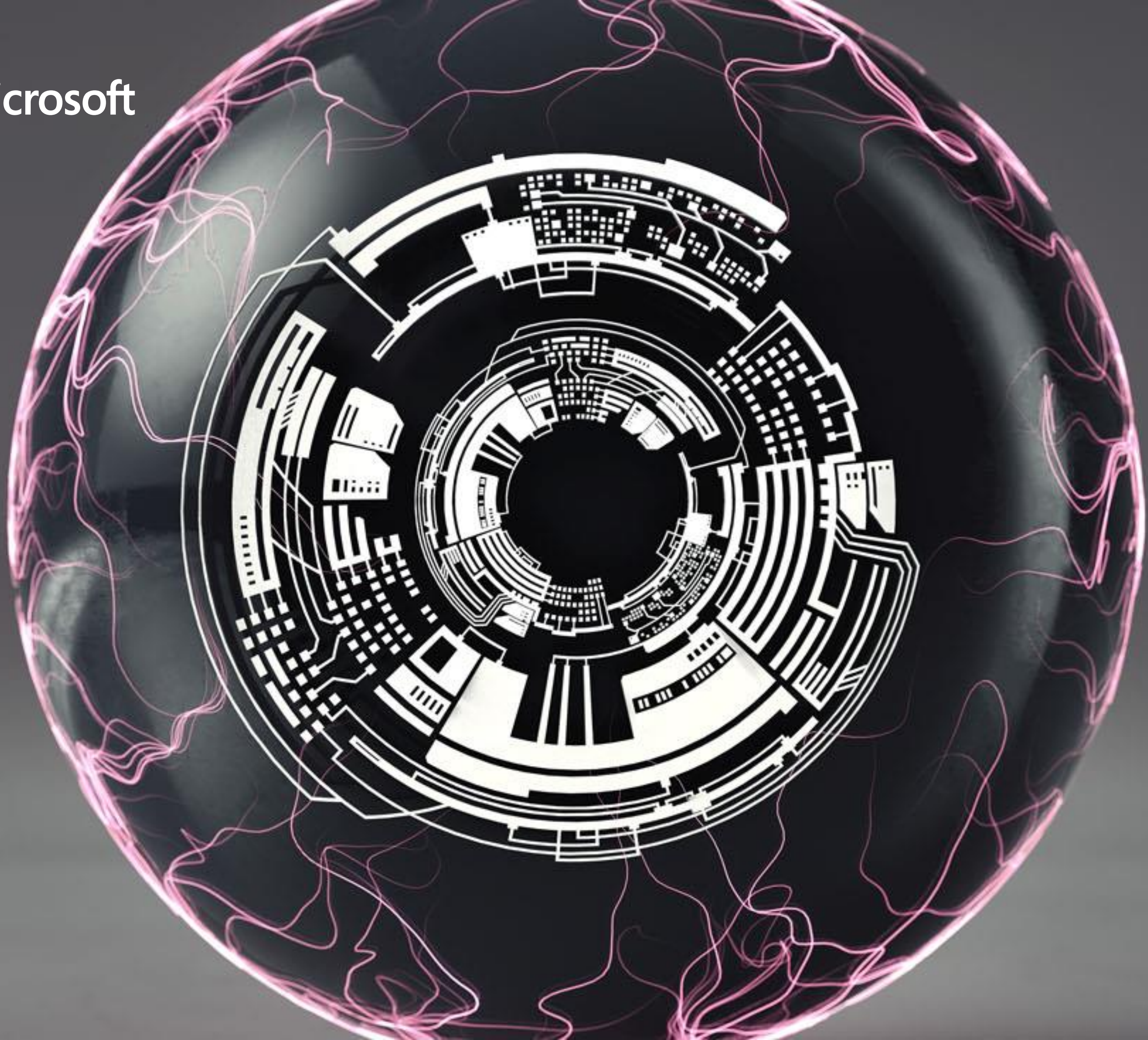
Microsoft



Microsoft Research Faculty Summit 2016

Large-Scale Silicon Photonic Switches

Ming C. Wu

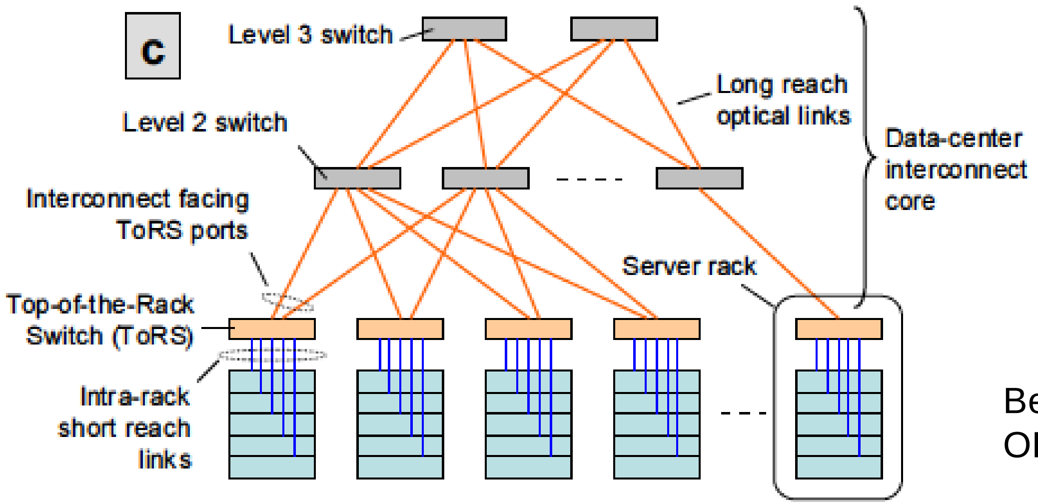
University of California, Berkeley

Electrical Engineering & Computer Sciences Dept.
& Berkeley Sensor and Actuator Center (BSAC)

Outline

- Optical switches for data centers
  - Why do we need it?
  - What's available now?
  - What's needed in the future?
- Silicon photonic switch
  - Is it a game changer?
  - How does it scale?
- Silicon photonic MEMS switches
  - 64x64 switch on 1-cm^2 chip
  - Technology scaling
- Discussion about packaging needs
- Summary

Challenges in Datacenter Networks

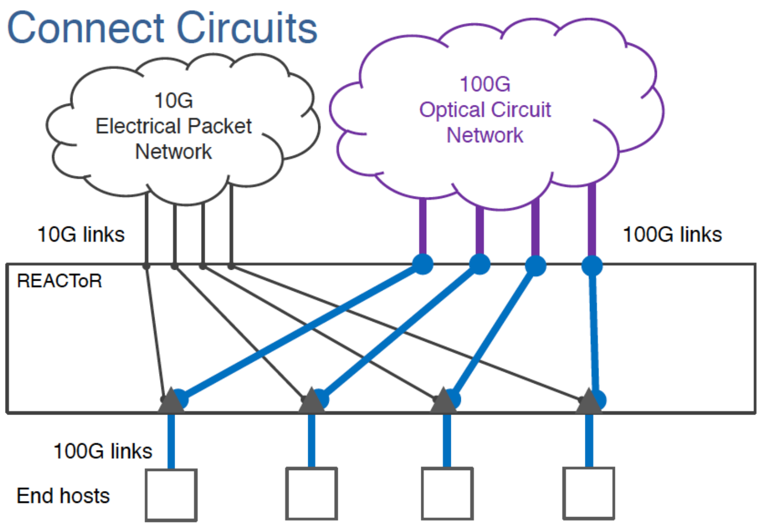


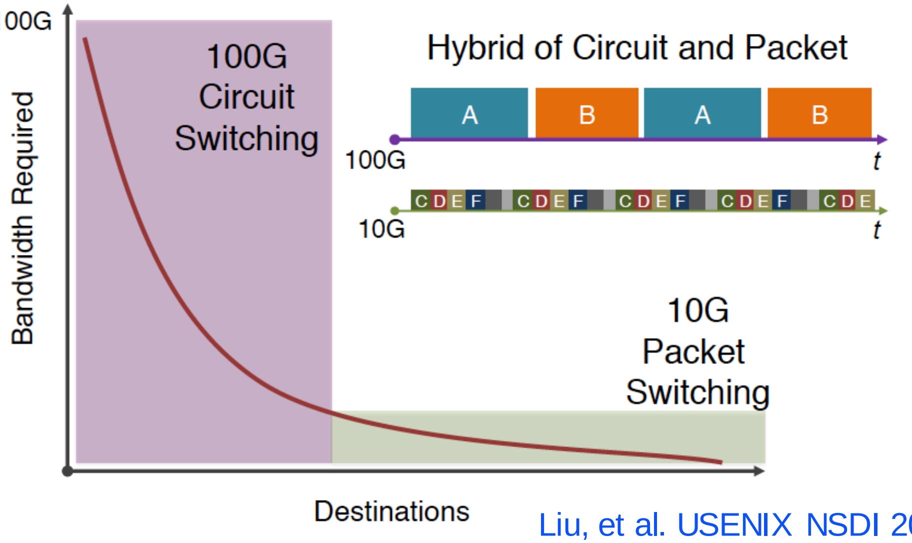
Bergman, Rumley, OECC 2016

- **Link rate continues to increase (40G, 100G, 400G, ..)**
- **Cannot rely on continual scaling of CMOS**
  - **Switch bandwidth-portcount limited by thermal issue, die size, pin count**
- **Optical switching can facilitate scaling out datacenters**
  - **Reduce number of hops, transceivers, power consumption**
  - **Critical issues: switching time, arbitration, cost, power consumption, scalability**

Berkeley
UNIVERSITY OF CALIFORNIA

Hybrid Data Center Networks Example (1): REACToR (UCSD)





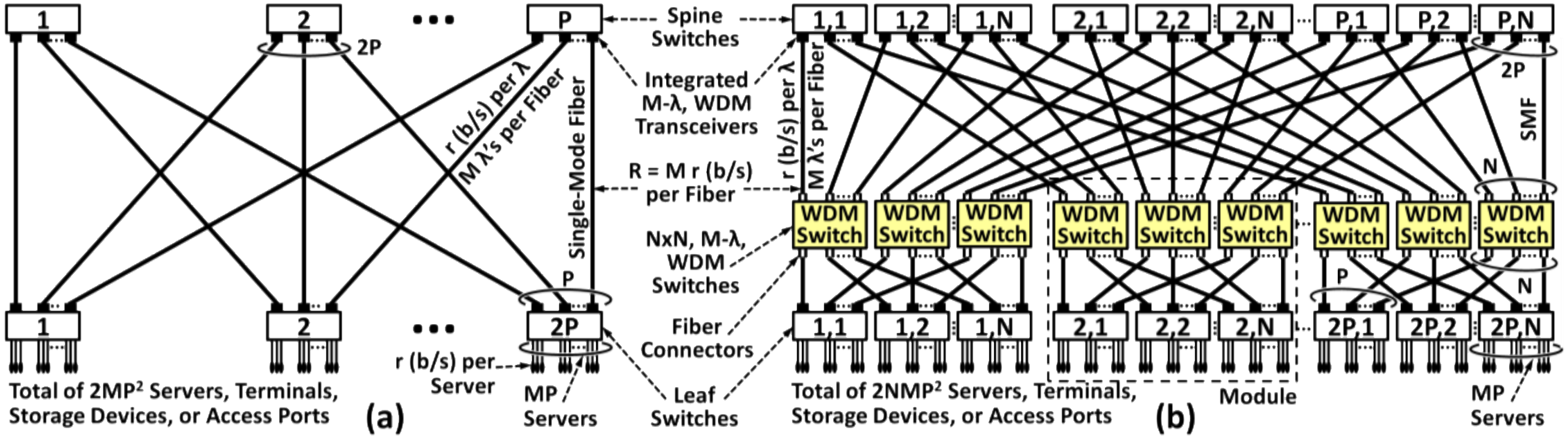
Liu, et al. USENIX NSDI 2014

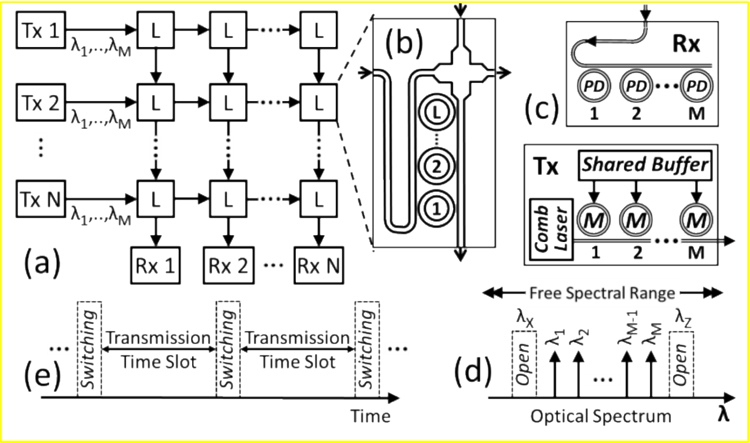
- **Optical circuit network**
  - **High bandwidth,**
  - **Bufferless TDMA network**
  - **Tx when circuit connects**
- **Electrical packet switch networks:**
  - **Low bandwidth, small packets**
  - **Buffered all the way**
  - **Tx all the time**
- **Implementation:**
  - **Using Nistica MEMS WSS**
  - **30 us reconfiguration time**
- **Performance**
  - **10G EPS + 100G OCS ≈100G EPS**

H. Liu, F. Lu, A. Forencich, R. Kapoor, M. Tewari, G. M. Voelker, G. Papen, A. C. Snoeren, and G. Porter, “Circuit Switching Under the Radar with REACToR,” USENIX Conference on Networked Systems Design and Implementation, 2014

Berkeley UNIVERSITY OF CALIFORNIA

Hybrid Data Center Networks
Example (2): Elastic WDM Switches (UCSB)





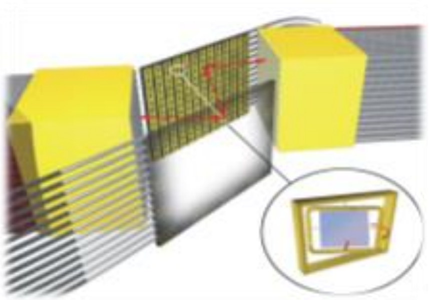
A.A.M. Saleh, A.S.P. Khope, J.E. Bowers, R.C. Alferness, "Elastic WDM Switching for Scalable Data Center and HPC Interconnect Networks", OECC 2016

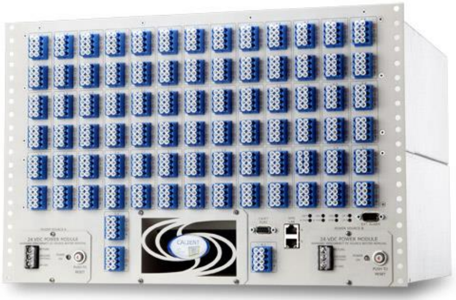
- **Adding an layer of optical switches between spine and leaf greatly expand the scale of network (number of servers)**
- **Can be space switch or wavelength switch**
- **Wavelength routing also investigated by many other groups (Columbia, UCD, Nagoya U, NTT, ..)**

Commercially Available Switches

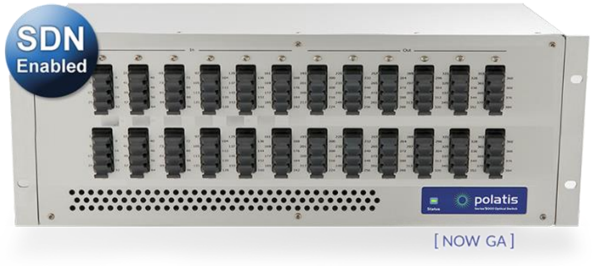
3D (Free-Space) Switch

Calient: 320x320





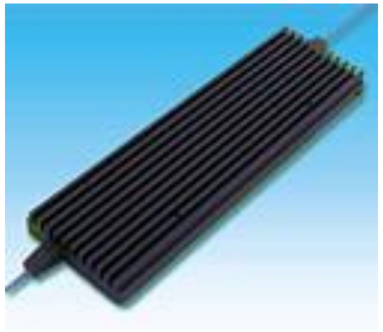
Polatis: 384x384



+ High port count
+ Low loss: < 3 dB
– Slow (10 to 25 ms)
– High cost ($100’s /port)

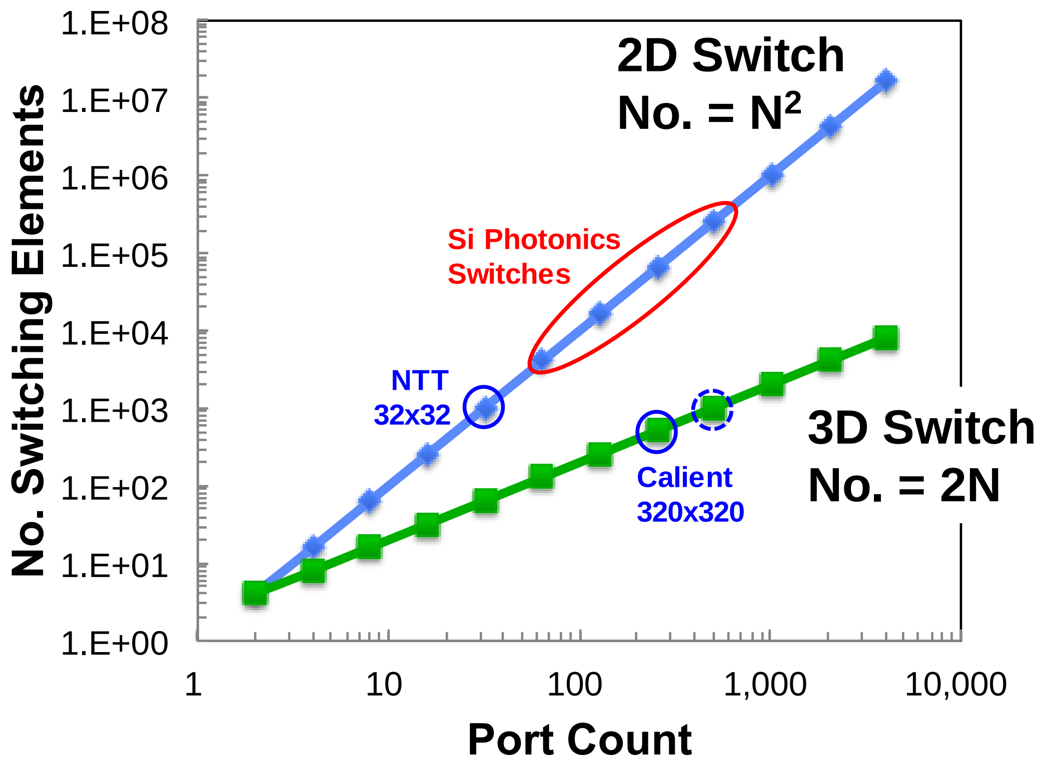
2D (Integrated) Switch

NTTElectronics: 16x16 (commercial)
32x32 (publication)



– Limited port count
– Higher loss: 5 dB
– Slow (3 ms)
+ Low cost (~ $10’s /port ?)
+ Fully integrated

Scaling of Optical Switches: 2D vs 3D



Largest (commercially available) Switch:

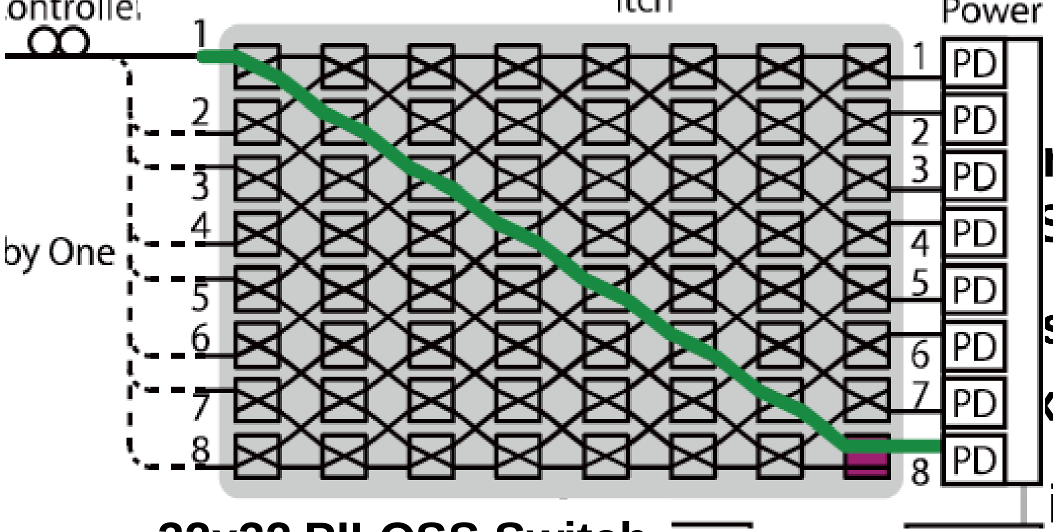
- 3D MEMS: 320x320 (< 3dB)
- 2D PLC: 32x32 (6.6 dB)

- 3D switch uses 2N analog mirrors
  - Complex control
- 2D switch uses ~ N^2 digital switching elements
  - but simple control
  - Monolithic integration
- Si photonics can be a game changer
  - High integration density
  - Tight bending radius

Berkeley
UNIVERSITY OF CALIFORNIA

32x32 Non-Blocking PILOSS Switch (AIST, Japan)

8x8 PILOSS Switch



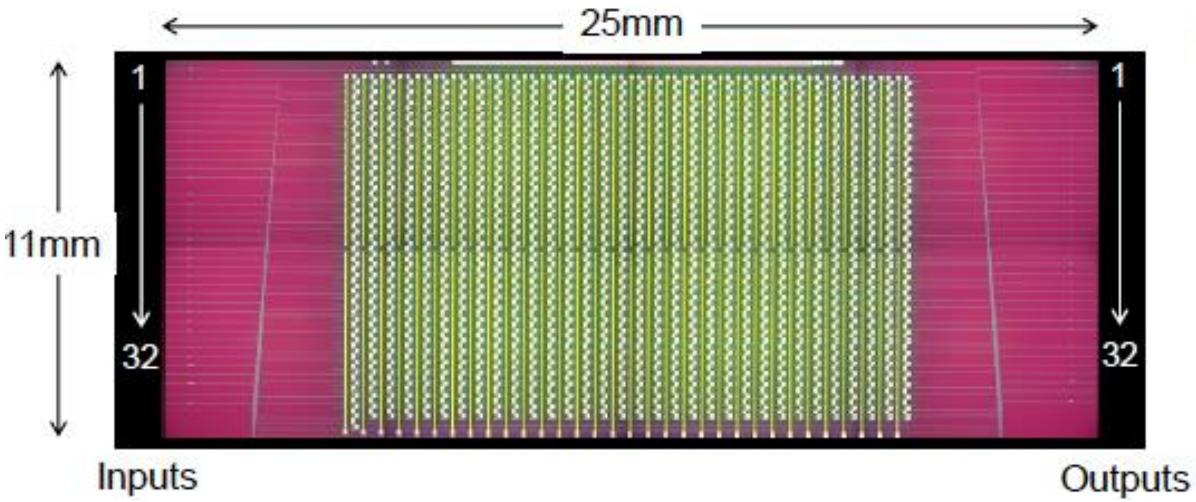
switch on 11x25 mm²

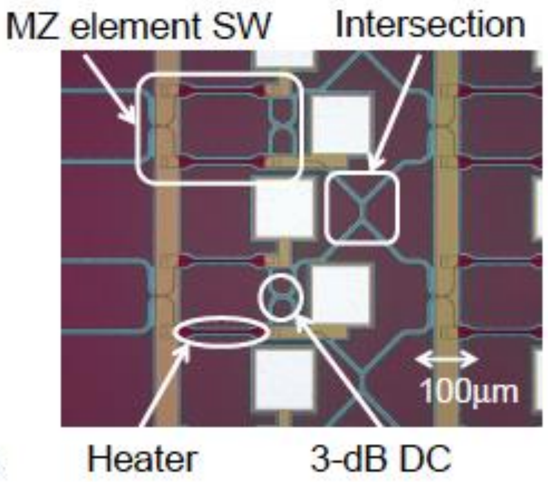
ndependent insertion loss
SS) architecture

stage topology:
xN switch has N stages

ip loss: 15.8 dB

32x32 PILOSS Switch

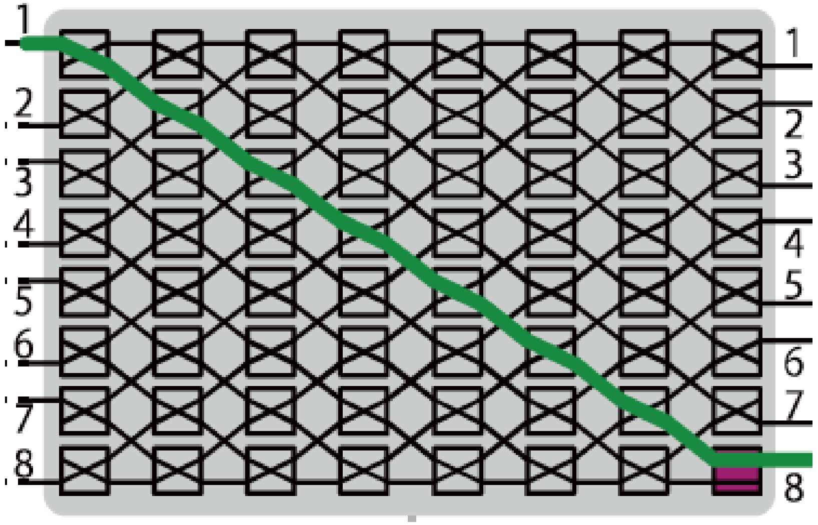




K. Tanizawa, et al, "Ultra-compact 32 × 32 strictly-non-blocking Si-wire optical switch with fan-out LGA interposer," Optics Express, 2015.

A Different Approach for High Radix Switch

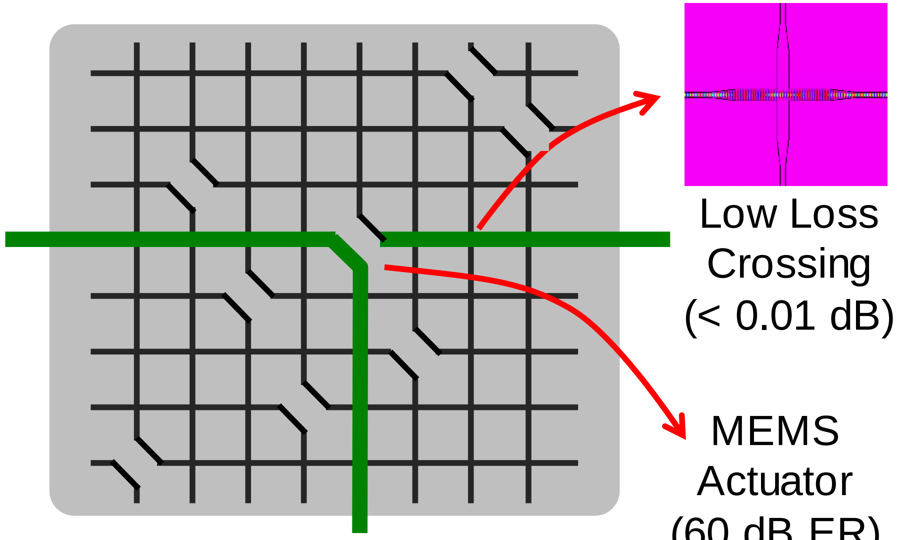
Traditional Approach:
Active Crossbar + EO (or Thermo) Switching



(PILOSS Switch from AIST, Japan)

- Many cascaded 2x2 elements
- Lossy in both Bar and Cross states
- High cumulative loss (~ 0.5·N dB for NxN)
- Largest switch demonstrated: 32x32 with 16dB on-chip loss

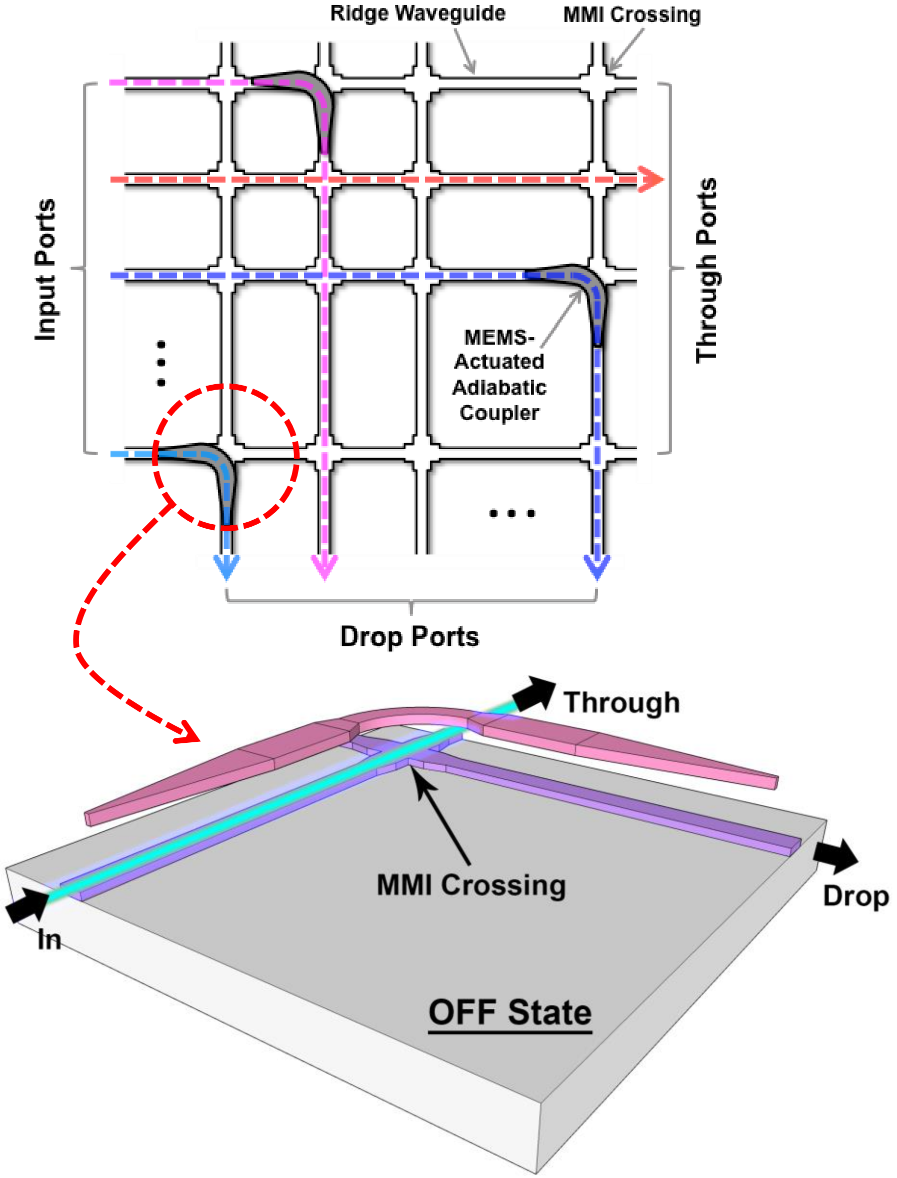
Berkeley Approach:
Passive Crossbar + MEMS Switching



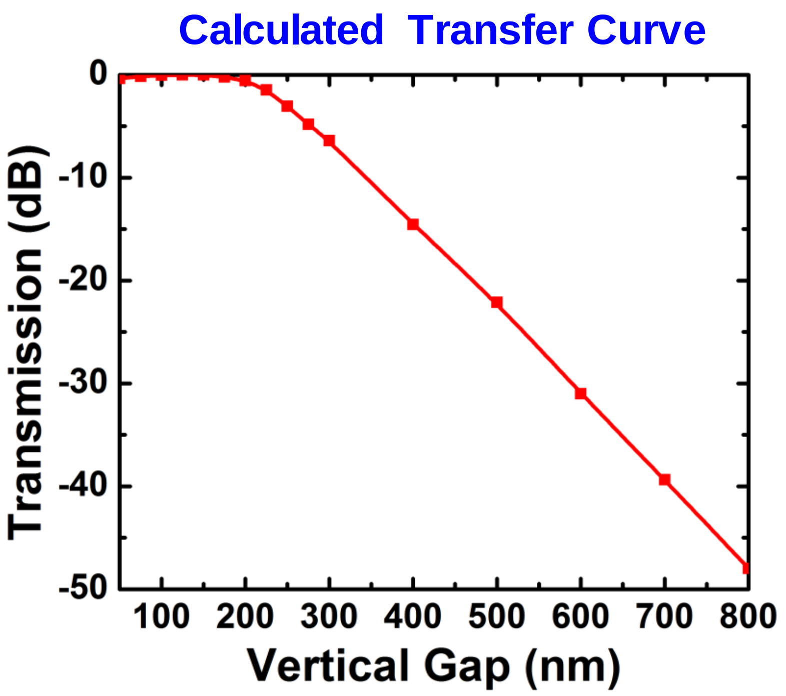
- Single stage switching
- Nearly zero loss in BAR state
- No cumulative loss
- Largest switch demonstrated: 64x64 with < 4dB on-chip loss
- Sub-microsecond switching time

Berkeley UNIVERSITY OF CALIFORNIA

MEMS-Actuated
Vertical Adiabatic Coupler Switch



- Nearly zero loss at OFF state
- Extremely high ON/OFF ratio (60 dB)
- Broadband operation: > 300nm
  - Covers S, C, L bands



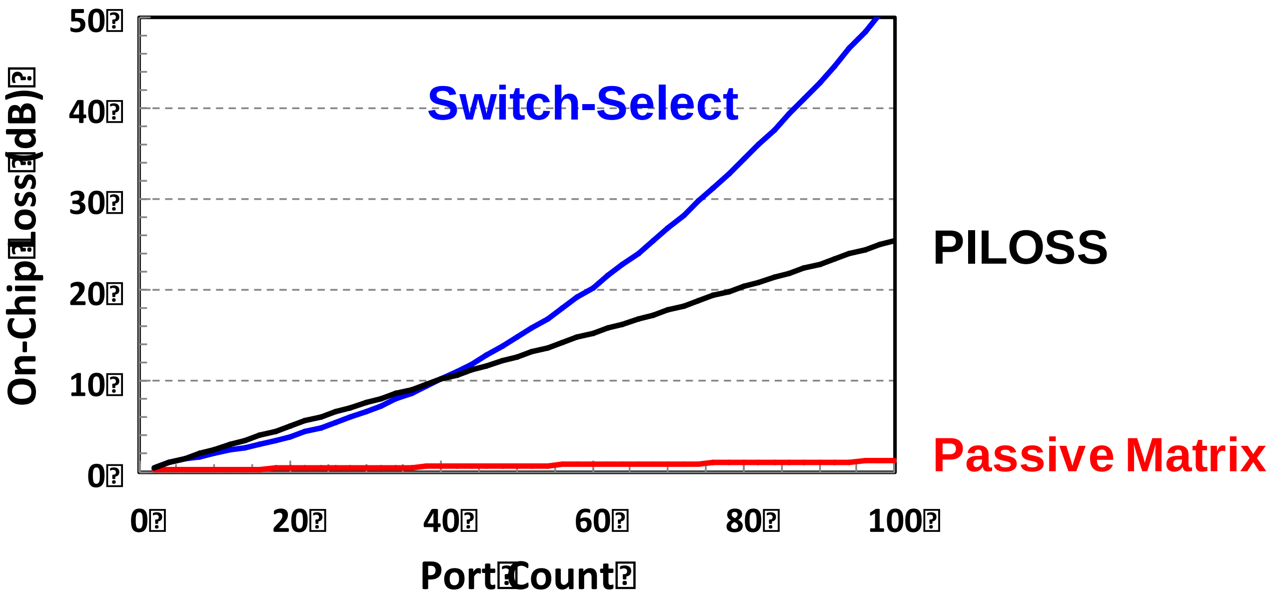
Berkeley UNIVERSITY OF CALIFORNIA

Scalability of 2D Si Photonic Switches

NxN Switch	Number of Switches	Number of Crossing
PILOSS	N	N-1
Switch-Select	2 $\log_2 N$	$(N-1)^2$
Passive Matrix	1	2N-1

Assumption:

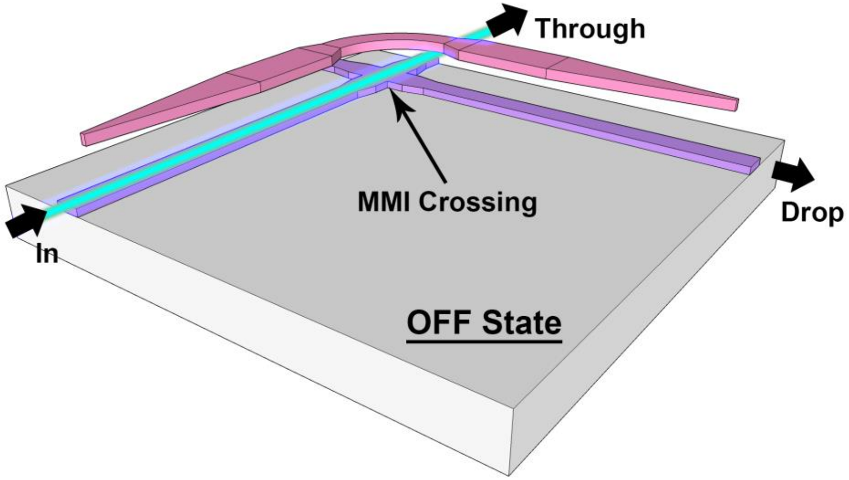
0.25 dB/switch
0.005 dB/crossing

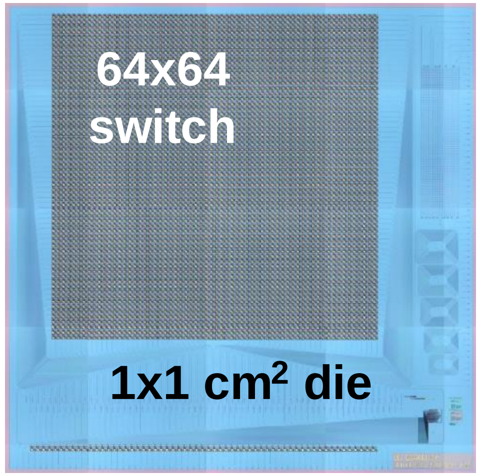


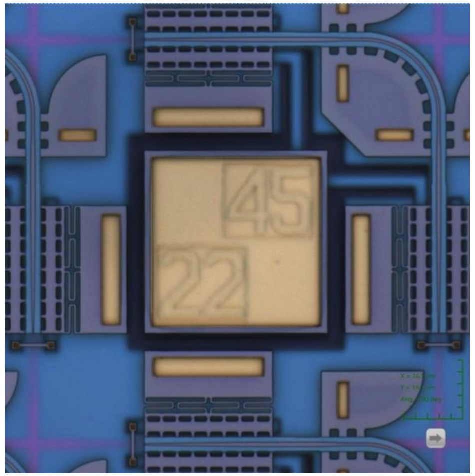
- **On-chip loss < 1.25 dB possible for 100x100 Si Photonic 2D MEMS Switch**

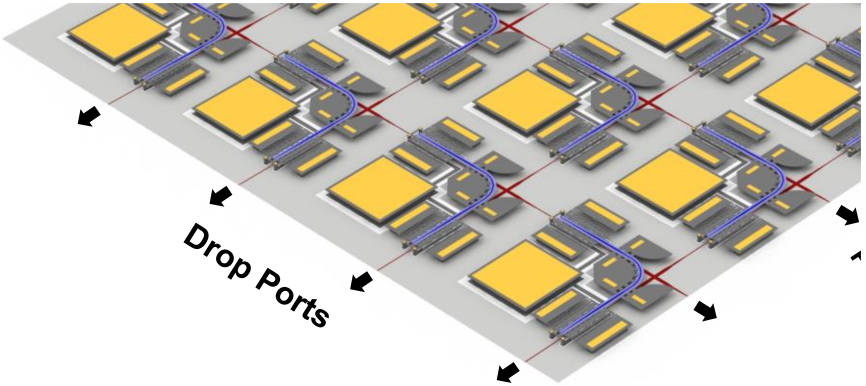
Berkeley
UNIVERSITY OF CALIFORNIA

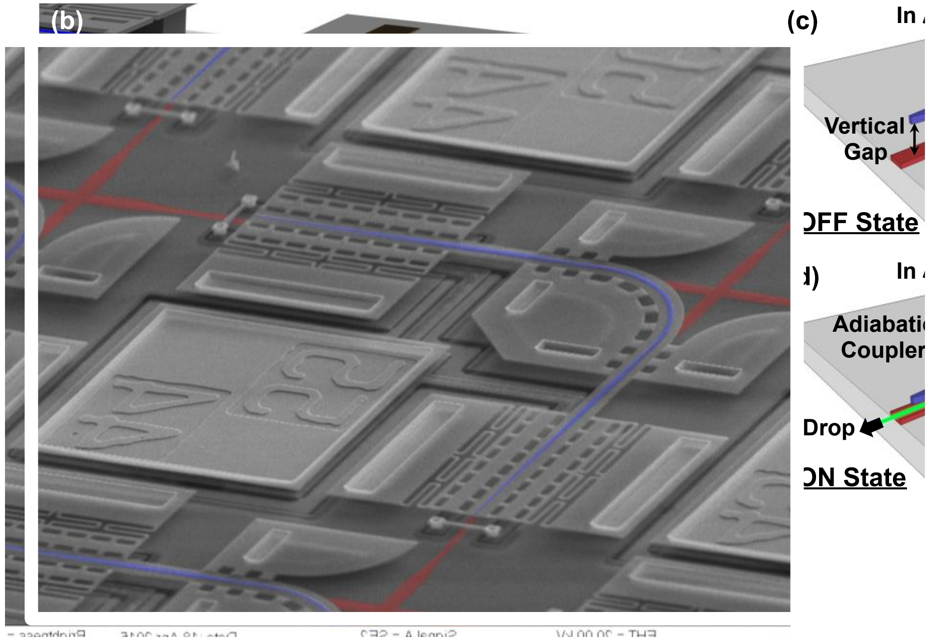
MEMS Crossbar Switch: Experimental Implementation





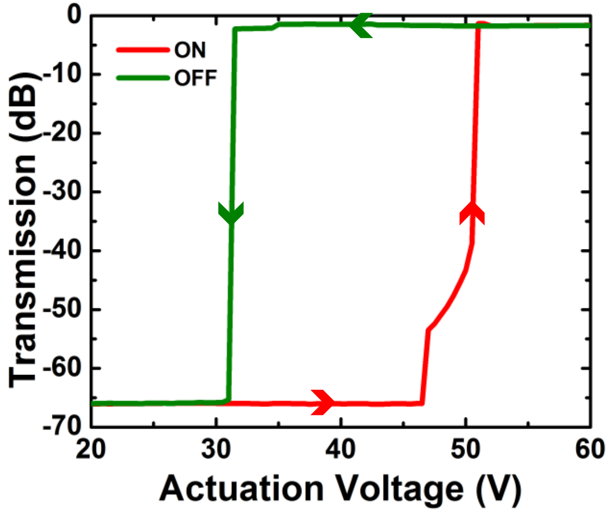




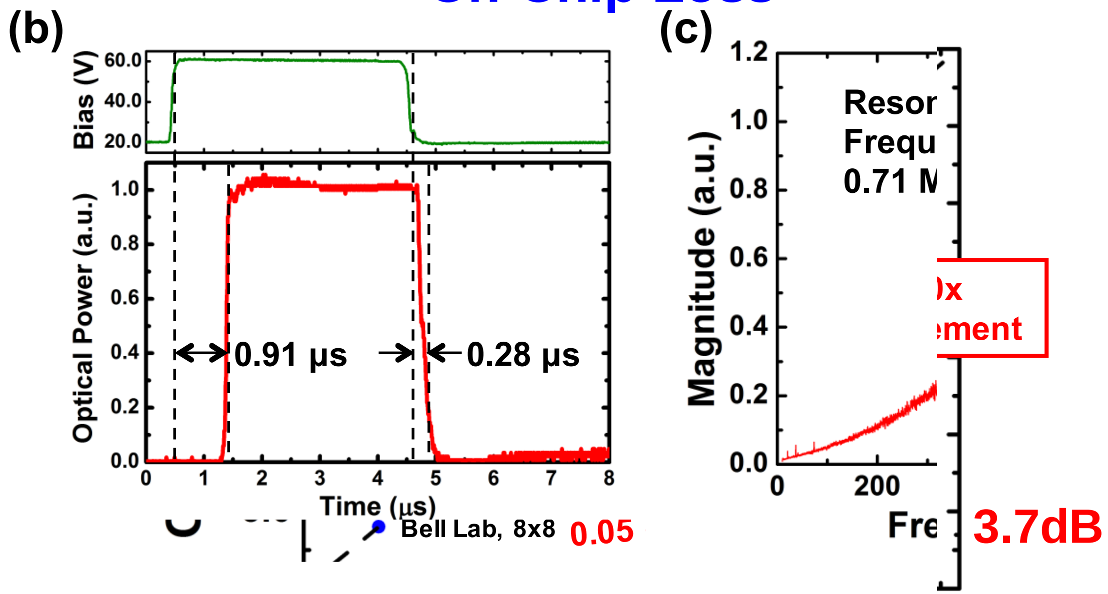


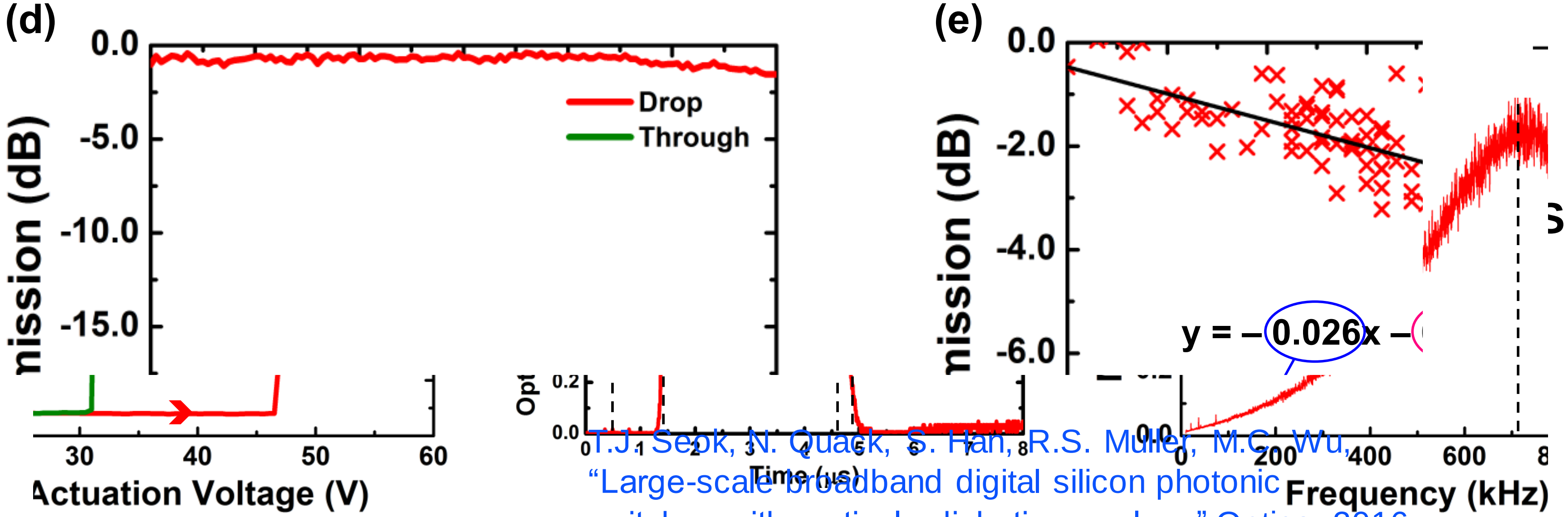
Switch Performance

Digital Switching



On-Chip Loss



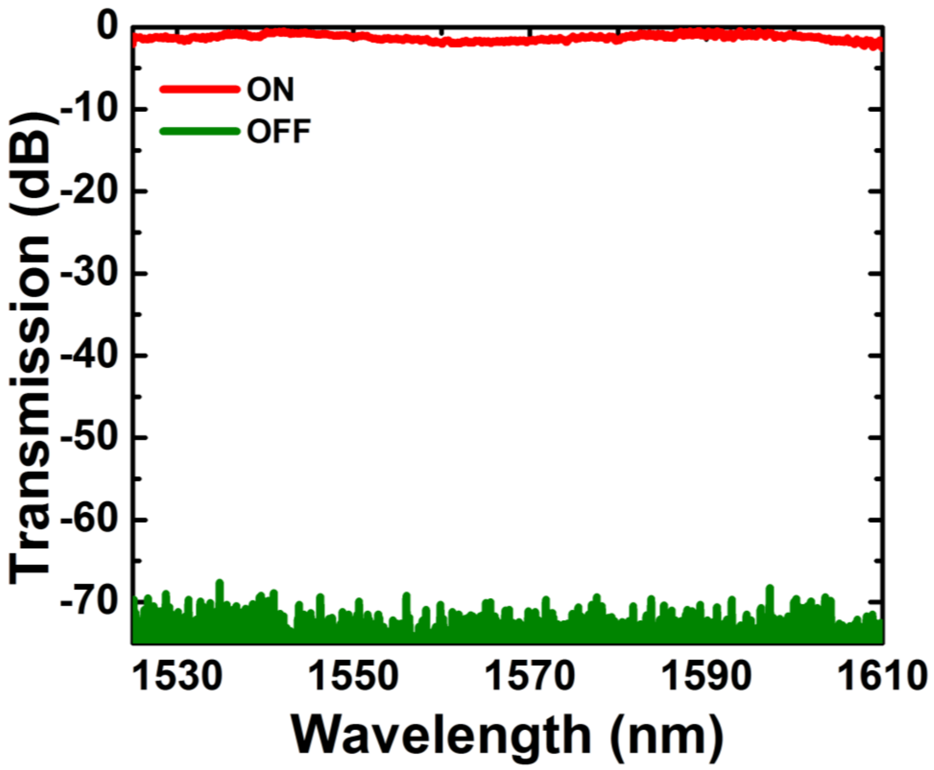


T.J. Seok, N. Quack, S. Han, R.S. Muller, M.C. Wu, "Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers," Optica, 2016

Berkeley
UNIVERSITY OF CALIFORNIA

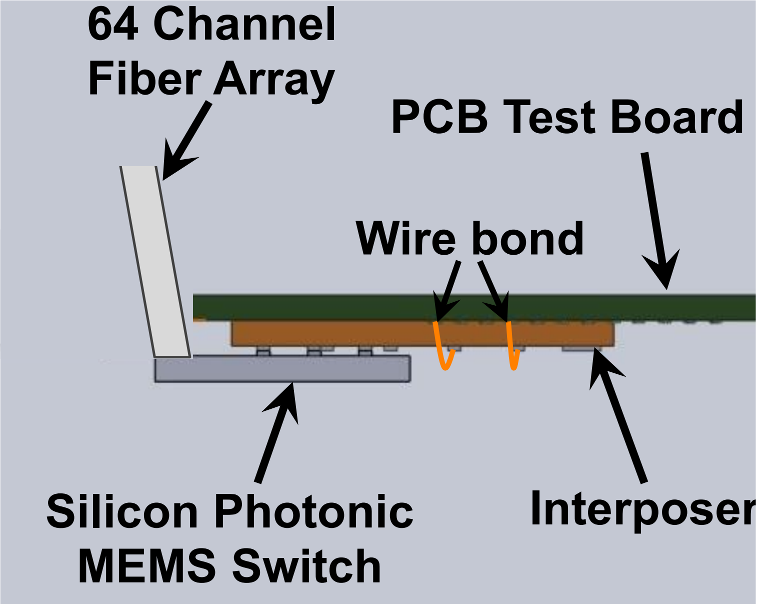
Broadband Operation

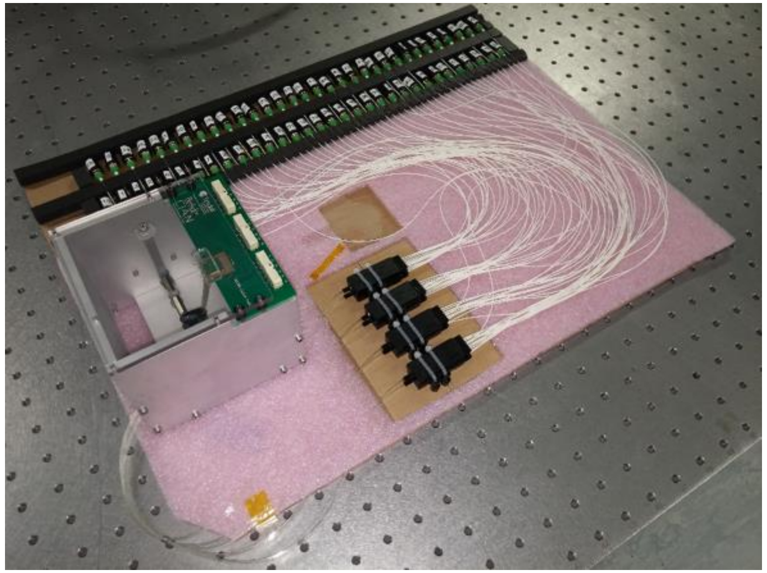
- Measured bandwidth > 120 nm (limited by range of tunable laser)
- Simulated bandwidth > 300 nm

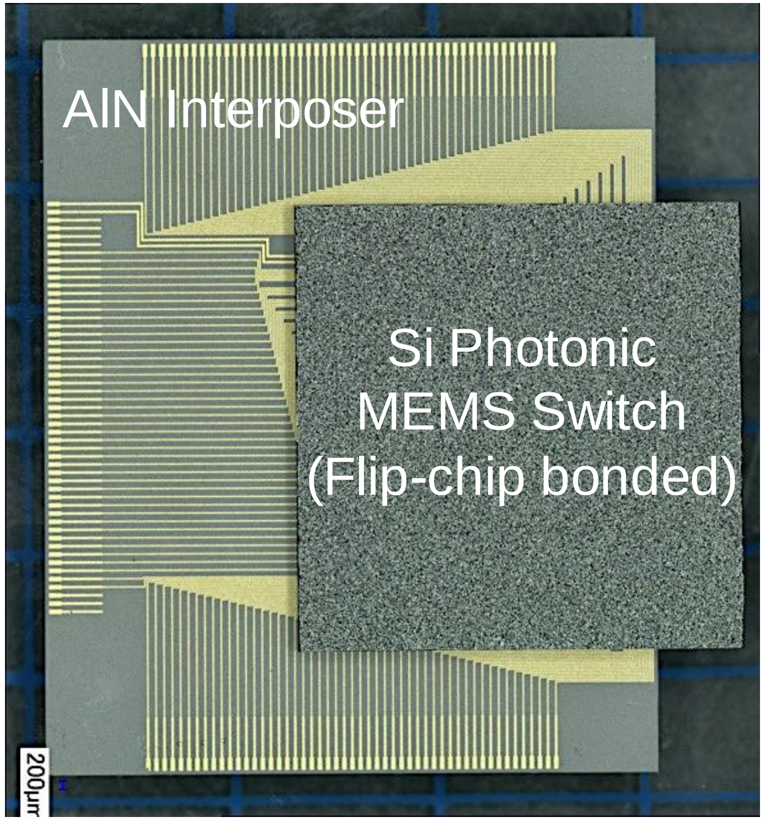


Berkeley
UNIVERSITY OF CALIFORNIA

Switch Packaging

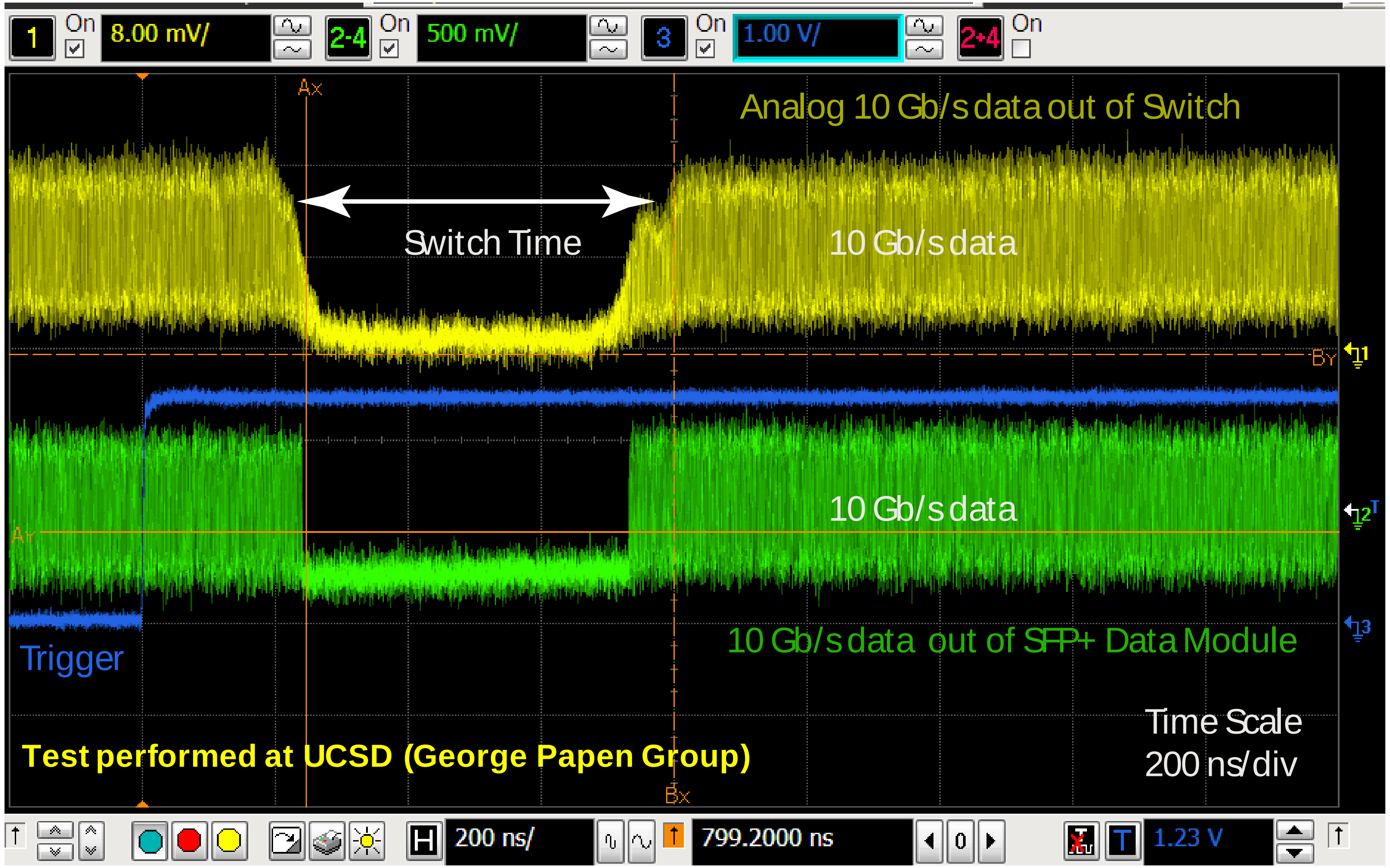






Packaging performed at Tyndall Institute

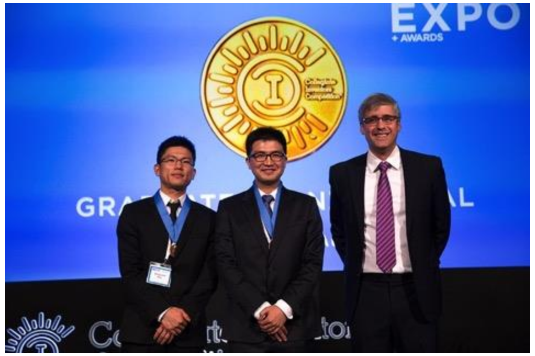
System-Level Testing of Packaged Si Photonic MEMS Switch



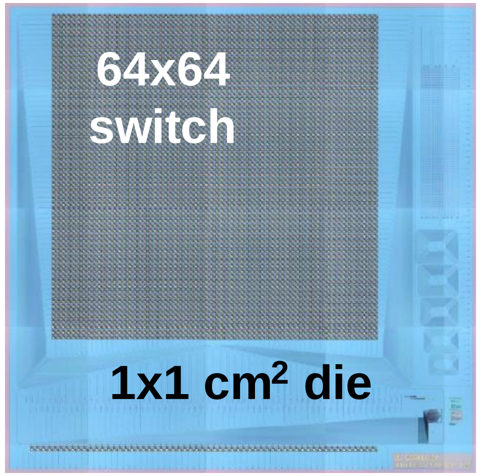
Berkeley
UNIVERSITY OF CALIFORNIA

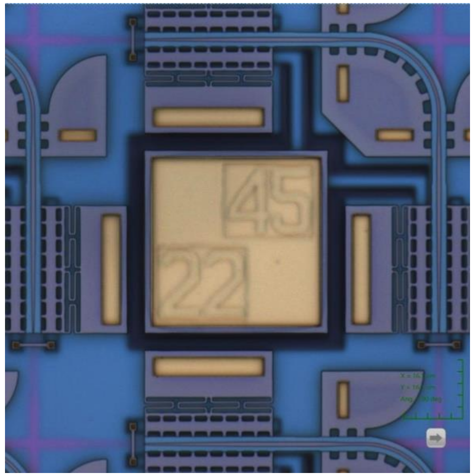
Summary of Si Photonic Switches

- **Highly scalable matrix switch**
- **64x64 switch integrated on 1 cm^2 die**
  - Largest integrated switch in any technology
- **Lowest on-chip loss (3.7 dB, or 0.05 dB/port)**
- **Sub-microsecond switching time**
- **Scalable to 100's of ports**

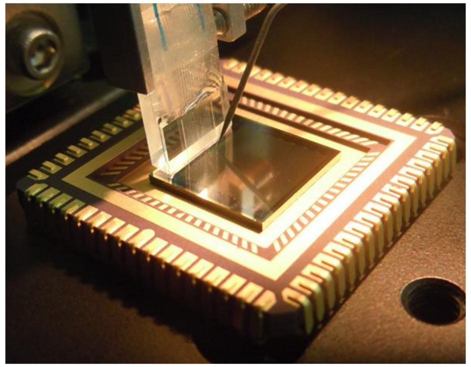


Bronze Medal, 2015 Collegiate Inventors Competition

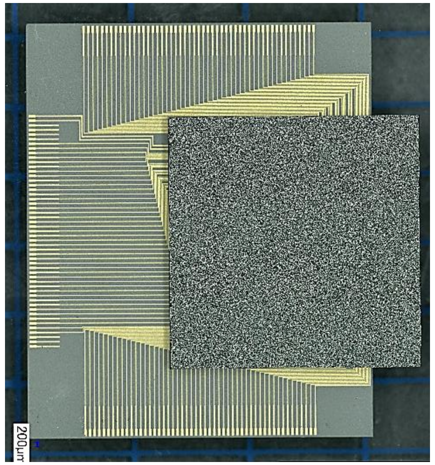




Bench-Top Probing



Tyndall Packaging



- T. J. Seok, N. Quack, S. Han, R. S. Muller, and M. C. Wu, *Optica*, p. 64, Jan. 2016.
- S. Han, T. J. Seok, N. Quack, B.-W. Yoo, and M. C. Wu, *Optica*, p. 370, Apr. 2015.

Berkeley
UNIVERSITY OF CALIFORNIA

Acknowledgment

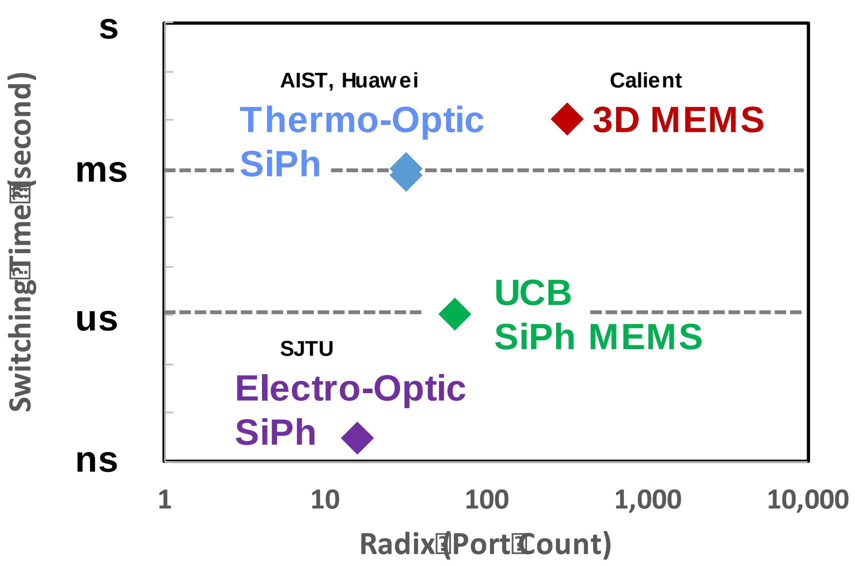
- **Berkeley Switch Team**
  - Dr. Tae Joon Seok, Sangyoon Han, Dr. Niels Quack
  - Prof. Richard Muller
- **Tyndall Institute at Ireland**
  - Drs. Peter O’brien, H.Y. Hwang, J.S. Lee, L. Carroll
- **UCSD**
  - Prof. George Papen
- **Funding support**
  - National Science Foundation (NSF) Center for Integrated Access Network (CIAN) #EEC-0812072
  - Defense Advanced Research Project Agency (DARPA) Electronic-Photonic Heterogeneous Integration (E-PHI) program

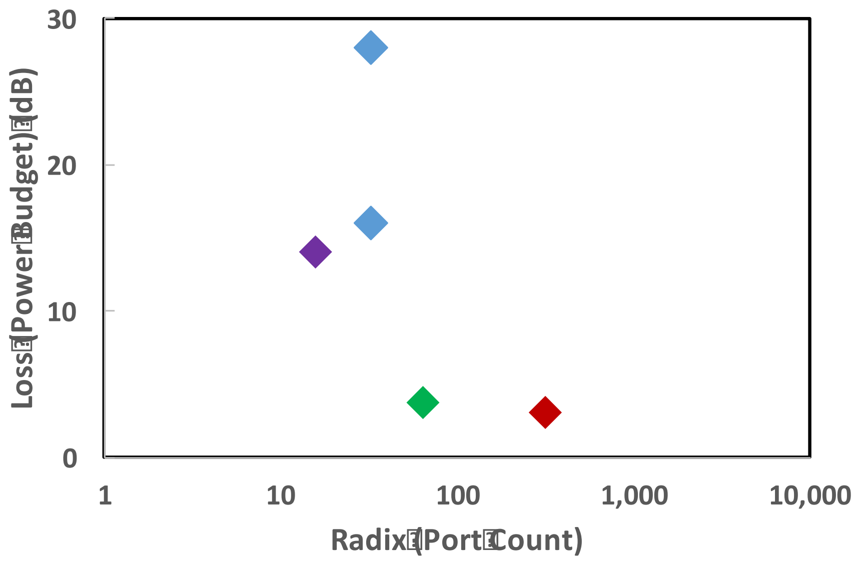
CIAN
CENTER FOR INTEGRATED ACCESS NETWORKS
NSF

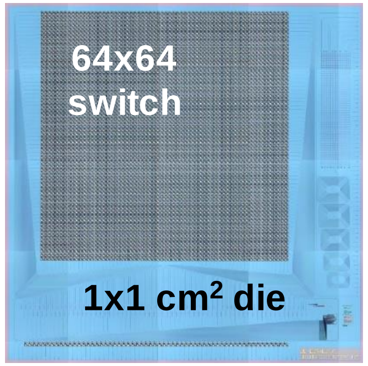
Berkeley
UNIVERSITY OF CALIFORNIA

Additional Slides

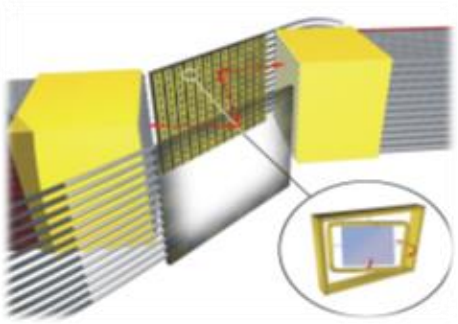
Current and Emerging Optical Space Switches



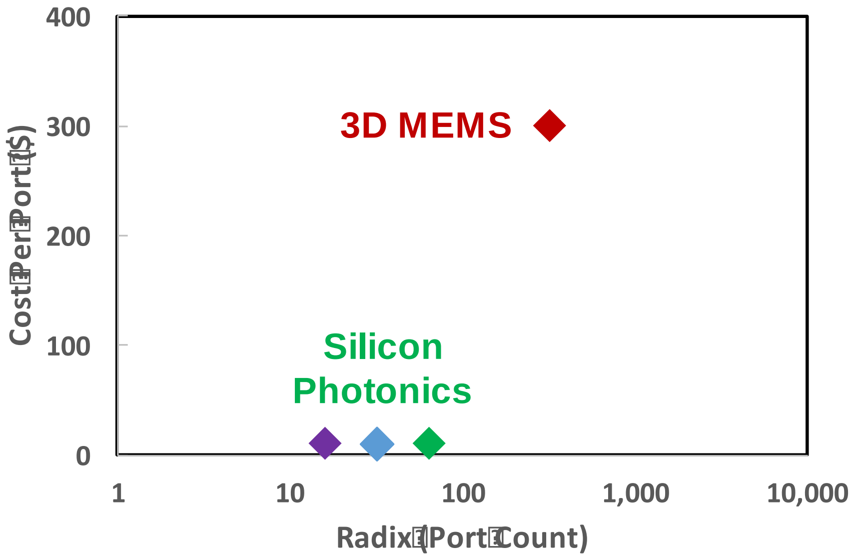




64x64 SiPh Switch
(UC Berkeley)

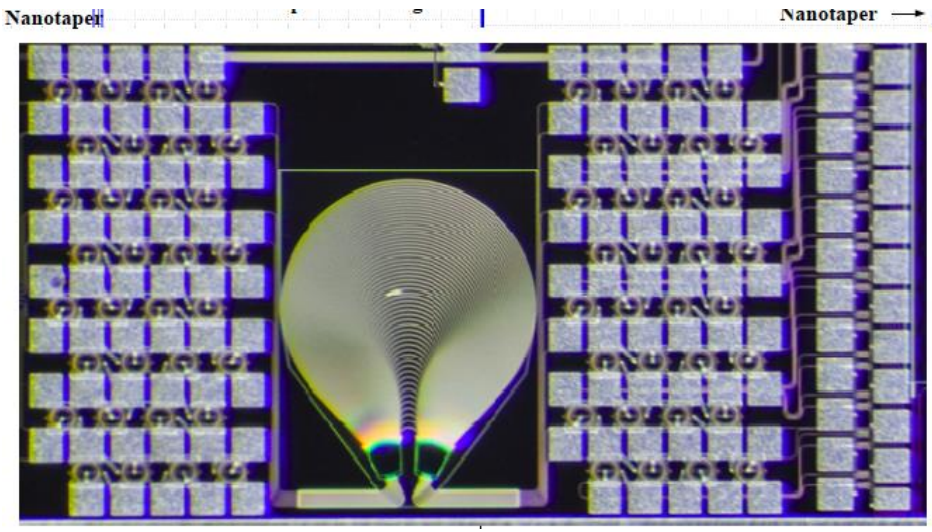


320x320 3D MEMS
(Calient)



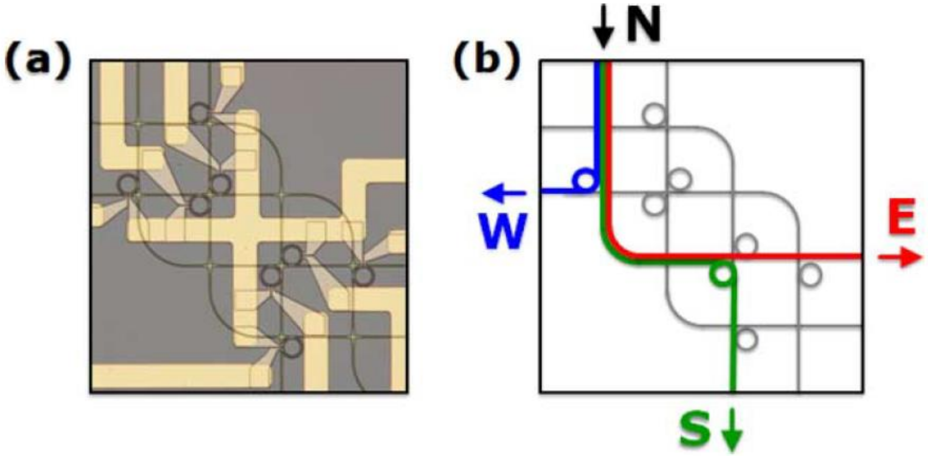
Berkeley
UNIVERSITY OF CALIFORNIA

Wavelength-Domain Switches/Routers

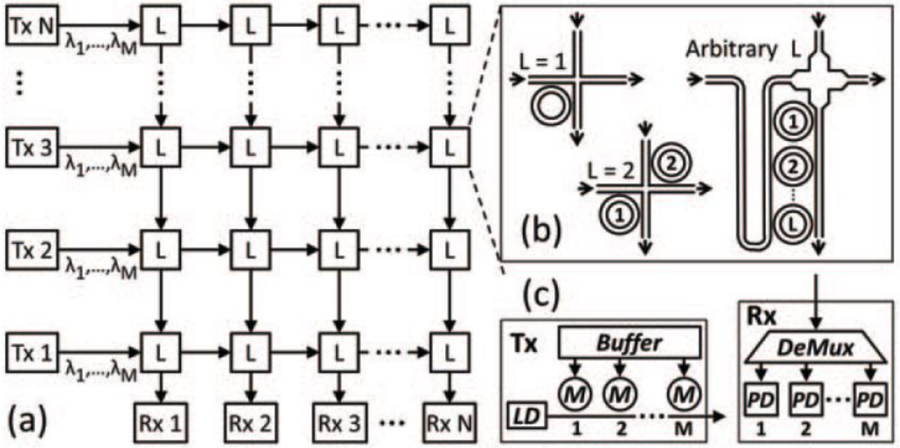


- **Array waveguide grating router (AWGR) with tunable lasers**
- **Microrings with fixed or tunable lasers**
- **(subjects of many papers in this conference)**

R. Yu, S. Cheung, Y. Li, K. Okamoto, R. Proietti, Y. Yin, and S. J. B. Yoo, Optics Express, 2013.



A. Biberman, B. G. Lee, N. Sherwood-Droz, M. Lipson, and K. Bergman, IEEE Photonics Technology Letters. 2010.



A. S. P. Khope, A. A. M. Saleh, J. E. Bowers, and R. C. Alferness, 2016 IEEE Optical Interconnects Conference (OI)

Berkeley
UNIVERSITY OF CALIFORNIA