HOW TO SET UP YOUR OWN HPC

Colocation choices UK

HOW TO SET UP YOUR OWN HPC
(AND WHY COLO MEANS YOU ABSOLUTELY SHOULDN'T)

High-performance computing is the latest highly desirable addition to enterprise IT, but should you run it in-house or use colocation facilities?

High-performance computing (HPC) has well and truly left the pages of science fiction to become a mainstay of many organizations' business intelligence functions. As hardware continues to shrink in footprint, and the advent of machine learning heralds a world where years of unstructured data can be uploaded and turned into a past, present and future forecast for almost any aspect of the enterprise, more and more companies are taking the plunge with their own in-house supercomputer in a bid to tackle ever-more complicated algorithms and to increase profitability. But is that really the best way?

Smaller, they may be, than in even the recent past, but supercomputer installations still require large amounts of space, as well as particular power and cooling requirements. Increasingly, therefore, businesses are turning to colocation, as a way of gaining access to HPC systems while someone else takes care of the computers' physical needs. But which is right for you?

My space

The first consideration, of course, is whether you have sufficient space to house a high-performance system. Given that these machines will be many times more powerful than anything else in the corporate data center, it's not just a matter of racking another box in the server room. For a start, HPC rigs create a large amount of heat that needs to be managed carefully. Depending on the power of the HPC system, existing cooling systems may be sufficient, but if it requires an upgrade, that's going to mean a lot of work to remove the large volumes of heat generated.

But what if an air cooling upgrade isn't enough? The more advanced systems get, and the more affordable they become, the more likely it is you'll need some form of liquid-cooled system. In reality, the liquid that circulates and recirculates within water-cooling systems.

These systems, which typically snake in and out of the building, are worth considering even if your HPC needs are relatively modest, because trying to retrofit them once a system is in operation is even more complicated, costly and will involve a significant amount of downtime.

Colo removes almost all of these headaches. By purchasing equipment for installation within a colocation data center, such as Vantage's CWL1 South Wales campus, you can be sure that you'll always have enough rack space kept at the right temperature in a perfect environment.

Adrian Wander is a well-known name in the field of HPC with more than 30 years of specialist experience in HPC procurement and tenders. He warns that there's much more to powering a hungry HPC than simply plugging it in. "Don't underestimate the amount of ancillary expense involved in HPC systems. It should always be remembered that HPC requires significant mechanical and electrical kit."

You've got the power

The next consideration is ensuring the system stays powered up. HPCs need a lot of electricity. For example, a system running at around 40 teraflops will require a current of about 2MW, the same as a modest, but not insignificant data center.

This isn't an insurmountable challenge: you can arrange with the provider of your local electricity grid to have a dedicated 'pipe' for your systems. However, that's likely to take some time with the bureaucracy alone and will then involve work on up to three local substations, and cables to be laid.

"HPC systems are power hungry," says Wander. "The current number one supercomputer in the world is running at 414 teraflops and that's drawing nearly 30 megawatts of power. There are very few places in the UK that have that kind of power available. Supplying it to existing machinery may require a significant upgrade to, or extensive updates to, the electrical

infrastructure, not just in the machinery, but downstream through upgrades to local substations".

By now, you're probably starting to add these costs up and accurately drawing some rather expensive conclusions.

With colo, you don't have to worry about any of that. Colo data centers are built from the ground up with both cooling and energy already taken care of and, because your system is sharing floorspace with those of other businesses, the aggregate power usage per machine is reduced, as is the shared cost of cooling, all of which means your carbon footprint is lower, too.

Plus, a colo data center will have been built with future-proofing in mind. So as your energy usage increases, they can simply turn up the power. At Vantage, this is made even easier with a direct, 400kV connection to the UK's SuperGrid.

Good with people

Once you have your HPC environment, you'll need skilled staff to operate it. It's not enough to ask your existing IT department; maintaining a supercomputer is a full-time job. In fact, it's several full-time jobs, because you'll need to have a team on-site, 24 hours a day, seven days a week because when you have a multi-million pound computer system, the last thing you want is for it to be lying idle. You'll want data engineers to check any glitches in the programming, and you'll want hardware engineers to fix any physical faults.

That means you'll be adding at least six highly skilled and specialised, highly paid employees to your headcount. Given that most of the time HPCs don't require much human intervention, to avoid the machine lying idle you'll need staff that are, for the most part, lying idle instead, but ready to deploy at a moment's notice. These staff could be among some of your highest paid, yet they'll be acting as security guards.

"HPC systems are rather specialized," says Wander, "and you'll need to hire and train staff to operate and maintain them. In addition, the mechanical and electrical kit is more complicated, typically involving water cooling. Everything requires significant skills in the infrastructure support team."

Colo solves this problem. The data center will have a whole complement of highly skilled technicians on-site who can get to a malfunctioning element of your system in minutes, meaning that if anything goes awry, you've got plenty of backup. As specialists in the field, colo providers can attract the best talent and put them to work for their clients, including you.

And service level agreements means they'll have a vested interest in fixing all outages and glitches as quickly and thoroughly as possible. Meanwhile, you're saving on wages, pension, holiday pay and uncertainty, while guaranteeing that someone is looking after your interests.

Adrian adds: "What you get is economies of scale. If you have one engineer who is servicing 15 generators for multiple clients, they have more work to do than your in-house engineer serving just one generator, but you still have to pay him the same salary."

The law of surprises

When you have a state-of-the-art system in your building, you don't want any surprises.

That means not only having fail-safes and redundancies built into your system, it also means testing them. After all, you don't want to find out that your backup power supply doesn't work because the primary one failed and your entire workplace has ground to a halt, taking out your email system and all your cloud storage with it.

Yet regular testing has its own challenges. A planned failure is still a failure and so any downtime on your system has to be accommodated and planned for. If you work in a 24/7 business, and most people do, that means planned outages

for everyone, which can be disruptive and expensive, especially if you find a fault that needs addressing.

"With operational HPC systems, people are often reluctant to run fire drills, because in the event that the fire drill fails, you'll generally drop operational workload," says Wander.

By now, you should be starting to realise that taking on your own HPC is not a decision to be taken lightly. But you may be thinking that it's more important for you to have complete ownership and control afforded by an on premise system.

Modern communications make for blazing fast data speeds but, regardless, the further your data has to travel, the higher the level of latency. You could consider a dedicated 'pipe' between the data center and your offices but that will add to costs. "You have to worry about network connectivity – what's the latency? What's the security and diversity of routing? If you go somewhere like the Nordics, how many undersea cables are there between there and here," says Wander.

"In the case of Iceland, it's two, so there's still potential for congestion, and if someone puts something through one of those cables, you've got problems because you've not got a diversely routed network. For a facility in the UK, it's almost certainly going to be fine – I'd be amazed if we ever did have an issue."

A colo arrangement in the UK, for a UK-based enterprise, provides the best of all possible worlds. Suppliers such as Vantage offer instant superfast data transfer – from virtually any communications provider – giving you performance imperceptible from an on-site solution.