

Understanding Liquid Cooling

The CDU as a Key Component of Modern Data Centers

Rising power densities require new cooling concepts

Modern IT infrastructures—particularly in the fields of AI and GPUs—are achieving ever-higher power densities. Traditional air-based cooling is increasingly reaching its physical limits.

The solution lies in **liquid cooling**.

Instead of air, water is applied directly to the hottest components. This allows heat to be dissipated much more efficiently and precisely. Especially at high power densities, this approach enables stable and energy-efficient cooling.

At the heart of this technology is the **Coolant Distribution Unit (CDU)**. It serves as the central link between building cooling and the IT system, handling the controlled distribution and regulation of the coolant.

Dynamic load profiles as a key challenge

Modern AI and HPC (High-Performance Computing) applications do not follow a constant load pattern. Unlike traditional IT workloads, they are characterized by extreme dynamics: power consumption and thermal load can surge dramatically within seconds. These rapid load transients and unpredictable GPU utilization lead to short-term thermal peaks that the cooling system must immediately handle.

For stable operation, it is precisely this dynamic nature that poses the greatest challenge. Inaccurate or overly sluggish control quickly leads to unstable temperatures and pressure conditions in the secondary circuit. In the worst case, the IT system responds with “thermal throttling” (power throttling), resulting in the loss of valuable computing time.

Modern cooling systems must therefore:

- ▶ Enable rapid control responses through frequency-controlled pumps and precise valves.
- ▶ Prevent pressure surges that could stress the piping during abrupt load changes.
- ▶ Efficiently manage partial load ranges, as AI clusters do not run at full load continuously.

Why Cooling is Undergoing a Fundamental Shift

For many years, data centers have primarily relied on air-based cooling concepts. These concepts, such as cold/hot aisles, operate reliably and energy-efficiently at moderate power densities. However, as power demands increase, these systems reach their physical limits.

The key difference lies in the physical properties of the cooling media. Water has a significantly higher density and heat capacity than air and can therefore absorb many times more heat.

While air requires ever-increasing airflow rates and higher fan power as the load increases, water enables significantly more efficient heat dissipation at much lower flow rates.

Liquid cooling is therefore increasingly becoming the technical standard, particularly in AI and GPU environments with high power densities.

As the heat load increases, either more coolant must be circulated or a greater temperature difference must be utilized. In air-cooled systems, this leads to high airflow rates and, consequently, increased energy consumption.

In air-cooled systems, this leads to:

- ▶ very high airflow rates
- ▶ increased fan power
- ▶ increased energy consumption
- ▶ increasing noise levels

Liquid cooling is becoming increasingly relevant from a technical standpoint, particularly at high power densities, such as in AI and GPU environments with 60-100 kW per rack. It enables significantly higher heat removal at the same temperature spread with substantially lower airflow rates and considerably reduces the energy consumption of the transport medium.

Physical Limits - Air vs. Water

Water has a significantly higher density and specific heat capacity than air. As a result, water can carry the same amount of heat in a fraction of the volume.

Taking both density and heat capacity into account, water can transport approximately 3,000-3,500 times more heat per unit volume than air.

Feature	Air cooling	Liquid cooling (DLC)
Specific heat capacities	Air: approx. 1,0 kJ/(kg*K)	Water: approx. 4,2 kJ/(kg*K)
Density	1 m ³ Air approx. ca. 1,2 kg	1 Liter Water approx. 1 kg
Power Density	Up to approx. 20-30 kW/rack	60 - 1000+ kW/rack
Energy Consumption for Transport	High (fans)	Low (pumps)
Space Requirements	High (air ducts / plenum)	Low (coolant lines)

How does Direct Liquid Cooling (DLC) work?

With Direct Liquid Cooling, heat is dissipated directly at the source—the CPU or GPU. It is transferred immediately to a cooling medium via cold plates.

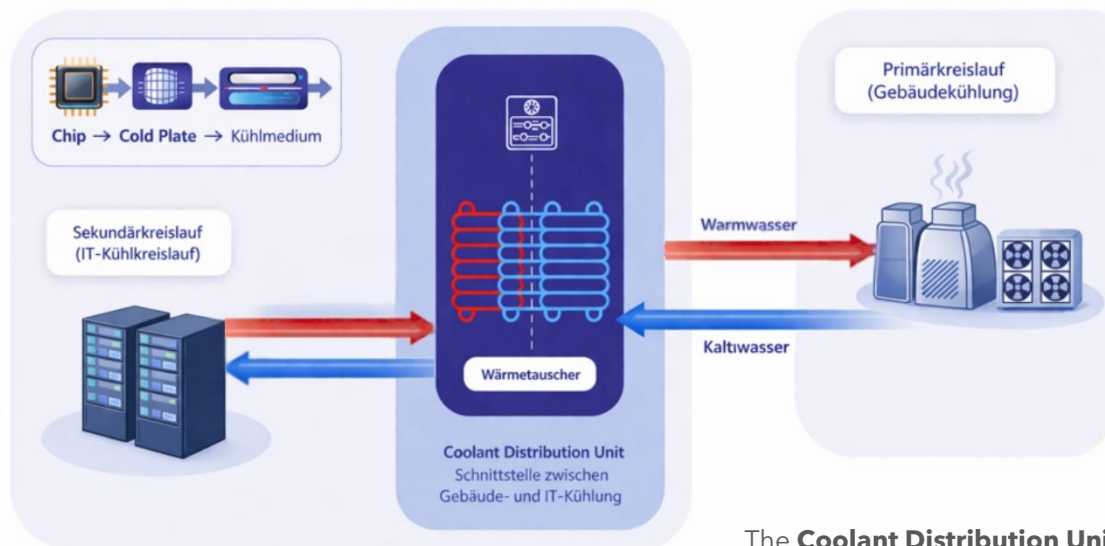
The CDU plays a central role in the overall system. It connects the primary circuit of the building cooling system with the secondary circuit of the IT cooling system and ensures that the cooling medium reaches the components in a controlled manner and according to demand.

A key component is the hydraulic separation of the two circuits, typically via a heat exchanger. This reliably protects sensitive IT systems from the effects of the building's cooling system.

At the same time, the CDU handles key control functions: it regulates temperatures, stabilizes pressure conditions, and dynamically adjusts the flow rate to the actual load.

How Direct Liquid Cooling (DLC) Works

In Direct Liquid Cooling, heat is dissipated directly at the source—that is, at the CPU or GPU. The heat is transferred directly to the coolant via so-called cold plates.



The **Coolant Distribution Unit (CDU)** connects the primary circuit (building services) to the secondary circuit (IT cooling circuit).

It ensures that the cooling medium is delivered to the IT components in a controlled, stable, and demand-driven manner. This protects the IT systems from the effects of building cooling while simultaneously ensuring defined operating conditions in the IT cooling circuit.

Hydraulics and System Architecture as the Key to Operational Reliability

Performance alone is not enough to ensure stable cooling. A well-designed hydraulic configuration of the entire system is crucial.

Modern liquid cooling systems operate with very fine structures in the cold plates and correspondingly high pressure losses. To ensure stable operating conditions, a uniform flow rate, reliable venting, and appropriate filtration are essential. Only in this way can stable operating conditions and high availability be ensured in the long term.

System architecture also plays a decisive role. Decentralized CDUs enable flexible and modular scaling with short piping runs. Centralized systems, on the other hand, bundle high power capacities and reduce complexity in the IT sector.

Which solution makes sense depends largely on power requirements, the redundancy concept, and the existing infrastructure.

Centralized or Decentralized Cooling Architecture—Choosing the Right Topology

As power density increases, the choice of system architecture becomes a critical factor.

- ▶ Decentralized in-row CDUs offer high flexibility and enable incremental scaling directly within the IT area. At the same time, however, the number of components increases, and with it, system complexity.
- ▶ Centralized distribution units consolidate high cooling capacities in a single location and facilitate maintenance work outside the IT area. In return, the demands on planning, pressure management, and hydraulic balance increase.

The optimal solution always results from the interplay of power requirements, scaling strategy, and infrastructural conditions.

Focus on Efficiency, Monitoring, and Availability

As modern applications become increasingly dynamic, the demands for energy efficiency and transparency are also rising. Cooling capacities must be adjusted quickly without causing system instability.

At the same time, regulatory requirements such as the Energy Efficiency Act (EnEFG) mandate continuous collection of operational data. Modern CDUs provide the necessary measurement data for this purpose and enable targeted analysis and optimization of cooling operations.

Availability is also becoming increasingly important. Depending on requirements, redundant systems and multi-stage power supply concepts are implemented, often based on standards such as DIN EN 50600.



Operational Reality: Water Quality and Scaling

As power density increases, operational details are coming into sharper focus. A critical aspect is water quality and filtration.

The fine microchannels of modern cold plates are sensitive to particles or deposits. Even the smallest contaminants can significantly impair cooling performance.

Modular CDU designs offer advantages here. By connecting multiple units in parallel, cooling capacity can be scaled flexibly while simultaneously increasing operational reliability.

Rated Capacity and Actual System Performance

The maximum cooling capacity of a CDU specified in the datasheet is not a fixed value. It depends heavily on the specific operating conditions.

The following factors are particularly critical:

- ▶ Temperature differences (ΔT): Performance depends on the temperature difference between the primary circuit (building) and the secondary circuit (IT).
- ▶ Cooling medium: The use of glycol additives reduces heat capacity compared to pure water.
- ▶ Hydraulic integration: The actual flow rate and pump characteristics determine the achievable waste heat removal in the field.

For planning and procurement, therefore, it is not the peak value that is decisive, but the actual performance in real-world operation.

Availability and compliance with DIN EN 50600

Liquid cooling systems are now a central component of modern data centers and must meet high availability requirements.

The DIN EN 50600 standard defines four availability classes (VK1-VK4) that determine the level of protection for the IT infrastructure:

- ▶ VK1 & VK2 (Basic to Enhanced Security): These classes offer no or only partial redundancy. Maintenance work can lead to service interruptions here.
- ▶ AC4 (High Availability): This is the standard for professional data centers. Through N+1 redundancy (e.g., for pumps or CDUs), individual components can be maintained during operation without interrupting cooling.
- ▶ VK4 (Very High Availability): A multi-path architecture is used here. The system is fully fault-tolerant; even the failure of an entire supply path does not lead to an interruption in IT power supply.

For the technical implementation, this means that CDUs must not only have redundant pumps and a dual power supply but must also be integrated into separate hydraulic loops. Continuous monitoring of leaks, pressure, and flow rates is an integral part of this process to guarantee the targeted availability class.

Legal Requirements and Monitoring

In addition to operational safety, regulatory compliance is becoming increasingly important. The Energy Efficiency Act (EnEfG) requires operators to comprehensively record and document their energy consumption. Modern CDUs serve as a central data source for this purpose by continuously recording parameters:

- ▶ Thermal dynamics: supply and return temperatures as well as the resulting temperature spread (ΔT).
- ▶ Hydraulics & energy: flow rates, differential pressure, and the electrical power consumption of the pumps.

This data not only enables compliance with legal requirements but also allows for targeted optimization of energy efficiency and the overall system.

Conclusion

Liquid cooling is a central component of modern data centers. The Coolant Distribution Unit (CDU) plays a key role as the interface between building services and IT.

With increasing demands from AI and GPU infrastructures, precise control, stable hydraulics, and a well-designed system architecture are crucial for efficiency, availability, and future-proofing.

The right design is not only a technical decision but also a strategic one.

- ▶ Scalability: Systems can be expanded step by step - investments remain flexible.
- ▶ Transparency & Regulations: Operational data is automatically recorded and helps meet legal requirements.
- ▶ Operational Safety: The separation of circuits protects IT from malfunctions and contamination.
- ▶ Energy Efficiency: Heat can be better utilized, and energy consumption drops significantly.

Checklist: Key Planning Questions for Designing a CDU in a Data Center

The following questions help with the initial technical assessment:
When is a Coolant Distribution Unit needed, what capacity is required,
and what conditions must be taken into account?

1. Capacity Requirements

- ▶ Maximum IT load per rack (kW)?
- ▶ Number of racks per CDU?
- ▶ Required total capacity, including reserve?
- ▶ Is future scaling planned?

2. Load Profile & Dynamics

- ▶ Constant load or dynamic AI/GPU workloads?
- ▶ Are rapid load spikes expected?
- ▶ Is stable differential pressure and flow rate control required at partial load?

3. Hydraulic & Thermal Conditions

- ▶ Primary supply temperature available?
- ▶ Desired temperature range in the IT circuit?
- ▶ Differential pressure requirements?
- ▶ Is hot water operation planned?

4. Availability & Redundancy

- ▶ What availability class (DIN EN 50600) is targeted?
- ▶ Is a redundant pump or power supply required?

5. Energy Efficiency & Monitoring

- ▶ Is integration into DCIM or building management systems planned?
- ▶ Must operating data be documented in accordance with the EnEFG?
- ▶ Is maximizing the proportion of free cooling planned?

6. Infrastructure & Security

- ▶ Is electrical connection capacity available?
- ▶ Requirements for filtration and water quality?
- ▶ Is leak monitoring required?

→ Detailed planning checklist available as a PDF.