JUNE 3 2025

More for Less: Integrating Capability-Predominant and Capacity-Predominant Computing

Zhong Zheng[1], Michael E. Papka[1,2] and Zhiling Lan[1,2]
[1]University of Illinois Chicago, Chicago IL, USA
[2]Argonne National Laboratory, Lemont IL, USA

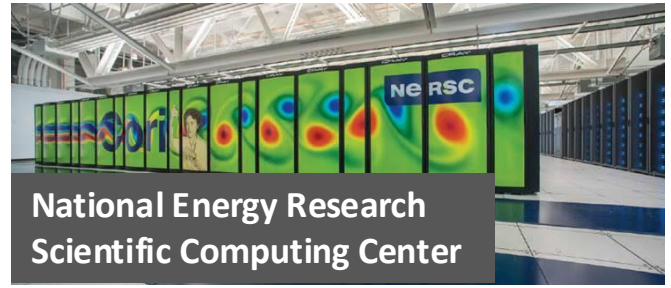
JSSPP 2025

Department of Energy ASCR Facilities

- **Enable World-Class Scientific Discovery** - Provide cutting-edge supercomputing, data, and networking (via ESnet) that empower breakthrough research across disciplines, from materials science to climate modeling.
- **Advance Computational Science and Algorithms** - Drive the development of next-generation computational methods, algorithms, and software to solve complex scientific and engineering problems.
- **Deliver Capability and Capacity for National Priorities** - Balance leadership-class facilities (LCFs) for the largest, most complex simulations (capability) with centers like NERSC supporting broad, high-throughput workloads (capacity), all underpinning U.S. national security, energy, and innovation.
- **Support and Grow Diverse Research Communities** - Serve thousands of users across national labs, universities, and industry, accelerating discovery through open access, partnerships, and community engagement.
- **Push the Frontiers of Exascale, AI, and Quantum** - Lead the exploration of exascale computing, artificial intelligence, quantum computing, and ultra-fast networking to shape the future of scientific computing.

Argonne Leadership Computing Facility

Oak Ridge Leadership Computing Facility



National Energy Research Scientific Computing Center


Energy Sciences Network

Capability versus Capacity Predominant Systems

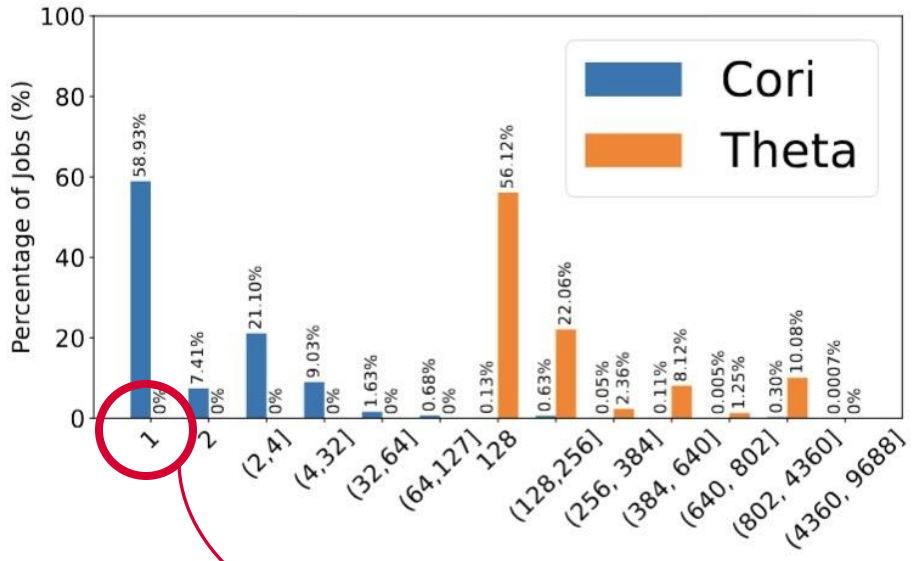
Capability systems (large-sized, long-running jobs)

- Theta @ ALCF

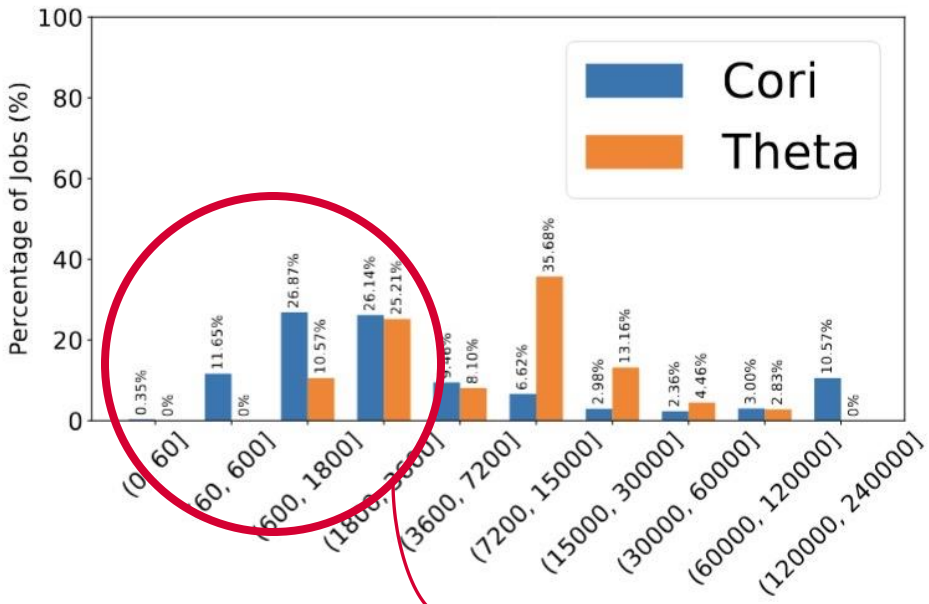
Capacity systems (small-sized, short-running jobs)

- Cori @ NERSC

Theta@ALCF versus Cori@NERSC



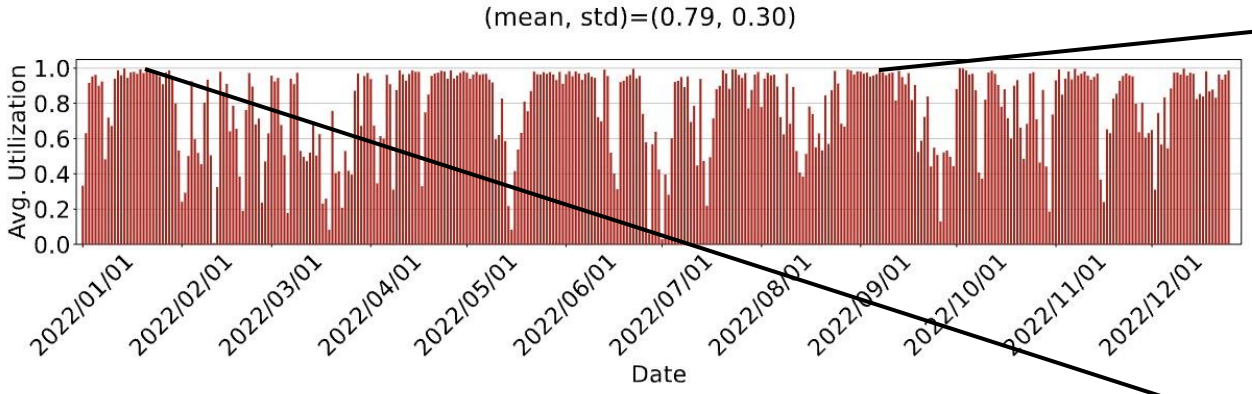
(a) Job distribution by job sizes (number of nodes)



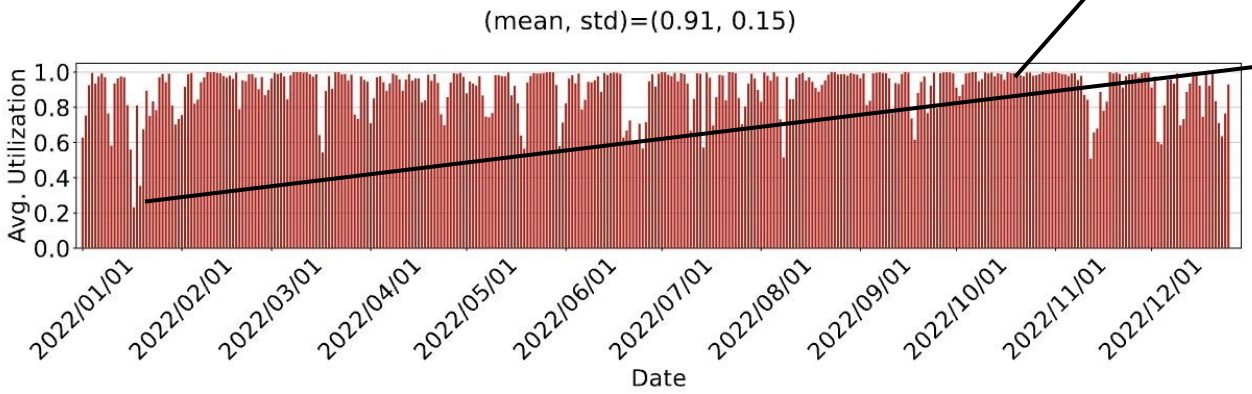
(b) Job distribution by job runtimes (seconds)

- Around 59% of Cori jobs request only 1 compute node
- 60% of Cori jobs run for less than 1 hour
- Jobs on Theta are large (more than 128 nodes) and long-running (64% of jobs run for over one hour)

Resource Utilization



(a) Theta @ ALCF



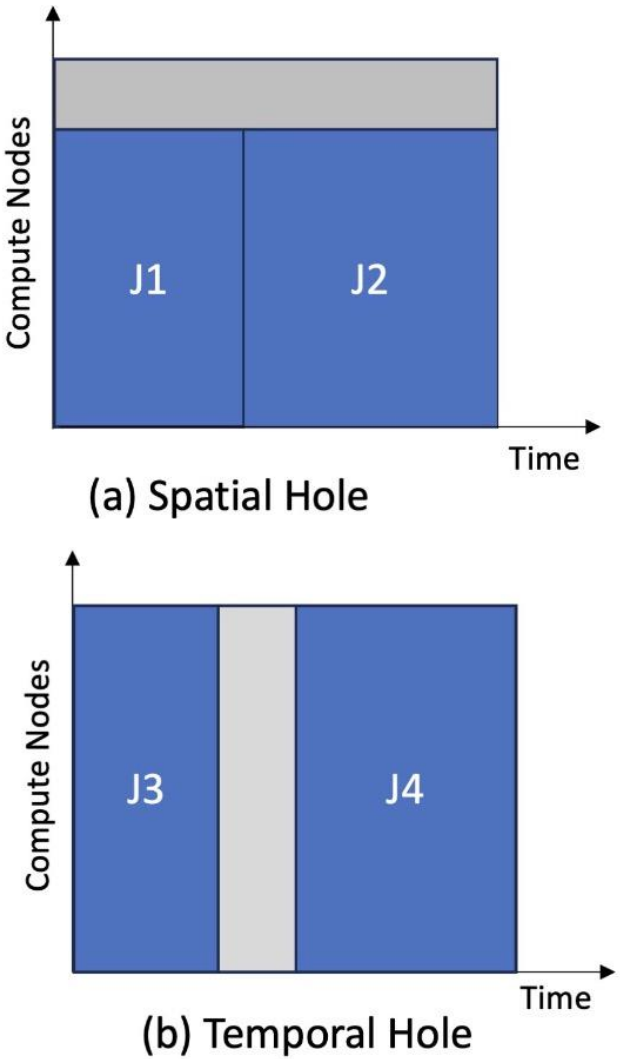
(b) Cori @ NERSC

Both capability and capacity systems can achieve high system utilization

High usage on one system may coincide with low usage on another

Resource Utilization

- Resource fragmentation in capability systems
  - In (a), J1 and J2 occupy most of the computing nodes for a given period, leaving only a small chunk of computing resources
  - In (b), there is a short period between the execution of J3 and J4



Problem Statement

- Existing approaches
  - Enhancing scheduling policies
  - Migrating high-throughput jobs or Function-as-a-Service (FaaS) jobs to capability computing systems
  - Adopting cloud bursting to address escalating workloads
- This study explores **what-if analysis,** co-running capability and capacity jobs within a unified system
  - **Workload fusion**
  - **Workload injection**

UIC

What-If Analysis via Simulation

- We extend the open-source simulator **CQSim**
  - A discrete-event driven scheduling simulator
  - Real job traces were fed to CQSim based on job arrival times
  - Scheduling policy: FCFS with EASY backfilling

- Analysis metrics
  - Resource utilization (system —— level metric)
  - Job wait time (user — level metric)

CQSim Link: https://github.com/SPEAR-UIC/CQSim

Workload Fusion

- A hypothesized unified system (14,048 nodes)
  - Trace-based simulation
  - Consolidate two workloads based on job arrival time
  - Downsize the unified system to 13,345 (95%), 12,643 (90%), 11,940 (85%), and 11,238 (80%) nodes, while keeping the workloads unchanged
- Research questions:
  - What impacts could arise from accommodating both capability and capacity computing workloads within a unified platform
  - To what extent can the size of the unified system be reduced to lower total ownership costs without compromising computational capabilities for a mix of jobs

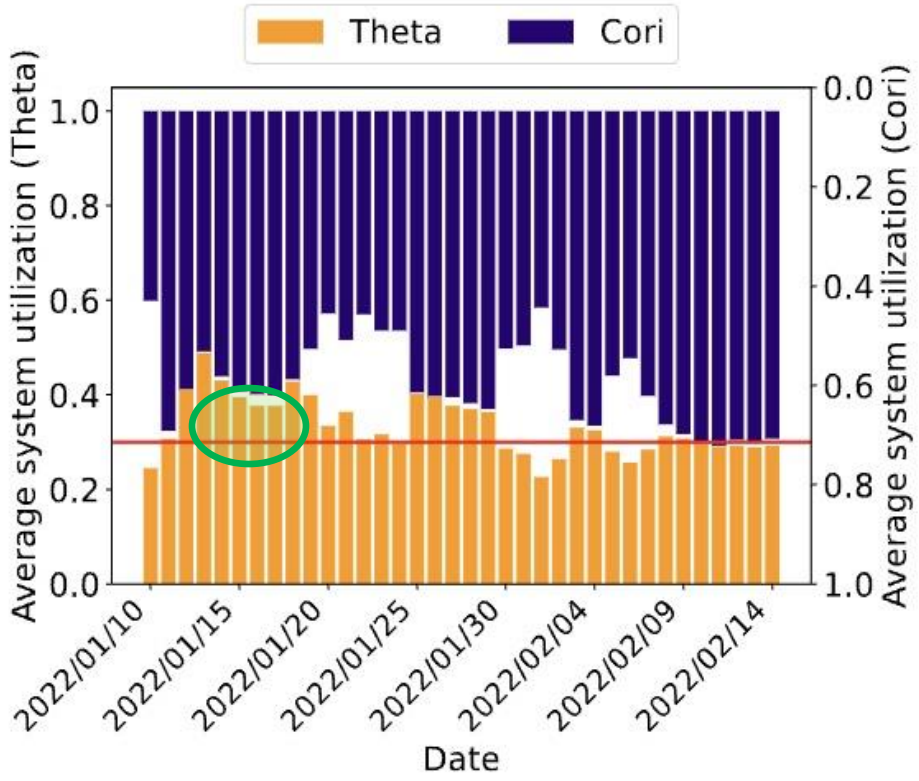
TABLE I: Theta and Cori Workloads

	Theta	**Cori**
Location	ALCF	NERSC
Category	Capability	Capacity
Scheduler	Cobalt	Slurm
Machine Type	Cray XC40	Cray XC40
Compute Nodes	4,360	9,688
Processor	Intel Xeon Phi 7230 64-core	Intel Xeon Phi 7250 68-core
Interconnect	Dragonfly	Dragonfly
Trace Period	1/1/2022-12/31/2022	1/1/2022-12/31/2022
Number of Jobs	23,911	2,349,370
Min Job Size	128 nodes	1 node

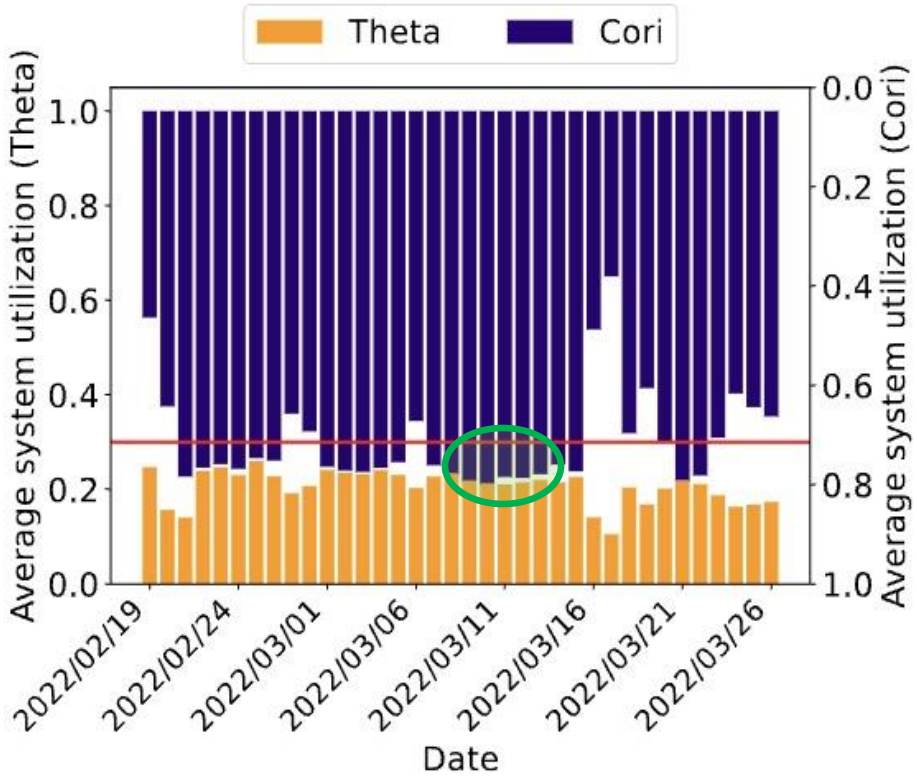
	a hypothesized unified system
Category	General purpose
Compute Nodes	14,048
Trace Period	1/1/2022 – 12/31/2022
Number of Jobs	2,373,281
Min Job Size	1 node

Workload Fusion

- The red horizontal line marks the boundary between the two systems

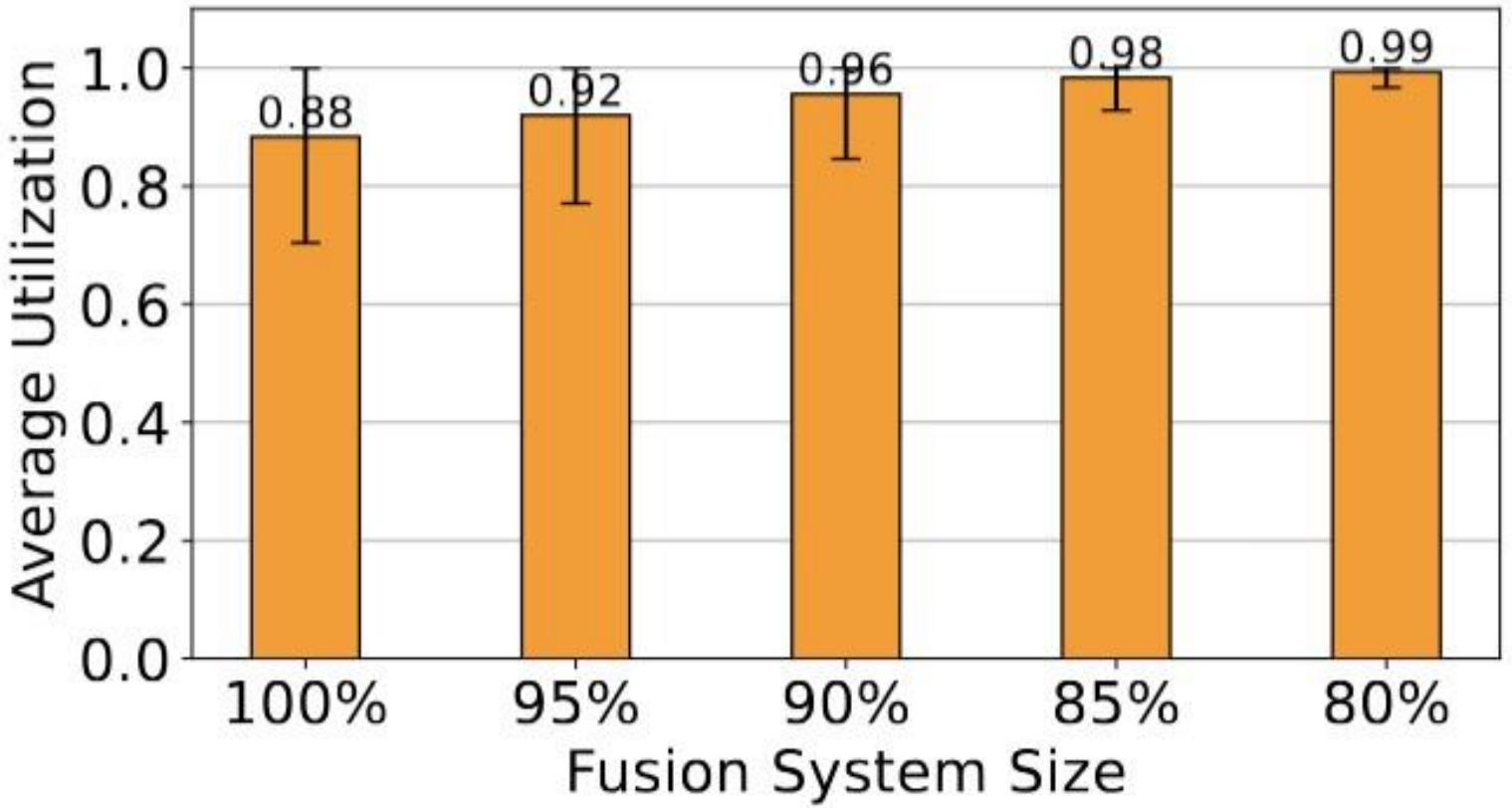


Theta jobs benefit from the unified system.



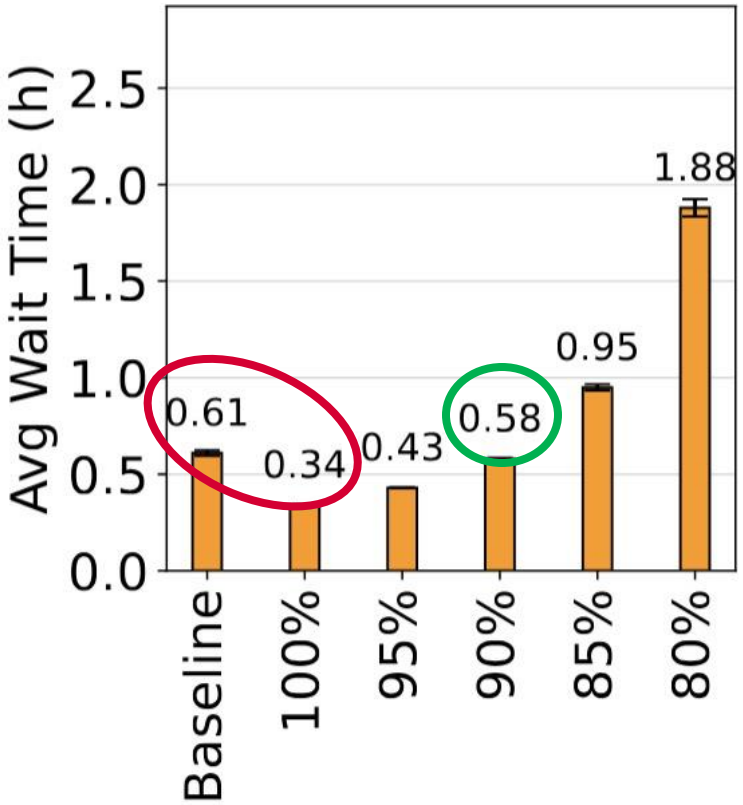
Cori jobs benefit from the unified system.

Impacts on Resource Utilization

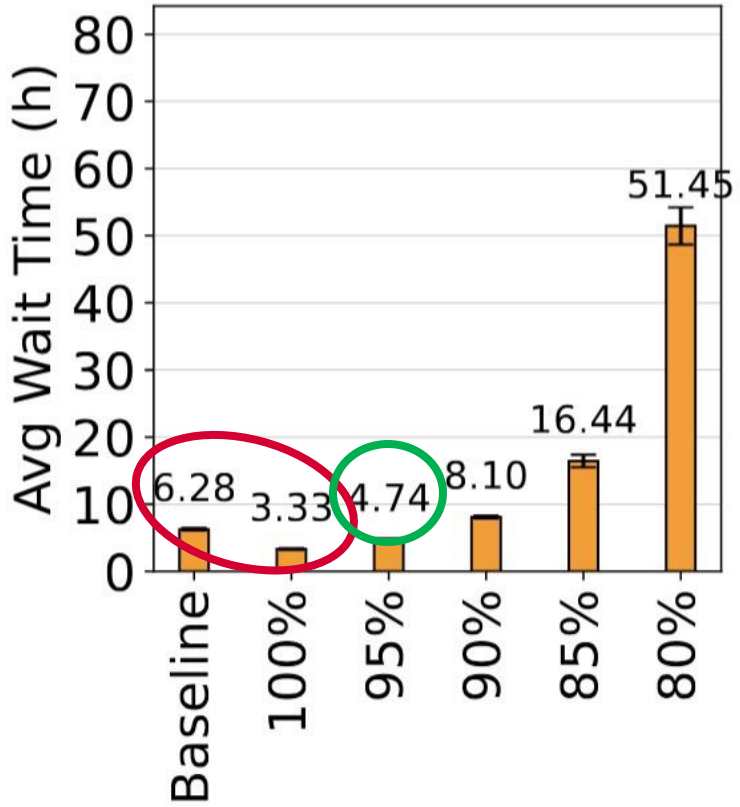


By reducing the system size, the system utilization is improved approaching to **99%**

Impacts on Job Wait Time



Cori job wait time

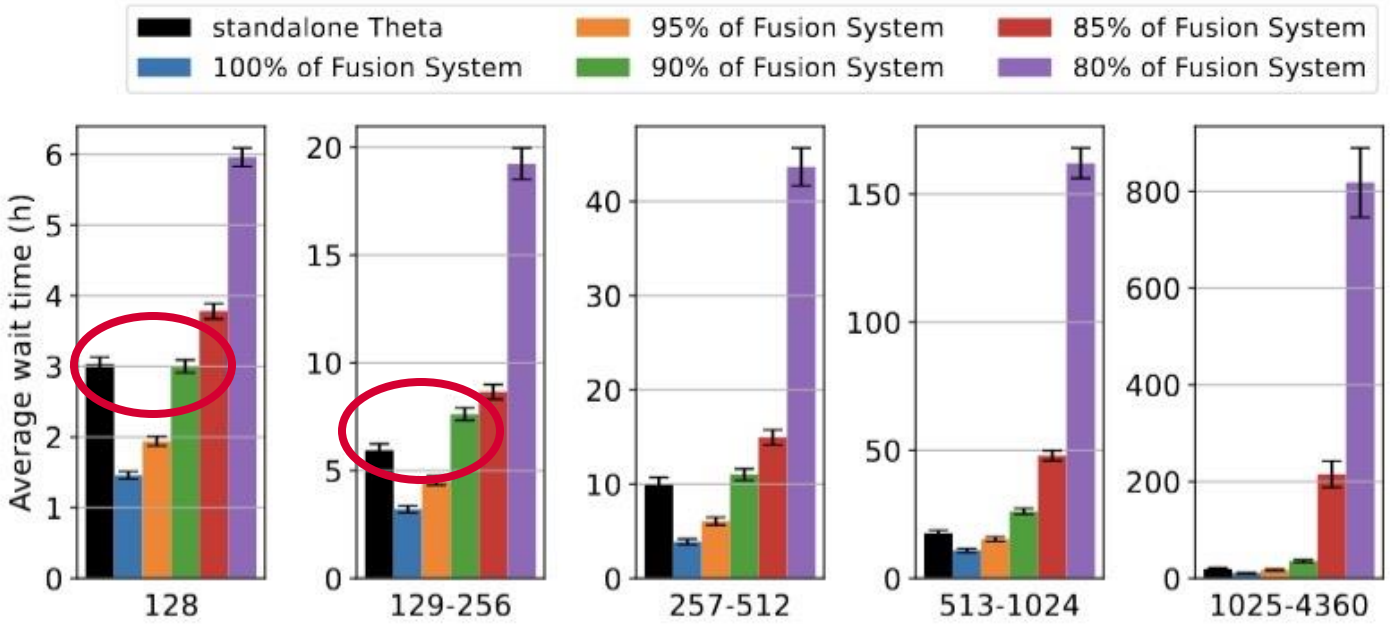


Theta job wait time

1. Integration significantly reduces job wait time — up to twice as fast

2. Cori and Theta Job wait time, remains unaffected compared to baseline when the unified system is downsized by up to 10%, and 5% respectively

Impacts on Job Wait Time



Theta workload by different job sizes

Jobs of varying sizes exhibit different degrees of sensitivity to changes in the size of the unified system

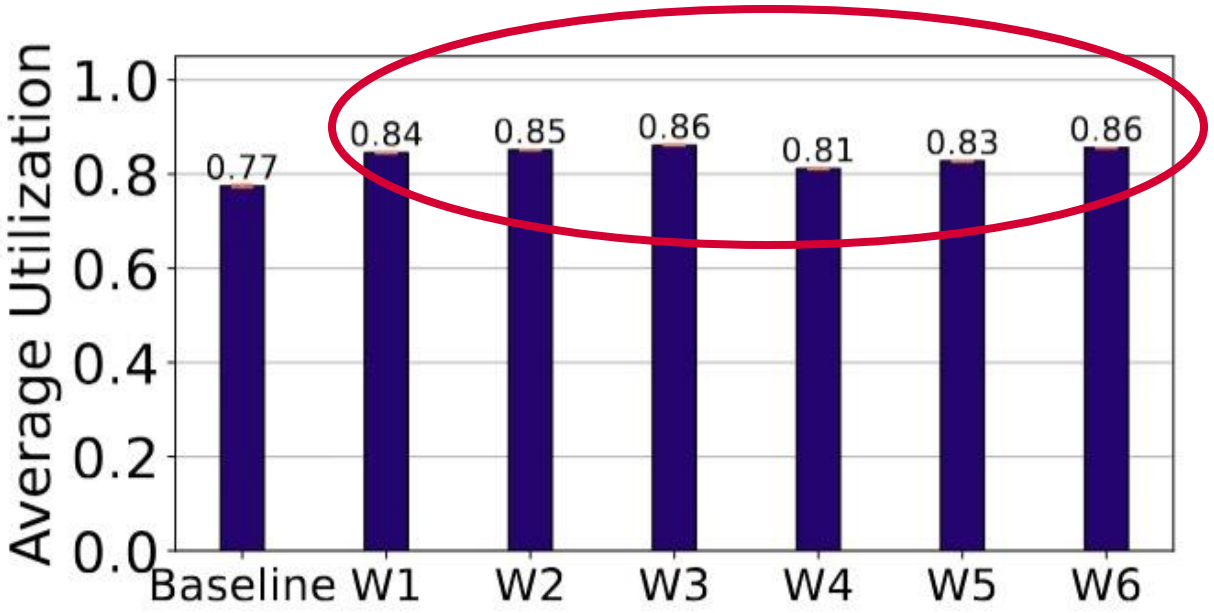
Smaller jobs are less sensitive to changes in the size of the unified system, while larger jobs are more affected

Workload Injection

- Injecting Cori jobs to Theta
  - Select **small-sized** (W1-W3) and **short-running** (W4-W6) jobs from Cori as backfilling jobs
  - Study the impact of workload injection on Theta
- Research questions:
  1. What are the potential impacts of accommodating additional capacity jobs on a capability computing system
  2. How can we strategically select capacity jobs without compromising the performance of capability computing jobs

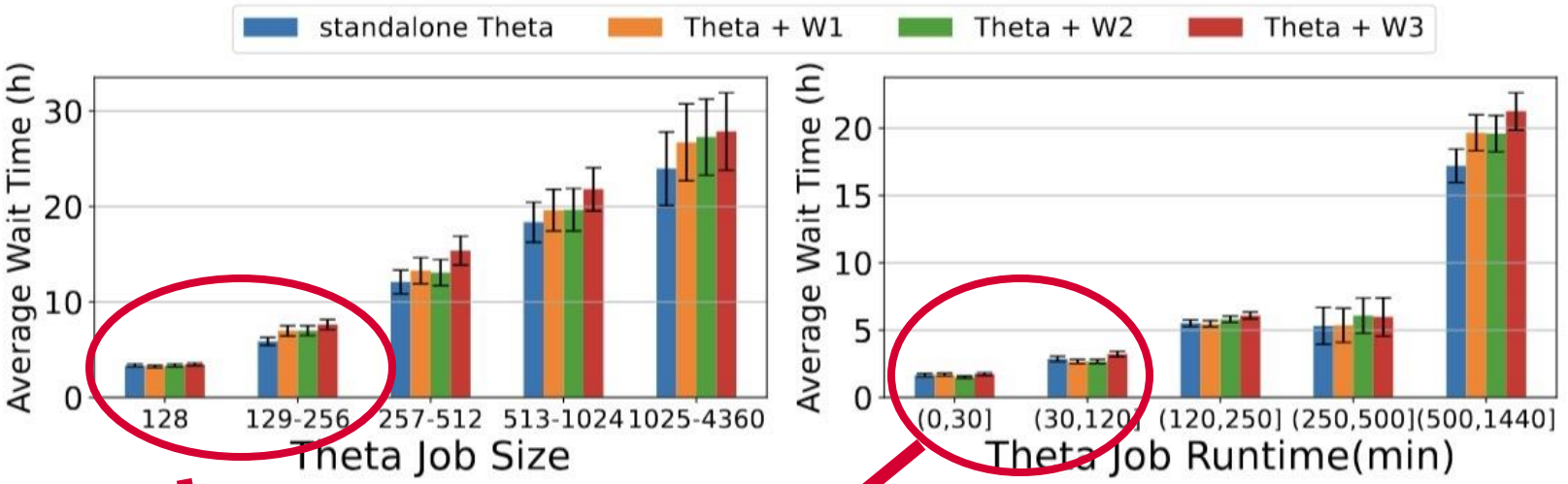
	Job size (nodes)	Job runtime (min)	Job counts
W1	≤ 128	≤ 30	521,574
W2	≤ 128	≤ 45	629,475
W3	≤ 128	≤ 60	803,285
W4	≤ 4096	≤ 10	133,914
W5	≤ 4096	≤ 29	268,959
W6	≤ 4096	≤ 30	528,243

Impacts on Resource Utilization



Both small-sized and short-running capacity jobs can improve the capability system resource utilization by up to 9%.

Impacts on Job Wait Time



(a) Job wait time under W1,W2,W3

Capability jobs on Theta, with varying sizes and runtimes, respond differently to the injection of capacity jobs.

Small size and short running jobs are impacted minimally by injected Cori jobs

Conclusion

- Workload fusion
  - Integration significantly reduces average job wait time — up to **twice** as fast
  - There is no adverse impact on job wait time for either Theta or Cori workloads compared to the baselines with reducing the unified system size to **90%**

- Workload injection
  - Both small-sized and short-running capacity jobs can improve the resource utilization by up to **9%**
  - Backfilling **133K** shorting running Cori jobs into Theta system introduce negligible Theta job average wait time increasing, meanwhile increasing the Theta utilization by **4%**

UIC

Thank you!

This work is supported in part by US National Science Foundation grants CCF-2413597, OAC-2402901, and the US Department of Energy under Contract No. DE-SC0024271.