Photonics technologies for low-energy networks at AIST

Shu Namiki
National Institute of Advanced Industrial Science and Technology (AIST)
Network Photonics Research Center

ISUPT 2013, Oct. 21, 2013
At University of Rochester

Part of this work is supported by
- Special Coordination Funds for Promoting Science and Technology of Ministry of Education, Culture, Sports, Science and Technology (MEXT).
- New Energy and Industrial Technology Development Organization (NEDO)
Team

- **Network Architecture**
  - Kiyo Ishii
  - Junya Kurumida
  - Tomohiro Kudoh

- **Silicon Photonics and Devices**
  - Hitoshi Kawashima
  - Keijiro Suzuki
  - G. W. Cong
  - Ken Tanizawa
  - Sang-Hun Kim
  - Haruhiko Kuwatsuka
  - Hiroyuki Matsuura
  - Aaron Albores-Mejia

- **Transmission and Node Subsystems**
  - Takayuki Kurosu
  - Takashi Inoue
  - Hung Nguyen Tan
  - Karen Solis-Trapala
  - Mingyi Gao
Outline

• Introduction:
  – Scalability, and now is endangered!

• Opportunity for optical switches
  – Distinct energy scaling and bandwidth
  – AIST project called “VICTORIES”
  – Development of 32 x 32 silicon switches

• Dynamic Optical Path Network

• Inter-Data Center Network

• Summary
Network traffic keeps on increasing

But, how sustainable is it for the future?

Source: Fortune 500

After T.J. Xia (Verizon) at WIN 2012, Inuyama
Data-centricity of Traffic

And, most of the contents are video.

How have we scaled the transport capacity?

• 1960s
  – Copper wire for telephony

• 1970s
  – Use optical fibers, LDs, PDs

• 1980s
  – Use ETDM (Sonet/SDH)

• 1990s
  – Use EDFAs, Internet revolution!!!

• 2000s
  – Use WDM, Raman amplifiers, FEC, the Bubble!!!

• 2010s
  – Use advanced modulation formats, digital coherent technology

• 2020s
  – Use Space and/or Mode-Division-Multiplexing

• We’ve been so acclimated to scalable technology, but…
What we are learning from HPC

- Because chips are no longer as scalable, Moore’s law may finally be ending unless the network scales.

Can data centers scale out?

- The role of network gets more important.

Figure 8: Future “optimized network era” will follow current “East-West era”

Data Center Cost/Performance

North-South Era
- Moore’s law driven
  - Chip cost/perf decreasing
  - High single proc performance
  - Moore’s law driven scaling

East-West Era
- Moore’s law slowing
  - Compute cost/perf driven by chip + scale out
  - TCO: HW % of cost decreasing →
  - Utilization increasing w/modest intra-DC BW growth (virtualization)
  - SW model for reliability, failover, redundancy, distributed storage, SDN

Optimized Network Era
- Moore’s law slowing & BW scaling continues
  - Utilization → diminishing returns
  - Network Virtualization, Multipath networks
  - Aggregate HW cost/power becoming significant
  - Distributed switch architectures
  - Disaggregation
  - Mega DC economies of scale
  - Inter-DC optimization
  - New optics enabled networks?

Source: Marc Taubenblatt, IBM

OIDA Workshop Report April 2013
Fundamental limits being faced!

- **Quantum Limit**
  - Distributed Raman amplifiers with 3.1 dB noise figure
  - E-FEC allows error free transmission with only several photons per bit at receiver

- **Rayleigh Scattering Limit of Fiber**
  - ~0.15 dB/km

- **Shannon Limit**
  - A few dB away from E-FEC technologies
  - Non-linear Shannon limit is about 100 Tbps/fiber

- **Electrical Bandwidth Limits = The Limit of Parallelism**
  - Radiation loss
  - Thermal jitter
  - Bandwidth density = energy consumption / bit

- **TCP/IP Limits**
  - Increasing inefficiency for larger Delay-Bandwidth products
  - QoS limitations
  - IP routers consumes energy proportional to the amount of traffic
Energy issue will eventually take over everything

- Green of ICT, then Green by ICT

- Disaster proof
  - Communications is the ‘Lifeline’.
  - Energy supply will be limited when disasters occur.

- The current technologies can’t scale to the increasing traffic in future.
- 3-4 digit energy saving is necessary, which means we need a new paradigm.
Almost all of the traffic will be due to video transfer.

Traffic Growth Projected by CISCO

Source: Cisco Visual Networking Index: Forecast and Methodology, 2008-2013
Packet switching versus circuit switching

IP: Process every packet for routing

Which suits better for video?

Optical Path: Set up end-to-end paths
# Switches: photonics vs. electronics

<table>
<thead>
<tr>
<th>Type of switching</th>
<th>Optical switches</th>
<th>Electrical Switches</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Circuit switching (requires a control plane)</td>
<td>Packet switching</td>
</tr>
<tr>
<td>Granularity</td>
<td>Coarse and inflexible</td>
<td>Fine and flexible</td>
</tr>
<tr>
<td>Bandwidth scaling</td>
<td>Transparent</td>
<td>Transceivers and LSI</td>
</tr>
<tr>
<td>Physical I/F</td>
<td>Optical connectors</td>
<td>Transceivers</td>
</tr>
<tr>
<td>Energy scaling</td>
<td>&lt; W/port (incl. amp)</td>
<td>0.3 nJ/bit (Infiniband) 6 nJ/bit (IP router)</td>
</tr>
<tr>
<td>Energy consumption for 1,000 ports x 1Tbps</td>
<td>&lt; 1 kW</td>
<td>300 kW (Infiniband) 6,000 kW (IP router)</td>
</tr>
<tr>
<td>Scalability issue</td>
<td>Number of ports</td>
<td>Moore’s law</td>
</tr>
<tr>
<td>QoS</td>
<td>Physically guaranteed</td>
<td>Depends</td>
</tr>
<tr>
<td>Suitable apps</td>
<td>Video</td>
<td>Packetized data</td>
</tr>
</tbody>
</table>

Optical switch can save energy **by several orders of magnitude!!**

The hybrid use of packet and optical switching will be the key.
Difference at the network level..

- Conventionally exchange digital packets: Aggregate and Transport

  - 1-10 GbE
  - 10-100 GbE
  - Statistical Multiplexing
  - A few lines between Routers
  - Many CRS-3 required!!

→ A New Scheme: Exchange end-to-end numerous fibers in parallel!!

- N users
- 10-100 GbE
- m paths
- N x m Matrix Switch
- Optical switches: Transparent and Green
- Transmission lines in parallel space
- SDM rather than TDM
- Full mesh connections rather than Aggregate and transport

S. Namiki et al., ECOC 2010, Mo.2.A.4.
S. Namiki et al., JSTQE, 17, 446 (2011).
It calls for new network technologies

- Energy and Cost Efficiency
- Agility Resilience

L3 (IP) Routing

L2 (Ethernet) Switching

Dynamic Optical Path Network (DOPN)

Good Control Plane at upper layers & Good Optical Devices
Toward Optical Cut-Through of Everything (OCE)

• Today
Toward Optical Cut-Through of Everything (OCE)

- Tomorrow

![Diagram of optical cut-through network](image)

Optical Cut-through of Core Routers

OLT

Splitter

ONU
Toward Optical Cut-Through of Everything (OCE)

• Ideally Energy-Saving DOPN

Global Control Plane and Resource Management are necessary.
Jump from Optical Cut-through to DOPN

Model for Dynamic Optical Path Network (DOPN)

Calculation by K. Ishii (AIST),
Under the VICTOIRES project.
Details in Photonics West 2013.

Extrapolation of the present model

The energy reduction by optical cut-through is not large enough.
The energy consumption at edge routers is dominant.
Some initiatives have already started

- GreenTouch Consortium
- VICTORIES

http://www.greentouch.org/

IEC’s Megatrend: Energy Sustainability of optical communication networks and their driver: Dynamic Optical Path Network

http://www.aist-victories.org/

http://www.aist.go.jp
Digital Optical Path Processor using Silicon Photonics

• 32x32 Matrix Switch
  – Energy: Need only a fewWs
  – Footprint: Within 2 cm x 1 cm Chip

• Switching time
  – ~μsec by thermo-optical switching
  – ~nsec by plasma-effect switching

• Integration with control circuits for black-box operations

Courtesy of Y. Shoji, K. Kintaka, et al.
Why 32 x 32 matrix switches?

- 32 x 32 SWs make a 512 x 512 Clos SW.
- 100,000 CPU nodes can be accommodated by 675 450x450 (simplex) switches. A 450x450 SW consists of 96 32x32 SWs, leading to a total of 64,800 32x32 SWs.
- Each port can carry a bandwidth > Tbps.

2-stage Fat tree Switch Fabric for 100,000 nodes

Total: 101250 CPU nodes
Roadmap of Si-photonics matrix switch at AIST

Mach-Zehnder Interferometer based TO switching

2010
2 x 2

2011
4 x 4

2012
PILOSS 4 x 4

2013
PILOSS 8 x 8

2014
32 x 32

Larger scale
~20 mm
Fabricated 8×8 PILOSS Switch

K. Suzuki et al., ECOC 2013, PDP

- 64 Mach-Zehnder Switches
- 49 Intersections
- TO phase-shifter
- Designed for TM-mode
- Footprint: 2.4 × 3.5 mm²

- Electrical Connection:
  - Wire Bonding
- Optical Connection:
  - Butt-Coupled Fiber and Spot-Size Convertor
  (Coupling Efficiency: 3.6 dB)

Switch Elements

Loss

- Input-Output Connection:
  - 1-1, 1-2, ..., 8-7, 8-8
  - Path ID: 1, 2, 63, 64
- PILOSS Switch Topology
  - Small Loss Variation
- Breakdown:
  - MZ Switch: 0.6 dB
  - Intersection: 0.2 dB
- Low losses enabled by the use of the directional coupler and intersection
2 × 2 Polarization diversity

Kim et al., in ECOC2013

Orthogonal SOP Crossing (JP 2011-158030 )

S. Namiki, ISUPT2013, Oct. 27, Univ. of Rochester

http://www.aist.go.jp
Si-ph. SW integrated with CMOS driver circuits

- OXC: 2x2 TO switch with MOSFET
  - G. W. Cong et al., OE March 2013 (AIST)

- Interconnect SW: 4x4/8x8 p-i-n switch with CMOS driver
  - B.G. Lee et al., OFC NFOEC PDP 5C3, 2013 (IBM)
Realizing 32 x 32 chip scale

Drawing time using Spot-Beam EB Lithography System (JEOL6000FS/E (50kV))

- 32 x 32
  - 20 mm x 10 mm
  - 128 hrs. (~5 days) by EB

- 8 x 8
  - 11 mm x 7 mm
  - 8 hrs.

- 4 x 4
  - 6 mm x 3 mm
  - 2 hrs.

EB → DUV photolithography

ArF Immersion Lithography facility
Super Clean Room (SCR) 12 inch line @AIST
Hierarchical multi-granularity in routing

- Various path granularities
  - From fine to coarse
  - node that can handle such

- Network topology based on hierarchical routing
  - Low energy consumption
  - Scalability to national scale

K. Ishii et al., Photonics West 2013

S. Namiki, ISUPT2013, Oct. 27, Univ. of Rochester
CDC-ROADM using Silicon Photonics (NEC+AIST)

T. Hino et al., ECOC2012 Invited: Tu.3.A.5
Feature of ODU XC (Fujitsu+AIST)

- ODU XC (Cross Connect) is a digital cross connect for ODU (Optical Data Unit) that defined by ITU-T G.709.
- ODU XC has capability of switching traffic with 1.25Gbps granularity. It is very efficient to switch 1Gbps-40Gbps stream.
- ODU XC has 1/3 power consumption and 1/100 latency compared to core router.
- AIST and Fujitsu are developing the ODU XC chip and this chip will be available in 2013.

OTN is effective and efficient for operating >1 GbE. Fujitsu plans to incorporate the OTN switch fabric into Fujitsu products.

Slide courtesy of H. Honma and H. Onaka, Fujitsu
32 x 32 SWs make up a nationwide Dynamic Optical Path Network

Replace IP routers by optical switches!!

- Hierarchical multi-granular node
- Sub-λ and wavelength cross-connect node
- Multi-granular aggregating node
- Content server
- Sub-λ aggregating node
- Wavelength path terminal
- Sub-λ path terminal

Regular mesh in Core; Full Mesh in Metro; Complete bipartite graph in Access
Flex-Grid compatible, ODU-XC for sub-λ granularity

K. Ishii, et al. (AIST), Photonics West 2013

S. Namiki, ISUPT2013, Oct. 27, Univ. of Rochester

http://www.aist.go.jp

Victories
Inter-datacenter network: challenges

- Chunk of continually generated data
- Ultra-coarse granularity
- Low latency data transfer
- High capacity and efficiency
- Strict QoS with scheduled connections

- Seamless connections to intra-datacenter network
- Resilience for disaster
Multi-Tbps Elastic Transmitter: WDM vs. OTDM

- **Super-channel Nyquist/OFDM WDM**

- **Nyquist-OTDM WDM**
Optical comb generator (OCG) for Tbps elastic Tx/Rx

- High spectral efficiency through superchannel (OFDM or Nyquist)
- High quality optical comb/ultra-short pulse through highly nonlinear fibers
  - V. Ataite et al. (UCSD), OFC NFOEC 2013, PDP5C.1
  - C+L-band
    - 40 dB OSNR
  - < 400 fsec
    - Tunable

- R&D on Tbps elastic transceivers using OCG is highly expected.
Research on Nyquist OTDM-WDM signals

Features of Nyquist OTDM
- High spectral efficiency up to 1 symbol/s/Hz
- Ultrahigh yet flexible baudrates (up to 1T)
- Generation in optics, high energy efficiency
- Eliminating guard-bands in WDM system
- Better transmission perf. than OTDM
  - K. Harako et al., OFC 2013, JW2A.38.
  - H. Hu et. al., CLEO 2013, CTh5D.5.

Nyquist OTDM-WDM in network
- Granularity will grow to be more than ~100G in the future big-data centric era
- Especially, traffic between data-centers
- Highly efficient, ultra-coarse yet flexible granular inter-datacenter network based on all-optical elastic network

→ Nyquist OTDM-WDM signals
Nyquist filtering using Waveshaper

- Ultra-coarse granular Nyquist signals (>40G) can be generated by Waveshaper.
- Waveshaper-based WSS can efficiently add-drop WDM channels of Nyquist OTDMs with almost no guard-band.
- Pass-drop over cascaded WSSs: *H. Nguyen Tan et al., OFC 2013, JW2A.50*
  → This paper: transmission and pass-drop operations in full elasticity.
Experimental setup

- Baudrate/pol.: 43Gbaud, 86Gbaud, 172Gbaud, and 344Gbaud
- Transmission 4 spans of 80km SLA+IDF and WSS nodes
- WSSs elastically pass-drop WDM channels of Nyquist OTDMs
H. Nguyen Tan et al., OECC 2013, PDP / ECOC We.1.C.5
Details to appear in OE.

- Simple generation of different baudrate signals by Waveshaper
- WSS efficiently add-drop WDM channels of Nyquist OTDMs with almost no guard-band
BER performance of Nyquist OTDMs
Oct. 2014 in Tsukuba, Japan
Co-located with a 3-day international symposium
Demonstrate 4k/8k video apps over DOPN
Summary

- The energy consumption of the IP network is becoming a serious bottleneck, calling for a paradigm shift.
- For disaster survivability, resilience under limited energy supply is of critical importance.
- Dynamic Optical Path Network may be the only option in the sense of both energy, capacity, and suitability for video services, thus has been recognized by IEC as a driver of this megatrend.
- In order to realize this, some R&D initiatives have been put into action.
- We have reviewed some of the activities under VICTORIES project at AIST.
- An extensive effort on developing 32 x 32 silicon photonics switches was reviewed.
- All optical Nyquist filtering works well higher bandwidths than 40 Gbaud signals.

- Jump from ongoing optical cut-through to Dynamic Optical Path Network must be assisted through academic activities and standardization!
- Orchestration of different layers and/or standardization bodies is essential.
Thank you!

shu.namiki@aist.go.jp