



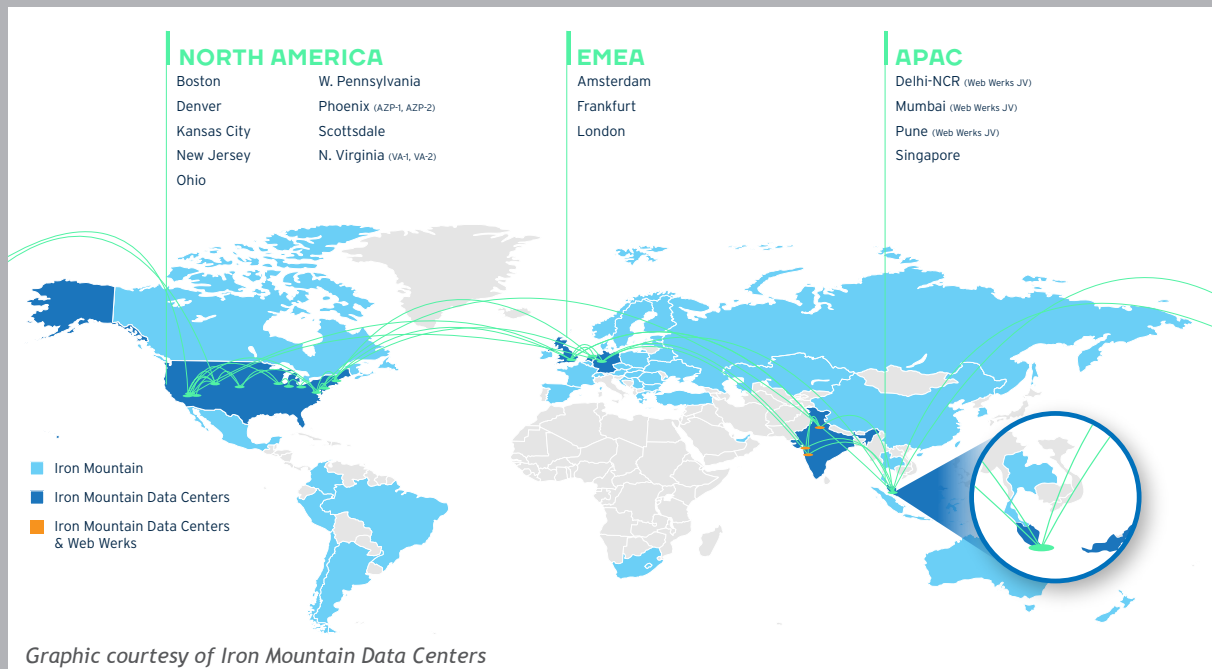
DATA CENTER

Frontier Special Report

Ecosystems & the Edge

Where the Data Center Becomes a Transactional Marketplace

Written by Paul Gillin



brought to you by



Contents

Ecosystems & the Edge: Where the Data Center Becomes a Transactional Marketplace	2
A New Breed of Transaction	3
A Three-Tier Architecture	3
Global Potential	4
Interconnection is Key	4
The Colocation Option	5
The Value of Colocation Providers	6
Cloud vs. Colocation.....	6
Q&A: Iron Mountain Executives on Edge Computing	7
Sponsor: Iron Mountain Data Centers	9

Ecosystems & the Edge: Where the Data Center Becomes a Transactional Marketplace

Rapidly evolving edge computing architectures are often seen as a way for businesses to enable new applications that require low latency and place computing close to the origin of data.

While those are important use cases, what is less often discussed is the opportunity for businesses to leverage the edge to spawn ecosystems that generate new revenue. To realize this value, companies must think of the edge as more than just a collection point for data from intelligent devices. They should

Edge computing is an expansion of cloud infrastructure that moves data collection, processing and services closer to the point at which data is created or used.

broaden their vision to see the edge as a new business hub. These small data centers can evolve into full-fledged service providers that attract local businesses, generate e-commerce transactions and enable interconnections that never touch the central cloud.

Edge computing is an expansion of cloud infrastructure that moves data collection, processing and services closer to the point at which data is created or used. It is the fastest-growing segment of the cloud category with the total market expected

Global Power Footprint Growth Infrastructure Edge Deployment

2019 | 1 GW

2028

ESTIMATED GROWTH

40+ GW

Source: The Linux Foundation

to expand 37% annually through 2027, according to Grand View Research. The Linux Foundation estimates that the global power footprint for infrastructure edge deployment will grow from 1 GW in 2019 to over 40 GW by 2028.

Edge computing is also the subject of many misconceptions, however, led by the belief that its principal value is to manage intelligent devices like cameras and sensors. That is only one scenario; the potential to fuel digital transformation is much greater.

The bigger picture sees edge as an extension of cloud hubs and corporate data centers into regions and applications that can't be served adequately by today's cloud infrastructure. That opens up a host of new business opportunities.

A New Breed of Transaction

Distributed cloud infrastructure enables new types of transactions that are impractical today. Most of these center on partner ecosystems that develop around edge data centers. A few examples are:

- ▶ Streaming video services in the cloud can interconnect with content delivery networks near the edge for caching and last-mile delivery. Streaming video can be delivered that's customized to the device level.
- ▶ E-retailers will tap into local networks of suppliers, order fulfillment companies, warehouses and delivery networks to enable just-in-time delivery and tracking.
- ▶ Healthcare insurers entering a new region will use interconnection to link up to local care providers, transportation networks, equipment providers and pharmacies to guarantee quality service delivery.
- ▶ Package delivery services will enter new markets by interconnecting with regional logistics firms and transportation fleets.
- ▶ Green energy providers will find contractors, local generation facilities and even drone operators to install, inspect and maintain equipment.
- ▶ Real-time ad brokers will customize messages to local languages and cultural preferences.
- ▶ Sensor-enabled predictive maintenance of equipment in the field will enable businesses to identify problems based on sensor readings and dispatch repair crews before equipment fails.

Redeploying compute power closer to the point of the transaction can cut response times by 90%.

Many of these ecosystems will be based on low-latency connections that edge facilities can deliver. Applications like augmented reality, remote vehicle control, real-time gaming and point-of-sale credit card fraud detection can't tolerate 100-millisecond response times from a cloud server. Redeploying compute power closer to the point of the transaction can cut response times by 90%. In scenarios like autonomous vehicles that is the difference between life and death.

Edge data centers can support compute-intensive mobile services like autonomous vehicles and real-time fleet management without the latency and IT infrastructure requirements of a central cloud.

A Three-Tier Architecture

Evolving edge strategies center on a multi-tiered architecture and distributed intelligence. A global end-to-end solution will have three components:

- ▶ Public and private cloud services, internet exchanges and software-defined network platforms will consolidate data, provide ecosystem access and support large third-party marketplaces.
- ▶ Core data centers will provide time-sensitive compute and data analysis services, interconnection and cost-effective routing to public and private clouds.
- ▶ Edge data centers will process data in motion, connect to internet service provider networks and deliver experiences to individual devices and users. Content delivery network caching servers connected by high-speed networks will facilitate data exchange with core data centers while supporting high-speed service delivery to endpoints.

This architecture limits round trips to the central cloud, thereby reducing latency. Edge data centers can support compute-intensive mobile services like autonomous vehicles and real-time fleet management without the latency and IT infrastructure requirements of a central cloud.

Edge nodes will be part of what the [Linux Foundation](#) called "an internet of systems, where devices serving different vertical applications need to communicate directly to exchange knowledge, autonomously and securely with no single point of failure." Far beyond supporting just intelligent devices, these edge environments will enable new types of transactions. Driving this activity closer to the edge of the network and locating compute to deliver single-millisecond latency will reach new customer bases, particularly in remote regions and underdeveloped parts of the world.

Global Potential

Edge computing is usually framed in the context of improving what is already in place, but the bigger opportunity is elsewhere. Although people in developed economies take high-speed internet for granted, the reality is:

<p>MORE THAN 40% of the world's population has no reliable internet access <i>Statista 2021</i></p>	<p>LESS THAN 1/3 of internet users have high-speed broadband <i>Statista 2021</i></p>	<p>MORE THAN 50% of overall web traffic comes via mobile phones <i>Oberlo</i></p>
--	--	--

Even in areas with high internet penetration, bandwidth is often spotty and unpredictable, making the task of serving millions of users from cloud regions or zones with sub-second response times difficult to impossible.

These disparities won't last, however, particularly with the impending arrival of large-scale 5G wireless deployments. Consumer devices of all kinds now come with IP addresses and people will want to connect them when bandwidth is affordable. Businesses will also want to connect their many sensors, environmental controls, factory-floor devices and other "always-on" equipment to management hubs and central control centers.

This will drive the explosive growth of many edge platforms into full-blown data centers, supporting transactions, e-commerce and interconnection. Ecosystems that were previously concentrated in major urban data centers will become more distributed and closer to the points at which data is collected and delivered. Regional servers will orchestrate clusters of devices and deliver data downstream to endpoints and upstream to the cloud over high-speed connections. Cloud regions will aggregate data from regional devices for analytics and planning.

This transformation can happen quickly. For example, India's colocation data center market size is expected to grow 21% annually through 2025, at which time the Asia Pacific region will account for half of the global colocation market. The number of mobile banking users in India, Indonesia, the Philippines and Vietnam is expected to grow 75% between 2019 and 2023, pressuring banks in those regions to quickly put into place distributed architectures that can scale rapidly. Companies like Web Werks,

one of India's top colocation data center providers, are aggressively building out their edge architectures with global partners like Iron Mountain to realize this opportunity.

Centralized data centers and cloud servers will evolve to become points of connectivity between cloud servers, endpoint devices and the partner ecosystems that enable digital transformation. They will be supplemented by smaller regional processing hubs that provide a subset of services focused on metro areas or even neighborhoods.

Regional data centers will provide multi-tenant colocation services in places such as stadiums and urban rooftops, delivering peering and interconnection services with extremely low latency to nearby customers.

5G networks will present a host of new business opportunities in the coming years as high-speed mmWave 5G service delivers wireline-like bandwidth to untethered devices. Widespread adoption of these latency-sensitive applications could overwhelm data center infrastructure that is unprepared to accommodate it.

Interconnection is Key

Interconnection will be a critical consideration for enterprises building out edge data centers. It enables data exchange between two or more entities or partners at the fastest possible speed by combining high-performance networks with physical proximity. Customers avoid the unpredictability and cost of the public Internet.

Interconnection enables business partners to rapidly connect to speed new services to market and to improve performance, availability and functionality. Customers using multiple cloud providers can enjoy access speeds equivalent to dedicated on-ramps along with the flexibility to switch sessions easily between cloud providers and transfer data with lower egress charges. It's not surprising that worldwide interconnection bandwidth capacity more than quadrupled between 2017 and 2020 to 5,000 terabits per second.

Interconnection today is principally the domain of large companies like financial services institutions, manufacturers, cloud computing providers and telecommunications firms, but the growth of edge infrastructure will draw in smaller firms and local use cases. This is a major business opportunity for edge data center operators.

The Colocation Option

Edge strategies should start small and grow incrementally using infrastructure that can scale to full-blown data center functionality over time. Because edge applications are, by their nature, latency-sensitive, organizations must be prepared to expand capacity without time-consuming and expensive acquisition of new real estate and equipment.

Most edge computing architectures will make use of colocation partners like Iron Mountain that already have facilities in place, the ability to turn current facilities into full or partial data centers and knowledge of local markets. Once the domain of primarily large enterprises, colocation data centers have evolved to provide a variety of customizable services that support a wide range of businesses. These include on-site assembly and testing, inspection and maintenance, inventory analysis and 24-hour services.

Many colocation providers offer redundant HVAC systems and multiple cooling options such as centrifugal chillers, adiabatic cooling solutions and water loop pumps.

Colocation providers typically provide reliable power with redundancy and battery backup or uninterruptible power supplies. They maintain physical security controls that include guards, full-time surveillance and multiple forms of authorization. Many also offer redundant HVAC systems and multiple cooling options such as centrifugal chillers, adiabatic cooling solutions and water loop pumps.

Perhaps most importantly, colocation providers know the territory. “The best space, power and connectivity in every geography everywhere in the world is owned and operated by a data center company in that place,” wrote Jonathan Seelig, CEO of the globally distributed cloud platform Ridge in The Linux Foundation’s State of the Edge report.

Not all colocation partners are alike, however. The factors that should influence a business’ choice of edge partners include:

Interconnection capacity - This includes the amount of bandwidth available for interconnection services, the types of available interconnection (peering,

Once the domain of primarily large enterprises, colocation data centers have evolved to provide a variety of customizable services that support a wide range of businesses.

cross-connect, internet-site connectivity, blended IP) and service-level agreements.

Room to expand - In addition to existing colocation facilities, operators should have facilities or real estate that can be transformed quickly into data centers on a global scale in second- or third-tier locations to serve new edge customers.

Partner ecosystem - Look for co-location providers that have existing partnerships with a wide range of carriers, cloud providers, CDNs and managed service providers to offer their customers the widest variety of options.

Workload capacity - The service should have the physical space available to expand without compromising on security, environmental controls or service levels.

Bandwidth - An edge configuration is likely to involve a connection to an internet service provider. Ensure that ample bandwidth and multiple redundant connections are available. Bandwidth should be over-provisioned so that there is no risk of running up against peak usage thresholds.

Backup/business continuity - Redundant power sources, bandwidth and HVAC systems are factors here. The provider should have plenty of UPS capacity and be willing to provide service-level guarantees.

Ability to support devices in the field - “Remote hands” support in the data center should also extend to the devices you plan to deploy in the field. Make sure the provider has the necessary skills and equipment to fulfill your needs on a schedule that works for you.

Security - Because co-location facilities are shared by their nature, equipment must be secure. Look for features such as video surveillance, keycard or biometrics access, mantraps and security personnel to regulate site access and locked cages for your owned equipment.

Administration & monitoring - You should have full visibility into the status of your equipment at any time.

The Value of Colocation Providers

One of the greatest values colocation providers bring to the edge equation is an understanding of the local regulatory environment. Look for partners who understand and can also guide you about rules in such areas as:

- ▶ Privacy protection
- ▶ Data security
- ▶ Data locality
- ▶ Data retention and disposal
- ▶ Power utilization effectiveness
- ▶ Emission controls
- ▶ Water usage and recycling
- ▶ Accessibility
- ▶ Seismic protection
- ▶ Fire controls

Cloud vs. Colocation

The major cloud computing providers are all aggressively deploying edge infrastructure in similar ways. Their objective is to extend their existing platforms to core and edge data centers so that customers have identical functionality regardless of location. To do this, the big providers are pursuing a two-pronged strategy of installing their infrastructure inside customers' data centers and also in local telecommunications facilities. There are several advantages to this approach, chief among them being the ability for cloud providers to quickly provision infrastructure that fully compatible with the cloud regions.

However, limitations on size, bandwidth and compute capacity may challenge some to provide their customers with the full range of capabilities needed to build scalable edge architectures. Regional service breakdowns show that service availability can vary significantly by location.

None of the major public cloud providers has a network of wholly-owned colocation facilities, so they must rely on partnerships and workforces they don't control. These relationships will take some time to develop and can be fragile.

Edge infrastructure that isn't designed to be managed remotely or that must be administered

by contract personnel can be a risky proposition in far-flung locations. And self-contained miniature data centers are prone to theft, weather damage and unplanned outages. Most companies building out edge networks will want infrastructure that can't be taken offline by a traffic accident.

Over time, cloud providers will no doubt address these structural limitations, but some customers may not want to wait. The fastest, safest and most flexible option for them is to locate edge infrastructure in established colocation regions that are fully equipped and staffed by trained personnel and that already support local ecosystems.

- ▶ ROI is faster because regional colocation providers already have the facilities and relationships to support edge buildouts.
- ▶ Regional colocation centers can scale to meet capacity demands and, in most cases, are fully compliant with relevant regulations.
- ▶ The staff is trained and familiar with the equipment they work with every day.
- ▶ Facilities are secure and reliable with power, environmental and seismic controls that already meet local requirements.
- ▶ The staff speaks the local language and understands the culture and expectations of the ultimate end-users of edge services.
- ▶ Connections to nearby services can be quickly facilitated through peering and interconnection.

Most colocation providers also have existing ecosystems of customers and partners to whom new customers can connect. For example, Iron Mountain Data Centers' rapidly growing ecosystem with hundreds of third-party partners provides specialized network connectivity, access to fiber networks, software-defined networks, streaming services, specialized peering solutions, vertical market expertise and access to data centers in specific geographic and remote locations.

Many established co-location providers have also adopted public application program interfaces for rapid onboarding of customers and partners and have existing backend integration with a wide variety of other providers, including telcos, cloud platforms and colocation partners.

Q&A: Iron Mountain Executives on Edge Computing



RICH MILLER
Editor, DCF



PAUL GILLIN
Writer, DCF



MARK LEWIS
Senior Director,
Ecosystem Development,
Iron Mountain



MARK KIDD
EVP and General
Manager, Data Centers,
Iron Mountain

Q RICH MILLER: A lot of the edge action we've seen has been in two places—major markets where we're seeing more density, and regional data centers where there are new services coming in that hadn't been seen before. Does that resonate with what you guys are seeing and thinking at Iron Mountain?

A MARK LEWIS: Whenever you talk about edge, the first question that comes up is “that's lovely, but what exactly do you mean by edge?” Pretty much everybody has a slightly different definition of what edge means for them.

For us, the importance of location is that five milliseconds (of latency). That's important. The movement of compute right out to the edge—to the device and to the edge of the network—is definitely something that's going to happen. There may be small deployments of half a container here or there in Walmart parking lots and such. Is that really a data center? I think the key differences is the extension of data services and the ability to deliver transactions all the way down to that sub-five milliseconds range near a major metropolitan area.

There are definitely use cases for the things out in the middle of nowhere at the base of cellphone towers. But fundamentally, most of the drivers that have moved the data center industry have been about how you get closer to the people who are doing the transactions, because that's where the business is.

As devices generate more data over time, data itself is likely to become a commodity. The nature of data matters, and we think the ability to transact is the thing that's likely to drive the direction that you expand around the edge.

Q RICH MILLER: From a business standpoint, it seems everyone within the data center industry is starting to focus on building those ecosystems where folks can connect with one another, and are able to offer services to one another. One of the things it seems (Iron Mountain) has been working is interconnection and building customer density in your hubs.

A MARK LEWIS: Definitely. One of the things that's quite interesting about how the majority of data center companies have started to approach edge is the tendency to look at the use cases.

Things like healthcare, for instance, are very much focused on maintaining data locally. There's also the customer experience with streaming, so CDNs want to make sure they're developing a better customer experience by getting closer to the edge. Alongside that, there are use cases driven by industrial devices with the CCTV camera or pricing sensor on the shelf of the local store.

This type of data generation is where these marketplaces become more important. There will be people trying to analyze that data or sell that data. There will be a natural flow back to the data center from the edge.

For other people, it's the ability to connect. They can do compute, they can do analysis, and they can also interconnect with other suppliers, partners or network to complete transactions and connect applications.

Q PAUL GILLIN: What is special about what Iron Mountain is doing in this area? You obviously have many locations and reach a lot of countries. Is there anything you're doing from a technology standpoint, that's distinctive?

A MARK LEWIS: I think there's two things are distinctive here. We're probably the only data center global data center provider who could reach that many metros in the sub five milliseconds range, and connect it back to a core data center facility. With 91 million square feet of space in 1,470 data center facilities in 695 metropolitan areas across 50 countries, we have a unique global reach.

The Iron Mountain company has 225,000 customers, including 95% of the Fortune 1000. With our core data center ecosystem, our ambition is to convert that whole customer base into a single marketplace.

There's quite a bit of crossover already. For instance, in the document storage side there's a very strong digitalization business that's taking documents and insight and information to be put into the cloud.

A MARK KIDD: Every customer we have now is set up for the access and retrieval of information. Enterprises are actually giving us access to a digital transformation story. Our core (data center) business is being driven from a colo business into an integrated platform and connectivity story.

We have a view of how transaction-driven interconnection is going to evolve, and bring that model out to the edge to enable digital transformation. We believe that's happening, but we think it's happening slowly. Our question is "how do we play into that trend and evolve with it?"

You can see that in the acquisition of Web Werks, whose facility is the most peered location in Mumbai. We don't think we have magic sauce yet. But we think we have a good understanding how the world is evolving, and we think we're positioning ourselves to take advantage of that.

Q RICH MILLER: What are the applications, use cases and verticals that you think in the near term are telling the most interesting story about where this starts to happen?

A MARK LEWIS: There are going to be a variety of different use cases that get deployed locally. There will be workload deployments like AWS and Google that are an extension of their existing world down to a level much closer to their customers.

But alongside that, we also see media and gaming growing rapidly, so those clusters are going to start to move out of the core data centers—which tend to be largely interconnection points—and into cities where they can get large quantities of eyeballs and reduce latency to improve gaming quality.

We have a lot of medical companies in our ecosystem. In Europe there's an increasing concern around privacy and GDPR. That's a strong reason why a lot of people will want to maintain data residency in their local facility in the same region or country. I think for regulatory reasons there are going to be communities that are going to cluster around an edge facility.

And then there's the industrial component, which is those data-generating sensors that are going to become commercialized. That may be bespoke customized deployments for specific large vendors, and then deployments that allow access to eyeballs, which won't be exclusive to one on customer.

Q PAUL GILLIN: AWS is going after this business, as is Microsoft and everybody else. Where can you fill in that gap?

A MARK LEWIS: We have the potential to reach an awful lot of cities that the large cloud providers are likely to want to get to. That's especially important in countries outside of North America and Europe because things like permitting and land and energy availability become quite problematic. It's very difficult to build a new facility.

Also, a lot of customers for these edge zones have the opportunity to colocate their own legacy applications next to those newer applications.

A MARK KIDD: Some of those new places will very quickly become core locations. Four or five years from now, cloud players won't be there with a single node; they'll be there with their standard availability zone architecture. The application movement is happening.

The key is to get started now. See the early demand because the application changes your workload, and then that grows as it expands.

Q RICH MILLER: Can you comment about something that I've been hearing a lot more lately, which is edge as a global phenomenon?

A MARK KIDD: I completely agree. Asia, South America and Africa are all on the radar.

We will continue to raise awareness from Fortune 1000 about what's possible, where the world is headed and the type of applications and deployments they'll be able to do in the next 12, 24 or 36 months, and how that can improve their business.

I think there's much more receptivity to faster change on the heels of 12 months of COVID, where we've all seen the aggressive changes and rollout of technology.

Q RICH MILLER: How has the way that companies are using technology in the pandemic altered the course of how you see the edge trend, develop?

A MARK KIDD: Here's one example that I think we've talked about publicly. CVS is a very big client of ours on the physical side of the business. When you look at what they've done in response to the COVID rollout, you've see a complete re-architecture of their systems and processes over 90 days. I think the pace of change is extraordinary right now.

Edge is more like urban development than anything else. You see a city center core and it starts to sprawl out, and then you see the evolution associated with it. So, if the current edge becomes the future core, we will see growth and usage.

Sponsor: Iron Mountain Data Centers

ironmountain.com/data-centers

Iron Mountain Data Centers, a division of Iron Mountain Incorporated, is a leading provider of data center and colocation services. Our global portfolio consists of hyperscale-ready, strategic edge, and underground data centers comprising over 3.5 million square feet across fifteen locations in five countries.

More than 1,200 customers including cloud providers, global enterprises, and local market organizations choose Iron Mountain Data Centers for our thirty-year proven track record, risk mitigation, and operational efficiency. Iron Mountain colocation facilities are sustainable and provide the most logical venue for hybrid IT with easy access to the carriers, cloud providers, exchanges, and IT services necessary for digital transformation.

Comprehensive compliance support ensures highly regulated enterprise and public sector customers are protected as they change their organizations to thrive in a multi-cloud world.

Iron Mountain offers some of the world's most secure data centers. Our multi-layered approach to security in our highly-protected facilities includes a combination of technical and human security measures.

