COLOCATION ON PREMISE

VELOCITY data centers

THE CONUNDRUM OF CHOICE

Over the last several years, an entire data center industry has arisen allowing users to park their data compute assets in someone else's facility, namely colocation. Some colocation tenant customers state that they want to be out of the data center business, others move due to capacity issues, real or perceived (see Software Defined Power below) while others park redundant assets in a colocation facility as part of their disaster recovery/business continuity needs. Some tenants simply choose to diversify their assets into multiple locations.

Cloud providers and cloud data centers arose out the ability to spin up IT resources either in in someone else's facility using the provider's compute and hardware or using in house compute hardware for private cloud. Public cloud providers can be smaller or fall into the other category of hyperscale providers the likes of Amazon®, Google®, Microsoft®, and Facebook®. The compute giants provide cookie cutter racks filled with compute at a massive density in their facilities. These giant data centers tend to be more regionally located as opposed to locally owned and operated. But most companies don't build to that scale. Many companies that do take advantage of the quick spin-up of cloud, do so in a hybrid approach (some public, some private) of cloud compute working in concert.

While the industry would like you to believe that "everyone is going colo" and "everyone is going hyperscale" or "everyone is going cloud" the truth remains that many organizations still retain their internal data centers or at least a portion of those assets on site. Some retain data centers exclusively on site, and are committed to doing so long term. In reality, cloud, colo, hyperscale, and entity owned assets should all be viewed as tools in the toolkit of productivity and data asset management. Each has its strengths and weaknesses for any given application. Company strategies and applications will change multiple times over the life of the company and its data properties.

In a recent survey conducted by AFCOM and reported by Data Center Knowledge, "not only are companies not rushing to replace their own data centers with Amazon's or Microsoft's but many are expanding their own footprint." The average number of data centers per company responding was 12 but expected to grow to 17 over the next 3 years. Across these companies, an average of 1.8 were slated for immediate renovation with 5.4 expected to renovate over the next three years. In fact, 45% said that they were planning a new data center build. That clearly shows that there are companies that don't fit the "everyone" mold.

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075

The best CIOs and CTOs evaluate the technologies available and determine the best location for compute based on a variety of factors a few of which are:

- o Cost of location power
- o Renewable strategies and micro-grids
- o Smart city initiatives and latency adverse applications
- o Tax benefits of OPEX versus CAPEX
- o Depreciation of existing facilities assets and asset recovery periods
- o Lifecycle of existing compute assets
- o Local talent versus remote travel to data costs
- o Data sovereignty needs
- o Entity preferences as to data location
- o Security requirements (i.e. DoE and Homeland Security)
- o Cost of other/remote asset locations (i.e. circuits, insurance, etc.)
- o Cloud strategies
- o Availability of space in a given market
- o Overall bandwidth and the speeds needed across and between applications
- o Changing needs of a business

When a company spins up a corporate office building and services including a data center, their capital asset depreciation cycle starts for their larger equipment like buildings, generators, chillers and computer room air handling/conditioning equipment. Companies that lease their space may or may not realize all of these costs. But suffice it to say that there will be some capital outlay for equipment. Some assets will depreciate over a shorter period of time and many will operate under a longer depreciation schedule whether or not they in use. This makes abandoning them expensive and not suitable for many companies due to tax consequences.

For those enterprises that have already made capital investments, it is difficult and cost prohibitive to simply cast the investments aside. Data centers include compute, storage, and circuits in addition to capital power, cooling, and space. And let's not forget investments in talent.

For many, it's just not easy to pick up and move. This is particularly true for entities like hospitals, broadcast, and smart cities that will continue to maintain the capital equipment for the remainder of operations regardless if it is utilized by the data center or not. Some organizations receive tax incentives to locate in their jurisdiction and could lose those incentives if they move. Some have aggressive energy conservation plans that do not lend well to capabilities of others controlling the power side of the equation. Software defined power and microgrids are a great example.

With respect to power, aging assets is a problem in and of itself. Older assets are not environmentally friendly. In fact, a 2020 Uptime Institute study of over 300 data centers suggests that 66% of power and cooling are being consumed by aging data equipment with that power consumption is only supporting 7% of the actual compute happening in those same facilities. There is a lot of room for efficiency improvement. Entities may think they are at capacity and move when other options are available to reevaluate those theoretical capacity limits. So, what is a company to do in light of the conundrums?

THE MOVE TO COLO

Companies that picked up and moved to colocation facilities a few short years ago were hit with a surprise. It was estimated by the government tax and accounting office that roughly $3 trillion was going unreported in corporate lease obligations. For that reason, the tax laws changed, and the full obligation of leases now must be reported on the balance sheet and no ramp up periods are allowed. This hurt many companies with large "ramp up" leases for large data halls in colocation facilities and many contracts were quickly renegotiated. In some cases, companies decided to look towards hybrid (part onsite, part in the cloud or colo) solutions. In others, companies decided to move their assets to smaller spaces or back to their own locations removing the obligation for open spaces.

When determining the location of data and applications, an enterprise or public entity has a few choices. They can house it all internally, house it in the cloud, house it in a colocation facility, or pick some combination of the three, which is generally the norm. There was a time when companies like AWS predicted that all compute would happen in their cloud. While there are several cloud players today in addition to AWS, the cloud only prediction quite simply didn't hold true. Some companies exclusively utilize the public cloud, some partially utilize the public cloud, but by and large companies have either built their own cloud or use some hybrid combination of public and private cloud for their applications and data. Cloud providers are working with entities to provide onsite onramps or outposts that connect to their cloud for hybrid computing. Quite simply, moving everything to a colocation provider is not the only answer.

THE APPLICATION FIRST STRATEGY

Application first strategies are certainly a great place to start. In fact, rightsizing redundancy and compute needs is paramount to any efficiency plan. Uptime has long been guaranteed by adding additional power feeds and cooling paths N+1, 2N, etc. However, the piece that has largely been ignored is the resiliency of the IT resources. If a single application is made redundant to multiple sites, do all sites really require dual power connections, dual network connections and dual storage connections? The answer is quite simply, no. Further peak utilization is a poor factor for overall power needs as peak loads are rarely realized. Multiply this by the number of power feeds and the waste multiplies.

Aging assets and capacity limits are not necessarily a reason to move to a colocation facility. Some towns have very few colocation options and the ones that are there are very expensive due to high occupancy and demand over low supply. Rightsizing became a mantra. PUE (Power Usage Effectiveness) became the metric to measure waste and efficiency of data center equipment.

At some point redundancy creates waste. In fact, estimates are that roughly 65% of energy is wasted through transmission and the inability to accurately predict IT load. In a colocation/occupant scenario, the colocation controls the facilities (power and cooling) and the occupants control the IT load portion of a PUE equation. Therefore, not all PUE reported from providers is the same. In general, PUE reported from providers report power in/power delivered to IT kit, not consumed as in the original PUE equation. According to Brady, Gemma, Nikil Kapur, Jonathan Summers, and Harvey Thompson. "A Case Study and Critical Assessment in Calculating Power Usage Effectiveness for a Data Centre

> "It is crucial that an accurate IT load is used for the PUE, and that it is not based upon the rated power use of the equipment. Accuracy in the IT load is one of the major factors affecting the measurement of the PUE metric, as utilization of the servers has an important effect on IT energy consumption and hence the overall PUE value".

From a colocation perspective, those numbers are impossible to control as they are tenant controlled inside of the space. Colocation facilities power control stops at various points in the power chain depending on the facility. Therefore, PUE numbers are difficult at best to compare in any apples to apples comparison. Some are modeled, some are averaged, some are snapshots in time, and some are peak. There is not an enforced standard for reporting. This makes it difficult for companies that wish to be green to accurately predict and manage power costs, sourcing (green or not), and usage. Incentives from power companies are generally granted to the provider and not the tenant. Those incentives may alter the power costs but not pass through to the tenant.

THE TIME FOR MODULAR DATA CENTERS

Data centers can be purpose built, enclosed within a building, containerized (in a shipping container) or modular (not to be confused with a container). Modularly built data centers are pre-engineered. The components are crafted offsite and then the building itself and critical services (power, cooling, etc.) are assembled on site. The practice of crafting walls offsite mean that the facility can be constructed far more quickly than a bespoke data center. The savings of not having to re-engineer each site directly translates into savings. The preconstructed components translate into rapid, predictable deployment. In fact, with modular buildings, the lengthiest part of construction can be the permitting. The building is "live" 90 days after site approval, sometimes faster depending on other factors.

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY
data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY
data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075

According to Markets and Markets "The modular data center market size is estimated to grow from USD 13.07 billion in 2017 to USD 46.50 billion by 2022, at a Compound Annual Growth Rate (CAGR) of 28.90%. Growing requirement for scalable data center, reduced CAPEX, and growing need for green data centers are the major growth drivers of the market. In this study, 2016 has been considered as the base year and 2017-2022 as the forecast period…. such data centers comprise components with a standardized design that allows for rapid construction and deployment and are more economical to build than the traditional data centers. Portability and scalability benefits offered by modular data centers have helped in propelling the demand for these solutions across most of the organization sizes."

According to Global Market Insight in their "Industry Trends"

Modular Data Center Market size valued at over USD 13 billion in 2018 and will grow at a CAGR of around 23% from 2019 to 2025.
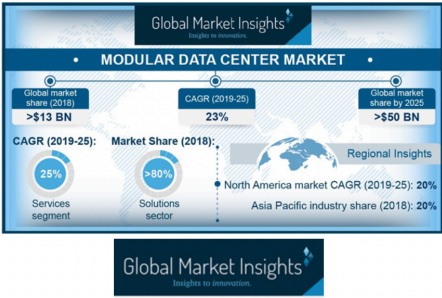
Driven largely by
- Reduced CAPEX
- Demand for scalable and cost-effective data center facilities
- Rising demand for edge computing operations in industrial sectors in the U.S.
- Growing need for energy-efficient IT infrastructure facilities

The modular data center building itself can be used for a variety of data center types from enterprise to smaller colocation. As it is modular in nature, it can grow into one larger building, or several small buildings can be clustered together in a campus footprint. For cloud onramps and hybrid environments, the nature of modular makes perfect sense, as data center on site assets have been shrinking in footprint in general. Smaller footprint customers don't always receive the same attention in colocations as an enterprise customer that will become an anchor tenant and put an entire data hall to use. And for some customers, there will ALWAYS be a portion of data assets managed and housed on site either due to latency, protection, compliance, trade secrets, or simply preference.

On premise data centers can provide better financial control for the end user/data owner. The decision isn't an all or nothing decision any longer. The argument that OPEX is better than CAPEX and therefore colocation is the only way to go is simply no longer true. Modular data centers can be built on site and then turned over to a lease management company in a colo on premise model allowing the owner to take advantage of an OPEX spend for roughly the same fixed monthly cost as occupying colocation space(s). Beyond space costs, the difference with colo on premise lying in the consolidation of some expense spend. For instance, companies may be able to take advantage of a single insurance policy with the modular data center on site. Whereas if a company has assets in multiple locations, that economy might not exist through the multiple policies needed.

The same can be said for power circuits, generator feeds, UPS and power transmissions, carrier circuits, SD-WAN, and communications that the company is already paying for on a single site. For off-site locations, it may be necessary to work with different suppliers for connectivity to various facilities which can limit bulk buying power and add a layer of complexity for interoperability across the protocol stack. The age-old finger pointing can be a deterrent. Most importantly, the local area higher speed connections will not carry a wide area provider monthly fee. The local connection will simply be a switch connection to a local switch port. Most enterprises will have some combination of local and wide area, but for lower latency, lower cost, higher speed needs local port speeds win. That lower cost, low latency connection is a driver to local workloads for hospitals, pharmaceuticals, life sciences, higher education, disaster recovery, manufacturing and the like.

With respects to power, several of the power providers internationally are providing grants for corporate consumers to "green up" their facilities. These can be used for microgrids and other energy savings measures. These measures may carry a term of use to be valid requiring a company to keep assets on site for some period. In a colo facility, credit is not always given for lowering your power footprint as the power costs are a pass thru to the tenant. Colos also must over-provision power to assure that all tenants can increase loads up to their contracted limits. Lack of power control and costs can be a significant deterrent if a tenant is planning to move to cloud services or move some assets out of the space during the lease period. Contract negotiations with respect to power and overall footprint are critical prior to occupancy of any space.

Measures to become more energy efficient may fall by the wayside if there isn't a financial benefit to becoming a better environmental steward. The "if I have to pay for it anyway…I might as well use it" attitude could stifle efficiency measures. Certainly not seeing the actual power usage on a monthly bill can render power "out of site out of mind." The type of power being utilized can also matter to some enterprises. As some companies look to become better environmental stewards, they prefer renewables or non-coal generated power. Microgrids are becoming increasingly popular as a means to utilize solar, wind, hydro and other cleaner forms of energy. Managing the source of power at various times of the day can be a significant factor in power cost savings.

THE CONTROLS

Almost all data center contracts have a prefixed price increase for each year of the contract. These range from 3-15% of the overall contract price. An on-premise data center will generally carry a fixed cost with limited increases as it is a single use space as compared to multi-tenant space with a ramp up period. Space tenant/customer owned and operated in an OPEX model, can be lower than for a larger operator. The contract term can be fixed to a longer period of time, if desired, for increased cost savings. If management contracts are in place, the same talent can be used to support more than the data center. For companies that are committed to supporting their local communities, on site is a step in the right direction. Local talent can be used and dedicated to the sole tenant as opposed to being spread across multiple tenants.

Upgrading to an onsite modular data center allows an entity to repurpose older data center space. In many cases, the resultant revenue can offset the cost of the modular data center. For instance, hospitals estimate revenue of roughly $2,000 per square foot. Modular data centers can also be a place to consolidate rogue computer rooms while providing redundancy, resiliency, and security. As security requirements increase for entities that connect to homeland security, these modular buildings can become a hardened SKIF. Entities can build several onsite with various levels of protection saving the costs of putting non-essential assets in SCIF (Sensitive Compartmented Information Facility) rated facilities that are ICD-705 compliant with a UL-752 ballistic rating. For power companies, the buildings can be NERC-CIP-014 v5 in addition to the ICD-705 compliance.

Risk aversion is a consideration for the location of assets, as well. Companies used to perform risk calculations to determine data center locations. Velocity's engineering has provided a solution for risk averse locations. The secure modular buildings are safe from tornadic, hurricane, earthquake and damage up to their ratings. Each tightly controlled facility that can be constructed on site in little time.

THE SOFTWARE DEFINED POWER ADVANTAGE

With an application first mentality, an entity can determine the needs, resiliency and uptime needs for each application, not the overall data center. Reviewing which IT resources self-heal through intelligent software and hardware saves on the need for overly redundant and overly provisioned power resources that are applied to any one application. Orchestration becomes intelligent and power aware. Of course, the buildings can be highly redundant as needed. However, there is much work towards rightsizing the solution across both facilities and IT. With the ability to add software defined power to the mix, the efficiency of the solution can be exemplary. Software Defined Power (SDP) truly brings the best of IT and power management together. SDP allows data centers to orchestrate workloads where and when they are needed and adds the missing link of efficiency for a truly software defined data center.

As mentioned, a benefit of owning both the power distribution and the workloads mean that the two factions can work together in a more purposeful manner. With the power being IT aware, peak shaving occurs and real sizing is intelligence based. Servers that are idle can be put in standby or orchestrated to other servers in some cases shutting down entire racks in real time without loss of functionality based on that dynamic, real time intelligence. When high peak usage occurs, hardware provides the "spike" power flattening out the demand and allows the use of standby power when needed.

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY
data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY
data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075

loss of functionality based on that dynamic, real time intelligence. When high peak usage occurs, hardware provides the "spike" power flattening out the demand and allows the use of standby power when needed. Artificial intelligence is adaptive. This allows facilities that thought they were at capacity to operate well past calculated "just in case" limits. For facilities that thought they were at capacity for their power feed, the intelligence and hardware together find that lost capacity often resulting in no additional power needed for the same compute loads. Modular on site data centers can be erected preserving capital already spent and transferring dollars to predictable OPEX spend.

Orchestration helps data centers separate their power infrastructure availability from the application availability allowing them to create a wider spectrum of SLAs out of their existing redundant infrastructure. In fact, the software can completely move loads to another facility based solely based upon power costs should that need arise.

THE COLOCATION ON PREMISE ADVANTAGE

Whether the desire is to have control of the power and cooling, IT and data, latency, or all of the above, the solution is no longer either colocation or on premise. Those that wish to have a predictable monthly fixed OPEX spend can do so with a Velocity modular build data center or two. Those looking for onsite smart hands have long had that capability. Smart cities can spin up multiple buildings around the city footprint to handle low latency traffic. Construction time is within 90 days of permitting, sometimes faster. The concrete data centers can be erected through capital expenditure or as data center as a service OPEX spend. In fact, if it is of interest, Velocity has partners that will take over the entire management of the both the facility (building, power and cooling) and the IT kit inside. Velocity partners can act as a broker for fiber, telecommunications, cloud outposts and other services that will be utilized. It is a true DCaaS (Data Center as a Service) model backed by construction as a service. By adding SDP, the facility(s) can be a mastery of energy efficiency.

THE MONETIZATION OPTION

Another option to offset the cost of the colo on premise, is to turn spare capacity into a revenue driver. Enterprises that require a slightly smaller footprint can lease out spare capacity to edge providers. EDJX is one such provider that can monetize the extra space by placing their cabinets and connectivity in the space. Their model creates an edge mesh network of sorts in other owners' spaces creating a distributed edge computing platform, allowing industry-first access to a secure platform for IoT, M2M, and Mobile edge apps. The software platform coupled with EDJX Nanoserver Infrastructure enables developers to write, test and deploy their own computing to the near-edge, and to meet the demand for data processing that will serve billions of IoT and industrial IoT devices. This technology can be used for video-as-a-Sensor, Audio-as-a-Sensor and Smart Intersections.

Serverless Edge allows customers to build serverless IoT applications at the near-edge, at ultra-low latencies to IoT devices and connected things. Edge devices will be everywhere bringing data centers to non-traditional data center cities. Adding services and providing space at the edge is going to be an attractive business model offsetting the cost of the buildings themselves. Early adopters are going to be the first reaping the rewards through partnerships. This makes colocation on premise particularly attractive for smart cities and retail locations that are seeking additional revenue streams in smaller fixed configurations.

Typical modular configurations range from 4-40 cabinets per building with power ranging from 18-400kW of critical power. Due to the efficient size, several buildings can be located throughout a municipality serving both the original tenant and other monetized occupancies. The buildings are fully engineered with power, cooling, fire suppression and physical security systems included in the factory manufactured on site assembled building. Even the landscaping and security fencing are included. For more information, visit www.velocitydatacenters.com or call 734-323-3075. Begin your colocation on premise project today.

Corporate Offices
102 E Middle St
Chelsea, Michigan 48118

VELOCITY
data centers

Contact Information
Steve.Jacobs@Velocitydatacenters.com
(734) 323-3075