

2019

Top 10 Data Center Industry Trends



Uptime Institute Research

December 2018

Overview

The data center industry is changing and growing at a striking pace. Large build-outs are causing supply-chain strains and skill shortages, while innovations in networking, power, and resiliency are creating new opportunities, but creating challenges as well. Overall performance, service availability, and security are top concerns for operators, and these are compounded by the increasing interdependencies and complexity across ecosystems. Strategies to balance risk mitigation against cost controls are key.

What changes should data center owners, managers, and operators expect? Which innovations will make a difference? Where are problems and challenges likely to emerge? Let's take a look...

Lead analyst

Andy Lawrence (alawrence@uptimeinstitute.com)

Contributing analysts

Rhonda Ascierio (rascierio@uptimeinstitute.com)

Kevin Heslin (kheslin@uptimeinstitute.com)

Uptime Institute Research recommends data center owners and operators pay close attention to ten key areas in 2019 due to their potential to transform traditional data center designs and operational processes:

- **Big Cloud builds push the ecosystem to its limits**
The accelerating demands of the big cloud operators for more data center capacity are distorting and straining the ecosystem of suppliers, builders, operators, and power companies.
- **Worried governments step up oversight and regulation**
Governments around the world are becoming more concerned about the profits and power of large IT companies – and about societies’ dependency on invisible infrastructure.
- **The transition to distributed resiliency will not be smooth**
Disruptive and often high-profile outages will continue as operators grapple with the complexities of deploying distributed, hybrid systems across multiple data centers and services.
- **Edge data center hype outruns deployment**
The demand for small, edge data centers is coming more slowly than predicted. Issues with security, costs, business models, integration, networking, and 5G rollout will constrain adoption.
- **Connectivity is king: Operators work to build the network fabric**
Demand for fast and secure network connections to both trading partners and cloud operators continues to grow as the network becomes the critical component of the hybrid infrastructure.
- **Skills shortage will force new workforce strategies**
Even with automation and AI, the data center sector’s staff shortages are set to intensify. To keep pace with demand today and to avoid a precipitous shortfall tomorrow, data center operators will work to diversify the talent pool with new initiatives, hiring strategies, and workforce training.
- **Climate change forces fresh review of resiliency planning**
The risks associated with climate change may be more urgent, varied, and extensive than IT planning had previously anticipated. Data center resilience assessment is more critical today than ever.
- **Economics will drive acceptance of data center AI...eventually**
AI-based approaches to analyzing data center risk and efficiencies, including via new cloud services, will be proven at scale, driving mainstream acceptance and, over time, high levels of adoption.
- **Growing threats will necessitate new ‘zero-trust’ approaches**
Security vulnerabilities now affects mission-critical facilities. Organizations will need to adopt more stringent policies regarding data center equipment, services, contractors, suppliers, and staff.
- **Programmable power unlocks new efficiencies and agility**
A combination of Lithium-Ion batteries and other new forms of energy storage gives data center operators a new set of levers to help improve data center performance.

1. Big Cloud builds push the ecosystem to its limits

The accelerating demands of the big cloud operators for more data center capacity are distorting and straining the ecosystem of suppliers, builders, operators, and power companies.

The requirements of hyperscale data centers have already re-shaped the industry's supply chain. Demand for wholesale data centers has boomed, forcing data center builders and operators to pivot to attract the new business. They have customized their designs, moved contract facilities staffing, deployed more custom (and low cost) IT, and collaborated on new co-engineering approaches with facility equipment makers. Hyperscale customers are already driving supplier relationships, in terms of costs, designs, and production.

Total hyperscale capacity growth has long outpaced the rates of other data center segments, with global take-up in 2018 expected to be double that of two years ago. Forecasts for future hyperscale capacity vary, but all predict an exponential pace, with demand uneven across global regions.

In 2019, we expect operators and suppliers to focus on more standardization – of data center and equipment designs, of build approaches, of incremental power requirements, and so on. Meanwhile, the bigger suppliers are pushing for earlier and better visibility into the expected future demand from hyperscalers. Smaller suppliers will still thrive – but they will find that their margins are pushed down as they must accept a limited role, with bigger companies getting stronger across the supply chain.

2. Worried governments step up oversight and regulation

Governments around the world are becoming more concerned about the profits and power of large IT companies – and about societies' dependency on invisible infrastructure.

In recent years, governments viewed "Big IT," and the wholesale of adoption of IT generally, as a positive force in almost every way. IT promotes innovation, productivity, and a positive trade flow (for some), and creates jobs, wealth, and investment in infrastructure. Any downside, on privacy, on monopoly power, on energy use or carbon emissions, on employment practices, or on tax avoidance, had been largely downplayed. But the pendulum has swung: The Internet giants in particular – but many other operators too – can expect a more resistant, skeptical, and occasionally frosty climate in the years ahead.

Governments across the globe have already enacted an array of new regulations and taxes, and this will have the effect of increasing their control and influence over the Internet. These regulations are necessary, the argument goes, because large IT companies playing key societal and economic roles have been granted too much freedom, strain local infrastructure, and lack sufficient procedures to ensure the availability of the necessary services they provide. New regulations and taxes are intended to protect individual privacy, improve cybersecurity, increase infrastructure investment, mitigate against downtime incidents, and/or to limit social disruptions. Expect more in 2019 and beyond.

3. The transition to distributed resiliency will not be smooth

Disruptive and often high-profile outages will continue as operators grapple with the complexities of deploying distributed, hybrid systems across multiple data centers and services.

The IT industry is in the middle of a large and difficult transition – from single, secure, and tightly managed data centers, to a network of distributed, dynamically interconnected systems deploying clouds, microservices, and software-defined networks (SDN). The new architecture is emerging over time, being constructed partly with new data centers, networks and systems, and partly by building on top of existing infrastructure. The result is a complicated distributed grid that supports an array of applications and services.

Evidence suggests that the best – but not necessarily the cheapest – way to ensure resiliency is to combine site and network level redundancy with distributed IT and resilient architectures using cloud technologies. As the industry moves in this direction in 2019, some painful and sometimes high-profile lessons will be learned.

4. Edge data center hype outruns deployment

The demand for small, edge data centers is coming more slowly than predicted. Issues with security, costs, business models, integration, networking, and 5G rollout will constrain adoption.

A resurgence in distributed IT is underway from the Internet of Things, the emergence of mobile edge computing (MEC), and other new approaches. Of these, IoT is having the earliest impact; a proliferation of connected devices, sensors, meters, mobile phones, and medical devices are some examples of technologies now deployed at scale, with others such as virtual reality and augmented reality being introduced.

The data generated is driving demand for data center of all types: those nearby ('local edge' or 'edge') for first-line processing, analysis, and routing; those within a local area ('near' or 'regional' edge) to connect, integrate, and re-route; and those far away – such as economical hyperscale facilities, ('core') for further processing, analysis, and archiving.

Uptime Institute Research, along with some suppliers, has forecast significant new edge data center demand, particularly for micro data centers. These are fully encapsulated, integrated, remotely-managed and small-footprint facilities, oftentimes mesh-connected and typically <150 kW in IT capacity. We believe micro data centers, which can be deployed as pockets of discrete capacity or as modular, incremental blocks of large deployments, will find many use cases: for efficient upgrades of network closets and for supporting IoT appliance data analysis, storage, and resiliency in smart factories, buildings, and elsewhere.

This new wave of edge demand will be transformative and warrants some of the current attention and hype it is attracting. But, as is characteristic of large new markets, adoption at scale takes time. 2019 will not be the breakthrough year, with business and technology issues yet to be fully resolved.

5. Connectivity is king: Operators work to build the fabric

Demand for fast and secure network connections to both trading partners and cloud operators continues to grow as the network becomes the critical component of the hybrid infrastructure.

SDN fabrics are secure online platforms that allow virtual private interconnections to others on the platform to be rapidly provisioned, so that organizations can easily connect to public clouds, service providers, partners, and suppliers in different data centers and regions. This should extend the ability to route traffic securely and predictably between different data centers and, ultimately, drive down organizations' costs to connect. With the range of applications and partners growing, some requiring low latency or high bandwidth connections for applications such as IoT or distributed resiliency, good connectivity now commands a premium for the colocation sector.

As organizations' digital footprints continue to sprawl (both geographically and by the number of interdependent IT services they consume), demand for and reliance on SDN fabrics will grow. In 2019, we believe more SDN fabrics across colocation providers' data centers will be established. How the various suppliers will interconnect with and compete with each other is still unclear: customers may push for more openness.

6. Skills shortage will force new strategies

Even with automation and AI, the data center sector's staff shortages are set to intensify. To keep pace with demand today and to avoid a precipitous shortfall tomorrow, data center operators will work to diversify the talent pool with new initiatives, hiring strategies, and workforce training.

Uptime Institute Research survey results suggest that data centers will continue to struggle to recruit and retain sufficient qualified staff to maintain and grow reliable operations. In our 2018 survey, 45% of respondents said that a shortage of data center facilities staff will limit the data center industry's growth in the coming five to seven years. In another survey, 38% of data center operators said they were having difficulty finding qualified candidates for open jobs and 17% were having trouble retaining staff.

Uptime Institute advises that the skills shortage will be felt throughout the industry, with companies struggling to fill traditional facilities positions such as operations and management, security, network connectivity, cloud provisioning, and mechanical engineering but also compliance, contracts/SLAs/vendor management, software, financial management, supply-chain management, and corporate and social responsibility.

7. Climate change forces fresh review of resiliency planning

The risks associated with climate change may be more urgent, varied, and extensive than IT planning had previously anticipated. Data center resilience assessment is more critical today than ever.

No one could accuse the data center industry of complacency or ostrich-like behavior. Conservative risk assessment lies at the core of data center design and operations. Even so, in the field of climate change, which has become politicized in some countries, many data center owners and operators are advised by Uptime Institute to take a second look. Risk assessments of three or five years ago may be outdated.

Rising seas, higher and frequent floodwaters, more violent storms, and other effects caused by climate change may pose an unexpected and potentially large challenge to data center operators. According to Uptime Institute, even where a facility rides out the challenges unharmed, damage to local infrastructure can deprive a facility of staff, utilities and fuel, and access to telecommunications.

Data center operators can mitigate these consequences by conducting frequent facility availability assessments, with particular emphasis on how flooding or drought, high winds, and warmer temperatures would affect the facility's availability to cope with new extremes. Assessments should involve local government, utilities, and telecommunications companies with information about local disaster recovery plans and include specific actionable items.

8. Economics will drive acceptance of data center AI...eventually

AI-based approaches to analyzing data center risk and efficiencies, including via new cloud services, will be proven at scale, driving mainstream acceptance and, over time, high levels of adoption.

Data center operators are often perceived to be wary of AI, and indeed of advanced software in general, but this is somewhat misleading. They are wary of unproven or risky technologies and strategies. Machine learning (ML), for example, has been adopted by data centers for at least a decade, most commonly in the form of dynamic cooling optimization software that automates cooling units' operational status, variable frequency drives, and temperature setpoints. However, the launch of the first data center management as a service (DMaaS) offerings in 2016 signaled a new wave of AI proliferation in data centers.

With DMaaS, large sets of monitored data about equipment and operational environments from different facilities (and different customers) are encrypted, pooled in data lakes, and analyzed using ML, anomaly detection, event-stream playback, and other approaches. ML-driven outcomes have largely focused on improving existing processes such as shortening alarm lead times and root-cause analyses, improving PUE, and optimizing utilization levels – use cases that deliver tangible operational savings.

In 2019, most DMaaS will be extended proof-of-concepts with a growing number of full adopters. We also expect that new insights from ML approaches in 2019 will lay the foundation for increased automation in the future, proving the technology has moved beyond just hyperscale facilities.

9. Growing threats will necessitate new ‘zero-trust’ approaches

Security vulnerabilities now affect mission-critical facilities. Organizations will need to adopt more stringent policies regarding data center equipment, services, contractors, suppliers, and staff.

Data center operators are becoming ever more aware that, while IT systems make obvious, high-profile, and probably well-understood targets for hackers and criminals, they need to invest more money and attention on the physical infrastructure. Physical-perimeter security, data center access, and internal private security and governance continue to be a focus of attention – yet other areas are often neglected. A persistent and vexing issue for many operators, for example, is adequately monitoring vetted third-party people and their devices once inside a facility, including the increasingly digitized work they carry out. There are growing concerns about the need to vet and watch not only the people but all the equipment that is brought into a data center.

As data centers become smarter and more connected, using IP-based controls and equipment, the ways in which they can be compromised grows. Organizations often connect to controls through private networks and grant access to the outside world for vendor support of building management/automation systems and equipment. While IT departments are gravitating toward security approaches such as zero-trust, whereby all access attempts to IT systems, both internal and external, are verified before granted, there are relatively few data center departments that are similarly equipped to mitigate a growing list of threats.

In 2019, more data centers are likely step up their approaches and move toward conditional-access management policies, whereby access to specific network resources is granted only if certain conditions are met, including time of day, location of access, and so on. If any condition is unmet or if there is unexpected behavior, access is revoked. The threats are growing, ranging from third-party contractors to operational cloud services and increasingly automated, software-driven equipment – security investment and focus will need to keep pace.

10. Programmable power unlocks new efficiencies and agility

A combination of Lithium-Ion batteries and other new forms of energy storage give data center operators a new set of levers to help improve data center performance.

As data centers get larger and larger, the cost of power distribution equipment and the power itself become ever greater. Reducing these costs, especially in an environment where cloud operators are trying to drive them down, is becoming a management priority. A big challenge for operators is to cut power costs without increasing risk.

Power costs can be reduced in three areas:

- Operations: where energy overhead (i.e. PUE) can be reduced
- IT load: where utilization can be driven up or gear power consumption be driven down
- Supply: regarding utility costs, storage and service commitments

Innovators have been working on cost reduction for nearly two decades, especially on the IT side, but progress in the facilities realm has been limited. Uptime Institute Research believes this may be about to

change. These include the slowing of Moore's Law, the rise of cloud, improved management tools, the integration of IT and facilities, and the rapid adoption of Li-Ion batteries.

Conclusion

The data center industry is changing at a pace never seen before. Technologies and best practices that have been accepted for years are now being challenged. Technical approaches such as cloud and edge, automation, security, and network are all becoming the focal points of modern high-performance infrastructure. And business issues including regulatory intervention, staffing, and climate change are now part of our daily vernacular. In 2019, infrastructure owners and operators would be wise to look deep within their portfolios, their processes, and their people and assess the effectiveness of each in delivering business services based on their existing hybrid infrastructures.