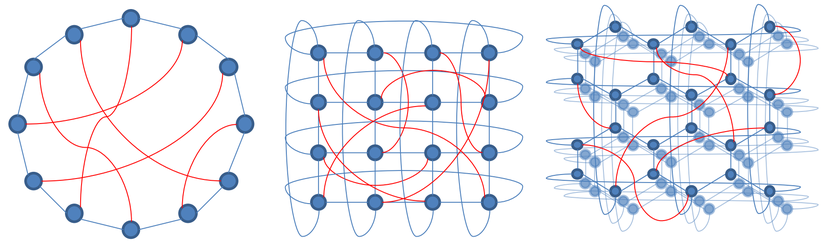
Small-World Datacenters

Ji-Yong Shin*, Bernard Wong+, and Emin Gün Sirer*
{jyshin, egs}@cs.cornell.edu, bernard@uwaterloo.ca
*Cornell University +University of Waterloo

1. Small-World Datacenters

- Connect servers using links created at random!
  - On top of a simple underlying grid
    - Supports topology-aware applications
  - Random links provide
    - Low network diameter
    - Strong connectivity
- Characteristics
  - High bandwidth
  - Fault tolerant
  - Scalable

2. Design



Small World Ring (2 regular + 4 rand) | Small World 2D Torus (4 regular + 2 rand) | Small World 3D Hexagon Torus (5 regular + 1 rand)

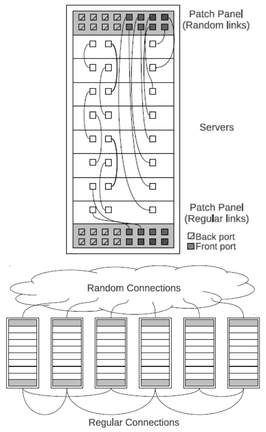
- Direct connections from server to server
  - No need for switches
  - Software routing approach

3. Routing

- Shortest path
  - Link state protocol (OSPF)
  - Expensive due to TCAM cost
- Greedy geographical
  - Find min distance neighbor
  - Coordinates in lattice used as IDs
  - Maintain info of 3 hop neighbors
- Efficient content routing
  - Logical and physical topologies can match
  - Random shortcuts accelerate routing

4. Packaging and Scaling

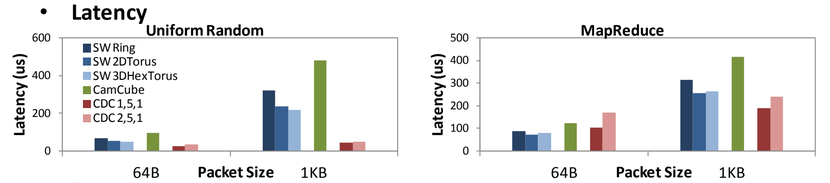
- Supports preconfigured, reusable, scalable components
- Reusable racks
  - Regular links
    - Only short cables necessary
  - Random links
    - Predefined Blueprint
    - Random number generator
    - Pre-cut wires based on known probability



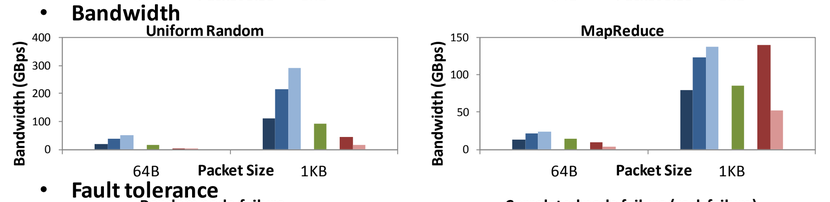
- Ease of construction
  - Connect rack-> cluster (or container) -> datacenter
  - Switches, repeaters, or direct wires for inter-cluster connections

5. Evaluation

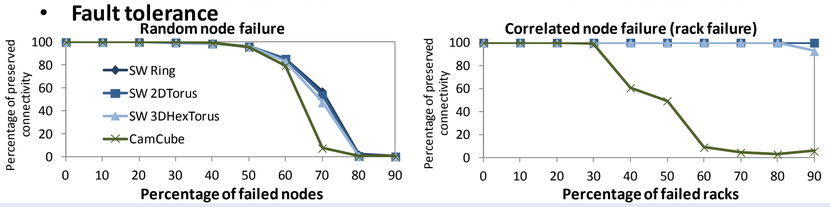
- **Latency**



- **Bandwidth**



- **Fault tolerance**



6. Conclusion

- Unorthodox topology comprising a mix of regular and random links can yield:
  - High performance
  - Fault tolerant
  - Easy to construct and scalable
- Small-world datacenters provide:
  - Higher latency and bandwidth than CamCube
  - Higher bandwidth than CDC
  - Flexible and easy-to-build network infrastructure that enables efficient content routing