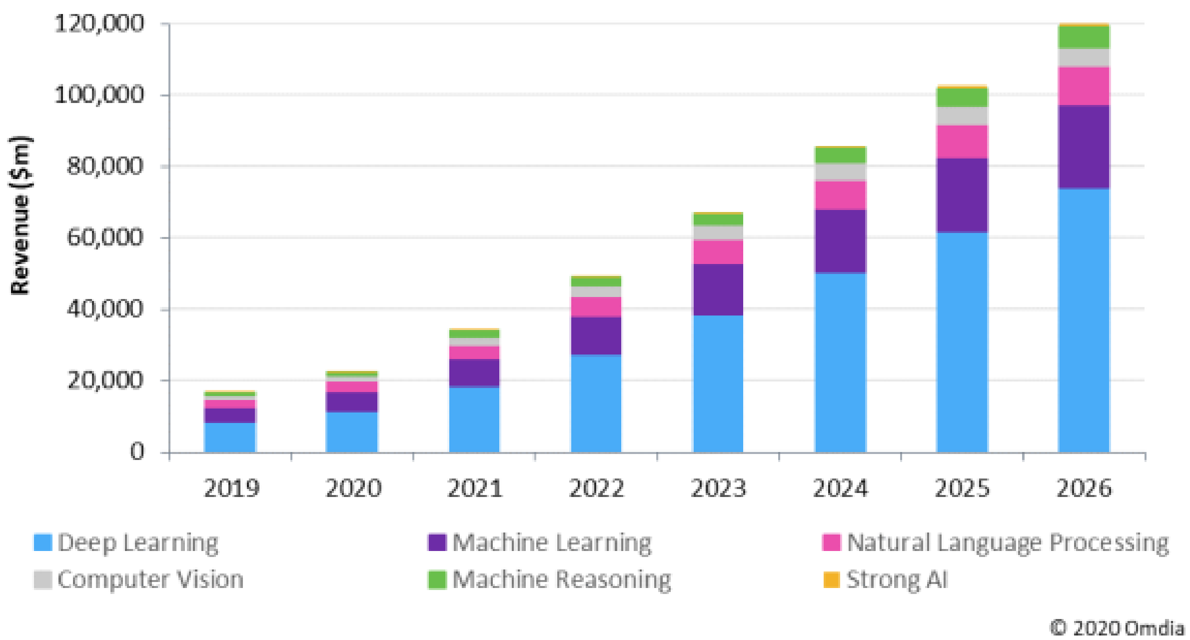
Publication date:
29 January 2021
Author:
Wang Shen

Next Generation of Full-Stack Data Centre: Lead the digital transformation of operators

Brought to you by Informa Tech

With 5G, cloud computing, AI, big data, and other technologies gradually maturing after several years of development, various industries have quickly entered the era of interconnection. At the same time things are becoming more digitalized, connected, and intelligent, and both the depth and the breadth of cloud computing are growing rapidly. Attracted by this flexible operation model, many enterprises are setting out on the journey of digitalization and cloud migration, with the enterprise workloads in the cloud expected to increase from 35% in 2020 to around 50% in 2022. This leads to an exponential increase of data in the cloud. In the meantime, with large amounts of data for deep data mining and processing, computing technologies such as AI pose greater challenges for the computing power of cloud and network throughput. Large-scale data centres are required to meet these rapidly growing demands. As the most concentrated node for massive data exchange, the data centre serves as a platform for information data storage and system operation and also plays a key role as a data circulation centre. The explosion of demand for data brings challenges to the storage, computing power, and network performance of operators' current data centre infrastructure.

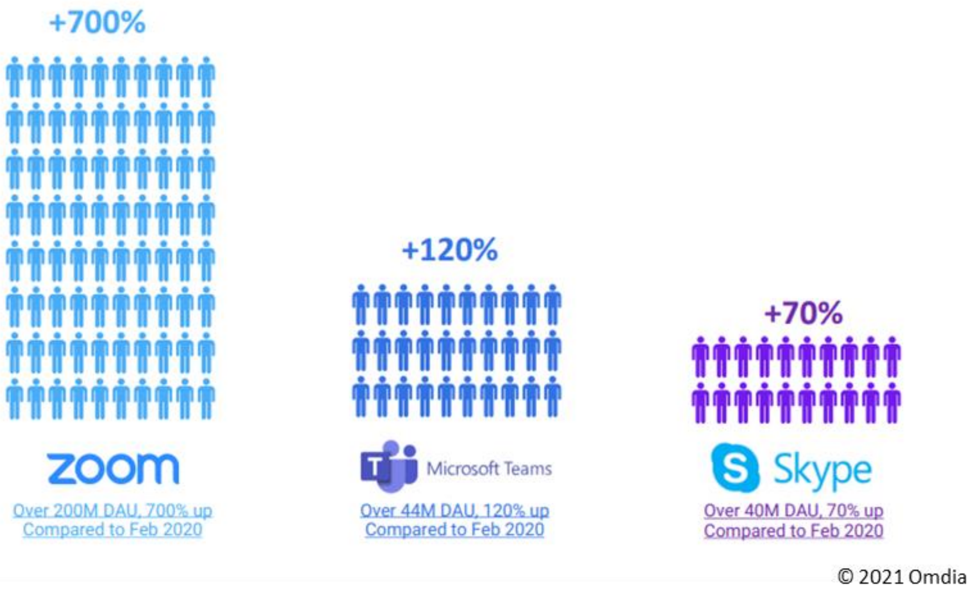
Figure 1: Global AI software revenue trend by technology, 2019–26



Source: Omdia

The outbreak of the COVID-19 pandemic has had a profound impact on the global economy, causing a majority of companies across the world to reduce costs in the face of an uncertain future. On the other hand, under the infection prevention requirements of avoiding physical contact wherever possible, network and remote interaction have become the new norms in people's life and work. As a result, demands for live streaming, cloud gaming, and remote collaboration have experienced a significant increase. According to Omdia's 2020 Enterprise Cloud Migration survey, 32% of the respondents admit that COVID-19 is driving their cloud migration. Compared with February 2020, the number of daily active users (DAU) of videoconferencing software Zoom increased by 700%, while Teams increased by 120%, and Skype by 70% (see Figure 2). The network data traffic across the world has generally gone up by about 20% during the pandemic, especially for online gaming (increased by more than 200%). The proliferation of online traffic puts huge pressure on bandwidth as well as data processing and storage, and users are deeply affected by latency, especially when using time-sensitive remote collaboration applications. Latency requirements and data sovereignty regulations led to the hierarchical structure of today's data centres. The implementation of this hierarchical structure from edge to local, regional, and global levels, means that more data centres are needed to serve this growing market.

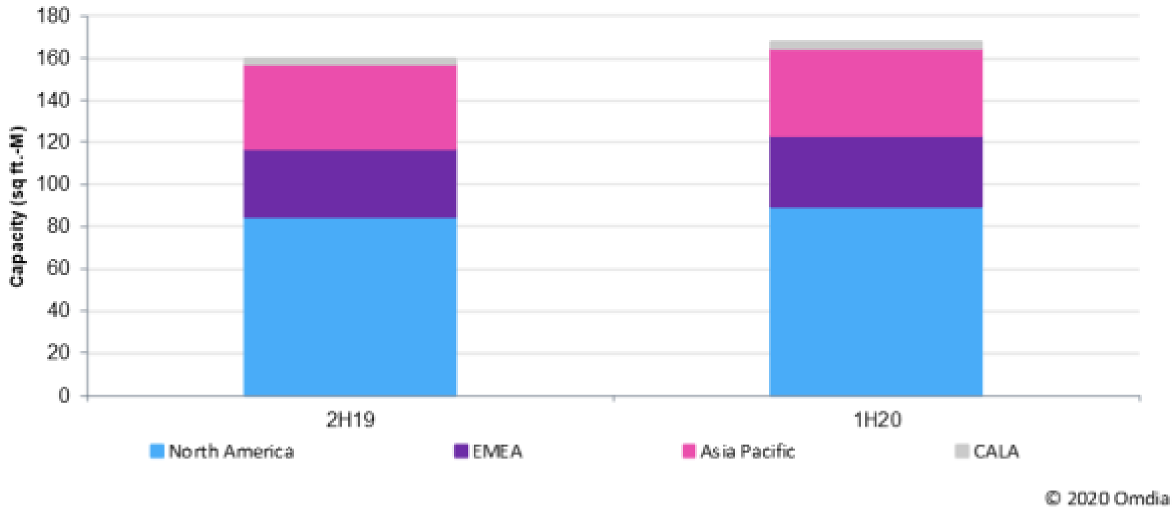
Figure 2: COVID-19 caused a surge in DAU of videoconferencing software



Source: Omdia

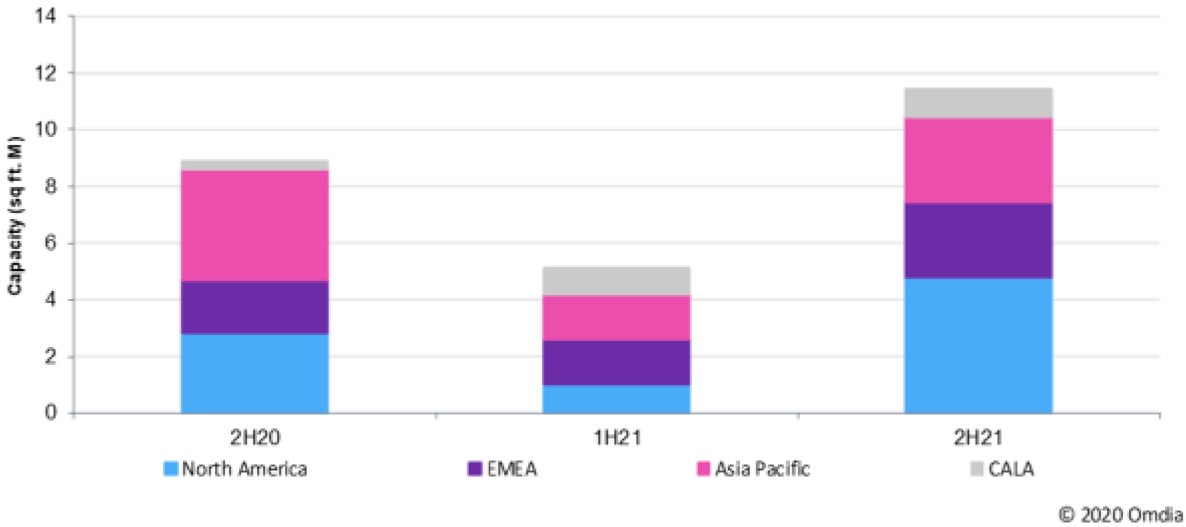
When the growth of data traffic gradually slows as it reaches its peak, data centres and network facilities will face pressure from end-point applications for a long time. Data and data centres have therefore gained widespread attention. Cloud service providers across the world tend to build their own data centres when possible. In recent years, many projects have been carried out in North America, Europe, Asia Pacific, and other regions, particularly Latin America. The growth of the cloud services market and the strong demand for network density drive cloud service providers to quickly fill in the gaps with colocation data centres and to enhance data caching and reduce latency with edge computing. Meanwhile, there are many cloud service providers whose physical data centre infrastructure is mainly or completely dependent on colocation service providers.

Figure 3: Operating area of data centres by region, 2H19–1H20



Source: Omdia

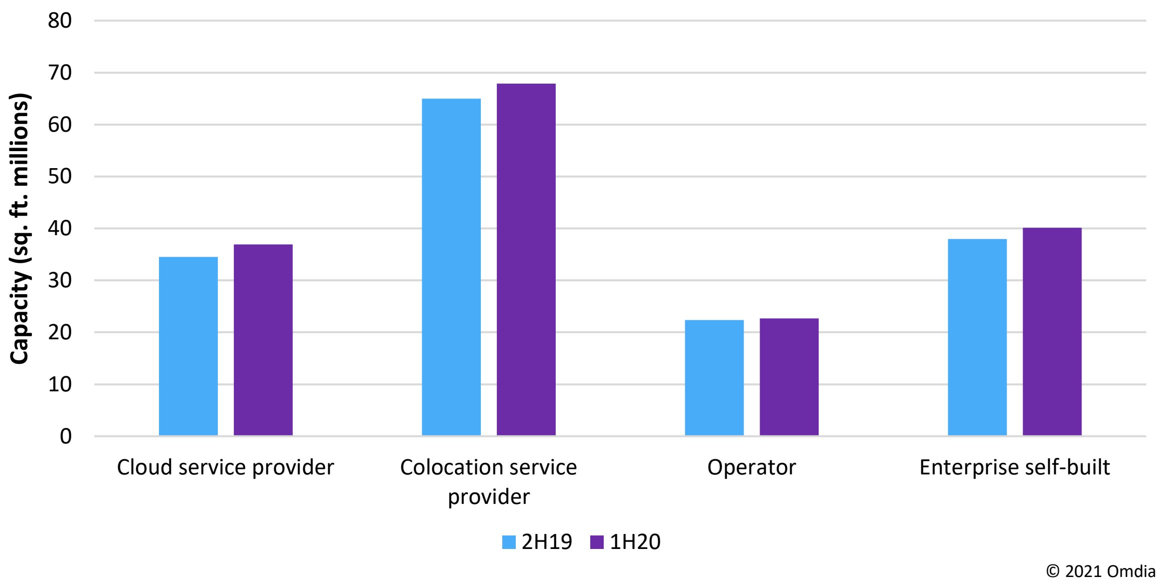
Figure 4: Operating area of data centres under construction by region, 2H20–1H21



Source: Omdia

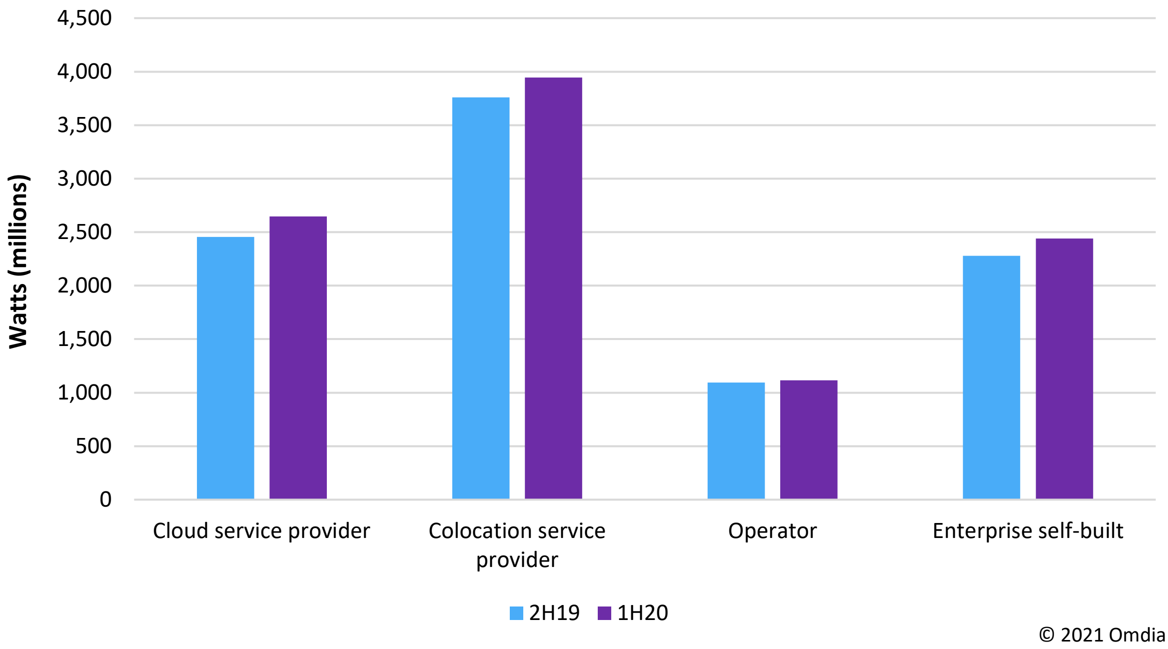
According to Omdia's 2020 database, by the end of 1H20, the power capacity of self-owned data centres globally (including those from cloud service providers, colocation service providers, operators, and enterprises' self-built data centres) increased by 559MW, or 6%, compared to the end of 2H19, reaching 10.1GW. Among these, the newly added capacity of cloud service providers and colocation service providers reached 190MW and 186MW respectively. Although the added capacity of operators' data centres is lower in comparison, the number still reached 21MW. At the same time, the power per unit area of the data centre has also increased significantly. Even if the operators' data centre business is dominated by the colocation service and the power density is low, the number has steadily increased by 10.8W/m2 to 1.05KW/m2 in the past six months. Compared with the other three types of data centre owners, the low power density for data centres owned by operators suggests that there is huge room for improvement.

Figure 5: Total area of data centres by type of data centre owner



Source: Omdia

Figure 6: Total power of data centres by type of data centre owner



Source: Omdia

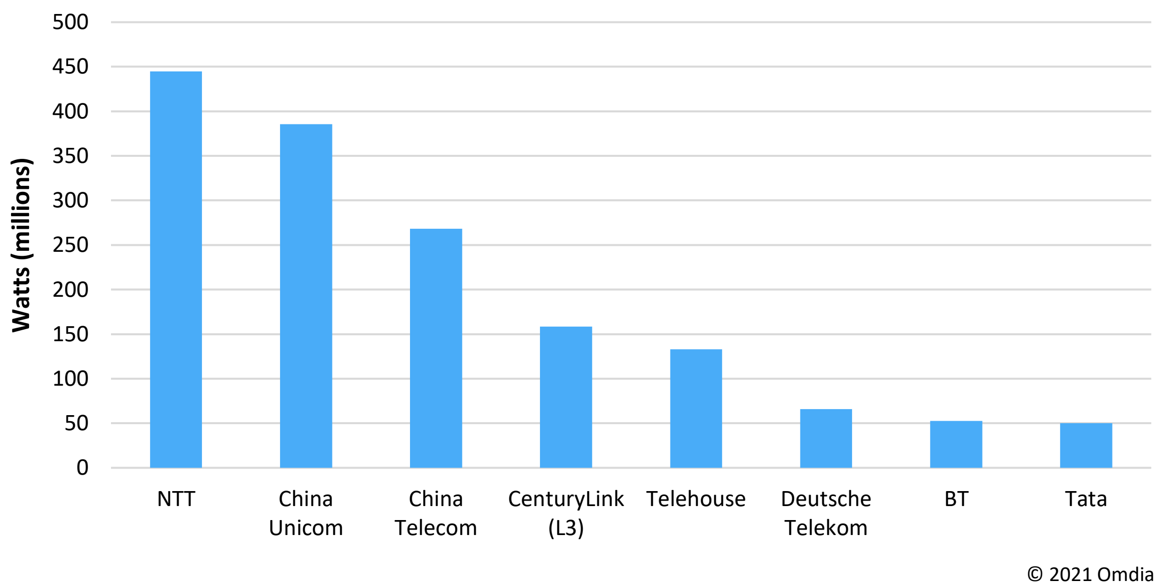
In 2H20, 63% of the new data centre projects in the world were from colocation service providers, 25% were from cloud service providers, and 12% from operators. In terms of operating area, projects from colocation service providers account for 43% of the total area, followed by cloud service providers (40%) and operators (17%). The number of colocation service provider projects under construction is more than twice the number of construction projects for cloud service providers, but the overall area across the world is roughly the same. The reason behind this apparent inconsistency is that colocation service providers featuring retail colocation business have a lot of smaller expansion projects underway, while cloud service providers are building data centres on a larger scale. Operators are also more inclined to build bigger regional data centres.

Operators are driving the substantial growth of data centre colocation service providers for two main reasons: Firstly, the ecosystem of multiple operators' data centres coexisting attracts more customers to the flexibility offered by colocation services. Secondly, due to the new demands for the network and online content of working from home and remote learning, the hybrid multicloud and edge-side connectivity provided by operators through cloud, colocation, and their own data centres have become critical. Operators' enterprise customers are increasingly taking hybrid multicloud as an operating model. This model of selecting suppliers based on requirements opens a new opportunity for operators. To meet this demand, operators must provide data centre capacity options at the edge, local, and regional levels and develop relationships with users of hyperscale data centres across the world. The key to this new business model is network resources and how enterprises will use 5G networks and existing backbone networks.

Meantime, with the saturation of the global telecommunication market, operators' traditional telecommunication services enter the plateau phase and new opportunities are needed to expand their business. The digital transformation of downstream traditional industries and the explosion of new network applications accelerate the transformation of operators, driving them to transform from network providers to diversified-service carriers and shift their business attention and direction to internet services. Data centres owned by operators have a unique mix of use cases. Apart from the data centre capacity required for network operation, part of its capacity is used for cloud or colocation services targeting enterprise customers. Therefore, operators' enterprise-oriented business will be more reliable with data centres as the physical infrastructure for network storage and computing data. With the explosion of digital applications, the consumption of existing data centre resources is accelerating and bandwidth resources are stretched. Consequently, operators are forming a new trend of investing in data centres, which is expected to reach a larger scale in the next few years. Demands for and discussions on digital sovereignty around the world are also expected to favour operators' data centres.

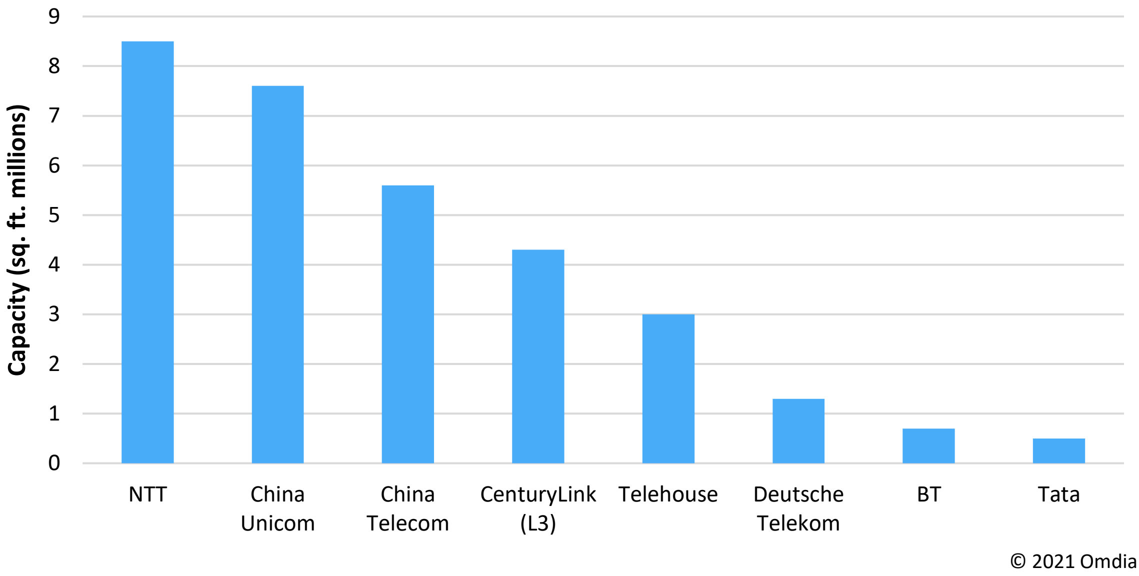
Among the major operators globally tracked in the Omdia 1H20 survey, NTT ranks top with a total data centre power capacity of 444.8MW, followed by China Unicom and China Telecom with 385.5. MW and 268.2MW, respectively (see Figure 7). Among the top operators, CenturyLink owns and operates 178 data centres after acquiring Level 3 Communications (L3), but more than 90% of these data centres are relatively small, resulting a small total operating area (see Figure 8). China Telecom ranks second worldwide with 157 data centres, of which more than half (84) are located in China with the rest spread across Asia Pacific, Europe, the US, and the Middle East & Africa.

Figure 7: Power capacity of selected operators' IT infrastructure in 1H20



Source: Omdia

Figure 8: Area capacity of selected operators' IT infrastructure in 1H20



Source: Omdia

In the market, operator data centres face competition from hyperscalers' self-built data centres; the difference is that the latter have a clear business orientation. Operators often invest in planning and construction first, and then gradually expand to include enterprise customers from different fields. This wholesale business model has its special competitive advantages and demands. Due to the huge investment required for large-scale data centres, and their relatively high operation and expansion costs, reducing the capital investment in initial construction, shortening the planning and construction cycle, and saving operation & maintenance (O&M) costs are top priorities for operators to provide high-quality data centre services to future customers.

As a result, it is necessary for operators to develop a continuous improvement plan, adopt new technologies, and seamlessly retire old technologies with lower efficiency and higher costs. To this end, a global perspective is needed to reevaluate each step in the entire lifecycle of the data centre in terms of time and space; optimize resources in alignment with cost, risk, environment, and other factors; improve the efficiency of managing existing resources; maximize cost-efficiency; and provide the best customer experience. In the full lifecycle of the data centre, when looking to shorten time to market (TTM), improve power usage effectiveness (PUE), improve computing-storage-network performance, and optimize O&M, all steps from land construction through cooling and power supply to network computing and storage are consistently planned, designed, constructed, and delivered. This means greater efficiency is achieved to improve competitiveness and gain a higher return on investment.

TTM means the time it takes to build data centres. Shortening the cycle of a series of processes from planning to construction helps operators to take the lead in the market. Among the variety of ways in which TTM can be shortened, modular construction is important. Treated as separate building blocks, land construction and room construction could be started at the same time to shorten the whole process. The modularization could start with the server components and then move on to the architecture and room and finally to a full modularization of the main systems in the data centre, such as power supply, cooling, storage, and network. In this way, the current industry-wide 24-month TTM could be reduced to under 10 months, or even less.

Apart from reducing the cost, another critical consideration is the environmental factor of data centres. Green and energy saving are also key requirements for next-generation data centres. In 2007, the average PUE of global data centres was as high as 2.5, which has been reduced to less than 1.6 in 2018. In the future, PUE will generally be required to be lower than 1.25. In order to reduce the large amount of heat generated by the operation of server equipment, cooling energy consumption accounts for more than 30% of the total energy consumption of traditional data centres. And as the installed density increases, cooling energy consumption will further increase, so the key to reducing PUE lies in cooling technology innovation. In addition to using environmental resources for natural cooling and green energy generation, leveraging new technologies such as liquid cooling and indirect evaporation to avoid the use of air conditioning can also reduce PUE to make significant savings in energy investment and operating costs.

Given the huge pressure that the wave of digital transformation has put on networks, computing power, and storage, a strategic goal for operators during the next phase should be to improve both the computing power and storage of the server and the efficiency and capacity of network equipment. Advances in cooling technology can achieve higher computing power density, which can save a lot of cabinets and floor space. The increase in computing power density will inevitably promote the growth of data throughput. For this reason, the density and throughput of storage devices will also inevitably achieve a higher level, and the trend of shifting to all-flash storage and resource pooling is set to grow. In addition, technologies such as IP-based networks enable a more intelligent allocation of network resources, ensuring zero packet loss while supporting high throughput. The complementary relationship between computing power, network, and storage is a necessary driver for the improvement of data centre capabilities.

To better ensure the long-term stable performance of these technical capabilities, a smarter and more efficient O&M capability is another key demand. With the next-generation data centre becoming more complex and larger, it is necessary to move away from an O&M approach that relies on personal experience and passive responses and instead shift toward a comprehensive digital fine management that is intelligent and responsive, minimize the intervention of unstable factors such as manual work, and improve the visibility of each operation area in the data centre in order to save O&M costs while improving quality and efficiency.

In a future that will see a massive explosion of online data, expponential growth in computing power, and intelligence everywhere, the needs of enterprise customers will become more diverse, and the requirements of structural flexibility have therefore become the new focus. Facing challenges such as how to better improve computing power, resolve cooling challenges brought by high density computing innovation, make further improvements in saving energy and increasing efficiency, and use smarter operation and maintenance to meet business SLAs has become a key issue for operators when facing the market. In order to realize the long-term leading advantage of operators' data centre business in the future, it is no longer sufficient to seek solutions solely from any level of L0, L1, L2, and L3. It is necessary to connect all levels vertically and to comprehensively rethink about how to face the future requirements of digitalization development from a new height, aiming for the design goal of digitalization of the data centre's full lifecycle in order to realize the flexible deployment, high efficiency and energy saving, computing power increase, storage optimization, network intelligence, and controllable operation and maintenance costs of data centres. This is the essence of the so-called "full stack" of the next-generation data centre.

Appendix

Author

Wang Shen
Senior Consultant, Global Enterprise & IT
askananalyst@omdia.com

Get in touch

www.omdia.com
askananalyst@omdia.com

Omdia consulting

Omdia is a market-leading data, research, and consulting business focused on helping digital service providers, technology companies, and enterprise decision-makers thrive in the connected digital economy. Through our global base of analysts, we offer expert analysis and strategic insight across the IT, telecoms, and media industries.

We create business advantage for our customers by providing actionable insight to support business planning, product development, and go-to-market initiatives.

Our unique combination of authoritative data, market analysis, and vertical industry expertise is designed to empower decision-making, helping our clients profit from new technologies and capitalize on evolving business models.

Omdia is part of Informa Tech, a B2B information services business serving the technology, media, and telecoms sector. The Informa group is listed on the London Stock Exchange.

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help your company identify future trends and opportunities.

Copyright notice and disclaimer