Design and Performance Evaluation of Small-Buffered Optical Packet Switching Networks

Onur Alparslan

Advanced Network Architecture Laboratory
Graduate School of Information Science and Technology

Osaka University

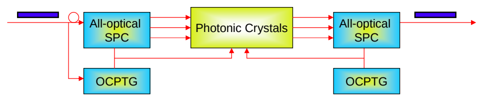
Outline

- Problem Statement
- Objective
- Proposals
- Conclusions
- Future Research

Problem Statement

- **Major differences between optical packet-switched (OPS) networks and electronic packet-switched (EPS) networks.**
- **In EPS networks, contention is resolved by**
  - Storing the contended packets in a random access memory (RAM)
- **Buffering limitations in optical domain,**
  - The well-known rule-of-thumb link buffer size requirement (*B = RTT x BW) is too big for optical buffering*
  - Conversion from optical to electronic domain in order to use electronic RAM is not a feasible solution, because of the processing limitations of EPS, so processing and switching in the optical domain is necessary.
- **Buffering in the optical domain**
  - Fiber Delay Lines (FDL)
    - FDLs require very long fiber lines, which cause signal attenuation, inside the routers.
    - There can be a very limited number of FDLs in a router due to space considerations, so they can provide a small amount of buffering
  - Optical RAM
    - Still under research
    - Not expected to have a large capacity, soon
- **TCP has low throughput due to burstiness, when buffer is very small**

NTT's All-Optical RAM



- **All-Optical RAM is better than Electronic RAM, because**
  - Requires much less power
  - Size is smaller
  - Integration is better

Objective

- **Designing a new OPS network architecture that can achieve high utilization by using small buffers**
  - Fiber Delay Lines
  - Optical RAM
- **Showing the buffer requirements**

Advantages

- **Decreasing the buffer requirements in the core**
- **Realizing all-optical high-speed OPS networks**

Research

- **Proposed and evaluated two different OPS network architectures for very small buffered networks**
  - Rate-based XCP controlled edge pacing
  - Buffer occupancy based edge/core node pacing

Outline of Thesis

1. Introduction
2. Rate-based Pacing for Small-buffered Optical Packet Switched Networks
3. Switch Architectures for Small-buffered Optical Packet Switched Networks
4. Pacing for Optical Packet Switched Networks with Very Small Optical RAM
5. Node Pacing for Optical Packet Switching
6. Conclusions

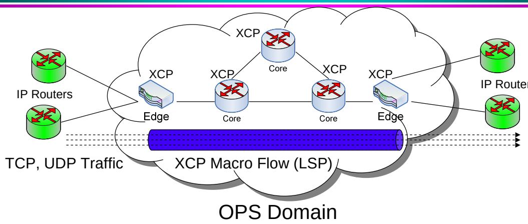
Rate-based XCP Pacing for Small-buffered Optical Packet Switched Networks

XCP-Based Proposed Solutions 1/2

- **Preventing wavelength over-utilization**
  - Apply XCP-based congestion control
    - » XCP is a new congestion control algorithm specifically designed for high-bandwidth and large-delay networks.
    - » Network layer control
    - » Nodes exchange probe packets in order to learn link information
    - » Uses an efficiency controller for high link utilization and fairness controller for high fairness among flows
  - Carefully select XCP parameters
  - Control maximum wavelength utilization ratio by XCP

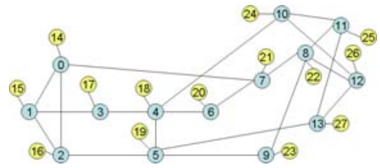
D. Katabi, M. Handley, and C. Rohrs, "Congestion control for high bandwidth-delay product," in *Proceedings of ACM SIGCOMM*, 2002, pp. 42-49.

XCP-Based Proposed Solutions 2/2



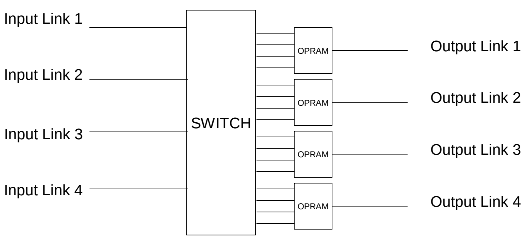
- **Burstiness**
  - Establish macro flows between edge nodes
  - Assign incoming TCP, UDP traffic to macro flows (similar to XCP-CSFQ, TeXCP)
  - Apply leaky bucket pacing to macro flows according to XCP flow rate at edge node
  - Possible to use LSPs for controlling macro flows if GMPLS is available

NSFNET Simulations

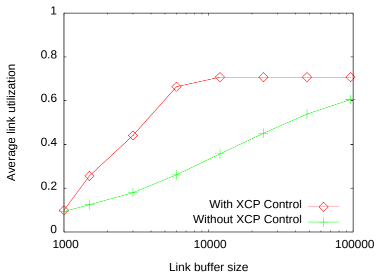


- 28 nodes (14 edge + 14 core) and 35 links (21 core + 14 edge)
- Wavelength speed 1Gbit/s
- 40 seconds simulation
- Around 1500 TCP flows
- Output buffering with optical RAM
- XCP's alpha, beta and gamma parameters are 0.2, 0.056 and 0.05 respectively

Output Optical RAM Buffered Switch

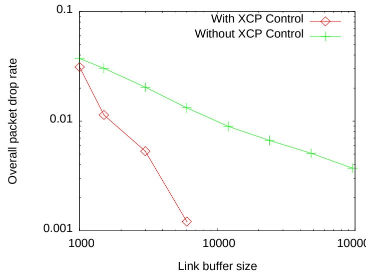


Link Utilization Comparison



- **Due to traffic matrix, some links are congested and some links are underutilized, so overall average link utilization converges to around 0.7**
- **Rule-of-thumb requires around 3 Mbytes buffer per link**
- **Only 6 Kbytes of buffering is enough for our architecture!**

Drop rate Comparison

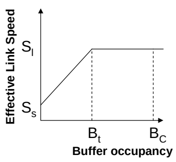


- **Only around 6 Kbytes of buffering is enough for achieving very low drop rate in our architecture**

Node Pacing for Optical Packet Switching

Buffer Occupancy Based Edge/Core Pacing

- No XCP Control, so it is simpler and easier to implement
- Uses part of the node buffer for smoothing the traffic
- Applies pacing at the output link by spacing between packets



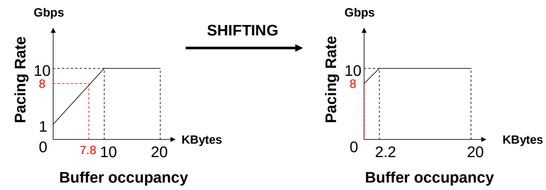
$$\text{Rate} = \begin{cases} S_S + B\,\frac{S_l - S_S}{B_t} & B < B_t \\ S_l & B \geq B_t \end{cases}$$

- Updates effective link speed based on buffer occupancy
  - **Rate** : Effective link speed
  - **B** :Buffer Occupancy
  - $\mathbf{B_t}$:Buffer Threshold
  - $\mathbf{B_C}$:Buffer Capacity
  - $\mathbf{S_S}$:Initial Pacing Rate
  - $\mathbf{S_l}$:Link Capacity

Adaptive Control

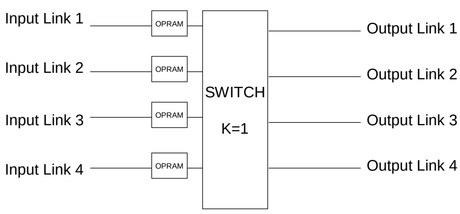
- **Use both buffer occupancy and average arrival rate for calculating pacing rate in order to decrease average buffer occupancy**

Example

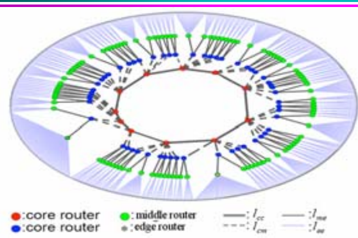


Input Buffered Switch

- Small switch size
  - We need a switch with a size of NxN
- Apply Virtual Output Buffering (VOB) for solving head-of-line blocking



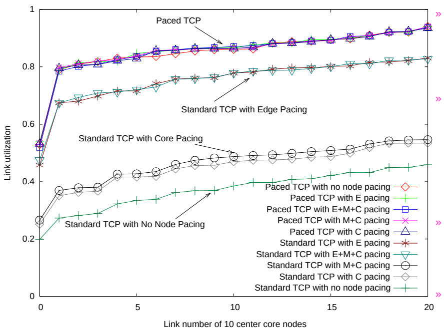
Sparse Abilene Topology



- 1Gbit/s link speed
- Simulate TCP Variants
  - **Paced TCP**
  - **Standard TCP**
- Simulate Node Pacing
  - **Edge node (E) pacing**
  - **Middle node (M) pacing**
  - **Center node (C) pacing**

- 698 edge nodes (E)
  - Grey nodes (Traffic enters the network)
- 106 middle core nodes (M)
  - Green nodes (Routers that are connected to edge nodes)
- 65 center core nodes (C)
  - Red and Blue nodes (Routers that are only connected to core nodes)
- 10 center backbone node (C)
  - Red nodes

Simulation Results (50Kbits Input buff.)



» **Comparison of sorted average utilization of links between backbone nodes**

» **When utilization is low and burstiness is high, core pacing brings a considerable throughput improvement.**

» **Additional edge pacing can almost double the throughput**

» **When Paced TCP is used, additional E, M or C pacing does not make a difference**

Conclusions

- **Proposed and simulated two different OPS network architectures for very small buffered networks**
  - Rate-based XCP Pacing for Small-buffered Optical Packet Switched Networks
    - » Simulations show that our architecture can decrease the buffer requirements by 500 times
    - » High link utilization and low packet drop rate with very small optical buffers
  - Node Pacing for Optical Packet Switching
    - » Our node pacing algorithm achieves considerably high link utilization with an easier deployment without requiring XCP control

Future Work

- **Hybrid OPS/OCS integrated architecture**
  - Support of optical circuit switching relaxes the traffic and buffer occupancy
    - » Circuit switching requires no buffering inside the network as there is no multiplexing.
  - ORION architecture can be a good hybrid architecture for further decreasing buffer requirements

Thank you