

SUPERMICRO RACK SCALE LIQUID COOLING SOLUTIONS

LIQUID COOLING ADVANTAGES



REDUCTION
Electricity Costs of
Cooling Infrastructure
in Server



REDUCTION
in Electricity Costs for
Entire Data Center



REDUCTION
in Datacenter
Server Noise

SUPERMICRO DIRECT TO CHIP LIQUID COOLING SOLUTIONS

Liquid cooled rack solution that delivers superior performance and efficiency for large scale AI and cloud scale compute infrastructure

- Full turn-key single source solution optimized from proven total solution blueprints of compute, GPU, storage, networking and power and cooling reference designs, with integrated power management tools
- Support highest densities and highest TDP CPUs and GPUs with up to 100KW power and cooling per rack
- Fully validated and tested at system (L10), rack (L11) and cluster (L12) levels
- Accelerated lead times based on in-stock inventory with deployment in weeks versus years
- Enterprise grade redundant cooling pump and power supplies, leak-proof connectors and leak detection



DATA CENTER SCALE LIQUID COOLING REDUCES COSTS AND INCREASES PERFORMANCE

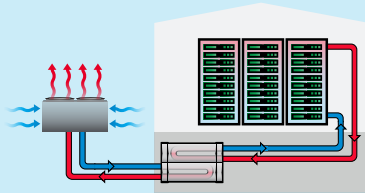
The latest servers with multiple CPUs and GPUs are creating a significant challenge for data center operators. High end servers are now generating up to 10kW of heat, which must be removed from the server. Traditional air cooling through CRAC units, even with hot and cold aisle separation, are expensive and inefficient.

Servers that are application optimized for AI, HPC, and Analytics require the latest in CPU and GPU technologies, which run hotter than previous generations. Multiple CPUs and GPUs per server are needed for performance intensive computing, driving up the electricity demands for the server as well as at the rack level.

RACK SCALE SOLUTION

Rack Level

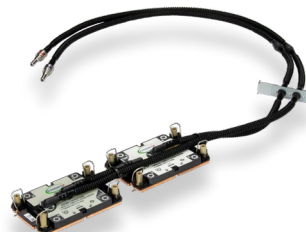
- Design
- Assembly
- Configuration
- Testing
- Logistics



Horizontal CDM

Hoses and Connectors – Flexible hoses are used to bring the cooler liquid to the CPUs and GPUs cold plates and return the hot liquid to the CDMs. The Supermicro connectors are single handed, 0 drip, quick disconnectors, they allow operators to service liquid cooled systems safely and efficiently.

Cold Plates – The cold plates are placed on top of the CPUs and GPUs; flowing coolants through their micro sized channels cools down the chips very efficiently. The Supermicro cold plates are designed to reduce hot spots on the chip and reach ultra-low thermal resistance.



CPU Cold Plate



GPU Cold Plate

COMPONENTS OF AN EFFICIENT LIQUID COOLING SOLUTION

Supermicro's liquid cooled rack solution is made of several components that are designed in house to achieve the highest level of performance and reliability. All the components are integrated as a rack level solution providing a true one-stop shop customer experience. The critical components of Supermicro's liquid cooled rack solution are:

Coolant Distribution Unit (CDU) – contains the pumping system that circulates the coolant to the cold plates cooling down the CPUs and GPUs. The Supermicro CDU integrates 2 hot-swappable and redundant pumping modules and power supply modules guaranteeing nearly a 100% uptime to the operator. The CDU cooling capacity is up to 100kW enabling extremely high rack densities. The CDU also offers an easy-to-use touch screen to monitor and control the rack operation with WebUI access and integrated in Supermicro's Super Cloud Composer datacenter management software. The CDU control system optimizes power consumption while ensuring efficient cooling is delivered to all the CPUs and GPUs. An effective anti-condensation strategy is adopted to prevent any hardware degradation.



Supermicro CDU

Coolant Distribution Manifold (CDM) – The CDM are the distribution pipes that supply coolant to each server and collect the hotter coolant back to the CDU. There are two types of CDMs:

Vertical – Vertical manifolds are placed at the back of the rack and directly connected to the CDU with hoses. They deliver coolant to the cold plates on systems with inlet and outlet hoses at the back of the rack.

Horizontal – Horizontal manifolds are placed at the front of the rack in a 1U rack mount space. They connect the vertical manifolds at the rear of the rack to cold plates on systems with inlet and outlet hoses at the front of the rack (SuperBlades and 8U GPU servers)

RACK SCALE ADVANCED ENGINEERING

As the Rack becomes the unit for scalable computing, engineering a liquid cooled rack requires careful planning and expert assembly. A liquid cooling system needs to be flexible to handle a wide range of servers, while being able to cool an entire rack of high performance systems.

RACK SAMPLE CONFIGURATIONS



SRS-48UDTN-SKU1-L1-SMCI

Up to 4 GPU Servers (8U, 8 NVIDIA H100 HGX GPUs) per 42U Rack (32 NVIDIA H100 HGX GPUs total)

Product	Qty.
Supermicro GPU Server SYS-821GE-TNHR	4
CDU	1
Vertical CDM	1
Horizontal CDM	4
SWITCH	1



SRS-48UBTW-SKU1-L1-SMCI

Up to 76 Servers Nodes / 19 Systems in a 48U Rack

Product	Qty.
BigTwin® (2U4N) or 2U Hyper or 1U Hyper (shown)	19 21 44
CDU	1
Vertical CDM	1
SWITCH	1



SRS-48UBLD-SKU1-L1-SMCI

Up to 80 Server Blades / 4 Systems in a 48U Rack

Product	Qty.
SuperBlade® Enclosure (8U20N)	4
SBI-421E-1T3N	80
CDU	1
Vertical CDM	1
Horizontal CDM	8
SWITCH	1



GPU

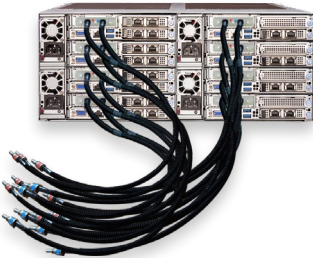
***SYS-421GU-TNHR, SYS-421GE-TNR & SYS-821GE-TNHR
AS-8125GS-TNHR & AS-4125GS-TNRT***

Supermicro GPU systems are at the heart of today's AI and HPC excitement by combining the fastest processors, memory, and GPUs in a family of systems for AI/ML, Inferencing, and HPC. The 2U, 4U or 8U GPU systems support 4 or 8 NVIDIA® H100 GPUs together with NVLink® and NVSwitch respectively and are powered by up to the 4th Gen Intel® Xeon® Scalable processors or up to the AMD EPYC™ 9004 Series processors. In addition, up to 32 DIMMs of DDR5 memory can be installed, providing an extremely compact and powerful AI or HPC system. Finally, D2C coolers are mounted on each of the processors and GPUs within the GPU system and routed through CDM loops to the Liquid Cooling CDU.

BigTwin®

SYS-221BT-HNTR & SYS-221BT-DNTR

The Supermicro BigTwin represents flagship performance for the most demanding applications and HCI environments. The innovative design supports up to four nodes in a 2U enclosure with no-compromise support for processors, memory, and I/O. Each node can support dual 4th Gen Intel® Xeon® Scalable processors, up to 16 DIMMs of DDR5 memory, and up to twelve high speed NVMe drives. AIOM (superset of OCP 3.0) networking options include 10GbE, 25GbE, 100GbE, and InfiniBand (200 Gb HDR per port). Shared power and cooling maximize the resource savings of the multi-node design. D2C coolers are mounted on the processors within each BigTwin node and routed through a CDM loop to the Liquid Cooling CDU.



FatTwin®

SYS-F511E2-RT & SYS-F521E3-RTB

The Supermicro FatTwin® are high-density systems offering advanced multi-node 4U twin architecture with 8 or 4 nodes (single processor per node). These systems with a front-accessible service design allows cold-aisle serviceability, with highly configurable systems optimized for data center infrastructure with compute and storage density and options. In addition, the Supermicro FatTwin supports all-hybrid hot-swappable NVMe/SAS/SATA hybrid drive bays with up to 6 drives per node (8-node) and up to 8 drives per node (4-node).

SuperBlade®

SBI-421E-1T3N

A shared cooling, power, and networking infrastructure is key to the high density and server efficiency offered by the SuperBlade. Supermicro's high performance, density optimized, and energy-efficient SuperBlade supports up to 20 blade servers in an 8U chassis, with a choice of the 4th Gen Intel® Xeon® Scalable processors or 4th Gen AMD EPYC™ processors. With advanced networking options, including 200G HDR InfiniBand, Supermicro's new generation blade product portfolio has been designed to optimize the TCO of critical criteria for today's data centers, e.g., power efficiency, node density, and performance. A D2C cooler is mounted on each processor within the SuperBlade system and routed through a CDM loop to the Liquid Cooling CDU.



Hyper

SYS-221H-TNR & SYS-121H-TNR | AS-2125HS-TNR & AS-1125HS-TNR

Supermicro Hyper servers are designed to deliver the Hyper Family – The X13 Hyper series brings next-generation performance to Supermicro's flagship range of rackmount servers, built to take on the most demanding workloads along with the storage & I/O flexibility that provides a custom fit for a wide range of application needs. Supermicro Hyper systems are available in 1U or 2U versions, with up to 32 DIMM slots. With the cooling capacity to accommodate the highest performing CPUs, the Supermicro Hyper product family is optimized for maximum compute performance.

Product Family	Server	Description
GPU	SYS-821GE-TNHR	Dual 4 th Gen Intel® Xeon® Scalable Processor 8U, 32 DIMMs HGX H100 8-GPU SXM5 Multi-GPU Board
	AS -8125GS-TNHR	Dual 4 th Gen AMD 9004 Series Processors 8U, 24 DIMMs HGX H100 8-GPU SXM5 Multi-GPU Board
	SYS-421GU-TNXR	Dual 4 th Gen Intel® Xeon® Scalable Processor 4U, 32 DIMMs HGX H100 4-GPU SXM5 Multi-GPU Board
	SYS-421GE-TNR (PCIe)	Dual 4 th Gen Intel® Xeon® Scalable Processor 4U, 32 DIMMs GPU-NVH100-80,GPU-NVA100-80-NC
	AS -4125GS-TNRT (PCIe)	Dual 4 th Gen AMD 9004 Series Processors 4U, 32 DIMMs Up to 8 Double-Width/Single-Width Cards (Full Height Full Length) NVIDIA H100 and AMD MI200 series
BigTwin®	SYS-221BT-HNTR	Dual 4 th Gen Intel® Xeon® Scalable Processor 2U, 4-Nodes, 16 DIMMs
	SYS-221BT-DNTR	Dual 4 th Gen Intel® Xeon® Scalable Processor 2U, 2-Nodes, 16 DIMMS
FatTwin®	SYS-F511E2-RT	Single 4 th Gen Intel® Xeon® Scalable Processor, 4U, 8-Nodes, 16 DIMMs
	SYS-F521E3-RTB	Single 4 th Gen Intel® Xeon® Scalable Processor, 4U, 4-Nodes, 16 DIMMS
SuperBlade®	SBE-820C/J/J2/L/H-820	8U Enclosure
	SBI-421E-1T3N	Dual 4 th Gen Intel® Xeon® Scalable Processor 16 DIMMS
Hyper	SYS-221H-TNR	Dual 4 th Gen Intel® Xeon® Scalable Processor 2U, 32 DIMMs
	SYS-121H-TNR	Dual 4 th Gen Intel® Xeon® Scalable Processor 1U, 32 DIMMs
	AS -2125HS-TNR	Dual 4 th Gen AMD 9004 Series Processors 2U, 24 DIMMs
	AS -1125HS-TNR	4 th Gen AMD 9004 Series Processors 1U, 24 DIMMS



GPU Server
8U 8 H100 SXM GPU

SYS-821GE-TNHR
AS -8125GS-TNHR



GPU Server
4U 4 H100 SXM GPU

SYS-421GU-TNXR



Universal GPU
4U 8/10 GPU PCIe

SYS-421GE-TNR
AS -4125GS-TNRT



BigTwin®

SYS-221BT-HNTR
SYS-221BT-DNTR



SuperBlade®

SBI-421E-1T3N



Hyper

SYS-121H-TNR
AS -1125HS-TNR



Hyper

SYS-221H-TNR
AS -2125HS-TNR



4U 8 Node FatTwin®

SYS-F511E2-RT



4U 4 Node FatTwin®

SYS-F521E3-RTB

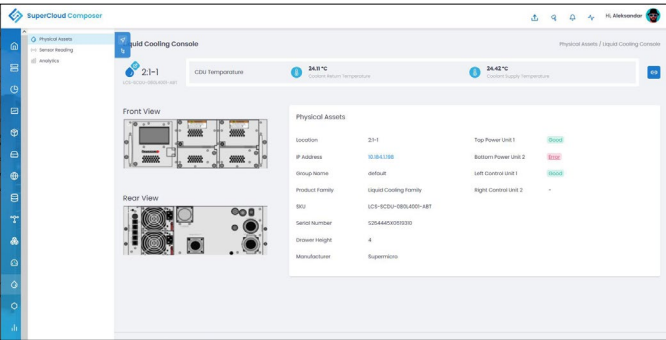
Liquid Cooling Tower

The Supermicro Liquid Cooling Tower solution is versatile and energy efficient at removing the heat produced by today's latest servers. The system is optimized to transport and remove the heat from today's most powerful AI servers. Supermicro is pioneering a way to obtain a complete AI solution, from the servers to the cooling infrastructure.



Supermicro SuperCloud Composer for a Liquid Cooled Data Center

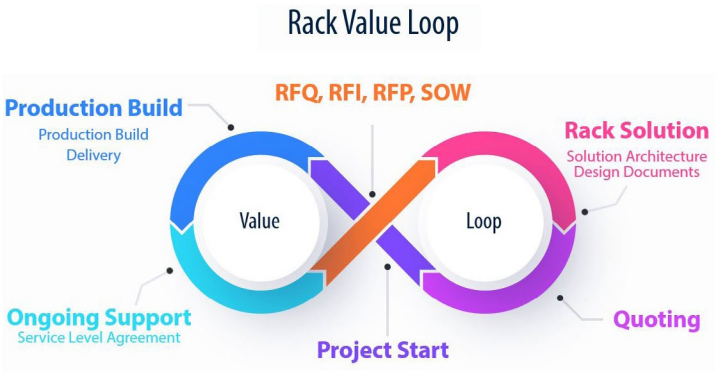
SuperCloud Composer's LCCM (Liquid Cooling Consult Module) is a powerful tool to collect vital information on physical assets and sensor data from a CDU (Cooling Distribution Unit), including pressure, humidity, pump and valve status, and more. CDU data is presented in real-time, enabling users to monitor operating efficiency of their liquid cooled racks. Using these insights, SuperCloud Composer (SCC) helps the user to set up alerts, manage firmware updates, and more.



Supermicro's Rack Integration Services, Turnkey Cluster Level Liquid Solutions

Supermicro's Rack Integration Services leverage application-optimized motherboards, chassis, cooling subsystems, networking components, cluster management tools, energy-efficient power supply technologies, and compact enclosures to design and develop customized and enterprise solutions. Supermicro understands the importance of today's fast pace business problems and customer requirements; therefore, we offer an end-to-end integration service that helps customers reduce overhead, maximize efficiency and quality, making this a competitive strategy and a quick go-to-market advantage.

Supermicro works with leading organizations in all geographies to design, install, and test various liquid cooling solutions. The Supermicro process involves a rigorous set of phases that ensure the most optimized and tested solution for environments where liquid cooling is required for maximum performance.



Supermicro Rack Scale Liquid Cooling Solutions

ABOUT SUPER MICRO COMPUTER, INC.

Supermicro® (NASDAQ: SMCI), the leading innovator in high-performance, high-efficiency server technology is a premier provider of advanced server Building Block Solutions® for Data Center, Cloud Computing, Enterprise IT, Hadoop/Big Data, HPC and Embedded Systems worldwide. Supermicro is committed to protecting the environment through its “We Keep IT Green®” initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

www.supermicro.com/liquidcooling

No part of this document covered by copyright may be reproduced in any form or by any means — graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system — without prior written permission of the copyright owner.

Supermicro, the Supermicro logo, Building Block Solutions, We Keep IT Green, SuperServer, Twin, BigTwin, TwinPro, TwinPro², SuperDoctor are trademarks and/or registered trademarks of Super Micro Computer, Inc.

All other brands names and trademarks are the property of their respective owners.

© Copyright 2023 Super Micro Computer, Inc. All rights reserved.

Printed in USA

 Please Recycle

02_LiquidCooling-2023_R06

