Cloud DC Networking: Are You Selecting the Right Network Device?

Ray Mota
ACG Research

EXECUTIVE SUMMARY

Cloud providers are faced with multiple challenges: On one hand, there is the growing challenge of scaling their data centers (DCs) because of an exponential increase in bandwidth. On the other hand, they must deal with the growing CAPEX/OPEX associated with their expansions. This increase in bandwidth comes from the introduction of new bandwidth-intensive services based on AI/ML and big data; an increase in the demand for high-definition video content/online gaming; and the need to establish DCs closer to the users to offer bandwidth-intensive and latency critical services such as AR/VR.

The networking layer that includes switches and routers is part of the DC expansions; however, a wrong selection in a network device can increase the CAPEX/OPEX and hinder network agility. It is very important to evaluate and select the right network device at times of expansions and technology refresh. The following are the five critical choices and recommendations for each:

- Sustainability through power and space savings
- Scalability for higher speeds
- Openness, consistent, and programmable
- Agility through zero-touch-automation
- Simplicity through IP optics integration for DCI

Cloud provider's traffic is growing, but not without challenges

Cloud providers are facing scalability challenges like never before: and need to scale their networks to address the exponential growth in the bandwidth. This increase in the data is triggered by the growing use of bandwidth-intensive applications such as streaming video/online gaming, the need to offer new services such as AR/VR and CDNs at the edge, and, more significantly, the rise of AI and

ML based applications. The high volume and fast-moving data sets of these AI/ML applications are compute-intensive, which trigger expansion in servers as well as in the networking layer (routers and switches). With all these scaling needs, the cloud providers have to address the growing costs of CAPEX/OPEX.

At the same time, it is also true the CAPEX/OPEX of the networking layer is viewed differently from servers. Network owners view servers that run services as part of the profit center as they have tangible ROI and, therefore, investment in them can be justified easily. The networking layer is generally considered as part of the cost center, which is perceived to **always** add to the CAPEX/OPEX. It becomes very critical to select the right network gear to address the concerns related to CAPEX/OPEX.

With this in mind, the focus of this white paper is how to help network planners and decision makers of cloud providers make the right choice of networking gear to reduce costs and, at the same time, increase network agility and programmability. We list key recommendations to follow regarding network expansions and technology refresh. (We will refer to the switch and/or router as a network device for simplicity).

1. Sustainability through power and space savings

Because DCs consume considerable power, network devices need to be highly power and space efficient. There are two reasons why power and space efficiency are important.

First, it results in energy savings for the cloud providers, which is a direct financial benefit. DCs consume a significant amount of power, which results in high energy bills. Adding network devices that consume more energy to the DC means less energy left for servers. The same applies to space also. Bigger equipment also requires additional cooling, which results in additional power consumption. In edge DCs, the space and power are already limited, so a smart and power-efficient network device is ideal for such environments.

Second, there are worldwide initiatives to save the environment by reducing carbon emissions through sustainable green DCs that use power efficiently. For example, the Net Zero initiative by ESG aims at zero carbon emissions by using renewable energy in DCs.

DCs are power-hungry:

- They are projected to consume one-fifth of world energy by 2025
- According to ResearchGate, data centers will use approximately 3–13% of global electricity in 2030 compared to 1% in 2010

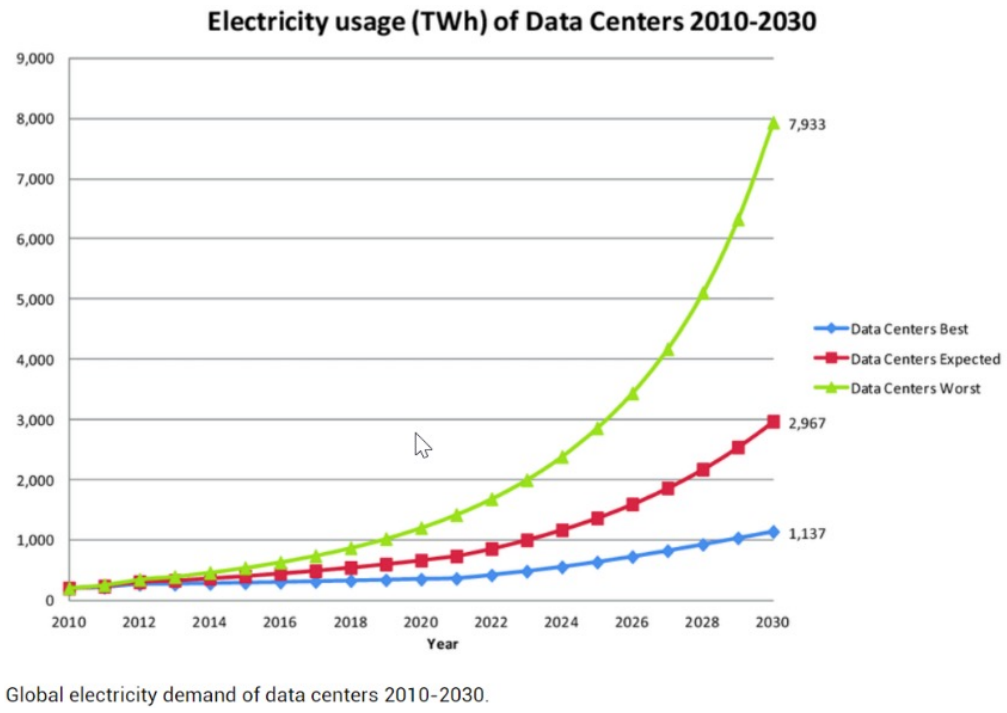


Fig 1. Global Electricity Demand of Data Centers (2010–2030) Ref: ResearchGate

Some vendors have already innovated in designing a very power-efficient device where each component of the network device, from ASICs to transceivers, from CPUs to I/O cards, are optimized for power consumption. Innovations have been made in the cooling and fan airflow design. All these initiatives lead to a smarter and greener DC.

We recommend opening a discussion with the vendor to understand if the vendor has done any innovation in this area and then calculate the power and space efficiency by using Watts/Gbit and RU/Gbit to select the best and most efficient network device.

Recommendation: For benchmarking the devices, Watt per Gigabit (W/G) and RU per Gigabit (R/G) can be used respectively for power and space. That is dividing the total power consumption and RU size for a fully loaded device divided by the total max capacity. This way, the power and space efficiency among devices can be compared, and the one providing the least values offers the most efficiency.

2. Scalability for higher speeds

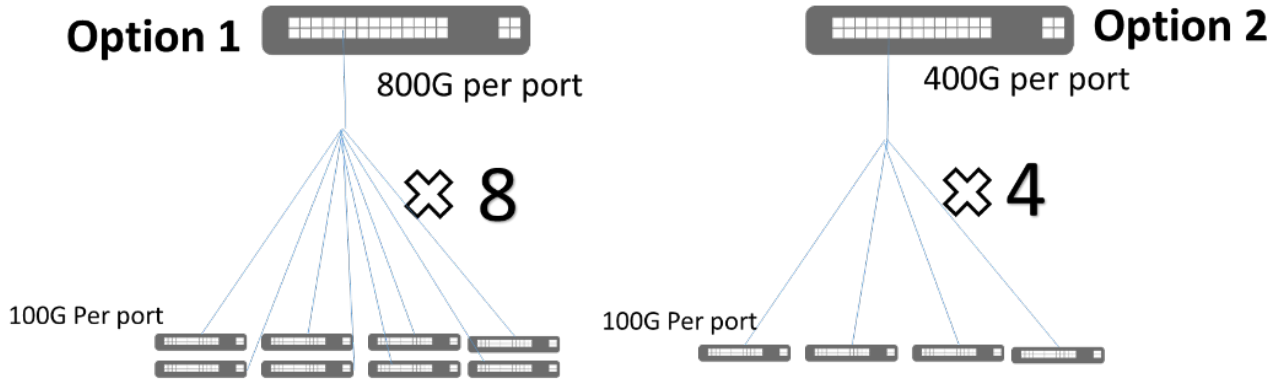
Every cloud provider is looking for a scalable network. However, this boils down to how much a network device is scalable.

Recommendation: Think beyond the current or near future requirements and go with a device that can provide maximum speed for the use case you need. If a port is used at maximum speed, the device is utilized at maximum efficiency.

Is it the right time to move to 800G? We do not see any reason why a cloud provider should not use 800G per port in the high-speed inter-switch links as that way, the device port efficiency and scalability can be used most efficiently.

The 800G per port is expensive, but it is less than eight times the cost of 100G or less than two times the cost of 400G. Effectively, this means that cost per bit decreases when you move to a higher speed and is also the case for power usage.

A good way to utilize 800G is using breakout to low speed, such as 100G. Figure 2 compares the same switch (1 RU) in two configurations. The switch uses a 25.6T ASIC and 32 ports by using 800G transceivers (Option 1) versus 400G transceivers (Option 2) fully loaded and breaking out each port to 100G ports for connecting to a lower speed.



Breakout Comparison for same Switch ASIC

	Option 1	Option 2
Switch ASIC	25.6 Tbps	25.6 Tbps
Number of ports	32	32
Operating mode	800G per port	400G per port
Total Breakouts	32 x 8 = 256 (100G ports)	32 x 4 = 128 (100G ports)
Power Consumption (Watt/ Gbit)	Lower	Higher
CAPEX /OPEX (Per Gbit)	Lower	Higher

Fig 2. CAPEX/OPEX Comparison of 800G versus 400G Ports

Comparing the two options, it is clear that Option 1 is highly efficient in power and CAPEX/OPEX:

- Option 1 can connect to more switches using the same bandwidth, which means more servers can connect
- The CAPEX/OPEX and power consumption per port for Option 1 is higher but dividing by capacity results in lower cost and better power consumption per Gbit

It is recommended that a scalable cloud provider data center procure switches with high-speed ASICs and utilize them to the full capacity.

Recommendation: For better scalability, utilize the full throughput of the switch ASIC available at hand. For example, a switch that runs a 25.6T ASIC can provide 32 x 800G/400G ports using compatible transceivers; use it at 800G per port instead of 400G per port. Inter switch links are the most suitable use case.

3. Openness, consistent and programmability

Openness and programmability are indispensable for a scalable and agile DC. Openness comes at multiple layers:

- An open system allows a cloud provider to choose the hardware or software. The network vendor should be able to offer both options to a cloud provider, hardware only and hardware combined with the vendor's OS. This kind of platform is called a disaggregated platform and has the benefit of reducing dependency on a single vendor for everything. With the hardware-only option, the end user can choose to run any third-party OS, including an open-source OS such as SONiC/FBOSS. Among them, SONiC (created by Microsoft) has become a popular open-source choice for NOS.

- The second layer of openness is how open and well defined is the network configuration APIs for the device. This kind of openness at the higher layer and related to network configuration; the device should be able to offer standard-based YANG models based on IETF and OpenConfig to make configuration from a management layer easier and quicker.

However, openness is not enough; it needs to be augmented with consistency. The SDK must be unified across all domains: DC fabric, core, and DCI. A vendor that offers a unified and consistent SDK exposing its hardware to run any OS brings operational simplicity and integration simplicity. In addition, consistency at the NOS level is essential. It is an operational nightmare to keep different operating systems (NOS) in the network across switches and routers. Upgrading firmware and troubleshooting issues can be complex and take up many resources. With the same NOS, the network operator can leverage existing automation tools and scripting and re-use the technical knowledge of network engineers across domains.

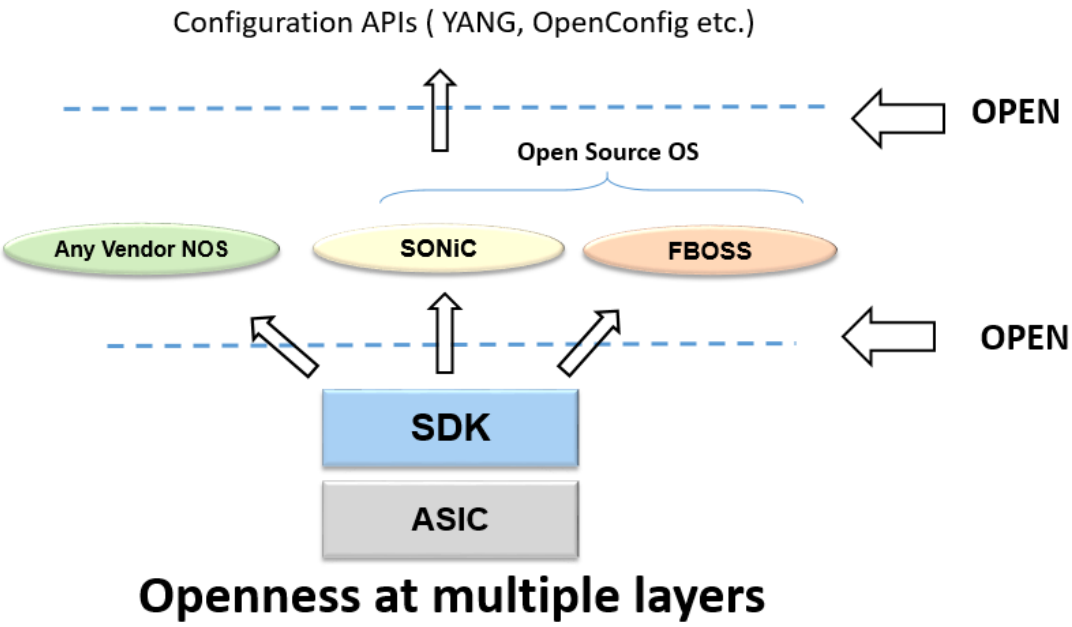


Fig 3. Openness in Network Devices

Another popular feature is P4 programmable hardware. With P4 programming, the processing pipelines in the network ASICs can be programmed to have new functions without waiting for the vendor to develop the next generation of ASIC. This is in contrast to fixed functions ASIC where pipelines are predefined and cannot be changed. P4 opens the door for innovation and creates new use cases such as inline telemetry, traffic engineering, and others without the need for a new ASIC.

Recommendation: A disaggregated platform with a unified and abstracted SDK that can integrate with any NOS is recommended. At a lower layer, it should be possible to program the ASIC using P4 programming. At a higher layer, it should be possible to program using open YANG and OpenConfig. Both SDK and NOS should be consistent in all parts of the domains to simplify the operation.

4. Zero-touch automation

The network of a DC is complex. With disaggregation, it becomes more complex as the cloud provider is expected to integrate new software independently. Doing such integration requires specific skill sets, which not every cloud provider may have. The device vendor shall provide all tools to help the end user simplify everyday operations. Zero-touch automation (ZTA) can help in such cases.

- With ZTA, network owners can automate all manual tasks, resulting in agile network operations

- ZTA is of two types: Zero Touch Provisioning (ZTP)and Zero Touch Operation (ZTO)

- With the ZTP, devices are provisioning automatically with zero-touch without waiting for weeks and months. This eliminates the need to send the installation team to the site. The device communicates with the central placed provisioning server. Upon authentication, it can start downloading a new image and initial configuration through a secure channel. All this happens without manual intervention.

Once a device enters the operational phase, it should be possible to automate life -cycle management tasks such as software upgrades, bug patching, auto tests.

Recommendation: Using ZTA expedites network provisioning and operation. Look for a vendor that provides comprehensive tools to facilitate ZTA.

5. Simplicity and cost saving with coherent optics integration for DCI

When it comes to Data centers, the lesser, the better. Less equipment and fewer interfaces save both CAPEX and OPEX.

Data centers are distributed; they require direct DCI connectivity using DWDM. Until now, the only way to connect DCs using DWDM was through the transponders, which is additional layer.

With the introduction of coherent ZR/ZR+ optics, a DCI router can provide a built-in coherent interface facing WAN, thus eliminating the need for DWDM transponders. This also simplifies the network operation.

The added advantage of using DWDM optics in routers is unified management of IP and optical domains, which brings operational simplicity.

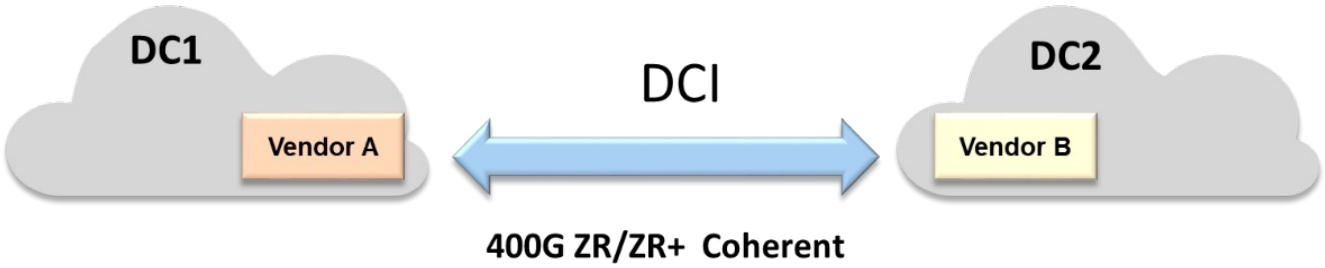


Fig 4. DCI Connectivity Using Coherent Optics

Recommendation: Using coherent interfaces on DCI routers reduces hardware inventory and results in cost savings and operational simplicity.

Conclusion and Take Away

When it comes to choosing the right network device, following some basic guidelines may help. We have seen cloud providers' quest to scale has resulted in networks that are poorly designed, OPEX intensive, and a roadblock to agility. Cloud providers can avoid this situation early by selecting the right networking infrastructure. The following summarizes the key features, benefits of the right network device, and how its impacts the cloud provider.

Feature	Benefits	What to Look for?	Impact
Sustainability	• Savings on power and space can result in a greener DC, better environment, and overall costs savings	• Minimum Watts/Gbit and minimum RU/Gbit for maximum loaded device	• OPEX savings
Scalability	• Use of max port rate results in cost savings • Lower power/bit • Lower space/bit • Lower OPEX/bit	• Port scalability of the switch matters, the higher the better	• OPEX savings
Open, Consistent and Programmable	• Open programmable and unified SDK allows easy and quick integration with any NOS and eliminates vendor lock in • A programmable and consistent NOS facilitates operational simplification across multiple domains	• Unified and consistent SDK with any NOS • Consistent NOS across any domain • Support of YANG and OpenConfig	• Agility and OPEX savings
Zero-Touch Operation	• Provisioning agility • Operational agility	• Tools provided by vendor to facilitate zero-touch provisioning and zero-touch operation	• Agility • OPEX savings
DCI IP Optics Integration	• Helps reduce hardware inventory, power consumption, and results in OPEX/CAPEX savings	• Availability of coherent, ZR and ZR+ optics on the router	• CAPEX/OPEX savings • Simplicity

ACG Research delivers information and communication technology market share/forecast reports, consulting services, and business case analysis services. Copyright © 2022 ACG Research. The copyright in this publication or the material on this website (including without limitation the text, computer code, artwork, photographs, images, music, audio material, video material and audio-visual material on this website) is owned by ACG Research. All Rights Reserved.