

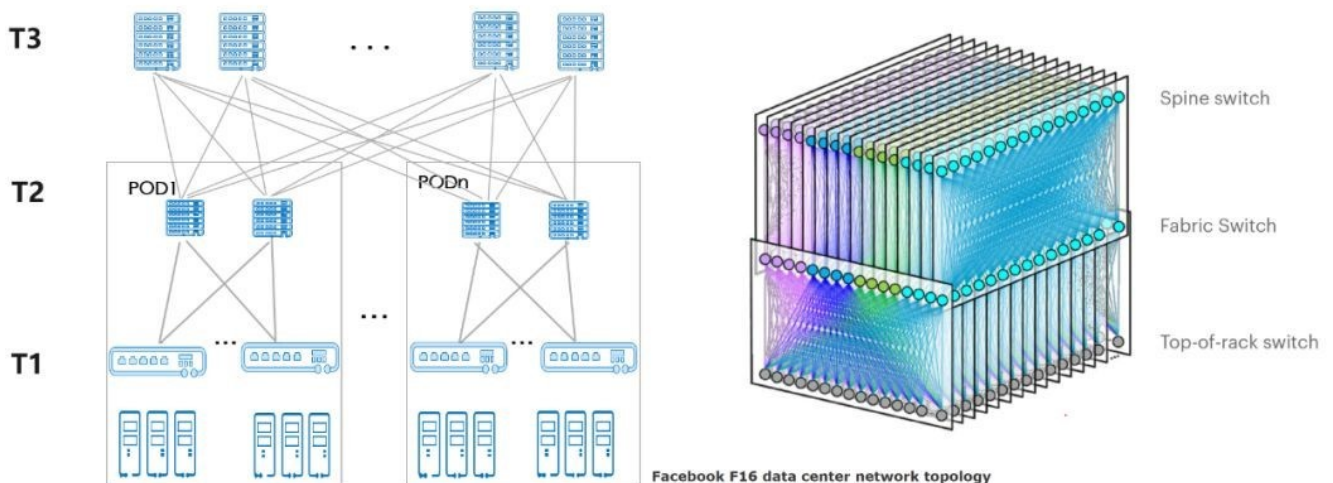
200G vs 400G: Who is the next choice for data center networks?

The Internet connects more than 4 billion users around the world, supporting an endless stream of digital applications such as VR/AR, 16K video, automatic driving, Artificial Intelligence(AI), 5G, Internet of Things(IOT), etc. The combination of online and offline users such as education, medical treatment and office is affecting and changing all aspects of people's life.

As the infrastructure for the survival and development of Internet business, data center network has already moved from the initial Gigabit and 10 Gigabit network to the stage of scale deployment of "25G access + 100G interconnection".

1. 100G Interconnect: Full-box architecture valued by large Internet companies

Under the architecture of "25G access + 100G interconnection", the data center network realizes large-scale access through three-level networking, and the scale of single cluster server can exceed 100,000. As shown in the figure below, the pod based on T1 and T2 layers can be flexibly expanded like Lego blocks and built on demand.



As the capacity of high-capacity forwarding chips increases and the cost of 100G optical interconnects decreases, a full-box device networking solution for building 100G interconnects

with single-chip switch devices has emerged in the market. This single-chip multi-plane interconnection solution is typically represented by a 12.8T chip, which can provide 128x100G port density on a single chip, and a single POD can provide access to 2,000 servers.



Typical 128 port 100G high density box switch

Compared with the traditional box device networking scheme, although the number of network nodes and the optical interconnection modules between devices have increased, resulting in the increase of operation and maintenance workload, the introduction of high-performance forwarding chip has effectively reduced the single bit cost of the network port of the data center, which is very attractive to large Internet enterprises. On the one hand, large Internet enterprises quickly introduce 100G full box architecture to reduce the cost of network construction. On the other hand, based on their strong R&D ability, they improve the level of network automation

deployment and maintenance to meet the challenge of increasing operation and maintenance workload.

Therefore, large Internet enterprises tend to have the same ideas for 100G network solutions, and the full box equipment networking has become the base for the evolution of 100G network architecture.

2. 200G or 400G

The "25G access + 100G interconnection network" solution has led to the unification of chip selection and rapid uptake, fully illustrating the rapid evolution of IDC network architecture driven by the technology. With the launch of single-chip network products, the 100G technology has also been fully captured. In the current context of continuous and rapid business development, bandwidth upgrade has become inevitable. A selection problem is placed in front of enterprises: choose 200G or 400G?

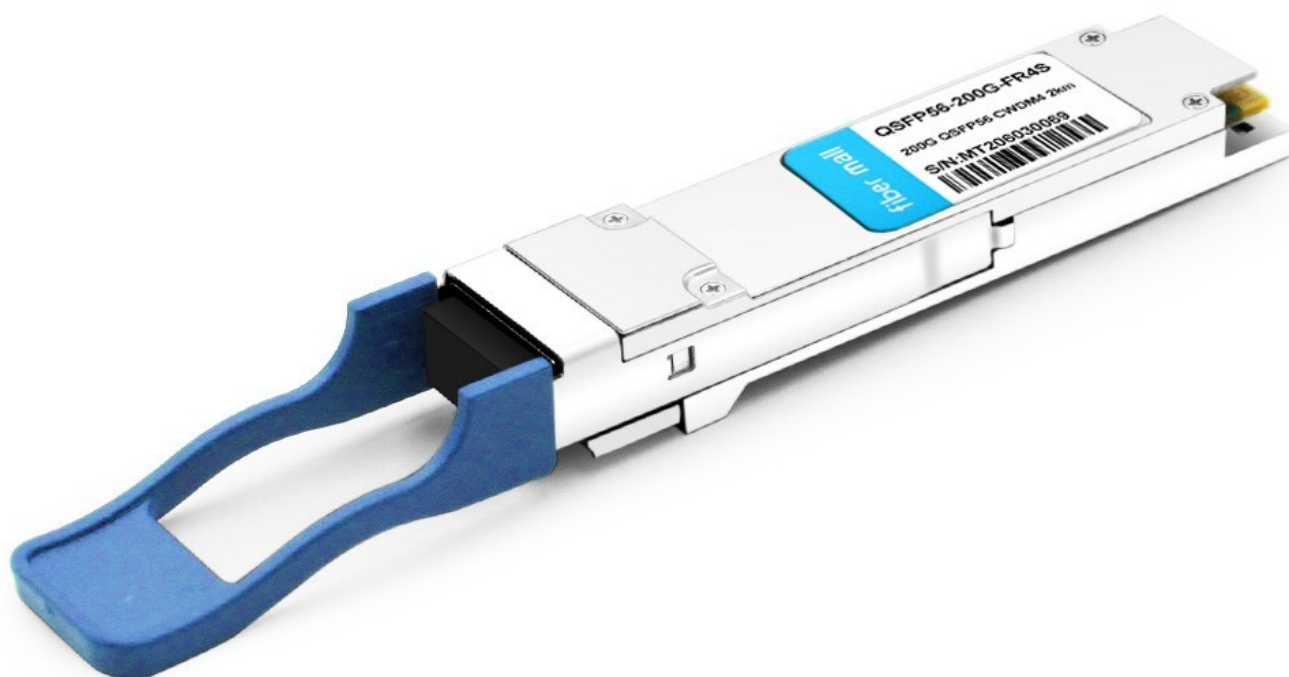
The network never exists in isolation, and the environment of the industry is the big soil that determines whether the technology can grow and mature. Let's look at the current status of the 200G and 400G industries from three aspects: **network standards**, **servers** and **optical modules**.

? 200G vs 400G: the protocol standards are mature

The 200G standard was launched later than the 400G standard in the evolution of IEEE protocol standards.

The IEEE 802.3 Ethernet Working Group established the 400G standard in 2013 after completing the BWA I (Bandwidth Assessment I) project investigation. 2015, in order to further expand the market scope to include 50G server and 200G switch specifications, the IEEE established the 802.3cd project to initiate the development of the 200G standard.

Because of the relevance of 200G and 400G specifications, the 200G single-mode specification was eventually incorporated into the 802.3bs project. By then, 400G had basically completed the main design of PCS, PMA, and PMD, and the 200G single-mode specification was generally based on the 400G single-mode specification by half, such as 200G QSFP56 optical transceiver.



Fiber Mall 200G QSFP56 FR4

On December 6, 2017, IEEE 802 finally approved the IEEE 802.3bs 400G Ethernet standard specification, containing both 400G and 200G Ethernet single-mode. IEEE 802.3cd defines the standard for 200G Ethernet multi-mode, which was officially released in December 2018.

As seen in the table below, 400G has achieved standard support for the full range of application scenarios, including 100m, 500m, 2km and long range 80km.

| Distance | Standard | Name | Rate | |
|----------|-----------------|------------|---------------|----------|
| | | | Electric port | SFP port |
| 100m | IEEE 802.3cd | 200G SR4 | 4 x 56G | 4 x 50G |
| | IEEE802.3cm | 400G SR8 | 8 x 56G | 8 x 50G |
| | | 400G SR4.2 | 8 x 56G | 8 x 50G |
| 500m | IEEE 802.3bs | 400G DR4 | 8 x 56G | 4 x 110G |
| 2km | IEEE 802.3bs | 200G FR4 | 4 x 56G | 4 x 50G |
| | IEEE 802.3bs | 400G FR8 | 8 x 56G | 8 x 50G |
| | 100G Lambda MSA | 400G FR4 | 8 x 56G | 4 x 100G |
| 6km | 100G Lambda MSA | 400G LR4 | 8 x 56G | 4 x 100G |
| 10km | IEEE 802.3bs | 400G LR8 | 8 x 56G | 8 x 50G |
| 80km | OIF | 400G ZR | 8 x 56G | DP-16QAM |

? 50G vs. 100G servers: 100G servers will become mainstream

By the end of 2019, both 50G and 100G NICs have started shipping. there was an industry-wide swing in 2018 and 2019 on next-generation upgrade options for 25G NICs. In 2019, the shipments of 50G and 100G server had a reversal. But after 2020, the orders and shipments of 100G servers have surpassed 50G servers, and the industry has finally regained the confidence in 100G servers.

Turn to the CPU chip, the two mainstream manufacturers “company I” and “company A” have launched new products one after another. The new chips of “company I” supporting PCIe 4.0

has been launched in Q3 2020 while I/O reaches 50G for mainstream applications and 100G/200G for high-end applications. From what we can see, the QSFP56 is on mass production now. The two giants are expected to launch chips supporting PCIe 5.0 respectively in mid-2021, which will again raising mainstream I/O to 100G and high-end applications to 400G.

Therefore, both CPU chips and server shipments show that 50G is a flash in the pan, and 100G servers are rapidly becoming the mainstream.

? 200G vs 400G optical module: 400G has better cost and more mature industry

The data center access server evolves from 25G to 100G, so should the current 100G Internet choose 200G or 400G optical modules?

Usually, when data centers evolve from 10G servers to 25G and network interconnection is upgraded from 40G to 100G, the network bandwidth doubles, but the interconnection cost and power consumption remain the same, i.e., the Gbit interconnection cost and power consumption drop by half. So, 100G replaces 40G as the mainstream network interconnection solution in 25G era.

200G and 400G optical modules are not quite the same as others. The conventional optical modules use NRZ (Non-Return-to-Zero) signaling technology, which uses two signal levels, high and low, to represent the 0 and 1 of digital logic signals, and can transmit 1 bit of logic information per clock cycle. The 200G and 400G optical modules both use the higher-order modulation technology, namely, PAM4 (Pulse Amplitude Modulation 4). The PAM4 signal uses four different signal levels for signaling, and can transmit 2 bits of logic information per clock cycle, i.e., 00, 01, 10, and 11.

| Optical module | 200G | 400G |
|--------------------------|-------------|-------------|
| Modulation mode | PAM4 | PAM4 |
| Implement scheme | 4 x 50G | 4 x 100G |
| Cost | Large | large |
| Power consumption | Medium | Medium |

Therefore, under the same baud rate, the bit rate of PAM4 signal is twice that of NRZ signal, the transmission efficiency is doubled, and the Gbit cost is effectively reduced. From the perspective of optical module composition, 200G and 400G modules adopt the mainstream architecture of 4-lane, so these two's design cost and power consumption tend to be the same.

There are only two types of modules for 200G: 100m SR4 and 2km FR4, of which 100m 200G SR4 is currently available in the market from only a few suppliers. On the contrary, 400G has 5 types of modules, like QSFP-DD. and the top 8 vendors in the market have laid out 100m, 500m and 2km modules. 400G has a much more mature industry than 200G, and customers have more choices. This result illustrates the technical costs of cost and power consumption for 200G optical modules due to the introduction of PAM4 technology. The industry is eager to move beyond 200G to 400G to reduce this cost in the data center network segment. Consequently, 400G module with the same technology and cost composition is more competitive in terms of evolution.

On the whole, 400G has obvious momentum, and 200G may become a temporary transition or be skipped directly.