High-Performance Multicast Router using an Optical Switching Core

Prof. Mohammed N. Islam
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor
1301 Beal Avenue, Office 1110
Ann Arbor, MI 48109-2122 USA
Ph: (734)647-9700
Email: mni@eecs.umich.edu
OUTLINE

I. Motivation for Heavily Multicast Routers
II. Advantages of Optics for Multicast/Broadcast
III. Optical Switching: Traffic Aggregated into Superframes
   → Broadcast & Select Switching Core with Tunable Filters
   → Sub-system and Components for B&S Architecture
IV. Hybrid Switching: Arbitrary Frame Size
   → Optics for multicast, high-speed backplane, interconnection between racks
   → Electronic switching/multiplexing between detectors and Line Cards
V. Summary
Growing Need for Multicast

• Military Applications
  – War gaming: Generals commanding forces from computers
  – Immediate updates on everyone’s terminals

• High-end Network Users
  – Distributed data repositories and digital libraries
  – Real-time collaborations over advanced cyber infrastructure

• Commercial Applications
  – Playing games over the Internet
  – Teleconferencing, web-casting, etc.
  – Multicast demands increasing over time

• Today’s scene
  – No one will buy a router that does not support multicasting
  – Electronic routers can support small fraction of traffic that is multicast, but performance suffers greatly for large amounts of multicasting
Multicast in Electronic Routers

• Electronic routers support multicast by
  – Making multiple copies of input frames
  – Using multiple transmissions to send to each receiver

• For heavily multicast environment
  – Making copies strains limited memory bandwidth resources
  – Multiple transmissions reduce throughput efficiency

• In heavily multicast or broadcast environment, electronic routers become extremely inefficient!
Optics in Switching Core

• Use the natural parallelism of optics in a switching core to implement high-performance routers that are efficient for heavily multicast/broadcast applications

• Broadcast & Select optical architectures naturally provide multicast and broadcast capabilities

• Discussion here is in routers, but same advantages also apply to CATV networks or distribution networks and metro or access networks
Advantages of Optics for Multicast/Broadcast Applications

• Broadcast & select architectures provide multicast and broadcast capabilities without requiring retransmissions

• Optical backplanes connecting line cards (LC’s) or transponders can be very high speed
  – 10Gb/s, 40Gb/s or 160Gb/s relatively easy in optics
  – Electronic backplanes are ~5Gb/s over limited distances

• For multi-rack systems (typically > 16 LC’s) or spatially distributed systems, optical interconnects and switching fabrics are critical
  – Today’s multi-rack systems use optical fiber interconnects

• As router capacity increases, optics provides increasing potential for lower power, less heat dissipation, and possibly smaller size and lower cost
Single Rack Systems
(“3rd generation routers”)

Switched Backplane

Typically <160Gb/s aggregate capacity
Limitation: number of LC’s in a rack

(Nick McKeown, Stanford University)
Multi-rack Routers
(“4th generation routers”)

Optics inside a router for the first time

Switch Core
Optical links
100s of metres

Linecards

0.3 - 10Tb/s routers in development
(Nick McKeown, Stanford University)
5th Generation Routers/Switches
Switching Fabric and Fiber Interconnection Combined

Star Coupler

Fiber Interconnect

No Intermediate O/E/O for Interconnect

Linecards
Aggregation into Superframes for Backbone Network

- Traffic on backbone network requires highest capacity, but is highly aggregated
  - On Internet2, the superframe size is 9KBytes
  - At OC-192, this corresponds to a frame size of 7 micro-sec!!!

- Consider two designs of routers
  1. Highly Aggregated Backbone Routers
     -> USE OPTICAL SWITCHING AND INTERCONNECT OR BACKPLANE
     -> TUNABLE FILTERS PERFORM SELECTION AT OUTPUT
  2. Arbitrary Frame Size Routers
     -> USE HYBRID SWITCH WITH OPTICAL INTERCONNECT OR BACKPLANE AND ELECTRICAL SWITCHING
     -> PASSIVE OPTICAL INTERCONNECT OR BACKPLANE, ELECTRONIC SWITCHING BETWEEN DETECTORS AND LC’s
Broadcast & Select Switching Fabric

Tuning Schedule

<table>
<thead>
<tr>
<th>Line Card</th>
<th>Line Card</th>
<th>Line Card</th>
<th>Tunable Optical Filter (TOF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOF 1 Tune</td>
<td>TOF 1 Tx</td>
<td>TOF 1 Tune</td>
<td>TOF 1 Tx</td>
</tr>
<tr>
<td>TOF 2 Tune</td>
<td>TOF 2 Tx</td>
<td>TOF 2 Tune</td>
<td>TOF 2 Tx</td>
</tr>
</tbody>
</table>
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
- Also: tunable lasers much more difficult for multicast!

[ 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 34dB (~22dB loss budget and another ~8dB for OADM/DC, ~3-4dB EOL)
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

- Microsecond devices cost $<<$ nanosecond 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!
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

- Input fibers
- Power splitter
- Power combiner
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

P₀

L

(D)

Pulse
Chirping
(2m SMF-28)

L: only 2 meters long
Exemplary SC Spectrum

Experimental Parameters

\[ \begin{align*}
L &= 2 \text{ [m]} \\
\Box_0 &= 1539 \text{ [nm]} \\
P_{\text{avg}} &= 1100 \text{ [W]} \\
D &= 1.13 \text{ [ps/nm-km]} 
\end{align*} \]

-20 dB Bandwidth: 211 nm
Sec Filters using 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 $\lambda/4$
  – 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>
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>25 GHz</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>

---

**Diagram:**

- **Transmission (%)** vs. **Wavelength (nm)** for different filter types and shapes.
- **Transmission (dB)** vs. **Wavelength (nm)** for different channel counts and cavity numbers.
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.
Switching of Arbitrary Sized Frames

- Hybrid switch uses best of optics & electronics
- Optical advantages:
  - Simple distribution for multicast and broadcast
  - High-speed backplanes
  - Interconnect multiple racks of LC’s
- Electronics good for switching
  - High speed is possible with sub-nsec switching times
  - Does not scale well for very large number of LC’s
  - Localized switching is better than distributed switching
- To combine both:
  - Use WDM optics from input to output detectors
  - Use electronic switching between detectors and receiving LC’s
Hybrid Switching with Star Coupler Core

- Transmitters
- LC1, LC2, LCn
- m*n inputs
- Rack#1
- LC1', LC2', LCn'
- Rack#m
- m*n detectors
- m racks
- (m*n)*n fabric
- Scheduler (Controller)
- n LC’s per rack
- m racks
Exemplary Electronic Switching Fabric

- **Optional Electronic Amplifiers**
- **1:n electronic power splitter**
- **Electronic MUX**
- **Rack #i**
  - **LC1**
  - **LC2**
  - **LCn**

- **m:n Detectors**
- **m:n power splitters**
- **n multiplexers**
- **(m:n)n Electronic switching fabric**

- **n line cards per rack**
Optical Amps & Segmentation by Racks

Transmitters

Rack #1
LC1
LC2
LCn

Rack #m
LC1'
LC2'
LCn'

Power Splitter

Optional In-line Optical Amplifier

Scheduler (Controller)

Array of detectors

Optional Optical Pre-Amplifier

Electronic Switching Fabric

Optical Amps & Segmentation by Racks

Optional Booster Optical Amplifier
Advantages of Hybrid Switching System

- Heavily multicast/broadcast easily handled
- Passive optical interconnect distributes all LC inputs to each output
- Electronic switching fabric has better scaling
  - Optical distribution sends to m electronic switches
  - Each electronic switch is of order (mxn)xn
  - Cross-bar would grow as (mxn)^2
- Switching speed determined by electronic fabric
  - Needs only to run at packet or frame rate…nsecs or longer
- Can reduce required hardware (demux, detectors, electronic switches) by using fast optical filters
Summary

- Heavily multicast routers may be required for war gaming type scenarios, commercial applications anticipated
- Optics beneficial in multicast routers because of natural parallelism of B&S Architecture
  - Optics distributes signals for multicasting
  - High-speed backplane or interconnection between racks
- For traffic aggregated into superframes, use optical switching with B&S fabric and tunable filters
- For arbitrary sized frames, hybrid switching used
  - Optical distribution using B&S, optionally with amplifiers
  - Electronic switching between detectors and line cards
  - Passive optics partitions electronic switch between multiple racks