

A Case for Intra-Rack Resource Disaggregation for HPC

George Michelogiannakis,
Benjamin Klenk, Brandon Cook,
Min Yeh Teh, Madeleine Glick,
Larry Dennison, Keren Bergman,
John Shalf



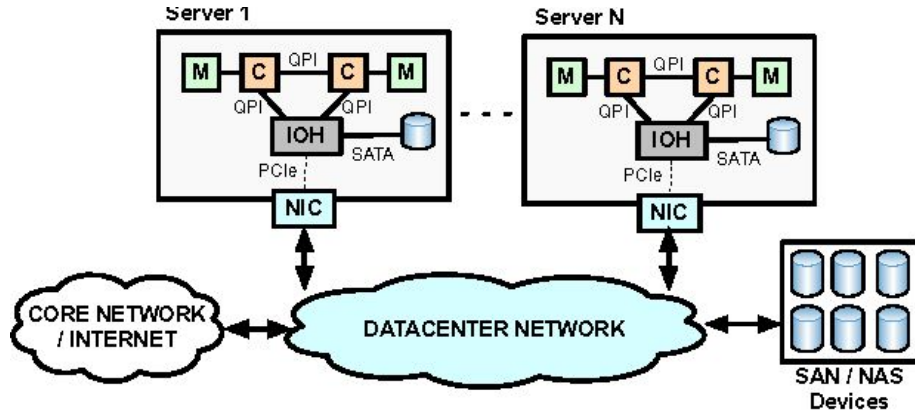
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK



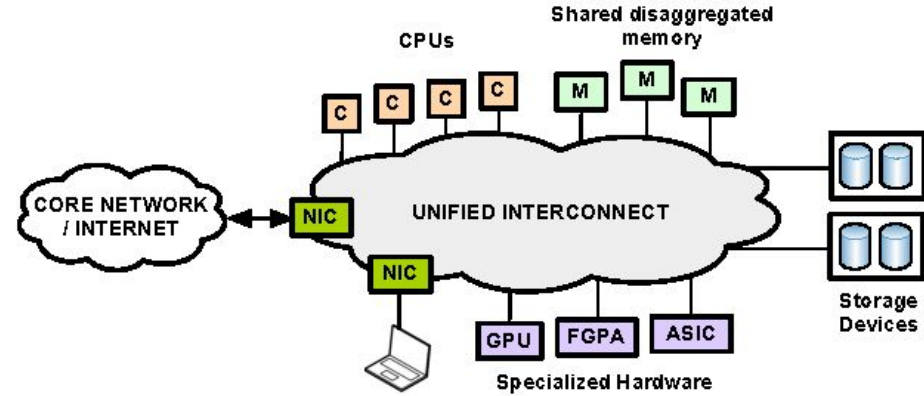
PINE



System-Wide Resource Disaggregation



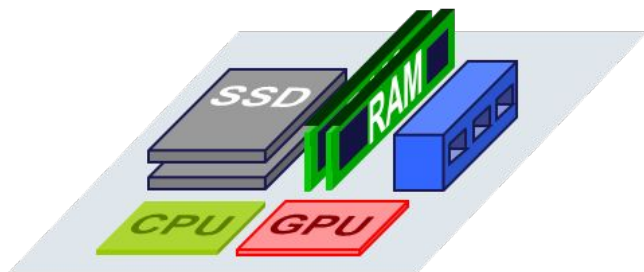
(a) Current datacenter



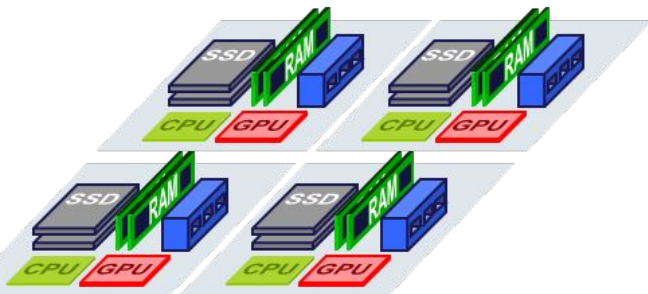
(b) Disaggregated datacenter

Intra-Rack Resource Disaggregation

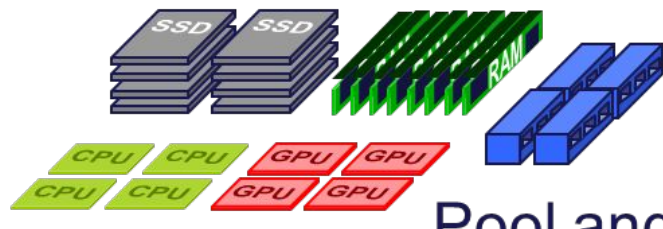
Current server



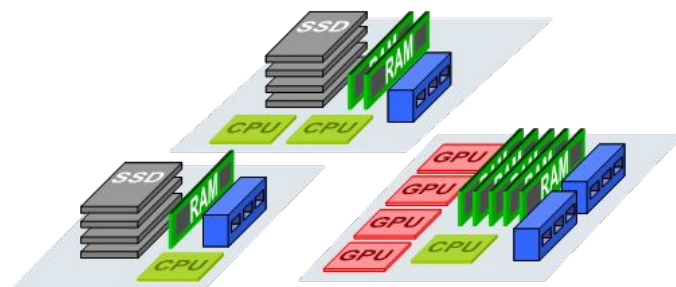
Current rack



Disaggregated rack



Pool and
compose



Questions

- Is resource disaggregation promising for HPC?
- If so, at what range?
 - Node, rack, system

We adopt a system analysis approach

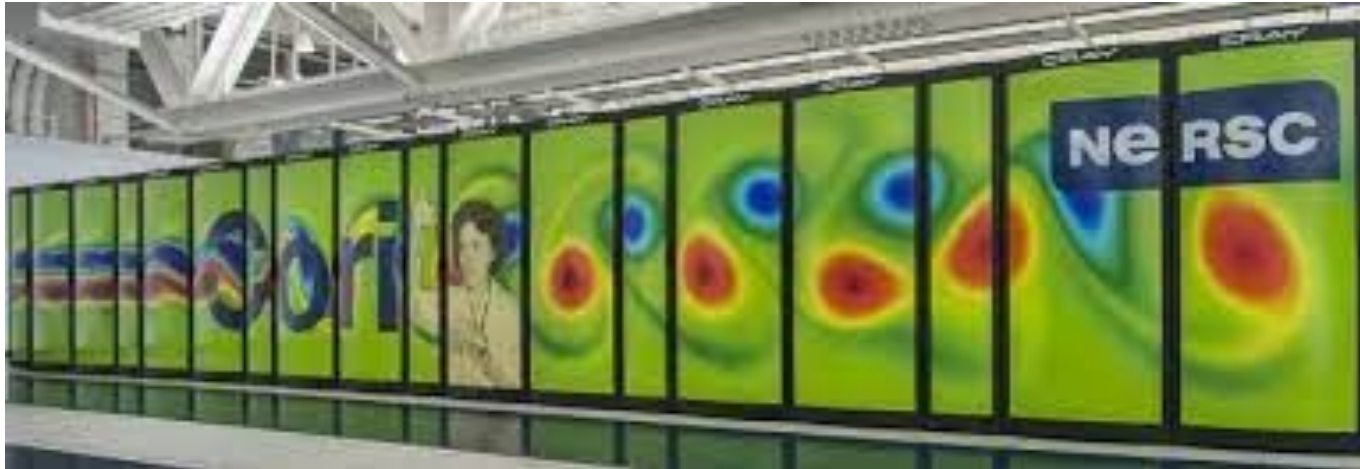
Resource Usage Analysis on NERSC's Cori

Cori: top 20 system. 192 nodes per rack

2,300 Haswell nodes. 128 GB mem

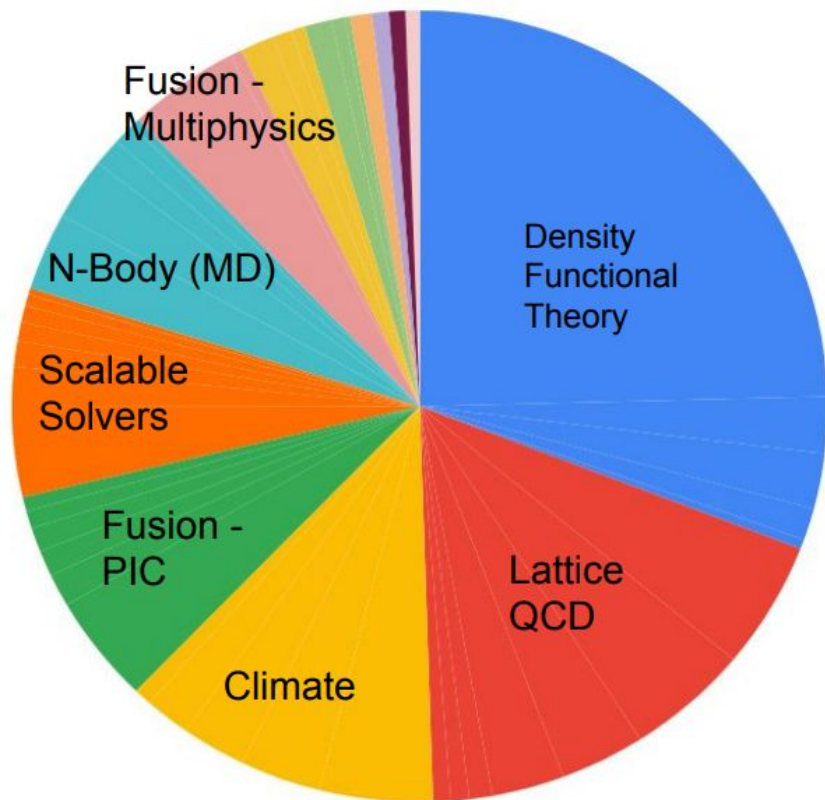
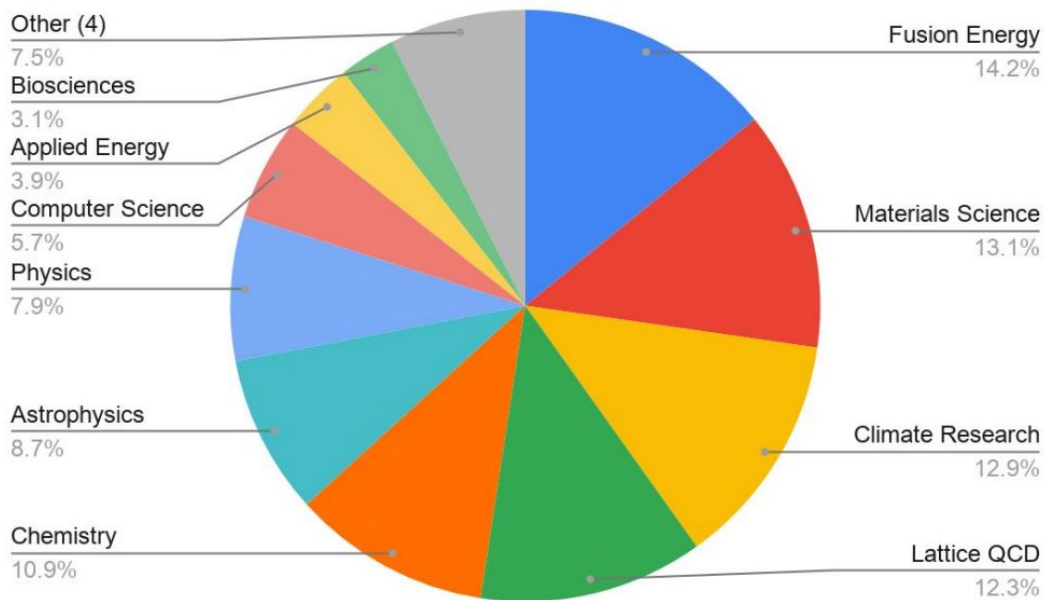
9,688 KNL nodes. 112 GB mem

Data over three weeks. Sample every 1s



Cori Ran the Entire NERSC Workload

NERSC: ~1k projects. ~8k users



Do jobs fit inside a rack?

Most Jobs Fit Inside a Rack

We sample the size of jobs an individual node is assigned to

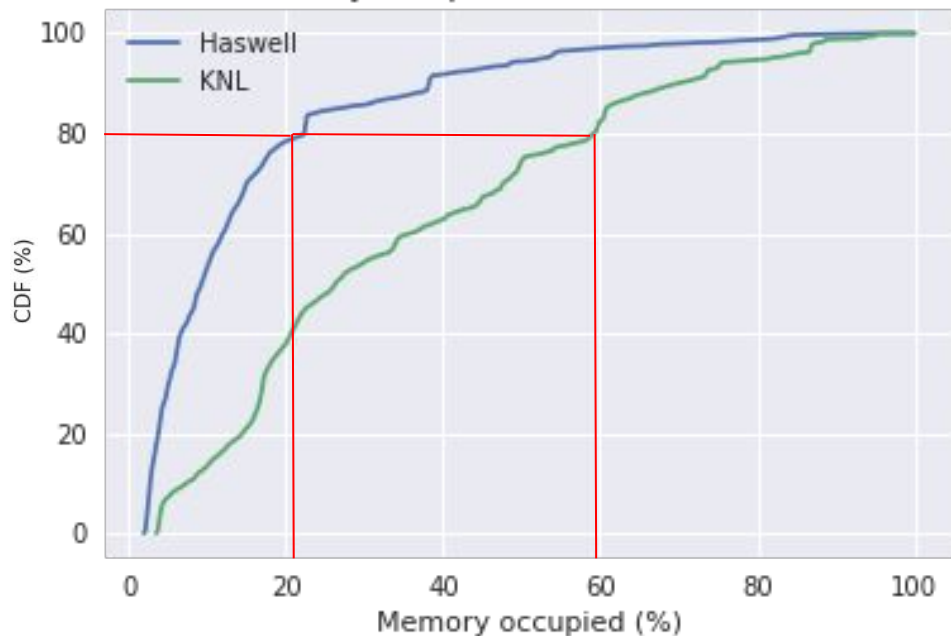


Size of job in nodes

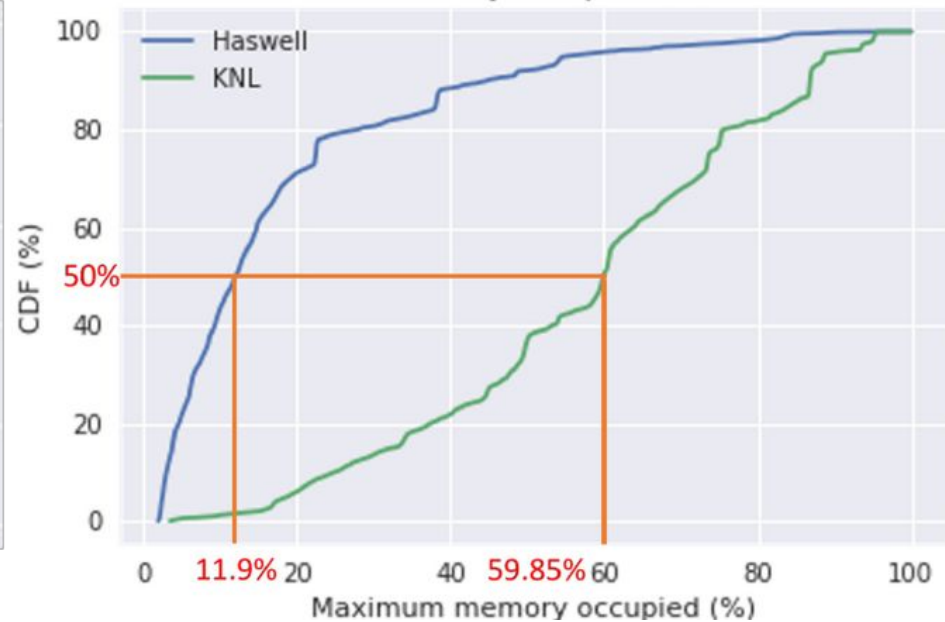
How intensely are resources utilized?

Memory Occupancy is Low Usually

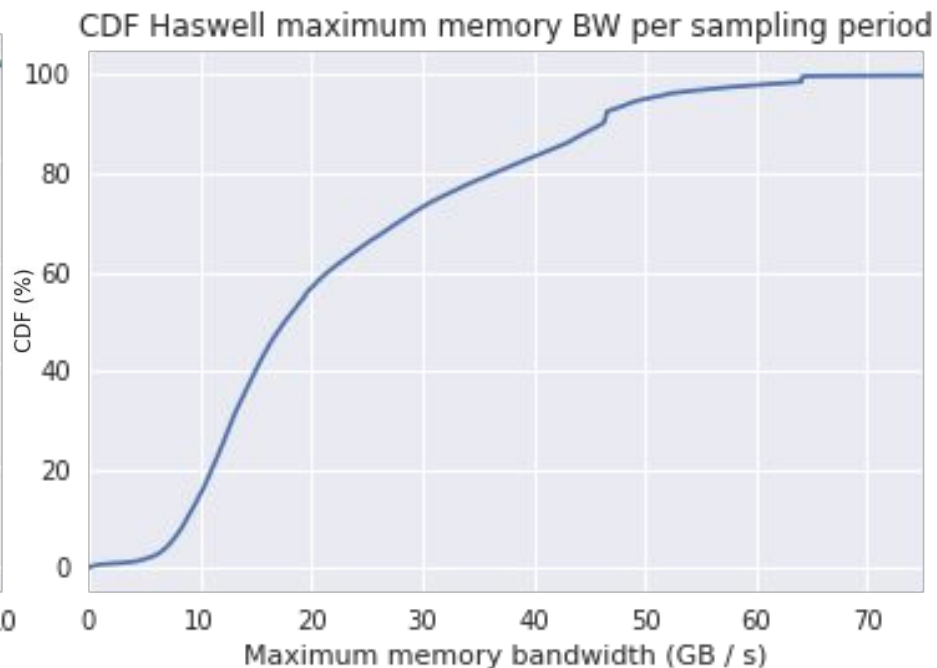
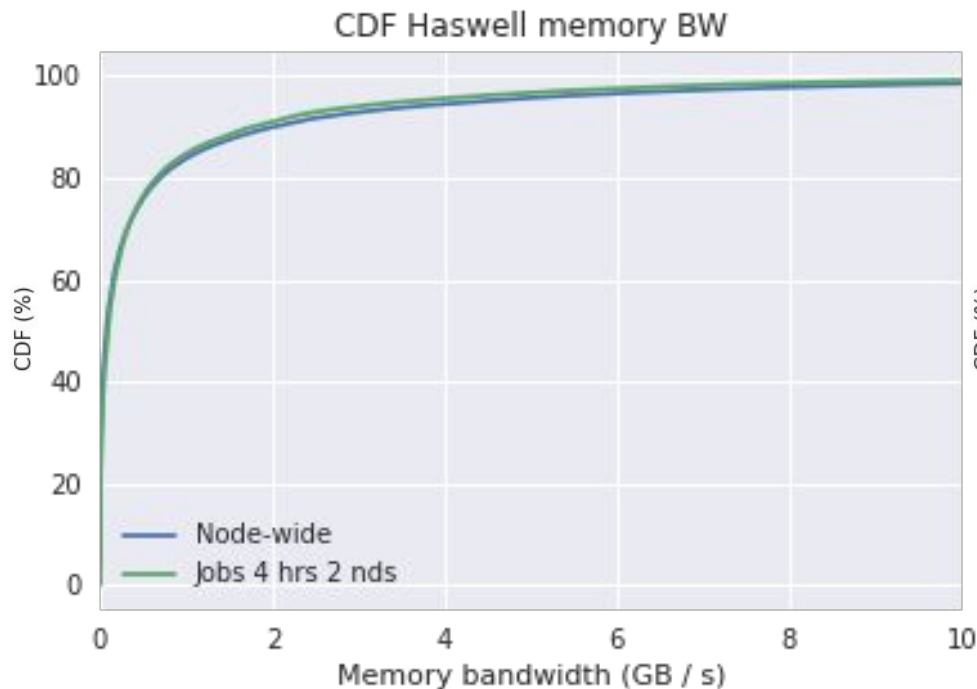
CDF memory occupied (%) node-wide statistics



CDF Maximum memory occupied (%) across nodes



Memory Bandwidth Utilization Also Usually Low



This does **not** mean reducing memory bandwidth won't have an impact to applications

Other resources

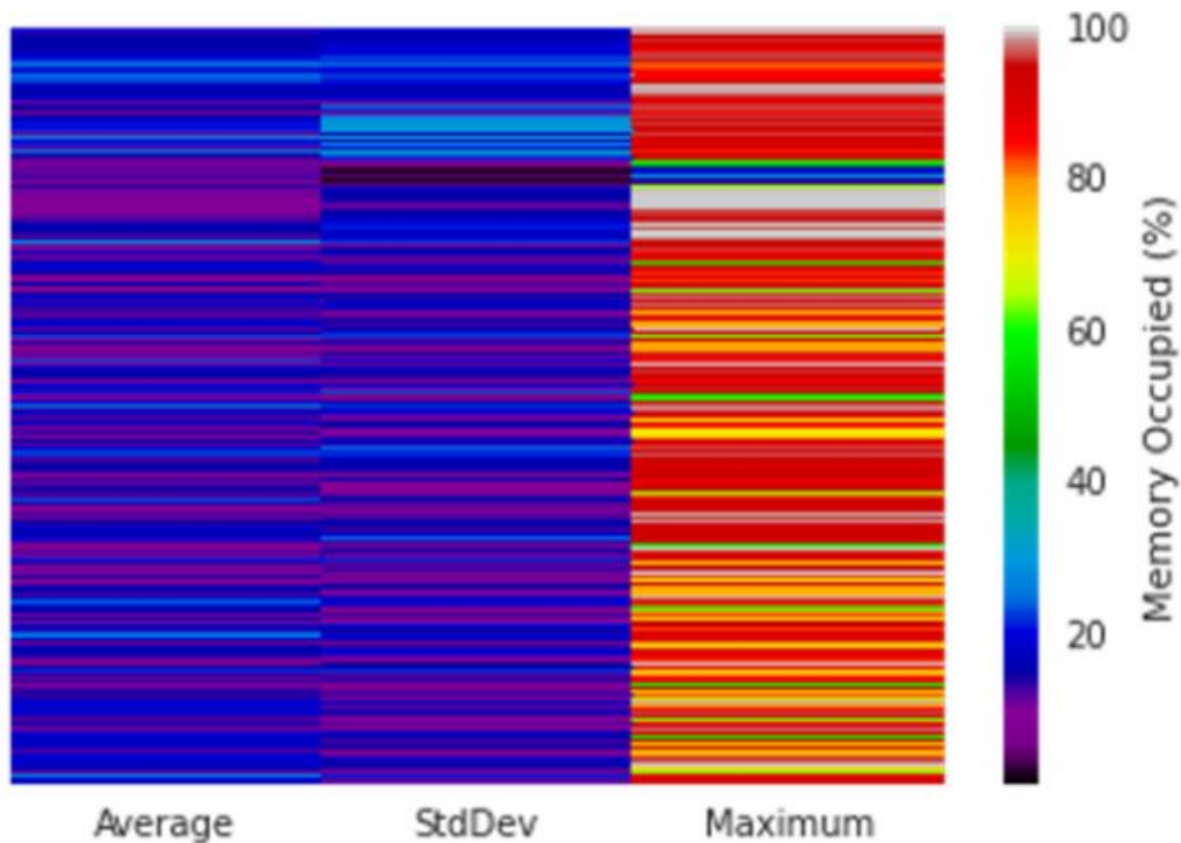
NIC bandwidth is even less utilized, but bursty

CPU idle % is high

Is job scheduling helping resource utilization?

There is Spatial Variation in Memory Usage

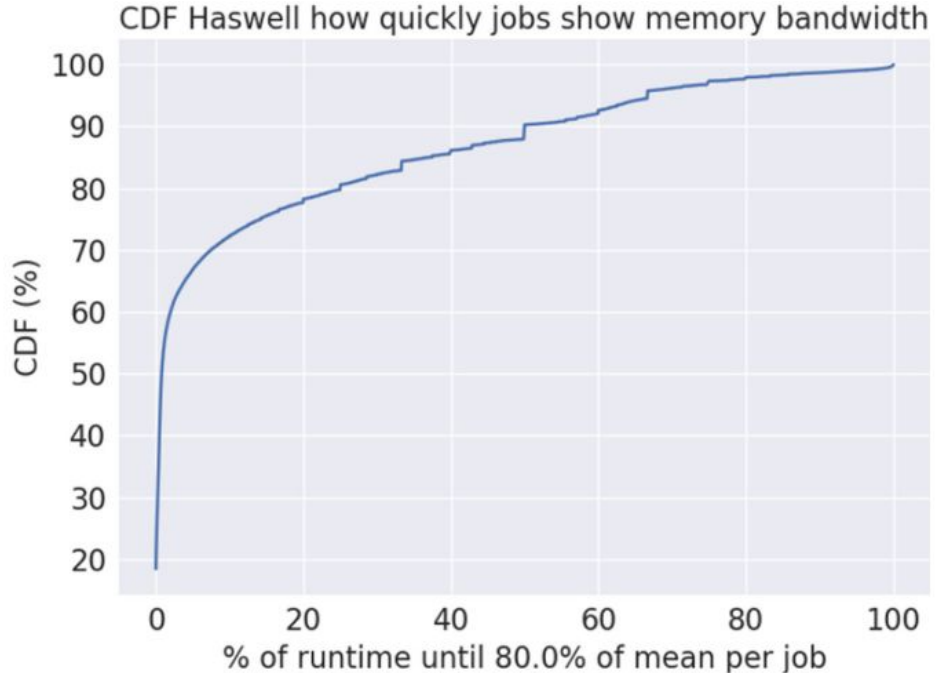
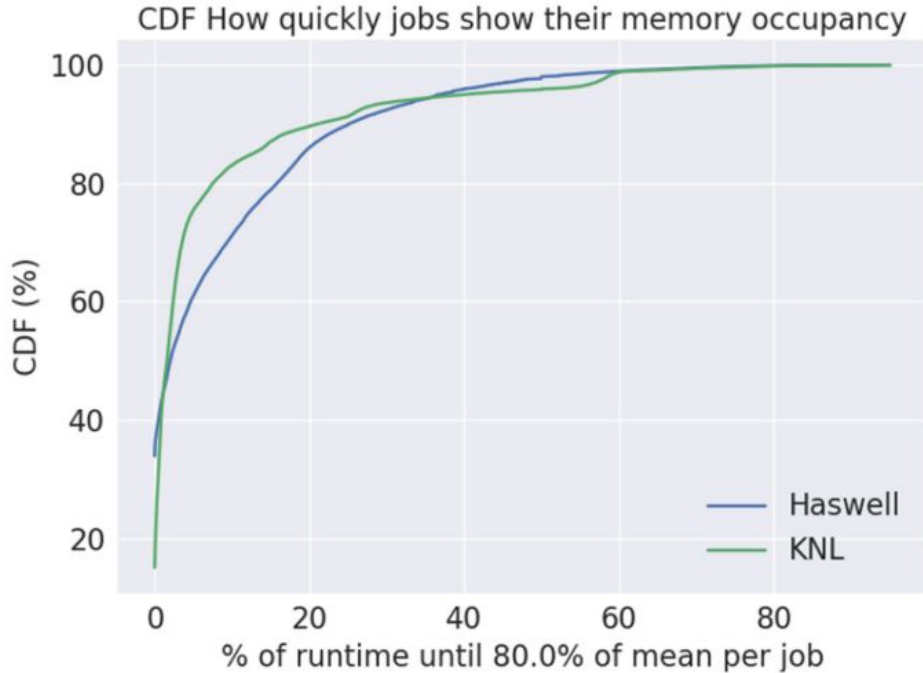
Haswell



When does a job become predictable?

Majority of Jobs Reach Steady State After 20% of Runtime

Definition: When does a job's utilization reach 80% of lifetime mean



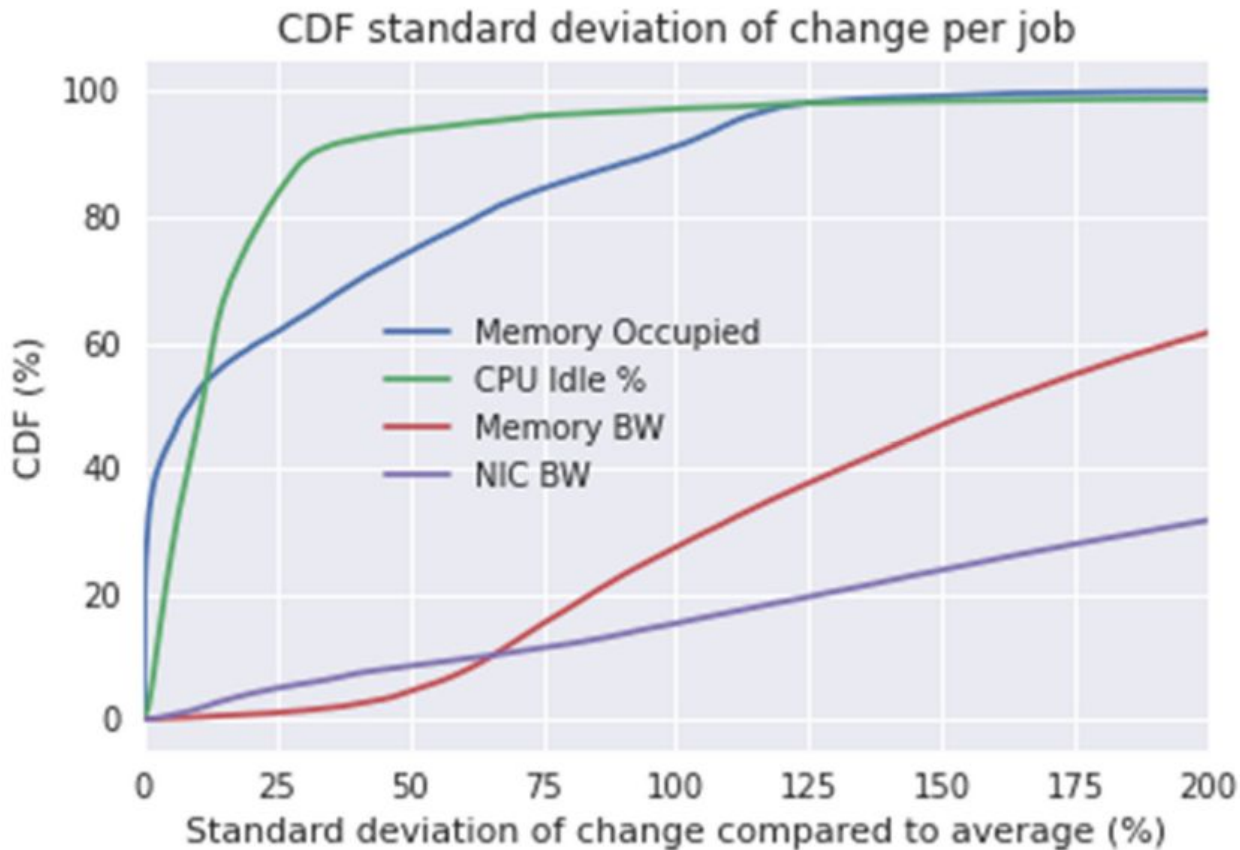
How much does utilization change?

CPU Idle And Memory Occupied Don't Change Much

$\frac{\text{StanDev}}{\text{Average}}$

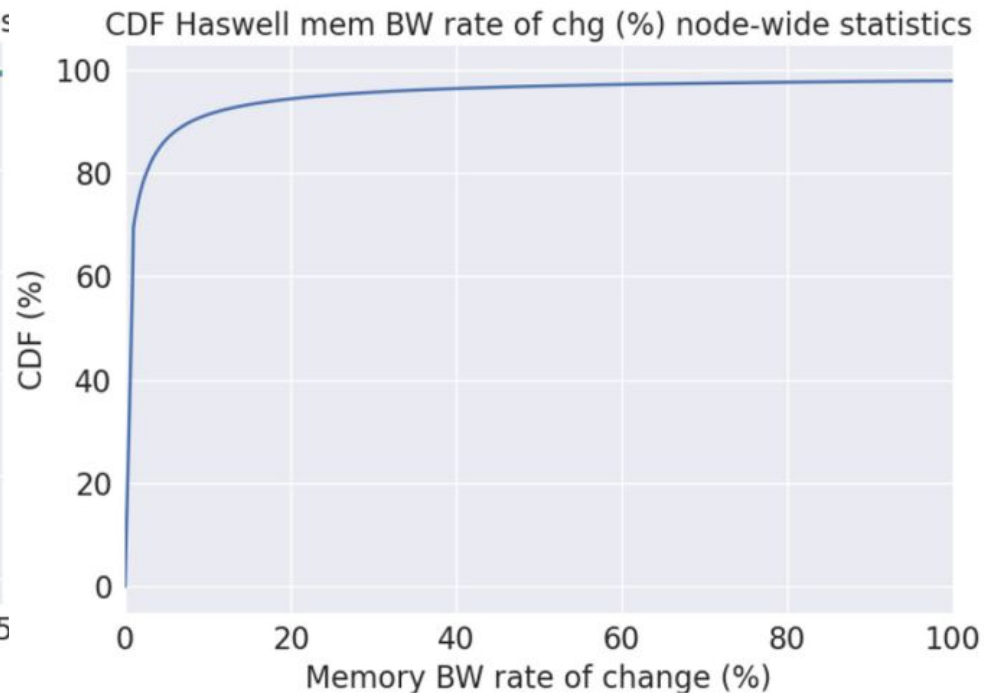
As expected, NIC and memory bandwidths are more bursty

This helps understand what utilization (average, max, eg) we should aim for



How quickly does utilization change?

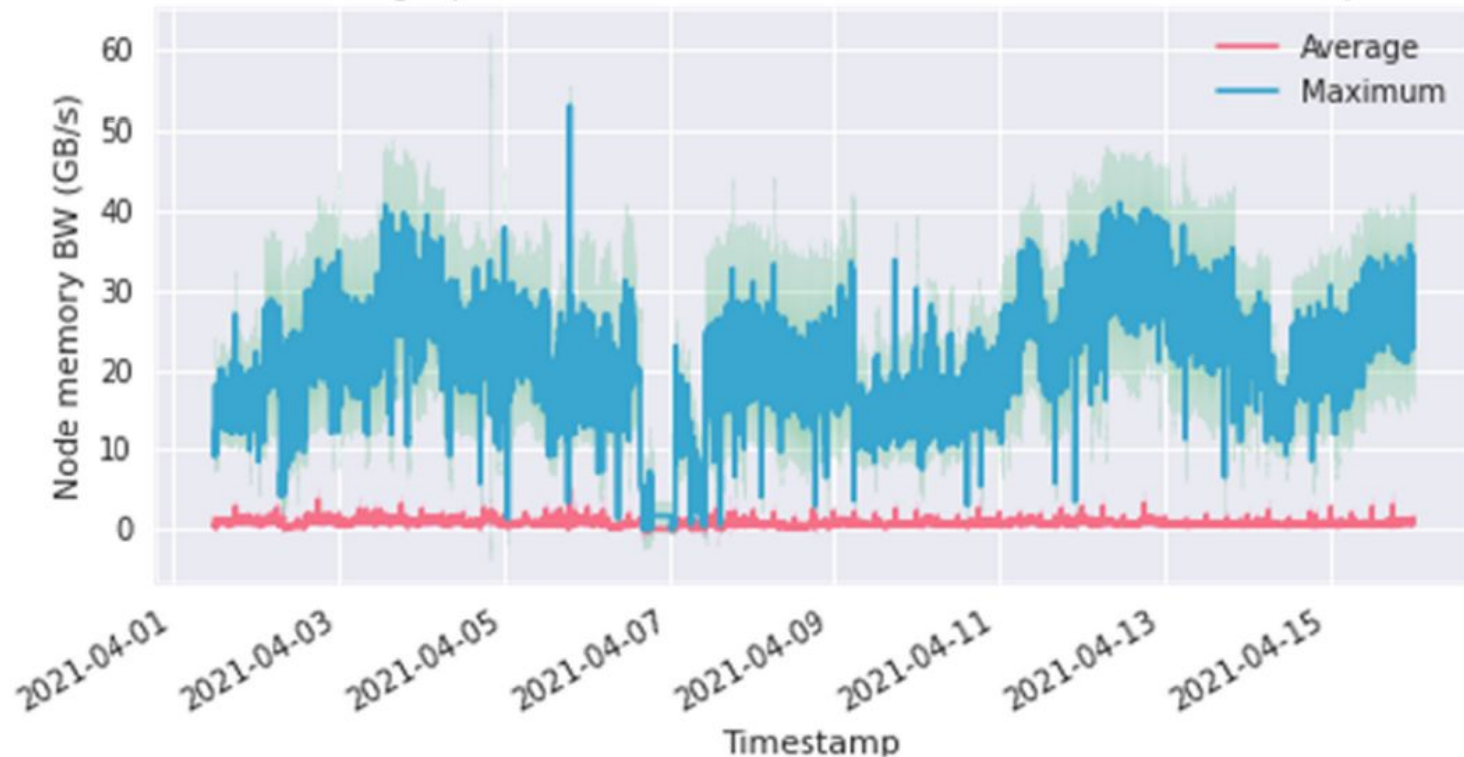
Large Changes Are Infrequent But They Happen



Temporal variation across the system

Memory Bandwidth Across Haswell Nodes

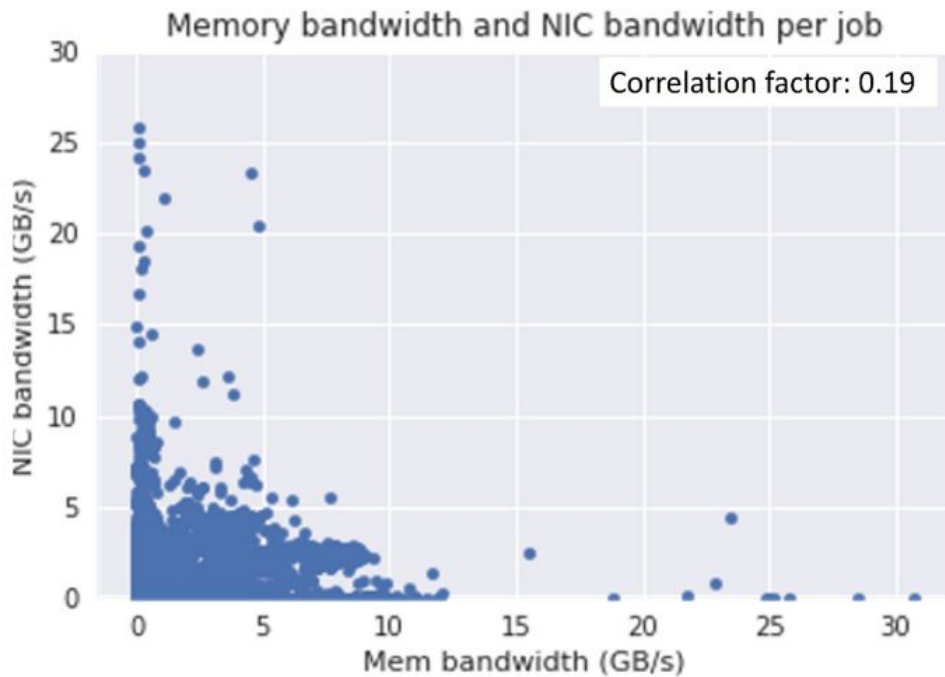
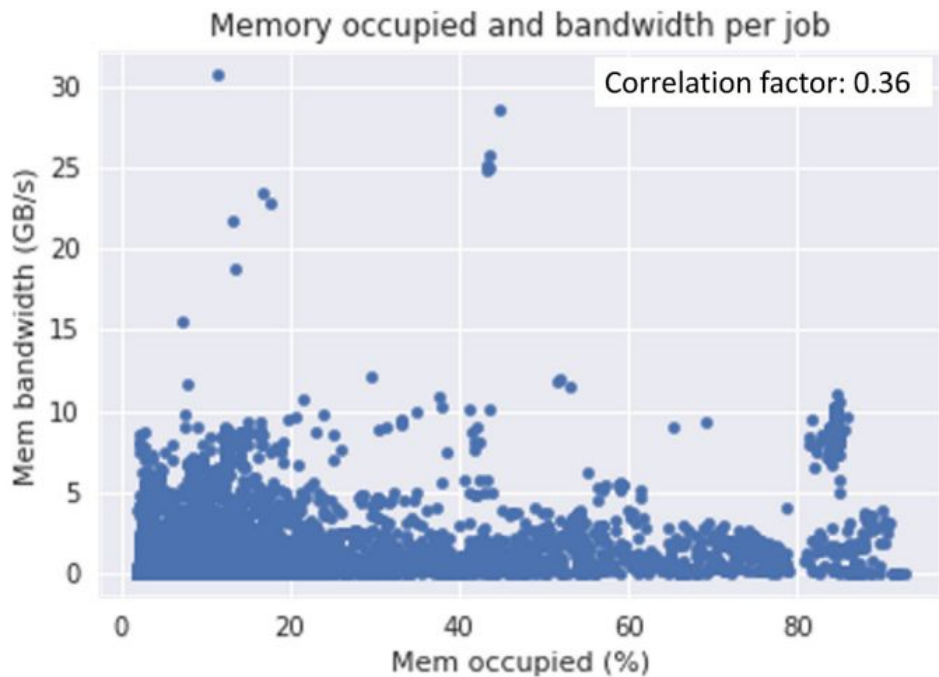
Time-series graph. Shaded areas show half a stdev in a 30s time period



Shaded area represents half a standard deviation of the value's range within 30 s of each sample

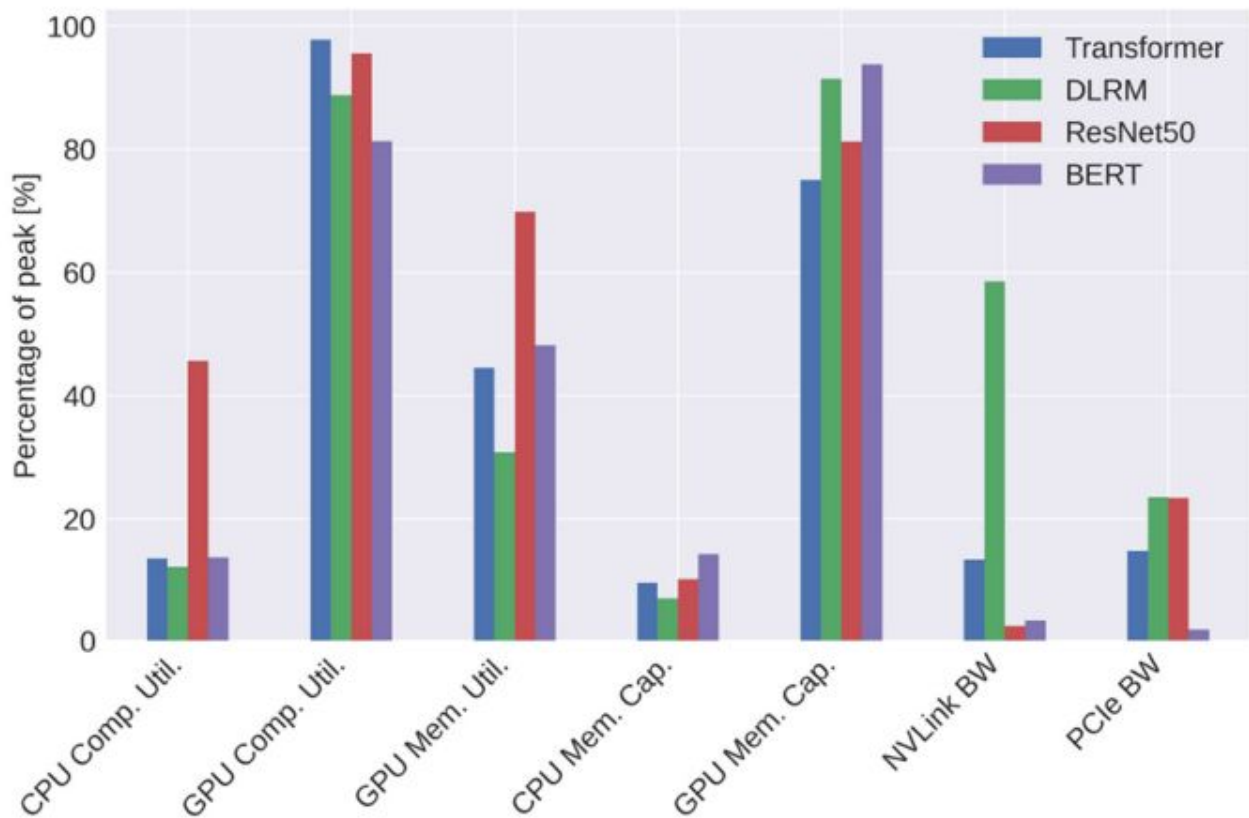
Does usage of resources correlate?

Only Two Resources Types We Noticed Weak Correlation



What about GPUs?

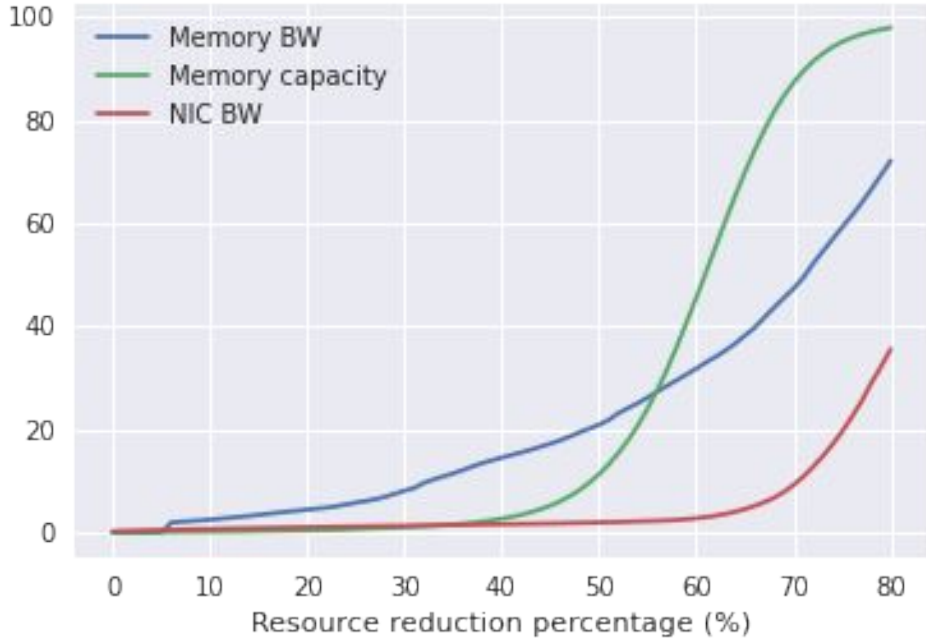
Single-Node Training. NVIDIA DGX1



What does this mean for disaggregation?

What It Means for Disaggregation

Probability for a CPU to cross rack boundary per resource



	Mem occupied	Mem BW	NIC BW
Haswell	39.95%	69.01%	59.17%
KNL	5.36%	-	43.35%

How much we can reduce each resource and still satisfy the worst-case rack average utilization

Questions?



Contact: mihelog@lbl.gov