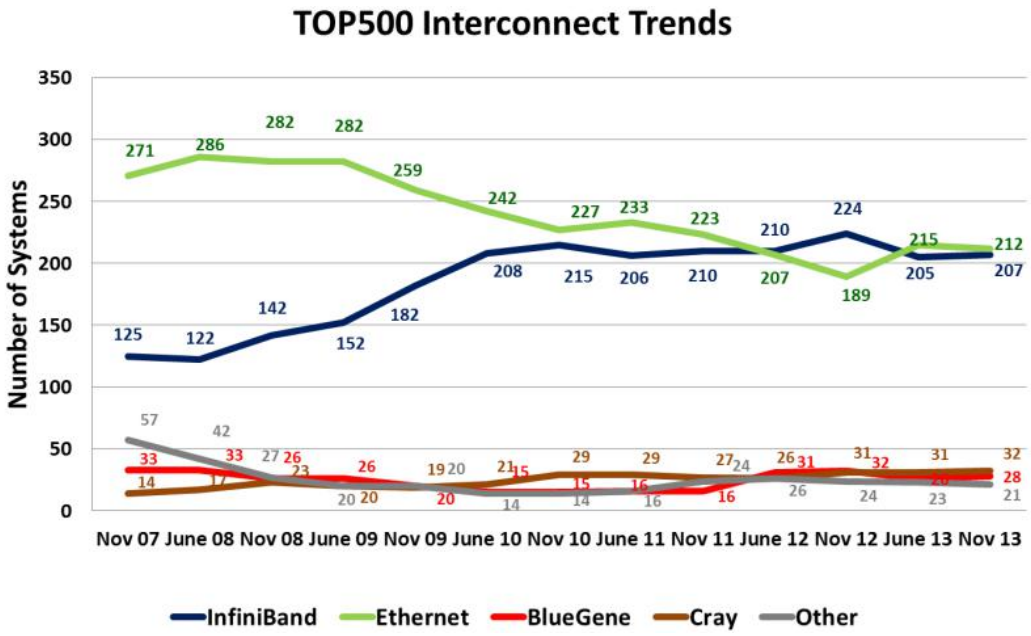
超算中心 InfiniBand 网络和易飞扬互连解决方案

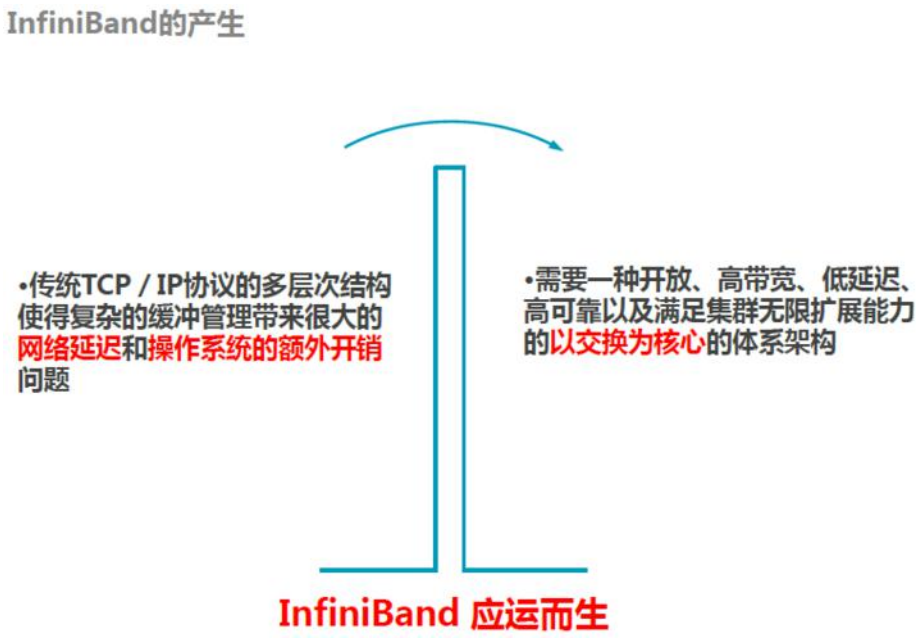
高性能计算(High Performance Computing，以下简称 HPC)，即超算中心 HPC 市场正在朝着使用异质计算系统和提高能效比的方向发展：GPU、DSP 和 ARM 处理器同时运行，以实现用更少的能耗带来更高的 Petaflop(1 千万亿次浮点计算/秒)值。IDC 预测，HPC 服务器市场要实现在 2017 年达到 150 亿美元销售额的目标，每年的增速就必须要保持在 7.3%左右。超级计算(Supercomputing)的应用范围非常广泛，包括汽车制造模拟、天气预报、分子生物学研究、地球物理学等，在这些领域中往往需要并行计算和处理大量的数据、进行复杂的运算。最近经常被讨论的大数据分析，也会用到超级计算。

高性能模拟需要最高效的计算平台。**计算集群（Cluster）**技术平台由于其出色的生产力和灵活性，现已成为 HPC 模拟最常用的硬件解决方案。而计算集群 **Cluster** 技术平台通常采用以下高速互联技术进行沟通，如：InfiniBand、高性能 Ethernet(精简过的帧结构)、BlueGene、Gray 等互联技术，其中 41.2%的超算中心采用了 infiniBand 互连技术。



超算中心 Top500 中还有不少系统采用了以太网或者 Cray 的互连技术，但正是这些独特的优势让 infiniband 在超算中大行其道，据统计，全球超算中心 Top500 榜单中，有 41.2%的系统采用了 infiniband 互连技术。

InfiniBand 诞生的缘由：

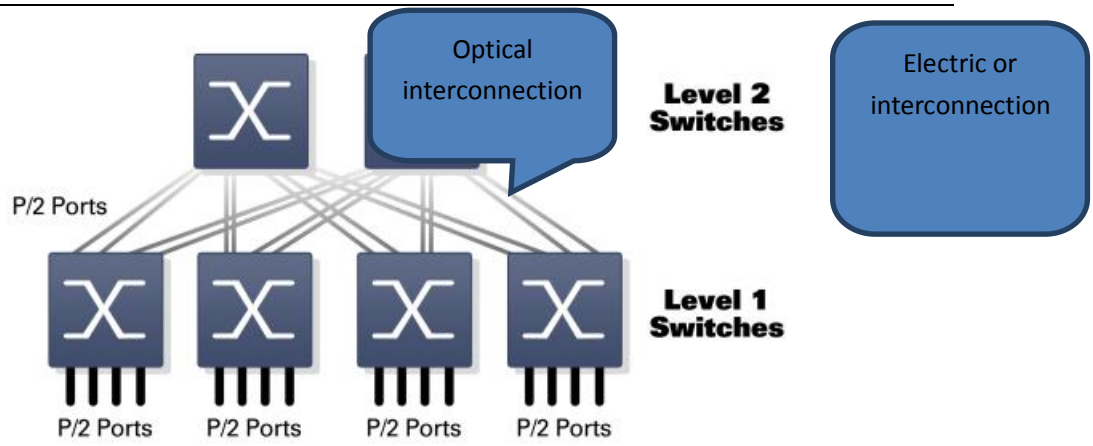


InfiniBand 是一种输入输出（I/O）宽带结构，可以提高服务器各设备之间、网络子系统之间的通信速度，为将来的计算机系统提高更高性能和无限扩展性的宽带服务。InfiniBand 技术不是用于一般网络连接的，它的主要设计目的是针对服务器端的连接问题的。因此，InfiniBand 技术将会被应用于服务器与服务器（比如复制，分布式工作等），服务器和存储设备（比如 SAN 和直接存储附件）以及服务器和网络之间（比如 LAN，WANs 和 the Internet）的通信。超算中心是一个庞大的系统，正因如此，超算系统的建设并非是简单的硬件堆砌，除了 CPU、GPU 等核心部件，网络互联技术也是决定超算计算能力和计算效率的重要部分。

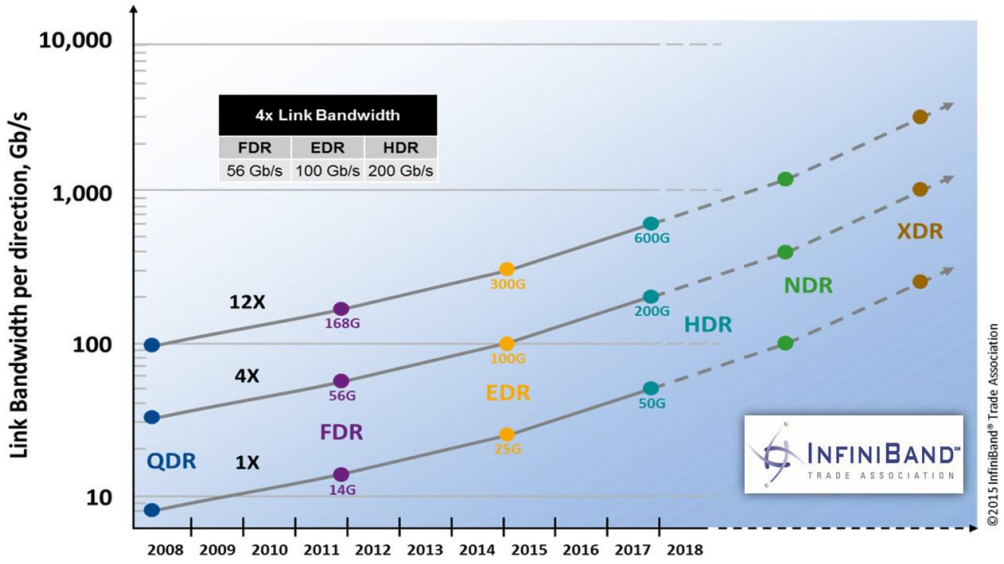
InfiniBand 的优势

对于 InfiniBand 架构来讲，提供比现有总线技术更高的数据传输速度是它相对于现有 PCI 总线架构的优势之一，它的另外一个优势就是采用了点对点的交换模式，而不是共享总线的模式对数据进行传送。共享总线模式的最大缺点在于当一个设备占据总线时，其它设备只能等待。目前的 PCI 设备大部分是 Master 设备，或称之为主设备，与 Slave（从）设备的区别在于，它会不断地发出询问信号探测总线是否有人占用，如果总线空闲，而它又恰好要使用，它会向总线控制器申请占据总线一段时间，降低了系统的工作效率。

InfinBand 通常采用非阻塞的 Fat-Tree 架构进行互连，从传统的 3-tier 数据中心架构变为 Top of Rack (TOR)和 aggregation layers（汇聚层）的 2 层架构。



Infiniband 技术发展线路图和 InfiniBand 交换机主流设备商名单，（易飞扬产品可以兼容 Mellanox, Intel, IBM 等主流设备厂家的 InfiniBand 交换机设备）



2016 年 02 月 23 日，位于比利时的佛兰德超级计算中心（VSC）已采用 Mellanox 的端到端 100Gb/s EDR 互连解决方案，将其与 NEC 公司的全新 LX 系列超级计算机相集成。该系统未来将成为比利时最快的超级计算机(Tier-1)以及第一个完整的端到端 EDR 100Gb/s InfiniBand 系统，同时它也是 EDR InfiniBand 技术在全球部署量日益增长的另一大例证。

面对超算中心 InfiniBand 互连应用，易飞扬可提供完整的产品线，工作速率覆盖 QDR、FDR 和 EDR 根据实际应用场景，可提供不同的方案组合（高性能电缆、光缆 AOC 和光模块都可选），为客户的方案选择提供了极大的便利，产品线如下所示：

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 850nm	100m(OM4)
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