SUSTAINABILITY AND ENVIRONMENT SUBGROUP
APRIL 2020 WEBINAR

CATHERINE HURLEY
Sustainability Program Manager
Argonne National Laboratory
churley@anl.gov
630-252-0237

SUZY BELMONT
Information Systems Coordinator |Site Operations
Site Operations
National Renewable Energy Laboratory
Suzanne.Belmont@nrel.gov
303-275-3885

April 16, 2020

SESG April Webinar Agenda		
11:00– 11:05	Welcome	Catherine Hurley, ANL Suzy Belmont, NREL
11:05 – 11:15	Sponsor Message	Steve Bruno, SPD Cate Berard, AU-21
11:15 – 11:50	HPC Energy Efficiency in the Exascale Era	Jeff Broughton NERSC Deputy for Operations at Lawrence Berkeley National Laboratory
11:50-12:25	Energy Efficient Design of Oak Ridge's Summit Supercomputer	Jim Rogers, Computing and Facilities Director of the National Center for Computational Science at Oak Ridge National Lab
12:25 – 12:30	Survey Launch - HPC and HPSB	Steve Bruno, SPD Suzy Belmont, NREL
12:30	Webinar wrap-up	Catherine Hurley, ANL

SPD UPDATE

Steve Bruno

Sustainability Performance Division, MA-53

U.S. DEPARTMENT OF ENERGY
Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Argonne NATIONAL LABORATORY

AU-21 UPDATE

Una Song

Office of Sustainable Environmental Stewardship (AU-21)

50 YEARS
U.S. DEPARTMENT OF ENERGY
EARTH DAY
ACTION FOR EARTH!
GET OUTSIDE
• Check out the Smithsonian Gardens virtual tours with Smithsonian horticulturists.
• Visit the cherry blossoms without leaving your home!
GET INVOLVED
• Sustainability Treasure Hunt
Go on a Sustainability Treasure Hunt at home! Learn more here.
• Earth Day Coloring Book
Be creative with kids to help them learn about Earth Day!
• Recycling Educational Banner
Learn how to take action for earth by recycling at home.
LEARN MORE
Earth Day 50 Video and Contest Winners Announcement Tune in on April 22nd at 10am EDT:
https://www.energy.gov/live-stream
• Migratory Birds at Pantex
Learn about Pantex's award winning work to help migratory birds!
EARTH DAY PROMOTIONAL ITEM
Hydrate with your DOE Earth Day 50 water bottle! Apparel is also available for purchase! Visit the DOE Earth Day Store
Visit the Earth Day Powerpedia page for more information.

Earth Day Video Screening

- AU-21 is compiling a video highlighting DOE's commitment through the years to preserve and conserve the Earth's natural resources
  - *Thank you to all the sites that submitted footage!*
- Screening to be held Wednesday, April 22, 10:00 am (Eastern)
  - via Webex.
  - Access Code: 900 117 678
  - Call-in Number: +1-415-527-5035
- Earth Day photo and haiku contest winners will also be announced on the Webex.

For More Information

- Send us your plans for Earth Day
- Please contact the Office of Sustainable Environmental Stewardship, AU-21

Una Song

(202) 586-4335

una.song@hq.doe.gov

https://powerpedia.energy.gov/wiki/Earth_Day

https://www.energy.gov/management/spo/sustainability-performance-division-media

HIGH PERFORMANCE COMPUTING WEBINAR

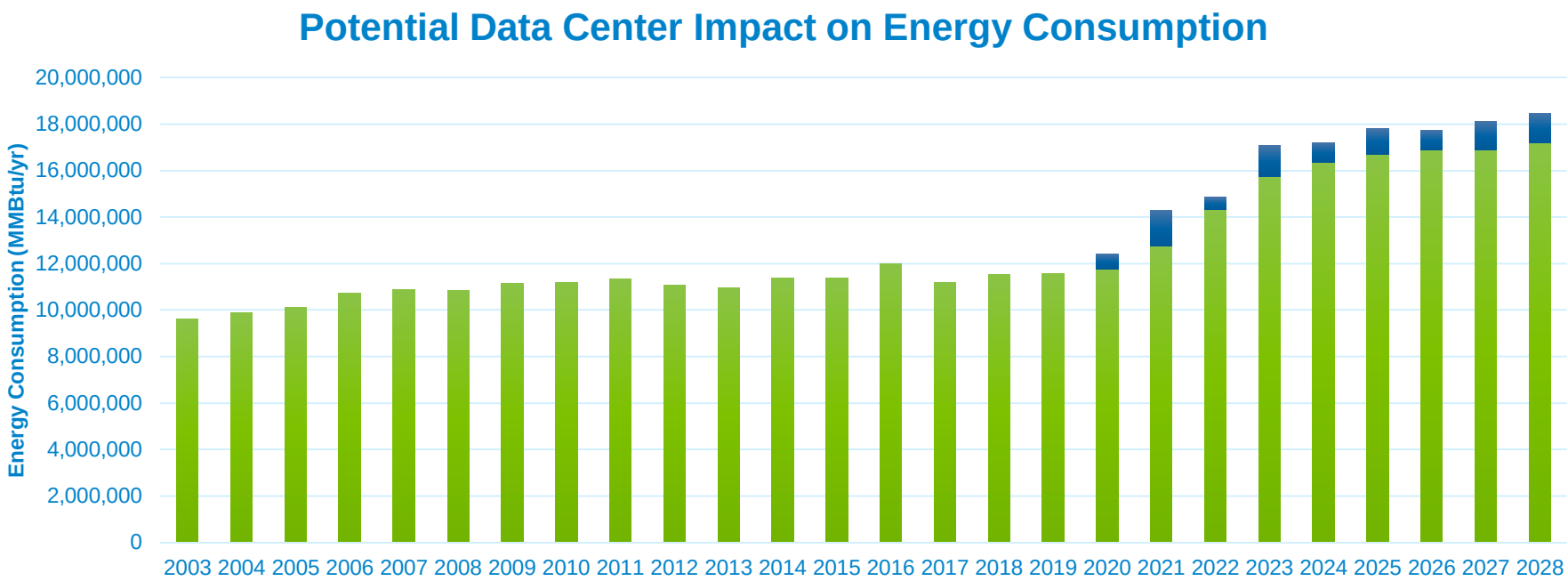
We would like to welcome our speakers!

- **Jeff Broughton: HPC Energy Efficiency in the Exascale Era**
  - NERSC Deputy for Operations, Lawrence Berkeley National Laboratory

- **Jim Rogers: Optimizing Performance for Summit – from Design to Operation**
  - Computing and Facilities Director for the National Center for Computational Science, Oak Ridge National Laboratory

PROJECTED PERFORMANCE - ENERGY

DOE sites with new HPCs, 2003-2028



Please note the energy consumption data in this chart is for sites with planned/new large data centers/HPCs only: ANL, LANL, LBL, LLNL, NREL, NETL, ORNL and SNL. Information is based on FY 2019 reported data and does not reflect current status. Increase due to HPC is in dark blue.

HPC ENERGY EFFICIENCY IN THE EXASCALE ERA

JEFF BROUGHTON
NERSC DEPUTY FOR OPERATIONS AT
LAWRENCE BERKELEY NATIONAL LABORATORY

ENERGY EFFICIENT DESIGN OF OAK RIDGE'S SUMMIT SUPERCOMPUTER

JIM ROGERS, COMPUTING AND FACILITIES DIRECTOR
OF THE NATIONAL CENTER FOR COMPUTATIONAL SCIENCE AT
OAK RIDGE NATIONAL LAB

FEEDBACK REQUESTS

- **Sustainable Buildings Feedback Request**
  - Let us know what areas you are struggling with for both new construction and existing buildings, and how we can help via the Sustainable Buildings Feedback Request.

- **HPC/Data Center Feedback Request**
  - HPC Working Group is recruiting team members interested in various aspects of energy and water efficiency in HPCs and data centers via the HPC/Data Center Feedback Request.

Please forward to your colleagues and respond by **May 5th (extended deadline)**

Questions? Please email sustainability@hq.doe.gov

SAVE THE DATE – JUNE WEBINAR

Please join us for our next webinar:

June 18, 2020

11am -12:30pm EST

Topic: Smart Labs

HPC Energy Efficiency in the Exascale Era

Jeff Broughton
NERSC Deputy for Operations
Lawrence Berkeley National Laboratory

EFCOHG/SESG Monthly Meetinng
April 16, 2020

Exascale by the Numbers

1,000,000,000,000,000,000 flops / sec

500 x U.S. national debt in pennies

100 x number of atoms in a human cell

2 x number of seconds since the Big Bang

1 x number of insects living on earth

Exascale by the Numbers

> 1,000,000,000,000,000,000 flops / sec

> $600,000,000 life cycle cost *

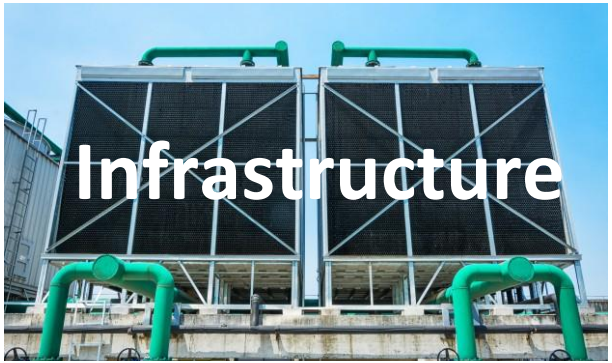
~ 200,000,000 kWhr / year

~ 150,000 $MTCO_2e$ / year

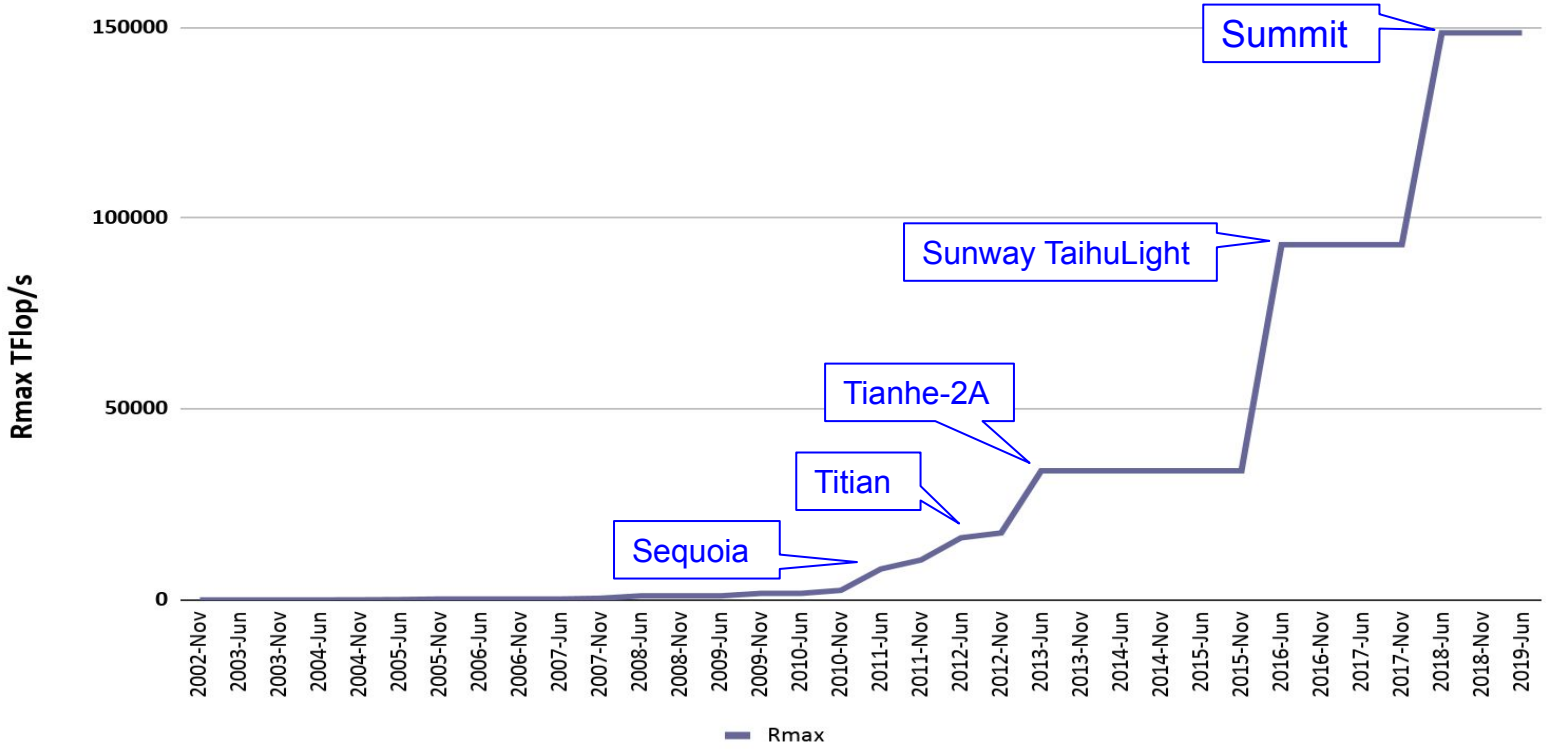
< 100 apps ready

(*) Not including NRE, construction or staff time

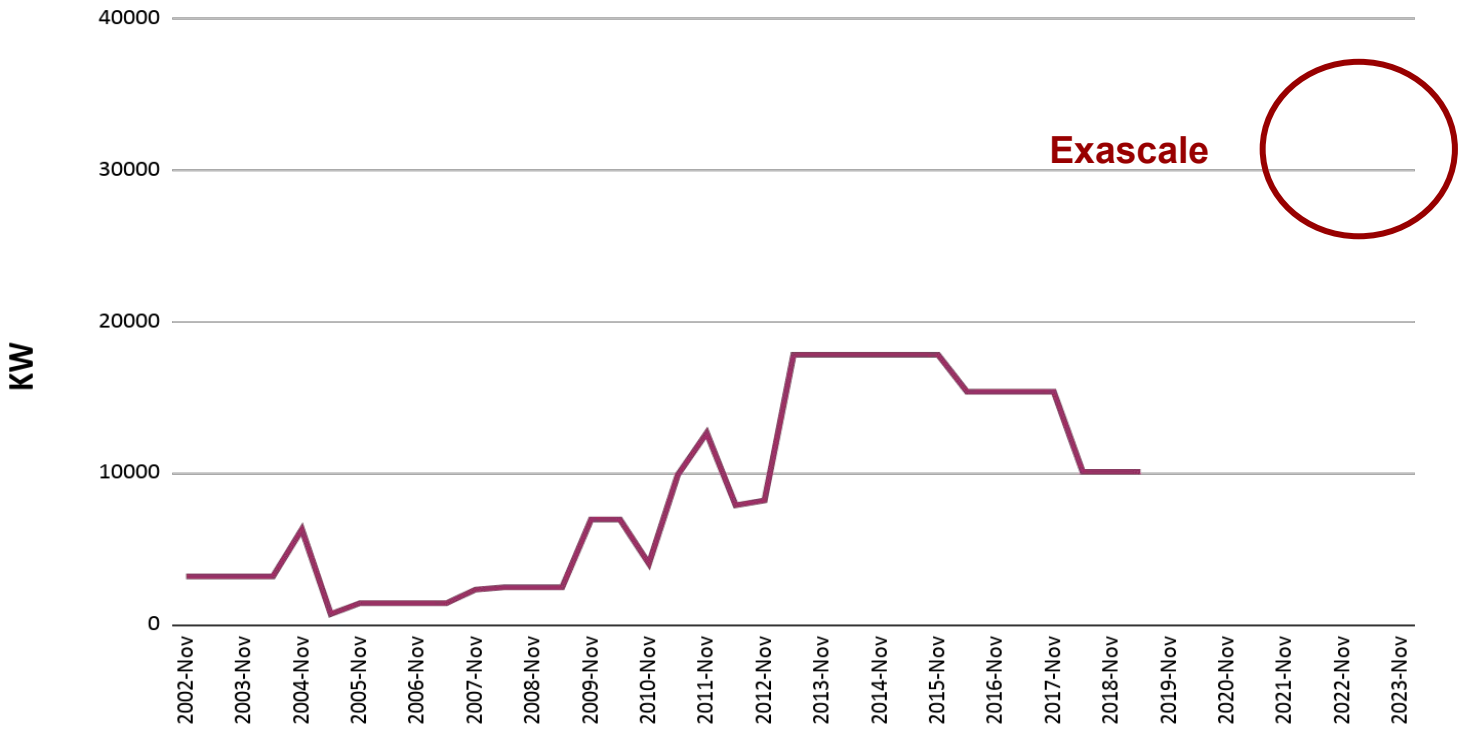
Four facets of energy efficiency



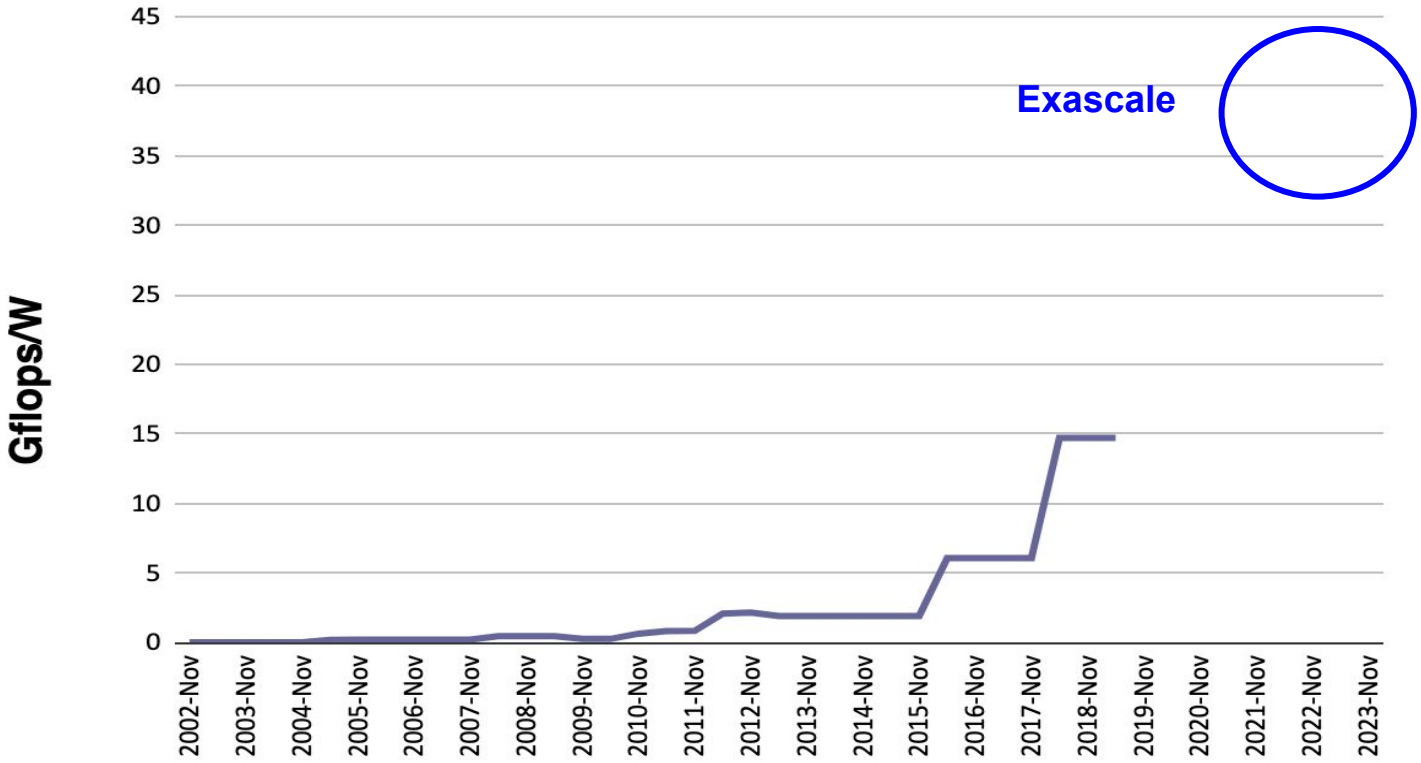
Performance of Top 1 System (Rmax)



Power consumption of Top 1 system



Power efficiency of Top 1 system



Technology evolution has delivered substantial energy efficiency improvements over time

Successes

- Moore's law
- Volume Server Technology
- GPUs!
- Incremental Improvements
- Bigger Budgets!!

Noble Attempts

- BlueGene
- Knights Family
- Hero systems

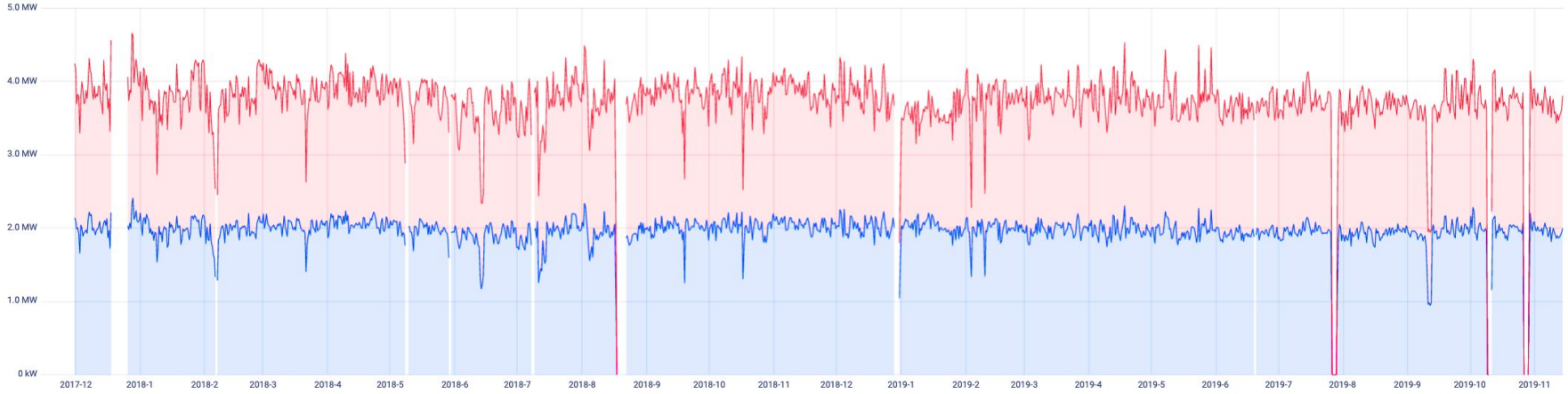
Near Future Wins?

- Flash and storage class memory
- Specialized accelerators

Four facets of energy efficiency

How much power does your system use?

2 year history of Cori power consumption



Cori rated 5.7 MW, observed peak 4.6 MW, typical 3.9 MW, idle 2 MW

Underused capacity is an economic and efficiency problem

Big Trends in Packaging and Energy Efficiency

Warm Water Cooling

~50F → 90-120F
ASHRAE Standards

More Efficient Component Cooling

Air Cooling → Hybrid/Indirect liquid cooling → Direct liquid cooling

Elimination of Chillers

Compressors → Part-time Economizers → Full-time Economizers

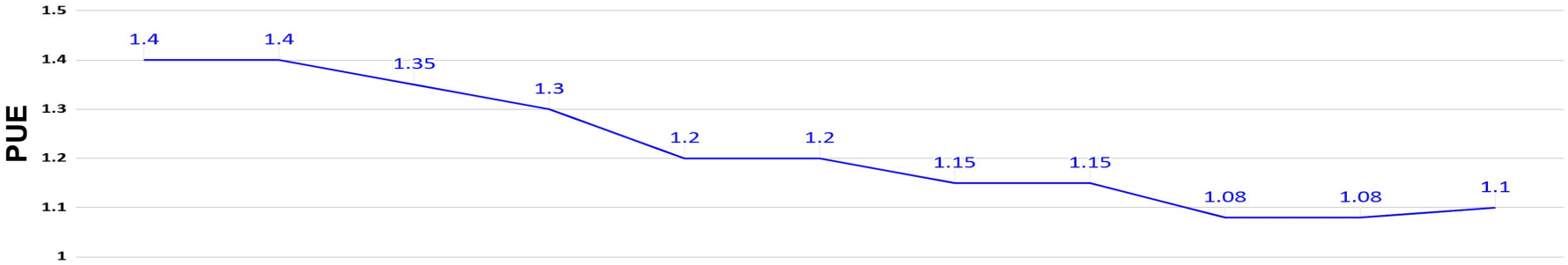
Control Sophistication

Thermostat → Data Analytics
System/Plant Coordination
Machine Learning

Denser Packaging

<1KW/sf
<200 lb/sf → 10KW/sf
350 lb/sf

Incremental improvements to energy efficiency over the years



Franklin

Air-cooled

Cray XT4
0.27 PF
2007

Hopper

Hybrid w/ chillers

Cray XE6
1.05 PF
2010

Edison

Hybrid w/ full-time economizers

Cray XC30
1.26 PF
2013

Cori

Hybrid w/ full-time economizers

Cray XC40
14.0 PF
2015

Energy Efficiency Optimization Project

2017 - ongoing

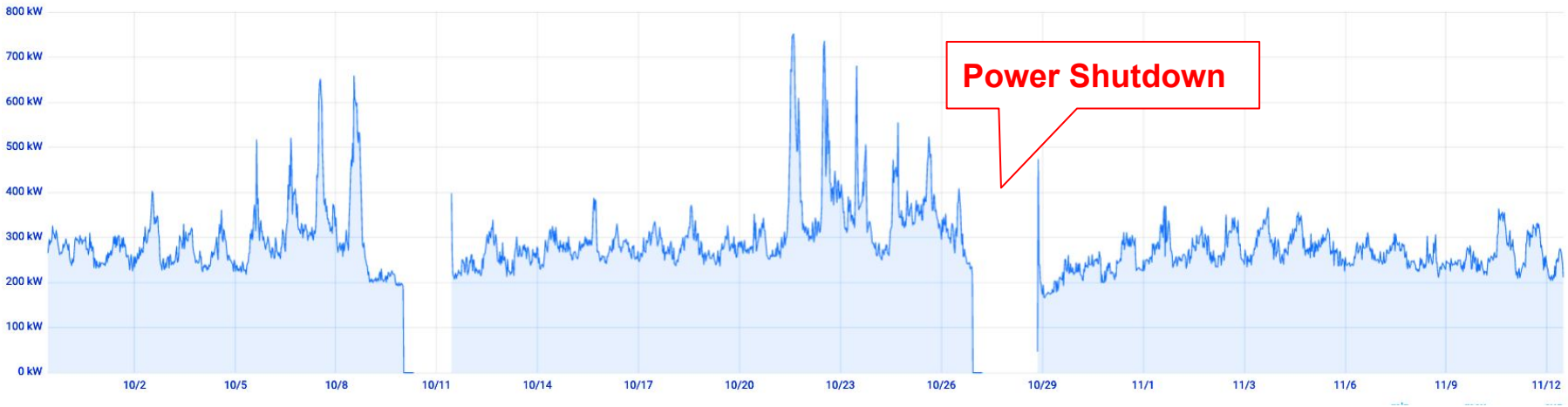
Edison Retirement

May 2019

Oakland → Berkeley

New cooling methods require sophisticated controls

45-day history of Cori physical plant power consumption



Responds to both system load and to weather conditions

Varies CW temp in response to wetbulb temp to optimize fan and pump energy

Varies CW pump speed in response to system load

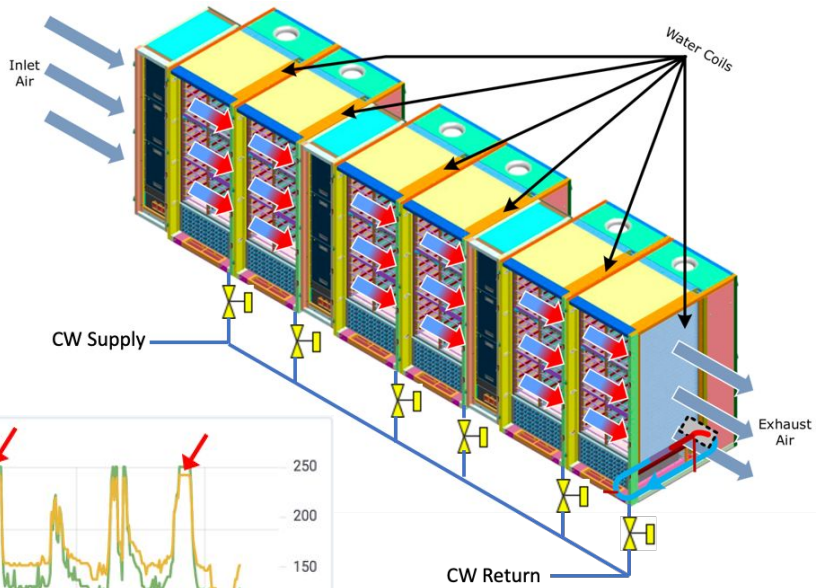
System and infrastructure must communicate

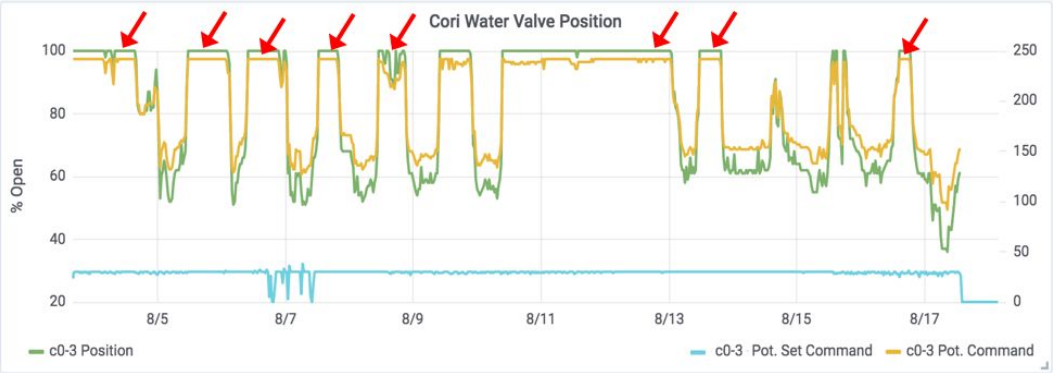
Out of the Box Operation: ***Static cabinet air temp setpoint***

- Can't set for all weather, therefore CW pumping and fan energy waste

Interactive Operation: ***Dynamic cabinet air temp setpoints***

- Cooler CW temps - Cooler cabinet air therefore fan speed turndown
- Warmer CW temps - Fan speeds compensate. Dynamic setpoint reduces excessive CW demand valve positions





Gather the data and the Q's will come

OMNI Data Collection System

Current Data Sources

- Substations, panels, PDUs, UPS
- Cray XC internal SEDC data
- Onewire Temp & RH sensors
- BMS through BACNET
- Indoor/Outdoor Particle counters
- Weather station

System Data Sources (WIP)

- Syslog
- Job Data
- Lustre + GFPS statistics
- LDMS

Rabbit MQ, Elastic, Linux

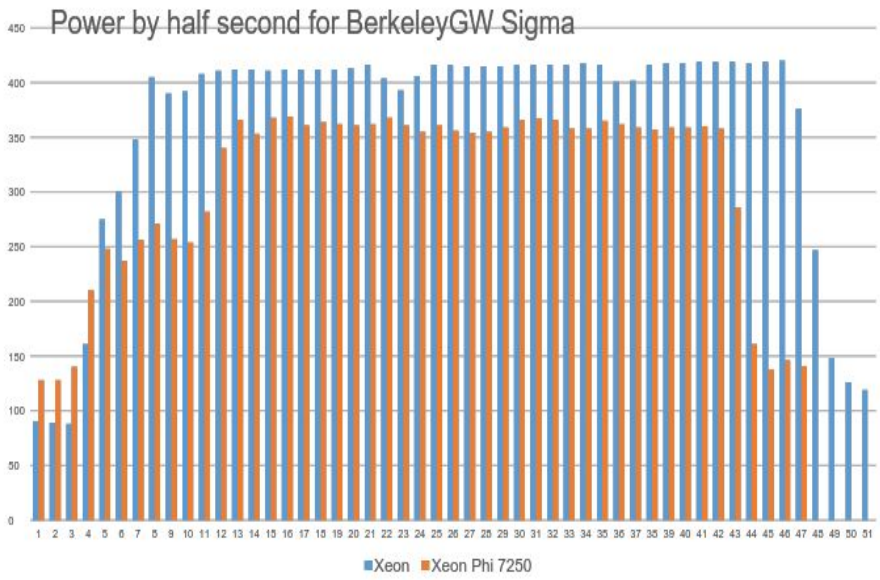
- Collects ~20K data items per second
- Over 100TB data online
- 45 days of SEDC (versus 3 hours on SMW)
- 180 days of BMS data (6X more than BMS)
- > 2 years of power data

Kibana, Grafana, Skyspark

Four facets of energy efficiency

Reducing time to solution is the first order approach to use energy efficiently

- Application Energy Efficiency ~ 1 / Wall-Time
- Potential application power savings << typical power consumption
- Goal of capability/capacity systems is more science for ~ same power



Software optimizations can improve performance and reduce energy

Energy-efficient processors have multiple HW features to optimize for both performance and energy efficiency

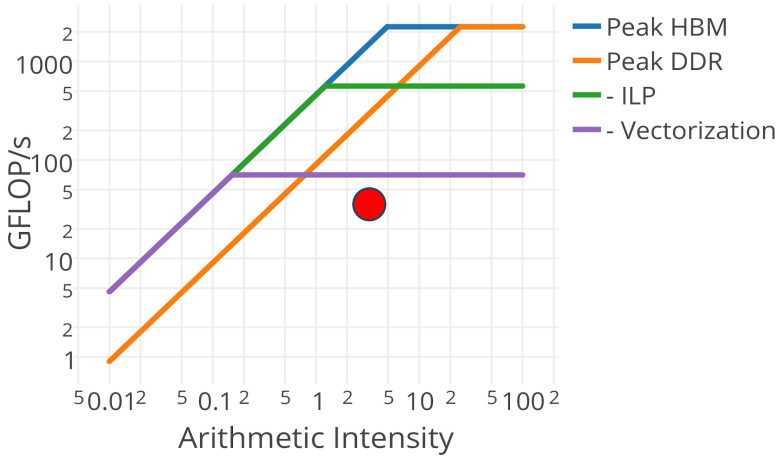
- Many (Heterogeneous) Cores
- Bigger Vectors
- New ISA
- Multiple Memory Tiers

Opportunities for energy reduction:

- Reduce CPU power consumption in Memory or IO intensive code-regions (and vice versa)
- Avoid highly masked or inefficient vector instructions
- Change algorithm to minimize data motion

Large gains possible by connecting users to actionable performance data

Roofline Model:
Visualization of Counter Data



Major efforts underway to optimize applications for exascale-class systems

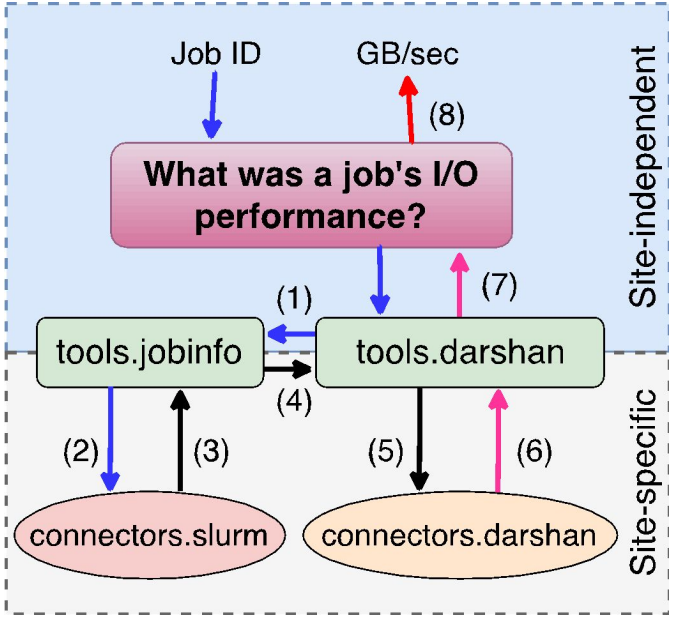
- Multi-year, cooperative efforts
  - DOE National Labs
  - Vendors
  - Application teams
- Approach
  - Domain and computer scientists
  - Focused optimization sessions
  - Dedicated app team support
  - Education, training and outreach
- Parallel effort on performance portability

GPU For Science Days

GPU Community Hack-a-thon

System software opportunities to improve operational efficiency

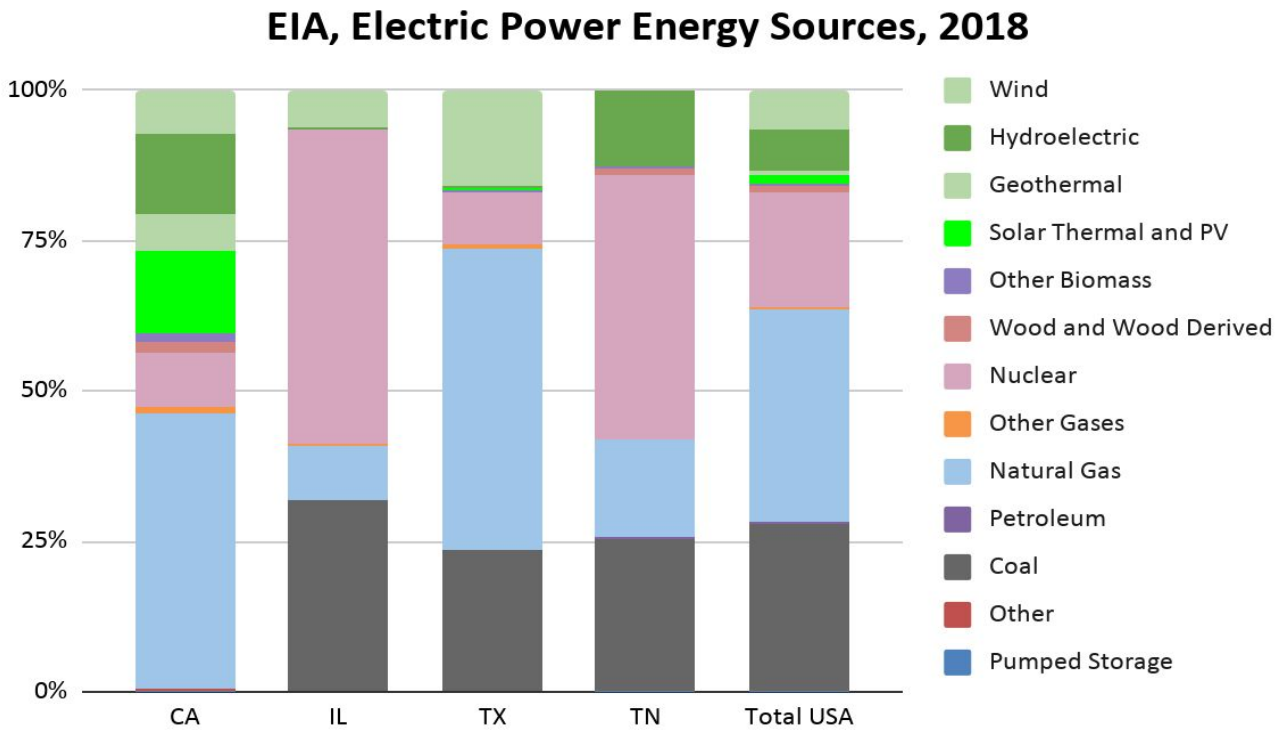
- Power saving
  - Idle time reduction
  - Detection of failing / aggressor jobs
- Power management
  - Power capping / band limiting
  - Automatic demand-response
- Data Collection
  - Job, application and I/O statistics



G. K. Lockwood et al, "TOKIO on ClusterStor: Connecting Standard Tools to Enable Holistic I/O Performance Analysis," in Proceedings of the 2018 Cray User Group. 2018.

Four facets of energy efficiency

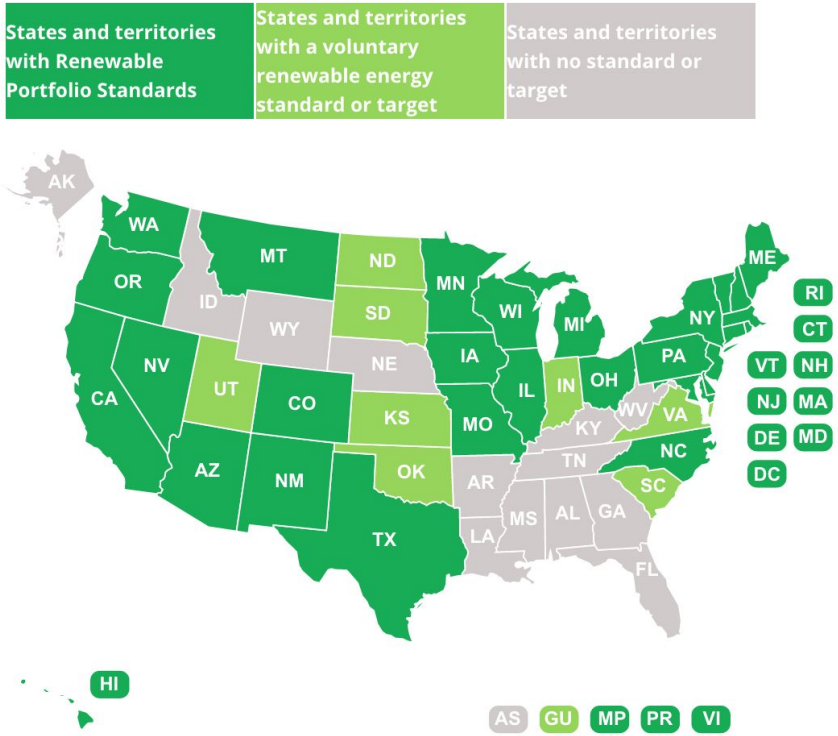
HPC needs to use greener energy



Greenhouse gas mitigation requirements are increasing

Renewable Portfolio Standards (RPS)

- California + 28 states + DC have mandatory programs
- 8 more have voluntary programs.
- ~ one half of renewable energy growth is due to these standards



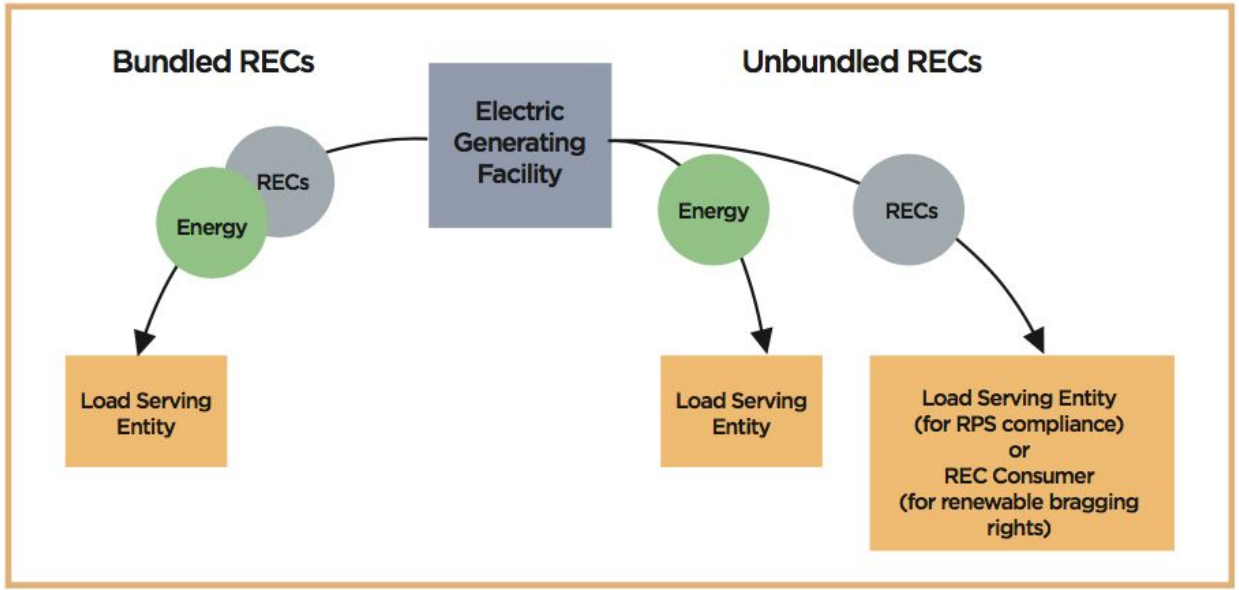
Source: National Conference of State Legislatures
http://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx

Greenhouse gas mitigation requirements are increasing

California Environmental Quality Act (CEQA) creates a legal obligation to meet GHG mitigation goals

- Must buy Renewable Energy Credits (RECs) to mitigate nearly all GHG for NERSC-9
- University of California developing renewable energy projects to achieve net zero GHG by 2025

RECs come in many forms and prices



Pinkel and Weinrub, 2014, What the Heck is REC

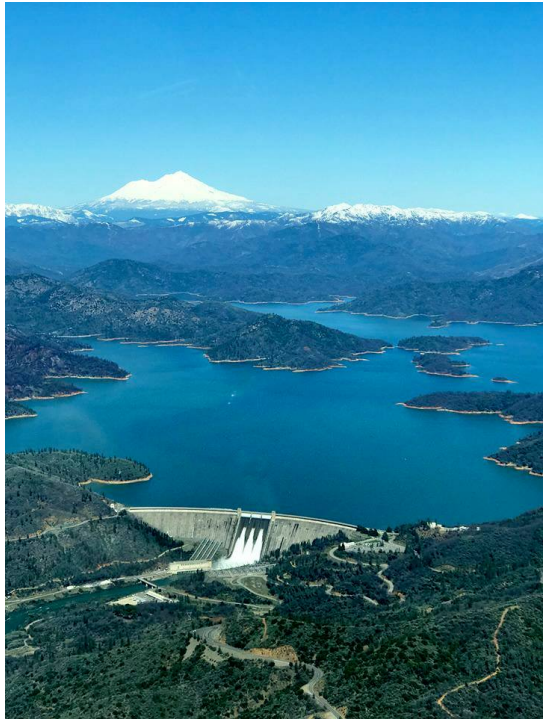
- Electricity and RECs are the same product, on the same grid
- Recent historical typical price: 1.5¢ to 4¢ per kWh

- Environmental attribute is separate from the electricity
- Recent historical typical price: 0.1¢ to 2¢ per kWh

Renewables are becoming less expensive than conventional power.

Can we save water too?

- Cooling towers consume up to 3.5M gallons per year per MW
- ~14 $MTCO_2e$ per MW from energy to deliver the water
- Alternatives to cooling towers
  - Dry coolers
  - Thermosyphons
  - Geothermal cooling
  - Deep water cooling



Final thoughts

Energy efficient improvements for HPC require:

- A holistic approach
- Some breakthroughs... but mostly incremental improvements
- Data analytics to drive and measure progress

Thank you!

Optimizing Performance for Summit – from Design to Operation

Jim Rogers

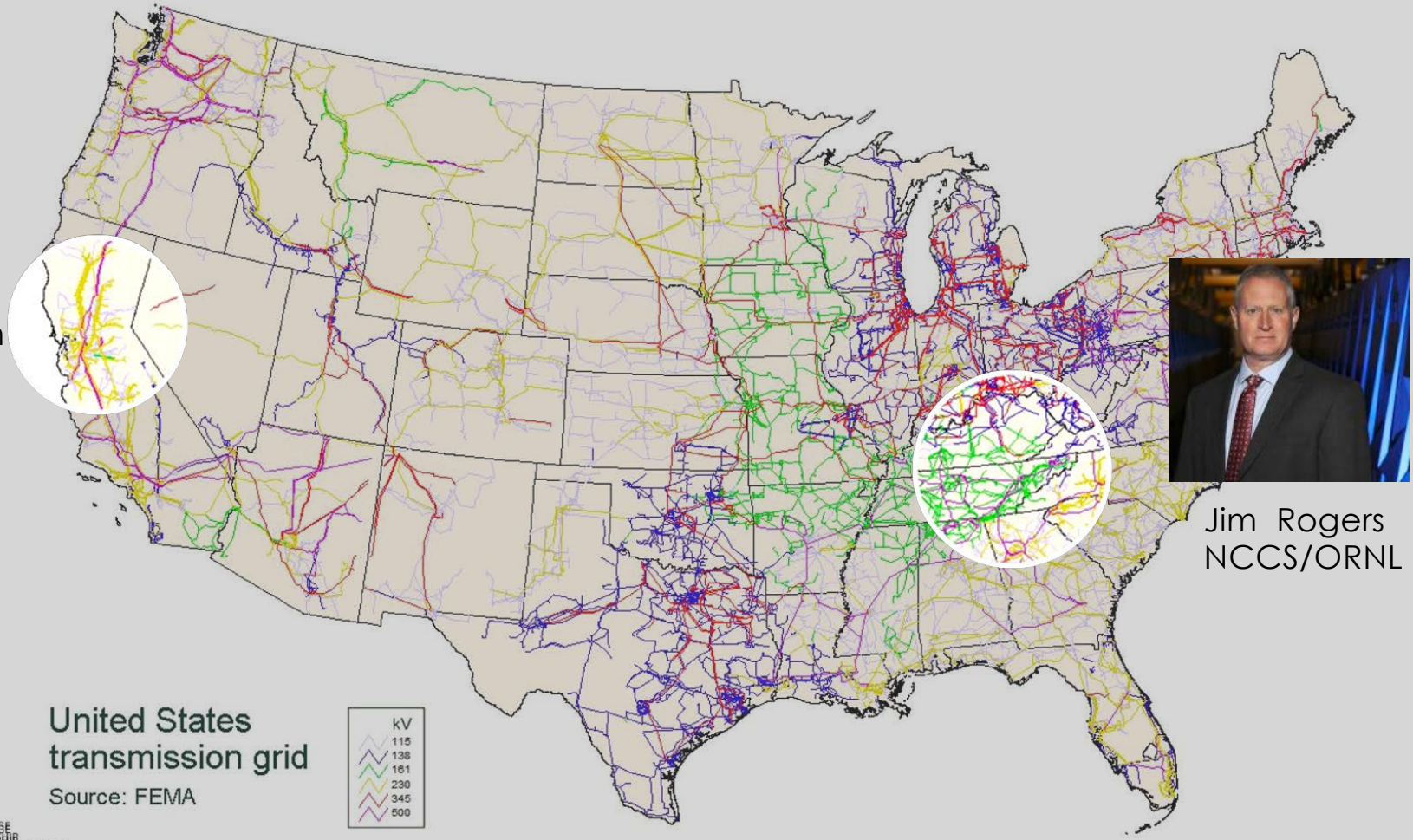
Director, Computing and Facilities
National Center for Computational Sciences
Oak Ridge National Laboratory

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

EFCOG/SESG Monthly Meeting
'April 16, 2020

Who are these guys?

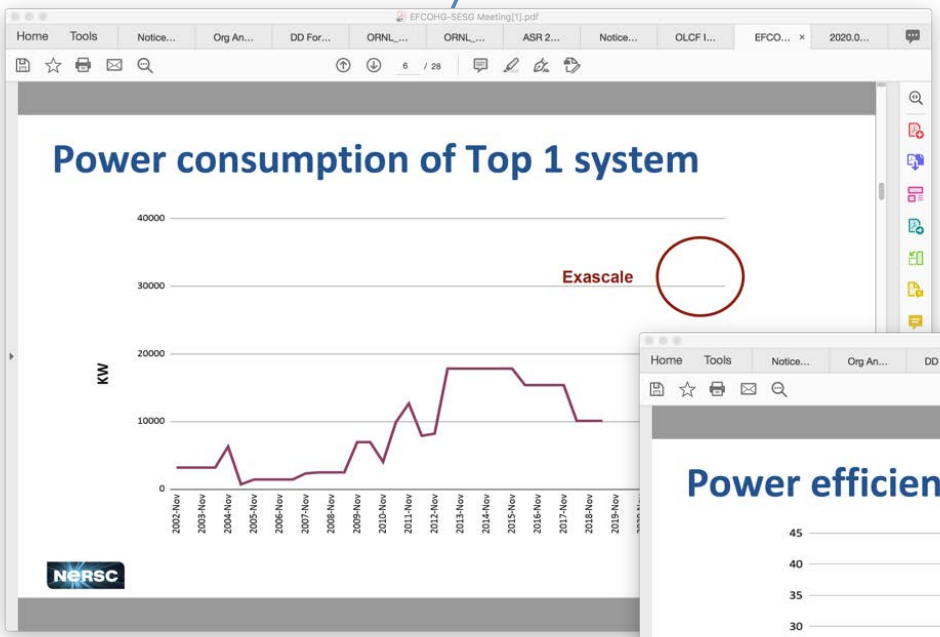
Jeff Broughton
NERSC/LBNL

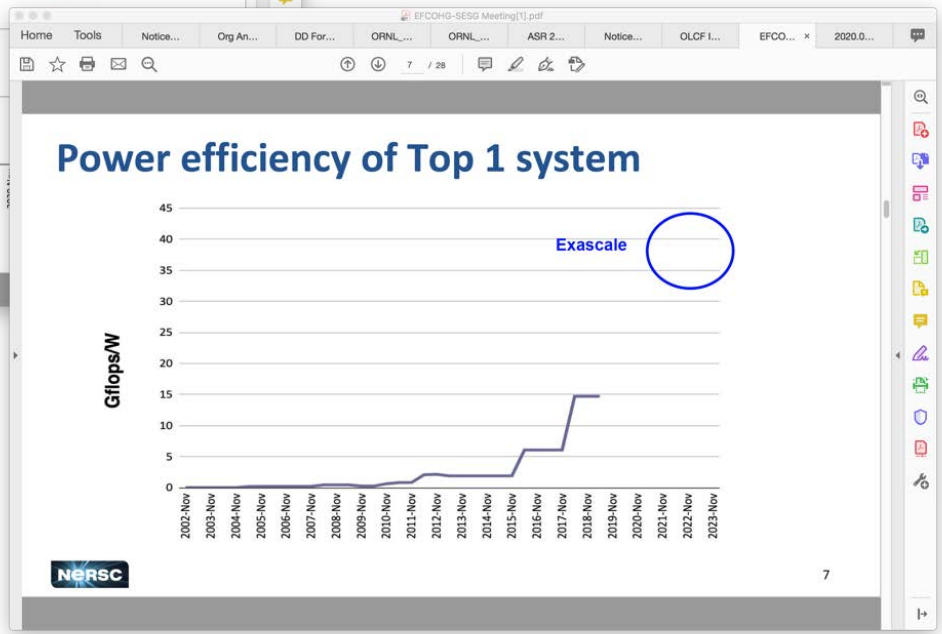


Jim Rogers
NCCS/ORNL

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Oak Ridge Leadership Computing Facility – a DOE Office of Science User Facility

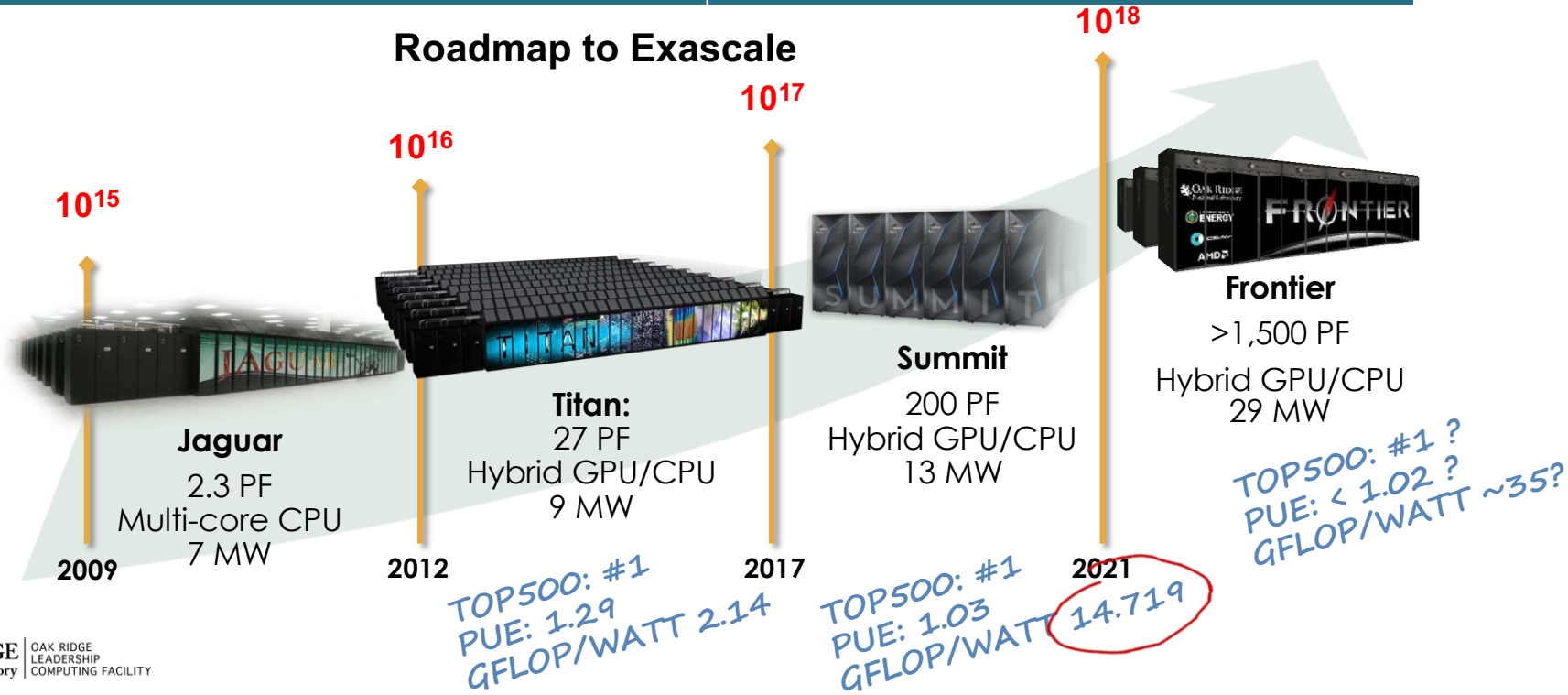




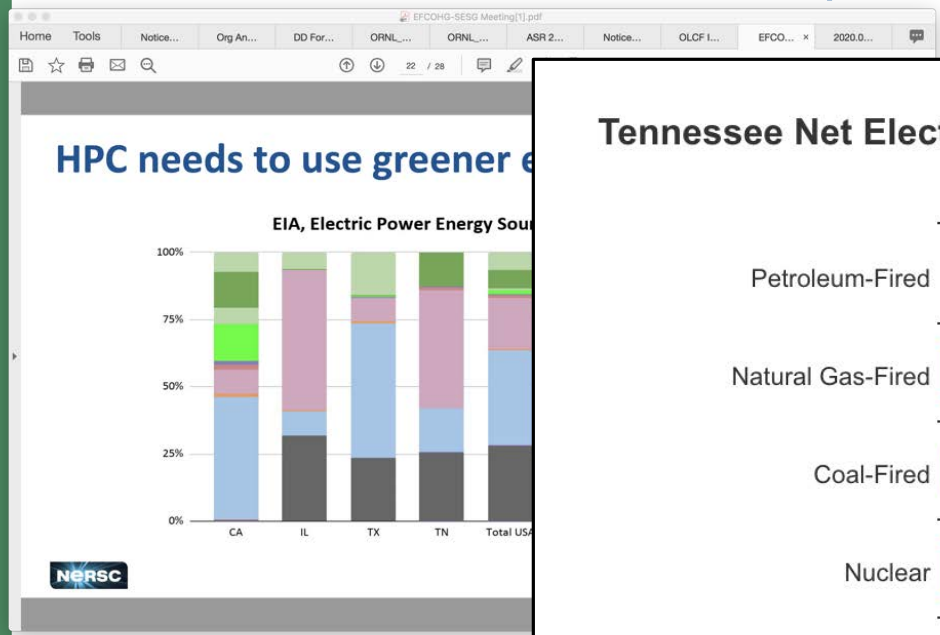
Following up from Jeff Broughton's presentation…

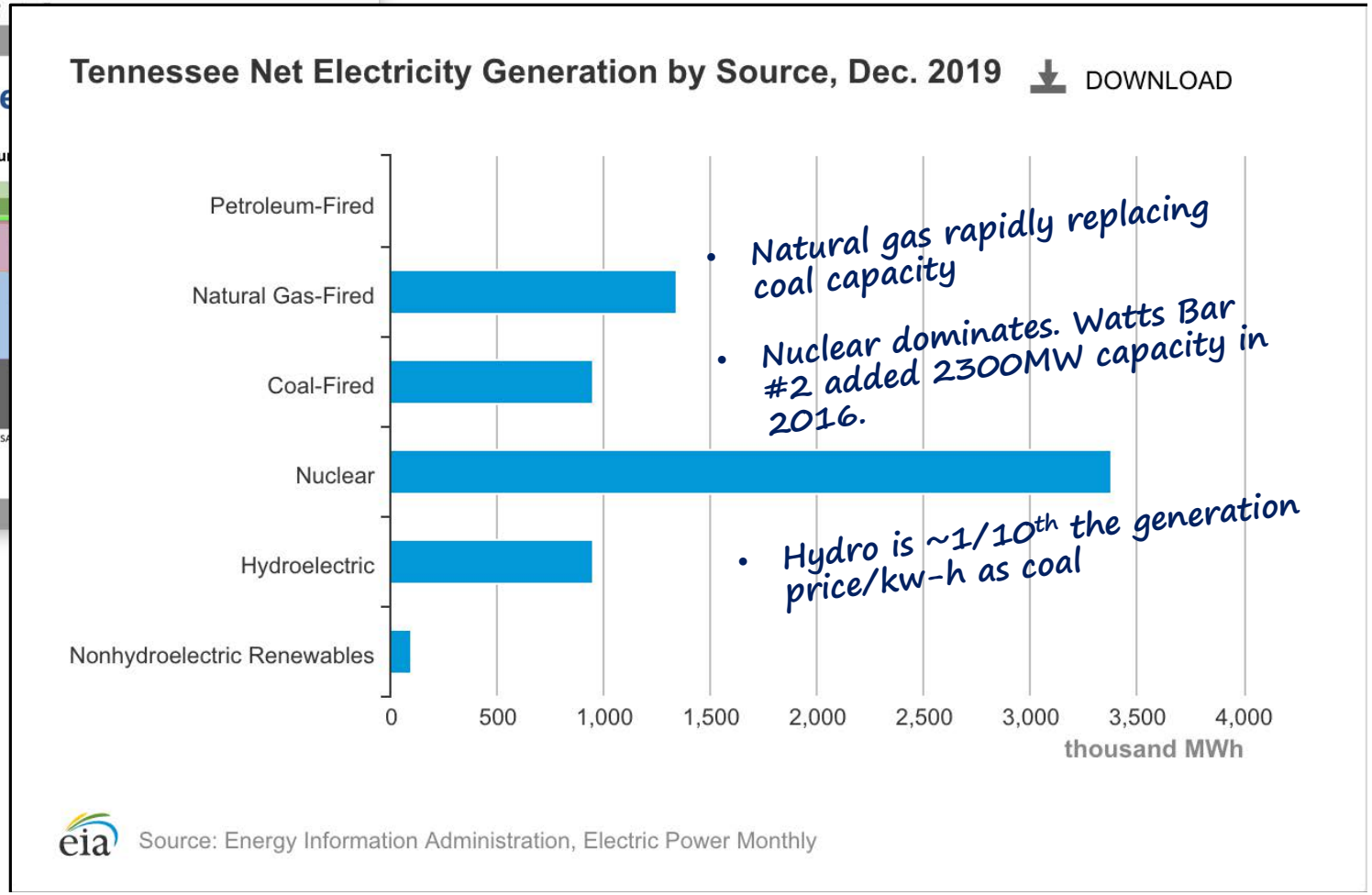
Oak Ridge Leadership Computing Facility – a DOE Office of Science User Facility

Mission: Providing world-class computational resources and specialized services for the most computationally intensive global challenges	Vision: Deliver transforming discoveries in energy technologies, materials, biology, environment, health, etc.



Electrical Generation by Source





Summit by the Numbers

Application Performance	200 PF
Number of Nodes	4,608
Node performance	42 TF
Memory per Node	512 GB DDR4 + 96 GB HBM2
NV memory per Node	1600 GB
Total System Memory	>10 PB DDR4 + HBM2 + Non-volatile
Processors	2 IBM POWER9™ 9,216 CPUs 6 NVIDIA Volta™ 27,648 GPUs
File System	250 PB, 2.5 TB/s, GPFS™
Power Consumption	13 MW

A **200-petaflop** machine, Summit can perform 200 quadrillion (peta-) floating point operations per second (flops). If every person on Earth completed one calculation per second, it would take **305 days** to do what Summit can do in **1 second**.

At over **340 tons**, Summit's cabinets, file system, and overhead infrastructure weigh **more** than a large commercial aircraft.

For some AI applications, researchers can use less precise calculations than flops, potentially **quadrupling** Summit's performance to **exascale levels**, or more than a billion billion calculations per second.

Occupying **5,600 sq. ft.** of floor space, Summit could fill **two tennis courts.**

Summit is connected by **185 miles** of fiber optic cables—or the distance from Knoxville to Nashville.

Summit's file system can store **250 petabytes** of data, or **74 years** of high-definition video.

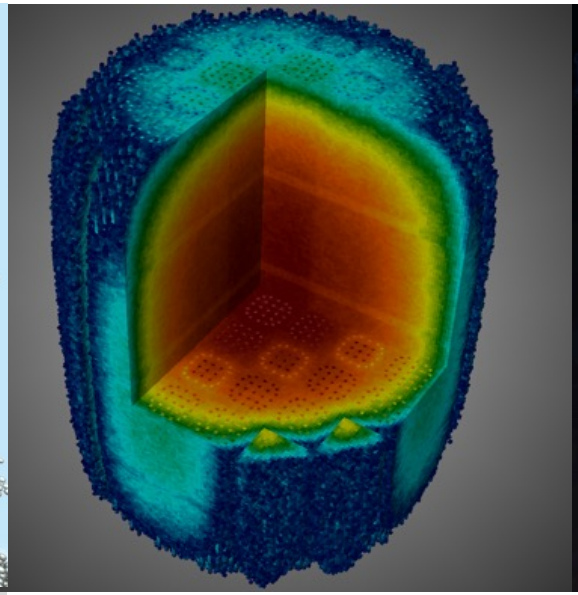
More than **4,000 gallons** of water pump through Summit's cooling system every minute, carrying away about **13 megawatts** of heat.

What Makes These Machines Interesting to Scientists

Selectively Killing Cancer Cells

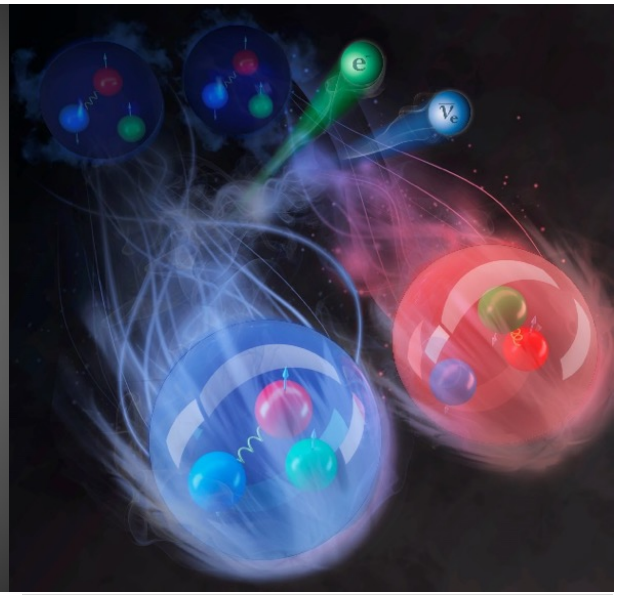
PI(s)/Facility Lead(s): Leonid Zhigilei
ASCR Program/Facility: INCITE/OLCF
ASCR PM: Christine Chalk
Publication(s) for this work: C.-Y. Shih, et al., "Two Mechanisms of Nanoparticle Generation in Picosecond Laser Ablation in Liquids: The Origin of the Bimodal Size Distribution." *Nanoscale* 10, no. 15 (2018):6900–6910, doi:10.1039/C7NR08614H.

Fusion Reactor Design



An illustration displays the total neutron interaction rate throughout an SMR core as simulated by the Shift code on Summit. Image Credit: Steven Hamilton, ORNL

Formation of Matter



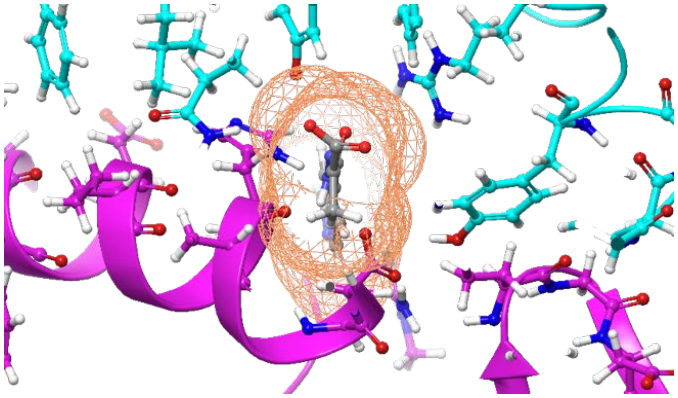
PI(s)/Facility Lead(s): Gaute Hagen
ASCR Program/Facility: OLCF
ASCR PM: Christine Chalk
Publication(s) for this work: P. Gysbers et al. "Discrepancy between experimental and theoretical β-decay rates resolved from first principles." Nature Physics. 11 March 2019.

ORNL Team Enlists World's Fastest Supercomputer to Combat the Coronavirus

Researchers amass list of drug compounds, providing a roadmap for future studies

The Science

Researchers at ORNL and the University of Tennessee, Knoxville, have used the world's most powerful and smartest supercomputer, the IBM AC922 Summit, to identify small-molecule drug compounds that might warrant further study in the fight against the SARS-CoV-2 coronavirus, which is responsible for the COVID-19 disease outbreak. The researchers performed simulations on Summit of more than 8,000 compounds to screen for those that are most likely to bind to the main "spike" protein of the coronavirus, rendering it unable to infect host cells. They found 77 compounds of interest that they believe might have value in experimental studies of the virus. They published their results on *ChemRxiv* in February.



Compound (gray) calculated to bind to the SARS-CoV-2 Spike Protein (cyan) to prevent it from docking to the Human Angiotensin-Converting Enzyme 2 (ACE2) receptor (purple). Image Credit: Micholas Smith, ORNL

The Impact

In the simulations, the compounds bind to regions of the spike that are important for entry into the human cell and, therefore, might interfere with the infection process. Computational screening allows researchers to quickly "shine the light" on promising candidates for experimental studies, which are essential for verifying that certain chemicals will combat the virus. The team is hopeful that the computational results may inform future studies and provide a framework that experimentalists may use to further investigate the 77 small-molecule compounds.

PI(s)/Facility Lead(s): Jeremy C. Smith
ASCR Program/Facility: DD/OLCF
ASCR PM: Christine Chalk
Publication(s) for this work: Micholas Smith and Jeremy C. Smith, "Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface," *ChemRxiv* (2020). doi:10.26434/chemrxiv.11871402.v3.

See the COVID Dashboard at olcf.ornl.gov

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY
THE UNIVERSITY OF TENNESSEE KNOXVILLE
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Optimizing Performance for Summit – from Design to Operation

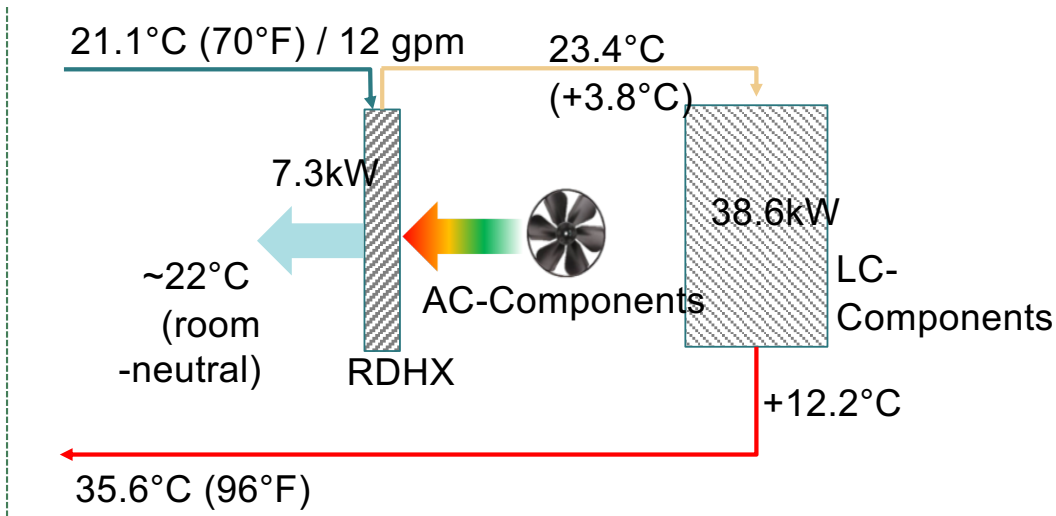
- Summit has 4,626 compute nodes
- Each node has 2 IBM POWER9 CPUs and 6 NVIDIA GV100 GPUs.
- Each of these components has a cold plate, with circulating water that removes no less than 75% of the waste heat, ejecting it directly to water.
- DIMMs and other components remain air-cooled.

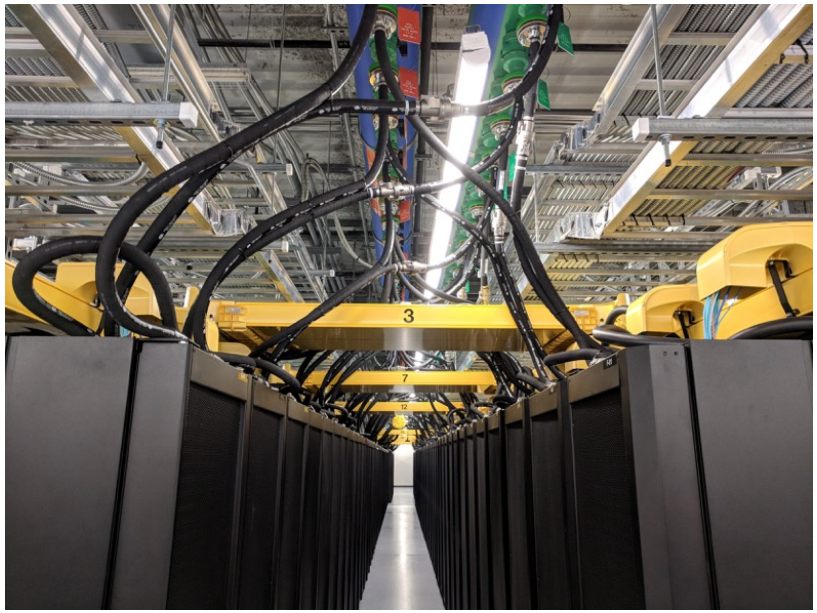
The Summit Compute Node

More than 100,000 water connections…

Summit's Mechanical Cooling Design – Single Cabinet

- 18 Summit compute nodes per cabinet
- Water supplied to each rack via a cabinet-level manifold
- 257 compute cabinets





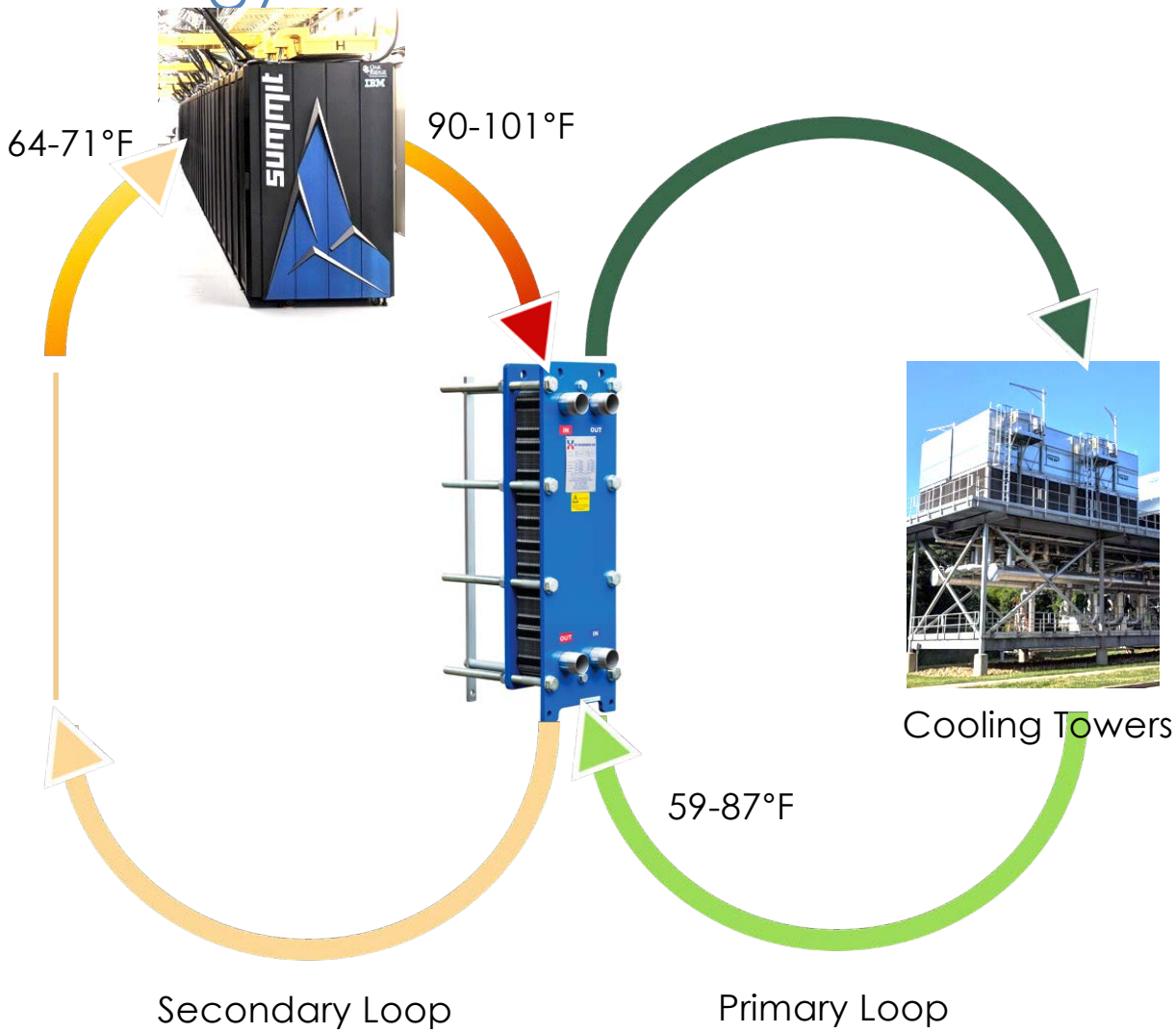
- 12-gpm/rack protects the equipment (18-gpm spec'd)
- Serial design eliminates second loop, increases d(t).

  No flow control

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Warm Water Central Energy Plant

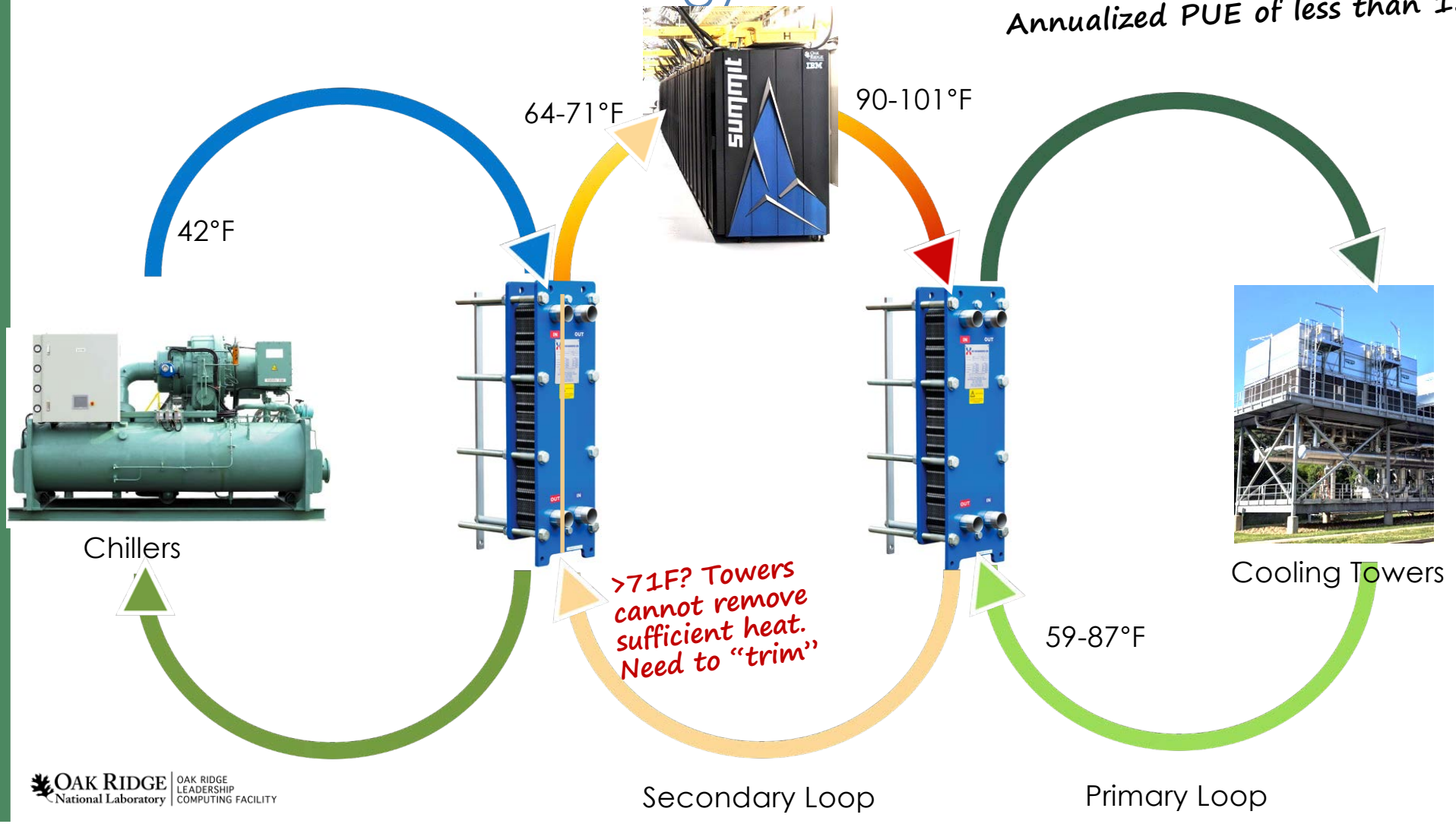
- More than 70% of the year, the Summit demand (to 13MW) can be managed using evaporative cooling only.
- PUE during these periods can be excessively low, below 1.03.



OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Warm Water Central Energy Plant

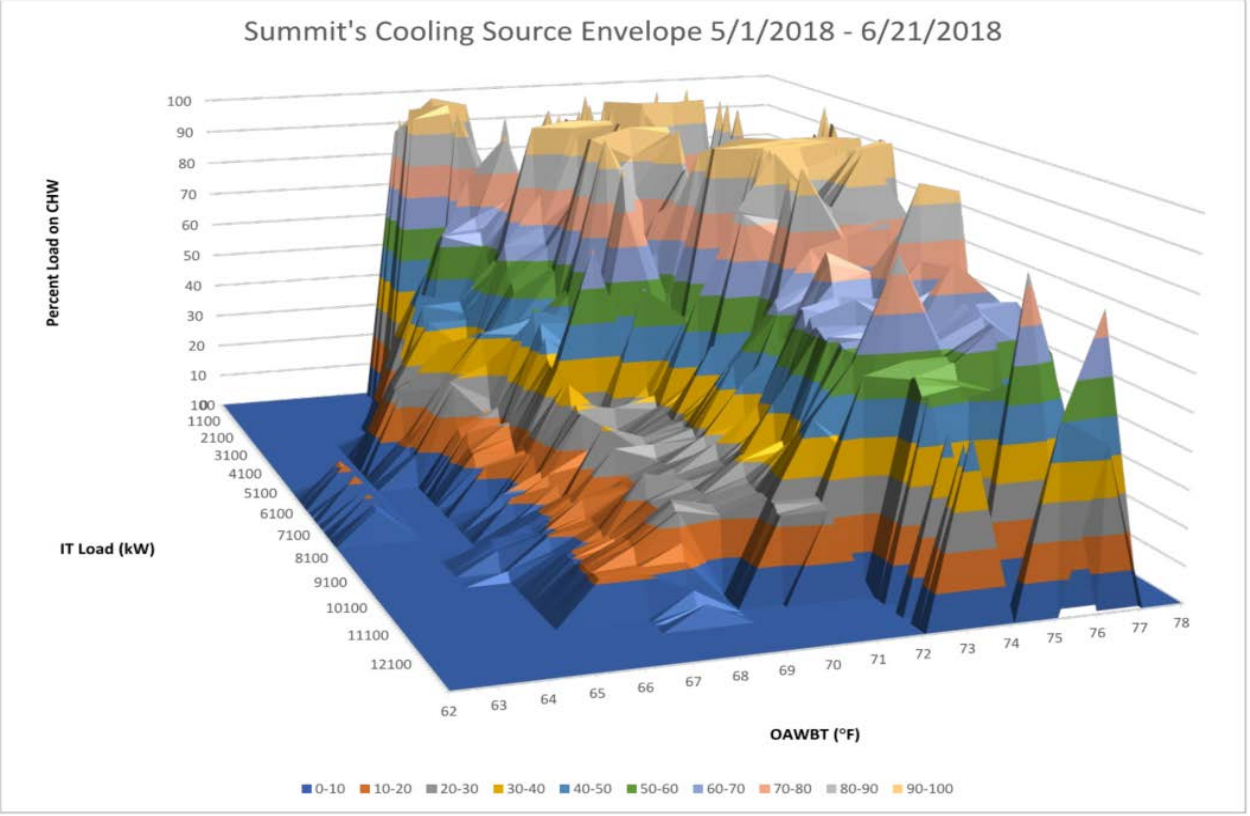
Annualized PUE of less than 1.10



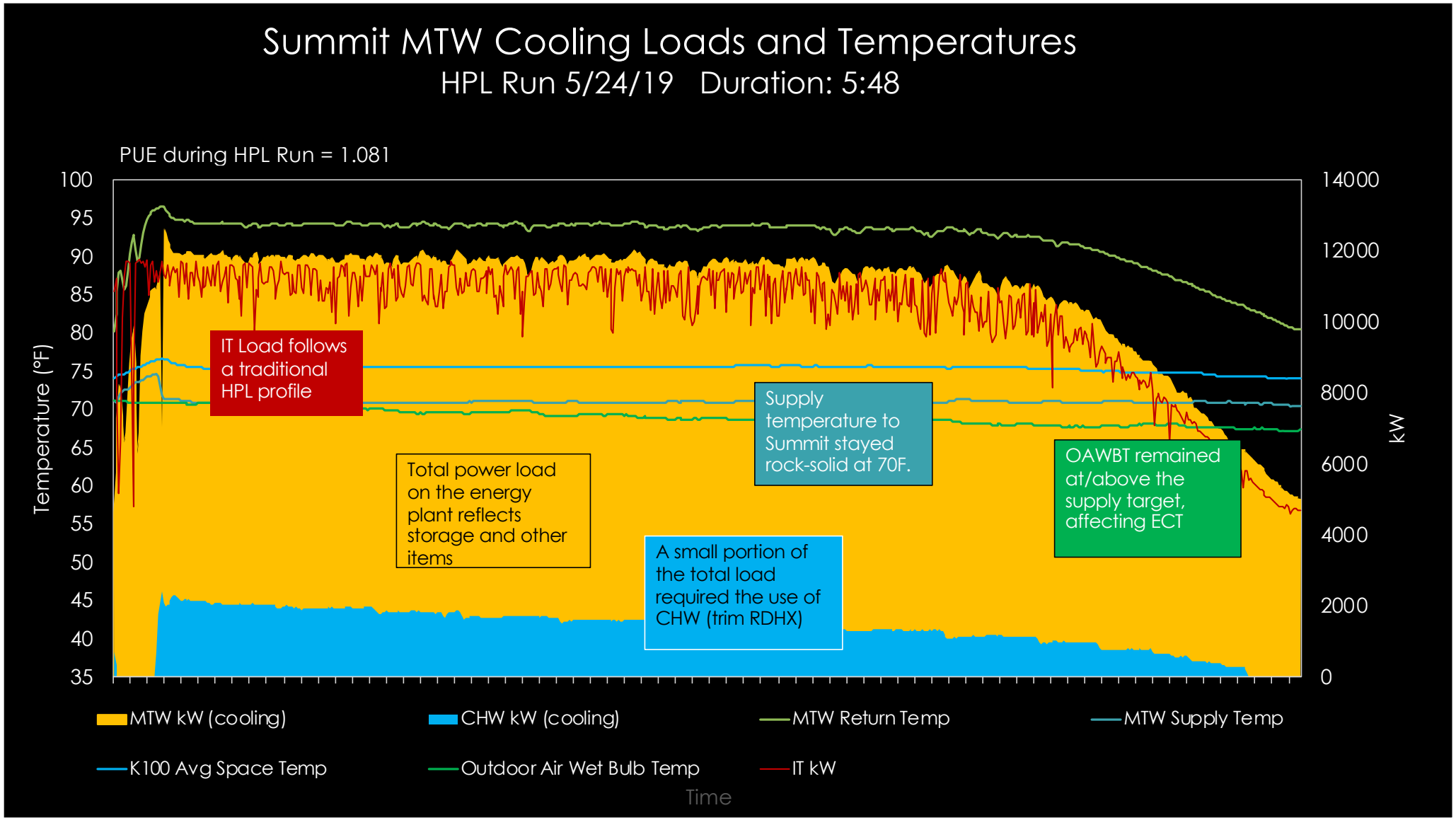
OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

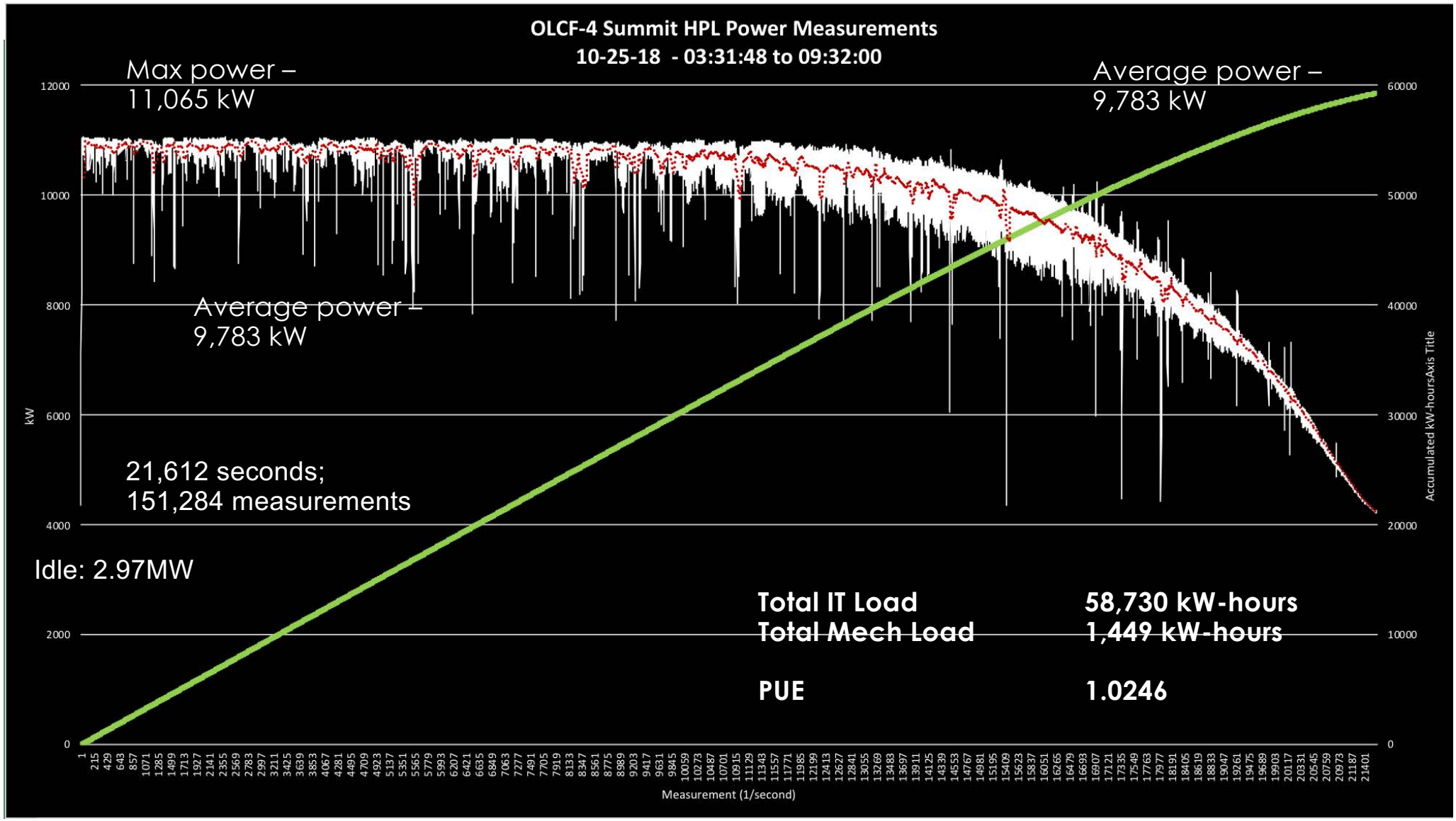
Optimization – Minimize CHW Use

- Summit to PLC interconnect
  - Machine Learning to inform the PLC (ongoing work)
  - Goals:
    - Push the MTW supply temperature higher
    - Reduce the MTW flow rate down to produce a higher dT
  - PLC will make decisions based on two envelopes
    - K100 Conditions (Full Range of 2015 ASHRAE Recommended Inlet Conditions)
    - Keep CPUs and GPUs from thermally throttling



Live Summit Cooling Metrics
Live Summit OpenBMC data stream scraped every 10 sec - All 4606 nodes
Active OpenBMC Datastreams
4606/4606
Active OpenBMC Datastreams
Active Data Streams Current: 4.606 K
Realtime Summit Power
3.36 MW
Supply Temperature
69.7 °F
Flow
3381 gpm
K100 PUE
1.057
Wetbulb (+0h)
15.1 °C
Wetbulb (+12h)
17.3 °C
Summit Total Power
Summit Total Power
Summit Power Breakdown
GPU Power Distribution (4606 x 6 : 5min)
CPU Power Distribution (4606 x 2 : 5min)
Node Usage By User (From Elastic Search)
Node Power Distribution (4606 nodes : 5min)
MTW Supply Temperature
Supply Temperature Current: 69.72 1h avg Current: 69.31
GPU Core Temp
GPU Mem Temp
CPU Core Temp
DIMM Temp
MTW Flow
MTW Flow (gpm) Current: 3.370 K 1h avg Current: 3.352 K
K100 PUE
K100 PUE Current: 1.0658 1h avg Current: 1.0575


Summit MTW Cooling Loads and Temperatures
HPL Run 5/24/19 Duration: 5:48
PUE during HPL Run = 1.081
Temperature (°F)
100
95
90
85
80
75
70
65
60
55
50
45
40
35
14000
12000
10000
8000
6000
4000
2000
0
kW
IT Load follows a traditional HPL profile
Total power load on the energy plant reflects storage and other items
Supply temperature to Summit stayed rock-solid at 70F.
A small portion of the total load required the use of CHW (trim RDHX)
OAWBT remained at/above the supply target, affecting ECT
MTW kW (cooling)
CHW kW (cooling)
MTW Return Temp
MTW Supply Temp
K100 Avg Space Temp
Outdoor Air Wet Bulb Temp
IT kW
Time


OLCF-4 Summit HPL Power Measurements
10-25-18 - 03:31:48 to 09:32:00
Max power –
11,065 kW
Average power –
9,783 kW
Average power –
9,783 kW
21,612 seconds;
151,284 measurements
Idle: 2.97MW
Total IT Load
58,730 kW-hours
Total Mech Load
1,449 kW-hours
PUE
1.0246
kW
Accumulated kW-hoursAxis Title
Measurement (1/second)

BANDWIDTH

The network bandwidth of the Frontier system is 24,000,000 times greater than the top home internet connection. With it, you could download

100,000
HD MOVIES IN ONE SECOND

HORSEPOWER

Frontier will have the performance of the **TOP 160** fastest supercomputers in the world **COMBINED**

SIZE

Frontier will cover over 7,300 square feet.

That's almost **2** **BASKETBALL COURTS**

SPEED

If all 7.7 billion people on earth each completed one calculation per second, it would take over

6 YEARS

to do what the Frontier system can do in

1 SECOND

CABLING

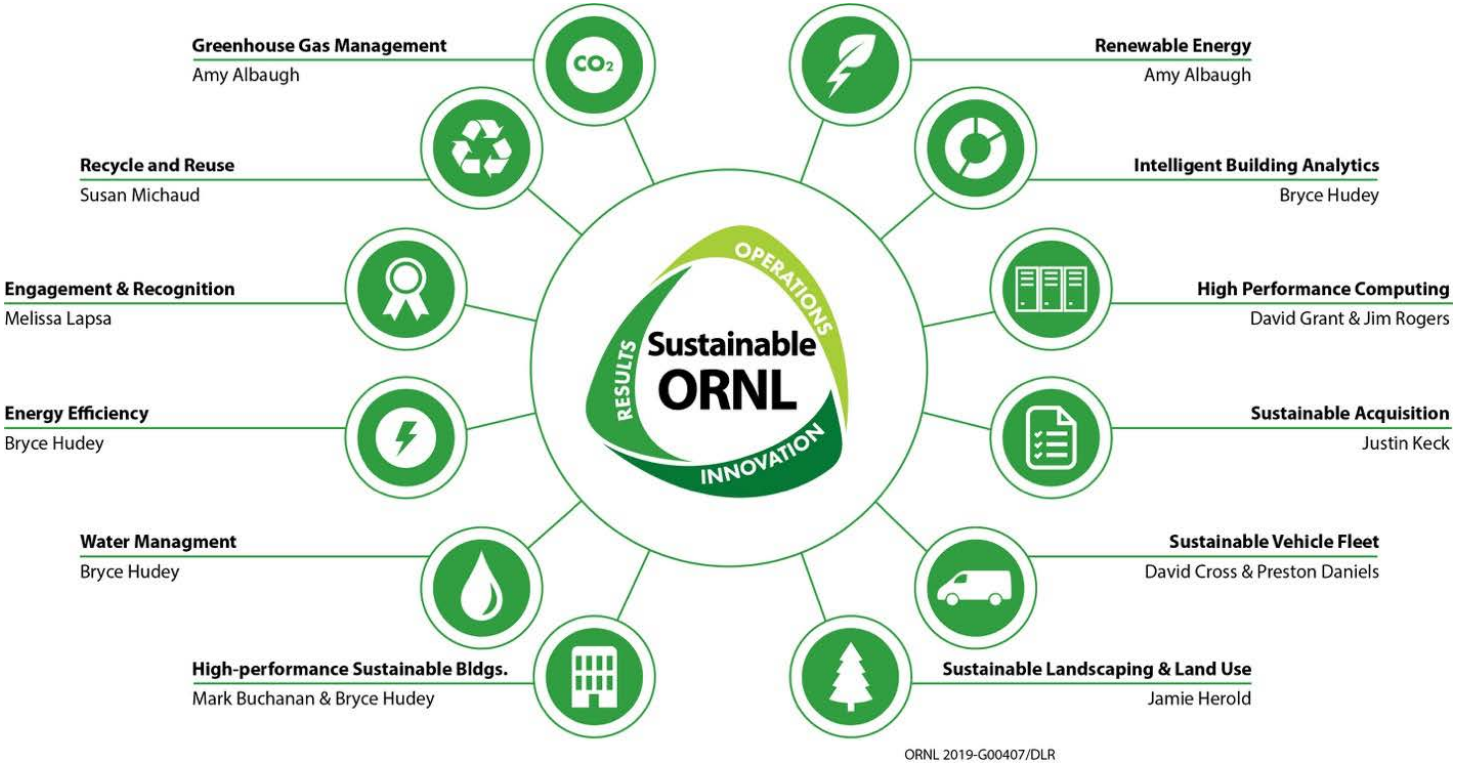
THE 90 MILES OF CABLES

IN THE FRONTIER SYSTEM would span the distance from Philadelphia to New York City

CRAY

© 2019 Cray Inc.

Sustainable ORNL



Thank you!

Sustainable ORNL

OPERATIONS · RESULTS · INNOVATION