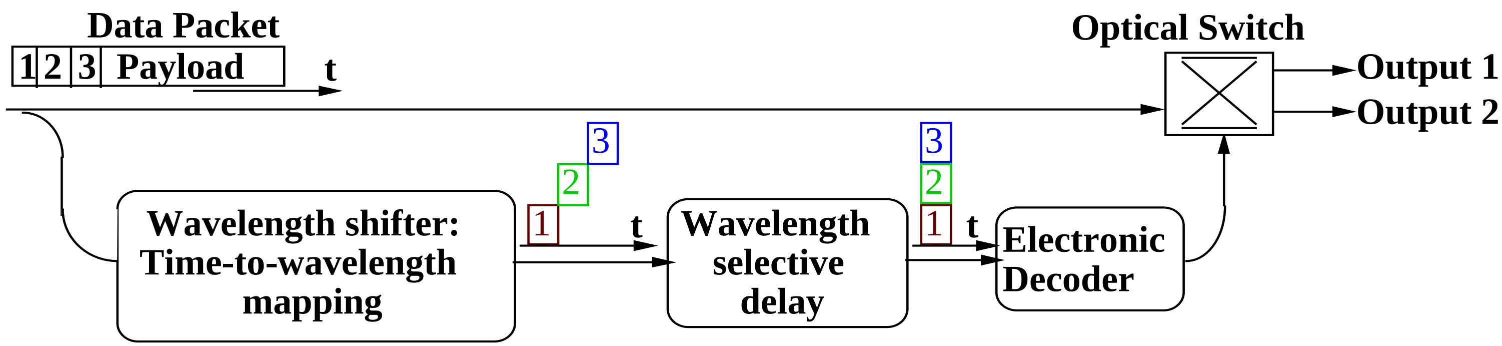
All Optical Identical-λ Header Replacement

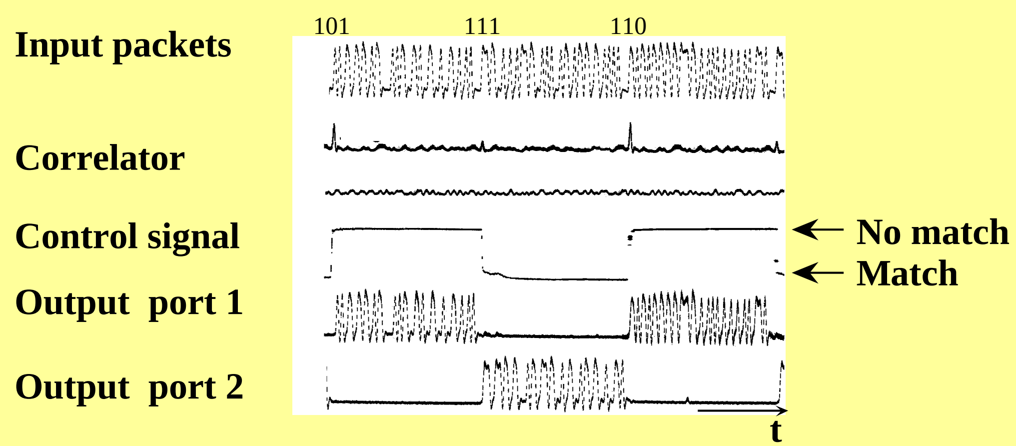


- Must ensure that output is at *exactly* the same λ as input

All-Optical Header Recognition & Switching

USC INFORMATION SCIENCES INSTITUTE

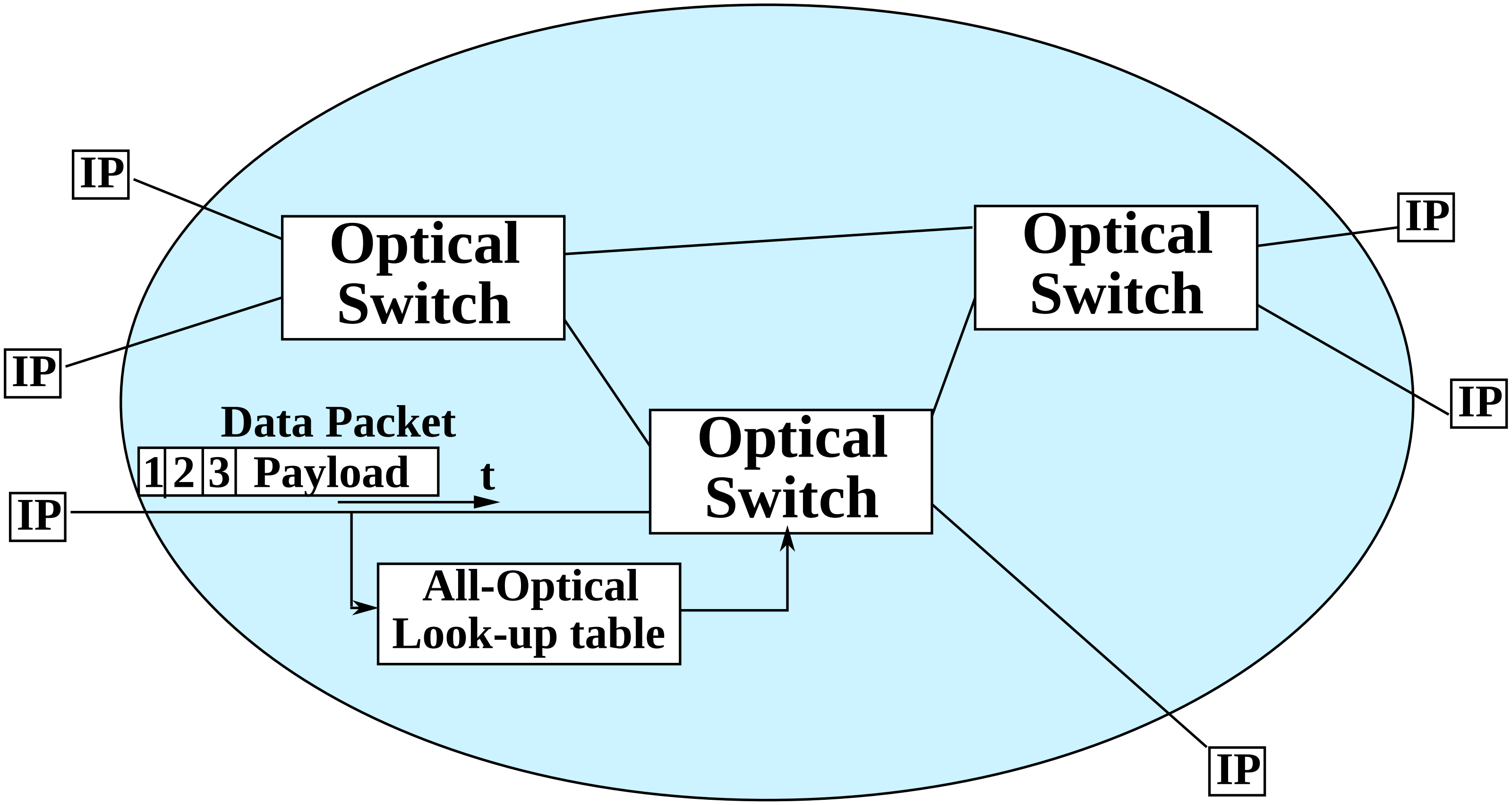




2.5 Gb/s packets with 1.6 nsec guard time

Accepted to OFC' 99

All-Optical Header Recognition & Switching

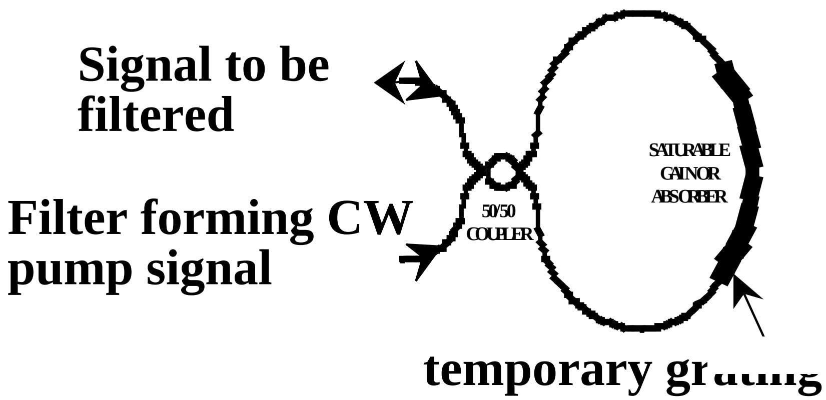


Transparent Optical Core

Tunable RF Fiber Loop Mirror Filter

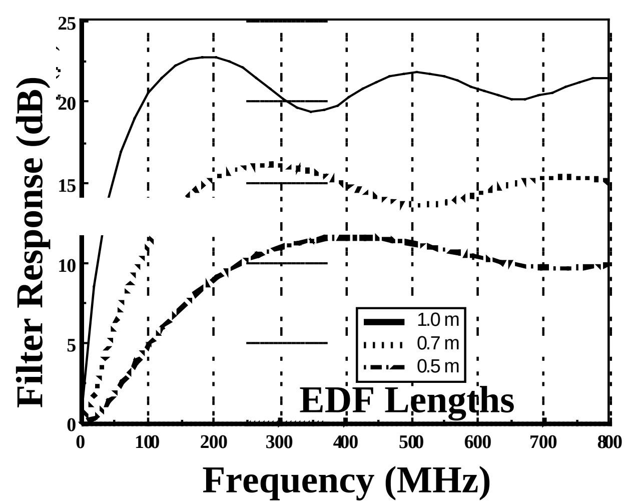
USC INFORMATION SCIENCES INSTITUTE

- Bandwidths from sub-MHz to beyond GHz
- Configurable to either bandpass or notch
- Built in interferometric phase alignment



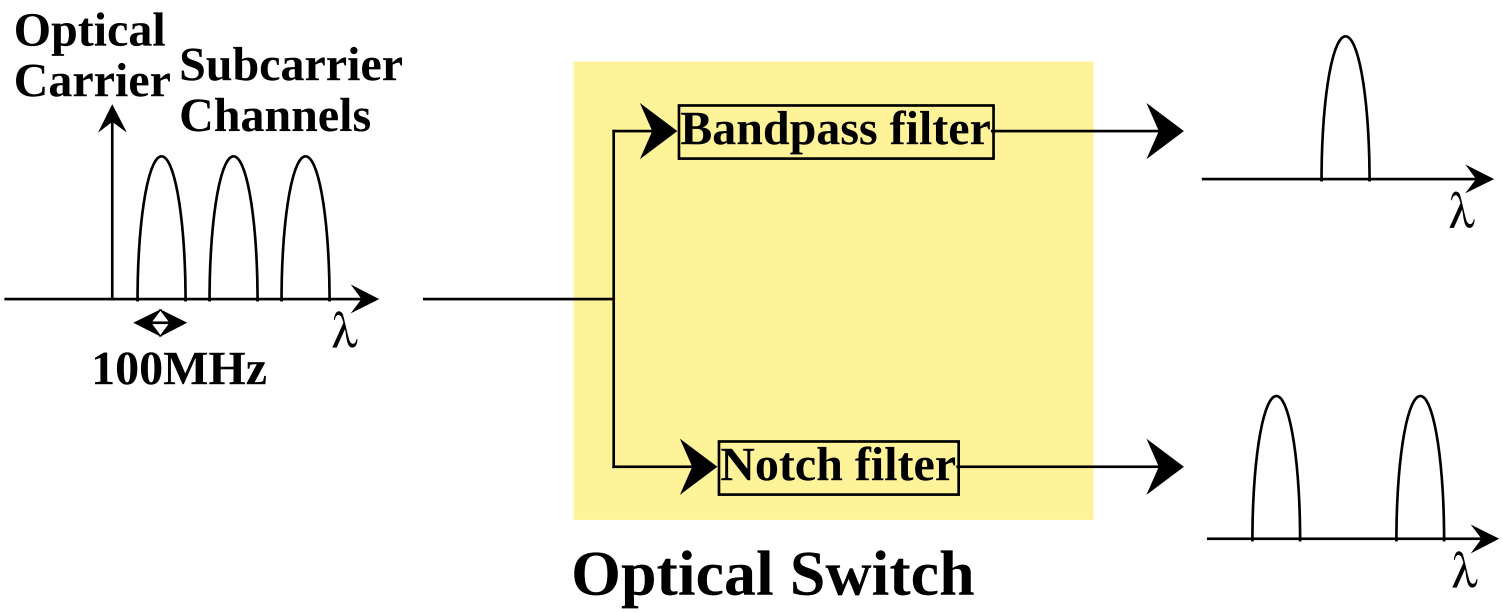
- Pump signal forms a temporary grating that filters signal

Theoretical response



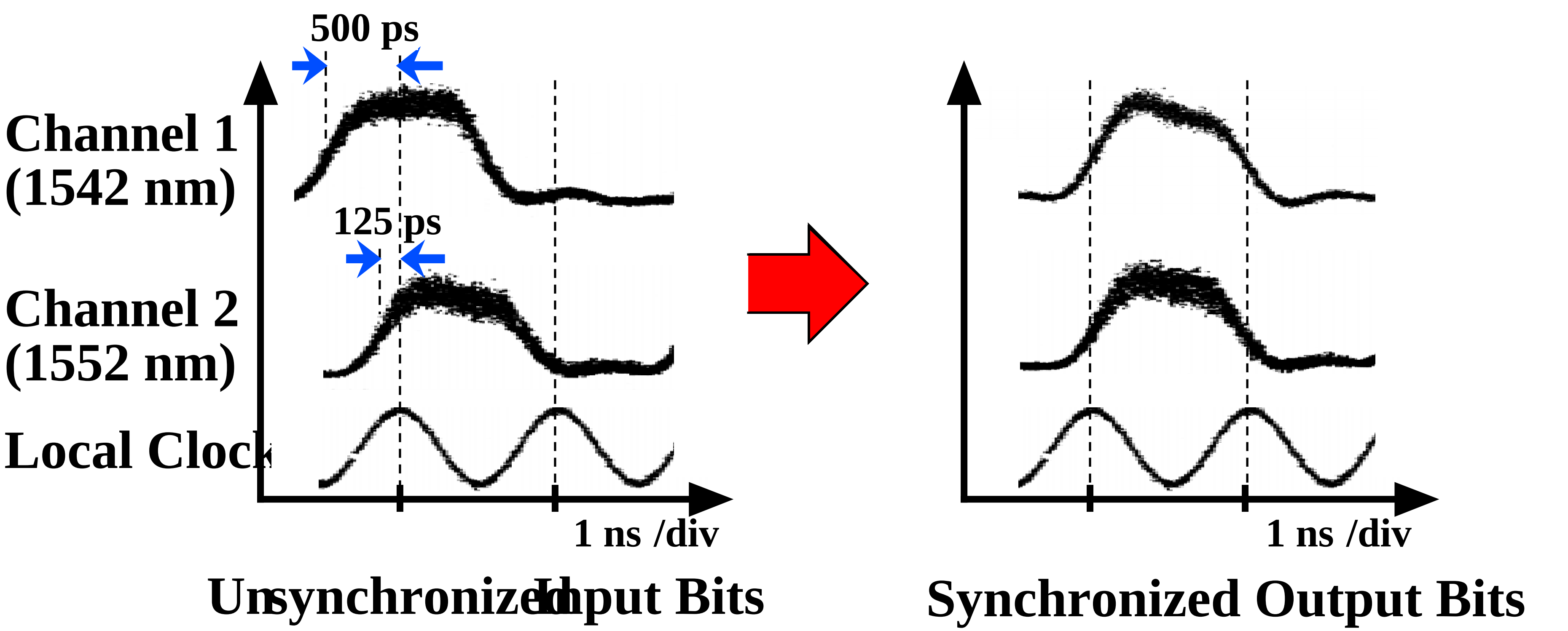
Accepted to OFC' 99

Tunable RF Fiber Loop Mirror Filter



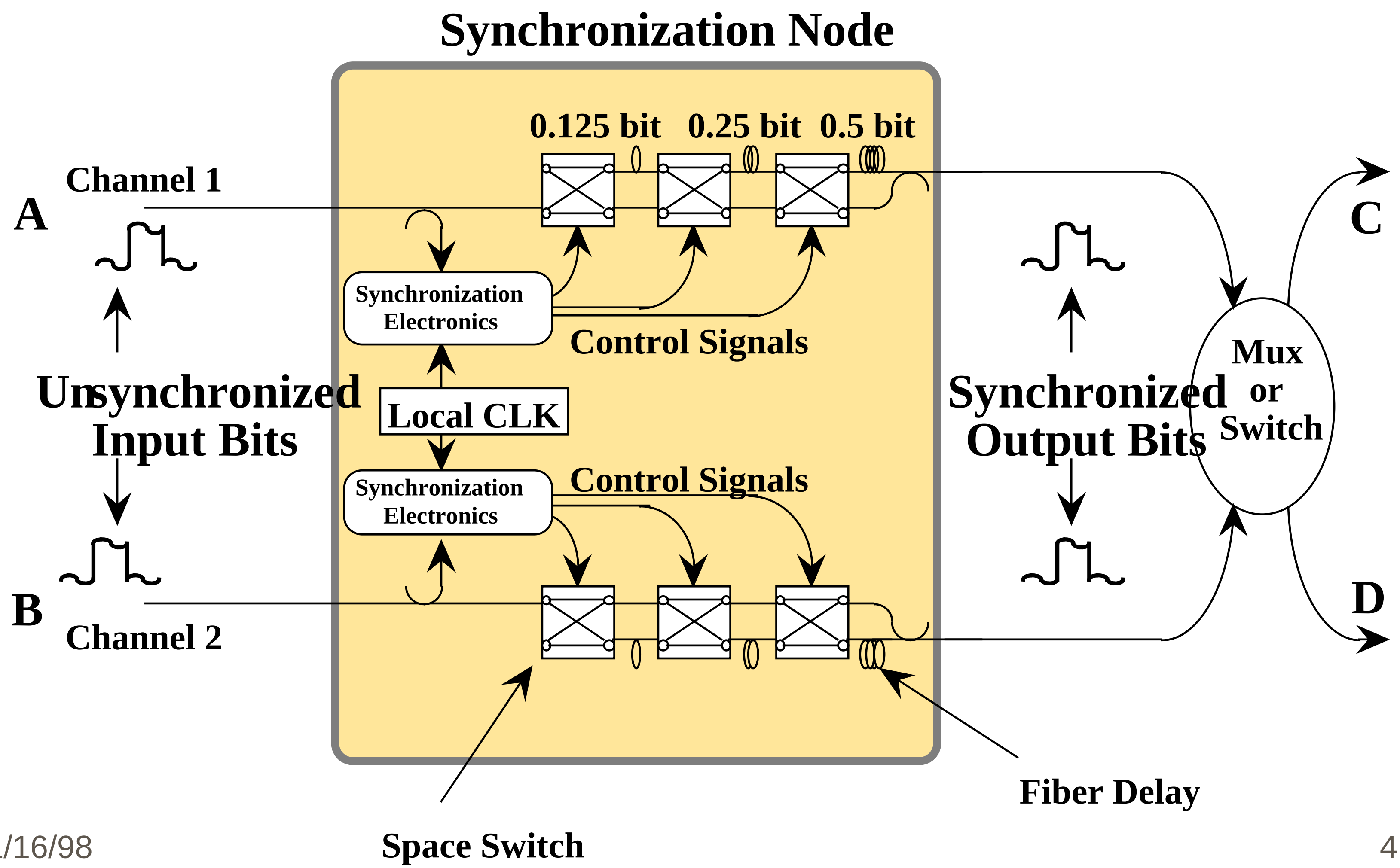
- Subcarriers used for routing in NGI
- Current Optical filters for RF signal processing:
  - Limited tunability
  - FWHM is too wide

Bit Synchronization Using Subcarriers



- Resolution <1/8 bit time

Bit Synchronization Using Subcarriers



Conclusion

- To perform well in a wavelength-limited ($\leq$ 64 λ) system, POW requires intensive flow aggregation
- Further study is required
  - Better traffic traces
  - Large-scale networks
  - Technology enhancements (e.g. λ-merger)