OAK RIDGE National Laboratory

Resource Management: Assessing the Warm Water Facilities and Systems Supporting ORNL's Summit

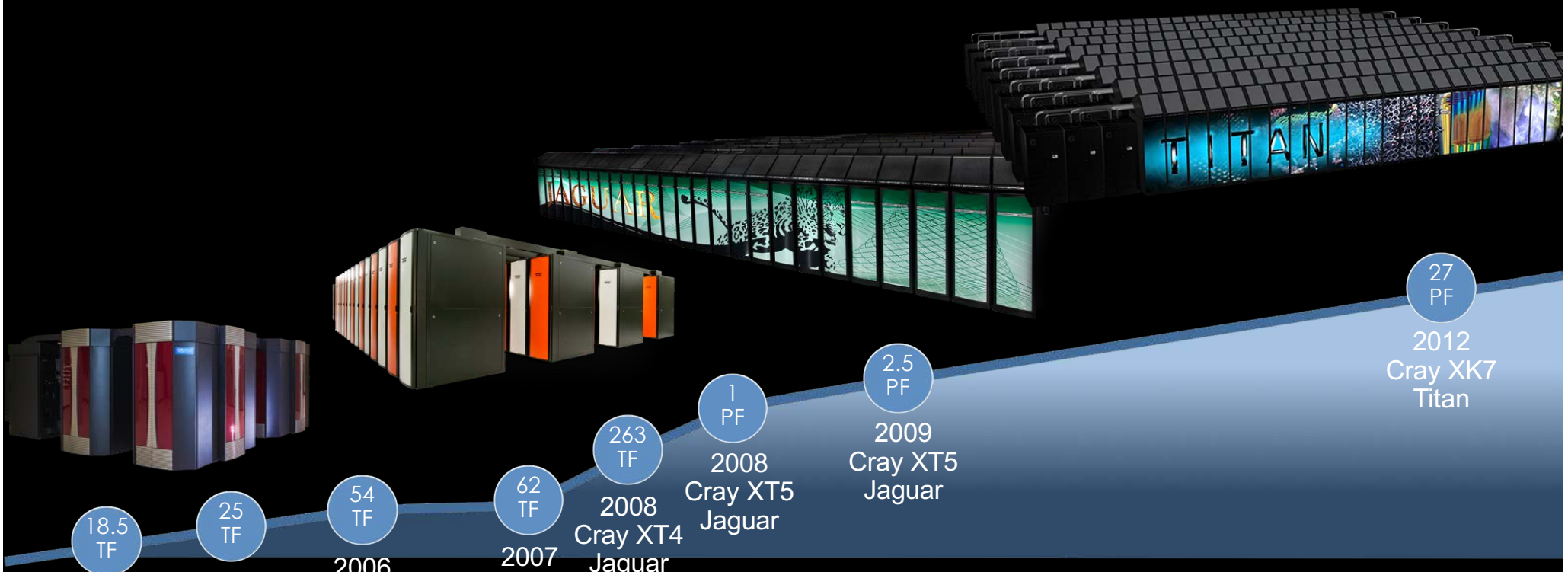
Jim Rogers

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

U.S. DEPARTMENT OF ENERGY

ORNL Leadership Class Systems 2004 - 2018



From 2004 – 2018, HPC systems relied on chiller-based cooling (5.5°C / 42°F supply) with annualized PUEs to ~1.4

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

ORNL's Transition to Warmer Facility Supply Temperatures

Titan: Refrigerant-based per-rack cooling with direct rejection of heat to cold 5.5°C water

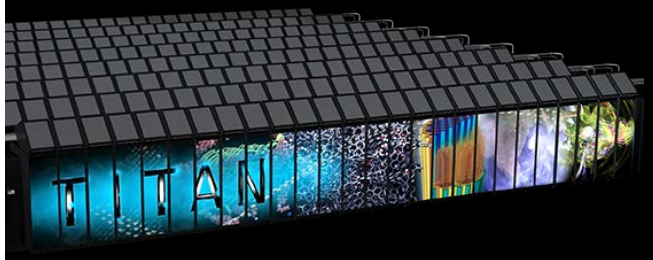
- ***Below dewpoint***
- ***100% use of chillers***

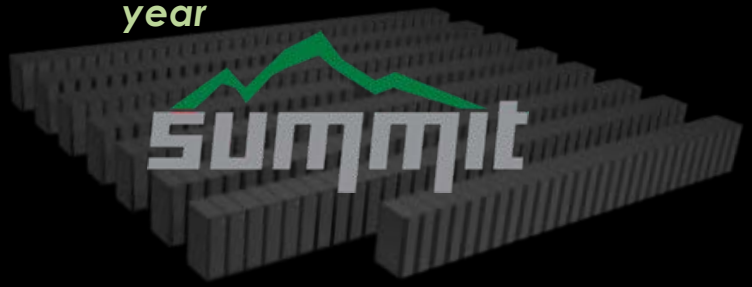
Summit: A combination of direct on-package cooling and RDHX with 21°C / 70°F supply is > 95% room-neutral.

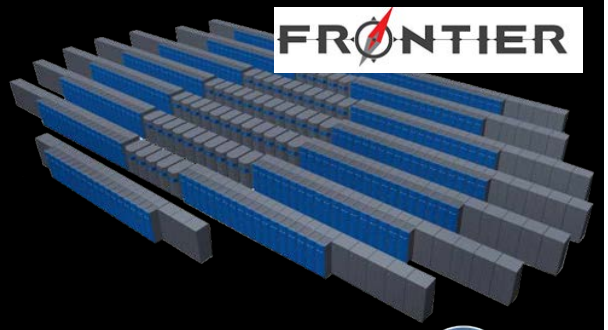
- ***Above dewpoint***
- ***Contribution by chillers ~20% of the hours of the year***

Frontier: Custom packaging is >95% room-neutral with a 32°C / 90°F supply.

- ***~100% Evaporative Cooling, with supplemental HVAC for parasitic loads***







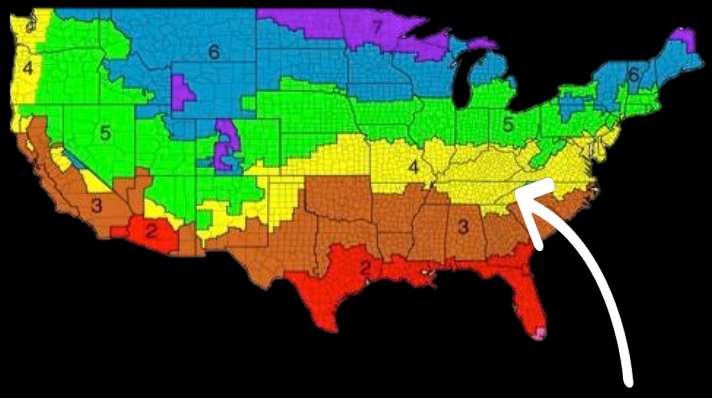
27 PF
2012
Cray XK7
Titan

200 PF
2018/2019
IBM
Summit

~1.5 EF
2021/2022
HPE/Cray
Frontier

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Motivation for "Warmer" Cooling Solutions Serving HPC Centers

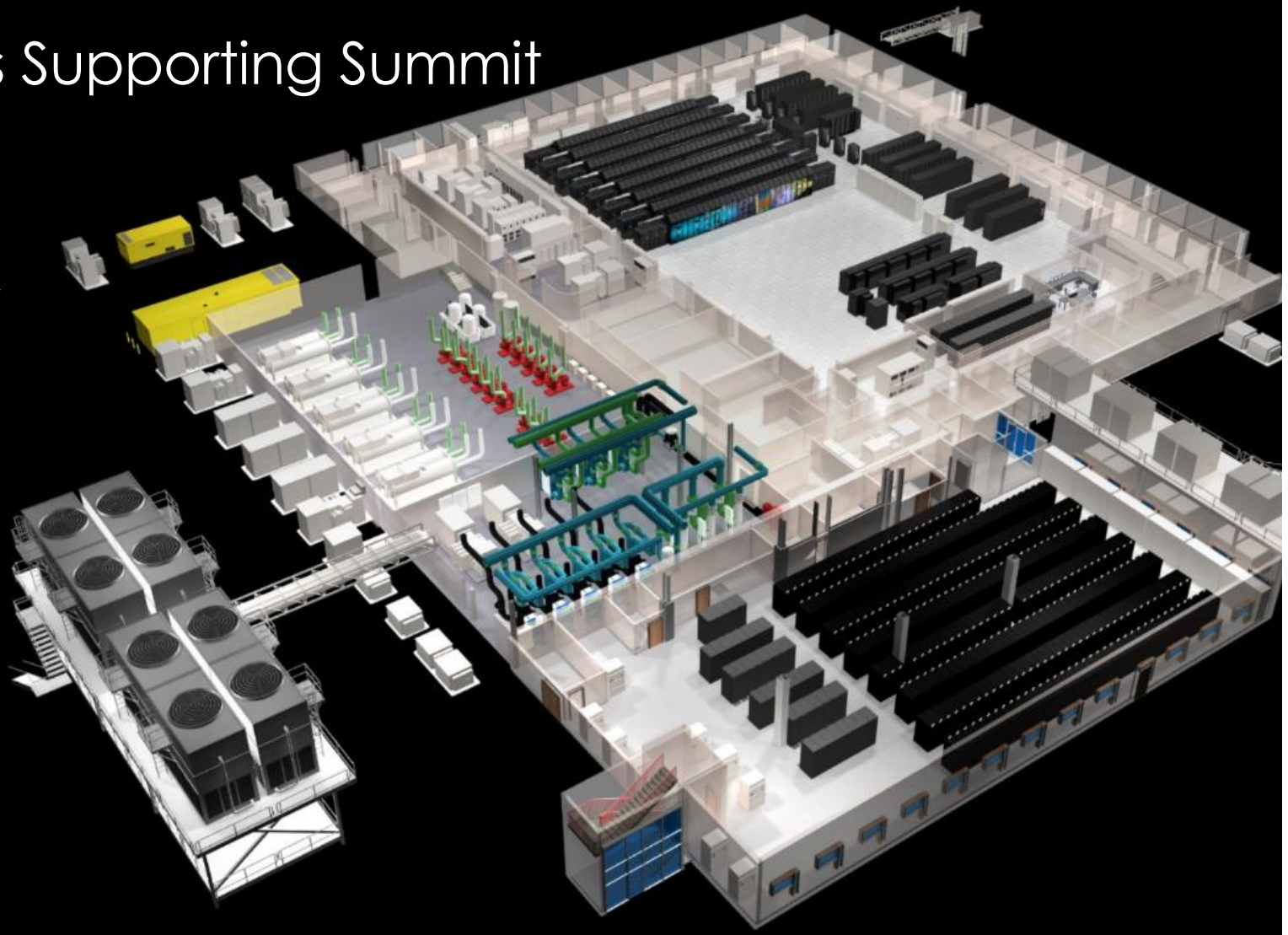


- Reduced Cost, Both CAPEX and OPEX
  - Reduce or eliminate the need for traditional chillers.
    - No chillers, no ozone-depleting refrigerants (GREEN)
  - Oak Ridge calculates an annualized PUE for air--cooled devices of no better than 1.4 (ASHRAE Zone 4A – Mixed Humid)
  - HPC power budgets continue to grow – Summit has a design point for >12MW (HPC-only). Minimizing PUE/ITUE is critical to the budget.
- Easier, more reliable design
  - Design is reduced to pumps, evaporative cooling, heat exchangers.
  - Traditional chilled water may not be necessary at all (NREL, NCAR, et al)

OLCF Facilities Supporting Summit

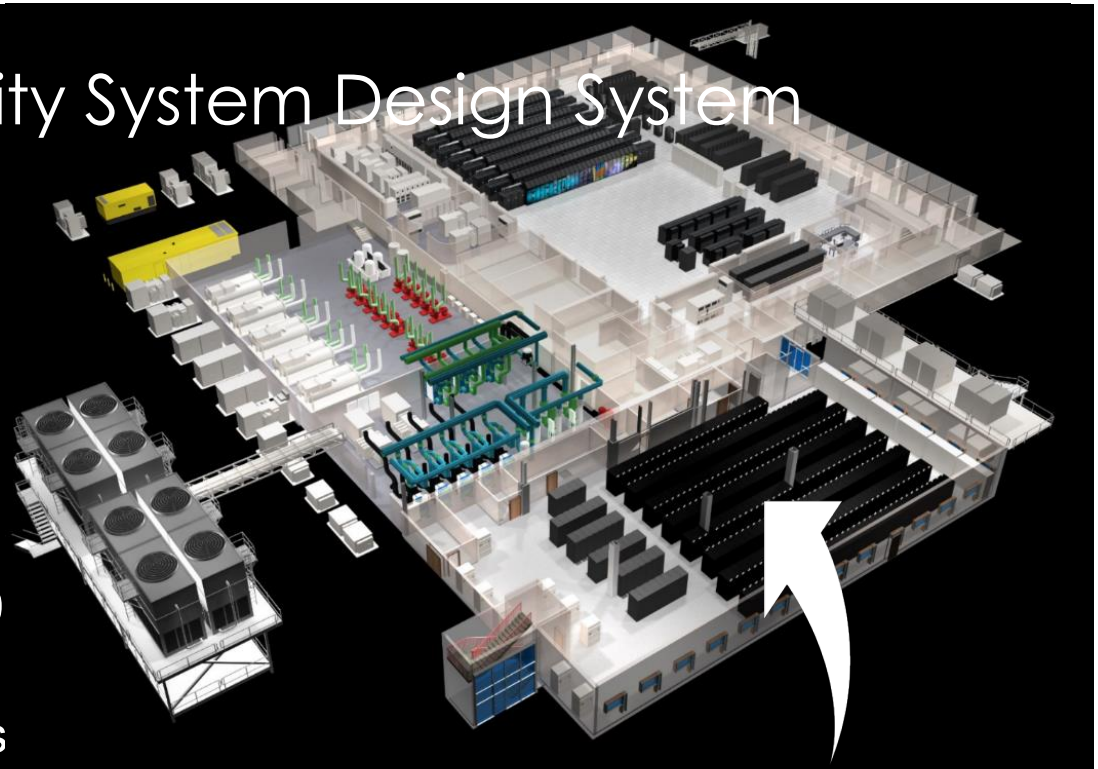
- Titan – 9MW @ heavy load
- Sitting on 250 pounds/ft2 raised floor
- Uses 42F water and special CDUs (XDPs)

- Summit – 256 compute cabinets on-slab
- 100% room-neutral design uses RDHX
- 20MW warm-water cooling plant using centralized CDU/secondary loop



21°C/20MW/7700 ton Facility System Design System

- Summit
  - Demand: 3.3 (idle)-11.5MW;
- Secondary Loop
  - Supply 3300GPM (12,500 liters/min) @ 21°C; Return @ 29 33°C
  - CPUs and GPUs use cold plates
  - DIMMs and parasitic loads use RDHX
  - Storage and Network use RDHX



Medium Water Temperature Cooling Design

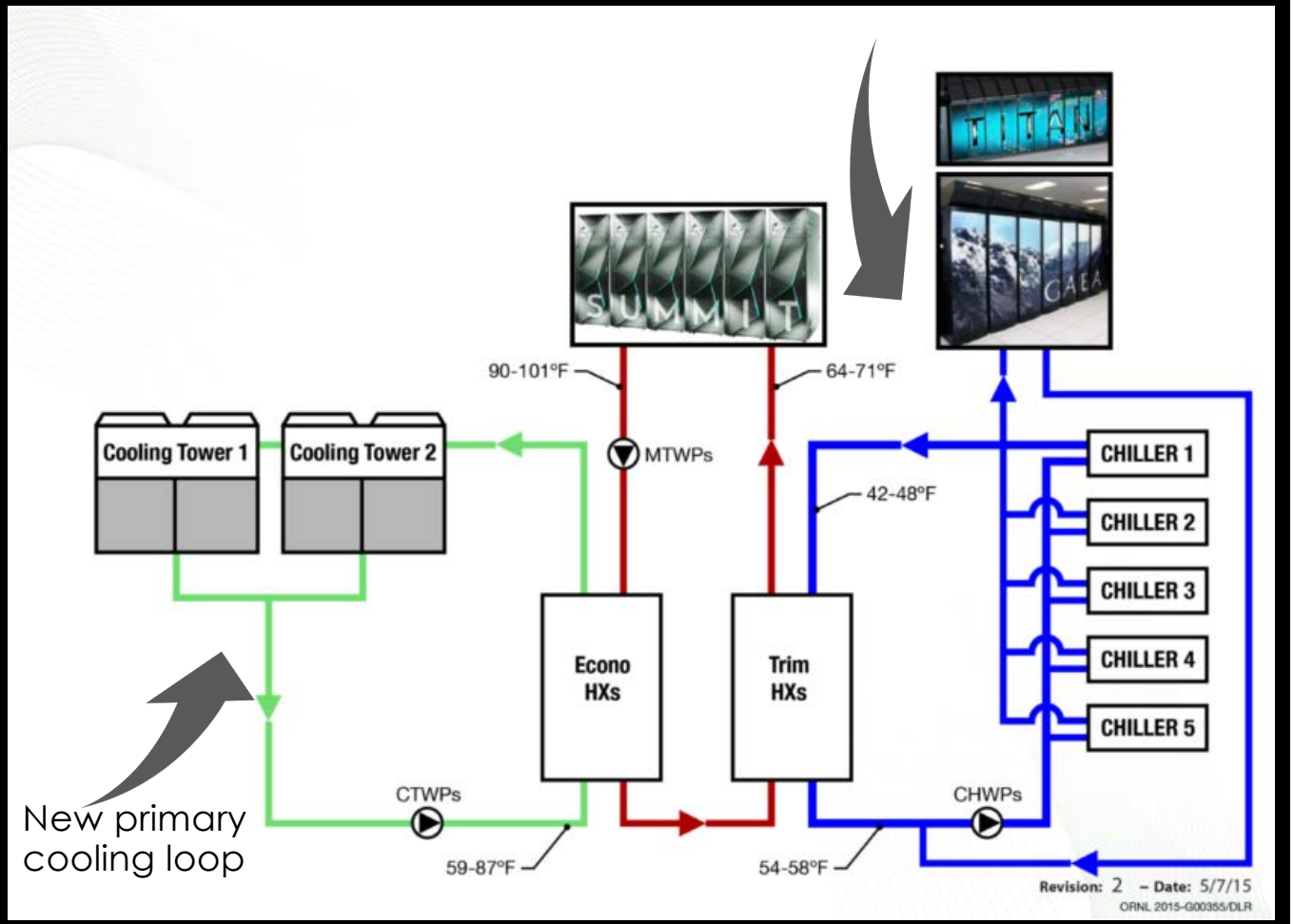
Primary Loop uses Evaporative Cooling Towers (~80% of the hours of the year)

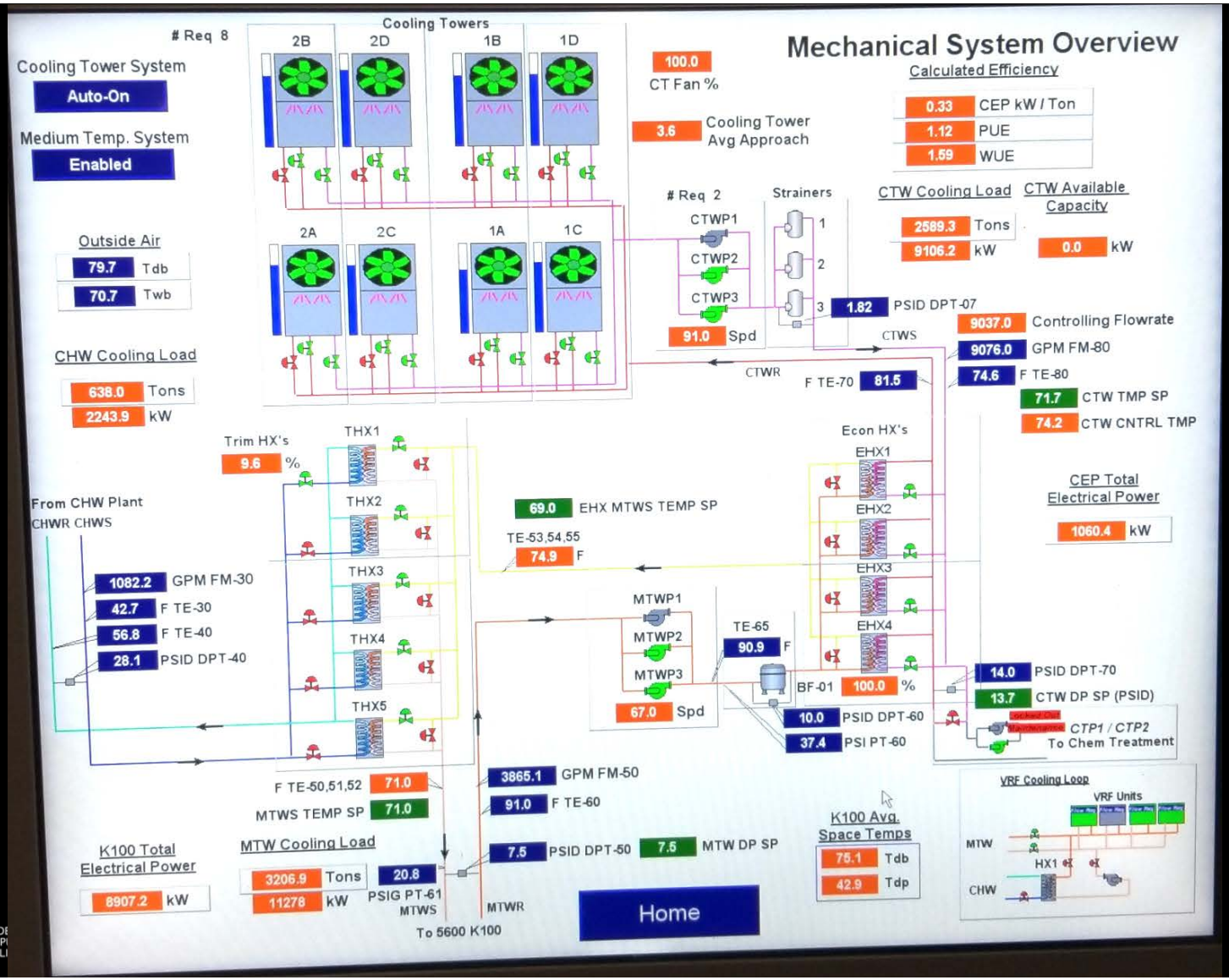
When the MTW RETURN is above the 21C setpoint, use a second set of Trim HX (with 5.5C on the other side) to drive MTW to the 21C setpoint.

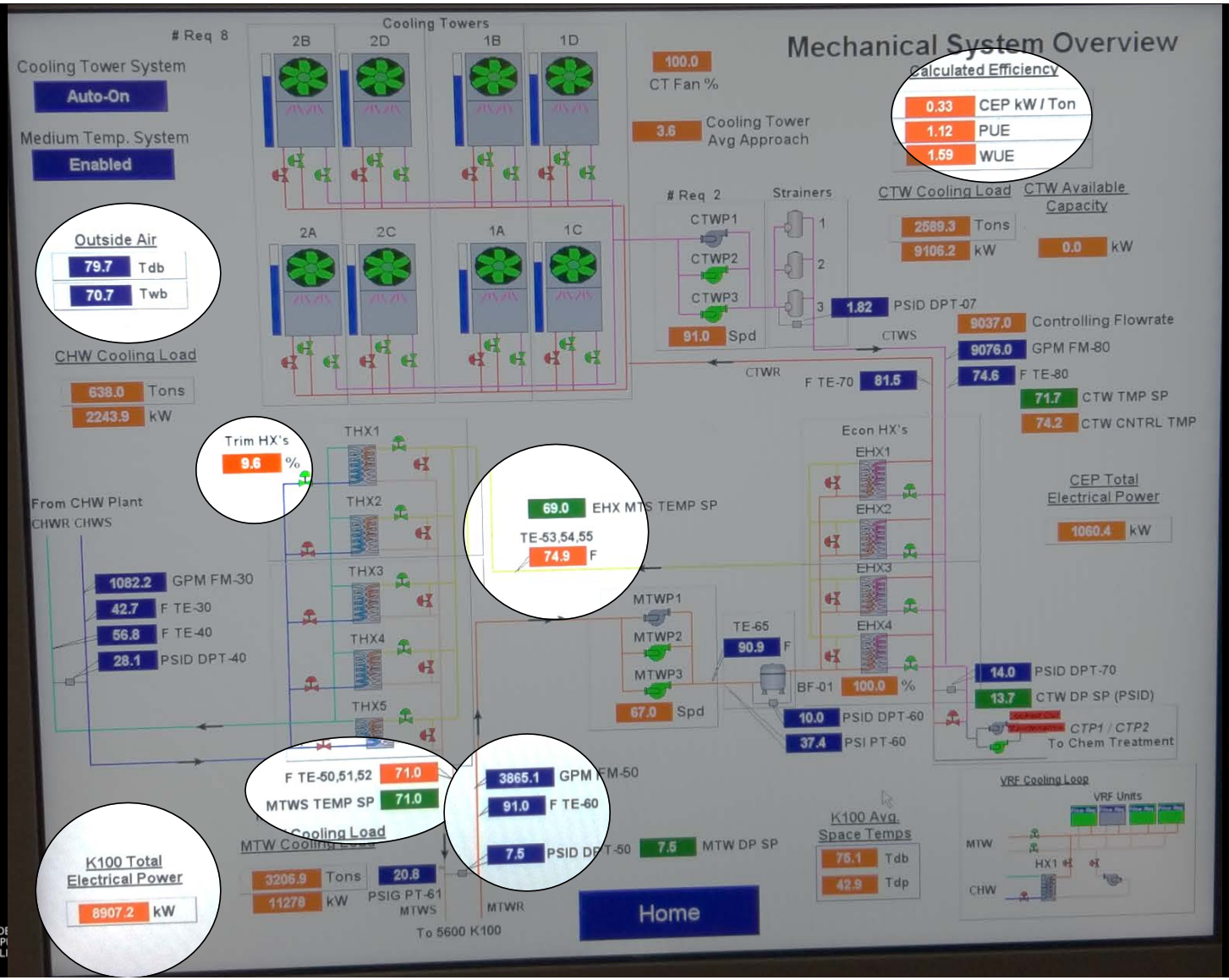
The need for the trim-loop is about 20% of the hours in the year, and can ramp 0-100% to meet the setpoint back to Summit

A New Challenge – Titan and rotated off-line, so there is no significant load on the Chillers[1-5]. NOAA moved to the warm water loop (~500GPM demand)

Existing chilled water cooling loop







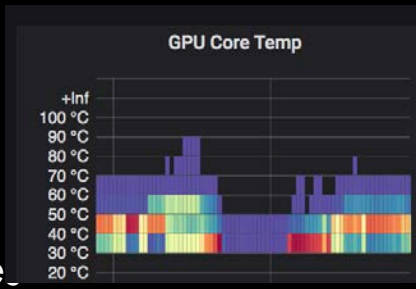
Benefits of ORNL's Warm-Water (21°C) Mechanical Plant

Warm Water

- Total Capacity to manage 20MW of IT load
- Warm Water allows annualized PUE of 1.1
  - For each ~$1M cost *per MW-year* for consumption on Summit;
  - A corresponding ~$100k cost *per MW-year* for waste heat management
- Titan was $6-7M/year *plus* $1.8-$2M to remove the waste heat
- Summit will have a similar "power bill", but closer to $500k annually for waste heat.

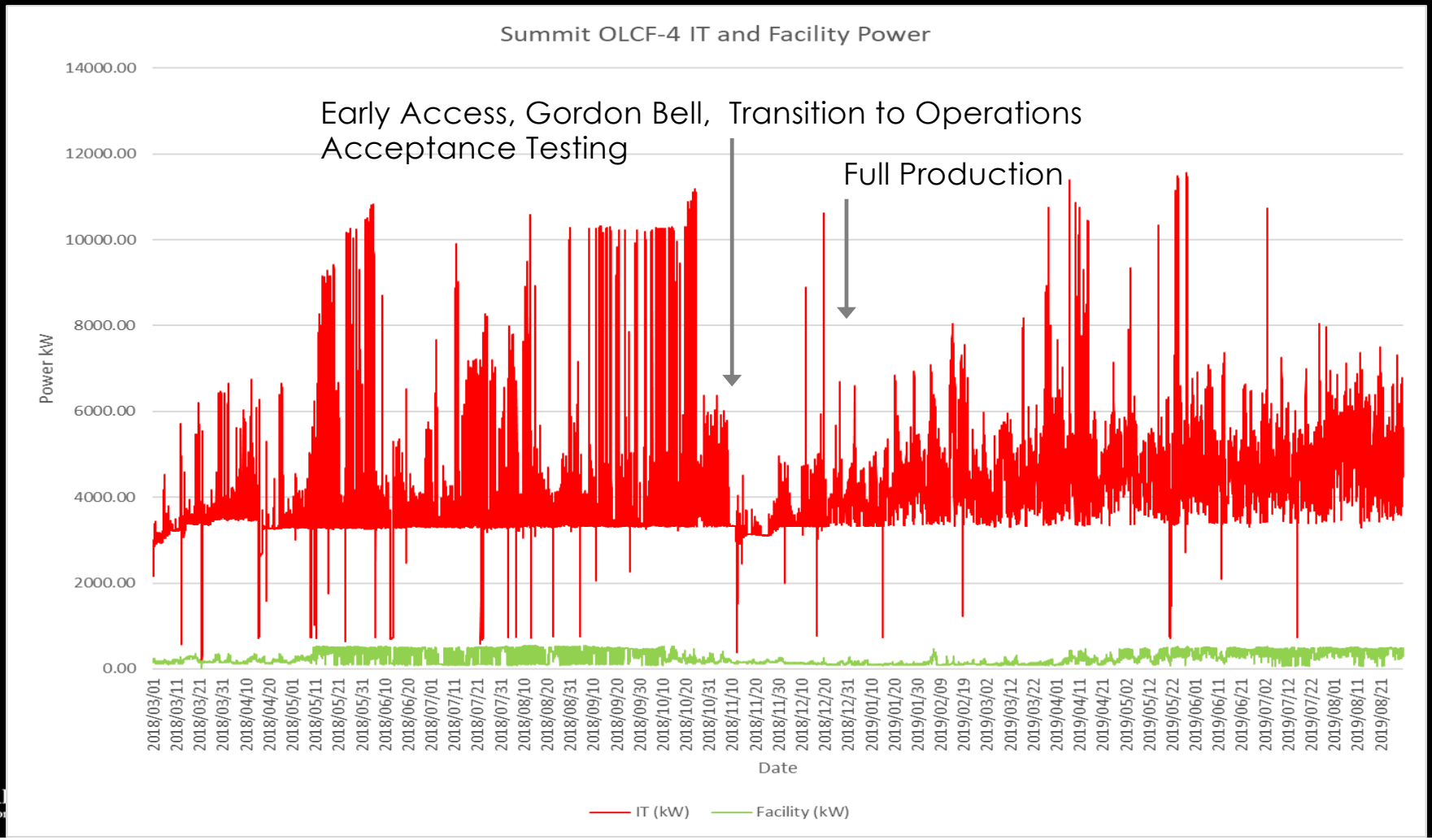
Integrated System/Facility Operation

- Integration with the PLC allows us to tune water pressure and flow
  - Better delta(t); less pumping energy
- Integration with IBM's OpenBMC provides information necessary to protect ~37K CPUs and GPUs from inadequate flow across the cold plate.

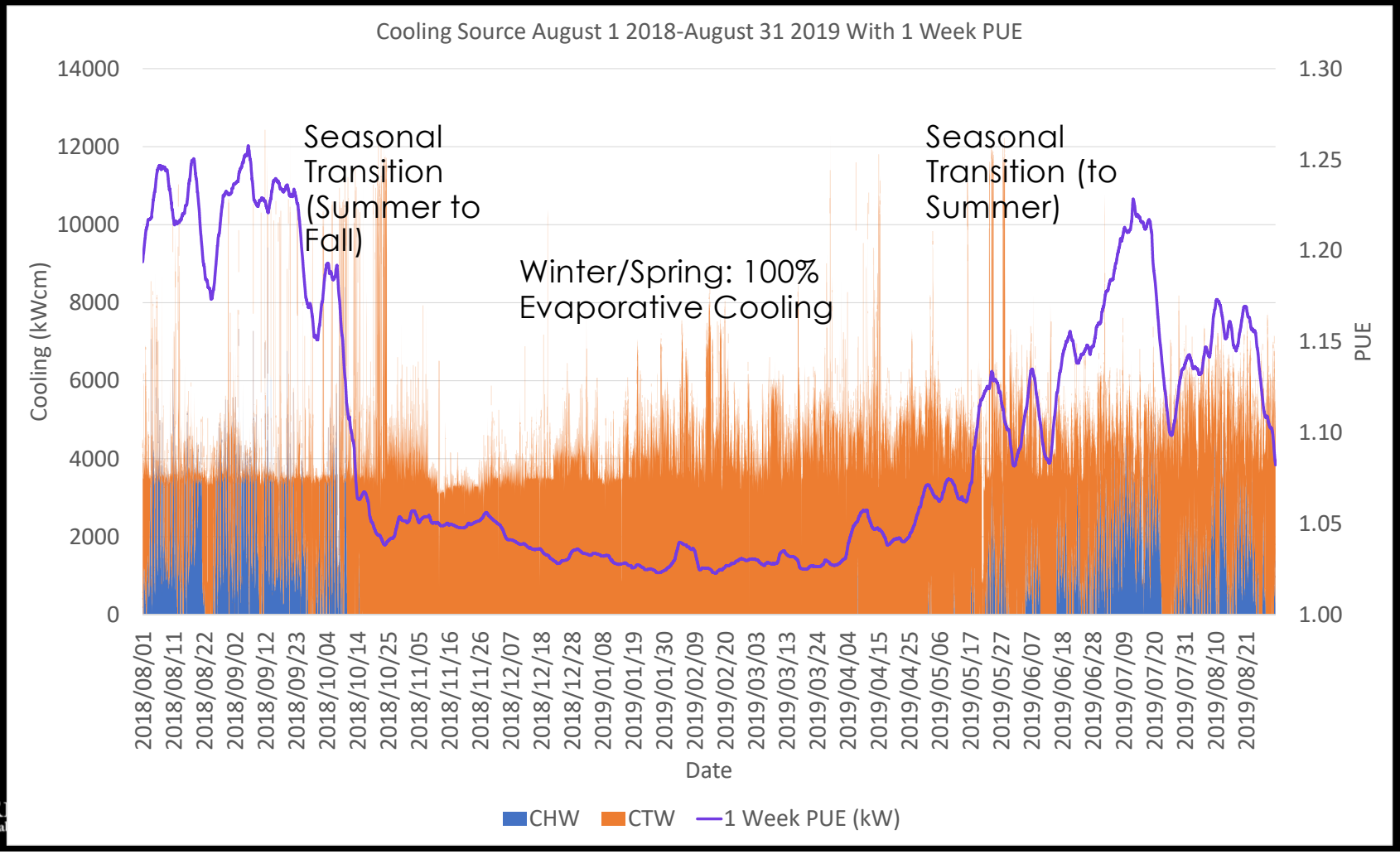


- Integration with the scheduler allows us to correlate power and temperature data with individual applications

Summit's Power Demand, Aug 2018– Aug 2019



Detailed Analysis of Summit's Annual PUE



Summit's RM Challenge – Data Volume

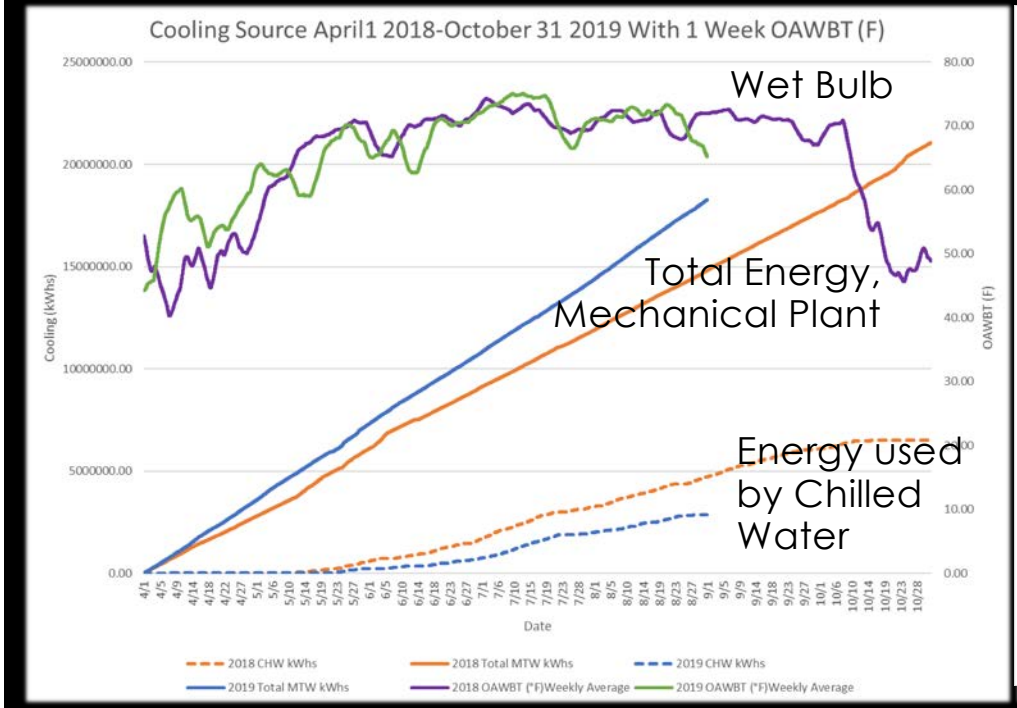
- Data Streams Include:
  - IBM OpenBMC framework (99 metrics/node/second x 4608 nodes = ~460,000/sec) - OOB
  - IBM LSF jobs data for running applications (~10sec update interval)
  - NOAA weather/wet bulb for Oak Ridge area (continuous, external)
  - Programmable Logic Circuit (PLC) water flow (continuous, protected as part of BAS)

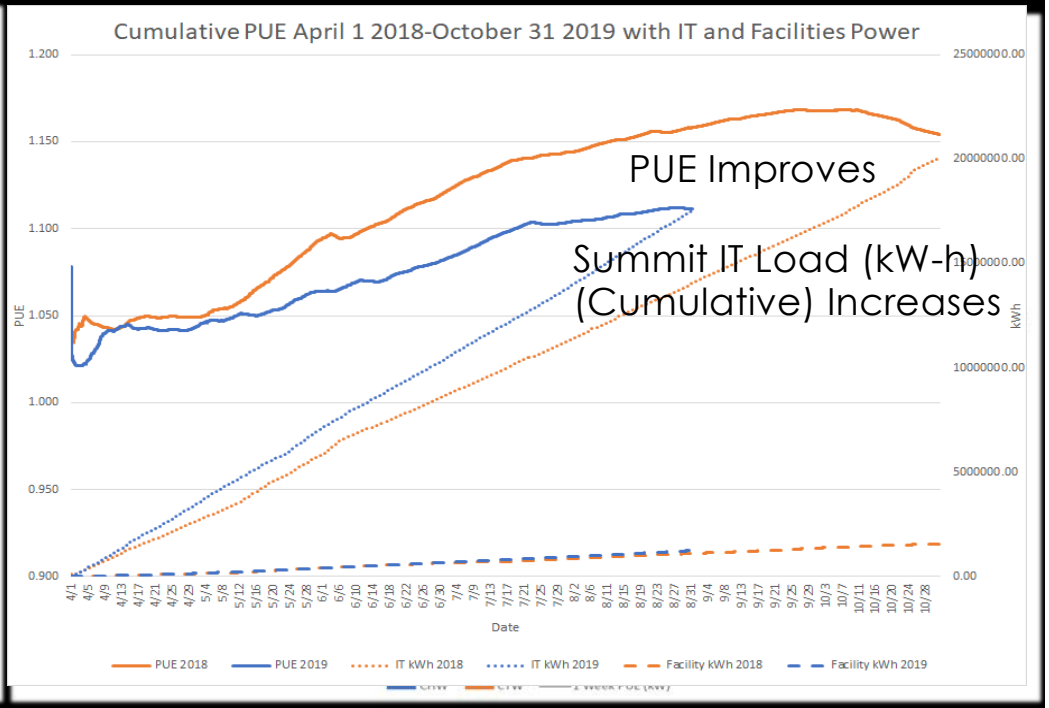
TOTAL ~470,000 metrics per second available for real-time analytics

Sample from Grafana Dashboard (live version shown)

MTW Performance

2018:Supply 70°F & ~4500GPM
2019:Supply 71°F & ~3300GPM





Year over Year, 2018 to 2019:

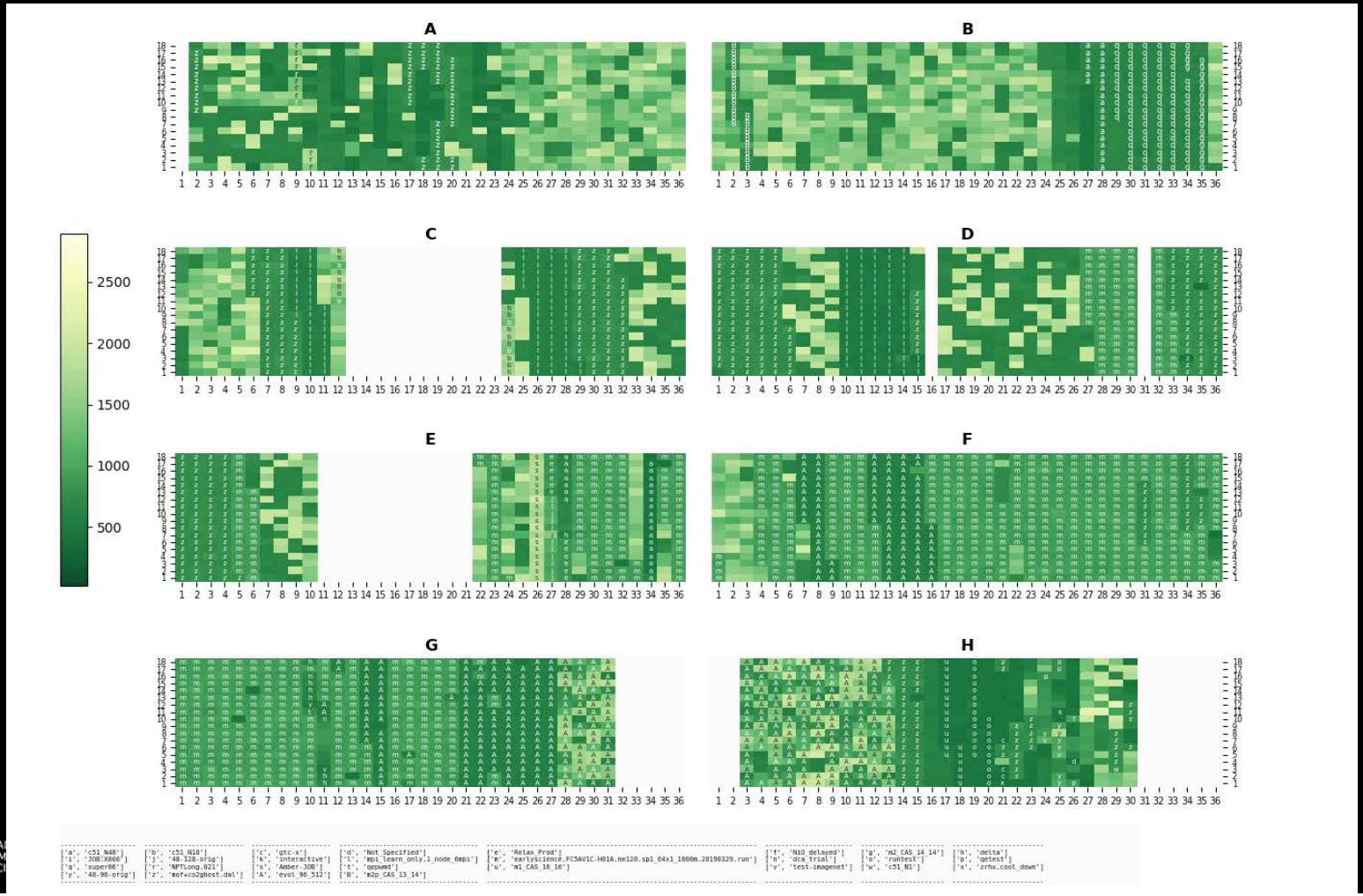
- IT Demand (kW-h) is up 19%
- CHW use is 39% LESS than in 2018
- The one-degree adjustment of supply temperature provided a savings of $40,000

Summit (System) Impact:

- No substantive impact to CPU or GPU operating conditions (junction temperatures)
- One degree increase in room temperature (worst case about 72.5° F)

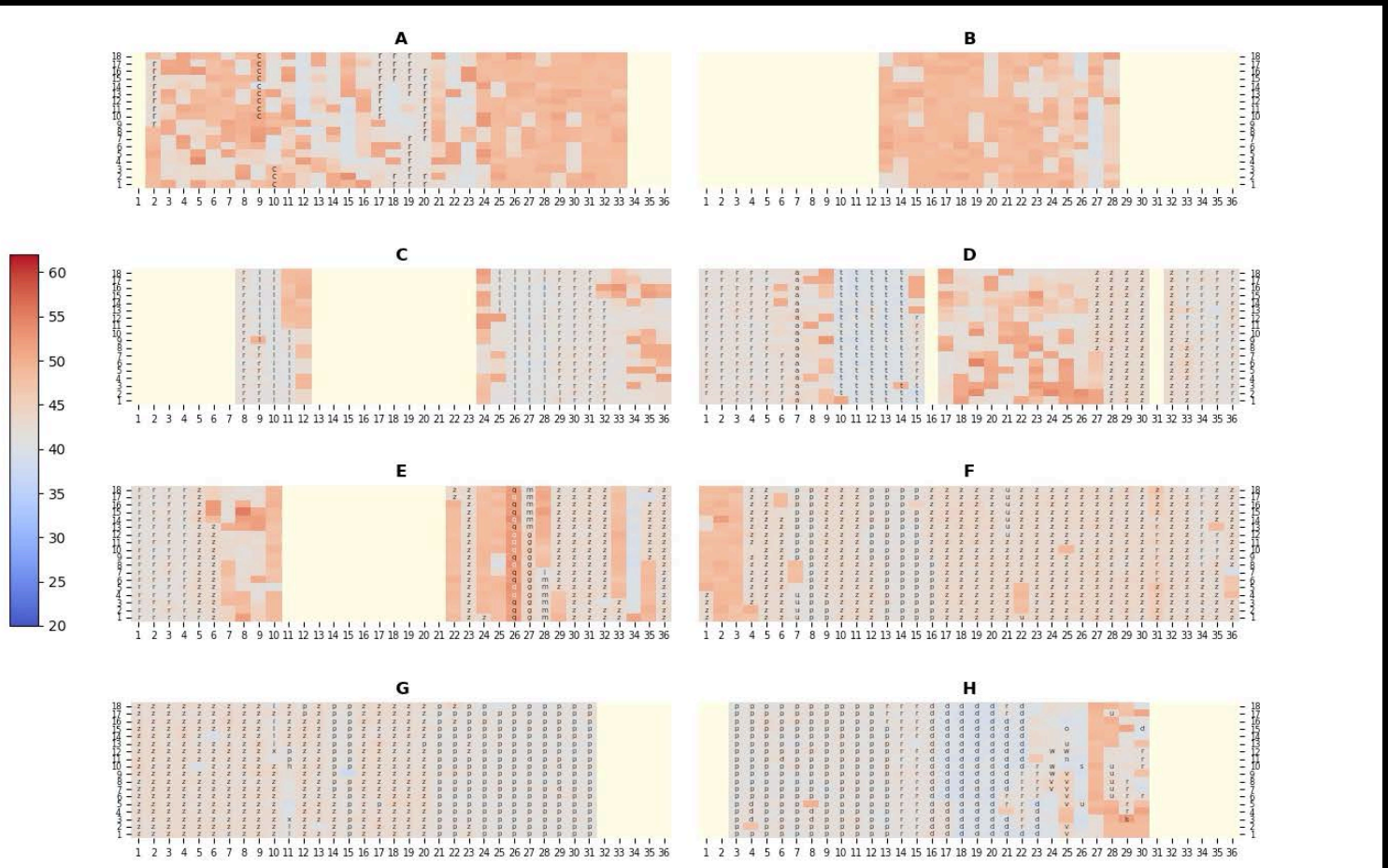
Summit Power/Jobs Every Second

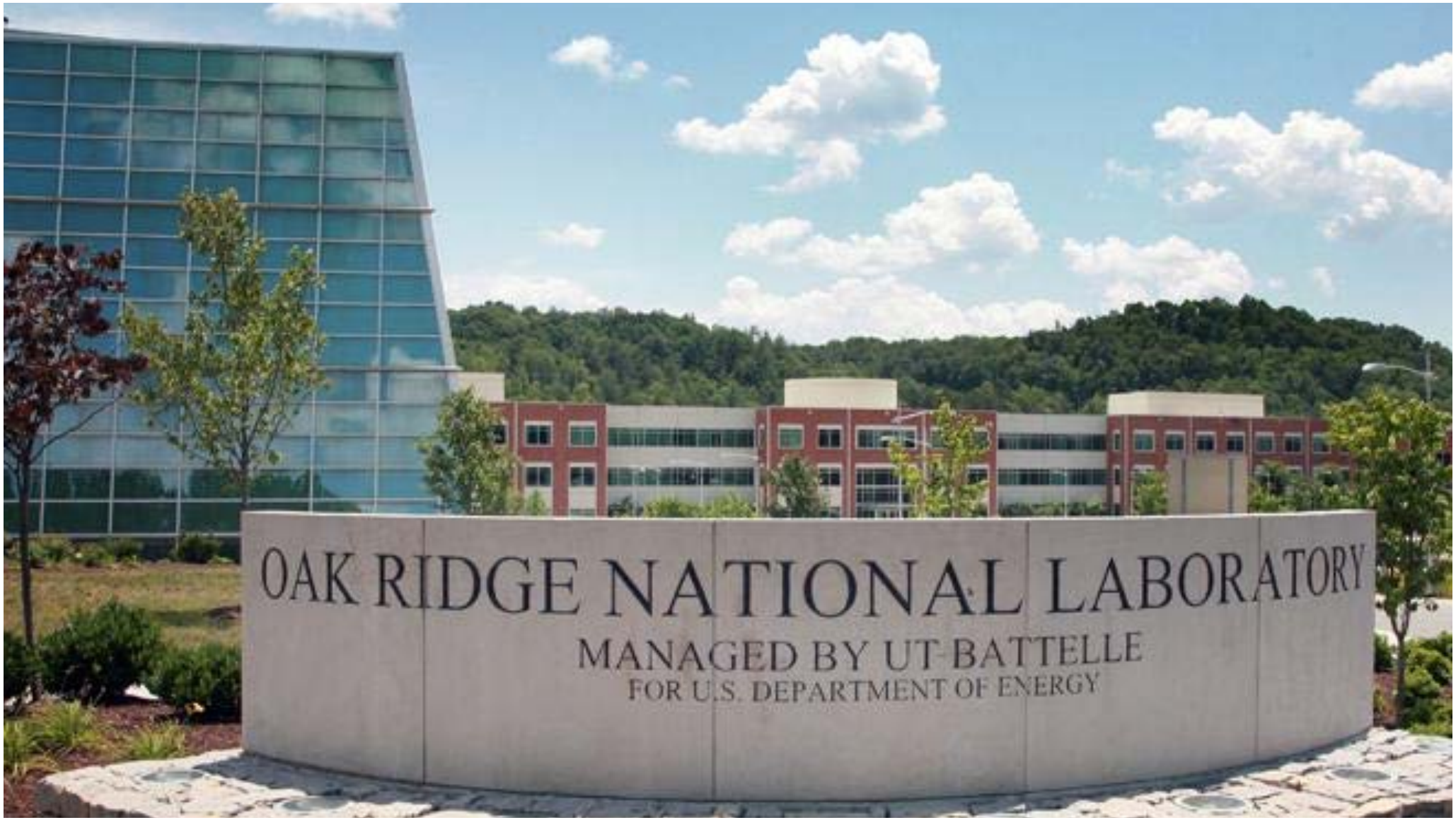
Click to Play Movie



Summit Temperature/Jobs Every Second

Click to Play Movie




OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR U.S. DEPARTMENT OF ENERGY