McKinsey & Company

Technology, Media & Telecommunications Practice

Investing in the rising data center economy

Private investors have snapped up data centers in recent years, but plenty of other potential investment opportunities in the sector's value chain may be going unnoticed.

This article is a collaborative effort by Srini Bangalore, Arjita Bhan, Andrea Del Miglio, Pankaj Sachdeva, Vijay Sarma, Raman Sharma, and Bhargs Srivathsan, representing views from McKinsey's Technology, Media & Telecommunications Practice.

January 2023

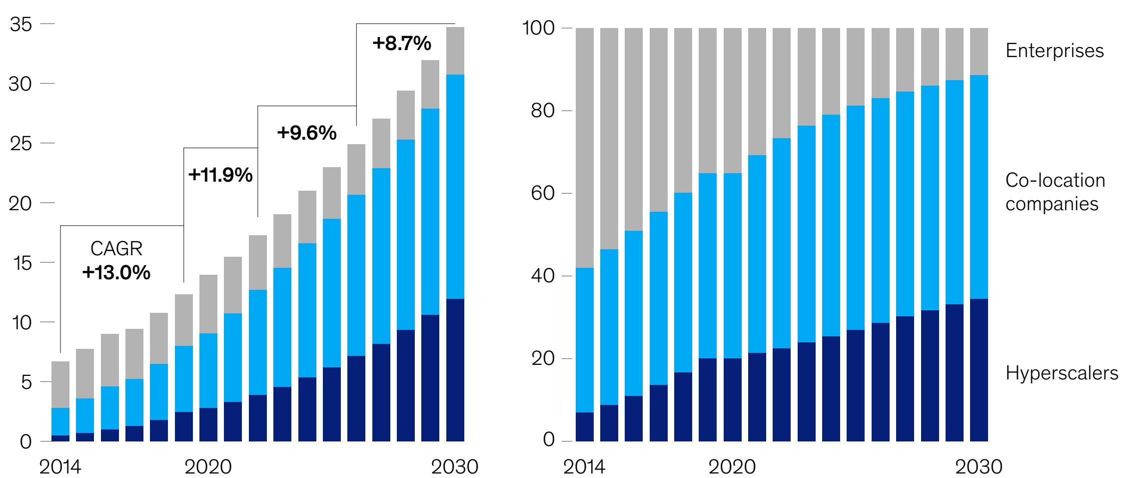
The explosion in demand for data centers has attracted the attention of investors of all types—growth capital, buyout, real estate, and, increasingly, infrastructure investors. In the US market alone, demand—measured by power consumption to reflect the number of servers a data center can house—is expected to reach 35 gigawatts (GW) by 2030, up from 17 GW in 2022, according to McKinsey analysis (Exhibit 1). The United States accounts for roughly 40 percent of the global market.

Data centers are typically owned and operated either by big companies (such as cloud vendors, banks, or telcos) for their own purposes or by co-location companies. The latter lease out the space and typically provide network capacity and power, as well as the cooling equipment that keeps down server temperatures. Tenants bring their own IT equipment. Data centers have attracted the interest of investors, often because of the steady, utility-like cash flows and risk-adjusted yields.[1] In 2021, there were 209 data center deals, with an aggregate value of more than $48 billion, up some 40 percent from 2020, when the deals were worth $34 billion. In the first half of 2022, there were 87 deals, with an aggregate value of $24 billion. From 2015 to 2018, private equity buyers accounted for 42 percent of the deal value. Their share increased to 65 percent from 2019 to 2021 and to more than 90 percent in the first half of 2022.[2]

Several factors could limit this trend, however. First, higher interest rates raise the cost of funding

Exhibit 1

US data center demand is forecast to grow by some 10 percent a year until 2030.



[1]Demand is measured by power consumption to reflect the number of servers a data center can house. Demand includes megawatts for storage, servers, and networks.

McKinsey & Company

[1] Relatively few data center operators remain publicly owned. According to NAREIT, as of October 31, 2022, the dividend yield stands at 3.14 percent, and the 2021 total return at 25.47 percent.

[2] "Private equity is driving a boom in data center M&A deals," Synergy Research Group, June 22, 2022.

deals, though this is not limited to only data centers. More specifically, multiples and forward price-to-earnings ratios for co-location companies have been relatively high, boosted by competition to snag the diminishing number of potential acquisition targets for private investors and for co-location companies keen to expand.

In addition, the operating margins of co-location companies are under pressure from prominent cloud vendors, including Amazon Web Services and Google Cloud. These have long been major customers for co-location companies but also continue to own and build their own world-class centers.[3] As a result, they are signing shorter-term contracts and are often in a strong position to demand favorable leasing terms. Moreover, if this strategy means that co-location companies will no longer be long-term owners and operators of data centers but more akin to developers, the market will open up for other competitors. Real-estate companies, for example, could increasingly build or lease out space for their tenants to equip and use.

But this possibility does not mean that the sector lacks value-creating investment opportunities. As Exhibit 1 shows, co-location companies will continue to have a strong position in the market. Hyperscalers still need them to meet fast-growing demand, and smaller enterprises depend on their specialist services. That helps explain why three investment groups are in the running to acquire Global Switch.[4] However, many investors may not have considered the upstream opportunities in a complex value chain. Here we explain where investors might start to look.

Data centers have four main components: the facility itself; the industrial equipment, including the mechanical, electrical, and plumbing (MEP) gear; the IT hardware; and the software. There are also opportunities in data center operations—the management of facilities and IT services, such as hosting and infrastructure as a service (IaaS). Other services, such as power and connectivity, present opportunities, too. Potential investments lie across this value chain, but in four areas demand is particularly high (and not matched by supply) or innovation is especially likely to create value.

Sustainable (or green) energy

Data centers are big energy consumers—a hyperscaler's data center can use as much power as 80,000 households do. Pressure to make data centers sustainable is therefore high, and some regulators and governments are imposing sustainability standards on newly built data centers.[5] This development gives investors opportunities to help data centers secure carbon-free energy supplies.

In fact, the use of renewable energy is a critical component of the hyperscalers' strategies. Thanks to carbon offsets, Apple, Google, and Meta, for example, were all carbon neutral by 2020. They and other hyperscalers have committed themselves to using only carbon-free energy by 2030. Co-location companies are also under pressure, not least from some of their customers, to meet sustainability goals. A former executive at a hyperscaler told us that the sustainability record of co-location companies was a significant consideration in deciding which ones to work with.

To reach carbon-free energy goals, data center owners are signing power purchase agreements (PPAs) with suppliers of renewable energy. Meanwhile, hyperscalers are starting to fund the building of renewable-energy plants in the face of soaring prices caused by supply shortages.[6] In the United Kingdom, for example, Amazon has supported Scottish Power's wind farm and is purchasing its entire 50-megawatt (MW) output.[7]

3 McKinsey analysis.
4 Dan Swinhoe, "Stonepeak and Gaw drop out of Global Switch bidding, three companies still in running," DCD, November 22, 2022.
5 Singapore and the Netherlands are examples.
6 Dan Swinhoe, "Power purchase agreement prices up nearly 50 percent in Europe over last year," DCD, July 14, 2022.
7 "Amazon's first Scottish wind farm project comes online," Amazon News, October 28, 2021.

Yet such moves will not suffice if using only renewable energy is the goal. The first problem is intermittency. Solar power is generated only in the daytime, and wind power depends on the weather, so fossil-fuel supplies often supplement power from renewable PPAs. One emerging solution is "24/7" PPAs, which commit themselves to matching each hour of electricity consumption with a combination of carbon-free supplies and, quite important, stored renewable energy. These contracts come at higher prices, however, not least because current storage technologies are expensive. The levelized cost of electricity from a system that combines wind, solar, and lithium-ion (Li-ion) battery storage typically exceeds $200 per megawatt-hour. Long-duration storage solutions that deploy hydrogen and green-ammonia energy could push that below $100, but these technologies remain at a relatively early stage of development.[8] Backup power is another issue, since many data centers still use diesel generators during power outages. Li-ion batteries are the most developed carbon-free backup solution but can prove expensive over long periods.

All of this presents opportunities for investors. Not all co-location providers have the scale to procure renewable power either through PPAs or investments in power plants. Investors with smaller data centers could aggregate their purchasing power to optimize energy procurement and storage. Some might also consider investing in renewable-energy plants that could supply consortiums of smaller players.

Other potential technology and R&D investments include more stable renewable-energy technology and sources (such as geothermal and wave energy) to facilitate zero-carbon backup power and storage solutions. Investors might even offer to pilot new solutions in data centers they own.

Cooling and energy consumption

Climate change and unpredictable weather events, pressure to decarbonize data centers, and increasingly powerful computers offer investment opportunities in cooling and energy efficiency technologies for data centers.

Data center equipment, often consisting of thousands of servers, must be cooled to work efficiently. Indeed, the capacity of a data center is dictated by how well it cools the servers—the more closely they can be stacked, the more productive the square footage. Efficient cooling is therefore a crucial driver of a data center's profitability. Cooling accounts for some 40 percent of a data center's energy consumption.[9] The cost of downtime from overheating can be high.

Cooling technology has improved rapidly over the past decade. Most large data centers have replaced old air-conditioning-like systems that keep entire rooms cool with in-row or rotodynamic heater-based cooling designs: heat emitted from servers is drawn away by fans and then cooled with water or a refrigerant. Yet even better performance is required because today's more advanced systems can struggle to control the temperatures associated with global warming. Google and Oracle, for instance, both faced downtime during a heatwave in Europe this past summer.

Higher computing power and innovative chip designs are also putting more demand on cooling systems by raising the power density: the energy consumption of the equipment stored in racks. Their average power densities have more than doubled over the past six to seven years and continue to rise.[10] The density of a single rack can be as high as 20 to 30 kilowatts in specific high-performance environments. Space constraints—especially for smaller edge-computing data centers in urban areas—also raise demand for systems with higher power densities.

---

8 "Decarbonizing the grid with 24/7 clean power purchase agreements," McKinsey, May 11, 2022.
9 McKinsey analysis.
10 Andy Lawrence, "Density is rising," Uptime Institute, December 7, 2020.

Data centers need to use energy more efficiently as well. Power usage effectiveness (PUE)—the amount of power the computing equipment in a data center uses relative to its total energy consumption—fell considerably from 2007 levels, but progress has flattened over the past decade (Exhibit 2).[11] Even the hyperscalers that deploy advanced cooling designs and technologies struggle to improve significantly: the reported PUE of Google's data centers has fallen only incrementally during the past seven to eight years, for example.[12]

To address these challenges, companies are developing and deploying several technologies, including immersion cooling, artificial intelligence and machine learning, and the use of waste heat (see sidebar, "Cooling and energy efficiency technologies"). Yet extensive investment is still required to make progress in R&D and in deployment at scale. In some instances, collaboration across the value chain will probably be needed: for example, full-immersion cooling, which submerges servers in a coolant, could require changes to chip packaging; the form factors of servers; building layouts, materials, and design; and MEP systems. An investor with data center assets might vertically integrate components of the value chain by, for example, investing in a cooling-technology company to prove a concept and test the solution.

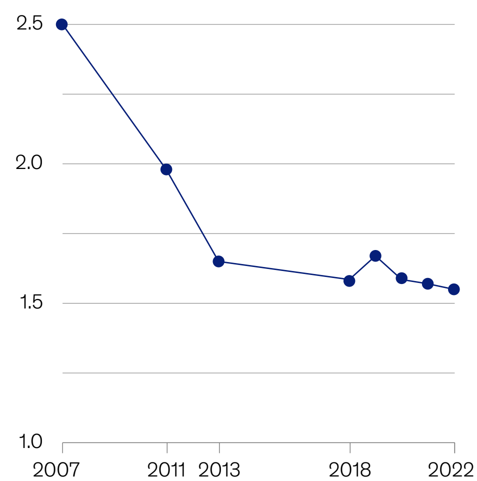
Constructing prefab and modular data centers

Rising demand for new data centers is also creating potential investment opportunities in the fragmented prefabrication and modular (PFM) sector. According to the Synergy Research Group, in 2022 hyperscalers alone allocated some

Exhibit 2

Gains in power usage efficiency have stalled during the past decade.

Power usage effectiveness (PUE)[1]



[1]A measure that shows the amount of power used by the computing equipment in a data center relative to its total energy consumption. The closer PUE is to 1, the more efficient a data center's power usage is.
Source: Uptime Institute Intelligence

McKinsey & Company

[11] *Uptime Institute Global Data Center Survey Results 2022*, Uptime Institute, September 14, 2022.
[12] "Efficiency," Google Data Centers.

$9 billion to build more capacity—a sum expected to grow by more than 4 percent a year until 2030 (Exhibit 3). Such plans face headwinds, however: the labor market is tight, commodity prices volatile, inflation high, and supply chains constrained, so global capital costs for construction projects have risen by at least 6 percent since 2020.[13]

As a result, hyperscalers have turned increasingly to PFM solutions that enable parts of the construction process to take place off-site. Done well, PFM not only cuts construction times but also reduces costs and improves safety, quality, and sustainability, since more work takes place in controlled manufacturing settings. One company recently cut the cost of building a 45-MW facility in Europe by 20 percent and slashed construction time to 11 months (from 17) by using prefabricated components for the building, as well as modularized components for the electrical and cooling systems.

There are four types of prefabricated or modularized solutions:

— the prefabrication of structural and architectural components, such as concrete beams, walls, slabs, facades, and precast underground culverts

[13] McKinsey Capital Analytics.

Cooling and energy efficiency technologies

To improve the efficiency of cooling systems and reduce energy consumption in data centers, several technologies are now under development.

Immersion cooling. New approaches include full-immersion and direct-to-chip/cold-plate cooling. In the former, IT equipment (such as a server) is immersed entirely in a nonconductive and nonflammable di-electric liquid that acts as a coolant and dissipates heat generated by the equipment. In the latter (and more targeted) approach, a metal plate (or heat sink) is used for high-thermal-emission components (such as chips) in the servers. This approach transfers the heat and then cools it using a liquid coolant. By maintaining consistent, uniform temperatures, full-immersion cooling can cope with higher power densities (upward of 100 kilowatts) and raise the average performance of central processing units by as much as 40 percent.[1] Riot Platforms uses immersion technology at a Bitcoin-mining farm in its Whinstone facility in Texas,[2] and hyperscalers are developing and testing it. But the widespread use of these technologies will require collaboration between IT manufacturers and infrastructure owners.

Artificial intelligence and machine learning. Hyperscalers such as Google have used artificial intelligence/machine learning (AI/ML) algorithms to focus cooling where it is most needed, depending on factors such as workload intensity and changing power loads across racks. Early adopters have reported 20 to 30 percent reductions in power usage effectiveness. AI/ML applications have also balanced the load on uninterruptible-power-supply units by changing power routes to servers throughout the day to optimize cooling and save energy.

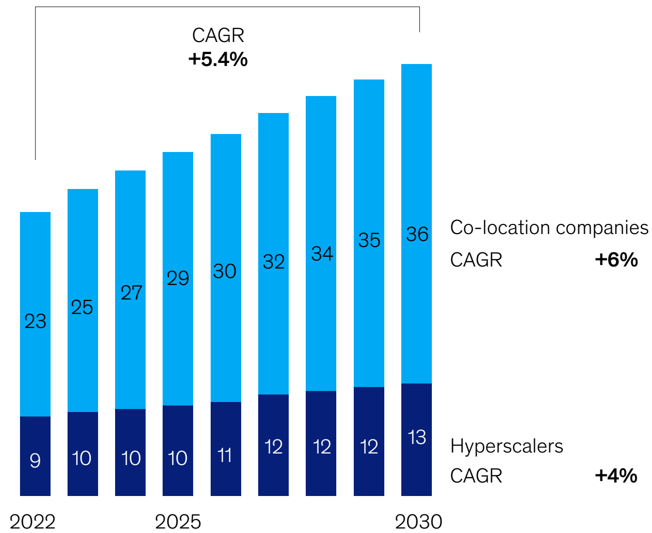
Waste-heat applications. To reduce a data center's carbon footprint, these applications use heat from data centers for other purposes, such as district heating. Amazon uses recycled heat from a data center in Ireland to supply district heat in Dublin, for example, and Facebook says that the heat from its Danish data center is warming 6,900 homes.[3]

[1] McKinsey analysis.
[2] "Riot announces first industrial-scale immersion-cooled bitcoin mining operation," Riot Platforms press release, October 19, 2021.
[3] Lauren Edelman, "Facebook's hyperscale data center warms Odense," Meta, July 7, 2020.

Exhibit 3

Global spending on the construction of data centers is forecast to reach $49 billion by 2030.

Data center construction spending,[1] $ billion



[1]Includes construction spending by providers. Excludes enterprise spending and any other capital expenditure outside of construction (such as equipment).
Source: Synergy Research Group

McKinsey & Company

— skid-mounted MEP equipment: metal racks for mounting and connecting mechanical, IT, power, and cooling components and systems

— enclosed MEP modules, which mount and connect the same kind of equipment, but in a cabinet

— all-in-one data centers: turnkey data centers, which are feasible only for smaller facilities of 1.0 to 1.5 MW.

Some companies manufacture parts for solutions and others integrate them. A few well-established players serve a broad range of industries; others work specifically with data center operators. But a long tail of start-ups focus on small slivers of this industry. They are too small to serve large companies, and that holds back demand. Investors in manufacturers or integrators of PFM components could help these companies expand their reach and improve economies of scale.

Edge computing

Although enterprises are rapidly shifting vast amounts of their work to the public cloud, they are also growing more knowledgeable about what *not* to store there. Applications (for instance, autonomous driving) that require real-time insights at very low latencies might be better conducted close to the data's source. The cost of transferring large volumes of data to and from the public cloud can also favor edge computing. So do data privacy and residency regulations that require certain types of data to be stored near their point of origin.

All this explains the growing size of the addressable market for edge computing. According to IDC, the worldwide spending of enterprises and service providers on hardware, software, and services for edge solutions was projected to reach $176 billion in 2022 (an increase of 14.8 percent from 2021) and $274 billion in 2025.[14]

Two types of investment opportunities related to data centers stand out. The first is real estate. Demand for edge computing must be met from smaller data centers, often located in urban areas close to customers, rather than huge facilities in far-off locales. The second is technology. The components of the edge-computing tech stack (the hardware, connectivity, platforms, and software) are not new. The challenge is how to deploy and architect them at the edge—how artificial-intelligence capabilities can be brought there, for example, or how to deploy and manage platforms.

Not surprisingly, the big edge providers (such as hyperscalers, telcos, and large systems integrators) are acquiring or investing in start-ups in the space. In March 2022, for instance, T-Mobile invested $40 million in Spectro Cloud to promote innovation in Kubernetes management. In April, Google Cloud acquired MobiledgeX, an edge-computing management specialist aiming to develop a standard orchestration layer for edge-computing assets. Private investors interested in data centers could also consider investments in edge technology.

As demand for data centers increases, investments in companies that operate them and in co-location companies remain an option. Yet potentially attractive opportunities lie in other parts of the data center value chain. Investors might focus on individual elements (such as green-power generation or immersion cooling) or invest where elements intersect—modular solutions for edge data centers, for example, or carbon-free edge data centers. Investors that already own data center assets could also consider vertically integrating critical elements across the value chain.

Investment activity in some areas, such as procuring green energy for data centers or cooling solutions, may be picking up. However, these are still far from the mainstream, leaving considerable untapped potential. Since the risk/return profiles differ from those for the acquisition of a data center, these areas will probably appeal to different sets of investors, depending on their investment objectives. With data centers now playing such a central role in the economy, it makes sense to consider the entire value chain.

Srini Bangalore and **Bhargs Srivathsan** are partners in McKinsey's Bay Area office; **Arjita Bhan** is a knowledge specialist in the Waltham, Massachusetts, office; **Andrea Del Miglio** is a senior partner in the Milan office; **Pankaj Sachdeva** is a partner in the Philadelphia office; **Vijay Sarma** is a senior expert in the London office; and **Raman Sharma** is an associate partner in the Toronto office.

The authors wish to thank Katy Bartlett, Luca Bennici, Marcel Brinkman, Vladimir Cernavskis, Rishi Gupta, Arpit Kaur, Matt Linderman, Andrew Mullin, Shaila Narang, Jakob Österberg, Nick Padgett, Jürgen Sailer, Lorraine Salazar, Nestor Sepulveda, Shashank Sharma, Even Skjellaug, Christoph Sohns, Adithya Sukumar, Shally Venugopal, and Philippe Wrobel for their contributions to this article.

Designed by McKinsey Global Publishing

[14] "New IDC spending guide forecasts double-digit growth for investments in edge computing," IDC, January 13, 2022.