Augmenting Low-latency HPC Network with Free-space Optical Links

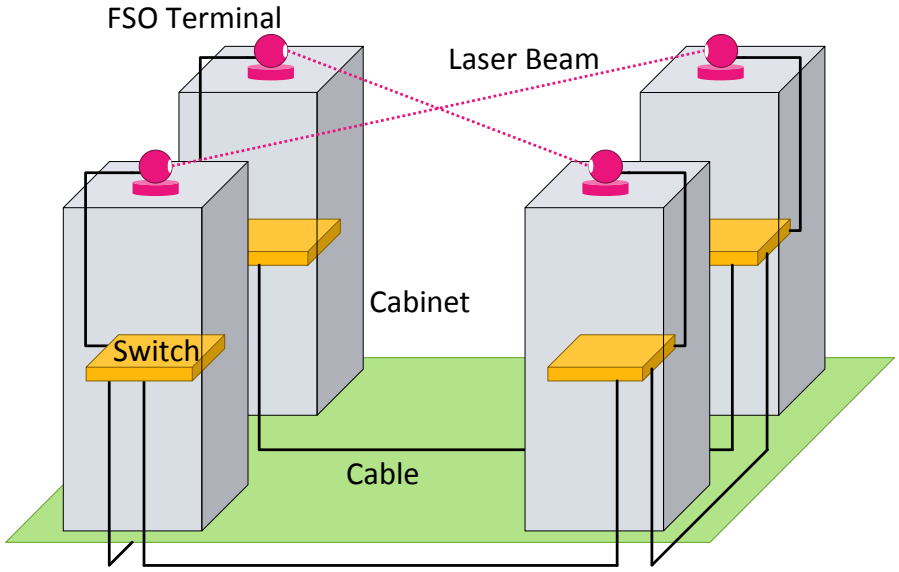
Ikki Fujiwara
Michihiro Koibuchi
National Institute of Informatics

Tomoya Ozaki
Hiroki Matsutani
Keio University

Henri Casanova
University of Hawai'i at Manoa

Story at a Glance

- What if **steerable wireless links** appear on top of cabinets?



Reduced cable length & latency

Topology optimization for diverse apps

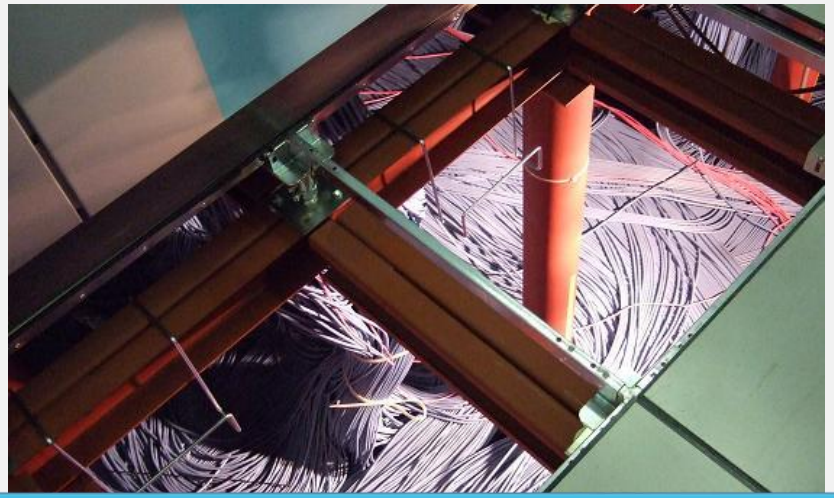
Efficient Power-aware On/Off Link Regulation

- **Motivation**
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - For Improved Topology Embedding
  - For Power-aware On/off Link Regulation
- Conclusion

Motivation 1/3: Cable Reduction

Earth Simulator, 1st gen. (crossbar)

83,200 cables
2,400 km
140 tons



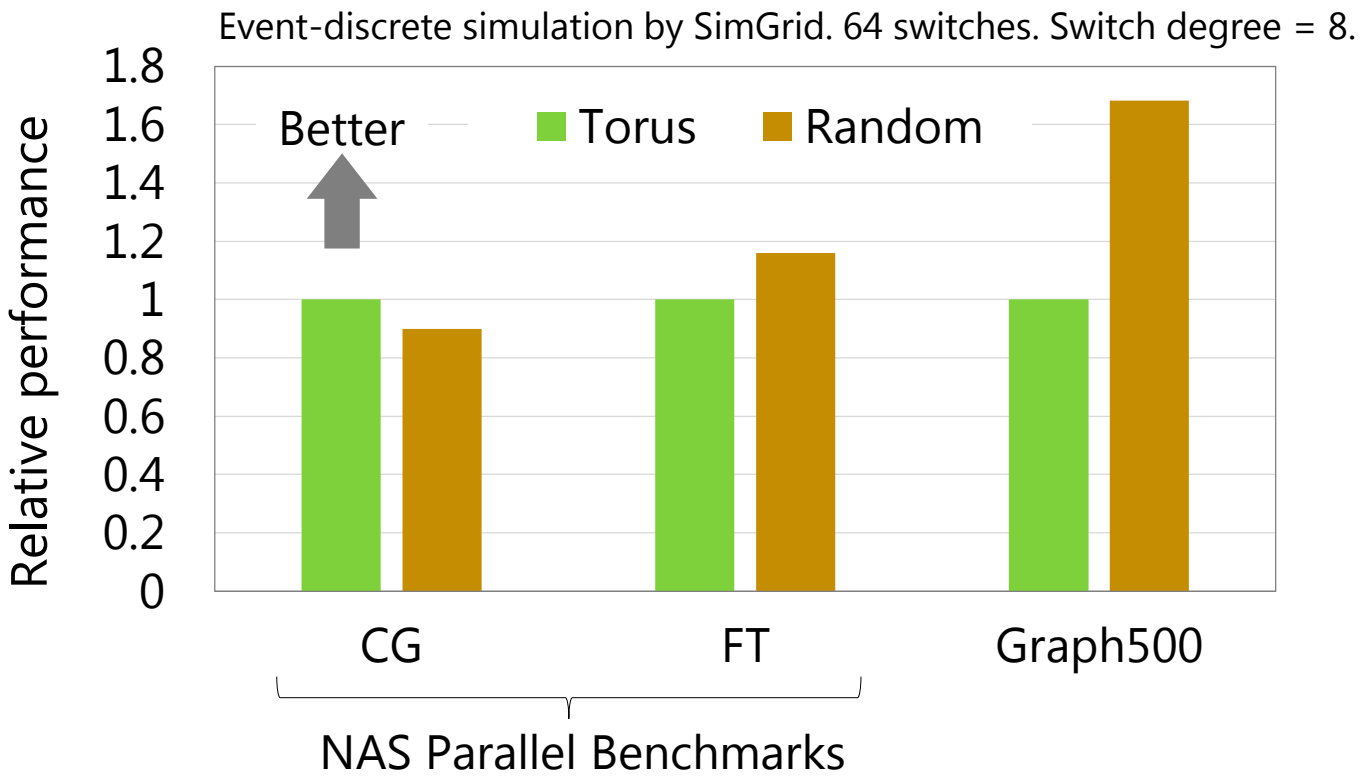
K Computer (6-D mesh/torus)

200,000 cables
1,000 km

FSO provides **shorter cable length** and **lower link delay**

Motivation 2/3: Topology Optimization

- Diverse parallel applications have each different preferable topology



FSO provides a **reconfigurable** network

Motivation 3/3: Leveraging Power-aware On/Off Link Regulation

- Link consumes power regardless of workload

e.g. Energy Efficient Ethernet

- Turned-off links saves link power, but causes a negative impact on performance in HPC use [1]

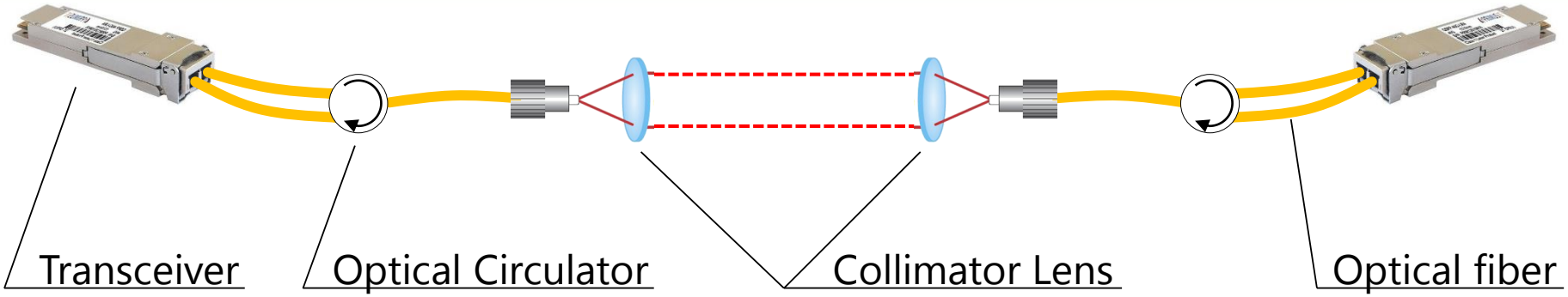
Performance loss is not acceptable for HPCs

- Let's turn off more links!
  - As long as the performance loss is compensated by replacing wired links with FSO-based shortcuts

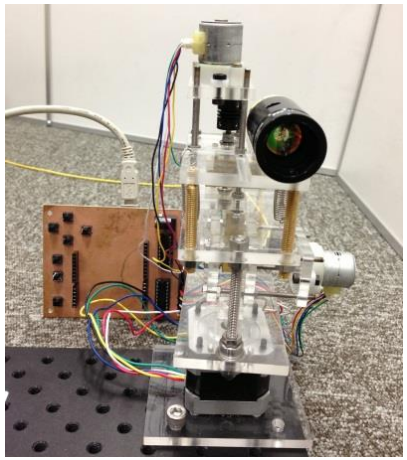
[1] Saravanan et al., "Power/performance evaluation of energy efficient Ethernet (EEE) for High Performance Computing", ISPASS 2013

- Motivation
- How to make Free Space Optics (FSO)
  - **FSO Terminal Devices**
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - For Improved Topology Embedding
  - For Power-aware On/off Link Regulation
- Conclusion

Free-space Optical Links

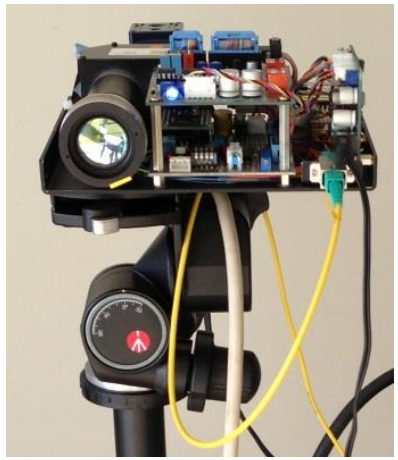


- 10–100 Gbps, 200m distance using commodity laser (e.g. 1310 nm)
- Negligible interference enables high-density layout on top of cabinets
- Terminal devices applicable to HPC use:



Our prototype

Hamedazimi's [2]

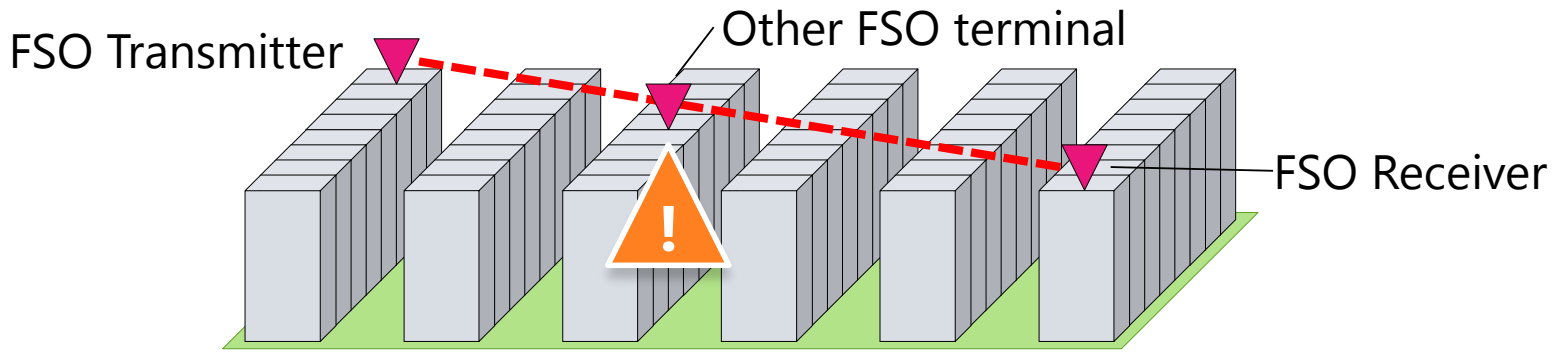


Arimoto's [3]

[2] Hamedazimi et al, "FireFly: a reconfigurable wireless data center fabric using free-space optics", SIGCOMM 2014
[3] Arimoto et al., "Wide field-of-view singlemode-fiber coupled laser communication terminal", SPIE 2013

- Motivation
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - **Layout of FSO Terminals**
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - For Improved Topology Embedding
  - For Power-aware On/off Link Regulation
- Conclusion

Line-of-sight Layout of FSO Terminals

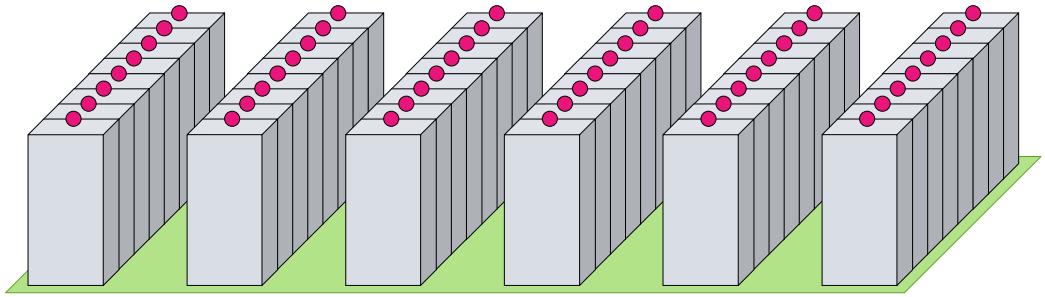


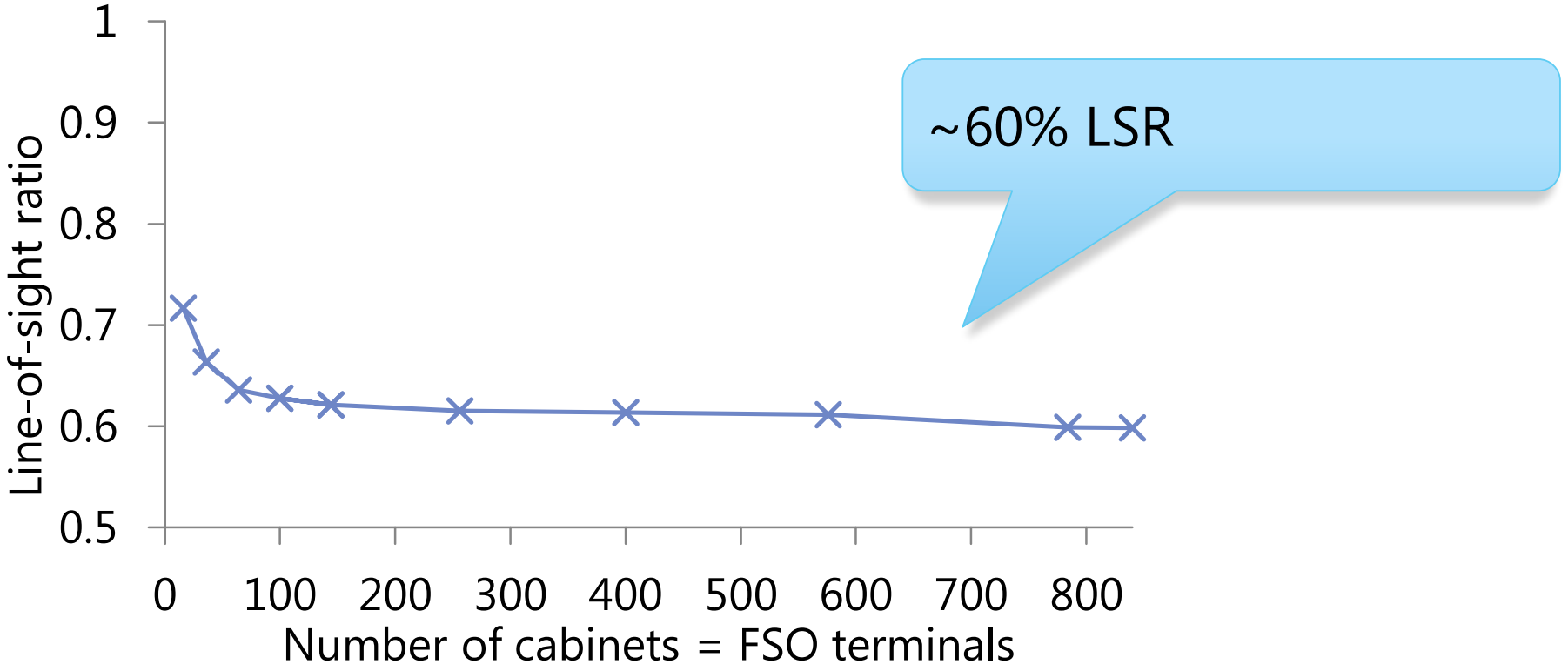
- No laser beam should be interrupted by the other terminals
- Want to layout FSO terminals so as to minimize the interruption

➡ Maximize the line-of-sight ratio (LSR) = $\frac{2L}{N(N-1)}$

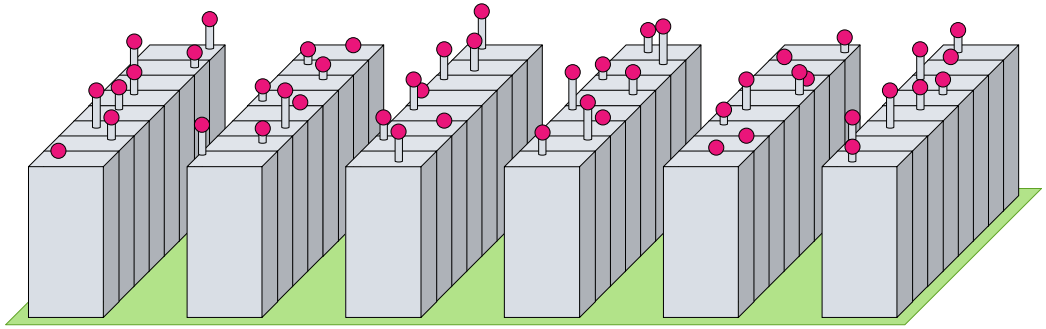
  - N = number of terminals
  - L = number of terminal pairs with direct line of sight
  - Calculated using a ray tracer

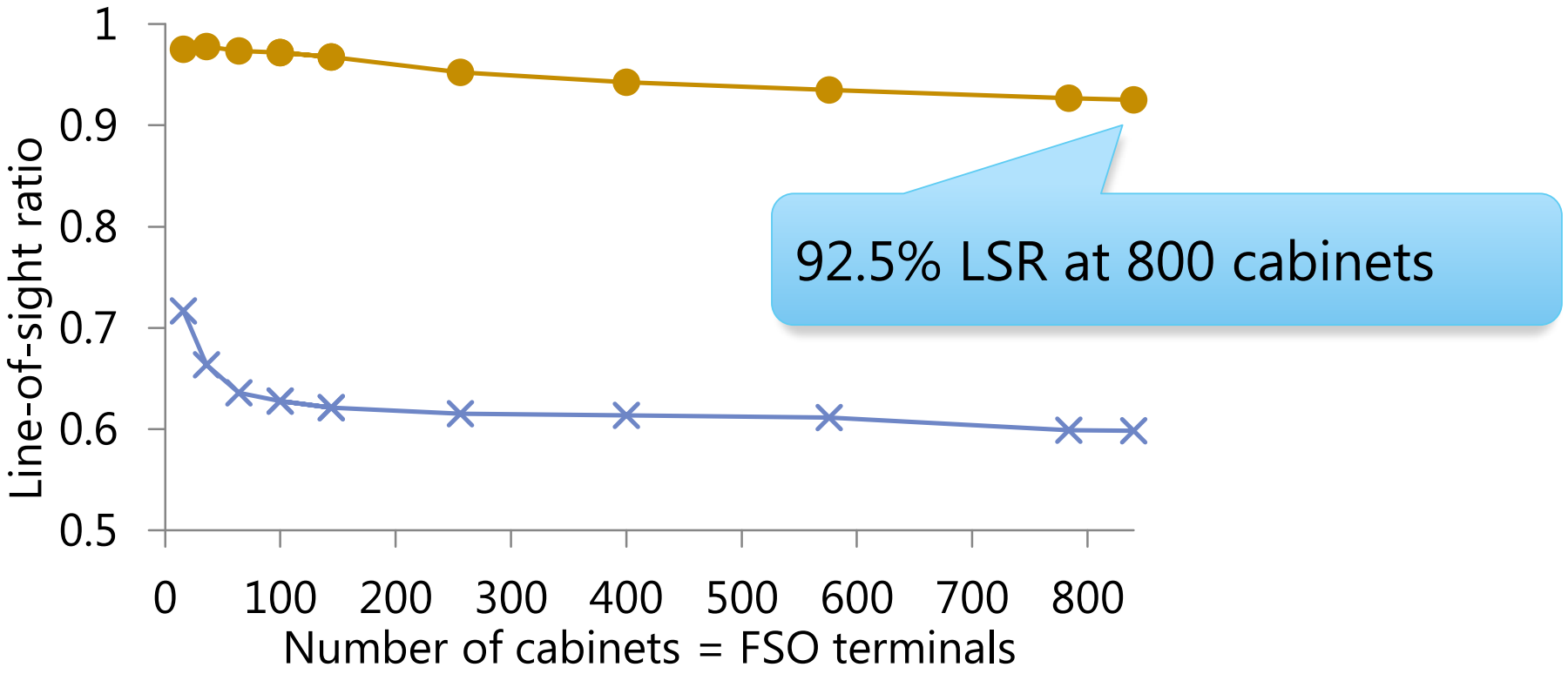
Straight Layout (Naive)



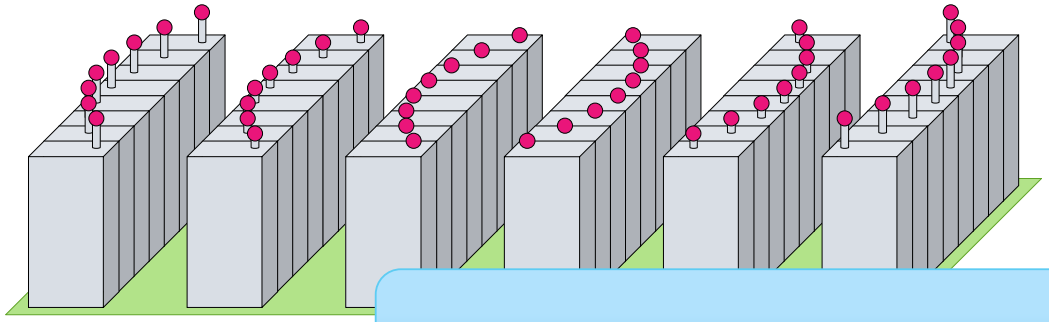


Random Layout

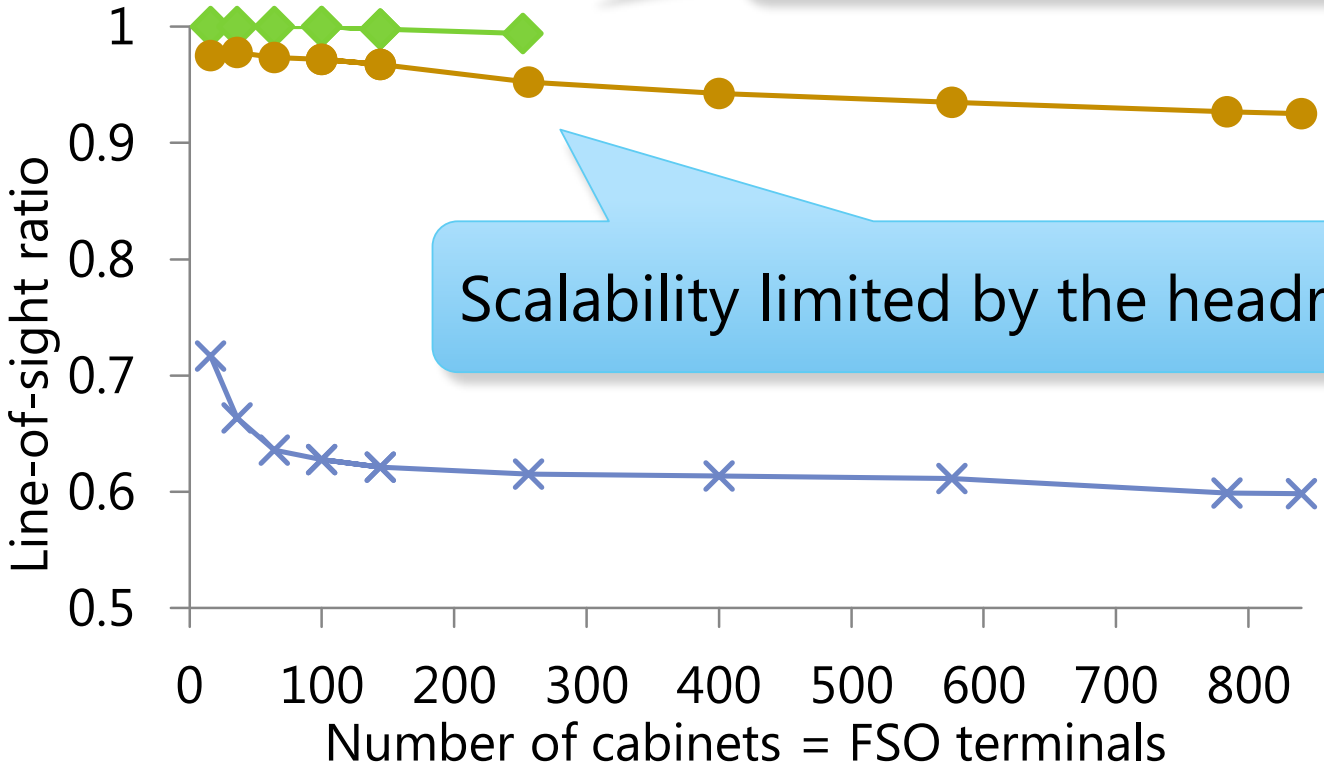




Theater Layout

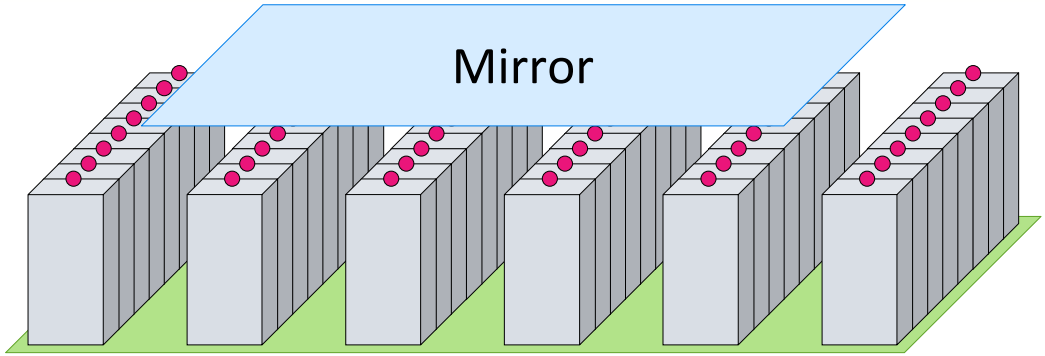


100% LSR up to 252 cabinets



Scalability limited by the headroom

Alternative Layout using a Mirror



100% LSR

Unlimited scalability (ideally)

- FSO beams can be reflected by a mirror
- Similar idea is used for 60GHz wireless [4]
- Hereafter we assume 100% LSR

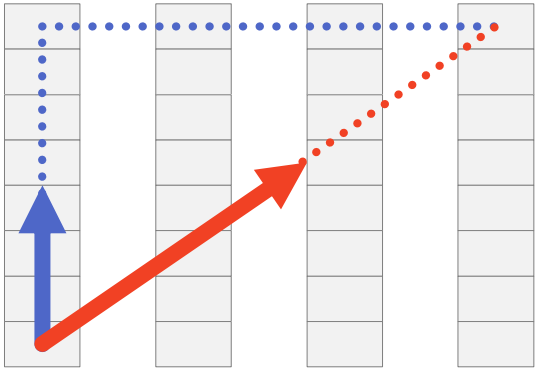
[4] Zhou et al, “Mirror mirror on the ceiling: flexible wireless links for data centers”, SIGCOMM 2012

- Motivation
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - **For Reduced Cable Length and Latency**
  - For Improved Topology Embedding
  - For Power-aware On/off Link Regulation
- Conclusion

Physical Merits of FSO links

- Reduced cable length
- Lower end-to-end communication latency
  - At most 53% lower latency (in theory)

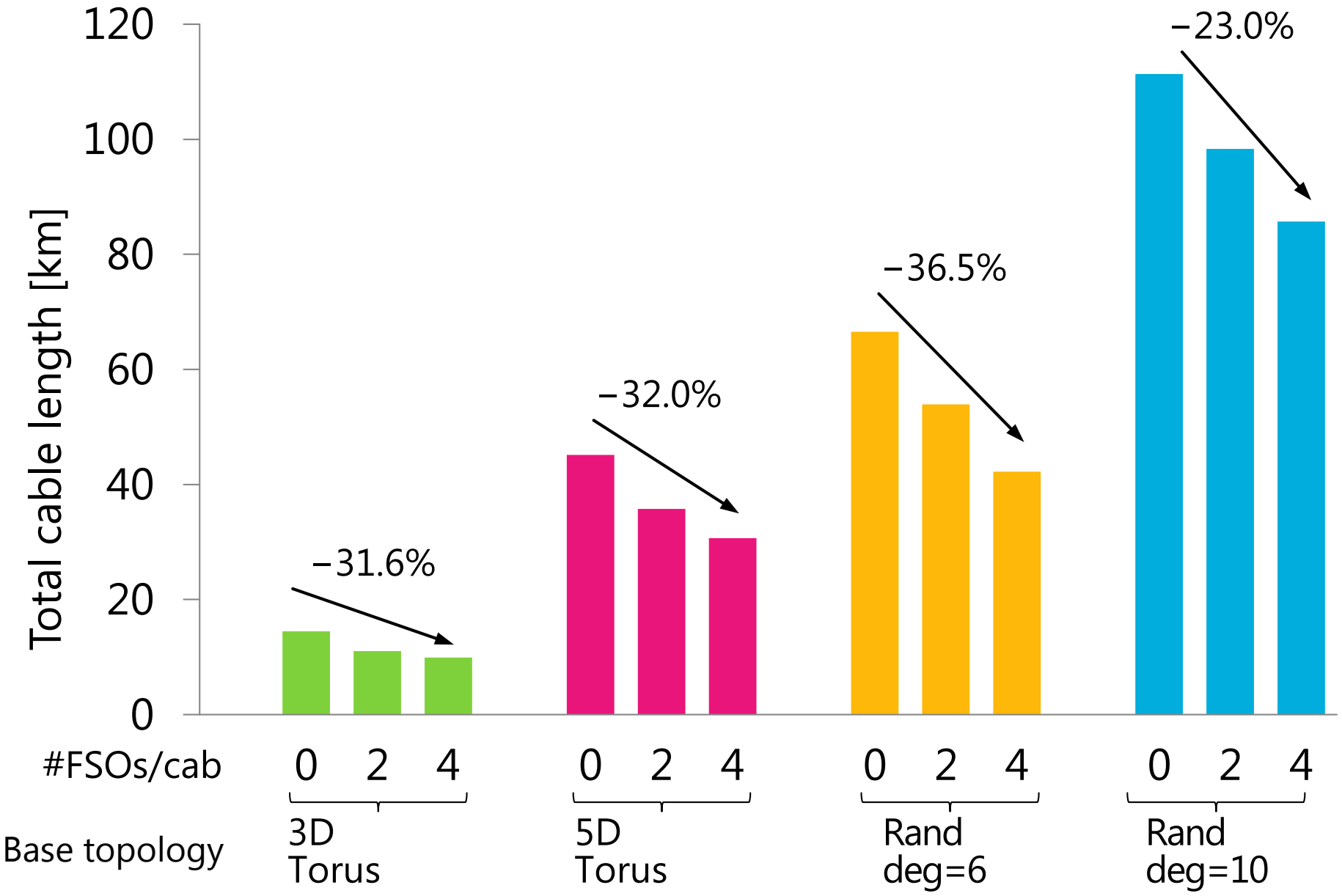
Cable: 0.2 m/ns
Manhattan distance
(Go orthogonal)



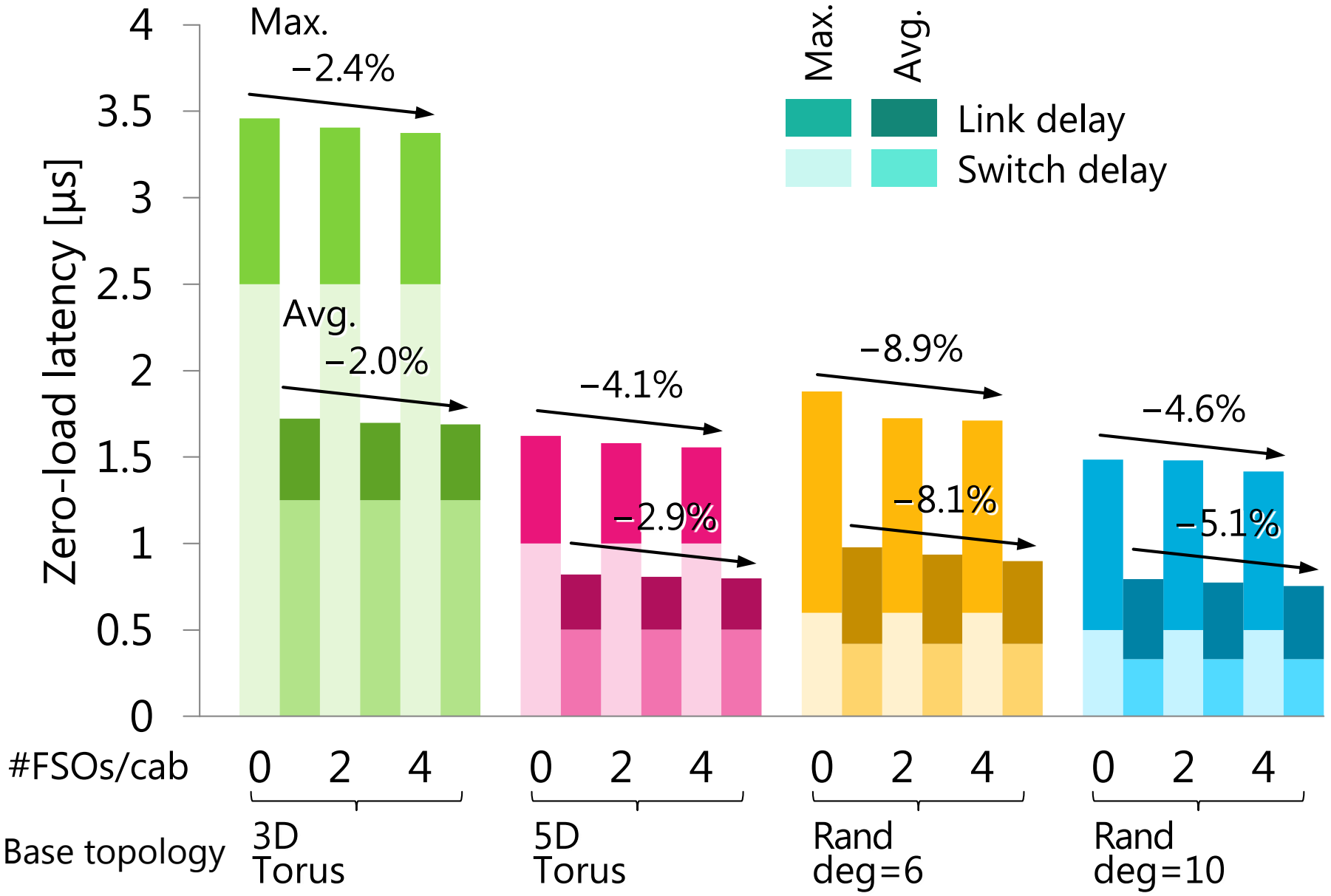
FSO: 0.3 m/ns
Euclidean distance
(Go diagonal)

- Calculated using graph analysis
  - When replacing long cables with FSO links
  - 1,024 switches; 512 cabinets; 1, 2, 4 FSO terminals/cabinet

Reduced Cable Length



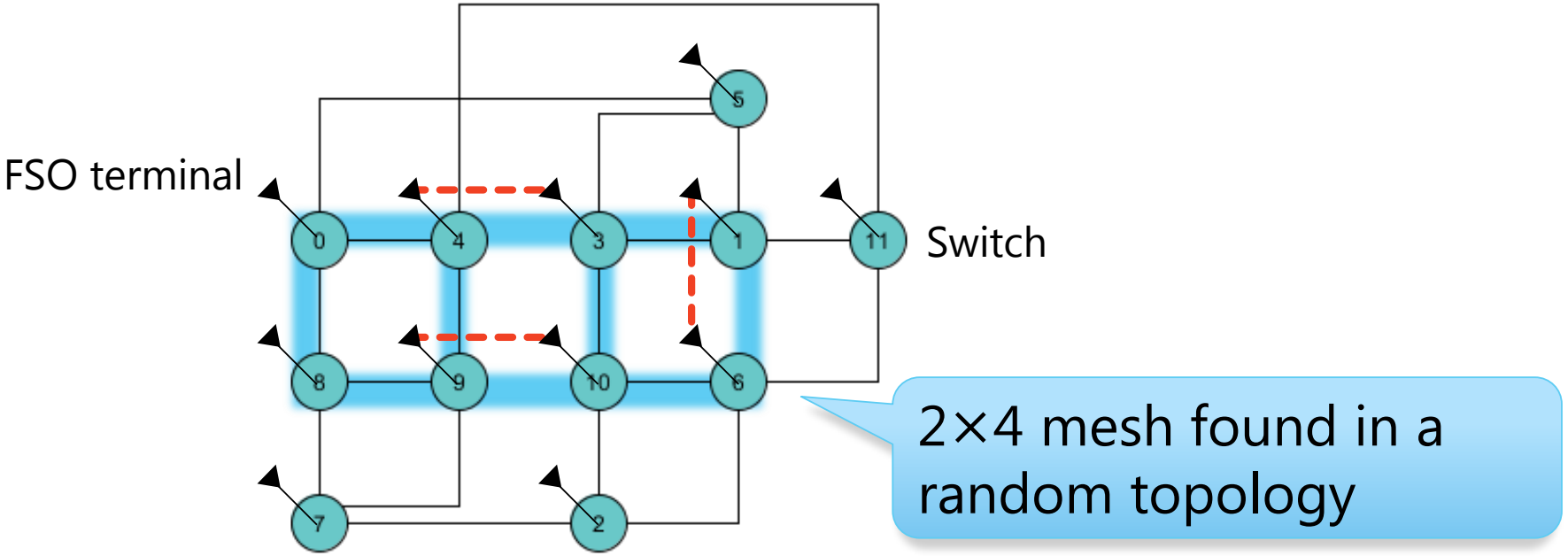
Lower End-to-end Latency



- Motivation
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - **For Improved Topology Embedding**
  - For Power-aware On/off Link Regulation
- Conclusion

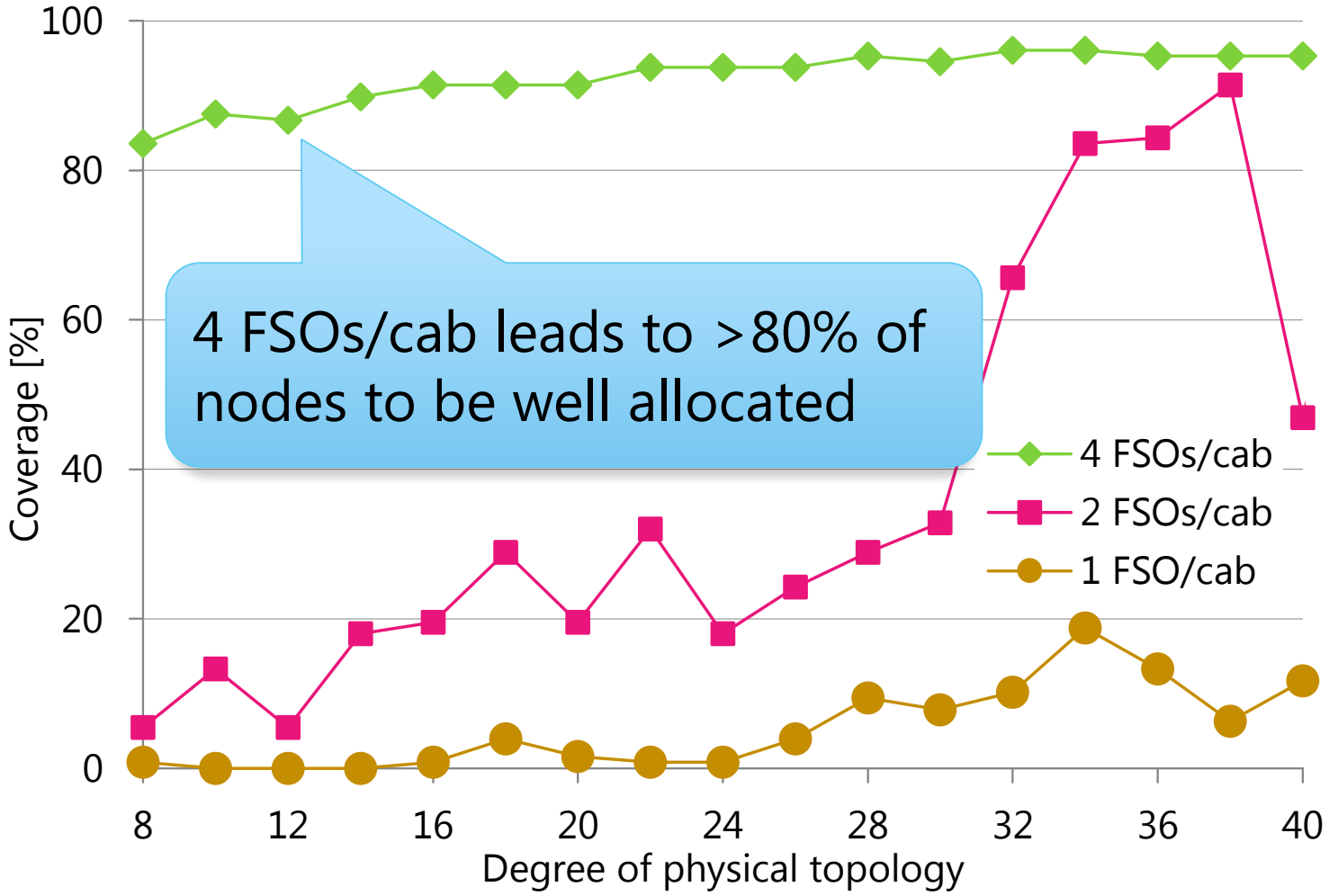
Topology Embedding

- Many small jobs run simultaneously in an HPC
- Want to efficiently allocate their preferable topology
  ➡ Graph embedding problem (NP-hard)
- FSO largely alleviates the embedding problem



- Optimized using a genetic algorithm
  - So as to maximize the number of embedded topology

2×4 Tori Found

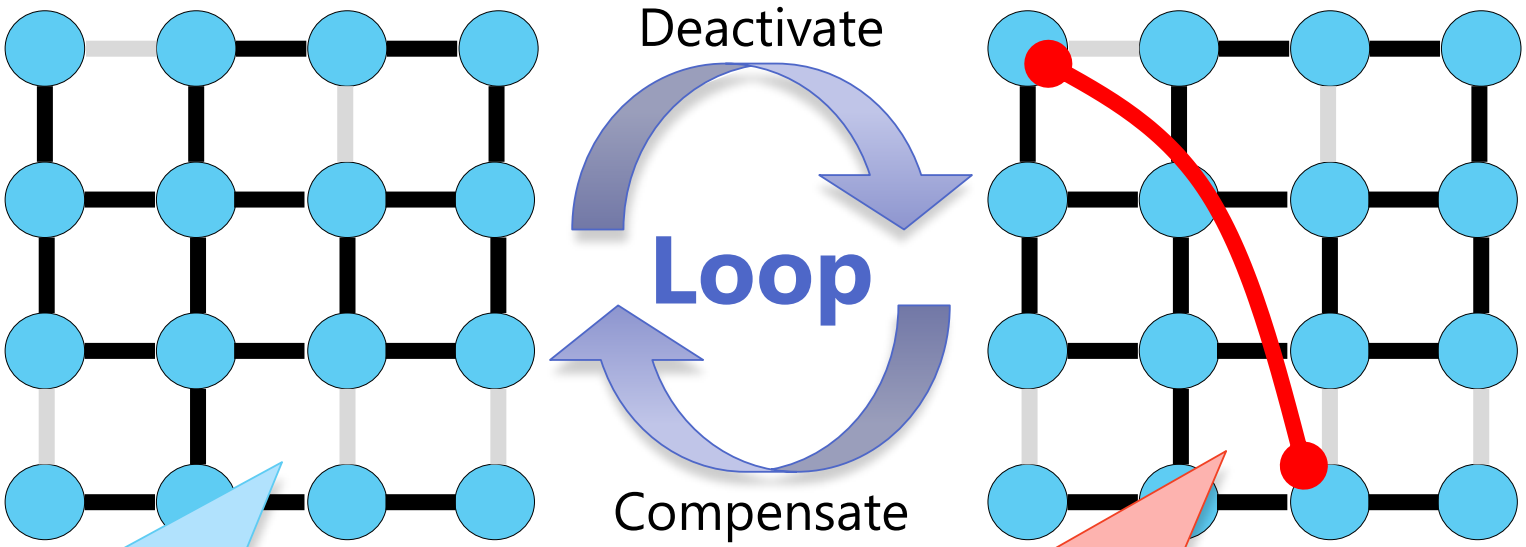


FSO opens a possibility for a better job allocation

- Motivation
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - For Improved Topology Embedding
  - **For Power-aware On/off Link Regulation**
- Conclusion

Power-aware On/Off Link Regulation

- Our idea: let's turn off more links!
  - As long as the performance loss is compensated by replacing wired links with FSO-based shortcuts depending on a given workload

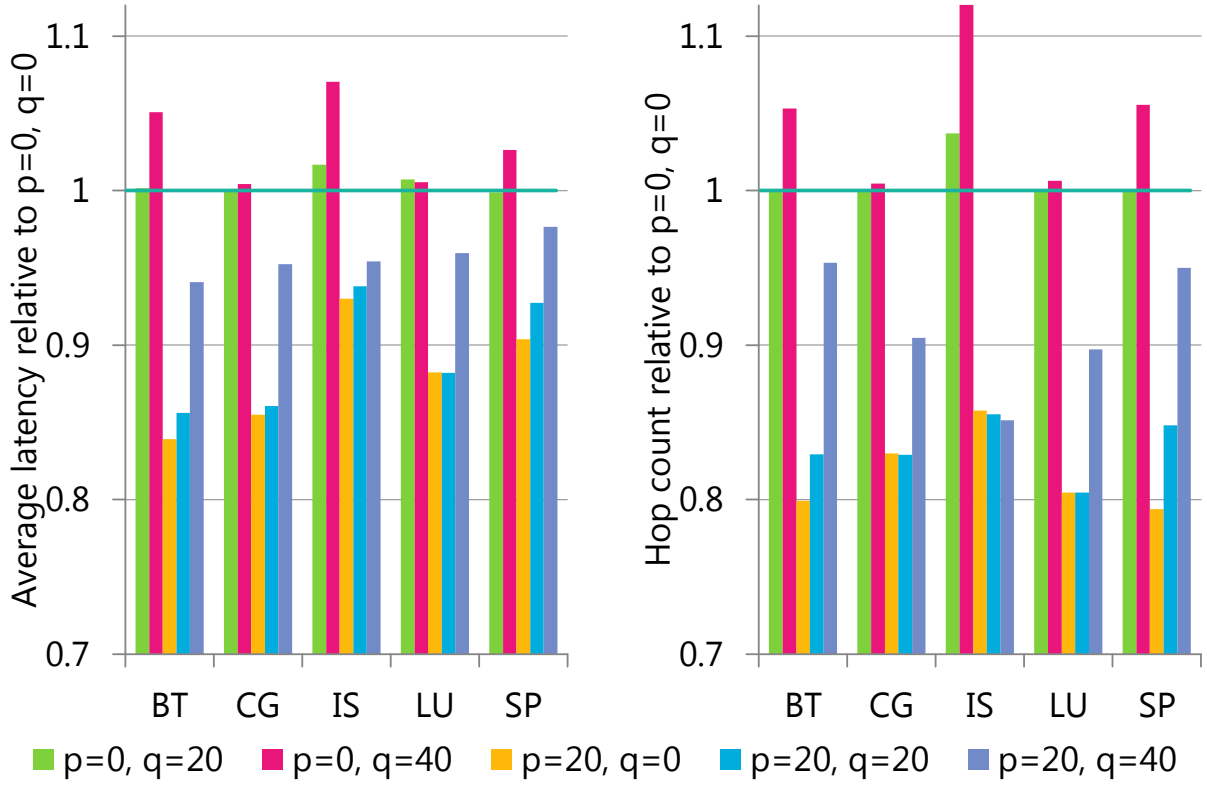


1. Deactivate wired links less contributing to avg path length

2. Insert an FSO shortcut to remedy the avg path length

Power-aware On/Off Link Regulation

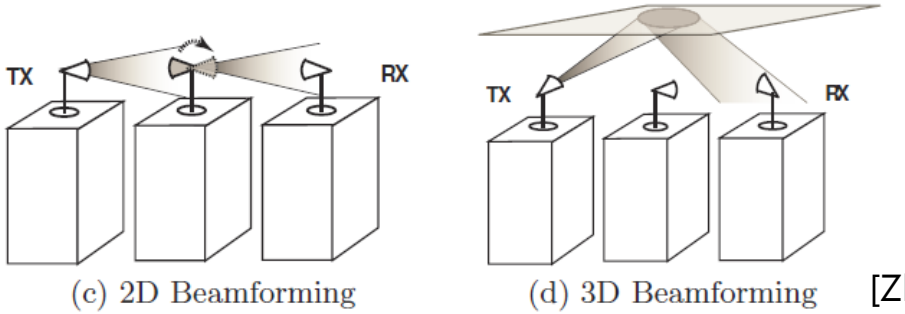
- Evaluation results using flit-level simulator
  - p percent of the wired links are replaced with FSO
  - q percent of the links are deactivated



FSO works well with a power-aware link regulation

Comparable Technologies

- 60 GHz radio wireless links
  - Larger interference than FSO



(c) 2D Beamforming (d) 3D Beamforming [Zhou et al., SIGCOMM 2012]

- Embedding using Optical Circuit Switches (OCS)
  - Wired links via an optical circuit switch can support partial reconfiguration
  - Its embedding capability is lower than FSO

Only FSO realizes our three objectives

- Motivation
- How to make Free Space Optics (FSO)
  - FSO Terminal Devices
  - Layout of FSO Terminals
- How to use FSOs in an HPC system
  - For Reduced Cable Length and Latency
  - For Improved Topology Embedding
  - For Power-aware On/off Link Regulation
- **Conclusion**

Conclusion

- Augmenting Low-latency HPC Network with Free-space Optical Links, we get…

Reduced cable length (−36%) & latency (−9%)

Topology optimization for diverse apps

Torus FatTree Random

Efficient Power-aware On/Off Link Regulation