CalNEXT
ENERGIZING CALIFORNIA'S FUTURE
Funding acknowledgement

Assessment of Cooling Strategies for High Compute Density Data Centers-Performance, Energy Use and Non-energy Benefits

Vinod Narayanan

University of California Davis

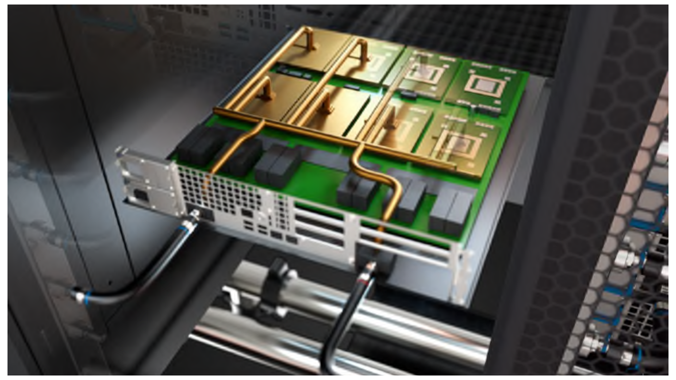
IPER Workshop February 10-11th 2026

Key contributors: Subhrajit Chakraborty, Ines-Noelly Tano, Ellian Eorwyn, Erfan Rasouli, Sarah Outcault

UCDAVIS
Western Cooling Efficiency Center

Background and Motivation

Liquid cooling is becoming mainstream for high density AI servers



Cold plate; Courtesy: Vertiv

L-L CDU; Courtesy: Boyd

Questions

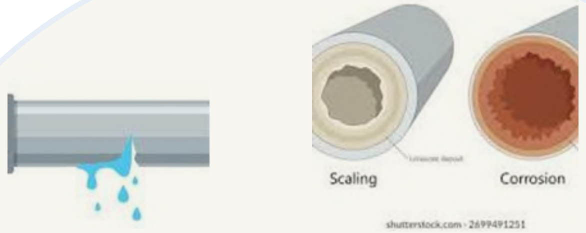
1. What are the different types & how can we classify them?
2. Which liquid cooling solution is the "best"?

RDHX; Courtesy: Data Center Dynamics

L-A CDU; Courtesy: Vertiv

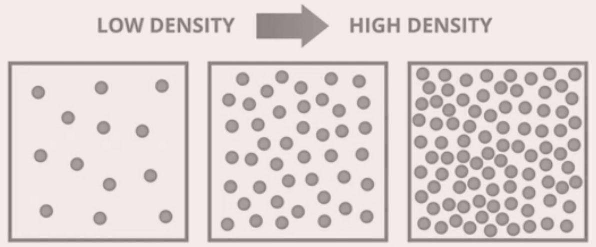
Reduced energy consumption

Ease of retrofit

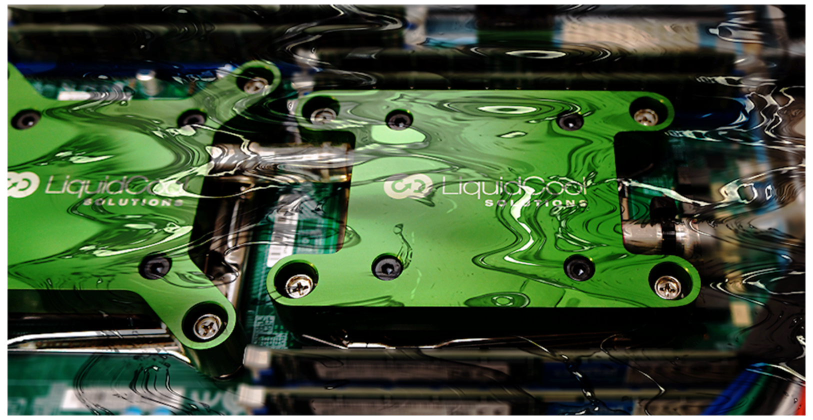


$ ⟷ $$$

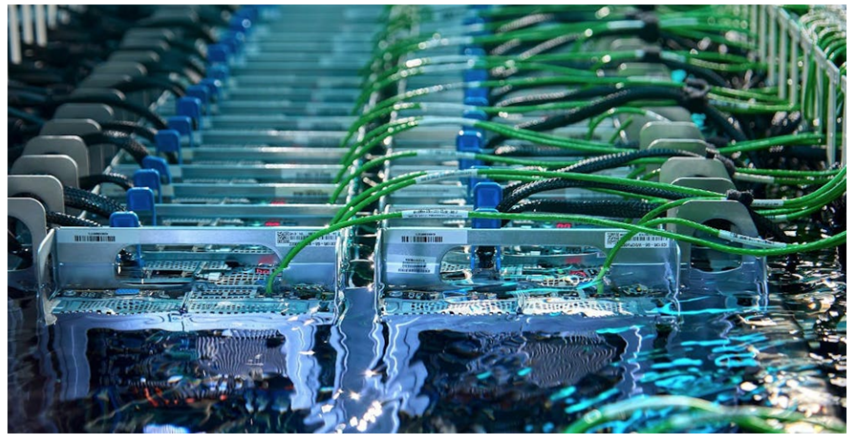
Corrosion, erosion, leaks, cost
(Non-energy considerations)



High compute density



Single phase hybrid immersion servers; Courtesy: LiquidCool Solutions Inc



Single phase immersion tanks; Courtesy: Data Center Dynamics

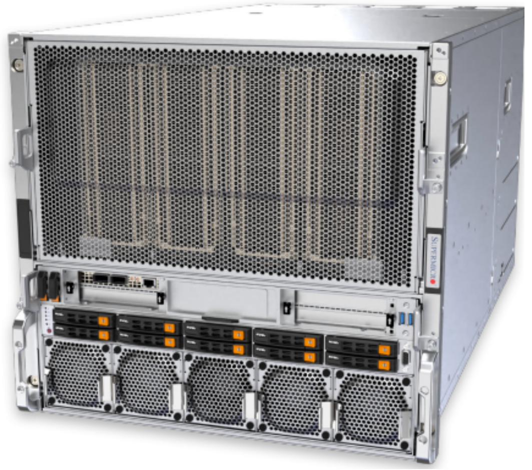
Case Study- Reference Architectures

Air-cooled 10 U server components and power density

Server components	TDP each (W)	Quantity (#)	Total power (kW)
Blackwell GPU/server	1000	8	8
CPU/server	300	2	0.6
NV link /server	35	8	0.28
DPU bluefield/server	75	2	0.15
SSD/server	15	36	0.54
DDR/server	15	24	0.36
Network switches/server	100	4	0.4
Total IT TDP/server			10.3

Air-cooled 48U rack components and power density

Rack components	TDP each (W)	Quantity (#)	Total power (kW)
Servers /rack	10330	4	**41.3 (Silicon Compute Power)**
PSU thermal dissipation/rack	2900	1	2.9
Total IT TDP/rack			**44**

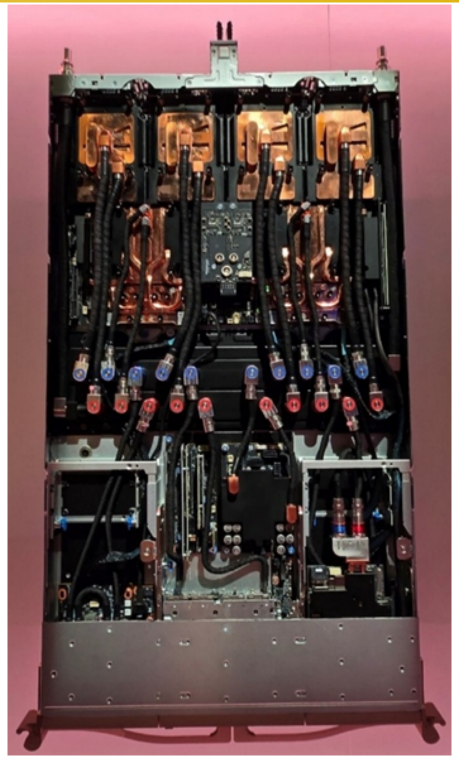


Liquid-cooled 1U compute/server tray components and power density

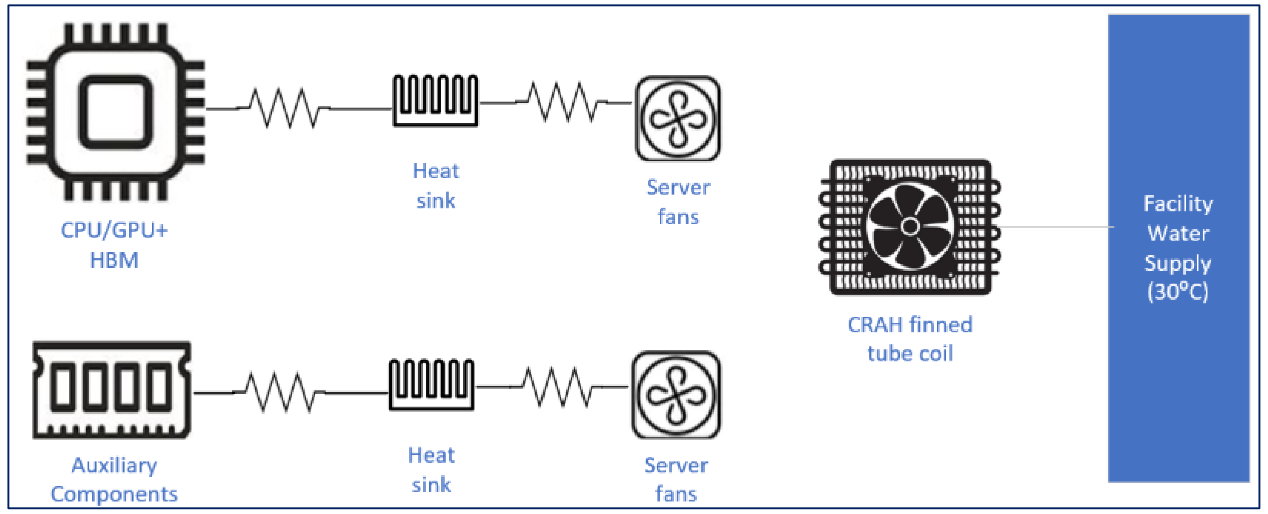
Server Tray Components	TDP (W)	Quantity (#)	Total power (kW)
Blackwell GPU/server	1400	4	5.6
CPU/server	350	2	0.7
NV link/server	35	4	0.14
DPU bluefield/server	75	1	0.075
SSD/server	15	16	0.24
DDR/server	15	16	0.24
Total IT TDP/server			7

Liquid-cooled 48U rack components and power density

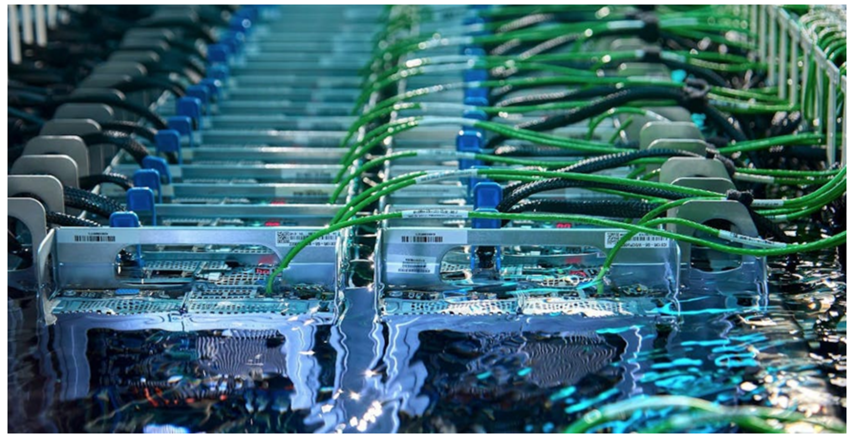
Rack Components	TDP (W)	Quantity (#)	Total power (kW)
Servers (1U)/rack	6995	18	**125.9 (Silicon compute power)**
NV link Switch trays (1U)/rack	2000	9	18
PSU thermal dissipation/rack	1680	6	10.1
Total IT TDP/rack			**154**

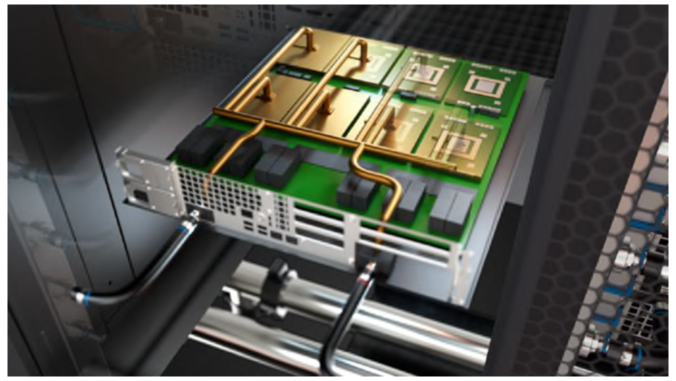


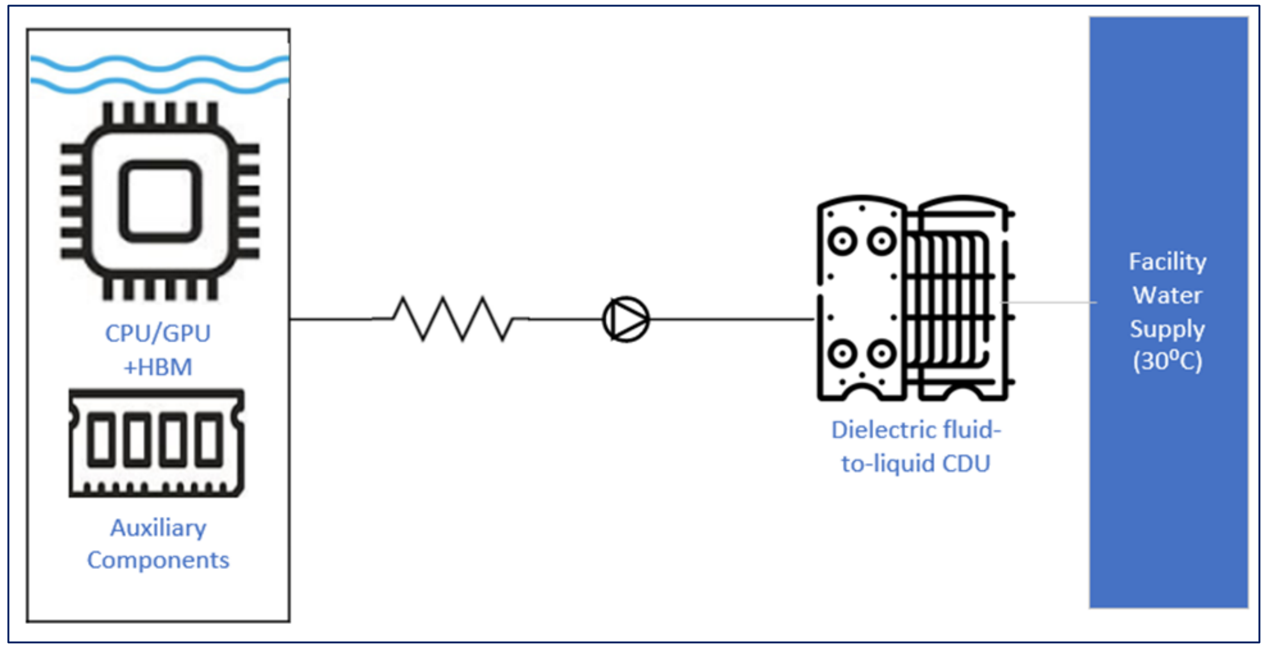
Classification of Liquid Cooling Technologies- Single mode

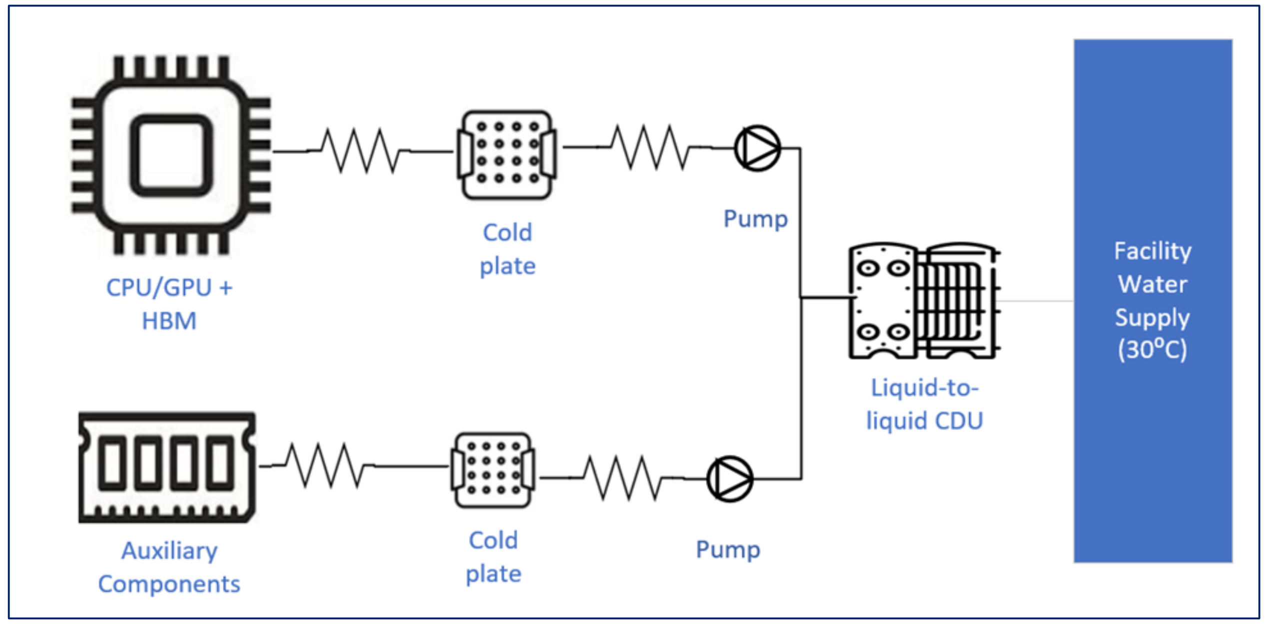


Type of cooling	Options
Traditional CRAC/CRAH air-cooling	• Without Rear Door HX • With Rear Door HX
D2C with cold plates	• Single phase liquid • Two phase refrigerants
Full-immersion cooling	• Single phase liquid • Two phase refrigerants

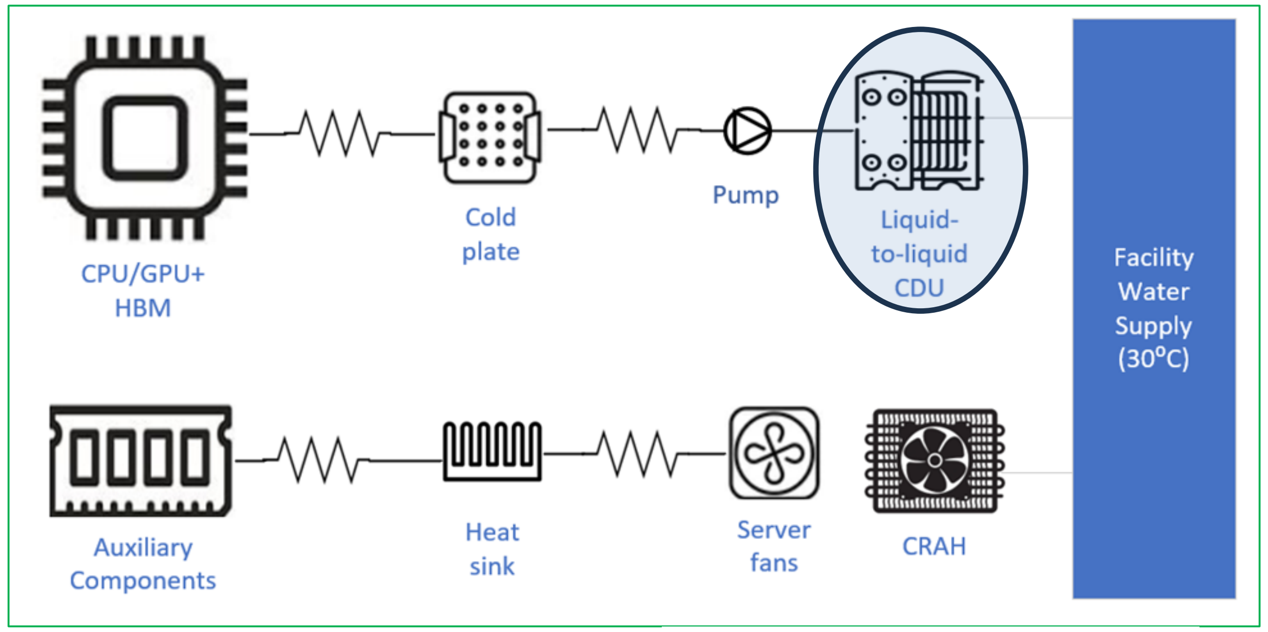




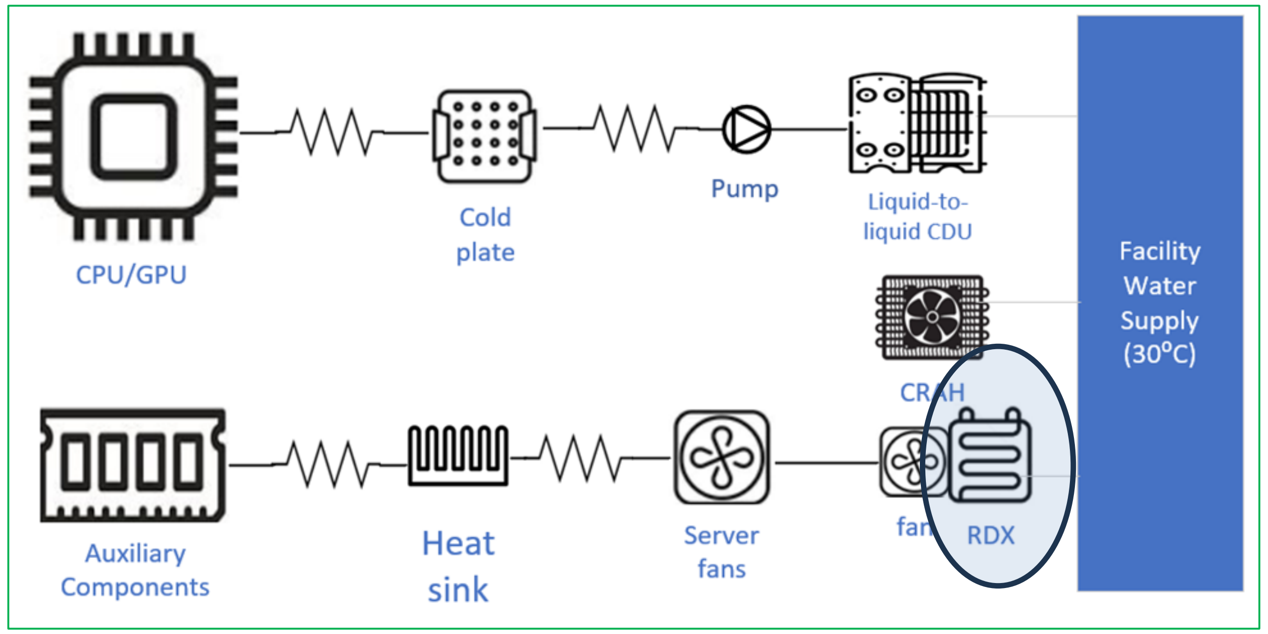




Classification of Liquid Cooling Technologies- Hybrid mode



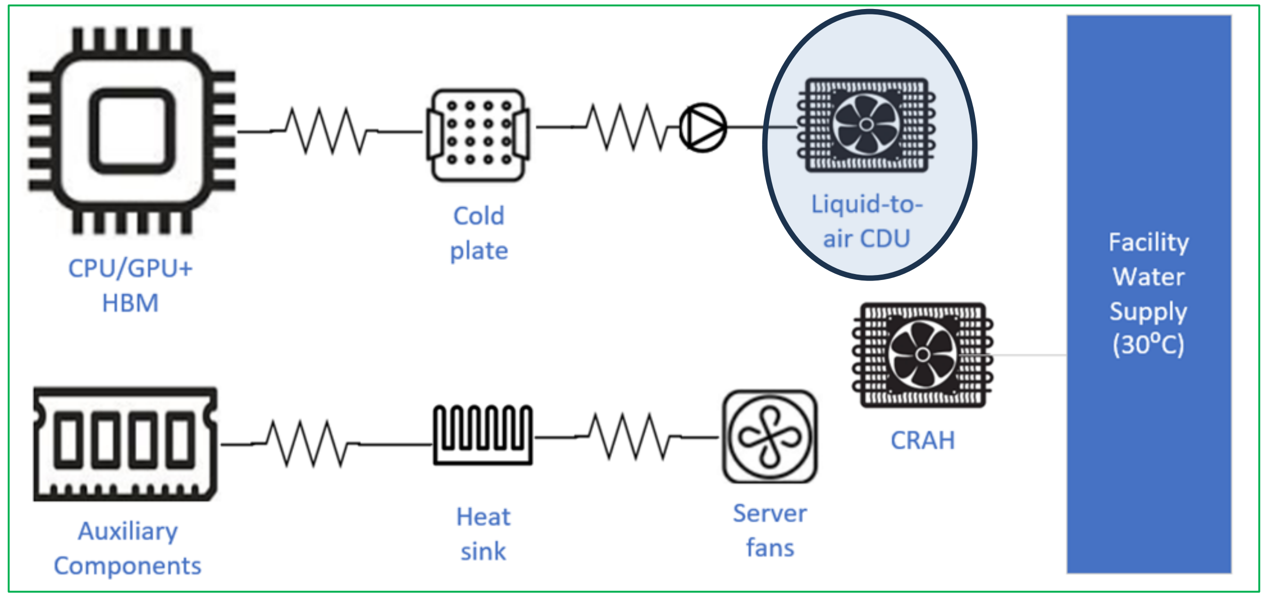
Type of cooling	Options
D2C cold plate + traditional air-cooling	• Single phase liquid • Two phase refrigerants
D2C cold plate + liquid-to-air CDU for traditional air-cooling	• Single phase liquid • Two phase refrigerants
D2C cold plate + Rear door HX for traditional air-cooling	• Single phase liquid • Two phase refrigerants
D2C cold plate + immersion cooling	• Single phase liquid • Two phase refrigerants

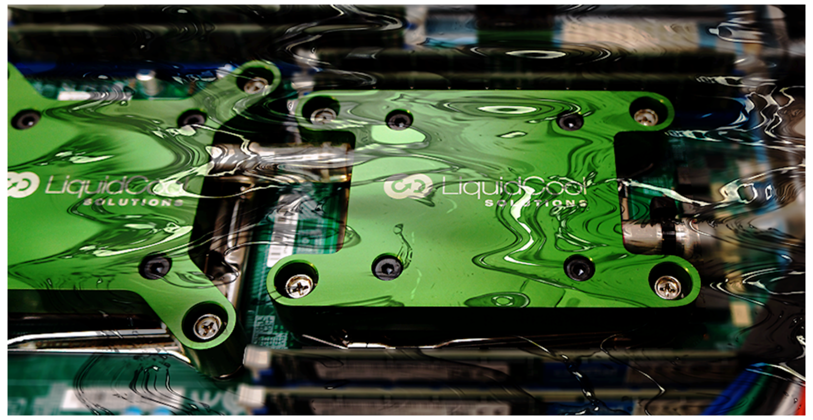


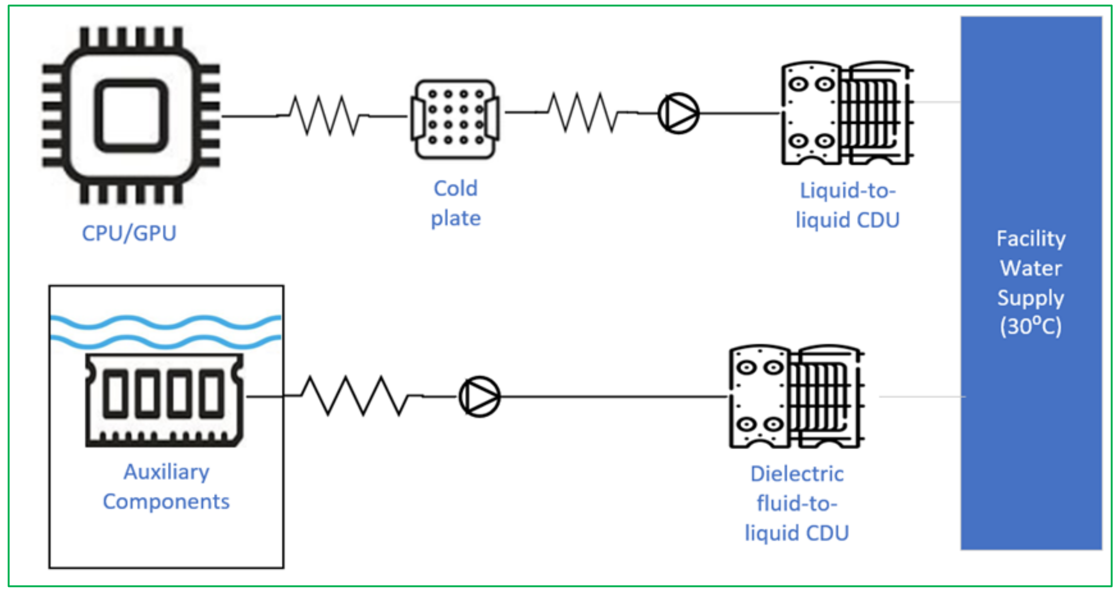

L-L CDU; Courtesy: Boyd


L-A CDU; Courtesy: Vertiv


RDHX; Courtesy: Data Center Dynamics




Single phase hybrid immersion servers; Courtesy: LiquidCool Solutions Inc



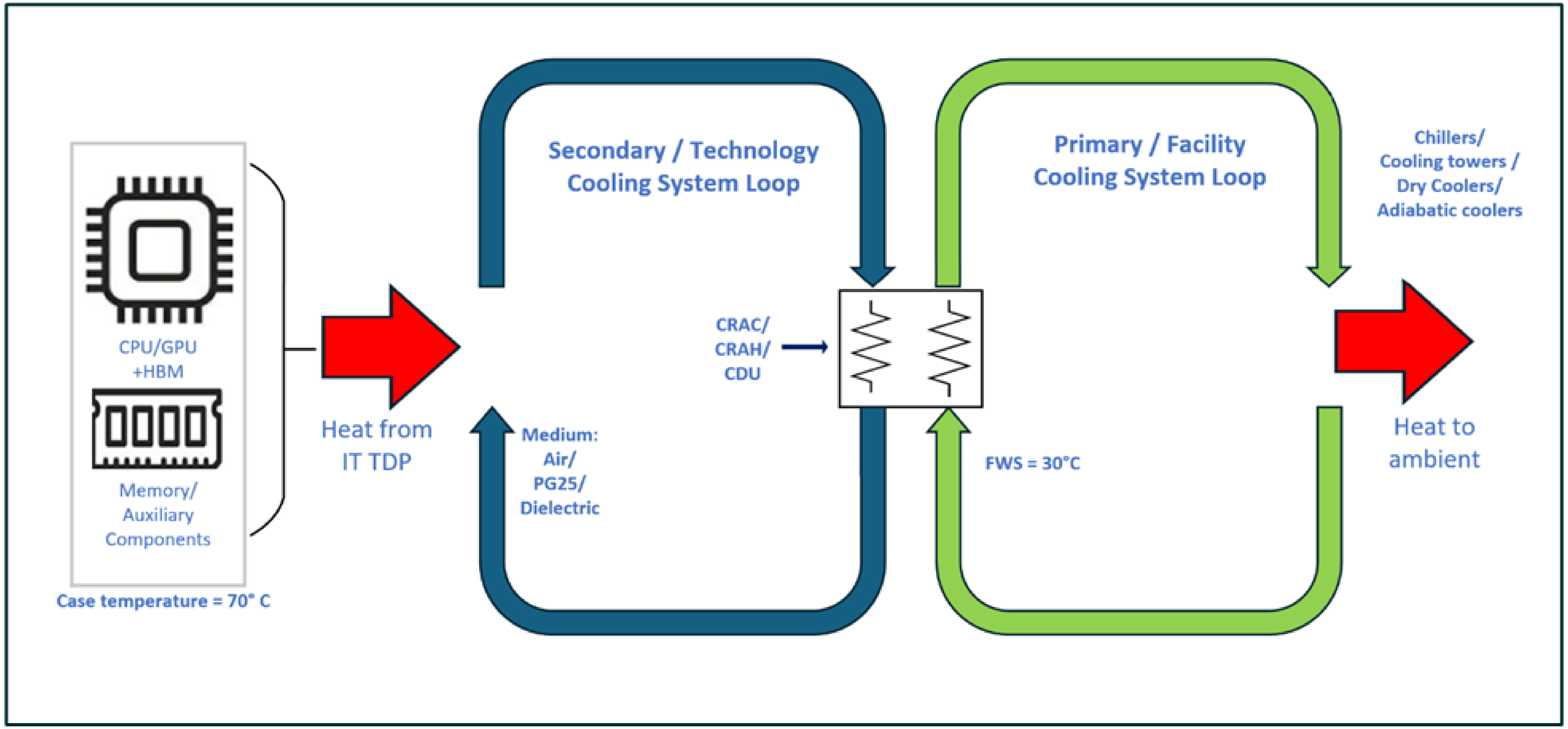
Single mode comparison table

Type	0. Full air-cooled		1. Full D2C Cold plate		2. Full Immersion	
Description	Air-cooled CRAH	Air-cooled CRAH + RDX	Single phase	Two phase	Single phase	Two phase
Retrofitting- relative comparison	need to upsize CRAH for higher density racks; need hot and cold aisle containment	existing CRAH can be used; but CDU and RDX needed	fans replaced by pumps; custom designs and increased cost	condenser needed in CDU; flow distribution needs careful design; potential use of PFAS and GWP refrigerants; higher cost than single phase	requires different architecture than racks; potential to get locked into a specific vendor	requires different architecture than racks; potential to get locked in with vendor; PFAS and GWP; significantly higher cost than strategy 2a
Reliability considerations (pros/cons)	• Mature and well-understood; easy to maintain • Redundancy (e.g., N+1 CRAHs, dual RDX fans) improves uptime • Less efficient at high rack densities • Susceptible to thermal hotspots if airflow is not well managed		• High thermal efficiency and predictable performance • Easier to control and monitor than two-phase systems • Requires robust leak detection and maintenance protocols	• Very high TDP chips can be cooled • More complex control and safety requirements • Sensitive to orientation and flow balance • Requires precise charge management and sealed systems	• Excellent thermal uniformity and simplicity • Reduced moving parts (no fans on servers) • Requires fluid monitoring and periodic replacement • Maintenance can be more involved (cleaning, fluid handling)	• Very high TDP chips can be cooled • Complex system design (condenser, vapor containment) • Requires precise thermal and fluid management • Limited field experience compared to other methods
Environmental Impacts	Low; may have higher land use		Low; traditional PG25 fluid is used	Medium; small refrigerant volume, GWP and PFAS concerns	Medium; uses dielectric oils which require proper recycling	High; limited refrigerant options and large refrigerant volume, GWP and PFAS concerns
Cost (relative to baseline), major cost components	(baseline)	$ RDX and CDU	$$$$ cold plates; fluid hose and fittings; CDU	$$$$$ cold plates; fluid hose and fittings; CDU; fluid	$$$$ Immersion bath; CDU	$$$$ Immersion bath; CDU; fluid

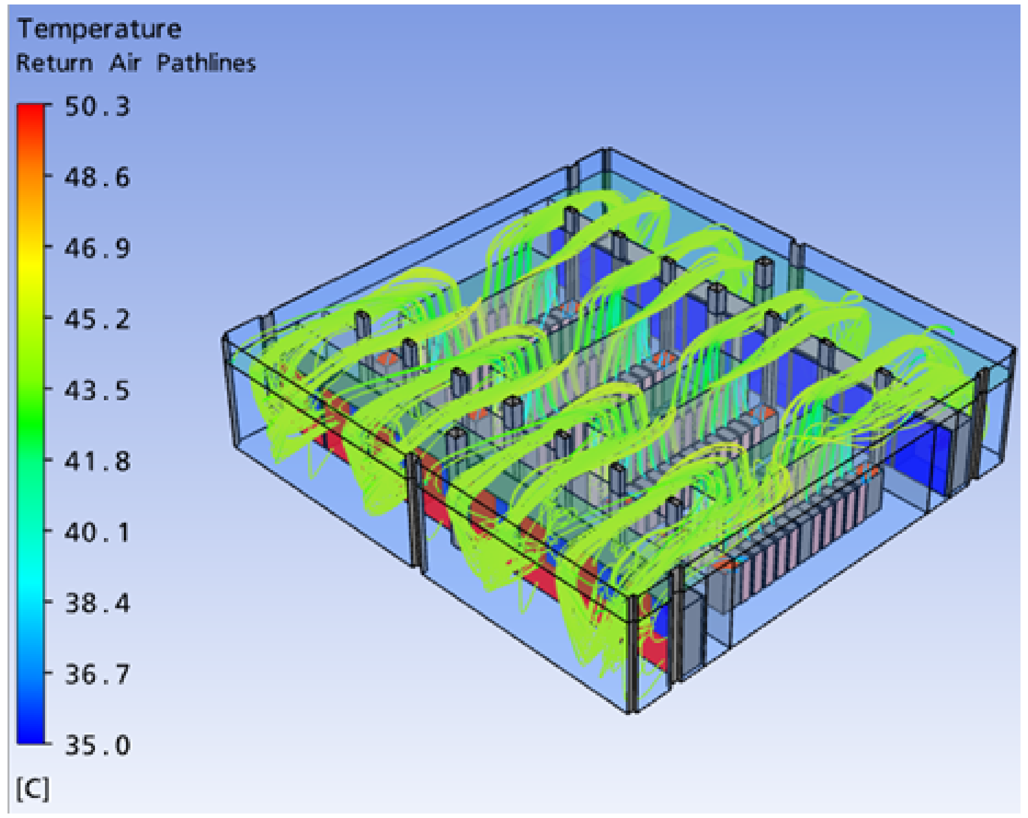
Hybrid mode comparison table

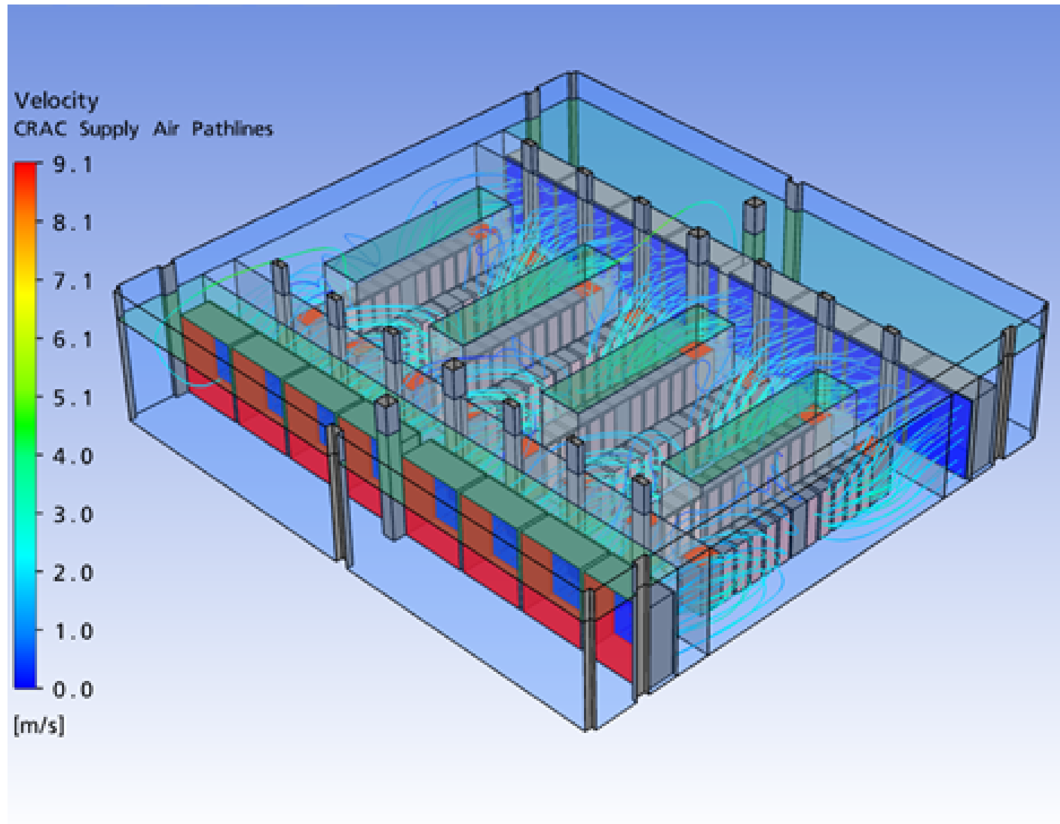
Type	3. Cold plate CPU/GPU + CRAH cooled AUX		4. Cold plate CPU/GPU with liquid-to-air CDU + CRAH cooled AUX	5. Cold plate CPU/GPU + CRAH with RDX for AUX		6. Cold plate CPU/GPU + Immersion cooling AUX	
Description	Direct Liquid Cooling (DLC) in cold plate	2-phase flow (2PF) in cold plate	AUX Air-cooled with CRAH	Direct Liquid Cooling (DLC) in cold plate	2-phase flow (2PF) in cold plate	Direct Liquid Cooling (DLC) in cold plate	2-phase flow (2PF) in cold plate
Retrofitting- relative comparison	CPU/GPU are DLC while rest are air cooled with fans; existing CRAH can be used; easier retrofit than strategy 1a	Condenser needed in CDU; flow distribution needs careful design; potential use of PFAS and GWP refrigerants; higher cost than strategy 1a	need to upsize CRAH; more expensive than strategy 0a; however, rack densities can be increased and there is no need for liquid lines into the datacenter	promising retrofit scenario— existing CRAH can be used; but CDU, cold plates and RDX needed	condenser needed in CDU; flow distribution needs careful design; PFAS and GWP of refrigerants; needs separate CDU for RDX; higher cost than strategy 5a	can be fitted into a rack architecture, however serviceability low; potential to get locked into a specific cooling solution and provider	rack architecture, serviceability low; potential to get locked into a specific vendor; PFAS and GWP refrigerants; significantly higher cost than strategy 6a, needs two CDUs
Reliability considerations (pros/cons)	• Uses traditional air-cooling components with higher level of understanding • Requires careful thermal zoning and airflow management • Dual maintenance protocols (liquid and air systems)		• Liquid is confined in a small loop, easy management • Requires robust leak detection and flow assurance • Hybrid airflow and liquid loop coordination is critical	• Rear door HX improves rack-level cooling redundancy but multiple cooling paths increase complexity • Requires coordinated control across three subsystems		• Immersion cooling offers uniform thermal management but fluid monitoring and maintenance are essential • Integration of two liquid systems requires careful isolation and control	
Environmental Impacts	Low; traditional PG25 fluid is used	Medium; small refrigerant volume, GWP and PFAS concerns	Low, may have higher land use	Low; traditional PG25 fluid is used	Medium; small refrigerant volume, GWP and PFAS concerns	Medium; uses dielectric oils which require proper recycling	High; limited refrigerant options and large refrigerant volume, GWP and PFAS concerns
Cost	$$ Cold plates on only CPU/GPU; liquid-to-liquid CDU	$$ Same as 3a	$$$ Cold plates on CPU/GPU; liquid-air CDU	$$$$ Cold plates on only CPU/GPU; RDX; liquid-to-liquid CDU	$$$$ Same as 5a	$$$$$ Costs of 3a + cold plates on CPU/GPU	$$$$$$ In addition to costs in 6a, two different fluids needed

Energy analysis- modeling system boundary

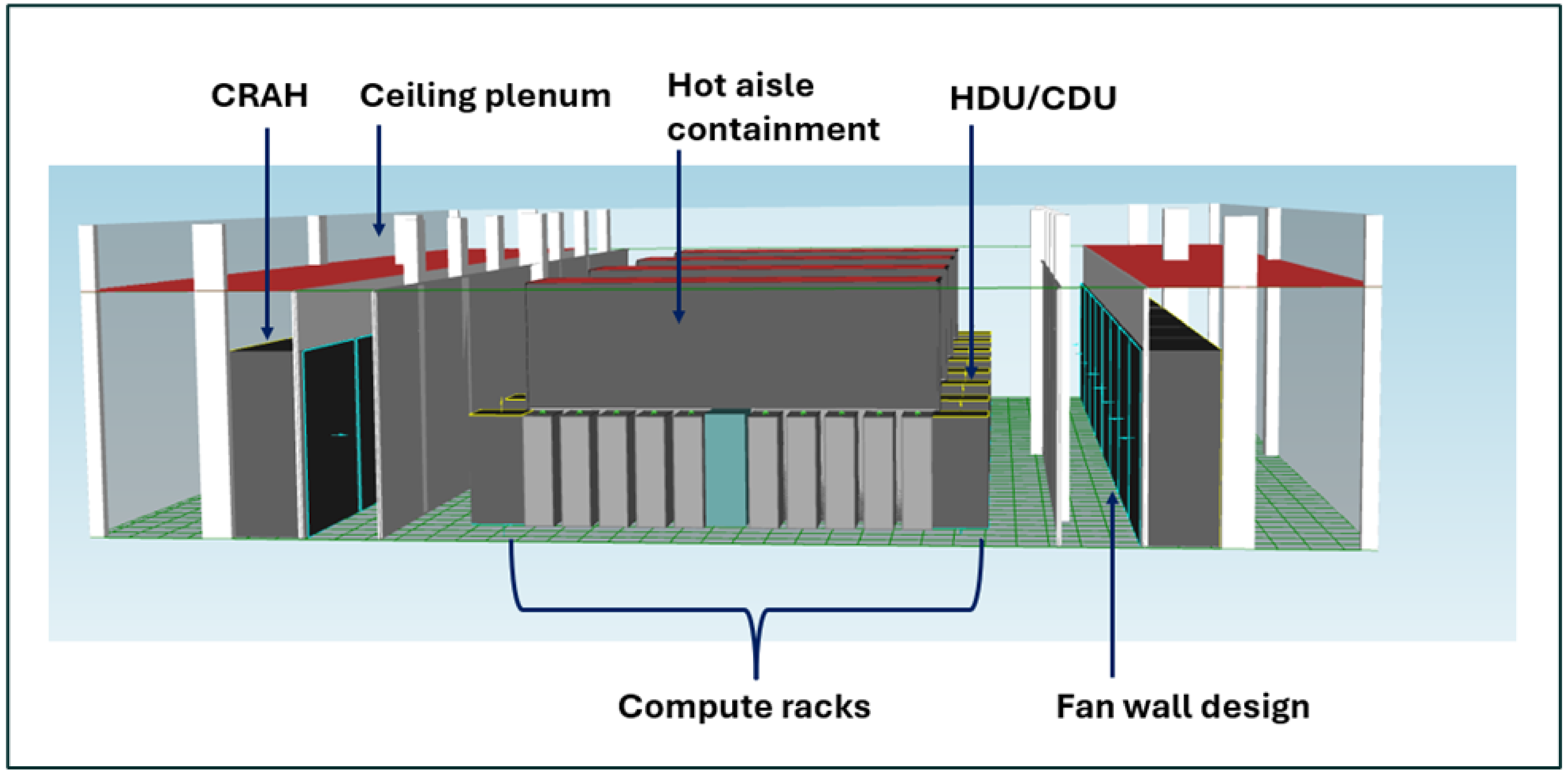


Energy analysis- Baseline Air-cooled DC simulations





- 80 racks @ TDP of 44 kW ea
- hot aisle containment
- 16 CRAHs each with capacity of 228.6 kW (65 Refrigerant tons)
- chilled water entering at 30 ºC and a leaving air temperature of 35 ºC.



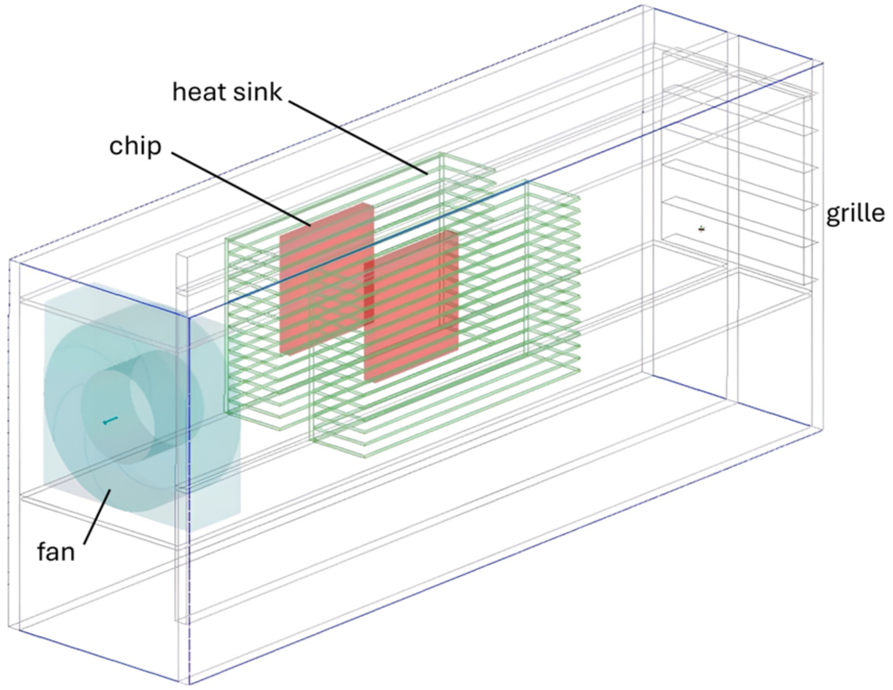
Coolsim

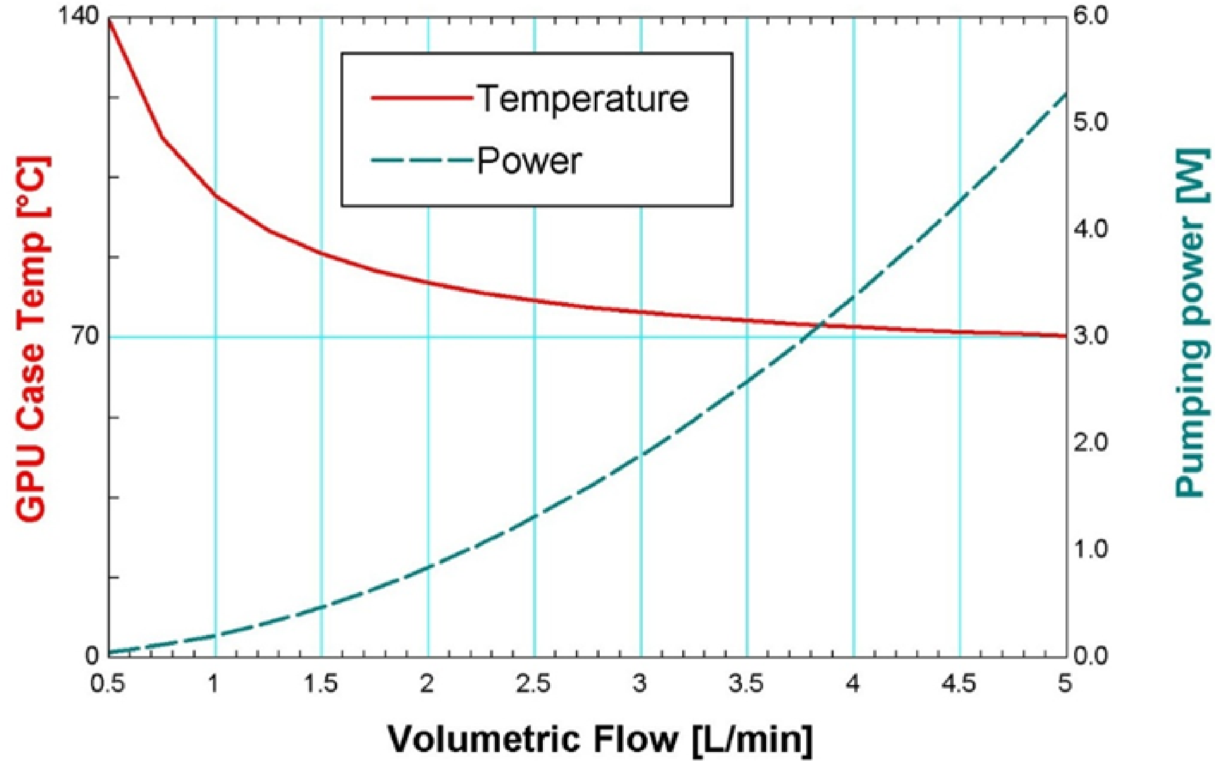
Energy analysis- Component models

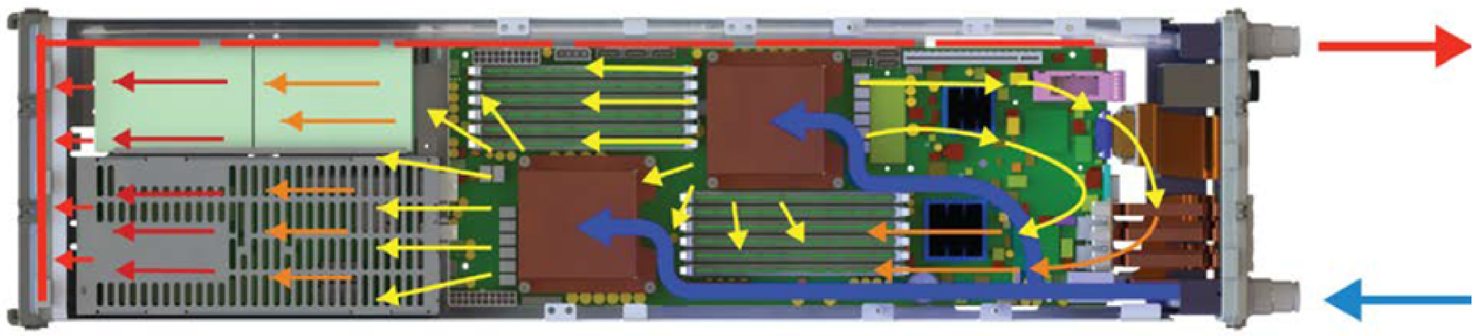
Component models developed for

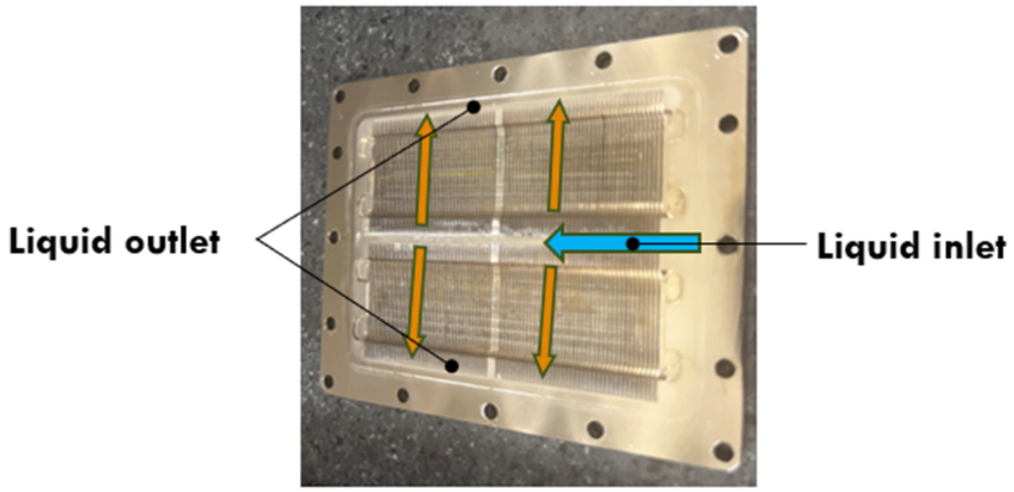
1. Cold plate (parallel microchannels) for PG25 (D2C) and Opticool 8725 (hybrid immersion)
2. Auxiliary air-cooled - finned heat sink + fan – 0.005 kW (172 CFM)/kW TDP of auxiliary components- Ansys Icepak
3. Auxiliary- immersion cooled- finned heat sink
4. L-L CDU- pumping power- plate heat exchanger + lines (assumed 2x GPU cold plate) + GPU cold plate
5. L-A CDU- pumping power above + fan power through liquid-air HX (A frame multiport microchannel with 4 rows)- CoilDesigner^R
6. RDX- fan power through liquid-air HX (multiport microchannel with 2 rows)- CoilDesigner^R
7. CRAH model for auxiliary component heat dissipation –Coolsim

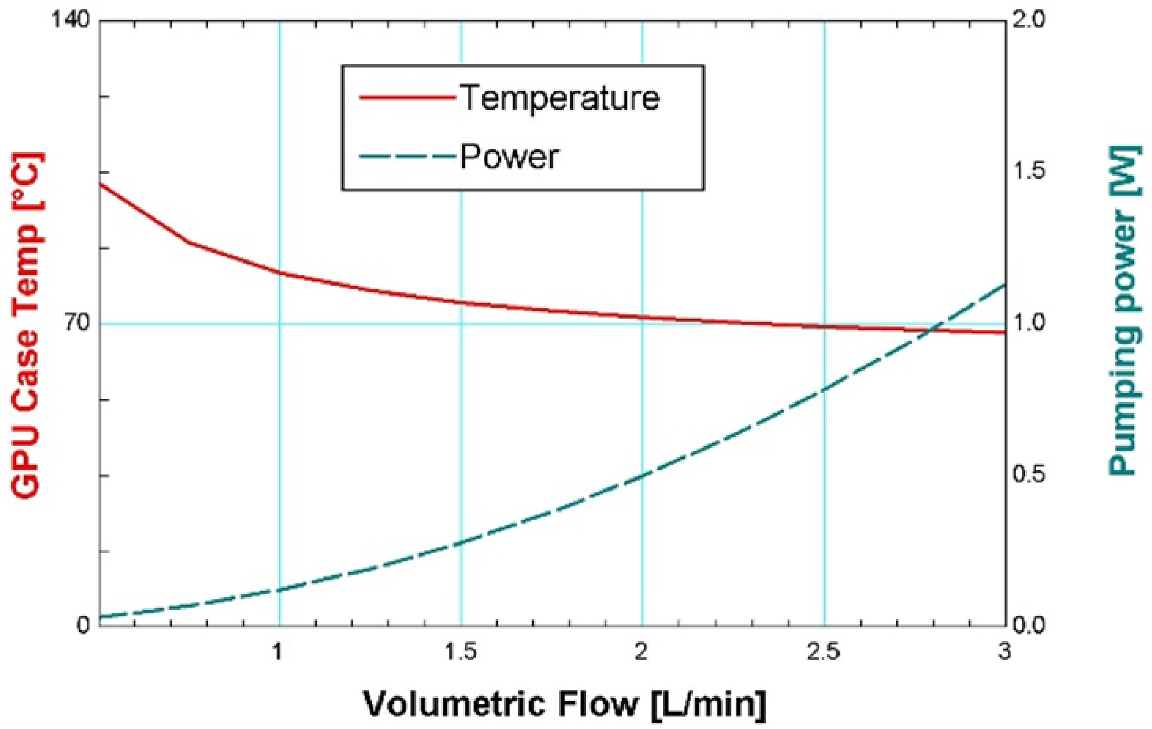
Single phase immersion and two-phase not modelled











Components models used for each cooling strategy

Cooling Technology	Model 1	Model 2	Model 3	Model 4
Traditional CRAC/CRAH air-cooling	Datacenter room and CRAH simulations	Server fan with heat sink analysis		
D2C with cold plates	D2C cold plate simulation for all components	Liquid-liquid CDU		
Full-immersion cooling	Not modeled. Manufacturer and published research data used for comparison with other technologies in analogous conditions			
D2C cold plate + traditional air-cooling	D2C cold plate simulation for GPU and CPU	Liquid-to-liquid CDU for GPU and CPU	Datacenter room and CRAH for AUX	Server fan with heat sink for AUX
D2C cold plate + liquid-to-air CDU for traditional air-cooling	D2C cold plate simulation for GPU and CPU	Liquid-to-air CDU for GPU and CPU	Datacenter room and CRAH for AUX	Server fan with heat sink for AUX
D2C cold plate + Rear door HX for traditional air-cooling	D2C cold plate simulation for GPU and CPU	Liquid-to-liquid CDU for GPU and CPU	RDX model for AUX	Server fan with heat sink for AUX
D2C cold plate + immersion cooling	D2C cold plate simulation for GPU and CPU with dielectric fluid	Dielectric oil flow through AUX with heat sinks	Liquid-to-liquid CDU	

Combined secondary loop estimates

Cooling mode	$SCCP_{ratio}$ (%)	$C_{density}$ (kW/m^2)
Traditional CRAH air-cooling (baseline)	2.00	5.4
D2C with cold plates	0.14	82.7
Full Immersion tank	1.08*	16.3**
D2C cold plate + traditional air-cooling	0.65	18.9
D2C cold plate and liquid-to-air CDU + traditional air-cooling	4.14***	5.4
D2C cold plate + RDX for traditional air-cooling	0.89	18.9
D2C cold plate + immersion cooling	1.08*	63.4

*Based on CDU pump power with secondary side dielectric oil flow and GPU throttled to 1100 W.
** Calculated from product brochure (Green Revolution Cooling 2025) using FWS of 32°C
*** GPU needed to be throttled to 1300 W to maintain case temperature of 70°C

$$SCCP_{ratio} = \frac{Secondary\ loop\ input\ electrical\ power}{Total\ IT\ compute\ power} \qquad C_{density} = \frac{IT\ compute\ power\ supported}{Total\ floor\ area\ of\ datacenter}$$

Key Takeaways

New builds- prioritize all-D2C liquid cooling as the baseline strategy for AI and HPC workloads.

Retrofits- adopt hybrid approaches tailored to facility constraints: CRAH-assisted D2C where auxiliary loads are manageable, RDX where CRAHs are insufficient, and liquid-to-air CDUs where liquid distribution is infeasible.

Lower-density or modular deployments- consider immersion cooling, with hybrid immersion-D2C for higher TDP chips.

Utilities and policymakers- align incentive programs with technologies that deliver low $SCCP_{ratio}$, higher facility water supply temperature, and higher rack density, and establish standards that reduce the risk of integrating two-phase cooling technologies.

Market- encourage vendor collaboration on interoperability, expand service models to address operational concerns, and accelerate the transition from hybrid to fully liquid systems in high-density environments.

Check out our research highlights