Impact of Failure on Interconnection Networks for Large Storage Systems

Qin Xin, Ethan L. Miller, Thomas J. E. Schwarz*, Darrell D. E. Long

Storage Systems Research Center
University of California, Santa Cruz

*Department of Computer Engineering
Santa Clara University

UC Santa Cruz

Problems

- ◆ Reliability is a real concern for large storage systems
  - A large number of components
  - Complex interconnections
  - Human errors
- ◆ Robust network interconnection is desired
  - Failures of switches, routers, and network links are common
  - Organization of nodes and links is crucial
    - *Scale* makes things complicated: petabyte storage system -- 10,000s disks, 100s routers,100,000s links
  - Robust network topology
    - Tolerant multi-points of failures without noticeable performance degradation

Contributions

- **Examine various network topologies**
  - Butterfly networks
  - Mesh structure
  - Hypercube
  - Tradeoffs among them: cost, robustness, and performance.
- **Estimate the impacts of link and node failures by simulation**
  - I/O path connectivity
  - Number of hops in varied degraded modes
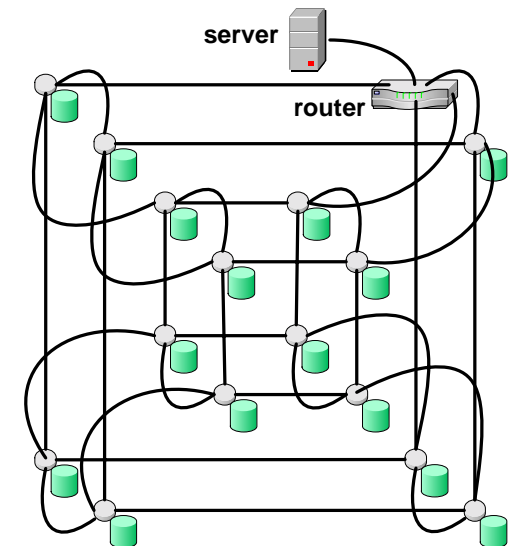 - Failure impact on its neighborhood
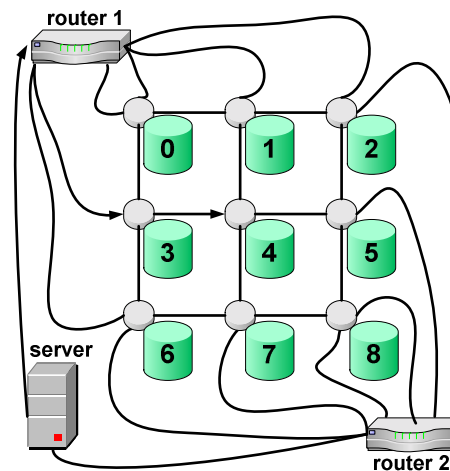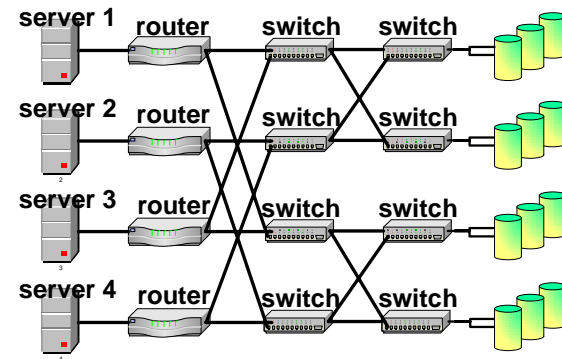
Network Topology for Storage Systems

- ◆ Butterfly networks
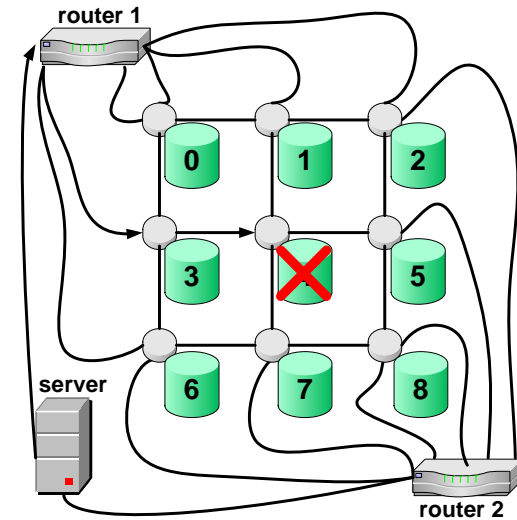 - Reasonable cost
  - Poor robustness
- ◆ Mesh
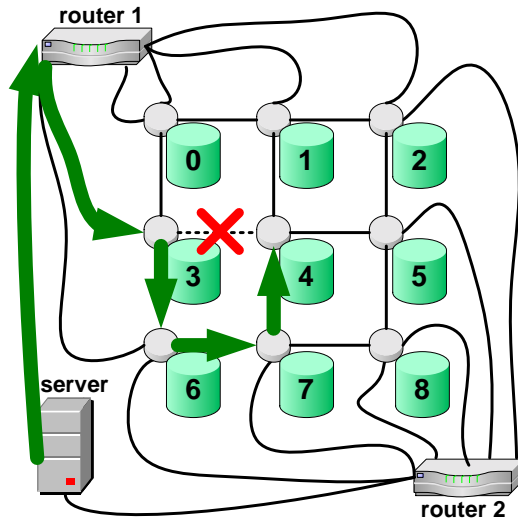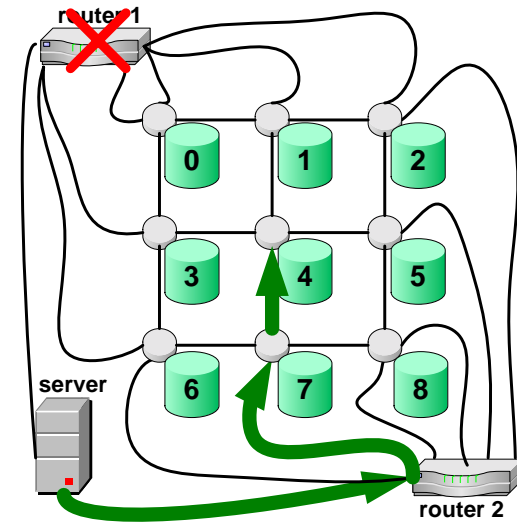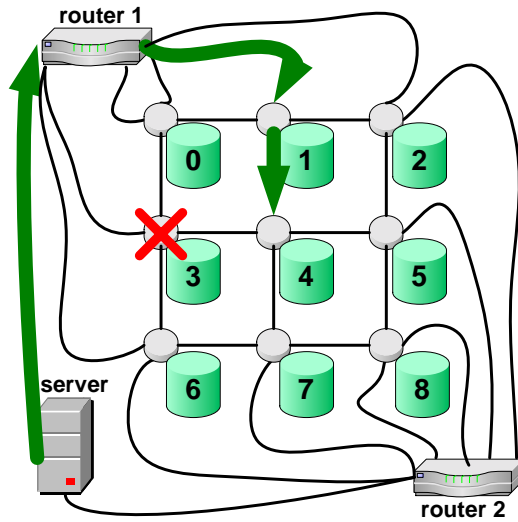 - Brick-structured, good robustness
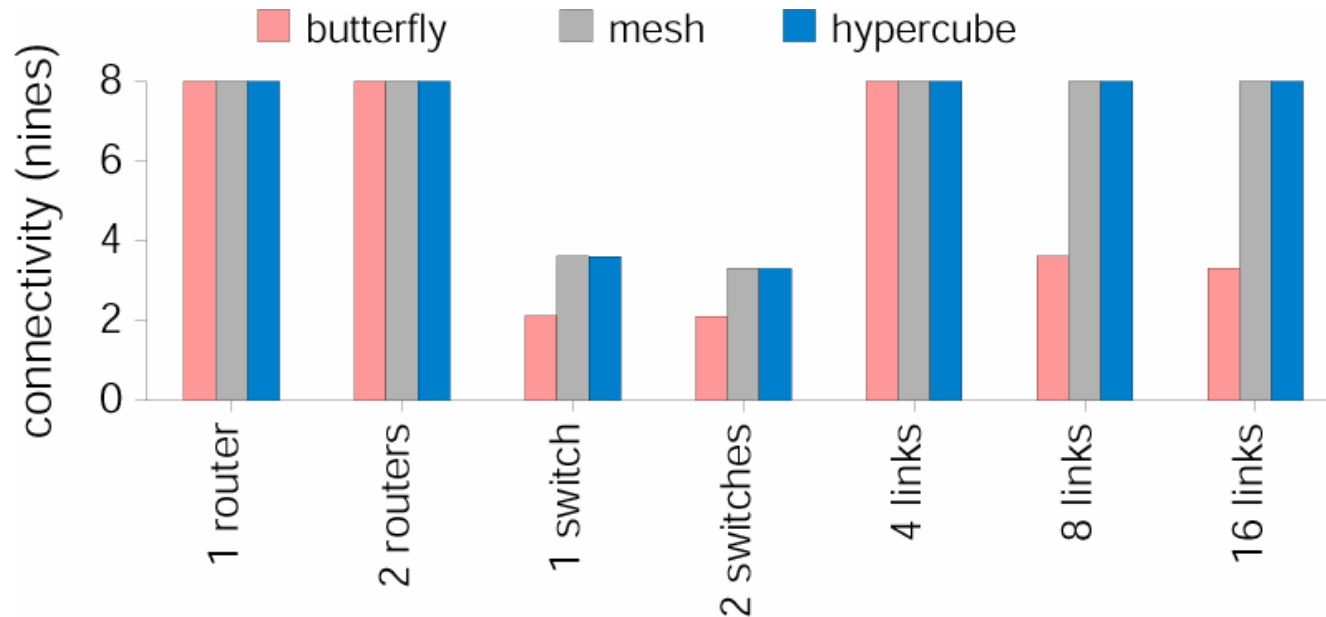 - Relatively long path
- ◆ Hypercube
  - Superior robustness
  - Complex, costly

A case study: Failure Scenarios
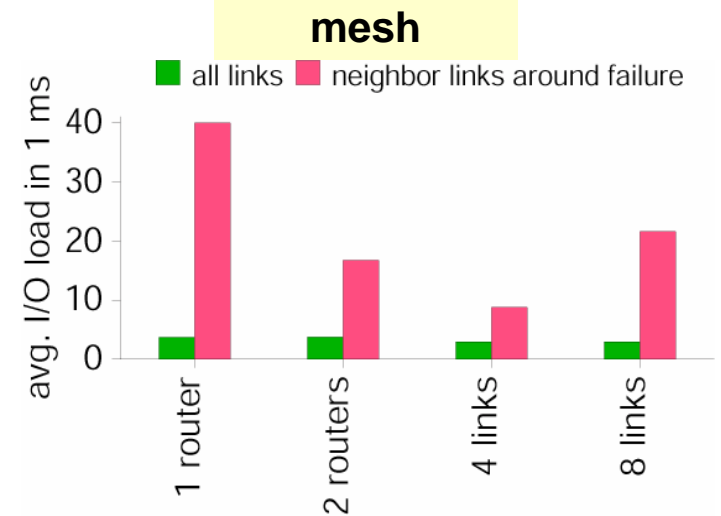
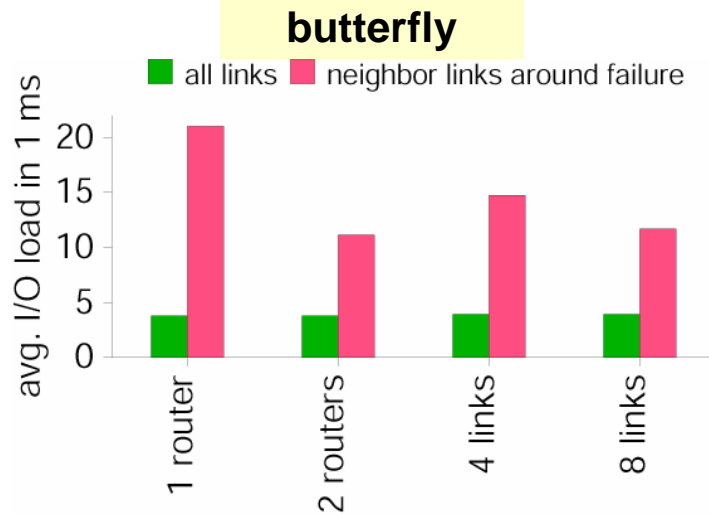# Simulated Results: Connectivity

- ◆ I/O path connectivity
  - Measured in "nines": $-\log_{10}(1-P)$, P: m/n, m: # of good requests, n: # of total requests. (3 nines = 99.9%)
  - Poor connectively when the connection nodes are lost.
  - Hypercube and mesh are more robust in connectivity than butterfly networks.

Simulation Results: Neighborhood

butterfly

mesh

hypercube

◆ Failure impact on network neighborhood

- Neighbors suffer in presence of failures.
- Hypercube outperforms butterfly & mesh structure in avg. I/O load.
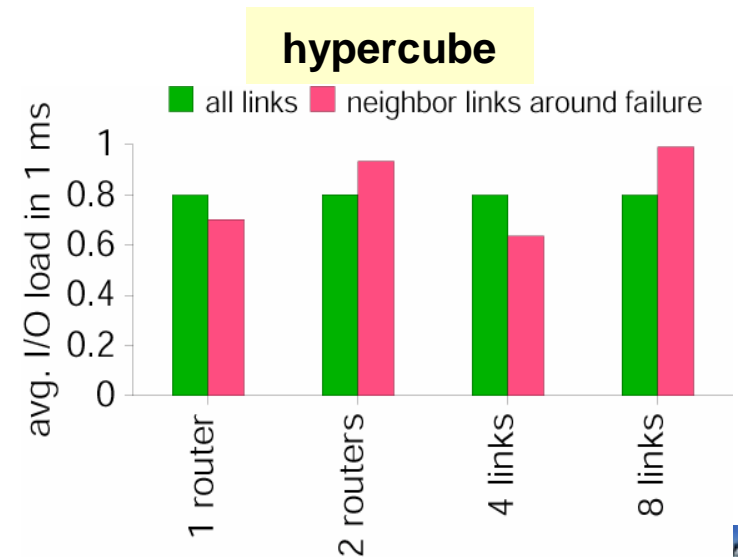
Summary

- A well-chosen topology can provide robust interconnection.

- Neighbors around failures suffer much more than average.

- Hypercube structure provides better interconnection in degraded modes than butterfly and mesh structure.

Questions?

- For further information, please visit:
  - http://ssrc.cse.ucsc.edu/
  - http://www.cs.ucsc.edu/~qxin/

- Thanks to SSRC faculty, students, affiliates, and alumni
- Thanks to our sponsors
  - Lawrence Livermore National Laboratory, Los Alamos National Laboratory, Sandia National Laboratory
  - Engenio Information Technologies, HP Lab, Hitachi Global Storage Technologies, IBM Research, Intel, LSI Logic, Microsoft Research, Network Appliance, and Veritas
  - NSF, USENIX

Big Picture: A Petabyte-Scale Storage System