Uptime Institute® | RESEARCH

EMERGING TECHNOLOGIES

UI Intelligence report 21

Ten data center industry trends in 2019

Key innovations, market developments, and challenges to watch

Lead Analyst
Andy Lawrence (alawrence@uptimeinstitute.com)

Contributing Analysts
Rhonda Ascierto (rascierto@uptimeinstitute.com)
Kevin Heslin (kheslin@uptimeinstitute.com)

UII-21 v1.1 published November 29, 2018 last updated August 6, 2019

Ten data center industry trends in 2019

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 opportunities but throwing up challenges. 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 should data center owners, managers, and operators expect in 2019? Which innovations will make a difference? Where are problems and challenges likely to emerge? Uptime Institute Research has identified 10 areas for 2019 where it is worth paying close attention.

Summary: Top 10 data center industry trends in 2019

Big Cloud builds push the ecosystem to its limits
The accelerating demands of the big cloud operators for more data center capacity is 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 result: more government vigilance, more oversight, more regulations, and more taxes.

The transition to distributed resiliency will not be smooth
Disruptive and often high-profile outages will continue as operators seek to support ever more applications and services on a low-cost base—and grapple with the complexities of deploying distributed, hybrid systems across multiple data centers and services.

Edge data center hype outruns deployment
The promised explosion in demand for small, edge data centers is coming—but not yet. Issues with security, costs, business models, integration, networking, and 5G roll out will continue to hold back large-scale deployments.

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. Big data center operators and suppliers of software-defined networking fabrics are working to establish themselves as essential providers in a software-driven, distributed world.

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 (especially the biggest ones) will work to diversify the talent pool, with new initiatives, hiring strategies, and new workforce training.

Climate change forces fresh review of resiliency planning
The risks associated with climate change may be more varied and extensive than IT planning had previously anticipated. Without reassessment of data center resilience plans, major service failures will be more likely.

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 to corporate IT now encompass mission-critical facilities. Organizations will increasingly realize they can take nothing on trust and will adopt more stringent policies affecting all data center equipment, services, contractors, suppliers, and staff.

Programmable power unlocks new efficiencies, agility
A combination of Lithium-Ion batteries and other new forms of energy storage, along with configurable, intelligent management tools, means data center operators have a new set of levers to help improve data center performance.

TREND 1:

Big Cloud builds push the ecosystem to its limits

The accelerating demands of the big cloud operators for more data center capacity is 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.

In some regions, demand will soon strain or outpace supply—for power, staff, real estate, connectivity, equipment, components, water, and other resources. Shortages and constraints will depend on the geography—real estate in London, power in India, components in Brazil, water in Southern Europe and the South East US, skills in almost all regions. This may affect data center design, encouraging automation or reducing the use of water.

Total hyperscale capacity growth has long outpaced the rates of other data center segments, and future growth is forecasted at an an exponential pace, with demand uneven across regions (see table below).

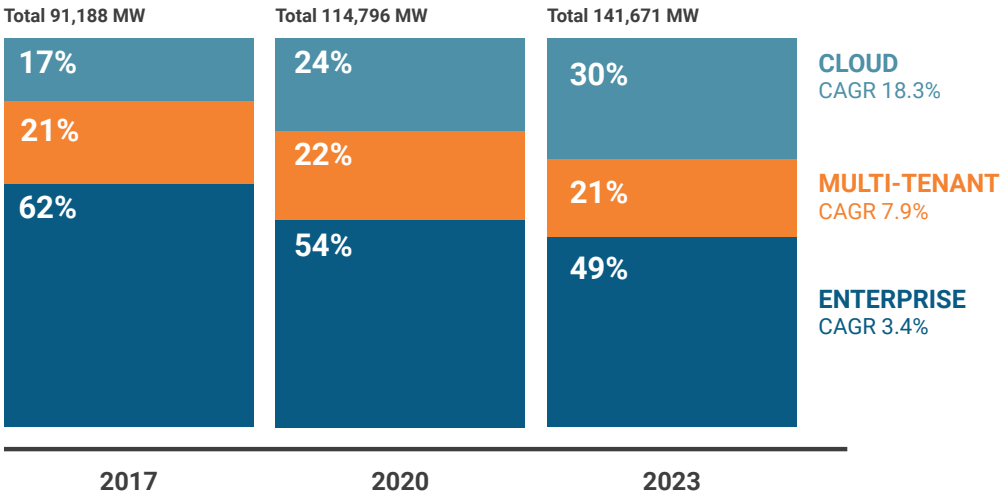
The big cloud companies are preparing on several levels. They are investing more in on-site and near-site power generation, and adopting real-estate alternatives, even considering underwater facilities. In spite of their competitive status, they are collaborating more, including on IT architectures and on staff pipeline development. Meanwhile, wholesale data centers are planning more speculative builds, assured that demand will materialize. Facility equipment makers are investing in new engineering, centers and test sites, where they are working on large-scale designs, value engineering, and rapid deployment. In 2018, these and other approaches are so far succeeding—supply and demand are generally in equilibrium. But the future looks less certain.

In 2019, we expect operators and suppliers to focus more on 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 hyperscalers' expected future demand. Smaller suppliers

will strive 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.

THE GREAT BUILD OUT

Data center UPS power (net MW, global)



Source: 451 Research Data Center Monitor, Q2 2018

Hyperscalers include:
Alibaba, Amazon, Apple, Facebook, Google, IBM, Microsoft, Oracle, SAP, Tencent

Hyperscalers lease capacity in multi-scale facilities:

- supplement growth
- rapidly establish presence
- create PoPs for customer on-ramps
- test new markets (some lease for all <1MW

Uptime Institute Research, 2018

TREND 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. The result: more government vigilance, more oversight, more regulations, and more taxes.

In recent years, governments viewed "Big IT," and the wholesale 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.

The rapidly growing legal departments at hyperscalers (Apple, Google, Facebook, Microsoft, Amazon, Tencent, Alibaba, and others) will work to deal with legislation and restriction across many fronts and geographies.

Governments across the globe have already enacted an array of new regulations and taxes, which 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 limit social disruptions. Expect more in 2019 and beyond.

Regulations like GDPR in the EU limit or regulate the movement of data, making it costly for IT companies to comply, and in some cases, requiring IT operations to locate new facilities in specific places to meet a government's data restrictions. Organizations face penalties for noncompliance with increased security and reliability directives, third-party audits, and incident reporting requirements. Recent downtime incidents in financial services has UK regulators considering new rules on resiliency, the European Union is taking another look at data center energy use, and fears over energy security and hacking have prompted renewed discussions on hardening critical infrastructure.

Similarly, national governments (in the UK and Spain, for example) are unwilling to wait for a global or regional consensus on taxation. They want tougher, international laws to force big operators to pay more. Expect that new taxes, including those levied on digitally fulfilled or ordered goods and services, will be introduced with the goal of protecting local business and sharing national infrastructure. Expect too, that large US-based tech giants and low-tax nations that currently benefit from the status will protest furiously and loudly that their investments are hugely beneficial regardless of direct taxes, energy use, or other negative impacts.

Governments extend their control of IT activities through regulation and taxes

JURISDICTION	RULE	INTENDED EFFECT	KEY DATE	NOTES
California (US)	Statewide Net Neutrality	Limits advantages and disadvantages "pipes" owners can provide content providers	Enacted September 30, 2018	Effective date is blocked by litigation filed by the US federal government
Russia	Draft Law: On an Autonomous Internet System	The bill calls for placing the domains .ru and .рф under government control and would make installation of the Russian state surveillance system SORM mandatory	Bill drafted May 27, 2016	Not yet advanced, subject to revision
UK	Discussion Paper: Building the UK financial sector's operational resilience	The discussion paper seeks input on ways to improve the effectiveness of IT resiliency in the UK financial services sector	NA	Intended to address IT resiliency shortcomings that have caused recent IT outages in the UK's financial sector, possibly via a framework
UK, US, NZ, EU, elsewhere	Digital services taxes	New funding, protect "bricks and mortar" retail	UK: Effective 2020 US: JUN 21, 2018 (Quill ruling, SCOTUS) NZ: Effective April 2020 EU: Initially proposed March 2018	UK: 2%, US: Varies by state and local taxing authority, NZ: TBD, EU: 3%
European Union	Network and IS Directive	Combat cybercrime	Enacted August 1, 2016	Member States have to transpose the Directive into their national laws by May 9, 2018 and identify operators of essential services by November 9, 2018
China	New Security Directive	Prevent "the spread of illegal and harmful information"	Effective November 30, 2018	Requires social media businesses, online forums, video services, and search engines to routinely collect detailed user information and to establish systems for reporting this information to the police when requested

Source: Uptime Institute Research, 2018

TREND 3:

The transition to distributed resiliency will not be smooth

Disruptive and often high-profile outages will continue as operators seek to support ever more applications and services on a low-cost base—and grapple with the complexities of deploying distributed, hybrid systems across multiple data centers and services.

In 2018, Uptime Institute published research that showed the number of outages affecting the data center sector was running as high as ever—and furthermore, the costs and impact of these outages were growing. The research was puzzling to many: Surely, better and newer technologies, and better processes, would have reduced the number and impact of outages?

What the findings pointed to is that 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. 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.

The end result should be a more resilient infrastructure that meets failures, incidents, and even regional disasters with a smooth and largely automated response. Cloud operators are able to offer a very high level of availability by replicating data and services across regions using availability zones. Many enterprises are doing the same. But all operators need to learn more, to invest more, and to build in processes and governance.

In 2017 and 2018, there were incidents or outages at organizations such as AWS, Google, Microsoft, and GitHub that clearly showed the complexity of managing distributed and replicated databases and global traffic management systems, and of embedding availability zones into service offerings. Other failures at colocation and enterprise companies showed how M&E and network failures could cause IT problems that cascade across a network.

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.

TREND 4:

Edge data center hype outruns deployment

The promised explosion in demand for small, edge data centers is coming—but not yet. Issues with security, costs, business models, integration, networking, and 5G rollout will continue to hold back large-scale deployments.

A resurgence in distributed IT is underway from the Internet of Things (IoT), 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
- Those far away—such as economical hyperscale facilities, ('core') for further processing, analysis, and archiving

The paths and volume of data generated are different for different edge applications—as is the latency and resiliency required. New distributed and cloudy application environments for the edge are emerging (though not yet at scale), as are new deployment models. Most of the major cloud providers are evolving their edge strategies, with a focus on hybrid (on- and off-premises) deployments.

Edge data will be managed by artificial intelligence (AI) and run in micro data centers—these will hand off data and state to core/cloud platforms running in larger facilities. Implementations of edge compute, storage, and control system stacks will continue to be varied but the deluge of edge data will benefit all types of venues.

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. (Some will need higher densities and will be liquid cooled.) 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
- For supporting IoT appliance data analysis, storage, and resiliency in smart factories, buildings, and elsewhere
- As nodes for very fast sub-4 ms latency at cell towers
- As next-generation nodes for 4G and 5G access and aggregation points

Industrial, retail, medical, and telco sectors are among today's buyers of micro data centers.

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. It is still very early days for 5G, for driverless cars, for immersive and augmented reality, and even for AI; many networking, energy, and resiliency technologies are vying for investment and take up; and the ownership and control of small data center facilities and edge networks has to yet settle into a clearly investible pattern.

A proliferation of production projects is inevitable but it will take time, which means new IT capacity in the form of micro data center deployments is still at the base end of the curve.

TREND 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. Big data center operators and suppliers of software-defined networking fabrics are working to establish themselves as essential providers in a software-driven, distributed world.

Software-defined networking fabrics (SDN) are secure online platforms that allow virtual private interconnections to others on the platform 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, with 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.

Colocation providers see the future as SDN fabrics, with the largest operators (such as Equinix) positioning themselves to serve as the single supplier of network configurations for customers, as well as a primary supplier of leased data center capacity. Their expansive portfolios and customer ecosystems will be difficult to replicate (if only from a real-estate standpoint), and it is likely that they will enjoy a market-leading position in interconnection SDN fabrics for some time—there is a view that the market will mature into a 'winner takes most' situation.

However, while most organizations are adopting hybrid or multi-cloud architectures across several different providers, many are still in the early stages of developing or executing their cloud strategies, while others are looking for a more diverse set of partners. There is still much to play for, especially as network operators without their own data centers (such as Megaport) offer alternatives.

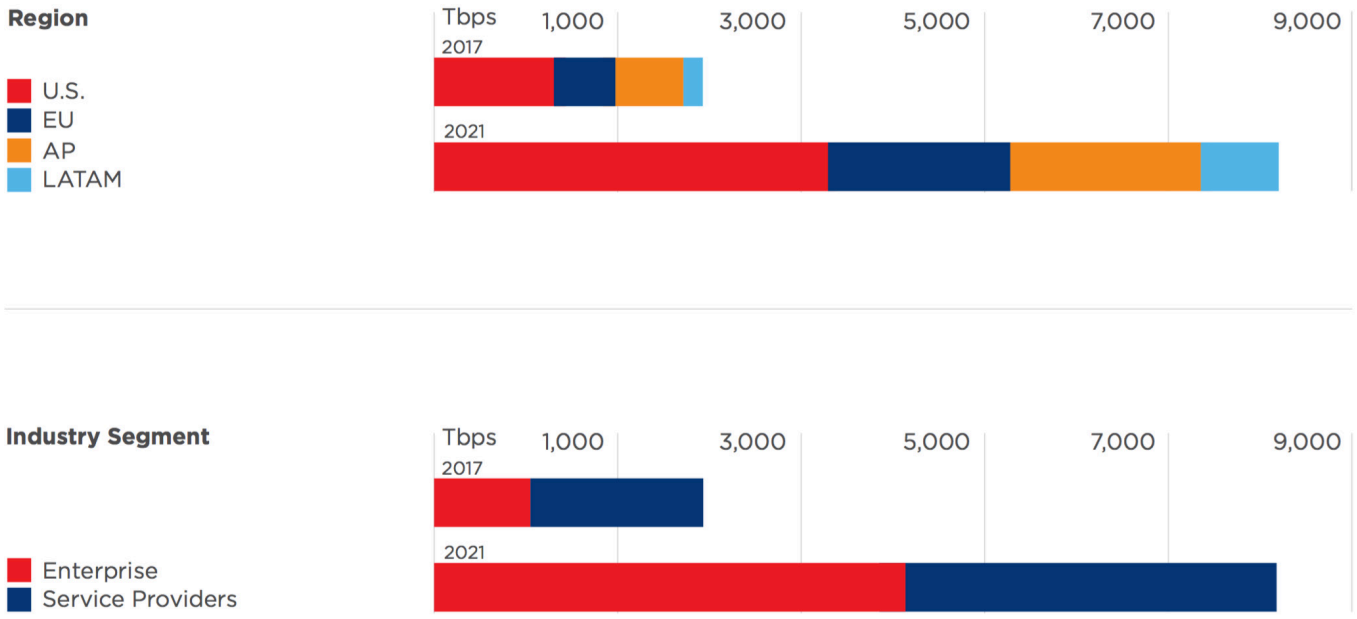
As organizations' digital footprints continue to sprawl (both geographically and by the number of interdependent IT services they consume), demand for and reliance on SDN fabrics will grow. In 2019, we believe more SDN fabrics across colocation providers' data centers will be established. How the various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

The overall opportunity for exchanges and network operators is vast and long term, and rapid growth is forecast through 2021 (see 'Interconnection Bandwidth Forecast' below), leading some

to speculate on potential network shortages. It seems likely that the market will support a variety of operations—and in time, many more services. SDN platforms, built on general purpose hardware, can be used to develop applications such as security management, efficiency and latency reporting, disaster avoidance, network path transparency, immutability-as-a-service, container orchestration, and many others.

To better compete, smaller colos with fewer data centers and/or smaller ecosystems are likely to form partnerships, including with vendor-neutral SDN network-as-a-service specialists, telcos, or other colo providers. But the largest suppliers are likely to be protective of their opportunity and offer ever more services to improve network performance.

Interconnection Bandwidth Forecast



Bandwidth = total capacity provisioned to privately and directly exchange traffic with a diverse set of counterparties and providers at distributed IT exchange points inside carrier-neutral colo data centers

Source: "Global Interconnection Index, Volume 2," Equinix, September 2018, Page 6

TREND 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 (especially the biggest ones) will work to diversify the talent pool, with new initiatives, hiring strategies, and new workforce training.

Uptime Institute Research survey results suggest that data centers will continue to struggle to recruit and retain sufficient qualified staff to comfortably maintain and grow reliable operations. In one of our 2018 surveys, 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.

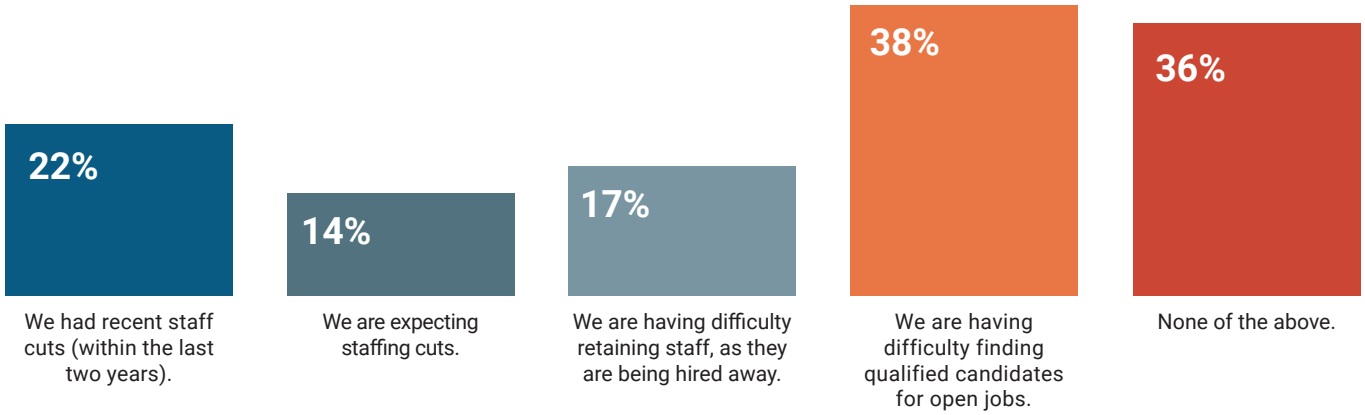
Uptime Institute advises that the skills shortage will be felt throughout the industry, with companies struggling to fill traditional facilities positions such as operations and management, security, network connectivity, cloud provisioning, and mechanical engineering but also compliance (22%), contracts/SLAs/vendor management (21%), software (21%), financial management (11%), supply-chain management (8%), and corporate and social responsibility (8%).

Industry may find customary recruitment tools to be relatively ineffective. *The Wall Street Journal* recently reported, "With relatively few unemployed Americans looking for work, employers are increasingly having to bid up wages to poach workers from other employers. That has been happening for several years for higher-skilled workers such as engineers and welders, but now it is occurring for relatively lower-skilled workers…."

With unemployment rates in countries that host the most data centers ranging from under 3% to about 5.5%, competition for workers should be intense globally. Furthermore there is a perception that salaries offered by the Internet giants are increasing wages across the board.

Resolving the gender imbalance in the industry would bring large numbers of qualified candidates who could help close the skills gap. Almost two-thirds of suppliers surveyed by Uptime Institute Research in 2018 believe that the industry will succeed in recruiting more women in the next three to five years. Employing more generalists and cross training existing staff are also options that will increase the flexibility of workforce schedulers.

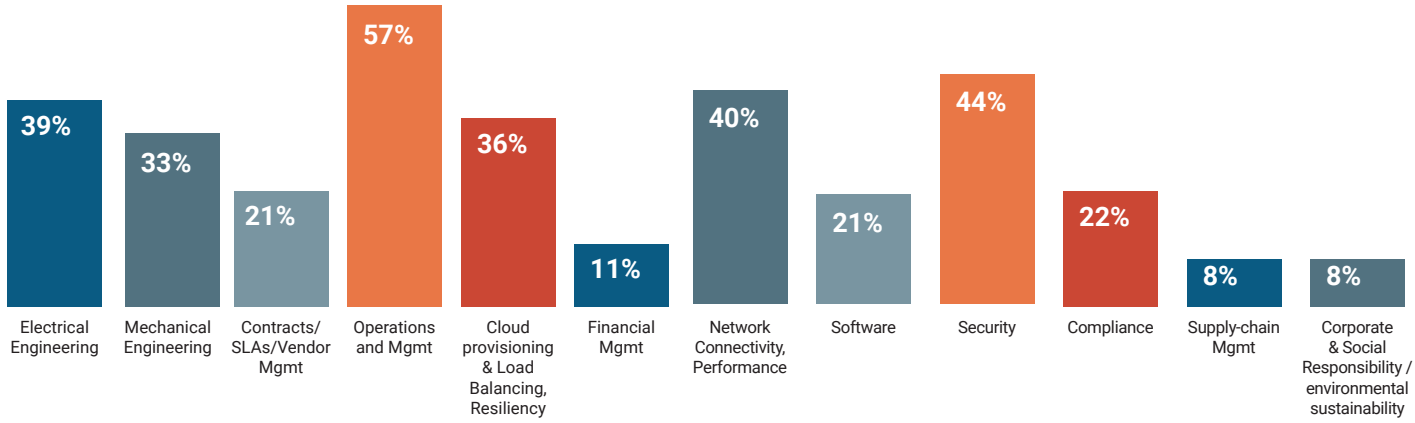
Finding qualified candidates is a challenge for many IT organizations



Staffing: Please select any of the following statements that apply

Uptime Institute's global annual survey of data center operators, 2018, n=538

Operations, management, and security roles are hardest to fill



What disciplines or skill sets are most needed in your organization to support IT Infrastructure? Select areas of expertise that are particularly critical and difficult to hire for:

Uptime Institute's global annual survey of data center operators, 2018, n=526

TREND 7:

Climate change forces fresh review of resiliency planning

The risks associated with climate change may be more varied and extensive than IT planning had previously anticipated. Without reassessment of data center resilience plans, major service failures will be more likely.

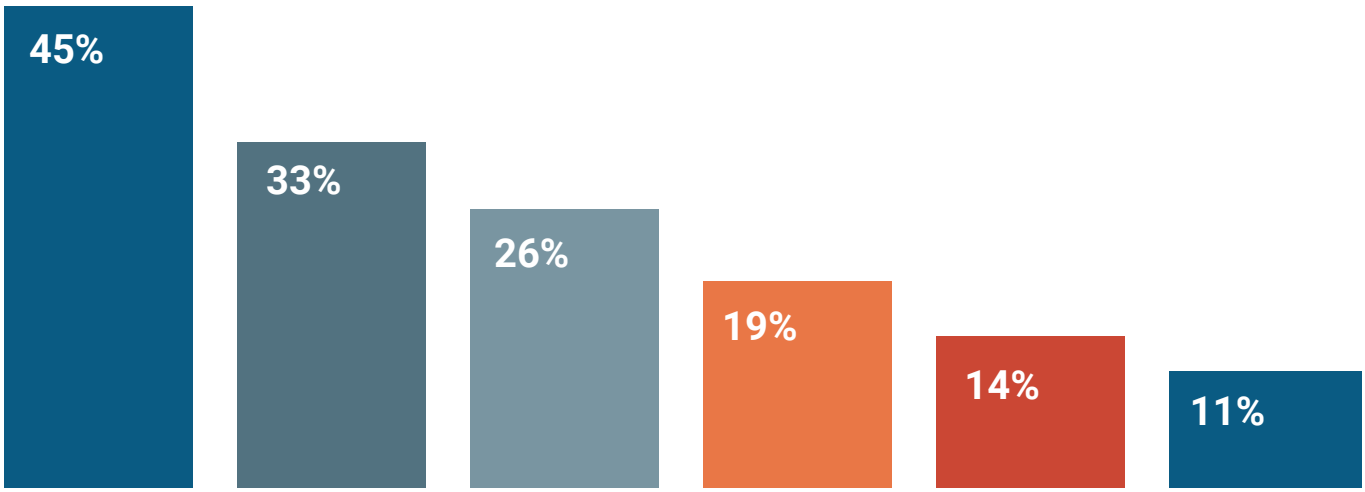
No one could accuse the data center industry of complacency or ostrich-like behavior. Conservative risk assessment lies at the core of data center design and operations. Even so, in the field of climate change, which has become politicized in many countries, many data center owners and operators are 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. Economic models can also be affected: in 2018, drought in Sweden, for example, forced hydropower plants to shut some turbines, increasing power costs and forcing some crypto-miners to suspend operations.

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.

In addition, IT operations should conduct multi-site IT resiliency assessments for extreme weather events, when essential services may be unavailable for an extended time. This assessment should ensure that an organization can shift IT load to another facility or a public or private cloud.

Many facilities have taken no action to mitigate climate change risk



Is your organization revising policies and/or planning to revise data center policies to adapt to climate change?

Uptime Institute's global annual survey of data center operators, 2018, n=709

TREND 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 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. Today, several suppliers offer DMaaS, paralleling a similar development at large data center operators, which use internal data from across their portfolios. The technology has moved beyond just hyperscale facilities.

So far, ML-driven outcomes have largely focused on improving existing processes such as increasing alarm lead times and decreasing root-cause analysis times, improving PUE, and optimizing utilization levels—use cases that deliver tangible operational savings. There have also been some standout applications, such as predictive maintenance and peer benchmarking, and in 2019 we expect there will be more. New insights are highly likely as suppliers and large companies apply ML to analyze a wider range of relationships and patterns among a vast range of variables, including resource use, environmentals, resiliency, and equipment configurations.

ML can be used to help data center operators reduce risks and improve overall responsiveness and costs by modelling and anticipating failure-rates, budgetary impacts, supply-chain needs, and the impact of design changes and configurations. Even data centers that have not been constructed can be modelled and simulated in advance (for example, to compare an operational/performance profile and total cost of ownership of a Tier II design data center vs a Tier III data center).

New insights will be drawn on large and/or more diverse sets of clean (prepared) data, driven in part by greater adoption of DMaaS—services that typically leverage low-cost cloud environments, making them inexpensive and easy to deploy (especially when compared to on-premises DCIM software). The usefulness of new insights will, at least initially, heavily depend on data center-domain knowledge from humans, including to validate results and to mitigate potential risk or stupidity.

In 2019, most DMaaS and other AI-driven management approaches will be extended proof-of-concepts with a growing minority of full adopters. We also expect that new insights from ML approaches in 2019 will lay the foundation for increased automation in the future (with automation management residing largely on-premises rather than as a cloud-delivered function), including for "self-optimizing autonomic data centers"—the highest level of Uptime Institute Research's data center management maturity model.

Data center management maturity model

LEVEL	DESCRIPTION	OPERATING EFFICIENCY	SOFTWARE
Level 5: Self-optimizing, autonomic	AI-driven integrated management software adjusts data center behavior & makes best use of resources according to goals, rules & service requirements throughout its lifecycle.	HIGH	AI-driven, integrated DCIM with automation
Level 4: Optimizing	Physical & virtual IT and data center subsystems integrated; models used for prediction, service management & multiple views, optimizing in near real time. AI is applied to DCIM-based data lakes for advanced analytics.	MEDIUM	AI-driven, integrated DCIM
Level 3: Proactive	Physical data center equipment characteristics, location & operational status is tracked. Energy & environmental data is used to reduce risks, waste.	MEDIUM	Integrated DCIM
Level 2: Reactive	Software installed to monitor environmentals and equipment power use. Ability to adjust basic controls (i.e. cooling) to demand.	LOW	DCIM monitoring
Level 1: Basic	No integration of infrastructure data. Basic monitoring supplied with equipment. A reliance on BMS data. Simple alarming, error messaging.	LOW	Ad hoc

Source: Uptime Institute Research and 451 Research, 2018

TREND 9:

Growing threats will necessitate new 'zero-trust' approaches

Security vulnerabilities to corporate IT now encompass mission-critical facilities. Organizations will increasingly realize they can take nothing on trust and will adopt more stringent policies affecting all 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 focuses 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.

Although there have been very few known incidents, it is unlikely that hackers and attackers are unaware of the damage they could cause by interfering with mission critical equipment. 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. And more cloud-delivered operational services are being adopted (see Trend 8), sending operational data to a provider's cloud, as well as to mobile devices used by on-site technicians and managers. All of these complicate and widen the "attack surface" of the data center and may require new policies and investment.

Other threats include the surfacing of site location details, both publicly (Wikileaks recently published a 2015 list of Amazon's, for example) and via the dark web. Then there are claims that Chinese "spy chips" have been found on Supermicro motherboards. Supermicro customers Apple and AWS have denied they were victims—but regardless of their validity, the claims highlight an important area of vulnerability.

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.

TREND 10:

Programmable power unlocks new efficiencies, agility

A combination of Lithium-Ion batteries and other new forms of energy storage, along with configurable, intelligent management tools, means data center operators have 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 becomes ever greater. Reducing theses costs, especially in an environment where cloud operators are trying to drive down costs, 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:

- In operations, where PUE can be driven down
- In IT, where utilization can be driven up or machine power consumption be cut or reduced
- In dealing with the utility—which includes selling power back

In all three areas, the most promising innovation at is present the application of intelligence through integrated systems or software—so that, for example, a management system can actively select when to reduce power supply, demand, to change balance of capacity or provisioning, or to switch supply sources.

Innovators have been working in this area for nearly two decades, especially on the IT side, but adoption, for many reasons, has been extremely limited. Even the category name (programmable power? software-defined power?) is unclear. But this, Uptime Institute Research believes, may be about to change—again for many reasons. 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 Lithium-Ion (Li-Ion) batteries.

Li-Ion may fundamentally change the way power storage is used in a data center, as it can be safely distributed in the white space, and batteries can be emptied and recharged hundreds or thousands of times without serious degradation. Major suppliers cite increasing adoption. For example, in terms of total power (kVA) it sells per annum, about 70% of Schneider Electric's three-phase UPSs are now Li-Ion enabled; and Vertiv says it shipped more than 100 MW of Li-Ion batteries into data centers in North American in 2018 to date, up 300% year-over-year (from a relatively low starting base).

In programmable power or software-defined power, Uptime Institute sees many areas of innovation in 2019, with early stage adoption beginning to gather pace. One area is the use of software and Li-Ion batteries to balance supply and demand from rack to rack, an area pioneered by Virtual Power Systems. Another area is the use of systems to manage a pool of shared UPS power, diverting capacity to where it is needed, and a third is the use of large racks of Li-Ion batteries and software-controlled UPS systems to identify over-provisioning or safe conditions, and then sell capacity back to the utility grid.

In all these areas, it is very early days and 2019 is likely to be a year of innovation and proof, rather than widespread take up. But the advantages in cost and possibly resiliency will likely to be big enough to warrant follow-on investment.

ABOUT THE LEAD ANALYST

Andy Lawrence is Uptime Institute's executive director of Research. Mr. Lawrence has built his career focusing on innovative new solutions, emerging technologies, and opportunities found at the intersection of IT and infrastructure.

ABOUT UPTIME INSTITUTE INTELLIGENCE

Uptime Institute Intelligence is an independent unit of Uptime Institute dedicated to identifying, analyzing and explaining the trends, technologies, operational practices and changing business models of the mission-critical infrastructure industry. For more about Uptime Institute Intelligence, visit https://uptimeinstitute.com/ui-intelligence.

All general queries:
Uptime Institute
5470 Shilshole Avenue NW, Suite 500
Seattle, WA 98107
USA
+1 206 783 0510
info@uptimeinstitute.com

