Technologies, Architecture and Services for the Next-Generation Core Optical Networks

Adel A. M. Saleh

DARPA
Strategic Technology Office
Arlington, Virginia, USA
Adel.Saleh@darpa.mil
Next-Generation, Global-Scale Core Optical Network

**Vision**: Increased Throughput, Robustness and Security with Reduced Latency, Capital Cost and Operational Cost

Undersea Cables
Current State of IP/WDM Network Node Architecture

Clients

IP Router

WDM Tx/Rx Cards

Direct OEO Connection

Optical Bypass and Ultra-Long-Reach Optics

Clients

IP Router

WDM Tx/Rx Cards

All-Optical λ Switch or OADM

25 March 2007 - Saleh

Approved for Public Release, Distribution Unlimited

OFC 2007 Workshop FON - 3
Current State of IP/WDM Network Node Architecture

- Direct OEO Connection
- Optical Bypass and Ultra-Long-Reach Optics

Clients

IP Router

WDM Tx/Rx Cards

Optical Bypass in the Core

25 March 2007 - Saleh

Approved for Public Release, Distribution Unlimited

OFC 2007 Workshop FON - 4
Possible Evolution of IP/WDM Network Node Architecture

From: Saleh and Simmons, “Evolution Towards the Next-Generation Core Optical Network”, JLT, Sept 2006

25 March 2007 - Saleh
Approved for Public Release, Distribution Unlimited

OFC 2007 Workshop FON - 5
Possible Evolution of IP/WDM Network Node Architecture

- Direct Connection
- WDM Tx/Rx Cards
- IP Router
- Optical Bypass in the Core
- All-Optical Switch
- Optical Bypass and Ultra-Long-Reach Optics
- Ethernet, OBS and other edge aggregation technologies can play a strong role here

From: Saleh and Simmons, “Evolution Toward the Next-Generation Core Optical Network”, JLT, Sept 2006
Meanwhile, everyone is saying that **Ethernet** is coming to the Core!

But, …
But what is the vision for the Ethernet architecture in the Network Core, and how does it relate to IP and WDM?

Is it going to be **IP** beside Ethernet, both over **WDM**? **OR** Is it going to be **IP over Ethernet** over **WDM**?

- **IP Router**
- **Ethernet Switch**
- **WDM Tx/Rx Cards**
- **All-Optical \( \lambda \) Switch or OADM**

---

**OR**

- **IP Router**
- **Ethernet Switch**
- **WDM Tx/Rx Cards**
- **All-Optical \( \lambda \) Switch or OADM**

---

25 March 2007 - Saleh

Approved for Public Release, Distribution Unlimited

OFC 2007 Workshop FON - 8
But what is the vision for the Ethernet architecture in the Network Core, and how does it relate to IP and WDM? Is it going to be a combination of the two architectures?

These are Interesting Questions!
### Optical Spectrum

<table>
<thead>
<tr>
<th>Mod Fmt</th>
<th>Gb/s per λ</th>
<th>Δλ GHz</th>
<th>SE b/s/Hz</th>
<th>Num λ's C-Band</th>
<th>Tb/s in C-Band</th>
<th>Tb/s in C,L-Band</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>100</td>
<td>0.025</td>
<td>40</td>
<td>0.25</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>0.1</td>
<td>40</td>
<td>0.4</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>0.2</td>
<td>80</td>
<td>0.8</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>0.4</td>
<td>160</td>
<td>1.6</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>100</td>
<td>0.4</td>
<td>40</td>
<td>1.6</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>50</td>
<td>0.8</td>
<td>80</td>
<td>3.2</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>25</td>
<td>1.6</td>
<td>160</td>
<td>6.4</td>
<td>12.8</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>2.0</td>
<td>80</td>
<td>8.0</td>
<td>16.0</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>4.0</td>
<td>160</td>
<td>16.0</td>
<td>32.0</td>
<td></td>
</tr>
</tbody>
</table>

**High-Spectral-Efficiency (SE) Fiber Transmission**

**Progress to Today**

- Binary OOK
- 10 Gb/s per λ, 100 GHz, 0.1 SE b/s/Hz, 40 λ's, 0.8 Tb/s in C-Band, 1.6 Tb/s in C,L-Band

**Today's Plans**

- Binary OOK
- 10 Gb/s per λ, 25 GHz, 0.4 SE b/s/Hz, 160 λ's, 1.6 Tb/s, 3.2 Tb in C,L-Band

**Near-Term Vision**

- Duo-QPSK
- 40 Gb/s per λ, 50 GHz, 0.8 SE b/s/Hz, 80 λ's, 3.2 Tb/s, 6.4 Tb in C,L-Band

**Future Vision**

- 100 Gb/s per λ, 50 GHz, 2.0 SE b/s/Hz, 80 λ's, 8.0 Tb/s, 16.0 Tb in C,L-Band

- 100 Gb/s per λ, 25 GHz, 4.0 SE b/s/Hz, 160 λ's, 16.0 Tb/s, 32.0 Tb in C,L-Band
Future Needs in Fiber-Optic Transmission

- Spectral Efficiency higher than 1.5 b/s/Hz
- 100 Gb/s (or higher?) per wavelength
- Regeneration-free optical reach greater than 1500 km
- Security “encryption” technique that is Low in Cost and in Size, Weight and Power (SWaP)
- Encryption may require a secure system to manage the encryption keys, e.g., something like quantum key distribution
Network Cost as a function of the Line Rate for various Aggregate Network Demands

From: Saleh and Simmons, “Evolution Toward the Next-Generation Core Optical Network”, JLT, Sept 2006
### Optical-Reach & Transponder-Cost Assumptions for the Network Study

<table>
<thead>
<tr>
<th>Line Rate</th>
<th>Regeneration-Free Optical Reach</th>
<th>Relative Cost of Transponder</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 Gbps</td>
<td>4,000 km</td>
<td>1</td>
</tr>
<tr>
<td>10 Gbps</td>
<td>3,000 km</td>
<td>2 x 1 = 2</td>
</tr>
<tr>
<td>40 Gbps</td>
<td>2,000 km</td>
<td>2.5 x 2 = 5</td>
</tr>
<tr>
<td>160 Gbps</td>
<td>1,500 km</td>
<td>2.5 x 5 = 12.5</td>
</tr>
<tr>
<td>640 Gbps</td>
<td>1,000 km</td>
<td>2.5 x 12.5 ≈ 31</td>
</tr>
</tbody>
</table>

*From: Saleh and Simmons, “Evolution Toward the Next-Generation Core Optical Network”, *JLT, Sept 2006*
A Grid-Computing Application, and its Requirement for Very Fast Connection Set-Up
### Where the top 500 supercomputers reside (June 2006)

<table>
<thead>
<tr>
<th>Country</th>
<th>Number</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>299</td>
<td>59.8%</td>
</tr>
<tr>
<td>U.K.</td>
<td>35</td>
<td>7.0%</td>
</tr>
<tr>
<td>Japan</td>
<td>29</td>
<td>5.8%</td>
</tr>
<tr>
<td>China</td>
<td>28</td>
<td>5.6%</td>
</tr>
<tr>
<td>Germany</td>
<td>18</td>
<td>3.6%</td>
</tr>
<tr>
<td>India</td>
<td>11</td>
<td>2.2%</td>
</tr>
<tr>
<td>Israel</td>
<td>9</td>
<td>1.8%</td>
</tr>
<tr>
<td>Australia</td>
<td>9</td>
<td>1.8%</td>
</tr>
<tr>
<td>France</td>
<td>8</td>
<td>1.6%</td>
</tr>
<tr>
<td>Canada</td>
<td>8</td>
<td>1.6%</td>
</tr>
<tr>
<td>Italy</td>
<td>7</td>
<td>1.4%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>New Zealand</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>South Korea</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>Brazil</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>Taiwan</td>
<td>3</td>
<td>0.6%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>3</td>
<td>0.6%</td>
</tr>
<tr>
<td>Singapore</td>
<td>2</td>
<td>0.4%</td>
</tr>
<tr>
<td>Ireland</td>
<td>2</td>
<td>0.4%</td>
</tr>
<tr>
<td>Belgium</td>
<td>2</td>
<td>0.4%</td>
</tr>
<tr>
<td>Sweden</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Spain</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>South Africa</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Russia</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Denmark</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Belarus</td>
<td>1</td>
<td>0.2%</td>
</tr>
</tbody>
</table>


Slide prepared by: B. Pierce (DARPA); W. Bengtson & E. Wolfer (BAH/DARPA Ctr.)

Approved for Public Release, Distribution Unlimited
Assume $N$ (say 100) super computers are connected to a national or global network to perform independent computing tasks by forming cooperative computational groups …
Let $M$ (say 10, which is $<< N$) be the average number of computers in a computation group.

$time = 0$
The computational groups change after each computation duration, $T$ (say minutes or more)

\[ \text{time} = T \]
A Grid Computing Scenario (4)

Let the bitrate per connection be $R$, the average burst duration be $\tau$ (milliseconds to seconds), the average number of bursts per computation be $\nu$, and the corresponding average burst communications duty ratio be $\rho = \nu \times \tau / T$ ($\rho << 1$)

\[
time = 2 \ T
\]
Consider a network with \( N \) distributed-computing nodes

Assume that the computing nodes need to be connected to each other when and if they have to exchange bursts of data

To speed-up computations, the bitrate, \( R \), per connection is assumed to be equal to the optical line rate, e.g., 40 Gbit/s

Not all \( N \) nodes need to be connected to each other to perform a single computation; thus assume that, on the average, \( N/M \) groups each with \( M \) nodes (where \( M << N \)) need to be formed, where communication is needed only within a group. *In general, the groupings change for each computation*

The duration, \( T \), of a computational task can be large, e.g., minutes

The lengths of the individual exchanged data bursts, \( \tau \), can be small, e.g., milliseconds to seconds

Let the average number of bursts per computation be \( \nu \)

Note that the average burst-communications duty ratio, \( \rho = \nu \times \tau / T \), can be considerably less than unity (\( \rho << 1 \))
Networks with Different Configurability Speeds ($S$) with Some Average Computation Duration ($T$), Average Computation Burst Duration ($\tau$), and Average Burst Activity Duty Ratio ($\rho$),

1) **Static Network**: (i.e., $S >> T$)
   In this case, all $N$ computing nodes need to be connected to each other (a total of $N(N-1)/2$ connections), all of the time
   Total Network Bandwidth Requirement:
   \[ BW1 = R \times N (N-1)/2 \]

2) **Slowly Configurable Network**: (i.e., $T >> S >> \tau$)
   In this case, one can slowly form $N/M$, fully connected $M$-computing-nodes groups on a computation-by-computation basis
   Total Network Bandwidth Requirement:
   \[ BW2 = R \times N (M-1)/2 = (M-1)/(N-1) \times BW1 \ (<< BW1) \]

3) **Rapidly-Configurable Network**: (i.e., $S << \tau$)
   In this case, one can quickly form any connection as needed on a burst-by-burst basis
   Total Network Bandwidth Requirement:
   \[ BW3 = R \times \rho \times N (M-1)/2 = \rho \times BW2 \ (<< BW2) \]
Assume the following network switching speeds:

- \( S = \text{Several seconds for Slow Optical Circuit Switching (Slow OCS)} \)
- \( S = 50 \text{ milliseconds for Fast Optical Circuit Switching (Fast OCS)} \)
- \( S = 50 \text{ microseconds for Optical Burst Switching (OBS)} \)

Also assume the following parameters:

- Total number of computational nodes, \( N = 100 \)
- Number of nodes in each computation group, \( M = 10 \)
- Bitrate per connection, \( R = 40 \text{ Gbit/s} \)
- Average computation duration, \( T = \text{minutes or more} \)
- Burst durations, \( \tau \), varies from 1 second down to 1 millisecond
- Burst-communications duty cycle, \( \rho = 10\% \)

### Approximate Total Network Bandwidth Requirement (Tbit/s)

<table>
<thead>
<tr>
<th>Switching</th>
<th>( \tau = 1 \text{ sec} )</th>
<th>( \tau = 100 \text{ msec} )</th>
<th>( \tau = 10 \text{ msec} )</th>
<th>( \tau = 1 \text{ msec} )</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>Required BW too high</td>
</tr>
<tr>
<td>Slow OCS</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>10% of BW of Static</td>
</tr>
<tr>
<td>Fast OCS</td>
<td>2</td>
<td>4</td>
<td>20</td>
<td>20</td>
<td>1% to 10% BW of Static</td>
</tr>
<tr>
<td>OBS</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1% of BW of Static</td>
</tr>
</tbody>
</table>

25 March 2007 - Saleh

Approved for Public Release, Distribution Unlimited

OFC 2007 Workshop FON - 22
Last August, DARPA issued the CORONET BAA (#06-29), entitled “Dynamic Multi-Terabit Core Optical Networks: Architecture, Protocols, Control and Management”
See: http://www.darpa.mil/sto/solicitations/CORONET/

This BAA is for Phase 1, which is for developing and testing (by simulations) architecture, protocols, and algorithms (18 month)

Phase 2 will be for developing and testing a compatible network control and management software suitable for transition to Government and commercial telecommunications carriers

Thus, the CORONET Program thrust areas are
- Architecture (of the Network, Network Node and Network Element)
- Protocols and Algorithms (for very-fast service set-up and restoration)
- Network Control and Management
- But No Hardware Development or Testing

Proposals for Phase 1 are currently being evaluated for selection
Overview of the CORONET Target Network

- Global core optical network
- IP over WDM architecture
- Network services
  - Predominantly IP services (with differentiated QoS)
  - Substantial amount of $\lambda$-services
- Scalable for up to 10x increase in aggregate network demand over today’s state-of-the-art networks
- Highly dynamic network with very fast service set-up and tear-down
- Resilient to multiple concurrent network failures
- Simplified network operation and increased security
## High-Level CORONET Program Goals

<table>
<thead>
<tr>
<th>Network Requirements</th>
<th>Today’s State-of-the-Art Networks</th>
<th>High-Level CORONET Program Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Network Demand</td>
<td>Up to 10 Tb/s</td>
<td>Up to 100 Tb/s</td>
</tr>
<tr>
<td>Maximum Fiber Capacity</td>
<td>Up to 1.6 Tb/s</td>
<td>Up to 20 Tb/s per Fiber</td>
</tr>
<tr>
<td>Bit Rate per Wavelength</td>
<td>10 or 40 Gb/s</td>
<td>40 or 100 Gb/s</td>
</tr>
<tr>
<td>Maximum Bit Rate per Stream</td>
<td>40 Gb/s</td>
<td>Up to 1 Tb/s</td>
</tr>
<tr>
<td>End-to-End Network Services</td>
<td>IP, SONET</td>
<td>IP (75%±) and λ-Services (25%±)</td>
</tr>
<tr>
<td>Optical-Layer Multicasting</td>
<td>Typically Not Possible</td>
<td>Basic Requirement</td>
</tr>
<tr>
<td>Performance Monitoring</td>
<td>Mostly in Electrical Layer</td>
<td>In Electrical and Optical Layers</td>
</tr>
<tr>
<td>Optical-Layer Configurability</td>
<td>Slow, Often Manual</td>
<td>Fast, Fully Automatic</td>
</tr>
<tr>
<td>Max Speed of Service Setup</td>
<td>Hours to Weeks</td>
<td>≤ 100 msec (CONUS)</td>
</tr>
<tr>
<td>Speed of Shared Restoration*</td>
<td>50 to 100’s msec (Ring)</td>
<td>≤ 250 msec (Global)</td>
</tr>
<tr>
<td></td>
<td>Sec’s to Min’s (Mesh)</td>
<td>≤ 100 msec (CONUS)</td>
</tr>
<tr>
<td></td>
<td>* From How Many Failures:</td>
<td>≤ 250 msec (Global)</td>
</tr>
<tr>
<td>IP Services with Differentiated End-to-End QoS</td>
<td>Very Limited</td>
<td>Basic Requirement</td>
</tr>
<tr>
<td></td>
<td>Typically, One Failure</td>
<td>Up to Three Failures</td>
</tr>
</tbody>
</table>
The Main CORONET Challenge

High Network Efficiency
(i.e., Low Cost, Size, Weight and Power)
The Main CORONET Challenge

Very Fast Configurability

High Network Efficiency
(i.e., Low Cost, Size, Weight and Power)
The Main CORONET Challenge

Very Fast Configurability

Not So High Network Efficiency?
(i.e., Not So Low Cost, Size, Weight and Power)
The Main CORONET Challenge

- Very Fast Configurability
- All-Optical Bypass

Not So High Network Efficiency?
(i.e., Not So Low Cost, Size, Weight and Power)
The Main CORONET Challenge

- **Very Fast Configurability**
- **All-Optical Bypass**

**Low Network Efficiency?**
(i.e., High Cost, Size, Weight and Power)
The Main CORONET Challenge

- Very Fast Configurability
- All-Optical Bypass

High Network Efficiency
(i.e., Low Cost, Size, Weight and Power)

The CORONET Program Asks: How to Push Back?