Sponsored by

VERTIV.

The Future of Cooling Supplement

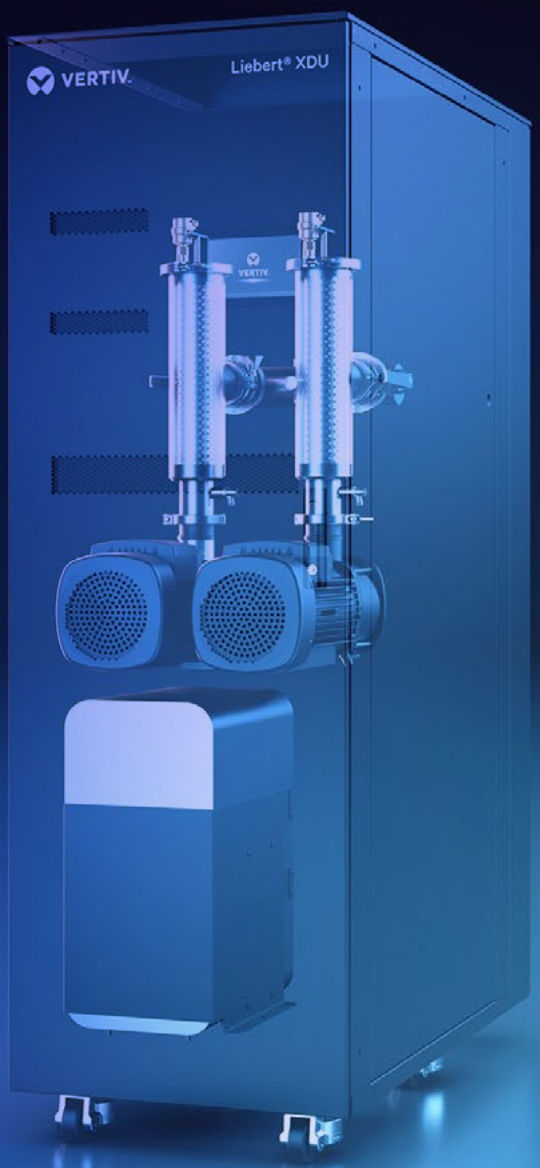
Preparing for a high-density data center

INSIDE

VERTIV.

Designed for density. **Ready for acceleration.**

Manage increased heat with liquid cooling technology.

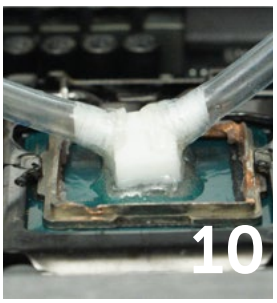


See how we can help you at **Vertiv.com/GetCool**

DCD

Sponsored by VERTIV.

Contents



The physics of cooling

Thermodynamics comes for us all sooner or later, and you can't negotiate with those laws. Data centers now use a significant amount of energy and generate a level of heat that impacts their surroundings.

To deal with this, we have to find most creative way to work within the rules decreed by physics.

This has implications at the macro level, for the whole building, and the surrounding area.

It also opens possibilities at the micro level, down to the processors, and even inside them.

This supplement shows how engineers can find a way forward for the next generation of data center technology.

Use everything twice

All the energy we use in IT systems turns into heat. The heat is the one inevitable output, and heat is a valuable commodity.

So what happens if you flip your view, and view data centers as a heat source which also provides compute as a byproduct?

This thinking leads to heat reuse, and that brings up the prospect of district heating systems, where data centers channel their warm air and hot water to their neighbors.

If we could make sure all our compute energy does two jobs, we could have all the energy we need - but there's more than one way to do this.

Getting inside the chips

As chips get more complex and include more transistors, it is increasingly hard to cool them. For years now, fast hot chips have been cooled by specialist heatsinks, in turn often cooled by liquids.

But what if we could channel fluids closer to the transistors, inside the chips?

It turns out that the processes which etch and print tiny electronic components on a substrate can be adapted to make microscopic channels that carry coolant to the heart of the processor.

Ways to CoolerChips

The silicon community is finding other, less invasive but equally radical cooling techniques to enable chips with ever higher thermal design points (TDPs).

How serious is this? The US Government's COOLERCHIPS program is gathering multiple approaches to cool systems at this level, including schemes like JetCool, which fires jets of coolant within direct-to-chip cooling systems.

The DOE has also roped in Intel. The silicon giant has realized that its schemes to ship ever more powerful silicon will hit a wall without effective ways to cool those new processors.

Intel's combines computer generated heatsinks and new materials.

Everyone is looking for radical techniques to unlock continued growth.

Combining heat and compute

Peter Judge
Executive Editor

Every Watt of data center power should be used twice

Thermodynamics says that all energy will eventually become thermal energy. If you use 1kWh of electrical energy to run a computer, it will emit 1kWh of heat.

Astrid Wynne of TechBuyer explained it like this in a 2022 article: "The first law of thermodynamics, also known as the law of conservation of energy, states that energy can neither be created nor destroyed, but changes from one form to another. The electrical energy that data centers draw through the servers is almost entirely converted to heat. Data generation is a byproduct of this, but accounts for less than 0.1 percent of the conversion."

To the data center engineer, that's a problem, because that heat has to be removed or the equipment will overheat.

To the urban planner, that's an opportunity. The community wants that heat.

Use your energy twice

Wynne says: "The term 'data center' is a misnomer in energy terms. It is more accurate to call them 'heat centers' which can also be used to process and store information. Flipping thinking in this way allows us to be more creative about how

we plan and manage data center builds now and in the future."

Communities need heat, and heating currently creates a lot of emissions. In the UK, for instance, heating makes up around 37 percent of total emissions. Some of that comes from high-temperature industrial heat, but more than half is simple space heating and hot water.

Right now, most of that heat is generated directly from fossil fuels, or else from electricity, much of which is also from fossil fuels.

To approach net-zero emissions, heating has to decarbonize, which is why governments are trying to phase out gas boilers, and look for alternative low-carbon heat sources. Electric ground source heat pumps are one such source, but they need to be powered by renewable energy.

Heat produced by computation is effectively zero-carbon, because it's a byproduct of energy used elsewhere.

And if the heat is used elsewhere, it arguably cancels out the burden of the computation on the energy infrastructure. You get 1kWh of computing from 1kWh of energy that was going to become heat anyway.

There is, in theory, almost no limit to the demand for waste heat.

As Mark Bjornsgaard. CEO of Deep Green Energy observed in a *DCD podcast*, the total requirement for building heating in Europe is vastly greater than the amount of energy used in data centers.

For instance, the UK uses around 434TWh of energy to make buildings warm each year. The total energy used in data centers is a tiny fraction of that, estimated at around two or three TWh per year.

That's 2TWh to 3TWh of heat produced each year in data centers. If that could all be used in heating then, on one level, those data centers would arguably have achieved net-zero in their operational energy use. And with more than 400TWh of heat still required, there's scope to power the projected AI boom as well.

That's a tantalizing prospect, but not so easy in practice - or is it?

The difficulty of heat networks

Data centers have been offering their waste heat for years. But very little is actually being successfully used.

Data centers offer a predictable supply of heat because they keep their servers running continuously. But the heat is "low-grade:" It is warm rather than hot, and it comes in the form of air, which is difficult to transport.

So, most data centers vent their heat to the atmosphere.

Sometimes, there are district heat networks, which provide warmth to local homes and businesses through a piped network. If your data center is near one of these, it is a matter of extending it to connect to the data center, and boosting the grade of heat.

But you have to be in the right place to connect to one. "There are certain countries that have established or developing heat networks, but the majority don't have a heat network per se, so it's going on a piecemeal basis," Neal Kalita, senior director of power and energy at NTT, told *DCD*.

You are unlikely to find one in the US, says Rolf Brink of cooling consultancy Promersion: "The United States is a fundamentally different ecosystem. But Europe is a lot more dense in terms of population, and there is more heat demand."

The Nordic countries have a lot of heat networks. Stockholm Data Parks is a well-known example - a data center campus in urban Stockholm, where every data center has a connection to the district heating network and gets paid for its heat.

Germany has some heat networks, and plans to expand them. The UK has some, but experience has been difficult so far.

"I think the challenge with data centers plugging their heat into heat networks is when is that heat network going to be operational?" says Kalita.

How good is your heat?

Another big challenge is the quality of the heat: "It's not what we call high-quality heat or high-temperature heat," says Kalita. "You can't run a whole heat network from one data center's heat. You either need to install a heat pump and put more energy in to increase the temperature, or the heating network needs to have some kind of heat source."

The heating network must be very close by, and the heat in the warm air has to be supplemented with other energy sources, such as burning biomass or using a heat pump to raise the temperature until it can be piped into the system.

In Sweden, for instance, computing expert Cloud&Heat has placed a data center consisting of two liquid-cooled 20-foot containers equipped with 1,600 high-performance graphics cards right at a combined heat and power (CHP) installation.

The CHP system burns biomass to make heat and electricity. The Cloud&Heat systems use some of that electricity, and their waste heat can go straight into the heat stream, so none is wasted.

Low-grade heat is less of a problem with newer "fourth generation" district heat systems, which have become better suited to handling lower-grade heat, through measures such as polythene pipes, which can hold the heat of warm water.

Bring in agriculture

Of course, homes in much of the world don't need so much heat in the summer. The answer to that is planning, and finding other uses. Food production has a lot of processes that need heat, from greenhouses to fish farms, fermentation, and cheese production, Even in countries with warm summers, says Wynne, heat can still be used: "When the weather gets warmer, for example, the energy can be diverted from warming houses to fermentation sites."

Green Mountain has been giving warm water from its data center in Norway to a fish farm, The underground Lefdal Mine elsewhere in Norway does the same, and the snow-cooled White Data Center in Japan uses warmth to farm eels.

Valuing your heat

Another problem is that, because data centers see heat as a waste product,

they don't value it. Heat networks are a business, and they expect to buy heat on a commercial basis from a supplier committed to a predictable supply. They want a power purchase agreement (PPA) for heat.

"If you enter into a heat PPA, then the expectation is that you will provide that level of quality of heat and that volume of heat consistently," says Kalita. "But there are occasions when data centers go down, or a customer leaves, or the facility needs to be refurbished, and then the heat falls away.

"Where you're committed to provide heat for 10 to 20 years, that can curb your flexibility."

For some, that is not insurmountable. Heat reuse consultant David Gyulnazaryan told a London data center audience that waste heat is "a highly sustainable and reliable source of heat" and, if necessary, it can be "backed up by power generation in case of failure, which gives a 99.996 percent level of reliability."

On the financial side, Gyulnazaryan says "the cost of converting low-grade heat from the data center is approximately €25/MWh, whereas the natural gas price is around €200/MWh.

The monetary saving potential is huge for public bodies, particularly when the natural gas supply is reducing."

Brink says it's a matter of planning: "When effectively considered during the build phase of a data center facility, heat can be a tremendous asset with fantastic reductions to your TCO.

"When choosing a location to build a data center, heat reuse is not factored in most of the time," he says. "If you do factor it in, you could locate a data center next to a swimming pool, or in the vicinity of something that would have a use for the heat."

He says: "The discussion around heat reuse is a little bit out of balance simply because as long as it's not fed into the site scoping, it's going to be meaningless to discuss it later - simply because you're not selecting a site for for the purpose."

Regulations promote heat sharing

The biggest problem with district heating systems could be that they take a massive investment and many years to implement. Some local and national governments are intervening to encourage the use of heat networks.

In the Amsterdam region of the Netherlands, data centers must be designed to be able to connect to district heat networks. Germany has similar rules across the whole country enshrined in its newly passed Energy Efficiency Act.

This has not been a simple matter: Earlier drafts of the bill required new facilities built from 2025 to reuse 30 percent of their heat. The German Datacenter Association objected, and the final act says data centers over 200kW need only to reuse 20 percent of their heat - and don't have to do it till 2028.

The GDA still objects, as this requirement effectively limits data centers to locations where there are heating systems in place.

Benjamin Brake, head of the government's Digital and Data Policy department at the Federal Ministry of Digital Affairs and Transport (BMDV), conceded to the GDA's conference that: "The use of waste heat as the sole criteria for choosing a location is problematic, especially when there are no requirements and obligation for waste heat users."

In the UK, there are around 14,000 mostly small heat networks, according to the *National Housing Federation*. The Fuel Poverty Action Group is *critical* of the state of Britain's heat networks, arguing that domestic customers have few rights,

since the networks are unregulated. They are also vulnerable to outages and steep price rises as most networks include heat input from natural gas whose price has shot up.

The Government has a *pipeline* of new district heating projects amounting to nearly £3 billion ($3.8bn). Only one project in Park Royal, with a £36 million subsidy, intends to make use of data center waste heat. It won't be switched on till 2040 (see side box).

Given the huge cost of the project, and its long time scale, many observers have expressed doubts to *DCD* that it will ever be completed.

"London absolutely hates roads being dug up," says Kalita. "Imagine having to dig up the road for a new heat network across the distances that it will have to travel."

Others have a more fundamental issue, and wonder whether it should be done at all.

"I am not close enough to the project to know, but when I saw the cost I did a double take," says Charlie Beharrell of Heata. "The connection cost per home, excluding any IT or cost overruns, is more than the total cost of installing one of our systems."

The systems he is talking about are variously called digital boilers and data furnaces.

Beharrell, along with Bjornsgaard and a coterie of other companies, are championing a different approach. It could be more effective than heat networks - but it rethinks the data center.

The data furnace

The idea of a data furnace was first proposed by Microsoft in a *paper* in 2011, authored by a group including Sean James. The basic idea is to put racks of servers in homes and offices where their heat is useful.

The development of the cloud means that useful jobs can be sent in encrypted form to the data furnace. Customers pay a price for the computing that covers the electricity used, and the host gets free heat.

"In the coldest climate, about 110 motherboards could keep a home as toasty as a conventional furnace does," said the paper's authors. "The rest of the year, the servers would still run, but the heat generated would be vented to the outside, as harmless as a clothes dryer's."

Various companies have developed businesses around the idea, including Qarnot in France and Cloud&Heat in Germany.

The latest generation of distributed heat and computing companies in the UK include Deep Green and Heata.

Deep Green puts compute-intensive servers in liquid-cooled immersion tanks, so their heat energy can be piped to nearby users. Its first customers are swimming pools. Heata offers heat appliances for domestic homes.

Both use the same cloud provider, Civo, to despatch jobs to their units,

The idea has the potential to outperform district heating networks, says Beharrell, because "it is much cheaper and easier to move the data than the heat."

Data furnaces could be put in every dwelling in a housing development for instance, with only minimal infrastructure needs. The homes already have power.

Some might need network upgrades, but that is a much smaller job than laying pipes from a large data center building, via the heating system, to each house.

Beharrell doesn't think the two approaches are at odds, he is happy to see projects like the Park Royal heat network go ahead: "I think it's good to be optimistic, and honestly, it is good for everyone if it goes ahead."

But he'd like to see more awareness of the possibility of actually distributing computing, rather than moving heat: "It would be nice to see greater investment behind more disruptive models rather than reinforcing the status quo."

Deciding which approach is best for a given situation requires genuinely strategic thinking,

"Achieving the carbon reductions necessary by 2050 means a whole systems approach," says Wynn. "Integrating data centers into this would go a long way towards achieving success."

Photo by Luke Stackpoole

DISTRICT HEATING IN LONDON

This year, the UK Government awarded £36 million ($44.5m) to a district heating system in West London, to be built by Aecom, which will share data center waste heat with up to 10,000 new homes.

Two data centers under construction have signed up to provide 98.7GWh of heat to the system, which is not expected to be operational until 2040.

The Department of Energy's Green Heat Network Fund (GHNF) will pay the money to the Old Oak and Park Royal Development Corporation (OPDC) to create the heating system to harness waste heat at between 20°C (68°F) and 35°C (95°F).

The location is significant, because this is precisely where a shortage of electricity distribution effectively "halted" home building projects in 2022, because data centers had reserved all the available power from local substations.

The project will also get around £400,000 ($495k) of technical expertise funding from the Mayor of London's Local Energy Accelerator (LEA) program, with David Lunts, OPDC Chief Executive, describing it as "an exciting and innovative example of OPDC's support for the Mayor's net zero ambitions."

When less isn't more

Data center operators will need to adopt hybrid cooling strategies to handle growing IT densities. What does that mean?

'Less is more' is one of those clichés that rarely stands up to real scrutiny, especially when it comes to data centers. Clients always want more power, more space, more of everything – year after year – and data center operators must meet these demands, without missing a step in their march towards Net Zero.

These demands are especially challenging when it comes to conventional server hall cooling, where more powerful CPUs and GPUs have been crammed into racks, generating more heat, and posing a conundrum for designers to maximize the air flow and thermal landscape of the data center.

Indeed, with demand for AI and machine learning applications only expected to intensify, racks with multiple GPUs on board will become more common in order to handle the increased workload – but can the data center take the heat all this powered-up silicon will generate?

Struggling for air

"All these new demands will add immense complexity to the thermal landscape," says Nigel Gore, senior director of liquid cooling at Vertiv.

While air cooling has evolved to handle the ever-greater challenges posed by each passing generation of technology, it faces challenges adapting to the next phase of compute.

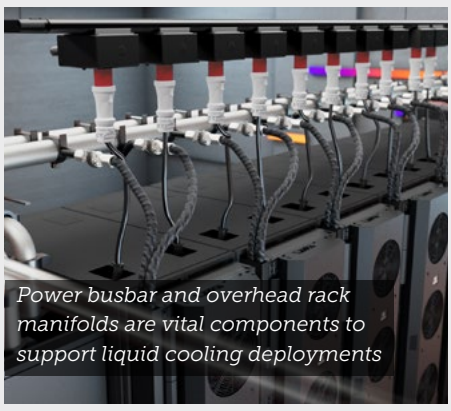
The industry solved High Density air cooling by moving it closer to the heat load in the racks in the form of row cooling and rack-based cooling. The next step involved managing the recirculation of exhaust air going into the hot aisle at elevated temperatures to avoid raising the ambient temperature of the entire server hall.

However, while air cooling technology and topologies were ingeniously adapted over the years as IT demands grew, the challenges are now so great that an inflection point has been reached where liquid technologies will need to be deployed to handle the increase in power at the processor level.

Data center operators will need to make a decision: how are they going to incorporate liquid cooling into their server halls in order to support high-performance IT and, therefore, attract and retain the well-heeled clients that need it?

It's critical that installing direct-to-chip cooling technology from companies like Vertiv in server halls is straightforward. The server racks have the same footprint as conventional racks and can be run alongside conventionally cooled IT in the same hall.

Moreover, the perceived biggest 'risk' posed by liquid cooling – that of downtime as a result of a leak – has been significantly reduced by a rack design with integrated liquid manifolds that incorporate high-quality plumbing. This plumbing includes multiple couplings, valves and fail safes, so that engineers can quickly and safely conduct maintenance or upgrades on racks or even individual rack-mounted servers.



Power busbar and overhead rack manifolds are vital components to support liquid cooling deployments

Indeed, if your abiding impression of liquid cooling has been drawn exclusively from the world of PCs, where enthusiasts compete to construct the most powerful, overclocked rig their money can buy (either by using a water-cooler, or even liquid nitrogen in some extreme cases) then the reality of a modern liquid-cooled server rack will blow you away.

Power and water

In a typical rack installation, suggests Gore, the first thing is to make sure that the power density is supported. "A lot of conventional data centers have been designed around 5kW per rack, but now we're looking at 25kW, 50kW, and we've even seen 70kW per rack," says Gore. Indeed, he adds, it's possible that power densities may soon reach 100kW or 150kW per rack as transistor population is expected to increase exponentially in future silicon.

Below the power busbar, the plumbing for the coolant is typically connected, in a similar style, to each rack. For this, Vertiv uses high-quality stainless-steel piping to ensure that galvanic corrosion doesn't contaminate the coolant and affect heat transference and flow. The company is also examining developments in high-heat plastics as a potential alternative.

This network of liquid-cooling pipes, technically referred to as a secondary fluid network (SFN), are designed with shut-off valves for maintenance, with the couplings from the racks back to the main pipework featuring in-line valves to further isolate the flow, if necessary. "These are manual couplings that you can disconnect, a design with a ball valve that automatically prevents any drips as you disconnect" says Gore.

Alternatively, there are electronic valves that facilitate orchestration on control of coolant flow in combination with the coolant distribution unit, he adds.

All this pipework and hardware connects to the racks on a supply (cold coolant) and return (warm coolant) basis,

and row manifolds enable individual racks to be isolated, if necessary. Hence, work can be carried out on one rack – to install upgrades, for example – while the others hum away.

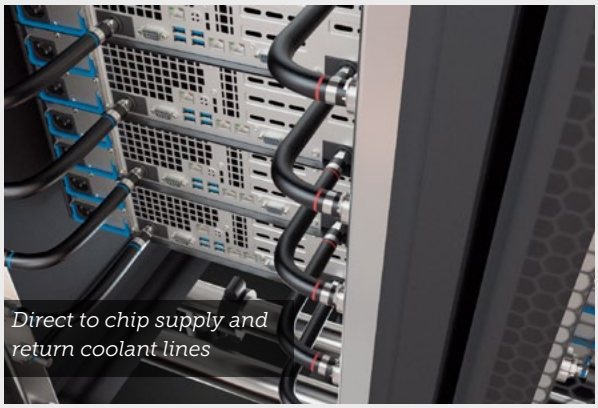
For raised floor data center halls, this pipe network could also be installed under the floor. "Personally, though, I feel it's easier to operate when the pipework is above, if for example, you need to shut it down and conduct maintenance or install upgrades in the rack," says Gore.

Regardless of where you place your pipework, it's crucial that the right pressure and temperatures are maintained. To do this, coolant distribution units – located either in the row, in-rack or perimeter – use intelligent flow monitoring to control the flow of liquid, as well as providing filtering, easily accessible water-fill ports, and drain locations.

"We've also seen deployments where facility teams want to automatically control the liquid to each rack so you can have group control and can shut down flow to individual racks, and not just shut down the flow, but control the volume of flow, as well," he adds. In other words, the cooling to the racks can be fine-tuned in ways that simply aren't possible with air cooling.

At the rear of the racks, each rackmount server has couplings for the liquid coolant – a supply and return water line, handily color-coded – while the servers have a standardized layout for the CPUs or GPUs, securely fixing the coolant plates directly over the chips.

Within the server, copper contacts provide the best heat-conductive material for transferring the heat from the CPU/GPU, to liquid and while the water is returned to the coolant distribution unit and heat exchanger, server fans expel the residual heat-to-air through the rear of the rack into the hot aisle of the server hall, away from the IT.



Direct to chip supply and return coolant lines

Fully immersed in cooling

Of course, chip-level liquid cooling is not the only liquid choice. Immersion cooling remains an option but will require a learning curve for many operators to make, with servers immersed in horizontal – rather than vertical – tanks in a non-conductive liquid covering all the components on the rack, not just the CPUs and/or GPUs. We have seen advances in specifications of dielectric liquid from the Open Compute Project contributed by Intel.

"It means putting the entire server into a dielectric liquid with a specific chemistry. Because every single component is going to be immersed – every transistor, capacitor, memory module, cable, connector, and so on – you need to be sure that there is no risk of contamination to make sure that your servers are compatible with the system," says Gore.

For example, PVC-jacketed cabling and plastics could dissolve into the dielectric liquids or oils and will need to be swapped out for plenum-rated cables or plastics. Reference material is available from vendors or Open Compute Project providing immersion requirements and material compatibility guides.

There will also be radiated heat losses from the tank, which will need to be supported by the facility. Operational guidance will also call out risk of contamination each and every time an engineer opens up the tank to conduct maintenance on a server, or to install upgrades.

Moreover, Gore adds, when comparing liquids, none of them can approach the efficacy of water when it comes to the physics of heat transference.

"When you examine the properties of a dielectric fluid compared to water by weight or volume, there is a significant difference in terms of their heat-carrying capacities and properties. With a dielectric fluid, any change in viscosity significantly affects the ability to transport heat. As with air cooling, there's also the tricky business of 'bypass flow,' where obstacles inside the server prevent optimal flow across the hotter components, so cooling can be less effective [than direct-to-chip].

"So there is a need to design around channelling the liquid to make sure that it remains an effective coolant," says Gore. Optimally capturing the generated heat as the liquid is pumped around the tank can therefore be a challenge in densely populated servers and tanks. This is well understood, says Gore, and development teams are advancing optimized heat sinks and internal flow distribution control.

But perhaps where direct to chip cooling scores highest over immersion cooling is simply the utilization of water, rather than a non-conductive mineral oil. "The ultimate liquid before shifting to fluids for cooling is essentially water because it has the highest heat-carrying capacity of liquids. With liquid cooling, what we're doing is getting the most heat-conductive liquid directly across the hottest components in the rack as quickly as possible. Fluid cooling is in our future as an emerging technology using phase change to accelerate heat harvesting," says Gore.

With the increase of racks supporting liquid cooling, the addition of liquid cooling to the data center is already underway. Vertiv is developing guidance and support to enable facilities to add liquid-cooled racks, which can co-exist in server halls with air-cooled racks, expanding as demand for liquid-cooled server hardware increases.

As the next generation of chip technologies arrive in the data center, the issue of cooling will only intensify. To put the heat-generating characteristics of imminently arriving CPUs and GPUs into perspective, 750W is sufficient to power one mini oil-filled heater that could keep a modest room warm and toasty throughout winter.

With eight 750W GPUs per rack-mounted server, and anywhere between 14 and 42 servers to a cluster, that's almost 340 heaters per cluster. Multiplied across a typical data center, those without an effective cooling solution risk meltdown.

>>To find out more about Vertiv's liquid cooling technology, please check out *Vertiv's dedicated cooling options website*

VERTIV.

Microfluidics: Cooling inside the chip

Peter Judge
Executive Editor

If you think immersion tanks are the end game for liquid cooling, think again. DCD hears from the engineers who want coolant to flow inside your chips

We all know that liquid cooling is the future for data centers. Air simply can't handle the power densities that are arriving in data halls, so dense fluids with a high heat capacity are flowing in to take over.

As the heat density of IT equipment increases, liquids have inched ever closer to it. But how close can the liquids get?

Running a water-circulating system through the rear doors of data center cabinets has become well-accepted. Next, systems have been circulating water to cold plates on particularly hot components, such as GPUs or CPUs.

Beyond that, immersion systems have sunk whole racks into tanks of dielectric fluid, so the cooling liquid can contact every part of the system. Major vendors now offer servers optimized for immersion.

But there is a further step. What if the fluid could be brought closer to the source of that heat - the transistors within the silicon chips themselves? What if coolants flowed inside processors?

Husam Alissa, a director of systems technology at Microsoft, sees this as an exciting future option: "In microfluidics, sometimes referred to as embedded cooling, 3D heterogenous, or integrated cooling, we bring the cooling to the inside of the silicon, super close to the active cores that are running the job."

This is more than just a better cooling system, he says: "When you get into microfluidics, you're not only solving a thermal problem anymore." Chips with their own cooling system could solve the problem at the source, in the hardware itself.

Birth of microfluidics

In 1981, researchers David Tuckerman and R F Pease of Stanford suggested that heat could be removed more effectively with tiny "microchannels" etched into a heatsink using similar techniques to those used in silicon foundries.

The small channels have a greater surface area and remove heat more effectively.

The heatsink could be made an integral part of VLSI chips, they suggested, and their demonstration proved a microchannel heatsink could support a then-impressive heat flux of 800W per sqm.

From then on, the idea has persisted in universities but only tangentially affected real-life silicon in data centers.

In 2002, Stanford professors Ken Goodson, Tom Kenny, and Juan Santiago set up Cooligy, a startup with an impressive design of "active microchannels" in a heatsink built directly onto the chip, along with a clever silent solid-state electrokinetic pump to circulate the water.

Cooligy's ideas have been absorbed by parts of the mainstream. The company was bought by Emerson Network Power in 2005. Its technology, and some of its staff, still circulate in Emerson's new incarnation, Vertiv.

The idea of integrating cooling and processing became more practical as silicon fabrication developed and went into three dimensions. Starting in the 1980s, manufacturers experimented with building multiple components on top of each other on a silicon die.

Making channels in the upper stories of a multi-layer silicon chip is potentially a quick win for cooling, as it can start simply by implementing tiny grooves similar to the fins seen on heatsinks.

But the idea didn't get much traction, as silicon vendors wanted to use 3D techniques to stack active components. That approach is now accepted for high-density memory, and patents suggest that Nvidia may be intending to stack GPUs.

In the microprocessor industry, cooling and processing were seen as separate disciplines. Chips had to be designed to dissipate their heat, but this was done by

relatively unsophisticated means, using thermal materials to siphon the heat to the big copper heatsink on the surface.

The heatsink could be improved by etching smaller channels, but it was a separate item, and heat had to cross a barrier of adhesive to get there.

But some researchers could see the possibilities. In 2020, Tiwei Wei, of the Interuniversity Microelectronics Centre and KU Leuven in Belgium, integrated cooling and electronics in a single chip.

Wei, whose work was published in Nature in 2020, did not think the idea would catch on in microprocessors, saying that micro cooling channels would be more useful in power electronics, where large-sized chips made from semiconductors like gallium nitride (GaN) actually manage and convert electricity within the circuits.

That possibly explains why Emerson/Vertiv wanted to get hold of Cooligy, but Wei didn't see the tech going further: "This type of embedded cooling solution is not meant for modern processors and chips like the CPU," he told *IEEE Spectrum*.

Digging into the chips

Already, by that time, researchers had been working on etching microfluidic channels into the surface of silicon chips for some years. A team at Georgia working with Intel in 2015 may have been the first to make FPGA chips with an integrated microfluidic cooling layer, on top of the silicon, "a few hundred microns [micrometers] away from where the transistors are operating."

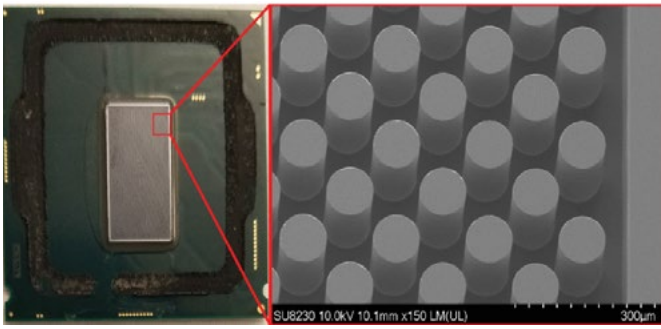
"We have eliminated the heat sink atop the silicon die by moving liquid cooling just a few hundred microns away from the transistor," team leader Georgia Tech Professor Muhannad Bakir said in Georgia Tech's press release. "We believe that reliably integrating microfluidic cooling directly on the silicon will be a disruptive technology for a new generation of electronics."

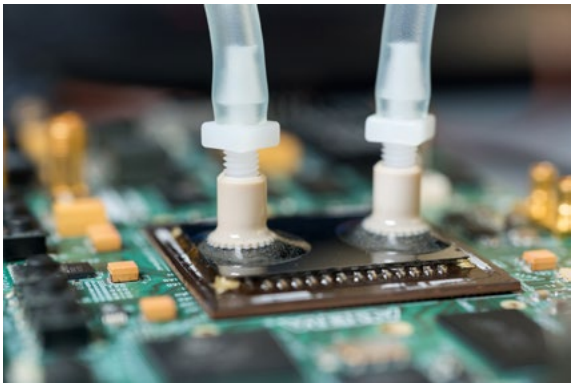
In 2020, researchers at the École Polytechnique Fédérale de Lausanne in Switzerland, took this further, actually running fluid in tunnels underneath the heat-generating transistors.

Professor Elison Matioli saw the opportunity to bring things even closer together: "We design the electronics and the cooling together from the beginning," he said in 2020, when his team's paper was published in *Nature*.

Matioli's team had managed to engineer a 3D network of microfluidic cooling channels within the chip itself, right under the active part of each transistor device, just a few micrometers away from where the heat is produced. This approach could improve cooling performance by a factor of 50, he said.

Matioli etched micrometer-wide slits in a gallium nitride layer on a silicon substrate, and then widened the slits in the silicon substrate to form channels that would be big enough to pump a liquid coolant through.





After that, the tiny openings in the gallium nitride layer were sealed with copper, and a regular silicon device was created on top. "We only have microchannels on the tiny region of wafer that's in contact with each transistor," he said at the time. "That makes the technique efficient."

Matioli managed to make power-hungry devices like a 12kV AC-to-DC rectifier circuit which needed no external heatsink. The microchannels took fluid right to the hotspots and handled incredible power densities of 1.7kW per sq cm. That is 17MW per sqm, multiple times the heat flux in today's GPUs.

On to standard silicon

Meanwhile, work continues to add microfluidics into standard silicon, by creating microfluidics structures on the back of existing microprocessors.

In 2021, a Microsoft-led team, including Husam Alissa, used "micropin" fins etched directly on the backside of a standard off-the-shelf Intel Core i7-8700K CPU.

"We actually took an off-the-shelf desktop-class processor, and removed the case," he says. Without the heat spreader cover and the thermal interface material (TIM), the silicon die of the chip was exposed.

"When that die was exposed, we applied etching methods to carve out the channels that we want to see," he continues. The back of the die was etched away selectively, to a depth of 200 microns, leaving a stubble-field pattern of rods 100 microns thick - the "micropins" that form the basis of the integral direct-to-chip cooling system.

That's a delicate task, warns Alissa: "You have to consider how deep you are etching, so you are not impacting the active areas of the silicon."

Finally, the back of the CPU die was sealed in a 3D-printed manifold, which delivered coolant to flow amongst the micropins. The chip was then overclocked to dissipate 215W of power - more than double its thermal design power (TDP), the energy it is designed to handle safely without overheating.

Surprisingly, the chip was able to perform at this level using only room-temperature water. Delivered through the manifold. The experiment showed a 44 percent reduction in junction-to-inlet thermal resistance and used one-thirtieth the volume of coolant per Watt than would have been needed by a conventional cold plate. The performance was evaluated with standard benchmark programs.

This was the first time microfluidics channels were created directly on a standard consumer CPU and achieved the highest power density with microfluidic cooling on an active CMOS device. The results show the potential to run data centers more efficiently without the need for energy-intensive refrigeration systems, the group reported in *IEEE Xplore*.

All that would be needed would be for the chip maker to mass-produce processors with etched micropins, and sell them packaged with a manifold attached in place of the usual heatspreader cap.

If foundries like TSMC could provide their chips with built-in liquid cooling, that would change the dynamics of adoption. It would also allow the technology to push

boundaries further, says Alissa.

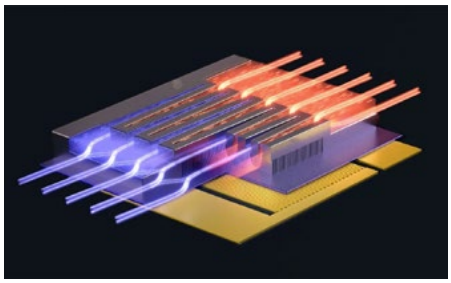
"With cold plates, you might get water at 40°C (104°F) but with microfluidics you could probably have 80°C (176°F) and higher coming out of these chips, because the coolant is so close to the active cores," he says. "This obviously enhances the efficiency and heat recovery benefits, paired with lower requirements for flow rate."

The future of microfluidics

"There are two main flavors of microfluidics," says Alissa. The lighter touch option, which he says could be deployed "in a couple of years," is the approach his team showed - to etch channels in commercial chips: "Go buy chips, do the etching, and you're done."

A more fully developed version of this approach would be for the foundries to do the etching before the chip reaches the consumer - because not everyone wants to lever the back off a processor and attack it with acid.

Beyond that, there is what Alissa calls the "heavier touch" approach. In this, you "intercept early at the foundry and start building 3D structures." By this, he means



porous chips which stack components on top of each other with coolant channels in the layers between.

That's a development based on the approach used by Matioli in Lausanne. As Alissa says, "That promises more but, obviously, it's more work."

Alissa has a goal: "The North Star we want to get to is where we're able to jointly optimize this chip for cooling and electrically at the same time, by stacking multiple dies on top of each other, with [microchannel] etching in between."

Cooling would allow multiple components to be stacked connected "through chip vias" (TCVs) which are copper connections that travel through the silicon die. These tower chips could need lower energy and work much faster, as the components are closer together: "Overall, you're gaining on performance, you're getting on cooling, and also on latency because of the proximity."

There's another benefit. If microfluidics allows chips to go to a higher thermal design point (TDP) this could remove one of the hurdles currently facing silicon designers.

The difficulty of removing heat means that today's largest chips cannot use all their transistors at once, or they will overheat. Chips have areas of "dark silicon" (see box), and applying microfluidics could allows designers to light those up, boosting chip performance.

But don't expect microfluidics to solve everything. Back in 2012, Professor Nikos Hardavellas predicted the next problem: "Even if exotic cooling technologies were employed, such as liquid cooling coupled with microfluidics, power delivery to the chip would likely impose a new constraint."

Once we work out how to get more heat off the chip, we will have to develop ways to deliver a large amount of power, that can provide signal integrity at the low voltages required by the transistors.

Are we ready for that one?

DARK SILICON

Current and future generations of chips have a fundamental problem. Performance has always increased, as more transistors are packed into a single processor. But now, there are so many that they cannot all be used at once, without the chip overheating.

Processor makers publish a thermal design power (TDP) for each chip, which is the amount of energy it can handle and dissipate safely - and will assume there is a good heatsink on the chip.

TDPs have grown very high. For example, the H100 SXM5 Nvidia GPU has a 700W TDP, which is massive compared with standard CPUs like Intel Xeons, which consume around 130W.

But how much can you do with this power? Currently, transistors fabricated at 4nm consume a tiny 10 attoJoules (10-18 Joules) each, so if one of them switched at 1.8GHz, it would consume 18 microWatts (18 x 10-9 W).

That is tiny, but today's processors have colossal numbers of transistors. Jon Summers of the Swedish research institute RISE calculates that the Nvidia H100 GPU, which has 80 billion transistors, would generate 1,440W - more than double the TDP Nvidia publishes for it.

"With a TDP of 700W, it must mean that 51 percent of the chip is dark silicon," Summers told an audience at DCD Connect London in November 2024.

Continued miniaturization won't fix the situation. Smaller transistors have a lower switching energy, so more can be lit up within the TDP envelope, but the number of transistors is also going up.

Summers says that Intel plans to have a trillion transistors on a chip by 2030, each using around 1aJ per switch. If the clock speed has gone up to 4GHz, the chip is 1,000 sq mm, and thermal flux, then that means 40 percent of the transistors must remain dark.

Now, TDPs are based on a maximum heat flow (or flux) that can be removed from a chip. The Nvidia H100 has an area of 814 sq mm, so the heat flux is 860kW per sqm. That is comparable to the levels that are seen in nuclear fusion demonstrations, and Summers expects Intel to push on to 2.4MW per sqm.

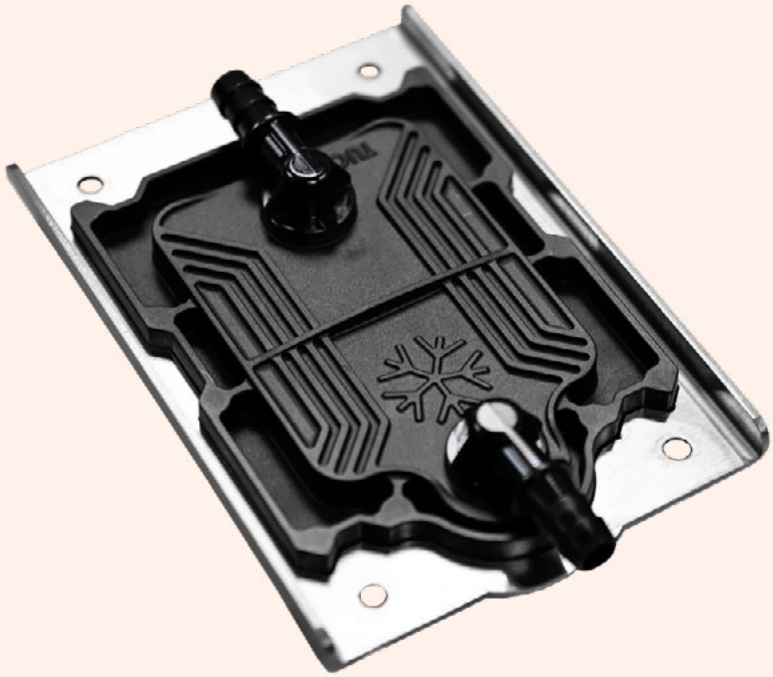
The issue of dark silicon has been known about for a long time: In 2012, Professor Nikos Hardavellas of Northwestern University, said in the magazine of the Advanced Computing Association, Usenix: "Short of a technological miracle, we head toward an era of "dark silicon," able to build dense devices we cannot afford to power. Without the ability to use more transistors or run them faster, performance improvements are likely to stagnate unless we change course."

There have been plenty of approaches to the problem, most notably, increasing the use of specialized cores within chips, that are only used when needed.

But maybe a way to reduce dark silicon would be if fluids could flow inside the chip itself, where they can remove more energy, and allow more heat flux.

Precision cooling at the chip level

Sebastian Moss
Editor-in-Chief



JetCool's CEO on how thousands of tiny jets could be coming to a data center near you

As chip temperatures and rack densities rise, a plethora of companies have come forward to pitch their vision of the future.

Cooling demands of artificial intelligence and other high-density workloads have outstripped the capabilities of air systems, requiring some form of liquid cooling.

"When you think about the landscape of liquid cooling, we see three different technical categories," JetCool CEO Bernie Malouin explained.

"There's single phase immersion, dipping in the oil. And that's interesting, but there are some limitations on chip power - for a long time, they've been stuck at 400W. There are some that are trying to get that a little bit better, but not as much as is needed."

The second category is two-phase dielectrics: "We see those handling the higher [thermal design point (TDP)] processors, so those can get to 900-1,000W. Those are fit technologically for the future of compute, but they're held back by the chemicals."

Many two-phase solutions use perfluoroalkyl substances (PFAS), otherwise known as forever chemicals, which are linked to human health risks, and face restrictions in the US and Europe. Companies like ZutaCore have pledged to shift to other solutions by 2026, but the shift has proved slow.

"It's a concern for a lot of our customers, they're coming to us instead because they're worried about the safety of those fluids," Malouin said. "They're concerned about the continued

availability of those fluids.

And then there's the third category: Direct Liquid Cooling (DLC) cold plates. "We're one of them," Malouin said. "There are others."

DLC cold plates are one of the oldest forms of IT liquid cooling - simply shuttling cold liquid to metal plates mounted directly on the hottest components. They have long been used by the high-performance computing community, but JetCool believes that the concept is due for a refresh.

Instead of passing fluid over a surface, its cooling jets route fluid directly at the surface of a chip. "We have these arrays of a thousand tiny fluid jets, and we work directly with the major chipmakers - the Intels, AMDs, Nvidias - and we intelligently landscape these jets to align to where the heat sources are on a given processor."

Rather than treating the entire chip as a whole with a singular cooling requirement, the microconvective cooling approach "tries to balance the disparate heat loads, disparate thermal requirements of certain parts of that chip stack," Malouin said.

"When you start thinking about really integrated packages, the cores themselves might be able to run a little higher temperature, but then you might have high bandwidth memory (HBM) sections that aren't as power hungry, but have a lower temperature limit."

Instead of trying to design for the high-power cores and the temperature-sensitive HBM, each section can be cooled at a slightly different rate. "This allows you to decouple those things and allows you to have precision cooling where you need," Malouin said.

While Malouin believes that facility-level liquid cooling is the future of data centers, the company also has a self-contained system for those looking to dip their toe in cooler waters., with a Dell partnership focused on dual socket deployments.

Two small pumping modules provide the fluid circulation and an air heat exchanger ejects heat at the other end of the Smart Plate system.

"When we add these pumps, you add some electrical draw, but you don't need the fans to be running nearly as hard, so it makes it 15-20 decibels quieter - and in net, we pull out about 100W per server after we've taken the penalty off of the pumps," Malouin claimed.

When you go to 10 racks or more, going to the facility level makes more sense, he said. Asked about the preferred inlet temperature, Malouin said the system was flexible but added, "we actually really like the warm fluids."

He said: "We have facilities today that are feeding us inlet cooling temperatures that are 60°C (140°F) and over. And we're still cooling those devices under full load." That's not common just yet, but Malouin believes that warmer waters will grow in popularity in places like Europe due to the heat reuse potential.

Back in the US, the company is part of the Department of Energy's COOLERCHIPS project, aimed at dramatically advancing data center cooling systems.

The focus of JetCool's $1m+ award is not just on the cooling potential, but a tantalizing secondary benefit: "We have instances where we've made the silicon intrinsically between eight and 10 percent more electrically efficient," Malouin claimed.

"That has nothing to do with the cooling system power usage, but with leakage."

Malouin doesn't mean leakage of the cooling system, but rather the quantum phenomenon of semiconductor leakage currents that can significantly impact a chip's performance.

The recent history of data center cooling has tended to assume that allowing temperatures to rise higher will save energy because less is used in cooling. Results, including research by Jon Summers at the Swedish research institute RISE, is finding that leakage currents in the silicon limit the benefits of running hotter.

"A big part of our COOLERCHIPS endeavor is to substantiate that through more rigorous scientific evidence and extrapolate it to different environments to see where it holds or where it doesn't go."

Looking even further ahead, Malouin sees an opportunity to get deeper into the silicon. "In some cases, it might actually be integrated as an embedded layer within the silicon, and then coupling that to a system that's outside that's doing some heat reuse. When we think about that holistically, we think that there's a real opportunity for a step change in data center efficiency."

For now, the company says that it is able to support the 900W loads of the biggest Nvidia GPUs and is currently cooling undisclosed 'bespoke' chips that use 1,500W.

"Ultimately, you're really going to have to look at liquid cooling if you want to run not just the future of generative AI, but if you want to run the *now* of generative AI."

Why Intel is developing data center cooling tech

Sebastian Moss
Editor-in-Chief

2kW processors and the future of data centers

It used to be so easy.

Intel is not a cooling company, and it doesn't want to become one. But the company is being forced in that direction by the rapid rise of its processors' thermal design points.

The chip company has become increasingly concerned that tomorrow's chips will overload tomorrow's data centers, and is researching how it can help.

"We want to be ready: As processor power keeps on increasing, the cooling challenges grow exponentially," Intel supercompute platforms group principal engineer Tejas Shah explained. In response, Intel is trying to upend data center cooling

Increasing power densities have pushed developers to use immersion cooling, and some are looking at two-phase immersion where boiling fluids remove even more heat. Shah is leading an effort to these systems with ultra-low-thermal resistance, coral-shaped heat sinks integrated with a 3D vapor chamber cavity.

The project was one of 15 chosen by COOLERCHIPS, a program to dramatically change how data centers are cooled, run by the US ARPA-E (Advanced Research Projects Agency-Energy).

But before we get into the tech, let's address the elephant in the room: Two-phase immersion fluids have proven controversial: "We are not developing the fluid, we will be using industry-standard fluids," said Shah.

"We are very aware of the issue with all the PFAS and the regulations surrounding it," he admitted, referencing the toxic 'forever chemicals' used by most two-phase solutions. "3M are basically exiting the business of two-phase immersion; we are working with some of the fluid vendors."

The other issue is the global warming potential (GWP) of many of the fluids, which might outweighs any energy efficiency savings that immersion may bring. "Getting the PFAS out is the biggest challenge, but we are pretty confident we can meet the GWP requirements," Shah said.

Now back to the tech. The company is building coral-shaped heatsinks using additive manufacturing, aka 3D printing.

"It's basically optimizing the design to follow a specific mathematical function," Shah said. "So it's topology optimization - it's an iterative mathematical procedure to optimize for a certain cost function, whatever that is. In this case, it's reducing the thermal resistance and making sure that we hit the right hotspots."

The heatsinks are coupled with 3D vapor chamber cavities: sealed, flat metal pockets filled with fluid, to spread the boiling capacity of the two-phase system.

"The coral shape is increasing the surface area, but we also want to reduce the spreading resistance, which is why you have an embedded vapor chamber inside the 3D structure," Shah said.

When current flows through a semiconductor material, it encounters "spreading resistance" where the electrons spread across a wider area than initially anticipated. Lowering a chip's temperature helps reduce the spreading, and therefore the resistance.

The new approach also uses an improved coating which enhances boiling by promoting nucleation. "So pretty much it helps with the boiling happening at the surface," Shah said.

With the three approaches combined, the company can cool processors all the way up to a thermal design point of 2kW. This is beyond what is on the market today, or likely to come out over the next few years, although Nvidia's Grace Hopper Superchip has a TDP of 450W to 1kW by combining a CPU and GPU.

The immersion system can support "at least eight processors" consuming 2kW each, Shah said. Increasing the number in a rack further becomes more of a power limitation than a thermal one, he said.

While tests have proved promising, overcoming every issue is not a foregone conclusion - especially if the immersion sector can't get past the PFAS problem (vendor ZutaCore says it will eliminate PFAS in 2026, but has yet to provide specifics, while rival LiquidCool is diversifying into single-phase immersion).

"It's not a slam dunk, it is very challenging," Shah said of the project and Intel's other efforts. "But the whole goal is to reduce the thermal resistance and target the next generation of processors with the best thermal solution out there."

If it does work, however, Intel could find itself in the interesting position of being a cooling vendor. "What we might end up doing - it's a big might - but we might even sell the thermal solution with the processor initially," Shah said.

"That will build the ecosystem, and then the next generation might catch up to our processors."

VERTIV.

AI is hot.
Be sure your data center is cool.

Adopt hybrid cooling strategies today to handle growing IT densities.

See how we can help you at **Vertiv.com/GetCool**