INTEGRATED SOLUTIONS INSIGHTS

/ LIQUID COOLING: YOUR GUIDE TO IMPLEMENTATION

Liquid cooling has come in from the cold.

You've already heard the buzz about liquid cooling and understand that when you're dealing with high-powered computing, it's a game changer. It's no longer a niche technology used by the few. It's now a necessity for businesses that need performance at scale and it's predominantly driven by the increasing demands of Artificial Intelligence (AI) workloads.

But what you may not know is how to move from your current solution to liquid cooling in a way that satisfies both your performance needs and energy limitations. You may also be concerned about energy consumption, server density or efficiency, sustainability, or be looking to reduce your carbon footprint.

In this guide, find out more about scaling your high-performance solution including:

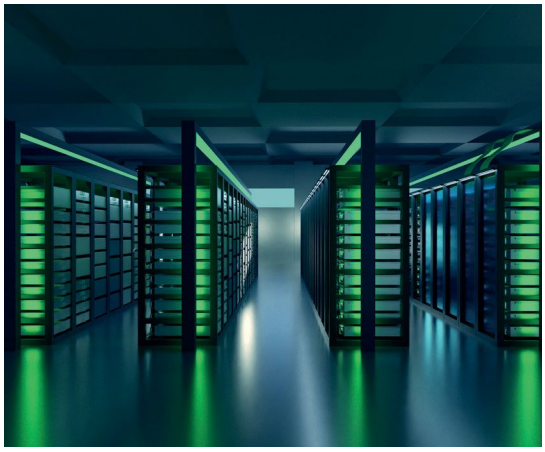
- What's involved in converting your existing solution to liquid cooled
- Operating with air cooled and liquid cooled environments in the same data center
- What you need to know for harsh, edge environments
- How to keep your tier one partners' warranties in place as you utilize liquid cooling

We will cover the two most common scenarios on your journey to implement and realize the power of liquid cooling. We've shared our industry-leading expertise which you can lean on without hiring new resources. The insights you can unlock here will help you implement these powerful solutions today to help you future proof your cooling solutions.

With Integrated Solutions and Lenovo OEM Solutions working together to design, build and deploy at scale, we're the coolest partnership on the block.

"It takes expertise to build liquid-cooled servers in volume, and to support and manage these servers in the field. Avnet is investing time and resources with our OEM partners to get these solutions to market successfully."

Andrea Tsapralis
Director
Solutions Development & Innovation
Avnet



/ WITH GREAT POWER COMES LIQUID COOLING

Liquid cooling is becoming more popular for a whole range of different, business-driven reasons.

Efficiency in power consumption and rising energy bills, the need to conserve water, reduce noise, or have servers working in inhospitable environments are all driving interest in more efficient cooling. But one of the biggest drivers is the increasing performance capability of semiconductors, including CPUs, GPUs and memory, as the demand for faster, higher performance computing continues.

The meteoric rise of generative AI is driving an unprecedented demand for power at the server and rack level to deliver the performance businesses require. For example, in March 2024 NVIDIA launched its Blackwell platform to enable AI training and real-time Large Language Model (LLM) inference for models scaling up to 10 trillion parameters.*

A recent report by Accenture noted that 98% of companies state that AI will play an important role in their strategy in the next five years. McKinsey analysts say that 65% plan to increase their investment in AI in the next three years.**

These changes at the business level are leading to increased interest in liquid cooling for 3 main reasons:

1. Increasing compute demand

AI-driven workloads demand huge volumes of compute ability. Applications like large language models, generative content and precision medicine are integrating with complex, distributed platforms and driving an ever-increasing need for high performance which in turn need high performance semiconductors – and these run HOT.

2. Energy efficiency

Energy costs are rising and the demands on energy are growing exponentially with new semiconductors driving higher-powered compute capabilities. Efforts to reduce power, water and noise along with a focus on cost reduction adds even more tension. A drive toward sustainability, a focus on environmental issues, and organizational pressures are driving interest in more effective cooling.

3. Compute density and location

Hotter chips, increased compute density, and a need to deliver high performance, increase the demands on space. But it is not always possible to expand on the same site. Liquid cooling can bring increases in efficiency and CPU density without a loss of performance or need for physical expansion.

*https://nvidianews.nvidia.com/news/nvidia-blackwell-platform-arrives-to-power-a-new-era-of-computing

**https://www.forbes.com/sites/forbestechcouncil/2023/09/11/ai-liquid-cooling-and-the-data-center-of-the-future/

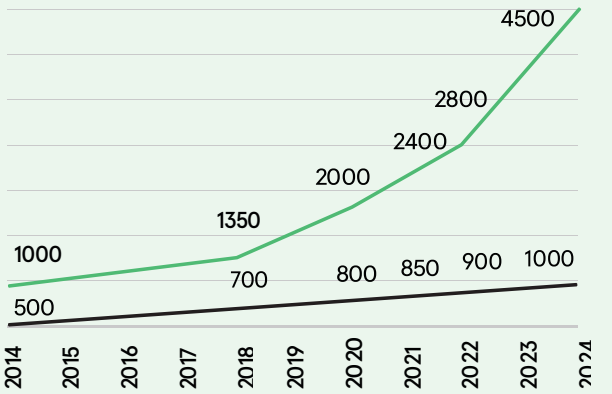
/ POWER AND COOLING TRENDS

In the data provided by Lenovo, you can see the sharp increase in the number of liquid-cooled 1U nodes over the last decade. As the power increases, liquid-cooled solutions are gaining traction as a way to solve issues.

"We started to see trends in power and cooling steadily rising as the need for high performance computing came in with early adoption of AI. As the number of thermal design points goes up, you have more transistors in a small area, which generate more heat. Then you start to have a density issue and thermal issues where there's high concentrations of heat which starts to affect the performance of the silicon. More and more power starts to become allocated to cooling to maintain performance. Without more efficient cooling, your footprint starts to shrink because more space and power are given over to cooling. That's where liquid cooling comes in."

Matt Ziegler
Director, Neptune
Lenovo

1U Power/Cooling Trend (Watts)*



CPU Thermal Design Power Wattage Trend*

	<205W	<205W-240W	<250W
1U Half Wide	Air/Hybrid	Water	Water
1U Full Wide	Air/Hybrid	Air/Hybrid	Water
2U Half Wide	Air	Air	Water
2U Full Wide	Air	Air	Air

*Lenovo Partner Technical Webinar – Neptune Liquid Cooling, 2023-10-24, Course Code 102023, Matt Ziegler

/ WHAT ARE THE DEMANDS OF TOMORROW'S COMPUTE WORKLOADS?

As the race for higher performance and faster chips continues to facilitate the development of massive and complex computing systems, the demand for cooling capabilities to match is building.

Liquid cooling can solve the high-density, server-cooling problem, because liquid can conduct more than 3,000 times as much heat as air and requires less energy to do so, according to Gartner®.*

So it's easy to see why liquid cooling is gaining ground. But what are the other considerations?

Future compute needs

Quite simply, air-cooled technology cannot keep up with the performance demands of new chips required for systems using AI. Traditional form factors will be challenged with air cooling.

The need to reduce 'lag' between the compute power and workstations on power-hungry workloads will also start to mean that more compute moves out of remote data centers and into shared workspaces. Also, as edge computing can often be in hostile, dirty, or harsh environments, efficient cooling of edge systems that handle more data need to be considered.

Sustainability and efficiency business drivers

As businesses look to recycle heat and/or water, liquid cooling offers clear benefits. Reuse of water within business environments where it is already available on site makes better financial and environmental sense. This drive not only saves money but helps with meeting net-zero targets, corporate social responsibility commitments, and sustainability ambitions.

Thermals capping performance

Keeping electronic components at optimal temperatures isn't just about immediate performance—it's also about longevity. Every 10°C rise in temperature can halve the life of electronic components. By more effectively managing heat, liquid cooling can extend the operational life of these components, reducing replacement costs and downtime. Also as chips reach peak temperature, self-limiting 'throttling' starts to affect processing speed, as chips slow down their processing to generate less heat.

Data center limitations

Companies are looking for ways to continue to optimize the physical space in data centers, whereas liquid cooling can maximize TFLOPS per square inch.

Increasing electricity costs

As power costs rise, air cooling becomes inefficient and more costly to maintain. As chips get faster and hotter, more power is required to cool them. It can be difficult to bring more power into existing locations and the cost of that power can become prohibitive.

*https://www.gartner.com/en/information-technology/glossary/liquid-cooling

/ HOW DOES LIQUID COOLING WORK?

A huge proportion of energy used by servers is involved in cooling them down – incredibly sometimes this can be almost half the energy used. Cooling occurs through fans and heat sinks at the CPU level, and facility cooling at a more macro level. The technical way to measure this is through a ratio called Power Usage Effectivity, or PUE. In a perfect world, 100% of the energy supplied to a server would be used for its computing tasks, but a proportion is always wasted because the server produces heat while it works. The PUE is the measure used to capture this.

As chips provide more processing power, they get hotter than ever. This means the space and energy needed to cool them keeps rising. While traditional air cooling blows cold air over the chipset, liquid cooling works by taking heat away from the hot, hardworking components of a server via a liquid coolant. The heat is transferred to the liquid, which is transported away to be cooled by air or another coolant and then recirculated.

In particular, because of the closed nature of partial or fully immersed liquid-cooling solutions, they are especially suitable for any environments with

- **High temperatures**
- **Low latency demands**
- **Dirty or inhospitable environments**

Liquid-cooling systems can be more expensive than air-cooled systems to deploy initially, but the benefits are clear when you look at the total cost of ownership (TCO). Also, effective cooling means less equipment is actually required and systems can be denser, taking up less space and with less mechanical infrastructure.

According to Paul Coates, "Liquid only SKUs are coming into play with tier one providers such as Lenovo, Intel, AMD and NVIDIA. This is a trend we expect to see building in the future."

"Liquid cooling isn't just a trend or a mere alternative – it's the embodiment of innovation, sustainability and the very essence of adapting to changing technological needs."

Paul Coates
Solutions Manager
Avnet

/ AN INTRODUCTION TO LIQUID COOLING

from Paul Coates, Solutions Manager at Avnet and Author of "Cooling with Liquid: Electronics Take the Plunge"

There are four types of liquid cooling. Here's an introduction to all of them and we'll provide details of two of them to show you steps you may decide to take on your liquid-cooling journey.

Rack Solutions

Usually referred to as "Chilled Doors" or "Rear Door Heat Exchangers." Enhanced cooling of the servers is provided by replacing the rear door of the rack with one that has an integral heat exchanger and fans. In this case, warm air in the facility passes through a radiator (much like the one in your car), cooling it. The cooled air then is used in exactly the same way to cool the servers.

Direct-to-Chip

Direct-to-chip liquid cooling replaces traditional heatsinked components with a copper block. Cool water (typically used) flows through a closed circuit, entering through an inlet pipe and transfers heat from the device. The heated water is then removed through an outlet pipe and cooled, usually by a chiller, before being recirculated. This method predominantly uses water as the coolant, distinguishing it from most other liquid cooling solutions.

Precision Cooling

Precision cooling in dielectric liquid cooling involves the submersion of components or sections of servers in the liquid coolant while allowing other parts to typically remain exposed to air or other cooling methods using liquid. Different providers of liquid-cooling solutions offer their own variations of this approach. Depending on which is adopted, the server may be orientated either vertically or horizontally.

Full Immersion

In fully immersed dielectric cooling solutions, servers are completely submerged in a vertical orientation in dielectric liquid. The vast majority of these solutions are realized as "Tank Solutions". While offering exceptional power efficiency, this method faces constraints regarding space, weight and cost. Typically deployed in new or refurbished areas of data centers, full immersion stands as the primary option for both single and two-phase cooling solutions.

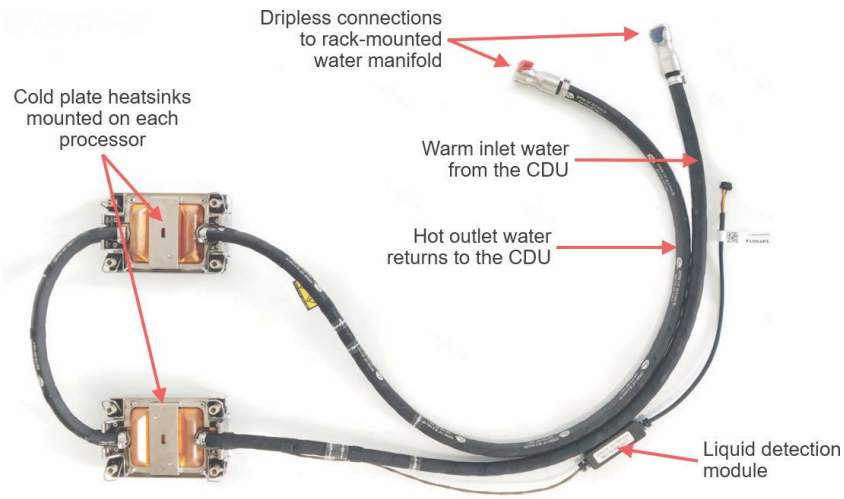
/ IMPLEMENTATION WITH LIQUID COOLING – SCENARIO 1: DIRECT-TO-CHIP

The journey to liquid cooling can be daunting. Most often people don't start with a greenfield data center, unlimited budget and full immersion. But if one thing is clear, it's that the demand for more, better and faster chips is here to stay.

Liquid cooling can help where energy pricing is a driver, where you are pushing for more sustainability, or you just need to manage bigger workloads.

Perhaps your first step toward liquid cooling is direct-to-chip.

In each CPU, chips have traditionally been air cooled through fans and heatsinks designed to allow air to flow across the chips and remove heat. With direct-to-chip solutions, liquid is piped to a cold plate (which replaces the heatsink) on the surface of the chip. This removes heat much more effectively and the warm water is then siphoned away (and often re-used).



Lenovo Neptune Processor DWC Module

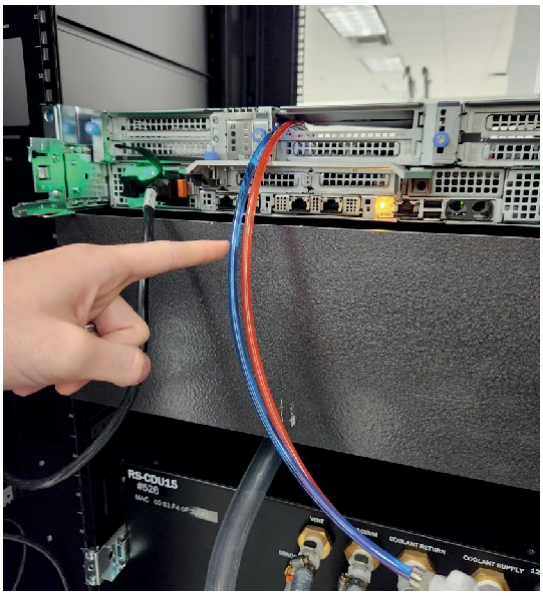
Here's how direct-to-chip helps:

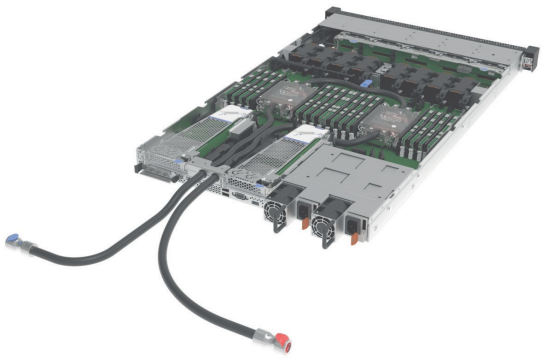
Your business is investing in AI and upgrading to servers that have high performance capabilities and power needs to match. Hotter chips are throttling performance and you can't pull more air through the systems to cool them down.

In this case, with direct-to-chip liquid cooling, you can expect a relative reduction in your power needs for cooling of between 20-40%. The solution can be retrofitted in your existing data center and won't take up more space.

This solution can be relatively cost effective and straightforward to install, giving you reduced power needs and allowing your chips, racks and servers to work toward their capability.

/ IMPLEMENTATION WITH LIQUID COOLING – SCENARIO 1: DIRECT-TO-CHIP (CONTINUED)





Lenovo Neptune Processor DWC Module installed in the SR630 V3

Advantages:

Targeted cooling: Hotspots, common in high-performance computing, can be effectively managed by directly targeting the primary heat producers.

Integration with existing systems: Often, direct-to-chip cooling can be integrated into existing server setups with fewer modifications than immersion setups.

Space efficiency: These systems are compact and can be more easily integrated even in space-constrained environments.

Disadvantages:

System complexity: Ensuring proper and uniform flow to all chips requires intricate design and accurate pump systems.

Leakage: There's a potential risk of leaks, which can lead to damage if not promptly detected. Negative pressure technology is available that can eliminate this threat.

Targeted cooling: Only those chips that are directly cooled benefit since the rest of the server is cooled with air.

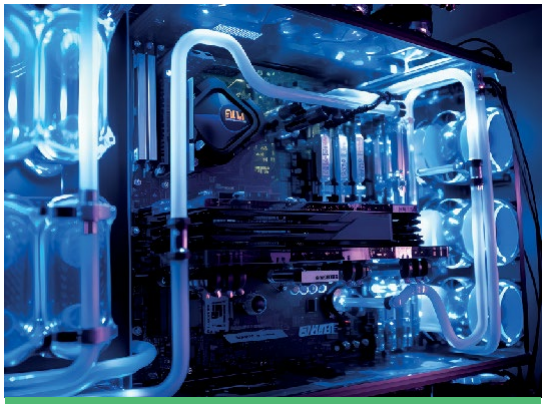
/ GETTING STARTED WITH LIQUID COOLING – SCENARIO 2: PRECISION COOLING

Another common step into liquid cooling is the use of partial, precision cooling.

As in direct-to-chip, this can mean a range of interim measures that can be taken that don't necessitate full immersion.

Precision cooling in dielectric liquid cooling involves the submersion of components or sections of servers in the liquid coolant while allowing other parts to typically remain exposed to air or other cooling methods using liquid.

Different providers of liquid-cooling solutions offer their own variations of this approach. Depending on which is adopted, the server may be orientated either vertically or horizontally.



Here's how precision cooling helps:

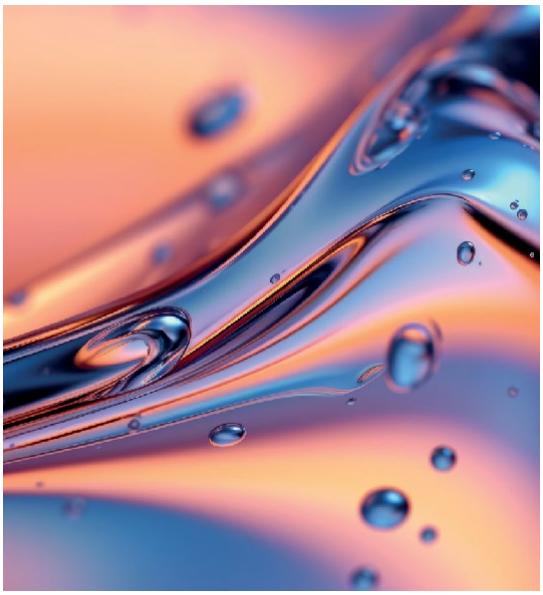
You are looking to reduce the amount of power being used in your data center because electricity costs are rising and you know that almost half of your energy costs are being spent on air cooling.

Precision cooling means using a liquid-cooled solution through partial immersion of some parts of the server in the cooling liquid.

This works more effectively than air cooling and so can bring down your power consumption costs.

It is also uses space more efficiently than simply adding more air cooling within the servers themselves or the space they are stored in.

/ GETTING STARTED WITH LIQUID COOLING – SCENARIO 2: PRECISION COOLING (CONTINUED)



Advantages:

More comprehensive cooling: Precision cooling can limit the need for infrastructure changes.

Scalability: Suitable for a wide range of applications, from small-scale setups to vast data centers.

Reduced acoustic noise: With fewer fans immersion-cooled systems are much quieter.

Lower maintenance: Fewer moving parts mean potential reductions in wear and tear and resultant maintenance.

Enclosed chassis: Sealed chassis enables the server to operate in harsh, edge environments.

Disadvantages:

Lead time: Introducing a new server type can take longer due to the optimization required.

Component accessibility: Servicing submerged components can be challenging.

Weight and handling: Liquid is heavy and may necessitate special handling and placement considerations such as reinforced floors.

/ HOW YOU WIN IN LIQUID COOLING WITH AVNET INTEGRATED SOLUTIONS

Avnet is one of the few global suppliers in this space that can deliver an end-to-end solution. We have the expertise you need at every stage of your liquid-cooling journey to ensure success.

We take time to ensure we understand the specific needs of your project and what's driving your decision to consider liquid cooling. We can recommend the best solution objectively, depending on your needs. Whether it's more efficient power consumption, a harsh environment, the need to reduce noise, space or just remove heat from faster, hotter chips, we take time to understand the problem you are solving for and recommend accordingly.

Avnet begins by helping you select the best tier one provider for the hardware you need to power your solution. Because we are a supply chain leader and have long-lasting relationships, we can get the systems you need to get what work you need done. Working with these partners, we help you come up with the custom-built, integrated solution you need.

Then we add in the layering of the liquid-cooled solution. We recommend the optimal cooling method using the best partner to get you the easiest and best solution at scale.

We can help you deploy and service these solutions all around the globe. Choosing Avnet to integrate your solution means less risk for your business.

Technology Enablement

- Industry-leading technology partners
- Direct, OEM relationships
- Experienced experts

Solution Development for Scale

- Global SKU optimization, sustainability options, product lifecycle management
- Regulatory and test automation
- POC planning and management

Systems Integration

- Servers, clients, storage arrays, networking
- Basic to complex
- Stand-alone, node, rack

Supply Chain

- Collaborative, hands-on supply chain orchestrators
- Transparent and direct communication
- Material planning and staging
- Build-strategy options

Solution Deliver

- Site surveys
- Field installation at scale
  - Node to rack
  - Application load and configuration
  - Segment specific expertise (i.e. healthcare, retail, industrial)
- Remote engineering
- Technical project management

Warranty and Support

- Forward stocking locations
- Technician dispatch
- Parts dispatch
- Service parts planning
- Reverse logistics
- RMA Management
- Repair in and out of warranty

/ DESIGN TO SCALE

Solution	Development	Integrate	Orchestrate	Deploy	Support
We help you bring your IP to market on world-class technologies	Enhance your solution's performance, reliability and profitability	Advanced system integration node to rack, and regional to global	On-demand expert resources to deploy to your locations worldwide	Technology, delivery, warranty and support services	Tools, processes and talent to perfectly align supply to demand

Solution development
We will work with you to understand your business drivers and identify an appropriate solution. This might be for a new site, or you might be looking to retrofit liquid cooling in your existing operation. Working together, Avnet will draw on our experience in this space to support you in developing a solution that works for you.

Systems integration and manufacturing
Once designed, we can help bring the entire solution together. Avnet's engineering organization has been working with IP providers to provide scale and production recommendations. We are engaging with industry certification organizations to provide CE, UKCA, UL.

Financial transactions
Setting up accounts, managing order entry, quoting, credit and contracts can all be facilitated by Avnet. This allows you to leverage our extensive financial expertise to engage directly with your customers or end users and use otherwise tied up capital to extend your advantage in the market.

Solution deployment and support
Our Services organization is creating programs for liquid cooling end-to-end support (replacement parts programs, entitlement management, forward stocking locations, technician and parts dispatch, and on-site support). And of course, we can manage all the elements of transportation, distribution, deployment and installation globally.

/ LENOVO'S LIQUID COOLING PORTFOLIO

Neptune™ from Lenovo enables performance without compromise. It delivers:

Sustainability

Helping to meet environmental goals and align with carbon reduction initiatives without sacrificing the ability to deliver the compute power your organization needs to compete and thrive.

Density

Neptune™ delivers more computing power in a compact footprint to support your most demanding workloads.

Competitive edge

Whether your organization is competing for faster time-to-market, scientific breakthroughs or research grants, you can benefit from super-efficient, high-performance computing that delivers results more quickly.

"Our industry-leading Lenovo Neptune™ liquid-cooling technology can be applied on CPUs and GPUs that scale from a single server to solutions with hundreds of racks, driving energy and performance efficiency that can be directed back into the business to help accelerate innovation and drive competitive outcomes."

Scott Tease
Vice President and General Manager
Lenovo

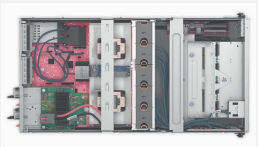
40%
Lower power consumption

95%+
Heat removal efficiency

Lenovo has been providing smarter liquid-cooling solutions for a decade. Neptune™ provides a holistic approach to cooling that goes beyond cooling just the CPU. It includes

Rack rear door water cooling

Direct water cooling to each node



Liquid-assisted cooling in an air cooled system

Lenovo's ThinkSystem and Neptune™ Portfolio

Lenovo ThinkSystem SD665 V3

The ThinkSystem SD665 V3 delivers maximum AMD CPU performance, extreme density, and leading energy efficiency.

ThinkSystem SD665-N V3 Neptune DWC

The ThinkSystem SD665-N V3 Neptune™ DWC implements Lenovo Neptune™ Direct Water Cooling (DWC) technology to optimally support workloads from technical computing, grid deployments, analytics, and is ideally suited for fields such as research, life sciences, energy, simulation, and engineering.

Lenovo ThinkSystem SR675 V3

From single-node enterprise deployments to the largest supercomputers in the world, the SR675 V3 server can scale to meet any performance demand.

ThinkSystem SD650 V3 Neptune™ DWC Server

The ThinkSystem SD650 V3 Neptune™ Direct Water Cooled (DWC) node is the next-generation high-performance server based on the fifth generation Lenovo Neptune™ direct water-cooling platform.

ThinkSystem SD650-I V3 Neptune™ DWC Server

The ThinkSystem SD650-I V3 Neptune™ DWC node maximizes compute power by integrating 4th Gen Intel Xeon Scalable processors, Intel Xeon CPU Max Series processors and Intel Data Center GPU Max Series accelerators.

Our team of Avnet and Lenovo OEM Solutions specialists can help with all aspects of hardware, software, integration, product design, development and deployment. Put our experience to use on your liquid-cooling project.

Let's talk!

Contact our team at integrated@avnet.com to find out more and discuss how we can develop and deliver your liquid-cooling solution, globally.