# Solution for Supercomputing Center InfiniBand Network and

Gigalight Interconnection

High Performance Computing (hereinafter referred to as HPC), namely Supercomputer Center HPC, its market is being developed toward the direction to use heterogeneous computing systems and improve energy efficiency ratio: GPU, DSP and ARM processors run at the same time to achieve higher Petaflop(1 quadrillion floating-point calculations/second) values with less energy consumption. IDC forecasted that the annual growth rate of HPC sales amount would have to remain at around 7.3% if HPC server market would realize the goal of sales amount of 15 billion dollars in 2017. Supercomputing has a very wide range of applications, including automobile manufacturing simulation, weather forecasting, molecular biological study, geophysics, etc.; in these areas, parallel computation, plenty of data processing and complex operations are always needed. Big data analysis which is often discussed recently also uses supercomputing.

HPC requires the most efficient computing platform. **Computing Cluster** technology platform has become the most common hardware solution of HPC simulation due to its excellent productivity and flexibility. The Computing **Cluster** technology platform often uses the following high-speed interconnection technologies for communication, such as: InfiniBand, high-performance Ethernet (simplified frame structure), BlueGene, Gray and other interconnection technologies, and also the infiniBand interconnection technology has been adopted by 41.2% Supercomputer Centers.
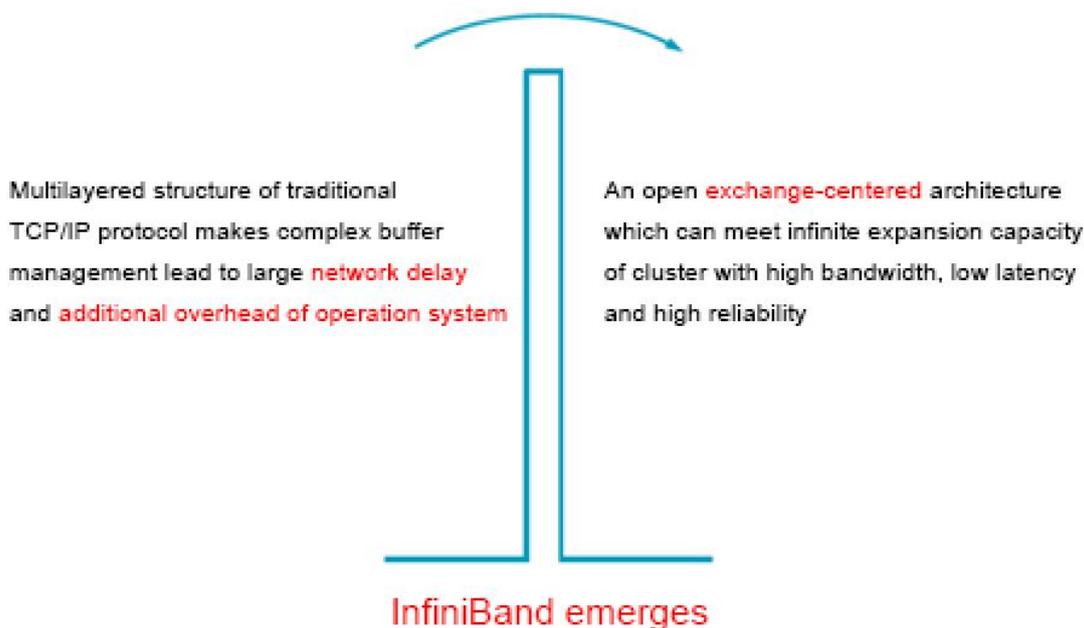

TOP500 Interconnect Trends
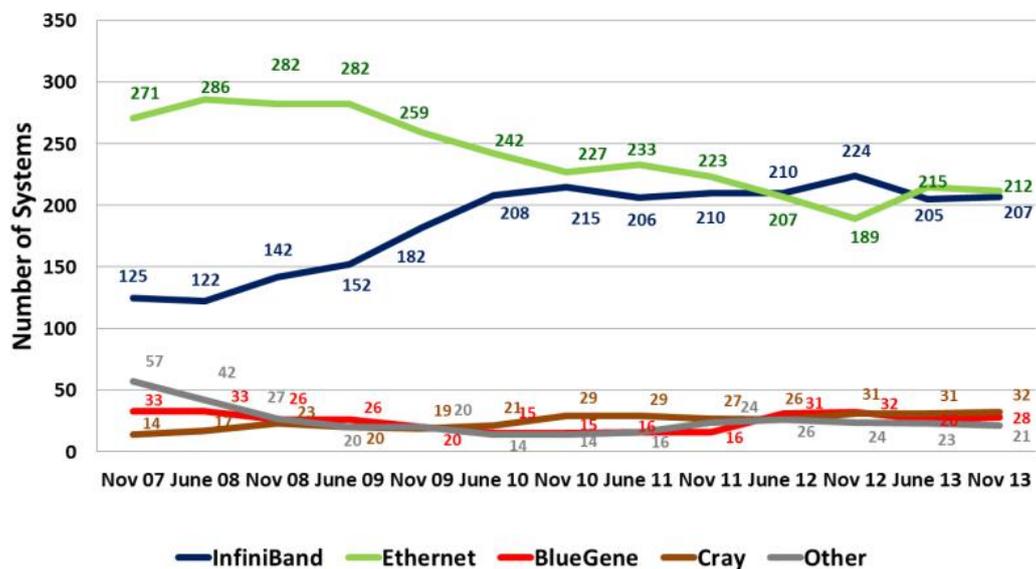
Ethernet or Cray interconnection technology is used by many systems in Top500 of Supercomputer Center and these special advantages make infiniband become popular in supercomputing; according to the statistics, the infiniband interconnection technology is used by 41.2% systems in Top500 list of global Supercomputer Center.

Generation cause of InfiniBand:

Generation of InfiniBand

Multilayered structure of traditional TCP/IP protocol makes complex buffer management lead to large network delay and additional overhead of operation system

An open exchange-centered architecture which can meet infinite expansion capacity of cluster with high bandwidth, low latency and high reliability

InfiniBand emerges

As an input-output (I/O) broadband structure, InfiniBand can improve the communication speed among various devices of server and network subsystems and also provide the broadband services with higher performance and infinite expansion to future computer systems. InfiniBand technology is not used for general network connection, and it is designed mainly for connection problems of server side. Therefore, InfiniBand technology will be used for communication between server and server (such as copying, distributed work, etc.), server and storage device (such as SAN and direct storage attachment), and server and network (such as LAN, WANs and the Internet). Since Supercomputer Center is a huge system, the construction of supercomputing system is not a simple packing of hardware and the network interconnection technology also plays an important role in deciding computing capacity and efficiency of supercomputing in addition to CPU, GPU and other core components.
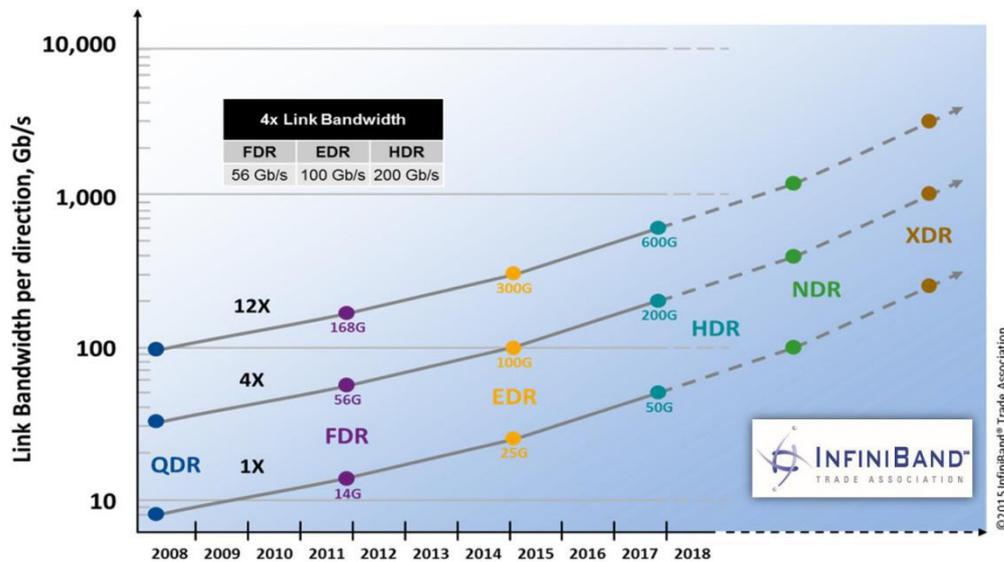
Advantages of InfiniBand

For InfiniBand architecture, one of its advantages is that the higher data transmission speed than that of existing bus technologies is provided in contrast with existing PCI bus architectures, and another advantage is that the peer to peer exchange model is used for data transmission instead of the shared bus model. The biggest disadvantage of shared bus model is that when a device occupies the bus, other devices can wait only. Most of current PCI devices are Master devices, or so-called Masters; the difference between Mater and Slave devices is that Master will send out inquiry signals constantly to detect whether the bus is occupied, it will apply to the bus controller to occupy the bus for some time if the bus is free and it just wants to use the bus, which reduces operation efficiency of system.
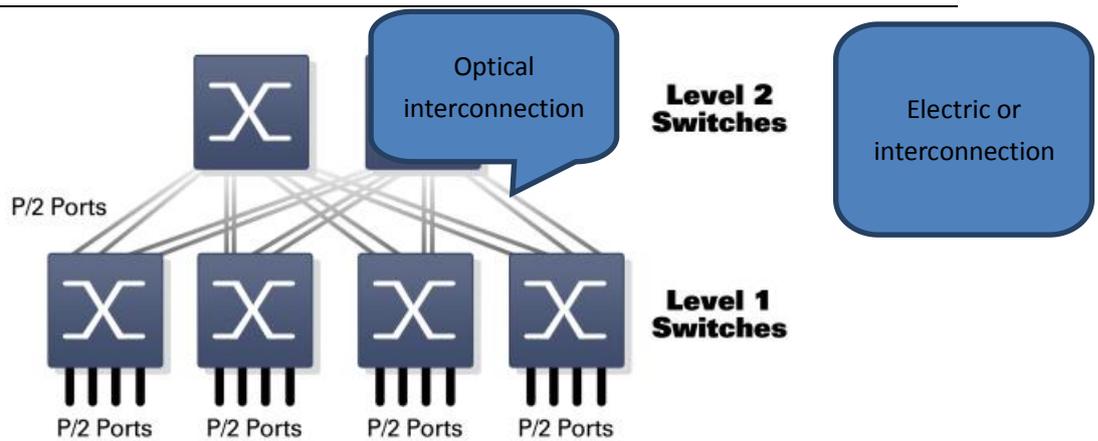
InfiniBand usually uses non-blocking Fat-Tree architecture for interconnection and the traditional 3-tier data center architecture is become 2-tier architecture of Top of Rack (TOR) and aggregation layers (convergence layer).

Infiniband technology development roadmap and main equipment supplier list of InfiniBand exchange (Gigalight product can be compatible with Mellanox, Intel, IBM and other InfiniBand exchange equipment of main equipment suppliers)

深圳市易飞扬通信技术有限公司
Shenzhen Gigalight Technology Co., Ltd
Optical interconnect technology innovator
www.gigalight.com

On February 23, 2016, the Flemish Supercomputer Center (VSC) in Belgium adopted Mellanox end-to-end 100Gb/s EDR interconnection solution and also integrated it with new LX supercomputers of NEC Company. The system would be the fastest supercomputer (Tier-1) and the first complete end-to-end EDR 100Gb/s InfiniBand system of Belgium in the future and it also proved that the global deployment amount of EDR InfiniBand technology had grown increasingly.

Facing Supercomputer Center InfiniBand interconnection applications, Gigalight can provide a complete product line and a operating rate covering QDR, FDR and EDR and provide different plan combinations (high-performance cable, optic cable AOC and optical module are optional) based on actual application scenarios, which can provide great convenience to the solution selection of clients; the product line is as follows:

Gigalight product model	Operating rate	Encapsulation mode	Type	Working distance
GQS-PC400-0xC	40GE, QDR	QSFP+ DAC	Copper Cable	5m
GQS-AC400-0xC	40GE, QDR	QSFP+ Active Copper Cable	Copper Cable	10m
GQS-MDO560-XXXC	56GE, FDR	QSFP+ AOC	Optical fiber 850nm	150m(OM4)
GQS-MDO400-XXXC	40GE, QDR	QSFP+ AOC	Optical fiber 850nm	400m(OM4)
GM-SDO400-XXXC	40GE, QDR	QSFP+ AOC	Optical fiber 1310nm	2km
GQS-MPO400-SR4	40GE, QDR	QSFP + Optical module	Multiple-mode 850nm	400m(OM4) 8-core
GQS-SPO400-LR4C	40GE, QDR	QSFP + Optical Module	Single-mode LR4	10km 2-core
GQM-SPO400-IR4C	40GE, QDR	QSFP + Optical Module	PSM Single-mode 1310nm	2km 8-core
GQM-SPO400-LR4C	40GE, QDR	QSFP+ TransceiverQSFP + Optical Module	PSM Single-mode 1310nm	10km 8-core
GQS-PC101-0XXC	100GE, EDR	QSFP28 DAC	Copper Cable	3m
GQS-4P28+PC-XXC	100GE, EDR	QSFP28 to 4x SFP28 DAC	Copper Cable	3m
GQS-MDO101-XXXC	100GE, EDR	QSFP28 AOC	Multiple-mode 850nm	100m(OM4)
GQP-MDO101-XXXC	100GE, EDR	QSFP28 to 4x SFP28 AOC	Multiple-mode	100m(OM4)

			850nm	
GQS-MPO101-SR4C	100GE, EDR	QSFP28 Transceiver QSFP28 Optical Module	Multiple-mode 850nm	100m(OM4) 8-core
GQM-SPO101-IR4C	100GE, EDR	QSFP28 Transceiver QSFP28 Optical Module	Single-mode PSM 1310nm	2km 8-core
GQS-MPO101-IR4C	100GE, EDR	QSFP28 Transceiver QSFP28 Optical Module	Single-mode CLR4	2km 2-core
GQS-MPO101-LR4C	100GE, EDR	QSFP28 TransceiverQSFP28 Optical Module	Single-mode LR4	10km 2-core