

An organisation's guide on when to pick colocation

...and how to choose the right colocation partner

Supernatural Power

It's easy to dismiss colocation – when organisations host their own servers and other IT equipment in specialist data center facilities – as old hat, after all, hasn't the rapid development and adoption of cloud-based services rendered colocation obsolete?

But don't be too hasty in sounding colocation's death knell. For a large proportion of organisations, colocation is still a key – often growing – element in their overall IT operations; complementing or even replacing on-premise data center facilities and cloud-based services alike.

In this Verne Global eBook, you'll learn why colocation is gaining in popularity as an alternative to both building your own data center and to the cloud. You'll also find out what to look for in a colocation provider, as well as how to pick the right partner to host your own IT hardware and applications.

Why colocation should play a central role in the future of your IT strategy

Colo defined

In simple terms, colocation (colo) providers offer space in their specialist data center facilities, which businesses and other organisations can rent on an ongoing basis to host their servers and other hardware.

Along with this space, the colocation provider will also offer access to power, cooling, and connectivity services, as well as ensure the physical security of the building. Some colocation providers will also offer 'remote hands' support, so simple maintenance jobs can be taken care of by their team, on behalf of your organisation.

The amount of space you rent will depend on your individual needs. You can usually start with as little as a cabinet or rack, graduating to a dedicated aisle, caged area, or private room as your needs grow. The data center facility will be full of servers and hardware from multiple organisations – hence the word 'colocation.'

Colocation providers will, of course, differentiate their services depending on what levels of security they provide, the uptime guarantees they offer, availability of power, network connectivity options, and the facility's location, as well as the fit-out of their data center. As such, there are a myriad of colocation options now on offer, with one to meet the exact requirements of nearly every organisation.

Crucially, colocation services allow you to treat your data center operations as operational expenditure (OPEX), where you pay for what you need on an incremental, easy-to-manage, and easy-to-predict basis. There's no upfront capital expenditure to fund or worry about.

Colo: The alternative to build-your-own

Colocation started to gain in popularity in the 1990s as a more affordable and flexible alternative to on-premise data center facilities.

It's easy to understand why; as IT became more complex and strategic to everyday business operations, and we all started to become hooked on the internet, colocation providers offered access to state-of-the-art facilities, with none of the capital outlay or hassle associated with designing, building, maintaining, and upgrading your own purpose-built facility.

The reasons to pick colocation over build-your-own are manifold and perhaps even more valid today than they were when the industry first established itself all those years ago.

Here are some of the main benefits of choosing colocation over your own in-house data center facility.



CAPEX avoidance

Designing, constructing and then equipping a data center is, quite frankly, eye-wateringly expensive. In addition to the not-so-little matter of acquiring land or a suitable building, there's the cost of project managing the build and the expense of fitting out the facility so it's optimised to host your equipment and the applications that run on it. That means negotiating contracts for reliable power supplies and connectivity, plus investing in the necessary cooling, fire suppressant, and physical security equipment that's needed for a facility to run around the clock.

This all contributes to a long and pricey shopping list, and businesses often have to take on substantial amounts of debt well in advance of the facility being up and running. For many firms, these figures simply don't add up.

By choosing to collocate your IT equipment in a specialist provider's premises, you're eradicating this CAPEX from your balance sheet. Here, you pay a monthly fee to keep your servers and other equipment with a third party, and this typically includes power, security, and cooling, as well as any other factors that are included in the contract.

This makes the cost of managing your data center portfolio completely predictable and manageable. There's less debt and less complication. Rather than becoming a data center operator by proxy, your organisation can remain focused on its core business activities.

Colo: the build-your-own alternative

Ability to innovate

The rate of technological advancement in the data center has probably never been faster. More efficient cooling, more sustainable and predictable power supplies, and lightning fast network connectivity have never been more in demand, as organisations pursue their digital transformation plans, make greater use of high-density IT equipment, and explore the capabilities of compute-intensive applications such as Machine Learning and Artificial Intelligence.

If you're running your own facility, you're faced with a constant cycle of upgrades in order to support the latest in IT hardware and application developments, while it's likely that you'll also be under pressure to minimise the amount you spend on power. You may even need to reduce your power usage in order to meet sustainability and environmental objectives.

What's more – in order to future-proof your operations – your data center must be able to support your compute requirements of tomorrow, not just today. This requires taking a punt on the right equipment in advance of needing it; a decision that is not without risk.

Third-party colocation providers find it easier to innovate. First of all, they are data center experts, so are perhaps better placed to navigate the various technologies that promise to make facilities faster, more efficient, less impactful on the environment, and secure. In addition, they enjoy economies of scale that organisations running their own operations could simply never achieve. If they invest in the latest liquid cooling solution, they can balance the cost of this equipment against all future customer contracts; making the upgrade more accessible, more affordable, and less risky.

Scalability and flexibility

As we've already touched on, organisations that run their own data centers are continually forced to predict their future requirements. If they get it wrong, their operations could be in danger of outgrowing their facilities; this makes it necessary to build data centers with more space and access to power than they need at that very point in time. Of course, there's always the possibility that they will never use all this spare capacity and that some previous investments will be wasted. For locations where land is at a premium, this approach makes little, if any, sense. There are very few CFOs who will sign off on an investment for high specification data center space in metropolitan areas 'just in case' their organisation needs it at an unspecified date in the future.

Colocation is much simpler to scale. An organisation can rent as little as one rack in a data center, taking on more space on an incremental basis as their operations grow.

In the event that their needs change, they can also – subject to the contractual small print – reduce the amount of space they rent.

This degree of flexibility is far more in keeping with today's 'on-demand' culture for procuring IT services. What's more, it mitigates risk and alleviates the need to build out facilities that are over-capacity.

Facilities that match individual applications

If an organisation operates its own facility, it's likely that most - if not all - of its applications and equipment will be housed there. After all, who wants to pay someone else rent when they have plenty of room at home? There is, however, a hidden cost to this approach. Different applications require different data center environments, all of which come with different price tags.

Take, for example, a retailer that is capturing personal data, taking payments and tracking deliveries on its customer-facing website. This activity is highly sensitive and closely regulated. The data that's being collected and stored needs to be held in a highly secure facility, while it's equally vital that its website is available around-the-clock, otherwise it could hit sales and profits. This type of activity requires high-spec data centers with stringent compliance controls and guarantees around uptime and security. You can read more about this class of facility – known as a tier 3 data center – on XXXX. The same company might also be running a number of back office applications that aren't quite so mission critical. For example, this could be batch processing of quarterly stock levels, or a payroll application which runs overnight, once a month. These types of applications can run optimally without the bells and whistles of a top tier data center, indeed, they could run just fine in a tier 2 or even tier 1 facility.

The problem that on-premise data centers have is that they have to provide an environment that is capable of supporting the company's most demanding application, even if the majority of their apps can function without this level of provision. You wouldn't pay for a Formula 1 pit crew to change the tyre on a family car, so why take this approach with your applications?

If you take the colocation option, you can elect to host different applications at different third-party facilities, each located in an optimal environment. A few colocation providers, like Verne Global, provide different tiers of service within the same facility, albeit in different data halls. This is more convenient still, as organisations can get access to different tiers of service, all from a single supplier.

Connectivity options

Being well-connected is a prerequisite of any data center. High-capacity, reliable connectivity is vital to the smooth running of an organisation's operations, while it's also critical to have both primary and secondary network connections in the event of a planned or unexpected outage.

This is possible to achieve in any data center, but colocation providers are at an advantage; the sheer number of organisations that choose to host their equipment at their facilities means it's lucrative for multiple network operators to establish Points of Presence (PoPs) at these facilities. As a result, colocation customers can pick from a wide range of connectivity options, selecting the providers and services that best suit their needs in terms of capacity, contention rates, routes, service level agreements (SLAs), and, of course, tariffs.

Think of a colocation facility as a readymade market place. Multiple tenants attract multiple connectivity providers, attracting more tenants and so it continues. This effect keeps network costs comparatively low, particularly when compared to connecting your own in-house data center.

This diversity of network operators – and the routes they offer – is particularly beneficial to growing organisations. The networking options available in typical colocation facilities will make it easy for them to reach new markets when they expand their operations internationally, either by setting up new offices or by winning new customers.

Colo vs Cloud

Many of the advantages of colocation are equally - if not even more - true of the cloud, which poses the question: Why not miss out colocation entirely and just migrate your on-premise operations directly to the cloud?

It's easy to see why this argument holds water. Cloud-based services offer unparalleled flexibility and scalability as they enable organisations to procure software, hardware, and platforms – rather than just space, power, and cooling – on an 'as-a-Service' basis.

However, hold your horses....in certain circumstances, colocation remains a better option. The 'cloud-first' mantra being heavily promoted by cloud providers, both large and small, isn't right for all organisations all of the time.

Here's when colocation still has a vital role to play.

Dedicated equipment for peace of mind and greater security

This might seem like an emotional reason but for some organisations there's something to be said for being able to see their own equipment. This provides a hard-to-define level of reassurance that they're in charge of their operations. It just feels right.

While some companies will like the idea and simplicity of handing over the management of their entire IT estate to a third party, for others this is too big a leap of faith. If your business is risk averse, running mission-critical applications, or if your whole business model depends on the apps running on these servers, would you really want to entrust them to a cloud provider, particularly a hyperscaler, where you are just one of millions of customers?

Of course, colocation providers are third party suppliers too, so you do have to be prepared to put your trust in someone from outside your own organisation. However, they will provide you with a highly secure and certified environment, including round-the-clock video surveillance, secure loading bays, 24/7 staffing, and segregated areas which mean only you can access your equipment. You will be allowed access in order to troubleshoot problems, patch equipment, undertake maintenance, and of course undertake upgrades. As mentioned previously, some colocation firms will even take care of some of this 'remote hands' maintenance for you.

On a technical level, there's another factor in colocation's favour. Only your software runs on your server – you're not sharing it with other unidentified customers who – for all you know – could be ripe for a denial of service attack that could bring down the entire stack. Plus, your entire suite of applications can run on a single or cluster of servers. It's not virtualised across multiple servers, which for all you know, could be in different data centers, different countries, or even different continents. When it comes to more specialised compute, such as high performance computing (HPC) and grid computing, dedicated equipment has a particularly important bearing on application performance.

You choose the location

As the name suggests, the cloud is nebulous. Your cloud-based applications and hardware could conceivably be anywhere in the world. While cloud service providers are prone to make a big deal of promoting that they have data center estates in, say for example, the EU (where data protection laws are comparatively tough), many remain opaque about where your apps, equipment, and data will actually reside.

The trend towards virtualisation means that the infrastructure underpinning your operations is likely to be split between various locations, each with their own exposure to risks. Some of your business-critical applications could be hosted on servers in areas exposed to natural disasters, or on the edge of a city with an aging and unreliable power grid. With the cloud, these risks are often unknown, yet they could endanger the resilience of your IT infrastructure and, with it, your business.

With colocation, you know exactly where your software and hardware reside because you pick the exact spot. While there will of course be risks, at least you can assess what they are and mitigate against them in an informed way.

Data sovereignty

On a connected point, data sovereignty is also an important factor to consider. If you're a business operating in the EU, you must comply with the General Data Protection Regulation (GDPR). This sets

stringent controls over how your data is secured and shared – if your apps find their way to data centers outside the EU, can you be 100% sure they're compliant with GDPR?

With colocation, you can be confident in your answer.

Optimised environments

Colo providers are also better placed to offer data center environments that are optimised to your specific applications.

Much like on-premise facilities offer a one-size-fits-all solution, so too does the cloud, where your applications and software will end up sitting on homogenised servers that are designed to fit the generic hosting needs of all types of applications – from email right up to specialist high performance computing (HPC) workloads. In contrast, colocation allows you to run your own hardware, so you know it's optimised for each of your individual applications.

Dedicated support

When it comes to cloud support, prepare to feel like a very small fish in a vast ocean. That's unless you're a massive multinational organisation with enough budget to pay for your own dedicated account team. Everyone else could find themselves 10 or even 20 steps away from an actual human being who is on hand to help if something goes amiss. Be prepared for a mystifying array of online FAQs, bots, user forums, and countless other online rabbit holes of information and misinformation.

With colocation, you're much closer to real help. While there will of course be automated processes to help you troubleshoot problems or get advice, you should expect to be able to reach a person 24/7 every day of the year. This means you'll receive much more tailored support, no matter how small your operations.

Environmental transparency

Be wary of cloud providers that say they use certified green energy, or can at least offer that option. In reality they could still be connected to grids powered by fossil fuels and will revert to this dirty energy if the need arises. And it will. After all, the sun doesn't always shine enough and the wind doesn't always blow.

There's more transparency with colocation providers. You can find out which grids feed their facilities and how they generate their energy. If sustainability is a key goal of your organisation, you'll be able to find locations that guarantee they only use 100 percent renewable energy.

Achieving the right mix

While colocation offers considerably benefits when compared with both in-house data centers and cloud-based services, many businesses use these different solutions in combination.

Some may have invested heavily in their own facilities so it simply isn't economically viable to migrate their equipment to a colocation provider, at least until these investments near end-of-life. Others – traditional banks and financial services companies are good examples here – may consider their IT estates so fundamental to their operations that they simply can't entertain the thought of passing control to any type of third party. Onerous compliance obligations strengthen their argument to keep things in-house.

On the flipside, some organisations will have unpredictable compute demands; perhaps their business is seasonal in nature or they might be growing at such a rate it makes it difficult to forecast their future needs. Here, the cloud might be a more attractive proposition because they can burst or grow their capacity without investing in hardware.

But, more often than not, an organisation will be running different software applications that are best supported in different environments; on-premise, colocation, and cloud.

Truly business critical workloads could remain on-premise, gradually migrating to colocation facilities as they outgrow their original homes, or in order to take advantage of particular benefits, for example, access to renewable energy sources or specialist environments that are optimised for HPC, AI, machine learning or grid computing.

If these workloads become less predictable because there's been a peak in activity or demand, organisations can then make use of the cloud. Indeed, cloud-bursting is becoming an increasingly popular way to supplement in-house and colocation based IT infrastructures.

A combination of on-prem, colo and cloud provides ultimate flexibility to organisations, allowing them to choose the optimal data center environment for each application.

Picking the right colo partner

If you have decided that colocation should be part of your IT infrastructure strategy, the next decision to make is which colocation provider to partner with. There is considerable choice. To help you navigate the options, here are some of the key factors to bear in mind.

Data center tiers

The global data center industry uses a standardised tiering system defined by the Uptime Institute which accurately describes the resilience of the infrastructure provided in each facility. Every data center is categorised in one of four tiers, with tier 1 being the most basic and tier 4 the most sophisticated. You should select the tier that best meets your requirements for resilience and uptime.

Lower tiers might represent too big a risk to your business, while higher tiers could be an over-investment. In practice, tier 3 data centres are largely considered 'enterprise-class', and suitable for the majority of mission-critical applications running in organisations today. Tier 4 facilities are highly specialist and designed to support extraordinary compute. As such, they are only necessary in a very small number of instances, for example, to support highly classified government or military operations.

Data center tiers explained

Tiers 1 & 2

These data centers have limited redundancy, so if something goes wrong, you can expect a service outage. Some may only have a single supply of power and cooling, although there may be an uninterruptable power supply (UPS), provided by onsite batteries or a generator. As a result, they offer very affordable colocation services.

* Expected uptime: 99.671% - 99.741%

* 22 to 28.8 hours of downtime per year

* Ideal for industrial-scale high performance computing, including: data analytics, cryptocurrency mining, machine learning workloads, AI models, and film rendering

Tier 3

These have multiple paths for power and cooling. These will either be described as N+1, which means there is one redundant path, or N+2, which has two back-up paths for additional resilience. These facilities will also have systems in place to update and maintain data center operations without taking it offline.

* Expected uptime of 99.982%

* 1.6 hours of downtime per year

* Suitable for the vast majority of applications run by enterprises today, including companies with a worldwide, 24/7 business or web presence; organisations that could be financially penalised for downtime; e-commerce firms; call centers; and VOIP providers

Tier 4

Tier 4 data centers are the gold standard of the industry, and are built to be completely fault tolerant and have redundancy for every component. This makes them comparatively expensive; for most enterprises, they represent an over-investment.

* Expected uptime of 99.995%

* 26.3 minutes of downtime per year

* Suitable for highly-classified applications such as those used in military, intelligence and government



Location: convenience vs risk

Data center location is key. You may want to regularly visit your equipment to perform maintenance or upgrades; as such it might be good to select a location that's proximate to your business.

However, there are more things to consider than your engineering team's commute.

Data centers in overcrowded, metropolitan areas – which are handily located for major business districts – are comparatively expensive, and could also be relying on outdated, overloaded power grids. This is a particular challenge in some Western European and North American cities, which are overcrowded and reliant on aging infrastructures. The upside is these metro area facilities are likely to have plenty of connectivity options.

By contrast, if your equipment requires less day-to-day maintenance, it could be hosted in a more remote facility. These are less convenient for everyday access and travelling, but have other advantages which we'll cover in the following sections.

Finally, consider the inherent risks of choosing a particular location: is it prone to natural disasters or extreme weather conditions? Could the area be subjected to a terrorist attack or civil unrest? These are big questions to answer, but it's important that you weigh up all the risks, however unlikely.



Networks and latency

As mentioned above, colocation facilities in or close to urban areas – often known as edge data centers – tend to offer a wide range of connectivity options and their proximity to major business districts means that latency is low; this is ideal for time sensitive applications, such as high frequency trading (HFT), video streaming, or gaming. That said, most colocation providers will be able to provide access to multiple, diverse routes and the vast majority of applications (probably 80 to 90 percent) are not latency sensitive, making distance less important.



Contract terms and lock-in

An advantage of colocation is its flexibility, but make sure it's not just theoretical. Be sure to check just how flexible your contract terms are in the event that you want to scale your operations up or down. If you do grow, will your colocation provider be able to facilitate that growth?



Power sources

Power costs make up a large proportion of the overall cost of colocation services. In countries or localities where power is expensive, you can expect the costs of colocation services to be higher too.

This might not be a big issue if you are using colocation to support low-to-medium grade enterprise applications and workloads, however, if you are an HPC user, power costs will represent an important variable in your purchasing decision. You might want to consider colocation providers based in countries utilising large amounts of lower-cost renewable energy, such as hydro-electric and geothermal. For example, colocation in Iceland – with its abundance of renewable energy – is typically 70%-80% cheaper than in equivalent facilities in the UK, Germany and other Western European countries. Because HPC applications aren't usually latency sensitive, there's no problem basing them in Iceland or in remote data centers in other countries.

More than cost, there's also the reliability of the grid to consider, as well as the availability of renewable energy. While the data center industry as a whole is committed to being as efficient as possible, there's still a very real chance that your facility could be drawing power from plants burning fossil fuels. This could be a deal breaker for many organisations, particularly those committed to sustainable practices.



Climate and cooling

The energy that data centers consume doesn't just power the IT equipment, it's also used to power auxiliary services, including the cooling systems that these facilities need in order to stop servers from overheating and malfunctioning. While data center design has improved no end – for example, by separating hot aisles from cold aisles – nothing is as important as climate.

If your data center is located in a hot country, expect an uplift in what you pay to account for the extra cooling that's required. To counter this, many organisations are starting to migrate their data center operations north, with the Nordics a popular choice because of its year-round cold climate. This means facilities can be cooled naturally, using nothing but filtered fresh air.

Taking an application-centric approach

This ebook should have given you a good understanding of the benefits of colocation services, and how they stack up against emerging and established cloud-based services, as well as against traditional, on-premise data centers.

For most organisations, a mix of colocation, cloud and on-premise solutions will be the appropriate way forward, with the exact combination depending on how risk averse you are, the nature of your business, and the regulatory framework you must adhere to. But perhaps the most important thing to remember is to pick the right hosting solution for each of your applications. As we've explained, some apps operate just fine in the cloud but others you might want to keep in house, with colocation providing a compelling third option, keeping CAPEX low and providing flexibility, without forcing you to cede control of your IT operations and equipment.

Moreover, you can even pick different colocation providers and locations for each your applications. If they are latency sensitive and need to be proximate to where you operate, you'll find there are plenty of options. If you just need a reliable, affordable power supply to run non-time sensitive AI models for a research project, that's not a problem either. The breadth of offerings in the colocation market means there will be a solution for pretty much every organisation, together with every one of their applications.

About Verne Global

Verne Global delivers data center solutions for high intensity computing, engineered for optimal high performance compute and built upon 100% renewable energy. Our clean grid and stable climate cuts costs and energy usage, and our expert team provides on-site, around-the-clock support to maximise performance and flexibility for customer workloads. Founded in 2012, our Icelandic data center campus powers some of the world's most innovative and demanding industries, including manufacturing, financial services, earth sciences, life sciences, engineering, scientific research and AI.

**Make an impact
from the ground up.**

info@verneglobal.com | verneglobal.com