PLANNING
& STRATEGY
UI INTELLIGENCE REPORT 51

# Uptime Institute Global Data Center Survey 2021

## Growth stretches an evolving sector

Uptime Institute's annual survey, now in its eleventh year, is the most comprehensive, longest-running study of its kind in the data center sector. The findings reveal significant growth but also increasing complexity and challenges for owners and operators of mission-critical digital infrastructure, and for the suppliers that serve them.

Uptime Institute® | INTELLIGENCE

**AUTHORS**
**Daniel Bizo,** Research Director, Uptime Institute
**Rhonda Ascierto,** Vice President of Research, Uptime Institute
**Andy Lawrence,** Executive Director of Research, Uptime Institute
**Jacqueline Davis,** Research Analyst, Uptime Institute

UII-51 V1.1M PUBLISHED SEPTEMBER 1, 2021 LAST UPDATED DECEMBER 10, 2021

35-45 MINUTES TO READ

# Synopsis

Uptime Institute's 2021 Global Data Center Survey reveals a vibrant sector that is growing and adapting to rapid change on multiple levels. It is also entering a phase of re-assessment — of infrastructure and service accountability, in terms of resiliency, and of material outcomes toward environmental sustainability and efficiency.

Owners and operators are taking many different approaches to a range of different challenges, such as improving resiliency, being more sustainable, maintaining staff levels, using the cloud, or building out edge data center capacity.

- Further industry-wide efficiency gains will be found at the IT rack level.
- Data center sustainability will grow in importance, yet practices are lagging; most organizations do not closely track their environmental footprint.
- Roughly half of all data center outages cause significant revenue, time and reputational damage.
- Data center staffing and supply chains are being stretched by continued capacity growth.
- Lithium-ion battery use in centralized uninterruptible power systems is now common.

# Contents



## 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 digital infrastructure industry. For more about Uptime Institute Intelligence, visit uptimeinstitute.com/ui-intelligence or contact research@uptimeinstitute.com.

## Illustrations

# Introduction

The Uptime Institute Global Data Center Survey, now in its eleventh year, takes snapshots from various angles to create a current profile of the mission-critical digital infrastructure sector and a sense of future direction.

The view is a complex one, as it captures operational performance achieved through efficiency, availability and sustainability practices, as well as trends in capacity growth, capital spending, technology adoption, staffing, cloud usage and edge computing.

The findings in this report are primarily based on the practices and experiences of IT and data center managers globally. Additional insight is provided by their equipment and engineering suppliers. The survey was conducted online and by email during the first half of 2021. For more details, including demographics, see the **Appendix**.

# Efficiency metrics stutter

Certain metrics, though imperfect, are widely used to define data center efficiency. Among them are power usage effectiveness (PUE) and rack densities. Despite improvements in both over the past decade, it seems clear the industry needs to now focus on measures at the IT rack-level for greater efficiency gains.

## PUE still flatlined

While PUE does not capture the efficiency of IT, it remains a useful proxy for how much power data center facilities use in addition to that used by the IT equipment. In 2021, survey respondents' average annualized PUE globally was 1.57, meaning data center facility functions added nearly 60% to the energy use of IT. This is a notch below last year's 1.59, and in line with longer term directions.

Overall, the picture is clear: After large efficiency gains through the first half of the 2010s, average PUEs have remained relatively stable for the past five or so years, as shown in **Figure 1**.

There is a clear explanation for this. Even as a growing number of new builds sport design PUEs of 1.3 or better, it is not economically or technically feasible for many operators to perform the major overhauls needed for better efficiency in many older facilities. Across much of this large population of older data centers, the easy gains from better airflow management, optimized controls and replacement of aging equipment have already been achieved. Further improvements will require significant change. In existing facilities this may include retrofitting with highly efficient cooling; in new data center designs it may require adopting innovative approaches or technologies, such as direct contact liquid or solid-state cooling.

# Density rises, slowly

The power density of IT racks influences infrastructure engineering choices in many ways, such as layout, cooling architecture and power distribution. When operators considerably underestimate density, the facility may struggle to deploy the latest IT systems or run out of power before the data hall is full. When they grossly overestimate the density, the penalties include sunk costs that cannot be recovered.

Our 2021 survey shows rack density remains relatively moderate — typically well under 10 kilowatts (kW) per IT cabinet — for most respondent data center owners and operators, even at their flagship sites. Over a third say their most common rack density is under 5 kW (modal average), while for nearly half it is 5 kw to 10 kW.

Our data hints at a link between power density and the size and efficiency of a data center. There is a shift toward more powerful racks, between 5 and 10 kW, in facilities larger than 3 megawatts (MW) of maximum IT load supported, compared with smaller sites. Rack densities tend to be even higher in data centers with capacities above 5 MW.

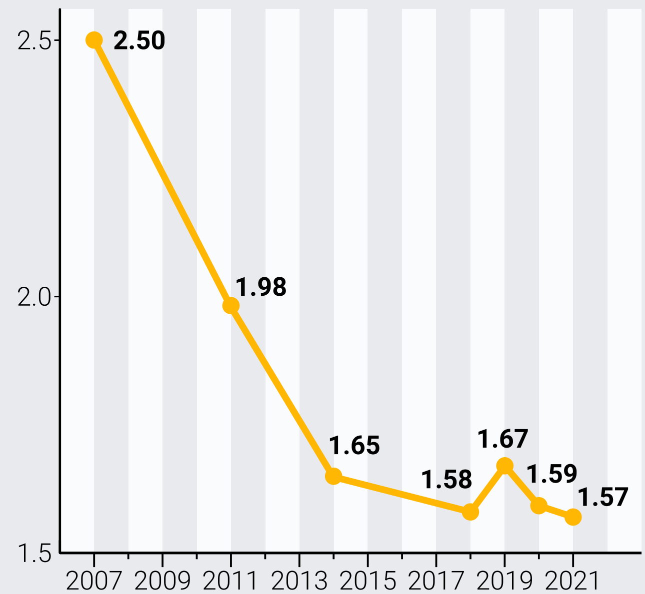
Another shift happens in the largest facilities, above 30 MW: racks in the 10-20 kW range are far more common than those under 5 kW. A likely explanation is that facilities at this scale tend to serve cloud and internet companies (be it their own site or a leased data center), which typically prefer cabinets fully racked with high-performing IT.

A similar correlation emerges between density and PUE — the more efficient a facility is by design, the more likely it is to have higher modal average power density. This is likely because larger data centers are typically newer, as well as better optimized for efficiency.

FIGURE 1

**PUE gains have stalled**

What is the average annual PUE for your largest data center?




UptimeInstitute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2007-2021 (n=566)

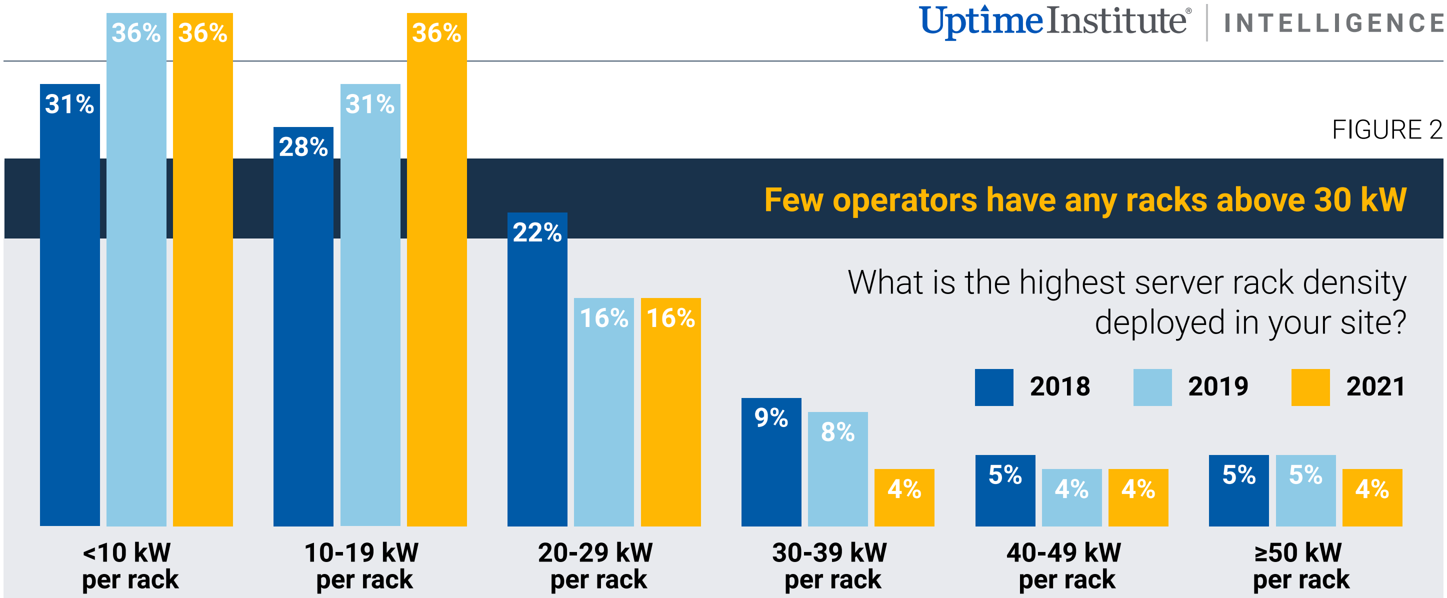
Data centers with annualized PUEs above 1.6 are dominated by relatively low-density racks of 5 kW and below. Those with PUEs in the 1.3 to 1.6 range have racks mostly between 5 and 10 kW. More than one in six in this lower-PUE group are running racks of 10 kW or greater as standard practice.

There is a yet more pronounced shift in sites with annualized PUEs of 1.3 or better: low rack densities under 5 kW are uncommon, while racks above 10 kW power are the norm for over a third.

FIGURE 2

**Few operators have any racks above 30 kW**

What is the highest server rack density deployed in your site?




UptimeInstitute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2018 (n=431), 2019 (n=462), 2021 (n=553)

When we asked owners and operators about their highest density racks — and exclude high performance compute (HPC) operations — few (one in 10) have any above 30 kW. About one in six have maximum densities of 20-30 kW racks.

These numbers follow the pattern for modal average densities: the bigger the data center, the more likely it is to have some higher density cabinets. About half of those in our study that operate data centers above 5 MW have some racks over 20 kW, and a quarter of those above 30 MW report some racks over 30 kW (**Figure 2**).

Data centers that run a meaningful number of high-density cabinets are candidates for step changes such as direct liquid cooling (cold plates or immersion). However, many may not be ready to do so because they can already deliver large volumes of cold air to their high-density racks — the incentive for change, often regardless of cost and environmental benefits, can be lacking.

# Sustainability and measurement

The global growth of data center capacity has led to higher scrutiny of resource use by customers, pressure groups, media and, more recently, policymakers. Data center environmental sustainability is now a concern for more senior executives and also for more data center equipment makers, some of which are responding by changing their approaches and products.

Data center sustainability is multifaceted and requires operational efficiency — and efficiency starts with measuring resource use. More than four in five owners and operators surveyed track data center power consumption, as shown in **Figure 3**. Most also calculate PUE.

The data confirms what we typically see in the field: power usage metrics are commonly tracked in facilities larger than 100 kW and reporting on PUE is prevalent in those larger than 1 MW. This is likely to change as more smaller edge facilities are deployed, especially as many micro data centers incorporate advanced monitoring (see **The data center edge expands**).

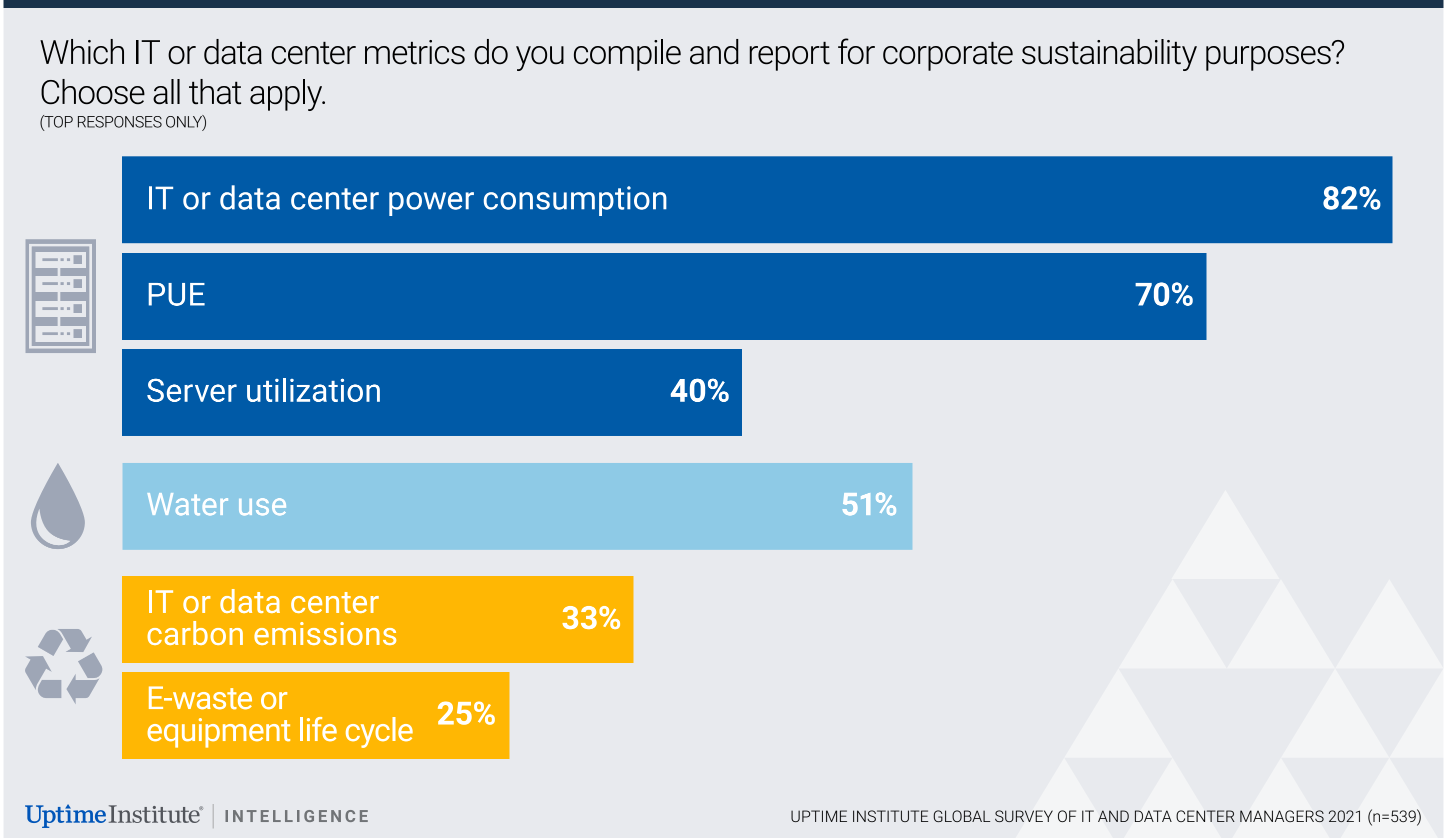
Electricity is the single largest operational facility cost, so it is not surprising larger data centers closely watch infrastructure power consumption rates and usage effectiveness. Even so, it is still common at smaller and privately owned enterprise data centers that the electric utility bill is paid by property management, which may have little interest or say in how the compute infrastructure is built or operated. This disconnect could explain why most still do not track server utilization, arguably the most important factor in overall digital infrastructure efficiency. Even fewer operators track emissions or the disposal of end-of-life equipment, which underscores the data center sector's overall immaturity in adopting comprehensive sustainability practices.

"Most still do not track server utilization, arguably the most important factor in overall digital infrastructure efficiency."

Many data centers consume large volumes of water, a specific (localized) environmental concern in water-stressed areas. Only about half of owners and operators surveyed say they track water use in some way, as shown in **Figure 4**. Those who do track water use

FIGURE 3

**Power consumption and PUE are top sustainability metrics tracked**



mostly monitor at the site level only and do not aggregate and track use across their fleet of data centers (just one in 10 respondents do).

Why don't more track water use? Most say it's because there is no business justification, which suggests a low priority for management — be it cost, risk or environmental considerations. Yet even some of those who do not track say they want to reduce their water consumption (**Figure 5**).

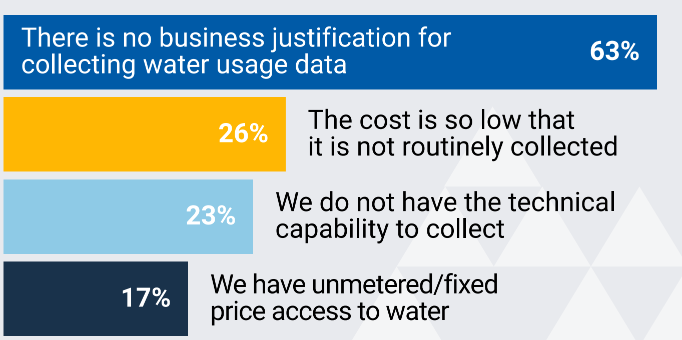
> "External and regulatory pressure may soon begin to drive down water use."

External and regulatory pressure may soon begin to drive down water use. A growing number of municipalities will permit new data center developments only if they are designed for minimal or near-zero direct water consumption. These types of rules will heavily influence facility design and product choices in the future, mandating cooling equipment that uses water sparingly (or not at all).

FIGURE 5

## Many still consider water an inexpensive commodity

Why doesn't your organization track data center water use? Choose all that apply.

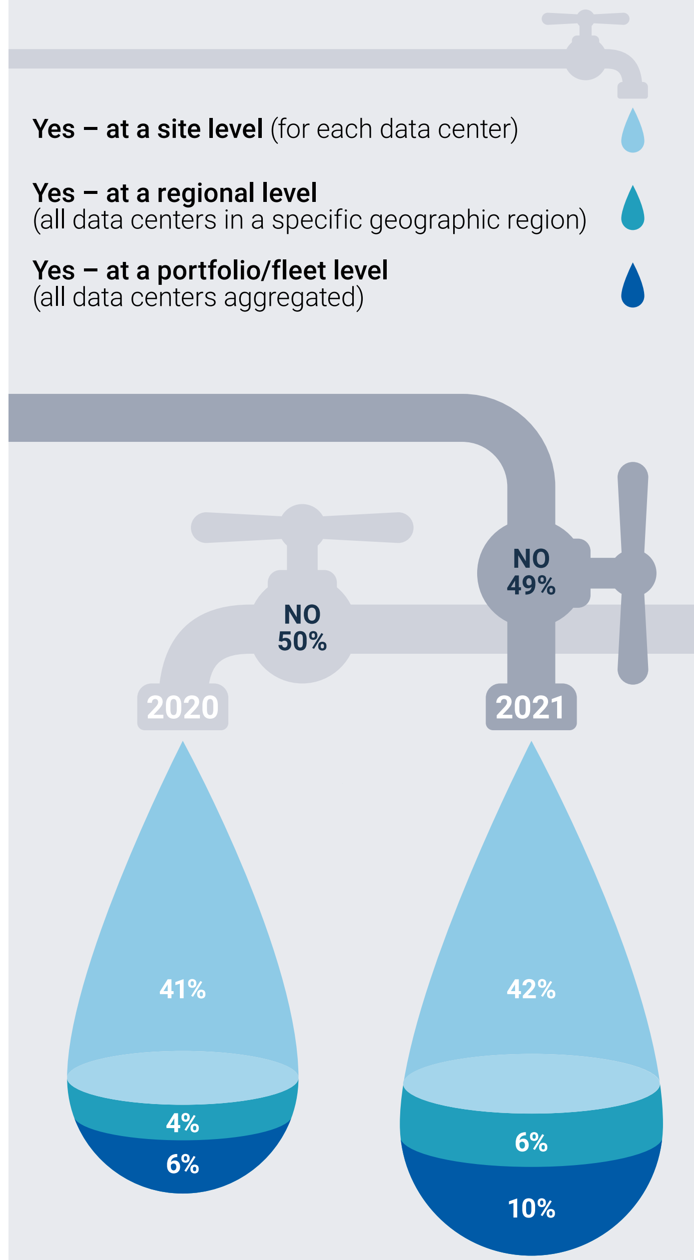



UptimeInstitute® | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2021 (n=268)

FIGURE 4

## Tracking water use across a fleet remains uncommon, but is increasing

Does your organization track water usage data for your IT/data center operations? Choose all that apply.




UptimeInstitute® | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2020 (n=431) and 2021 (n=552)

# Fewer outages, but impacts rise

Among survey respondents, progress toward reducing downtime and managing the impact of outages is mixed. The use of modern systems and processes has improved reliability overall. But many customers and regulators, concerned about the growing impact of big failures, and the dangers of workloads being concentrated in fewer large sites, want to see more evidence of good practices.

## Outage severity holds steady

If data center owners and operators have one overriding KPI (key performance indicator), it is availability: they must keep IT services running (regardless of the workload's venue) and, critically, prevent failures in their own facilities that could lead to an outage. The fact that executive management and governments are increasingly concerned about resiliency shows the goal of achieving continuous availability remains elusive — at least on an industry-wide level.

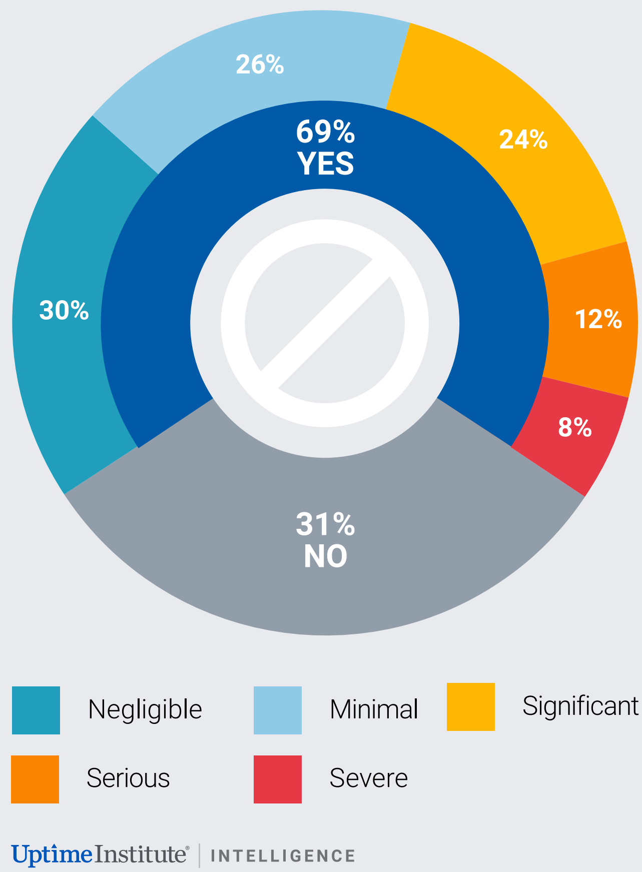
Tracking this KPI, however, is increasingly difficult. Applications are often composed of many disparate, distributed components and services. As a result, IT outages have become less binary — failures are often partial and dependent on user configurations. While a significant data center outage almost always has expensive consequences, the many other (usually smaller) interruptions vary in type and severity. Visibility into outages also depends on the type of outage and who is affected. (For a longer discussion on the changing nature of outages see **Annual outage analysis 2021: The causes and impacts of data center outages**.)

In our annual survey, we take an overview approach, asking about the number of and seriousness of outages over a three-year

FIGURE 6

**Most operators report no or negligible outages in the last 3 years**

On a scale of 1 (negligible) to 5 (severe) how would you classify the most noteworthy outage your organization has had in the past 3 years, either in your own site or a third-party provider site?




UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2021 (n=645)

TABLE 1

**No reduction in outage severity from 2019 to 2021**

| | 2019 | 2020 | 2021 |
|---|---|---|---|
| **Severe** | 8% | 6% | 8% |
| **Serious** | 10% | 14% | 12% |
| **Significant** | 23% | 24% | 24% |
| **Minimal** | 32% | 28% | 26% |
| **Negligible** | 26% | 29% | 30% |

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019 (n=448), 2020 (n=494), 2021 (n=645)

period. In 2021, 69% of owners and operators surveyed had some sort of outage (however serious) in the past three years (see **Figure 6**), a fall from 78% in 2020.

This large but slightly improved number confirms a trend recorded by Uptime Institute through various other sources: The recent improvement may be partially attributed to the impact of COVID-19, which, despite expectations, led to fewer major data center outages in 2020. This was likely due to reduced enterprise data center activity, fewer people on-site, fewer upgrades, and reduced workload/traffic levels in many organizations — coupled with an increase in cloud/public internet-based application use.

Perhaps a more a telling indicator is that when an outage occurs, about a fifth are classified as severe or serious, meaning there were big financial, reputational and other consequences; a further quarter were classified in the next most impactful category — significant, with lesser but still major consequences (**Table 1**). In other words, while just over half of all service interruptions are quickly and quietly dealt with and have few consequences, the remaining half have a significant impact in the form of cost, time and reputation.

> “While just over half of all service interruptions are quickly and quietly dealt with and have few consequences, the remaining half have a significant impact in the form of cost, time and reputation.”

The lessons from this data have not varied much for the past several years. Data centers are becoming more reliable — the amount of IT activity is growing much faster than the recorded rate of outages. There is, nevertheless, still a troubling number of outages and other major failures and interruptions. A proportion of these cause significant disruption and are costly (see **Outage costs are rising**). As the world becomes more dependent on IT services, reliability will receive greater scrutiny and calls for further improvements.

# Management and staff to blame?

Uptime Institute's experience is that data centers with adequate staff, and with staff that follow best practices and procedures, experience fewer failures.

Managers do have the ability to significantly reduce the likelihood of outages through investment and implementing good practices.

Our survey suggests a high and growing percentage of failures could have been prevented by better management and processes (**Figure 7**). The findings, based on the opinions of managers

FIGURE 7

**Most operators think their most recent downtime was preventable**

Would your organization's most recent significant downtime incident have been preventable with better management/ processes or configuration?

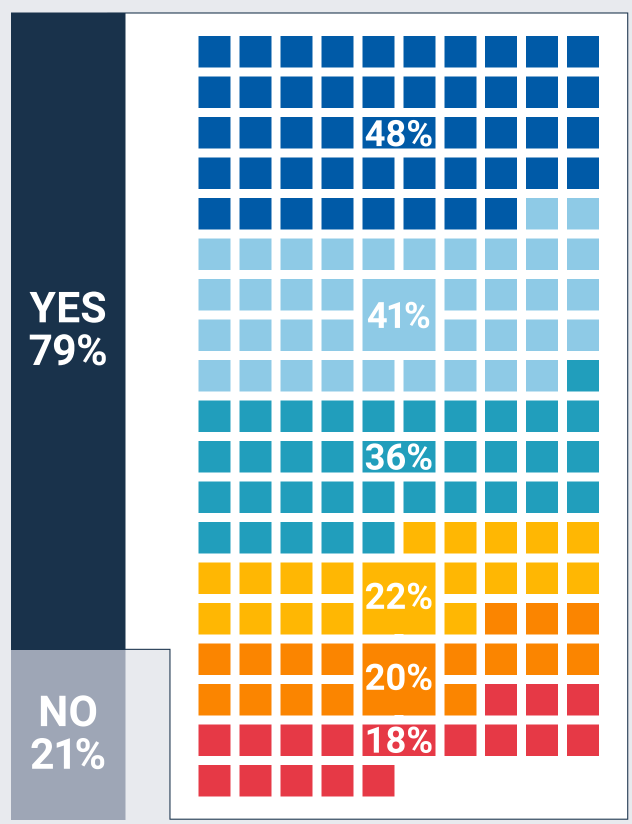

UptimeInstitute | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019-2021

FIGURE 8

**Flawed procedures play a role in most human error-caused outages**

Human error is often implicated in data center outages. In the past three years, has your organization had an outage in which human error played a role? If so, which of the following applies? Choose all that apply.

- Data center staff execution (e.g., failure to follow procedure)
- Incorrect staff processes/procedures
- In-service issue (e.g., inadequate maintenance or equipment adjustment)
- Data center design issues or omissions
- Preventative maintenance frequency issues
- Insufficient staff




UptimeInstitute | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2021 (n=179)

and operators at sites with outages, may reflect two viewpoints: that of managers and executives making an honest self-assessment of the failings; and that of their operational staff, who think executive under-investment may play a role. Either way, both are based on the opinion that the outage was preventable.

To better understand the prevalence and type of human error, we also asked data center owners and operators if they had an outage in the past three years in which human error played a role.

TABLE 2

**Power remains the leading cause of outages**

| | 2020 | 2021 |
|---|---|---|
| **Power** | 37% | 43% |
| **Network** | 17% | 14% |
| **Cooling** | 13% | 14% |
| **Software/IT systems error** | 22% | 14% |
| **Software as a service or hosting** | 3% | 5% |
| **Info security-related** | 2% | 4% |
| **Fire** | 0% | 3% |
| **Third-party cloud provider** | 2% | 2% |
| **Not known** | 1% | 1% |
| **Fire suppression** | 4% | 0% |

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2020 (n=152), 2021 (n=187)

As shown in **Figure 8**, nearly four of five said yes. Two common issues stand out: failure to follow procedures, and having incorrect processes and procedures. These findings point to a clear opportunity. With more investment in management, process and training, outage frequency would almost certainly fall — significantly.

# Outage causes are largely consistent

Corporate secrecy, blame cultures, and sometimes a failure to carry out a proper root-cause analysis can make sharing of insights into failures more difficult — despite the upside of lessons learned.

In 2020 and 2021, we asked about the primary cause of owners' and operators' most recent major outage (outages with negligible/minimal impact were not included). The results show that on-site power problems remain the single biggest cause of significant site outages, by a large margin (**Table 2**).

Three other causes stand out as particularly troubling: cooling failures, software/IT system errors, and network issues. All other outage causes are rare, although

> "With more investment in management, process and training, outage frequency would almost certainly fall — significantly."

the frequency of problems at third-party providers (e.g., software as a service, hosting and cloud providers) is creeping up.

The continued prevalence of power-related outages shows that the industry's ongoing concern with electrical systems and power distribution is justified, as is the high investment. Several decades of modern data center development have not eliminated problems — and the growing use of renewable energy in the electric grid is likely to increase instability, not reduce it.

Our **annual outages report** shows that failures of uninterruptible power systems (UPSs), transfer switches and generators are the three most-cited causes of data center power outages.

# Outage costs are rising

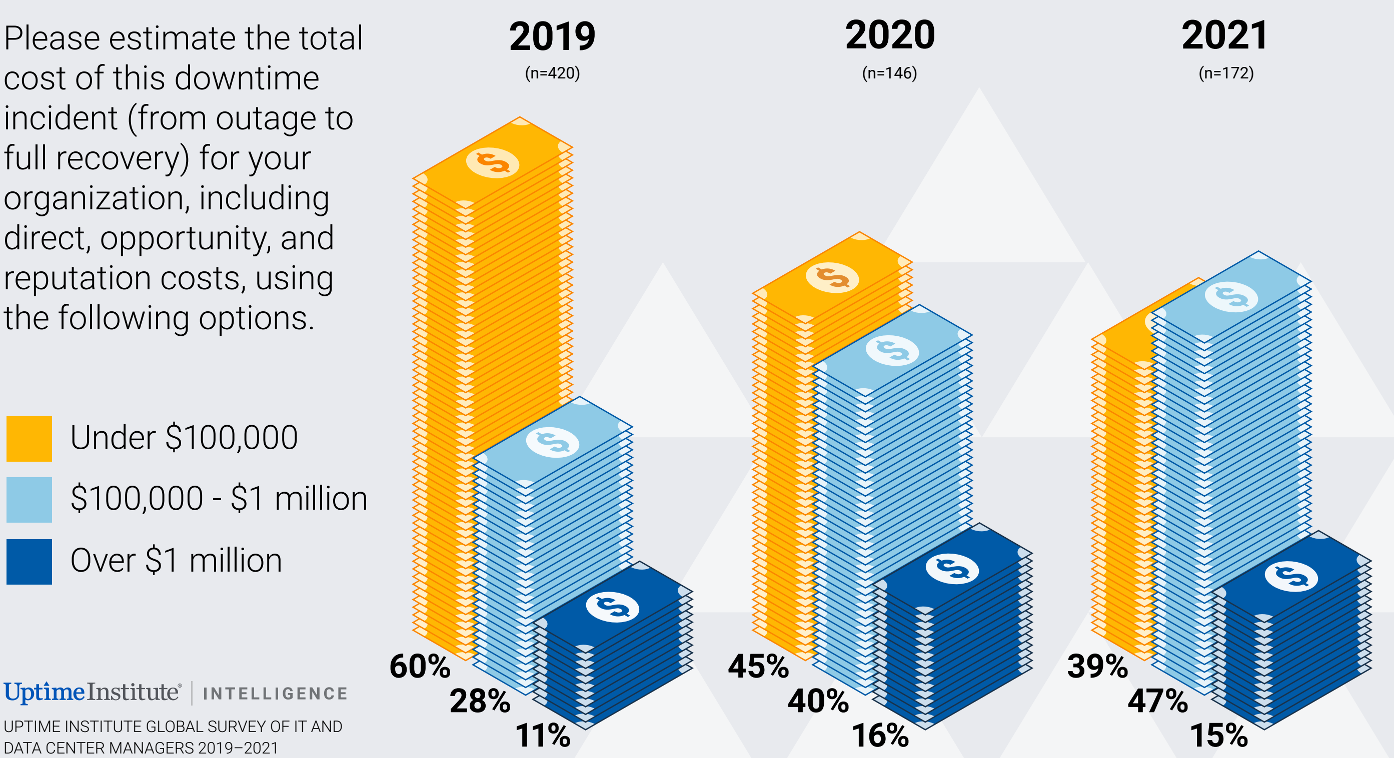
Outages can be expensive and, as stated earlier, about a fifth of all outages are classed as severe or serious by the organizations that suffer them. Uptime Institute's research suggests the cost of outages has been steadily rising in recent years, with over six in 10 major outages in our 2021 survey costing more than $100,000 (**Figure 9**).

We do not calculate an average cost of outages, because the insights gained are rarely useful — businesses and outage impacts vary widely. Each year, a few large outlier outages are so costly they can distort the overall picture. Some result in compensation, fines and lost business, with costs adding up to millions or even tens of millions. In our

FIGURE 9

**The proportion of single major outages costing over $100,000 is increasing**

Please estimate the total cost of this downtime incident (from outage to full recovery) for your organization, including direct, opportunity, and reputation costs, using the following options.




Uptime Institute | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019–2021

2019 survey, there were 10 incidents that led to losses of over $25 million; in 2020, there were three; in 2021, there were six.

The trend toward higher costs when outages occur will likely continue as dependency on digital services increases. Stronger service level agreements (SLAs), expected by some in reflection of this growing reliance, could make outages even more costly, as will more and higher regulatory fines and compensation for customers who experience a service disruption. This, in turn, justifies better analysis of the causes and costs of outages, and continued or increased investment in resiliency.

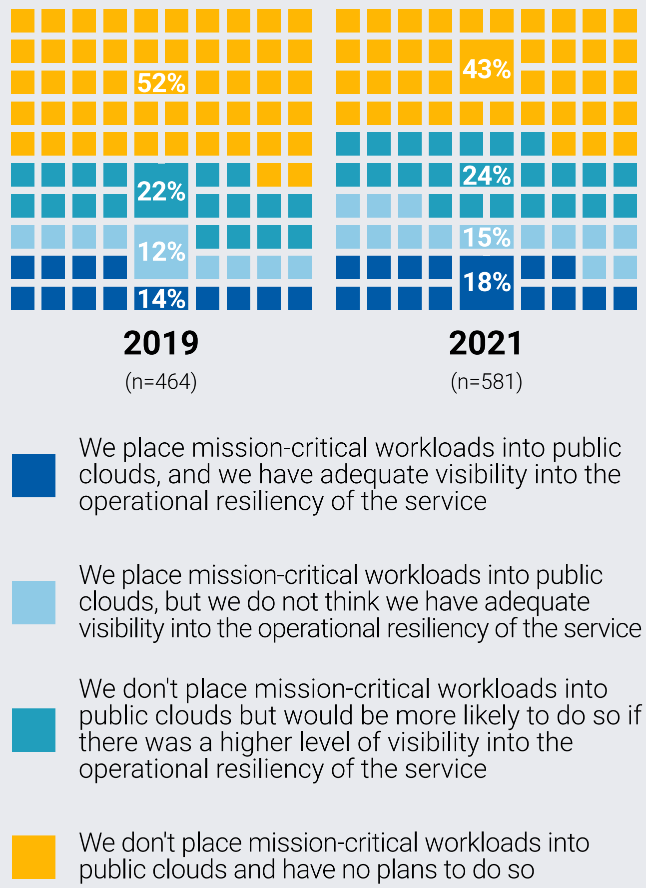
## Cloud providers lack transparency

In May 2021, a technical problem at one relatively little-known internet service company (Fastly) led to a global disruption in internet services, with thousands of sites and services — including many public cloud-based services — affected. It was the latest of several widespread outages in recent years. For many businesses, it demonstrated the lack of adequate visibility they have into their cloud providers' platforms and the relative powerlessness of organizations that are dependent on cloud services to assess their supplier's resiliency or gain a better understanding of potential vulnerabilities. In certain industries — such as financial services — this lack of transparency is a major issue that has attracted regulatory action.

FIGURE 10

**More mission-critical workloads in public clouds, but visibility issues persist**

Does your organization have adequate visibility into the resiliency of public cloud operations (e.g., AWS, Azure, Google Cloud Platform) in terms of architecture, availability record, management processes, and full understanding of options?




UptimeInstitute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019, 2021

Uptime's research (see **Figure 10**) shows owners and operators are increasingly moving mission-critical workloads to a public cloud — and that it also means the level of visibility is satisfactory to them. An additional quarter of our respondents hesitate to move critical workloads to public clouds but would be likely to do so if visibility of the operational resiliency of the service were better.

These numbers pertain to the presence of mission-critical workloads in the cloud for those surveyed. They do not reflect overall cloud capacity — many of the largest workloads, including social media, video and mobile applications, for example, run in cloud data centers, but size of capacity is not reflected in the charted data.

“Over time, more enterprise workloads will run in public clouds, but lack of operational transparency remains a hurdle for many managers with critical applications.”

Public cloud data centers may have adequate risk profiles for mission-critical enterprise workloads, but details about the infrastructure and its risks are often inadequate for regulators or auditors. Cost, performance and security can also play an outsized role in workload venue decisions. Over time, more enterprise workloads will run in public clouds, but lack of operational transparency remains a hurdle for many managers with critical applications.

Cloud companies are beginning to respond to the requirement for better visibility as part of their efforts to win over more large commercial customers. This trend toward improved access and auditability, still in its early days, will likely accelerate in the years to come, as competition and compliance requirements increase. Even among the owners and operators that do place mission-critical workloads in the cloud, nearly half want more transparency.

Visibility aside, most users of public clouds and of colocation data centers think distributing their workloads across these venues has increased their resiliency (see **Figure 11**). This is despite concerns that complexity in networks, orchestration and software has led to more failures in some areas.

About one in 10 think their organization has become less resilient using these architectures. It is notable that confidence in the resiliency of distributed hybrid IT has not risen over a three-year period (from 2018 to 2021), as might have been expected.

FIGURE 11

**Hybrid IT improves resilience for most – but some experience degradation**

Has having workloads spread across private on-premises, cloud and colocation data centers made your overall IT more resilient or less?


Uptime Institute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2018 (n=475), 2021 (n=572)

# Rapid expansion — growing pains

Data center capacity growth is expected to continue. Vendors of data center equipment and engineering/consulting services surveyed in 2021 say most customers' spending is at or above normal levels. Most agree that capital spending on data centers will grow in the next three to five years. However, scaling will be a challenge.

## Sector scaling stresses suppliers

In recent years, the pandemic, extreme weather and political change/instability have caused interruptions to global supply chains. This looks set to continue with tightness of supply for many key components, from chips to power electronics and even major electrical equipment.

Responding to our survey, which ran parallel to this year's owner and operator study, most suppliers expect problems with the supply chain of critical data center products and services in the coming two years. They predict it will affect either capital expenditure projects or general availability of IT equipment, or both — with only one in four expecting no delays or impacts.

There are signs of an unusually strong capacity uptake at major leased data center providers. Major publicly traded US data center real estate investment trusts had backlogs (contracted capacity not yet delivered) at the end of 2020 that were nearly double the dollar value of those at the start of 2020.

> "63% of suppliers say that in the next three to five years, large cloud and internet companies will likely restrict or reduce competition among equipment suppliers."

Much of this growth is powered by large, global technology companies. As cloud and internet giants grow, big (20 MW or larger) sites will become more common, and it is likely their needs will be at the forefront of demand. Hyperscale facilities have an outsized influence on the mechanical, electrical and plumbing supply chain, both in the availability and design of equipment.

Nearly a third of suppliers surveyed expect most of their customers will own data centers 20 MW or larger within five years. The number of these sites, and the volume of sales they represent, means equipment suppliers are often incentivized to prioritize the needs of giant customers. Many in the supply chain have some concerns about this: 63% of suppliers say that in the next three to five years, large cloud and internet companies will likely restrict or reduce competition among equipment suppliers.

Half of suppliers in our study say these high-volume customers often seek projects to be delivered on timelines, budgets, or at a scale that prove challenging for them. Among this group, about 60% were smaller businesses with annual data center-related revenues below $10 million; some (17%) were very large with sector revenues of more than $100 million, including some $1 billion-plus businesses.

The buying power of these internet giants is expected to change the ecosystem dynamics. Hyperscalers have their own technical expectations and price points that suppliers may not be able to meet. This may lead hyperscale operators to innovate more and to integrate products more closely with equipment in their data centers — a potential threat to incumbent makers of power and cooling equipment.

Half of suppliers surveyed say that in the next three to five years, large data center operators will likely take more control of their custom designs and will also create their own supply chains for the manufacture of power and cooling systems, bypassing traditional equipment vendors.

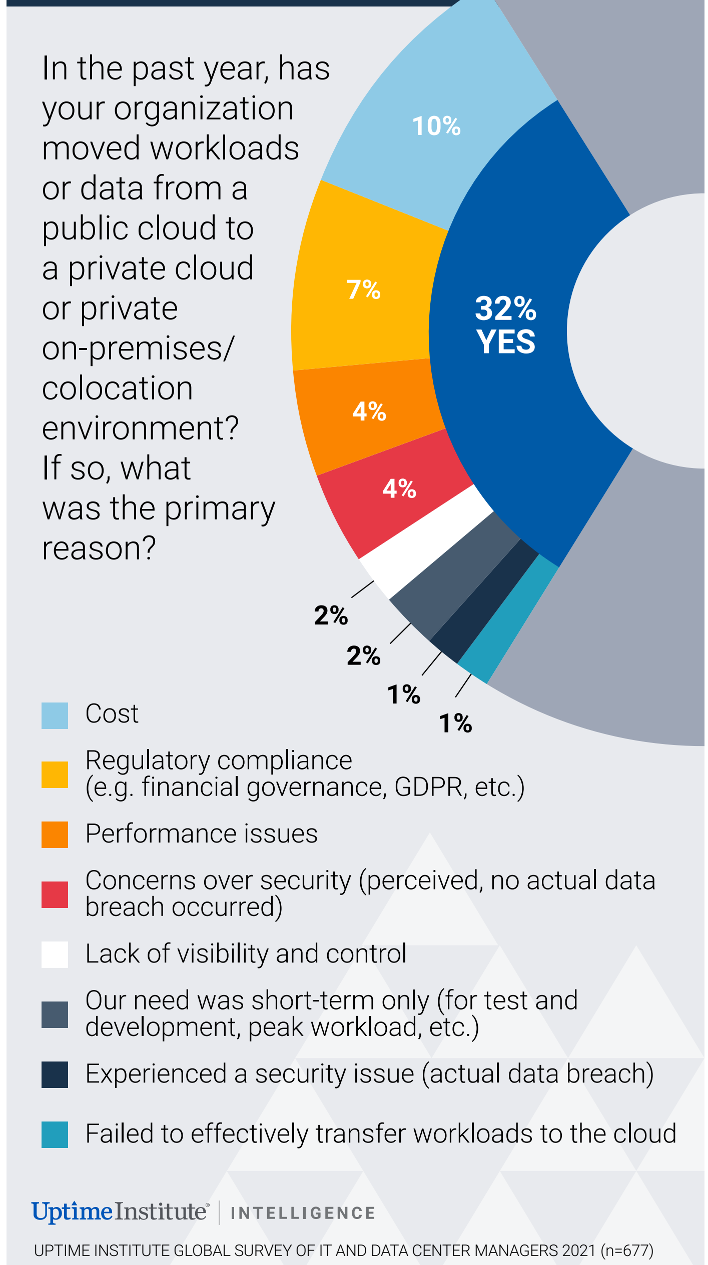
## Public cloud repatriation (or not)

When organizations deploy in a public cloud, their workloads typically remain there, despite speculation about an imminent wave of cloud repatriation. About 70% of owners and operators surveyed have not repatriated any workloads from a public cloud, as shown in **Figure 12**. This finding is in line with our previous years' survey results.

The most common primary reason for repatriation was cost, followed by regulatory compliance (including financial governance or compliance with privacy laws). Other reasons included performance issues and perceived concerns over security. Actual security breaches were responsible for just a tiny portion (1%) of repatriated workloads.

FIGURE 12

**Workload repatriation from public cloud remains low**




UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2021 (n=677)

# The data center edge expands

As operators prepare for future capacity requirements, many expect to build or operate more edge data centers to process data closer to local users and local devices. Today, demand for edge sites of varying sizes and capacity is largely driven by internet of things (IoT) and artificial intelligence (AI) workloads, as well as by mostly speculative deployments of 5G networks.

As shown in **Figure 13**, over half of the owners and operators surveyed say their organization's demand for edge computing will increase in 2021, which is the same as a year ago. In 2021, one in four think this increase will be significant – a higher proportion than reported last year (roughly one in five).

> "Expectations are likely to change as more edge workloads are deployed."

Suppliers of data center equipment and engineering and consulting services are similarly bullish: Over half expect that most of their customers will own edge micro data centers (<150 kW) within five years.

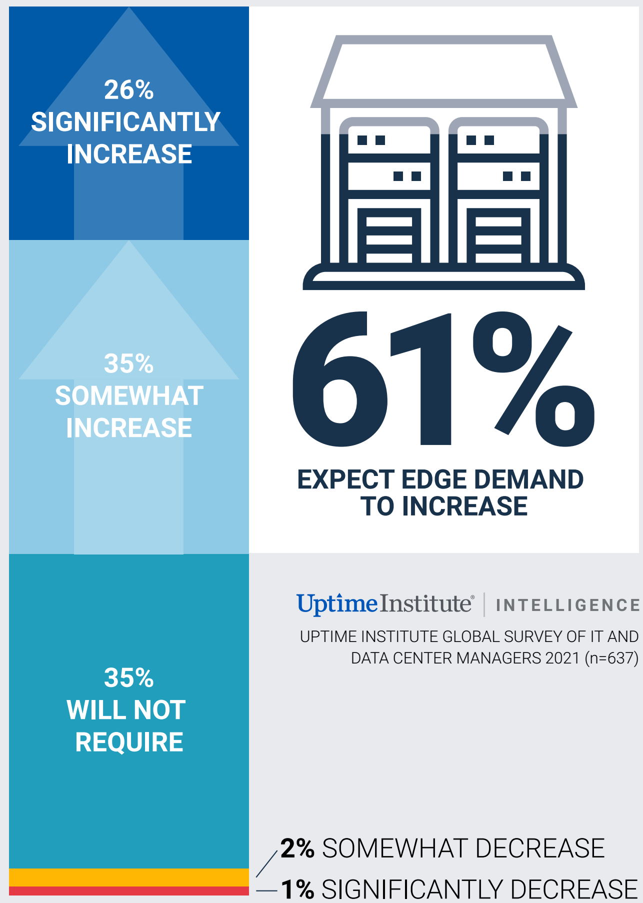
Will organizations prefer to use their own, private edge data centers or shared facilities, such as colocation, public cloud or those owned by a network operator? Many, 40%, of operators expect to use mostly their own private facilities, followed by 18% who expect to use a mix of private and colocation. Few say they will rely mostly on colos (7%) or mostly on public cloud or outsourcing (2% and 3%, respectively).

Yet it is still early days for this next generation of edge buildout, and expectations are likely to change as more edge workloads are deployed. Shared facilities may well play a larger role as demand grows over time. In addition to improved business flexibility and low capital investment, the benefits of shared edge sites will include reduced complexity compared with managing multiple, geographically dispersed sites. (For additional analysis, see our report **Demand and speculation fuel edge buildout**.)

FIGURE 13

**Most expect edge demand to increase, significantly in some cases**

What does your organization expect in terms of edge computing demand for 2021?



# Staffing shortages continue

The struggle to attract and retain staff persists for many data center owners and operators. The number and size of facilities have grown rapidly, creating jobs at a rate recruiters find hard to match. When asked about their single largest challenge, one in five survey respondents pointed to the lack of qualified staff.

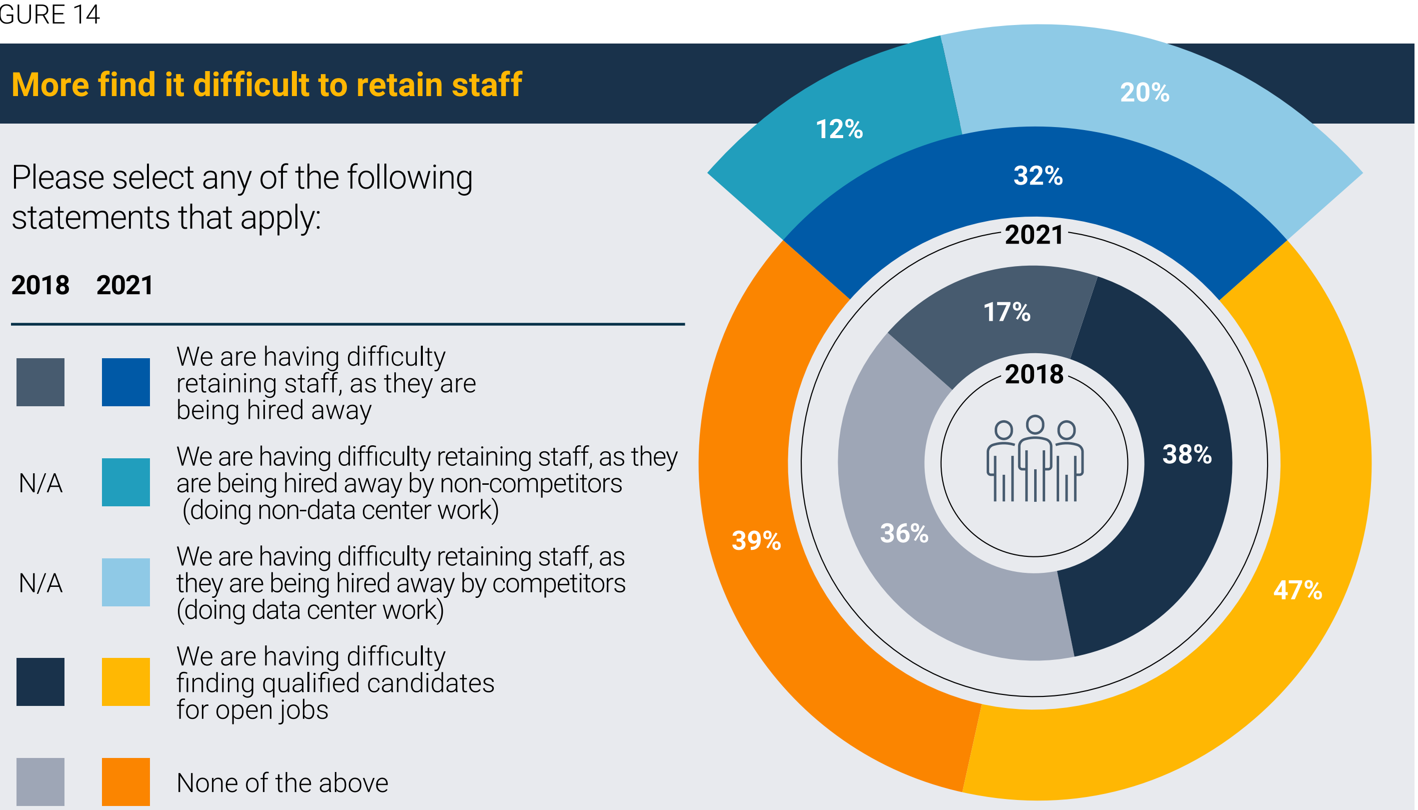
## Help wanted, even with AI

In 2021, Uptime Institute published the first forecast of data center workforce needs, reported by region, by data center type, and by education requirements. We estimate staff requirements will grow globally from about 2.0 million full-time employee equivalents in 2019 to nearly 2.3 million in 2025. (See our report **The people challenge: Global data center staffing forecast 2021-2025**.)

New staff will be needed in all job roles and across all geographic regions. In the mature data center markets of North America and Europe, there is an additional threat of an aging workforce, with many experienced professionals set to retire around the same time — leaving more unfilled jobs, as well as a shortfall of experience. An industry-wide drive to attract more staff, with more diversity, has yet to bring widespread change.

FIGURE 14

**More find it difficult to retain staff**




Uptime Institute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2018 (n=538), 2021 (n=547)

As shown in **Figure 14**, nearly half (47%) of operators report difficulty finding qualified candidates for open jobs, an increase from 38% in 2018. A third face an additional challenge of employees being hired away, a significant increase from 17% a few years ago — and most leave to join competitors.

Given the long-term career opportunities, it is not surprising that few say staff will leave the field altogether, choosing to do non-data center work. Three of four operators agree that most people working in the sector have job security in the long term. The one of four who are skeptical about job security may, among other things, be concerned about staffing requirements slowly changing. For example, AI, with its long-standing promise to augment or replace humans, could potentially decouple data center growth from job growth.

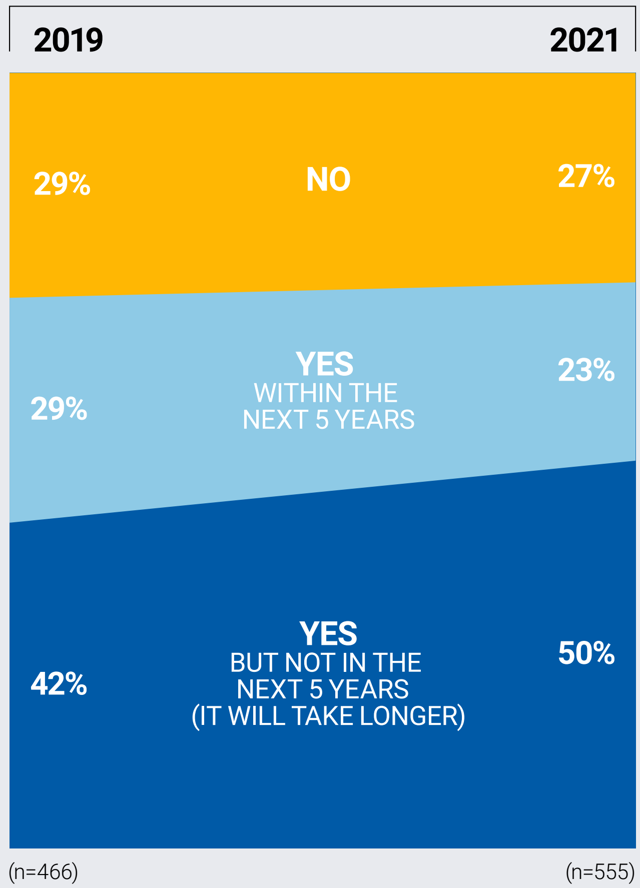
Most, however, do not think AI will replace humans in the short term. While three of four owners and operators say AI will reduce their data center operations staffing levels, most think it will take longer than five years, as shown in **Figure 15**. This is similar to previous years' findings — although today a lower proportion (23% in 2021, down from 29% in 2019 and 34% in 2020) expect AI to impact data center job roles within five years.

> "Three of four operators agree that most people working in the sector have job security in the long term."

FIGURE 15

**Most operators think AI will reduce staffing levels — eventually**

Do you believe artificial intelligence (AI) will reduce your data center operations staffing levels in the next 5 years?




UptimeInstitute® | INTELLIGENCE
UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019 AND 2021

Any replacement of data center staff with AI will require higher trust in the technology. Most operators view AI and its risks — some still unknown — with caution. When asked if they would trust AI to make operational decisions, nearly half of the owners and operators reported that they would, but only if the AI had been adequately trained on historic data.

For more on AI in data centers, see our report **Very smart data centers: How artificial intelligence will power operational decisions**.

# Diversity: Intent outstrips action

Although recruitment needs are expected to rise steadily, the current or expected shortage of staff does not need to be a crisis. Individual employers can take steps to address the issue, and the sector can act collectively to raise the profile of opportunities and to improve recruitment and training. These steps include initiatives to attract a more diverse talent pipeline.

> "Just 5% of operators report that about half of their data center staff are women."

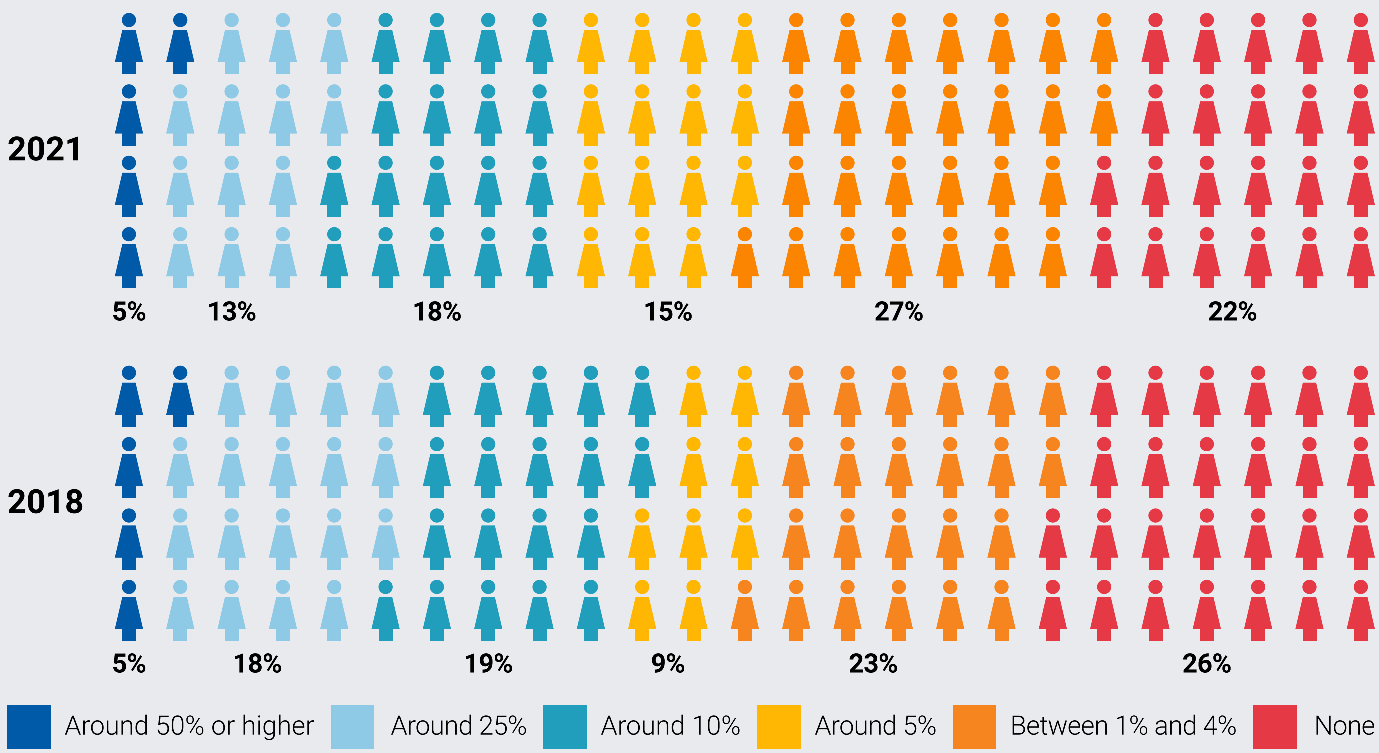
For example, there is a lack of women working in data centers' design, build, and operations staff. Since Uptime began tracking it in 2018, the industry's gender demographics have not changed materially.

As shown in **Figure 16**, just 5% of operators report that about half of their data center staff are women. More than three-quarters report that their data center workforce is around 10% women or less. Almost one in four have no women as part of their design, build and operations staff.

FIGURE 16

**The proportion of women in the data center industry remains low**

What portion of your organization's data center design, build, or operations staff is women?




UptimeInstitute® | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2018 (n=508), 2021 (n=550)

It remains to be seen when this imbalance will change. A large majority, 88%, of owners and operators think the industry will recruit significantly more diverse staff in the next three to five years.

Given slow progress to date, it seems clear that expectations and intentions have yet to translate into widespread action.

The COVID-19 pandemic has exacerbated the gender imbalance, as more workers were tasked with caretaking responsibilities outside of work. Across many industries, women represented a disproportionately large share of COVID-related job loss, as they often bore more responsibility for caregiving.

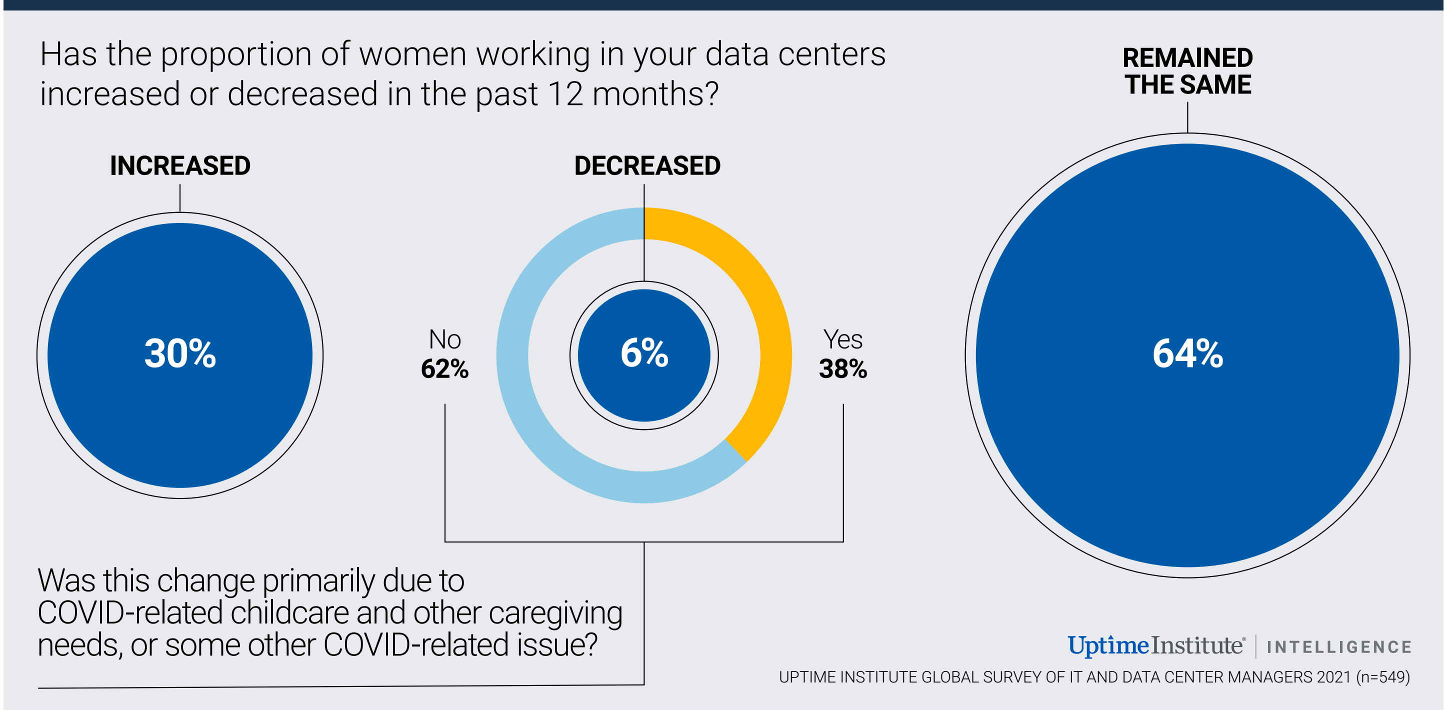
We asked owners and operators if they had seen an increase or decrease in the proportion of women working in their data centers — and if so, was the change attributed to COVID-19, including caregiving needs? While just a small portion (6%) reported a decrease, over a third of that cohort attributed the loss to COVID-related childcare and other caregiving needs, or some other COVID-related issue (**Figure 17**).

> "The COVID-19 pandemic has exacerbated the gender imbalance, as more workers were tasked with caretaking responsibilities outside of work."

With staffing shortages persisting in many markets, and with ongoing capacity growth, data center organizations are advised to actively expand their pool of potential job candidates, do more on-the-job training, and revisit minimum requirements when hiring.

FIGURE 17

**More are hiring women, despite the toll of COVID-related caregiving issues**



# Lithium-ion batteries charge on

UPS systems can be difficult and expensive to maintain. Lead-acid batteries represent the single biggest pain point for many operators. Users take issue with the space they require, their cooling and maintenance requirements, and their relatively short lifespan (four to six years, depending on cell design and operating conditions).

> "Nearly half of the owners and operators surveyed have adopted Li-ion for at least some of their centralized UPS systems."

Batteries using lithium-ion (Li-ion) chemistries are the front-runners for displacing lead-acid technology, thanks to fast-dropping costs and ever improving performance characteristics. As of 2020, Li-ion battery packs were typically more than twice as expensive as lead-acid for the same capacity. Yet they last longer, require less cooling, and are smaller for the same capacity. They are also suitable for novel uses, such as rapid load shedding in response to peaks in electricity pricing throughout the day.

Nearly half of the owners and operators surveyed have adopted Li-ion for at least some of their centralized UPS systems, up from 28% in 2019, as shown in **Figure 18**. Far fewer, about one in six, are adopting distributed UPS systems using Li-ion, compared with one in five in 2019. Those deploying Open Compute Project (OCP) designs and similar open-source IT hardware initiatives in large data centers are often most likely to adopt distributed UPS topologies.

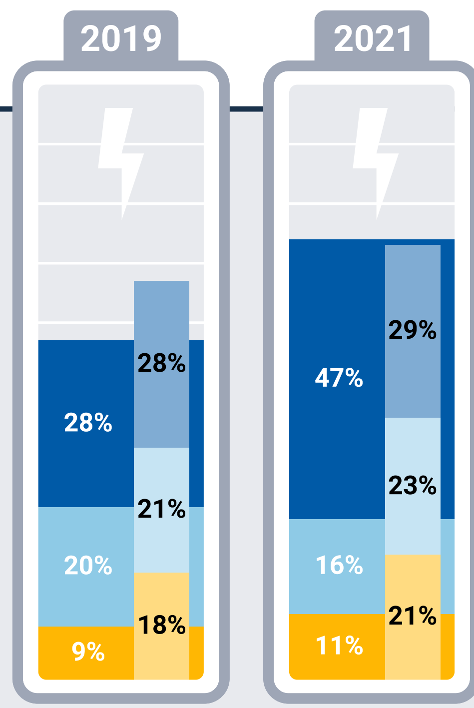
Adoption of fuel cells is also concentrated in larger OCP-shops, as well as some colocation providers. Interest in fuel cells is low but is growing as the industry increasingly explores options to reduce reliance on diesel generators.

FIGURE 18

**Li-ion use grows in centralized (but not distributed) UPS; fuel cell use remains niche**

Has your organization adopted or is it considering the following power technologies in its data centers?

| POWER TECHNOLOGY | ADOPTED | CONSIDERING |
|---|---|---|
| Fuel cells | | |
| UPS – Distributed Li-ion in server racks | | |
| UPS – Centralized Li-ion batteries | | |




Uptime Institute | INTELLIGENCE

UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2019 (n=464), 2021 (n=559)

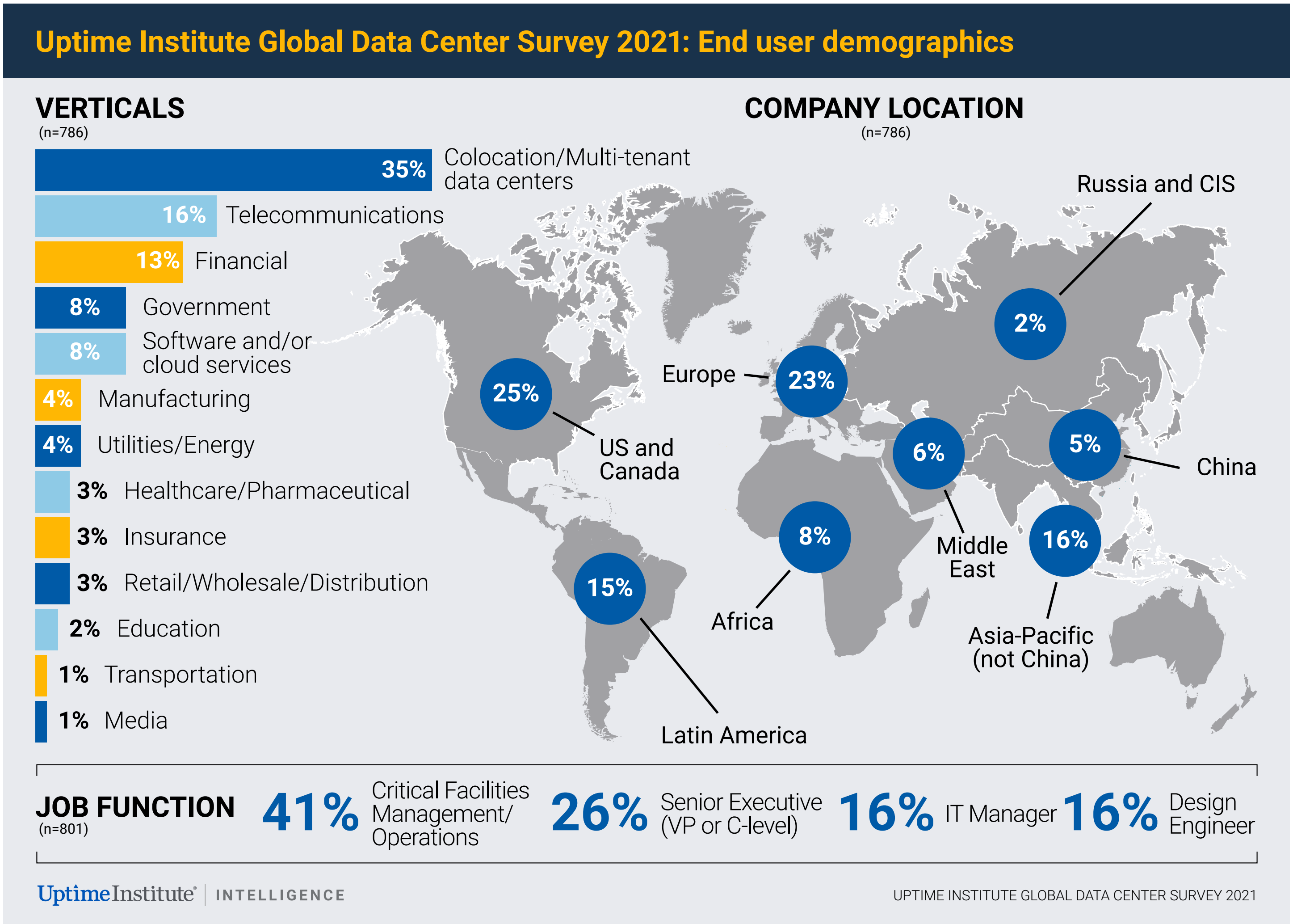
# Appendix: Survey methodology and demographics

Uptime Institute's Global Data Center Survey, now in its eleventh year, is conducted annually online and by email. The 2021 survey was conducted in the first half of the year.

Respondents are separated into two groups: data center owners and operators (Uptime Institute Global Survey of IT and Data Center Managers; 2021 demographics shown in **Figure A1**) and data center suppliers, designers and advisors (Uptime Institute Global Survey of Data Center Designers, Consultants and Vendors; 2021 demographics shown in **Figure A2**).

This report focuses on responses from the owners and operators of data centers, including

FIGURE A1



those responsible for managing infrastructure at the world's largest IT organizations. Job titles include senior executive, IT management, critical facilities management and design engineer.

The participants represent a wide range of industry verticals in multiple countries. Just under half are in North America and Europe. Approximately 40% of respondents work for professional IT/data center service providers — that is, staff with operational or executive responsibilities for a third-party data center, such as those offering colocation, wholesale, software, or cloud computing services.

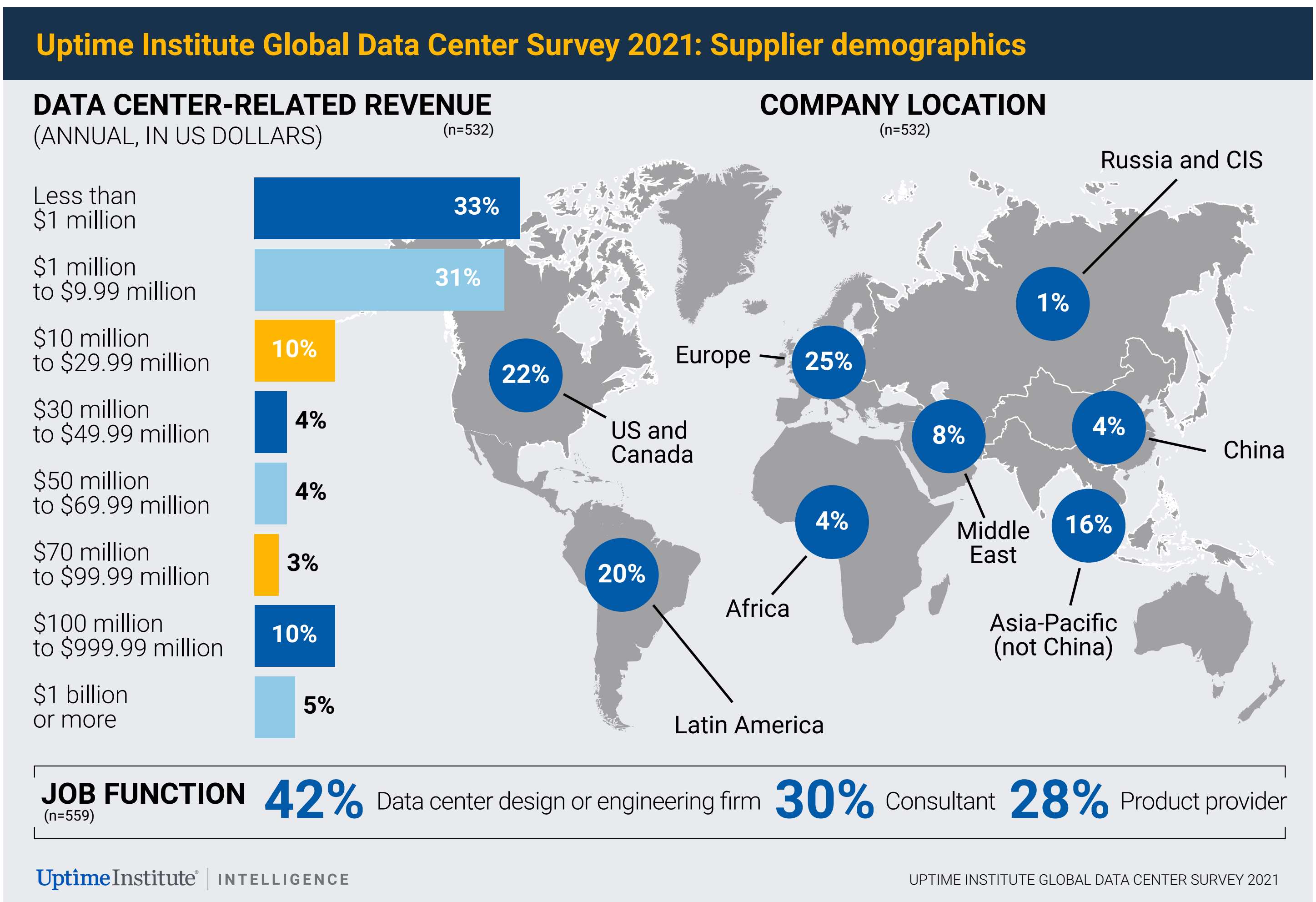
A total of 801 end users registered for the survey and answered at least one question. The number of respondents ("n") varies between individual questions because respondents are not required to answer every question.

Findings of previous surveys are available **here**.

A report focusing on the designer/supplier/advisor survey will be available later in 2021. We will also examine and publish some slices ("cuts") of findings — for example, by region or by industry — in the form of **Uptime Institute Intelligence Notes** in the coming weeks.

If you have questions, comments or seek further insights, please contact **research@uptimeinstitute.com**.

FIGURE A2



Uptime Institute Global Data Center Survey 2021: Supplier demographics

# ABOUT THE AUTHORS

**Daniel Bizo**
Research Director
Uptime Institute

Mr. Bizo has covered the business and technology of enterprise IT and data center infrastructure in various roles for 15 years, including a decade as an industry analyst and advisor.

**Contact:** dbizo@uptimeinstitute.com

**Rhonda Ascierto**
Vice President
of Research
Uptime Institute

A founding member of Uptime Institute Intelligence, Ms. Ascierto has spent two decades at the crossroads of IT and business as an industry leader, keynote speaker and executive adviser focusing on resource efficiency, innovation and disruptive technologies in data centers and critical infrastructure.

**Contact:** rascierto@uptimeinstitute.com

**Andy Lawrence**
Executive Director
of Research
Uptime Institute

A founding member of Uptime Institute Intelligence, Mr. Lawrence has spent three decades analyzing developments in IT, emerging technologies, data centers and infrastructure; and advising companies on their technical and business strategies.

**Contact:** alawrence@uptimeinstitute.com

**Jacqueline Davis**
Research Analyst
Uptime Institute

Ms. Davis covers the global trends and technologies that underpin critical digital infrastructure. Ms. Davis's background includes environmental monitoring and data interpretation in the compliance and health and safety fields.

**Contact:** jdavis@uptimeinstitute.com

## ABOUT UPTIME INSTITUTE

Uptime Institute and the company's Tier Standard is the globally recognized digital infrastructure authority known for the creation and administration of the world's most adopted standards for data center performance and resilience. For over 25 years, Uptime Institute has been providing customers with the assurance that their digital infrastructure can perform at a level consistent with their business needs, across a wide array of operating conditions. With its data center Tier Standard & Certifications, Management & Operations reviews, Digital Infrastructure Resiliency Assessments and other services, along with our accredited educational curriculum for data center professionals, Uptime Institute helps organizations optimize critical IT assets while managing costs, resources, and efficiency. Today, thousands of companies rely on Uptime Institute to enable their digital-centric business success.

---

Uptime Institute - The Global Digital Infrastructure Authority®, a division of The 451 Group, has office locations throughout the world, including North and South America, Europe, UAE, Russia, and Asia. Visit uptimeinstitute.com for more information.

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