Design the Plant and the Control Together

Shrenik Jadhav
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, Florida, USA
jadhavs2@my.erau.edu

Zheng Liu (advisor)
Department of Industrial and Manufacturing Systems Engineering
University of Michigan-Dearborn
Dearborn, Michigan, USA
zhengtl@umich.edu

Vidhyashree Nagaraju (advisor)
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, Florida, USA
nagarav1@erau.edu

Abstract

We built and validated a digital twin of the liquid cooling plant of Frontier, the exascale supercomputer at Oak Ridge National Laboratory, using one year of operating data, and used it in three linked studies. Setpoint changes alone save 27.8% of annual cooling energy on installed hardware, and counting carbon and capital over the service life selects a different plant than energy does.

Keywords

data center cooling, digital twin, exascale computing, life cycle assessment

1 Introduction

Cooling energy in a supercomputer depends on how the plant is controlled and on design choices fixed at construction: how many parallel subloops it has and which coolant distribution units (CDUs) each subloop serves. A public dataset of the Frontier plant [6] lets us study these choices on a real exascale system, where earlier work found about 85 MWh per year of excess cooling energy [4]. We ask three questions. How much theoretical saving survives real actuator limits? Which layout uses the least energy? Which plant should be built once carbon and cost are counted?

2 Study 1: Twin and Control

We built a Modelica model of the Frontier high temperature water loop: three parallel duty subloops, each with a plate heat exchanger and a variable speed pump, plus a standby pump (Figure 1). We validated it against one year of data, 47,186 records, under ASHRAE Guideline 14 [1]; its return temperature error per subloop is 1.96%, 1.97%, and 2.67%, against a 5% target.

Against the measured 1,814 MWh of annual cooling energy, optimizing pump flow alone saves 20.4%, and adding supply temperature reset reaches 30.1%. Limiting changes to 50 kg/s and 1 °C per ten minutes, rates real actuators can follow, keeps 27.8%, or 92.4% of the unconstrained saving. Cooling tower fans use 1,325 MWh of this, and the plant pumps a median 1.5 times the flow the heat load needs, about 3.4 times the necessary pump power. Details are in [3].

3 Study 2: Plant Layout

Frontier has 25 CDUs. We enumerated every way to divide them into two to six subloops, 611 layouts, and evaluated each over 49,353 ten minute steps. For each layout the optimizer sets the share of flow per subloop, and at every step the total flow and supply temperature, with each subloop return kept at or below 42 °C.

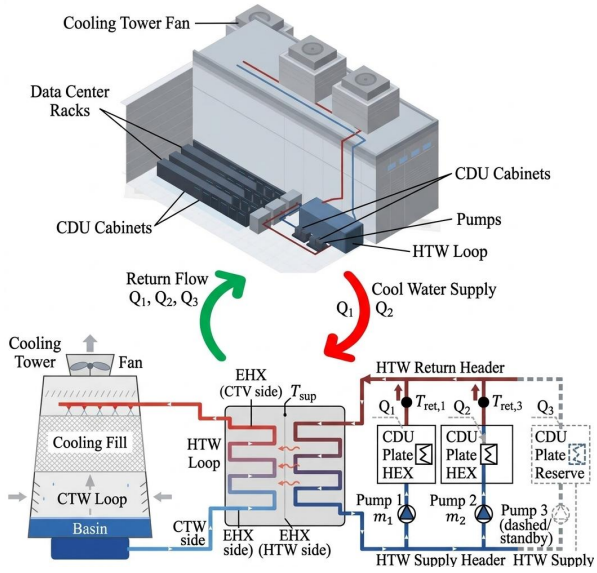


Figure 1: Frontier cooling plant reproduced by the twin, with three duty subloops, their pumps, and a standby pump.

Enumeration guarantees the best layout is found, and sequential least squares programming solves the continuous part.

In this setting the best layout, two subloops carrying 19 and 6 units, saves 35.48%, while the plant as built, carrying 14, 6, and 5 units, saves 35.30%. The layout decision is worth only 0.18 percentage points. The flow split matters more, cutting the penalty of an uneven unit assignment by 93%, from 96 to 7 MWh per year. Frontier already runs its subloop pumps independently, so near optimal operation needs only a software change. Details are in [2].

4 Study 3: Life Cycle

We then priced every layout over seven years, adding embodied carbon, capital, maintenance, and downtime to the Study 2 operating energy.

Counting duty subloops only, the best plant has two subloops carrying 14 and 11 units: 3,320.7 t CO_2e and \$3,987k over seven years, 35.7 t and \$63k below the plant as built. The gap is embodied, not operational, because each added subloop repeats trunk and header piping. The energy optimum from Study 2 uses 0.11 MWh less per year but emits 71.8 t more. It stays ahead of every design with three or more subloops across a ±40% range in embodied carbon factor and a factor of two in grid intensity.

A standby train changes the answer. Sized to the largest duty subloop, it costs most when subloops are few, so with one standby priced in the best designs have four or five duty subloops. The plant as built, at 3,807.6 t, has a duty count close to this, but its uneven split needs a large spare; the best three subloop split is 163.8 t lower. The standby cuts expected unplanned downtime from 4.90 to below 0.015 hours per year. Details are in [5].

5 Conclusion

Each study supplies the next: the validated twin makes the 611 layout search affordable, and its optimizer supplies the operating energy for the life cycle study. Three lessons follow. Improve control first, since setpoints alone capture most of the saving on installed hardware. Without redundancy, prefer fewer, larger duty subloops, with flow split control to remove assignment risk. When availability matters, get it from a spare, and keep the duty split even so the spare stays small.

Acknowledgments

This work was supervised by Zheng Liu and Vidhyashree Nagaraju. Following the SC26 policy on AI generated text, AI tools were used for language and quality refinement of this text, then reviewed and edited by the author, who is fully accountable for its content.

References

[1] ASHRAE. 2014. *Guideline 14-2014: Measurement of Energy, Demand, and Water Savings*. Technical Report. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, USA.

[2] Shrenik Jadhav and Zheng Liu. 2026. Co-Design Optimization for Data Center Cooling System via Digital Twin. arXiv preprint arXiv:2605.15516. doi:10.48550/arXiv.2605.15516

[3] Shrenik Jadhav and Zheng Liu. 2026. Digital Twin-Based Cooling System Optimization for Data Center. arXiv preprint arXiv:2603.01198. doi:10.48550/arXiv.2603.01198

[4] Shrenik Jadhav and Zheng Liu. 2026. Machine Learning Guided Cooling System Optimization for Data Center. arXiv preprint arXiv:2601.02275. doi:10.48550/arXiv.2601.02275

[5] Shrenik Jadhav, Vidhyashree Nagaraju, and Zheng Liu. 2026. Life Cycle Co-Design of the Cooling Subloop Architecture for Data Centers. Manuscript. An earlier version is available as arXiv:2606.15408.

[6] Jian Sun, Zhiming Gao, David Grant, Kashif Nawaz, Pengtao Wang, Cheng-Min Yang, Philip Boudreaux, Stephen Kowalski, and Shean Huff. 2024. Energy dataset of Frontier supercomputer for waste heat recovery. *Scientific Data* 11 (2024), 1077. doi:10.1038/s41597-024-03913-w