Jinshu SU, Baokang ZHAO, Yi DAI, Jijun CAO, Ziling WEI, Na ZHAO, Congxi SONG, Yujing LIU, Yusheng XIA, 2022. Technology trends in large-scale high-efficiency network computing. *Frontiers of Information Technology & Electronic Engineering*, 23(12):1733-1746. https://doi.org/10.1631/FITEE.2200217

Technology trends in large-scale high-efficiency network computing

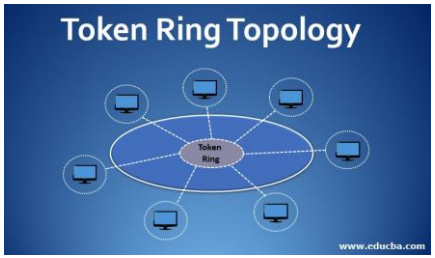
Key words: Supercomputing; Cloud computing; Network technology; Development trends

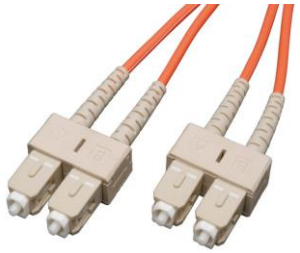
Corresponding authors: Jinshu SU, Baokang ZHAO
E-mail: sjs@nudt.edu.cn; bkzhao@nudt.edu.cn
ORCID: https://orcid.org/0000-0001-9273-616X
https://orcid.org/0000-0001-9200-9018

Background

- Networking technology has three important characteristics:
  - Networking technologies are changing rapidly.

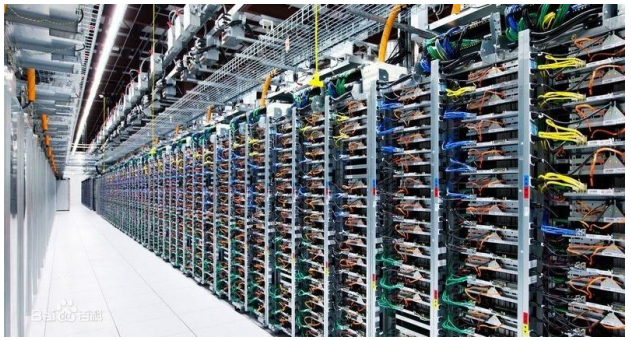
Ethernet has developed into a rich family





Some technologies (e.g., Token Ring, FDDI) disappear

  - There is a rapid development of data centers.
  - The computing power is developing at an unimaginable speed.



Background

- It is necessary to explore new models and new paths for the development of network technologies.

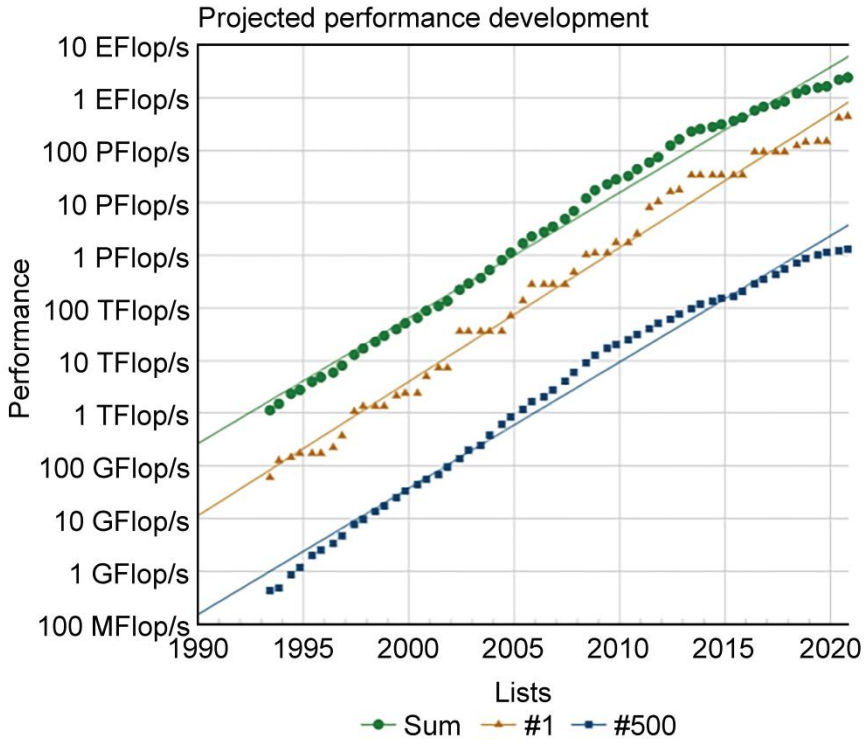


Fig. 1 Diagram of supercomputing TOP500 performance from 1993 to 2021

The graph is generated from the www. top500. org website. References to color refer to the online version of this figure

Three trends

- Integration: The performance demands are similar under different scenarios.
  - ➢High throughput and low latency are the main goals for both the data center and supercomputing.
- Differentiation: Many new technologies have evolved according to the goals pursued by different applications.
- Optimization: To satisfy the high performance requirements, optimization is a permanent topic.

1) Integration

- Heterogeneous network complementarity
  - The inherent networking requirements are multifaceted, and heterogeneous network integration is a trend.

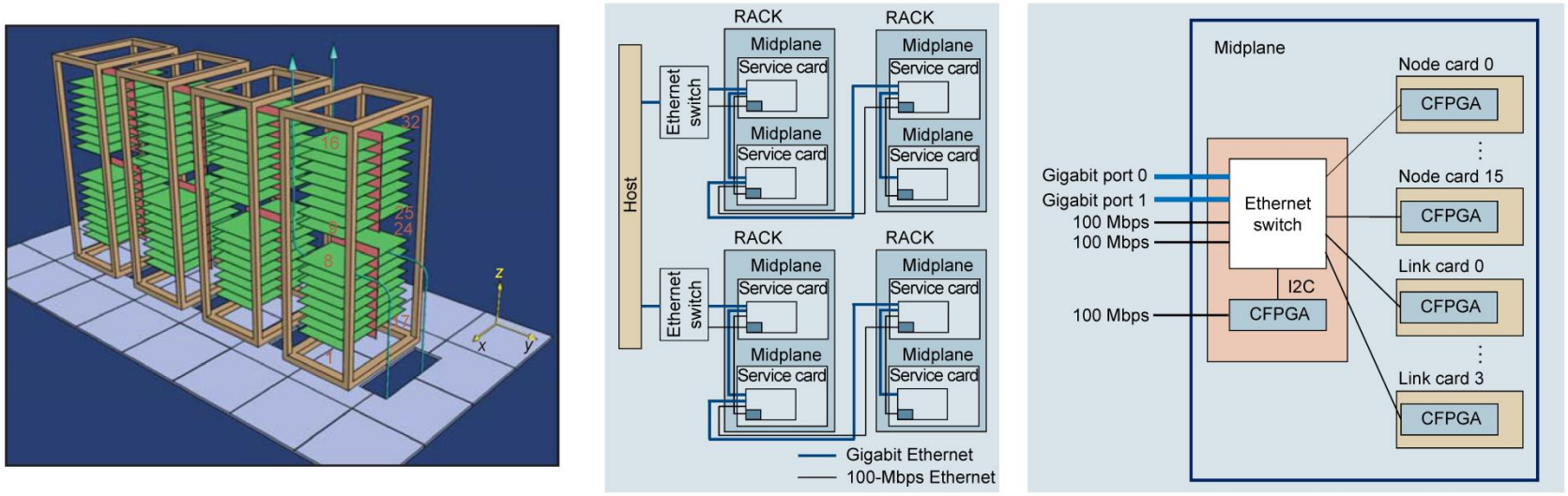


Fig. 2 Diagram of IBM Bule Gene Q series heterogeneous networks: (a) diagram of four RACKs; (b) top-level view of the Blue Gene L control network; (c) configuration diagram of half RACK (Coteus et al., 2005)

1) Integration

- Generalization of link technology
  - Improve the link speed
  - Shorten the point-to-point delay
- Link transmission unification
  - SerDes technology is widely used.

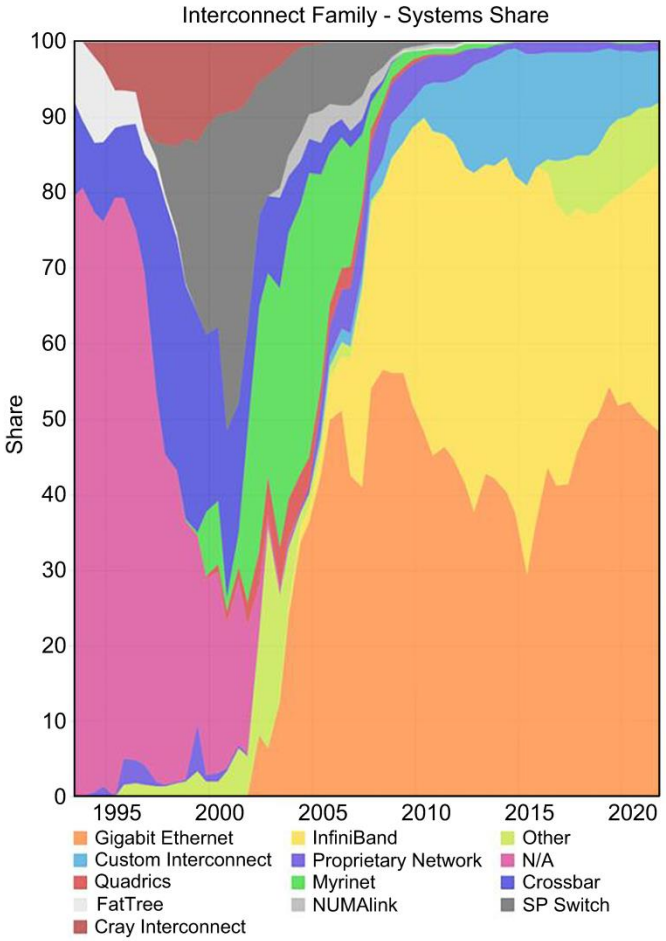


Fig. 3 Diagram of the network technology trend in TOP500 supercomputing

The graph is generated from the www.top500.org website. References to color refer to the online version of this figure

2) Differentiation

- High-end system customization
  - Advanced enabling technologies are widely adopted in high-end systems and customized interconnection networks.

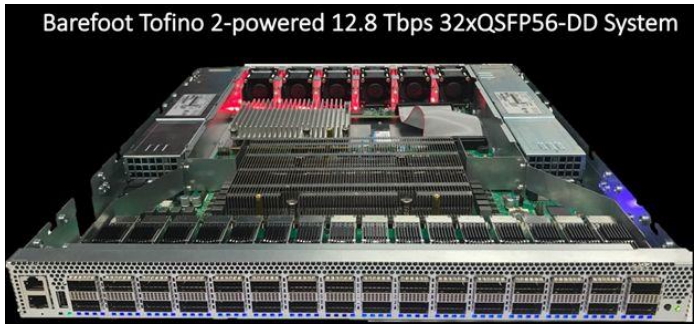
- Integration design for the network and interface
  - Data center networks generally use a single network technology.

- Transmission protocol customization
  - In the long term, QUIC instead of TCP may be the mainstream.

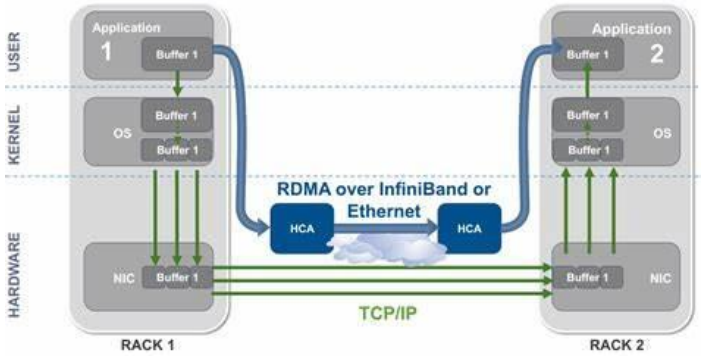
QUIC

3) Optimization

- In-network computing in data centers
  - The innovation of software and hardware has promoted the rapid development of in-network computing.



- Collective computing for supercomputing
  - Collective computing makes the CPU-centric processing mode evolve into a data-centric processing mode.
- Protocol offloading
  - Protocol offloading achieves acceleration using relatively complex protocols as individual components.
- Cross-layer optimization



Summary

- Driven by the development of the Internet, supercomputing, cloud computing, big data, artificial intelligence, and AR/VR of the Metaverse, increasing demands on network technology have been proposed.
- From the perspective of the overall development of network technology, this paper summarizes three development trends of network technology, i.e., integration, differentiation, and optimization.
- It aims to provide useful guidance for system designers and key technical points for researchers.

References

[1] TOP500, www.top500.org

[2] Coteus P, Bickford HR, Cipolla TM, et al., 2005. Packaging the Blue Gene/L supercomputer. IBM J Res Dev, 49(2-3):213-248.

[3] Liao XK, Pang ZB, Wang KF, et al., 2015. High performance interconnect network for Tianhe system. J Comput Sci Technol, 30(2):259-272.

[4] Iyengar J, Thomson M, 2021. QUIC: a UDP-Based Multiplexed and Secure Transport. RFC 9000. Internet Engineering Task Force.

Jinshu Su received his BS degree in mathematics from Nankai University, Tianjin, China, in 1985, and his MS and PhD degrees in computer science from National University of Defense Technology (NUDT), Changsha, China, in 1988 and 2000, respectively. He is a professor in the College of Computer, NUDT. His research interests include Internet architecture, network security, and AI for networking.

Baokang Zhao received his BS, MS, and PhD degrees from the College of Computer, NUDT, all in computer science. He was a visiting scholar at Hong Kong Polytechnic University, and is currently an associate professor with the College of Computer, NUDT. His main research interests include network optimization, AI for networking, high-performance networking, space-air-ground integrated networking, and distributed computing.