Opinion Paper

Key Trends in Data Centre Architecture & Technology

Alan Flatman

Principal Consultant, LAN Technologies

Nexans

Nexans Cabling Solutions
March 2015

Key Trends in Data Centre Architecture & Technology

Alan Flatman, Principal Consultant, LAN Technologies

This paper has been prepared by Nexans with the permission of Dr Alan Flatman and is based on a presentation made at the Nexans Technical Conference, Paris, December 2014. Alan is the Principal Consultant for LAN Technologies and has over 40 years experience in the electronics and computer industries, advising on network technology and strategy since 1980. He represents the UK at both International & European cabling standards committees and is an active contributor to IEEE 802, providing the essential link between cabling and network technology groups as their liaison officer.

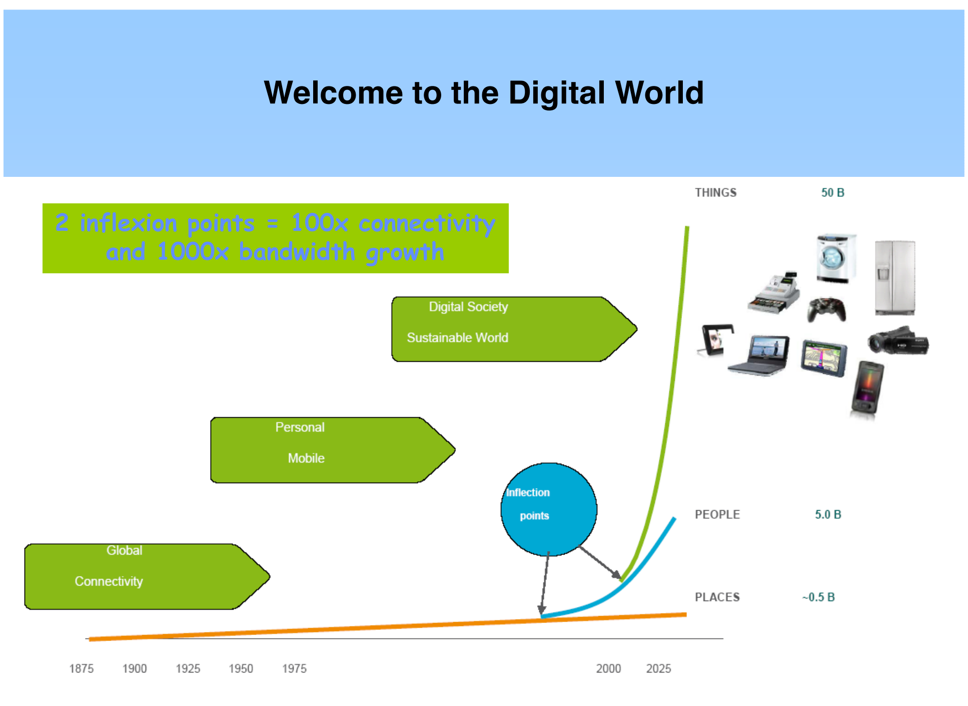
Some fifteen years ago, I was working on Ethernet - or 'High Speed Data Networks' as we called them at the time - and was first introduced to the concept of Data Centres. Since then, Data Centres have evolved enormously and are currently going through marked changes. In fact, I'd say Data Centres are in the middle of a storm. We're currently seeing tremendous growth, as well as vast opportunities. This places a great deal of strain on Data Centres when it comes to speed, power consumption, size, density and cost.

Today, I'd like to cover a number of topics related to these developments, starting with some primary market drivers. This will be followed by a brief overview of Data Centre types, shapes and sizes and a brief overview of high-level technology advances. Then, we'll look at some specific obstacles and examine how engineers need to balance speed, power consumption and size.

The forecasting challenge

First of all, I'd just like to illustrate how tough forecasting can be by sharing two quotes. The first is from1943, when Thomas Watson, chairman of IBM predicted: "I think there is a world market for maybe five computers." In 1977, during a talk to the World Society in Boston, DEC Chairman Ken Olson claimed: "There is no reason for any individual to have a computer in his home." Reasonable assumptions at the time – but wrong, as history has shown.

Now, I would like to present an overview of the 'digital world'. This image has actually been around for quite a while. It originally came from a self-titled 'futurologist' at Eriksson. This overview describes how the Internet has evolved from a limited number of physical places which were simply joined up, to a world where everybody can be wirelessly networked all the time. Interestingly – and accurately - he also predicted that - after places and people - there would be a network of 'things'.



Source: www.emc.com/leadership/programs/digital-universe.htm

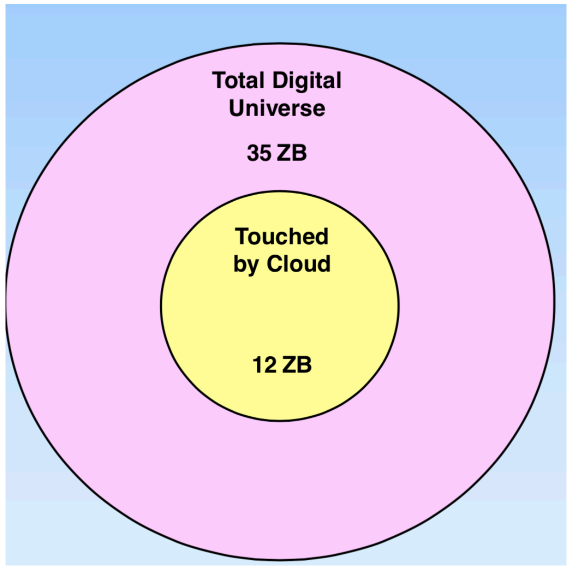
'The Digital Universe' study, which I'd like to look at now, represents five years of work looking into digital information growth to 2020. The research, sponsored by storage system supplier EMC and conducted by market analyst IDC, predicts data volume growth of 40-50% per annum.

Such shifts, or discontinuities, represent an enormous change in connectivity. This has been boosted by a factor 100 and bandwidth has grown by a factor 1000. That's absolutely terrific, in my opinion! The growth expected between 2009 and 2020 obviously puts tremendous pressure on existing infrastructure.

A new digital language

The development of the 'Digital Universe' has also brought us a new language. We used to think an Exabyte - 1,000 Petabytes - was a huge volume of information – but in 2010, global data production was already exceeding 1,000 Exabytes. A thousand Exabytes is one Zettabyte - or 2^{70} or 10^{21} Bytes. That's 143GB for each of the 7 billion people on the planet. Or 75 billion 16GB iPads – enough to fill Wembley stadium 41 times! Of all this data, 75% was produced by individuals and 25% was produced by machines.
Another interesting fact: by 2020, more than one third of the Digital Universe will either live in or pass through the cloud.

SI decimal prefixes – short scale			Binary usage	IEC binary prefixes	
Common Name	Name (Symbol)	Value		Name (Symbol)	Value
Thousand	kilobyte (kB)	10^3	2^{10}	kibibyte (KiB)	2^{10}
Million	megabyte (MB)	10^6	2^{20}	mebibyte (MiB)	2^{20}
Billion	gigabyte (GB)	10^9	2^{30}	gibibyte (GiB)	2^{30}
Trillion	terabyte (TB)	10^{12}	2^{40}	tebibyte (TiB)	2^{40}
Quadrillion	petabyte (PB)	10^{15}	2^{50}	pebibyte (PiB)	2^{50}
Quintillion	exabyte (EB)	10^{18}	2^{60}	exbibyte (EiB)	2^{60}
Sextillion	**zettabyte (ZB)**	10^{21}	2^{70}	zebibyte (ZiB)	2^{70}
Septillion	yottabyte (YB)	10^{24}	2^{80}	yobibyte (YiB)	2^{80}



Source: IDC Digital Universe Study sponsored by EMC (May 2010)

Let's look at growth rates as described in the IDC Digital Universe Study. The number of information containers - files, packets, images – which are managed, protected and stored in the Digital Universe is expected to grow 67 times, and the total amount of information will grow 44 times. The number of IT professionals in the world, however, will grow only 1.4 times.

According to the study, a huge amount of information is expected to be protected at some point, but there are no plans to do so right now. In 2020, 8 Zettabytes is expected to be protected, and 8.5 Zettabytes will be requiring protection. Note that the amount of unprotected data in 2020 equals the size of the entire Digital Universe in 2018! The decreasing cost of managing information will be an incentive to create more information, claim IDC. In 2009, protecting a Gigabyte of data cost around 5 cents. In 2020, that will be close to nothing. Hardly surprising that global data exchange will expand to some 35 Zettabytes!

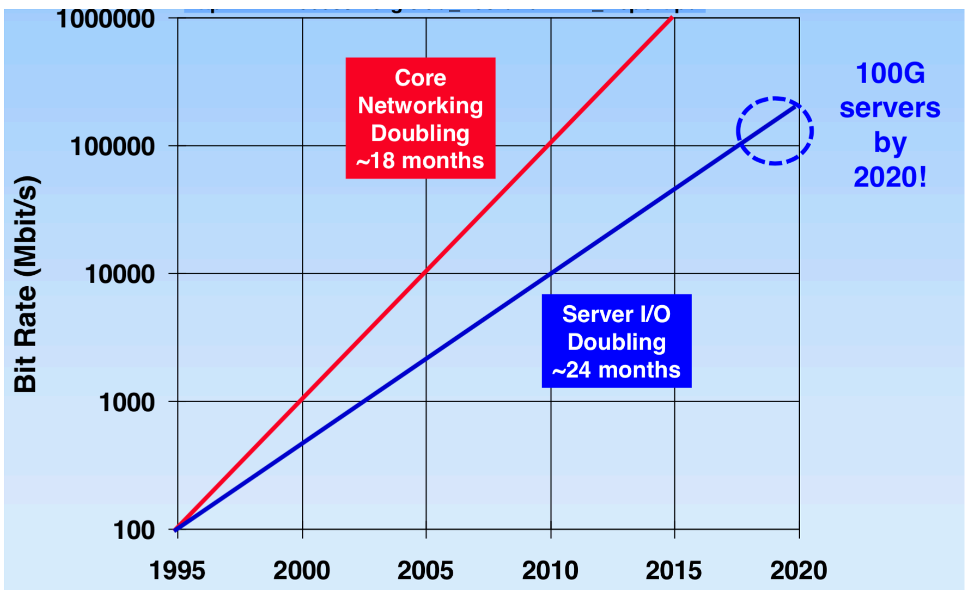
Growth in global IP Traffic and Users

My IEEE colleagues published a study some years ago, for which we spoke to end users in major sectors worldwide, such as science, healthcare, government and business. We mapped their predictions up to 2020 and placed these side by side. This allowed us to identify two major growth lines; core networking such as carriers and ISP links with doubling of bandwidth consumption every 18 months, and server I/O links with doubling of bandwidth consumption every 24 months.

That brings us to another highly interesting five-year study: Cisco's Visual Networking Index. This offers fixed and mobile data projections for the business and consumer segments, as well as global, regional and country-level usage trends. Global IP traffic is expected to reach 1 Zettabyte in 2015, which is eight times the amount of 8-times IP traffic generated in 2008 and equivalent to 28 million DVDs per hour.

Cisco sees a quadrupling in IP traffic from 2010 to 2015. There are several causes for this: a dramatic increase in connectivity, more Internet users, faster broadband speeds and richer media content. Data Centres are doubling at a slower rate than the number of users and the amount of data. That means we really have to work quickly to realize the deployments we need to satisfy demand by 2020.

Bandwidth Consumption Forecast



Source: Ethernet Bandwidth Assessment (IEEE 802.3, July 2012)

http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf

Data Centre Types, Shapes & Sizes

All data centres have three essential components: servers, which are computing engines with application software, storage devices, which can be tape- or disk-based or solid state, and a communications network. This last item can include structured cabling or point-to-point links as well as switches and different storage types: Local Area Network, Storage Area Network and Network Attached Storage.

Over time these elements have become more integrated and might fit into a single cabinet, whereas some years back, they could take up a whole building. There's a lot going on in the world of data centre architecture today. Just look at developments in colocation, cloud, in-house data centres and scientific systems. We're seeing the number of huge mega-DCs growing, alongside an increasing proliferation of small, fully functional, self-contained DCs. We're seeing different types of Data Centre for different applications: Public Cloud Providers (Amazon, Google), Scientific Computing Centres (National Labs) Colocation Data Centres (Private Clouds) and In-house Data Centres (Enterprise DCs) are a few examples. In 2008, there were practically no Cloud servers, and all servers were used for enterprise/premise purposes. By 2018, this division will be 50%-50%: there will be just as many cloud servers as enterprise/premise servers.

We're seeing greater diversity, as well as developments in the area of standardization. Cabinets have been made in standard sizes for years and server rows are strongly standardised. I've conducted a number of end user surveys to understand the physical dimensions of data centres and interconnections and discovered that rows always seem to be within 30 metres in length, which has become the 'de facto' standard for copper and fibre links. Inside the cabinet and between adjacent cabinets, links are typically limited to 5m.

In 2011, IEEE carried out a survey regarding link requirements in next generation data centres. Cable length isn't just increasing, speed requirements are, too. In DCs measuring up to 1 million square metres, such as Google's Mayes County facility, links might span 500 metres to 1 kilometre. Recently, we did a survey looking at Ethernet and the very largest data centers, up to 2.2 million square metres. Here, distance and speed requirements shot up even further with network links of 1000 metres to 2000 metres. The SuperNap, Las Vegas, Data Centre actually measures 2.2 million square metres.

How does this affect physical infrastructure? Looking at 'traditional' Data Centre Layout, we see storage is concentrated in a discrete SAN area and processing takes place in a separate server area. These connected to the main distribution area patch panels via switches. With the emergence of modular as well as Mega Data Centres operated by companies like Microsoft, facebook, IBM, Yahoo, Apple and Amazon, that's all changing.

Technology Advances & Obstacles: Speed Trends in Enterprise Server

I'd like to examine some trends in the technology of the server– the working engine of the DC. Between 2000 and 2015, central processor speeds increased. 45 times, DRAM storage 32 times and Network Ethernet link speed increased 400 times. The technology connecting components in the box became 60 times faster.

This brings us to today's biggest bottleneck: spinning disk technology, with its fairly long seek times long. This is a major obstacle to moving forward and must be replaced by Solid State Drives (SSDs), with seek times of microseconds(10^{-6}) compared with milliseconds (10^{-3})

Trends in Server Technology

	2000	2005	2010	2015
CPU	1 x Pentium 4 1.5 GHz	5 x Pentium D 2.6 GHz	15 x Nehalem Quad 2.6 GHz	45 x? Haswell 2.6 GHz?
DRAM	1 x DDR1	4 x DDR2	8 x DDR3	32 x? DDR4?
Network	1 x 100Mb Ethernet	10 x Gigabit Ethernet	100 x 10 Gigabit Ethernet	400 x 40 Gigabit Ethernet
Bus	1 x PCI 32-bit/33 MHz	15 x PCIe Gen1 x8	30 x PCIe Gen2 x8	60 x PCIe Gen3 x8
Fibre Channel	1 x 1GFC	4 x 4GFC	8 x 8GFC	32 x 32GFC
Disk	1 x 15K rpm hard drive	1 x 15K rpm hard drive	1 x 15K rpm hard drive	1 x 15K rpm hard drive

Source: Scott Kipp, Brocade

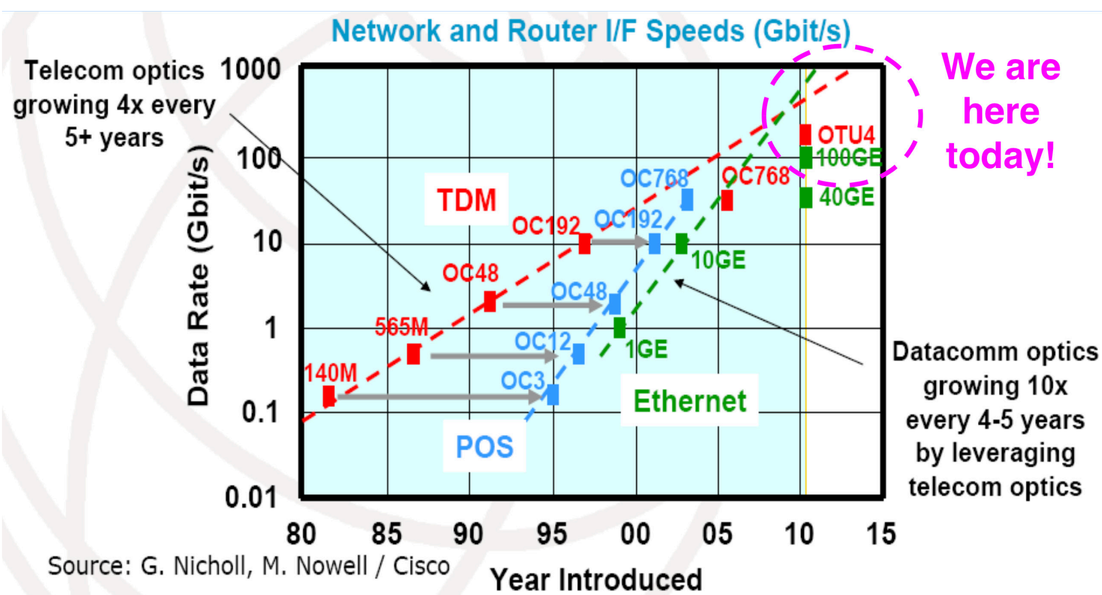
Trends in Data Centre Architecture

The pace of innovation in DC infrastructure is increasing and we've seen more innovation in the past five years than we've seen in the previous 15 years. These developments are having a marked effect on Data Centre infrastructure.

The three-tier architecture of a traditional DC works well for the equally traditional repetitive 'request for data – reply' model. This 'Tree' architecture is optimised for such 'North-South' traffic, with internet connectivity at the top, servers and switches at the bottom and a layer of switches in between.

In MSDCs, or Massively Scaled Data Centres, like those of Google and Facebook and cloud servers, a vast amount of traffic goes from server to server – what we call east –west traffic. As a result, MSDCs and cloud are moving toward a two-tier model which does away with the middle layer and is better suited to a combination of North-South and East-West traffic. Compressing the hierarchy from three to two tiers can bring massively improved speed and throughput.

Technologies For Next Generation Networks – and Technology *Exhaustion*



Today, we're seeing bandwidth demand exceeding technology innovation. Speed and density need to increase, but power usage and cost need to decrease, In the past, IEEE has often been able to reuse telecom technology, but this is no longer possible. We need to work together and find out how to keep up with today's and tomorrow's bandwidth demand. However, the type of risk normally associated with adopting totally new and untested advanced technology is not something we want data centres to take.

The short-term solution adopted at present is basically expansion of parallel connectivity. By using four lanes at 10G per lane, we can get 40 G. Four 25 G lanes allow us to reach 100G. 16 of these lanes get us to 400G. There are limits to this approach, though. We're waiting for a real breakthrough in optical technology, which will allow us to handle greater requirements. We need step-change improvements in optical technologies, enabling higher speeds of single and multiple wavelength laser devices. These must be translated to silicon photonics, or integrated optics which can evolve into optical backplanes and printed circuit boards. The required chips can be made, but this technology isn't yet ready for high yield, low cost, wide deployment manufacturing.

So how long must we wait? Please don't quote me on this, but I think that although we're more than a year away, we're not ten years away, either. I think this new technology will be deployed right throughout data centres, not just in server motherboards, but also in the system buss and in interconnects between boxes.

Balancing speed, cost, power and size

As engineers, we struggle to balance cost and size against speed and density. There often seems to be a tradeoff – you boost one, and you lose on another. Often, we end up making compromises on all four.

If we can get all this functionality on one switch, we have a solution. Switch faceplate density– physical AND bandwidth - is what it's all about. This must be approached with great care, though. We can combine different forms of transceiver modules on a single switch faceplate. Four ports CFP at 100G per port amounts to 400 Gb/s. 48 ports SFP+ at 10G per port gives us a total capacity of 480 Gb/s. 44 ports QSFP at 40G per port brings us 1,760 Gb/s. A CFP4 switch which will be available soon can get 32 100 G ports on a switch faceplate, giving us 3200 Gb/s

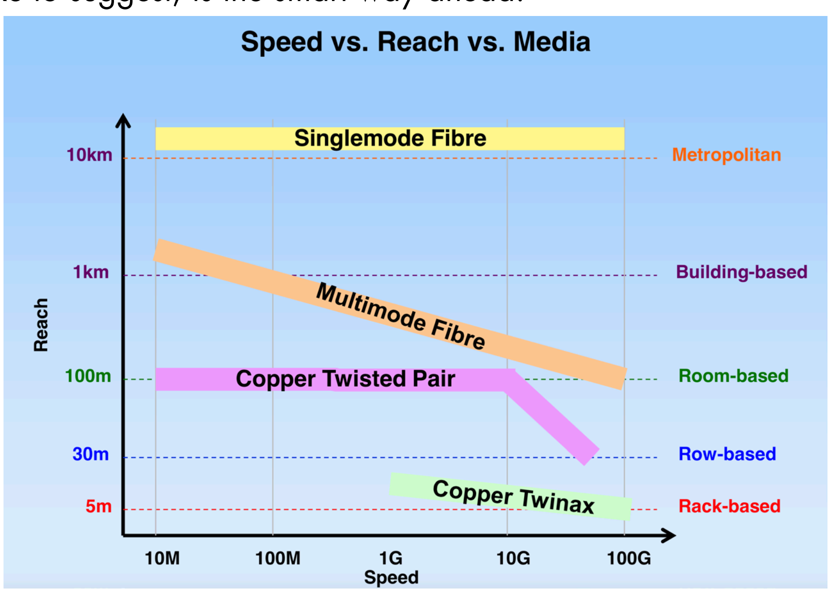
Data Centre Cabling Infrastructure

We can connect between cabinets in different ways, and we tend to use the lowest cost, most power-efficient, easy to use solutions. Inside cabinets, copper is mainly used. Copper twinaxial is perfectly good for short distances and bandwidth up to 100 Gb/s. The end of the road for that technology is in sight, though.

We usually employ twisted pair or optical fibres between cabinets. Billions of these links are deployed worldwide, but they're having a hard time when it comes to supported reach, as higher bandwidths have restricted their range from 100 metres to just 30 - which is still OK for end-of-row switching. Multimode is also taking a beating – every time required speed goes up, length goes down. Single mode fibre is a practical alternative but is considerably more expensive today.

So, to summarise: there are five things we want to keep as high as possible: speed, density, resilience (uptime), operational life and parallelism. However, there are four things we want to keep as low as possible: CapEx, OpEx, power consumption and latency. We should never underestimate the need for speed – especially in DC specs.
High density, small size and robustness are also vital to DC longevity. Low latency is also key, the need for the smallest possible delays between interconnect end points.

A survey I carried out last year focused on what big DCs wanted. During lengthy discussions with Microsoft, they explained they were implementing parallel single mode as far as possible, with short copper links. High capex, low opex and maximum lifespan. That, I would like to suggest, is the smart way ahead.



Nexans Cabling Solutions
Alsembergsesteenweg 2, b3 - B-1501 Buizingen
Tel: +32 (0)2 363 38 00 - Fax: +32 (0)2 365 09 99
Nexans Cabling Solutions UK and Intelligent Enterprise Solutions Competence Centre
2 Faraday Office Park - Faraday Road - Basingstoke - Hampshire RG24 8QQ
Tel: +44 (0)1256 486640 - Fax: +44 (0)1256 486650

www.nexans.com/LANsystems - info.ncs@nexans.com