Introduction to Research

Xin Yuan (xyuan@cs.fsu.edu)
Professor and Chair

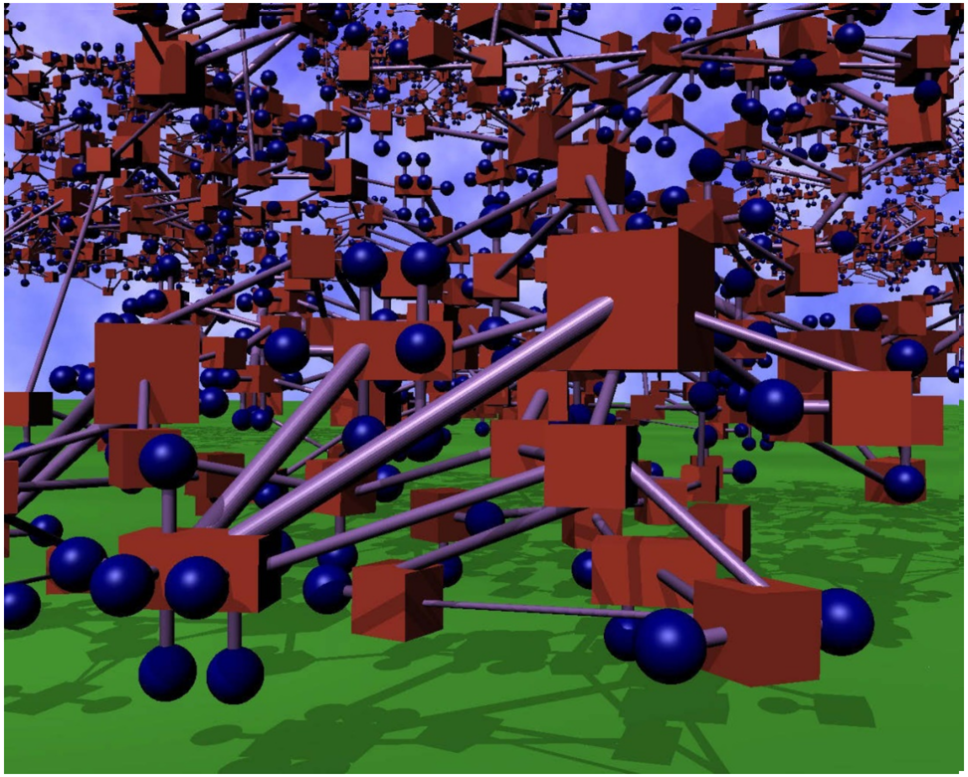
Research website: http://explorer.cs.fsu.edu

Research interests:

- Parallel and distributed systems
- Interconnection networks
- Software defined networks

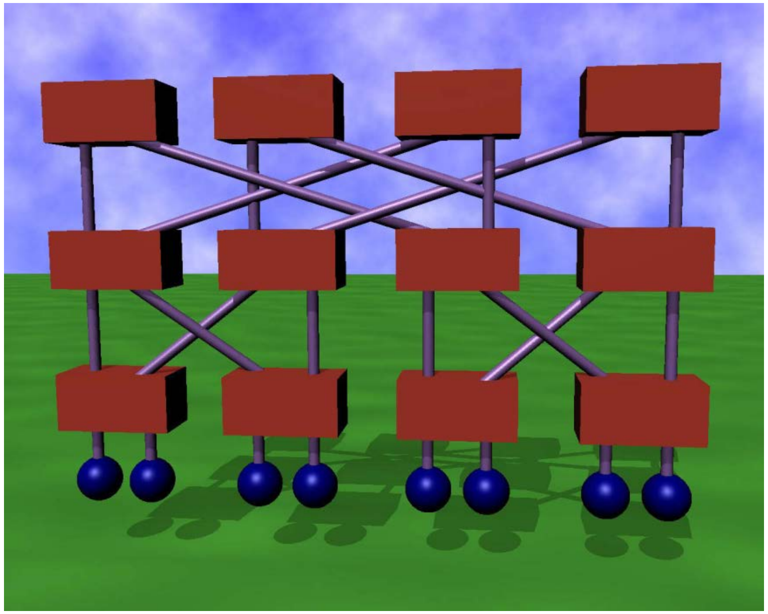
Large scale parallel and distributed systems and networks

- Internet, Internet of Things (IoT) : arbitrary topology, shortest path routing, complicated routing protocols



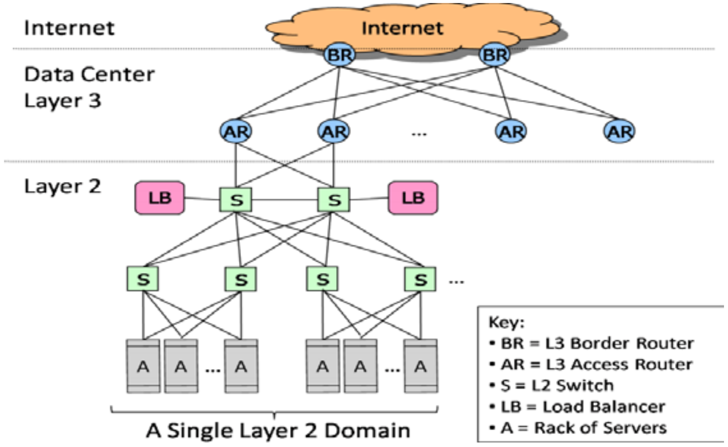
Large scale parallel and distributed systems and networks

- Supercomputers: Interconnect, regular topology, routing
  - Every new generation supercomputer uses a new network topology



Large scale parallel and distributed systems and networks

- Datacenters: Amazon EC2, cloud computing
  - As the size scales, the data center networks run into cost performance issue.



Large scale parallel and distributed systems and networks

- Internet, Internet of Things (IoT), supercomputers, cloud computing data centers

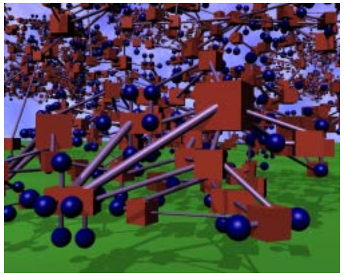
- Many issues in communication and networking:
  - How to make the network work at a very large scale?
  - How to make the network work better (efficiency in performance or power)?
  - How to make the network more reliable and secure?
  - How to evaluate the network design effectively?

Current research projects

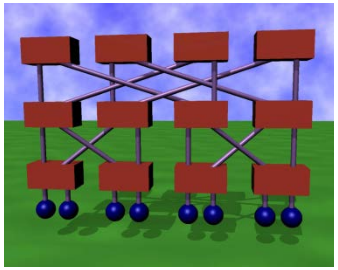
- Interconnection Networks for Future Large Scale Data Centers and HPC Systems
  - topology, routing, modeling and simulation
- Software Defined Networking for HPC systems
- Machine learning based approach for network design

Towards Optimal Network Connectivity

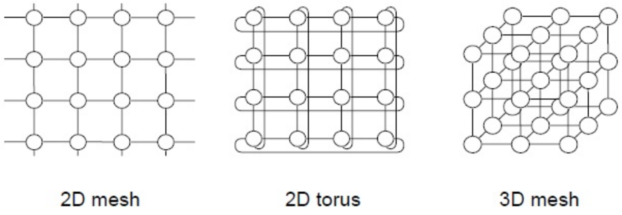
- The problem: given a (large) number of switches/routers with a fixed port count (N r-port switches) , what is the best performance that one can get?



Internet topology



Fat-tree



Low dimensional topologies

Background

- Supercomputers
  - The most cutting-edge computing technology
  - A new network topology in every generation.
- Data Centers
  - Need innovation at all levels of networking, including new topologies.

- How to effectively connect devices is still an unsolved problem.

Problem refinement

- Any topology formed by N r-port switches, (DRG(N, r))
- k-shortest path routing
- Desired properties:
  - Large aggregate capacity – high bisection bandwidth
  - Low latency and minimal resource usage – small diameter and average path length
  - Load balancing – paths for carrying packets must be evenly distributed among the network

Performance metrics

- Diameter
- Average k-shortest path length
- Load balancing
  - Maximum probability that a link is used to carry a random packet assuming the k paths are used evenly.
    - Equivalent to maximum link load for all-to-all communication.

Theoretical bounds for DRG(N, r)

Lemma 1 (diameter): The diameter of any $DRG(N,r)$ is at least

$$\lceil log_r(N(r-1)+1)\rceil - 1.$$

Theoretical bounds for DRG(N, r)

Lemma 2: For any $DRG(N,r)$, the average path length of all k-shortest paths between all SD pairs, $AKH(N,r,k)$, is

$$AKH(N,r,k) \geq \frac{\sum_{j=1}^{h-1} jr^j + hR}{k(N-1)}$$

where

$$R = k(N-1) - \sum_{j=1}^{h-1} r^j \geq 0$$

and h is the largest integer such that the inequality holds.

Theoretical bounds for DRG(N, r)

Lemma 4: For any $DRG(N,r)$, the lower bound of the maximum link load for all-to-all communication is

$$ML(N,r) \geq \frac{LK(N,r,1) \times (N-1)}{r}.$$

LK(N, r, 1) is the lower bound of the average shortest path length for *DRG(N, r)*.

Generalized De Bruijn Graph (GDBG(N, r))

- The outgoing link cycling through the N nodes in a round-robin fashion.
  - Node 0 connects to 0, 1, 2, …, r-1
  - Node 1 connects to r, r+1, …, 2r-1
  - …….

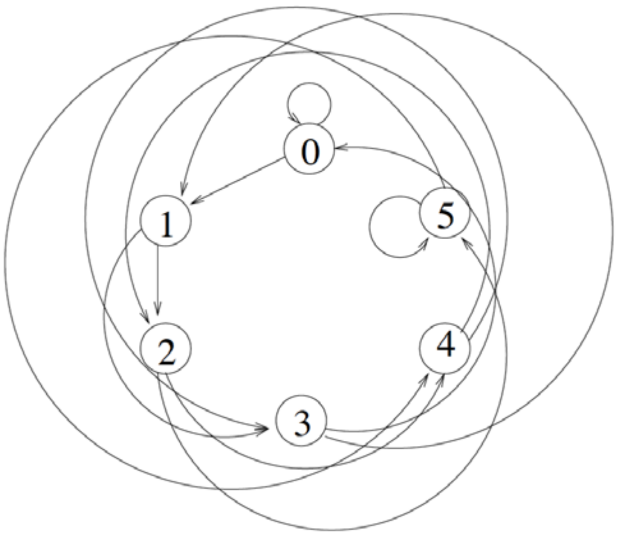


Figure 1. A GDBG(6,2) topology

GDBG properties

Lemma 7 (GDBG Diameter): The diameter of $GDBG(N, r)$ is no more than D such that $r^D \geq N$.

Optimal diameter satisfies $1 + r + \cdots + r^{D_{opt}} \geq N$

GDBG property

Lemma 8 (average k-shortest path length):
When $\frac{r^{H-1}}{N} < k \leq \frac{r^H}{N}$,

the average k-shortest path length of GDBG(N,r) is no more than

$$LK(N,r,k) + \frac{H \times (r^H - 1)}{(r-1)^2 \times k \times (N-1)}$$

GDBG property

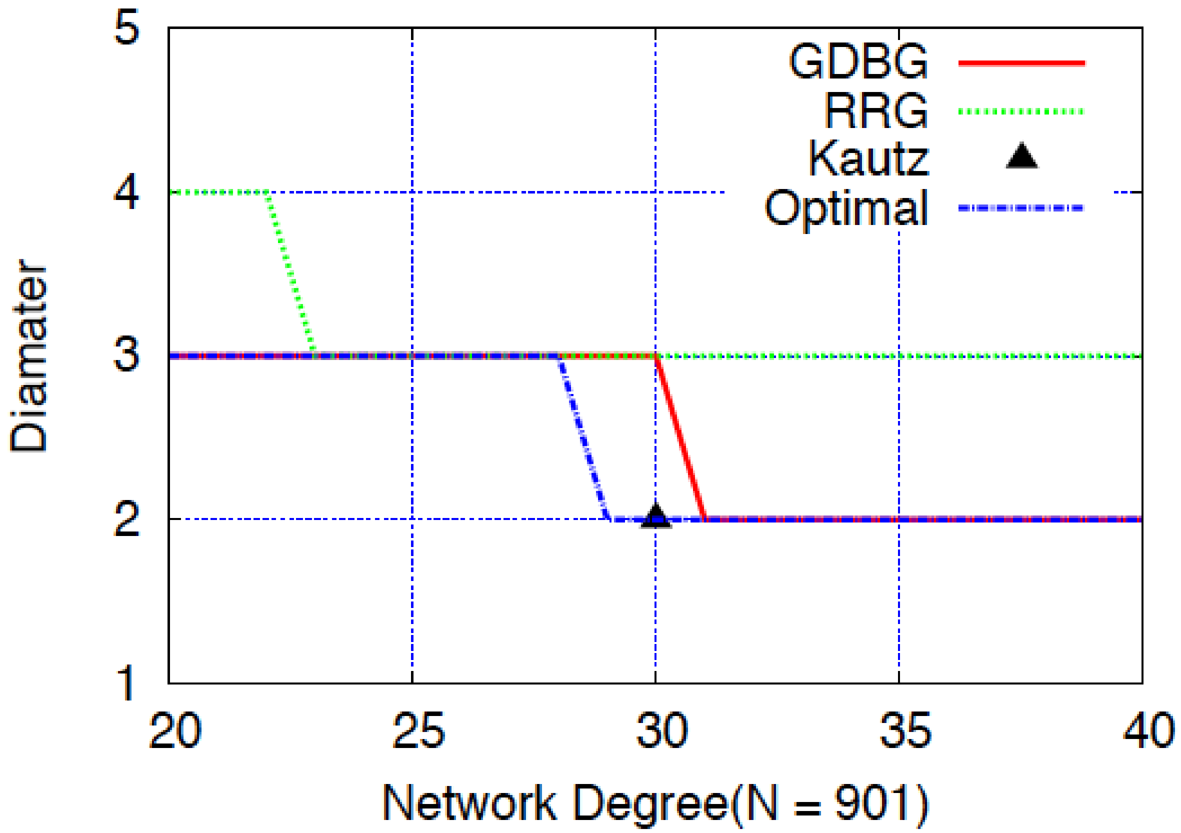
Lemma 10 (load balance): For $GDBG(N,r)$ with ALL-PATH(H), each SD pair will at least have $\lfloor \frac{r^H}{N} \rfloor$ paths, and the maximum link load for all-to-all traffic is no more than

$$\frac{(H)r^H - \frac{r^H - 1}{r-1}}{r-1} \times \frac{1}{\lfloor \frac{r^H}{N} \rfloor} .$$

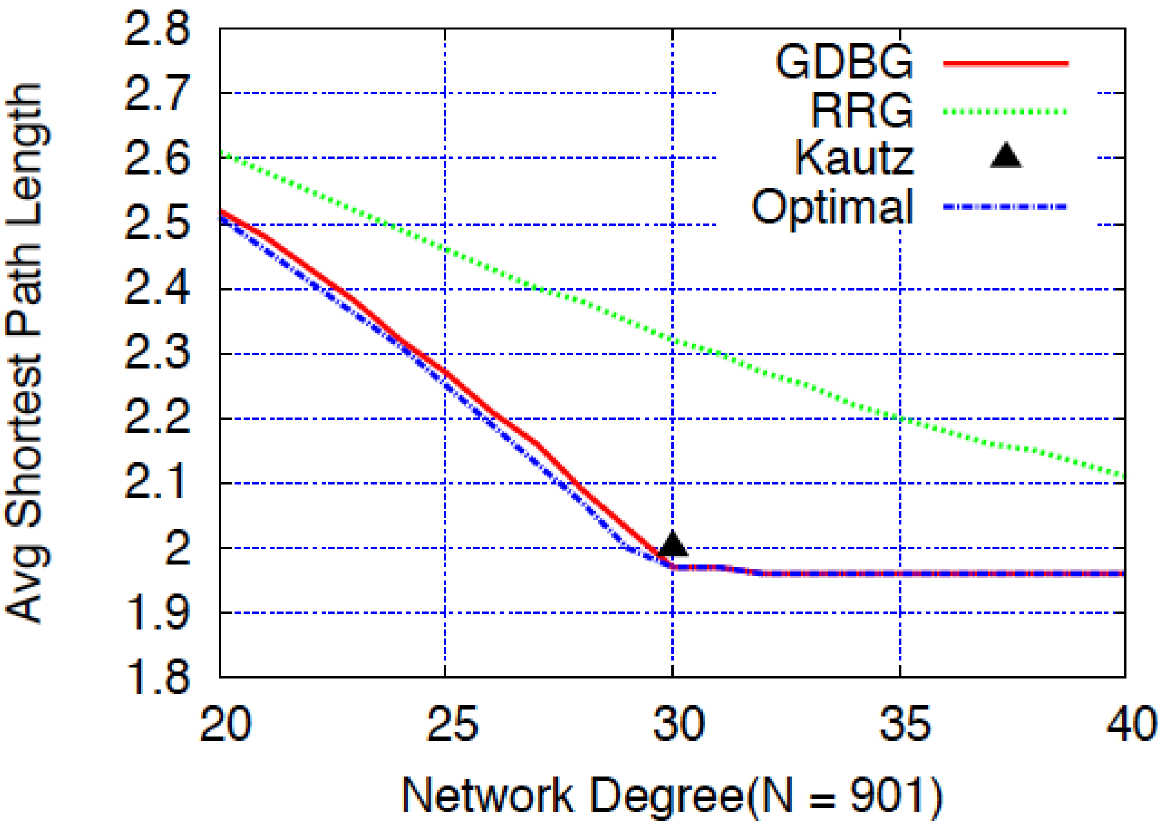
$$\frac{(H)r^H - \frac{r^H - 1}{r-1}}{r-1} \times \frac{1}{\lfloor \frac{r^H}{N} \rfloor} \approx \frac{H \times N}{r-1}$$

$$ML(N,r) \geq \frac{LK(N,r,1) \times (N-1)}{r} \approx \frac{(H-1)(N-1)}{r}$$

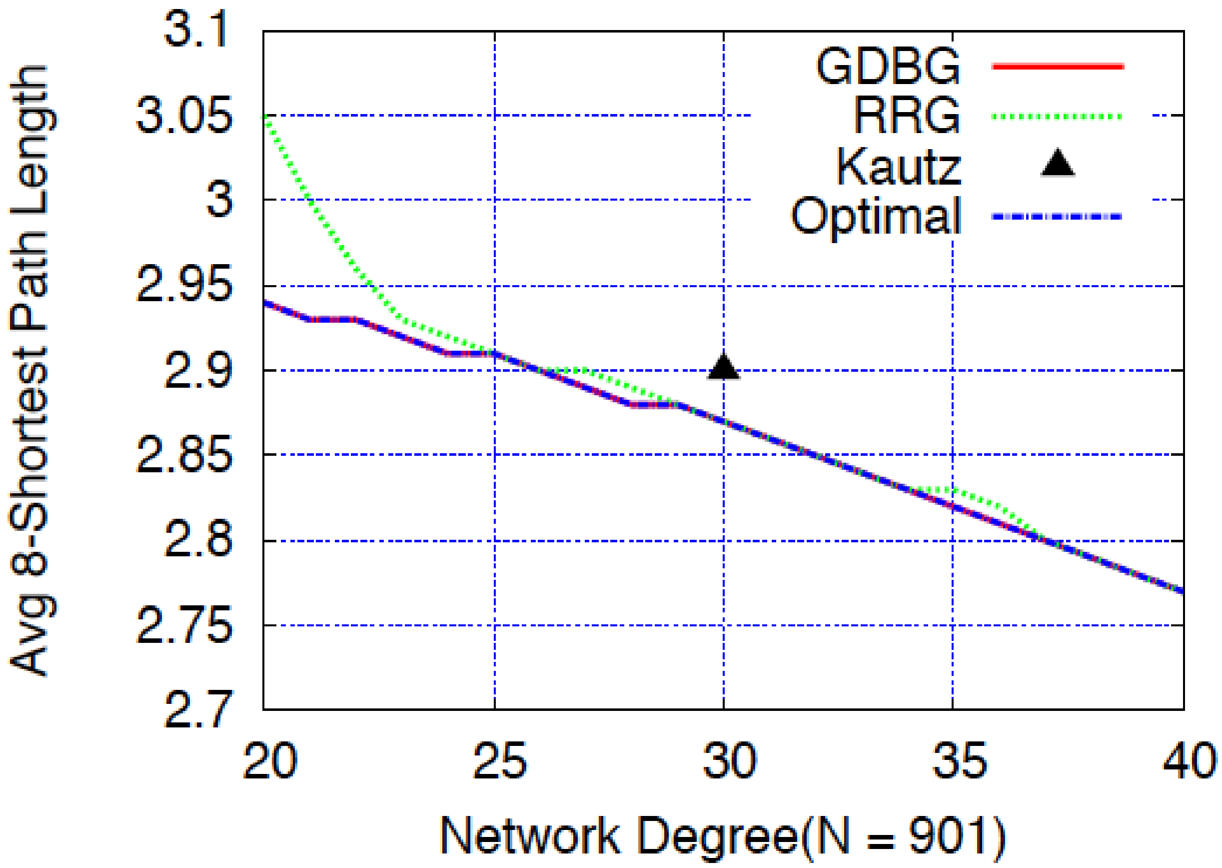
Some plots (diameter)



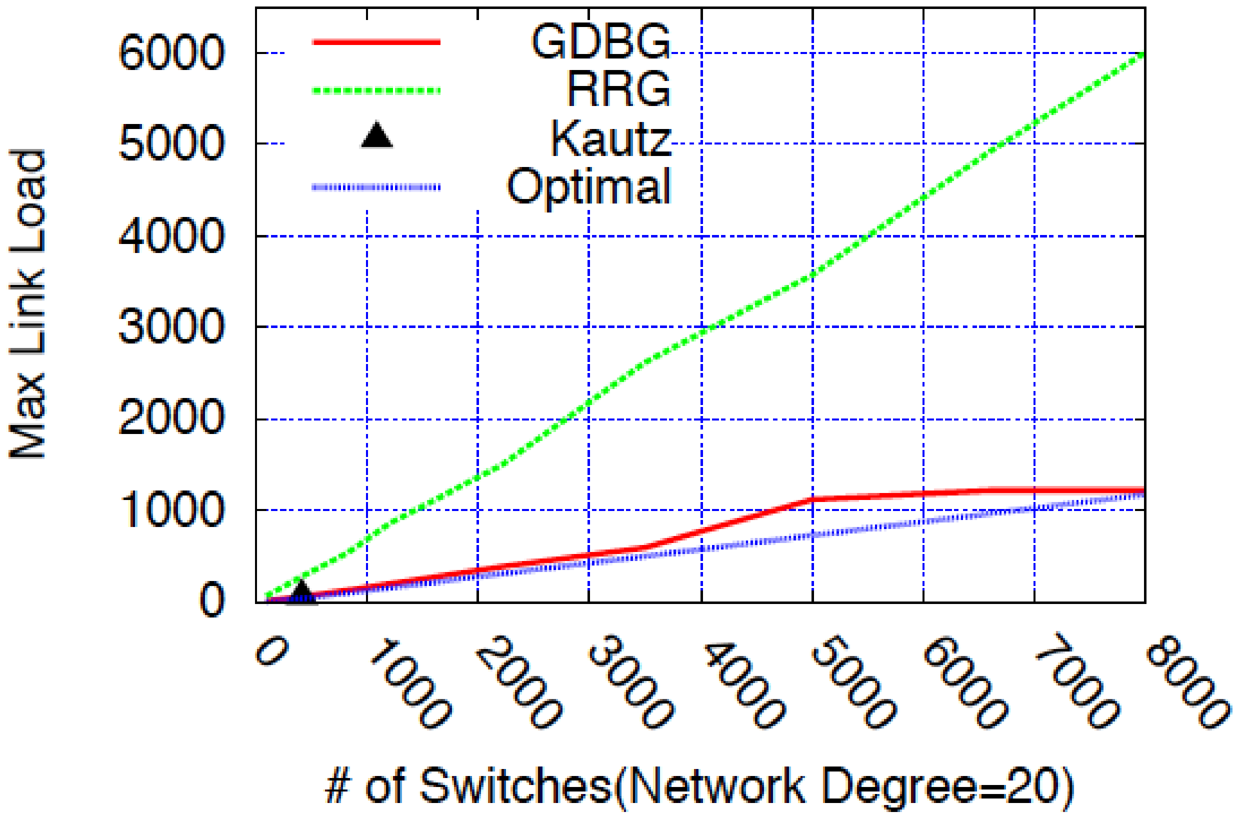
Average shortest path



Average 8-shortest path length



Maximum link load (load balance)



Conclusions

- We established the performance bounds that is possible for any DRG(N, r).
- We showed that with uni-directional networks, GDBG is almost a perfect network in our metrics.

Other research

- Design practical networking techniques that have theoretical backing.
  - Design optimal topologies
  - Design optimal routing schemes
- Software development projects
  - Using SDN to control InfiniBand networks
  - Re-develop the system software stack to take advantage of SDN capability for HPC systems

For more information

http://explorer.cs.fsu.edu