An Optically Turbocharged Internet Router

CCW 2001, Charlottesville, VA, Oct. 15

Joe Touch

Director, Postel Center for Experimental Networking

Computer Networks Division

USC/ISI

Information Sciences Institute

Outline

- Optical vs. Internet
- Goals of an Internet optical router
- Architecture
- Benefits and Issues
- Simulation results
- Remaining work

Motivation

Need Tbps true routers, but…

- Routing tables growing exponentially
- Lack of scaling (past 15 years):
  - Links increased 1,000,000x
  - CPUs increased 1,000x
  - RAM access increased 10x

Internet vs. Optical

- Internet
  - Longest-prefix search of 200,000 entries
  - Decrement TTL
  - Update checksum (IPv4)
  - Buffering for statistical muxing
- Optical
  - Too little time to do full lookup (160ps/pkt@1Tbps)
  - Per-packet switch setup *maybe*
  - Can't compute checksum
  - No storage

Current optical focus

- WDM as a bonus
  - Needed to overcome dispersion
  - Can be used to partition (benefit?)
  - Can be used to route (benefit?)
- Connection-based/-like traffic
  - ATM/MPLS flow-based setups (MPλS, SWAP)
    - BUT: Setup doubles connection latency
  - Packet-train setup on-the-fly (OBS, TBS)
    - BUT: Setup requires large gap after first packet
  - BUT: Both expect long flows or flow aggregation

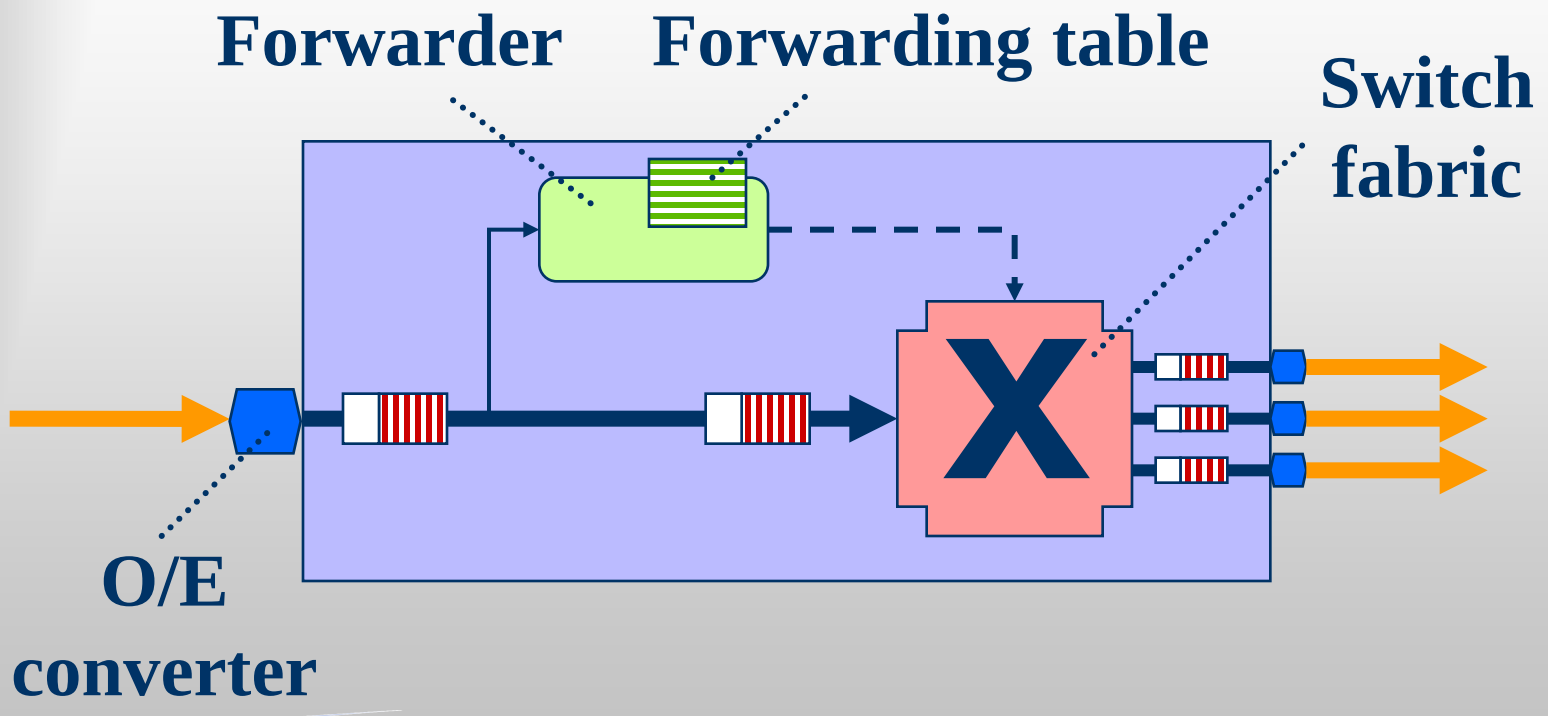
Goal – Optical Internet

(the rest of this talk)

- IP over light
  - No setup
  - Single terabit channels (***no WDM***)
  - Works for short flows, or for single packets

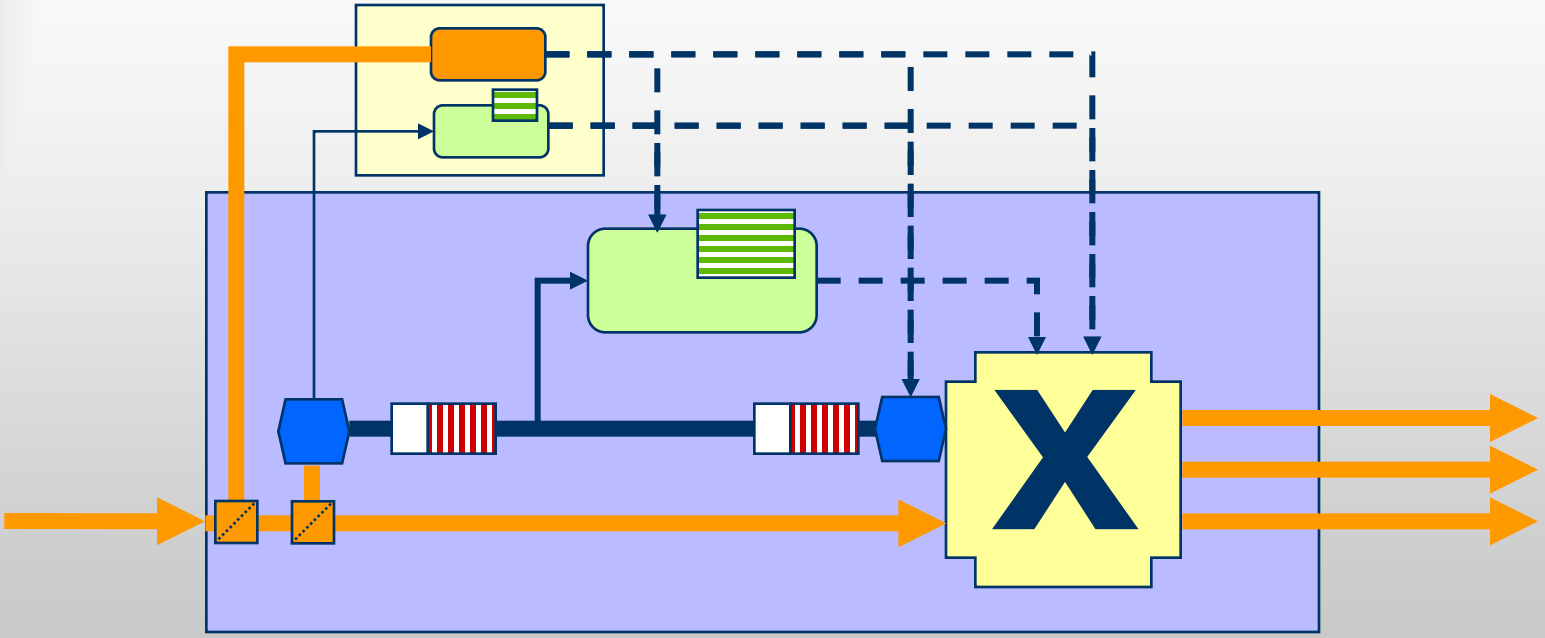
Inside current routers

- Forwarder + switch fabric

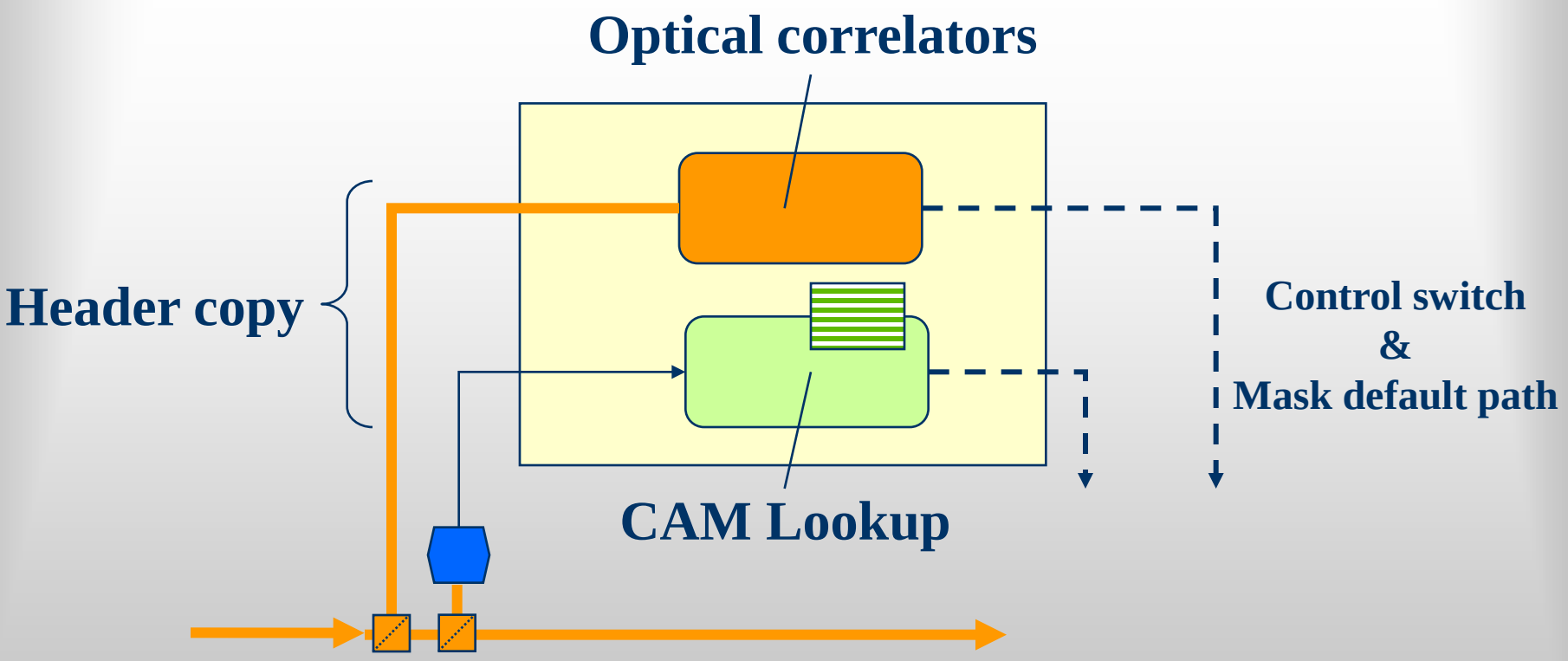


Optical Assist

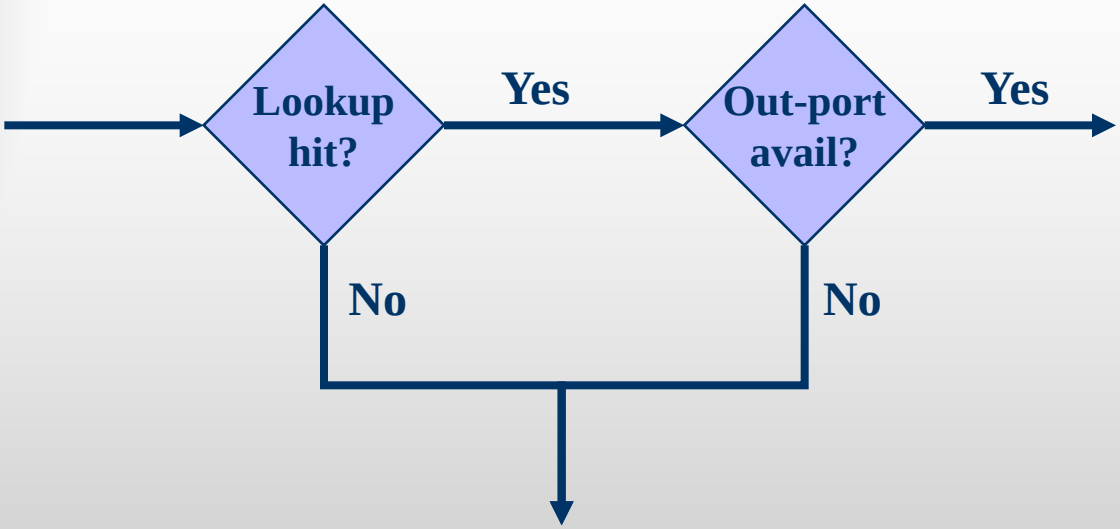
- Optics as ‘fast path’; electronics as backup



Bypass Design



Bypass decision tree



Key components

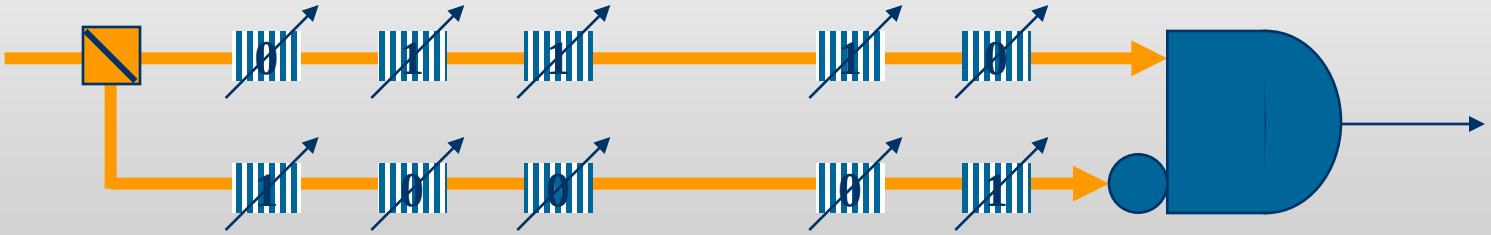
- Accelerator lookup
  - All-optical via correlators
  - Electronic via CAM
- Electronic backup path
  - Full lookup when correlator/CAM fails
  - Queue when output port is busy
- Optical switch
  - $LiNbO_3$ elements

Benefits of Accelerator

- Incremental deployment
  - No “cloud” required
- Partial solution
  - Electronic is full backup
- Lower latency
  - ns per hop, not ms
- Higher throughput
  - Offloads electronic path

Optical correlator

- Sequence of Bragg filters
- Tuned to match 0,1,X
  - 0,1 requires pairs, X is pass-through

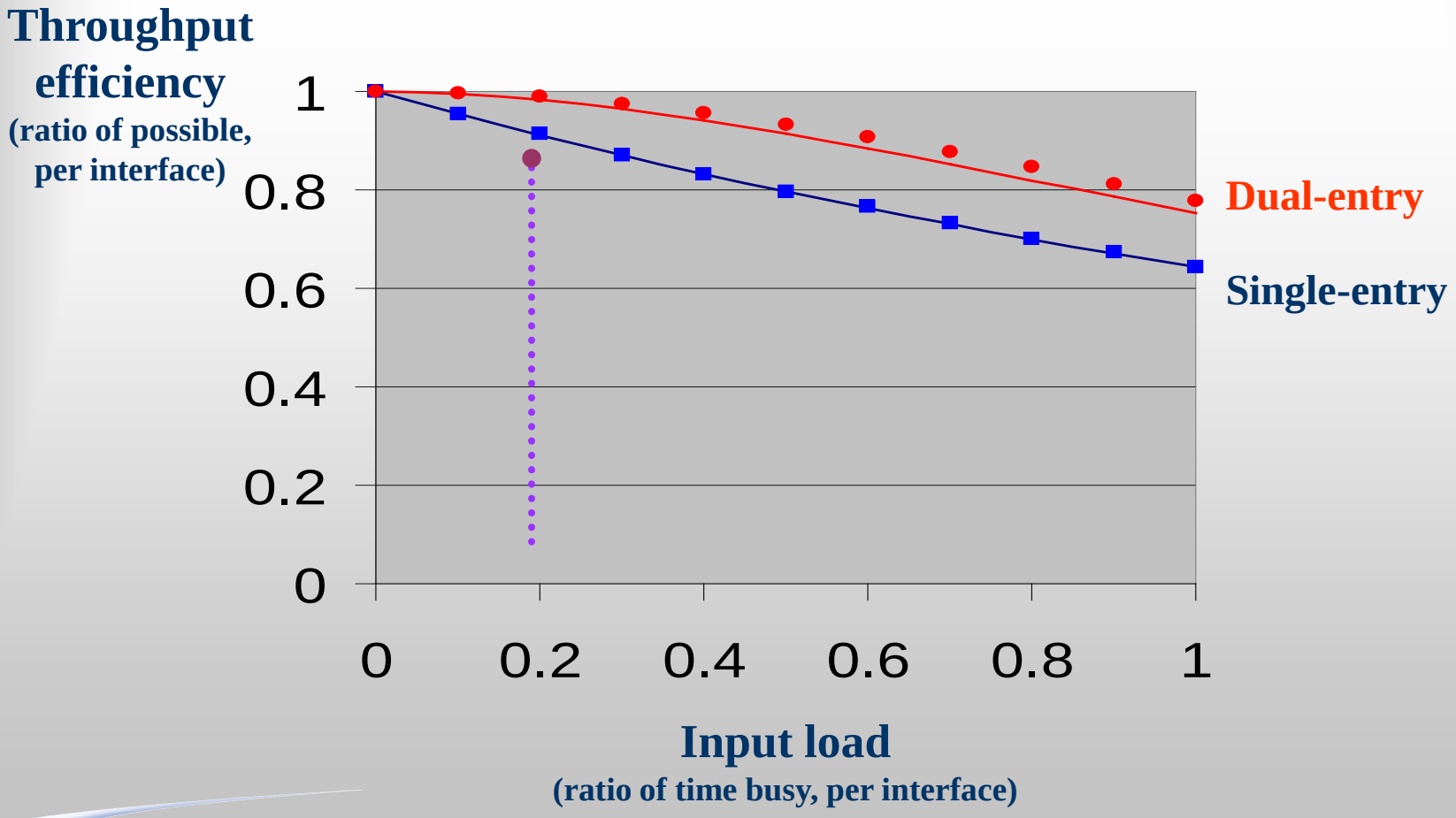


Some limits

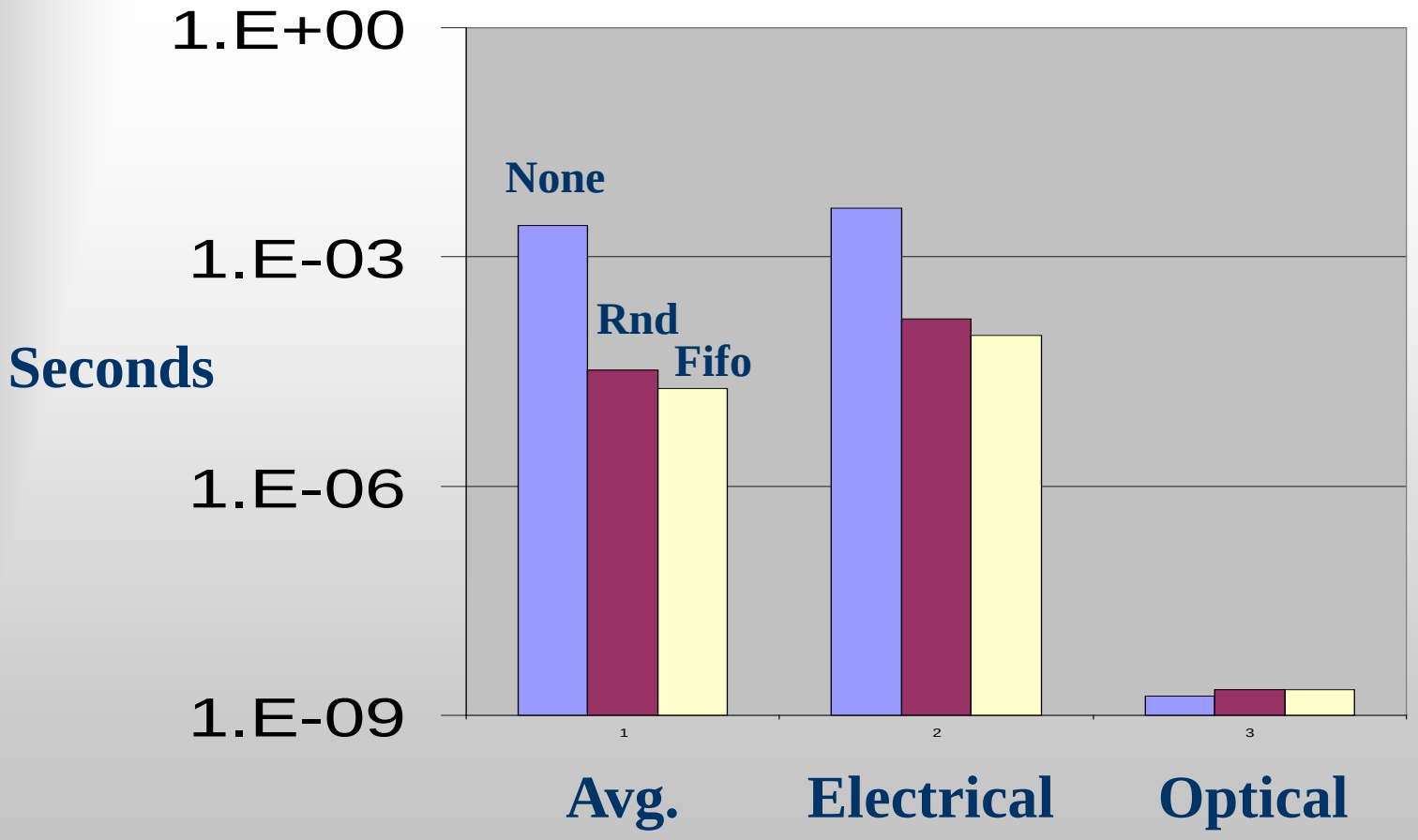
- Correlators
  - ~8 correlators
  - ~8 bits per correlator
- CAM
  - 1-8K entries, 1ns lookup (7x pipeline)
- Switch
  - $LiNbO_3$ at 1ns currently (parallelize?)

How useful?

Switching gain



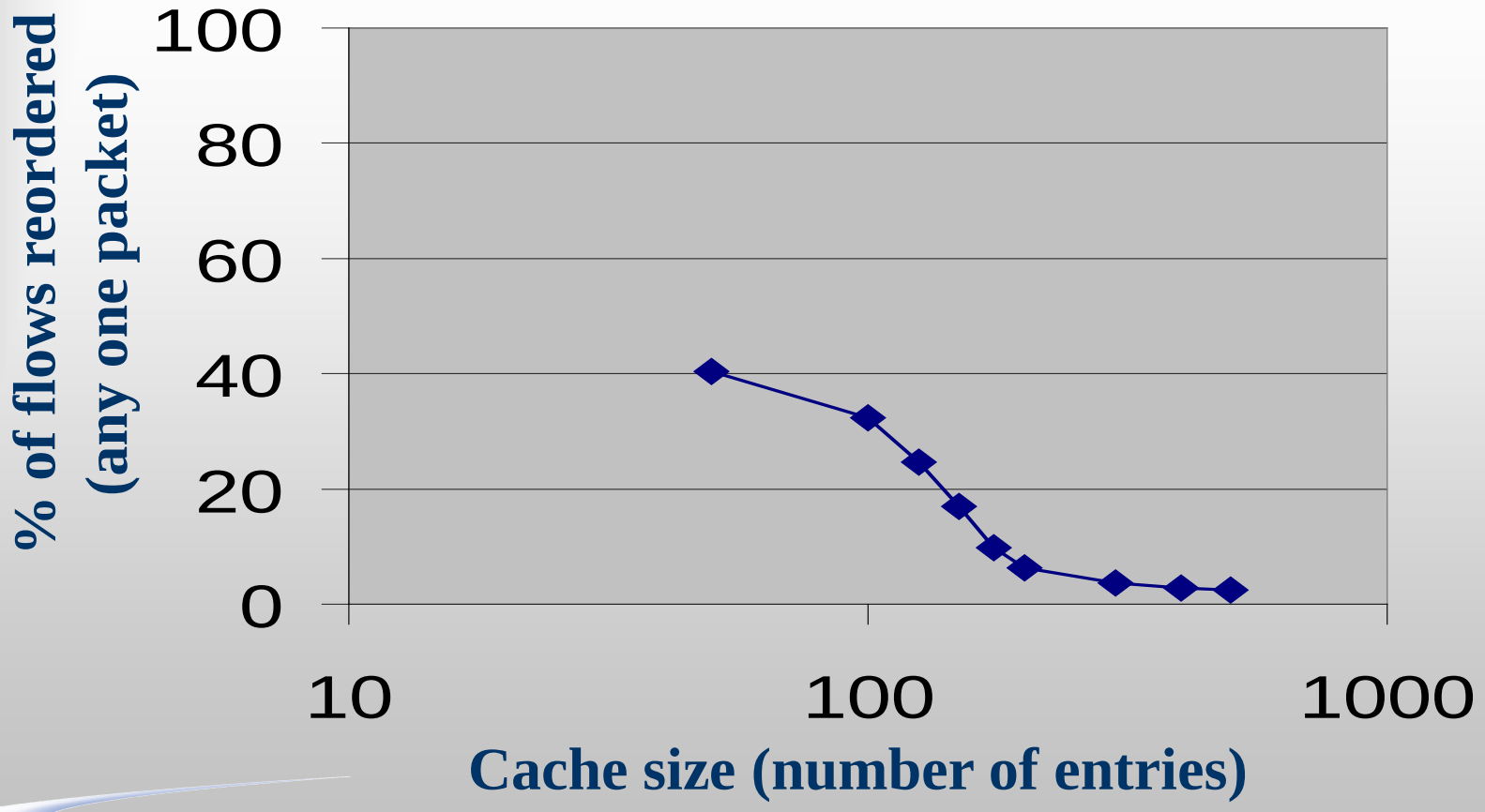
Latency



Cost

- Packet reordering
  - Two separate paths,
    each path non-reordering,
    but together can reorder
- HOL blocking
  - Optical has priority on outputs
  - Use second-best paths

Reordering



Partition function

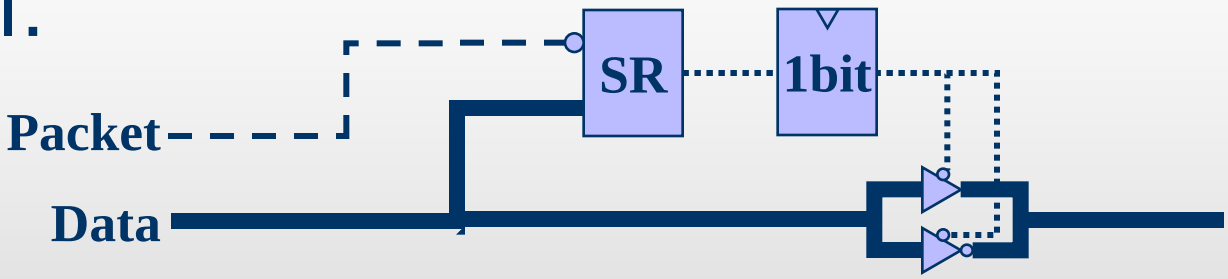
- Take advantage of optics
  - Easier to match 0's than 1's
  - Hard to match all 24 bits, easy to pick movable subset
- Incomplete function is OK
  - Avoid false positives
  - False negatives just reduce efficiency

Missing links

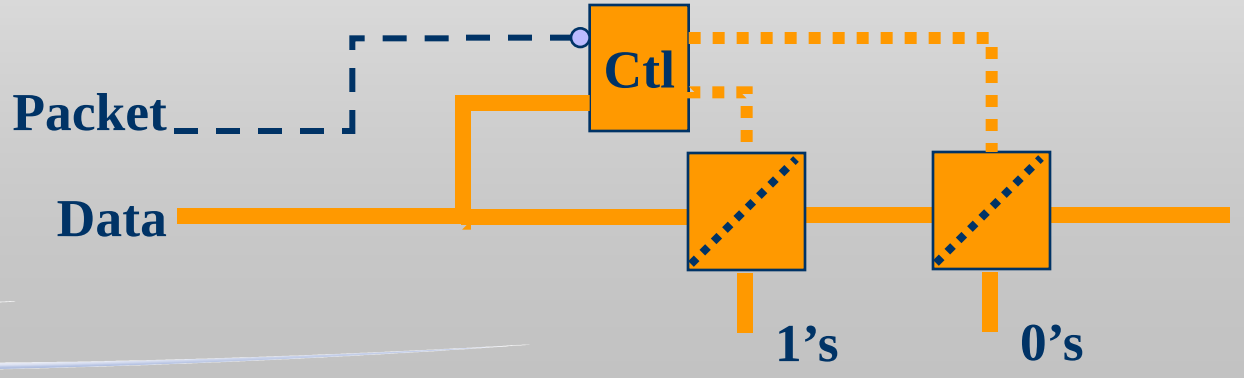
- TTL decrement-and-test
  - Trivial in electronics
  - All-optical design underway with USC
- Checksum
  - Update via optical header replacement
  - “Deferred maintenance” approach
  - Use IPv6? ☺
- High-speed traces and routing tables

Optical TTL

- LSB first encoding
- Electrical:



- Optical:



Current status

- Simulation analysis
  - What percent of traffic will be optically switched?
  - How much reordering?
- Optical correlator design
  - Implemented at USC campus
- Partition algorithm
  - Under development

For more info:

- http://www.isi.edu/pow
  - Joe Bannister, Joe Touch
  - GRAs Purushotham Kamath, Aatash Patel
  - GRA Stephen Suryaputra (alumnus)
- Papers:
  - **An Optical Booster for Internet Routers,** J. Bannister, J. Touch, P. Kamath, A. Patel, Invited Paper, Proc. Eighth International Conference on High Performance Computing, Hyderabad, India, Dec. 2001.
  - **Simple Wavelength Assignment Protocol,"** S. Suryaputra, J. Touch, J. Bannister, in Terabit Optical Networking: Architecture, Control, and Management Issues, John M. Senior, Sudhir Dixit, Chunming Qiao, Editors, Proceeding of SPIE Vol. 4213, pp. 220-233 (2000) (Proc. Photonics East).
  - **How Many Wavelengths Do We Really Need? A Study of the Performance Limits of Packet Over Wavelengths,** J. Bannister, J. Touch, A. Willner, S. Suryaputra, Optical Networks, April 2000, pp. 17-28.
  - **Dynamically Reconfigurable All-Optical Correlators to Support Ultra-fast Internet Routing,** M. Hauer, J. McGeehan, J. Touch, P. Kamath, J. Bannister, E.R. Lyons, A. Willner, submitted to OFC 2002.

Advance Q&A

- Why accelerate in parallel vs. pipelined?
  - Most current caches are in parallel
  - May ease incremental add if pipelined
- Why not correlate in electronics?
  - Electronic path can't run at 1 Tbps
  - Still need accelerator