Data Center Evolution - Beyond 100G!

DIGISOL SYSTEMS LTD

This paper discusses both the need for high capacity data center infrastructures as well as the science that satisfies this insatiable hunger for bandwidth. It also presents IEEE's roadmap ahead that envisions ethernet as the future for networks of all sizes - even WAN.

Table of Contents

Data Center Evolution - Beyond 100G

Introduction

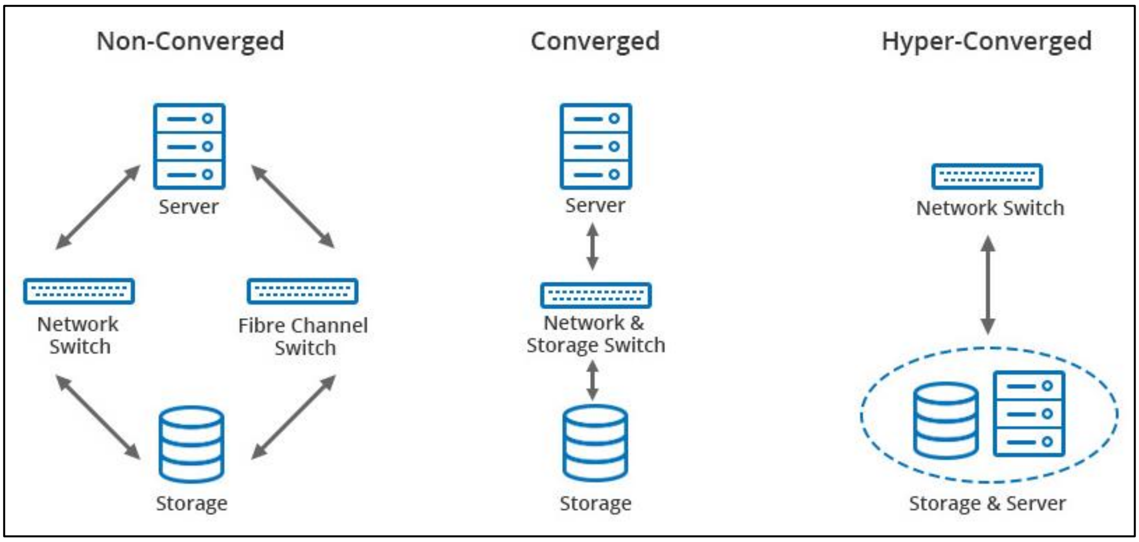
Over the past decade, Data centers have perhaps undergone the biggest and most dynamic of transformations. With over 8.5 million data centers worldwide as on 2017[1], everything we know about data center planning and strategy, the value stack they bring, and how they serve requirements is dramatically changing for the better.

Exponential, demand-driven increase in requirements for Cloud computing, Server Virtualization, Software Defined Networks, Network Functions Virtualization & more have catalyzed wholesome changes.

Converged and Hyperconverged Infrastructures

As virtualization and cloud computing grew, so did the need to move from a legacy (traditional) multi-vendor hardware environment to a more converged stack (single vendor) and finally to a hyperconverged environment that is now entirely operable through software.

Fig 1. Data center topology types[2]

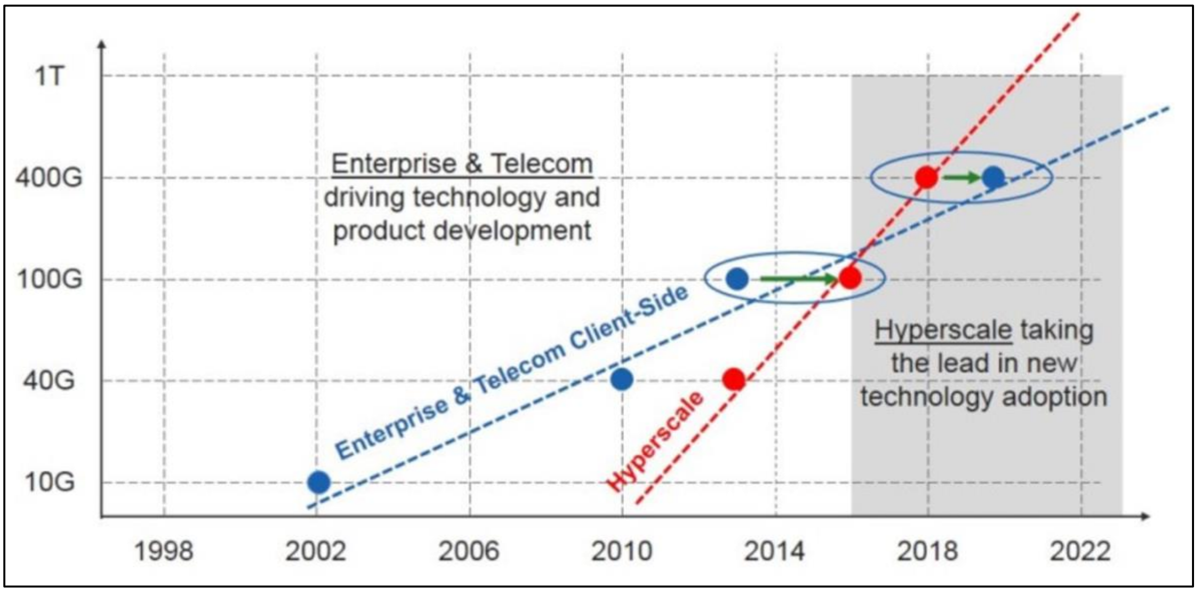


Converged Infrastructures (CI) refers to a hardware-oriented environment in which compute, storage, and networking all bought from a single vendor as a pre-packaged turnkey solution. Hyperconverged Infrastructure (HCI), on the other hand, uses a software-defined approach to decouple IT service delivery from its underlying hardware.

Both Converged and Hyperconverged setups are massive improvements over the legacy predecessors, and play a vital role in centralizing data center and server management. These also simplify the support, troubleshooting, and maintenance aspects of your data center investments.

Need for High Capacity Data Center Infrastructure

Fig 2. Telecom and Datacom adoption timelines[3]



The Global Cloud Index predicts that overall data center traffic will triple between 2016 and 2021[5]. Some of the significant factors that have the data centers of today as power-hungry as they are:

- Virtualization & Hyperconvergence - Server and Desktops
- Enterprise On-prem to Cloud Migration of hardware and applications (ERP, CRM)
- Vendor shift from one time buys to Subscription-based models for Cloud computing applications (for e.g., SAAS) - including HD streaming, Video conferencing, and Social media.
- Server hardware improvements: Improvements to server hardware such as faster processors, smarter NICs, FGPAs also drive up server capacity to handle[4].
- Hyper-scale architectures: For applications where compute and storage are independent variables and are modifiable to handle elastic workloads.
- The emergence of newer technology advancements such as Augmented Reality (AR), Telemedicine, IoT, Big Data, Smart Cities, Smart Vehicles - strongly data-driven requirements.
- Changed approaches: With the growing amount of data transferred across networks, more data centers are deploying storage on networks vs. servers. Ethernet technologies are being leveraged to attach storage to the network instead of legacy storage interconnects as the data center transitions from a traditional server model to networked storage.

Traffic Flow Analysis

An awful lot of bandwidth is required not only between the users and the data center but also in-between servers (East-West direction traffic) in the same data center or many a time, between data centers.

It is estimated that by 2021:

- Traffic from data centers to end users will comprise close to 15% of overall data center traffic.
- Traffic between data centers (replication, inter database links) will comprise close to 13.6% of the overall data center traffic.

- A staggering 72% of data center traffic will still be coming from within the data center e.g., storage, production, development, and authentication traffic.

There is expected to be 3X growth in global data center traffic[5].

Status Quo Bandwidth Capacities

Higher adoption of 25GbE over 40GbE

25GbE was an obvious choice considering lesser power requirement, less bulkiness, and overall cost-performance tradeoffs. It also assured the possibility of bundling multiple 25GbE channels to provide a combined bandwidth of 100Gbps.

100Gbps Capacities

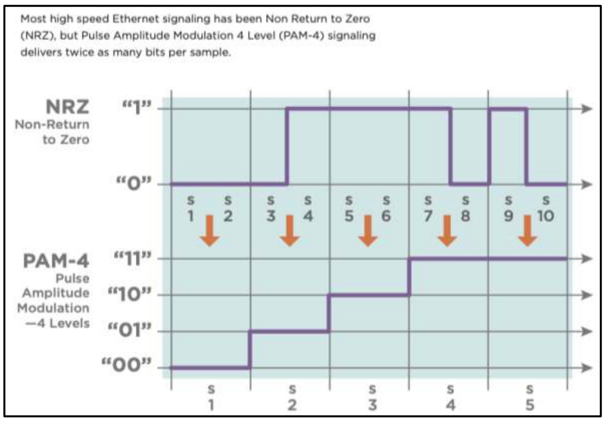
100Gbps technology was first implemented in data centers through QSFP28 modules. Standard 100G data center switches accommodate 32 QSFP28 modules in 1RU space, aligning with a 3.2 Tbps switching fabric ASIC.

200 & 400Gbps Capacities

Breaching the 100Gbps barrier needs a multifaceted approach as under:

- Higher-order modulation: moving from the present-day NRZ (Non-Return to Zero) modulation that transmits 1 bit per symbol to PAM4, which helps transmit 2 bits per symbol; effectively doubling the data rate without doubling the required overall bandwidth over conventional NRZ.

  Fig 3. NRZ vs. PAM4 Modulation.



  This change in modulation warrants an increased requirement for SNR (Signal to Noise Ratio) and, therefore, better linearity in PAM4 components in the modulated laser source and detector, higher extinction ratio, low RIN (relative intensity noise) excellent output power and higher responsivity. All this bearing in mind that hyper scale applications need enhanced loss budgets as opposed to the standard 3-5dB range.

- Components & Packaging: Existing setups use Directly modulated lasers (DML) and Externally modulated lasers (EML), and this works pretty well with NRZ modulation. However, in the process of moving from NRZ to PAM4 modulation, it is strongly recommended to use proven technologies that use Photon-Photon Resonance (PPR) effects between Distributed Bragg Reflector (DBR) and Distributed Feedback (DFB) structures.

- Conducive to Volume Manufacturing: 400G transceivers have to support low cost per bit and be able to scale to high-volume manufacturing. TOSA (Transmit Optical Sub-Assembly) and ROSA (Receive Optical Sub-Assembly) designs lend themselves to realize these benefits but with a caveat - only with the use of silicon photonics and complementary techniques within indium phosphide (InP). Additionally, silicon photonics are self-hermetic - hermetic packaging is a sealing to prevent environmental contamination.

  The overall idea is to merge industry-proven InP laser/modulator sources with wafer-level silicon photonics integration and packaging techniques to provide the required performance and low cost, power, and size. It will enable a business case for low-cost, large-scale 400 Gbps transceiver manufacturing.

Terabit Ethernet

Abbreviated as TbE, Terabit Ethernet refers to implementations exceeding 100 Gbps. Presently two standards are approved for implementation worldwide by IEEE P802.3bs Task Force.

- 200G/ 200GbE - 200 Gigabit Ethernet
- 400G/ 400GbE - 400 Gigabit Ethernet

IEEE 802.3cd standards provide detailed specifications for 200GbE.

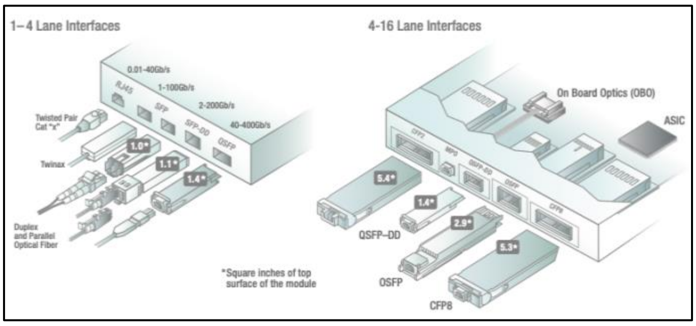
Fig 4. 200GbE Port types[7]

Name	Medium	Reach
200GBASE-CR4	Twinaxial copper cable	3m
200GBASE-KR4	Electrical backplane	
200GBASE-SR4	Multimode fibre 850nm laser	OM3: 70m OM4: 100m
200GBASE-DR4	Single-mode fibre WDM 1304.5–1317.5nm	500m
200GBASE-FR4	Single-mode fibre WDM 1271–1331nm	2Km
200GBASE-LR4	Single-mode fibre WDM 1295–1309nm	10Km
200GBASE-ER4		40Km
200GAUI-8	Chip-to-module/ Chip-to-chip interface	0.25m
200GAUI-4		

Fig 5. 400GbE Port types[7]

Name	Medium	Reach
400GBASE-SR16	Multimode fibre 850nm laser	OM3: 70m OM4: 100m
400GBASE-SR8		OM5: 100m
400GBASE-SR4.2		Uses 4 pairs of MMF: OM3:70m OM4: 100m OM5: 150m
400GBASE-DR4	Single-mode fibre WDM 1304.5–1317.5nm	500m
400GBASE-FR8	Single-mode fibre WDM 1273–1309nm	2Km
400GBASE-LR8		10Km
400GBASE-ER8		40Km
400GBASE-ZR		80Km
400GAUI-16	Chip-to-module/ Chip-to-chip interface	0.25m
400GAUI-8		

Fig 6. Common Form Factors in Ethernet Ports[8]

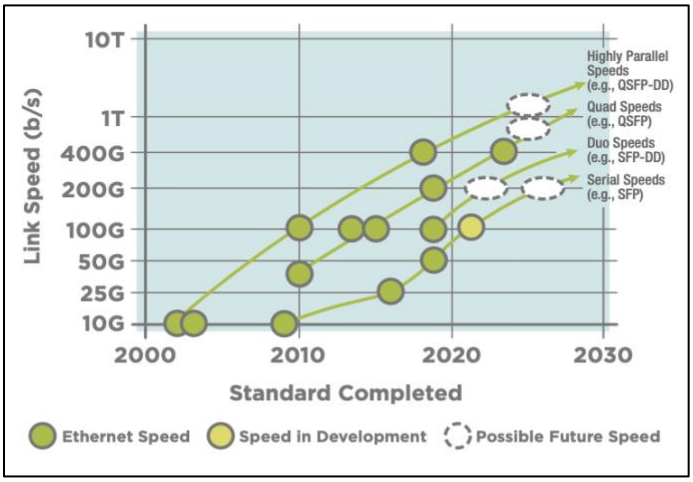


Hundreds of millions of RJ45 ports are sold a year, while tens of millions of SFP and millions of QSFP ports ship a year.

New form factors initially designed for 100GbE and 400GbE Ethernet ports are as shown. All have 4 or 8 lanes, and the OBO has up to 16 lanes. The power consumption of the modules is proportional to the surface area of the module.

IEEE's Long-Term Vision for Ethernet[8]

Fig 7. Ethernet Roadmap from Ethernet Alliance Group



Common Ethernet Interfaces Names & Media[8]

Fig 8. Ethernet Interfaces Depiction by Ethernet Alliance Group

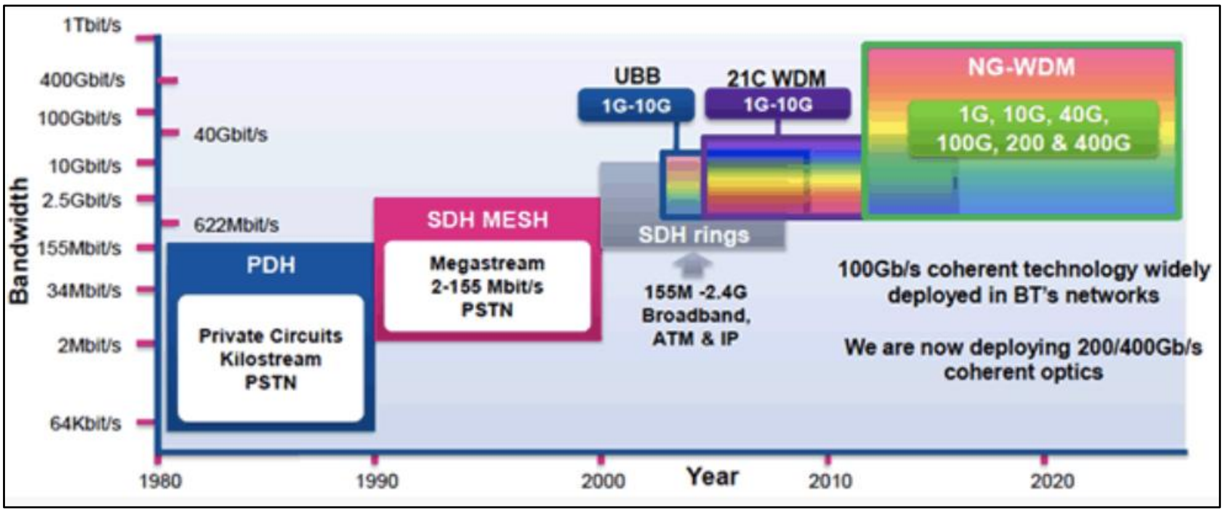
	Electrical Interface	Backplane	Twinax Cable	Twisted Pair (1 Pair)	Twisted Pair (4 Pair)	MMF	500m PSM4	2km SMF	10km SMF	20km SMF	40km SMF	80km SMF
10BASE-		T1S		T1S/T1L								
100BASE-				T1								
1000BASE-				T1	T							
2.5GBASE-		KX		T1	T							
5GBASE-		KR		T1	T							
10GBASE-				T1	T				BIDI Access	BIDI Access	BIDI Access	
25GBASE-	25GAUI	KR	CR/CR-S		T	SR			LR/ EPON/ BIDI Access	EPON/ BIDI Access	ER/ BIDI Access	
40GBASE-	XLAUI	KR4	CR4		T	SR4/eSR4	PSM4	FR	LR4			
50GBASE-	LAUI-2/50GAUI-2 50GAUI-1	KR	CR			SR		FR	EPON/ BIDI Access LR	EPON/ BIDI Access	BIDI Access ER	
100GBASE-	CAUI-10 CAUI-4/100GAUI-4 100GAUI-2 100GAUI-1	KR4 KR2 KR1	CR10 CR4 CR2 CR1			SR10 SR4 SR2	PSM4 DR	10X10 CWDM4/ CLR4 100G-FR	LR4/ 4WDM-10 100G-LR	4WDM-20	ER4/ 4WDM-40	ZR
200GBASE-	200GAUI-4 200GAUI-2	KR4 KR2	CR4 CR2			SR4	DR4	FR4	LR4		ER4	
400GBASE-	400GAUI-16 400GAUI- 8 400GAUI-4	KR4	CR4			SR16 SR8/SR4.2	DR4	FR8 400G-FR4	LR8 400G-LR4		ER8	ZR

Gray Text = IEEE Standard Red Text = In Standardization Green Text = In Study Group

Blue Text = Non-IEEE standard but complies to IEEE electrical interfaces

Real-World Implementations

Fig 9. British Telecom's view of the evolution of the core network. It is representative of worldwide trends[6]



While 100 Gbps is likely to fulfil LAN aspirations for several years to come, 200Gbps and 400Gbps are quickly being deployed in the WAN.

Multivendor Concerns

Ethernet Alliance, a global non-profit consortium dedicated to the continued success and advancement of Ethernet technologies, has conducted extensive testing across the 25GbE to 400GbE range as part of its High-Speed Networking (HSN) Plugfest in April 2019.

A total of 13 vendors participated in this showcasing their products ranging from electrical and optical interconnects, switches and NICs and cables across a range of form factors such as OSFP, QSFP, QSFP-DD. Results included Frame Error Rates (FER) tests with a 100% pass rate and functional interoperability tests scoring an aggregated 97.5% pass rate. It augurs very well for the future of high-speed ethernet, especially in a data center context.

Summary

Data centers continue to grow at an exponential rate, driven by business requirements and technology advancements. High capacity data center infrastructures are not just desirable but needed to support business aspirations today. Being able to roadmap this growth is the best way to plan and support its contribution.

References

1. Data Centers Worldwide - Statista
2. https://community.fs.com/blog/fs-s5800-8tf12s-switch-key-choice-for-hyper-converged-infrastructure-solution.html
3. https://www.lightwaveonline.com/data-center/article/16654650/data-center-interconnects-the-road-to-400g-and-beyond
4. https://www.lightreading.com/enterprise-cloud/network/always-hungry-feeding-data-center-bandwidth-demands-in-the-cloud-era/d/d-id/742592
5. Source: Cisco Global Cloud Index

6. https://www.connectorsupplier.com/networking-equipment-100g-and-beyond/
7. Wikipedia - Terabit Reference
8. Ethernet Alliance Ethernet Reference Document
9. Ethernet Alliance HSN Plugfest