Memorandum

To: Jessica Aiona, Bonneville Power Administration

From: Marjorie McRae, Research Into Action; Fred Schaefer, Cadeo

Date: March 19, 2018

Subject: TO007 Data Centers Market Actor Research

This memo summarizes findings from research conducted by the Cadeo and Research Into Action team (the team) on behalf of Bonneville Power Administration (BPA). The team conducted 23 interviews with data center market actors and experts, most exceeding 30 minutes in length, to investigate the research areas described in Table 1. Appendix C provides a complete list of questions this research addressed.

This memo uses common acronyms and industry terms. Appendix A presents a full glossary; the team defines two terms here:

- **IT:** information technology; in this report, refers to the energy-using hardware (computing equipment) that runs software.
- **Infrastructure:** the basic, underlying framework or fundamental facilities and systems; in this report, both the building housing the computing equipment and the building systems, including cooling, power conditioning, and other energy-using equipment.[1,2]

Table 1: Areas of Inquiry

Market Aspect	Key Research Question
Market trends	• What are the drivers of market growth? • How is the market changing?
Data center types	• What types of data centers are there?
Market actors	• Who are the market actors? How do they influence data center energy use?
Data center design and operations	• What are trends in data center design and operations? How are these trends influencing energy efficiency?
Energy efficiency and renewables	• What efficiency metrics are used? • What importance do market actors place on energy efficiency and renewable energy?

[1] The infrastructure houses, powers and cools IT equipment and supports connections with telecommunications infrastructure. For the largest data centers, electricity demand can be on the scale of that of small cities and cooling loads can dwarf those of large-scale manufacturing facilities. https://wiredre.com/data-center-broker/

[2] In contexts other than this report, speakers use the term to describe such things as the infrastructure of IT resources, which are as computing (conducted by servers), storage, and networking.

Market Aspect	Key Research Question
Market size and expected growth	• How large is the market? • What growth is projected?
Pacific Northwest considerations	• What are Pacific Northwest market characteristics?
Role of utility programs	• What utility programs are affecting data center energy use?

Key Findings

These findings reflect the perspectives of subject matter experts, market actors, and secondary research on the current state of the data center market, trends and expected growth, and new insights into the data center supply chain and customer decisions.

Data centers – both IT and infrastructure – are increasingly efficient, yet market growth overall and of applications requiring higher server densities outpace efficiency gains, resulting in rising energy consumption. Both interview contacts and secondary research confirm the data center market has been and continues to grow rapidly. Secondary research sources commonly forecast data center compound annual growth rates exceeding 10% for five or more years to come. One subject matter expert suggested that the seven largest data center companies, those whose names consumers commonly recognize, will be responsible for 80% of data center construction.

Contacts explained that many newer applications have high processing requirements, including applications that require graphical processing units (GPUs), such as machine learning; cryptocurrency mining, the activity of a highly volatile but burgeoning submarket; and many scientific and data analytic applications. Higher processing requirements, reflected in what the industry terms higher disk or server densities, use more electricity per square foot of IT-occupied data center space.

Although computing equipment and the infrastructure that houses it continues to be increasingly efficient, the overall market growth and, within the market, the growth of applications with high processing requirements applications, is outstripping efficiency gains, resulting in increasing data center electricity loads.

Contacts and secondary sources expressed the view that many cryptocurrency miners are insensitive to power costs and thus do not invest in energy efficient computing or infrastructure equipment. Contacts employed by hosted facilities (which the industry terms colocation facilities, where the "host" provides and operates the building and its cooling and power conditioning equipment), reported that their firm serves cryptocurrency miners. The significance of miners choosing hosted facilities is that these facilities are typically quite efficient.

> *Takeaway: Data centers constitute a large and growing electricity-using market. A few large firms that build and operate highly efficient facilities dominate market growth.*

Sources suggest the Pacific Northwest data center market will grow at roughly the same rate as, if not faster than, the national market. Contacts reported that data centers locate in areas that offer high connectivity (ability to connect to multiple telecommunication network carriers and cloud service providers), lower power costs, lower land prices, and water availability (for water cooled mechanical systems). Some of the large corporations cultivate a sustainable brand image and seek renewable power. The Pacific Northwest offers all these features. However, some projected data center growth owes to increases in edge computing, which places data centers close to their markets. The region has a relatively small population and thus BPA might anticipate that firms will build relatively fewer edge data centers in the Pacific Northwest than in regions with higher populations.

> *Takeaway: The region's electricity providers should prepare for the possibility of annual growth in data center loads of more than 10% for the next five years.*

The markets for most data center types are growing, with a few exceptions. Most industry contacts use a data center typology based on size and ownership structure. Sources report growth in all data centers owned and operated by service providers, spanning the size gamut from micro to hyperscale. Companies are currently spending more than ten billion dollars a year constructing hyperscale facilities in the US, a figure that market researchers expect will grow substantially.[3]

For internal data centers, those located within a business's facilities, sources anticipate a decrease in smaller facilities (those under 20,000 square feet) due to migration to the cloud and hosted facilities and anticipate low growth for facilities larger than 20,000 square feet.

> *Takeaway: The service provider market will grow rapidly, large internal data centers will grow slowly, and the remaining markets (smaller internal data centers) will shrink, increasing the efficiency of the overall data center market.*

Businesses increasingly use a mix of data center assets. Contacts described how businesses (as distinguished from service providers) are increasingly moving their data center assets to the cloud and to hosted facilities. Exceptions to this are firms that believe they need to control their business-critical applications (such as a credit card company's transaction processing or a healthcare firm's patient records).

Both contacts and secondary sources indicate that data center advisors have recently begun to advocate that the optimal approach for many businesses is to maintain a mix of data center assets, with data centers located in the business's facility (internal data centers), in the cloud, and in hosted facilities. Our interviews with two financial companies running business-critical applications indicated high efficiencies rivaling all but hyperscale facilities.

> *Takeaway: Business computing is becoming increasingly efficient as measured by energy use per square footage. Business data center assets as measured in individual facilities are and will continue to include a substantial proportion of inefficient assets.*

Subject matter experts believe that most smaller internal data centers use three times the power per server than the most efficient data centers. Multiple factors limit the ability and willingness of these businesses to upgrade these smaller (less than 20,000 square feet) internal data centers: The physical space in which the data center is located seldom supports appropriate air flow management. The businesses typically lack a point-person with responsibility for all computing assets; no one person

---

[3] The US comprises over 40% of a global hyperscale market that researchers expect to grow from $25 billion in 2017 to $80 billion in 2022. https://www.marketsandmarkets.com/PressReleases/hyperscale-data-center.asp; https://www.srgresearch.com/articles/hyperscale-data-center-count-approaches-400-mark-us-still-dominates

may be aware of all assets. The businesses typically do not track assets, nor asset utilization, and thus miss opportunities to consolidate loads onto fewer servers and remove unneeded assets. Energy efficiency investments can be expensive; contacts reported that such investments in smaller data centers do not provide a rate of return that meets corporate requirements.

> *Takeaway: Numerous barriers limit businesses' interest in upgrading the efficiency of their smaller facilities.*

Data center market organization is dynamic and market actors typically play multiple and shifting roles. Sources described a market whose organization and players change rapidly with technology advances and economic trends such as industry consolidation. For example, among the largest IT companies, some solely design, build, own, and operate their facilities, some solely outsource the building and operation of their facilities (which they may or may not design), and some do a mix of these approaches. Most cloud services are located on hyperscale facilities, most of which the largest IT companies own, but there are some exceptions to both assertions.

The industry uses the term colocation service provider to describe firms that host their clients' IT equipment in the host's facilities. The putative defining characteristics of these hosting firms are that their facilities host multiple clients (hence the term colocation), that they do not provide IT equipment (the clients install their own), and that they speculatively build their facilities to meet anticipated market demand. However, exceptions abound.

The industry uses the term managed service providers to describe firms that take responsibility for a client's IT assets, whether in-house, co-located, or cloud-based. Some of these firms also offer colocation facilities and cloud resources. The larger firms that advertise that they broker colocation facilities also describe themselves as advisors. Some of the advisory services overlap those of managed service providers. Some of these larger firms also engage in investment advising and matching investors to properties.

> *Takeaway: For the data center market, it appears less useful for analysts to think in terms of market actors and more useful to think in terms of services offered.*

Many types of firms influence the efficiency of computing resources. According to subject matter experts, the largest data center firms commonly design their own servers and may be working on developing application-specific processors. Given the vast numbers of servers they order annually (tens of millions) and huge annual electricity costs, these firms design or procure state-of-the-art efficient computing resources. Large firms that are not among the seven that dominate the market procure highly efficient equipment for their clouds although without the design advantages that the behemoths can deploy. Both types of firms – the large and the very largest – operate their assets efficiently.

The interviewed representatives of firms providing managed services reported that they rarely discuss energy efficiency with their clients but increase the efficiency of their clients' data center assets through virtualization, removal of redundant assets, and encouraging appropriate migration to the cloud. These firms also recommend new equipment, which overall has reaped the efficiency gains of prior innovations.

> *Takeaway: Most new computing resources are highly efficient. However, multiple factors are causing per-server energy usage to increase.*

Firms with data center specialization influence the efficiency of data center facility designs; they may or may not source mechanical and electrical equipment from manufacturers with data center specialization. All contacts agreed that owners need firms with a data center specialization to construct their facilities. Contacts described that the data center market is large, and there are numerous mechanical and electrical equipment manufacturers that either exclusively serve the data center market or have product lines geared to the data center market. In addition, builders of data centers source mechanical and electrical equipment that manufacturers do not tailor to the data center market. For example, an interviewed financial end user reported sourcing industrial-grade UPS (uninterruptible power supply).

Contacts anticipated that service providers' smaller data centers[4] are unlikely to have the very high efficiencies that typify service providers' larger data centers. Service providers may build out smaller data centers within existing facilities, such as cell towers, rather than in purpose-built space, the term for facilities designed specifically for their intended use (in this case, designed specifically for computing).

Contacts anticipate only the largest segment of the internal data center market will grow. These facilities are internal to the organization, but typically housed in stand-alone buildings, which will be purpose-built. Firms will engage data center professionals to design and construct these purpose-built facilities. These facilities typically will be efficient, although will lack efficiency measures whose paybacks exceed three years, which contacts thought was a typical investment criterion.

> *Takeaway: Most new computing facilities will be highly efficient. New data centers that are not highly efficient are nonetheless more efficient than older data centers.*

Methodology

This research involved a multi-stage data collection approach. The team first conducted secondary research with online resources. These resources provided background on data centers, characteristics of data center and equipment types, energy efficiency strategies, and current market trends. The team located sources that described the supply chain and identified and characterized the market actors.

The team used this information to augment the project's Task 1 deliverable (a slide deck describing the market, its equipment, and efficiency strategies) and to develop in-depth interview guides for subject matter experts. The team identified a potential list of experts to interview, prioritized the list in collaboration with BPA, and completed five interviews.

The team developed market actor in-depth interview guides based on findings from the expert interviews and secondary research, as well as insights from a data center conference one team member attended.[5] The team generated a potential list of market actors based on suggestions from the subject matter experts and other interviewed contacts, professional contacts of team members, and web searches. The web searches sought firms and contacts within the nine interview groups who have a high degree of data center knowledge and experience, and (as possible) are familiar with Pacific Northwest data center market.

---

[4] Smaller data centers include closets, rooms, and localized facilities, which range from under 100 to under 2,000 square feet.

[5] EUCI Utilities and Data Centers Conference, Austin Texas, January 30 to 31, 2018.

The team conducted 23 in-depth interviews with 10 types of contacts (Table 2).[6] This sample provides insights into the market; it is neither representative of the general population nor statistically reliable. The team augmented the interview findings with secondary research to provide more complete technical explanations than would be possible from the interview findings alone.

Table 2: In-Depth Interviews Completed

Contact Type	Number Completed
Subject matter experts	5
Enterprise (end user)	2
Colocation service providers	2
Colocation services brokers	3
Managed service providers	3
Hyperscale data centers	1
Designers, contractors, manufacturers*	4
Professional service organizations	1
Utility program implementers	1
Rocky Mountain Institute	1
Total	**23**

** We refer to this group collectively as "designers." All interviewed firms work internationally as well as regionally. They include two "engineering solutions" firms, which manufacture, design, build and service critical infrastructure, an architecture firm, and a design consultant.*

Detailed Findings

This section begins with a discussion of market trends, spanning both drivers of growth and market characteristics, an understanding of which provides a foundation for other study findings.

Market Trends

Contacts described the data center market as one that has grown and continues to grow rapidly due increasing demand for computing services that owe to many sources, as discussed in the next subsection. Ongoing innovation creates new applications for computing services, driving new sources of demand and more growth (while also replacing prior applications). Market actors (discussed in a subsequent section) and relationships change, as well, to exploit the opportunities innovation offers.

Contacts unanimously believed that a two-year horizon is about as far out as knowledgeable professionals can hazard a prediction. Market innovation and disruption are occurring so quickly as to render further speculation useless, with one exception. Contacts agreed that the explosive growth in the data center market shows no signs of abating and is likely to continue throughout the coming decade, if not longer.

[6] This memo uses the term "contacts" to describe the interviewed industry experts and actors.

Drivers of Growth

Digitization

Digitization is the process of capturing analog information in a digital format and is a key trend identified by contacts. The term applies to the capture of information previously in hard-copy format, as well as transforming business processes to create and record process metrics and analyze them to identify ongoing opportunities to increase efficiency or effectiveness. Medical organizations, for example, have been digitizing patient records to increase administrative efficiency and improve health outcomes through coordinated care.

Organizations have been pursuing digitization to varying degrees from the time they acquired computing resources, yet as recently as February 2017, *McKinsey Quarterly* reported that "on average, industries are less than 40 percent digitized, despite the relatively deep penetration of these technologies in media, retail, and high tech."[7]

The article encouraged companies to pursue digitization, concluding that "bold, tightly integrated digital strategies will be the biggest differentiator between companies that win and companies that do not, and the biggest payouts will go to those that initiate digital disruptions. Fast-followers with operational excellence and superior organizational health will not be far behind."

Digitization is implicit in the marketing term "**customer engagement**," another trend some contacts mentioned. The "engagement" is largely digital. While a ubiquitous term, adoption lags behind familiarity. Service providers are moving into this space, offering business customers suites of tools for digital engagement and customer service efficiencies described collectively as Unified (Customer) Communications as a Service—UCaaS.[8]

Blockchain is a specific type of digitization, the development of which has attracted increasing numbers of professionals across many and diverse industries in the last several years, as well as media attention. Many contacts mentioned blockchain or cryptocurrency mining (discussed below) as a key driver of growth.

Blockchain is a type of computer code designed to provide an open, transparent, distributed ledger of transactions that is secure and unalterable, enabling parties unknown to each other to trust the accuracy of the recorded transaction without an intermediary (such as a bank) acting to assure the transaction's legitimacy. The technology's name derives from the blocks of data that it links into a chain through encryption code.

The Energy Web Foundation, a co-creation of the Rocky Mountain Institute (RMI) and Grid Singularity, is developing a blockchain technology and platform to support energy trading among owners of distributed energy resources (DER), from the smallest to the largest resources. Some blockchain

---

[7] Jacques Bughin, Laura LaBerge, and Anette Mellbye, The case for digital reinvention. McKinsey Quarterly, February 2017. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-case-for-digital-reinvention

[8] A cloud-based service currently smaller than UCaaS, yet closely related, is CCaaS—Contact Center as a Service.

technologies are highly energy intensive, such as cryptocurrency mining[9] (although the energy intensity of mining differs among the specific processes for verifying transactions employed by the various currencies). Given the potential for extravagant energy use, the Energy Web Foundation is simultaneously working to develop a blockchain technology that uses minimal energy.

The team interviewed an RMI Associate and founding member of the Energy Web Foundation to explore blockchain applications and potential energy use. The contact confirmed that potential applications for blockchain are wide-ranging, not limited to financial and property transactions. Uses currently under development include such applications as government documentation, electronic health records, and productivity analysis.

This contact believed that, while the blockchain technology is highly likely to transform a substantial portion of digital information, an efficient blockchain technology should not use any more computing resources than the applications it replaces. Thus, according to this contact, the anticipated blockchain revolution is unlikely to result in increased energy use, exclusive of some cryptocurrencies, a topic discussed further below.

Cloud Computing

All contacts referenced the emergence and ongoing expansion of **cloud computing** as a key driver of data center growth. The complexity and scale of cloud computing is increasing. Contacts and secondary sources describe how the structure of cloud computing services continues to evolve, with new structures spawning new service providers and terms. International Data Corporation (IDC) estimates that 60 to 70 percent of all software, services, and technology spending by 2020 will be for cloud-based resources.[10]

The cloud is a network of remote servers, each with a unique function—and thus unique characteristics—hooked together via networking software to operate as a single ecosystem. The physical storage of information "in the cloud" spans multiple servers in various locations typically owned and managed by a hosting company.

Most consumers familiar with the term are referring to the **public cloud**, whose services include applications, resources, and storage for consumer and many business activities, for which the public internet provides the networking. The concept also includes **private clouds** that offer services to users within an enterprise over a network internal to the enterprise; these systems comprise private IT architecture dedicated to a single organization. Community clouds share resources among governmental agencies and related organizations.

Each cloud type includes multiple clouds, and users may use cloud services from one or more providers. Software developers choose among, for example, those of Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (the "Big 3," per Gartner, Inc.'s 2017 annual scorecard), as well as many others. The large cloud computing services companies have multiple hyperscale data centers (a term described subsequently).

---

[9] Cryptocurrency is a digital currency for which encryption regulates the generation of units of currency and verifies the transfer of funds, operating independently of a central bank.

[10] IDC is a market research firm that specializes in information technology, telecommunications, and consumer technology markets. It is the industry-leader in the provision of current and forecasted estimates of equipment and services supplied to the data center market.

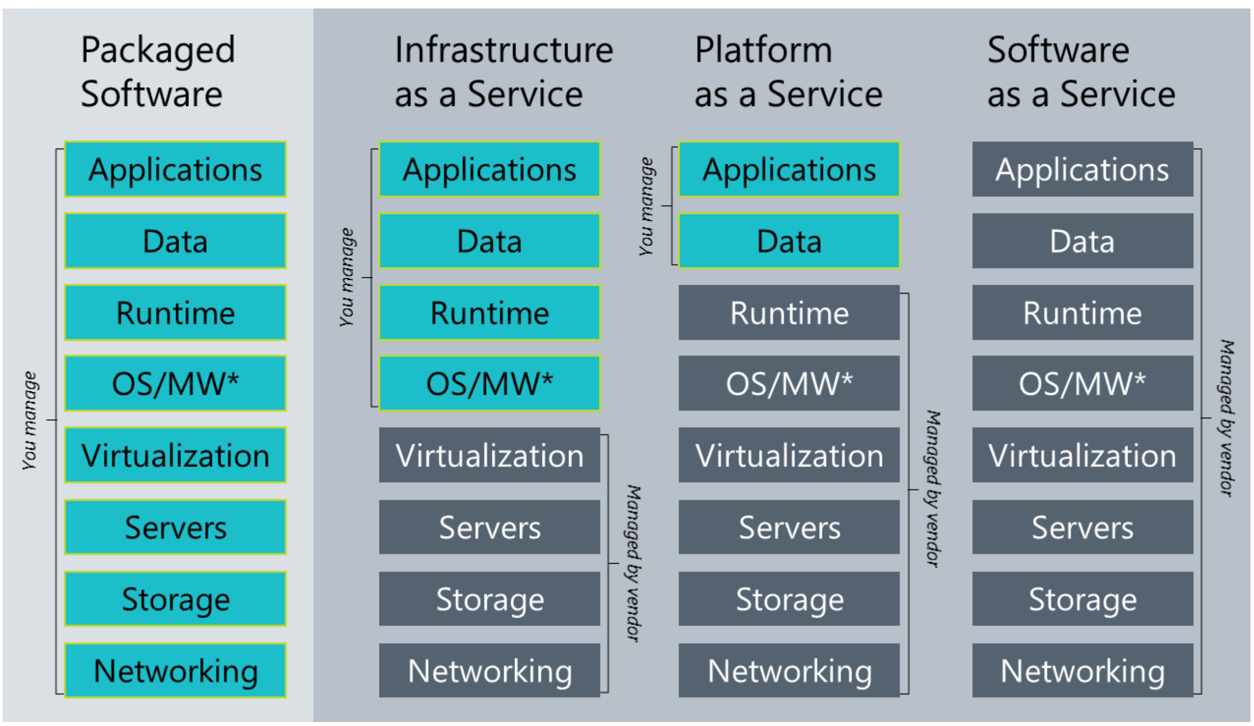
Cloud computing includes internet enabled services such as social networking and content streaming that IT nonprofessionals are aware of, as well as application development infrastructure, application deployment (use), storage, networking, and other services. For example, Microsoft Azure describes itself as "a comprehensive set of cloud services that developers and IT professionals use to build, deploy, and manage applications through our global network of datacenters…to build anything from simple mobile apps to internet-scale solutions."[11]

Cloud service providers offer services for one or more of the following types of activities, delivered on public and private clouds, depending on the product:[12]

- **Internet enabled services** (include social networking, e-commerce, email, gaming, and storage)
- **Software as a Service** (SaaS; subscription software;)
- **Infrastructure as a Service** (IaaS; virtual data center)
- **Platform as a Service** (PaaS; application development resources)

Figure 1 illustrates the meanings of SaaS, IaaS, and PaaS.

Figure 1: Cloud Computing Taxonomy



** OS/MW is operating system and middleware.* Source: Microsoft 2013

[11] https://azure.microsoft.com/en-us/overview/what-is-azure/

[12] These four are the largest product types. There are others, and seemingly more every day, such as Unified Communications as a Service (UCaaS; customer engagement resources).

Contacts spoke of the **hybrid cloud** as an increasingly prominent technology that shares services between public and private clouds, which operate independently and communicate over encrypted connections. For example, Microsoft first offered Azure as a public cloud and then introduced its private cloud platform in 2016 to enable its customers to incorporate features from the Azure public cloud with those of their own data centers.

The distinction between public and private clouds and the profusion of clouds has led to a new industry of cloud connectivity services. These firms have their own data centers, with networking equipment to link firms' internal data centers (of all sizes and including their access to the public web) with private clouds.

Connected Devices

Connected devices are becoming increasingly pervasive and contacts describe them as a primary driver of growth in data center market. According to Gartner, in 2017 there were over eight billion connected devices worldwide, a figure that Gartner expects to roughly double by 2020.[13] The incorporation of sensors into everyday items is driving the growth in connected (networked) devices and computing resources, a phenomenon termed the **internet of things (IoT)**.

Consumers may or may not be aware of their many connected devices, including mobile phones and tablets, wearables, and smart home technologies such as appliances, lighting, and security. More than three-quarters of IoT smart devices are in the commercial and manufacturing sectors (40%) with machine-to-machine (M2M) communications enabled by building and equipment automation and controls; in healthcare (30%), notably for portable health monitoring and electronic recordkeeping; in retail (8%; inventory tracking, smart phone purchasing, consumer choice analytics); and security (8%; biometric and facial recognition locks, remote sensors).[14]

Contacts described current trends leading to the proliferation of IoT devices include **smart cities** and **autonomous vehicles (AVs)**. Smart cities use internet-connected sensors to collect and analyze data to inform cities' provision of infrastructure services and manage strain on urban resources. Cities are incorporating internet-enabled sensors into transportation infrastructure, lighting, waste systems, and buildings at increasing rates.

The advent of the IoT and smart cities is also leading to, and concurrent with, increased automation in the transportation system. In addition to automating public transportation infrastructure, major car manufacturers and technology companies are experimenting with automating and connecting personal and commercial vehicles. Navient Research projects that the number of vehicles sold globally with at least Level 2 automation (automated steering, acceleration, and breaking in certain circumstances) will be over 90 million by 2026 – up from 250,000 in 2017.[15]

---

[13] Gartner, 2017. Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. https://www.gartner.com/newsroom/id/3598917. Other estimates vary widely. Some reports estimate up to 200 billion connected devices by 2024.

[14] https://www.intel.com/content/www/us/en/internet-of-things/infographics/guide-to-iot.html

[15] Navigant Research, 2017. Automated Driving Vehicle Technologies: Lidar, Radar, Vision Sensor, Ultrasonic Sensor, and Compute Platform Technologies: Global Market Analysis and Forecasts. https://www.navigantresearch.com/research/automated-driving-vehicle-technologies

Connected devices making up the IoT, smart cities, and AVs all rely on decentralized, edge computing (discussed further below) to reduce latency – the delay in data transfer. Many of these devices rely on wireless communications, and data transfer speeds currently limit some applications, particularly AVs. The upcoming rollout of **fifth generation (5G)** wireless technologies promises reduced network latency and increased ease in adding internet-enabled devices and sensors.

Yet data transfer is only one portion of the computing implications of IoT. "The true value of IoT is being realized when the software and services come together to enable the capture, interpretation, and action on data produced by IoT endpoints," according to Carrie MacGillivray of IDC.[16] As the number of connected devices grows, so will the computing resources companies dedicate to analytics, along with networking and storage.

Artificial Intelligence, Virtual Reality, Augmented Reality, Cryptocurrency

New computing demands in the form of **artificial intelligence (AI)**, including machine and deep learning, virtual and augmented reality (VR, AR), and cryptocurrency mining have increased the need for both hyperscale data centers and edge computing. With digitization and flows of data from IoT devices, AI is increasingly necessary for processing the vast amount of data generated.

Statista estimates that the global AI market will be about 60 billion dollars by 2025, up from about two billion dollars in 2017.[17] Industry analysts expect increasing automation to be a primary driver of AI technologies. MarketsandMarkets, a prominent market research firm, attributes its estimated 47% compound annual growth rate (2017 to 2022) of AI in the telecommunications market to increased industry adoption in general and use of AI-enabled smartphones specifically.[18]

Growth in **virtual and augmented reality** technologies follows a similar trend, with global market estimates of 215 billion dollars in 2021, up from nine billion dollars in 2017.[19,20] The general term for VR/AR is immersive technologies. Technology leaders are applying immersive visualization to optimize the functioning of business assets, and linking this visualization to context-aware computing, which uses computer technology to collect and analyze data about its own surroundings, either within the enterprise or the user's device.

Several contacts included **cryptocurrency**, which runs on blockchain technology, as a driver in data center growth in regions with low electricity costs. The process of cryptocurrency mining is an energy intensive and involves adding transactions to the blockchain by solving increasingly complex

---

[16] https://www.idc.com/getdoc.jsp?containerId=prUS42799917

[17] Statista, 2018. Revenues from the artificial intelligence (AI) market worldwide, from 2016 to 2025. https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/

[18] MarketsandMarkets™, 2018. AI in Telecommunication Market by Technology, Application (Network Optimization, Network Security, Self-diagnostics, Customer Analytics, and Virtual Assistance), Component (Solutions and Services), Deployment Mode, and Region - Global Forecast to 2022". The report identifies the top US firms in the AI telecommunications market as: IBM, Microsoft, Intel, Google, AT&T, Cisco Systems, Nuance Communications, Sentient Technologies. H2O.ai, Salesforce, and NVIDIA.

[19] Virtual reality is the computer-generated simulation of a three-dimensional image or environment with which the user can interact is a seemingly real or physical way. Augmented reality provides the user with a composite view of a computer-generated image with the user's unmediated view of the world.

[20] Statista, 2018. Forecast augmented (AR) and virtual reality (VR) market size worldwide from 2016 to 2021. https://www.statista.com/statistics/591181/global-augmented-virtual-reality-market-size/

computational problems. As of February 2018, the Bitcoin Energy Consumption Index estimated annual energy use of Bitcoin (the prominent cryptocurrency) to be about 50 terawatt-hours.

Contacts agreed that cryptocurrency market is highly volatile and difficult to predict, but there is rough consensus that growth in this area will continue, at least in the next few years. New cryptocurrencies that rely on proof-of-stake (PoS) or proof-of-authority (PoA) to verify transactions rather than the energy intensive process of proof-of-work (PoW) have the potential to significantly decrease the energy intensity of mining operations.

All the new computing demands of AI, VR/AR, and currency mining rely heavily on servers with **graphics processing units (GPUs)**, in addition to central processing units (CPUs), as discussed by some contacts. GPUs are better suited for multiple, simultaneous operations; CPUs are better suited to perform calculations in sequential order. GPUs, initially used for graphics and gaming, have recently become integral in many data center servers.

Computational tasks such as natural language processing, voice search (searching the web using voice commands), image recognition, and cryptocurrency mining all require the quick processing of vast quantities of data simultaneously. Due to the high demand for GPUs, Nvidia, the largest manufacturer of GPUs, reported supply constraints in early 2018, as well as a 34 percent increase in revenue.[21]

Market Characteristics

Challenges

The practical and strategic challenges faced by firms and service providers drive current and trending market characteristics. **Reliability**, also termed **uptime**, is arguably the single most important driver, as loss of service means costly loss of revenue and reputation, especially among service providers and firms with large internal data centers.[22] Reliability drove Oracle to invest in the development of what it terms "the next generation of the industry-leading database, the Oracle Autonomous Database Cloud," with "autonomy" now in the lexicon with uptime.[23] Oracle accomplished its objectives through AI, using "ground-breaking machine learning to enable automation that eliminates human labor, human error and manual tuning, to enable unprecedented availability".[24]

As cloud computing in its varied manifestations grows, **bandwidth** (capacity; the amount of information that can be transmitted simultaneously) becomes a limiting factor. Fiber optic cables support much of the nation's (and the world's) internet systems, as well as cable television and telephone services. Long-haul fiber-optic infrastructure carries digital signals from coast to coast and border to border.

---

[21] Fortune, 2018. Here's Why Nvidia's Stock Is Booming—Again. http://fortune.com/2018/02/09/nvidia-stock-data-center/

[22] Each year, even the most advanced service providers experience outages. In 2017, the ten largest outages, according to CRN and arranged chronologically, were IBM, GiLab, Facebook, AWS, Microsoft Azure, Microsoft Office, Apple iCloud, AWS, and Google Docs. https://www.crn.com/slide-shows/cloud/300097151/the-10-biggest-cloud-outages-of-2017.htm

[23] The research team is indebted to information presented by Jeff Brown, Director at Windstream Communications, at the EUCI Utilities and Data Centers Conference on January 31, 2018. His presentation, The Emergence of the Hybrid Cloud, has informed this and previous section on cloud computing.

[24] https://www.oracle.com/corporate/pressrelease/oow17-oracle-autonomous-database-100217.html

A 2015 study created a map of US long-haul networks (see Figure 4) and determined that most of them are co-located with some combination of major roadway and railway infrastructure

Latency, or data transfer speed, and bandwidth are interdependent; network congestion leads ISPs (internet service providers) to rout information over longer distances, increasing the delay. However, the speed of data transmission when the network is not congested is determined by the technology, not be the degree of congestion.

The industry measures latency in milliseconds; all users want fast internet. However, latency drives profitability and market share, especially for financial transactions, content streaming, and immersive/contextual applications.

Cybersecurity, the exclusion of unauthorized access, continues to be a challenge that assails both internal and service provider data centers. More than forty organizations had data breaches in 2017, according to IdentityForce, exposing data on over 100 million users. [25] These organizations included financial institutions (including Verifone, Equifax, and Deloitte), universities, hospitals, state agencies, medical insurers, fast food and hospitality companies, retailers (including eBay), and gaming firms, among others.

The data center market is nothing if not dynamic, experiencing constant growth and change. Companies need **flexibility** in data center deployment to seize unfolding market opportunities and meet technical challenges and demands. Flexibility includes **scalability**, which describes the ability to quickly add or decrease capacity as market conditions change. Businesses may also need **mobility** to redeploy in new locations previously built-out assets. They need flexibility at the speed of business.

Contacts described the challenge posed by the **high cost** of data centers and market changes to address this. Contacts referred to high **CapEx** (capital expenditures) to describe the cost, meaning high capital expenditure needed to build a data center.

Energy (both electricity and any back-up generation fuels) constitutes the largest operating cost. Energy to power IT equipment typically (over the recent past) comprises more than half of the energy use, with the remainder used to power the cooling and power conditioning equipment.[26]

Facility operations require a team of dedicated, specialized engineers and technicians to maintain system reliability. This team adds to the operating costs, yet labor costs are less than energy costs for midsize and larger data centers, and only a fraction of the energy costs of large centers.

Challenges facing suppliers in this rapidly changing market are intense **competition**, industry **consolidation**, and corporate behemoths. The next section discusses responses to these challenges.

Responses

Contacts described **edge data centers** as the market's response to challenges of bandwidth and latency. Historically, the internet exchange points were located at the largest (tier-one) markets, such as New

---

[25] https://www.identityforce.com/blog/2017-data-breaches.

[26] See Energy Efficiency section, below. PUE – power usage effectiveness – is the industry's energy efficiency metric. PUEs less than 2.0 indicate that IT loads use more than half of total facility energy use. Current PUEs are below 2 for the vast majority of computing. Five years ago, or so, PUEs commonly were higher than 2.

York, Chicago, and Los Angeles. Internet-based content provided to users in smaller cities came from these centralized exchange points. Edge data centers expand the "edge" of the internet to smaller **tier-two** markets. Edge data centers are especially critical for content streaming and immersive applications.

The industry has responded to the rapidly evolving market and burgeoning demand with **rapid** construction times, which contacts termed **time-to-market**. In 2013, according to AWS, the industry average data center build time was 24 months.[27] AWS EdgeConneX, as one example of increasing speed, built 40 edge data centers in four years (Q2 2013 to Q2 2017), 35 of which were built within a time span of 38 to 108 days (most within two to three months). As multiple contacts said using similar phrasing, "time-to-market is critical because you are trying to stay ahead of the competition."

Service providers have developed **modular** data centers to meet market needs for speed of deployment, scalability, and mobility, as well as offering cost advantages (considerably reducing CapEx). They can be placed anywhere capacity is needed. Modular data centers can provide stand-alone capability, can be networked together to form a new data center larger than the modules, or can be used by owners to incrementally grow or decrease existing data center capacity without disturbing its existing functionality. Contacts frequently mentioned the importance of the flexibility that modular designs provide.

Containerized data centers are the most common type of modular data centers. The service provider fits the IT equipment into a standard shipping container and transports it to the desired location. Containerized data centers typically include cooling equipment. Another modular approach uses prefabricated infrastructure components of shipping-container dimensions that the owner builds out on site. Still other approaches deliver a structure complete with IT equipment, such as offered by Hewlett Packard Enterprise (HPE).

Modularity refers to all elements of the data center – modularity in deployment of IT racks, modular options for optimizing airflow management in a rack or cabinet, skid-mounted **power modules**, and modular facilities of containing both IT and infrastructure.

Combining both the edge and modular developments, service providers are responding to the opportunities and challenges of the IoT with **micro data centers**, a new, rapidly growing sector. Service providers are building micro data centers in cell towers, among other locations. These micro data centers collect and analyze local IoT data and send the analysis to larger data centers, where it feeds into other applications.

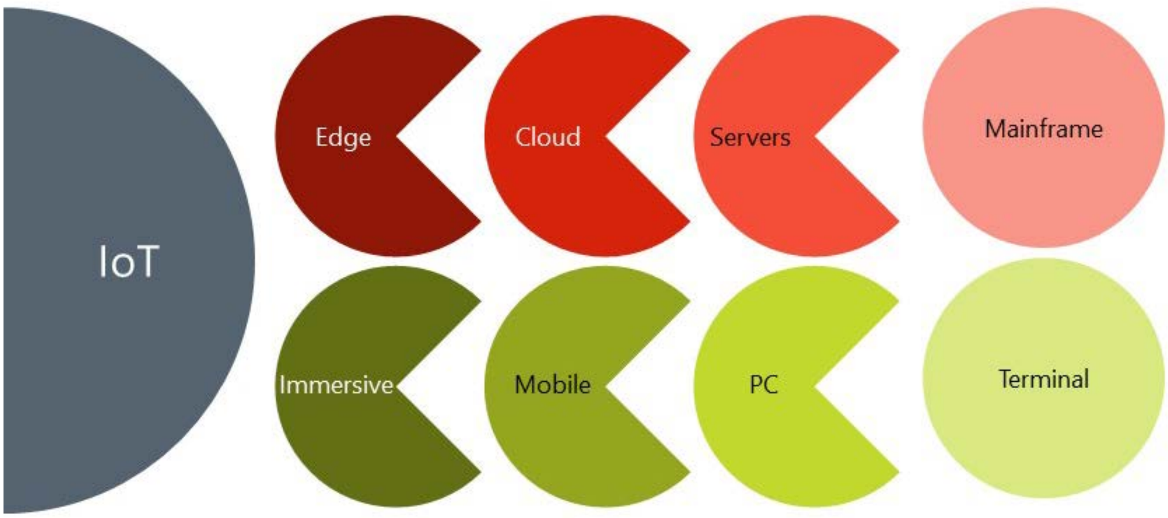
Most businesses are responding to the evolving data center opportunities and challenges by having a **mix of data center assets**, including internal (on-site), IT systems collocated with those of other firms in data center facilities (infrastructure) owned by service providers, private cloud and public cloud (or perhaps hybrid cloud services). This market response in turn has led to a niche for service providers to offer **data center infrastructure management (DCIM)** that combines the management of both facilities and software.

Figure 2 illustrates the evolution of the data center market (top row) and the associated user equipment (bottom row) from the market's inception (right side) to the IoT revolution (left side). The data center market began with mainframe computers located at large institutions such as universities that users

[27] http://www.edgeconnex.com/

accessed via terminals located in their offices. The data center market transitioned to on-premise servers supporting numerous personal computers (PCs). The next transition was to cloud computing facilities that users accessed from their PCs and mobile devices. Service providers then built midsized edge data centers to meet local markets' latency needs (primarily for streamed content) and micro edge data centers to handle large streams of localized IoT data. These facilities support users' immersive devices, as well as their PCs and mobile devices.

Figure 2: Evolution of Data Center and Computing Markets



Source: Windstream Communications, EUCI Utilities and Data Centers Conference on January 31, 2018

Market actor interviews suggest the cloud hosting market is very competitive, with firms with small market share increasingly acquired by larger firms or dropping out of the cloud market, resulting in further market **consolidation**. These findings coupled with the fact that hyperscale data centers are the most efficient (according to interviewed subject matter experts, as well as numerous published documents) suggest that as the cloud market burgeons, its energy intensity will decrease. Given the typical efficiencies reported by contacts for the different types of data centers, the percent decline in energy intensity will be small, but absolute energy savings will be sizable given the magnitude of the cloud market.

The trend to cloud computing is not distinguishable from the trend to hyperscale data centers. Both reap economies of scale. Accordingly, interviewed contacts who addressed the economics commonly mentioned that cloud computing is cheaper, a conclusion fully consistent with the marketing messages of cloud computing service providers.

However, several contacts expressed the view that cloud computing is not necessarily cheaper, for multiple reasons. One, its cost is driven by the resources one acquires and an end user can be overconsuming cloud resources just as it can overconsume and underutilize its own data center resources.[28] Two, as one contact pointed out, end users can migrate to the cloud at a relatively low cost (from $100,000s to $1,000,000s depending on the size and complexity of the migration), but exiting from the cloud and migrating to one's own IT equipment (whether in colocation or internal facilities) is a much more expensive undertaking.

Contacts universally expressed the opinion that the data center market continues to burgeon. Within that market, the larger data centers are getting larger.[29] As one contact described this, "Everything is getting bigger. Hyperscale and colocation facilities are getting bigger." While all contacts agreed this is true, many contacts also described tremendous growth in both midsize and micro service provider facilities in response to high demands for content streaming, increasing IoT, and other trends.

Data Center Types

Contacts and secondary research suggests that the universities, market research firms, and other organizations that estimate market size and growth appear to use a single data source, that of IDC. Consequently, most analysts use IDC's typology of 11 data center types.

The IDC typology defines two ownership categories: **internal** – data centers located in the commercial end-user's facilities, and **service provider** – data centers located remote to the end-user, owned and operated by what are termed data center service providers.

Embedded data centers is another term for internal data centers. Embedded facilities are not **purpose-built** facilities like those constructed by service providers but rather are located within the organizations' facilities. One contact observed, "Internal data centers are everywhere and nowhere. Everybody has one, but we don't know where they are."

IDC defines six size categories (Table 3), the largest of which (**hyperscale**) only service providers have. IDC also provides descriptions of the typical infrastructure system characteristics, which vary from minimal dedicated cooling and redundant power in server closets to advanced cooling and power systems in hyperscale.

Table 3: IDC Size Categories

Size Name	Typical Square Feet	Size Name	Typical Square Feet
Server closet	< 100 ft^2	Mid-tier	2,000-19,999 ft^2
Server room	100-999 ft^2	High-end	>20,000 ft^2
Localized	500-1,999 ft^2	Hyperscale	Up to over 400,000 ft^2

Source: IDC

Note: Two size ranges overlap: server rooms and localized.

[28] The end user overconsuming cloud resources pays more than necessary for services received. But there is no underutilization of cloud resources nor opportunity for energy efficiency gain because the cloud services provider does not hold the additional computing capacity in reserve. Instead, it constantly adjusts its resources to every changing demand.

[29] The large data centers are comprised of multiple facilities. Owners can add additional buildings to an existing data center complex.

The industry also commonly describes data center size in terms of capacity (the megawatts or kilowatts drawn by a data center). Through the team's contacts and secondary research, it was unable to discover a size threshold denoting hyperscale, however contacts expressed the view that "you know it when you see it." According to one internet website, the world's largest data center, at 150 MW, is in Inner Mongolia; the US has two data centers in the top 10, a 90 MW NSA data center in Utah and an 85 MW data center in Illinois.[30] Synergy Research Corporation published in December 2017 its estimate of about 400 hyperscale data centers by the end of 2017, 44% of which were in the US.[31] The industry uses **mega** as an alternative term for hyperscale.

According to a colocation space broker, high-end internal data centers are commonly between 100,000 and 200,000 square feet. Colocation service providers (discussed below) commonly construct facilities in the 30,000 to 40,000 square feet range, although they may also have mid-tier and hyperscale centers. They build out the infrastructure in phases, at about 5,000 or 10,000 square feet at a time.

Edge data centers come in all sizes up to the lower range of high-end. These larger data centers serve Tier 2 metro areas, such as Portland and Seattle.[32,33] **Micro** data centers are the smallest of data centers and can be in such locations as cell towers and specially designed or repurposed space.[34]

Market Actors

The term "market actors" describes types of individuals or organizations that conduct transactions in a market. We can use this term for the data center market, but the term falls short in a literal sense because the organizations frequently vary their roles according to the transaction. Most accurately, the terms the industry uses for market actors describe the types of transactions and not actors whose role in the market remains the same from one transaction to the next.

Enterprise

In the data center market lexicon, **enterprise** frequently refers to businesses whose exclusive role in the market is that of consumer of data center services. Yet in the data center market broadly defined, enterprises may build and operate their own data centers (what the team terms "internal" data centers), which bends the definition of consumer; the team is unaware of any other market in which one describes an organization's relationship with its own assets as that of being a "consumer" of those assets.

The team found it most helpful to interpret Interview and secondary sources as typically using the term enterprise to describe the end user, to use a term common to the energy efficiency industry.

Most accurately, enterprise describes a type of data center *use* – a business's use of computing services for its internal management and administrative activities as well as those related to its core business, such

---

[30] http://worldstopdatacenters.com/power/

[31] https://www.srgresearch.com/articles/hyperscale-data-center-count-approaches-400-mark-us-still-dominates. This publicly available source did not define hyperscale size.

[32] Tier 1 metro areas are the very largest, such as New York and Los Angeles.

[33] See http://www.edgeconnex.com/ for the locations of edge data centers AWS built from 2013 to 2017.

[34] Micro data centers fall in the service provider server closet category.

as the data centers Netflix uses to stream content to subscribers. “Enterprise” would *not* refer to the core business of such firms as AWS, whose core business is to provide users with web computing services.

Both data center service providers and end users engage in enterprise computing on equipment they may or may not own and operate.[35]

The major enterprise end user segments are: **BFSI** (banking, financial services, insurance), healthcare, and government. Some sources also include retail and/or travel and hospitality. These are also the major end user segments of the data center market generally, along with IT, telecommunications, and media firms, which are service providers.

Firms in the BSFI industry commonly have large data centers that they tightly control. Interviewed end users mentioned that their professional in-house staff operate and upgrade, commonly termed “refresh,” their data centers to minimize the possibility of data breaches. They have vendor partners that advise them on equipment and designs, but they stay abreast of innovative technologies and have the expertise to have confidence in their decision making.

These end users do not typically use service providers, with one exception. They often have an additional, smaller data center in colocation facilities (a term described below) located in a different region, to provide resiliency and enable disaster recovery.

Service Providers

Service providers are, unambiguously, organizations whose role in the market is that of supplier of data center services. However, that is not necessarily these organizations’ sole role. They may also be a consumer of data center services provided by other service providers.

To reiterate this key point, these different “types” of market actors are not distinct players in the market. The elements discussed in this section are different services, not different actors. For example, although some contacts categorically stated that hyperscale facilities are owner-operated, this appears to be a generalization. Interviewed colocation service providers reported that the large colocation providers own hyperscale facilities.

One of these contacts reported, based on information from a source he deemed reliable, that Microsoft outsources all its new data centers (that is, it relies on wholesale colocation service providers), Google and Facebook both own and outsource, and AWS has exclusively owner-builder-operator facilities.

Outsourcing is attractive to even these data center behemoths because, as our contacts point out, infrastructure design, build, and operation is its own core competency, unrelated to IT competency. The large firms that continue to have owner-operator facilities have, according to this contact, “a division [to handle the infrastructure] that is essentially a separate company.”

Managed Services

Managed service providers manage end users’ IT assets, including procuring those assets. Asset procurement encompasses purchasing IT equipment (or recommending equipment to purchase) and

[35] For example, an interviewed colocation space broker described “enterprise” data centers in Washington owned by Dell and Intuit.

establishing and managing cloud computing resources. The latter includes orchestrating end users' migration from internal data centers to the public cloud and creating private and hybrid cloud systems.

Managed service providers serve, or can serve, all a clients' data center assets: internal data centers, hosted or colocation data centers (see below), and cloud assets. Some colocation providers offer managed services as well.

Cloud Services

Cloud service providers provide, not surprisingly, cloud services. Contacts reported that these providers may be builder-owners-operators and/or operate clouds from colocation facilities, according to our contacts.[36,37] The largest data center market actors, which our contacts referred to variously as the Big 5 or Big 7, and include AWS, Google, Microsoft, and Facebook, are all cloud service providers. Contacts described that end users can place any software – such as Microsoft's – on AWS, whereas at the other end of the spectrum, Facebook's cloud solely uses its own proprietary software.

Cloud data centers include networks of large data centers located across the globe. Contacts agreed that cloud services are not provided regionally. Cloud owner-operators geographically distribute the loads to enable optimal operation of their assets while best meeting user demands.

Hyperscale

As described in the previous section, hyperscale data centers are a "type," not a service. The team mentions them here to show how they fit into the service provider market. Most cloud service providers deliver those services from hyperscale data centers. Most of the computing conducted by the large IT and telecommunications firms on behalf of their clients occurs in hyperscale data centers. These firms bypass original equipment manufacturers (**OEMs**), preferring to design proprietary equipment to achieve their computing and efficiency objectives. These firms work with equipment and design manufacturers (**EDMs**) that manufacture to their specifications and provide them with what the industry terms **unbranded equipment**.

An interviewed subject matter expert reported that the Big 7 largest IT companies construct 80% of new data centers, most of which is hyperscale.[38]

Colocation service providers, discussed next, may have hyperscale data centers among their assets, especially when they also offer cloud computing.

---

[36] Different entities may build a cloud-hosting data center, own the data center, and operate it. The possible configurations are three entities, one for each role; two entities (three possible combinations); and one entity. "Builder-owner-operator" refers to a data center for which one entity has all those roles.

[37] In the words of one interviewed broker, "The cloud is a data center. People talk as if they are different things. The only difference is who controls the equipment."

[38] Sources typically identify the Big 7 companies as Apple, Alphabet, Amazon, Facebook, Microsoft, Yahoo, and IBM, although the various size metrics employed by different analysts yield somewhat differing rankings.

Colocation Services

Colocation service providers supply colocation services to the market. But our interviews and secondary research confirms that these service providers frequently provide other types of services, as well. Thus, more accurately, "colocation services" has a single meaning, whereas the term colocation service providers does not describe a distinct, unique set of market actors.

Colocation services are offered by firms that design and construct data center facilities – the infrastructure, not the IT equipment – and lease the space in a given facility to multiple businesses. The colocation services thus include the ongoing operation and maintenance of the facility. By leasing the data center facilities, firms eliminate the CapEx associated with these facilities and increase their operating costs.

Contacts commonly used the term "**hosted facilities**" to refer to colocation data centers. Contacts mentioned it is most common for end users to seek hosted facilities in geographical regions removed from their headquarters. They want geographically diverse data center assets to best serve their customers and to minimize the likelihood that multiple data centers would be inoperative at once.

To complicate matters, some contacts distinguished between wholesale and retail colocation services. In this terminology, **retail colocation** services describe the situation where one facility houses multiple businesses while **wholesale colocation** services describe the situation where a single user occupies a single facility. Other contacts describe the latter situation as "**build-to-suit,**" where the supplier builds a data center facility – infrastructure only – and sells to a single user, or possibly leases the space to that single user and continues to operate the facility. Some colocation firms commonly build-to-suit, others do it only for their colocation customers, and some do not offer this service.

Colocation service providers do not provide the IT equipment the facilities' occupants use, except in the uncommon but not rare cases when they do. The industry terms these distinctions "**managed colocation**," where the service provider is responsible for the IT equipment, and "**non-managed colocation**," the more common arrangement. As another variant, some end users employ managed service providers to manage their hosted data centers, although according to our contacts, most do not.[39]

Although the term "colocation" suggests that multiple end users is the defining characteristic of these services, perhaps more importantly colocation services describe the ongoing provision of highly reliable infrastructure to house end users' IT-only data centers.[40] According to our interviews and secondary research, colocation firms are staffed with engineers and technicians to ensure power reliability and appropriate indoor temperatures. Upstream from operations, these firms have established design and procurement specifications for building their facilities.

[39] Most commonly, the clients send their own IT staff to the colocation facility as needed, which is they add or remove equipment or, much less frequently, a problem arises that they cannot resolve through a remote connection to the server. An interviewed end user whose colocation facility is over 500 miles from its headquarters explained that his firm uses "remote hands" as needed, and provides those "hands" with extremely detailed instructions, such as "disconnect this cable from spot x and reconnect to spot y."

[40] From BPA's perspective, the colocation facility is a single data center. From the end users' perspectives, each of their colocation-hosted data centers is simply another data center. An end user the team interviewed, for example, spoke of having two data centers, one internal and one at a colocation facility. We have coined the term "IT-only data center" to describe an end-user's data center within a colocation data center.

Another contact thought the defining characteristic of colocation firms is that they are speculative builders. They identify locations for data centers that they believe can attract a certain number of end users and build the data center without firm commitments from any client. Consequently, colocation data centers may take one to three years, contacts suggested, to fill up.

Colocation services brokers, as the term suggests, are real estate brokers that include a specialty in colocation services. Some of these firms exclusively broker what one of their websites' terms "IT infrastructure options" (colocation facilities), while other brokers explicitly identify an IT specialization among several or many market niches in which they operate. Some of these firms serve only a given region, including firms serving the Pacific Northwest; others may serve properties nationally or internationally. Most of these firms have staff that work directly with end users; a few brokers have an internet-only platform and provide all services through virtual assistants.

Advisory service providers combine brokerage services, for both colocation facilities and single-occupant data centers available for lease or sale, with other services. One example is a firm mentioned by one of the interviewed subject matter experts that provides data center strategic advising, design, planning, research, financial advising, and financing, all of which services span cloud, colocation and data center assets. They describe themselves as both "market participants and dealmakers… Recent projects include just-in-time (JIT) acquisition, client acquisition, rapid deployment of modular data center solutions, joint ventures, M&A [mergers and acquisitions], cloud, colocation, and data center research."

Design and Construction Services

Contacts agreed that it is "imperative" that facility **design engineers** or **solutions architects** have experience with and a specialization in data centers to adequately address the centers' complex requirements, including large interactive effects among building systems and resiliency. Contacts explained that the design of cooling systems requires expertise in computational fluid dynamics to understand air flow conditions, needs, and cooling-system-specific effects within a data center.

Designers said only specialized professionals can keep up with rapidly changing technologies. Said one designer, "If you are not specialized and focused specifically on the data center, you don't know, for example, that indirect evaporative cooling systems exist or that liquid cooling systems might be most appropriate for various applications." Another designer noted only specialized practitioners can appropriately pursue uncommon technologies and gave the examples of a Kyoto wheel (a type of air-side economizer) and UPS (uninterruptible power supply), for which the team may need to bring in the specialization of the vendor.[41]

Contacts reported there are not submarkets requiring differing specialization; "at the end of the day, a computer is a computer" is how one designer phrased it. This designer was referring to the physics of IT equipment, which needs a large amount of reliable power, much of which transforms into waste heat.

According to an interviewed subject matter expert, most data center design firms work regionally, although there are "a couple" of firms working nationally in data center design. He summarized, "There

[41] The contact specifically mentioned diesel rotary UPS and line interactive UPS, and that drop system UPS work differently than double conversion UPS. The team did not investigate these technologies but report them for their possible relevance to other facets of BPA's data center modeling effort.

are a couple of national [design] players and equipment players, but after that, it's very much a regional business."[42]

Ideally, the engineers are designing for specific IT equipment and thus know the power and cooling needs. The amount of interaction between IT and infrastructure teams varies by size and ownership of facility. The IT and infrastructure teams for large builder-owner-operated facilities (or build-to-suit or wholesale colocation facilities) coordinate closely. "Every time the IT side introduces a slight change, that will affect their power and thermal design," is how one contact phrased it.

Service providers build retail colocation facilities speculatively, and the design teams are unable to coordinate with as-yet-unknown end users. Contacts described two approaches these firms take to "matching" infrastructure capability to the IT needs. One, they typically build to the highest-need case, so that they can serve any interested end user. Two, and increasingly, they build very large buildings (and clusters of buildings) that they partition into data halls (subsets of the space) and build out the data halls as they perceive demand. This enables them to ramp up their build out (spacing out their CapEx) as they acquire customers and to configure differing cooling systems for the differing data halls.

For construction and renovation of all but the largest internal data centers, contacts described little to no coordination between the IT and infrastructure teams. One contact characterized the teams involved in internal data center construction – often the real estate, facilities and IT teams – as typically operating in "completely separate silos…there's kind of a brick wall between them." He continued, "The IT team goes off and buys what it wants and think it needs regardless of what the facilities actually are, or the facilities management team thinks."

According to interviewed subject matter experts, electrical contracting (responsible for taking power from the grid, power conditioning, including voltage regulation and back-up power, and power distribution) comprises about 50-60% of the labor cost of building a data center. Mechanical contracting (responsible for the cooling system) makes up about 25% of the cost. The general contractor, some specialty subcontractors, and some standard subcontractors (such as steel and concrete workers) receive the remaining share. As a subject matter expert summed up, all these fields, including design services, "come together to design, build, permit, and commission data centers. And then there is a maintenance contractor, although some or all of this can be handled in-house."

IT Manufacturers

Contacts described the data center IT equipment manufacturing market as consisting of manufacturers who focus on four equipment types: 1) servers, 2) server processors (a principal server component), 3) networking equipment, and 4) storage equipment. As discussed above (Hyperscale section), data center owners can source IT equipment from OEMs, which sell products under their brand name, or from EDMs that manufacturer equipment designed by (large) service providers.

[42] This expert added, "The big guys [the firms that dominate the hyperscale space] do it themselves." However, the team has learned there appear to be exceptions to every definitive statement one can make about the data center market. We learned across the breadth of our interviews that some "big guys" indeed do everything in-house, with data center construction constituting an essentially separate line of business, while other "big guys" use independent specialists with which they typically have long-term relationships.

According to subject matter experts, the largest data center firms commonly design their own servers and may be working on developing application-specific processors. These firms source their custom servers and other equipment from EDMs.

The IT equipment market is fluid and constantly evolving. As with other market actors, OEM and EDM are more accurately roles than firms. As examples, Intel (an OEM) processors are used in servers branded by other OEMs; EDMs, such as Foxconn, both manufacture unbranded equipment and serve as OEM for branded equipment.

Consolidation is also a key characteristic of the IT equipment market, where larger firms purchase smaller firms.

There is often overlap between the type of IT equipment manufactured by OEMs and EDMs, however, equipment brands generally do not span equipment types (such as servers and storage) apart from some large firms (for example, Cisco, Dell, and HP). The following are some examples of OEMs and brands by equipment type mentioned by interviewed contacts. Note that secondary analysis revealed numerous firms in addition to these top-of-mind mentions by contacts.

- **Servers:** Cisco, Dell, HP, IBM, ASUS, and Lenovo
- **Server processors:** Intel, ARM, AMD, and Nvidia
- **Networking equipment:** Cisco, Juniper Networks, HPE, Dell, and Netgear
- **Storage equipment:** Samsung, Seagate, Hitachi, and NetApp

Infrastructure Manufacturers

Contacts grouped infrastructure manufactures into two categories: HVAC (or cooling equipment) manufactures and power conditioning. Cooling infrastructure includes such equipment as computer room air conditioning (CRAC) units, computer room air handling (CRAH) units, air and water-side economizers, and chillers. Prominent cooling equipment manufactures identified through secondary research include 3M, 4Energy, Alfa Laval, Daikin, and Eaton.[43]

Power conditioning equipment can include UPS,[44] power distribution units (PDUs), generators, rack power distribution units, power supply units (PSUs), and electric switch gear. Prominent power conditioning equipment manufacturers identified through secondary research include ABB, Black Box Network Service, Schneider Electric, Caterpillar, and Eaton.

Not all infrastructure manufacturing firms work exclusively in the data center market nor do they manufacture one type of infrastructure equipment. Schneider Electric and Eaton, for example, are involved in the design and manufacturing both data center cooling and power conditioning equipment

[43] Source: https://www.technavio.com/report/global-data-center-data-center-cooling-solutions-market

[44] ENERGY STAR describes UPS as follows: In the event of a power failure, UPS provide emergency instantaneous power to critical devices… through energy that is typically stored in a battery. UPS temporarily provide this power to allow for proper equipment shut down (e.g., computers) or for a standby power generator to start up (e.g., at data centers). UPS are able to protect against power surges, voltage drops, and frequency distortions. https://www.energystar.gov/products/office_equipment/uninterruptible_power_supplies

in addition to equipment for other industries. Such firms nonetheless advertise a specialty in data center equipment.

Professional Services Organizations

The **Uptime Institute**, referenced by many of our contacts, has established standards for "the proper design, build and operations of data centers" (infrastructure).[45] From its origins in investigating the infrastructure conditions that led to a loss of network reliability, it developed a system of standards for design, construction, management, and operations relating primarily to power and cooling by which it certifies data centers reliability as corresponding to one of four tiers. Other organizations have developed standards that have not attained the market dominance of Uptime's Tier Standards.[46]

The **Standard Performance Evaluation Corporation** (**SPEC)**, mentioned by one subject matter expert, created the **Server Efficiency Rating Tool (SERT)** which is a tool used for benchmarking the energy efficiency of servers. The ENERGY STAR program supports the SERT tool, designed to be used by server manufactures to better understand the energy usage of their products.

The **Association of Computer Operations Management (AFCOM)** is a network of data center IT infrastructure professionals. The organization includes the Data Center Institute, a think tank focusing on "emerging trends around innovation, technological change, macro-economic shifts and workforce dynamics" in the data center industry.

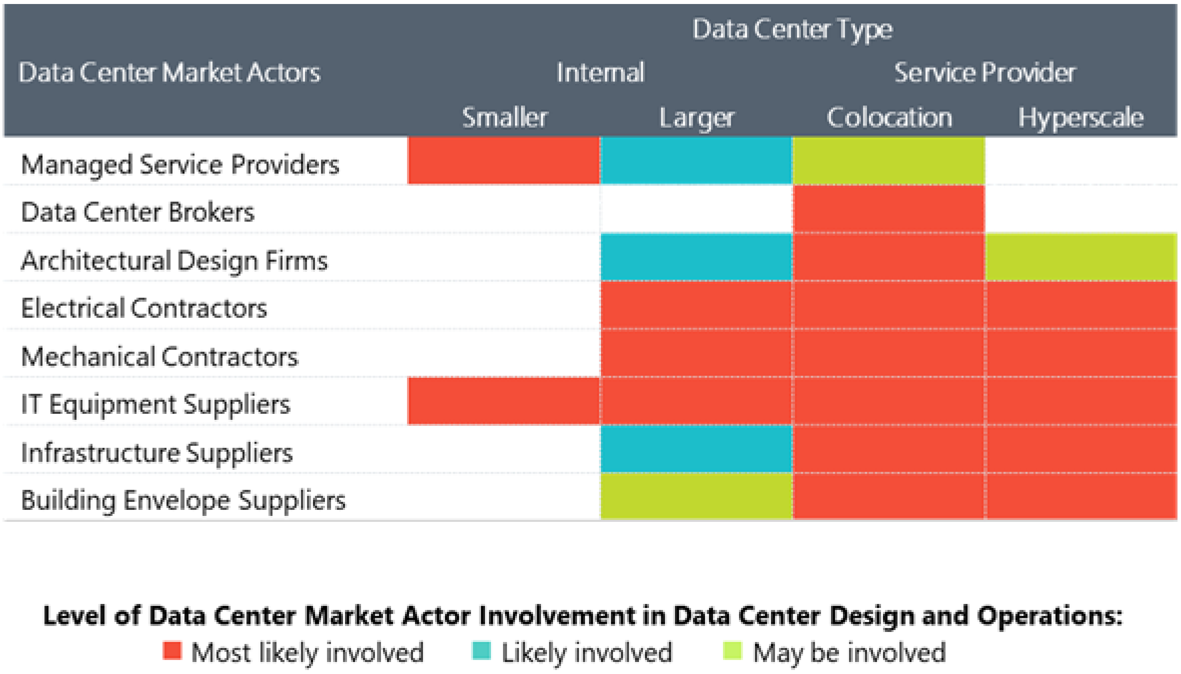
Summary of Market Actor Relationships

Figure 3 provides the team's synthesis of market actor relationships with data centers by data center type, as we understand them from our interviews. Explanations follow the figure.

[45] https://uptimeinstitute.com

[46] http://searchdatacenter.techtarget.com/opinion/Uptime-Institute-tier-classification-system-faces-new-rivals

Figure 3: Relationships between Market Actors and Data Center Types



Source: Research Into Action

To elaborate on the relationships illustrated in the figure:

- Managed service providers
    - Smaller internal data centers frequently use managed service providers.
    - Larger internal data centers typically use managed service providers, with the very largest using in-house staff.
    - A few clients of colocation facilities use managed service providers.
- Data center brokers
    - Colocation clients typically use brokers.
- Architectural design firms
    - Larger internal data centers commonly use design firms, although the smaller among these data centers typically are located in the firm’s existing facilities.
    - Colocation firms are most likely to use design firms, although they may develop standardized designs based on the firms’ work.
    - Hyperscale data centers use design firms, commonly executing standard specifications; some hyperscale data centers use in-house staff.
- Electrical and mechanical contractors

- All types of data centers use electrical and mechanical contractors except smaller internal facilities, which rely primarily on a general contractor, who likely subcontracts with these contractors.

- IT equipment suppliers
  - Smaller internal data centers use local computer suppliers.
  - Larger internal data centers and service provider data centers have long-term relationships with IT equipment suppliers and OEM manufacturers.
  - Hyperscale data centers purchase unbranded equipment directly from EDM manufacturers.
- Infrastructure suppliers
  - Smaller internal data centers use retailers for any thermal equipment needs that exceed the capacities of the buildings' HVAC systems.
  - The largest internal data centers use long-term suppliers.
  - Colocation and hyperscale data centers use long-term suppliers.
- Building envelope companies
  - Smaller internal data centers do not use building envelope companies; they are located within firms' existing facilities, not in purpose-built facilities.
  - Larger internal data centers use building envelope companies when they need to add square footage or new stand-alone structures.
  - Colocation and hyperscale data centers use building envelope companies.

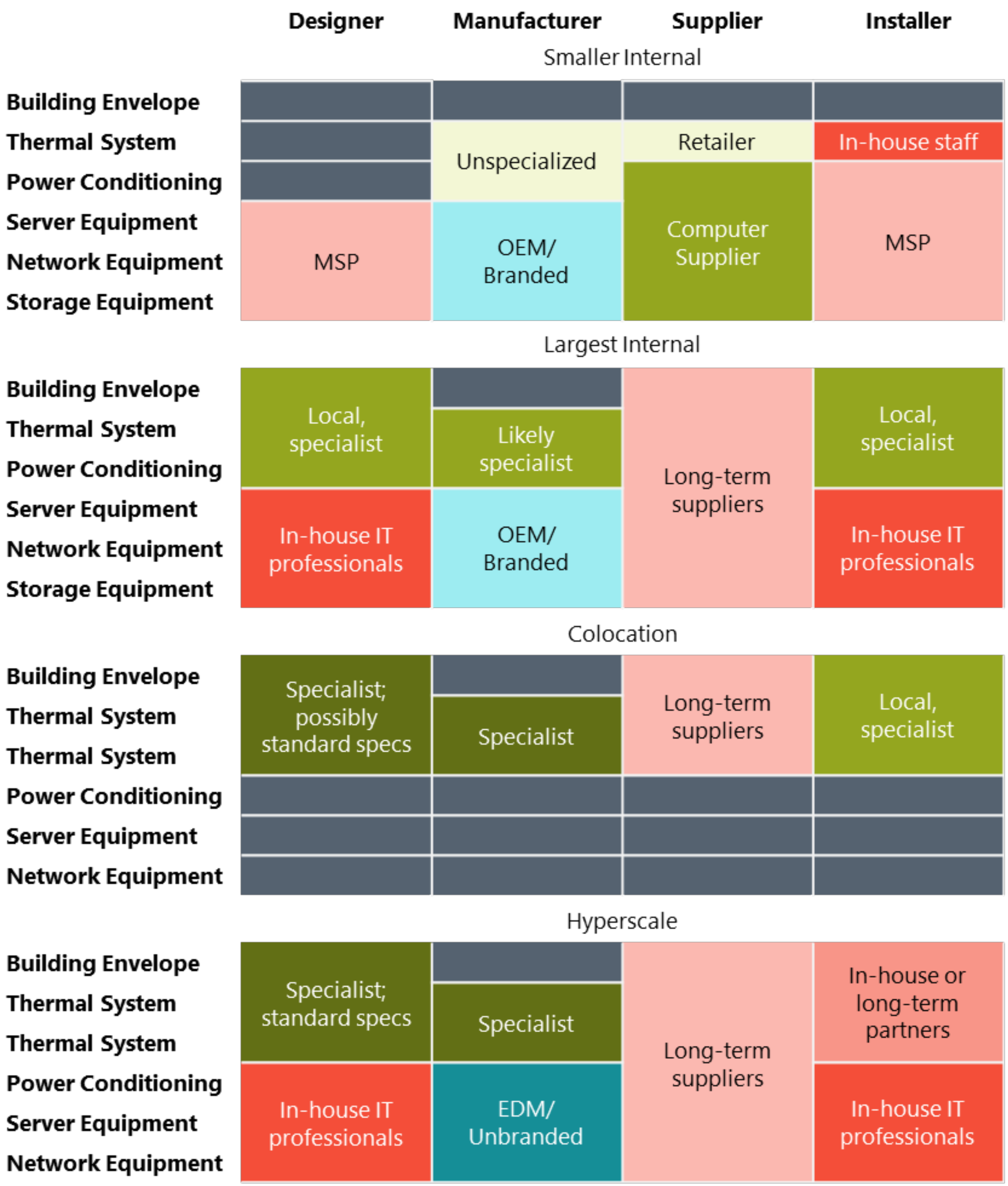
Supply Chain Characteristics

Figure 4 provides a characterization of the supply chain (as shown in the columns: designer, manufacturer, supplier, and installer) for six end uses/equipment types (the rows). The figure provides characterizations for four data center types (shown top to bottom): smaller internal, largest internal, colocation, and hyperscale. The team chose to illustrate the largest internal data center as it is distinct from the more general set of "larger" data centers, whose characterization resembles that of the smaller centers shown in the figure.

The team terms the figure a "characterization," rather than a definitive chain, as it describes widely varying market relationships. The figure illustrates common, but not universal, practices and relationships. The figure illustrates three types of variation across the data centers and market actors:

- Shades of green illustrate variations in the specialization of the market actor providing the service.
- Shades of red illustrate variations in degree of stability of client-market actor relationships.
- Shades of blue illustrate, for IT manufacturer, the variation of branded/unbranded equipment.
- Grey indicates areas that are not applicable to the combination of data center types and supply chain market players.

Figure 4: Data Center Supply Chain Characteristics



Source: Research Into Action

A Note on Cryptocurrency Miners

Cryptocurrency miners are the wild card in the data center market. Cryptocurrency miners engage in hypercomputing functions, with need for connectivity but little need for storage. As one contact said, “I don’t think of them as data centers at all. They are an industrial application that is housed in facilities that look like data centers.”

Cryptocurrency mining is highly energy intensive. The industry is speculative, and 2017 saw currency price increases exceeding the most egregious of past economic speculative bubbles. High currency prices and very low barriers to entry lead many individuals to start mining facilities.

The industry consensus is that miners do not invest in energy efficiency. Although they seek out low energy prices, they are not interested in investing in facilities or operations to minimize energy costs.

Interviewed colocation service providers reported have both current mining clients and startups inquiring about space availability. These service providers considered miners to be risky clients; one colocation contact had already experienced a miner that exited the business. Contacts and secondary sources indicated that miners also build their own facilities.

Although many contacts included cryptocurrency mining among the trends they identified, none speculated on the anticipated size of this data center submarket.

Data Center Design, Operations, and Trends

Ensuring Uptime

The fundamental driver of data center design is reliability. Concerns with what factors lead to failure and how to address those failures led to the formation of the Uptime Institute and its development of Tier Standards for the certification of reliability. IT equipment rarely fails due to component failures; vastly more common is the failure of IT equipment due to overheating, interrupted power, or voltage fluctuations. Thus, the engineering focus on data center cooling and power conditioning equipment. The IT equipment generates copious quantities of waste heat. The more power supplied, the more waste heat, which the facility must remove.

Engineers design the electrical and mechanical systems for reliability and resiliency; both the power supply and cooling systems need to be continuously performing. Previously, the only method the industry had for ensuring reliability was **redundancy**. According to all contacts addressing the topic, the industry is reducing redundancy in infrastructure equipment.

Initially, the industry had **2N** redundancy, which describes a fully redundant, mirrored system (the "N" in the term refers to the number of pieces of equipment, the "2" signifies 2xN or twice N). For example, if the IT devices require four UPS, then a 2N system would have eight UPS. With modularity in facility design, data centers now typically have **N+1** redundancy, where each properly sized module (such as supported four UPS) has one added component.

Contacts mentioned the ongoing trend to reduce infrastructure equipment also reduces capital and operating expenditures. This trend is especially evident among cloud technology providers, who for some recently constructed data centers have eliminated cooling equipment through a complete reliance on outside air (using fans only or fans with evaporative cooling) and eliminated backup power (UPS and generators) by relying on the redundancy within the cloud (N+1 redundancy at the data center, rather than processor, level). Subject matter experts reported that this "no infrastructure" approach is not the norm, although it likely is increasing because "cloud technology providers are always looking for how to reduce capital and operating expenditures by removing some of the redundant components and/or [removing some of] the physical and generative components" when they are able to accomplish their objectives some other way.

Our contacts suggested that the cloud and edge computing environment has enabled large service providers to commonly have N+1 redundancy at the data center level, not the modular level. When a data center has operating challenges, the operator shifts the entire workload to the remaining data centers in the network. Data center owners that previously would require the reliability of a Tier 3 of Tier 4 facility now might seek "a good, solid" Tier 2 design.[47]

Contacts made clear, however, that N+1 redundancy at the data center level is not without risk. Said a hyperscale contact, "It's like a high wire. You have a net, but still, you don't want to fall."

Basic Facility Needs

Designing for resiliency may be the largest part of the challenge, but not the sole part. In the words of one interviewed designer:

> *From the building standpoint, the designer needs to understand the functions of how the facility will operate. That includes understanding how equipment will be brought in and out of the building and the sizing of equipment, having a spatial understanding of how the facility works, understanding how the power distribution works from the rack level back to the source, conducting the mechanical flow and water flow analyses. All this coupled with resiliency of the systems.*

Contacts distinguished between a data center's **white space**, or the floor space dedicated to IT equipment (positioned in racks or cabinets), and its **gray space**, which supports the cooling and power conditioning equipment. According to an interviewed colocation space broker, data centers typically have 50% more space dedicated to infrastructure than computing (a ratio of 60/40 gray to white space).

Design Trends

Contacts frequently mentioned the importance of the flexibility that modular designs provide. Modularity in both IT and infrastructure is a key trend in data center configuration, as discussed previously.[48]

The use of **standardized** or "**precanned**" data center designs now dominates the market for hyperscale and colocation facilities.[49] Standardized designs provide owners with many advantages. They enable owners to replicate data centers quickly, reducing time-to-market, and consistency, reducing the risk of encountering problems. Equally importantly, owners can "line up the supply chain and do long-term deals on equipment," said a colocation services provider. This lowers their costs, speeds the construction, and "doesn't keep our vendors jumping through hoops."

---

[47] The Uptime Institute has established standards for data center infrastructure related to four levels of reliability, termed Tiers. A data center certified at Tier 4 ensures the highest reliability; the Tier 1 certification corresponds with the lowest certified reliability rating. https://uptimeinstitute.com.

[48] See the section "Responses" (referring to "Challenges," both in the Market Characteristics section).

[49] Previously, design was a custom process. "It used to be the design professional would get a project, write specifications, do the design documents, do the calculations and come up with the sizing of the infrastructure equipment, and then submit it for procurement."

According to an interviewed data center design professional,

> *"The data center market has become **commodified** from a design and even from a construction standpoint. The revenues and margins are very small because the effort is not the designers' anymore. This approach also reduces the designer's risk, because the risk is now on the technology company that is providing the design. The only risk to the project comes from site adaptation and implementation."*[50]

A designer described a process whereby the colocation services firm with which his company works contracts separately with design-build firms, engineering firms, and mechanical, electrical, and plumbing (MEP) firms to get the data center built. The designer described the opposing trend in internal data center construction. For internal data centers, his firm is experiencing increasing interaction between the facilities (infrastructure) and IT teams as they strive to optimize their capital and operating expenses. This designer also described the construction of hyperscale data centers as involving close communication between infrastructure and IT teams, when the owner does not use standardized designs.

Even with standardized designs, "the reality is that it is much more complex in terms of what gets built," said a subject matter expert. He continued,

> *There is a desire to be standardized, but companies are very dynamic. There are mergers and acquisitions, and transfers from one location to another.*[51] *If you build all data centers at once, yes, it will be more standardized. But you build everything over time, so you end up with a fair amount of diversity in your data center portfolio.*

Said a designer, "I've been in this business for more than 20 years and no two data centers are alike."

Until the last few years, the trend in IT energy use per unit of computation or storage was rapidly downward, as computing devices became more powerful and their size and power demands shrank. During that time, the associated cooling loads also reduced, commensurate with the reduced waste heat. More recently, however, trends in computing applications, such as GPUs (described above) are necessitating more power per unit (**U**) of rack or cabinet space.[52]

The two competing forces – miniaturization due to technology advances and increased power requirements for some computing applications – are now resulting in growing IT loads per U, referred to as higher **processing densities**. Contacts explained that a few years ago, cabinets commonly drew two to three kilowatts (kW); now, low density cabinets run 3 to 5 kW, medium density run 5 to 10 kW, and high density run 10 or 15 kilowatts or more.[53] High performance computing occurs at even higher densities. Contacts report that higher density cabinets need fewer circuits and PDUs. The appropriate server density

---

[50] As noted elsewhere, location is a driver of data center design.

[51] As described elsewhere, a data center's location affects its design, especially the design of the cooling system and the electrical equipment that connects to the grid.

[52] The industry measures the amount of IT equipment in relationship to the height of the rack or cabinet. The standard unit of measurement is known as the rack/cabinet Unit, referred to as "U," which is 1.75 inches high and full depth (commonly, 32 inches) and usable internal width (fixed at 19 inches). End users can lease from some colocation providers as little as one U of computing resources..

[53] Source: Interviewed contacts and http://searchdatacenter.techtarget.com/answer/How-do-I-estimate-server-power-consumption-per-rack

is a function primarily of the application, however owners can lower their capital costs by increasing cabinet density.

Refresh Cycles, Equipment Lifetimes

A design contact estimated that hyperscale facilities are "always on the leading edge" and so **refresh** (upgrade) their IT and infrastructure equipment typically within 6- to 12-month timeframes. Contacts noted that IT technology performance and efficiency improve so rapidly that replacement at about two years frequently is cost effective.

End users with large internal data centers refresh their equipment around two to three years, depending on the equipment. The contact thought end users with smaller internal data centers typically update them in the three- to five-year timeframe, as follows:

- Servers: 3 to 5 years
- Storage arrays: 4 to 5 years
- Network equipment: 4 to 5 years (diminishing availability of vendor support may influence replacement decisions)
- Infrastructure equipment: replace at manufacturer's recommended useful life

There are laggards in the data center market, of course. Contacts have observed end users and even colocation facilities running 15- to 20-year old infrastructure equipment.

Innovations

Contacts described **deployment** of rapidly **innovating technology**. They gave examples of liquid cooling systems and capture of waste heat for other, non-data center applications (commercial and industrial) that are located on the same campus. The National Renewable Energy Laboratory (NREL), for example, has a high-performance data center (super-computer) on its campus and is exploring using waste heat for nearby buildings.

The building shell can also contribute to efficiency. One designer reported using smaller fans and passive technologies such as chimneys to remove the heat. Many contacts reported common use of outside air and economizers.

Contacts identified the following **airflow management** activities they commonly employ to minimize energy use:

- Using outside air, **evaporative cooling** (direct or indirect), "free cooling," - a technology common in homes but relatively new to data centers
- **Hot and cold aisle containment**
- Eliminating hot spots
- Reducing mechanical cooling equipment and/or its operating times

- Deploying rack-specific **cable management**, because poor cable placement can significantly impede air flow through the rack
- Right-sizing equipment, which can be challenging for colocation services providers as they build without knowledge of the future IT loads
- Using modular designs to match the airflow design to the IT equipment[54]
- Using modular designs to match computing power to computing needs

Contacts discussed hot and cold aisle containment, where the front of the rack is served cool air and the back of the rack exhales hot air, and the need to prevent leakage between these spaces. One colocation services contact reported, "We use a hot aisle containment model in all our facilities, whether our clients like it or not. We believe we may be the only US-based company that does that."

Older designs, yet still advocated by some efficiency experts, have **raised floors** with the infrastructure positioned beneath the cabinets. This design offers easy access to deliver and remove air, although some contacts noted that the operator then needs to be vigilant to prevent "hot spots." There has been a movement away from raised floors to "**on-slab containment**," which is the hot and cold aisle containment.

Because of differences in computing equipment, the racks or cabinets that hold the equipment can have differing amounts of power needed and waste heat generated. Containment designs can address these differences in cooling needs.

Innovative manufacturers incorporate monitoring and control software into cooling and power conditioning equipment to ensure efficient operation and that no circuits are overloaded. One contact explained that UPS operate on an efficiency curve dependent on load; the monitoring enables the UPS to share loads across servers and thereby operate most efficiently.

Other contacts describe the importance of reducing the transformational losses associated with UPS systems, which convert AC power to DC for storage, then back to AC for use.[55] A hyperscale contact described reducing transformational losses by providing rack-level DC power.[56] A designer reported, "We use high efficiency DC motors, linear motors, and high efficiency compressors."

In the IT arena, contacts described the following common efficiency actions:

- Server **virtualization**
- Server **decommissioning**¸ which removes what the industry terms **comatose**, **zombie**, or **ghost** servers

---

[54] This contact noted that there is lower limit to modularity, as small spaces would need variable airflow and variable compressors to operate efficiently. He also said that "there is a bit of debate as to whether modularity gains you much for UPS."

[55] One contact mentioned that Mitsubishi now offers a UPS with 1% rectifier and inverter losses, compared to the 3% current industry standard, "but it costs a lot."

[56] For example, see: http://www.datacenterdynamics.com/content-tracks/design-build/dc-distribution-is-not-just-for-the-giants/95037.fullarticle

- Replacing spinning disk storage devices with **flash** or **solid-state drive (SSD) arrays**, which are a bit more expensive than spinning disks, but have much lower latency
- Using modular designs to match infrastructure equipment to computing needs

The industry's most common efficiency metric, PUE (discussed in the next section) does not vary with server utilization; however, the energy to both power and cool the equipment is wasted when the server is largely idle, as measured by **utilization** rates. Data center operators can remedy low server utilization by virtualization and decommissioning, as well as migration to the public cloud. The hyperscale contact estimated that "for hyperscale in general, I'd say conservatively that utilization rates are north of 70%." The contact also reported that average utilization rates vary by owner-operator.

An end user noted that his aggressive virtualization and use of only flash storage have resulted in measurable reductions in facility energy use. "And that's key for us – measurable. Our vendors know this." Similarly, an interviewed managed service provider reported that his firm aggressively pursues server virtualization. "I used to buy five physical boxes, now I buy one box and put five virtual servers on it." Eliminating servers also eliminates battery backup devices, which are "horribly inefficient." Air conditioning loads go down as well. "I've cut power consumption by 80%."

Another contact noted that "the cloud is really nothing but virtualized servers, virtualized storage… all of which can be allocated to a customer in seconds instead of the hours or days to build out from physical hardware."

An interviewed subject matter expert believes that increasing utilization rates among all data center assets is the greatest energy efficiency opportunity. The team did not ask other contacts to weigh in on this contention to corroborate or refute it, but forecasts of market growth suggest that the market share of cloud and hyperscale computing will continue to grow; these assets, as the expert noted, have high utilization rates.

Small internal data centers lack standard efficiency measures, much less innovative ones. The physical space in which the data center is located seldom supports appropriate air flow management. The businesses typically lack a point-person with responsibility for all computing assets; no one person may be aware of all assets. The businesses do not track assets, nor asset utilization, and thus miss opportunities to consolidate loads onto fewer servers and remove unneeded assets.

Some contacts mentioned that the industry is looking again at **fuel cells** for its power requirements, after eBay's initial foray into the technology in 2013. *The Seattle Times* reported in September 2017 that "Microsoft makes a 'crazy' bet on fuel cells to feed power-hunger data centers."[57] Natural gas-powered fuel cells, if proven viable at scale, would enable data centers to eliminate electricity use. The article explained, "By generating electricity close by — literally on top of the computing hardware — Microsoft's new design eliminates the inefficiency of producing electricity at a distant power plant and transporting it long distances to data centers." Our contacts explained that most of the power failures result from disruptions to the electricity transmission and distribution systems; natural gas fuel cells would eliminate that vulnerability.

[57] https://www.seattletimes.com/business/microsoft/microsoft-makes-a-crazy-bet-on-fuel-cells-to-feed-power-hungry-data-centers/

Energy Efficiency in Data Centers

Power Usage Effectiveness (PUE)

The industry most commonly uses a single metric to convey data center efficiency, that of power usage effectiveness, or **PUE**. PUE is the ratio of total data center energy consumption to IT equipment energy consumption. PUEs less than 2.0 indicate that IT loads use more than half of total facility energy use. Similarly, a PUE of 1.2, for example, indicates that for every kilowatt consumed by the IT equipment, the facility consumes 20 watts to provide power conditioning, cooling, and all other non-IT services. Contacts agreed that cooling efficiencies largely drive PUE.

Contacts provided somewhat varying estimates of PUEs for the data centers with which they are familiar but agreed that the clear majority of computing occurs at PUEs less than 2.0. Five years ago, or so, PUEs commonly were 2.0 or higher and some older, primarily internal, data centers may still be in operation with PUEs of 2.5 or above.

One contact mentioned, “The average PUE that I see is probably 1.5.” Another contact mentioned “a good PUE would be 1.5. If you are approaching 2.0, you need to make improvements.”

Large data centers are trending toward the most efficient platforms they can design or acquire. As a **profit center** for service providers (in contrast to end users, for which data centers are primarily a cost center), large data centers continually seek to reduce their costs. There is a lower limit to that of 1.0 PUE. A hyperscale contact offered, “I think that a PUE of 1.2 at the real state-of-the-art facilities is good, but lower than that is possible… I think Google’s average is in the 1.1 range.”

Contacts clarified that they are describing the current vintage of data centers. Even the largest service providers have facilities that are 10 to 20 years old. The service providers regularly update these facilities and their equipment, but older data center PUEs remain higher than newer ones. However, at the rapid pace by which the data center market is expanding, with each year older data centers comprise a smaller proportion of the total.

Data center PUE varies over time, even within a single day, as IT loads are added or removed, short-term operational changes (such as swapping in and out equipment) occurs, and both external temperatures and humidity vary. According to one designer with a representative answer, “My data centers today are 1.3, and peak as much as 1.7.”

Contacts described that retail colocation facilities typically have higher PUEs than builder-owner-operator. Like the latter, colocation facilities are also profit centers and to be competitive in their markets they need to offer their tenants low operating costs. However, as contacts described, the colocation service providers need their facilities to be able to accommodate unknown users, and thus need to be able to meet the needs of their most demanding potential customer, which one contact described as “building to the maximum common denominator.” In contrast, new wholesale colocation facilities have PUEs comparable to those of builder-owner-operator facilities.

At the opposite end of the spectrum, smaller internal data centers have higher PUEs, perhaps greater than 3.0, due primarily to the inability to optimize the air flow through the space. Often these data centers are in the interior of the buildings, with staff assigned to the perimeter. As one designer phrased it, “Hands down, physical location dictates efficiency.”

Contacts described that many factors influence PUE, including: age and location of data center, IT and infrastructure design, IT and infrastructure equipment selection, proportion of space in use,[58] monitoring of system conditions, and operations. Contacts described that a **holistic**, purpose-built design achieved through collaboration between IT and infrastructure teams contributes the most to energy efficiency. The efficiencies of individual pieces of equipment contribute to, but do not drive, overall facility energy efficiency. "It's different to talk about component efficiency than about the efficiency of the entire facility," said a designer.

In discussing PUEs, one contact noted, "You can do things that drastically improve the IT load efficiency that makes your PUE worse [because IT load is the metric's denominator]. So, it's a tough metric in the absence of understanding how much the servers are consuming." Similarly, a subject matter expert said:

> *Efficiency is not just PUE – we need to correct that notion. The most important thing we can do for efficiency is to measure it in the right way and that's energy use per computation and then to drive that down by making utilization as high as possible. Google, Facebook, and Microsoft are good at driving up utilization. That to me is a more important and fundamental indicator and driver of efficiency in data centers. The higher utilization, the lower emissions and more computations per energy used.*

Contacts also noted that some data center operators manipulate PUE reports, for example by not counting infrastructure loads housed in separate buildings from IT loads. To compare PUEs across data centers, "You need to understand how the data are being collected and compare over time to understand the data center's performance."

One interviewed end user tracks the PUE of his four data centers daily, as does one interviewed colocation services provider. Other end user and colocation contacts reported tracking PUE monthly and still other contacts reported that their organizations are preparing to track PUE, but do not currently do so.

Energy Efficiency and Renewable Energy

All contacts emphasized that no data center owner or manager will put energy efficiency or "green" power ahead of reliability, and few will put it ahead of cost considerations. One contact colorfully described the situation:

> *"If you are a data center manager who is trying to build your reputation on saving money or getting a green award, you will get a pat on the back. If you take down your company's home page, you're going to get fired."*

Although second to reliability considerations, large data center owners are highly motivated to invest in energy efficiency because their facilities are huge power consumers. One contact mentioned, "A one percent reduction in energy use can correspond with hundreds of thousands of dollars annually, and millions and millions across a service provider's assets." In contrast, energy efficiency investments considered by end users with smaller internal data centers do not provide the organization with a

---

[58] Colocation data center minimize their costs when, during occupancy ramp-up, they can contain or limit the cooling and power conditioning to the occupied portion of the space. But not all colocation facilities adopt a containerized approach to the build out.

favorable return on investment; the additional first costs and relatively low energy costs translate into paybacks exceeding three years, said contacts that discussed this issue.

A designer reported that “every client in the last 24 months has had a corporate initiative driving them towards a certain PUE or certain level of energy efficiency performance.” He qualified this statement by saying, “There always has to be a business case for energy efficiency. That’s one of the things we preach. Efficiency for efficiency sake is not a smart business choice.”

Another designer made a similar point about the business case:

> *“Keep in mind that servers are not a 10- or 20- year investment. They are two- to three-year investments. So, they are really like a commodity. You are not looking to pay a 10% uplift to get a certification that’s not really getting you anything in terms of server performance.”*

Designers stated that energy efficiency must reduce clients’ **total cost of ownership**, a metric they provide their clients. In some cases, “the efficient solution requires more floor space,” or has other unintended consequences. The total cost of ownership estimate enables an economically rational design.

Managed service providers reported they may never explicitly propose high-efficiency equipment. However, they encourage their clients to increase virtualization, and note that current IT equipment is more efficient than older equipment. Thus, energy efficiency “is a bit inherent in the process. It’s going to happen whether we propose it or not.”

Contacts thought that clients with smaller internal data centers are often unwilling to pay a premium for energy efficient equipment and that such investments rarely make financial sense given the small scale of their operations. According to one designer, small-scale still can be quite large. “Anything under 1 MW won’t have much client interest in “green” facilities.”

West coast clients, according to one contact, are “always looking for some type of green component.” The hydro power provides these clients with a “green story.” With that exception, being green “is a secondary consideration at best” for all but the largest data centers, according to these contacts.

One designer explained that cooling system upgrades provide a significant efficiency opportunity. However, “in general, you don’t change cooling technologies in a data center unless you are willing to shut down part of the data center to do that. It’s just hard to do. Very difficult. Not impossible, but a challenge.”

Contacts explained that the largest companies commonly pursue a brand image of responsible corporate management and these companies seek renewable energy. These providers may rank access to **renewable energy** high among their criteria for data center location and are willing to pay a small cost premium, if necessary.

Colocation providers noted that potential clients commonly include renewable energy in their “wish list,” but do not base decisions on it. This group commonly is quite pleased simply to locate in the service territory of a utility that has some portion of renewable energy in its fuel mix. Contacts characterized the clients’ seeking of renewable energy as “lots of talk.”

ENERGY STAR®

Subject matter experts and designers expressed the view that **ENERGY STAR** certification of servers and UPS was not an important energy efficiency consideration. The interviewed experts reported that most servers on the market exceed the ENERGY STAR criterion (by 50%, according to one contact), although they thought the designation could once again be important were the criterion raised.[59]

Similarly, they did not think that ENERGY STAR UPS provide "significantly better" performance than nonqualifying UPS.[60] Said one designer, "We don't have discussions about ENERGY STAR UPS because our efficiencies are much higher," a practice echoed by another designer. The two other interviewed designers reported that the energy savings from ENERGY STAR UPS do not warrant the cost premiums. An interviewed end user said, "Our UPS are industrial strength. We don't look for ENERGY STAR, we look for what works for our facility."

About half of the contacts said they recommend ENERGY STAR servers and UPS as a matter of course, to ensure higher efficiency choices, but also said that the final decision always rests with the client. As described in a different context above, contacts also pointed out that efficient servers are more important than efficient UPS, as a reduction in server energy saves both server and infrastructure energy.

One designer estimated that perhaps 80-90% of hyperscale servers are ENERGY STAR, perhaps 40-50% of colocation servers, and perhaps 50-60% of internal data center servers (a figure driven by the larger data centers). They also estimated that perhaps 40% of the data centers with which they have been involved have used ENERGY STAR UPS. They qualified that these numbers were "best guesses" and that figures were difficult to estimate.

One contact noted that their firm had designed a data center that achieved Platinum LEED certification, and all relevant equipment was ENERGY STAR. They also noted that this effort seemed strange as "few people ever enter a data center" and so few would ever learn of the certification.

Market Size and Expected Growth

Market Size

Multiple factors made it difficult for the team to develop an estimate of current market size, including the proprietary nature of market data in this highly competitive industry, secondary sources that report figures globally but not nationally, or that include spending on IT services in their market size estimates rather than reporting spending on data centers exclusively. Interviewed contacts provided estimates that appear to the team to provide their recall of publicly available information, yet details such as whether the statistic refers to the entire data center market or exclusively hyperscale data centers are unclear.

According to statista.com, in 2017 global data center market spending reached $3.5 trillion dollars, including real estate, infrastructure, IT equipment, software, and IT services.[61] CBRE, a leading commercial

---

[59] ENERGY STAR issued its server specification in 2013.

[60] ENERGY STAR issued its UPS specification in 2012.

[61] https://www.statista.com/statistics/314596/total-data-center-systems-worldwide-spending-forecast/

real estate firm, reported 2017 spending on data center acquisition (the land and infrastructure that house computer servers) of about $20 billion.[62] Global IT spending in 2017 added $178 billion dollars.[63] Secondary sources put the US share of the market at one-third or more.[64,65] Sources also note that a significant proportion of data center spending outside of the US serves the US market.

An article by Synergy Research Group stated that total spend on hardware and software from Q4 2016 to Q3 2017 was about $80 billion, split evenly between public and private clouds, with public cloud spending growing faster.[66] The article continued, "In aggregate cloud service markets are now growing over three times more quickly than cloud infrastructure hardware and software," and 2016 spending on services exceeded that on hardware and software for the first time.

A subject matter expert reported that the largest firms are collectively spending ten billion dollars a year constructing hyperscale facilities in the US, consistent with secondary sources.[67] The team needs to emphasize, however, that the publicly available sources consulted for this research do not make clear whether these market size numbers describe market spending on cloud services or on hyperscale construction.

Growth

Contacts and secondary sources anticipate annual growth rates exceeding 10% for all sizes of service provider data center assets.[68] Although contacts and secondary sources predict growth for all data center sizes, not all size-owner types are growing or growing rapidly (Table 4).

Industry forecasts predict comparatively little growth in high-end internal data centers, those greater than 20,000 square feet. Forecasts indicate a decline on all other internal data center assets. One subject matter expert estimated that the market for smaller internal data centers has contracted more than 10% for the last year or two, due to migration to the cloud and colocation facilities.

In contrast to internal data centers, secondary sources anticipate the growth in the global hyperscale market to exceed 24% per year. Contacts emphasized that "everything" is getting larger; hyperscale and colocation facilities are getting larger and owners continue to build them at a rapid pace. Growth in smaller service-provider data centers is accelerating as firms invest in edge data centers (which range in size from micro to midsized) that locate computing capacity closer to users and IoT data generation. As one designer anticipated, "It's analogous to the movement from mainframes to the distributed

---

[62] https://www.cbre.us/about/media-center/data-centers-2018-outlook

[63] https://www.statista.com/statistics/314596/total-data-center-systems-worldwide-spending-forecast/

[64] See, for example, this source for an estimate of one-third: http://www.ironpaper.com/webintel/articles/it-market-statistics-and-trends/

[65] See this source for an estimate of 40%: https://www.srgresearch.com/articles/hyperscale-data-center-count-approaches-400-mark-us-still-dominates.

[66] https://www.srgresearch.com/articles/2017-review-shows-180-billion-cloud-market-growing-24-annually

[67] The US comprises over 40% of a global hyperscale market that researchers expect to grow from $25 billion in 2017 to $80 billion in 2022. https://www.marketsandmarkets.com/PressReleases/hyperscale-data-center.asp; https://www.srgresearch.com/articles/hyperscale-data-center-count-approaches-400-mark-us-still-dominates

[68] One "outlier" source provided an estimate of a compound annual growth rate of 6.3% through 2021. https://www.prnewswire.com/news-releases/data-center-construction-market-in-the-us-to-grow-at-a-cagr-of-634-by-2021---tax-incentives-for-data-centers-and-a-reduction-in-electricity-costs-to-drive-growth---research-and-markets-300456668.html

architecture of personal computers. We've gotten centralized again with cloud computing and now I think we will see the pendulum swing back the other way to a very distributed type of architecture where we will see a few very large data centers but many more very, very small data centers."

Table 4: Expected Market Growth by Data Center Size

Data Center Size	Internal	Service Provider
Server closet	Decreasing	Rapid increase
Server room	Decreasing	Rapid increase
Localized	Decreasing	Rapid increase
Mid-tier	Decreasing	Rapid increase
High-end	Slight growth	Rapid increase
Hyperscale	Not applicable	Rapid increase

Source: Research Into Action

Contacts expected the large internal data center market would continue to grow in the coming years due to increased digitization records and of processes (as discussed above). Some internal data and processes will inevitably move to the cloud, but security and confidentially requirements will prevent a total migration – particularly within healthcare and BFSI market segments. One subject matter expert said, "There is a security concern about not having that data in your control and domain. There are also privacy concerns and legal requirements so that data has to be kept in-house, particularly with healthcare and financial records." Another subject matter expert echoed this sentiment, saying,

> *"There is a transition, but it is a slow transition, and I don't think on-site computing is going to go away soon. It may not be 100% transition because of performance, confidentiality, and cultural reasons where people want to manage their own data because of misperceptions around security and confidentially and reliability."*

An author summarizing the proprietary *IT Spending and Staffing Benchmarks 2017/2018* by Computer Economics reported the study's top-line findings of "essentially flat" IT capital spending as reported by more than 200 surveyed IT organizations.[69] However, the article's author took some exception to the top-line finding, noting that 49% of respondents reported planning to increase their spending on infrastructure, equipment, or major system development and implementation.[70]

Pacific Northwest Considerations

Contacts clearly described the drivers of data center location decisions. Data center builders seek (possibly in decreasing order):

- High connectivity, which is the ability to connect to transcontinental and intercontinental fiber optic trunk lines and to multiple network carriers and cloud service providers,

[69] https://www.networkworld.com/article/3219727/data-center/on-premises-data-center-spending-drops-in-priority.html

[70] Twenty-eight percent said their spending would remain unchanged; 23% percent said their spending would decline.

- Low power costs,
- Low land prices,
- Renewable power,
- Climate considerations, including temperature, humidity, and water availability, which determine cooling system options and thus cooling costs, and
- Policy incentives and costs, such as sales and use tax exemptions, and Enterprise Zones.[71]

Contacts reported that owners considering several locations within a region consider factors such as distance of the site from the transmission line (to lower their power costs) and variations among utility rates.

These factors drive the location of service providers large data centers. Service providers locate smaller data centers (mid-tier and below) and edge data centers close to the markets they intend to serve. End users locate stand-alone data centers (both internal and collocated) either close to their existing facilities or some distance away (hundreds to possibly more than a thousand miles) to ensure reliability and disaster recovery capability.

The Washington State Department of Commerce issued in 2018 its *State of the Data Center Industry.*[72] The study reports that although through 2011 Washington had the largest concentration of data centers in the Pacific Northwest, Oregon now has that distinction. Overall, growth has been "on the low end of the market," with the Seattle market growing about three percent year-over-year, which the authors conclude is essentially no growth.

The authors identify "three probable causes for Washington's lagging growth…: (1) lack of aggressive promotion of the state's data center economy and opportunities compared to other states; (2) historic confusion in the market about Washington's data center incentives, which may no longer be that competitive; and (3) concession of the urban data market to Oregon because the Seattle market is not competitive on the basis of sales tax.

The team did not identify a comparable market analysis for Oregon. The team found reference to a 2017 *Greater Portland Data Center Summit* sponsored by CAPRE Media whose publicity article stated that "350+ senior-level data center real estate and technology infrastructure executives" would be attending.[73] The article quoted an IT manager as saying:

The Portland market will continue to grow based on the new connectivity to Asia, renewable power at reasonable prices, no sales tax, a tech friendly local and state government and a well-trained technical workforce. …The greater Portland area market is approaching critical mass with large enterprises like NetApp and Adobe choosing Portland for dedicated facilities. In addition, leading wholesale providers like Dupont Fabros, Digital Reality Trust, T5 and Infomart are making substantial investments in the area.…The State of Oregon has continued to attract major hyperscale providers such as Facebook,

---

[71] For example, sales taxes add considerable cost to data centers' enormous capital expenditures.

[72] http://www.commerce.wa.gov/wp-content/uploads/2018/01/Commerce-Data-Center-Study-and-appendices-2017.pdf

[73] https://www.capremedia.com/close-infomarts-jim-linkous-capre-greater-portland-data-center-summit-speaker-discusses-rise-west-coast-market

Google, Apple, and Amazon. We expect more large enterprise companies like LinkedIn, Comcast, and NetApp will select Oregon to support their expanding IT Operations. [74]All contacts anticipate growth in the Pacific Northwest, although contacts differed in their assessments of whether the rate of growth would mirror, exceed, or lag that of the nation. Some contacts foresaw "a whole bunch of additional capacity coming on in the next three years," or foresaw "enormous growth for hyperscale and colocations." Other contacts thought the region would lag the country somewhat, noting two reasons. One, the region already has substantial data center capacity built out over the last several years. Two, colocation facilities, which are located close to the populations they serve, might shrink slightly in response to increase cloud migration.[75]

Two contacts also offered differing assessments of the status of the Pacific Northwest as a data center market, with one contact placing the region in the top five nationally (making it a Tier 1), and the other contact describing it as a Tier 2.[76]

All contacts agreed the data center market in the Pacific Northwest is growing, due to its high concentration of fiber optic cables; its cool climate east of the Cascades and low humidity west of the mountains; the cost of electricity, especially among the public utilities; the prevalence of low carbon hydro power; and the tax structure for some locales.

The region is "definitely on people's radar when they are looking to expand," said a colocation provider. Several contacts anticipated that the large companies with data centers located in the region will be looking to expand their facilities at existing sites rather than to acquire new sites.

The Pacific Northwest has a high concentration of submarine cables linking the US to Asia and Pacific countries (the APAC region). Oregon has five trans-Pacific cable landing stations along the coast and in Hillsboro. Oregon also has multiple submarine cables to Alaska, two to Canada, and one to Los Angeles, which also as a trans-Pacific landing station. Washington has one trans-Pacific landing station, in Seattle, and submarine cables to Alaska and Los Angeles.[77]

The region also has a dense network of transcontinental fiber optic trunk lines, or long-haul cables, which are analogous to the interstate system for vehicular traffic (Figure 5). Washington has considerable long-haul infrastructure crisscrossing the state; Oregon's infrastructure lies primarily on the coast; Montana and Idaho have major east-west routes connecting the east and west coasts.[78]

According to a secondary source:

> *The latest generation of cables are owned by consortiums of Asian carriers and US corporations (Google, Microsoft and Amazon), which all require access to the submarine*

---

[74] Op cit.

[75] One contact identified three areas within the region home to the most colocation facilities: Greater Portland, Greater Seattle, and Eastern Washington. This contact noted Hillsboro cornered most of the region's colocation growth in the past two years.

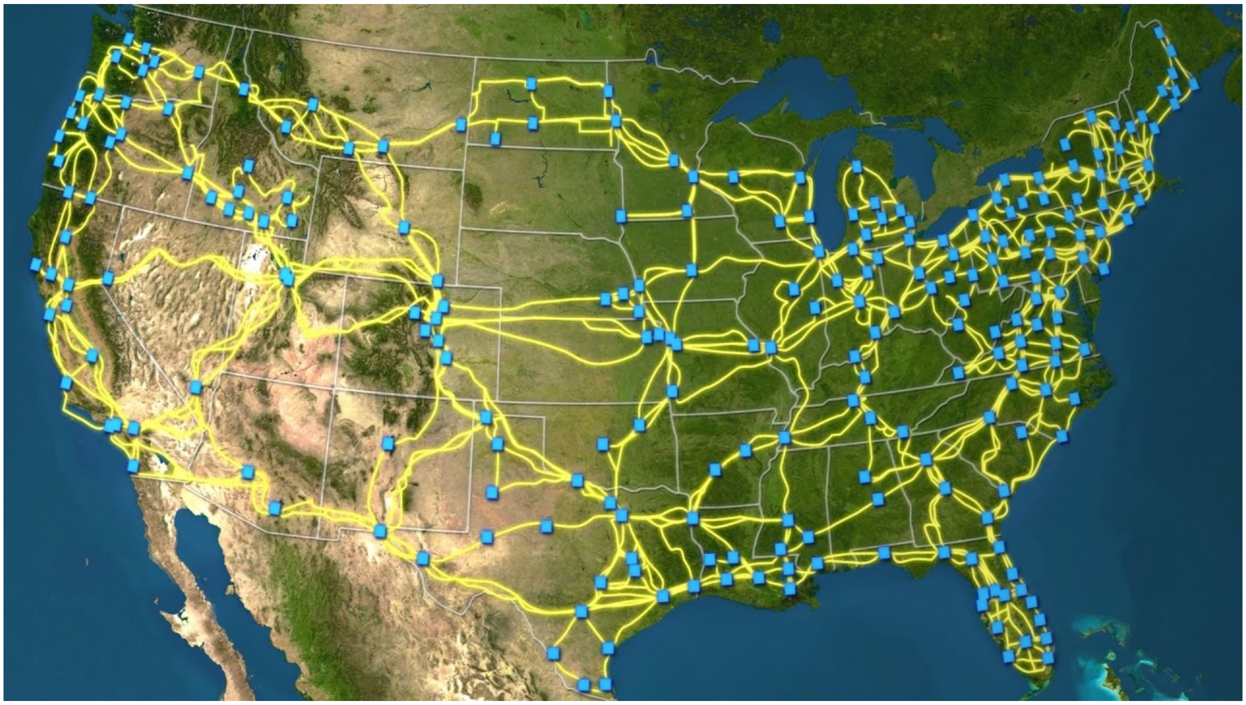
[76] The region is definitely a Tier 2 in terms of population; New York, Los Angeles, and other regions are Tier 1. The team speculates that the tier designation may have a meaning that differs by context, and that the region is a Tier 1 in terms of data center build out, for reasons described in this section.

[77] https://www.submarinecablemap.com/#/

[78] R. Durairajan, P. Barford, J. Sommers and W. Willinger. "InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure", In Proceedings of ACM SIGCOMM, August 2015. http://pages.cs.wisc.edu/~pb/tubes_final.pdf

cables at the landing sites in Hillsboro. This provides Portland with very low latency access to Asia. It is also prompting Asian carriers to build PoP's [point of presence, locations for interconnection between long-haul and local networks] in Hillsboro.... Hillsboro is rapidly becoming the new "Digital Gateway to Asia Pacific."[79]

Figure 5: Map of Long-Haul Fiber Optic Cables in the United States



Source: *https://meridianintl.co/us-fiber-optic-map.html*

Contacts described how the region's cool climate provides advantages for decreasing data center cooling costs, typically the single largest operating cost, by allowing the implementation of various technologies and operational procedures that take advantage of outside air (free cooling). These strategies also reduce or eliminate the need for water-based cooling systems, which use tremendous amounts of water given the large cooling loads.

Another cost-saving aspect for data centers locating in the Pacific Northwest is to take advantage of the regions heavy hydro-electric fuel mix. Contacts mentioned that numerous service providers have built and are building in Eastern Oregon and Washington, where hydro-electric rates can be as low as 2 cents per kWh. Contacts indicated that locations near dams along the Columbia River attract data centers due to lower transmission costs.

Contacts mentioned that Eastern Washington is experiencing a rapid expansion of cryptocurrency mining facilities, a finding corroborated by numerous stories in the popular and business press published during the course of this study. Today, the area is one of the largest cryptocurrency markets in the world.

---

[79] https://www.capremedia.com/close-infomarts-jim-linkous-capre-greater-portland-data-center-summit-speaker-discusses-rise-west-coast-market

According to a recent study in Politico magazine, by 2018 Eastern Washington will account for an estimated 15-30% of the entire worlds bitcoin mining market.[80]

Tax-free equipment procurement and tax breaks for data centers operators are helping to increase growth in the Pacific Northwest data enter market. Contacts suggested that data centers operators who frequently replace equipment factor Oregon's lack of a sales tax into their decisions to locate data centers in the region. Contacts noted that Hillsboro, Oregon is seeing an increase in growth due to no sales tax and the areas proximity to trans-Pacific cables. Areas across the region also offer tax-credits and business incentives for data center owners. For example, Oregon and rural areas of Washington offer enterprise zone programs which gives property tax relief to businesses locating or expanding in certain areas across the state. When the Hillsboro, Oregon qualified as an enterprise zone area, out of state data center providers began to come to this area to take advantage of these tax incentives.

An interviewed broker noted that many of the large service providers have acquired land they have yet to develop; the broker anticipated the companies will build data centers on this land in the next few years. The broker specifically mentioned land purchased in Hillsboro and Central Washington.

Contacts agreed that these factors (high connectivity, low power and land prices, renewable power, climate, policies, and incentives) have helped to make the Pacific Northwest – specifically, Washington and Oregon – one of the leading markets for data centers today and attractive for future growth.

The region's energy codes discourage some data centers from locating in the region, according to a few contacts; another contact thought that some data centers recognized the value of the codes because they drive lower cost cooling solutions. Similarly, the very low population densities east of the Cascades limit both the market for colocation services and the trained labor pool for such a facility. One contact noted that although the region does offer some tax abatement, other parts of the country more aggressively court data centers through tax relief.

One contact spoke of hydro power as a reliability risk, referencing the 2014 crack in BPA's Wanapum Dam (Grant County, OR) that necessitated lowering the water level.[81] This contact also described extended drought as a risk data centers face when choosing to locate in the region.

Risks to Regional Growth Projections

The team does not anticipate any factors that would lead the regional growth in larger internal data centers to lag (or outstrip) that of the nation. Colocation market growth might be limited by unanticipated growth in cloud migration, but contacts' assessments factored in cloud growth. Hyperscale market growth might fall short of expectations due to an increasing trend to locate hyperscale data centers in Asia and other countries. Secondary sources reported that hyperscale markets are growing in other nations but did not draw any inferences for US growth.

Cryptocurrency mining appears to be rapidly expanding in the region. No data center sources drew inferences about mining for data center market growth.

---

[80] https://www.politico.com/magazine/story/2018/03/09/bitcoin-mining-energy-prices-smalltown-feature-217230

[81] https://www.seattletimes.com/seattle-news/lsquoserious-problemrsquo-65-foot-crack-found-in-columbia-river-dam/

Utility Considerations

The team bases this discussion of utility programs on an interview with a utility program representative, comments of other contacts related to utilities and efficiency programs, and secondary research. The interviewed utility program representative team is responsible for Energy Trust of Oregon's New Buildings program implementation. Program offerings through the New Buildings Program included incentives for individual measures such as energy efficient UPS and HVAC components as well as assistance with energy savings analysis. The contact reported working with a range of data center types, from small internal closets to large enterprise data centers. The contact reported specializing in customer engagement, and reported engagement is often difficult due to confidentiality issues.

Utilities often service colocation data centers, and large data centers generally, at their industrial rates. Colocation hosts may double or triple this rate with their markup for power conditioning, which includes the purchase and operation of all the voltage regulation and power reliability equipment.

Contacts described that most clients of hosted facilities, and certainly the large clients, pay for the electricity they use, which the host independently meters. The host passes through to the end user the utility cost at a markup to cover the facility's power conditioning costs. The lease contract specifies the terms for electricity purchases. For smaller clients, the host may bundle the electricity service into the lease, with a certain number of kilowatt hours and price per kilowatt hours specified. For the very smallest clients, the host may simply include electricity in the square footage, rack, or U price governing the lease.[82]

Data centers consume a significant amount of energy which has led utilities to offer energy efficiency solutions. The DSIRE program database shows 28 data center programs offering one or more of the following: incentives (offered by 21 of the programs), loans/financing, trainings/workshops, and energy analysis.[83] These programs include a few exclusively for data centers; most utilities serve data centers in programs targeted to nonresidential new construction and equipment upgrades.

Utilities offer prescriptive rebates for UPS's, servers, storage devices, routers, fans, PDUs, and HVAC equipment and custom incentives, including performance-based, for data center improvements and new facility construction. Some programs offer support for an integrated approach that may involve energy modeling, load forecasting, engineering support, and educational opportunities for data center owners, operators, and designers.

Contacts mentioned challenges utilities face in offering data center programs. Perhaps most significantly, data center owners do not want involvement from "outsiders." The data center market is highly competitive; all players in the market search for any advantage they can attain over their competitors and do not want to risk exposure of their activities, which might occur during program participation. Program staff and sponsors need to treat their interactions with data centers confidentiality. Contacts reported that some data center owners are open to collaboration and see benefits in participating in programs whereas others do not.

---

[82] As defined in Design Trends, the standard unit of measurement is known as the rack/cabinet Unit, referred to as "U."

[83] DSIRE is the Database of State Incentives for Renewables & Efficiency, considered by the energy efficiency industry to be the definitive source of incentives offered by utilities and governments (states, federal).

One contact described difficulties utilities face in designing cost-effective, attractive programs that can work with diverse types and sizes of data center facilities and businesses. Each data center is unique; any single data center program is unlikely to be effective with each of the various data center customers in a service territory.

- Small data centers use relatively little energy and paybacks from efficiency investments typically exceed three years, which contacts described as the typical investment criterion. Small data centers are cost centers, not profit centers, and business needs drive facility decisions, with energy efficiency at best a secondary consideration, said contacts.
- Large data centers have significant motivation to minimize energy use as energy is the largest share of very high operating costs. These owners have capital and recognize the need to innovate and adopt new strategies for decreasing their power consumption.

Regardless of the size of the data center, utility incentives typically are too small relative to equipment and design costs to influence many of the decisions. Utility programs need to avoid free-ridership and the interviewed program representative reported that determining baselines and claiming energy savings is a challenge. Utilities may question whether designing a program for the largest and most energy-efficient data center facilities is worthwhile.

Utilities are thinking about new strategies and ways to engage data center owners. To improve program participation, a contact reported that some utilities are collaborating directly with designers and other data center professionals to bring in efficiency projects. Utilities need to demonstrate that company and facility information collected through a program will remain confidential and secure. The program representative recommended that utilities develop a straightforward process to negotiate non-disclosure agreements with potential participants.

Several subject matter experts believed that the largest efficiency potential comes from systematic and institutional changes made in data center operations. These contacts recommend utilities to develop strategies and solutions to improve a data center power management and utilization rates. Owners of the smaller internal facilities have a sizeable portion of underutilized and comatose servers (servers that are on and consuming power, but not operating). No one person in these organizations may have complete knowledge of all assets owned.

According to these subject matter experts, companies care most about a data center's reliability and consequently leave all server equipment running. Often companies lack a power management system. Many organizations have loads that vary seasonally. Outside of peak periods, organizations may need minimal capacity and operate their equipment at very low utilization rates. These contacts thought utilities should continue to consider load shifting and other strategies to regulate power in a data centers during off-peak times.

Interviewed managed service providers expressed a different view, however. They reported their clients, which are small to medium size firms – the market served by these contacts, have centralized facilities. These contacts' clients do not have comatose servers among their assets and the managed service provider plans for high utilization rates when advising their clients on replacing equipment.

Typical server utilization rates thus appear to be an open question that BPA might want to investigate.

One contact mentioned that large companies such as Google are looking into arbitraging power rates by shifting their loads among data centers.

Contacts reported that cooling provides the opportunity for the largest efficiency gains, and thought utilities might incentivize conventional evaporative cooling systems, as well as more innovative approaches such as chilled door rack cooling systems. UPS systems consume the most power after chillers, making them appropriate for incentives. Contacts reported that fans and fan controls, and server power management and virtualization also offer opportunities for incentive programs. Utilities might also explore battery storage options.

Utilities need to provide unique program offerings specific to the types of data centers found in the market today. Contacts interviewed suggested that a one-size fits all program for all data centers types is inadequate. Instead, contacts believed that providing custom approaches for each unique type of data center would be beneficial. Some of the feedback provided by contacts on program design suggestions include:

- Contacts thought utilities will be more effective targeting internal data centers and smaller colocation data centers. Incentives to hyperscale data centers risk high free-ridership.
- For small data center types, one contact believed that a midstream design approach would be the most effective way to motivate these hard to reach decision makers. A midstream program may change the current business model as most value-added resellers are not currently selling equipment directly to smaller data centers and managed service providers.
- One contact suggested that developing an upstream program for data centers would be the most beneficial. Contacts reported that IT manufacturers are developing new and innovative data center equipment that are driving efficiencies in the market. Utilities should consider collaborating with manufacturers who can help to push more efficient equipment across diverse types of data centers.[84]
- Utilities could address colocation data centers with support for energy modeling and design strategies, to help owners make more informed decisions.
- Utilities should consider offering trainings and workshops to data center owners and professionals. Educating data center decision makers on best practices and design strategies will be a benefit to both the customer and the utility.

[84] Xcel and Efficiency Vermont are two examples of utilities which offer upstream programs to data centers.

Appendix A: Glossary

Blockchain	Blockchain is a type of computer code designed to provide an open, transparent, distributed ledger of transactions that is secure and unalterable, enabling parties unknown to each other to trust the accuracy of the recorded transaction without an intermediary (such as a bank) acting to assure the transaction's legitimacy. The technology's name derives from the blocks of data that it links into a chain through encryption code.
Branded equipment	Computing equipment sold by an original equipment manufacturer (OEM; see below) under its logo.
Cabinet	A cabinet holds the computing equipment in a data center. A cabinet without sides is a rack. The industry uses the terms rack and cabinet in discussions of quantities of computing equipment and data center floor space. See "U."
CapEx	Capital expenditures.
Cloud computing services	Computing services made available to users on demand via the Internet from a service provider's servers as opposed to being provided from a company's own on-premise servers.
Colocation service providers	Colocation service providers host their client's IT equipment in their facility. Stated differently, colocation service providers rent clients space in their data center facility (a building with complete infrastructure services) but do not provide their clients with IT equipment, except in the uncommon but not rare cases when they do. The putative defining characteristics of these hosting firms are that their facilities host multiple clients (hence the term colocation), that they do not provide IT equipment (the clients install their own), and that they speculatively build their facilities to meet anticipated market demand. However, exceptions abound. For example, the industry distinguishes between "managed colocation," where the service provider is responsible for the IT equipment, and "non-managed colocation," the more common arrangement.
Cryptocurrency	Cryptocurrency is a digital currency for which encryption regulates the generation of units of currency and verifies the transfer of funds, operating independently of a central bank. Cryptocurrency miners in engage in hypercomputing functions using blockchain technology, with need for connectivity but little need for storage.
Digitization	Digitization is the process of capturing analog information in a digital format.
Edge data center	A data center located close to the market served. Edge data centers range in size from micro to midsized.
Embedded data center	Another term for internal data center.
End user	The term used in this report to describe the clients of data center service providers; not a term used by the computing industry.
Enterprise data center	Another term for an internal data center.
Equipment and design manufacturers (EDMs)	Equipment and design manufacturers are firms that manufacture to their clients' specifications and provide their clients with what the industry terms unbranded equipment. EDMs may add value to the product by providing the design specifications followed in manufacturing or adding software or hardware.

Graphical processing unit (GPU)	A GPU is a programmable logic chip (processor) specialized for display functions. The GPU renders images, animations and video for the computer's screen. Graphical processing units are better suited for multiple, simultaneous operations than central processing units (CPUs), which are better suited to perform calculations in sequential order.
Hosted facility	Facilities offered by colocation service providers.
Hyperscale	Hyperscale (also termed mega) data centers are the very largest data centers. IDC, an industry-leader in the provision of current and forecasted estimates of equipment and services supplied to the data center market, defines hyperscale facilities as "up to over 400,000 square feet;" however, they can be entire large campuses comprising many buildings.
Information technology (IT)	Information technology, in this report, refers to the energy-using hardware (computing equipment) that runs software. It comprises servers, server processors, networking equipment, and storage equipment.
Infrastructure	Infrastructure is the basic, underlying framework or fundamental facilities and systems; in this report, infrastructure describes both the building housing the computing equipment and the building systems, including cooling, power conditioning, equipment to support connections with the telecommunications system, and other non-IT energy-using equipment.
Internal data center	Internal data centers are those located within a business's facilities. All but the largest of internal data centers are not purpose-built facilities like those constructed by service providers but rather are located within the organizations' existing structures.
International Data Corporation (IDC)	IDC is a market research firm that specializes in information technology, telecommunications, and consumer technology markets. It is the industry-leader in the provision of current and forecasted estimates of equipment and services supplied to the data center market.
Internet of Things (IoT)	The Internet of Things refers to the interconnection via the internet of computing devices embedded in everyday objects, enabling them to send and/or receive data.
Latency	Latency refers to data transfer speed.
Managed service providers	The industry uses the term managed service providers to describe firms that take responsibility for a client's IT assets, whether in-house, co-located, or cloud-based.
Micro data center	The smallest of data centers owned by service providers, under 100 square feet.
Original equipment manufacturers (OEMs)	A company that sells computing equipment that does not have its own manufacturing facilities; OEMs buy computing products and resell them.
Power usage effectiveness (PUE)	Power usage effectiveness is the ratio of total data center energy consumption to the energy consumption of the IT equipment.
Rack	A rack holds the computing equipment in a data center. An enclosed rack is a cabinet. The industry uses the terms rack and cabinet in discussions of quantities of computing equipment and data center floor space. See "U."
Redundancy	Redundancy refers to the duplication of equipment or functionality. Initially, the industry had 2N redundancy, which describes a fully redundant, mirrored system (the "N" in the term refers to the number of pieces of equipment, the "2" signifies 2xN or twice N). For example, if the IT devices require four UPS, then a 2N system would have eight UPS. With modularity

	in facility design, data centers now typically have N+1 redundancy.
Reliability	Reliability describes the proportion of time the IT equipment is functioning; the proportion is also described by the term uptime.
Retail colocation services	Retail colocation services describe the situation where one facility owned by a colocation host houses the IT equipment of multiple businesses.
Server density	Server density describes the amount of energy needed to power the server; higher densities use more electricity per square foot of IT-occupied data center space.
Service providers	Service providers are organizations whose role in the market is that of supplier of data center services. However, that is not necessarily these organizations' sole role. They may also be a consumer of data center services provided by other service providers.
U	U is a measurement of rack or cabinet space. The industry measures the amount of IT equipment in relationship to the height of the rack or cabinet. The standard unit of measurement is known as the rack/cabinet Unit, referred to as "U," which is 1.75 inches high and full depth (commonly, 32 inches) and usable internal width (fixed at 19 inches). End users can lease from some colocation providers as little as one U of computing resources.
Unbranded equipment	The equipment provided by equipment and design manufacturers (EDMs).
Uninterruptible power supplies (UPS)	UPS is a device that provides battery backup to critical equipment when the electrical power fails or drops to an unacceptable voltage level.
Utilization rate	Utilization rate describes the proportion of the computer's processing resources used during a given period. Low utilization rates describe computing capacity in excess of demands.
Wholesale colocation services	Wholesale colocation services describe the situation where one facility owned by a colocation host houses a single company's IT equipment.

Appendix B. Interview Strategy and Methodology

The research team collected 81 web total sources. The team then consolidated and reviewed these studies, publications, and findings. The team used this research to develop a presentation on data centers for Bonneville Power Administration. In addition, information collected supported the development of in-depth interview guides explained in greater detail below.

The team collected and consolidated phone numbers and email addresses for potential market actors through company websites and other online searches. The team also used Hoovers, a business research database, when necessary to collect contact information on other potential interviewees.

Research Questions

Table 5 provides the research questions identified by BPA prior to beginning this research. The team identified several additional research questions during secondary research for Task 1 and Task 3 as indicated by an asterisk in the table below. (Table 1 provides a summary of the elements listed here.) The team used these research questions to guide the creation of interview questions for subject matter experts and market actors.

Table 5: Market Aspects and Related Research Questions

Market Aspect	Research Question(s)
Market trends	What are market trends in computing? In IT? In facilities?
Data center types	What factors affect use of Internal versus service provider data centers?
Market actors	What is the supply chain for IT equipment that goes into data centers?
	What is the supply chain for non-IT electrical equipment (lighting, uninterruptible power supplies, HVAC, etc.) that goes into data centers?
	Do HVAC systems that are installed in data centers get distributed and installed by conventional HVAC market actors, or is there a niche data center HVAC market?
	Are there niche services certain market actors provide? If so, what are these services?
Market size, capacity, and expected growth	Do ‘cloud computing’ adoption rates vary by data center or business type? Are efficiency gains greater for certain business sizes?
	What is the forecasted growth of enterprise data centers in the Pacific Northwest? What is driving or hindering that growth?
	What is the expected growth in colocation from those moving from enterprise and small internal?*
Data center operations	What is the approximate turnover rate for IT equipment? How does it vary by equipment type?*
Decision-making, including energy efficiency and	How do efficient practices and the adoption of efficient equipment vary by data center type (server room, enterprise, etc.)? Are enterprise data centers already highly efficient?

renewable energy	What are the reliability requirements for data centers? How do these vary by business type or size? How do reliability requirements affect efficiency?
	What is the market adopt/rate of use of ENERGY STAR servers? What factors influence customers' decision to purchase ENERGY STAR IT equipment?
	Do ENERGYSTAR rated products, or high-efficient IT equipment provide any downsides or side-effects that may make them unappealing to some market actors or in certain applications?
Efficiency metrics	What PUE is current (average) PUE for data centers? How does PUE differ for new vs. older data centers?*
	What are the trends in server utilization?*
Data center design	What are the trends in new data center construction and design?
Pacific Northwest considerations	Approximately how many data centers are in the Pacific Northwest?*
	What proportion of computing is used by Pacific Northwest?*
Role of utility programs	What is the current state of utility programs in the Pacific Northwest that are focused on data centers?
	Where are utility programs focusing their efforts for data centers?

** Additional research questions identified through the team's secondary research for Task 1 and Task 2.*

Selection of Contacts

Through interviews with subject matter experts and secondary research, the team identified eight types of market actors to interview, as well as a utility program representative and a representative of Rocky Mountain Research Institute, to discuss blockchain technology. Table 2 identifies the types of contacts interviewed.

Within the market actor types, the team used the following approach to identify firms and individuals to interview. The team had received at the outset of the projects suggestions of a few firms, some with named individuals, from the study's stakeholders and advisors. Subject matter experts similarly offered a few suggestions during the interviews.

The team conducted internet research to identify firms within the market actor categories that appeared to be industry leaders and/or served the regional market. The team brainstormed the types of positions within each market actor type that seemed appropriate to the study and agtain conducted internet research on those firms to identify individuals to contact. When the team lacked identified individuals, team members cold-called into the firms to identify appropriate contacts.

Appendix C provides the interview guides for each of the market actor types.

Appendix C: Instruments

Subject Matter Experts In-Depth Interview Guide

Introduction

Thank you for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

General Questions

Q1. Can you start by giving me a description of your role within your organization and your experience with data centers?

Q2. [*If unclear*:] Is your work focused on one specific area of the data center market or do you have experience with a broad range of data center or equipment types?]

Q3. What types of firms or organizations focused on data centers do you work with? We are not looking for you to identify clients, but rather the network of service providers you engage with.

Defining Data Center Submarkets

Q4. First, we'd like to discuss variation among data centers and how they might be categorized into homogeneous, yet distinct submarkets with respect to the types and characteristics of IT and infrastructure equipment. Based on business type, ownership, and equipment we believe that four buckets do a pretty good job. I'd like to get your feedback on our thinking and ask you to consider these four data center types when responding to our remaining questions. Our preliminary groups are:

- **Small Internal** – embedded server closets and server rooms typically 100-1000 sqft., which lack or have minimal dedicated power and cooling systems,
- **Enterprise** – medium to large-scale embedded data centers dedicated to a single business or customer (such as a hospital or university),
- **Colocations or Service Providers** – large-scale data centers serving multiple businesses who rent space within a single facility, and
- **Hyperscale** – high-end and the largest data centers in the market, which host larger collocation and cloud service providers.

In your opinion, are these appropriate submarkets to capture large patterns in baseline equipment, equipment sales, and equipment characteristics?

[If no:] What submarkets would you consider to reflect the key differences among data centers?

*[**Probe for both Yes and No responses:** Are there variations or major divisions within these four submarkets that you believe are important to distinguish? For example, ownership, business types, technology, branded/unbranded equipment, etc.?]*

Q5. How does the market for small internal data centers compare to the market for general office IT equipment? Would you say small internal data centers are roughly similar to, and trend with, the overall market for office IT equipment?

Q6. Are there any data center types where there is a lack of available information or where we will have difficulties in obtaining energy related information? For example, hyperscale data centers might be too proprietary to gain meaningful insights.

Market Actors, Equipment, and Practices

For the following questions, please think about the data center submarkets we discussed and please differentiate your responses by data center submarket when applicable.

IT Equipment

Q7. Let's talk about the IT equipment. What types of individuals and organizations are involved in the IT design, acquisition, and installation, and how are they involved in the process, from conception and design through installation and any final testing? And how does that vary by submarket, if at all? *[**If needed:** IT equipment includes, servers, storage devices, and networking equipment.]*

Q8. What equipment, design, or operational aspects would you say have the largest impact on IT energy use?

Q9. And thinking about those aspects, which types of organizations involved have the most influence on energy use?

Q10. What factors influence data center owners, IT, or operational mangers to purchase energy efficient equipment or implement energy savings operational practices? And how does that vary by submarket, if at all?

Q11. What factors limit the incorporation of energy saving IT equipment or IT related operations? And how does that vary by submarket, if at all?

Q12. When a data center is being constructed, expanded, or renovated, how integrated are decisions about the IT equipment? For example, does the owner or owner's representative contract separately for the design, acquisition, and installation, or is it typically done under the direction of one or two lead firms? And does that vary by submarket?

Q13. What would you say are the adoption rates for ENERGY STAR servers for each data center submarket?

Q14. What proportion of data centers pursue energy savings through server virtualization? And does that vary by submarket?

Q15. Our understanding is that larger data centers sometimes work directly with manufactures to produce customized server equipment. In your opinion, how does the energy efficiency of this customized server equipment compare to ENERGY STAR servers?

Q16. What would you say are key IT energy saving operational practices? For example, virtualization, decommissioning, purchasing energy efficient servers and networking equipment, or power management? And does that vary by submarket?

Q17. Who are a few of the most prominent data center IT equipment suppliers and contractors in the northwest, or nationally if there are not regional distinctions?

Infrastructure Equipment

Now I have some questions about data center building infrastructure equipment, which includes both HVAC equipment and electrical systems

Q18. What types of individuals and organizations are involved in data center HVAC equipment and electrical systems and how are they involved in the process, from conception and design through installation and commissioning? And does that vary by submarket?

Q19. Would you say data center specialization is required or beneficial for mechanical, electrical, and plumbing contractors who are involved in data center projects? If so, why and does it depend on data center submarket?

Q20. Are there any niche markets or niche services offered by any of the contractors we just discussed that are specific to each type of data center?

Q21. What design aspects would you say have the largest impact on the infrastructure energy use?

Q22. And thinking about those aspects, which types of organizations involved have the most influence on energy use?

Q23. What factors influence data center owners or operational mangers to purchase energy efficient infrastructure equipment or implement energy savings operational practices? And how does that vary by submarket, if at all?

Q24. What factors limit the incorporation of energy saving equipment or operations? And how does that vary by submarket, if at all?

Q25. What would you say are the adoption rates for ENERGY STAR UPS (uninterruptable power supplies) for each data center submarket?

Q26. What would you say are key infrastructure energy saving operational practices? And does that vary by submarket? *[**If needed:** Operational practices can include airflow management, server power management, cable management, or adjusting temperature and/or humidity levels]*

Q27. Who are a few of the most prominent data center infrastructure equipment suppliers and contractors in the northwest, or nationally if there are not regional distinctions?

Building Shell

Q28. Thinking about data center building shells. To what extent do we need to be thinking about the building shell for data center new construction? What building design aspects have the largest impact on energy use?

Q29. And thinking about those aspects, which types of organizations involved have the most influence on energy use? To what extent are they motivated to reduce energy use? And does that vary by submarket?

Q30. What factors limit the incorporation of energy saving features and/or equipment into data center building shells?

Q31. Who are a few of the most prominent builders of data centers in the northwest, or nationally if there are not regional distinctions?

Data Center Trends

We have a few questions related to trends occurring with data centers.

Q32. What would you say is the current best practice in terms of data center design and construction? For example, how common are design-build approaches, integrated project delivery, and other methods? Does that vary by IT versus infrastructure or by data center type? [If Needed: Design-build entails a single contractor managing the design and construction services (as compared to design-bid-build). Integrated project delivery involves collaboration between owners, designers, and contractors from the project beginning.]

Q33. How commonly are established designs - in contrast to customized designs - followed, such as by design-build firms or large clients? Does that vary by IT versus infrastructure or by data center type?

Q34. Migrating data to the cloud is becoming increasingly prevalent for large and small organizations. Does cloud computing vary by business type? How do you see this movement to the cloud impacting data centers? Will hyperscale data centers become more prevalent in the future?

Q35. What would you say are major trends in data center design, IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, use of fuel cells or renewables, increased use of graphics processing units (GPUs), change in uptime or targeted reliability tiers.]*

Q36. What impact do you see these trends having on energy usage? How do these trends vary by data center type?

Q37. Thinking specifically about data centers in the Pacific Northwest, what trends to you see occurring within the next 5 to 10 years?

Q38. Do you anticipate growth in the number of data centers in the Pacific Northwest? **[If Yes:]** Do you expect to see higher growth in one data center submarket over another?

Q39. In your opinion, what role can utilities play in encouraging energy efficient data center design and construction practices as well as IT equipment selection? Does this role vary by data center submarket?

Conclusion [ASK ALL]l

Q40. Now that you've learned the topics we are interested in understanding, what else do you think is important for us to understand about the data center market?

Q41. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

End User Enterprise In-Depth Interview Guide

Introduction

Hello my name is [**YOUR NAME**] calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

We will not be attributing any of your answers to you or [organization] in our reporting to BPA.

Is it alright if I record this call to assist my notetaking? (Nothing you say will be attributed to you or your organization, and only members of the research team will access the recording.)

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with your organization.

Q2. We'd like to get a general sense of your organization to help us put your responses into context. Very roughly, about how many data centers – of any size – is your organization responsible for? You can give me a range or an order of magnitude, if you like.

Q3. **[IF MULTIPLE DATA CENTERS:]** About what proportion, roughly, of these are in the Pacific Northwest states – that is Washington, Oregon, Idaho and Montana?

Q4. And very roughly, approximately what is the capacity of those data centers? **[If needed:]** By capacity I mean servers, storage, networking equipment.] You can give me a range or an order of magnitude, if you like.

Q5. In addition to your data main data center(s), do you also have servers spread throughout your facilities in multiple closets and server rooms?

Q6. Do you use managed service providers for any of your computing needs? [yes/no]

Q7. **[IF YES TO Q6:]** What services do they provide to your organization?

Q8. Do you use colocation data centers for any of your computing needs? [yes/no]

Q9. **[IF YES TO Q8:]** What colocation services does your organization use?

Q10. What cloud computing services does your organization you use?

Q11. What is the maximum proportion of your computing that could potentially move to the cloud?

Q12. When do you anticipate reaching that point?

Q13. To what extent are the cloud computing services you use located in the same geographical region as your firm?

Q14. Does your firm own the data centers hosting your clients' cloud computing applications?

Trends [ASK ALL]

Q15. What would you say are major trends in data center design or trends driving IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, increased use of graphics processing units (GPUs), AI, blockchain technology, cryptocurrency mining change in uptime or targeted reliability tiers.]*

Q16. What impact do you see those trends having on energy usage?

Decision Making [ASK ALL]

We are interested in how energy efficiency factors into [organization's] data center designs and operations.

Q17. What do you focus on to reduce energy use at your data center(s), if anything?

Q18. Do you strive for minimal possible energy consumption in your data center(s) regardless of other considerations, or are there trade-offs between energy efficiency and other operational considerations?

Q19. **[IF TRADE-OFF:]** What are those trade-offs?

Q20. Does your firm receive any outside assistance in making IT purchase decisions?

Q21. **[IF TO Q21 YES:]** From who? What type of assistance?

Q22. Can you tell me how frequently you typically replace the various types of key IT and infrastructure equipment?

Q23. And how frequently you replace the various types of key infrastructure equipment?

IT and Infrastructure Equipment [ASK ALL]

Next, I'd like to discuss IT and infrastructure equipment.

Q24. Does your organization track server utilization?

Q25. **[IF YES TO Q25:]** What would you say your average utilization rate is?

Q26. Does your organization track data center PUE or Power Usage Effectiveness?

Q27. **[IF YES TO Q27 AND MULTIPLE FACILITIES:]** Does the PUE vary much across facilities?

Q28. **[IF YES TO Q28, VARIES:]** What factors lead to PUE variation?

Q29. **[IF YES TO Q27, TRACK PUE:]** What would you say the average PUE is?

Q30. Does all of the capacity your firm adds at a given time – so capacity of roughly the same vintage – have roughly the same PUE?

Q31. **[IF NO TO Q31:]** What factors lead to the varying efficiencies?

Q32. Next, please use a 1 to 5 scale, where 1 is not aggressive and 5 is very aggressive to indicate how aggressively your organization pursues a few energy-conserving approaches for data centers.

Measure	Not	2	3	4	Very	N/A – does not	Additional

	aggressive	aggressive	have policy	notes
Server virtualization				
Server decommissioning				
Air flow management				
Power management				
Purchasing ENERGY STAR UPS				
Purchasing ENERGY STAR servers				
Something else (Please specify):				

Conclusion [ASK ALL]

Q33. Is there anything else you think is important for us to know and understand about data centers market and your involvement with data centers?

Q34. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Thank you for taking the time to speak with us today. Have a wonderful day.

Colocation Services Provider In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

We will not be attributing any of your answers to you or [organization] in our reporting to BPA.

Is it alright if I record this call to assist my notetaking? (Nothing you say will be attributed to you or your organization, and only members of the research team will access the recording.)

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with your organization.

Q2. How common is it for colocation firms to offer design-build services?

Q3. And does your firm offer design-build services? [yes/no]

Q4. **[IF Q3 YES, DESIGN-BUILD:]** Why do some firms seek design-build rather than colocation services; what characteristics drive them to seek design-build?

Q5. How common is it for colocation firms to offer cloud computing services?

Q6. And does your firm offer cloud computing services? [yes/no]

Q7. **[IF Q6 YES, CLOUD:]** To what extent are the cloud computing services you provide located in the same geographical region as your clients?

Q8. **[IF Q6 YES, CLOUD:]** Does your firm own the data centers hosting your clients' cloud computing applications?

Q9. **[IF Q7 YES, CLOUD:]** Is that sufficient capacity for all of the cloud-based loads you manage for your customers, or are you also accessing capacity owned by other providers?

About Your Clients [ASK ALL]

Q10. Are all of your clients looking for colocation services that are in the same general geographic area as themselves? [If no, explain; how common]

Q11. Do your clients typically pay separately for electricity use, or is that covered in the cost of service?

Q12. **[If Q11 pay separately:]** Do your clients' per unit electricity costs vary much among your locations in the Pacific Northwest?

Q13. To what extent are clients seeking "green" or energy-efficient facilities, **[IF DESIGN-BUILD]** and please distinguish between colocation and design-build clients?

Q14. To what extent are your clients seeking renewable energy?

Q15. Are your clients willing to pay a premium for energy-efficient facilities or renewable energy?

Market Size and Growth [ASK ALL]

Q16. What size range in terms of capacity describes the majority of colocation data centers?

Q17. Approximately how many colocation data centers are currently in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana? [If necessary: Please provide your best guess, perhaps a range.]

Q18. And what is the approximate colocation total capacity in the Pacific Northwest?

Q19. BPA, our client, is interested in identifying number and sizes of colocation data centers located in the Pacific Northwest. Do you know of any information sources that might help with the identification?

Q20. Since early 2016 – the past two years, how has the capacity of colocation facilities in the Pacific Northwest changed? [If necessary: At about what rate?]

Q21. **[If past growth or decline:]** What market conditions have driven this change?

Q22. What do you foresee for the next three years – through 2020 – in the capacity of colocation facilities in the Pacific Northwest? [If necessary: At about what rate?]

Q23. **[If foresee growth or decline:]** What market conditions have driven this change?

Q24. How does the growth of colocation facilities in the Pacific Northwest compare with that of other regions of the country?

Trends [ASK ALL]

Q25. What factors lead your clients to go with colocation rather than internal data centers?

Q26. Are there particular areas or organization types where you expect to see more movement to colocation data centers than other areas or types?

Q27. What would you say are major trends in data center design or trends driving IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, increased use of graphics processing units (GPUs), AI, blockchain technology, cryptocurrency mining change in uptime or targeted reliability tiers.]*

Q28. What impact do you see those trends having on energy usage?

Decision Making [ASK ALL]

We are interested in how energy efficiency factors into [organization's] data center designs and operations.

Q29. Do you strive for minimal possible energy consumption in your colocation facilities regardless of other considerations, or are there trade-offs between energy efficiency and other operational considerations?

Q30. **[IF TRADE-OFF:]** What are those trade-offs?

Q31. **[IF DESIGN-BUILD:]** How do you advise your design-build clients on the importance of energy efficiency compared with other considerations, if at all?

Q32. What do you focus on to reduce energy use?

Infrastructure [ASK ALL]

Next, I'd like to discuss infrastructure issues.

Q33. Can you tell me how frequently you replace the various types of key infrastructure equipment?

Q34. Does your organization track data center PUE or Power Usage Effectiveness?

Q35. **[IF YES TO Q34, TRACK PUE:]** Does the PUE vary much across facilities?

Q36. **[IF YES TO Q35, VARIES:]** What factors lead to PUE variation?

Q37. **[IF YES TO Q34, TRACK PUE:]** What would you say the average PUE is?

Q38. Does all of the capacity your firm adds at a given time – so capacity of roughly the same vintage – have roughly the same PUE?

Q39. **[IF NO:]** What factors lead to the varying efficiencies?

Q40. What would you say is the typical PUE for new colocation data centers?

Q41. And what is the typical PUE for older colocation data centers?

Q42. Next, please use a 1 to 5 scale, where 1 is not aggressive and 5 is very aggressive to indicate how aggressively your organization pursues a few energy-conserving approaches for data centers.

Measure	Not aggressive	2	3	4	Very aggressive	N/A – does not have policy	Additional notes
Purchasing ENERGY STAR UPS							
Power management							
Air flow management							
Something else (Please specify):							

[ASK IF TIME PERMITS]

Q43. We are aware of the following IT equipment manufactures and would like to know if you are aware of other players in these markets.

1. For servers, we've noted Cisco, Dell, HP, and IBM. Can you suggest others we might explore?
2. For processors, we've noted Intel and ARM. Can you suggest others we might explore?
3. For networking equipment, we've noted Cisco, Juniper Networks, and HPE. Can you suggest others we might explore?

4. For other electronic and design manufacturers, we've noted Foxconn and Quanta. Can you suggest others we might explore?

Conclusion [ASK ALL]

Q44. Is there anything else you think is important for us to know and understand about data centers market and your involvement with data centers?

Q45. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Q46. Thank you for taking the time to speak with us today. Have a wonderful day.

Brokers of Colocation Space In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with your organization.

Q2. Does your firm offer data center related services other than brokering colocation facilities? [If yes, describe]

About Your Clients [ASK ALL]

Q3. Are all of your clients looking for colocation services that are in the same general geographic area as themselves? [If no, explain; how common]

Q4. Are all of your clients providing their own IT equipment? [If no, explain; how common]

Q5. Do your clients typically pay separately for electricity use, or is that covered in the cost of service?

Q6. **[If Q5 pay separately:]** Do electricity costs vary much among colocation providers in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana?

[ASK ALL]

Q7. To what extent are clients concerned about their expected electricity costs?

Q8. To what extent are clients seeking "green" or energy-efficient facilities?

Q9. To what extent are your clients seeking renewable energy?

Q10. Are your clients willing to pay a premium for energy-efficient facilities or renewable energy?

Q11. Do the data centers you broker report PUE or Power Usage Effectiveness?

ASK IF YES TO Q11 PUE, ELSE SKIP TO Q14 (MARKET SIZE AND GROWTH)

Q12. Does PUE vary much among colocation providers in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana?

Q13. What would you say the average PUE is?

Market Size and Growth [ASK ALL]

Q14. Approximately how many colocation data centers are currently in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana? [If necessary: Please provide your best guess, perhaps a range.]

Q15. And what is the approximate colocation capacity in the Pacific Northwest?

Q16. BPA, our client, is interested in identifying number and sizes of colocation data centers located in the Pacific Northwest. Do you know of any information sources that might help with the identification?

Q17. Since early 2016 – the past two years, how has the capacity of colocation facilities in the Pacific Northwest changed? [If necessary: At about what rate?]

Q18. **[If past growth or decline:]** What market conditions have driven this change?

[ASK ALL]

Q19. What do you foresee for the next three years – through 2020 – in the capacity of colocation facilities in the Pacific Northwest? [If necessary: At about what rate?]

Q20. **[If foresee growth or decline:]** What market conditions have driven this change?

[ASK ALL]

Q21. How does the growth in the colocation market in the Pacific Northwest compare with that of other regions of the country?

Trends [ASK ALL]

Q22. What factors lead your clients to go with colocation rather than internal data centers?

Q23. Are there particular areas or organization types where you expect to see more movement to colocation data centers than other areas or types?

Q24. What would you say are major trends in data center design or trends driving IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, increased use of graphics processing units (GPUs), AI, blockchain technology, cryptocurrency mining change in uptime or targeted reliability tiers.]*

Conclusion [ASK ALL]

Q25. Is there anything else you think is important for us to know and understand about the data centers market?

Q26. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Thank you for taking the time to speak with us today. Have a wonderful day.

Managed Service Provider (MSP) In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

We will not be attributing any of your answers to you or [organization] in our reporting to BPA.

Is it alright if I record this call to assist my notetaking? (Nothing you say will be attributed to you or your organization, and only members of the research team will access the recording.)

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with your organization.

About Your Clients [ASK ALL]

Q2. Does your firm offer cloud computing services? [yes/no]

Q3. **[IF YES TO Q2:]** To what extent are the cloud computing services you provide located in the same geographical region as your clients?

Q4. To what extent do your clients engage in cloud computing?

Q5. What is the maximum proportion of your client's computing that could potentially move to the cloud?

Q6. When do you anticipate reaching that point?

Q7. Does your firm provide MSP services to client servers in colocation facilities?

Q8. **[IF Q3 YES:]** About what proportion of the installed capacity that you service would you say is in colocation facilities?

Q9. How common is it for your clients to have servers spread throughout their facilities in multiple closets and server rooms?

Q10. **[IF ANY:]** And for these clients with scattered servers, how common is it that your MSP contract is covering most but not all of these servers, possibly just the largest installations?

Q11. To what extent are clients seeking "green" or energy-efficient IT equipment and operation?

Q12. Are your clients willing to pay a premium for energy-efficient IT equipment and operations?

Market Size and Growth [ASK ALL]

Q13. What size range in terms of capacity describes the majority of MSP-managed data centers?

Q14. Approximately how many MSP-managed data centers are currently in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana? [If necessary: Please provide your best guess, perhaps a range.]

Q15. And what is the approximate total capacity of MSP-managed data centers in the Pacific Northwest?

Q16. BPA, our client, is interested in identifying number and sizes of MSP-managed data centers located in the Pacific Northwest. Do you know of any information sources that might help with the identification?

Q17. Since early 2016 – the past two years, how has the capacity of MSP-managed facilities in the Pacific Northwest changed? [If necessary: At about what rate?]

Q18. **[If past growth or decline:]** What market conditions have driven this change?

Q19. What do you foresee for the next three years – through 2020 – in the capacity of MSP-managed facilities in the Pacific Northwest? [If necessary: At about what rate?]

Q20. **[If foresee growth or decline:]** What market conditions have driven this change?

Q21. How does the growth of MSP-managed facilities in the Pacific Northwest compare with that of other regions of the country?

Trends [ASK ALL]

Q22. What factors lead your clients to go with colocation rather than internal data centers?

Q23. Are there particular areas or organization types where you expect to see more movement to colocation data centers than other areas or types?

Q24. What would you say are major trends in data center design or trends driving IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, increased use of graphics processing units (GPUs), AI, blockchain technology, cryptocurrency mining change in uptime or targeted reliability tiers.]*

Q25. What impact do you see those trends having on energy usage?

Decision Making [ASK ALL]

We are interested in how energy efficiency factors into [organization's] data center operations.

Q26. What do you focus on to reduce energy use, if anything?

Q27. Do you strive for minimal possible energy consumption in your MSP-managed facilities regardless of other considerations, or are there trade-offs between energy efficiency and other operational considerations?

Q28. **[IF TRADE-OFF:]** What are those trade-offs?

Q29. Does your firm assist your clients in making IT purchase decisions ?

Q30. **[IF Q25 YES ASSIST:]** When customers are purchasing new IT equipment, how do you advise them on the importance of energy efficiency compared with other considerations, if at all?

Q31. Does your organization track client's server utilization rates?

[IF YES TO Q28, TRACK UTILIZATION, ELSE SKIP TO IT EQUIPMENT:]

Q32. Do the utilization rates vary much across clients?

Q33. **[IF YES TO Q29, VARIES:]** What factors lead to variation in utilization rates?

Q34. **[ASK ALL IN THIS SECTION:]** What would you say your clients' the average utilization rate is?

Q35. How commonly do you advise your clients to increase their server utilization rates?

Q36. **[IF SOME, AND SOME CLIENTS HAVE SCATTERED SERVERS:]** When advising clients on utilization, how common is it that you are able to include all of their servers, rather than just those in the largest facility?

IT Equipment [ASK ALL]

Next, I'd like to discuss infrastructure issues.

Q37. Can you tell me how frequently you typically replace the various types of key IT equipment?

Q38. How frequently do you propose high-efficiency IT equipment to your customers?

Q39. Next, please use a 1 to 5 scale, where 1 is not aggressive and 5 is very aggressive to indicate how aggressively your organization pursues a few energy-conserving approaches for data centers.

Measure	Not aggressive	2	3	4	Very aggressive	N/A – does not have policy	Additional notes
Server virtualization							
Server decommissioning							
Purchasing ENERGY STAR servers							
Something else (Please specify):							

[ASK IF TIME PERMITS]

Q40. We are aware of the following IT equipment manufactures and would like to know if you are aware of other players in these markets.

1. For servers, we've noted Cisco, Dell, HP, and IBM. Can you suggest others we might explore?
2. For processors, we've noted Intel and ARM. Can you suggest others we might explore?
3. For networking equipment, we've noted Cisco, Juniper Networks, and HPE. Can you suggest others we might explore?
4. For other electronic and design manufacturers, we've noted Foxconn and Quanta. Can you suggest others we might explore?

Conclusion [ASK ALL]

Q41. Is there anything else you think is important for us to know and understand about data centers market and your involvement with data centers?

Q42. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Q43. Thank you for taking the time to speak with us today. Have a wonderful day.

Hyperscale In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of Bonneville Power's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

We will not be attributing any of your answers to you or [organization] in our reporting to BPA.

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with [organization].

Data Center Trends [ASK ALL]

Let's discuss data center trends.

Q2. Roughly, what capacity defines hyperscale?

Q3. For context, are you able to say about what proportion of [organization's] hyperscale capacity serves [organization's] enterprise operations, versus the proportion leased to other organizations to serve their cloud computing needs?

Q4. About what is the current total capacity of hyperscale data centers in the Pacific Northwest, that is Washington, Oregon, Idaho and Montana?

Q5. Thinking about total computing occurring in the Pacific Northwest at hyperscale data centers, approximately what proportion of that computing is actually for Pacific Northwest customers as opposed to computing outside of the Pacific Northwest? [*Would include enterprise plus PNW users of cloud resources.*]

Q6. And conversely, what proportion of Pacific Northwest demand is likely met by hyperscale data centers outside of the region?

Q7. Thinking specifically of the cloud computing services provided by hyperscale data centers, are those generally located in the same geographical region as the users?

Q8. What do you foresee as the growth in hyperscale data center capacity nationally over the next three years?

Q9. Would you expect the hyperscale data center market in Pacific Northwest states to grow at about the same or a different rate?

Q10. **[If different:]** Why do you say that?

Q11. What trends are driving the growth of the hyperscale data centers market? [**Probes** to spur conversation as needed:] [*Cloud computing, AI/GPU, blockchain, cryptocurrency, digitization/IoT, etc.*]

Q12. To what extent do you think the following data center market types will lose market share to hyperscale data centers over the next three to five years:

1. Enterprise (roughly 300 kW up to 3 MW, or capacity given in Q2)
2. Small internal data centers (less than 300 kW or so)
3. Colocation

Q13. Are there reasons those markets will persist and not fully give way to hyperscale?

Q14. What are trends in data center design or trends driving the specification of IT equipment and infrastructure? *[**Probes** to spur conversation as needed: Larger data centers, larger data halls, use of fuel cells or renewables, use of GPUs, change in uptime, etc.]*

Q15. Given these trends of increasing numbers of data centers, as well as increasing (or high) energy efficiency, what do you think are the implications for energy use for the entire hyperscale data center market over the next three years?

Decision Making [ASK ALL]

We are interested in how energy efficiency factors into [organization's] data center operations.

Q16. Do you strive for minimal possible energy consumption regardless of other considerations, or are there trade-offs between energy efficiency and other operational considerations?

Q17. **[IF TRADE-OFF:]** What are those trade-offs?

Q18. What do you focus on to reduce energy use?

Q19. **[IF NOT COVERED:]** How do you source increasing energy-efficiency – is the innovation internal to [organization] and provided to your suppliers, or are you expecting they will innovate to meet your requirements?

Q20. Does all of the capacity added at a given time – so capacity of roughly the same vintage – have roughly the same PUE?

Q21. **[IF NO:]** What factors lead to the varying efficiencies?

Q22. In your opinion, how does the energy efficiency of unbranded, customized server equipment compare with servers [organization] gets from OEMs?

Q23. Can you tell me how frequently you replace the various types of key IT equipment?

Q24. And how frequently you replace the various types of key infrastructure equipment?

IT Equipment and Infrastructure [ASK ALL]

Q25. What would you say is the typical server utilization rate for hyperscale data centers?

Q26. What would you say is the typical PUE for new hyperscale data centers?

Q27. And what is the typical PUE for older hyperscale data centers?

Conclusion [ASK ALL]

Q28. Is there anything we haven't covered that you think might be relevant to our investigation into the hyperscale market in general or specifically about how energy efficiency factors into [organization's] data center operations?

Thank you for taking the time to speak with us today. Have a wonderful day.

Contractors, Architects, and Designers In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest, the four states of Washington, Oregon, Idaho and Montana. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

We will not be attributing any of your answers to you or [organization] in our reporting to BPA.

Is it alright if I record this call to assist my notetaking? (Nothing you say will be attributed to you or your organization, and only members of the research team will access the recording.)

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with your organization.

Q2. Please describe what data center-related services your firm provides? **[IF NOT CLEAR:]** What proportion of your firm's work is in data centers?

Q3. What types of data centers has your organization been involved with; I'm thinking in terms of enterprise, colocation, or hyperscale?

Infrastructure Equipment and Building Shell [ASK ALL]

Next, I have some questions about building infrastructure equipment in data centers, which includes both HVAC equipment and electrical systems.

[IF INVOLVED IN MULTIPLE DATA CENTER TYPES]: For all questions, please differentiate your responses by data center type if possible.

[ASK ALL]

Q4. Thinking about building infrastructure equipment (HVAC and Electrical systems), what aspects would you say have the largest impact on the infrastructure energy use at data centers?

Q5. Thinking specifically about the design of the building shell, what aspects have the largest impact on energy use?

Q6. And what would you say are key infrastructure-related energy saving operational practices in data centers? *[**If needed:** Operational practices can include airflow management, server power management, cable management, or adjusting temperature and/or humidity levels]*

Q7. Do you recommend installing ENERGY STAR certified uninterruptible power supplies (UPS) to your clients? Why or why not?

Q8. **[IF YES TO Q10:]** What proportion of data centers your firm has been involved with have incorporated ENERGY STAR UPS?

IT Equipment [ASK ALL]

Now let's talk about IT equipment used at data centers.

Q9. What equipment, design, or operational aspects would you say have the largest impact on IT energy use?

Q10. What factors limit the incorporation of energy efficient IT equipment or the implementation of IT-related energy savings operational practices?

Q11. And when a data center is being constructed, expanded, or renovated, how integrated are decisions about the IT equipment? For example, does the owner or owner's representative contract separately for the design, acquisition, and installation of the IT equipment, or is it typically done under the direction of one or two lead firms?

Q12. Does that approach vary by type of data center?

Q13. Do you recommend installing ENERGY STAR certified servers to your clients? Why or why not?

Q14. **[IF YES TO Q16:]** Approximately what proportion of servers at data centers your firm has been involved with are ENERGY STAR certified?

PUE [ASK ALL]

Q15. Does your organization track data center PUE or Power Usage Effectiveness for your data center projects?

Q16. **[IF YES TO Q18:]** Does PUE vary much?

Q17. **[IF YES TO Q19:]** What factors lead PUE to vary?

Q18. What would you say the average PUE is?

Data Center Design and Construction Market [ASK ALL]

Q19. Would you say data center specialization is required or beneficial for mechanical, electrical, and plumbing contractors who are involved in data center projects?

Q20. **[IF YES TO Q22:]** Why is specialization required or beneficial?

Q21. **[IF YES TO Q22:]** Does specialization also vary by data center type?

[ASK AD AND DBM]

Q22. What would you say is the current best practice in terms of data center design and construction? For example, how common are design-build approaches, integrated project delivery, and other methods? *[**If Needed:** Design-build entails a single contractor managing the design and construction services (as compared to design-bid-build). Integrated project delivery involves collaboration between owners, designers, and contractors from the project beginning.]*

Q23. To what extent are clients seeking "green" or energy-efficient facilities?

Q24. To what extent are your clients seeking renewable energy?

Q25. Are your clients willing to pay a premium for energy-efficient facilities or renewable energy?

Q26. Do you strive for minimal possible energy consumption in your designs regardless of other considerations, or are there trade-offs between energy efficiency and other operational considerations?

Q27. **[IF TRADE-OFF:]** What are those trade-offs?

Q28. How do you advise your clients on the importance of energy efficiency compared with other considerations, if at all?

Q29. What do you focus on to reduce energy use?

Q30. What factors limit the incorporation of energy-saving infrastructure equipment, design features, or operational practices?

Q31. How commonly are established designs - in contrast to customized designs - followed by your clients? Does this vary by data center or client type?

Data Center Trends [ASK ALL]

We have a few questions related to trends occurring with data centers.

Q32. What would you say are major trends in data center design or trends driving IT equipment and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, increased use of graphics processing units (GPUs), change in uptime or targeted reliability tiers.]*

Q33. What impact do you see those trends having on building design or equipment selection?

Q34. What do you foresee for the next three years – through 2020 – in the capacity of data centers in the Pacific Northwest? [If necessary: At about what rate?]

Conclusion [ASK ALL]

Q35. Is there anything else you think is important for us to know and understand about the market for data centers and your involvement with data centers?

Q36. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Thank you for taking the time to speak with us today. Have a wonderful day.

Professional Services Organization In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

Roles and Responsibilities [ASK ALL]

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities with [organization].

Q2. And please briefly describe the data center-related expertise or services your organization provides.

Data Center Trends [ASK ALL]

Let's discuss data center trends.

[ASK ALL]

Q3. What would you say are major trends in data center design, IT equipment, and infrastructure? *[**If Needed:** Trends could include shifts towards larger data centers, larger data halls, use of fuel cells or renewables, change in uptime or targeted reliability tiers.]*

Q4. What impact do you see those trends having on building design or equipment selection?

[ASK UPTIME INSTITUTE]

Q5. Thinking specifically about reliability tiers, what trends have you noticed? For example, has reliability become less important due to networked data centers?

Q6. What impact do you see those trends having on energy usage?

Q7. How do reliability requirements vary by data center type and business type?

[ASK ALL]

Q8. What trends have you noticed involving increasing digitization, increased use of graphics processing units (GPUs), AI, blockchain technology, cryptocurrency mining, or changes in reliability requirements?

Q9. What effect do you anticipate those trends having on the data center market?

Q10. What impact do you see those trends having on energy usage?

[ASK GREEN GRID]

Q11. What is the average power usage effectiveness (PUE) found in today's data center's?

Q12. How does PUE vary between older and newer data centers?

Q13. How does PUE vary by data center type?

Q14. **[If not discussed:]** What factors lead to these variations?

Q15. What trends do you see with PUE?

Q16. **[If not discussed:]** What are the primary factors contributing to lower PUE?

Q17. **[If not discussed:]** What trends are you seeing with data center cooling? What affect do those trends have on PUE?

Q18. Do these PUE trends differ by data center type?

Q19. How receptive are data center owners and managers to monitoring their PUE?

Q20. How receptive are data center owners and managers to monitoring their utilization rates?

Q21. What other metrics do you feel that are important to pay attention to when thinking about the energy usage of data centers?

[ASK ALL]

Q22. Do you expect hyperscale data centers to become more prevalent in the future? If so, why?

Q23. And how much growth do you expect in the colocation market due to groups moving away from enterprise and small internal data centers?

Q24. At what rate do you think hyperscale data center capacity will grow in the next three years?

Q25. How about the rate of colocation data center capacity expansion in the next three years?

Q26. Finally in this series, what about the rate of enterprise, own-facility expansion?

Q27. Thinking specifically about data centers in the Pacific Northwest, what trends do you see occurring within the next 5 to 10 years?

IT Equipment [ASK ALL]

Next, I have some questions specifically related to IT equipment

Q28. What would you say are key IT energy saving equipment or operational practices used at data centers? For example, virtualization, decommissioning, purchasing energy efficient servers and networking equipment, or power management?

Q29. In your opinion, how prevalent is server virtualization? How does virtualization vary by data center type?

Q30. How prevalent is the practice of removing comatose or under-utilized servers? How does this vary by data center type?

Q31. What factors limit or prevent the incorporation of energy saving IT equipment or energy-saving operational practices at data centers? How do these limitations vary by data center type?

Q32. What is the typical turnover rate for servers?

Q33. How about the typical turnover rate for networking equipment?

Q34. And how about for storage equipment?

Infrastructure Equipment and Building Shell [ASK ALL]

Finally, I have some questions about building infrastructure equipment in data centers, which includes both HVAC equipment and electrical systems.

Q35. What would you say are key infrastructure energy saving operational practices? Do these practices vary by type of data center? *[**If needed:** Operational practices can include airflow management, server power management, cable management, or adjusting temperature and/or humidity levels]*

Q36. What factors limit the incorporation of energy saving equipment or operations at data center facilities? Does this vary by type of data center?

Q37. And thinking about the design of data center buildings, what aspects have the largest impact on energy use?

Q38. What factors limit the incorporation of energy saving design features into data center buildings?

Conclusion [ASK ALL]

Q39. Is there anything else you think is important for us to know and understand about the market for data centers and your involvement with data centers?

Q40. And can you recommend anyone for us to reach out to that you think can speak to any of the topics we discussed?

Thank you for taking the time to speak with us today. Have a wonderful day.

Utility Program Implementer In-Depth Interview Guide

Introduction

Thank you for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest. The research intends to better understand current market conditions and trends with respect to data center energy usage, and to support subsequent empirical modeling of data center energy usage and its changes over time.

General Questions

Q1. Can you start by giving me a description of your role within your organization and your experience with data centers?

Q2. *[If unclear:] Is your work focused on one specific area of the data center market or do you have experience with a broad range of data center types? What utility programs in the PNW have you worked with?]*

Utility Programs

Q3. What would you say is the current state of utility programs in the Pacific Northwest that are focused on data centers?

Q4. Where are utility programs focusing their efforts for data centers?

Q5. In your opinion, what role can utilities play in encouraging energy efficient data center design and construction practices as well as IT equipment selection? Does this role vary by data center submarket?

IT Equipment

Q6. Let's talk about the IT equipment. What types of individuals and organizations are involved in the IT design, acquisition, and installation, and how are they involved in the process, from conception and design through installation and any final testing? And how does that vary by submarket, if at all? *[**If needed:** IT equipment includes, servers, storage devices, and networking equipment.]*

Q7. What factors influence data center owners, IT, or operational mangers to purchase energy efficient equipment or implement energy savings operational practices? And how does that vary by submarket, if at all?

Q8. What factors limit the incorporation of energy saving IT equipment or IT related operations? And how does that vary by submarket, if at all?

Q9. When a data center is being constructed, expanded, or renovated, how integrated are decisions about the IT equipment? For example, does the owner or owner's representative contract separately for the design, acquisition, and installation, or is it typically done under the direction of one or two lead firms? And does that vary by submarket?

Q10. Who are a few of the most prominent data center IT equipment suppliers and contractors in the northwest, or nationally if there are not regional distinctions?

Infrastructure Equipment

Now I have some questions about data center building infrastructure equipment, which includes both HVAC equipment and electrical systems

Q11. What factors influence data center owners or operational mangers to purchase energy efficient infrastructure equipment or implement energy savings operational practices? And how does that vary by submarket, if at all?

Q12. What factors limit the incorporation of energy saving design features, equipment, or operational practices? And how does that vary by submarket, if at all?

Q13. What would you say are key infrastructure energy saving operational practices? And does that vary by submarket? *[**If needed:** Operational practices can include airflow management, server power management, cable management, or adjusting temperature and/or humidity levels]*

Q14. Who are a few of the most prominent data center infrastructure equipment suppliers and contractors in the northwest, or nationally if there are not regional distinctions?

Data Center Trends

We have a few questions related to trends occurring with data centers.

Q15. What would you say is the current best practice in terms of data center design and construction? For example, how common are design-build approaches, integrated project delivery, and other methods? Does that vary by IT versus infrastructure or by data center type? *[**If needed:** Design-build entails a single contractor managing the design and construction services (as compared to design-bid-build). Integrated project delivery involves collaboration between owners, designers, and contractors from the project beginning.]*

Q16. How commonly are established designs - in contrast to customized designs - followed, such as by design-build firms or large clients? Does that vary by IT versus infrastructure or by data center type?

Q17. What would you say are major trends in data center design, IT equipment, and infrastructure? *[**If needed:** Trends could include shifts towards larger data centers, use of fuel cells or renewables, increased use of graphics processing units (GPUs), change in uptime or targeted reliability tiers.]*

Q18. What impact do you see these trends having on energy usage? How do these trends vary by data center type?

Q19. Thinking specifically about data centers in the Pacific Northwest, what trends to you see occurring within the next 5 to 10 years?

Q20. Do you anticipate growth in the number of data centers in the Pacific Northwest? **[If Yes:]** Do you expect to see higher growth in one data center submarket over another?

Conclusion

Q21. Now that you've learned the topics we are interested in understanding, what else do you think is important for us to understand about the data center market?

RMI Blockchain-Focused Technology In-Depth Interview Guide

Introduction

Hello my name is **[YOUR NAME]** calling from Research into action.

Thank you again for taking the time to speak with us today. This interview is part of BPA's research on the data center market in the Pacific Northwest. I am reaching out to you to better understand the blockchain technology trend and its implications for data center computing.

We understand that EWF is working on multiple fronts to advance blockchain technology in support of transactive energy, including its recent release of Tobalaba. From your web information, we understand that Tobalaba is an open-source blockchain platform tuned to the needs of the energy sector that uses a highly energy-efficient validation approach.

I have about a dozen questions related to blockchains more broadly than Tobalaba and specifically concerning blockchains and implications for energy use. I do not have an IT background, so I'm seeking a layman's understanding of various issues.

May I record our conversation to assist with my notetaking?

Questions

Q1. Let's start with a bit about you. Please tell me your title, and briefly describe your role and responsibilities related to blockchain technology.

Q2. Blockchains grow over time, as do the networks to which they are connected. What are the limits to this growth in discrete blockchains, and in discrete networks?

Q3. We have read the technology is distributed, but lacking an IT background, do not understand the computing implications. To what extent is the information duplicated?

Q4. Do parties to a given transaction need access to the entire chain or entire network for computation or storage? Again, our question relates to the implication for computing resources; we are not seeking a highly technical answer.

Q5. Are there limits to the growth of the technology in general?

Q6. Are any unintended consequences evident?

Q7. We have read about blockchain technology in the context of commercial transactions. We have also read hints that the content of blockchains isn't limited to transactions. For example, it can include user characteristics that support transactions but are independent of transactions. Is that accurate, and are there other non-transactional uses? [If yes, what?]

Q8. Can blockchain technology support the digitization of previous analogue information that does not relate to transactions? [If yes, what types of applications, or how widespread are the applications?]

Q9. How can blockchain technology support the internet of things – the communication of monitored information to the web and intercommunication among devices, databases, and so on? [If yes, what types of applications, or how widespread are the applications?]

Q10. So the use of blockchains is growing, and the length of blockchains grows. We have read a lot about the energy consumption related to transaction validation. We haven't found much information on data storage. Has anyone examined the data storage implications? [If yes, what are the implications?]

[Blockchain size for Bitcoin is currently 149 gigs, up from 45 gigs in 2015.]

Q11. Hypothetically, how might the growth in data center loads be expected to compare under two scenarios – the world as it is with the increasing uptake of blockchains, digitization, and IoT phenomenon OR a hypothetical world of increased digitization and IoT *in the absence of* blockchain technology?

Q12. How pervasive do you think blockchain technology will be in the next ten years?

Q13. Might any of the issues we've discussed have specific relevance to the Pacific Northwest and the growth of data center loads in the region?

Conclusion

Q14. Is there anything else you think is important for us to know and understand about these trends in computing and data centers?

Thank you for taking the time to speak with us today. Have a wonderful day.