VERTIV WHITE PAPER

Develop Future-Ready Data Centers With Disruptive Technologies

Executive Summary

The inextricable rise of the large public cloud service providers (CSPs), together with high-profile M&A between colocation operators, has attracted a lot of recent attention. Arguably, this activity has been to the detriment of investment in enterprise-owned facilities, even though these sites still make up a large proportion of global data center capacity.

Enterprise-owned data centers are unlikely to go away any time soon, according to leading industry analyst, Gartner: "Many enterprises face the dilemma of what to do with their existing data centers. During the past few years, the outward focus of IT has been all about cloud migrations, edge strategies and getting workloads closer to the customer; however, a core set of workloads will remain on-premises," [1] the analyst group states.

Gartner further explains, "These workloads often remain on-premises, because they rely on mainframe or classic UNIX systems; have strict compliance, audit, or latency requirements; or are a set of tightly integrated production applications requiring significant investment (and potential risk) to re-platform." [2]

As well as the issue of what might be described as mission critical applications, new forms of high-density workloads have also emerged. In recent years, high-performance computing (HPC) technologies, including some forms of artificial intelligence (AI), have extended well beyond their traditional uses in scientific applications. HPC-type technologies are now deployed in data centers to support applications in the finance, online gaming, healthcare, film editing, animation and media streaming industries. And there appears to be no signs of slowing down.

While some of these new workloads may reside in colocation or cloud sites, enterprises may also be considering how to upgrade their existing data centers to accommodate the increased processing power and rack densities necessary to run these applications. A related challenge for enterprise data center operators is around what to do with data center capacity that may have been liberated by migrating workloads to cloud or colocation. One obvious option is to shutter legacy enterprise sites, as some companies are choosing to do, but there are also opportunities to retrofit these sites as some may also have the potential to be used as edge sites.

There are a number of retrofit solutions that we believe can effectively prolong the lifespan of existing data centers or future-proof these facilities. Gartner outlined these options in a recent report, "How to Turn Old Data Centers Into Critical IT Assets". As the report states, "Enterprises with existing workloads that will remain in their data centers must decide how best to restructure their physical infrastructures to improve efficiencies and extend the data center's useful life." [3]

Challenges Faced by Older Data Centers

Modern enterprises are tasked with difficult decisions around what to do with their existing data centers. For older data centers nearing occupational capacity, the primary limiting factor usually revolves around a lack of physical space to place more equipment, power to support additional equipment, or adequate cooling infrastructures to maintain the equipment within acceptable operating temperatures. While many data center planners prefer to build a next-generation data center to support long-term growth, or to use colocation, cloud, or hosting services, all the current options necessitate moving workloads away from the traditional on-premises operation, which can introduce risk and add complexity to the operating environment. As Gartner states, a better option may be to upgrade existing sites.

"Although continued investment in an older, more traditional data center may seem contradictory to current trends, if done wisely, it can yield significant benefits to short- and long-term planning," [4] the analyst group reports. When it comes to improving efficiency and extending the life of existing data centers, the following practices are recommended: enhancing delivery, reinvention, and maximizing space.

Enhance IT Delivery and Reinvent Infrastructure With Liquid Cooling and Rear-Door Heat Exchangers

According to Gartner, "The power for cooling a data center can take as much as 60% to 65% of the total power used. Higher density racks of 15kW to 25kW can often require more than 1.5kW of cooling load for every 1kW of IT load, just to create the cool air flow needed to support those racks." [5]

For those businesses like research organizations looking to retrofit data centers for extreme densities in a small footprint, often used for such applications as scientific computing, AI, or ML, liquid and immersive cooling have become viable options. Liquid cooling is extremely efficient as the cooling medium goes directly to the IT equipment rather than cooling the air around the equipment. Liquid cooling can enable the CPUs and GPUs in densely packed racks to operate continuously at their maximum voltage and clock frequency without overheating.

Gartner predicts that "by 2025, data centers deploying specialty cooling and density techniques will see 20% to 40% reductions in operating costs" [6].

That's because organizations looking to deploy extremely high-density racks (>30 kW) in their data centers will likely seek out better, more efficient alternatives to air cooling. No matter how the system is configured or optimized, air cooling will simply be unable to deliver the heat removal capacity required to maintain the energy efficiency of IT systems.

Liquid cooling technology available today has the capacity to efficiently and effectively cool racks of 50 kW and higher. This technology is available in a variety of configurations that use different technologies, including rear door heat exchangers (RDHx), direct-to-chip cooling, and immersion cooling. According to Gartner, "The use of liquid cooling can solve the high-density, server-cooling problem, because water (conductive cooling) conducts more than 3,000 times as much heat as air and requires less energy to do so. Liquid cooling enables the ongoing scalability of computing infrastructure to meet business needs." [7]

A RDHx is a mature technology that doesn't bring liquid directly to the server but does utilize the high thermal transfer properties of liquid. In a passive RDHx, a liquid-filled coil is installed in place of the rear door of the rack, and as server fans move heated air through the rack, the coil absorbs the heat before the air enters the data center. In an active design, fans integrated into the unit pull air through the coils to increase unit capacity.

"One benefit of the RDHx is that, not only do you have more-efficient racks, but much of the power once used for cooling becomes available for reuse by facilities to support other building systems or rerouted as additional IT load," [8] states Gartner. Common RDHx vendors cited by Gartner includes Vertiv.

In direct-to-chip liquid cooling, cold plates sit atop a server's main heat-generating components to draw off heat through a single-phase or two-phase process. Single-phase cold plates use a cooling fluid looped into the cold plate to absorb heat from server components. In the two-phase process, a low-pressure dielectric liquid flows into evaporators, and the heat generated by server components boils the fluid. The heat is released from the evaporator as vapor and transferred outside the rack for heat rejection.

As far as evaluating the effectiveness of incorporating this technology, Gartner says: "Every environment is different, so it's critical that I&O leaders use detailed metrics such as power usage effectiveness (PUE) or data center space efficiency (DCSE) to estimate the benefits and unique cost savings from such investments." [9]

With immersion cooling, servers and other components in the rack are submerged in a thermally conductive dielectric liquid or fluid. In a single-phase immersion system, heat is transferred to the coolant through direct contact with server components and removed by heat exchangers outside the immersion tank.

In two-phase immersion cooling, the dielectric fluid is engineered to have a specific boiling point that protects IT equipment but enables efficient heat removal. Heat from the servers changes the phase of the fluid, and the rising vapor is condensed back to liquid by coils located at the top of the tank. Examples of immersion cooling supplies cited by Gartner include two Vertiv partners: Green Revolution Cooling and TMGCore.

Vertiv's First Liquid Immersion Cooling Solution Was Developed With GRC (Green Revolution Cooling)

In addition to increased efficiency and reliability, here are other ways a data center can benefit from transitioning to a liquid cooling system:

- **Sustainability:** Not only does liquid cooling create opportunities to reduce data center energy consumption and drive PUE down to near 1.0, it provides a more effective approach for re-purposing captured heat to reduce the demand on building heating systems. The return-water temperature from the systems can be 140 degrees Fahrenheit (60 C) or higher, and the liquid-to-liquid heat transfer is more efficient than is possible with air-based systems.
- **Lower Cost of Ownership:** ASHRAE conducted a detailed cost of ownership analysis of air-cooled data centers versus hybrid model air- and liquid-cooled data centers and found that, while a number of variables can influence TCO, "liquid cooling creates the possibility for improved TCO through higher density, increased use of free cooling, improved performance and improved performance per watt."
- **Improved Utilization:** The use of liquid cooling should enable organizations to maximize the amount of compute that can be deployed efficiently into on-premises data centers by enabling increased density per rack, increased kilowatt per square meter, essentially getting more out of existing floorspace than previously thought possible.

Maximize Space With Micro Data Centers and Self-Contained Racks

The density enabled by liquid cooling allows a facility to better use existing data center space, eliminating the need for expansions or new construction, or to build smaller-footprint facilities. It also enables processing-intensive edge applications to be supported where physical space is limited. However, another solution for organizations dealing with HPC applications in tight spaces is a micro data center, which provides an affordable, reliable, IT solution that is ideal when critical, small-footprint compute resources are needed. Micro data centers are solutions that incorporate all the essential components you find in a typical data center, but in a scaled down deployment. Because the entire system is enclosed into the size of one standard IT rack, micro data centers can be used in existing network closets or small server rooms, as well as in open office spaces, retail stores, and clinics.

Vertiv™ Liebert® VIC

It appears that Gartner has a slightly broader definition of a micro data center, which it refers to as self-contained racks. These racks are manufactured enclosures that contain a group of racks (e.g., two, four or six) designed to support medium to high compute densities and will often integrate their own cooling mechanism.

Typically, micro data centers support critical loads of no more than 100-150 kW. Components within the rack usually include the following parts: uninterruptible power supply (UPS), rack power distribution unit (rPDU), rack cooling unit, and climatic controls with integrated heat rejection and remote monitoring sensors and software.

"Gartner has recommended this approach with many clients, where appropriate, and believes most data center planners should include these retrofit options on the list of candidate solutions for improving data center space," [10] the analyst group states. The group also identified Vertiv as one of the relevant suppliers of self-contained rack solutions.

The various micro data center form factors eliminate the complicated legwork involved in specifying the various components of an IT solution that can support smaller critical loads locally and at the edge. The pre-built solution enables fast deployment, usually in a matter of days. They also support standardization across multiple deployments where you have smaller critical loads to support.

Standardizing on a single style of rack simplifies the workload for your IT team and makes maintaining your distributed IT infrastructure much more efficient.

Conclusion

A phased retrofit approach, where appropriate, including technologies such as a RDHx, liquid cooling, and self-contained racks can help improve the lifespan and productivity of existing data center space.

Data centers that incorporate these technologies should experience a great number of benefits, including improved scalability, reduced cooling requirements, and reduced stranded power which can be used for additional IT workloads. However, embarking on a retrofit of a legacy site is probably best done with support from a trusted services partner.

[1, 2, 3, 4, 5, 6, 7, 8, 10] Gartner, How to Turn Old Data Centers Into Critical IT Assets, 12 March 2021, David Cappuccio
[9] Gartner, Your Data Center is Old. Now What?, May 03, 2021, Meghan Rimol

Vertiv.com | Vertiv Headquarters, 1050 Dearborn Drive, Columbus, OH, 43085, USA

SL-70914 (R02/22)