IEICE Journal (ED) Special Issue of CANDAR

- **Your efficient graph-generation method should be opened in a journal**

- **Submission deadline: Jan. 7, 2020.**
  (No Extension)

  https://review.ieice.org/regist/regist_baseinfo_e.aspx

Call for Papers

--- Special Section on Parallel, Distributed, and Reconfigurable Computing, and Networking ---

The IEICE Transactions on Information and Systems announces that it will publish a special section entitled " Special Section on Parallel, Distributed, and Reconfigurable Computing, and Networking" in December 2020.

The objective of this special section is to publish and overview recent progress in the interdisciplinary area of Parallel, Distributed, and Reconfigurable Computing, and Networking. This special section will include papers based on the presentation at the International Symposium on Highly-Efficient Accelerators and Reconfigurable Technologies (HEART'19) and the International Symposium on Computing and Networking (CANDAR'19) in addition to papers applied for this call for papers. All submitted papers are subjected to the same review process as those papers accepted for publication in the regular issues.

1

Why don't you use a better graph as network topology of supercomputers?
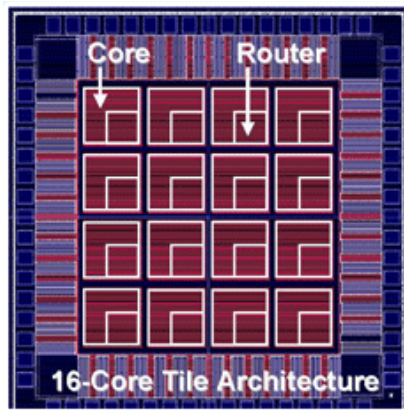
Michihiro Koibuchi

National Institute of Informatics (NII), Japan

2019/11/26
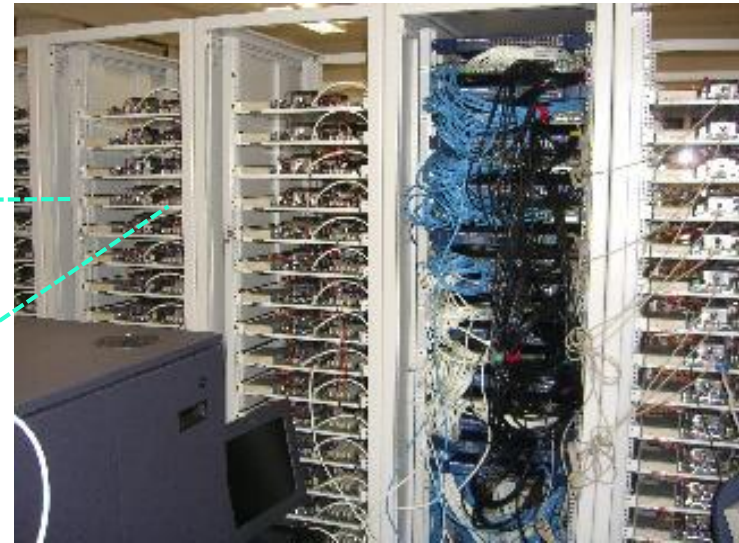
From Graph Golf to Real NW Topology on Supercomputers

Off-chip Network Topology

→ general graph

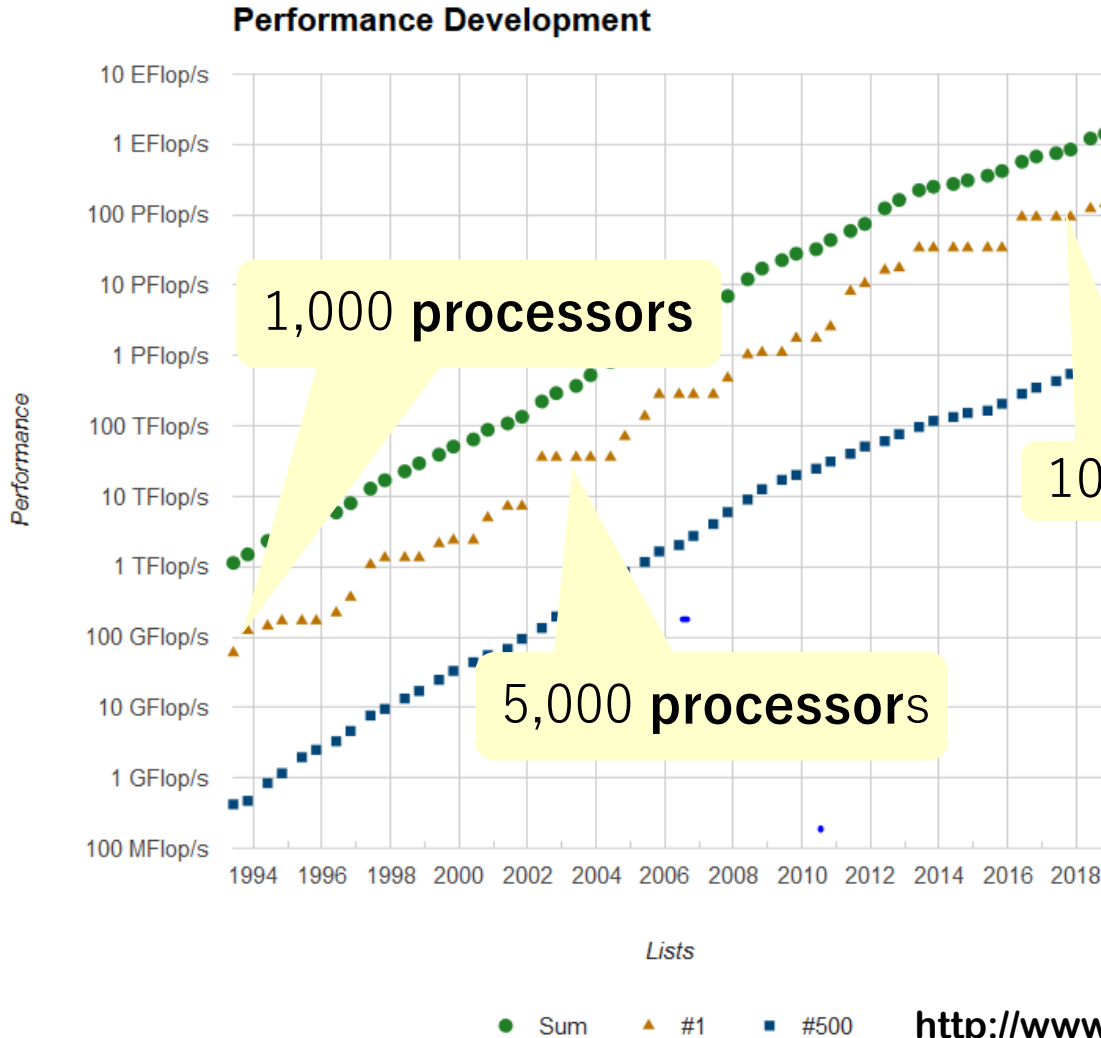**On-chip Network Topology**

→ **Grid Graph**

3

Index

- **Many-core is a game changer for network-topology design**

- **Recommendations of GraphGolf network topologies from different aspects**

Supercomputers @ Top500

- x1,000 times improvement over 10 years
- Recently, " **Many Cores**"

Performance Development

1,000 **processors**

5,000 **processors**

10M **Cores**

K-computer ©RIKEN

1 **bit** = **0/1**
Kilo=10^3
Mega=10^6
Giga=10^9
Tera=10^12
NOW ⇒ **Peta=10^15**
Exa = 10^18
Zetta=10^21

Performance

| 10 EFlop/s |
| 1 EFlop/s |
| 100 PFlop/s |
| 10 PFlop/s |
| 1 PFlop/s |
| 100 TFlop/s |
| 10 TFlop/s |
| 1 TFlop/s |
| 100 GFlop/s |
| 10 GFlop/s |
| 1 GFlop/s |
| 100 MFlop/s |

1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
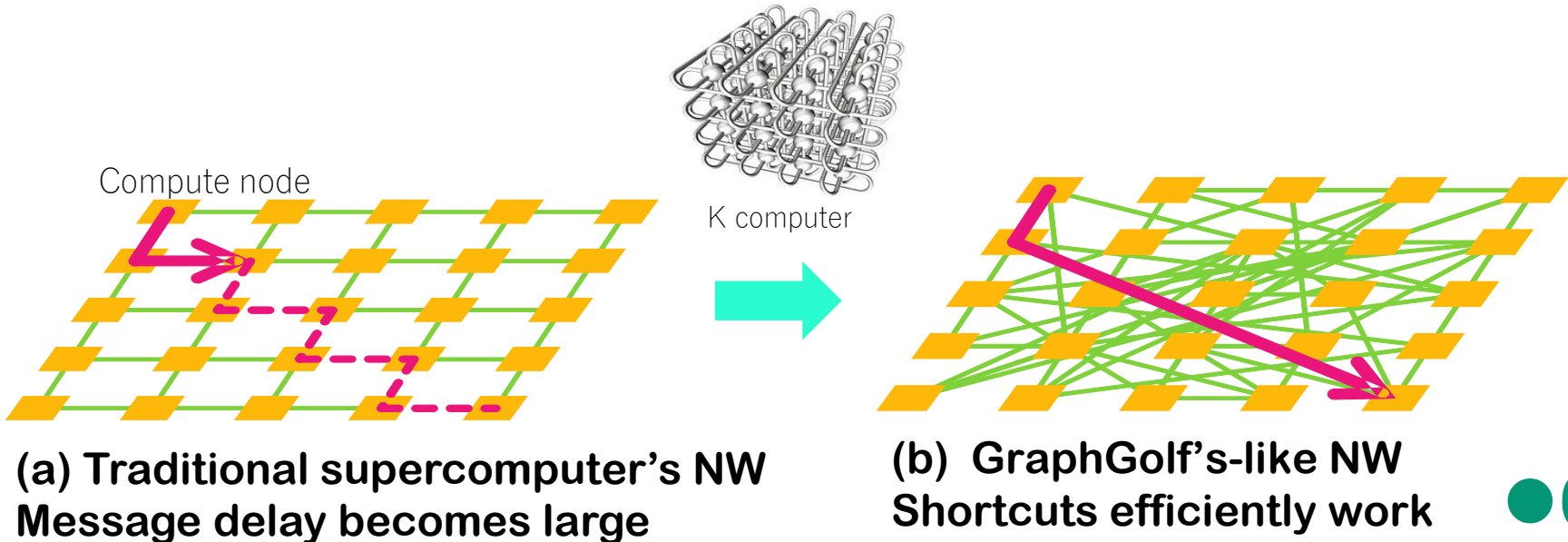
Lists

● Sum ▲ #1 ■ #500   http://www.top500.org

5

"Many-Core" is a game changer (1) → General Graph Category

- **Low Latency: 1 μs latency across system**
  - **Minimize latency for small message (< 3KiB) (fine-grained parallelism)**

[IAA Interconnection network workshop, final report 2008]

→ GraphGolf Competition
- **Low diameter of network topologies**
- **General Graph Category**

Compute node

K computer

(a) Traditional supercomputer's NW Message delay becomes large

(b) GraphGolf's-like NW Shortcuts efficiently work

●6

"Many-Core" is a game changer (2)
→ Grid Graph Category

- **Scalability: greater than 100,000 endpoints**
  - **1 endpoint includes 100~ cores**
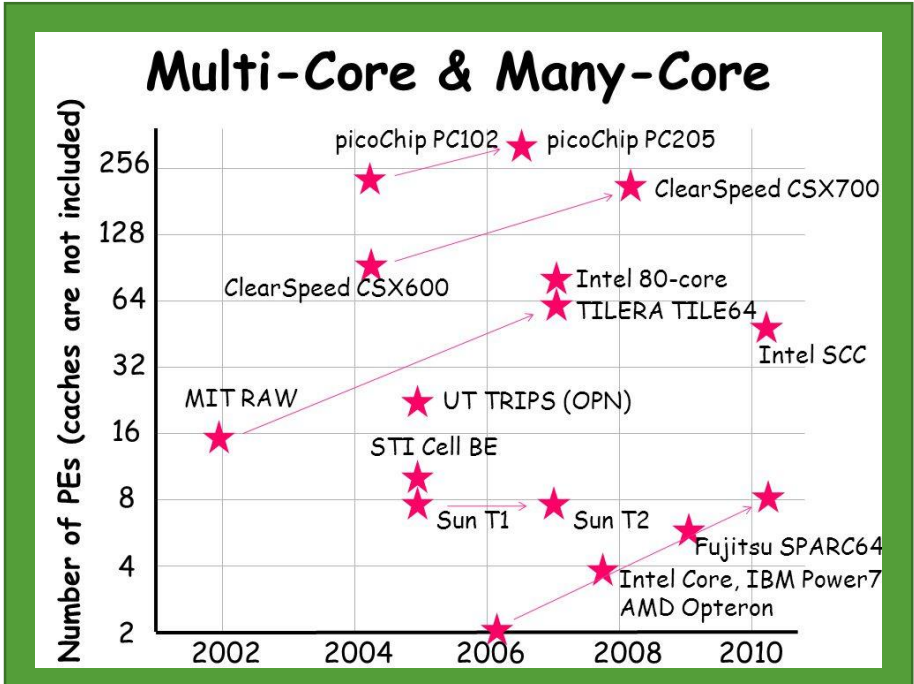
  [IAA Interconnection network workshop, final report 2008]

→ **GraphGolf**
  - Network topology within an **endpoint** [on-chip network]
  - **Grid Graph Category**
    - Short edge length
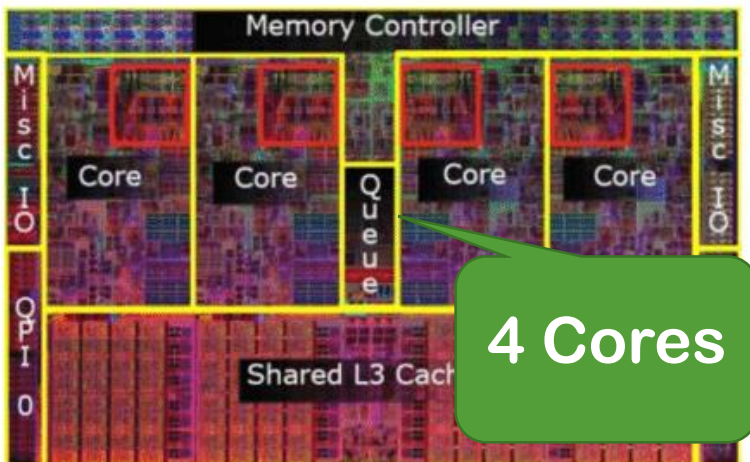
(a) Intel's Nehalem Architecture

4 Cores

Multi-Core & Many-Core

[H.Matsutani]

Index

- **Many-core is a game changer for network-topology design**

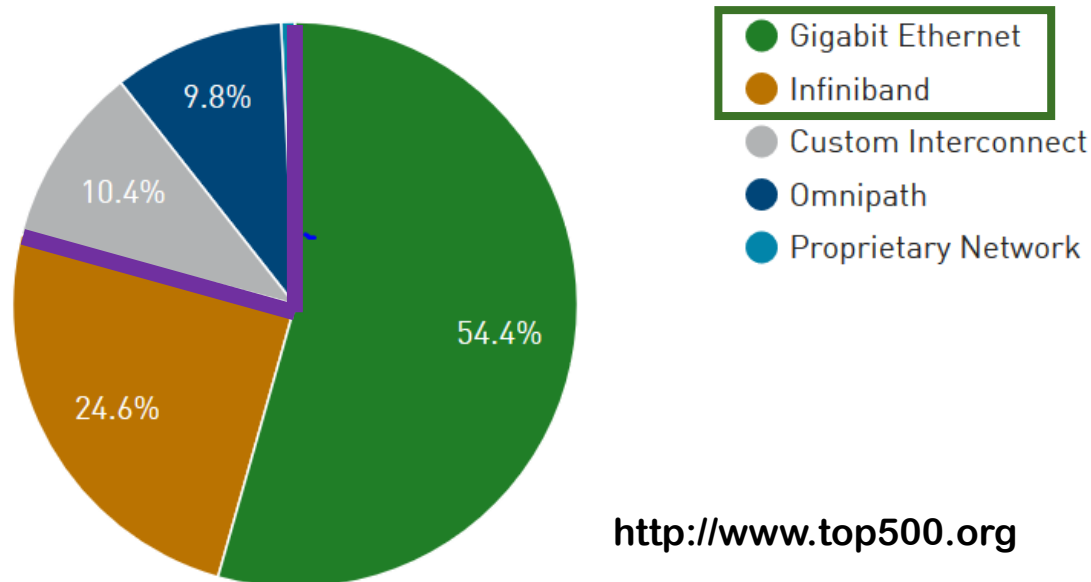- **Recommendations of GraphGolf network topologies from different aspects**

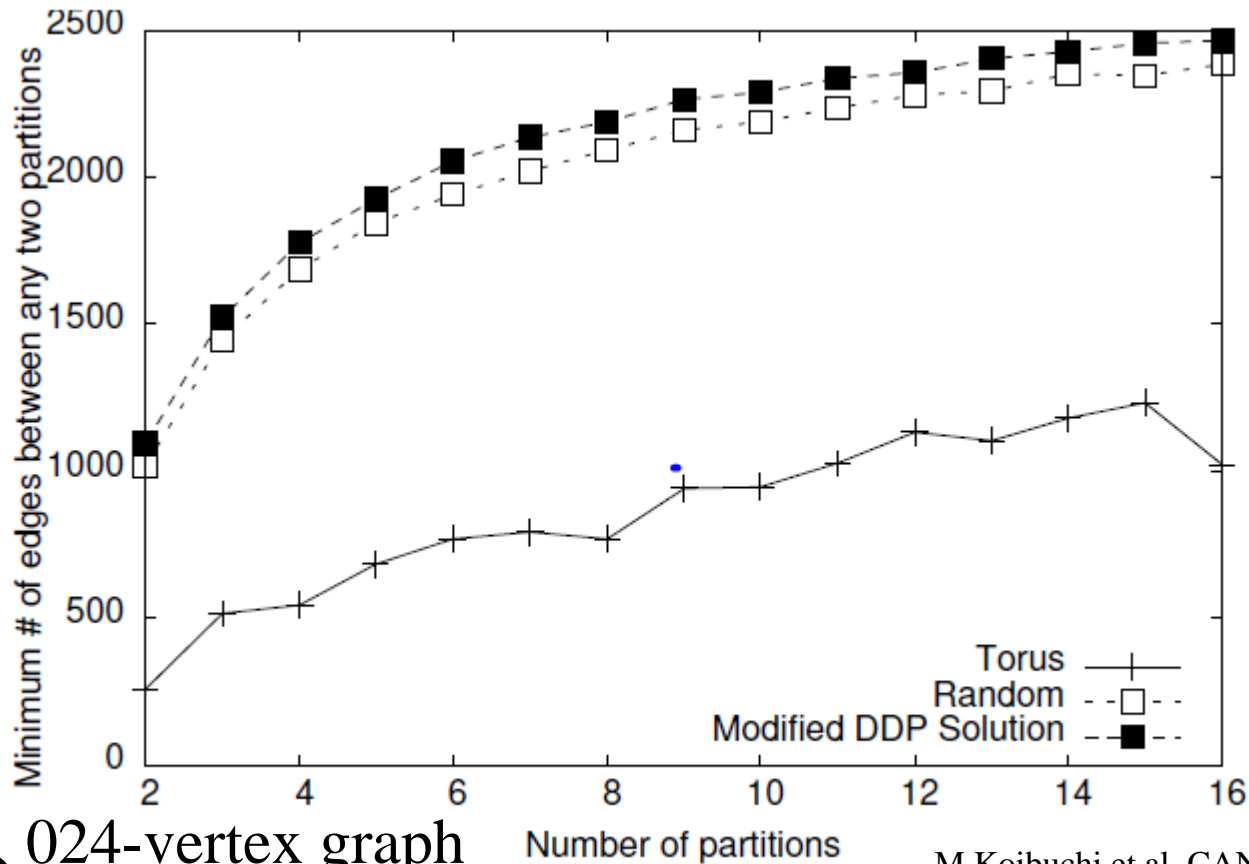# Recommend. 1: High Feasibility 80% of top500 supercomputers can use GG NW Topology

- **Ethernet + InfiniBand ≒ 80% in top500 supercomputers**
  - Support arbitrary network topology

Interconnect Family System Share

Legend:
- Gigabit Ethernet
- Infiniband
- Custom Interconnect
- Omnipath
- Proprietary Network

54.4%
24.6%
10.4%
9.8%

http://www.top500.org

●9

Recommend. 2: High Bandwidth

- **GraphGolf topology = tightly coupled network topology**

1、024-vertex graph

M.Koibuchi et al, CANDAR16

●10

Recommend. 3: Low cost of GG NW topologies

- **7 hops for 970ns**

970ns

150ns 100ns 100ns 100ns 100ns 100ns 100ns 150ns
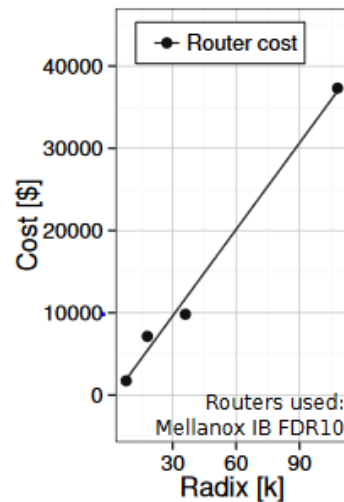
① ② ③ ④ ⑤ ⑥

10ns 10ns 10ns 10ns 10ns 10ns 10ns
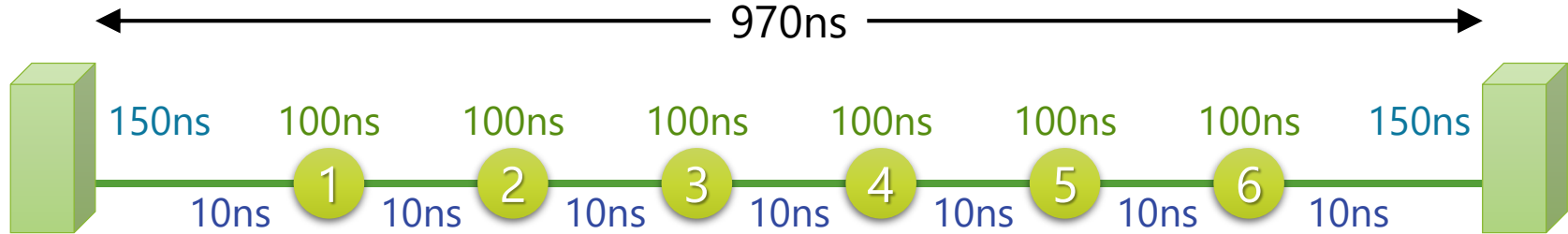
- Low switch degree (radix) leads to the low network cost
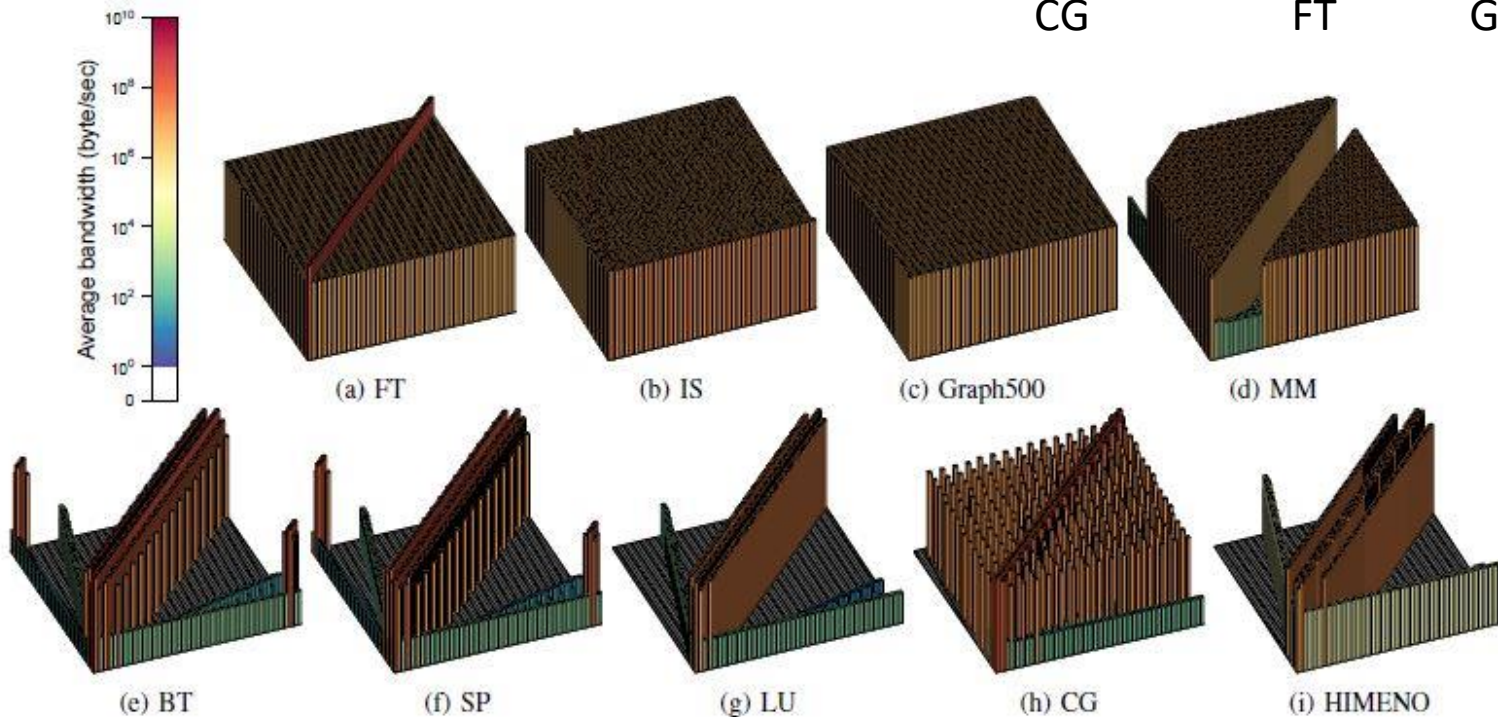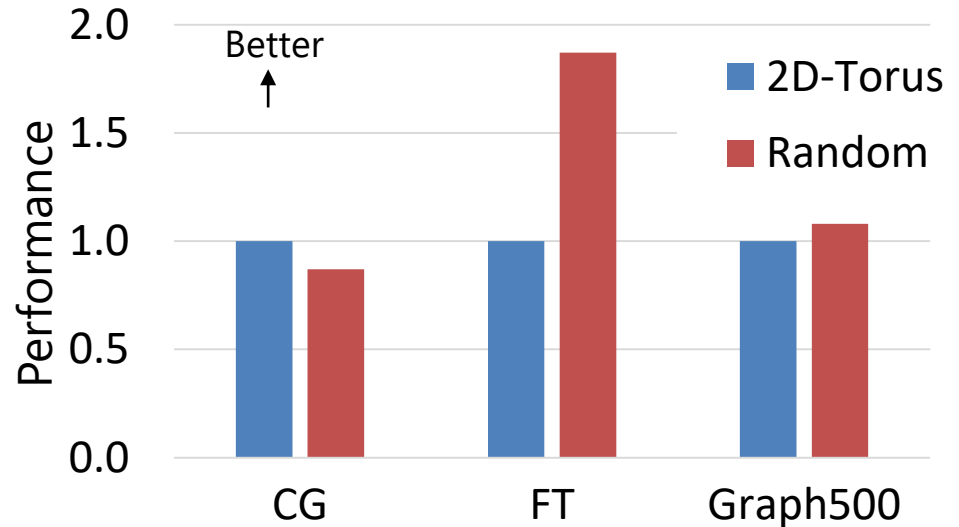 - **Graph Golf's network topology is also advantage in terms of cost**
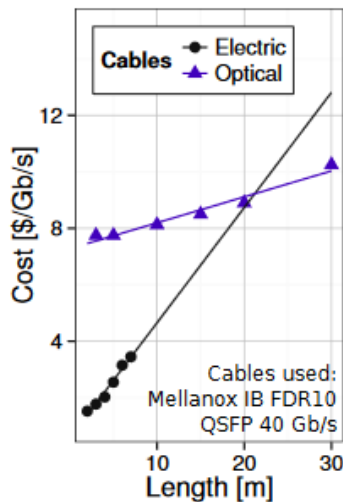
(a) Cable cost model.   (b) Routers cost model

M.Besta et al, SC14

Recommend. 3: Killer Traffic Patterns

- **All-to-all access**
  - **GG NW topology** is better
- **Pure stencil access**
  - Topology does not impact

(a) FT   (b) IS   (c) Graph500   (d) MM

(e) BT   (f) SP   (g) LU   (h) CG   (i) HIMENO

F.Chaix et al, PDP16

●12

Recommend. 4: High Fault Tolerance

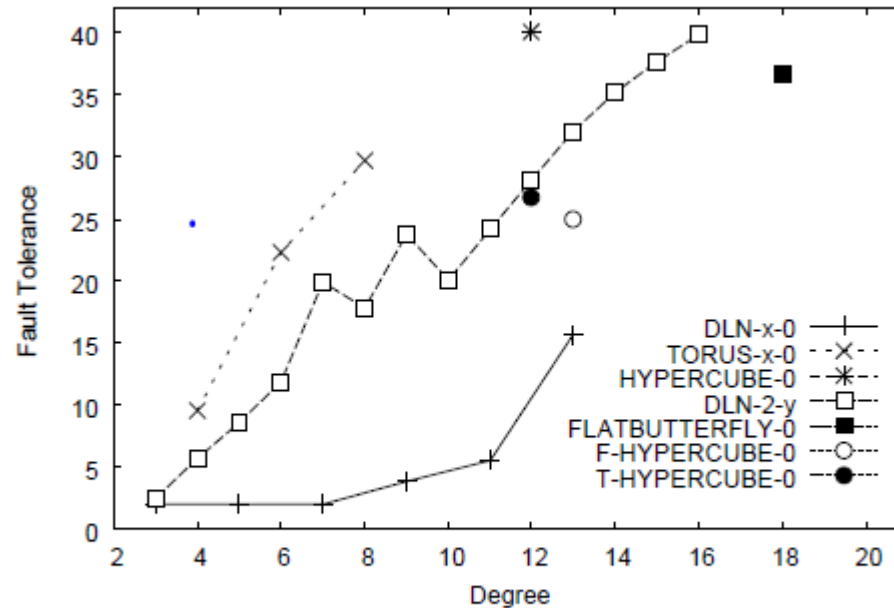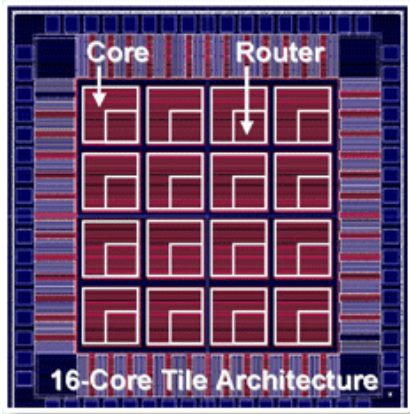- **Unstructured NW topology is still unstructured NW topology w/ miscabling/link failure**

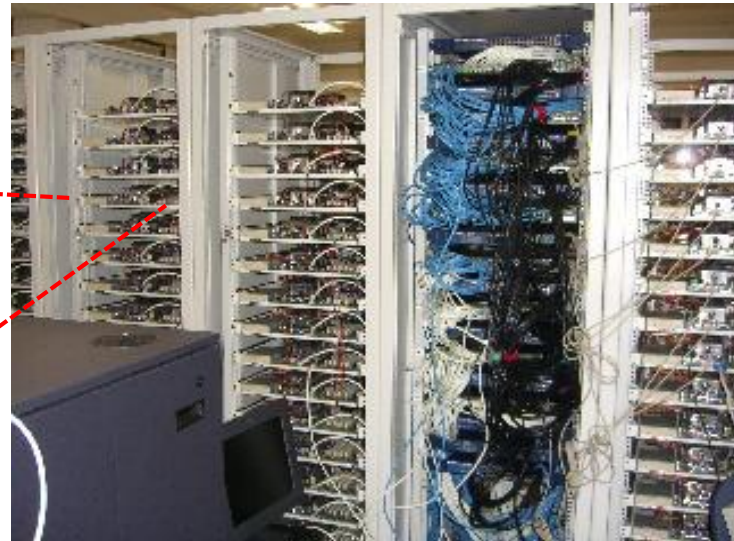Figure 6. Fault tolerance vs. degree ($N = 2^{12}$ vertices).

M.Koibuchi et al, ISCA2012

●13

Conclusions

- **Traditional NW topologies are no longer a best solution on many-core supercomputers**
  - → **GraphGolf graphs**

On-chip NW Topology

→ **Grid Graph**

Off-chip NW Topology

→ **General Graph**

14