Next-Generation Cooling For NVIDIA Accelerated Computing

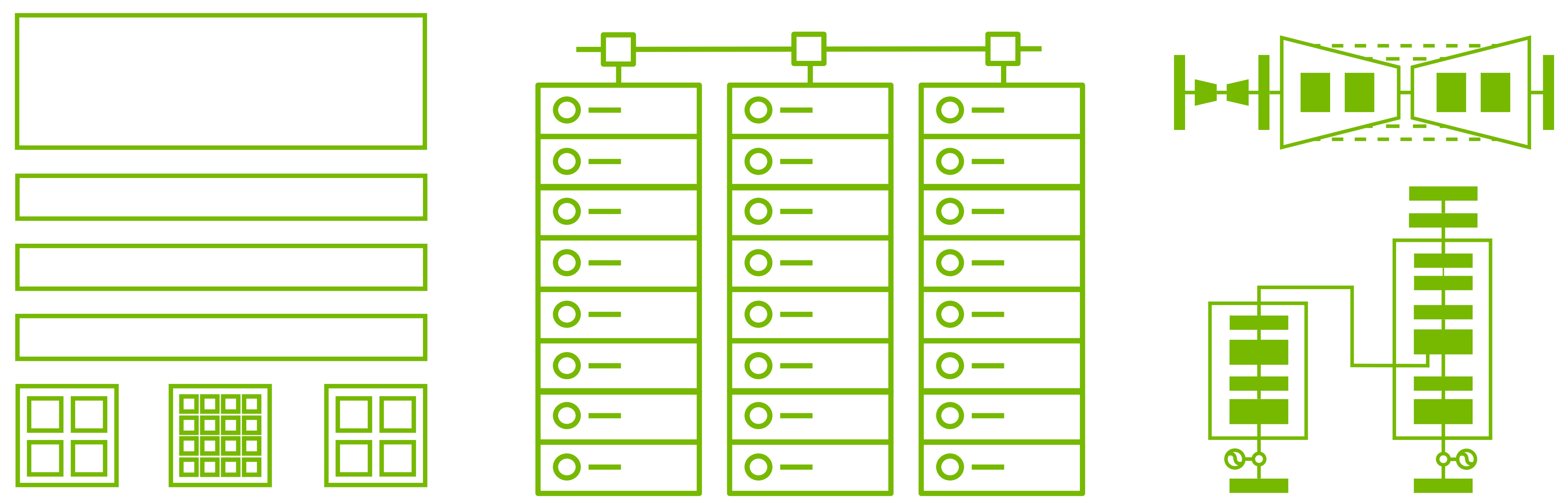
Ali Heydari, Director, Data Center Cooling & Infrastructure

Aug 25, 2024

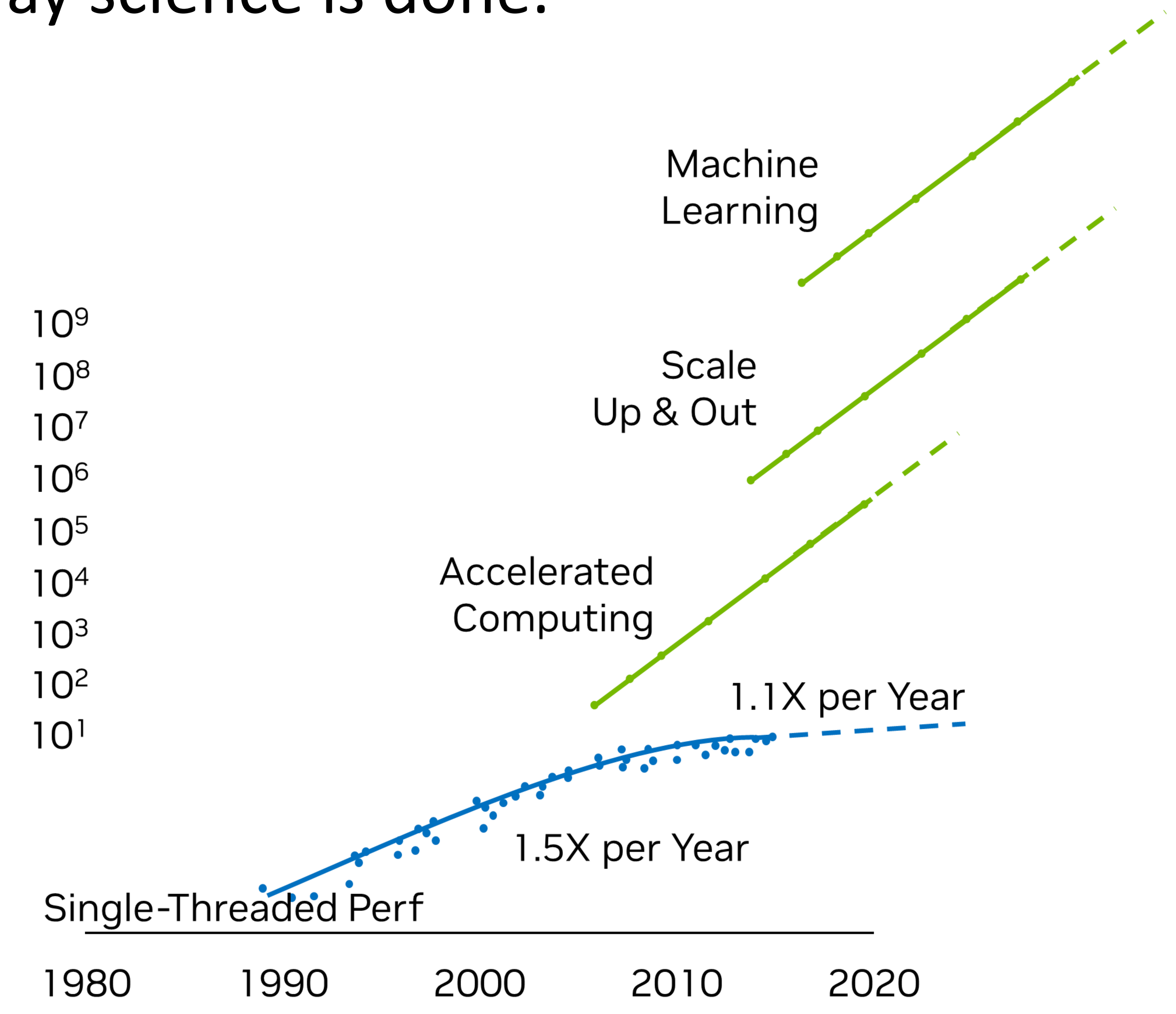
Era of AI

- The 3 connected dynamics – AI, Accelerated Computing and Data Centers are revolutionizing the way science is done.

ACCELERATED COMPUTING DATA CENTERS AI

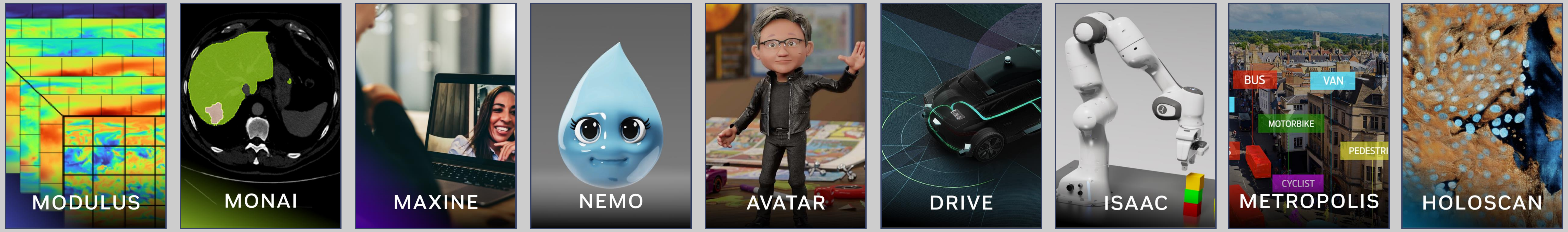


The 3 connected dynamics to give Million-X leap in computations



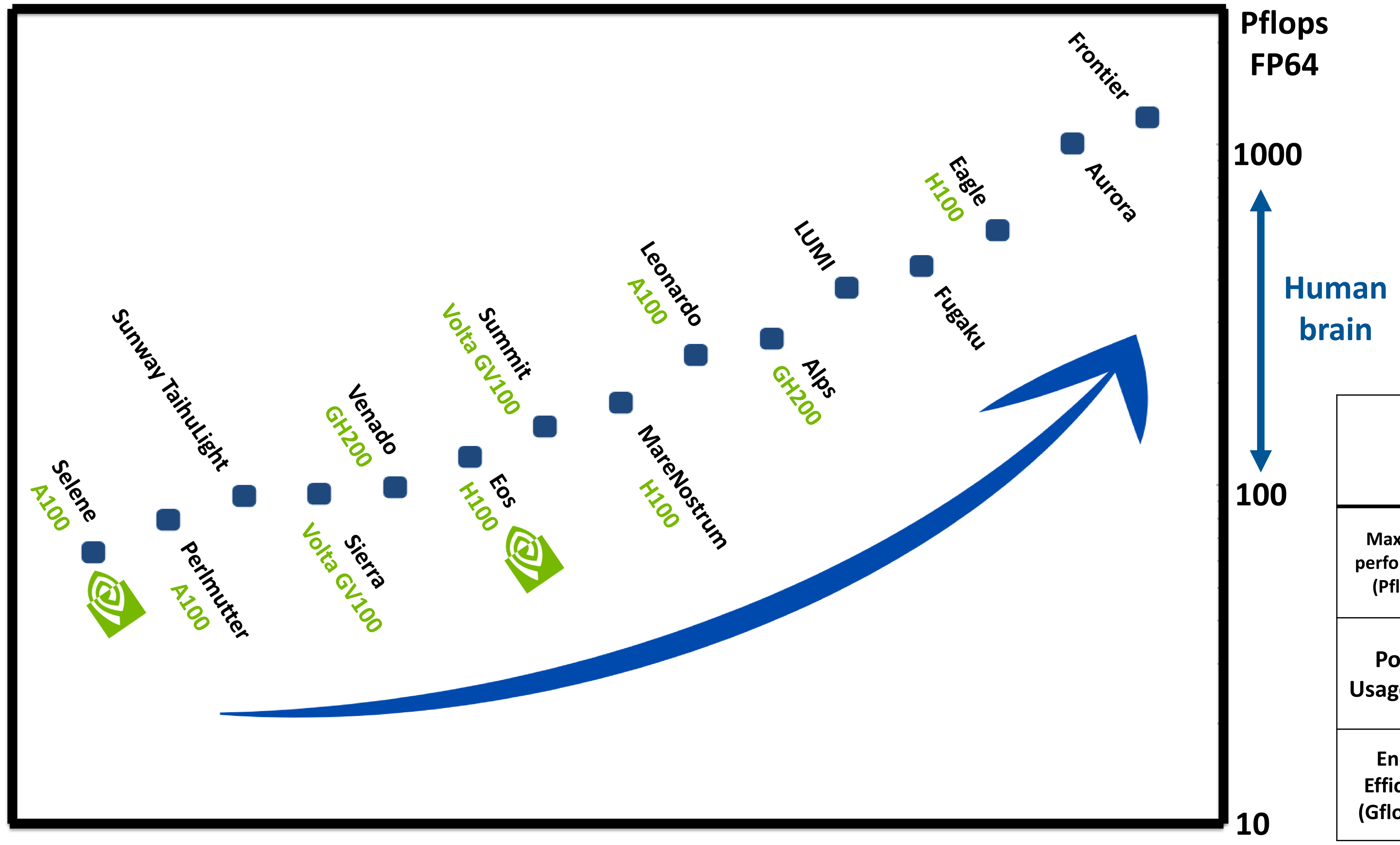
- The million-X leap can solve problems previously impossible, like in the areas of computational fluid dynamics, data science, climate, imaging, computational biology, computational lithography, quantum physics and many more.

Application Frameworks



Factories for Generative AI: Data Centers

- Generative AI models to drive the next industrial revolution
- Large AI factories required for:
  - Hosting multiple GPUs with high rack density
  - Providing higher throughput, low latency and running larger AI models in real-time
  - Driving growth in tokens/sec & revenue
- Complexity & scale of modern AI models require significant computational power.
- Multi GPU clusters in AI factories are the future for producing AI tools.
- Liquid Cooling to enable NVIDIA's next-gen chips like **Blackwell** push boundaries in both training and inference



	Human Brain	Frontier
Maximum performance (Pflop/s)	~ 1000	1206
Power Usage (kW)	~0.015	22,786
Energy Efficiency (Gflops/w)	~66,666,666	52.93

Building Bigger AI factories

Selene 2021	EOS 2023	Next AI Factory
4,480 A100 GPUs 3 EF AI Compute 112 TB/s BW	10,752 H100 GPUs 43 EF AI Compute 1100 TB/s BW	32,000 GPUs 645 EF AI Compute 58,000 TB/s BW

Productivity

Entertainment

Travel

Customer Service

Personal Assistant

Finance

Smart Home

Image and Video editing

Social Media

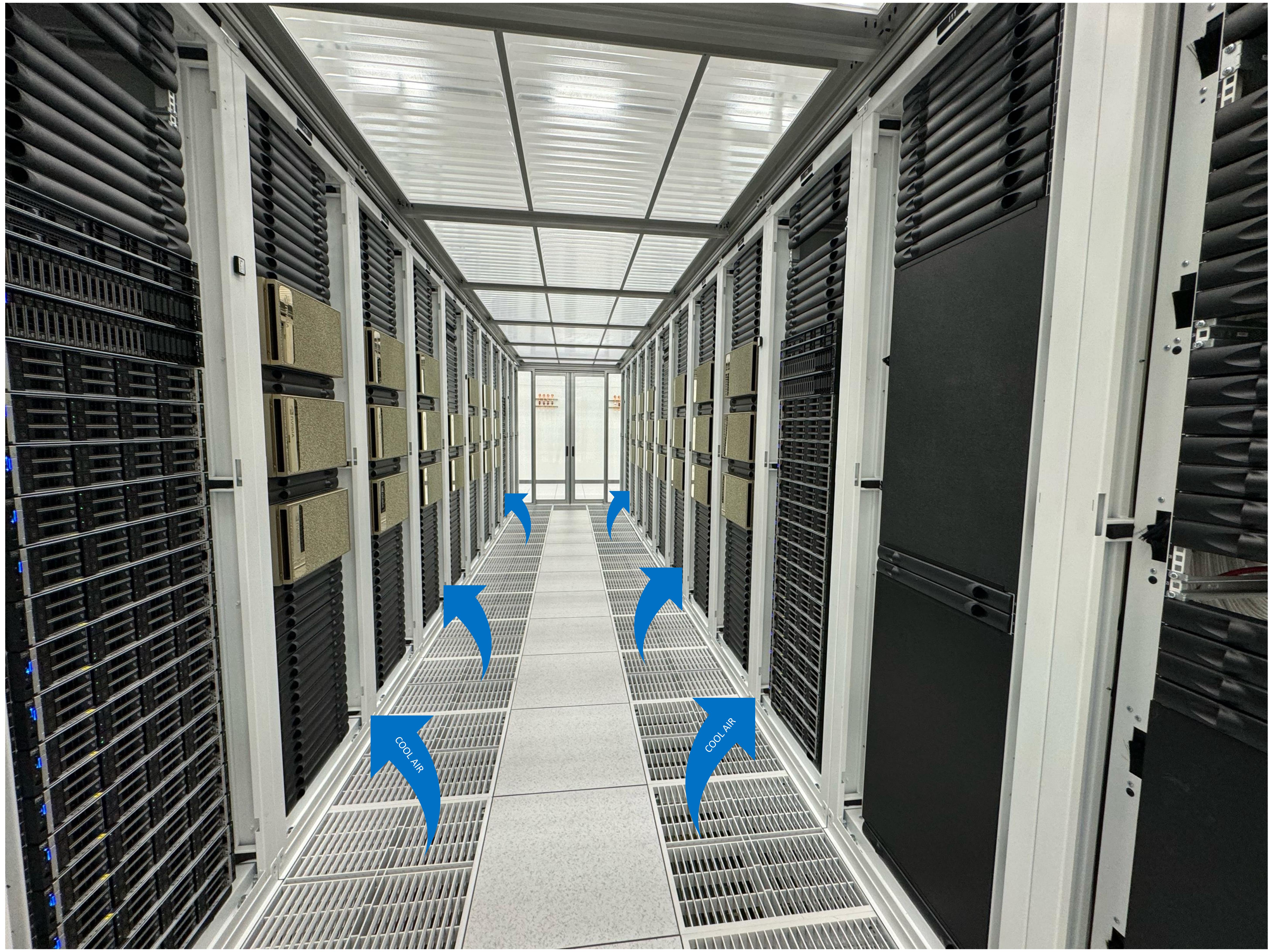
Healthcare

Language and Translation

E-commerce

EF- ExaFLOPs
BW- Bandwidth

Hybrid Data Centers (Air + Liquid Cooling)



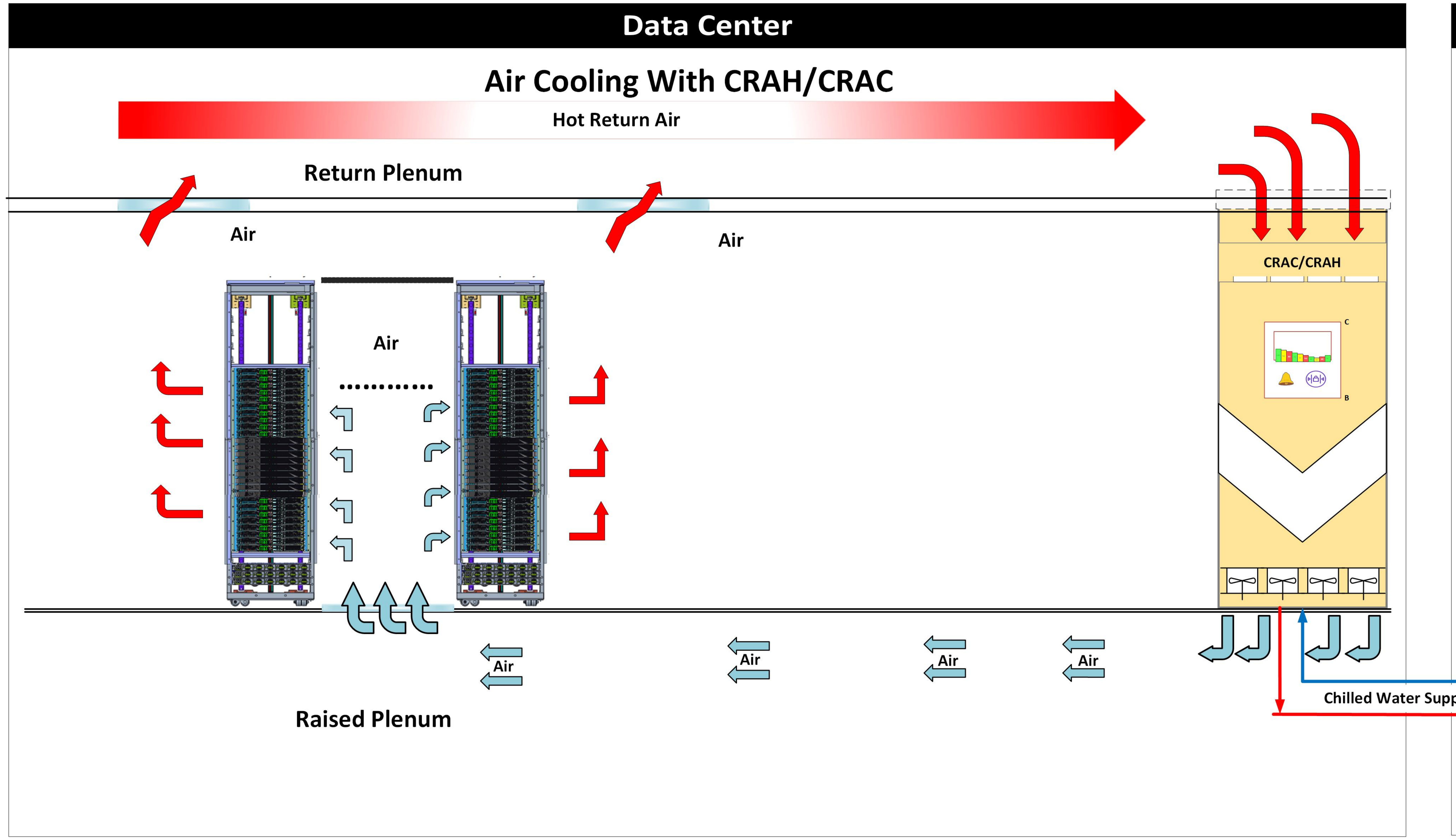
- Cold aisle containment with supply of air through perf tiles in the cold aisles, return of the air to the CRAHs in the hot aisle
- Liquid supply (PG25) to the rack manifolds using QDs to servers
- Flow distribution through the data center using row manifolds and headers from Cooling Distribution Units (CDUs) and Heat Exchanger Sidecars.

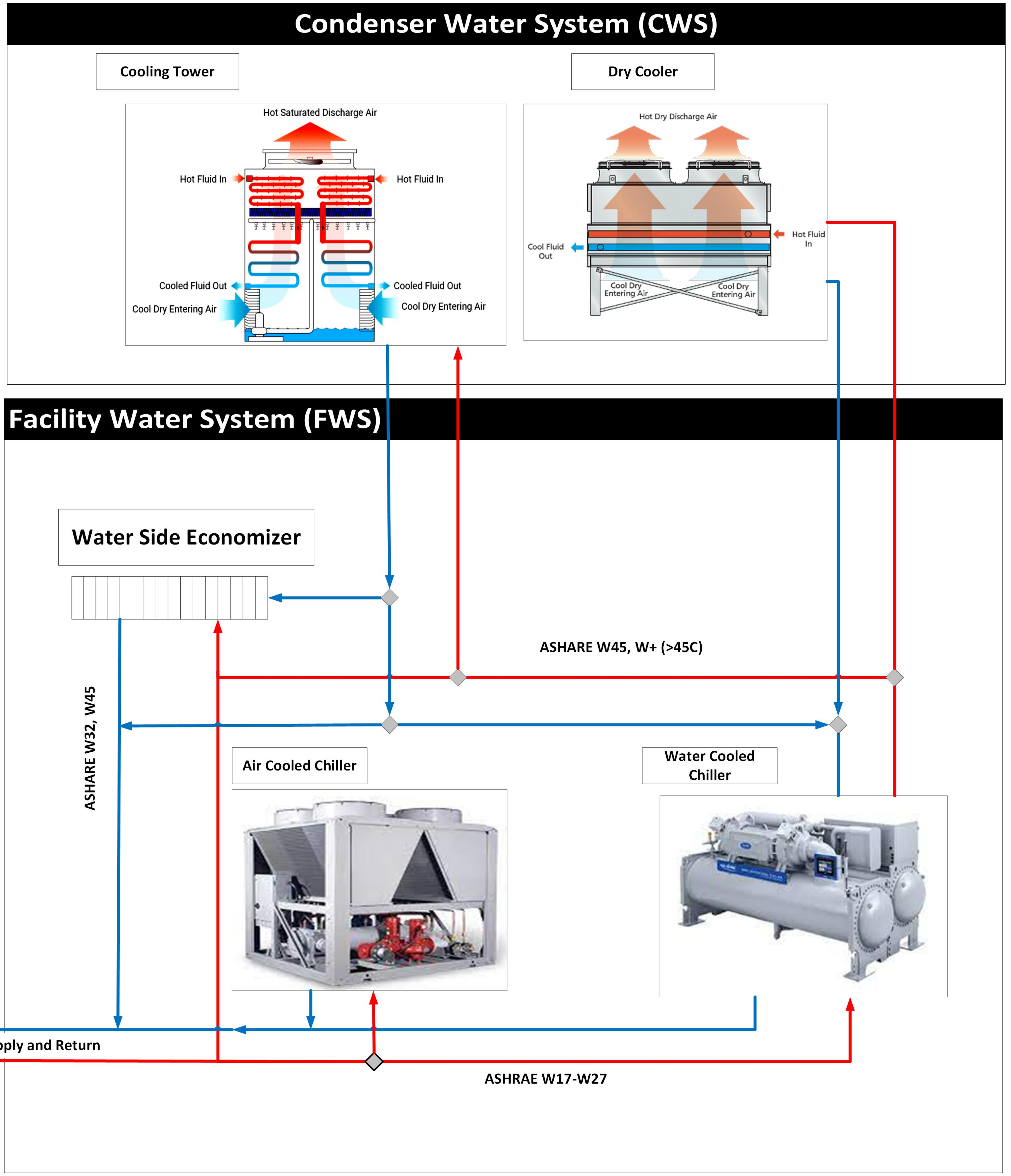
NVIDIA

Schematic of Data Center Cooling Technologies
Air Cooling

Air cooling with CRAH/CRAC

- Room-based cooling, suitable for low density racks
- Heat picked by air is transferred to facility chilled water in CRAH.
- Raised/ slab floor.
- Cold aisle or hot aisle containment.

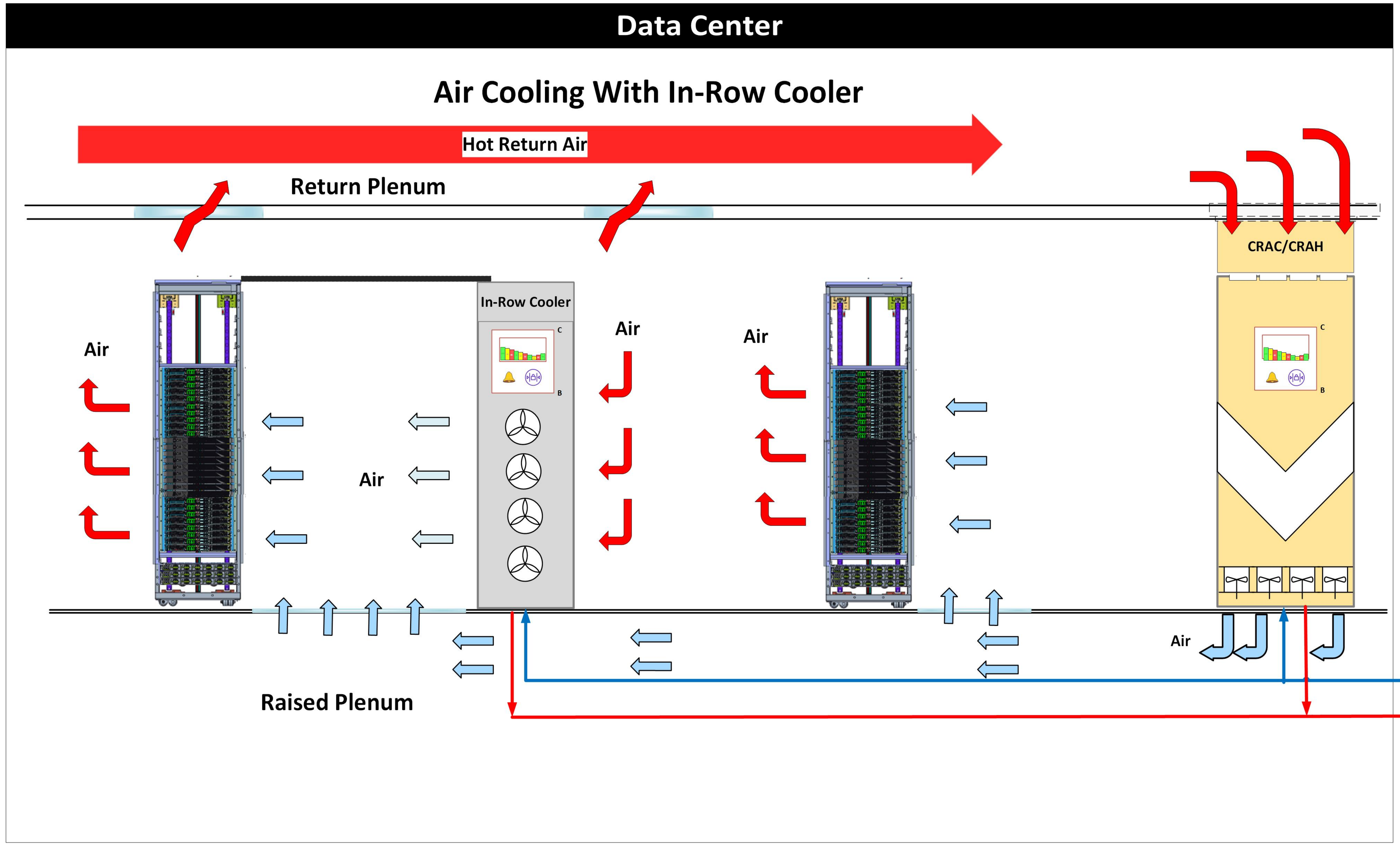


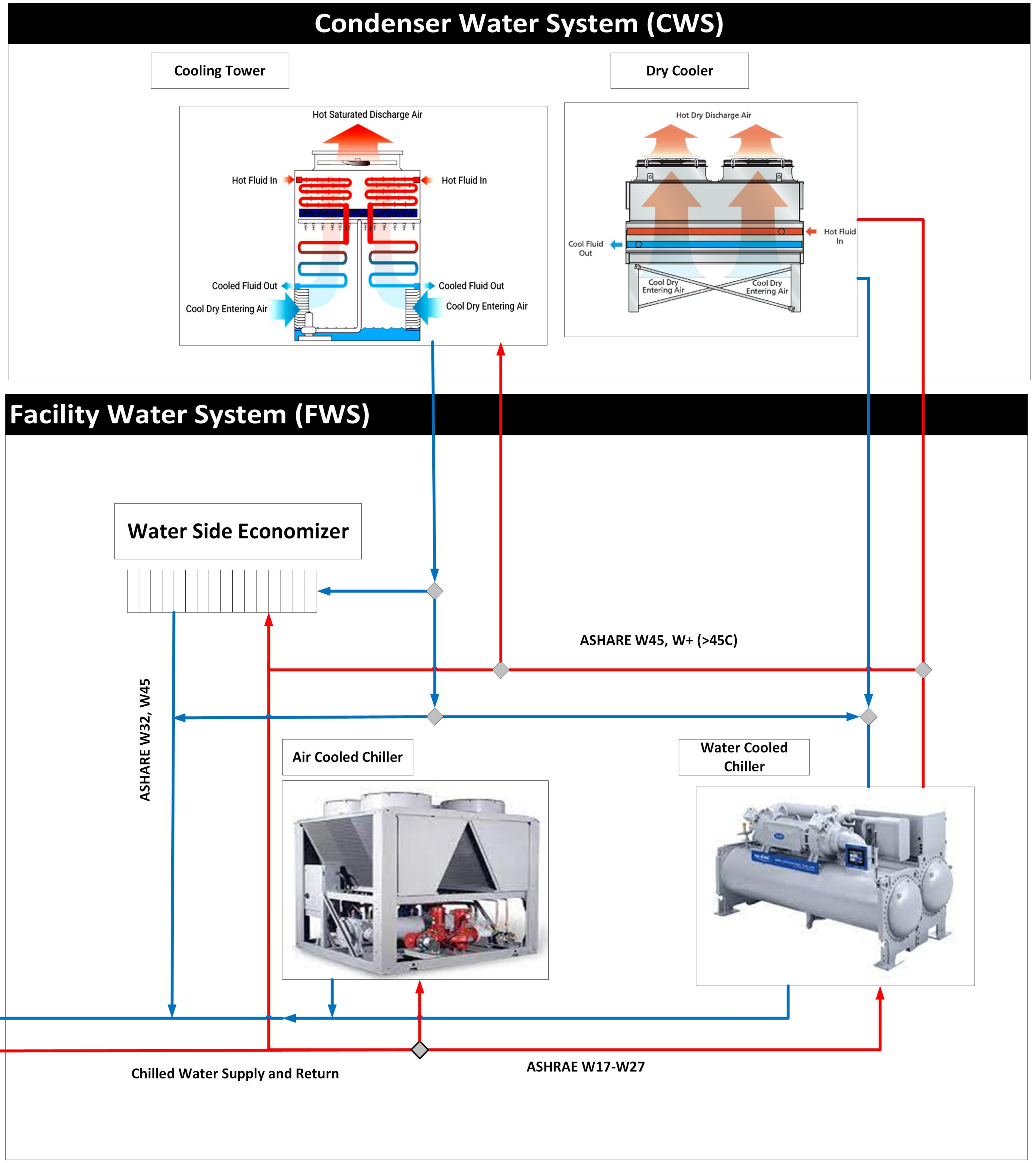


Schematic of Data Center Cooling Technologies
Air Cooling

Air cooling with In-Row Coolers

- Aisle-based cooling, suitable for medium density racks
- Heat picked by air is transferred to facility chilled water through In-Row Coolers and CRAH unit.
- Raised/ slab floor.
- Cold aisle or hot aisle containment

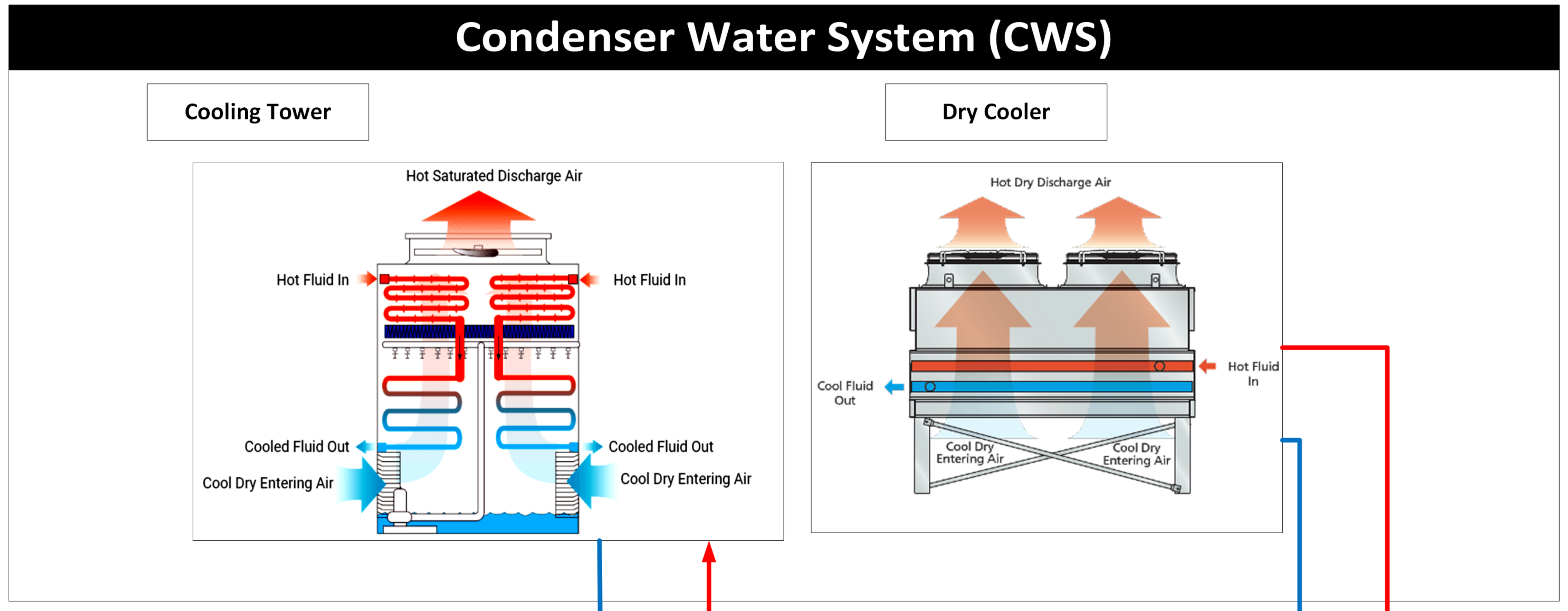




Schematic of Data Center Cooling Technologies
Air Cooling

Air cooling with Rear Doors Heat Exchangers (RDHX)

- Rack-based localized cooling; suitable for medium density racks.
- Heat picked by air and transferred to facility chilled water via RDHX.
- Raised/ slab floor.
- Could be used as standalone or as assisted cooling in conjunction with CRAH units.



Schematic of Data Center Cooling Technologies
Air and Liquid Cooling

Liquid cooling with Liquid to Air Side Car (L2A)

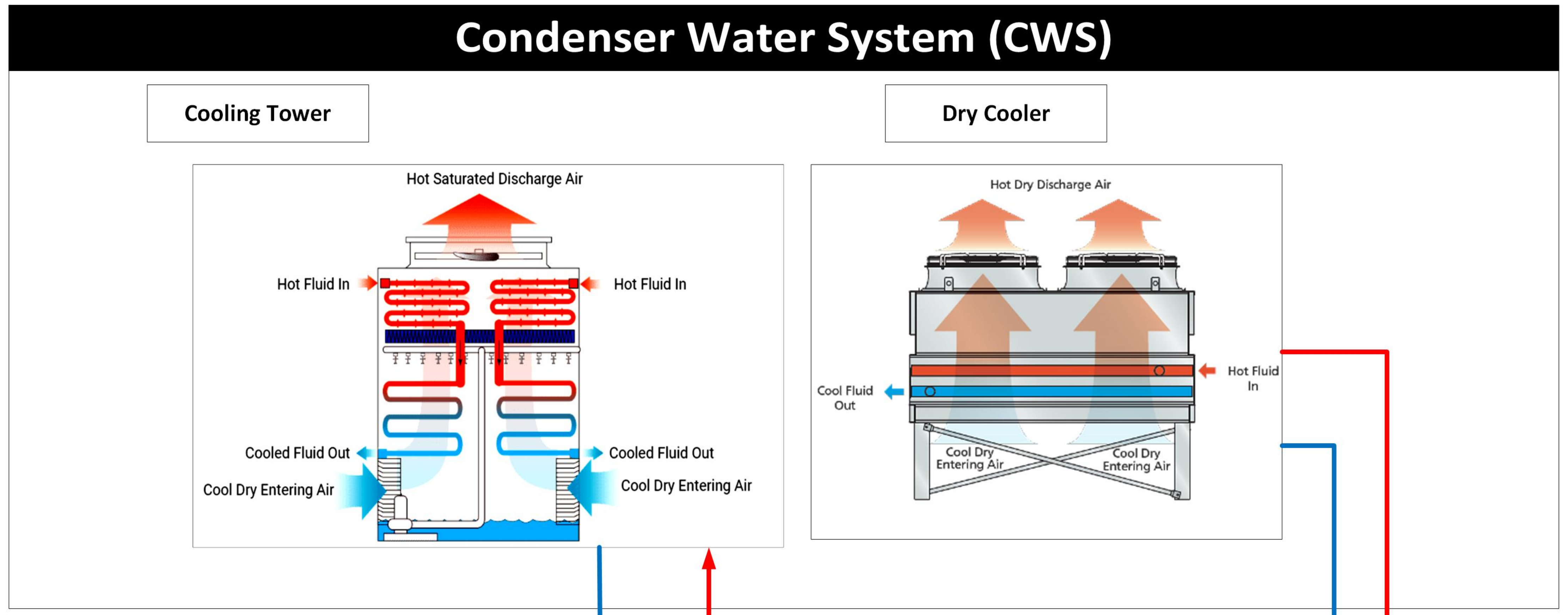
- Aisle-based hybrid of air/liquid cooling, suitable for high density racks
- Air Assisted Liquid Cooling suitable for legacy air-cooled data center.
- No additional liquid cooling Infrastructure required.
- Transitional solution - Limited cooling capabilities.

NVIDIA

Schematic of Data Center Cooling Technologies
Air and Liquid Cooling

Liquid cooling with Liquid-to-Liquid CDU (L2L)

- Aisle-based hybrid of air/liquid cooling, suitable for high density racks
- Row-based cooling distribution units can remove MWs of IT heat in a 4’x4’x6’ CDU unit.
- Additional liquid cooling Infrastructure required.
- Transitional solution – Limit of single-phase liquid cooling.



NVIDIA

Liquid-to-Air Cooling Solution

The need for L2A CDUs

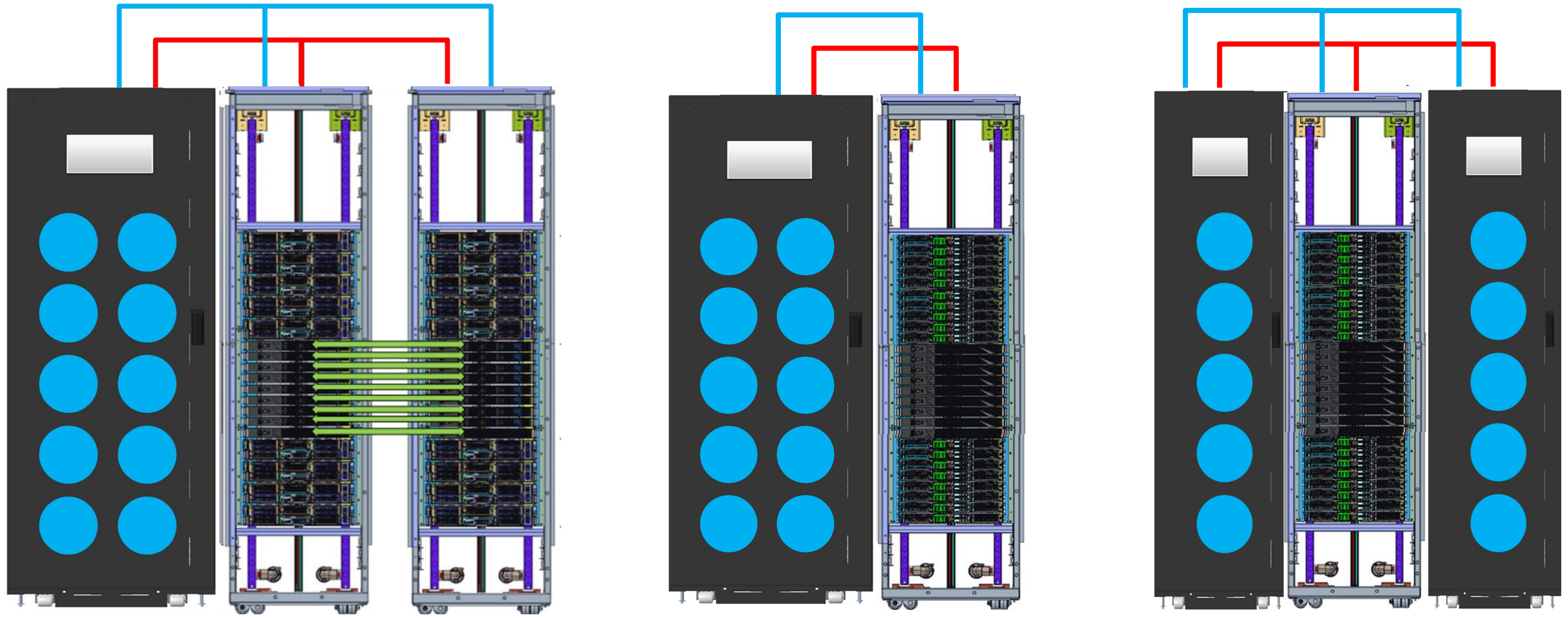
- Intermediate solution for legacy air-cooled data centers that utilizes existing infrastructure to deploy liquid cooled IT.
- Liquid-to-Air cooling cooling distribution units provide ~60kW cooling in a 2’x4’x6.5’ space has similar characteristics as IRCs.
- Power consumption of ~4% nominal cooling capacity.

L2A CDUs Evaluation Criteria

- Emulators design/build to simulate LC servers.
- Buildup of test lab-controlled environment
- Performance evaluation of L2A CDUs against product requirement.
- Digital Twin buildup of L2A CDUs with CFD/FNM

CDU-Rack Configurations

- Single rack-width vs double rack-width L2A CDUs
- Cooling capacity (kW) vs pumping capacity (LPM/kW) constraints. CFM/kW constraints, typical cooling capacity 0f ~60-120 kW per rack footprint.
- Data center air flow distribution balancing and CRAH return air temperature limitations.



Direct to Chip Liquid-to-Liquid CDU Solutions

L2L Cooling System Configuration

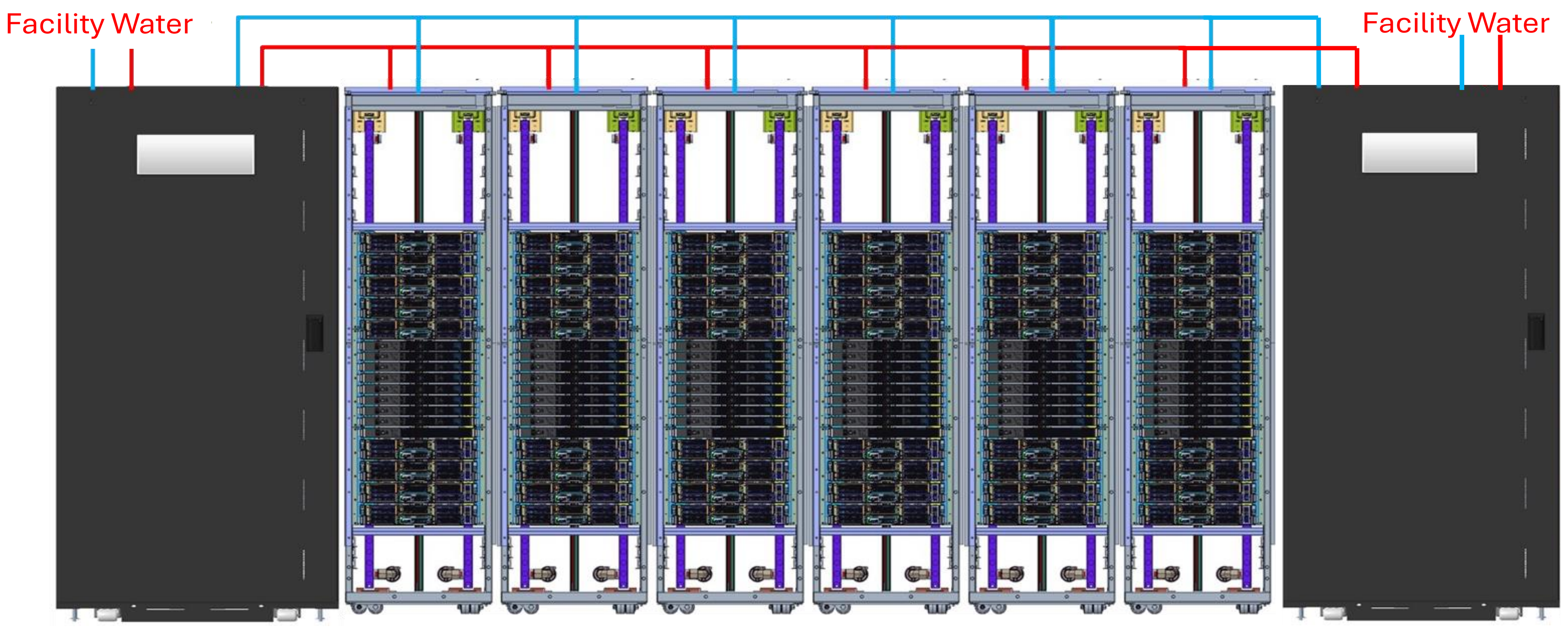
- Liquid-to-liquid cooling cooling distribution units provide ~2MW cooling in a 4’x4’x6.5’, in 2.75x less space produces 6.5x more cooling than CRAHs
- Power consumption of ~1% nominal cooling capacity

L2L CDUs Evaluation Criteria

- Emulators designed/built to simulate LC servers.
- Buildup of test lab-controlled setup.
- Performance evaluation of L2L CDUs.
- Physics aware Digital Twin model of L2L CDUs.

CDU-Rack Configuration

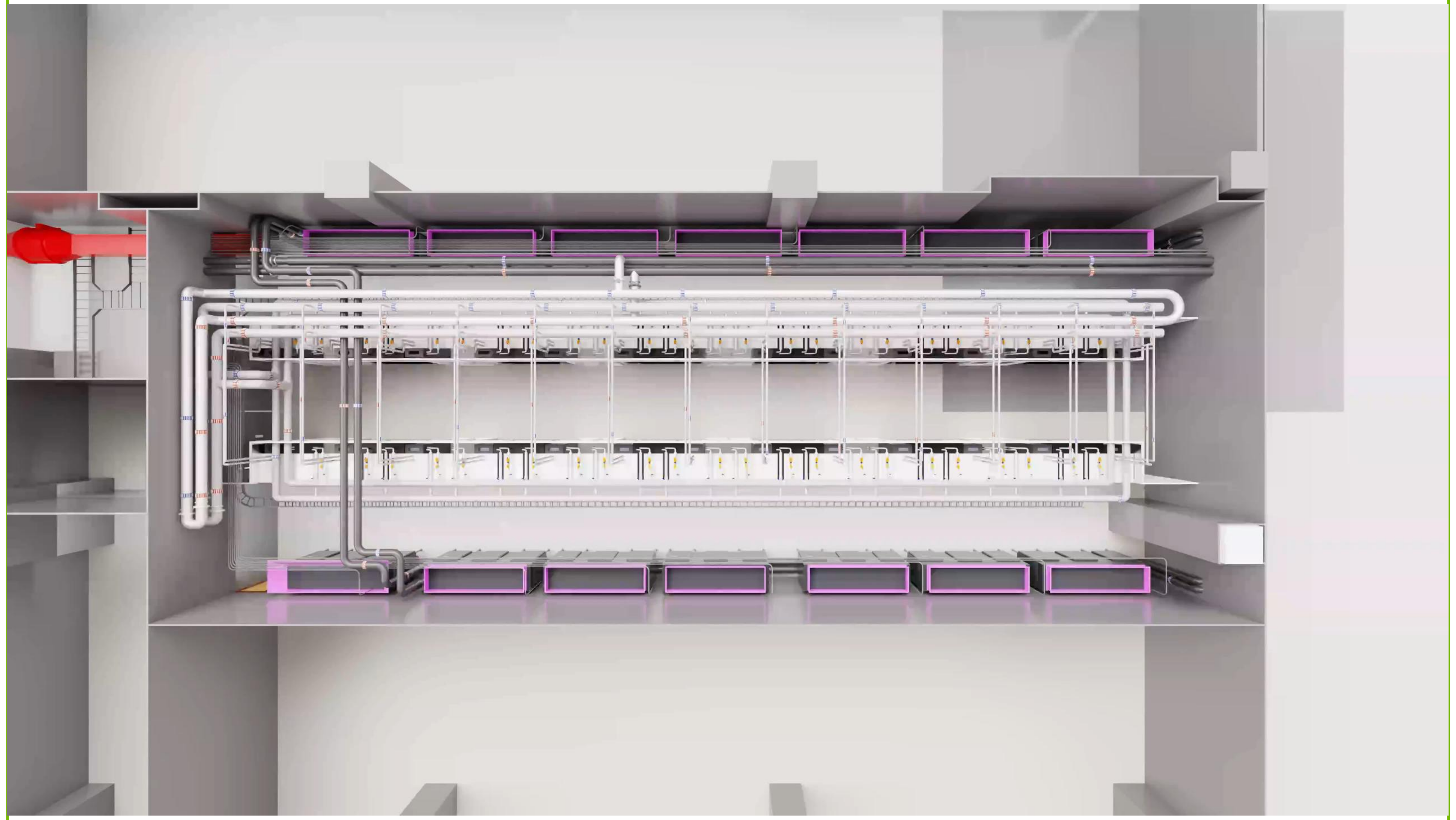
- L2L CDUs availability: Rackmount (60 KW – 200 KW), mid-range In-row (400 KW - 800 kW), high-range In-Row (1000 KW - 2400 kW) CDUs
- Rack-based fluids flow distribution controls, capability to establish pressure differential controls for each liquid cooled rack.



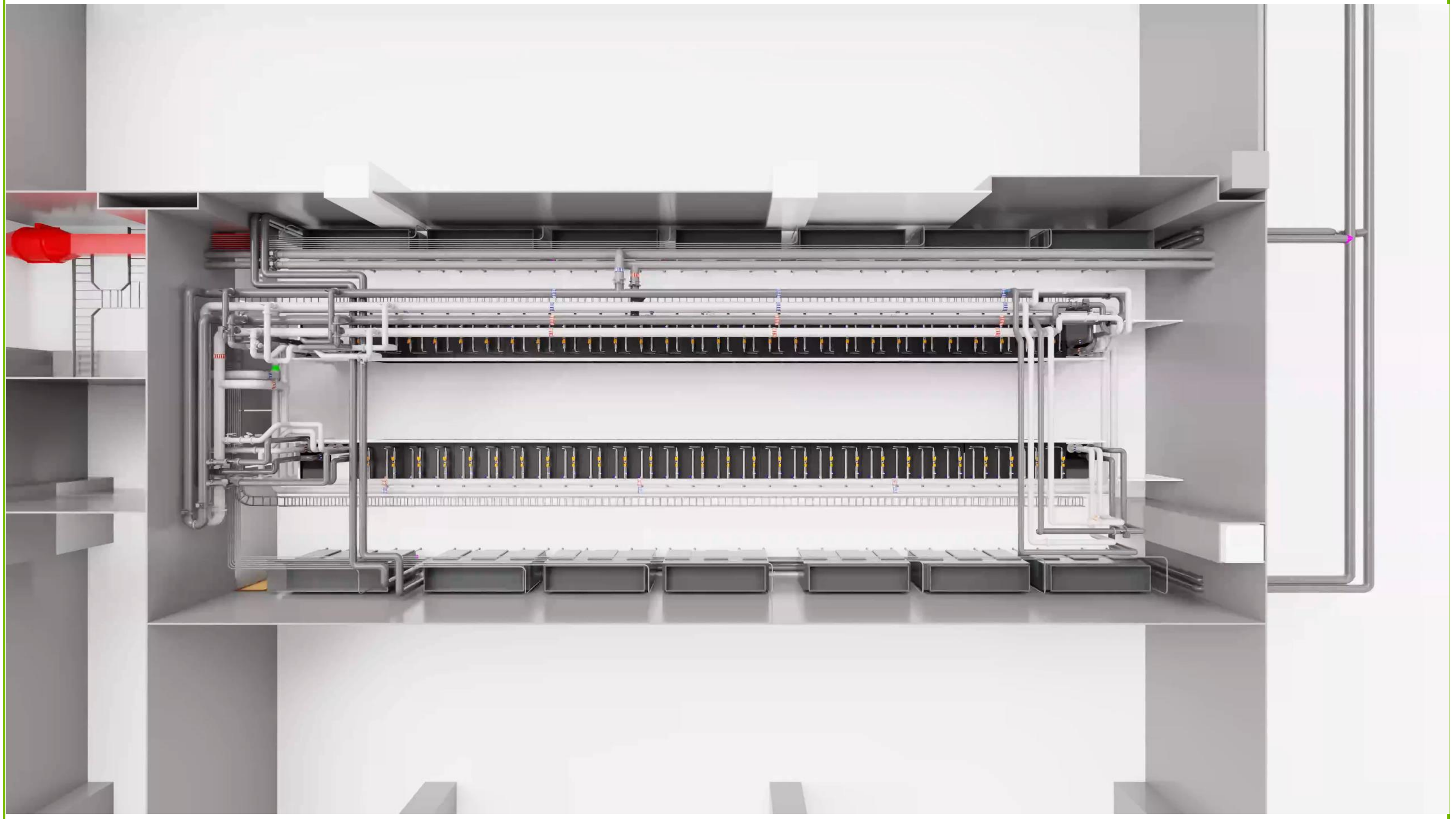
Omniverse Digital Twin of Liquid-Cooled Data Center

Liquid-to-Air and Liquid-to-Liquid Cooled Data Center

L2A Cooled Data Center



L2L Cooled Data Center

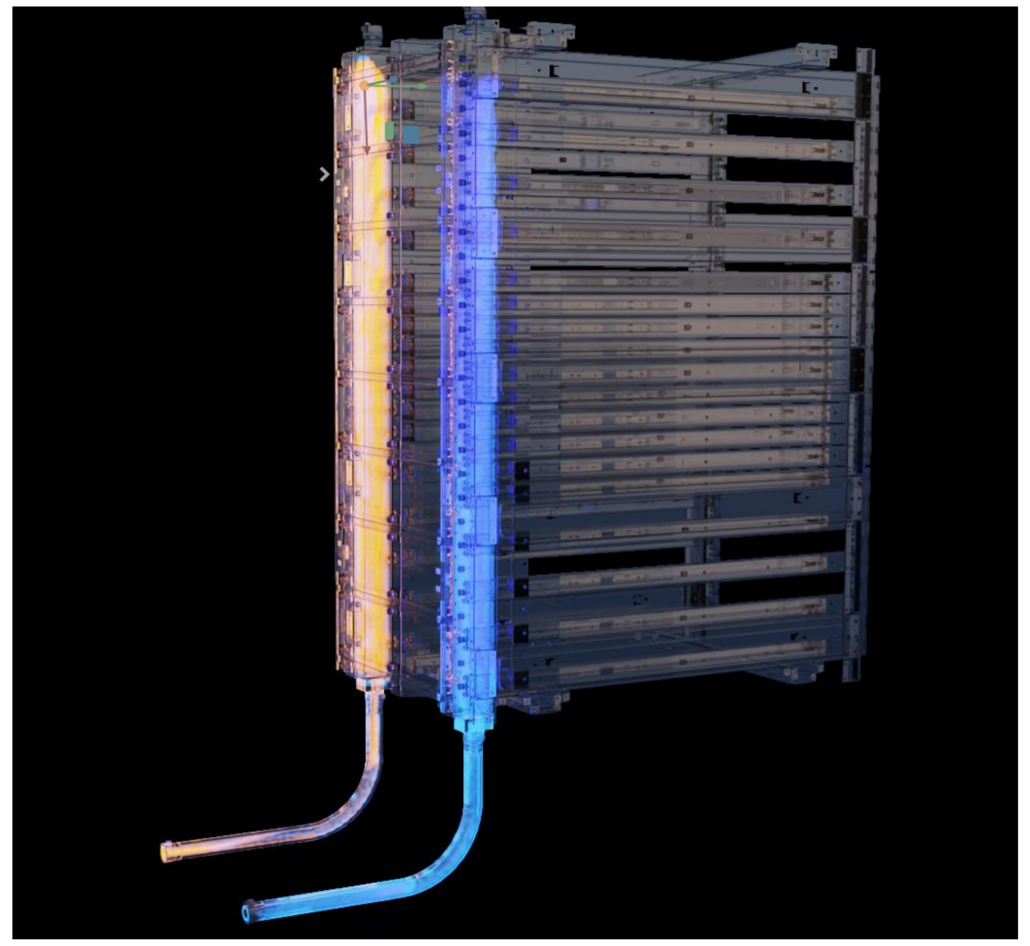


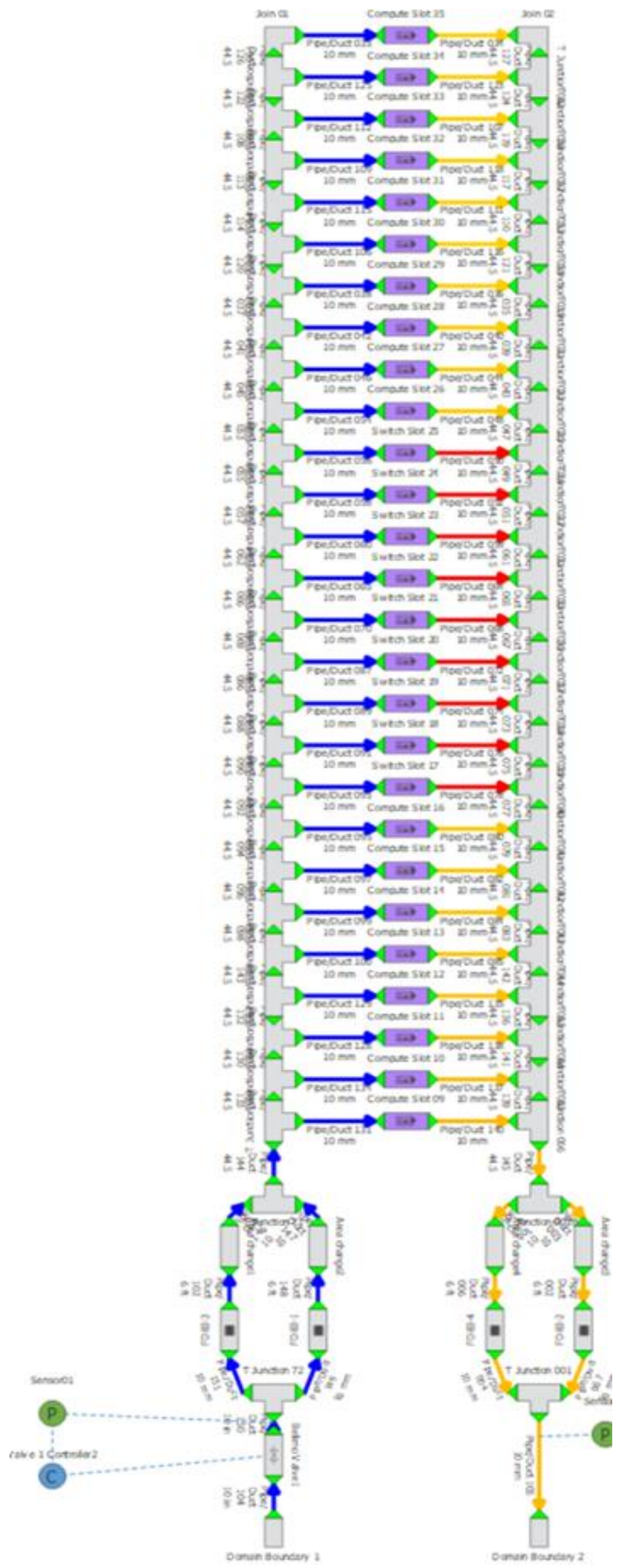
Liquid-Cooled Data Center Modeling

CFD (Computational Fluid Dynamics) & FNM (Flow Network Modeling)

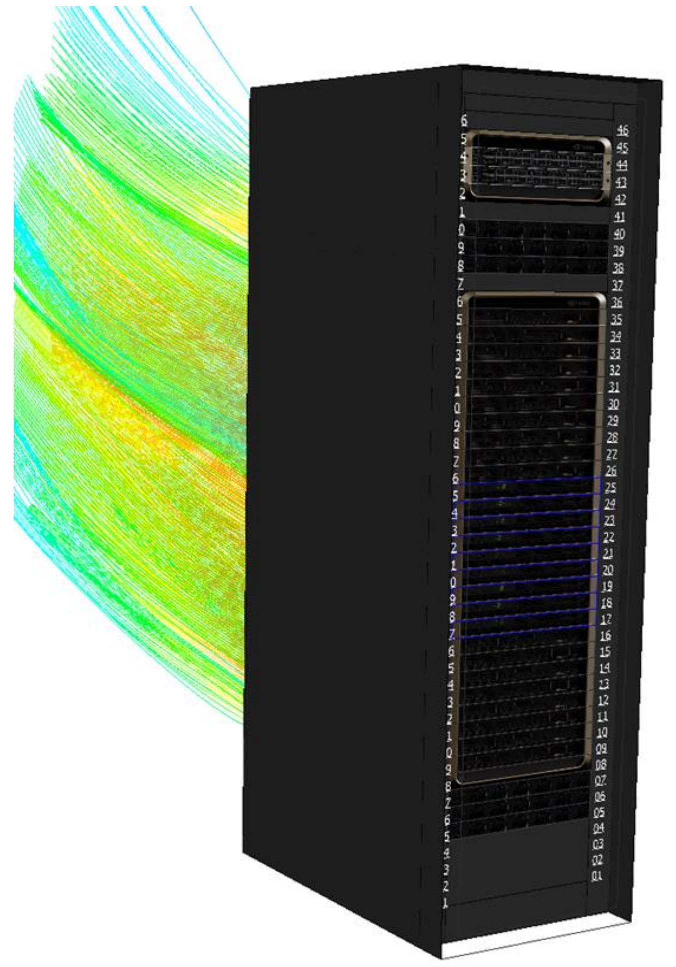
Rack Level Cooling

Flow Network Model for Rack Level Liquid cooling



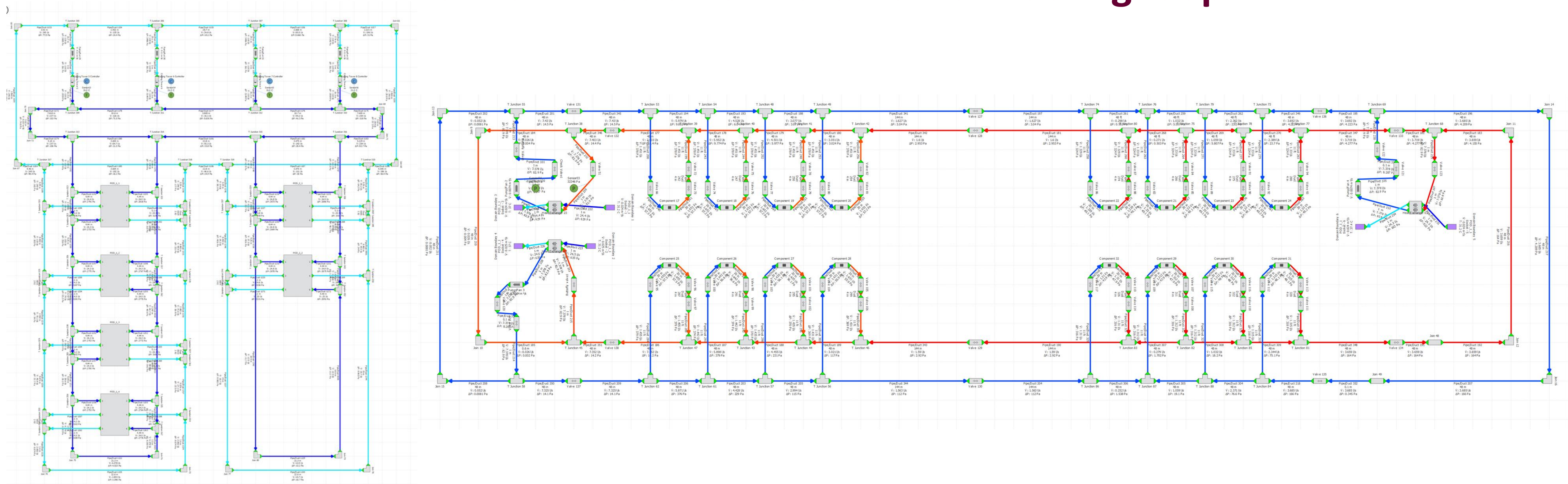


CFD for Rack Level Air Cooling



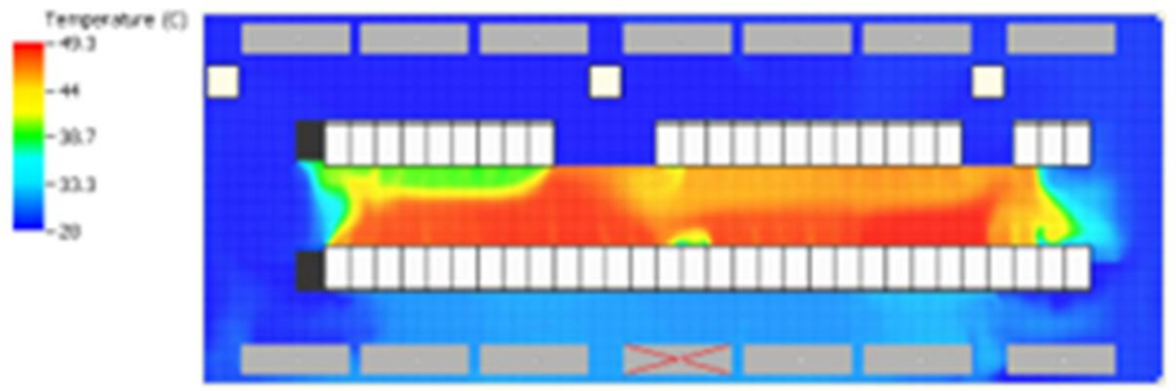
DC Level Cooling

Flow Network Model for Data Center Cooling loop

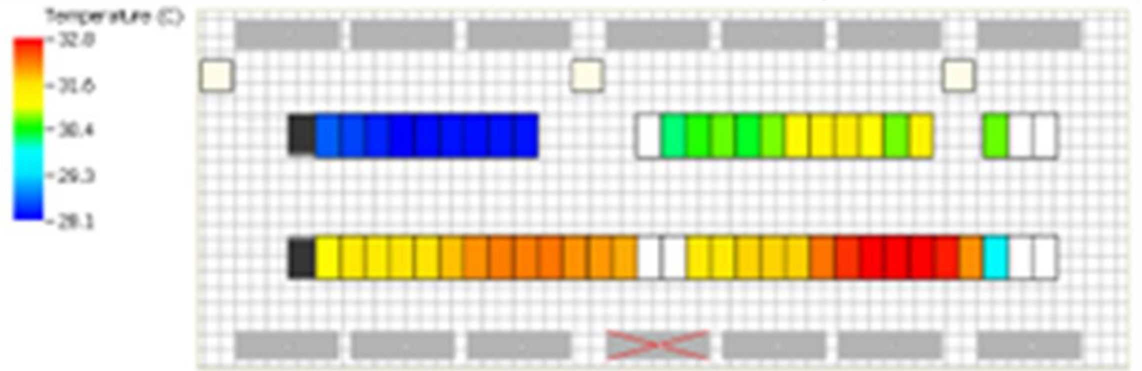


Simulation for Liquid and Air Cooling in Data center

Temperature Profile at halfway cabinet height



Maximum rack inlet temperature



AI-Accelerated Data Center Digital Twin

Generative AI for Data Center Design

Reinforcement Learning for Data Center Optimization

Physics-Informed AI Simulator

1200X CFD Speedup

Virtual Testing Environment

L2L/L2A CDUs

Dynamic Digital Twin

Industrial Metaverse
Nvidia Omniverse + Nvidia Modulus

Physical Twin

Data Center aka "AI Factories"

In Omniverse
- Designed
- Optimized
- Controlled

Energy Usage Prediction and Carbon Footprint Reduction

Predictive Failure Maintenance aiding Cost Savings

BCs to PINNs for Physics Simulation

IoT/Infrared Images

Data Center Digital Twin: AI-Driven DC Design & Operations

Features

Real Time Simulation

Inferencing and simulating the physics in real time

Design DT

Predicting product needs before manufacturing

Operation DT

Accruing **sensor data** and monitoring
Testing **failure scenarios**

Control DT

Tuning the control system dynamically
Response to system architecture change

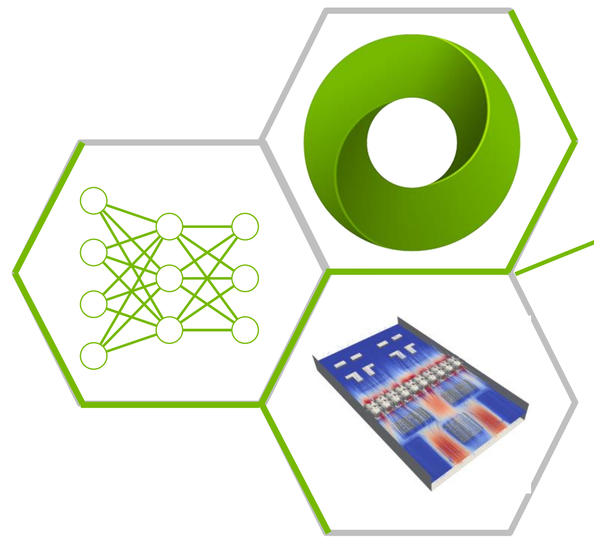
Requirements

Components

- CDU
- Servers
- Racks
- DC facility

Tools

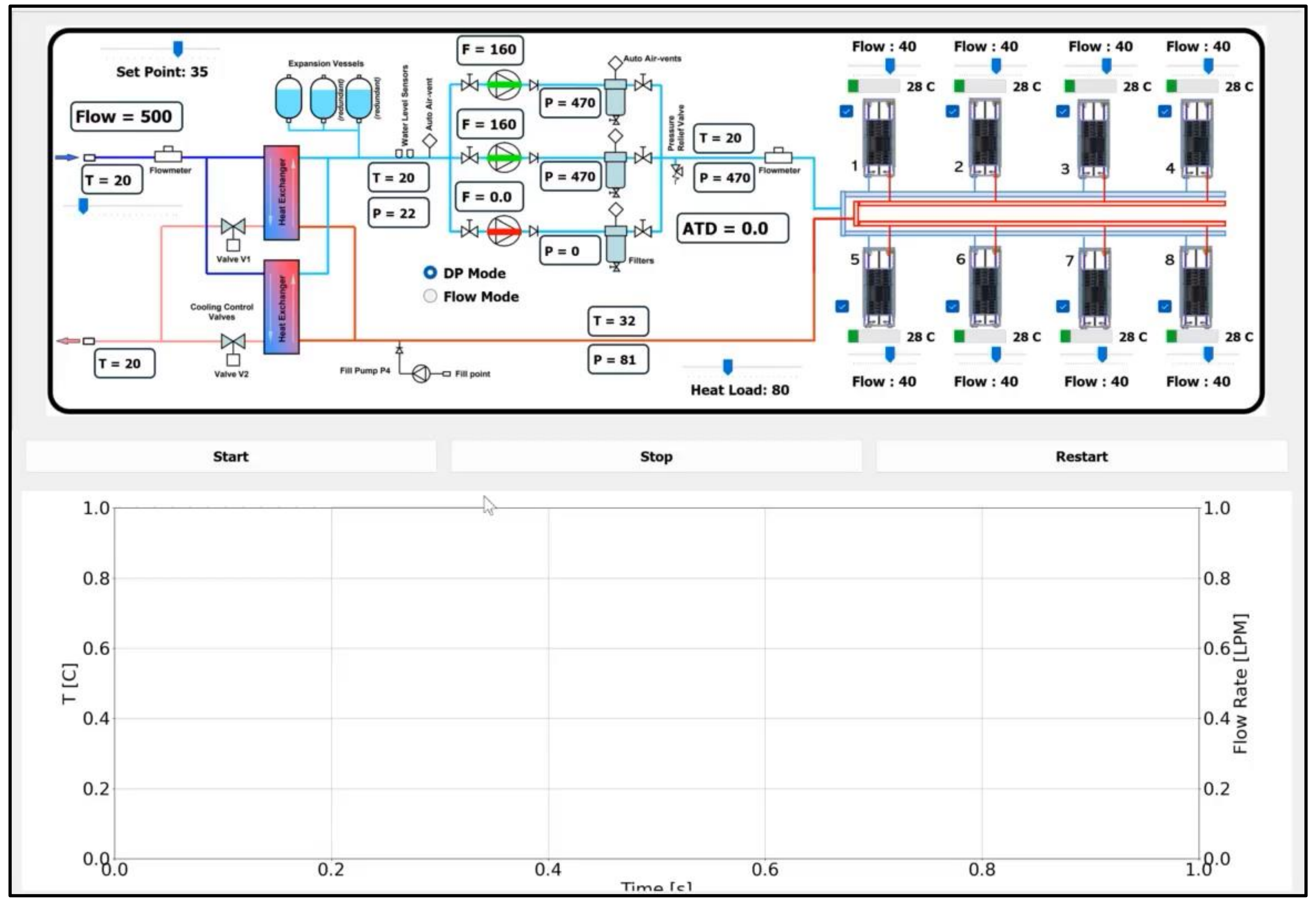
- Omniverse
- Modulus's PINN
- CFD
- FNM



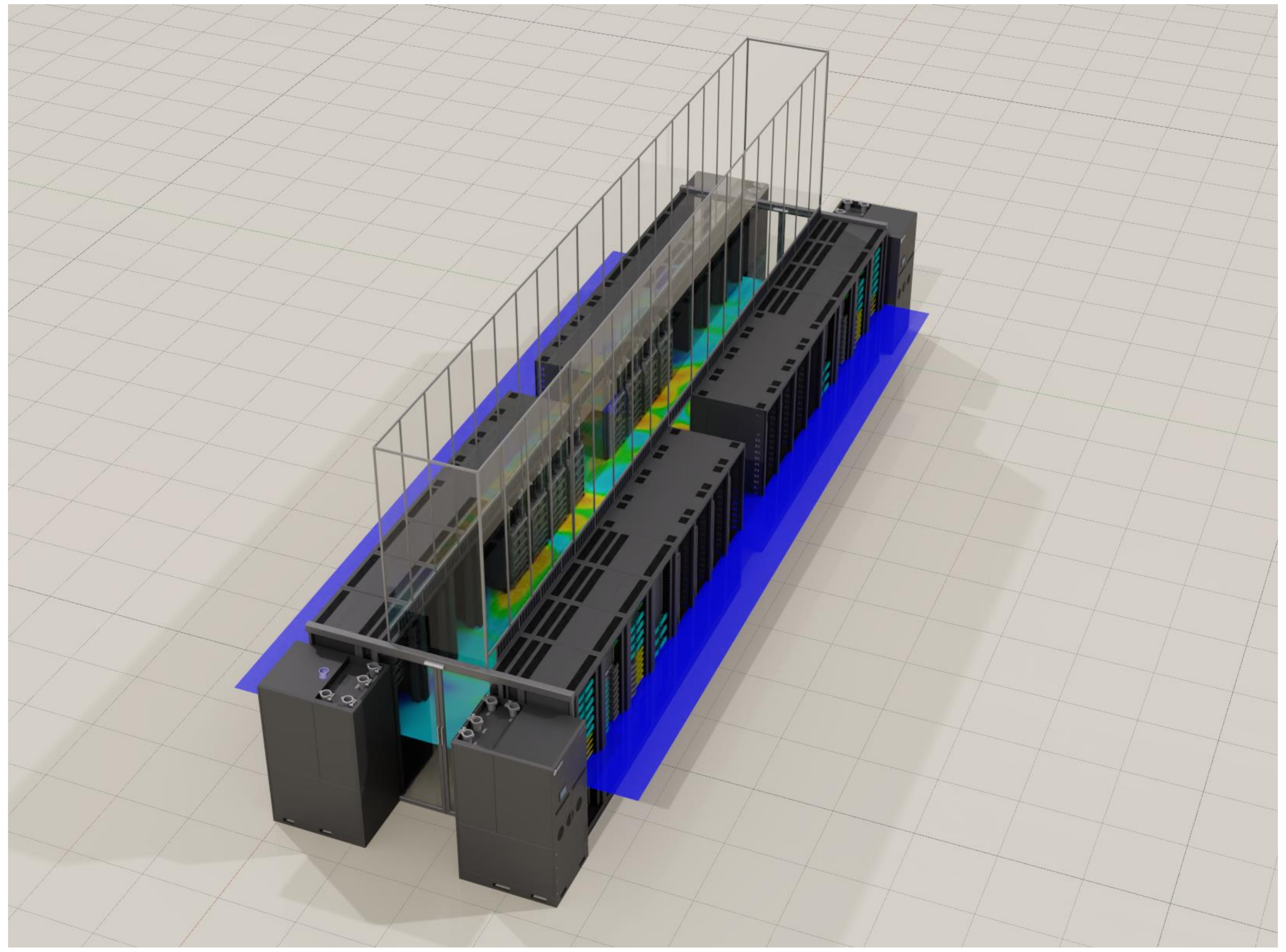
DC requirements

- Physical spec
- IT spec
- POD design & spec

Liquid Cooling Digital Twin



Air Cooling Digital Twin

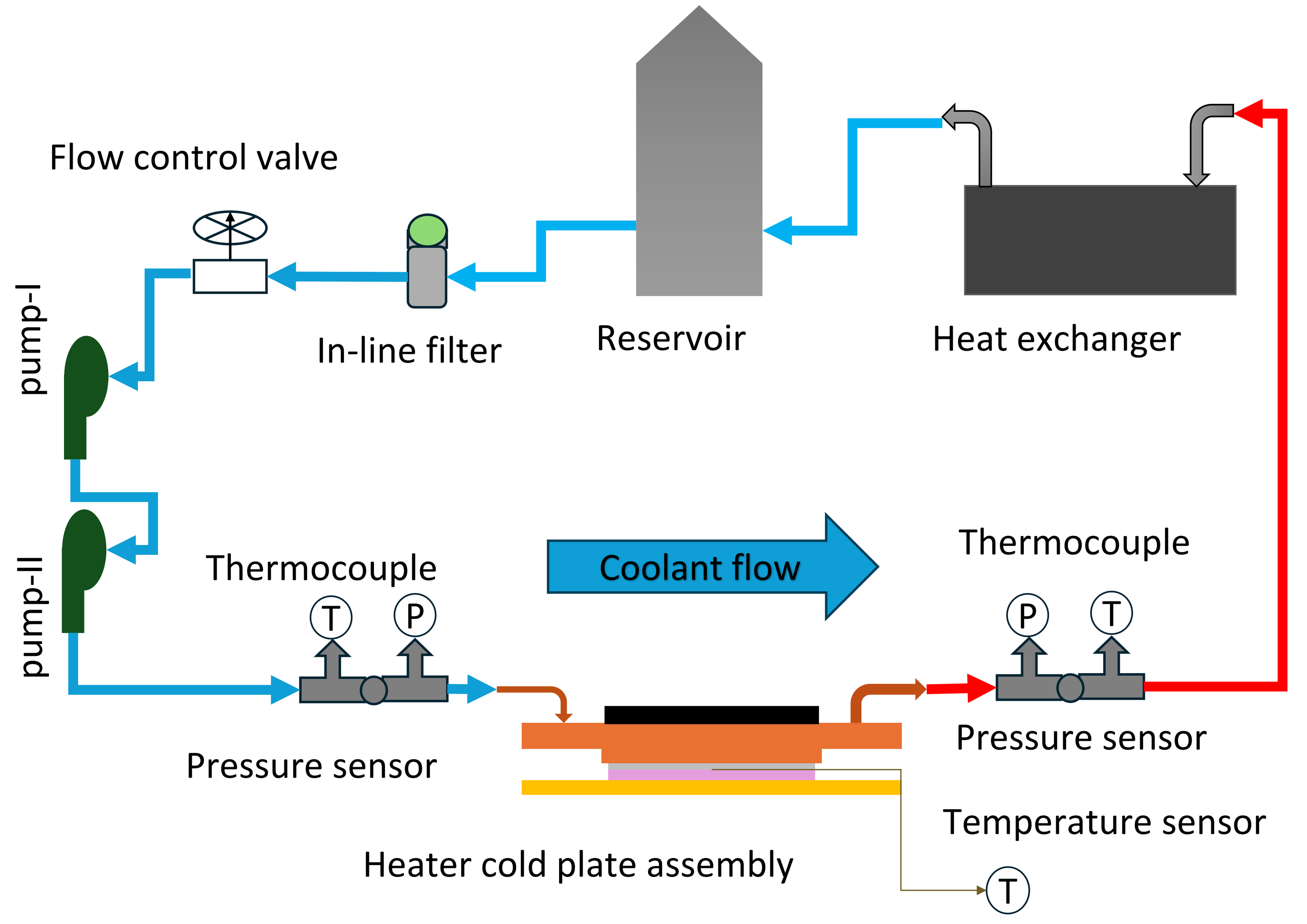


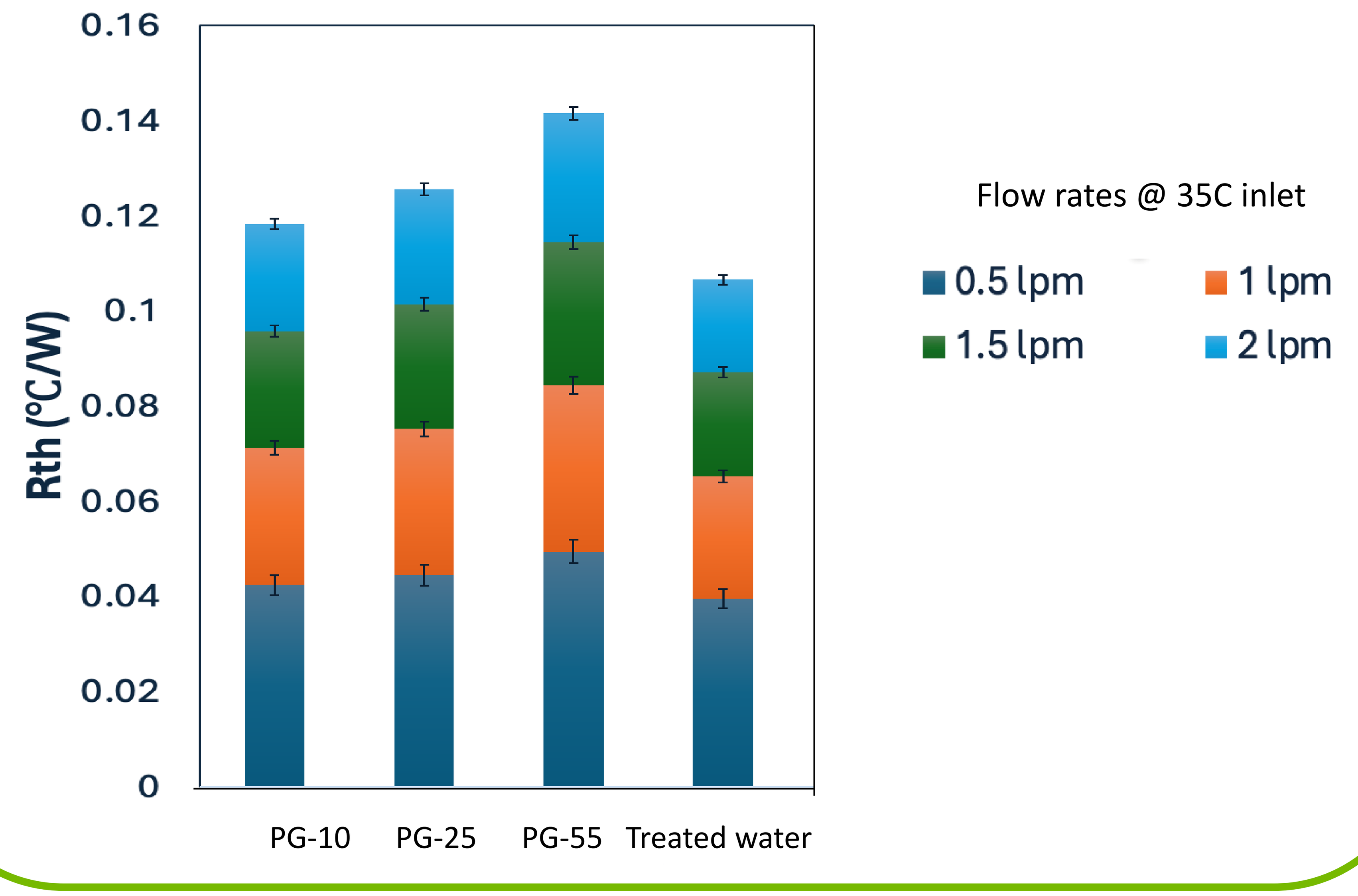
Real-time inference of thermo-fluid dynamics in a POD using NVIDIA Modulus and Omniverse.

Research Focus: Secondary Fluids, Corrosion, Erosion Studies

Performance

- Evaluating multiple coolants for thermal performance using thermal test vehicles.





Corrosion

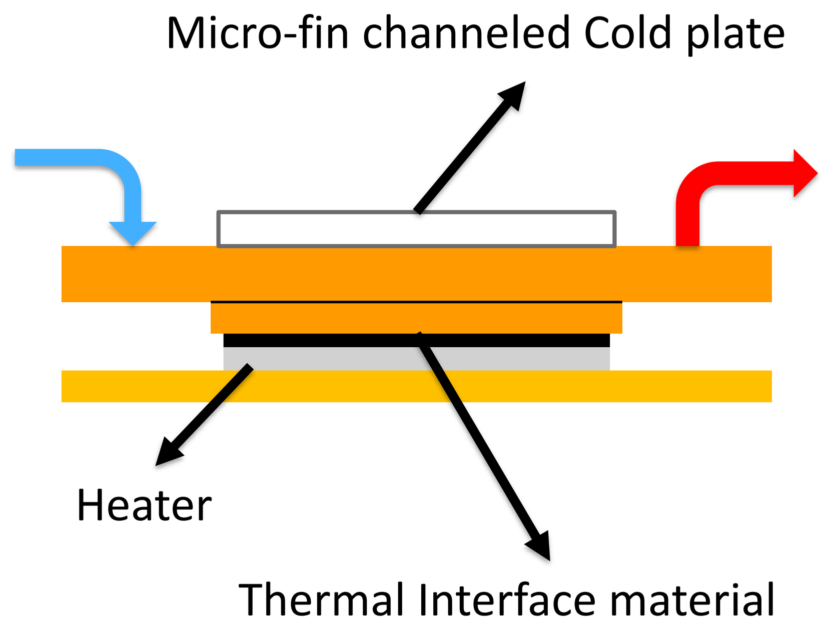
- Using ASTM-D1384 and D8040 based testing to evaluate corrosion of wetted materials.

Coupon Set

Thermal Bath

Erosion

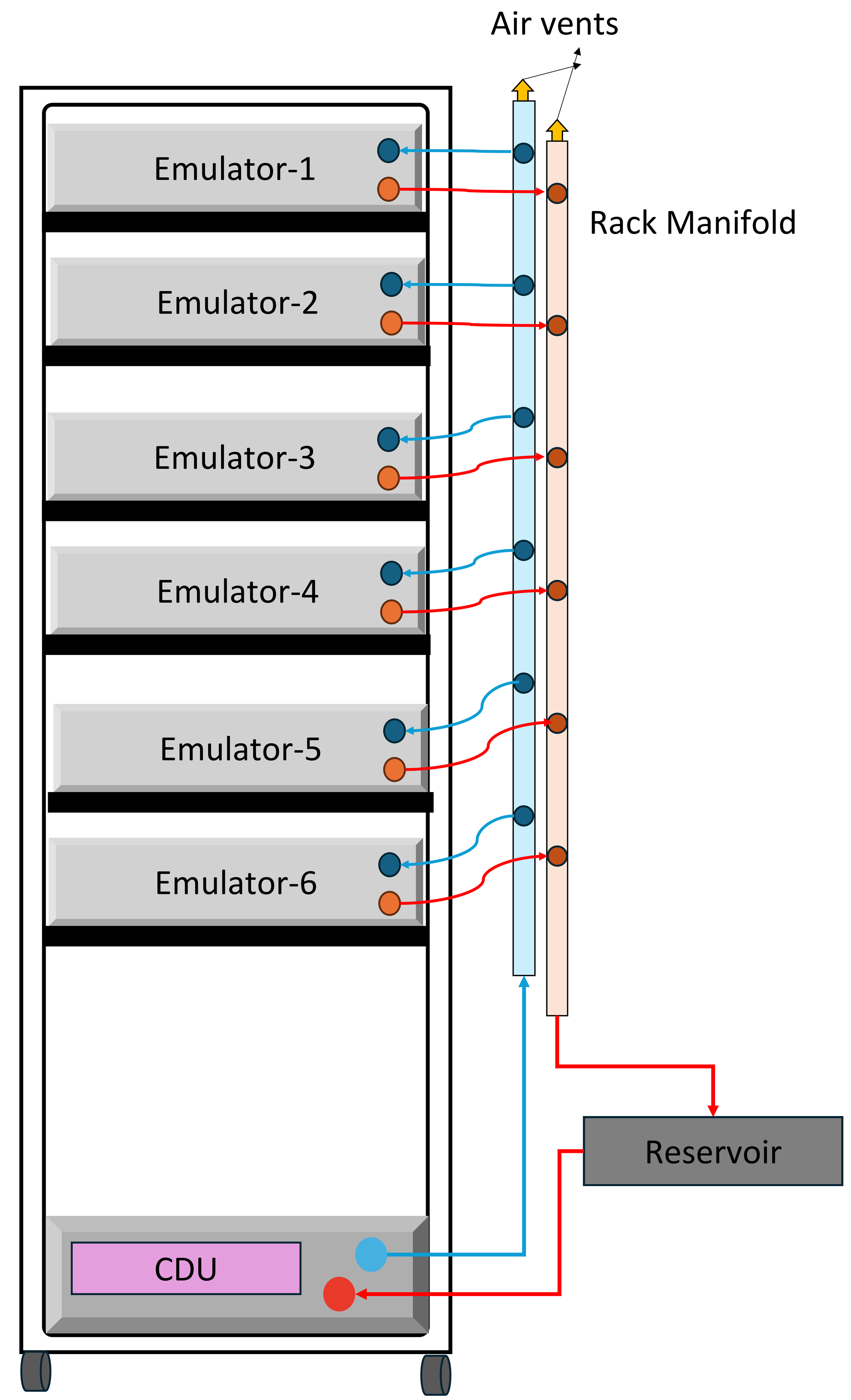
- Applicability of ASHRAE erosion fluid velocity limit of 1.5 m/s for liquid cooling applications



Micro-fin channels

Bio-growth

- Analyzing bio-growth in coolants for both stagnated and continuous circulation scenarios.



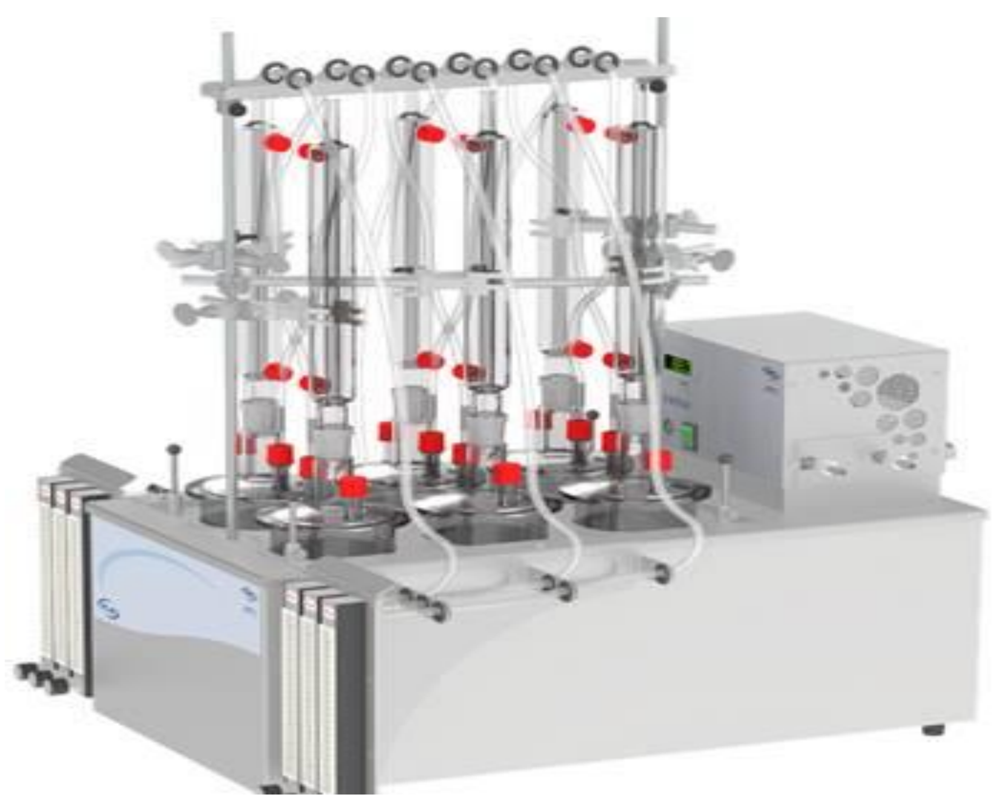
Rack Emulator

NVIDIA

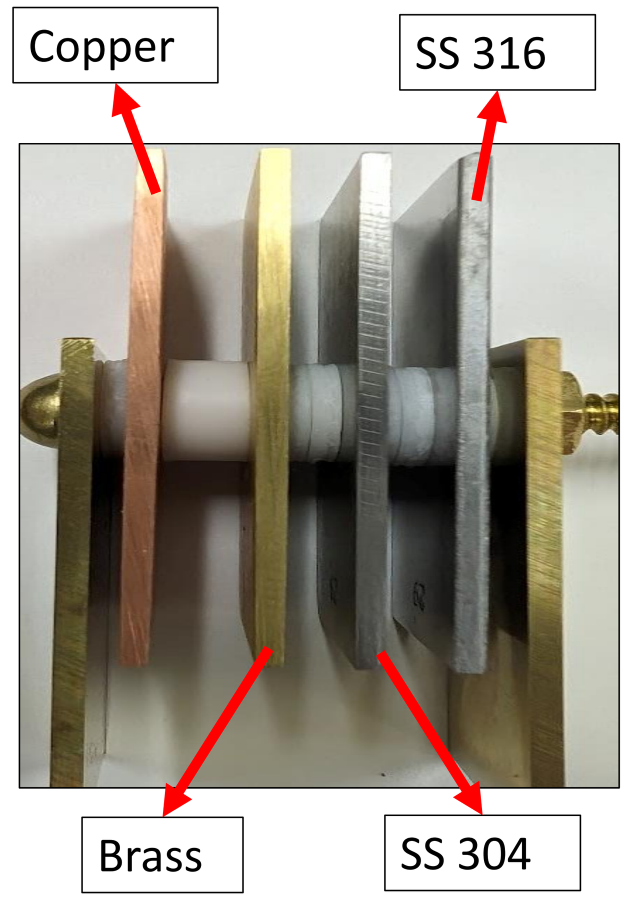
Coolant Reliability

Corrosion

- ASTM-D1384 and D8040 based testing to evaluate corrosion Inhibitor efficiency of Glycol based and water-based with respect to data center wetted materials.



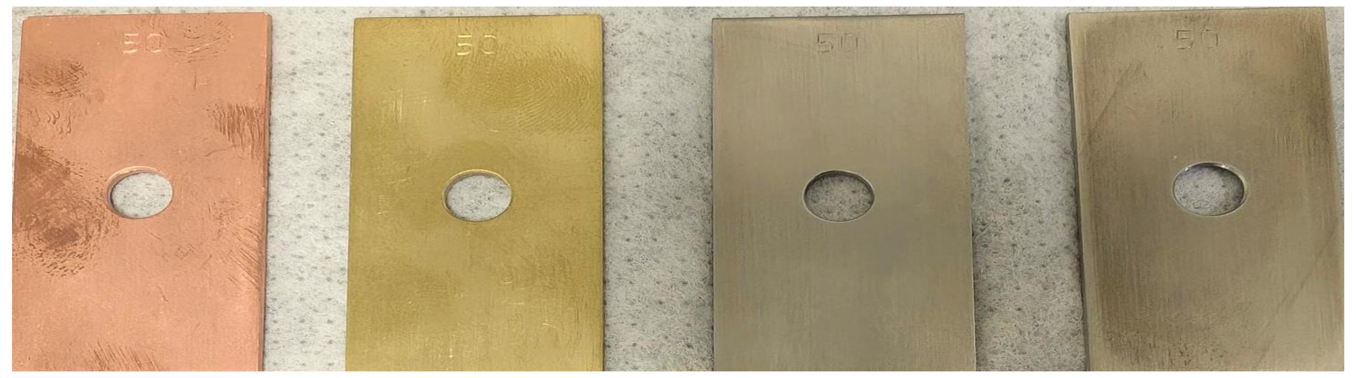
ASTM test setup



Coupon set

Copper Brass SS 304 SS 316

Test coupons before testing

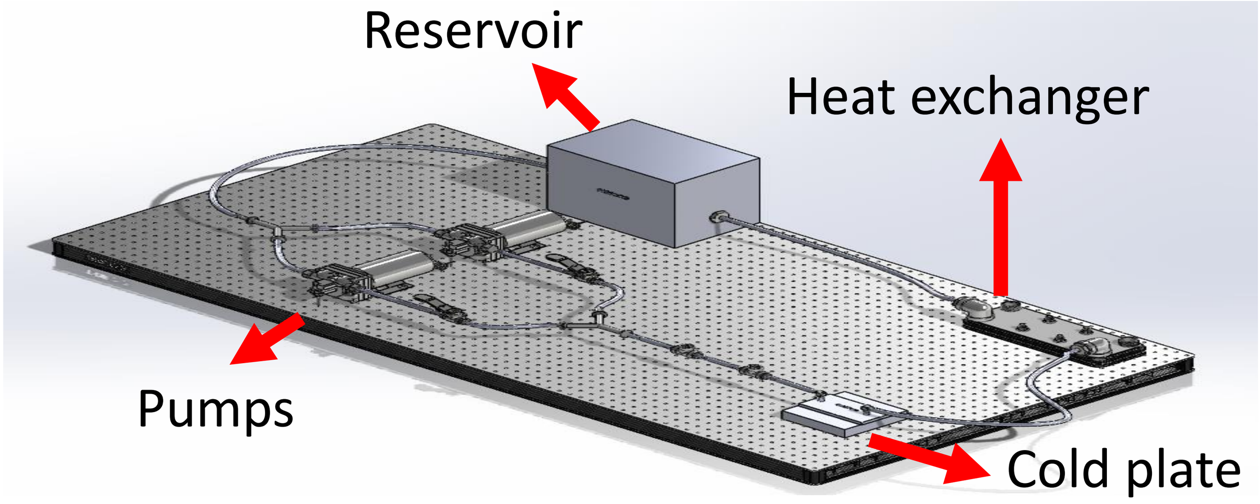


Copper Brass SS 304 SS 316

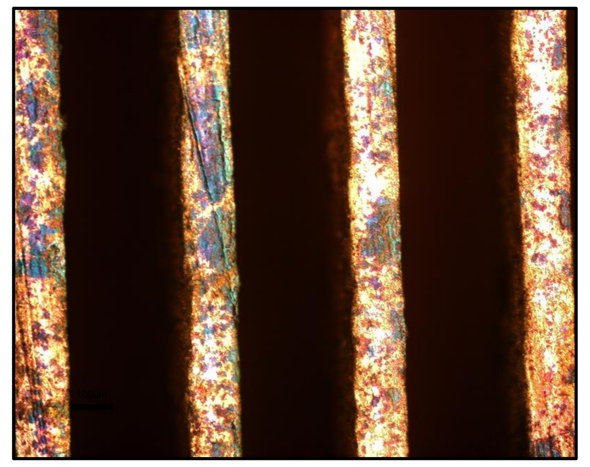
Test coupons after testing

Erosion

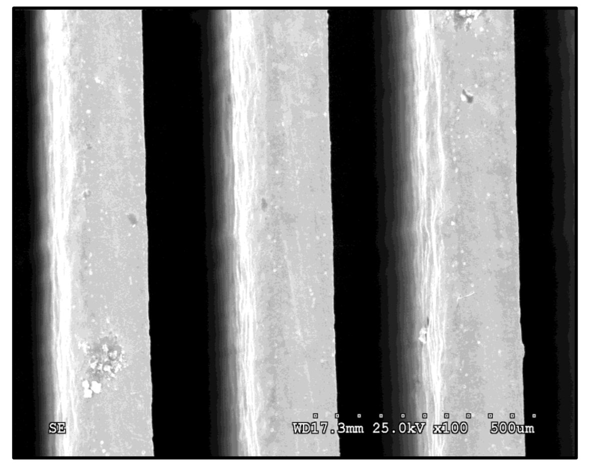
- Applicability of ASHRAE erosion fluid velocity limit of 1.5 m/s for liquid cooling applications in a TCS loop



CAD model of Test Setup

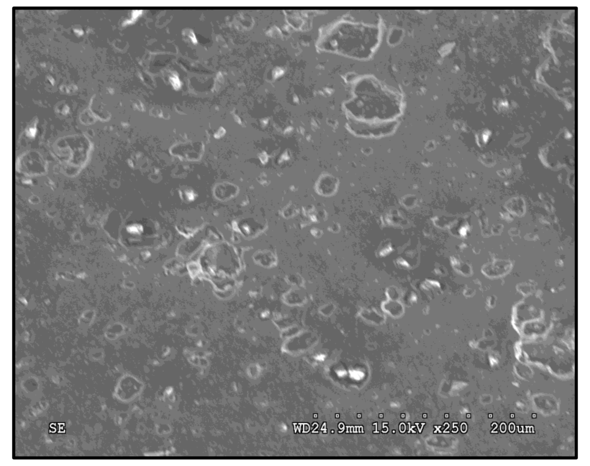


Optical image of micro fin channel



Micro-Fin Channel Surface on a SEM

Optical image of EPDM hose



EPDM Hose Surface on a SEM

<u>Future tests</u>

- EPDM erosion test at 6.5m/s
- Copper erosion - Shearing
- Copper erosion - Impingement

Bio-Growth Analysis of Secondary Fluids

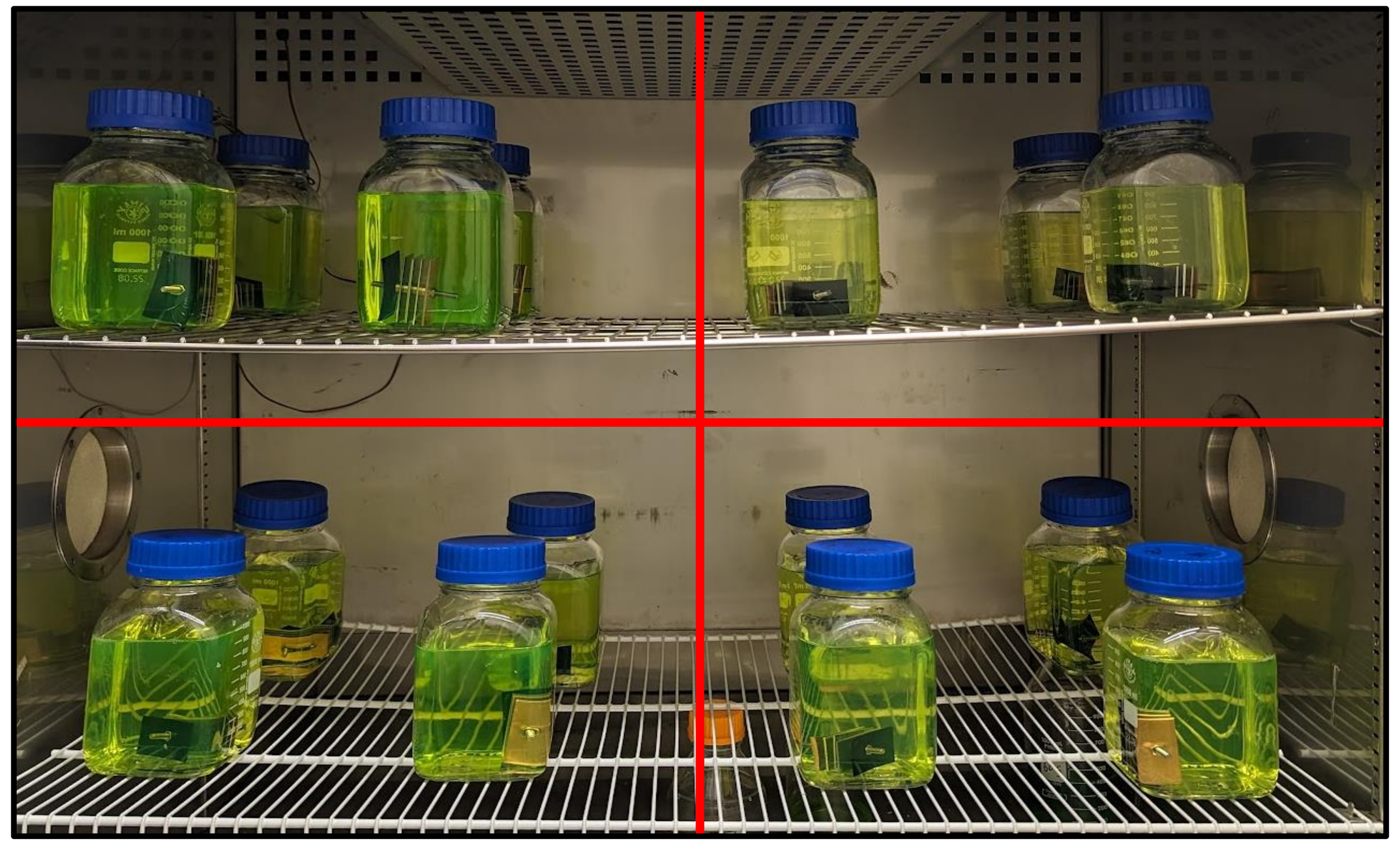
Stagnated

- Analysis of bio growth in stagnated scenarios of cross-mixed glycols and water-based coolants
- Anerobic Bacteria growth potentials require continuous fluids flow

1. PG 1.8vol % -> 925ml of Treated water (A) + 75 ml of PG-25 (B)
2. PG 2.2vol % -> 990ml of Treated water (A) + 90 ml of PG-25 (B)
3. PG 2.5vol % -> 900ml of Treated water (A) + 100 ml of PG-25 (B)
4. PG 3.1vol% -> 875ml of Treated water (A) + 125 ml of PG-25 (B)

30 Days

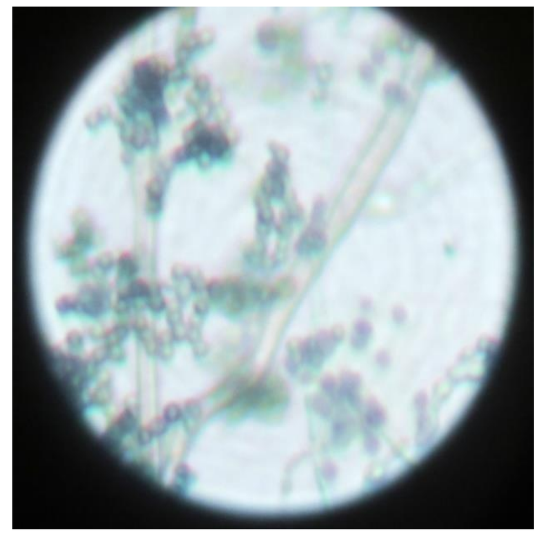
60 Days



90 Days

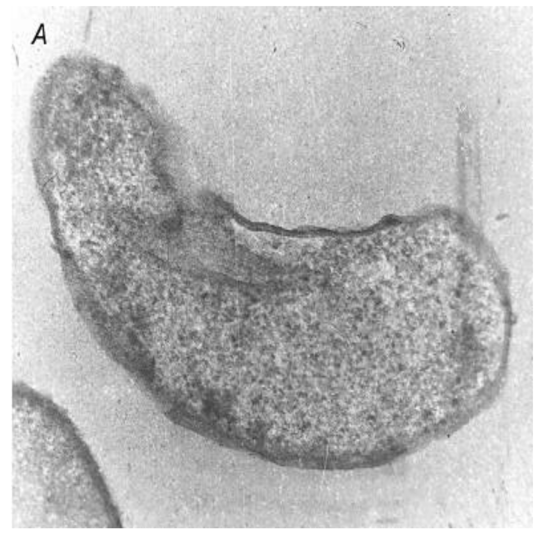
180 Days

Air-tight glass jars Kept in Environmental Chamber



Fungal bacteria[#]

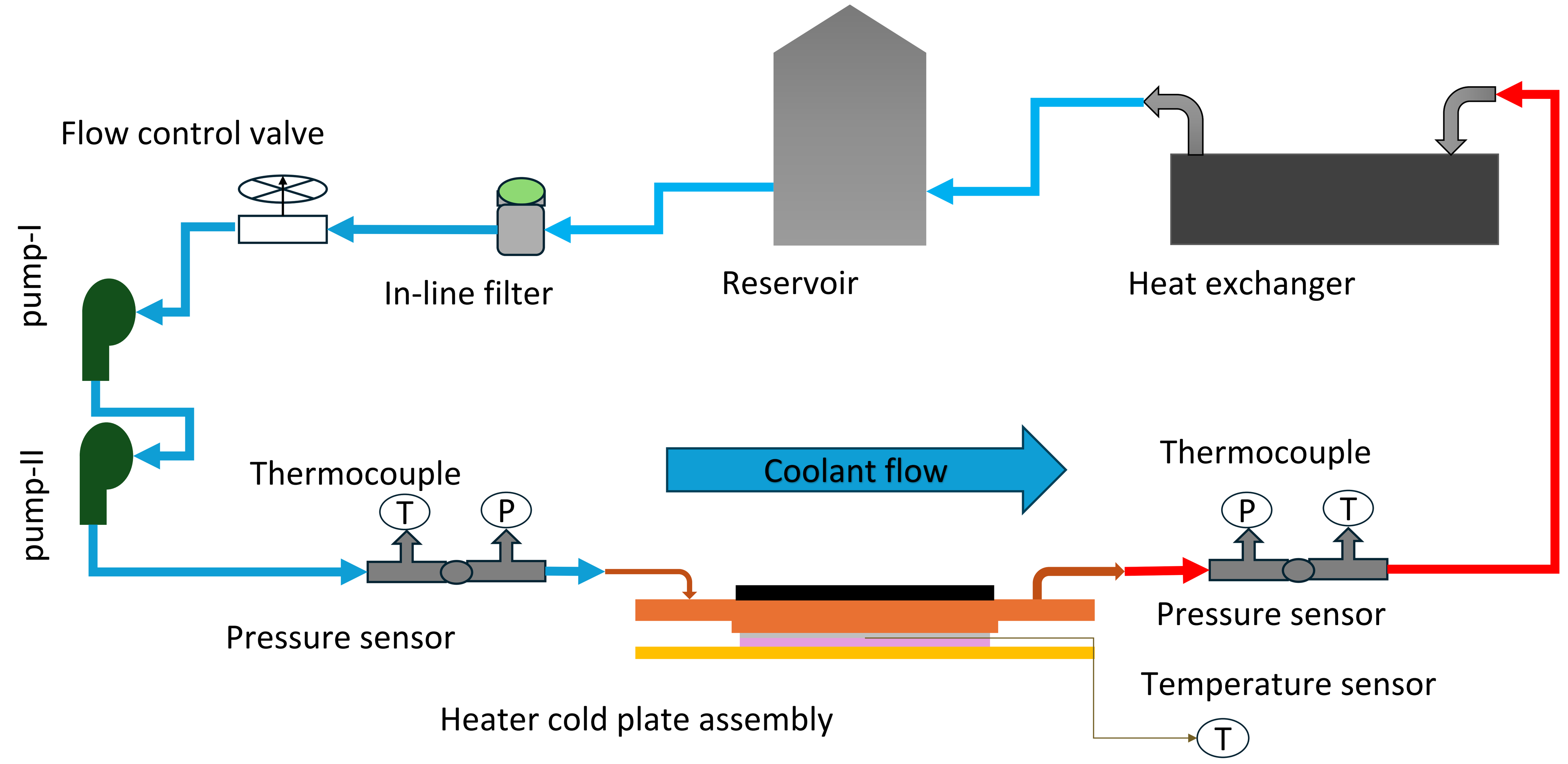
Aerobic bacteria[#]



Anaerobic Bacteria(SRB)[#]

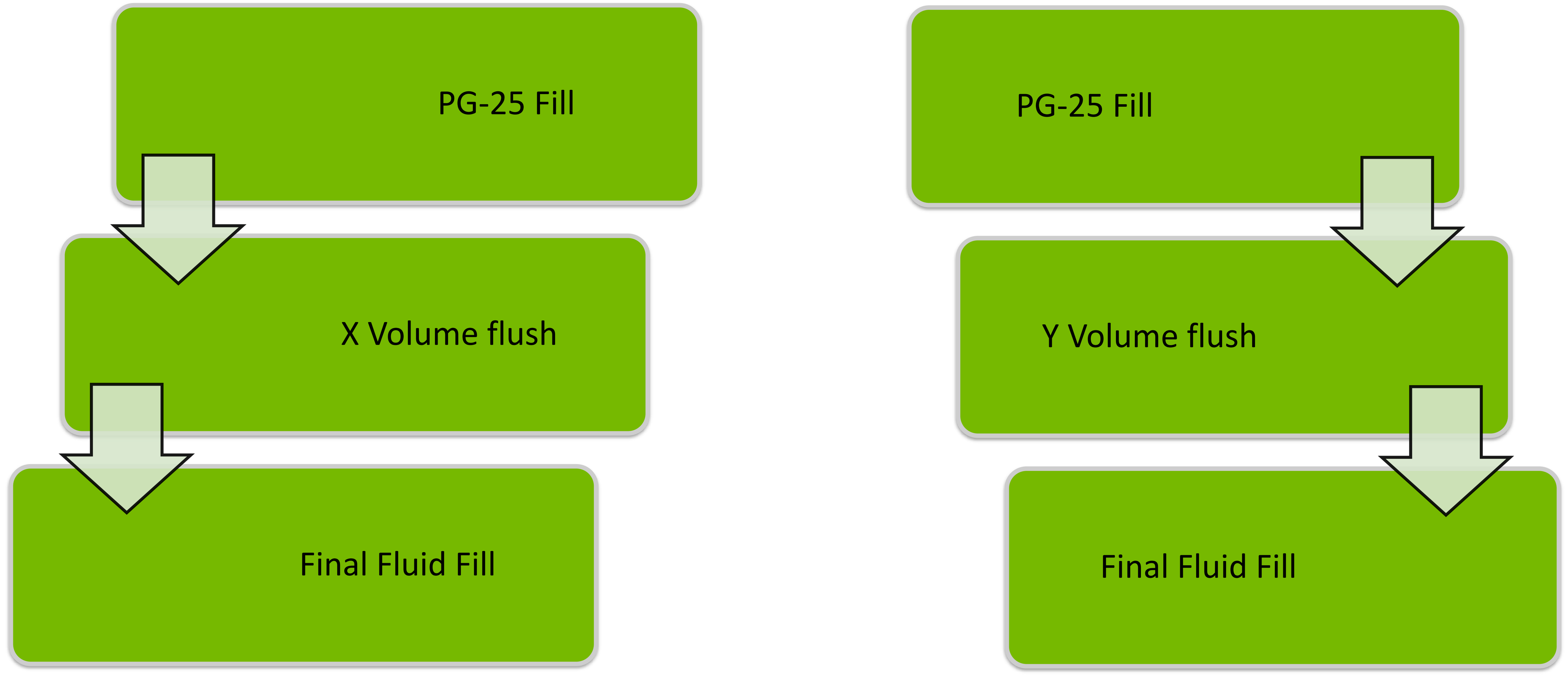
Continuous Circulation

- Evaluation of bio growth in continuous flow scenarios following various flushing conditions
- Aerobic bacteria detection requires prescribed flushing followed by liquid cooling comissioning.
- Inline fluids monitoring and predictive maintenance for data centers.



Test-I

Test-II



Research Focus: Rack Power Density vs DLC Cooling Technology Roadmap

Thermal Limits

- Allowable case temperature
- Best in-class cold plate
- Supply fluid temperature
- Maximum fluid flow

+

Mechanical Limits

- Rack dimensions
- Rack manifold sizing, Fluid velocity
- Rack connection sizing
- Cooling loop sizing

Cooling Technologies Limit Range

Heat can be managed within

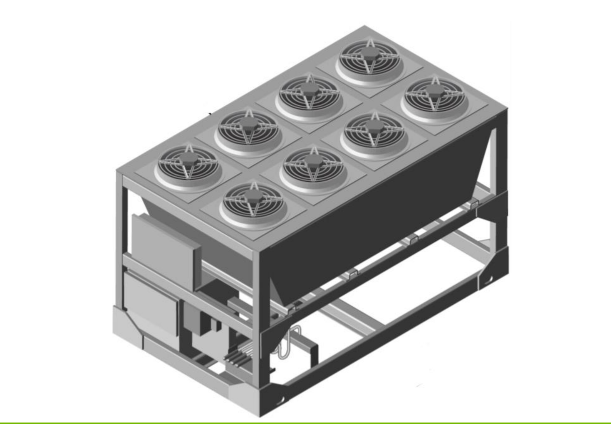
SERVER	Air Cooling	Higher limit: 2 kW/U
	Single Phase	Higher limit: X kW/U
	Advanced Cooling	Higher limit: XX kW/U
RACK	Air Cooling	Higher limit: 60 kW
	Single Phase	Higher limit: Y kW
	Advanced Cooling	Higher limit: YY kW

Research Focus: Data Center Efficiency & TCO Analysis

Energy Efficiency Main Contributors

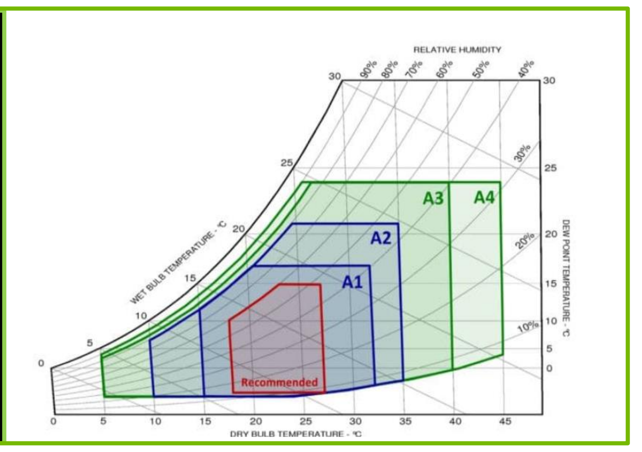
IT System Cooling Technology

Building Cooling Equipment & Performance Efficiency

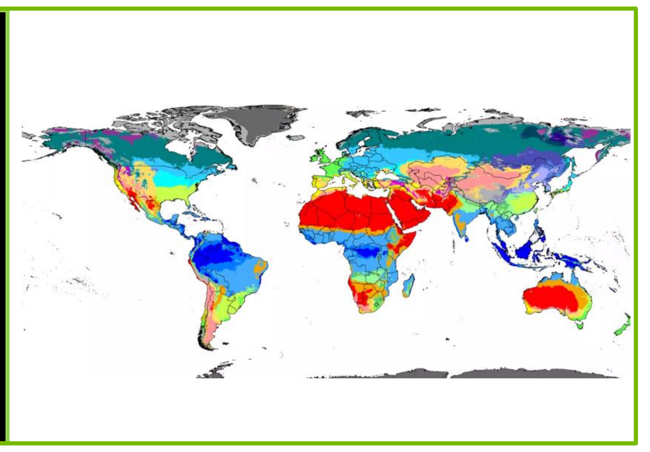


Power Transformation & Losses

Operating Specification



Location & Environmental Conditions



Air Cooling vs. Hybrid of Air & Liquid Cooling

15.6% reduction in **Total Data Center Power**

28% reduction in **Facility Power**

Further power reduction

12.4% reduction in **Server Power** due to **74%** lower **Fan Power**

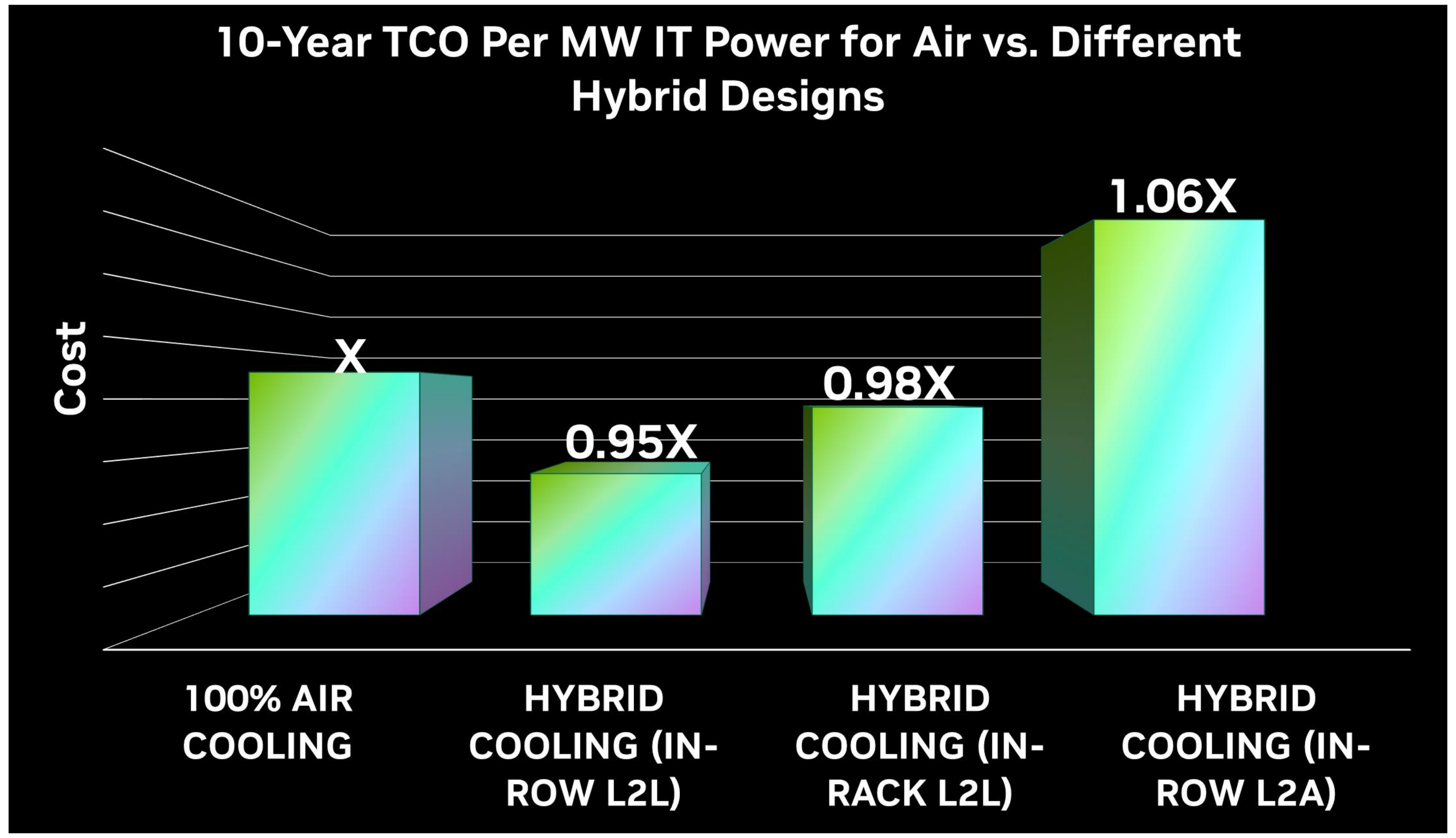
Air Cooling (100%) — Air Cooling + (39%) Liquid Cooling (61%) — Air Cooling + (31%) Liquid Cooling (69%) — Air Cooling + (25%) Liquid Cooling (75%) — Air Cooling + (15%) Liquid Cooling (85%)

Air Cooling (AC)

- Limited cooling capacity
- Lower energy efficiency
- Higher water usage
- Higher noise level
- Lower upfront cost
- Higher operation cost

Liquid Cooling (LC)

- Improved cooling efficiency
- Higher upfront cost
- Lower operation cost
- Increased complexity
- Risk of leaks
- Extended IT server life
- Heat Reuse possibility



➢ Both power efficiency and TCO analysis are based on mechanical cooling design only.

Ref: A Comparative Data Center Energy Efficiency and TCO Analysis for Different Cooling Technologies | InterPACK | ASME Digital Collection

NVIDIA

Data Center Sustainability

Data Center Sustainability Strategies

- Renewable Energy Integration
- Carbon Reduction
- Waste Heat Recovery
- Water Conservation
- Performance Metrics

Sustainable Design

- Digital Twin Integration
- Comprehensive Analysis
- Future Proof Design
- Energy Efficiency Optimization
- Lifecycle Assessment

Deployment & Operation

- Optimize Operations
- Maximize Resources
- Predictive Maintenance
- Ongoing Improvement
- Data-Driven Decision Making

Monitoring & Automation

- Data Analytics for Monitoring
- AI-Driven Efficiency
- Automated Management
- Real-time Data
- Key Metric Tracking

Lifecycle Management

- IT Asset Management
- Lifecycle Management
- Regulatory Compliance
- Documentation
- Disposal

Waste Heat Recovery Project

- Reducing power consumption by utilizing waste heat from IT equipment & design optimization
- Collaborating with research centers to explore possibilities to recover waste heat from liquid-cooled racks
- Developing and testing different prototypes to generate electricity from waste heat recovery unit (WHRU) to power cooling equipment such as adsorption/absorption chillers

ARPA-E COOLERCHIPS Project: *OMNICOOL*

The Department of Energy’s ARPA-E supports early-stage impactful energy technologies, with the COOLERCHIPS program focusing on efficient and reliable cooling for data centers, and Nvidia receiving the highest grant of $5 million from the program's $40 million fund.

Objectives

- Boosting **energy efficiency** by employing hybrid D2C two-phase and single immersion cooling systems.
  - Target: PUE <1.05
- Elevating **power density** challenge by utilizing a compact flow distribution and heat rejection systems.
  - Target: >160 kW/ rack
- Overcoming **geolocation and weather constraints** deploying >1 MW compute power in remote and harsh environments.
  - Target: fit within ISO 40’ container, Ambient temperature ≥40 °C
- Emphasizes **environmental impact** and **sustainability goals** by using Green refrigerants and achieving zero water consumption.
  - Target: GWP <1

Innovations

- Integrating two electronic cooling approaches
- An innovative two-phase porous metal cold plate technology.
- Non-orthogonal, interlocking coil arrays for compact free coolers.
- In-Rack Distributed Pumping and Flow Separation System.
- Multi-scale flow distribution systems utilizing hierarchical architectures.

NVIDIA