White Paper

# A Holistic Approach to Managing Liquid-Cooled AI Clusters

November 2024

## A direct liquid cooling management solution for enhancing performance, energy efficiency, and reliability of AI clusters

### Authors

**Ted Pang**
Senior Director, System & Software Engineering
Wiwynn

**Karl Chiang**
Director, System & Software Engineering
Wiwynn

**PJ Lin**
Senior Manager, System & Software Engineering
Wiwynn

**Rami Radi**
Sr. Product Manager and Solution Architect, Product Office, Strategic Market Development and Product Management Group
AMI USA

**CD Song**
Sr. Engineering Manager, R & D Engineering, Security Software Group
AMI China

**Kemp Ke**
Sr. Principal Engineer, R & D Engineering, Security Software Group
AMI China

**Muthukkumaran Ramalingam**
Sr. Engineering Manager & Product Architect, R & D Engineering, Security Software Group
AMI USA

### Summary

The integration of advanced liquid cooling systems in AI clusters is crucial for maintaining thermal stability and optimizing performance. Key components such as AI racks, Cooling Distribution Units (CDUs), Wiwynn UMS100 (Universal Management System), and AMI DCM (Data Center Manager) work together to ensure efficient cooling management. The Wiwynn UMS100 manages cooling at different levels, acting as an intermediary between individual servers and the overall cooling system, while the AMI DCM provides a centralized platform for monitoring and managing all servers and devices within the AI cluster. This integrated approach enhances performance, improves energy efficiency, and ensures reliability.

### Contents

# 1. Introduction

In the rapidly evolving landscape of data center technologies, efficient cooling solutions have become paramount to ensure optimal performance and longevity of hardware components. As data centers continue to scale up to accommodate the increasing demands of high-performance computing and artificial intelligence, traditional air-cooling methods are proving to be insufficient. This whitepaper aims to explore the integration of advanced liquid cooling systems in AI clusters, focusing on the architecture and management of such systems.

The primary components discussed in this whitepaper include AI racks, Cooling Distribution Units (CDUs), Wiwynn UMS100 (Universal Management System), and AMI DCM (Data Center Manager). Each of these components plays a crucial role in maintaining thermal stability of the AI cluster, ensuring that the servers operate within their optimal temperature ranges.

The Cooling Distribution Unit (CDU) is responsible for distributing the cooling liquid to the various components within the AI cluster, ensuring that each component receives the necessary cooling to function efficiently.

The Wiwynn UMS100 is a controller that manages the cooling status in different levels. It acts as the intermediary between the individual servers and the cluster's overall cooling management system. AMI DCM, on the other hand, serves as a management solution, providing a centralized and holistic platform for monitoring and managing all the servers and devices within the AI cluster.

This whitepaper delves into the functions of these components, illustrating how they work together to create a cohesive and efficient cooling management system. By leveraging the capabilities of these advanced cooling solutions, AI clusters can achieve higher performance, improved energy efficiency, and enhanced reliability.

# 2. Building Block of AI Cluster

The building blocks of a liquid-cooled AI cluster are explored using a basic configuration of the Wiwynn AI cluster with a direct liquid cooling system. The details of direct liquid cooling system can refer to another Wiwynn whitepaper [1]. The design features one CDU (800KW) serving up to 8 liquid-cooled AI racks (each 100KW) at maximum capacity as shown in Figure 1. However, the exact number of AI racks supported by one CDU depends on factors such as power consumption (KW per rack), flow rate & differential pressure per rack, water pressure per rack, facility temperature, water temperature, etc.

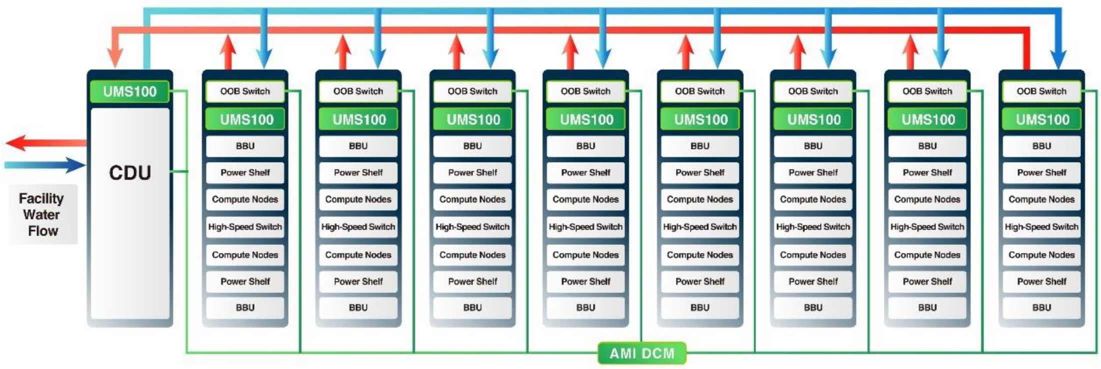


**Figure 1** A Basic Configuration of Wiwynn AI Cluster with Direct Liquid Cooling System

## 2.1 AI Rack

The AI racks are the core of the AI cluster, composed of GPU servers for processing AI training and inferencing tasks. These racks are designed to house high-performance GPUs, which generate significant amounts of heat. The inter-GPU connections are inside the rack and connect with a high-speed switch, ensuring efficient data transfer between GPUs. Power shelves provide the necessary power, and multiple Backup Battery Unit (BBU) ensures power backup during peak usage or power failures. Figure 2 shows the direct liquid cooling related components. Each AI rack has a manifold in the back to connect to the pipe, with water flow controlled by CDU. The drip tray is included to handle any potential leakage, ensuring the safety and stability of the system.

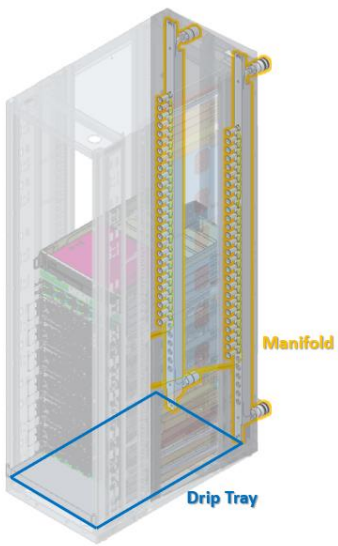


**Figure 2** Rack Manifold and Drip Tray

## 2.2 Cooling Distribution Unit (CDU)

The CDU is a critical component in the liquid cooling system. It is equipped with valves and flow meters that regulate the distribution of cooling liquid. By dynamically adjusting the flow rate, the CDU ensures that each rack receives the appropriate amount of cooling, optimizing energy efficiency and maintaining the optimal operating conditions for the servers.

## 2.3 Wiwynn UMS100 (Universal Management System)

Wiwynn UMS100 plays a dual role in managing the AI cluster’s cooling system. For in-rack management, the in-rack UMS100 protects individual servers by monitoring their leakage and the server status. It collects data from sensors within the servers, including temperature and other critical parameters. This data is then used to make real-time adjustments to the cooling system, ensuring that each server operates within its optimal temperature range. For in-row management, the in-row UMS100 monitors the flow meter and controls CDU's valves. By analyzing the data collected from the servers, UMS100 can make informed decisions about how to allocate cooling resources effectively. This level of precision in cooling management helps maintain the stability and performance of the data center.

## 2.4 AMI Data Center Manager (AMI DCM)

AMI Data Center Manager (DCM) is a powerful on-premise software solution designed to streamline the real-time management of high-density compute environments, such as AI clusters. DCM can be installed on physical or virtual machines, and allows administrators and operators to use its capabilities directly through a web console or programmatically via RESTful APIs, data streaming, or iFrame integration as shown in Figure 3.

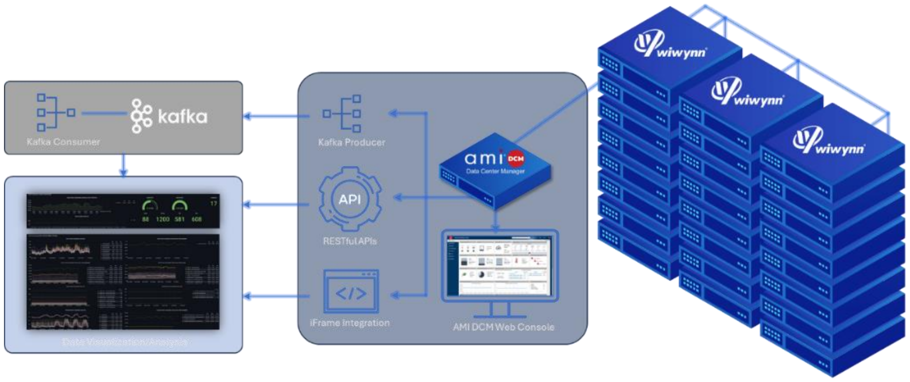


**Figure 3** Illustration of the Various Ways AMI DCM Can Be Utilized for Monitoring and Managing AI Clusters.

DCM collects, aggregates, and analyzes real-time data from a wide array of devices, including servers, switches, power units, and cooling systems.

DCM can communicate with the UMS100 universal management system to provide administrators a complete and detailed view of the inventory, energy consumption, thermal conditions, system health, utilization, and carbon emissions within the AI clusters as shown in Figure 4 and 5.

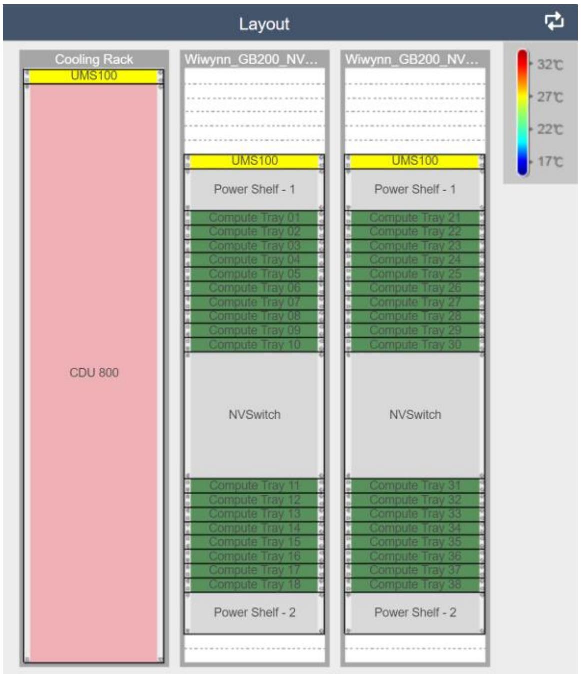


**Figure 4** AI Cluster Row Layout in AMI DCM, Color-Coded by Device Health.

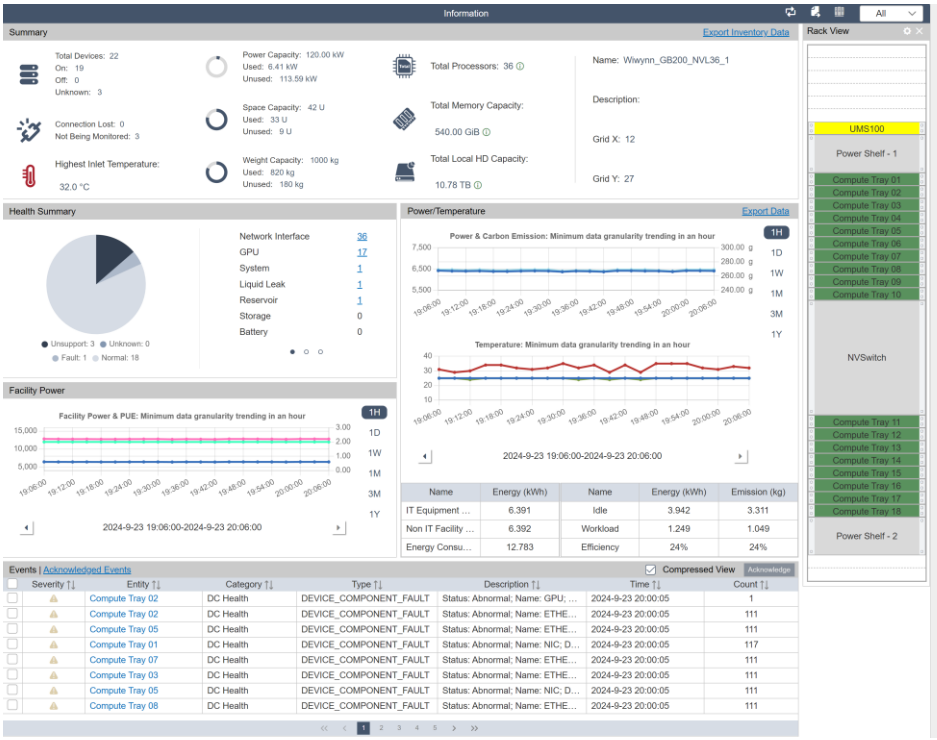


**Figure 5** High-Level Rack View Showing Aggregate Power Consumption, Carbon Emissions, and Temperature Over Time.

Through its advanced monitoring, modeling, reporting, alerting, and control features, DCM's web console offers critical insights that enable the AI cluster administrators and operators to detect and address potential issues proactively, ensuring continuous performance and availability while minimizing the cluster's environmental impact.

DCM can also manage the firmware versions for devices within the cluster and supports batch firmware upgrades for compute nodes with a single click. This ensures that all nodes run the same firmware version with the required performance, reliability, and security updates.

## 2.5 Network Architecture

The network architecture of the AI cluster features three networks: one for east-west traffic transfer between GPUs, one for north-south traffic transfer between different components of the data center, and an Out-Of-Band (OOB) network for management. The east-west network handles the data transfer between GPUs, ensuring efficient communication and processing of AI tasks. The north-south network manages the data transfer between different components of the data center, including servers, storage, and network switches. The OOB network is dedicated to management and monitoring, providing a secure and reliable channel for transferring event data and management commands.

Server BMC, UMS100, CDU and AMI DCM are connected to the OOB network for exchanging event data and management commands. The exchange of information in the OOB network determines the cooling efficiency of the AI cluster. In-rack UMS100 and servers/switches provide sensing data and status updates to the in-row UMS100 and AMI DCM, enabling higher-level management of the AI cluster.

This integrated cooling management ensures that the cluster operates within its optimal temperature range, reducing the risk of overheating and improving overall performance and energy efficiency.

# 3. Functionality and Management

Efficient management of the AI cluster's cooling system is crucial to ensure optimal performance and energy efficiency. This chapter delves into the functionality and management of the cooling system at different levels: server-level, rack-level, and cluster-level.

## 3.1 Server-Level Hardware Management

Wiwynn in-rack UMS100 plays a crucial role in monitoring and managing the server-level liquid cooling status. One of its primary functions is to detect and respond to leakage events within the servers and switches. The UMS100 achieves this by receiving Baseboard Management Controller (BMC) events, which provide real-time data on the status of the hardware components.

For servers and switches that do not support BMC-based leakage detection, Wiwynn has designed a Leak Sensing Adapter (LSA) which employs a leak sensing band with GPIO dry contact wrapped around the cold plate. LSA will detect any leakage from sensing band and send an alert to the UMS100. There are two LSA designs: one is a small module mounted to the motherboard, and the other is a PCI-E card designed to be installed directly in a server.

When the UMS100 detects a leakage event in a server/switch, it promptly sends a notification to the upper management system. Following this, the UMS100 initiates the power-off process to prevent any potential damage. The UMS100 performs a graceful shutdown of the server/switch by sending a Redfish command to ensure it is powered down safely and without data loss.

## 3.2 Rack-Level Hardware Management

At the rack level, in-rack UMS100 is responsible for monitoring the leakage of the manifold and drip tray within the rack. These components generate analog signals, which the UMS100 can support and interpret. Upon detecting a leakage, the UMS100 sends a notification to the upper management system and initiates the power-off process for the entire rack. This is similar to the server-level management process but on a larger scale.

In addition to powering off the rack, the UMS100 also turns off the valve of the rack to stop the cooling liquid flow into the rack. By managing the cooling liquid flow and ensuring a safe shutdown of the servers, the UMS100 maintains the stability and reliability of the rack. Figure 6 shows the in-rack UMS100's monitoring capability in server-level and rack-level.

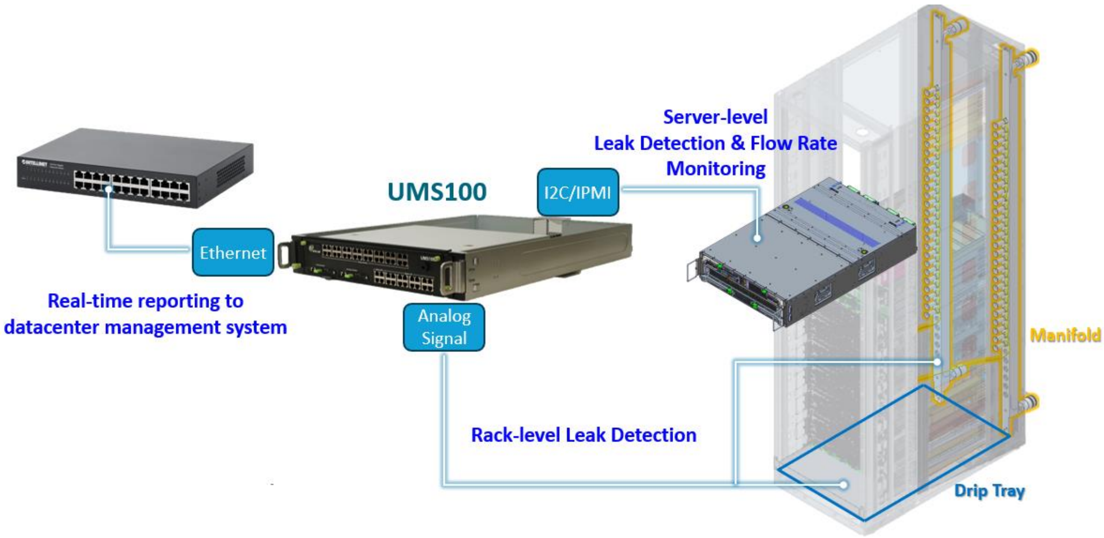


**Figure 6** In-rack UMS100 Protects Servers in Server-level & Rack-level

## 3.3 Cluster-Level Hardware Management

AMI DCM serves as a centralized platform for managing the entire AI cluster, including the compute nodes and their components such as the CPUs, GPUs, and memory modules. It also manages network switches, power shelves, and the UMS100 system, aggregating data from these different sources to deliver a comprehensive, real-time view of the cluster's thermal status, power consumption, and overall health as shown in Figure 7.


**Figure 7** Detailed View of a Compute Node Showing Various Node and Component-Level Telemetry

DCM allows administrators to set temperature thresholds that trigger alerts, enabling proactive measures to prevent overheating. Additionally, DCM sends alerts when critical events such as leaks, low reservoir levels, pressure imbalances, or pump failures are detected by the UMS100 system as shown in Figure 8, allowing for quick resolution to minimize downtime.

With DCM's real-time server-level thermal monitoring, administrators can safely raise CDU set point temperatures, avoiding unnecessary overcooling.

Adjustments to the CDU flow rate or coolant temperature are managed by the in-row UMS100. Slowing down the flow rate by 20% can result in a 50% reduction in the CDU's power consumption. By leveraging the capabilities of Wiwynn UMS100, CDU, and AMI DCM, AI clusters can achieve a high level of efficiency and reliability in their cooling management. This integrated approach not only enhances the performance of the servers but also contributes to significant energy savings, making it a sustainable solution for modern data centers.

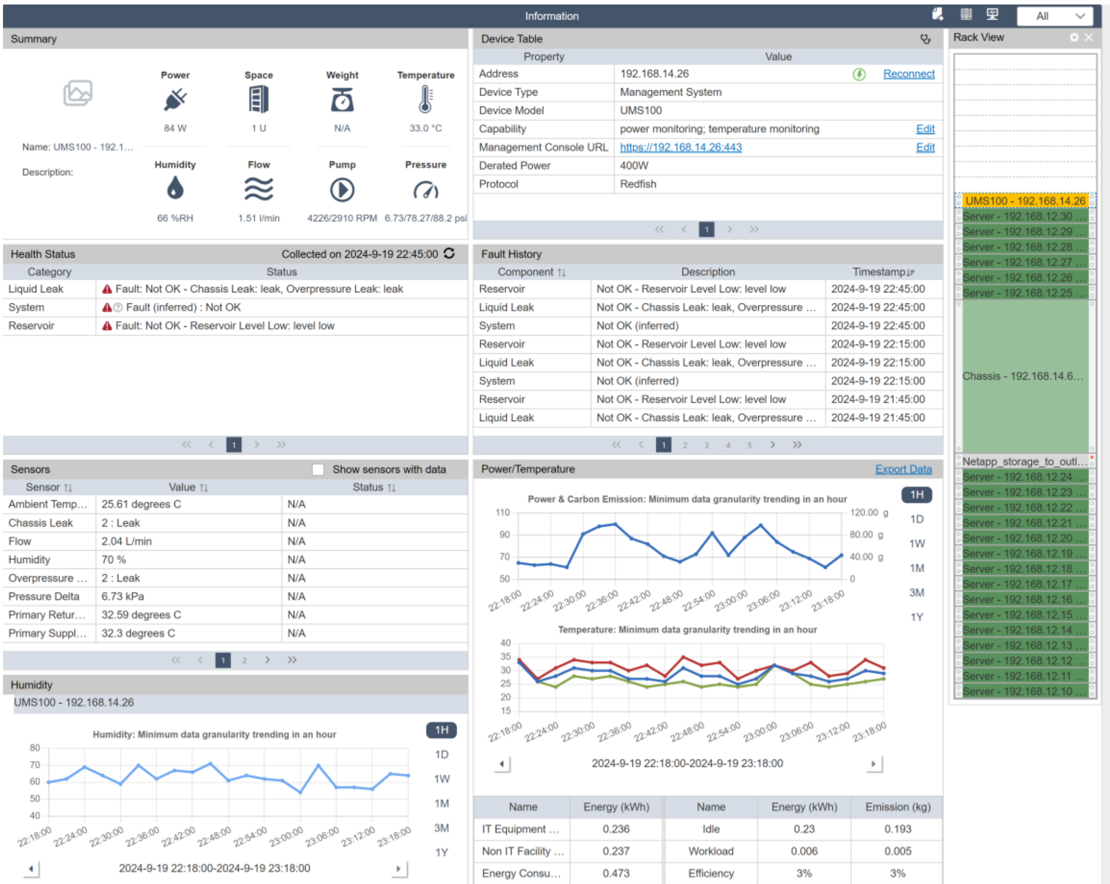


**Figure 8** UMS100 View Showing CDU Health Status, Faults, and Sensor Values

# 4. Wiwynn UMS100 and AMI DCM Features and Functions

## 4.1 Wiwynn UMS100 Features

The Wiwynn UMS100 (Universal Management System) is a versatile and robust solution designed to manage the cooling and operational status of servers and racks within an AI cluster. The UMS100 is available in both 1U and 2U form factors, providing flexibility to meet different deployment needs. It plays a dual role in managing the AI cluster’s cooling system. The in-rack UMS100 monitors individual servers for leakage and status, while the in-row UMS100 oversees the flow meter and controls the CDU's valves. This dual functionality ensures that each server operates within its optimal temperature range and that cooling resources are allocated effectively. Key features of the Wiwynn UMS100 include:

Redundancy for Continuous Operation

The UMS100 consists of two DC-SCM (Data Center Secure Control Module) controllers, a BMC-based system, that excels in energy efficiency, consuming significantly less power compared to traditional IPC systems. This redundancy is crucial for maintaining the stability and reliability of the AI cluster, ensuring that cooling management and monitoring continue without interruption.

Standard Northbound API for Easy Integration

To facilitate easy integration with other management systems, the UMS100 supports Redfish as standard northbound API. This API allows for seamless communication and data exchange between the UMS100 and upper management systems, such as the AMI DCM.

Support for Different Sensor Data Sources

The UMS100 is equipped to support various sensor data sources, providing flexibility in monitoring and managing the AI cluster. This versatility allows the UMS100 to collect comprehensive data from the servers and racks, enabling precise and effective cooling management.

Easily Expanded to Support Various CDUs

The modular design of the UMS100 allows for easy expansion to support various Cooling Distribution Units (CDUs). This scalability ensures that the UMS100 can adapt to the specific requirements of the data center, whether it involves adding more CDUs or integrating different types of CDUs.

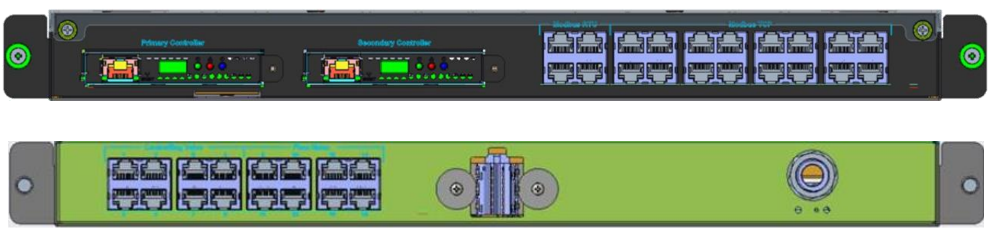

**Figure 9** Front and Rear Views of UMS100 1U System

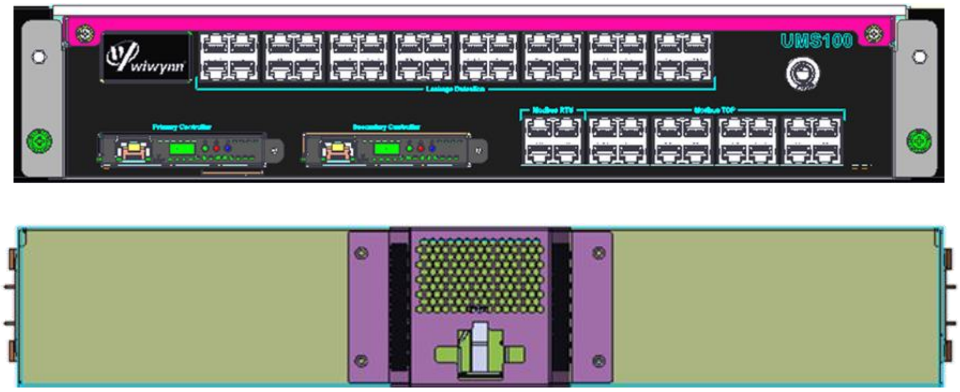


**Figure 10** Front and Rear Views of UMS100 2U System

**Table 1** Wiwynn UMS100 Specification

| | UMS100 1U | UMS100 2U |
|---|---|---|
| Processor | ● ASPEED AST2600<br>● Redundant Controller (Active-Standby) based on DC-SCM | |
| Management Interface | ● Redfish API<br>● Web UI | |
| Management Port | ● 1Gbps (RJ45) x1 per Controller<br>● MODBUS TCP 100Mbps (RJ45) x16<br>● MODBUS RTU 12Mbps (RJ45) x4 | |
| Power Source | ● 48V DC busbar<br>● CRPS 800W, 1+1 Redundancy (Reserved) | |
| Operating Temperature | ● 10 °C - 40 °C | |
| Special I/O Port | Front I/O | |
| | N/A | ● Analog Resistor input x4 for leakage detection<br>● I2C (RJ45) x32 for IT Device leakage detection |
| | Rear I/O | |
| | ● Rear Analog Voltage input/output x8 for controlling valve with 24V power supply (5W@24V per port)<br>● Analog Current input x8 for flow meter with 24V power supply (640mA@24V per port)<br>● Analog Resistor input x6 for leakage detection | ● Rear Analog Voltage input/output x8 for controlling valve with 24V power supply (5W@24V per port)<br>● Analog Current input x8 for flow meter with 24V power supply (640mA@24V per port) |

## 4.2 AMI DCM Features

As data centers expand in size and compute density, managing them becomes increasingly challenging. Additionally, environmental concerns are also becoming more pressing with new regulations, such as the Sustainable Finance Disclosure Regulation (SFDR) [2] and the EU Energy Efficiency Directive (EED) [3], requiring data centers to monitor and report energy consumption and emissions.

AMI DCM is a software solution that can be deployed on physical or virtual machines, that addresses these challenges by improving data center manageability and operational efficiency, as well as supporting compliance with environmental regulations. The top use cases of DCM are shown in Figures 11 and 12.


**Figure 11** AMI DCM's Key Manageability Use cases

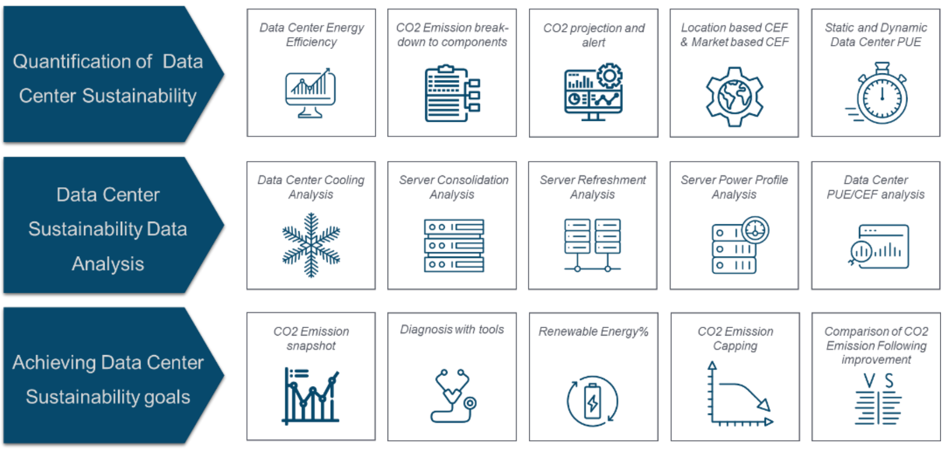


**Figure 12** AMI DCM's Key Sustainability Use cases

DCM supports servers from various vendors across different architectures, along with networking, power, storage, and cooling devices, including the UMS100 system. It continuously monitors different sensors across servers and components, such as CPUs, GPUs, memory, and other devices. This allows it to create detailed cooling analysis to identify hotspots and potential energy savings as shown in Figure 13.

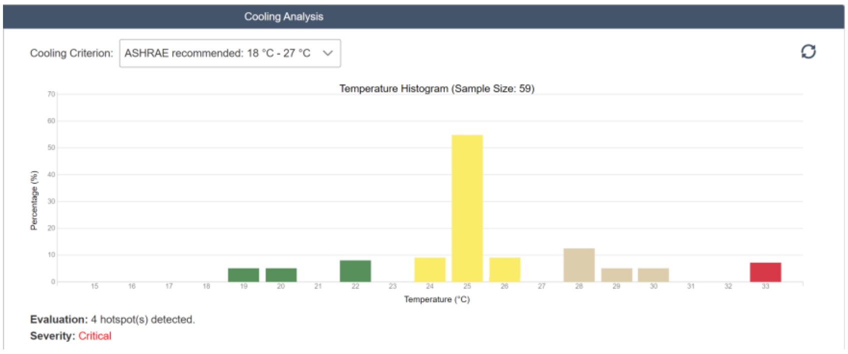


**Figure 13** Room-Level Cooling Analysis Indicating Potential Hotspots.

DCM's support of the UMS100 system in liquid-cooled environments plays key role in identifying and circumventing issues before they become catastrophic by setting thresholds for defined temperature levels, and alerts for critical parameters as previously shown in Figure 7. These capabilities ensure optimal performance and maximize uptime, reduce operational costs, and extend the lifespan of hardware.

DCM's real-time power consumption data when combined with metrics like Power Usage Effectiveness (PUE) and carbon intensity (CO2 emissions per kWh), allow DCM to generate detailed energy consumption reports as shown in Figure 14. These reports enable administrators to set carbon emission thresholds, track progress toward sustainability goals, and take precautionary actions such as server-level power capping when necessary. They also help identify areas for future improvement and ensure compliance with government regulations.

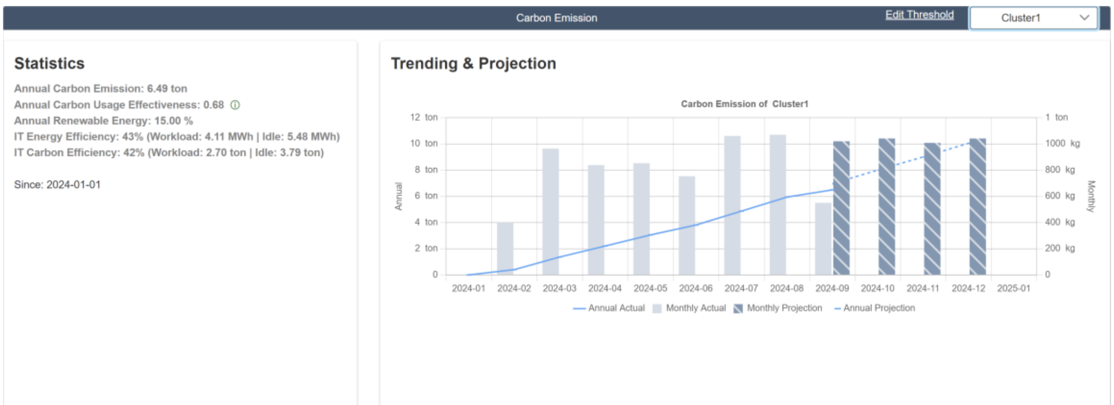


**Figure 14** Monthly and Annual Actual vs. Projected Carbon Emissions for an AI Cluster.

DCM can also perform What-if projections to calculate the impact that server replacement or consolidation, as well as changes in data center PUE or carbon intensity, can have on carbon emissions over time based on historical usage data as shown in Figure 15.

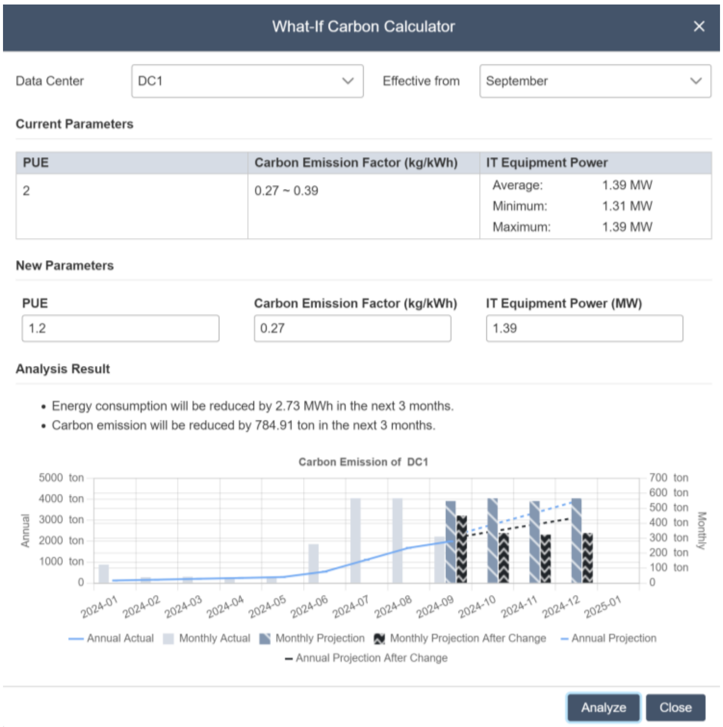


**Figure 15** What-If Carbon Calculator: Projected Reductions in Energy Consumption and Carbon Emissions from Improvements in PUE and Carbon Intensity.

To learn more about AMI DCM, please refer to [4].

# 5. Conclusion

This whitepaper has explored the integration of direct liquid cooling systems in AI clusters, focusing on the architecture and management of such systems. Key components include AI racks, Cooling Distribution Units (CDUs), Wiwynn UMS100 (Universal Management System), and AMI DCM (Data Center Manager).

The Wiwynn UMS100 provides a comprehensive solution for managing the cooling and operational status of servers and racks. Its modular design, support for various sensor data sources, and scalability make it essential for maintaining the stability and performance of the AI cluster. The UMS100 ensures continuous operation and easy integration with other management systems.

AMI DCM serves as the centralized platform for monitoring and managing the AI cluster. By consolidating real-time data from various components of the cluster, DCM ensures that the cluster is operating within the optimal temperature ranges, enhancing performance and energy efficiency.

As data centers scale up to meet the demands of high-performance computing and AI, efficient and flexible cooling solutions become critical. Advanced liquid cooling systems offer higher performance, improved energy efficiency, and enhanced reliability. Looking ahead, the continuous evolution of cooling technologies will pave the way for even more innovative and sustainable solutions, ensuring that data centers remain at the forefront of technological advancements. By embracing these adaptable cooling strategies, data centers can achieve unprecedented levels of efficiency and performance, driving the future of data center management.

# References

[1] Wiwynn whitepaper, “Analysis of Different Rack-Level Direct-to-Chip Liquid Cooling Solutions”, Nov. 2023

[2] European Commission, “Sustainability-related disclosure in the financial services sector”. Available: https://finance.ec.europa.eu/sustainable-finance/disclosures/sustainability-related-disclosure-financial-services-sector_en

[3] European Commission, “Energy Efficiency Directive”. Available: https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficiency-targets-directive-and-rules/energy-efficiency-directive_en

[4] AMI, “AMI Data Center Manager”. Available: www.ami.com/ami-dcm

# About Wiwynn

Wiwynn® is an innovative cloud IT infrastructure provider of high quality computing and storage products, plus rack solutions for leading data centers. We aggressively invest in next generation technologies for workload optimization and the best TCO (Total Cost of Ownership). As an OCP (Open Compute Project) solution provider and platinum member, Wiwynn actively participates in advanced computing and storage system designs while constantly implementing the benefits of OCP into traditional data centers.

For more information, please visit Wiwynn website or contact sales@wiwynn.com
Follow Wiwynn on Facebook and Linkedin for the latest news and market trends.

# About AMI

AMI is Firmware Reimagined for modern computing. As a global leader in Dynamic Firmware for security, orchestration, and manageability solutions, AMI enables the world's compute platforms from on-premises to the cloud to the edge. AMI's industry-leading foundational technology and unwavering customer support have generated lasting partnerships and spurred innovation for some of the most prominent brands in the high-tech industry.

Follow AMI on LinkedIn and Twitter/X to receive the latest news and announcements.