HOLISTIC RACK-TO-PROCESSOR POWER AND THERMAL CO-DESIGN FOR FUTURE SERVERS

Nenad Miljkovic (PI)
University of Illinois at Urbana-Champaign

Co-PI(s): William P. King (UIUC), Robert Pilawa (UCB), Sheng Shen (CMU), Kashif Nawaz (ORNL), David Huitink (UARK), Simon Chen (Wiwynn), Jiu Xu (Meta)

COOLERCHIPS™

Project Vision

Energy efficient and **reliable single-phase water-cooling** for next-generation data centers capable of reducing cooling energy consumption to **less than 5% of IT load** at **any location and any time**.

Total Project Cost:	$2.5M
Length	36 mo.

COOLERCHIPS Annual Meeting
December 11th, 2024

Project Overview

Architecting robust power-dense data centers anywhere anytime

Technology Summary

- Thermal management of 1 kW high performance multi-chip-module (MCM) based servers leveraging CTE matched materials, advanced TIMs, and advanced thermal integration
- Reliable and robust single-phase water thermal management readily integrated with existing and future data centers
- Primary side heat capture temperatures > 60ºC enabling efficient heat dissipation to the outside air independent of data center location, as well as potential re-use

Technology Impact

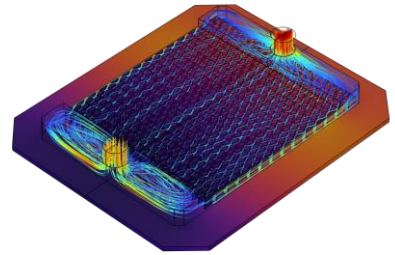
- New class of power-dense computational systems, data centers, and modular EDGE systems that are cooled using < 5% IT load at any location at any time.
- Mitigate energy use associated with AI applications in the network/datacenters by enabling heat reuse.
- Towards net zero networks and data centers with lower cost
- Drive US leadership in ICT and energy sustainability

Proposed Targets

Metric	State of the Art	Proposed	Unit
Chip-to-coolant thermal resistance	~ 0.1 (air + Cu heat sink)	0.008	K/W
Coolant Environmental impact	GWP > 0, ODP = 0 or > 0	GWP = ODP = 0	-
Compute node heat density, Q'$_{server}$	~ 80	93.3	kW/m^3
Total TUE (Primary + Secondary)	1.04	1.04	-
MCM TDP, Q$_{MCM}$	400	≥ 1000	W
System availability	>99.982% (Tier3)	>99.982% (Tier3)	-

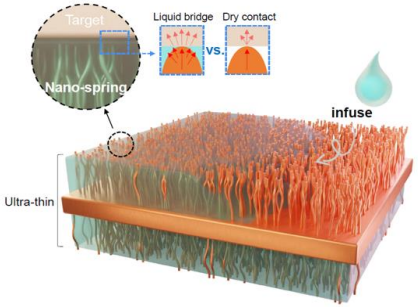
Cooler Development & System Integration

- CTE matched CuW cooler designed and fabricated
- Experimental setup for 4 kW 1U demonstration
- System dynamics study



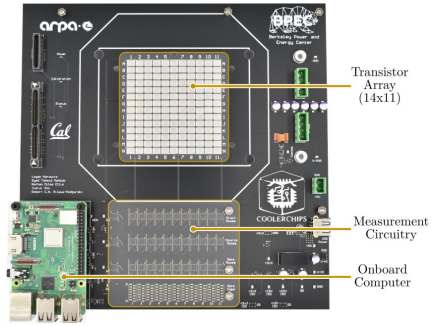
TIM Development and Reliability

- Thermally regulated nanostructured interface material
- Aging and durability testing & mechanical testing



MCM Development

- > 1 kW power array of packaged Si transistors with control and measurements
- Thermal testing vehicle integration



Award: DE-AR0001761

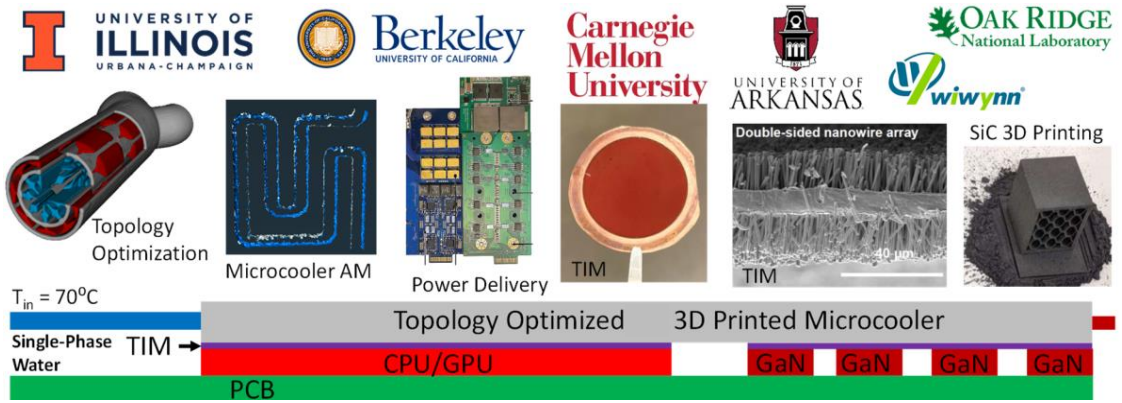
Brief COOLERCHIPS Project Overview

Fed. funding:	$2.5M
Length	36 mo.

Team member	Location	Role in project, core competencies
Prof. Nenad Miljkovic (PI)	UIUC	Cooler Design & Optimization, CDU Development, System Modeling, Server Integration, T2M
Prof. William P. King (co-PI)	UIUC	Cooler Design & Optimization
Prof. Robert Pilawa (co-PI)	UCB	Development of the Multi-Chip Module (MCM) Testbed, Server Integration
Prof. Sheng Shen (co-PI)	CMU	Design and Development of the Thermal Interface Materials (TIMs)
Dr. Kashif Nawaz (co-PI)	ORNL	Manufacturing of Cooler, Rack and System Level Modeling
Prof. David Huitink (co-PI)	UARK	Reliability Analysis of all Individual Technologies as well as System Reliability
Dr. Simon Chen (co-PI)	Wiwynn	Industry Partner and Consultant – Server Provider, T2M Analysis
Dr. Jiu Xu (co-PI)	Meta	Industry Partner and Consultant – T2M Analysis and Implementation Analysis

Context of the project

- High resolution user-friendly MCM
- CTE matched coolers (**< 0.008 K/W**)
- World leading TIMs (**< 0.001 K/W**)
- Reliable single-phase water cooling (Tier3)

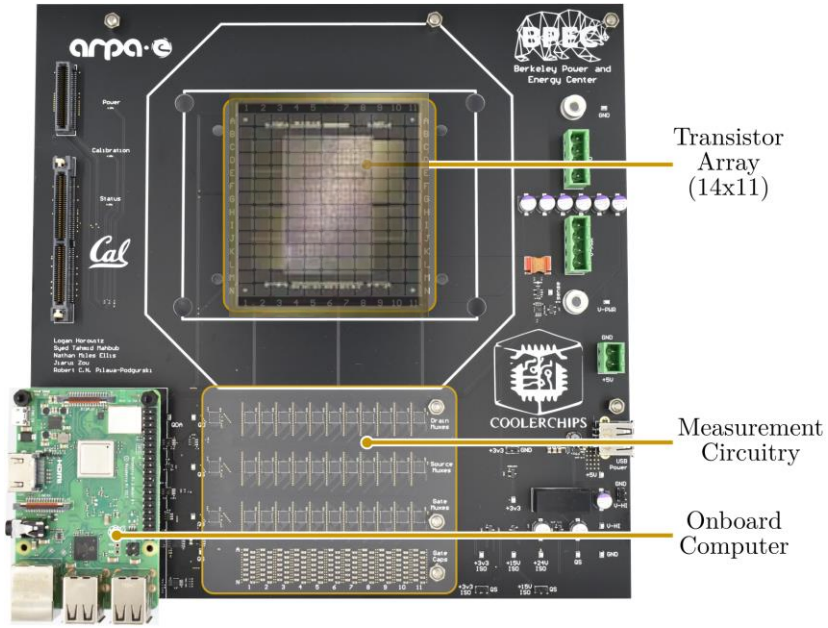


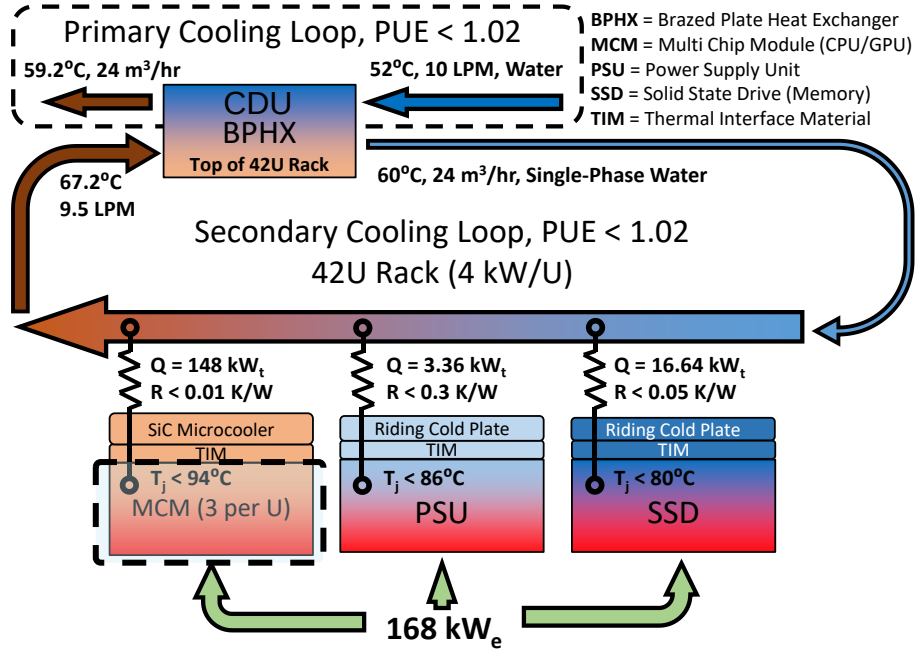
arpa·e CHANGING WHAT'S POSSIBLE

Concept Detail:

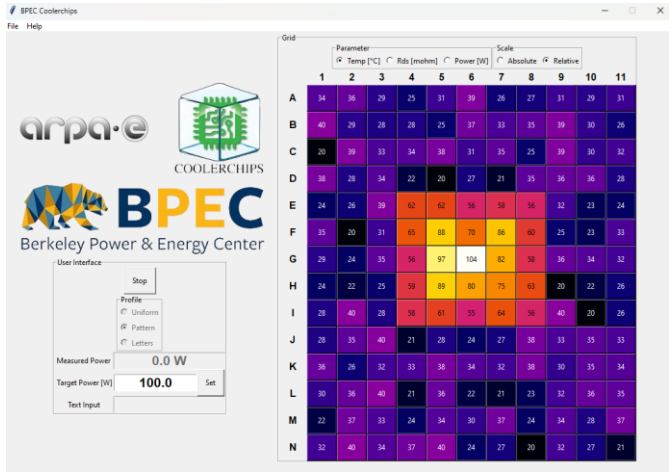
- *Multi-Chip-Module (MCM) Testbeds*

Specification	Proposed Solution
Maximum Power	2 kW
Number of Hotspots	Up to 154
Array Dimensions	75 mm x 75 mm
System Complexity	One standard power supply
Parts Cost	90% lower





- Designed, fabricated, and tested a high-performance TTV based on:
  - Transistor Array
  - Measurement Circuitry
  - Onboard Computer
- Developed a GUI for convenient test specification
- Can mimic heat generation of Nvidia H200 GPU at a fraction of the cost



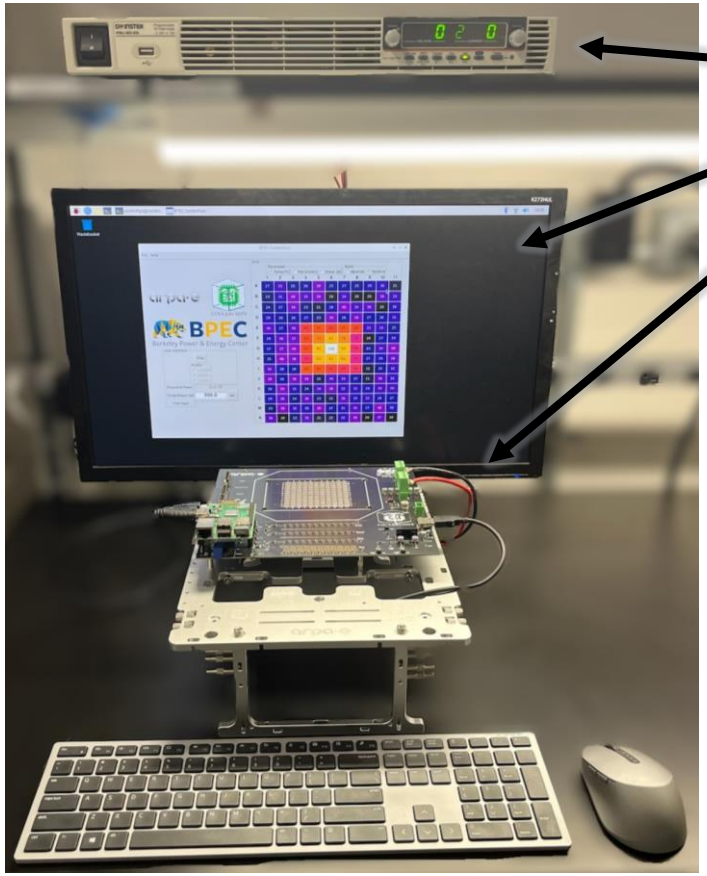
Patent Pending

FOA Metrics	Units
Chip-to-Coolant Resistance Target	< 0.008 K/W
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	**Custom MCM to accurately emulate relevant chipset**
Chip Power	**1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform**
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

arpa·e
CHANGING WHAT'S POSSIBLE

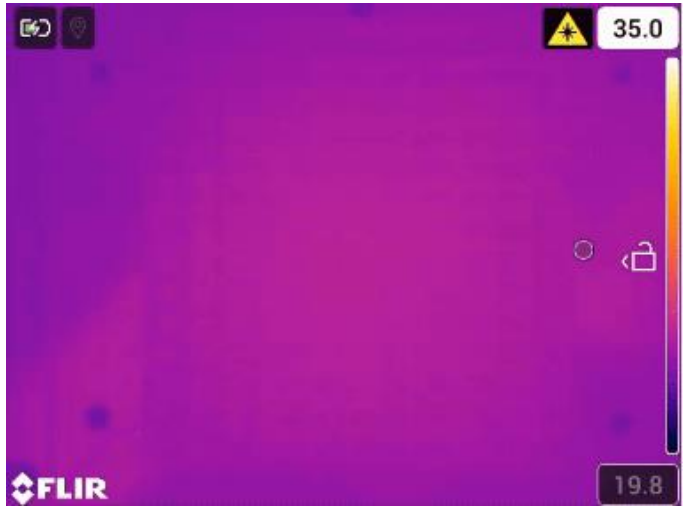
Concept Detail:

- *Multi-Chip-Module (MCM) Testbeds*



- Simple, affordable test setup
  - Single power supply
  - Standard computer monitor
  - TTV prototype

- ARPA-E Summit 2024 Demonstration
- Customizable hot spots including:
  - C-O-O-L-E-R-C-H-I-P-S



Patent Pending

Heat mapping example

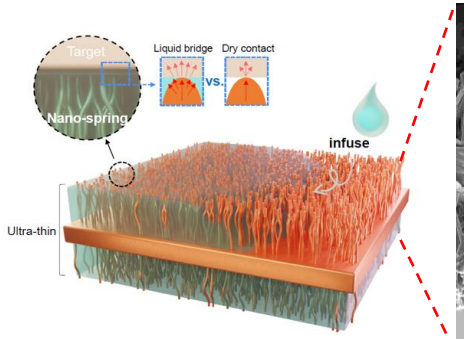
- Provide high resolution heterogeneous thermal profile for CPU/GPU tests
- Accurate local temperature measurements with uncertainty of lower than +/- 0.5 K
- Accurate power dissipation measurements with uncertainty of lower than 15%

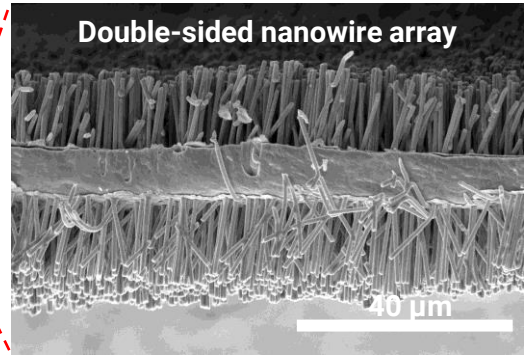
FOA Metrics	Units
Chip-to-Coolant Resistance Target	< 0.008 K/W
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	**Custom MCM to accurately emulate relevant chipset**
Chip Power	**1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform**
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

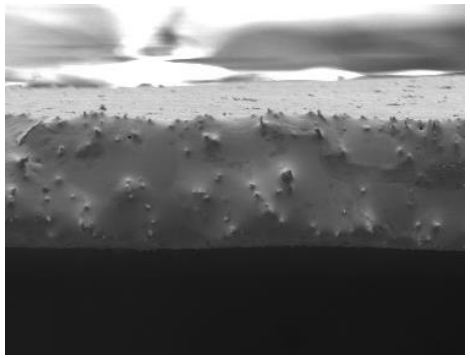
arpa·e
CHANGING WHAT'S POSSIBLE

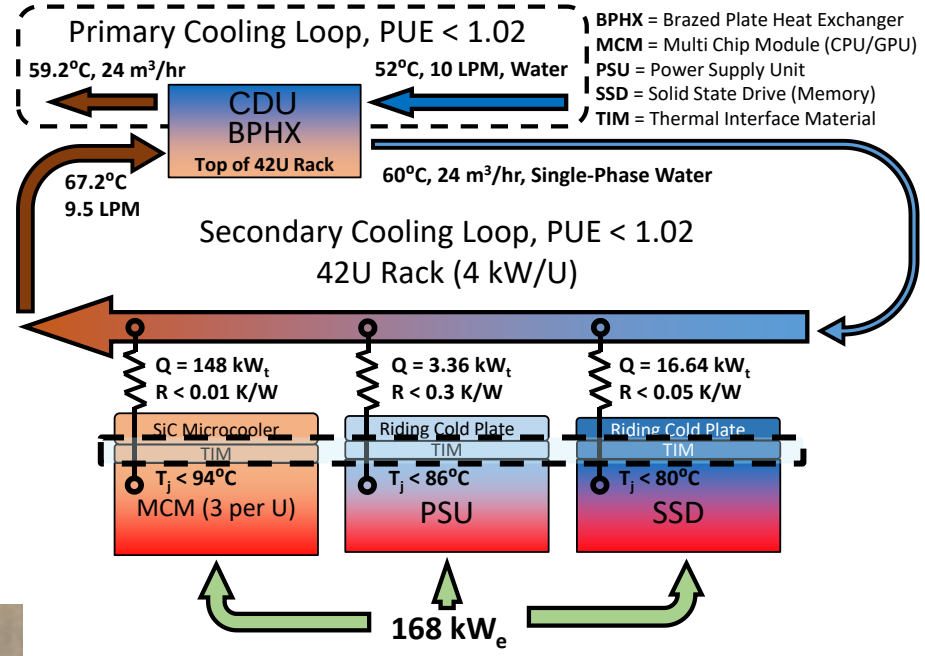
Concept Detail:

- *Reliable Thermal Interface Materials*



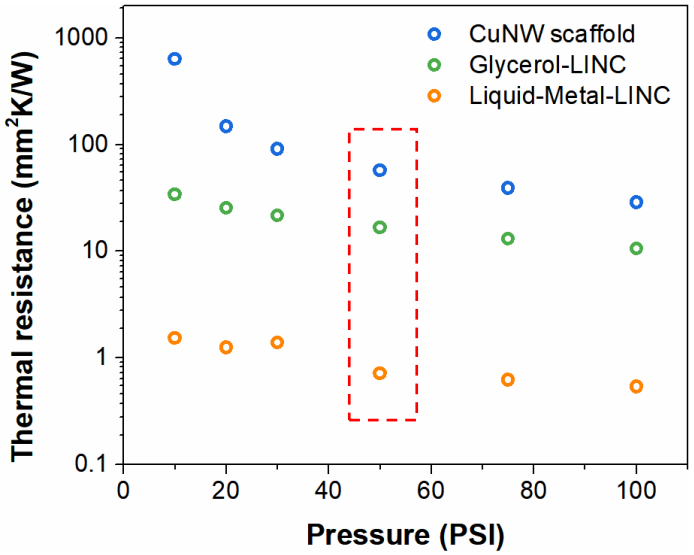


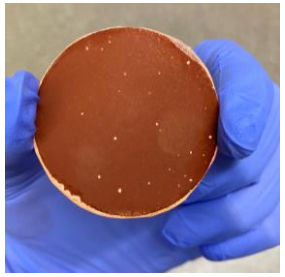


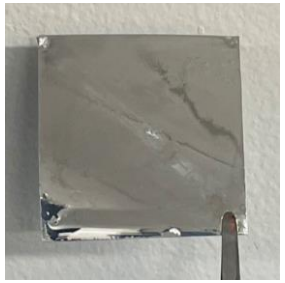


- Achieve uniform 75 mm x 75 mm fabrication with thermal resistance **< 0.001 K/W**
- High effective thermal conductivity up to **73 W/(m-K)**
- 75 mm x 75 mm fabrication available
- **< 0.001 K/W achieved**

Air (dry contact)	0.026 W/(m-K)
IPA	0.137 W/(m-K)
Glycerol	0.3 W/(m-K)
Water	0.6 W/(m-K)
Liquid metal	**73 W/(m-K)**







FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

arpa·e
CHANGING WHAT'S POSSIBLE

Concept Detail:

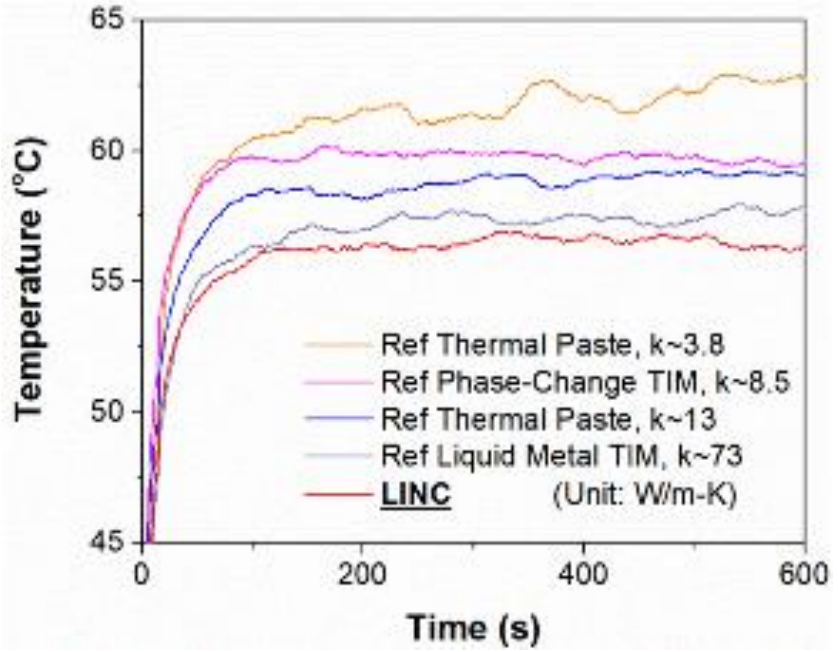
- *Reliable Thermal Interface Materials*

NL

https://www.novolinc.com/

https://www.linkedin.com/company/novolinc/about/

- Next-generation thermal interface material to resolve the heat dissipation challenge in energy-dense devices and systems
- > 5 times better thermal performance than the SOA commercial products at a comparable cost



OPEN Compute Project®

- Attended OCP 2024
- Greater performance, adaptability, and cost consideration

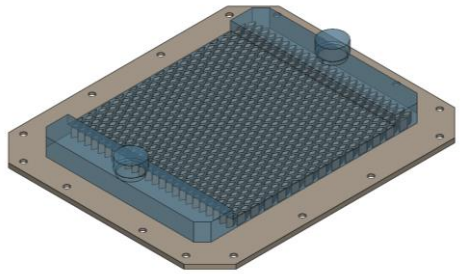
https://www.linkedin.com/posts/tdk-ventures_chipcooling-datacenterenergy-nextgenchip-activity-7257133448446783488-JFrO?utm_source=share&utm_medium=member_desktop

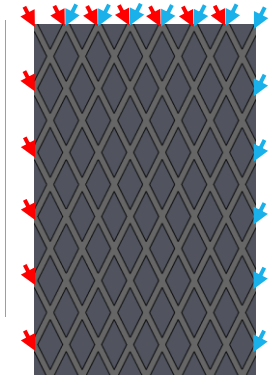
FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

arpa·e CHANGING WHAT'S POSSIBLE

Concept Detail:

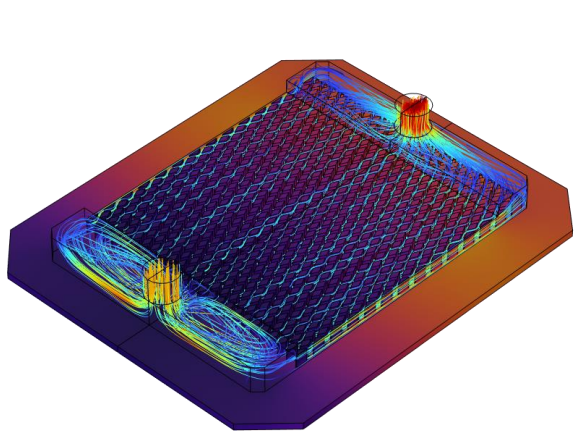
- *Diamond Pin Array Cold Plate*

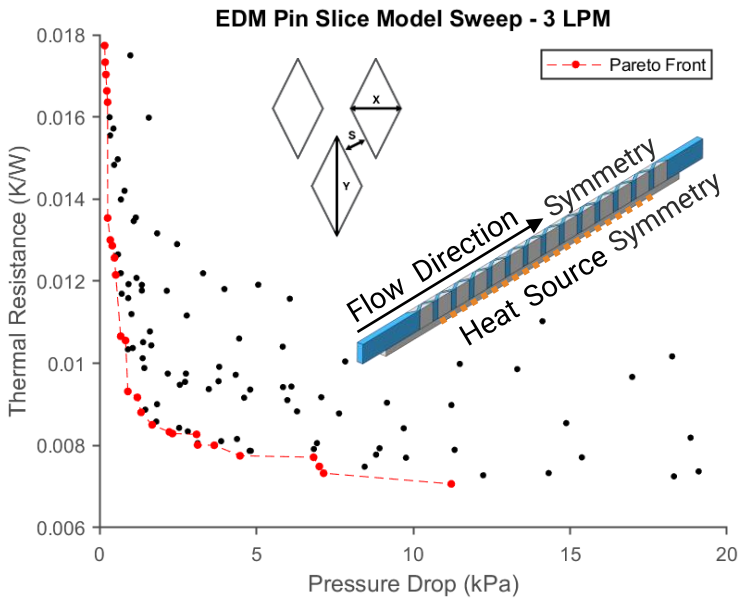


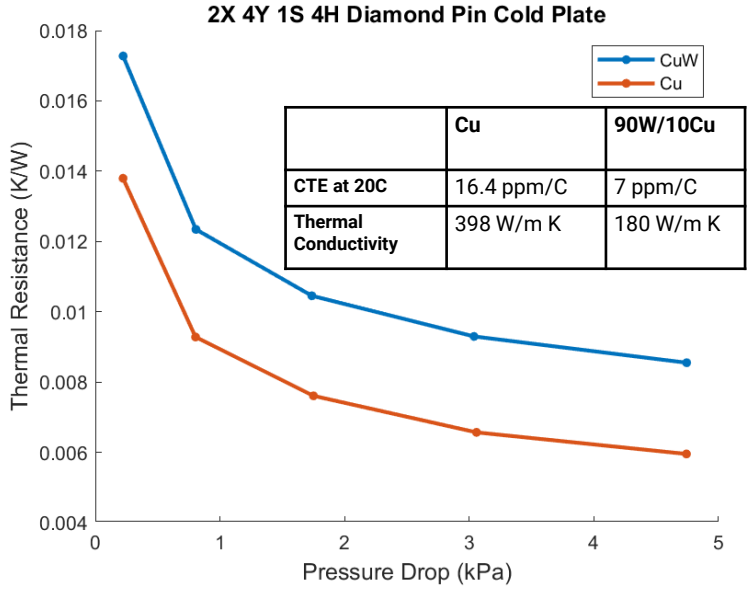


Cut Lines

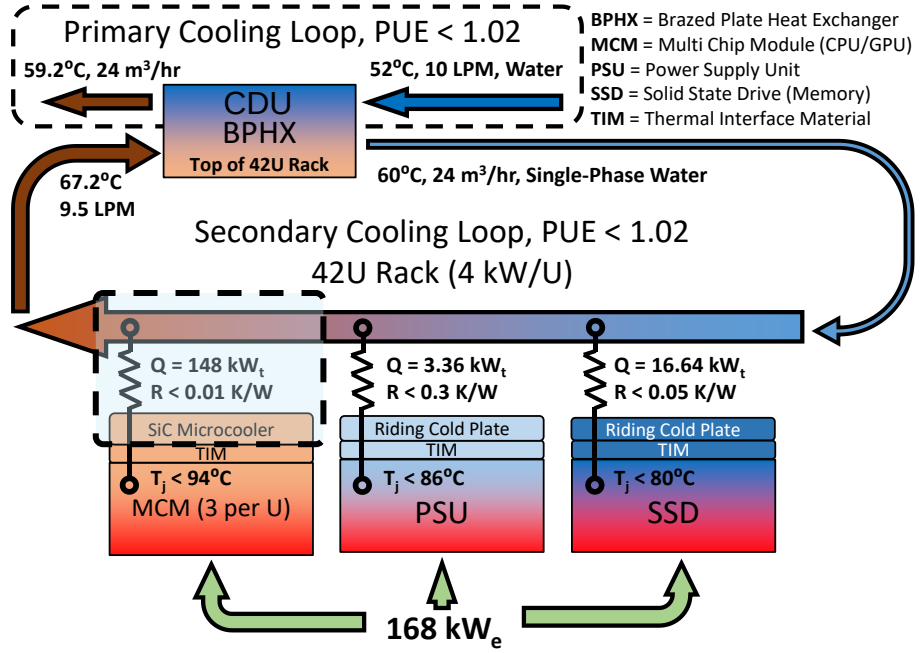
Diamond pins can achieve high performance while reducing manufacturing complexity







	Cu	90W/10Cu
CTE at 20C	16.4 ppm/C	7 ppm/C
Thermal Conductivity	398 W/m K	180 W/m K

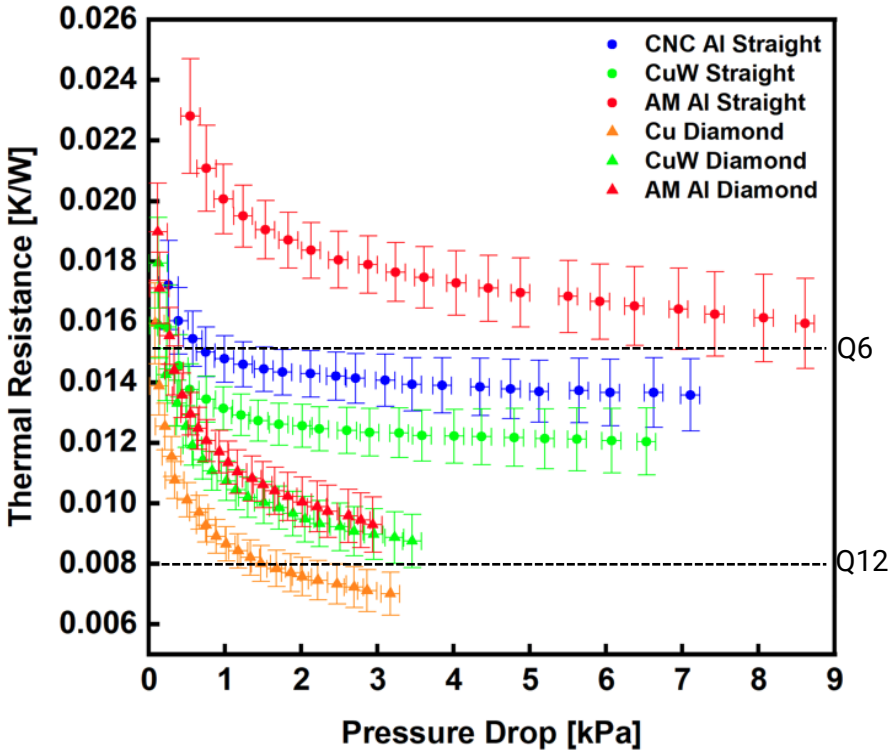


FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

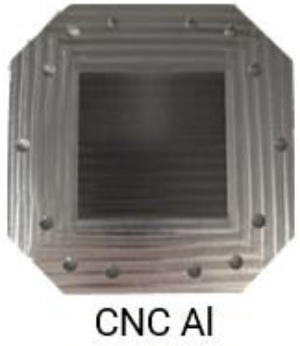
arpa·e
CHANGING WHAT'S POSSIBLE

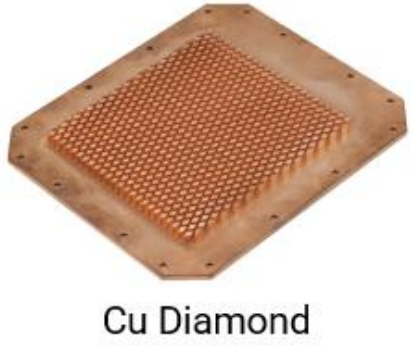
Concept Detail:

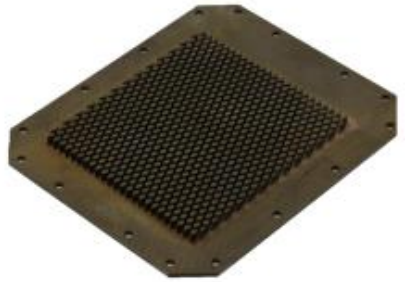
- *Diamond Pin Array Measured Performance*

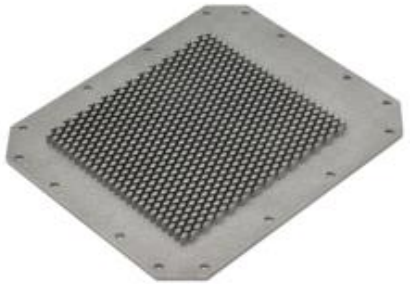


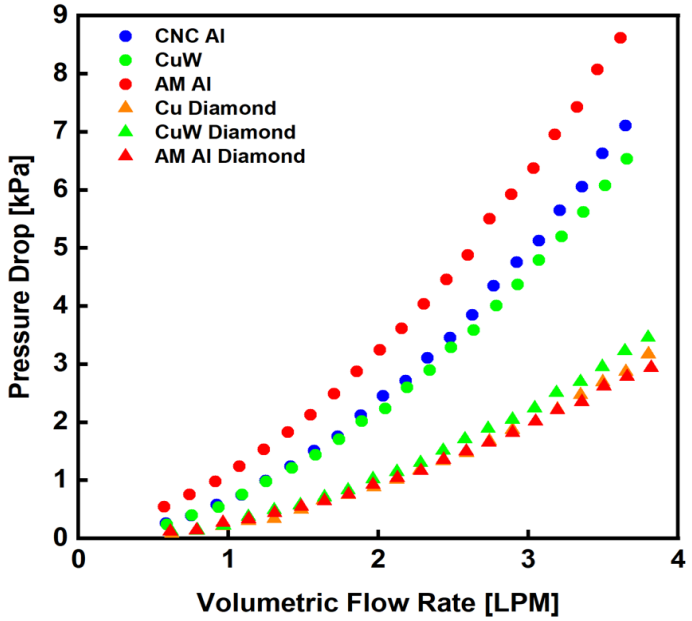
Diamond pin array outperforms benchmark microchannel finned cold plates and achieves **0.009 K/W** thermal resistance at **3.5 kPa** pressure drop for CuW









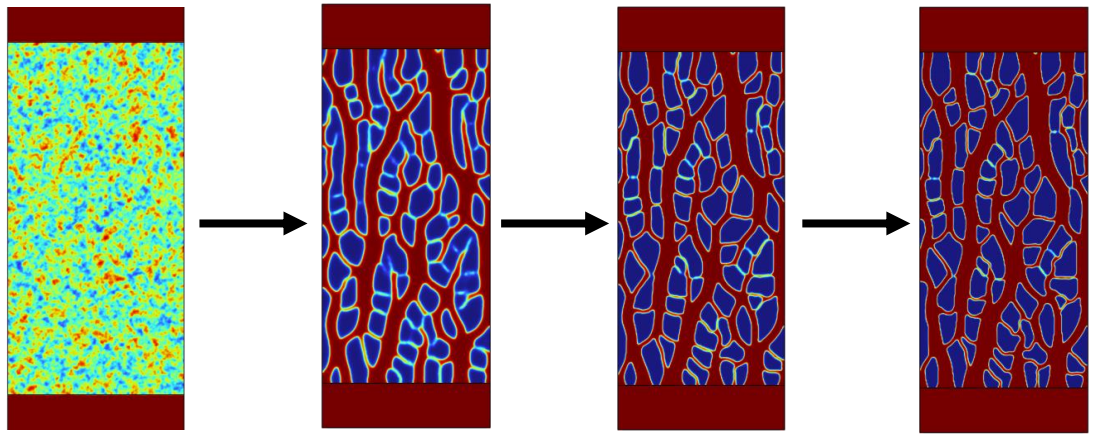


Optimized diamond pin cooler pressure drop is significantly lower than straight fin design

arpa·e
CHANGING WHAT'S POSSIBLE

Concept Detail:

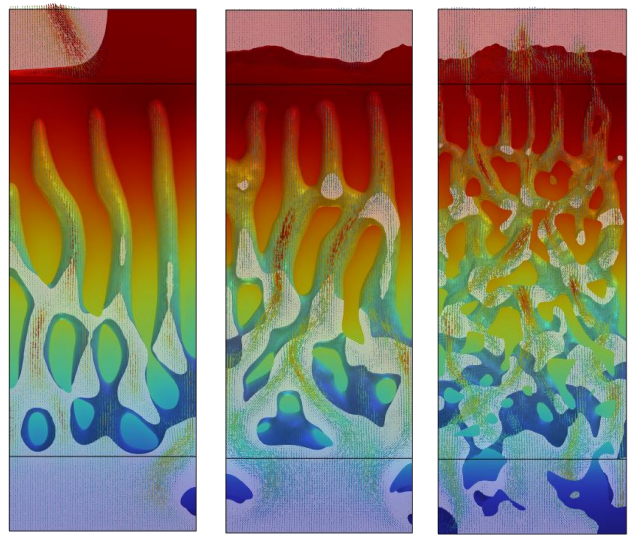
- *Topology Optimized Cooler with Extruded Pins*



Initial Density Field

Continuation approach applied over 150 iterations with weighted sum objective function

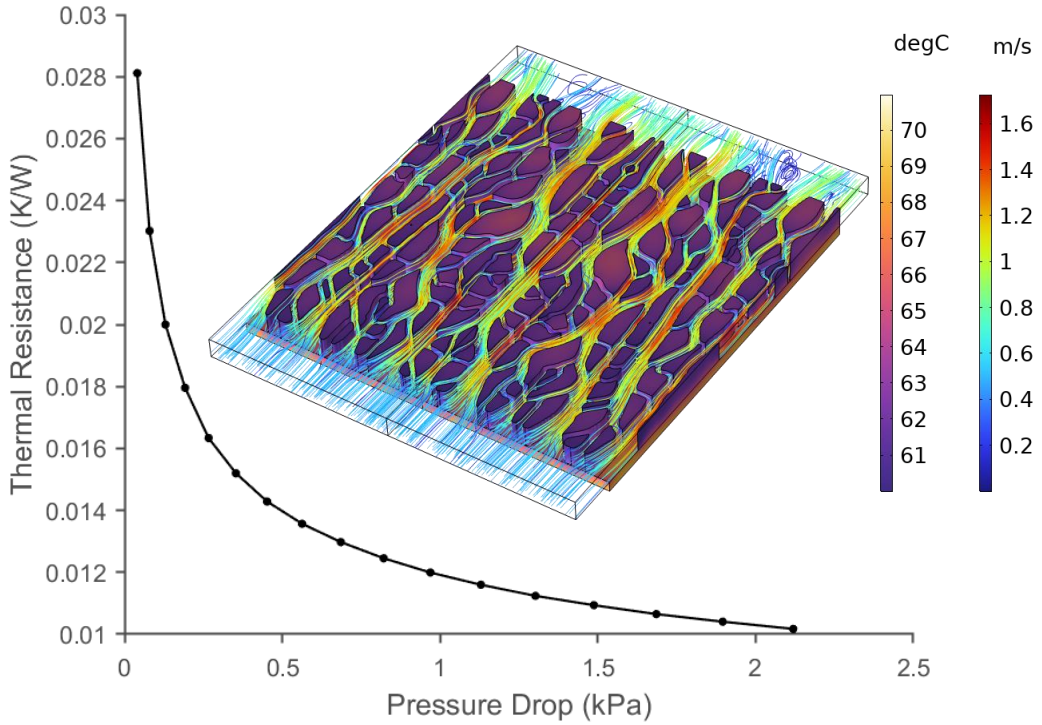
$$\min w_1 T_{av} + w_2 dP$$
$$s.t. \theta_{av} \geq 0.5$$



3 mm 2 mm 1 mm

Result is extruded and simulated in 3D

Geometry achieves 0.01 K/W thermal resistance at 2 kPa pressure drop



FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM)
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

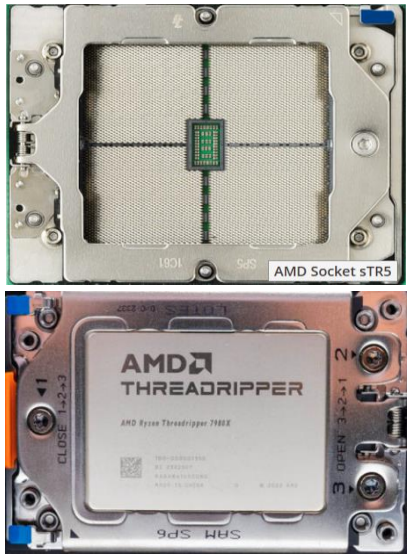
arpa·e CHANGING WHAT'S POSSIBLE

Concept Detail:

- *Single Server Prototype*

CPU platform

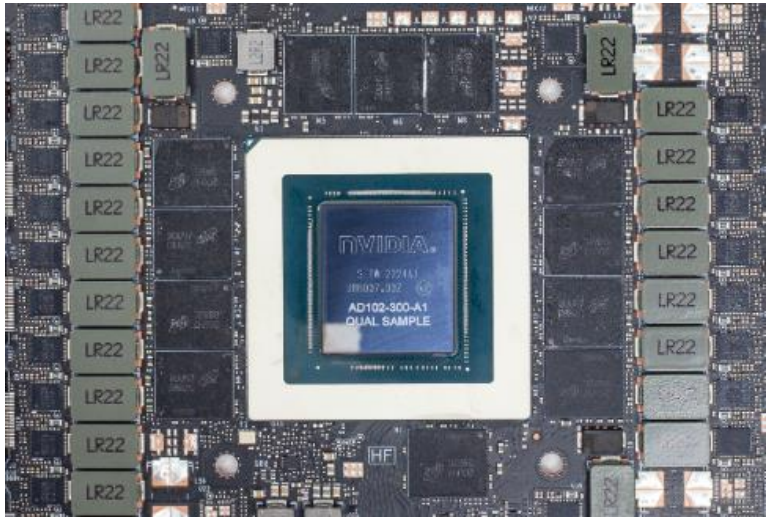
- Case integration
- **AMD Threadripper** 7980X – 70 x 56 mm
- 970 W TDP achieved
- 1U chassis with additional fan cooling for voltage regulators and other minor components



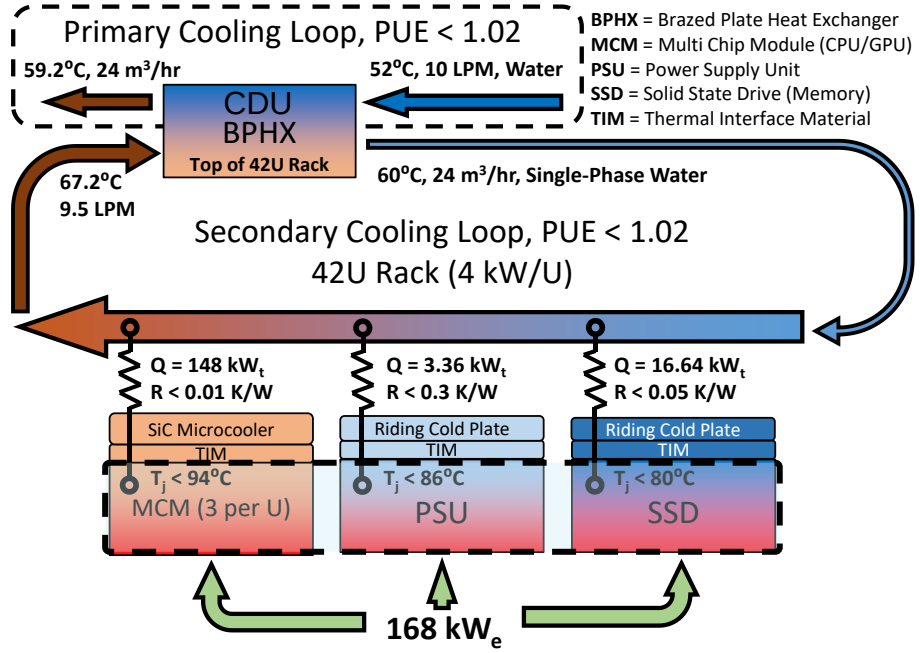
Surce: https://www.servethehome.com/amd-ryzen-threadripper-7980x-review-a-funky-workstation-cpu-some-will-love/

GPU platform

- Bare-die integration
- **Nvidia 4090** OC – 609 mm^2 + minor components
- 600 W TDP achieved
- Two GPUs to achieve 1200 W TDP



Source: https://www.techpowerup.com/review/nvidia-geforce-rtx-4090-founders-edition/4.html

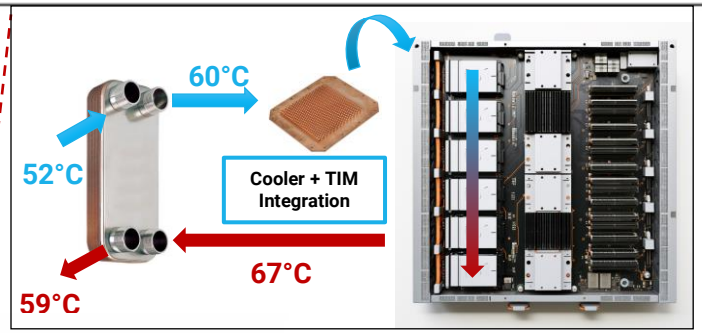


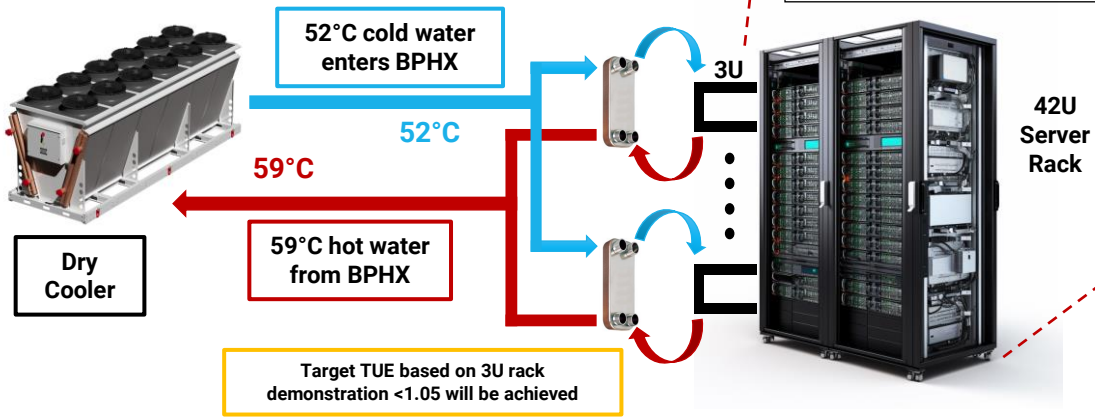
FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

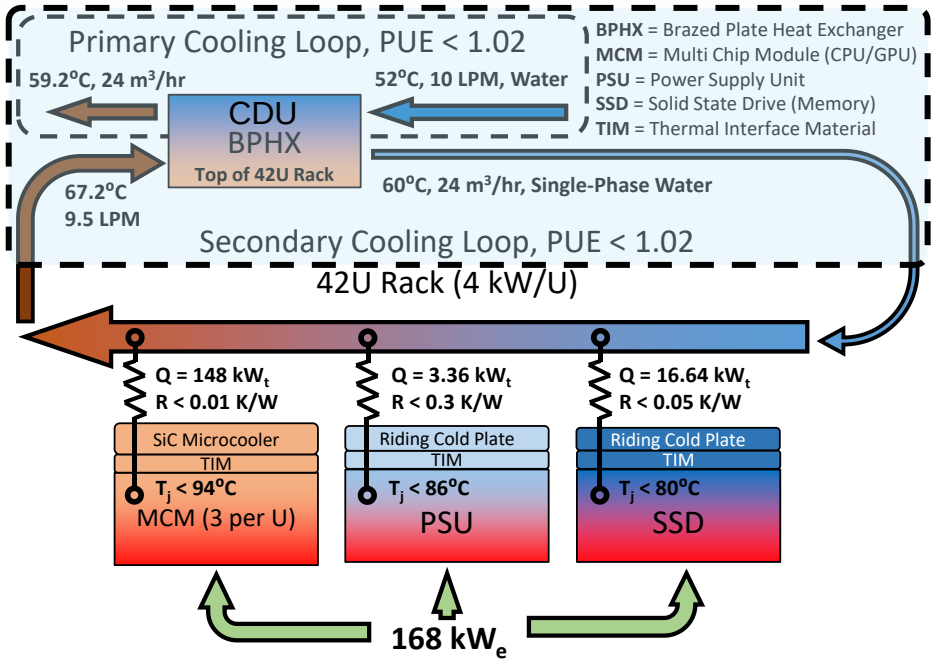
arpa·e
CHANGING WHAT'S POSSIBLE

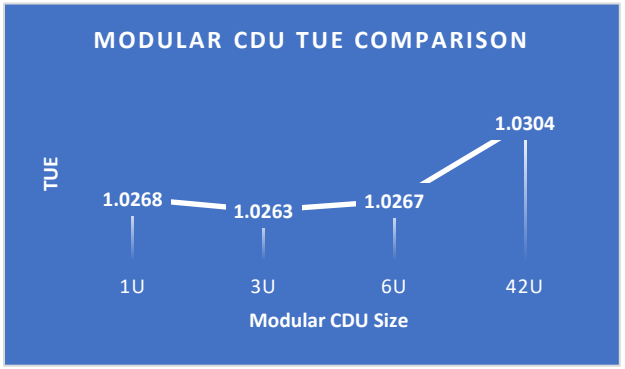
Concept Detail:

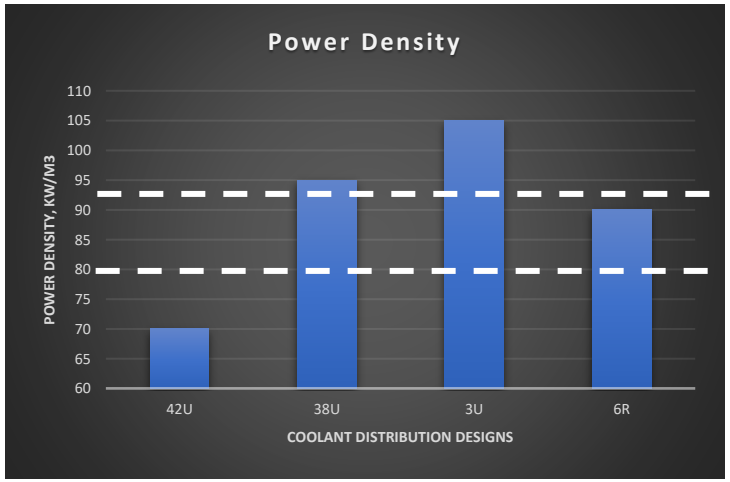
- *CDU: Low power consumption*
  - L2L cooling architecture
  - Modular CDU < 0.0006 K/W
  - System TUE < 1.027 (using 10 kPa)









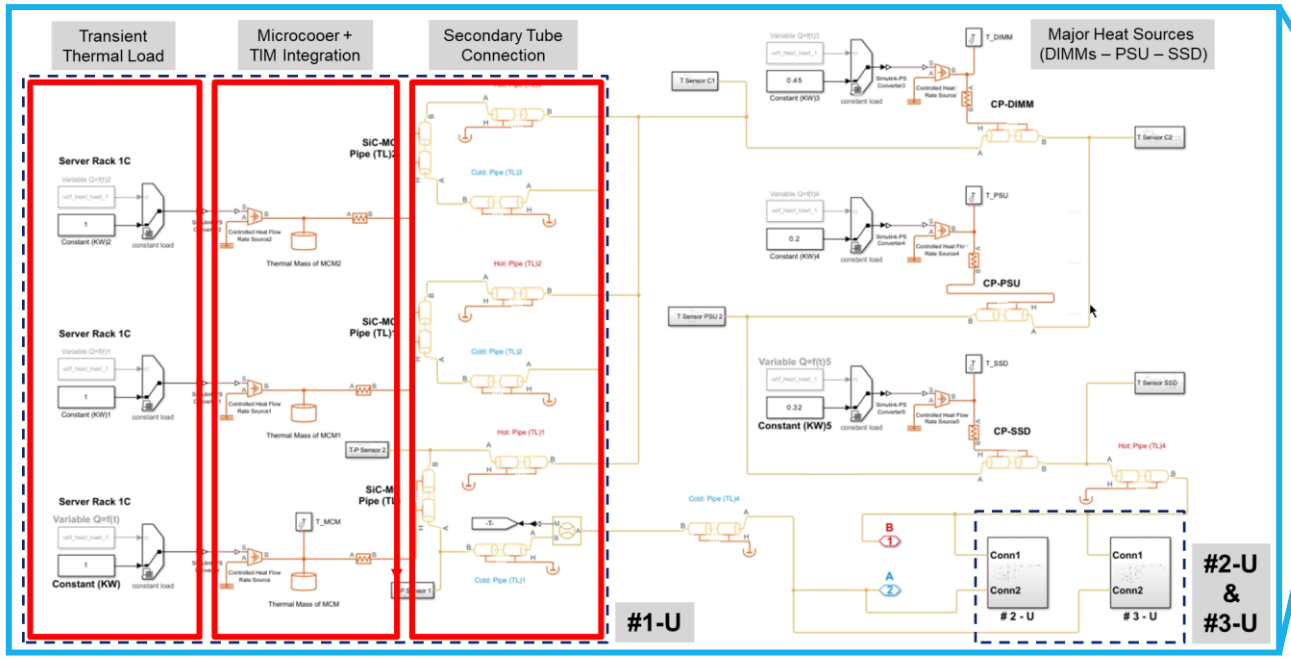


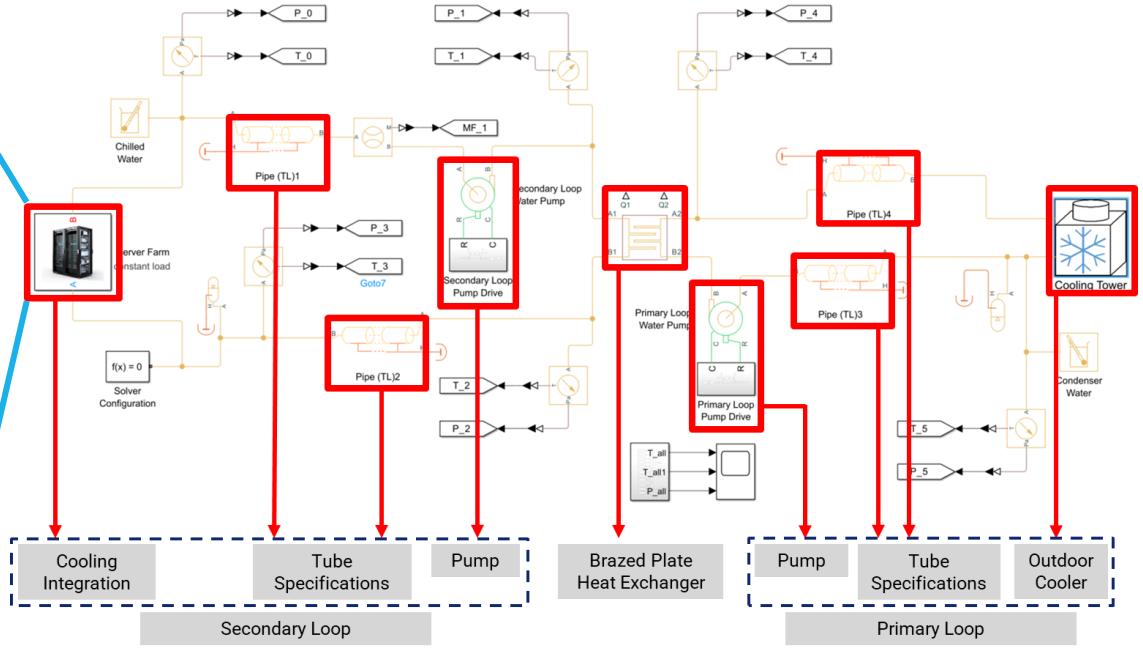
FOA Metrics	Units
Chip-to-Coolant Resistance Target	< 0.008 K/W
Cooling Power % of IT Load	**≤ 5 %**
System availability	>99.982 % (Tier 3)
CDU Metrics	**Proposed**
Coolant	WEG
Secondary Loop Cooling Power % of IT Load	2 %
Coolant Supply Temperature Range	42°C to 62°C
Coolant Return Temperature Range	47.5°C to 67.5°C

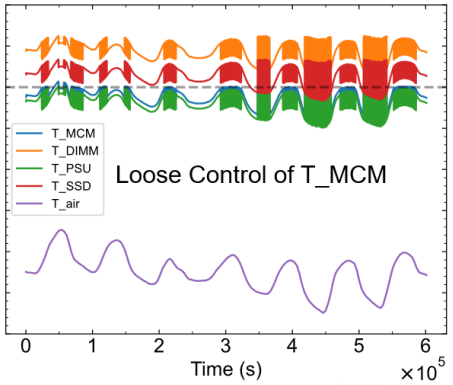
arpa·e
CHANGING WHAT'S POSSIBLE

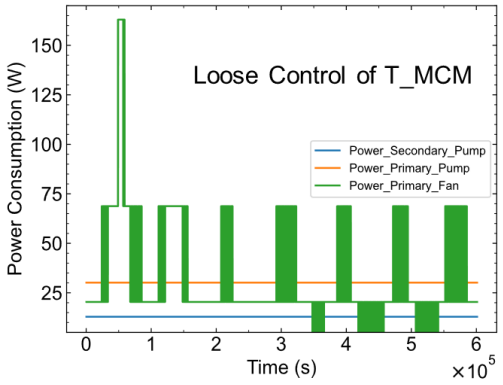
Concept Detail:

- *System Level Dynamic Study Model*









Hyperscale system level dynamic study model

- Active pump & dry cooler control following the optimized model based on transient parameters to achieve consistent chip temperature
- Current model: single rack
- Next step: hyperscale

FOA Metrics	Units
Chip-to-Coolant Resistance Target	**< 0.008 K/W**
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset	Custom MCM to accurately emulate relevant chipset
Chip Power	1 kW (75 mm x 75 mm MCM) Reference: Blackwell platform
Power per Rack	168 kW (42U Rack)
Demonstration power mid project	4 kW (1U Rack)

arpa·e
CHANGING WHAT'S POSSIBLE

Concept Detail:

- *Waterfall Analysis*

Temperature flow description:

❶85°C chip temperature →

❷TIM temperature →

❸server's liquid outlet temperature →

❹S.BPHX's liquid outlet temperature →

❺P.BPHX's liquid outlet temperature →

❻P.BPHX's liquid inlet temperature →

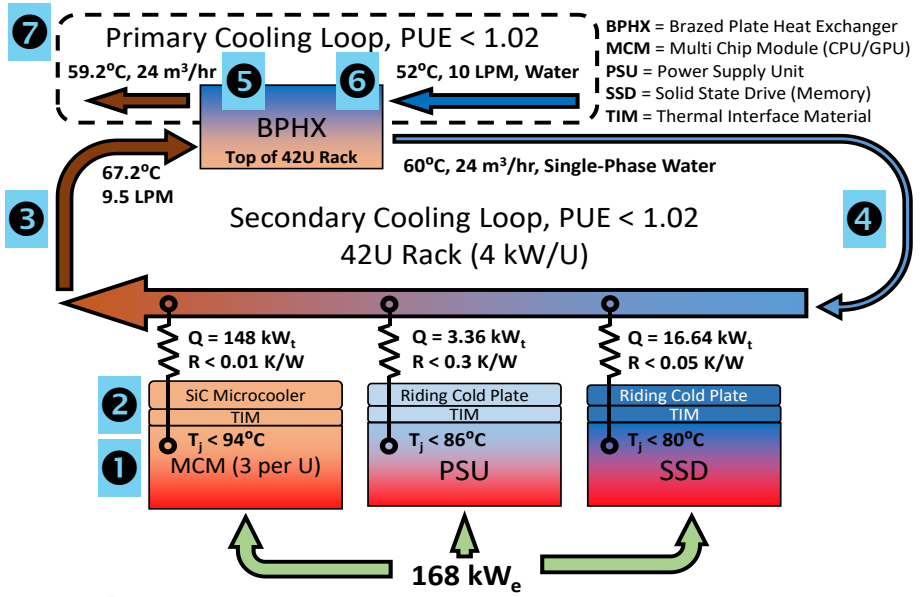
❼ambient air temperature

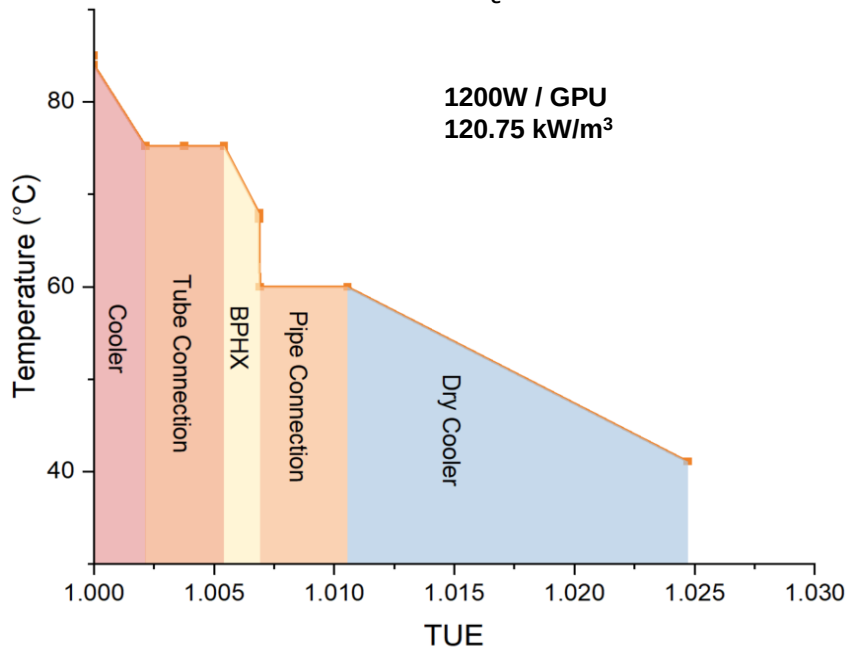
TUE flow description:

Microcooler loss → secondary QDs/manifold/tube loss → BPHX loss → primary manifold & pipe loss → dry cooler fan power

Phoenix, AZ - July

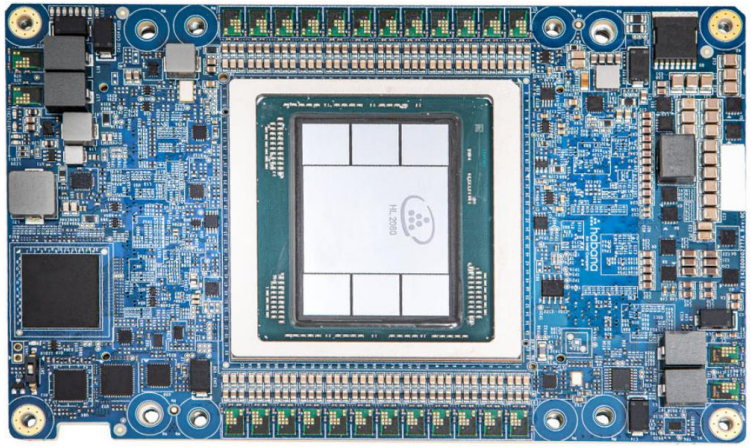
- Air Dry Bulb Temperature: 41.1°C
- Power density 120.75 kW/m^3 (1200W / GPU) : **1.0247 TUE**





Concept Detail:

- *Server: Wiwynn E-tron Training Server with Intel Habana Gaudi2*
- *Meta and Wiwynn trying to get a potential server platform for a demo of this project*



Key Parts	Description
System Dimension	3U 19”
CPU head node	2* INTEL Xeon Ice lake CPU, 270W
	32* DDR4 DIMM
	2* OCP 3.0 NIC
HIB / Switch Board	2* PCIe switch
UBB	OAI compliant, Up to 8 OAM 6 QSFP-DD 400G for scale out
OAM	8* Habana Gaudi 2 , 600W
Storage	4* PCIe Gen 4 U.2 NVME SSD
Cooling FAN	6056Fan MB x 6 ,UBB x12
PSU	4 (3+1) 4 kW PSU
Dimension	W 447mm, D 915x H 263mm

FOA Metrics	Units
Chip-to-Coolant Resistance Target	< 0.008 K/W
Cooling Power % of IT Load	≤ 5 %
System availability	>99.982 % (Tier 3)
Chipset (Q6)	**Intel Habana Gaudi2**
Chip Power	**> 1 kW**
Power per Rack	168 kW (42U Rack)
Demonstration power	12 kW (3U Rack)

arpa·e
CHANGING WHAT'S POSSIBLE

Concept Detail:

- *Technique to Marketing and Commercialization*

University of Illinois Urbana-Champaign — *Holistic Direct to Chip Liquid Cooling Architecture*

- Commercialization activities being pursued based on the competitive cooling performance advantage with our **High-Performing Microcoolers** for next-gen HPC.
- Building a **Holistic DLC Architecture Ecosystem** by partnering and engaging with key companies in the DC thermal management sector.
- Extend **Ecosystem** beyond DLC Architecture to **AALC Architecture** by working with **Meta**.

TycheTools — Datacenter Infrastructure Management Tools

Burns McDonnell — Integrated Solutions for Data Center Design and Construction

SemiAnalysis — Data Center Market Research

SABIC — High-Grade Insulations ULTEM™ 1000 and 2000

CPC — Everis® Quick Disconnects LQ4 and LQ6

KAORI — 3U Sized BPHX

Carnegie Mellon University — *High Performance and Reliable LINC TIM*

- Commercialization activities being pursued by forming a new company, **NovoLINC.**
- Preparation of **SCALEUP** program application through NovoLINC.
- Engaged with Google, Intel and Meta as potential clients through NovoLINC.
- Engaged with HP for future collaboration in testing of thermal performance, reliability and warpage conformability.

NovoLINC — Making Tech Cool™

Google, Meta, intel

HP — Collaboration for TIM Performance Testing

Berkeley University of California — *Low Cost and High Performance MCM TTV*

- Commercialization activities being pursued by forming a new company.
- Preparation of **SCALEUP** program application through the new company.
- Engaged with HP, JetCool, Google, Arista Networks, and Vertiv as potential clients.

HP, jetcool, Google, ARISTA, VERTIV™

Q & A

https://arpa-e.energy.gov