Technical white paper

Check if the document is available in the language of your choice.

COOLING STRATEGIES FOR THE MODERN DATA CENTER

CONTENTS

POWERING AND COOLING THE MODERN DATA CENTER

Today, IT leaders struggle to adequately power and cool IT equipment (ITE) in their traditional and colocated air-cooled data centers. Processors are now exceeding 200W, and graphics processing units (GPUs) are operating at over 300W with components continuing to get even more power-hungry. High-density rack configurations in the high-performance computing (HPC) space are moving from 20 kW to 40 kW—with estimates that they will reach beyond 70 kW per rack by 2022. It is not uncommon today to see custom liquid cooled HPC racks consuming 70 to 120 kW. In fact, power densities tend to double every seven years.

This technology brief first explains the limitations of traditional cooling practices. Then it describes a range of systems you can choose from to modify or supplement your existing cooling system to get the cooling capacity your data center requires.

CURRENT CHALLENGES WITH THE OPEN-AREA APPROACH

Enterprise data centers are most often designed with an open-area approach to cooling racks of servers and storage systems. With this approach, one or more computer room air conditioners (CRACs) are placed on the periphery of the data center. Racks are arranged in a cold-aisle/hot-aisle layout, as shown in Figure 1. Cool air is forced through a raised floor plenum and up through vented floor tiles in the cold aisle toward the front of the ITE racks.

The cool air is drawn through the ITE racks, and warm air is vented out the rear of the racks and up toward the ceiling. Air circulation works on the basic strategy of providing cool air at the floor level and collecting warm air near the ceiling.

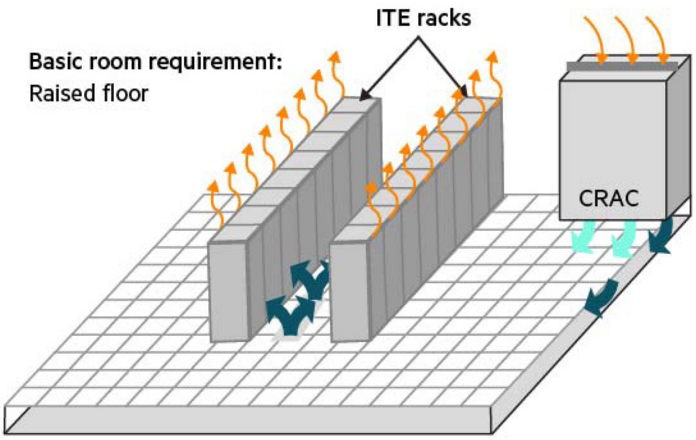

FIGURE 1. Traditional open-area data center aisle arrangement

The open-area strategy is generally adequate for racks using up to 10 kW of power and scales relatively easily to hundreds of racks. This approach can even support a few 30 kW racks, but only if other racks are much lower power or even vacant. Even with this potential, however, there are some important drawbacks. Some of the warm air mixes with the cool air, reducing cooling system efficiency, and other parts of the warm air can wrap back to the cold aisle, reducing ITE efficiency. The typical remedy has been to set the cooling system to run colder to compensate for the hot spot, or to add supplemental cooling or specific channeling.

Multi-node systems and 1U servers let you assemble high-density infrastructures. However, these systems create much more heat per square foot of floor space and more hot spots. In instances like this, the open-area approach cannot keep up with the demand for cool air for more than just a few racks in a row. With increasing ITE capability, now 1U and even 2U servers can start approaching densities that challenge the traditional open-cooling approach.

The increasing heat loads created by the latest server systems require more aggressive cooling strategies than the traditional open-area approach. Such strategies include the following:

- Air containment
- In-row cooling
- Liquid-assisted cooling (rear-door heat exchanger, close-coupled cooling)
- Component-level direct-liquid cooling (DLC)
- Fanless liquid cooling

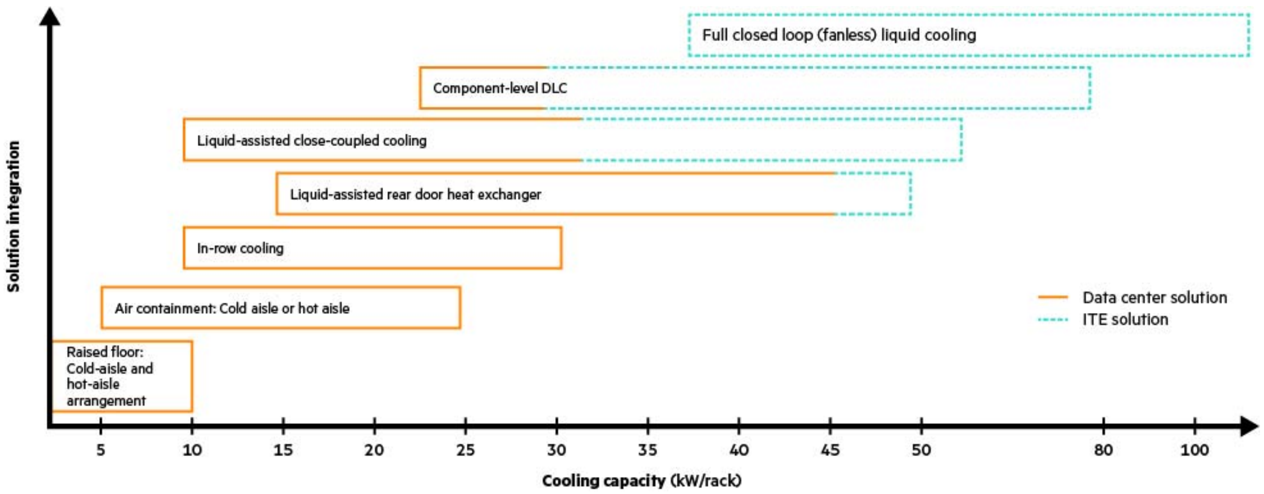

FIGURE 2. Observed cooling strategies and implementations based on kW per rack and on the solution integration

In Figure 2, cooling practices in solid line orange boxes at left work with air-cooled IT. Items in dashed turquoise boxes at right include IT specifically designed or optimized to be liquid cooled. Those items with both outline types support a mix of both technologies. Essentially, once you cross the 50 kW barrier, you will be implementing solutions that are specifically designed for liquid cooling.

Table 1 contains examples of each technology along with its benefits, considerations, and requirements. Subsequent sections provide more detail is on specific implementations.

TABLE 1. Benefits, considerations, and requirements of various cooling strategies

	Traditional air cooling	**In-row cooling**	**Liquid-assisted cooling**	**Component-level direct-liquid cooling**	**Fanless liquid cooling**
Benefit	• Air cooling with CRAC located outside of the rack row	• Moving liquid to the row • Can stay on 208 VAC power with 350 kW per row and typical data center pitch	• Recapture power density with racks that can support up to 55 kW each • No data center air containment required	• Data center voltages can remain the same • Transfer stranded chiller power to compute power • Increased power density • Top-end processor support • Lower operating expenses	• Ultimate performance density and efficiency • Large-scale units to amortize equipment
Considerations	• For the next five years, this will only be valid for low thermal design power (TDP) processor SKUs	• Perceived loss of node density since in-row coolers take up space • Mixing of data center air leads to inefficient utilization of the cooler	• Perceived loss of node density since in-row coolers take up space • Deeper racks or added row length for heat exchanger	• Row power capacity must be increased to fully take advantage of increased power density	• Row-based cooling distribution units (CDU) required • Data center must utilize high-voltage, three-phase A/C power
Requirement	• Most data centers as they exist today	• Requires air containment • Data center upgrade to add facility water	• May require data center upgrade • Requires facility water	• Requires facility water at each tile	• Requires facility water

AIR CONTAINMENT STRATEGIES

Air containment strategies separate the cold supply air from the warm return air to maximize air handler efficiency. When containment is used, it becomes easier to add more perimeter CRACs and use more open floor tiles and plenums. These benefits can extend air cooling capacity to 25 kW per rack. Selecting an air containment strategy generally depends on whether there is a raised floor or dropped ceiling, the ability to meet fire codes, and the temperature profiles of the cooling air entering and exiting the ITE. You can use either hot-aisle or cold-aisle containment.

- **Cold-aisle containment** is generally the easiest strategy to implement as it just requires a panel/door at the end of each row and a panel straddling the adjacent cold-aisle faces. This allows the air leaving the CRACs to feed the coldest temperatures to the ITE at peak pressure and with no hot air mixing, making the ITE run as fast and as efficiently as possible. As most of the supporting infrastructure—power cabling, Ethernet cabling, and high-speed cabling—is in trays over the hot aisle, cold-aisle containment also does not have to account for them.
- **Hot-aisle containment** does not mix hot and cold air, so that only exhaust air returns to the CRAC, allowing the air conditioning system to operate more efficiently than with cold-aisle containment systems. As ITE density has increased, rack exhaust temperatures have risen, making it impractical for extended periods in this type of heat, so most high-density deployments have standardized on hot-aisle containment. Good cold-aisle tile management and pressure combined with this approach can also deliver the coldest air to the ITE.

Figure 3 is a classic example of the simplicity of high-density hot-aisle containment. Note the space looks identical to the open raised floor environment, with the addition of doors on the end of row of adjacent hot-aisle faces, and a curtain to duct that flow into a dropped ceiling return air plenum going back to the CRAC. The frequently traversed cold-aisle environment is easy to work in at 75°F, and the >100°F hot-aisle environment is accessed only as needed.

FIGURE 3. Hot-aisle containment inside BP's state-of-the-art Center for High-Performance Computing (CHPC) in Houston

BRINGING COOLING CLOSER TO THE RACK

As the number of servers in data centers increases, IT professionals are looking for alternative cooling solutions to handle increasing power densities. As illustrated above, not every data center will take the same path, and it is unlikely that the journey will happen one step at a time with stops in between each step. For example, you might start with containment and supplement with in-row cooling to raise density, cooling capability, and efficiency. Or you could go straight to in-row cooling and essentially move the CRAC into the row, eliminate the CRACs at the perimeter, and thus eliminate the raised floor and its ducted supply and return plenums. Figure 4 is one such example.

FIGURE 4. In-row cooling at the Massachusetts Green High Performance Computing Center (MGHPCC)

LIQUID-ASSISTED COOLING

With liquid-assisted cooling, rear-door heat exchangers (RDHX) replace standard rear doors on rack enclosures. The heated air leaving the ITE is then passed through a liquid-to-air heat exchanger (L2A HEX), transferring heat to the liquid, while the RDHX fan delivers cool-neutralized air flow back into the data center.

Close-coupled cooling systems have separate cool air distribution and warm air return paths that are isolated from the open-room air. Closed-loop systems typically use L2A HEX that use chilled water for removing heat created by ITE. HPE offers a close-couple cooling system, called the HPE Adaptive Rack Cooling System (ARCS). Examples of both liquid-assisted cooling solutions are highlighted in Figure 5 below.

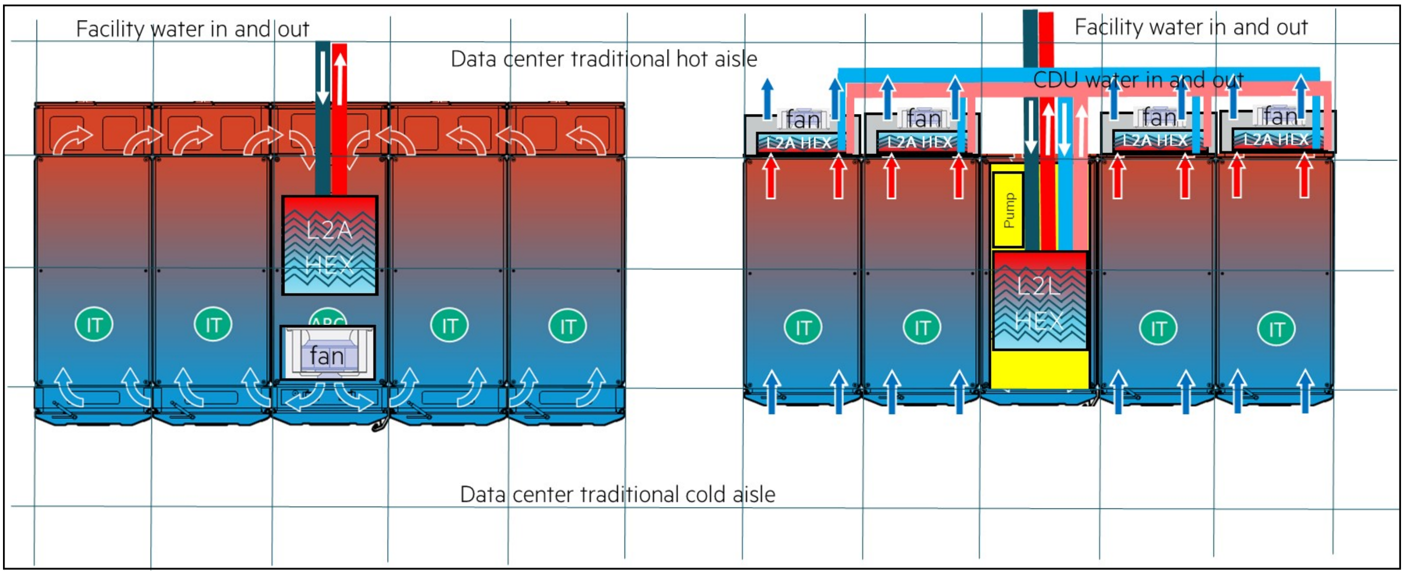

FIGURE 5. Liquid assisted cooling: ARCS at left compared to RDHX at right (Top view)

ITE requires cool air at the front and discharges heated air at the rear. In Figure 5, in the ARCS system, the air flow on the left follows front-to-rear, but all the ITE air does not require facility input air nor interact with facility return air. In ARCS, the entire heat load is carried by the water, thus no heat load enters the room. The RDHX example on the right also shows front-to-rear cooling of the ITE, but the noticeable difference is that facility air enters and leaves the rack. The heated air is cooled by the RDHX before leaving the rack so that it produces no net heat increase in the room. However, this solution requires room air movers to support the full air flow capacity of the ITE, whereas ARCS does not require air movers and can be installed in a non-data center environment.

Another differentiator between these two technologies is the placement of the exchanger in reference to the ITE. In an ARCS deployment, a wide range of facility liquid temperatures is permitted because the L2A HEX is not in the ITE rack. Conversely, with RDHX, it first requires facility water to go to a CDU liquid-to-liquid heat exchanger (L2L HEX), which then distributes secondary liquid network to each ITE rack's L2A HEX. This is significant because facilities often have below dew point water, which ARCS can now take directly, without requiring a secondary liquid network.

ARCS will always require one rack space to support two to four ITE racks in order to accommodate the L2A HEX, whereas certain facility liquid conditions may allow elimination of the RDHX L2L HEX. But they would still require fluid connections to every rack. In both scenarios, the rack is typically deeper, and this must be accommodated in the layout of the room. Another liquid-assist technology would put the L2A HEX overhead which could allow for different room arrangements if the space above the ITE does not already have the facility infrastructure installed.

DIRECT-LIQUID COOLING (DLC)

DLC systems are designed specifically for cooling ITE at the component level. They are the best solution for high-density and/or high-power systems where air cooling is not feasible. DLC systems typically use CDUs to redirect IT heat into the facility water supply.

HPE Apollo DLC system

HPE Apollo DLC systems provide a simple, fully integrated plug-and-play approach to optimizing processor performance, increasing data center efficiencies, and regaining data center densities. Factory-installed DLC concepts are available on the HPE Apollo 2000 Gen10 Plus System, which can be combined with HPE ARCS or rear-door heat exchangers to provide room-neutral solutions and further maximize data center efficiency.

HPE Apollo DLC consists of a cold plate and a tube running coolant over the components within the server, pulling the heat away from them. A manifold running the length of the rack brings the coolant supply to the servers, with the hot water traveling to a CDU contained within the rack. The CDU then connects to the facility water, up to 32°C (89.6°F), and water return from the servers. Additionally, the liquid-cooled infrastructure has been carefully designed to support multiple processor architectures and accelerator options while remaining forward compatible with next-generation CPU, GPU, and interconnects. Currently, Apollo DLC has CPU-only cooling on the HPE Apollo 2000 Gen10 Plus System with future support for CPU and memory cooling. Additionally, support for CPU, memory, and GPU cooling is planned on the HPE Apollo 6500 Gen10 Plus System.

FULL CLOSED-LOOP LIQUID COOLING

Deploying systems in this environment does not use data center ambient air. It is fully contained, as shown in Figure 6.

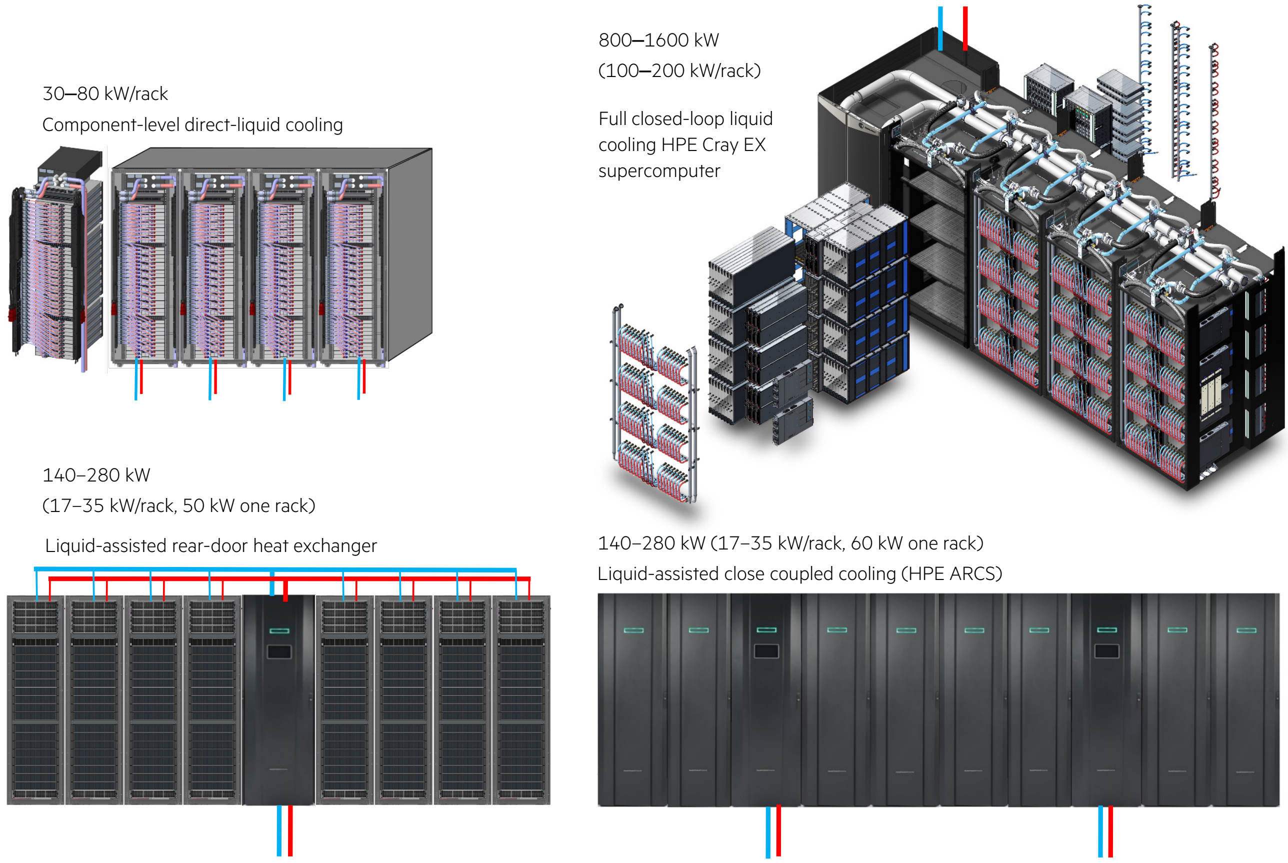

FIGURE 6. Examples of liquid assisted and direct liquid cooling implementations

HPE Cray EX supercomputer

For businesses requiring the greatest performance, density, and efficiency for large-scale systems, the HPE Cray EX supercomputer provides 100% solution-level heat capture, eliminating the need for external air cooling.

The HPE Cray EX architecture contains many innovative features to support the highest wattage CPUs and GPUs (in excess of 500W) using the efficiencies of warm water. The liquid-cooled infrastructure also results in a much more compact system architecture, and thus minimizes the use of more expensive optical interconnect cables over less expensive electrical ones.

HPE Cray EX and Apollo liquid assist

Full closed-loop liquid cooling doesn't use data center ambient air, so the site just needs to bring fluid to it from the facility cooling and doesn't have to provision for CRACs, raised floor, or any type of air moving equipment. Technically, certain in-row coolers and the HPE ARCS by itself can meet those conditions in some facilities. When that is combined with HPE Cray supercomputers and HPE Apollo solutions with DLCs, these can be installed in almost all data center environments.

Immersion cooling

Enterprise data centers have most often used an open-area approach to cool racks of servers and storage systems, as described earlier in this paper.

Immersion cooling takes a much different approach to solve the challenges of system cooling. Immersion cooling is the practice of submerging computer components or full servers in a thermally—but not electrically—conductive liquid, called dielectric coolant. IT hardware cooled in this manner does not require fans, and the heat exchange between the warm coolant and cool water circuit usually occurs through a heat exchanger, such as a heater core or radiator.

HPE servers have been engineered to support the traditional cooling practices deployed in most data centers. As such, there are some system modifications that need to take place in order to use HPE servers in an immersion cooling scenario. For example, thermal grease should be replaced with compatible thermal interface materials. The use of solid-state drives is required in place of hard disk drives for compatibility with the immersion fluid. System fans are removed to prevent impeding the natural flow of the coolant. Also, HPE does not offer a warranty or service products that are immersed.

All of the cooling strategies described above deliver similar or greater benefits to immersion cooling as it is today.

CHOOSING THE BEST COOLING STRATEGY

Which cooling strategy is best for your environment depends on a number of factors, such as facility characteristics, room mapping, power consumed per rack, and server density. As such, HPE does not recommend one specific cooling strategy as the best strategy for all data center environments.

Cooling decisions based on server density/power per rack

Your room characteristics and data center layout can suggest a specific cooling strategy, but equipment density and power consumption ultimately determine the best choice. We can make some general assumptions:

- Traditional data center cooling is adequate for racks using up to 10 kW.
- Racks using 10 kW or more will likely require some form of containment strategy.
- Liquid-assisted cooling can be a balanced approach starting at 10 kW rack, going up to 60 kW per rack, but above 45 kW per rack, purpose-built ITE that can air cool components at that density level will likely be required.
- Closed-loop cooling accommodates the widest range of traditional server/power densities—from 10 kW per rack with legacy IT that may be over 10 years old, all the way up to exascale supercomputers at hundreds of kW per rack.

DATA CENTER SERVICES

HPE Pointnext Data Center Technology Services (DCTS) offers a complete portfolio of consulting services from strategy and planning all the way to commission of a data center. As shown in Figure 7, DCTS offers services throughout the lifecycle of the data center. The services are grouped into five major portfolios, these are:

- **Advisory and strategy:** Services include road map planning, current state versus future state scenarios, IT sourcing options, and total cost of ownership (TCO) analysis.
- **Planning services:** A complete portfolio of conceptual planning services for mechanical, electrical, fire protection, fuel oil, and security systems. DCTS has also tremendous experience in planning HPC liquid-cooled data centers. Recently, DCTS designed some of the world’s largest HPC facilities in Asia.
- **Modular data center and professional services:** DCTS offers a variety of modular solutions through partnerships including secure edge data center (SEDC) which comes in several categories pertaining to capacity and redundancy. DCTS services include site evaluation, construction administration, and bid evaluation.
- **Data center management and operations consulting:** DCTS offers a complete line of services supporting data center management and operational efficiency. Runbook is a complete set of procedures, operational risk assessments, staff and organizational assessments, physical security assessment, continuous improvement program, ITSM service capability, forensic failure analysis, commissioning services, and DCIM consulting services.
- **Assessments and energy services:** A variety of space-power-cooling-air management services such as thermal quick and comprehensive assessments, infrastructure condition and capacity assessments, basic capacity assessments, energy efficiency assessments, water efficiency assessments, energy and sustainability management workshops, and green certifications.

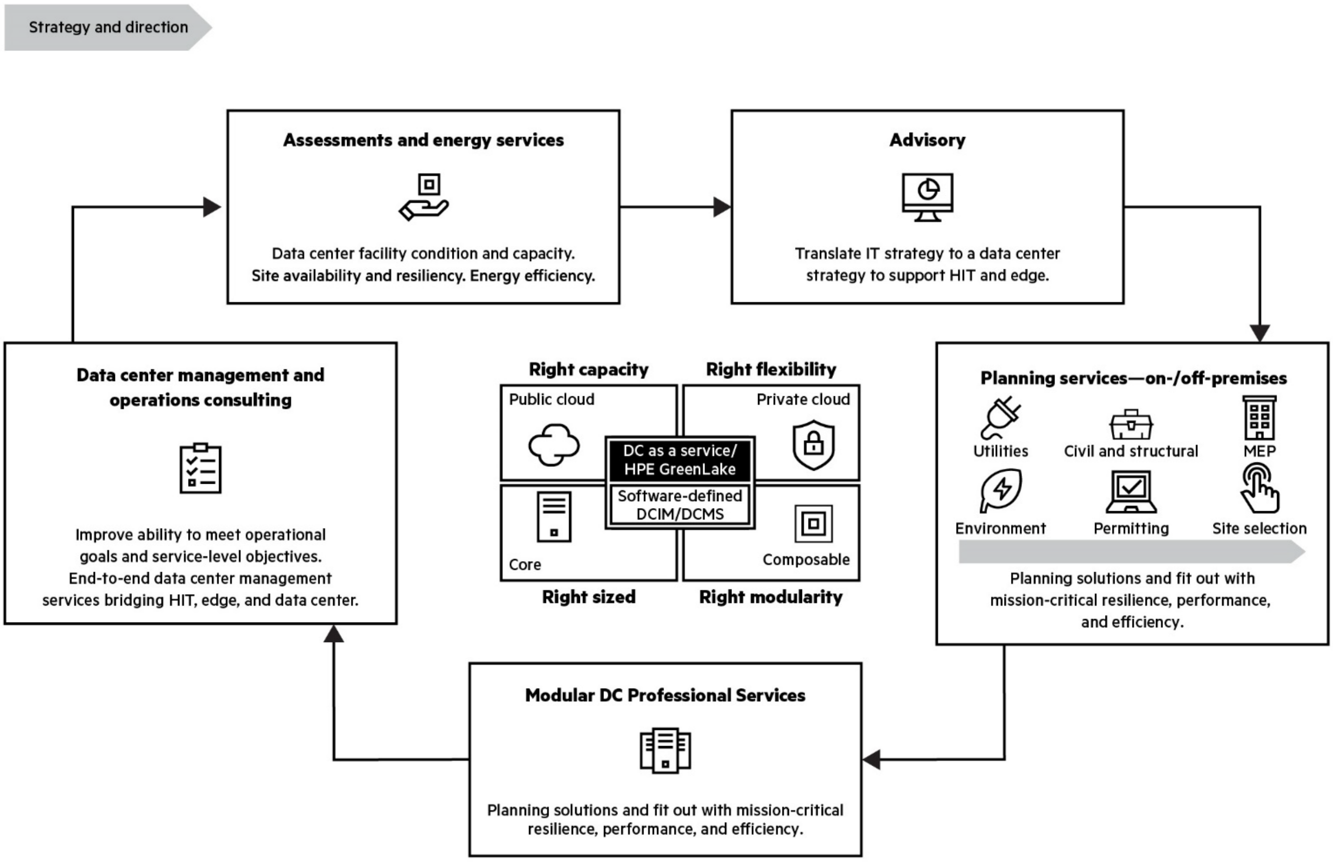

FIGURE 7. HPE Pointnext Data Center Technology Services (DCTS)

MANAGING POWER AND COOLING WITH HPE PERFORMANCE CLUSTER MANAGER

HPC systems require close monitoring and management to prevent system failures and to keep power and cooling costs under control. HPE Performance Cluster Manager is fully integrated system management software for HPE clusters offering administrator monitoring and management of all aspects of their cluster, including power and cooling. The software collects and analyzes power metrics for all hardware—CPU, GPU, rack, chassis, nodes, rack AC, bulk DC, and CDUs. The software also supports HPE ARCS, so users can analyze the power and cooling metrics and react to preconfigured alerts such as water leakage, power supply failure, or overheating.

Figures 9 and 10 show how HPE Performance Cluster Manager monitors thermal status and other indicators of CDUs and HPE ARCS and displays them in Grafana so that users can spot potential issues easily.

FIGURE 8. CDU monitoring screenshot over time displayed in Grafana

FIGURE 9. HPE ARCS metrics over time and displayed in Grafana

FIGURE 10. Cooling device detailed views

In the event of a cooling failure, HPE Performance Cluster Manager will shut the affected parts of the system down to safeguard it from the overheating components. The shutdown is topology- and protocol-aware and staged sequentially.

HPE Performance Cluster Manager also offers the ability to manage and monitor power consumption, which can be used to make decisions on rack/chassis/node power capping to fit within the power and cooling envelope. System administrators can set limits to trigger a power cap based on environmental thresholds such as power or thermal, data center power capacity, or for other reasons such as workloads, planned brownouts, time of day, and so on.

Power and cooling data are stored in a scalable database so they can be analyzed for capacity and planning purposes.

Figure 11 shows the user-friendly interface of HPE ARCS, with at-a-glance information about the system so users can clearly see issues such as water leakage, temperature changes, locked/unlocked doors, and power supply failure.

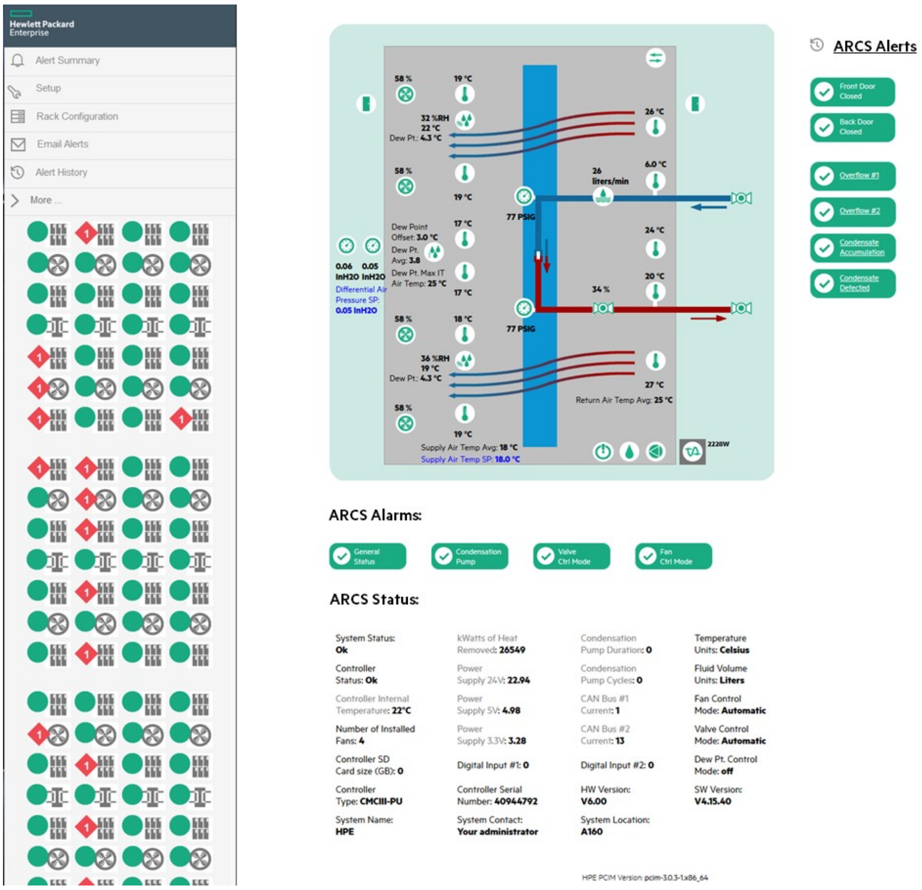

FIGURE 11. HPE ARCS at-a-glance user interface

HPE Performance Cluster Manager is fully integrated with HPE ARCS so users can monitor and analyze the power and cooling metrics and configure and react to alerts alongside the rest of the power and cooling information managed by the software. Please see the section on managing power and cooling with HPE Performance Cluster Manager for additional details.

CONCLUSION

As data centers require more efficient cooling solutions, HPE is working diligently to provide solutions to meet current and future needs. Since conditions vary from one data center to another, we provide a full range of solutions that you can choose from to cool your computing environment efficiently.

FOR MORE INFORMATION

For additional information, refer to the resources listed below.

Resource description	Resource links
HPE Pointnext Services: Consulting, operational, educational, and financial service to accelerate your digital transformation	HPE Pointnext Services
HPE Performance Cluster Manager	HPE Performance Cluster Manager
HPE Adaptive Rack Cooling System (ARCS)	HPE ARCS QuickSpecs HPE ARCS virtual product animation HPE ARCS webpage
HPE Cray EX	HPE Cray EX webpage
HPE high-performance computing (HPC)	HPE HPC webpage
HPE Apollo 2000 Gen10 Plus System (includes DLC)	HPE Apollo 2000 Gen10 Plus System webpage HPE Apollo 2000 Gen10 Plus System QuickSpecs HPE Apollo 2000 Gen10 Plus System virtual product animation

LEARN MORE AT

hpe.com/info/hpc

a00104981ENW, September 2020, Rev. 1