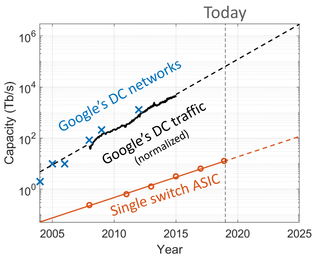
An Optical Fabric
for High-Bandwidth and Low-Latency Compute Clusters

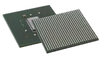
Max Mellette[1] and George Porter[1,2]
(max@infocusnetworks.com)

[1] inFocus NETWORKS
[2] UC San Diego JACOBS SCHOOL OF ENGINEERING

Challenges for Data Center Networking Infrastructure



1. ASIC scaling



Single chip I/O limits

Multi-ASIC chassis increases cost, power of fabric building block

2. OEO conversion

Electronic switching + Optical links = Large number of expensive optical transceivers in fabric

3. Upgrades

Upgrading to higher bandwidth requires replacing the entire network core

Optical Switching

Opportunities:

Transparent switching…

- ✓ Sidesteps ASIC scaling bottlenecks
- ✓ Reduces the number of OEO conversions
- ✓ Survives multiple link speed upgrades

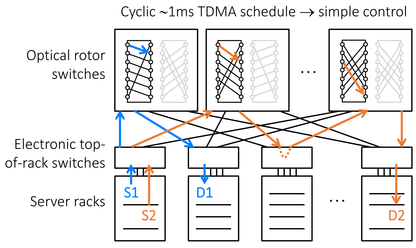
Challenges:

Transparent switching means…

- Routing and forwarding are external
  Control complexity a major issue
- No signal regeneration
  High radix, low-loss, fast switching are required, but are difficult to achieve

RotorNet: An Optical Cluster Network

Topology:

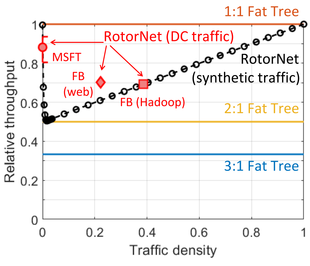


Blue: bandwidth-efficient direct path
Red: low-latency cut-through path

Cost:

(cost of optical port = 0.25× cost of packet port + TRX)

Throughput:



Mellette et al., SIGCOMM '17

Latency:

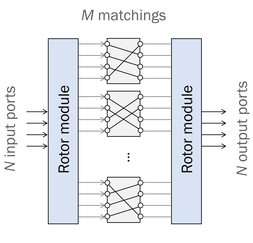
Simulated 95th percentile flow completion time (77 kB flows)				
Websearch traffic load:		1%	5%	10%
Network Architecture:	Fat Tree	71 µs	144 µs	172 µs
	RotorNet	80 µs	155 µs	287 µs

RotorNet enables (at similar deployment cost to a 3:1 Fat Tree):

- ✓ 1.5–3x higher throughput
- ✓ Similar flow completion times
- ✓ Optical link speed upgrades without replacing the network core

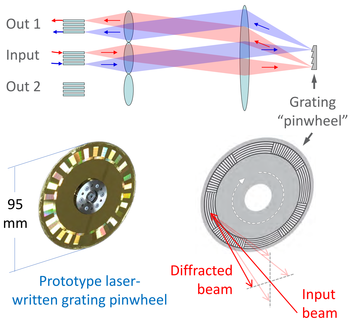
Rotor Switch: A Novel Optical Switch

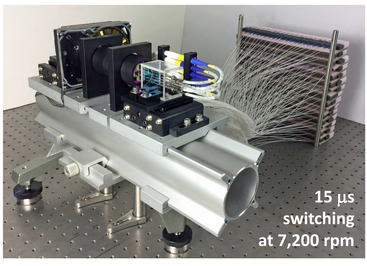
Rotor switch diagram:



Sequential switching between *M* *N* x *N* port matchings

Rotor module optical layout:





Prototype rotor module: 61 ports ×4 matchings
~4 dB loss* over 100 nm (2 dB loss designed)
**can improve grating efficiency by 2 dB*

→ Scales to 2,048+ ports with larger pinwheel

Acknowledgement

NSF
arpa·e CHANGING WHAT'S POSSIBLE
State of California Energy Commission