

Responsive CDUs for AI Factory Energy Efficiency

For every watt of power in, at least a watt of heat needs to come out.

Power densities of data centers are increasing at an alarming rate as AI is transforming business globally. AI server rack power has risen tenfold from 10kW to 100kW+ in a few short years, leading data center operators to be power constrained on the amount of performance they can deliver for their users.

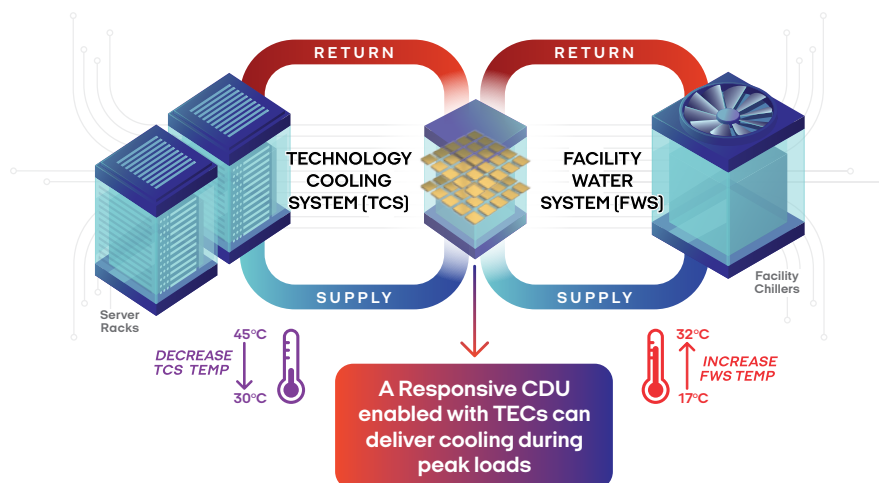
Most AI data centers have made the transition from traditional hot aisle / cold aisle air cooling, where high speed fans draw air through the servers, to Direct Liquid Cooling [DLC] with cold plates attached directly to CPUs and GPUs to draw the heat away. The cold plates are connected to Coolant Distribution Units [CDUs] via hoses and manifolds that pump the liquid to/

from the servers. The CDUs contain a heat exchanger at the intersection of the Technology Cooling System [TCS] loop that cools the servers and the Facility Water System [FWS] loop that evacuates heat from the building.

These two loops the interface inside the CDU have conflicting demands as higher power servers drive a decrease in TCS temperatures to keep them cool and data center facility operators want to increase FWS temperatures to save energy and avoid the costs of operating mechanical chillers like vapor compressors. Every 1°C increase in FWS temp results in a 2-3% decrease in energy consumption. Responsive CDUs can help solve this problem by utilizing passive cooling during periods of lower demand and activating thermoelectric cooling assistance during period of peak demand. By leveraging the responsive control of thermoelectric cooling, the TCS temperature can be held steady while the FWS temperature rises.

RESULTS:

A Responsive CDU enabled with TEMs can deliver ~25% decrease in cooling energy at the facility level.



PHONONIC RESPONSIVE COOLING SYSTEM™

Take Your Compute Performance to the Next Level with Phononic.

Thermal management has always been a constraint to achieving higher memory and compute performance, but it's increasingly becoming a bottleneck to the step-change progress required to support the future of AI. AI performance is delivered by faster GPUs and GPU performance is constrained by the power and cooling capabilities of the system. At the heart of the issue? The ability to respond immediately and precisely to changes in workload demand, at a system-level, not merely component to component. This stands in stark contrast to current approaches that over-provision cooling, resulting in energy 'wasted' on cooling that could have been deployed to compute.

Phononic's Responsive Cooling System™ approach to cooling throughout the data center integrates thermoelectric chips (TECs), existing cooling platforms, and proprietary control electronics to deliver distributed cooling that optimizes to temperature management needs. This approach decouples temperature where needed to eliminate hot spots, lowers total power requirements and unleashes meaningful TDP improvement quickly and reliably.

Scale your compute performance, with Phononic.

Multiply Your AI Performance with Phononic. Increasing future AI performance requires a fundamental change in cooling approach. Phononic's Responsive Cooling™ transforms cooling into a performance enhancing, intelligent platform.

Design Capabilities that Push the Cooling Boundaries to Enable AI

Phononic has a deep understanding of data centers and the components that are powering the future. With more than 30M+ devices in field today, deployed across all major US hyperscalers, Phononic's engineering team has been setting the standard for performance, efficiency and cost-effectiveness in our TECs.

Our IP library is robust and growing, with hundreds of patents covering materials, software integrations, thermal management approaches and more.

A wealth of reference design kits, along with the backing of ISO Quality Management Systems, IATF and Telcordia certifications ensures that our designs are consistently real-world predictive, reliable and deployable.

Learn more:
<https://phononic.com/datacenter-cooling>

Industry-Leading Design, Delivered Consistently and at Scale

At Phononic, as part of our design approach, we work closely with our customers to make system level tradeoffs that optimize not just the custom TEC we deliver, but the entire product for the end customer.

By leveraging our proprietary TEC technology and scalable device architecture, we are uniquely positioned to reliably deliver high performance cooling that meets the most rigorous demands, regardless the sector.

Together with our partners in Thailand, Phononic has the ability to scale availability of devices, and fully integrated solutions to our partners and licensees in a manner that both leverages the deep R&D and engineering expertise in HQ in RTP, NC while enabling full global scale and supply chain flexibility.