Uptime Institute INTELLIGENCE

NEW BUILDS & EQUIPMENT

UI Intelligence report 36

# Best-in-class data center provisioning

**Lead Author**
Dr. Tomas Rahkonen, Consulting Analyst, Uptime Institute

**Contributing Authors**
Andy Lawrence, Executive Director of Research, Uptime Institute
Rhonda Ascierto, Vice President of Research, Uptime Institute

In recent years, there has been a revolution in data center construction. Highly efficient, large-scale data centers are being provisioned faster and more cheaply than ever before, enabling a rapid build-out of IT and cloud capacity. This Uptime Institute report focuses on best practices and best-in-class provisioning: Which technologies and processes are being used? What are the best-achievable costs, efficiencies and provisioning times?

30-45 minutes to read

UII-36 v1.0 published June 8, 2020 last updated June 8, 2020

This Uptime Institute Intelligence report includes:



## 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 uptimeinstitute.com/ui-intelligence.

## KEY FINDINGS

- **Global averages for data center provisioning times have fallen dramatically. The average time to provision a new large data center (20 megawatts or more), following best practices, is just 9-10 months (global average). Some are provisioned faster.**
- **Achieving the best-case provisioning speed relies on the use of prefabricated systems in all areas and on having access to experienced builders doing repeat builds of standardized construction approaches with the support of a strong local supply chain.**
- **Globally, the average capital costs for data center provisioning have fallen dramatically. The money required to build a new large data center (20 megawatts or more), following best practices, has fallen to $7-8 million per megawatt (global average). Some are able to achieve this for less.**
- **Data center provisioning capital expenses will vary among projects, even when using best practices. This is due to differences in local building costs, local access to experienced builders, transport costs, IT rack power density, local compliance costs, client custom specifications, and other regionalized factors.**
- **Most or all new data centers, particularly large ones, are designed for energy efficiency. However, operational infrastructure energy efficiency, as demonstrated by a low power usage effectiveness value, can still be difficult to achieve in practice. Uptime Institute's 2019 annual industry survey shows a global average power usage effectiveness value of 1.67 for all data centers, and 1.51 for data centers 20 megawatts or more.**

# Introduction

Demand for IT capacity continues to grow rapidly across the globe, which has driven the need for more standardized, industrialized approaches to data center construction and component assembly. Large operators and their partners have scrambled to apply new processes and disciplines, expand and re-organize supply chains, deploy prefabricated components, and, where possible, reduce cost overheads, variation and complexity.

These approaches have led to dramatically shortened provisioning times in recent years. They have also significantly reduced the cost of building data centers and enabled the standardized adoption of energy-efficient designs.

Uptime Institute has researched this area by analyzing data points from real-world projects and by investigating novel approaches used by builders, including the use of prefabricated components to reduce data center provisioning times and costs. This report presents the results from our research.

# Methodology

The observations presented are based on data from global data center projects active between 2017 and 2019 and on benchmark data from leading data center builders. The projects represent a mix of colocation/leased and hyperscale data centers. Data is presented in aggregate form only; exact details of any given project are subject to confidentiality agreements and are not disclosed.

The report also draws on results from Uptime Institute's annual global survey of data center owners and operators, particularly with regard to efficiency. Our analysis is additionally informed by insights and data from across Uptime Institute.

## Scope and key assumptions

This report presents a global average colocation or cloud data center provisioning speed and capital expense (capex) analysis for "best-in-class" approaches, which implies:

- The adoption of strict processes and disciplines.
- A level of scale (5-20 megawatt [MW] or greater of total IT capacity).
- Repeatability and a high level of builder experience.
- A proven supply chain.
- The use of prefabricated modular data center components.

We are not, in this report, assessing traditional "bricks-and-mortar-only" builds. These approaches typically have longer provisioning times and, usually, higher costs.

We have analyzed provisioning times and capex for two broad market segments:

- Medium data centers between 5 MW and 19.9 MW in total IT capacity (uninterruptible power supply [UPS] power).
- Large data centers 20 MW or more in IT capacity (UPS power), usually defined as hyperscale data centers.

The scope of our analysis for data center provisioning times includes the construction/installation of the data center building foundation and shell, all facility power, an electrical management system, cooling systems, physical security systems and fire protection equipment — with all systems installed and tested, ready for use.

All of the data centers analyzed were configured for Uptime Institute's Tier III: Concurrently Maintainable Site Infrastructure for a constructed facility, or equivalent levels of power and cooling equipment redundancy.

IT and networking equipment and cabling are not included as part of the analysis.

Our analysis for data center provisioning times assumes all of the following elements:

- Suitable land has already been purchased, and planning approvals for the site have already been obtained.
- Extensive ground works like pilings are not needed.
- The site is not in a seismically active area nor at a high altitude that would require equipment de-rating.
- The construction site already includes the installation/availability of an on-site utility power base station and fiber connectivity.
- All data center design documentation is available and locally approved for compliance.
- A sourcing plan with the specifications needed to buy all required components is available.
- There are no seasonal factors, including inclement weather, or other events, such as cultural, religious or local community events, at the data center site or nearby that may slow construction.

For the purposes of this report, the term “provisioning” is used to include the entire process from when the builders first break ground to the successful completion of final, level 5 integrated systems testing.

# Speed and cost drivers

Time to market has always been an important metric for colocation companies/commercial data center operators. Such companies can only start making money when a new data center is ready and tenants or clients start deploying racks or provisioning IT services. Costs, however, begin a year or two earlier, with any financing/borrowing a major component.

The increased role of cloud computing is making rapid deployment more important. Cloud operators have become the biggest buyers of leased data center capacity, but they will often need multiple megawatts of capacity, sometimes in markets where colocation/leasing companies normally have spare available capacity for only smaller-scale enterprise clients. Securing a large cloud customer may require a new, fast data center build.

Serving large cloud providers as customers also means that data center leasors face increased price pressure: cloud customers may buy more capacity, but they will seek to pay less than smaller-scale clients. Leasing companies typically seek to reduce their data center build costs as one measure to help maintain profitability. Building larger increments of capacity often requires that the financial risk is lowered by building in as short a time as possible.

Large cloud providers themselves face competitive pressures that motivate fast data center construction in new markets. When entering a new market, cloud providers will often benefit by being a first mover, and first to scale; for example, government and other large clients typically require data to be stored locally and, for simplicity, will opt for a single local cloud provider.

Some cloud providers build multiple geographically separated sites in each country or region to increase resiliency. This leads to another challenge for both the cloud provider and their data center capacity company: the limited number of in-country skills — in other words, the availability of qualified people, which can sometimes mean key individuals — to manage their aggressive build-out plans. Here, standardization and prefabrication help the best-in-class cloud providers and data center provisioning companies to manage more projects, or larger projects, with fewer people.

It is particularly challenging to build at low cost in emerging markets that lack a proven supply chain to support the construction of a large multi-megawatt data center. In such cases, the best-in-class providers use remote prefabrication of modules or skids for power and cooling to help reduce project risk and costs.

The benefits of this approach have been illustrated by a large, best-in-class Beijing-based company that builds and operates hyperscale data centers in China, India and Malaysia. By having control over the lifecycle of the project, with on-staff employees who perform planning, design and construction management functions, the company is able to deliver at least 36 MW of capacity within 6 months, on a greenfield site. This is achieved by executing these functions in parallel, more or less, and by using prefabricated modular approaches, including for mechanical and electrical plant equipment; building and data hall construction; and fiber pipeline systems.

Note that in remote areas, while capex can be improved, projects may be impacted negatively by transportation cost and duties.

The economies of scale achieved by specialist data center builders (and operators) today, especially in North America, has severely dampened demand for enterprise data centers (those that serve a single organization; many are 1-2 MW or less in IT capacity). Specialist builders can apply best practices from previous projects and incorporate economies of scale into their model.

The availability of low-cost leasing and cloud computing options means that enterprises that provision new data centers only once every few years need good reasons to do so. It remains to be seen how newer prefabricated data centers and smaller edge data center designs will affect this trend.

# Data center provisioning speed

The time to provision data center capacity is largely determined by four main activities: ground works and foundations, material supply, building shell construction, and installation and commissioning (testing), as shown in Table 1.

Table 1. **Main activities that determine provisioning time**

| | |
|---|---|
| Ground works and foundations | A critical factor for speed is the ability to mobilize many construction workers for a rapid foundation work project. |
| Material supply | Standard lead times can be more than 6 months for equipment such as gensets, medium-voltage switchgear, etc., which is inadequate for fast provisioning. The material supply chain must be configured to handle fast delivery of equipment in high volume. |
| Building shell construction | A locally supported building shell type designed for rapid deployment is key, together with a team of experienced people that has already executed repeat builds. |
| Installation and commissioning | The designs for data center power and cooling must enable massive parallel installation on-site. Power and cooling equipment must arrive to the site as prefabricated, factory-tested skids or modules. Using experienced installation teams that have already completed multiple similar projects will increase speed of installation. |

*Source: Uptime Institute Intelligence, March 2020*

Uptime Institute® | INTELLIGENCE

## Ground works and foundations

Achieving a short provisioning time for ground works and foundations mainly comes down to a local capability to quickly mobilize many construction workers for a brief project at a reasonable price. This can be a management and logistical challenge and requires the availability of considerable local presence and expertise.

The extent of this challenge is illustrated using data from a large European construction company: To complete the ground works and concrete foundations of a 12,000 square-meter (sq. m.; 130,000 square-foot [sq. ft.]), 15 MW data center in 8 weeks, the following was executed:

- One week for earthmoving and truck transports, whereby 8,400 cubic meters (297,000 cubic feet) of earth was transported in 467 truck runs from the site. Assuming a 50-hour work week, this required 10 truck runs per hour.
- One week for graveling and packing, with portacabins (temporary housing barracks) for construction workers installed in parallel.
- Six weeks for concrete foundation works, which required 44 construction workers (assuming a regular 40-hour laborer work week). The supply of 2,740 cubic meters (97,000 cubic feet) of concrete needed was not a bottleneck in Europe — but the availability of a suitable labor force proved challenging.

This example shows how the costs of even the ground works/foundation may vary from country to country. Mobilizing 44 construction workers for a short 6-week project at a reasonable cost will be problematic in several European markets — but may be easier in, for example, some Asian markets. Factors will include local availability and size of construction companies, labor work-time regulations and overtime compensation, and the cost of bringing in additional construction workers from other regions, if necessary.

## Material supply

Achieving a best-in-class data center delivery time requires material supply to be organized as prefabricated systems, reducing the work hours needed for installation on-site. This applies to the building shell, power rooms, cooling systems, rack-ready pod systems and other components.

This may be more difficult for certain components than for others. Gensets and medium voltage switchgear, for example, can be subject to delivery times of 6 months or more. Reaching best-in-class provisioning times depends on working with suppliers to reduce lead times and standardizing these components over multiple projects. To reduce project delivery times, builders have been known to pre-order items with a long lead time based on speculation, being prepared to resell items or negotiate potential supplier penalties if projects are not scheduled or are significantly delayed.

Site location can significantly impact material supply times. For example, using sea transport for the delivery of cooling units or power systems from Europe or North America to Africa or South America will add a month or more to material supply times and, therefore, increase project delivery times. This gives an advantage to suppliers with in-region distribution or manufacturing hubs, and to builder/operators that buy at scale and have developed deep relationships with suppliers.

## Building shell construction

Achieving short provisioning times requires the use of a building shell construction type designed for rapid deployment. Further, it requires the mobilization of enough experienced construction workers at reasonable cost and fast material supply in adequate quantity. This means a locally established construction type should be used for the building shell. In some parts of the world, precast concrete constructions are standard, whereas steel constructions dominate in other regions. The workforce mobilized for building shell construction should ideally have previous experience in deploying the specific building system. Note that this is different from ground works and foundations, where a more generic construction workforce can be used.

## Installation and commissioning

The final main activity to achieving fast provisioning times is the installation and commissioning. This activity is closely related to material supply, in that prefabricated components are the key to reducing installation time on-site. As far as possible, these prefabricated systems, such as power rooms/skids and gensets, should arrive to the site commissioned and load tested from the factory. Installation on-site will then consist of limited site assembly and connection of the prefabricated systems.

The focus of commissioning on-site should be on the final level 5 integrated systems testing. To enable fast parallel installation, most builders have adopted a standard power increment, typically between 1.5 MW and 10 MW, that is repeated in multiples to achieve the total power delivered in a project. Repeating a standard power increment, as opposed to fully customizing the design per project, simplifies material supply and enables highly parallel work on-site. Reaching the fastest-possible provisioning time will further depend on mobilizing an installation team with previous experience from a similar project.

We have observed the following data center provisioning times using a best-in-class approach, as shown in Table 2. For large data centers, 20 MW or more, the provisioning times are about 6 months in the very best cases with ideal conditions. Achieving such fast delivery times relies on the use of experienced teams doing repeat builds, a high level of standardization, the use of prefabricated systems, and the availability of construction workers and installers who can mobilize at scale for short projects at reasonable cost.

Table 2. **Global best practice data center provisioning times***

| Data center size | Best case using best practices | Global average using best practices |
|---|---|---|
| Medium (5-19.9 MW) | 6 months | 8-10 months |
| Large (≥20 MW) | 6 months | 9-10 months |

**MW - megawatt*

***Note.*** *"Provisioning" includes all stages from breaking ground through testing/operations. These figures are global — local factors could extend the times given. All numbers assume adoption of best practices for medium/large-scale provisioning.*

*Source: Uptime Institute Intelligence, March 2020*

Uptime Institute | INTELLIGENCE

For medium-sized data centers, in the 5-19.9 MW range, similar best-case times of under 6 months have been achieved. Medium-sized projects will require fewer resources than large projects and thus will be possible at high provisioning speed in more geographical areas.

Next, looking at global average provisioning times using best-practice approaches for large data centers, 20 MW or more, 3-4 months may need to be added to the best case for medium-sized facilities, to accommodate a larger-scale build. This can be caused by limitations in local supply chains, long transport times to site and a lack of local experienced teams. We suggest a global average provisioning speed of 9-10 months for data centers of 20 MW or more, using a best-in-class approach.

For medium-sized data centers, again there is less strain on local resources, and fast delivery times are possible in more geographical areas. Our observations suggest a global average of 8-10 months provisioning speed for medium-sized, 5-19.9 MW data centers.

Note that it is not uncommon for data center projects to be delayed by external factors, such as delays in incoming power/fiber/water, changes in customer requirements or timeline, or complaints related to permits. Such external disturbances can quickly shift project focus from a need to achieve a fast provisioning time to (for example) optimization of project cash flows.

## Innovations that speed provisioning

Key innovations that enable fast data center provisioning have been made in the areas of prefabrication of different systems.

The data center building shell can be rapidly constructed using prefabricated insulated concrete sections, other self-sustaining construction elements, or modular building systems. Power modules, manufactured as steel modules or skids, are built and tested off-site in factories before arrival at the construction site and can contain transformers, UPS systems, switch gear and generators.

In the same way, cooling systems are commonly delivered as stand-alone modules. Self-supporting rack-ready pod systems can be directly assembled on a concrete floor, as a quick way to install power supply to racks, hot or cold aisle rack containment, and trays to host optical fiber and/or copper network cabling. All of this can be done simultaneously, in parallel, as an efficient way to provision many rows of server racks. (Of course, entire data centers can be prefabricated off-site but only in smaller sizes, which are not the subject of this report.)

A proven approach to saving time and money in construction is to use the same team of supervisors and installers to repeat builds. This can lead to innovations in terms of process improvements and in many small details, such as using the best-possible screw types to save installation time. Repeat builds using the same team can reduce data center provisioning time by more than 10%.

No company can achieve best-in-class provisioning times on its own. All parties in the supply chain must act in a coordinated way and have the flexibility to quickly deliver large quantities of materials and/or quickly mobilize a sufficient workforce.

# Data center capital expenses

As shown in Table 3, several aspects will impact the capex for a given data center project (at given specifications, as discussed above).

Table 3. **Project aspects that affect capex***

| | |
|---|---|
| Local construction costs | Local construction costs and labor costs vary significantly across regions and with local economic and macroeconomic cycles. |
| Access to specialist data center builders | Access to a specialized data center builder with the right experience and a local supply chain is key to low capex, especially if combined with fast delivery. |
| Material transport | Ready access to materials is important. The need for material transport to remote locations will increase capex. |
| Site conditions | Unfavorable site conditions that require piling works or removal of contaminated soil can cause large cost increases. |
| Site location | Compliance to strict local regulations (e.g., regarding noise) can drive extra costs, as will construction in a seismically active area or at high altitude. |
| Power density | A higher power density of IT racks allows a smaller building footprint and lower capex (per megawatt; see **Power density**). |
| Custom requirements | Custom client specifications can drive complexity and higher cost but, in some cases, can also enable a lower-cost cooling solution by relaxing the requirement on strict compliance with ASHRAE guidelines. |

**Capex - capital expenses*

*Source: Uptime Institute Intelligence, March 2020*

Uptime Institute | INTELLIGENCE

## Local construction costs

Local construction costs vary significantly between countries and regions of the world and with economic cycles, as construction companies adapt their prices according to the market dynamics of supply and demand. As discussed in **Innovations that speed provisioning**, the local availability of experienced project teams and installers can significantly impact project delivery time — experienced teams can help reduce risk and ensure cost-efficient project delivery without overruns. In cases where teams need to fly to remote locations, there will be additional costs related to travel and accommodations. Also, material transport costs will be higher.

## Unfavorable site conditions

Unfavorable site conditions that require pilings or extensive ground works (for example, to remove contaminated soil) can drive significant extra project cost. Some countries or regions have strict environmental regulations that can add additional costs, as do potential site-specific requirements such as achieving very low noise levels. In some geographies, the use of certain materials or components (for example, transformers), may limit choices to specific approved brands. Seismically active (earthquake-prone) areas call for stronger, more expensive, designs; at high altitudes, power and cooling equipment may need to be de-rated, resulting in installation of units with higher nominal capacity at a higher cost to reach target IT loads. Local extreme conditions, such as high humidity and heat, common in some Asian cities, can also restrict the use of some cooling systems and drive up costs (see discussion below).

## Power density

An important aspect that drives cost is power density of the IT racks. Higher density can drive down overall costs, as shown in Table 4. While there could be some trade-offs in electrical equipment and cooling, the main impact of higher density is reduced floor space/building shell, resulting in overall lower building costs.

Table 4. **Provisioning capex variation with IT rack power***

| | Low-power racks | High-power racks |
|---|---|---|
| Total IT power = 1 MW<br>IT power per rack<br>Number of racks for 1 MW | 1,000 kW<br>5 kW/rack<br>200 | 1,000 kW<br>10 kW/rack<br>100 |
| White space floor area per rack<br>White space floor area needed for 1 MW IT power | 3 sq. m. (32 sq. ft.)<br>600 sq. m. (6,460 sq. ft.) | 3 sq. m. (32 sq. ft.)<br>300 sq. m. (3,230 sq. ft.) |
| White space building shell cost per unit floor area<br>White space building shell cost per 1 MW | $1,000/sq. m. ($93/sq. ft.)<br>$0.6M/MW | $1,000/sq. m. ($93/sq. ft.)<br>$0.3M/MW |

**Capex - capital expenses; MW - megawatt; kW - kilowatt; sq. m. - square meter(s); sq. ft. - square foot(feet)*

*Source: Uptime Institute Intelligence, March 2020*

Uptime Institute | INTELLIGENCE

## Custom requirements

Custom requirements have long been known to drive up data center build costs, even when using traditional designs and processes. Custom requirements can force suppliers to deviate from proven designs and can cause additional costs in the supply chain.

In some instances, though, custom changes can also drive down costs. For example, most data center cooling systems are designed according to ASHRAE specifications, which set limits on heat and humidity in the data center white space. Some large cloud providers, designing their own server hardware, can deviate from ASHRAE rules, allowing a wider thermal envelope (usually hotter), to keep down cooling system costs. This can be the case where DX (direct expansion mechanically powered) trim cooling (i.e., in additional to free cooling) is needed, but only for shorter periods each year. With a wider range, the cloud provider can opt to forgo extra DX trim cooling to reduce costs, under the assumption that the yearly hours with a data hall temperature exceeding ASHRAE specification will not drive additional server costs — and that the same server replacement cycle (say, 3 years) is maintained. For colocation/leasors, the option to accommodate these types of approaches to cooling when providing capacity may be limited by their cloud provider customers' requirements to operate under strict service level agreements.

## Best-in-class costs

The capex for best-in-class builders of large data centers, 20 MW or more, as an aspirational target with ideal or very good conditions, is $6M per megawatt on average — and some projects can be lower; we are aware of exceptions as low as $3.6M per megawatt. This is a significant improvement compared with a decade ago, when $12M per megawatt was common for data centers typically in the 3-15 MW range.

For a 5-19.9 MW, medium-sized data center, a similar price point can be reached, because large data centers are often built by repeating a standard increment of 1.5 MW to 10 MW. Also ground works, foundations and data center shells are built using similar prices per square meter for both large and medium-sized data centers. Large data centers can reach slightly lower price points by purchasing equipment in larger volumes and by fixed project costs being split over additional megawatts.

The global average capex using best-practice approaches for large, 20 MW or more data centers can vary between projects, as shown in Table 5. Based on our observations, we suggest a global average for large data centers (20 MW or more) in the range of $7-8M per megawatt. For medium-sized data centers of 5-19.9 MW, we propose a global average of $8-9M per megawatt. The reason for the higher numbers in the medium-sized data center range is because medium-size data centers are often built in more remote geographies, which typically leads to higher costs.

Table 5. **Data center capex: Best case and global average using best practices***

| Data center size | Best case using best practices | Global average using best practices |
|---|---|---|
| Large (≥20 MW) | $6M/MW** | $7-8M/MW |
| Medium (5-19.9 MW) | $6M/MW | $8-9M/MW |

**Capex - capital expenses; MW - megawatt; M - million*
*** Some projects can be lower; we are aware of exceptions as low as $3.6M per megawatt.*

**Note.** *All monetary values given in US dollars.*

*Source: Uptime Institute Intelligence, March 2020*

UptimeInstitute | INTELLIGENCE

Note the potential dependency between capex and commissioning speed: targeting a short delivery time increases the sensitivity/exposure to project disturbances. This can result in significant capex overruns, including various project stand-still costs — for example, when the supply chain of materials to the site fails, there is a domino effect of workers being mobilized but unable to work.

# Regional provisioning speed and cost

The provisioning speeds and cost that can be achieved using best-in-class approaches vary globally, with many factors at play (see **Table 3**). Key factors include local availability of specialized hyperscale data-center builders, the establishment of local supply chains capable of handling large volumes, the costs and duration of equipment transport to the site, local labor costs, and local construction costs.

Our research shows that some of the fastest and lowest-cost deployments are in North American data center hubs with highly industrialized supply chains and in some major Asian markets supported by low construction costs and access to inexpensive labor at scale. The supply chain in Asia may also be relatively strong and more responsive, due to large, in-region equipment and component manufacturers in China and elsewhere.

In both North America and Asia, best-in-class provisioning times as short as 6 months have been observed, as well as provisioning capex of $6M per megawatt, on average. There are exceptions; we know of one Beijing-based company that builds and operates hyperscale data centers in China, India and Malaysia that has achieved an average capex of 24.5M RMB per megawatt (about $3.6M per megawatt) due to its customization approach and direct equipment procurement without distributors. Most of its data center campuses are in China.

Outside these areas and in Latin America, Europe, and the Middle East/Africa, the provisioning times and capex observed using best-in-class approaches are generally according to the global averages presented, with provisioning times of 9-10 months and capex of $7-8M per megawatt for 20 MW or more deployments and provisioning times of 8-10 months and capex of $8-9M per megawatt for data centers in the 5-19.9 MW range.

# Data center operating expenses

A data center's operating expenses (opex) largely consist of utility costs, including electrical power and water for cooling and/or humidification, and the costs of maintenance and replacement of equipment, data center staffing (including operations and security) and genset diesel fuel in areas with frequent utility power disturbances.

Data center opex is not the main focus of this report. There are usually many opportunities to reduce opex, but they often involve a trade-off, with decreased staff and some added risk. Automation can also play a role.

Power is a significant component in operating costs, and the largest technical opportunity to reduce opex lies in minimizing the electrical power used for cooling. The choice of cooling system will largely determine this, as will the choice of geographic location of the site. Making use of available cold or dry outside air (or seawater) is key to achieving a low power usage effectiveness (PUE; see **Industry average PUE**) and reducing opex. The choice of cooling system can determine capex, so the target electrical efficiency (PUE) needs to be factored into decisions about provisioning, build partners and operations.

Another area enabling reduced opex is the use of state-of-the art UPS systems that have low electrical losses (2-3%), including when they are operating at partial load — which is common, particularly for new facilities.

# Industry average PUE

Newly provisioned data centers are almost always designed to be as energy efficient as possible. The data center industry's de facto standard for measuring the energy efficiency of a data center facility (that is, excluding the efficiency of the IT equipment) is PUE.

Uptime's 2019 annual industry survey data shows a global average PUE of 1.67, an efficiency of around 60% (the lower the PUE number, the better). As shown in Figure 1, 2019 facility energy efficiency levels deteriorated slightly from an average PUE of 1.58 in 2018. The survey data represents a mix of data center types, ages and sizes.

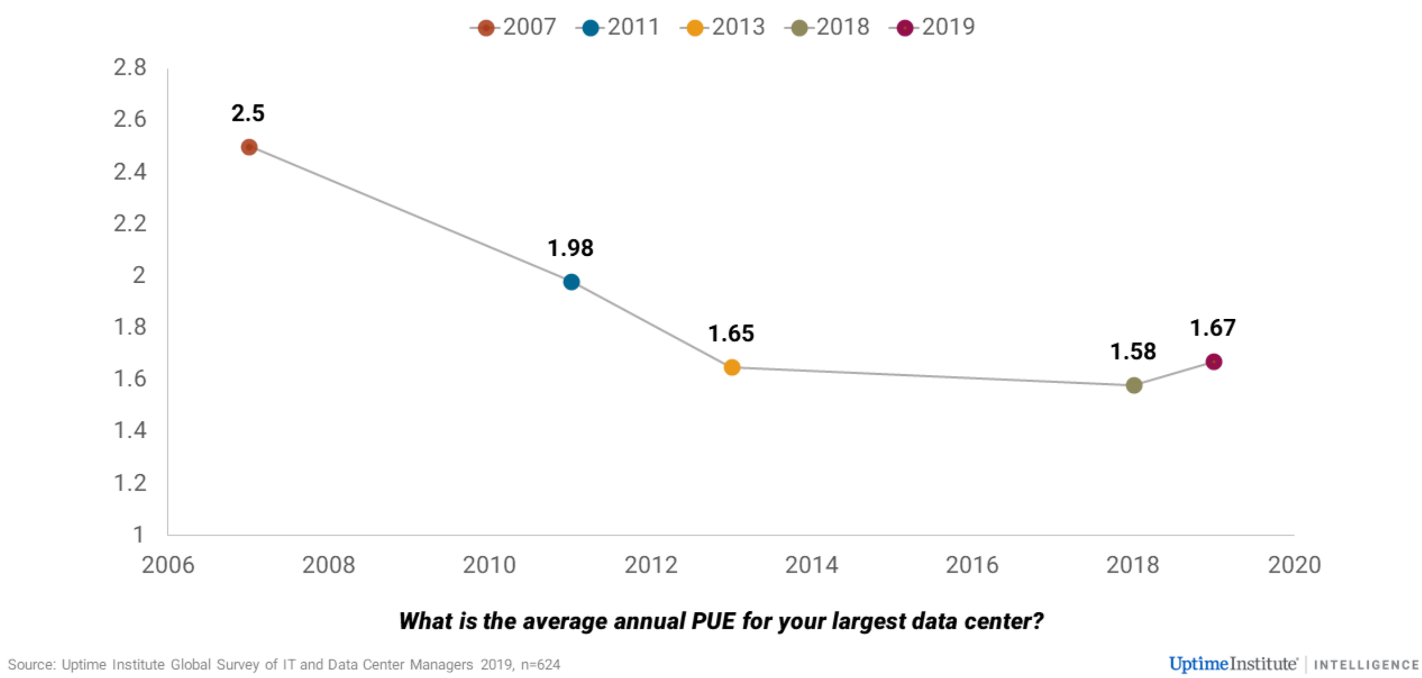



Figure 1. **PUE averages over time: Data center efficiency gains have stalled**

These PUE figures are far higher than many of those cited by, for example, leading hyperscale cloud operators, and demonstrate the difficulty of increasing the efficiency of data centers once they have been designed and provisioned. Furthermore, most data centers are significantly less efficient if they are not fully loaded with servers; in reality, many are not full, and consequently have higher PUE values.

Newer and larger data centers are usually more efficient — but again, it depends on the initial facility design, choice of cooling, geographic location and utilization (how full data halls are, as discussed below). The global PUE average for large data centers (20 MW or more) in Uptime's 2019 survey is lower: 1.51, as shown in Figure 2.

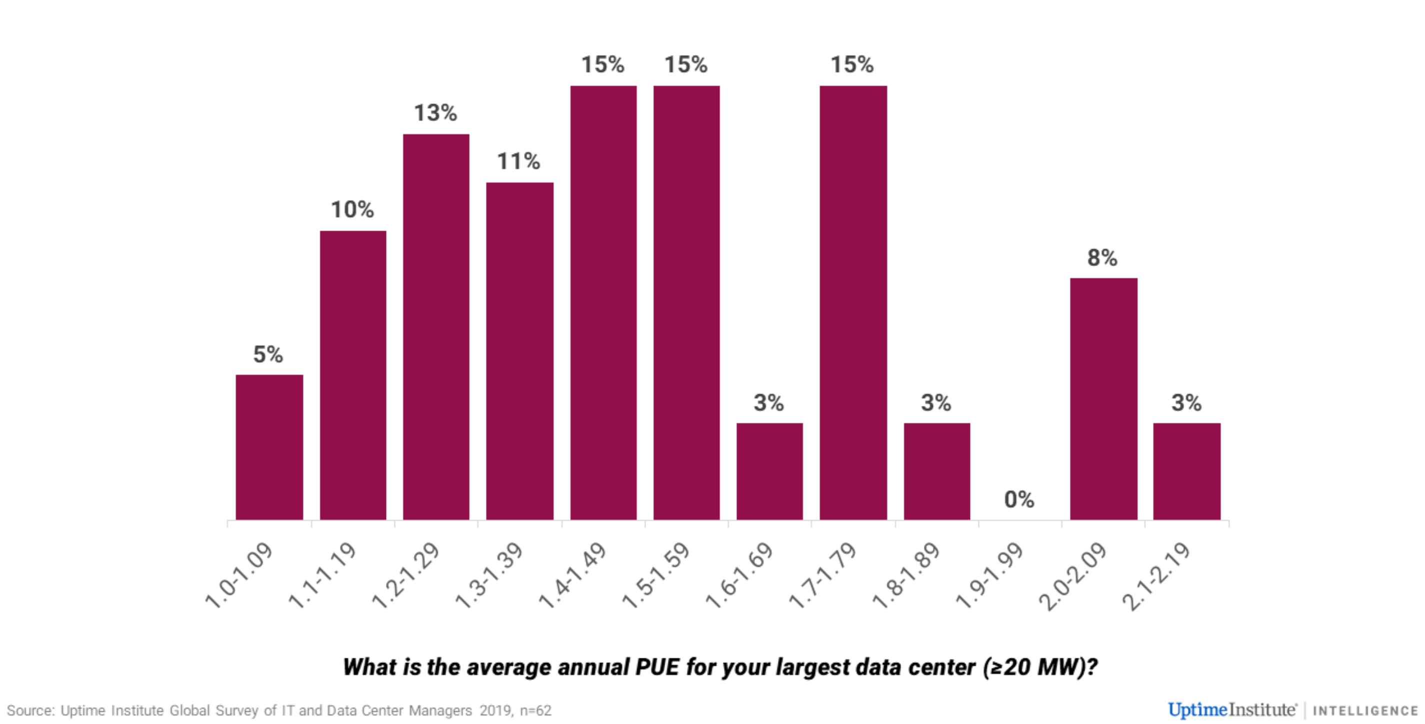



Figure 2. **PUE distribution in large (20 MW or more) data centers globally (average PUE = 1.51)**

The PUE achievable for a specific data center will depend on local climate, including outside air temperature, humidity and risk for airborne pollution. While a PUE below 1.2 may be achievable in ideal areas with a cold and dry climate, a PUE below 1.4 can be considered very good in climates with extended periods of high temperatures and high humidity, which includes large parts of the Asia-Pacific region, sub-Saharan Africa, and Central America. It is more difficult to achieve a low PUE figure at a data center in, for example, Singapore, than it is in Sweden.

It is, however, possible. One Beijing-based company that builds and operates hyperscale data centers in China, India and Malaysia, for example, has achieved an average annualized operating PUE of 1.17 across 12 hyperscale data center campuses; more than half of these data centers are near Beijing, where the climate is relatively cool for the Asia-Pacific region.

Note that the lowest PUE numbers (e.g., 1.0-1.09) can normally only be achieved in optimal geographical locations with access to clean, cold air or seawater throughout the year. It is not currently feasible to reach such low average PUE values at scale for data center suppliers that need to serve clients locally — they are typically required build data centers in locations that are less ideal from a cooling perspective.

# Conclusions

The explosive growth of large cloud data centers in recent years has been accompanied by — and, to a certain extent, made possible by — extraordinary advances in data center engineering, particularly in the field of provisioning. Large-scale data centers (20 MW or more) are now being built from start to finish in as few as 6 months, for as little as $6M per megawatt on average. Some projects today are delivered at even lower costs; we are aware of exceptions as low as $3.6M per megawatt. This is three or four times faster, and half the cost, of a decade ago.

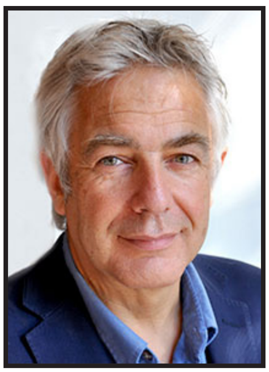
No single technology or practice can be credited — this is achieved through a combination of technologies and practices, with prefabricated components, strong local builders and operators, and highly organized supply chains working together, each playing a key role.

Data center capacity providers (such as colocation and wholesale leasing companies) that are rapidly building large data centers are competitively positioned to attract large-scale cloud customers — and are likely to continue to refine their processes and approaches to achieve shorter provisioning times, at even lower cost, in the future.

Uptime Institute® | INTELLIGENCE

# ABOUT THE AUTHORS

Dr. Tomas Rahkonen is passionate about driving thought leadership for tech companies — analyzing markets; identifying and developing new technologies; managing innovation and setting strategy. He has more than 20 years' experience from data centers, telecom and mobile consumer apps. Dr. Rahkonen holds an M.Sc. in Computer Science focusing on artificial intelligence and a Ph.D. in Electrical Engineering.



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. Contact: alawrence@uptimeinstitute.com

Rhonda Ascierto is Uptime Institute's Vice President of Research. She has spent two decades at the crossroads of IT and business as an analyst, speaker, adviser, and editor covering the technology and competitive forces that shape the global IT industry. Contact: rascierto@uptimeinstitute.com

### ABOUT UPTIME INSTITUTE

Uptime Institute is an advisory organization focused on improving the performance, efficiency and reliability of business critical infrastructure through innovation, collaboration and independent certifications. Uptime Institute serves all stakeholders responsible for IT service availability through industry leading standards, education, peer-to-peer networking, consulting and award programs delivered to enterprise organizations and third-party operators, manufacturers and providers. Uptime Institute is recognized globally for the creation and administration of the Tier Standards and Certifications for Data Center Design, Construction and Operations, along with its Management & Operations (M&O) Stamp of Approval, FORCSS® methodology and Efficient IT Stamp of Approval.

Uptime Institute – The Global Data Center Authority®, a division of The 451 Group, has office locations in the US, Mexico, Costa Rica, Brazil, UK, Spain, UAE, Russia, Taiwan, Singapore and Malaysia. 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