100 Tb/s Aggregate Capacity Router using an Optical Switching Core

Prof. Mohammed N. Islam
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor
I. Motivation and Router Background
II. Broadcast & Select Optical Switching Core
III. Sub-system Issues with B&S Architecture
IV. Key Enabling Technologies (Components)
V. Summary
Relative performance increase

- DWDM Link speed: \(x2/8\) months
- Internet: \(x2/yr\)
- Router capacity: \(x2.2/18\) months
- Moore’s law: \(x2/18\) m
- DRAM access rate: \(x1.1/18\) m
First Generation Routers

Typically <0.5Gb/s aggregate capacity
Limitation: buffer memory
Second Generation Routers

Typically <5Gb/s aggregate capacity
Limitation: bus interconnection
Third Generation Routers

Switched Backplane

Typically <160Gb/s aggregate capacity
Limitation: number of LC’s in a rack
4th Generation Routers/Switches
Optics inside a router for the first time

Switch Core

Optical links

100s of metres

Linecards

0.3 - 10Tb/s routers in development
Fundamental Problems

• We are designing for 100Tb/s next generation

• Fundamental Problems
  – Switching Core
    » Electronic switch requires a rack for 2Tb/s switching core
    » Size, power dissipation and cost prohibitive for 100Tb/s
  – Scheduler Algorithm
    » Algorithm typically grows as some power of N, as much as N^3
    » Takes more and more time to compute algorithm as N gets larger
100Tb/s Routers using DWDM?

- Reasonable bit rate: 40Gb/s
- Spectral efficiency: 4 b/s/Hz (using sophisticated coding, super-FEC, good filtering)
- Telecom window: 1400-1650nm (250nm)
- 100Tb/s = 40Gb/s x 2500 channels
- For channel spacing 12.5GHz (0.1nm), telecom window has 2500 channels
- Bottom line: doable, but tough
Goals of Design

• **Low cost, low power dissipation, small size**
  – Lowest cost using tunable filters
  – Capacitive devices
  – Packaging or array of devices

• **Combine interconnect with switching fabric**
  – Fiber and star couplers

• **Functionality of OXC and Router**
  – Same DWDM on outside and inside

• **Minimal impact on line cards**
  – Tunable filters in the common bay equipment

• **Hide switching time of devices**
  – Use architectural tricks to use slower, cheaper devices
5th Generation Routers/Switches
Switching Fabric and Fiber Interconnection Combined

No Intermediate O/E/O for Interconnect
Multi-Channel Fast Tunable Filters

Filter Array
Filters in Common Bay Equipment

Optical Switch

Power Supplies

Midplane

Control Boards
Advantages of B&S Switching Fabric

• Hitless reconfiguration and hide switching time
• Multicast & Broadcast traffic without copies
• OXC/Router functionality combined
• Passive, transparent switching fabric scalable
• LC’s minimally impacted because tunable filters and circuits in common bay equipment RU
• Low-cost approach
  – Power splitters/taps cheaper than WDM
  – Transmitters can be low-cost EAM/DFB lasers
  – All but fast tunable filters are off-the-shelf
Comparing with Tunable Lasers

- Significant space on Line Card!
- Switching time ~30-100nsec minimum
- External modulators required (~3-4GHz bw)
- Bottom Line: tunable filters much cheaper

[ Lucent, IEEE PTL, July 2001 ]
Issues for B&S Switching Core

- Scalability using amplifiers
  - Overcoming 1/N loss using WDM and amplifiers
- Emulating fast switching time using inexpensive devices
  - Speed-up, aggregation, ping-ponging between filters
- Simplifying scheduler for large N
  - Two stage switch using inexpensive switching fabric
- Continuum source for >100 channels of 40Gb/s
  - One expensive laser and SC copies to >100nm wavelengths
Scalability Issues

- Inherently have a 1/N loss with Star Coupler
- For 2500 channels, loss of 34dB
- If use multiple filters per LC, then 3-6dB additional loss
- By way of reference, a typical optical link will have a budget of about 30dB (~22dB loss budget and another ~8dB for OADM/DC, etc)
Broadcast & Select Switching Fabric

Tuning Schedule

TOF 1 Tune  TOF 1 Tx  TOF 1 Tune  TOF 1 Tx
TOF 2 Tune  TOF 2 Tx  TOF 2 Tune  TOF 2 Tx
By-pass traffic
Emulating fast switching using low-cost components

- 50-100ns devices cost << than sub-ns devices
- Speed-up
- Aggregation and superframes
- Ping-pong between different filters
Simplified Scheduler with 2-Stages

- Simple round-robin scheduler in two sections
- Middle stage may have additional memory to avoid mis-sequencing of packets
- 2x switching fabric, so fabric must be low cost!
- Questions to scheduler experts:
  - Does the switching fabric need to reconfigure every cell cycle (in which case our approach is important)
  - Are LC’s needed in all three locations, or can we get away with two (in which case fabric is simplified)
Two-Stage Switch

Switch gives 100% throughput for non-uniform, bursty traffic, without a scheduler or speedup!

[Nick McKeown, Stanford University, Opticomm 2001]
Exemplary 2-stage Switching Fabrics
Cost Estimate for 1Tb/s Core (2004)

- Assume 100 ch @ OC-192
- Transmitter: $1.7k \times 100 = 170K
- WDM: $0.1K\times100 = 10K
- Amplifier: $ = 50K
- 1xN splitter: $0.05K\times100 = 5K
- Tunable filter: $0.1k \times 100 = 10K
- Total: $245K
- Connectors, electronics, cases, software, extra
- Cost < $200K if we use broadcast & select without amplifier
Laser Sources in Large Router

- Many LD’s become expensive
- If channel spacing is close, then stabilizing wavelengths and maintaining channel spacing difficult and expensive
- For 40Gb/s (or 160Gb/s) per channel, light source can be expensive
- For many LD wavelengths, many part #’s and have to match LD wavelength per LC
- With 40Gb/s sources and stabilization circuits, significant space on LC will be used
Supercontinuum (SC) Source

- One modelocked source and a common SC set-up
- Channel spacing set by passive WDM demux, which is used to carve out channels from SC
- For 40Gb/s (or muxed to 160Gb/s), only one expensive ML source required. SC copies to many wavelengths
- Each LC can have a modulator, but individual LD’s are not required. If modulator broadband, few part #’s
- ML laser and SC set-up will be in common bay equipment. Only modulator placed on LC

**BOTTOM LINE:** SC less expensive for #’s >100 and 40Gb/s per channel or higher
Exemplary System

- SC source can all be placed in common equipment bay
- Modulator placed on line card
- WDM can be replaced by power splitters and fixed or tunable filters
SC Experimental Setup

Mode Locked Ring Cavity EDFL

15 dB EDFA

Pulse Chirping (2m SMF-28)

SC Generation

\( P_0 \) → \( L \) → \( (D) \)

Diagnostics

Optical Spectrum Analyzer

Autocorrelator

L: only 2 meters long

\( t = 400 \text{ fs} \)
Exemplary SC Spectrum

**Experimental Parameters**

- $L = 2$ [m]
- $\Delta_0 = 1539$ [nm]
- $P_{\text{avg}} = 1100$ [W]
- $D = 1.13$ [ps/nm-km]

-20 dB Bandwidth: 211 nm
SC Spectral Flatness

±0.5 dB maximum power fluctuation over 61 nm

± 0.1 dB power fluctuation over 35 nm
Coherence of Carved SC Spectrum

- Pulse width $t_{\text{FWHM}} \lesssim 500$ fsec can be carved from SC for high-speed TDM applications.
Timing Jitter of Source and Filtered SC Output

- Pulses carved from flat section of continuum have same timing jitter as the source
- Short SC fiber length minimizes additional timing jitter

Measurement Parameters
- 5 KHz Span
- 30 Hz Resolution
- 430th Harmonic of laser fundamental
Key Enabling Technologies (Components)

- Broadband Amplifiers
- Tunable Filters
- Surface Normal Modulators
All-Raman Broadband Amps

- Very low MPI level
- Dispersion and slope compensation provided by the gain fiber
- Large gain & low NF over a 100nm continuous spectral window
- Demonstrated the transmission feasibility of 240 OC-192 channels over > 1500km SSMF
Line amplifier flat operation over 100nm demonstrated at $+24\text{dBm}$ output power (corresponds to $0\text{dBm/ch}$ for 240 channels)
Electro-optic Tunable Filter

- Optical cavity with electro-optic material
  - Tune filter by voltage induced index changes of EO material

Filter characteristics

<table>
<thead>
<tr>
<th></th>
<th>1-Cavity Filter</th>
<th>3-Cavity Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1 dB BW</td>
<td>25GHz</td>
<td>25 GHZ</td>
</tr>
<tr>
<td>-30 dB BW</td>
<td>625 GHz</td>
<td>100 GHz</td>
</tr>
<tr>
<td>In-Band Ripple</td>
<td>&lt;0.25 dB</td>
<td>&lt; 0.25 dB</td>
</tr>
</tbody>
</table>
Pass-band shape vs. number of cavities

<table>
<thead>
<tr>
<th>Filter Type</th>
<th>Fabry Perot</th>
<th>Narrow Band</th>
<th>Wide Band</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter Shape</td>
<td>Lorentzian</td>
<td>Square</td>
<td>Square</td>
</tr>
<tr>
<td># Channels</td>
<td>1 channel</td>
<td>1 channel</td>
<td>4-16 channels</td>
</tr>
<tr>
<td># Cavities</td>
<td>1 cavity</td>
<td>3 cavities</td>
<td>8-10 cavities</td>
</tr>
<tr>
<td>Application</td>
<td>Channel Monitor</td>
<td>OADM, MUX/DeMux</td>
<td>OADM</td>
</tr>
</tbody>
</table>
A tuning range of 15 nm is obtained with electric fields of -1 MV/cm to 0.5 MV/cm across each cavity.

The pass-band ripple increases near the extreme ends of the tuning range:
- The transmission ripple remains less than 0.1 dB in the 15 nm tuning range.
- The group delay ripple increases by ~1 ps.
Alternate Approach– 1D MEMS

• Combine optical functionality with well-known electro-static actuators
• Challenge: combining MEMS actuator growth with optical coating technology

simple piston up-down motion
FIMS– Fast Interferometric MEMS Switch

• High speed achieved because
  – Interferometer: maximum displacement $l/4$
  – Stress: make a tight guitar string
  – 1D: simple 1D motion with simple control

• High reliability expected
  – Small motion without hinges
  – Electro-static actuators well-proven technology

• Low cost
  – Simple packaging
  – Standard processing steps with many devices on a wafer
## Predicted Performance of FPI

High speed while maintaining optical performance and low voltage

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuning Speed</td>
<td>60 ns mechanical, 100 ns electrical</td>
</tr>
<tr>
<td>Tuning voltage</td>
<td>40 V max</td>
</tr>
<tr>
<td>-3 dB bandwidth</td>
<td>0.1 nm</td>
</tr>
<tr>
<td>-30 dB bandwidth</td>
<td>1.5 nm</td>
</tr>
<tr>
<td>Channel Selectivity</td>
<td>12.5 Ghz (0.1nm)</td>
</tr>
<tr>
<td>Tuning range</td>
<td>100 nm</td>
</tr>
<tr>
<td>Finesse</td>
<td>6300</td>
</tr>
<tr>
<td>Insertion loss</td>
<td>1-3 dB</td>
</tr>
<tr>
<td>PDL</td>
<td>&lt; 0.1 dB</td>
</tr>
</tbody>
</table>
• Tunable filter functions as a surface normal modulator
  -- Operates on single channel/frequency to encode data on cw light
    » Filter is tuned in and out of channel frequency band to create high and low signal states
  -- Sharp transition from high to low transmission
    » Does not require phase shift for high contrast
  -- High-speed
    » Fast EO response + low capacitance
• Surface normal configuration is advantageous for building transmitter array
  -- Source array: Laser diode array, LED array or broadband light source with optical filter array
Summary

- Routers expected to be bottleneck in future systems
- 100Tb/s router project currently at UM
  - Strawman system design to understand limiting technologies
  - Limitations from switching fabric and scheduler
  - Broadcast & Select architecture using DWDM technology

- Sub-system Issues of B&S Architecture
  - Scalability using Broadband Amplifiers & WDM
  - Hide switching time by ping-ponging between filters
  - Two-stage switch for simplified scheduler
  - Broadband continuum source to simplify transmitters

- Key Enabling Technologies
  - Broadband Amplifiers are commercially available
  - Fast tunable filter
    » Two approaches: Electro-optic thin films and 1-D MEMS
  - Surface Normal Modulators can be made with fast filters