Life Is On | Schneider Electric

Navigating Liquid Cooling Architectures for Data Centers with AI Workloads

White Paper 133

Version 2

Energy Management Research Center

by Paul Lin
Robert Bunger
Victor Avelar

Executive summary

Many AI servers with accelerators (e.g., GPUs) used for training LLMs (large language models) and inference workloads, generate enough heat to necessitate liquid cooling. These servers are equipped with input and output piping and require an ecosystem of manifolds, CDUs (cooling distribution) and outdoor heat rejection. There are six common heat rejection architectures for liquid cooling where we provide guidance on selecting the best one for your AI servers or cluster.

RATE THIS PAPER ★★★★★

Introduction

AI training and inference servers use accelerators and processors with high thermal design power (TDP)[1]. Air-cooling these chips becomes less practical when considering heat sink dimensions, server airflow and energy efficiency, forcing a transition to liquid-cooling. Liquid cooling servers offer benefits including improved accelerator reliability & performance, increased energy efficiency, reduced water usage, and reduced sound level.[2]

There are two main categories of liquid cooling for AI servers – direct-to-chip and immersion[3]. There are slight differences in the heat rejection ecosystem that we will cover. Data center operators and IT Managers unfamiliar with deploying liquid-cooled servers will need to answer a few questions:

- How do I get cold water in and hot water out?
- What is a CDU, and do I need one?
- What steps do I take to select an appropriate liquid cooling heat rejection architecture?

There are three elements (i.e., heat capture within the server, CDU type, and method of rejecting heat to the outdoors) in a liquid cooling ecosystem. A CDU is a system used to isolate the IT fluid loop from the rest of the cooling system and is necessary to provide five key functions (i.e., temperature control, flow control, pressure control, fluid treatment, heat exchange and isolation). There are six common liquid cooling architectures each with advantages, disadvantages, and when to implement as shown in **Table 1**.

Table 1

Common liquid cooling architectures are comprised of a heat rejection method and a CDU type.

Heat rejection method		CDU type
Use existing facility heat rejection system	Reject heat into air in the IT space	using rack-mount CDU
		using floor-mount CDU
	Reject heat into existing facility water system	using rack-mount CDU
		using floor-mount CDU
Create dedicated facility heat rejection system	Reject heat into independent water system	using rack-mount CDU
		using floor-mount CDU

Describing liquid cooling architectures

The coolant distribution unit (CDU) is a key element in the architecture. The CDU is used to isolate the IT cooling fluid from the rest of the cooling system. **Figure 1** shows a simplified view of a liquid cooling architecture from ASHRAE. It shows three loops including the technology cooling system (TCS), facility water system (FWS), and condenser water system (CWS). The FWS loop is considered the primary loop while the TCS loop represents the secondary loop. This brief description provides a hint of how a liquid-cooling architecture discussion could become complex without a logical framework.

[1] See White Paper 110, *The AI Disruption: Challenges and Guidance for Data Center Design* for more information on this topic.

[2] See White Paper 279, *Five Reasons to Adopt Liquid Cooling* for more information on the benefits of liquid cooling.

[3] See White Paper 265, *Liquid Cooling Technologies for Data Centers and Edge Applications* for more information on liquid cooling methods.

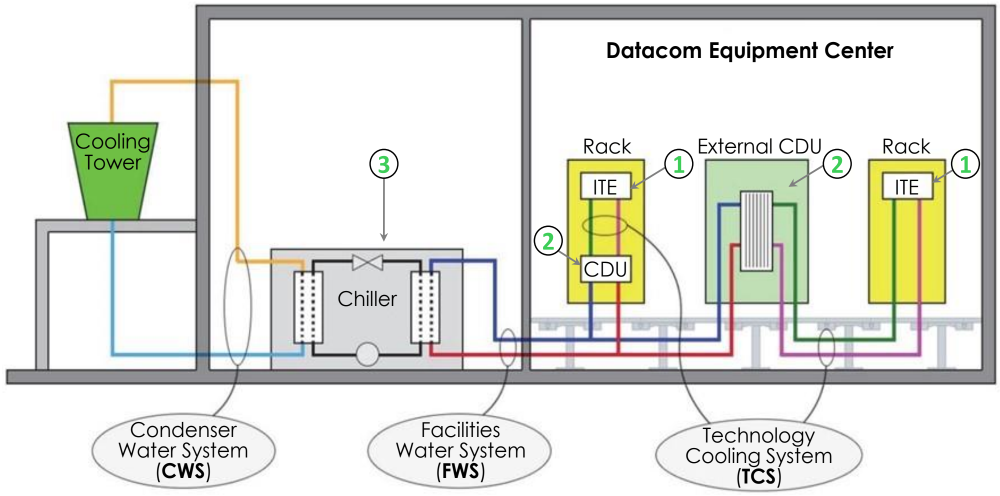


Figure 1

Simplified view of a liquid cooling architecture in a data center

Source: ASHRAE, [Water-Cooled Servers: Common Designs, Components, and Processes](), page 10

We propose that a liquid cooling architecture can be fundamentally described by:

1. Heat capture within the server (out of the scope for this paper)
2. CDU type
3. Method of rejecting heat to the outdoors

These three elements are labelled in Figure 1 and briefly described below.

1. Heat capture within the server

Heat is captured from the IT components using a liquid. This liquid can be a dielectric (commonly oil) in direct contact with the components, or a refrigerant or water pumped through cold plates attached to the heat-generating components. While this is an important part of the liquid cooling architecture, it is out of scope for this paper.

2. CDU type

As we indicated, a CDU is a system that isolates the IT fluid loop (TCS) from the rest of the cooling system. CDUs are typically a single enclosure with all parts integrated within it. They perform five key functions described below. It's important to understand these functions before we describe the types of CDUs.

- Temperature control – CDUs will precisely control the fluid temperature in the TCS loop. The TCS supply temperatures are dictated by IT vendors and are generally determined by the maximum case temperatures of the accelerators & processors and liquid cooling solution used.
- Flow control – To move the heat away from chips, CDUs must be able to provide sufficient flow through the rack manifolds, connectors, and cold plates across all the servers and racks supported. Immersion tanks also have flow requirements of the dielectric across the servers.
- Pressure control – There are two aspects of pressure a CDU manages. First is the maximum pressure allowed in the system and the second is the differential pressure (delta P) needed to provide the required flow. TCSs are commonly operated under a positive pressure, but alternative CDUs can pump water under a vacuum, or often called "negative pressure"[4], eliminating the risk of leaks in the TCS. This feature is commonly called a leak-prevention-system (LPS).

[4] Key feature of a negative pressure CDU is leak prevention but has other advantages of simplifying the connections & components of the TCS loop and the rack & servers, which can reduce the overall cost.

- **Fluid treatment** – The filtration and chemistry requirements of the fluids in the TCS are more stringent than facility systems. For water-based TCSs, propylene glycol-based water is a common fluid used to prevent biologic growth and maintain water quality.
- **Heat exchange and isolation** – Exchanging the heat out of the TCS is a fundamental function of a CDU. It must also isolate the fluid in the TCS loop, from the rest of the cooling system[5].

CDUs are fundamentally comprised of pump(s), heat exchanger, filtration system, and controls to perform these functions. There are many detailed attributes involved in ultimately specifying a CDU (e.g., controller, filter type, etc.). However, you don't need to specify all these attributes to choose the appropriate liquid cooling architecture for your facility. We can simplify this process by identifying only the critical attributes of a CDU. Critical attributes are those that, if you incorrectly choose, will force you to go back and change your fundamental architecture, wasting time and work spent on detailed designs.

We believe the CDU type must be based on two critical attributes:

- **Type of heat exchange** (liquid-to-air, liquid-to-liquid, etc.)
- **CDU capacity and form factor** (rack-mounted, floor-mounted)

We describe each attribute in detail below.

Type of heat exchange

There are six types of heat exchange seen in the liquid cooling industry:

- **Liquid to air (L-A)** – TCS liquid loop heat is pumped to a coil (i.e., radiator) where the heat is rejected directly into the data center air[6].
- **Liquid to liquid (L-L)** – TCS liquid loop heat is transferred to a facility water system.
- **Refrigerant to air (R-A)** – Two-phase direct-to-chip system rejects heat directly to air via a radiator. This operates like an air-based condenser.
- **Refrigerant to liquid (R-L)** – Two-phase direct-to-chip system rejects heat to a facility water system. This operates like a water-based condenser.
- **Liquid-to-refrigerant (L-R)** – A TCS liquid loop rejects heat to a facility pumped refrigerant system.
- **Refrigerant-to-refrigerant (R-R)** – Not typical.

CDU capacity and form factor

The pump size, heat exchanger size, and fluid type define a CDU system's overall capacity (kW). CDUs come in a wide range of capacities depending on form factor:

- **Rack-mounted** – CDU mounted within a rack provides a TCS loop for a single rack and can be pre-integrated with servers. Heat exchange can be L-A or L-L with CDU capacities ranging from 20-40 kW or 40-80 kW respectively. Rack-mounted refrigerant TCS loops will also have capacities in these ranges.

[5] Direct-to-chip liquid-cooled servers have stringent requirements in water temperature, flowrate, and chemistry. This means that water from facility systems (e.g., a chiller) cannot run directly through a chip's cold plate. Doing so can corrode metals and tiny fluid channels within the cold plate.

[6] Some server designs use cold plates, liquid loop, and L-A heat exchanger all contained within the server chassis. This is not considered a CDU since the server is operated and cooled just like a traditional air-cooled server.

Life Is On Schneider Electric

- **Floor-mounted** – CDU provides a TCS loop for several racks. Form factor may be similar to an IT rack or larger as capacities increase. These CDUs are usually located nearby or adjacent to liquid-cooled IT racks and for immersion systems can be integrated into the tanks. Floor-mounted CDUs can be L-A with capacities up to about 60 kW. L-L floor-mounted CDU capacities can range from 300 kW to greater than 1 MW. **Figure 2** illustrates some examples.

Figure 2

Examples of CDU form factors

(a): Floor-mounted (L-A)
(b): Floor-mounted (L-L)

(a) (b)

For this paper, we focus on two predominant types of heat exchanges (L-A and L-L). **Table 2** describes four common CDU types:

Table 2

Common CDU types

Type of heat exchange	CDU capacity and form factor
Liquid to Air (L-A)	Rack-mounted (20-40 kW)
	Floor-mounted (Up to 60 kW)
Liquid to Liquid (L-L)	Rack-mounted (40-80 kW)
	Floor-mounted (300 kW and over)

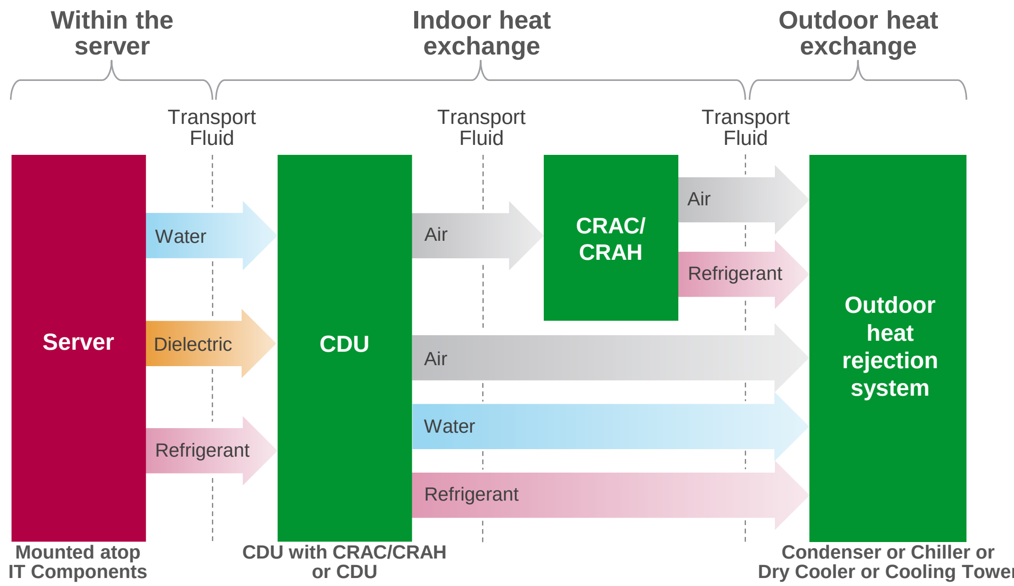
3. Method of rejecting heat to the outdoors

This is the third and final element for describing a liquid cooling architecture. Once heat from the IT equipment is captured by the TCS loop the question then becomes, how do I transfer this heat energy to the outdoors? The answer lies in the heat rejection system as described in **Figure 3**. There are three common methods:

- **Existing heat rejection system**
  - Reject heat in TCS loop to air in the IT space via liquid-to-air heat exchange (also known as closed-loop local heat rejection)
  - Reject heat in TCS loop to water in the facility systems via liquid-to-liquid heat exchange (tap into existing FWS or CWS loop)
- **Dedicated heat rejection system** – design a new independent heat rejection system for liquid cooling.

Figure 3

Simplified view of heat rejection in a liquid cooling architecture



Using combinations of the last two architecture elements ("CDU type" and "Method of rejecting heat to the outdoors") we can construct six common liquid cooling architectures seen in the industry. Table 3 illustrates these combinations. We describe how to choose the appropriate liquid cooling architecture in the next section.

Table 3

Common liquid cooling architectures are comprised of a heat rejection method and a CDU type.

Heat rejection method	CDU type[7]
Existing heat rejection system – Reject heat to air in IT space	L-A Rack-mounted
	L-A Floor-mounted
Existing heat rejection system – Reject heat to facility water systems	L-L Rack-mounted
	L-L Floor-mounted
Dedicated heat rejection system – Reject heat to independent water systems	L-L Rack-mounted
	L-L Floor-mounted

Choosing the appropriate architecture

In this section we simplify the process of choosing the most appropriate of the six common architectures by breaking it down into two steps.

- Step 1 - Choose the heat rejection method
- Step 2 - Choose CDU capacity and form factor

Notice in Table 3 that the type of heat exchange (e.g., L-A) listed under "CDU type" is dictated by the heat rejection method.[8] The second CDU attribute, "CDU capacity and form factor", is independent of the heat rejection decision. Therefore, you can determine the appropriate liquid cooling architecture in two independent steps. The selection of a *heat rejection method* coupled with *CDU capacity and form factor* depends on many factors including the four key ones below:

[7] R-A and R-L will have same decision process and L-A and L-L and thus are not added to the table.

[8] For example, the secondary air side of L-A heat exchanger can't interface with facility water system.

Life Is On Schneider Electric

- Existing cooling infrastructure compatibility – indicates the ease the existing cooling infrastructure can support new liquid-cooled servers.
- Deployment size – indicates the number of racks that can be supported by the liquid cooling architecture.
- Speed of deployment – indicates how long it takes the facility to deploy the liquid-cooled architecture from design to construction, to first operation.
- Energy efficiency – indicates the relative efficiency of the overall liquid cooling architecture. Note that all architectures will provide some improvement over equivalent air-cooled architectures.

You likely won't be able to maximize all these factors, but the idea is to make tradeoffs between them based on your priorities. For example, you're unlikely to get the highest efficiency with the liquid cooling architecture that is most compatible with your existing air-cooled system. This is because it's less efficient to transfer heat using air compared to water. The following sections comprehend these factors by describing each architecture choice in terms of *advantages*, *disadvantages*, and *when to implement*. Note that all the heat rejection system diagrams illustrate a floor-mounted CDU but could be replaced with rack-mounted CDU.

Step 1 – Choose the heat rejection method

Existing heat rejection system – Reject heat to air in IT space

This architecture allows you to design the TCS loop as a self-contained system within the IT space. The L-A CDU may be rack-mounted or floor-mounted. In this architecture, everything about the existing air-cooled facility infrastructure stays the same (as shown in Figure 4). This architecture is also known as closed-loop local heat rejection. Finally, all the heat in the IT room is rejected to the outdoors by the existing cooling infrastructure.

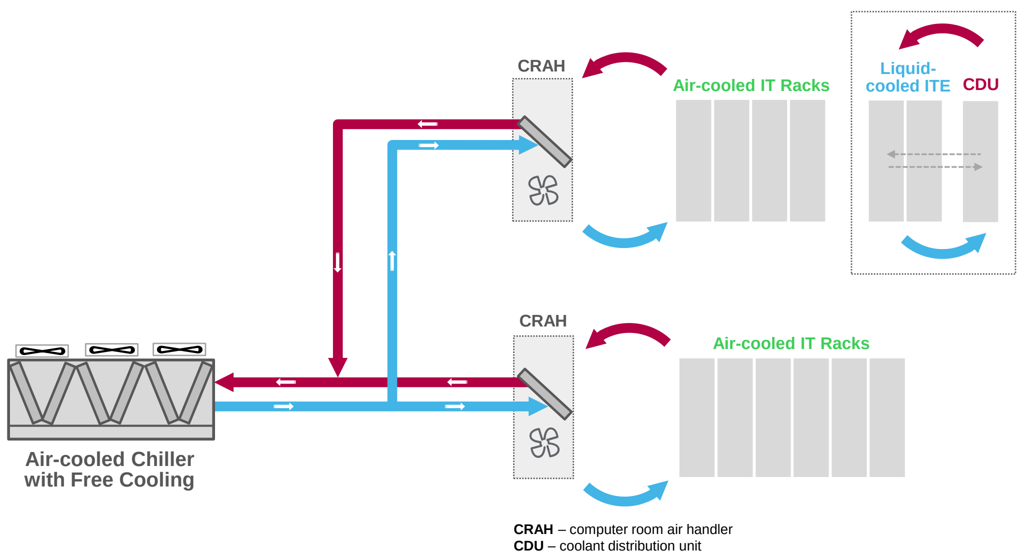


Figure 4

Diagram of "Reject heat to air in IT space" architecture

Advantages

- Compatible with most existing cooling infrastructure
- No need to modify existing cooling infrastructure
- Can be prefabricated for easier installation, standardization, etc.
- If there is a problem with the TCS loop, a small number of servers / racks would be affected

Disadvantages

- Reduced efficiency driven by more heat exchanges and fans compared to L-L CDU deployments
- Rack-mounted and floor-mounted L-A CDUs can take up rack or floor space
- Costly for large-scale deployments
- Many separate loops are required to monitor and maintain water quality
- More difficult to achieve full concurrent maintainability or full redundancy
- Small water based TCS loops have short thermal ride through, which means if the CDU fans fail, the volume of water in the loop won't have much cooling capacity to support the load compared to other architectures

When to implement

- Chilled or condenser water is not available or connecting to existing cooling infrastructure is not acceptable
- There is sufficient air-cooled capacity and airflow analysis or computational fluid dynamics (CFD) indicates the room can support the high densities.
- Supporting small scale liquid-cooled server deployments, from a single server to several racks
- Speed of deployment is a top priority

Existing heat rejection system – Reject heat to facility water systems

In this architecture, the TCS loop leverages an L-L CDU, to become an isolated loop fed off a chilled or condenser water loop. The server heat is transferred from the TCS loop to the facility loop via the CDU's L-L heat exchanger (as shown in Figure 5). The heat is then rejected to the outdoors or reused for other purposes (e.g., district heating). 60% to 90% of a liquid-cooled server's heat can be removed via the liquid, depending on the number of liquid-cooled components. The remaining heat is removed by air cooling (e.g., CRAC, CRAH[9], rear door heat exchanger).

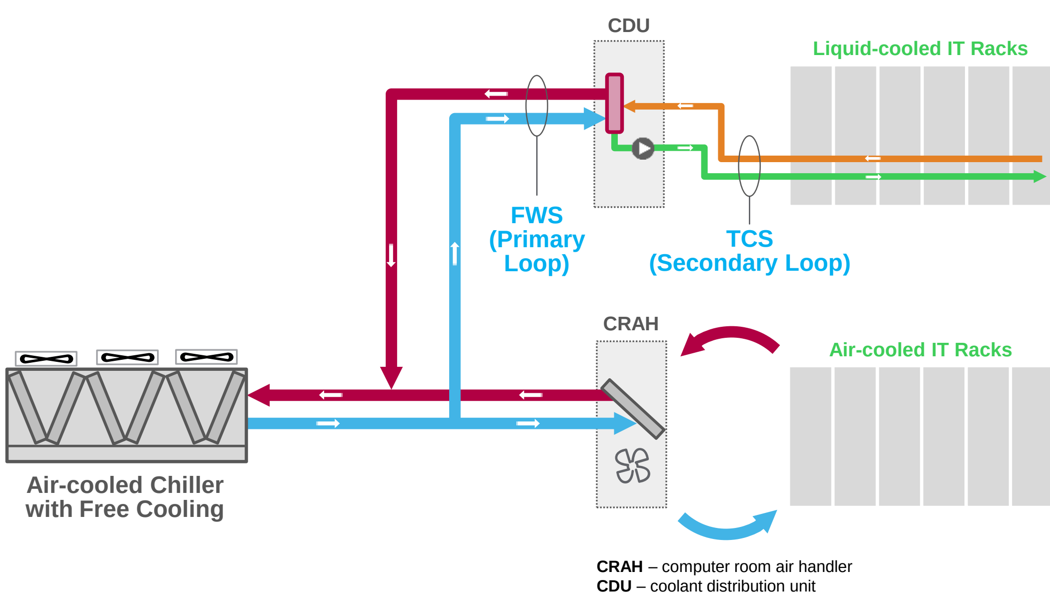


Figure 5

Diagram of "Reject heat to existing chilled water loop" architecture

[9] CRAC – computer room air conditioner, CRAH – computer room air handler

Advantages

- Reduced investment by using existing heat rejection system
- Higher liquid cooling heat rejection capacities, increased energy efficiency, and reduced sound levels (lower airflow rate across servers) compared to liquid-to-air “Reject heat to air in IT space” architecture
- For retrofits, the CDUs can recoup space previously occupied by CRAHs

Disadvantages

- More site installation work compared to L-A CDUs, including CDU connection to facility water systems and TCS piping to racks

When to implement

- Mid to large scale liquid-cooled server deployments in data centers with chiller plant
- When water loop connections or “tap-offs” already exist
- When energy efficiency is a higher consideration over speed of deployment (e.g., L-A CDUs)

Dedicated heat rejection system – Reject heat to independent water systems

In this architecture, a dedicated heat rejection system is designed for liquid cooling using a L-L CDU. This optimizes temperature and flow of the TCS and heat rejection loops in the most efficient way without the constraints imposed by a shared air-cooled heat rejection system. Figure 6 demonstrates an example of dedicated heat rejection systems for liquid cooling and air cooling. A dry cooler with trim compressor is used to provide a high supply water temperature (40°C/104°F) for liquid cooling, while a free-cooling chiller is used to provide a low chilled water temperature (20°C/68°F) for air cooling.

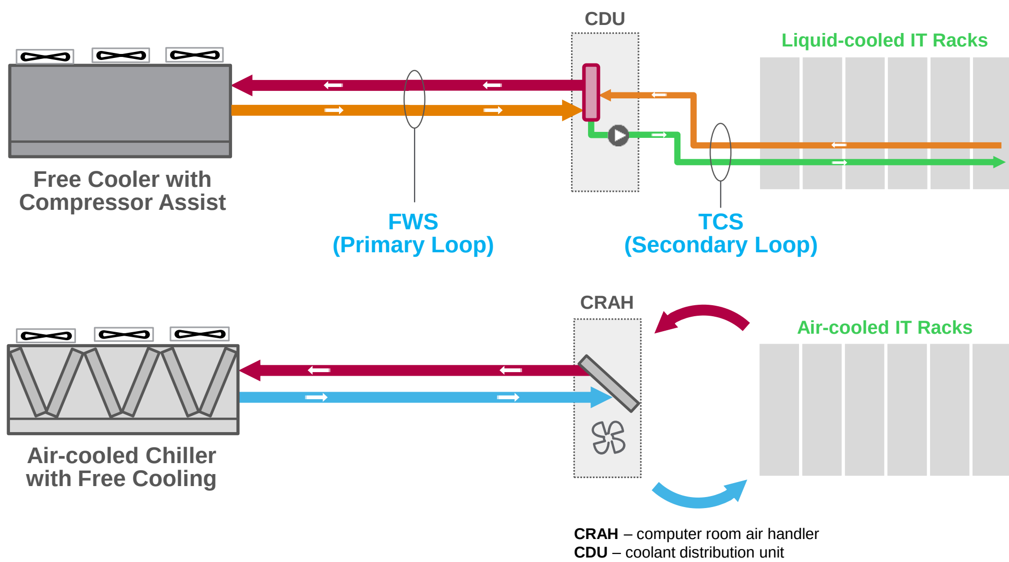


Figure 6

Diagram of dedicated heat rejection system architecture

Advantages

- High energy efficiency due to increased free cooling hours (mechanical cooling is not required in most cases except on hottest days)
- Higher return water temperatures offer the opportunity for reuse in space heating, preheating industrial process water, etc.
- Implementation does not disrupt existing cooling system

Disadvantages

- Requires design of additional facility-level piping system
- Requires investment in dedicated heat rejection system
- Longer deployment time compared to other architectures

When to implement

- Significant liquid-cooled server deployments are expected
- High energy efficiency is a priority

Step 2 – Choose CDU capacity and form factor

Rack-mounted

The CDU is dedicated to a single rack, which means each rack has its own TCS loop. Generally mounted at the bottom of the rack, the CDU includes a pumping unit, filtration, and controls. Heat is either transferred to the data center air via a fan-assisted rear-door heat exchanger (L-A) or to a facility loop via a L-L heat exchanger.

Advantages

- Can be pre-integrated and tested with the servers prior to install in the data center.
- Limits potential failure modes to a single rack (e.g., TCS leak or contamination)
- Redundancy (e.g., 1N vs 2N pumps) can be targeted for each rack
- Simple solution for a traditional data center with a few liquid-cooled racks

Disadvantages

- Cost per kW IT load becomes greater than floor-mounted CDUs as the number of racks increase
- CDU occupies IT server space
- Limits maximum rack density to about 40 kW (L-A) and 80 kW (L-L)
- As number of racks increase, total installation time is longer compared to a larger floor-mounted CDU (i.e., commissioning CDU in each rack)
- As number of racks increase, efficiency degrades compared to a single larger floor-mounted CDU

When to implement

- Speed of deployment is important (applies for few liquid-cooled racks)
- Limited liquid-cooled rack deployments are expected (1 to 10 racks)

Floor-mounted

The CDU is dedicated to a row or multiple rows of racks, meaning they all share the same TCS loop. It may be placed at the end of the row or further away from the AI cluster. Heat is either transferred to the data center air via a fan-assisted heat exchanger (L-A) or to a facility loop via L-L heat exchanger.

Advantages

- Lower cost per kW IT load for large deployments compared to rack-mounted CDUs
- In retrofits, CDUs can be selected based on the location and capacity of the CRAHs they replace, to minimize piping work

- No IT rack space occupied by CDU
- Higher rack densities can be achieved compared to rack-mounted CDU
- As the number of racks increases, installation time can be compressed compared to multiple rack-mounted CDUs
- As number of racks increase, efficiency improves compared to many rack-mounted CDUs running multiple pumps
- More thermal ride thru than rack-mounted CDUs

Disadvantages

- All racks on single TCS loop become susceptible to common cause failures (e.g., TCS leaks, contamination, controls, etc.)
- Takes up floor space

When to implement

- More than 10 liquid-cooled racks are expected
- Workloads (e.g., AI training cluster) can tolerate a common cause failure (e.g., loss of fluid flow) for all racks served by a single CDU
- There's no vertical space in IT rack to install rack-mounted CDUs

Life Is On Schneider Electric

Conclusion

An increasing number of servers require liquid cooling systems to support AI workloads. Depending on the scale of liquid-cooled server deployments, a data center can be cooled through exsiting or decidicated heat rejection systems. The terminology, architectures, and factors in choosing dicussed in this paper provide a starting point for data center operators to build out the ecosystem. We provide the following answers to common liquid cooling questions raised by data center operators:

- **How do I get cold water in and hot water out?** – Use three elements (i.e., heat capture within the server, CDU type, and method of rejecting heat to the outdoors).
- **What is a CDU, and do I need one?** – A CDU is a system used to isolate the IT fluid loop (TCS) from the rest of the cooling system, and is necessary to provide five key functions (i.e., temperature control, flow control, pressure control, fluid treatment, heat exchange and isolation).
- **What steps do I take to select an appropriate heat rejection architecture?** – This paper describes six liquid cooling architecture each with advantages, disadvantages, and when to implement.

About the authors

Paul Lin is the Research Director and Edison Expert at Schneider Electric's Energy Management Research Center. He is responsible for data center design and operation research and consults with clients on risk assessment and design practices to optimize the availability and sustainability of their data center environment. He is a recognized expert, and a frequent speaker and panelist at data center industry events. Before joining Schneider Electric, Paul worked as an R&D Project Leader in LG Electronics for several years. He is also a registered professional engineer and holds over 10 patents. Paul holds both a Bachelor's and Master's of Science degree in mechanical engineering from Jilin University. He also holds a certificate in Transforming Schneider Leadership Programme from INSEAD.

Robert Bunger is the Innovation Product Owner within the CTO office at Schneider Electric. In his 26 years at Schneider Electric, Robert has held management positions in customer service, technical sales, offer management, business development & industry associations. While with APC / Schneider Electric, he has lived and worked in the United States, Europe, and China. Prior to joining APC, he was a commissioned officer in the US Navy Submarine force. Robert has a BS in Computer Science from the US Naval Academy and MS EE from Rensselaer Polytechnic Institute.

Victor Avelar is a Senior Research Analyst at Schneider Electric's Energy Management Research Center. He is responsible for data center design and operations research, and consults with clients on risk assessment and design practices to optimize the availability and efficiency of their data center environments. Victor holds a bachelor's degree in mechanical engineering from Rensselaer Polytechnic Institute and an MBA from Babson College. He is a member of AFCOM.

RATE THIS PAPER ★★★★★

Life Is On Schneider Electric

Resources

[The AI Disruption: Challenges and Guidance for Data Center Design](#)
White Paper 110

[Liquid Cooling Technologies for Data Centers and Edge Applications](#)
White Paper 265

[Five Reasons to Adopt Liquid Cooling](#)
White Paper 279

Browse all white papers
whitepapers.apc.com

Browse all TradeOff Tools™
tools.apc.com

Note: Internet links can become obsolete over time. The referenced links were available at the time this paper was written but may no longer be available now.

Contact us

For feedback and comments about the content of this white paper:

Schneider Electric Energy Management Research Center
dcsc@schneider-electric.com

If you are a customer and have questions specific to your data center project:

Contact your Schneider Electric representative at
www.apc.com/support/contact/index.cfm

© 2024 Schneider Electric. All rights reserved.

Life Is On Schneider Electric