Linesight

Hyperactive **Hyperscale**

The next step of the digital revolution

Introduction

As we find ourselves in the thick of the digital era, our appetite for data continues to grow rapidly. Data traffic is increasing by 28% per annum, and in Cisco's 'Global Cloud Index' paper, it predicts that there will be in excess of one billion extra users on the internet by 2020. This brings the total number of users to 4.1 billion – just over 50% of the planet.

The technology space is being revolutionized by our use of data, and as a result, organizations are pouring millions, and even billions, into increasing their data facilities. Given the pace at which this growth is happening, it is not surprising that big cloud operators are seeking to build bigger and faster.

In recent years, a major trend has emerged towards hyperscaling, which refers to data centers on a much bigger scale; campuses whereby hundreds of MW are being planned, rather than multiples of 2 or 3MW. Hyperscale operators typically have thousands of servers, but large multinationals, such as Amazon, Microsoft, Google, Facebook, Apple and IBM, who collectively account for more than half of the cloud space, can have millions of servers across their network globally. Between the first three of these providers, US$26 billion was invested in 2015 (MarketRealist.com), largely in new hyperscale data centers – campuses with 60-100MW+ potential installed capacity.

As technology's use and capabilities continue to gather pace, traffic within these hyperscale facilities will increase fivefold (Cisco, 2016). These facilities currently account for 34% of all data center traffic, and are expected to account for 53% by 2020.

Hyperscale facilities are expected to account for **53%** of all data center traffic by 2020.

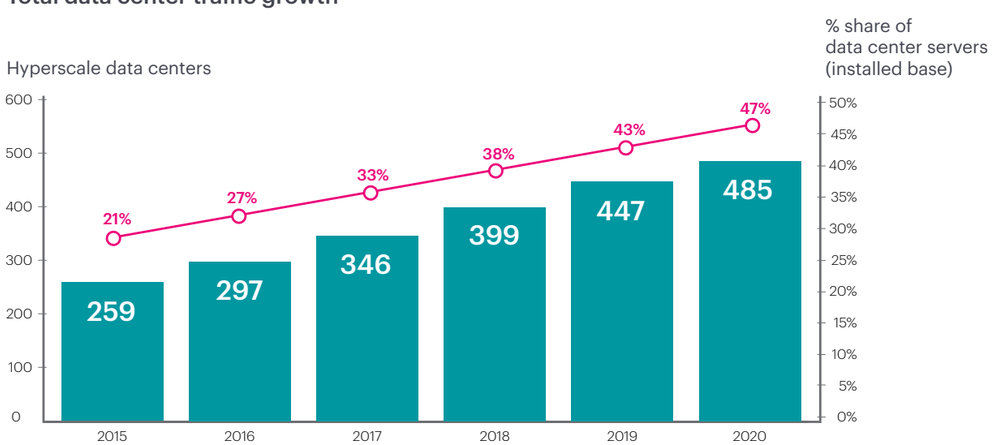
Growth

The Compound Annual Growth Rate (CAGR) of the global data center market between 2016 and 2020 is expected to be at a robust 11% (Technavio, 2016). The Asia Pacific region will be the fastest-growing contributor, with a CAGR of around 13%, and telecommunications organizations accounting for the majority of data center facilities in the region.

Looking specifically at the hyperscale market, it is expected to be a significant contributor to the sector growth, with a CAGR of 20.7% between 2016 and 2022, reaching a revenue of US$71.2 billion by 2022 (Allied Market Research, 2016). According to Cisco's Global Cloud Index, hyperscale data center numbers will increase from 259 at the end of 2015 to 485 by 2020, while Synergy Research Group estimates that 47% of all data centers will be hyperscale facilities by 2020.

Hyperscalers are dominating a series of key areas from cloud to networks, with Synergy Research Group finding that they accounted for 68% of cloud infrastructure services (Infrastructure as a Service (IaaS), Platform as a Service (PaaS), private hosted cloud services) and 59% of the SaaS (Software as a Service) market at the end of 2016. This was remarkable growth from the 2012 figures of 47% for each of those markets.

Total data center traffic growth



Source: Cisco Global Cloud Index, Synergy Research

Hyperactive Hyperscale White Paper

What is driving demand?

This extraordinary growth can be largely attributed to a few factors; increasing digitization, the shift towards cloud computing, the ever-increasing range of applications for these facilities from a broadening range of industries, the evolution of artificial intelligence (AI) and the growth of the Internet of Things (IoT).

Increasing digitization

As the appetite for data remains almost insatiable, there is a requirement for increasingly large facilities. With data storage accounting for between 5-15% of data center space, the expanding amount of data to be stored means that storage requirements are growing at 40% per annum (Cisco, 2016). Simply put, larger facilities are required to store these larger volumes of data.

The shift towards cloud computing

Cloud computing facilitates adaptability, scalability and agility, making it better suited to modern, fast-moving business environments. Therefore, the focus in recent years has been, and for the future will be, on design solutions that prioritize modularity, flexibility and speed.

Data storage requirements are growing at **40%** per annum.

However, as cloud technology continues to grow, there is increased pressure on businesses to distribute more information without a disruption to service. Because of this, data centers now need to be designed in a way that can maximize their capacity, while retaining enough adaptability to expand as the cloud gets bigger. This has fed into the hyperscaling trend.

JLL forecasts that cloud adoption acceleration will lead to a 100% increase in the size of the data center industry over the coming five years. To put that into context, the 3.9ZB (zettabyte) of annual global cloud IP traffic that was recorded in 2015 is expected to increase to 14.1ZB by 2020.

At present, additional capacity is being catered for by cloud providers via a number of means; building on greenfield sites, retrofitting existing facilities, colocation, deploying capacity at the edge by creating a distributed cloud network that can reduce network latency by managing loads closer to the consumer. This can lead to unpredictability, however, which manifests itself as challenges across the supply chain.

A broadening range of applications/uses

The critical importance of data centers is not limited to technology-related businesses, as more and more sectors now rely upon such facilities. The hyperscale data center is of ever-increasing importance across sectors, and is at the heart of the digital economy which now shapes everyday life, from banking to socializing.

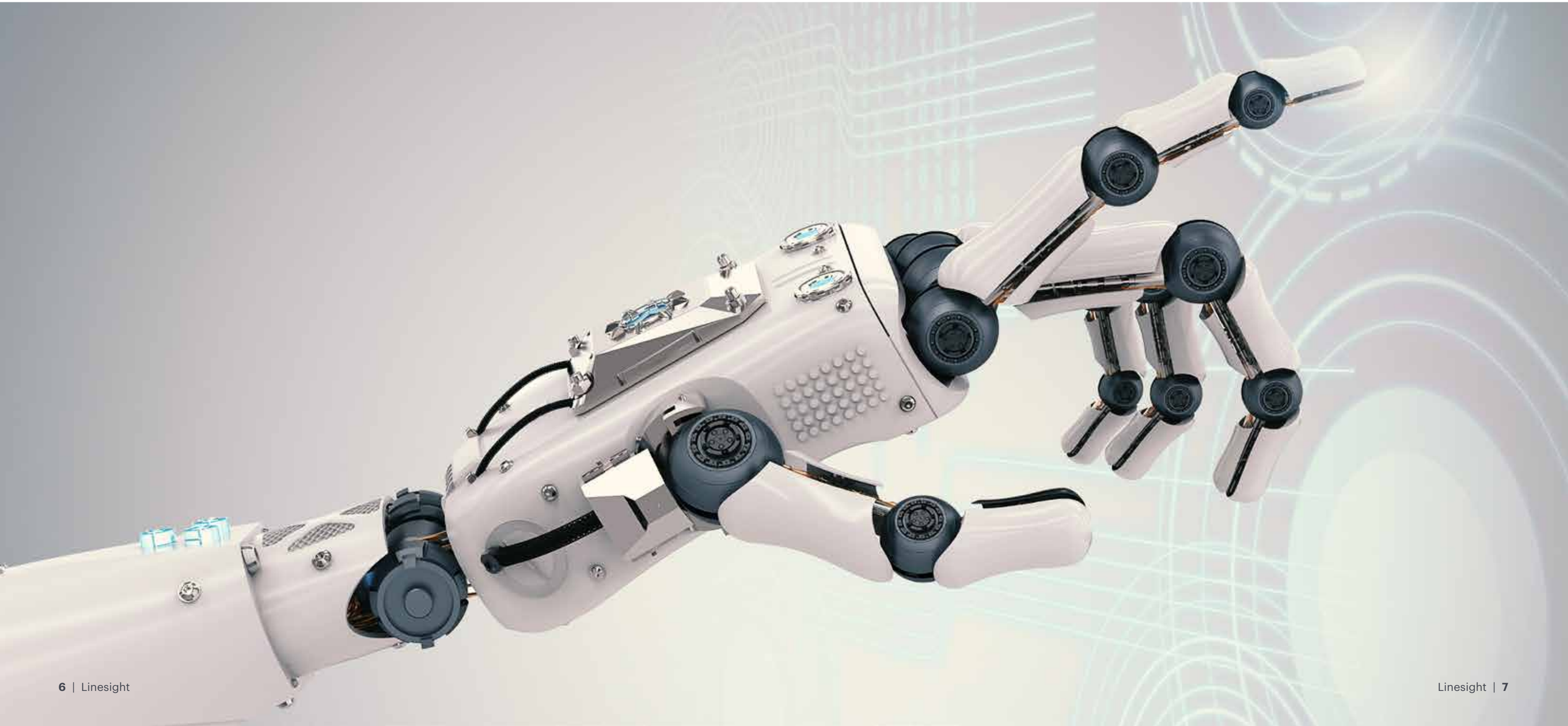
Increase in Artificial Intelligence

Gartner Research asserts that by 2021, AI will be an important consideration in data center design and architectures, but hyperscalers are amongst the early adopters. This is not surprising, given that AI typically requires significant data storage capacity, which at present is a constraint to its adoption for the vast majority of organizations.

Hyperactive Hyperscale White Paper

The growth of the Internet of Things

With Gartner putting the number of connected devices at 26 billion by 2020, IoT is a major talking point for the tech industry in recent times. As with the AI data storage requirements discussed previously, the sheer volume of data generated by the IoT necessitates huge storage capacity. Again, this requires hyperscale facilities to meet these storage requirements at present.



Challenges

Latency, security, power and fiber availability, and cost pose challenges to the hypersite, all intertwined and pulling in opposite directions.

Latency

If it weren't for latency, surely everyone would move to Iceland for free cooling, low-cost power and the security of a tiny population in the middle of the Atlantic? But the milliseconds add up and not all regions can hang off that fiber. So, for the perfect latency solution, should we have a myriad of mini data centers spread all over the globe, and wave goodbye to hyperscale? Pulling fiber and power to all of those latency-friendly sites, however, not to mention securing them, is prohibitively costly, and so the pull is back towards regional hyperscale, giving a balance of the opposing forces.

Security

Data security has never been more important, and is becoming an increasing driver towards in-country data hosting. Previously the lowest cost geography, which could connect across national boundaries, would have won out on site selection. But now countries which were previously overlooked because of rigid working directives and the high cost of construction labor have emerged onto the potential locations list, because of their domestic requirements.

Power

Power is also becoming a more complex factor. 'How much?' was previously the focus – how much does it cost, and how much is available? Now, if a hypersite exists it can power itself, with gas generation plants springing up in areas of need.

Sites with direct access to 400kV power circuits, the super grids of the host countries, are the smart places to be. The 220kV circuits in urban settings are no longer offering the security of supply at the required capacity that would have been perfectly acceptable five years ago.

Green power is also driving different solutions, whether that means locating the hypersite in a hydro or wind power-rich geography, or sourcing green energy from the network provider.

Ultimately, flexibility means that the lease market remains an important player, but lower cost, greater security, adapted power solutions and a global scale network giving suitable latency will all combine to drive the hyperscale data center towards an ever-bigger size – even if it is in smaller steps.

Cost

The heavy capital investment required, and resultant high level of depreciation, puts a strain on the P&L before income has been built up. So despite the push for bigger and faster, cost remains a significant factor, as each of the big players eye up their competitors' cost base. It is notoriously difficult to compare like with like, but whether the capex cost per MW is US$10 million or US$7 million, it will never be low enough.

Hyperactive Hyperscale White Paper

Trends

With the uptake of lease space being at a high, the resource base has become somewhat stretched, and so, the quicker time-to-market does not always materialize.

Buy versus lease

With a shift of some of the major cloud operators into lease capacity last year, and a slowdown of their own build programs, the future for hyperscale data centers has taken a bit of a wobble. However own-build construction appears to be ramping back up in 2017, showing a renewed appetite for build rather than lease.

On the cost front, modularization is the new Holy Grail to ever-lower capex cost per MW. Owners and OEMs are moving towards modular designs and equipment sets that can be built incrementally, to better match a P&L approach, as opposed to the traditional 'build it and they will come' capex splash. Once a large building shell with infrastructure, capable of hosting 20MW of servers goes live, a large depreciation charge hits the profit line. If there is a considerable lag in turning on revenue-generating servers in that space, profit is impacted. Building, and depreciating, in smaller chunks that better match the server build curve makes more sense. Lean construction strategies dovetail nicely with this off-site fabrication approach, so more of that will be seen, as well as a bigger role for the integrator of these different equipment elements.

In addition, prices being achieved for lease space have dropped, making it less attractive for the more peripheral lease players to invest in new space, hence a future reduced supply will initiate an upswing in lease price in the coming years.

This, in turn, means that ten-year discounted cash flows and lifetime cost models will be showing an upswing for build over lease, and 'build big' will give the lowest cost model.

In general, there is a swing towards leasing purpose-built facilities, or businesses moving their IT requirements to the cloud. Traditionally, lease space has been an integral part of the cloud operators' strategy, offering shorter time-to-market (TTM) and scalability, in an environment in which time is very much of the essence. While fewer internal resources become tied up in acquisition, as opposed to purchase, simpler, frills-free specifications will reduce infrastructure build costs. However, with the uptake of lease space being at a high, the resource base has become somewhat stretched, and so, the quicker time-to-market does not always materialize.

Second-tier markets

One of the biggest trends of 2017 has been a surge of interest towards more regional locations, which would enable the building of bigger and more scalable facilities, allowing businesses to grow to meet the voracious appetite for data. However, second-tier markets are still some way from adopting a hyperscale model approach. These markets remain immature with regards to data center demands. The requirements for data centers are smaller, the infrastructure is not in place to support such scale, and so hyperscaling remains some way off.

Hyperactive Hyperscale White Paper

Retrofit of existing facilities

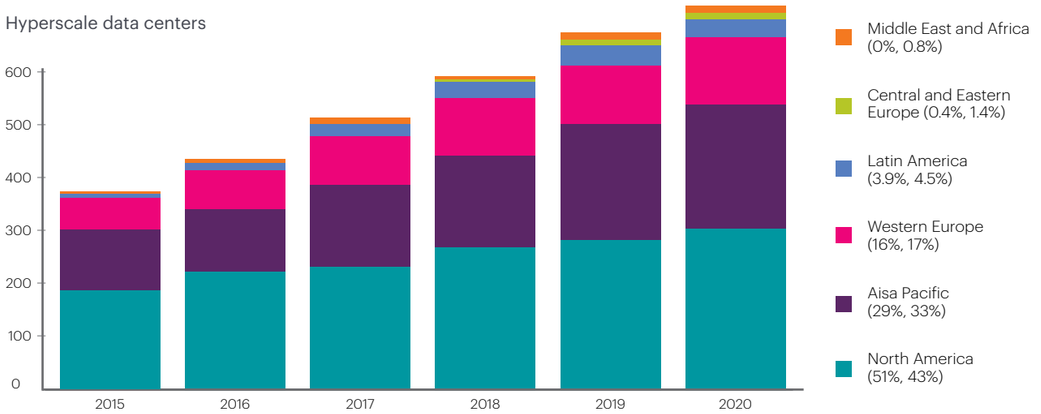
Data center consolidation remains a challenge globally, and as players continue to adapt to evolving requirements and fast-paced growth, retrofitting of existing facilities is becoming more commonplace. It can be the most cost-efficient method of upgrading and modernizing facilities and equipment, as well as being faster than a new build.

Hyperactive Hyperscale White Paper

Regional view

Regionally, the USA accounts for the vast majority of major cloud and internet data center sites, with 45% of such facilities currently located there (Synergy Research Group), despite an on-going push to locate new facilities in other countries around the world. This is followed by China (8%), Japan (7%), the UK (5%), Australia (4%) and Canada (4%).

Data center growth: regional view



Source: Cisco Global Cloud Index, Synergy Group

USA

The data center construction market in the USA is forecast to grow at a CAGR of 6.34% during the period 2017-2021. At the end of 2016, there was 271MW under construction, of which more than 160 MW was being delivered on a speculative basis. Hubs like Ashburn, Virginia, dubbed "Data Center Alley", are home to almost every major US data center provider, and could double from the current 10 million sq.ft. of data center footprint area to as much as 20 million sq.ft. by 2020.

Data center consolidation is a major challenge in the USA, and yet, the sector is expected to generate revenue of approximately US$13 billion between 2016 and 2021. Under the Federal Data Center Consolidation Initiative (FDCCI), the US Government's goals include cutting data center hardware, software and operating costs; moving more applications and services to the cloud; promoting energy efficiency in federal IT and boosting security. This had reportedly saved the US Federal Government over US$2.8 billion by early 2017. The move to consolidate or shut down federally-operated data centers is generating business in retrofitting existing facilities, as well as driving demand amongst data center providers to deliver cloud services for the US Government.

EMEA

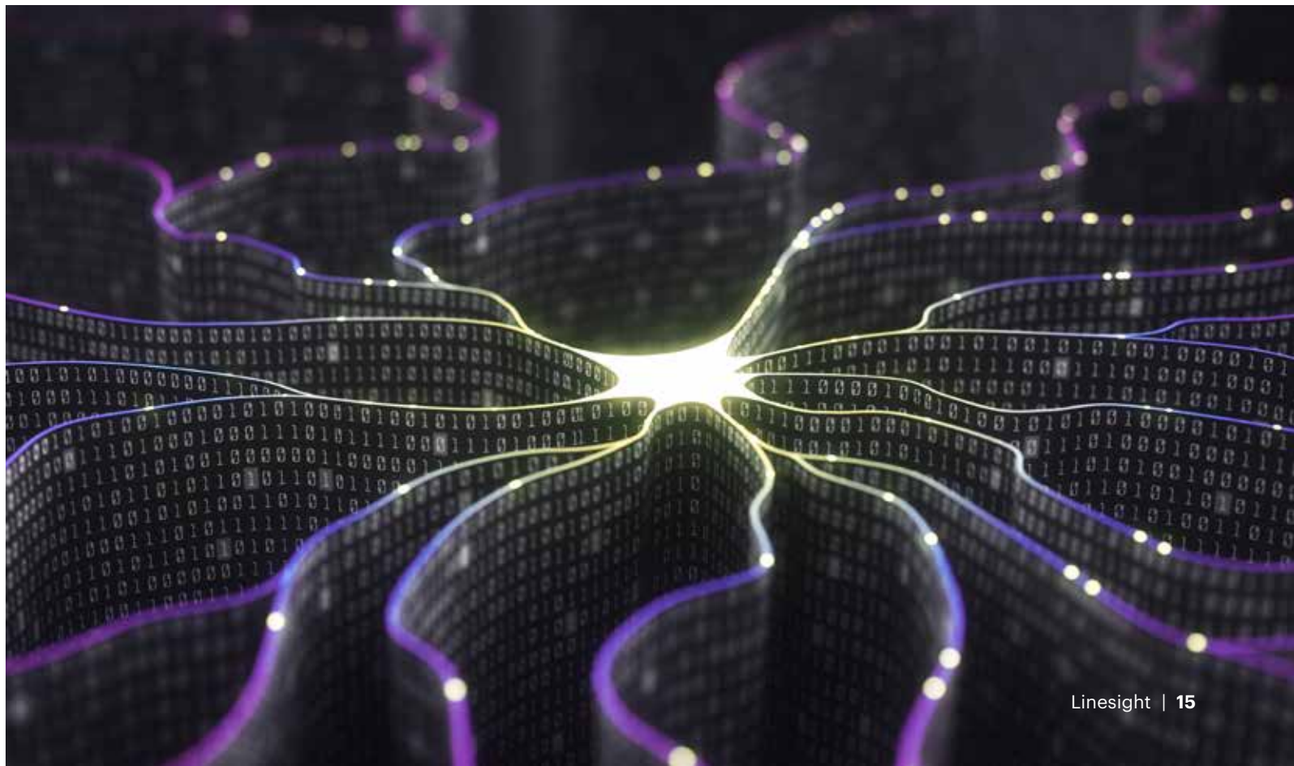
In Europe, market supply in the four core hubs of Frankfurt, London, Amsterdam and Paris hit 996MW at the end of Q1 2017, which represents an increase in supply of 20% over the previous 12 months. This growth exceeds the 8% CAGR that was recorded between 2012 and 2016. However, these 'Big 4' now account for only 60% of the European market, with outside markets increasing in importance.

Take-up over the last two years has been particularly robust across EMEA, and it is expected that in excess of 100MW will be taken up in 2017, for the second year in a row. Colocation market revenues are growing at a CAGR of 11.1%, while the average absorption rate of space continues to decline across the region. Cloud computing remains a key driver, with IaaS market revenues growing at a CAGR of 27.3%.

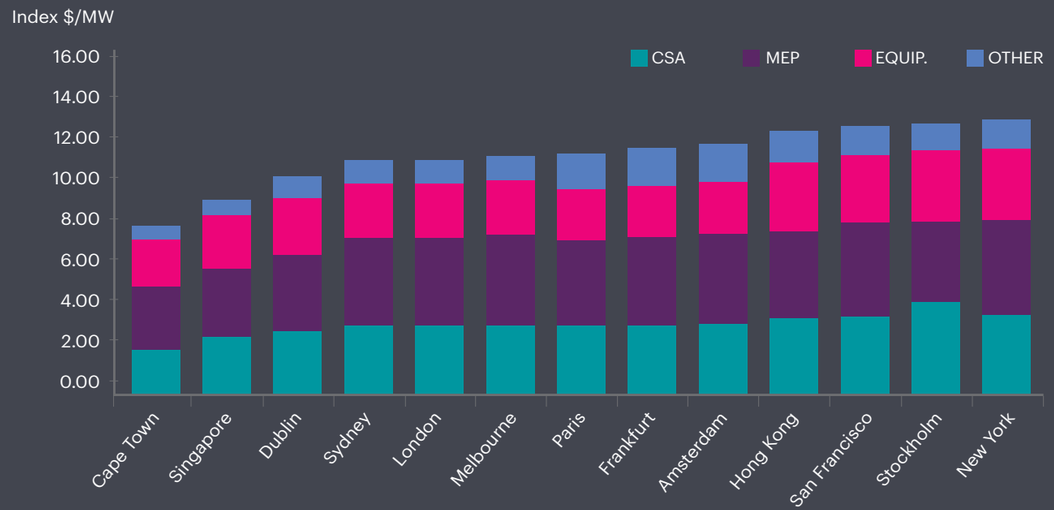
APAC

As mentioned above, APAC continues to be a significant contributor to the rapid growth being recorded, and this is expected to continue, with PWC anticipating that the region will surpass Europe by 2020. With an expected smartphone penetration rate of 66% by 2020, and broadband penetration expected to reach 40%, the data requirements in the region are voracious.
The market size in 2016 reached US$12 billion, driven by a surge in data consumption locally, a preference for outsourced data center services and regulations requiring local hosting for a number of key industries.

Within the region, the main data center markets include Singapore, Sydney, Hong Kong and Tokyo - all of which have contributed to a market growth over the last 12 month period of 17.6% based on total IT capacity, according to CBRE's 2017 MarketView. It is further envisaged that while markets in Singapore and Sydney will remain robust, and Tokyo will remain stable, Hong Kong, which has been relatively slow, will see a large increase over the next 12 months.



Global data center benchmarking



The above figures are based on a generic 32MW facility on a greenfield site, and would include all construction costs, all equipment, professional fees, plus IT and fiber provision, but would not include land costs, power, tenant fit-out and in-house client fees. It is assumed that power and fiber are available at the site perimeter. Statutory fees and contributions, permitting costs and applicable taxes are not included.

Costs are in USD and based on third quarter 2017 price levels, assuming competitive procurement. Adjustments should be made to account for differences in specification, program, location and procurement route.

Because of the high proportion of specialist services installations in data centers, local regional adjustment factors should be applied.

Summary

We see a rise in the demand and construction of hyperscale on a global basis, and this demand is set to increase further in the coming years. While it is evident that hyperscale is here to stay for the foreseeable future, it is not a solution for all locations, and there will always be a market for individual requirements with smaller data center deployments. With the abovementioned market developments, the rapid pace of growth and trends such as AI and IoT having a prolific effect on the data center market, it is an exciting time to be involved in Hyperactive Hyperscale.

Contacts

USA
Steven Cooke
Managing Director, West Coast
steven.cooke@linesight.com

EMEA
Paul Butler
Director
paul.butler@linesight.com

EMEA
Michael Riordan
Managing Director, United Kingdom
michael.riordan@linesight.com

APAC
John Butler
Managing Director, Asia Pacific
john.butler@linesight.com

linesight.com