OPS: Optical Packet Switches

Hiroaki Harai (harai@nict.go.jp)
National Institute of Information and Communications Technology

Sep 8, 2006
Optical Network Testbeds Workshop 3
Why do we Need OPS?

- Internet Traffic in Japan: approx. 500 Gbps
- Peta-bps backbone future: doubled per year $\xleftarrow{\text{500 Tbps in 10 years}}$
- Electronic packet switch
  - Year 2004: Throughput 640Gbps (16x40 Gbps)
- Lightpath networks
  - Need fully meshed connections/ feasible?
  - Need complex traffic engineering
  - Important technology for bandwidth-assured applications
- OPS networks
  - Provide extremely high-throughput
    - Much larger bandwidth for switching (> 40 Gbps)
    - O/E/O: 40Gbps $\rightarrow$ 64 x 622 Mbps bus, SERDES
  - May need MPLS-like control (labels can be merged)
  - Important to ubiquitous society
Optical Packet Switching

- Data-path is all-optical (No O/E/O)
  - Switch, Buffer
  - Increase data bandwidth
- Label lookup (i.e. forwarding)
  - Electronic parallel processing?
  - Optical processing
What should be Solved for OPS?

- OPS
  - Increasing number of ports of optical switch
    - Electronic: 16x16, 40Gbps → 640Gbps
    - Optical: 128x128, 160Gbps → 20.48 Tbps
      - 25 Waves → 500 Tbps
  - Increasing speed of label lookup and buffer management
    - Wire-speed operation
  - Increasing number of labels looked-up
    - Several thousands (New L2 possibility)
    - More (L3 switching)
  - Increasing buffer size
    - At least tens of fiber-delay-lines
  - Decreasing guard time between packets
    - Several nanoseconds
  - OPS Monitor/Analyzer
    - Bit error / Optical packet error

Under developing in NICT
Remaining Topics of This Talk

- OPS Prototype
- Optical label lookup
- Optical buffer
- Electronic buffer management
NICT’s 40Gbps-based OPS Experiment

N. Wada, H. Harai, F. Kubota, OFC 2003 (no. FS7).
Optical label has different modulation format with payload data

Optical label is physically distinguished from payload data

Optical hardware based label processing is available

Fully passive, ultra high-speed optical label processing
Time Domain Optical Code Processing
--Measured Waveform at 8-chip, 200Gchips/s

Optical Fiber-Delay-Line Buffer

- Different lengths of FDLs
- Need at least tens of FDLs
Optical FDL Buffering at 160Gbps

Switch 1 out
Packet collision!!

Switch 2 out

Buffer 1 out

Buffer 2 out

Output port

Avoidance of collision

Single Switch

Double Switches

Scheduler

Noise

Source: N. Wada (NICT)
High-Performance Buffer Management for Optical Fiber-Delay-Line Buffer

- Establish practical-scale high-performance management for FDL buffer

  1. Develop buffer management by parallel and pipeline processing
     - For number of ports, time complexity of each processor is $O(1)$
     - Parallel expansion of sequential (i.e. round-robin) scheduling
     - $N$-times higher throughput than sequential scheduling

  2. Confirm feasibility of support for 128x40Gbps packet switch by FPGA
     - 8 times higher performance than ASIC based router (16x40Gbps)
     - IP packet granularity (64byte or more; 10 Gpps), variable length

  3. Prototyping 8-port buffer management system

**Parallel and pipeline buffer management ($N=8$)**

---

# Performance Comparison

<table>
<thead>
<tr>
<th></th>
<th>NICT OPS Prototype</th>
<th>NICT’s Top Data (As of Sep, 2006)</th>
<th>Electronic Router</th>
</tr>
</thead>
<tbody>
<tr>
<td>IN/OUT ports</td>
<td>2</td>
<td>2 *</td>
<td>16</td>
</tr>
<tr>
<td>Bit rate</td>
<td>160 Gbps</td>
<td>160 Gbps</td>
<td>40 Gbps</td>
</tr>
<tr>
<td>Label processing</td>
<td>800 Mpps/port</td>
<td>10 Gpps (at 40Gbps)</td>
<td>125 Mpps/port **</td>
</tr>
<tr>
<td>Scheduling</td>
<td>4 Mpps</td>
<td>10 Gpps</td>
<td>2 Gpps **</td>
</tr>
<tr>
<td>Buffer</td>
<td>2/port</td>
<td>31/port</td>
<td>160000/port</td>
</tr>
</tbody>
</table>

* Can scale with nanosecond optical switches
** Estimated data: Assumption of wire rate processing of 40byte-packets
Conclusion

- We need high-throughput backbone network for ubiquitous society
- OPS will provide extremely high-throughput
  - Switching bandwidth is not limited
  - Buffer size is increasing
  - Electronic scheduling is fast
- NIC'T has developed OPS but,
  - Need more advanced devices (e.g., ns-switch) and systems

Thank you for your attention

Acknowledgment

- N. Wada, H. Furukawa of Photonic Network Group in NIC'T for valuable discussion, collaboration, and some slides in OPS