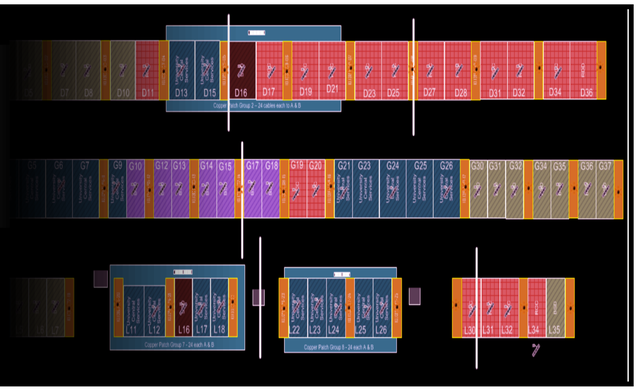
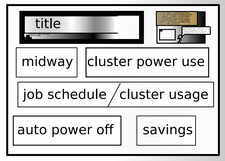
CONSERVING ENERGY IN HETEROGENEOUS CLUSTERS

W McFadden[1,2] SD Jacobs[1] R Parpart[3] A Nikolich[1] HB Runesha[1]
[1]University of Chicago, Research Computing Center
[2]University of Chicago, Biophysical Science Program
[3]University of Chicago, Data Center Operations



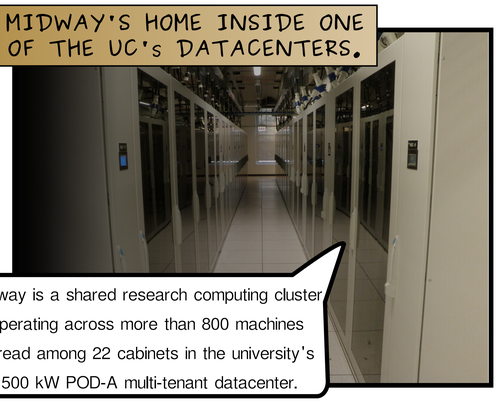
WE'VE COLLECTED AND ANALYZED THE POWER AND CPU USAGE DATA FROM THE UNIVERSITY OF CHICAGO'S MIDWAY CLUSTER TO SEE IF WE CAN REDUCE OUR ENERGY FOOTPRINT. WE'VE FOUND THAT THE CLUSTER IS OFTEN UNDERUTILIZED WITH MANY PROCESSORS IDLING. WE SIMULATE THE ENERGY SAVINGS OF SHUTTING OFF IDLE MACHINES.



Simply turning off idle machines can cut our energy use by 2%.

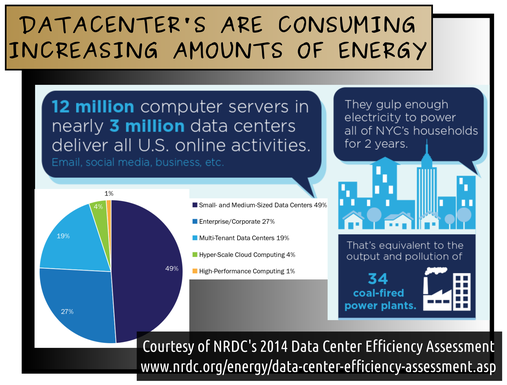
Developing better usage predictions may lower our energy use even more.

MIDWAY'S HOME INSIDE ONE OF THE UC's DATACENTERS.



Midway is a shared research computing cluster operating across more than 800 machines spread among 22 cabinets in the university's 500 kW POD-A multi-tenant datacenter.

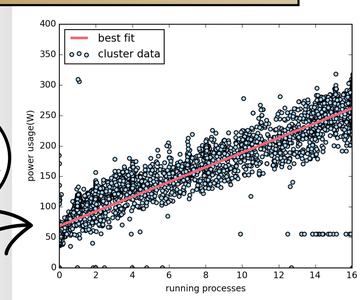
DATACENTER'S ARE CONSUMING INCREASING AMOUNTS OF ENERGY



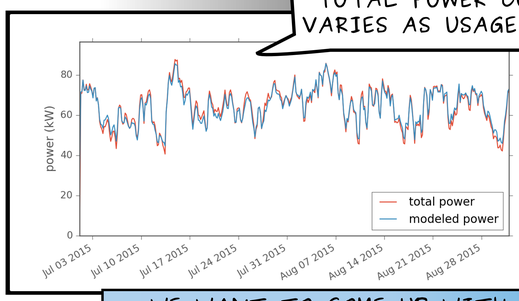
Courtesy of NRDC's 2014 Data Center Efficiency Assessment
www.nrdc.org/energy/data-center-efficiency-assessment.asp

POWER CONSUMED DEPENDS ON NUMBER OF PROCESSES RUNNING ON A NODE

BUT, even when a node is not working, it still uses a bit of energy to idle

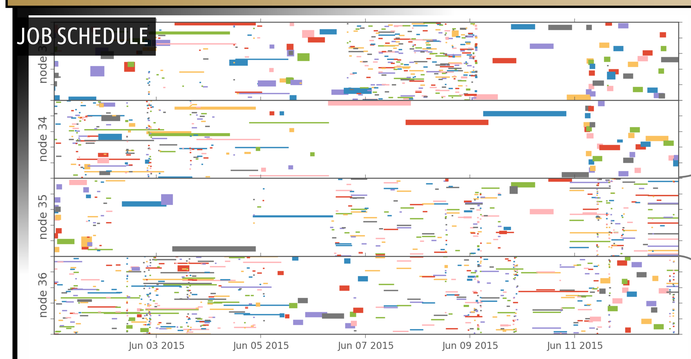


TOTAL POWER CONSUMPTION VARIES AS USAGE FLUCTUATES

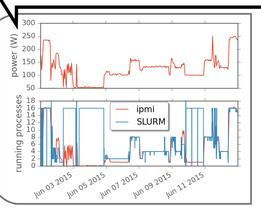


WE WANT TO COME UP WITH WAYS TO MANAGE JOB SCHEDULING SO THAT WE ONLY BURN ENERGY WHEN WE REALLY NEED IT?

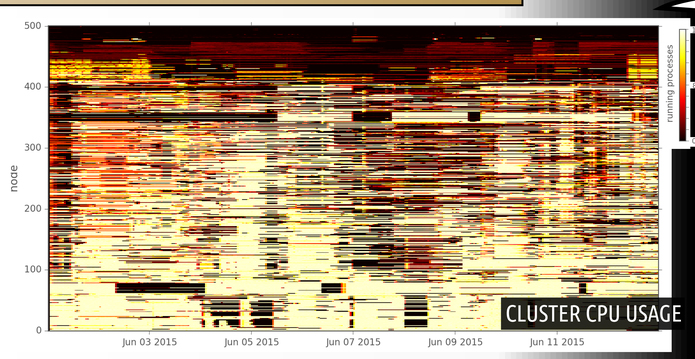
WE KEEP TRACK OF JOB SUBMISSION STATISTICS TO ANALYZE USAGE PATTERNS ACROSS THE CLUSTER



We do see some discrepancies between the node and scheduler data about the number of running processes. Power usage agrees with node stats.

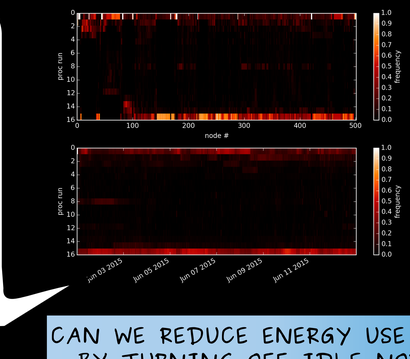


CLUSTER USAGE STATS REVEAL MANY NODES AREN'T BEING FULLY UTILIZED



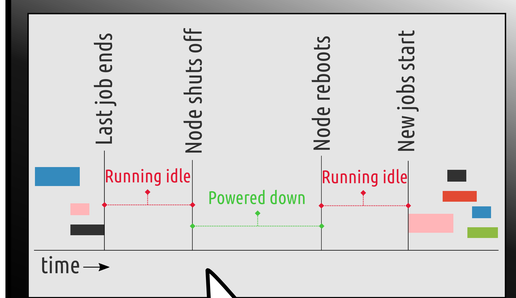
On the left, we see the number of running processes on every node over the course of 3 weeks. We can see that the nodes at the top are being used less often.

On the right two panels, we slice this raw data to reveal variation in usage over time or between nodes.

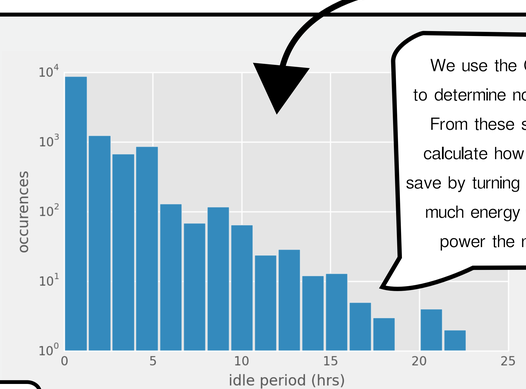


CAN WE REDUCE ENERGY USE JUST BY TURNING OFF IDLE NODES?

SLURM SCHEDULER CAN AUTOMATICALLY SHUT DOWN NODES WHEN NOT IN USE

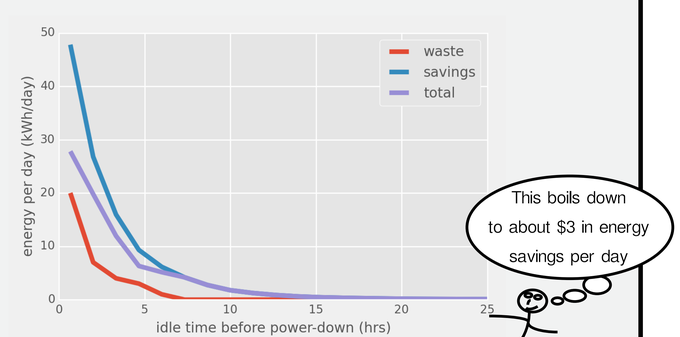


SLURM's power saver feature runs a shutdown script after a node has been idle for a preset amount of time and restarts the node when a new job is scheduled.



We use the CPU usage data to determine node idle time stats. From these statistics we can calculate how much energy we save by turning nodes off, and how much energy we waste as we power the nodes back up.

POWERING OFF IDLE NODES COULD REDUCE MIDWAY'S ENERGY CONSUMPTION BY 2%



This boils down to about $3 in energy savings per day

WE'D LIKE TO THANK…

We'd like to thank the University of Chicago and the RCC for supporting this research. As well as Anita Nikolich and Andy Wettstein for their advice and assistance. We'd also like to thank Pierre Delforge for permission to reuse his datacenter infographics.

AND WHAT'S NEXT?

We're currently building better usage models to fine tune our job and idle scheduling optimizations.

CAN BETTER PREDICTION OFFER EVEN MORE SAVINGS? TUNE IN NEXT TIME…