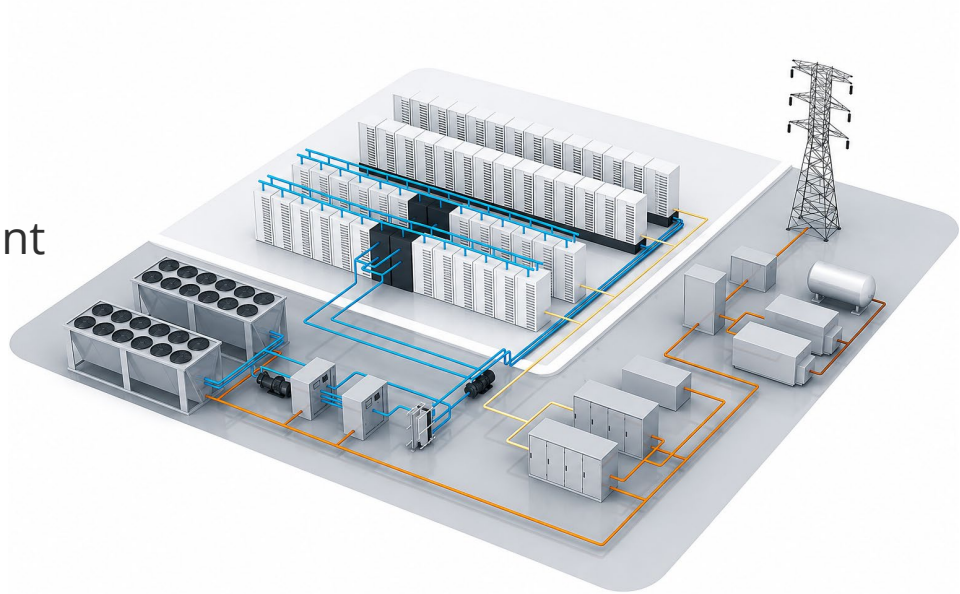
The cooling chain is one system.
It is time we designed it like one.

A system-simulation perspective on data center thermal management



Nithish Selvan
Senior Simulation Engineer · AI Data Center Thermal Management, Modelon
K8-02 Panel · ASME Summer Heat Transfer Conference · July 27, 2026

One thermal event, three response times

Chip

700 to 2000 W per processor.
Throttles in

milliseconds

designed by the semiconductor industry

Cold plate and CDU

Coolant loop and valve control.
Responds in

seconds to minutes

designed by the cooling hardware industry

Facility plant

Chillers, towers, energy supply.
Adjusts over

minutes to hours

designed by the facilities industry

An AI cluster steps from idle to full load in milliseconds. Each layer answers at its own speed, and each was designed by a different industry.

No single team sees the whole response. This is where co-design has to begin.

The question changed — not just the heat load

we are here

1990s–2012 **1–5 kW/rack**
air cooling

→

2012–2022 **10–25 kW/rack**
hybrid cooling

→

2023 on **40–130+ kW/rack**
liquid / two-phase

The old question was **“can we remove the heat?”**

The new question is **“will it hold — every load, every hour, all year?”**

You cannot answer that at a design point. You need the whole system, in motion.

What that experience taught us: aerospace and automotive solved this same class of problem — many components, coupled heat balances — with system simulation.

Progress accelerated once OEMs and suppliers standardized on shared model exchange, through FMI. Data center cooling now needs the same.

Data Center Library: the full chain in one physics

Data Center Library
in Modelon Impact
Physics-based dynamic system simulation

Modelon Impact

1 IT equipment
- ❑ Air and liquid racks
- ❑ High-density GPU

2 Cold plate and two-phase
- ❑ Direct-to-chip cooling
- ❑ Two-phase flow

3 Coolant Distribution Units
- ❑ ASHRAE calibrated
- ❑ Dynamic hydraulics

4 Cooling plant
- ❑ Vapor compression
- ❑ Plant optimization

5 Facility cooling network
- ❑ Pumps, valves, HX
- ❑ Flow balancing

6 Controls
- ❑ Fully dynamic controls
- ❑ Sensors and logic

7 Digital twin
- ❑ Operational analytics
- ❑ Virtual commissioning

8 Open standards
- ❑ Modelica and FMI
- ❑ OpenUSD

9 Sustainability
- ❑ Waste heat recovery
- ❑ Water, energy

Each layer has been modelled before. What is new is modelling them together.

Everyone is asking the same question: can I decide faster?

Chip, CDU, chiller, facility, commissioning — different teams, same pressure: decide faster, with more confidence, as the technology moves faster than the decision cycle.

Detailed modelers — the experts

High-fidelity models: chip, CDU, chiller

Rack-level hydraulic and thermal resistance
Flow balancing, manifold pressure drop
Two-phase behavior in cold plates and CDUs

Decision makers — operators, commissioning

Fast full-system models, lower fidelity

PUE, cooling distribution, capacity vs redundancy
Heat recovery: where to extract for efficiency
Answers in business terms: more AI load, more cost

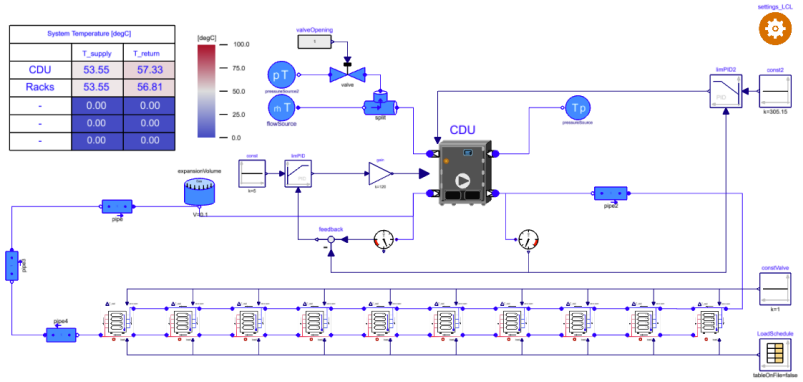
Same physics, different resolution — one consistent model serving both.

Experts build and interrogate it. Decision makers get answers in business terms. Increasingly, AI agents act on it directly.

One shared model is what lets all of them move faster — without losing the physics.

Caught in simulation, before hardware was at risk

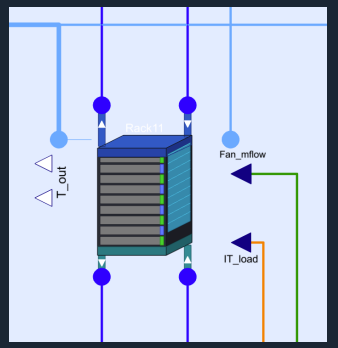
CDU control at 1 to 10% of rated load



At very low loads, secondary supply-temperature fluctuation can fail the primary-side 3-way valve. This failure mode is invisible at the design point.

A control strategy validated in simulation holds supply temperature within ±0.5 °C.

Two-phase direct-to-chip, as a system



Rack and cold-plate models across air, liquid, hybrid, and two-phase cooling.

"Modelon's system-level capabilities allowed us to understand two-phase direct-to-chip cooling as an integrated system and explore realistic operating conditions with more confidence."

Dr. Lingnan Lin, Mechanical Engineering, University of Maryland

Local optima are not good enough anymore.

Designing the cooling chain as one system needs a shared, open, dynamic model.

A shared dynamic model

The common technical substrate three industries can each trust and extend.

Open and AI-ready

Modelica, FMI, and OpenUSD as a common description layer — so experts, operators, and AI agents all act on one model.

The real barrier

Not capability. Whether the industry adopts a shared substrate, or stays siloed.

Simulation for experts, for non-experts, and for agents — on one open model.

That is what makes cross-industry co-design possible.

Modelon

Nithish Selvan · nithish.selvan@modelon.com