The Latest Data Center Procurement Trends

By Kirk Killian

May 2023
© *Mission Critical Magazine*

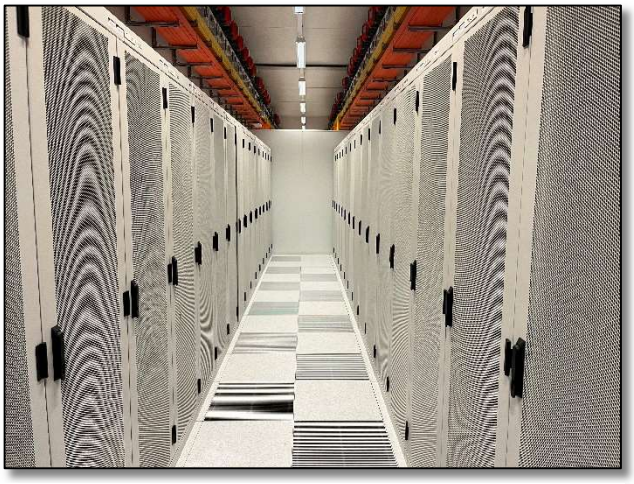
KEYWORDS colocation / data center / construction / supply chain

Companies are increasingly seeking data center facilities that are Dense, Hybrid, Green, and Available

For years, journals have published articles with general observations about "current IT trends," but this article highlights the impactful effects of four IT dynamics on data center facility procurement by enterprise users.

Dense

The quantity of power consumed within each data center cabinet has risen consistently for many years. According to Uptime Institute's annual data center survey, the average power density per cabinet for enterprise data centers rose from 2.4 kW in 2010 to 8.4 kW in 2020. The rate of increase has grown in the last three years, with 69% of enterprises reporting increasing densities and 25% of enterprises deploying some cabinets consuming more than 20 kW in 2022.



Users seek colo data centers offering hot- or cold-aisle containment to cool up to 25 kW/cabinet.

The push for higher densities is driven by multiple factors. Modern bleeding-edge processor chips produced by industry leaders NVIDIA, Intel, and AMD are now consuming 350-400 watts and many more processors can be loaded into multi-U server chassis than just a few years ago.

Artificial intelligence and machine-learning application use is exploding, and those applications tend to be especially power-hungry. Of global CIOs surveyed for a 2022 MIT Technology Review report, 78% indicated that scaling AI use cases to create business value is their top priority in the next three years.

AI tends to work optimally in high-density clusters of homogeneous hardware placement, whether in the public cloud, on-premises, or colocation. For many enterprises, AI has become a mainstream computing component. Router and fiber network efficiency also drives some users to place as much server and storage hardware as physically close as possible to the routers.

Related Data Center Procurement Effects:

- Users are incorporating higher heat rejection capabilities into their data center requirements. Enterprises have evolved from seeking a facility capable of hosting mostly 8-10 kW cabinets to requiring data centers that can effectively cool a cluster of (40) cabinets consuming 25 kW each – or one megawatt in a single cabinet pod.
- Because density growth is forward-looking, enterprises are selecting facilities for their 2023 deployments with densities that can cool intended cabinet densities for five to eight years into the future. For example, an enterprise averaging 12 kW across its cabinets today might select a colocation facility with a designed cooling capacity of 25 kW per cabinet for "future proofing" flexibility.
- Higher densities are forcing enterprises to seek colocation facilities that can easily and affordably support hot- or cold-aisle containment across an entire data hall. Users now prefer data centers with wider cold aisles supporting more high-flow perforated tiles in front of the highest density cabinets and facilities with extra mechanical room space for future increases in cooling plant capacity.
- Users with even higher densities are seeking data centers that can supply liquid refrigerant to the cabinets or immersion cooling options, including some users which may not yet embrace immersion technology. At the 2022 Open Compute Project Summit, Meta (Facebook and Instagram) announced a roadmap for its gradual shift to water-cooled heat rejection. Once liquid cooling becomes more widely deployed, users' apprehensions about "liquid on the floor" will dissipate.

Hybrid

The public cloud is now a widely established delivery model but still not "one size fits all" for some enterprises, especially those in heavily-regulated industries, many of which are choosing a hybrid approach.

Uptime Institute's 2022 Global Data Center Survey indicates that respondents choosing not to place mission-critical workloads in the public cloud have shrunk from 74% in 2019 to 63% in 2022.

A few high-profile public cloud outages have tempered enthusiasm for a "move it all to cloud" approach, instead encouraging a more segmented and strategic planning process.

Application development and cloud migration require effective teamwork and planning.

Most enterprises are evaluating which of their applications and IT workloads will work well in cloud, which could work acceptably after minor recoding, and which do not work well in cloud for technology, cost, or regulatory reasons, and should therefore be kept on-premises or moved to colocation.

Not surprisingly, government and industry regulators are often playing catch-up in approving data security, audit trail, and compliance procedures for public cloud use. A February 2023 U.S. Treasury Department report cited several risks to financial institutions in widespread public cloud adoption, including insufficient transparency to support monitoring, gaps in human capital and tools to securely deploy cloud services, and exposure to potential operational incidents systemically cascading across many financial users.

Related Data Center Procurement Effects:

- Because enterprises are steadily placing an ever-larger portion of overall compute into the public cloud, they are also seeking colocation facilities that enable easy future application migration to cloud. Many users are placing hardware in colocation buildings which also host public cloud installations or are near public cloud data campuses.
- Enterprises executing the hybrid approach are seeking colocation facilities that already have direct connect or software defined network options to support cloud adoption.
- Many enterprise users are planning to move applications back and forth over time between the public cloud and colocation or on-premises. Application placement near public cloud campuses facilitates easier repatriation of some applications from the public cloud if desired in the future to reduce costs or meet technology or regulatory requirements.
- Users are collaboratively working with regulators to expand approvals for new application and data migrations into the public cloud, but that is often an arduous process in an environment where many regulators favor transparency, control, and security over innovation.
- Anticipating continued migration of some applications to the public cloud, enterprises are seeking new colocation contracts with early termination rights, quantity reduction rights, and spend-shift rights to private cloud services.

Green

Many IT users are embracing a worldwide environmental responsibility agenda. There has been good visibility into efficient electricity consumption since the Power Usage Efficiency (PUE) measurement was introduced in 2007 by The Green Grid, a consortium of leading hardware and technology leaders. The data center industry has performed well in decreasing cooling and ancillary power loads that support IT loads, as measured by PUE.

Cloud and colocation data centers are aggressively reducing water usage in new data center designs.

However, water usage in data centers is not as proactively measured in data centers, although hyperscale users Google, Meta, AWS, and Microsoft announced initiatives to reduce their overall water usage over time. The Green Grid proposed a Water Usage Effectiveness metric in 2016, but this measurement hasn't gained the same extent of recognition as PUE. Many colocation companies are also using recycled or "grey water" when available from local water districts to reduce the load on municipal potable water systems.

Especially in Europe and parts of Asia, governments are expected to soon institute new regulations increasing environmental reporting and compliance efforts, such as the European Commission's Energy Efficiency Directive. These initiatives will improve visibility on environmental metrics but may also increase data center development costs, with such cost increases expected to be passed through to users.

Environmentalist campaigns have dramatically raised awareness among consumers and in the corporate boardroom over the resources consumed by data centers, both operationally and in all raw materials used in data center construction.

Related Data Center Procurement Effects:

- Users are very focused on improving efficiency, motivated in part by a surge in electricity rates resulting from higher natural gas prices and the Russian invasion of Ukraine in 2022.
- Users are increasingly seeking data center facilities with 100% renewable electricity sources and far less intensive water consumption.
- To improve electrical efficiency, many users are also much more willing than five years ago to accept an ASHRAE TC 9.9 Allowable temperature range (59.0 – 89.6 °F) in data centers than the previously sought ASHRAE Recommended temperature range (64.4 – 80.6 °F).
- Requests for Proposals for data center capacity are incorporating thorough questions around Environmental, Social, and Governance issues. Enterprises, especially public-facing companies with high consumer brand value, are weighting the providers' ESG responses heavily in data center selection scoring models.
- Data center users are seeking more comprehensive reporting from the public cloud and colocation providers on total environmental impact (not just PUE) and requesting a complete eco-friendly platform including recycling, materials sourcing, and social and governance issues. In addition to traditional focus on the cost reduction elements of operational efficiency, users are focusing on measurable reporting across all ESG goals.

Available

Supply chain disruptions have extended timelines for construction of on-premises expansions, new colocation buildings, and cloud campuses; however, some of the most acute delivery bottlenecks appear to be modestly improving since early 2022. Uptime Institute's 2022 Supply Chain Survey indicated that 52% of respondents indicated improvement in overall supply chain experience in 2022, while only 11% indicated worsening conditions.

Transformer delivery bottlenecks are delaying new substation energizations and data center expansions

Supply chain disruptions have also delayed utility substation and transformer installations, deferring new data center construction projects and enterprise IT project completions.

For many years local municipalities have encouraged new data center developments to increase their tax bases and to appeal to technology businesses needing local data center availability. Since 2019, some municipalities including Dublin, Frankfurt, Amsterdam, Ashburn, and Manassas have resisted or delayed approvals of new data center projects extending the NIMBY ("not in my back yard") concept to a traditionally attractive property development class.

The overall vacancy rate in colocation facilities across the "Big 6" largest US data center markets (Northern Virginia, Dallas, Chicago, Silicon Valley, Phoenix, and New York/New Jersey) has dropped from 8% in 2020 to 3% in early 2023, including a current availability rate of about 1% in Northern Virginia, dramatically decreasing the facility options for enterprises seeking "occupancy-ready" data center space.

Reduced availability of colocation data centers coupled with much higher construction costs due to labor and raw materials inflation have significantly increased colocation occupancy costs, with rental rates jumping as much as 40% since 2020 in some key markets.

Strong data center development in traditionally secondary markets like Atlanta, Portland, Nashville, and Austin is providing significant additional capacity to meet users' needs. Data center development is also booming in secondary international markets like Mumbai, Jakarta, Seoul, and Madrid. A combined 1,800 megawatts of current data center development is underway in these eight secondary markets, a staggering increase from a few years ago.

Supply chain disruptions in hardware availability have accelerated procurement schedules, forcing users to place orders for servers, storage, and routers further in advance of planned installation dates.

Data Center Procurement Effects:

- Faced with fewer options for immediate data center delivery in some locations, users are starting procurement projects sooner and executing new contracts further in advance of required delivery dates. Related, users are improving capacity planning tools and usage forecasting to avoid over- or under-procurement.
- Users are considering a broad spectrum of data center markets because options in the most popular locations have dwindled. Enterprises procuring data center capacity in non-preferred metro areas are also seeking early termination options and spend-shift rights to facilitate relocations later if availability across all major markets returns to traditional levels.

Record-setting leasing of colocation capacity has reduced the options for expanding enterprise users.

- Reacting to reduced data center availability and higher occupancy rates, many users are structuring new colocation contracts with incremental "must take" expansion phases, more closely matching contract capacities to projected needs, deferring some spending until later in the colocation contract term.
- Users are also improving flexibility in their data center delivery strategies. Users are shifting some IT spending into the public cloud or private cloud and bare metal services delivered by colo providers because some hardware categories have long lead times from order placement to customer delivery.

KEYWORDS colocation / data center / construction / supply chain

Kirk Killian

Kirk Killian is President of Partners National Mission Critical Facilities and has advised enterprises in project scoping, site selection, and contract negotiations for data center procurements worldwide since 1999. He serves on the Mission Critical Magazine Technical Advisory Board. Contact him at kkillian@pnmcf.com.