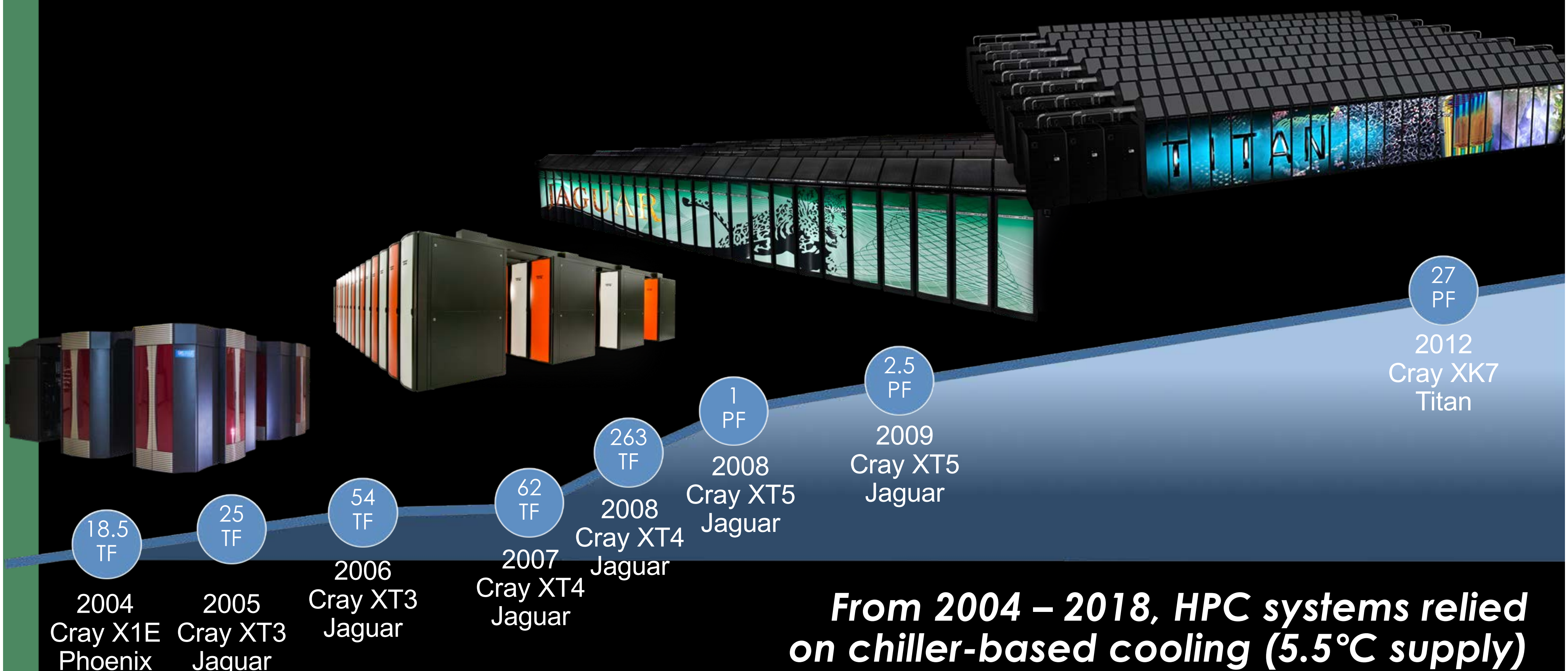
Summer 2019 - Assessing the Facilities and Systems Supporting HPC at ORNL

Jim Rogers

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

ORNL Leadership Class Systems 2004 - 2018



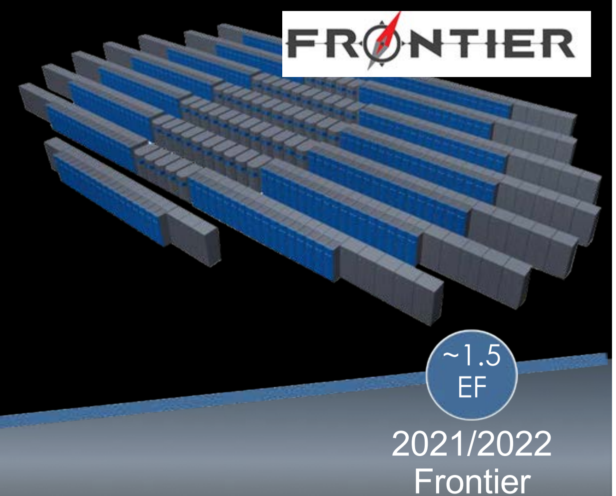
From 2004 – 2018, HPC systems relied on chiller-based cooling (5.5°C supply) with annualized PUEs to ~1.4

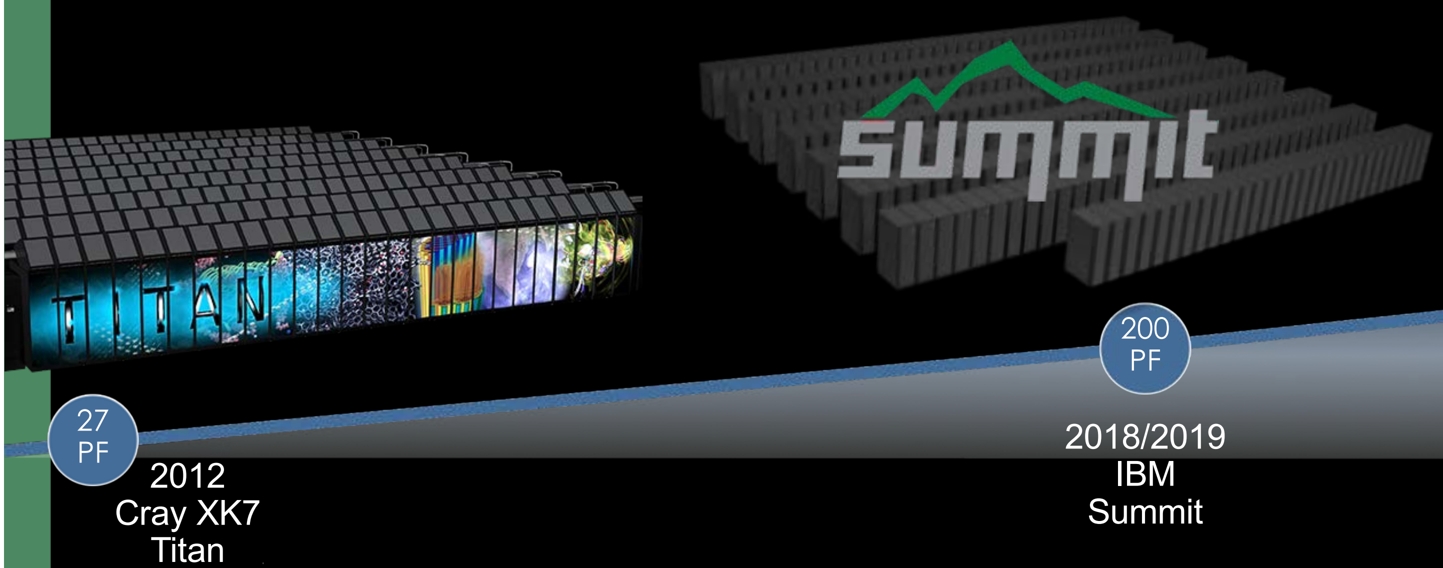
ORNL's Transition to Warmer Facility Supply Temperatures

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 supply is > 95% room-neutral.
- *Above dewpoint*
- *Contribution by chillers ~20% of the hours of the year*

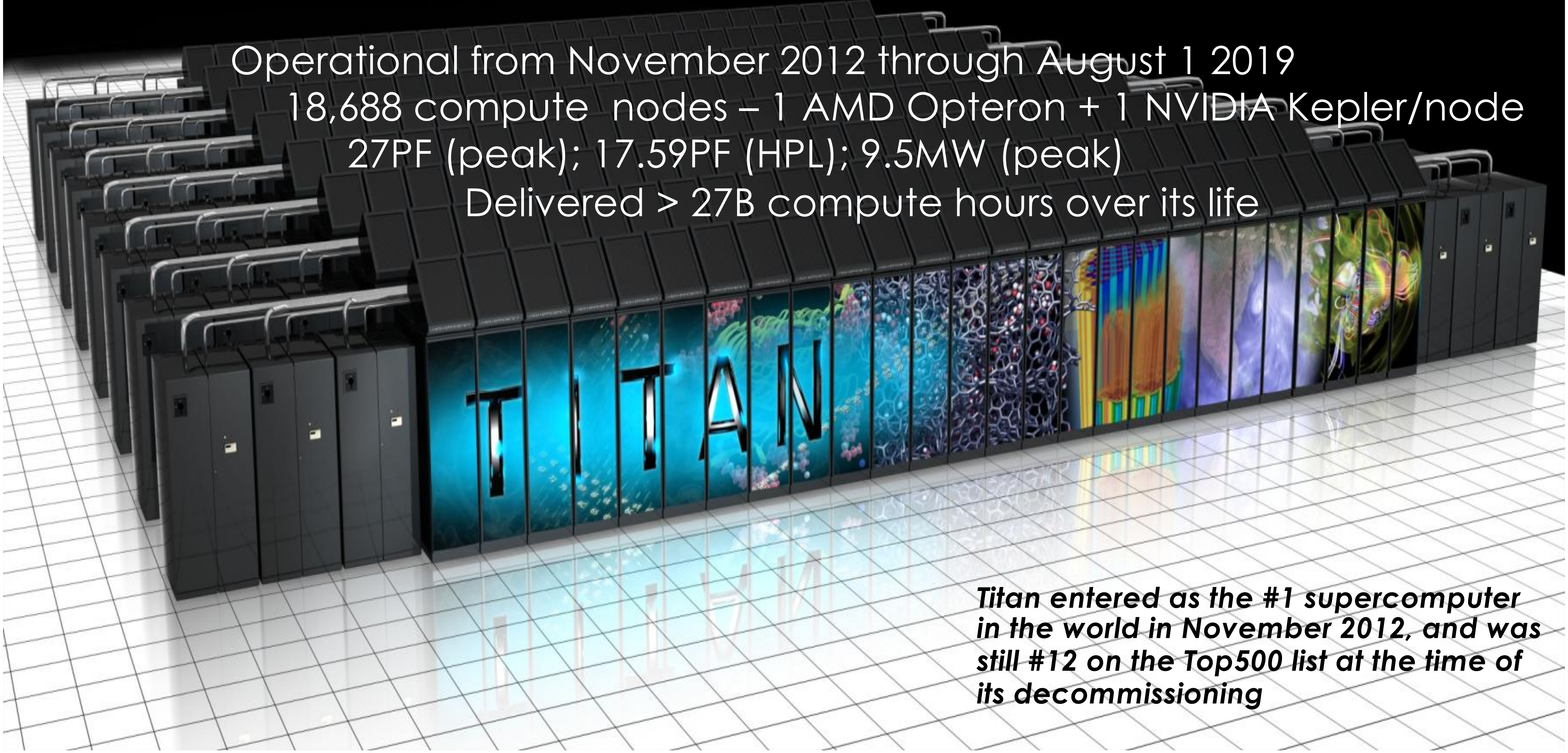
Frontier: Custom mechanical packaging is >95% room-neutral with a 30°C supply.
- *~100% Evaporative Cooling, with supplemental HVAC for parasitic loads*





Oak Ridge National Laboratory's Cray XK7 Titan

Operational from November 2012 through August 1 2019
18,688 compute nodes – 1 AMD Opteron + 1 NVIDIA Kepler/node
27PF (peak); 17.59PF (HPL); 9.5MW (peak)
Delivered > 27B compute hours over its life



Titan entered as the #1 supercomputer in the world in November 2012, and was still #12 on the Top500 list at the time of its decommissioning

Perspective on Titan as an Air-cooled Supercomputer

ORNL's Cray XK7 "Titan"

- 200 cabinets, each with a 3,000 CFM fan
    - 600,000 CFM (air volume)
- Dry air has a density of 13.076 cubic feet per pound
- Titan moves 600,000/13.076/60 -> **765 lbs/sec**

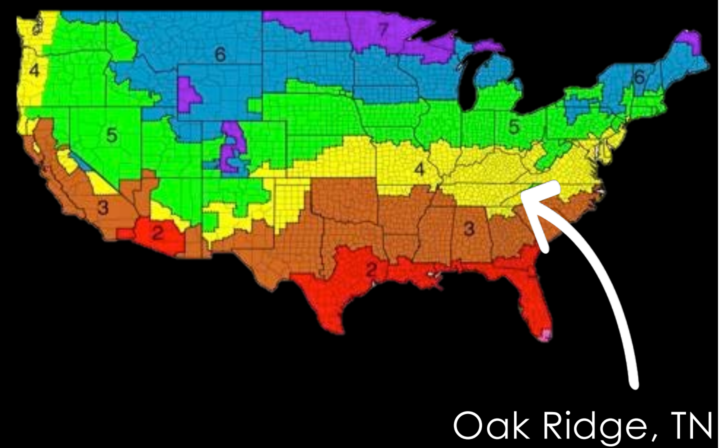
Wait. What?

Titan moves as much air as a long-range Airbus A340 at cruising altitude.

Airbus A340

- At takeoff, each of (4) engines generates 140kN of force and consumes **1000 lbs/sec of air**
- **At cruise, the A340's engines each produce ~29kN and consumes ~200 lbs/sec of air.**

Motivation for “Warmer” Cooling Solutions Serving HPC Centers

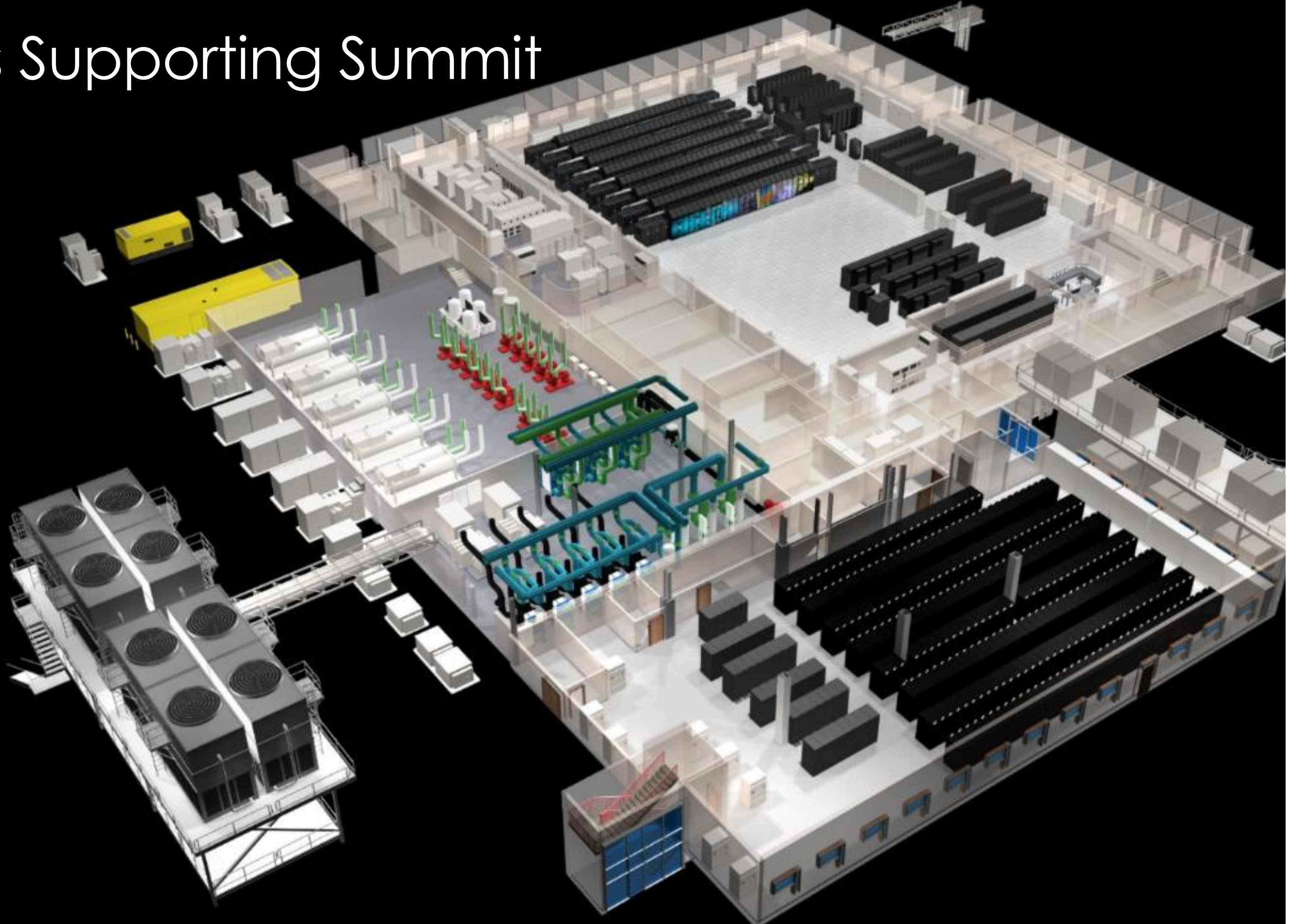


- Reduced Cost, Both CAPEX and OPEX
  - Reduce or eliminate the need for traditional chillers.
    - No chillers, no ozone-depleting refrigerants (GREEN)
  - Oak Ridge calculates an annualized PUE for air--cooled devices of no better than 1.4 (ASHRAE Zone 4A – Mixed Humid)
  - HPC power budgets continue 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

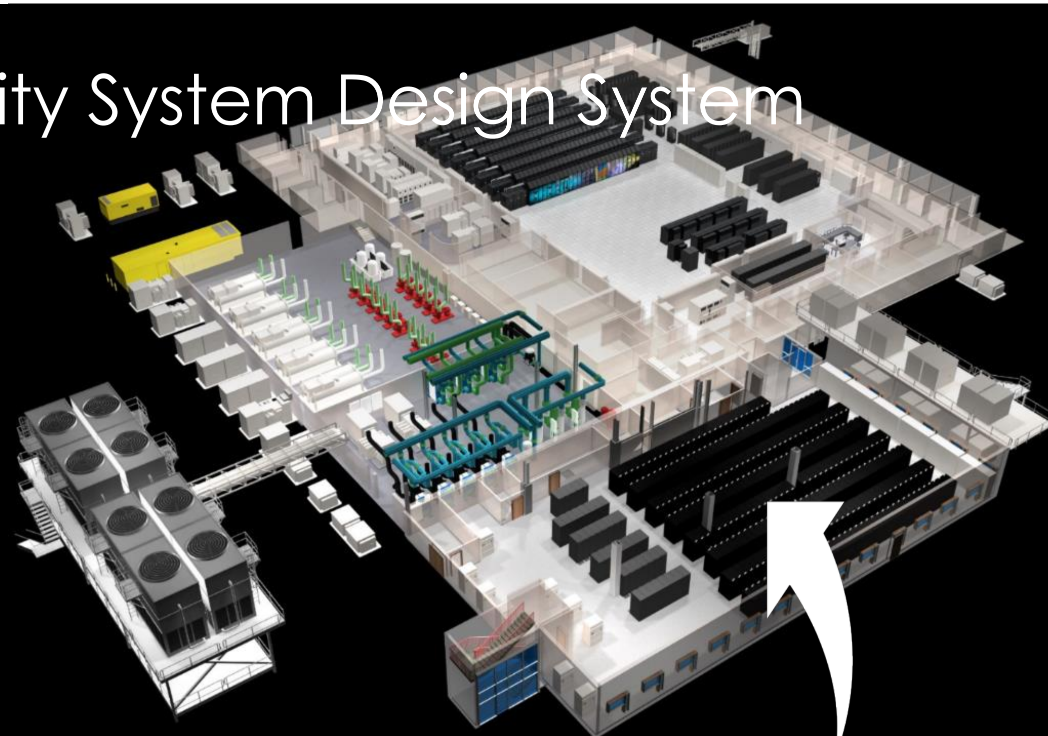
- Titan – 9MW @ heavy load
- Sitting on 250 pounds/ft2 raised floor
- Uses 42F water and special CDUs (XDPs)

- 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: 3-10MW;
- 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



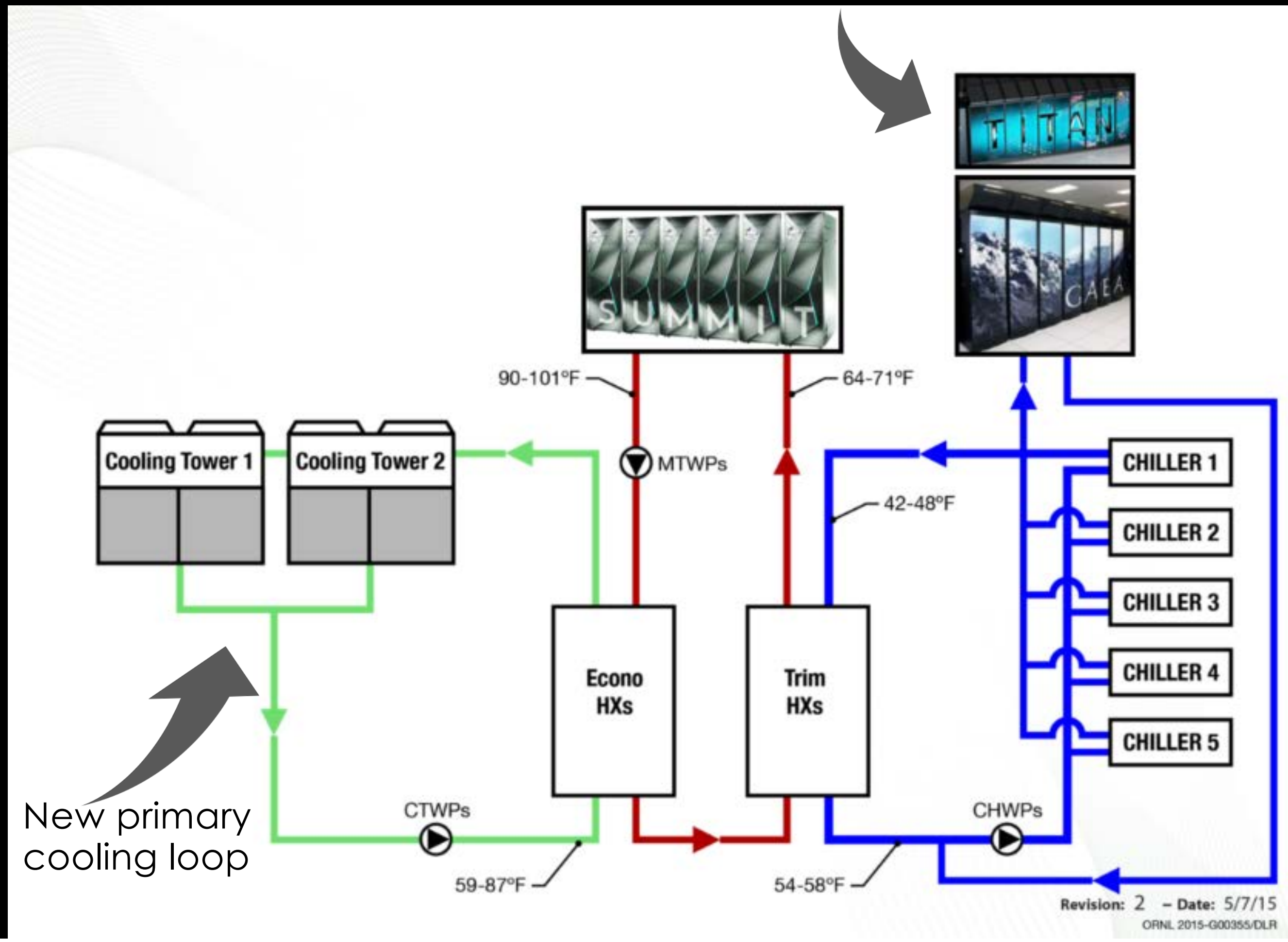
Cooling Design

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

Existing chilled water cooling loop



New primary cooling loop

Benefits of Warm Water + Operating Dashboards

- Warm Water allows annualized PUE of 1.1
  - ~$1M cost *per MW-year* for consumption on Summit;
  - ~$100k cost *per MW-year* for waste heat management
- Integration with PLC allows us to tune water flow
  - Better delta(t); less pumping energy
- Integration with IBM's OpenBMC allows us to protect these 40k components from inadequate flow across the cold plates
- Integration with the scheduler allows us to correlate power and temperature data with individual applications.
- Additional data streams to be added- most from the Facility PLC

Comparing Energy Performance - Titan to Summit

Demonstrated Performance on Titan (Oct 2012)

17.59 PF 8.9MW peak 8.3MW average

Demonstrated Performance on Summit (Oct 2018)

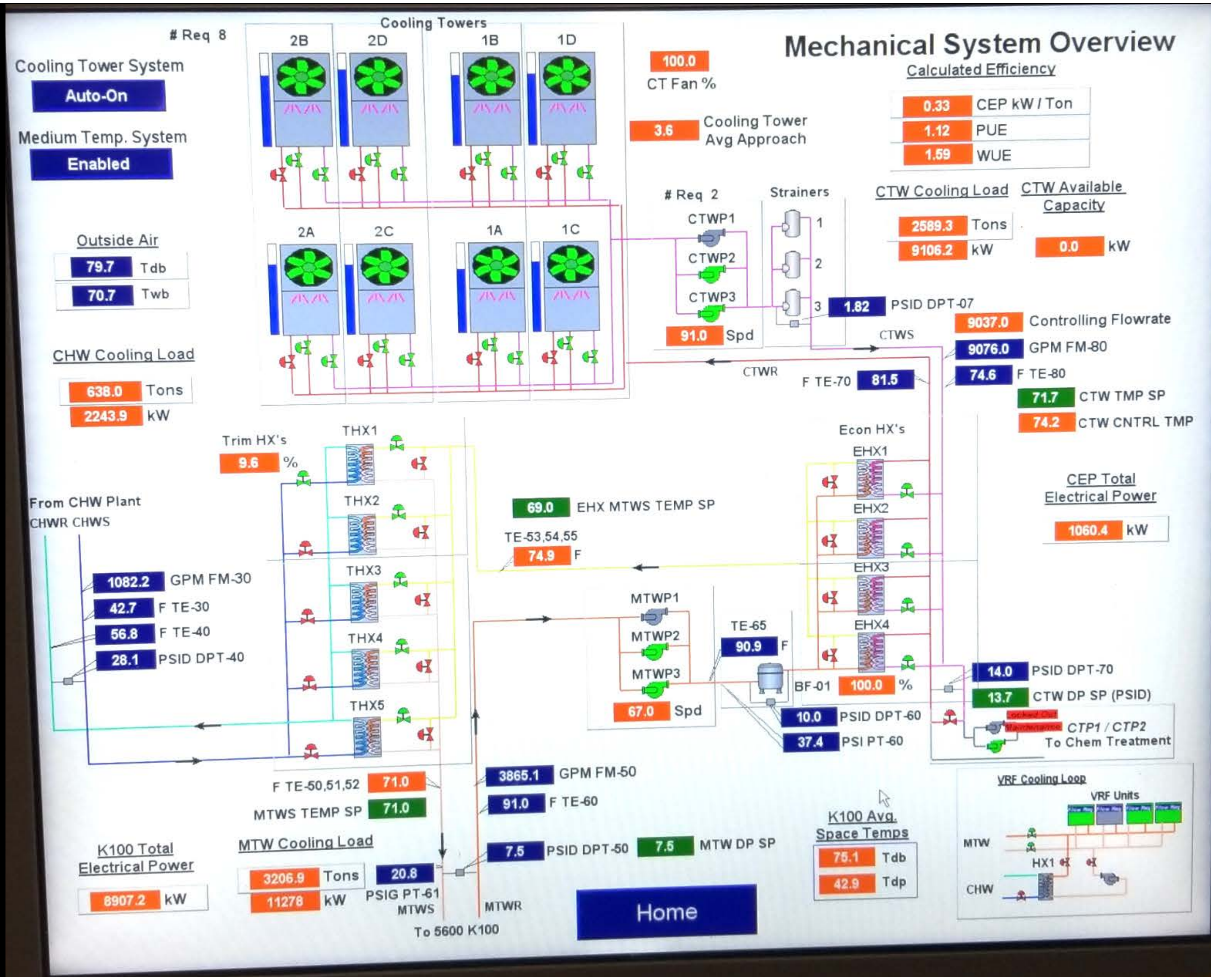
143.5 PF 11,065kW peak 9,783kW average

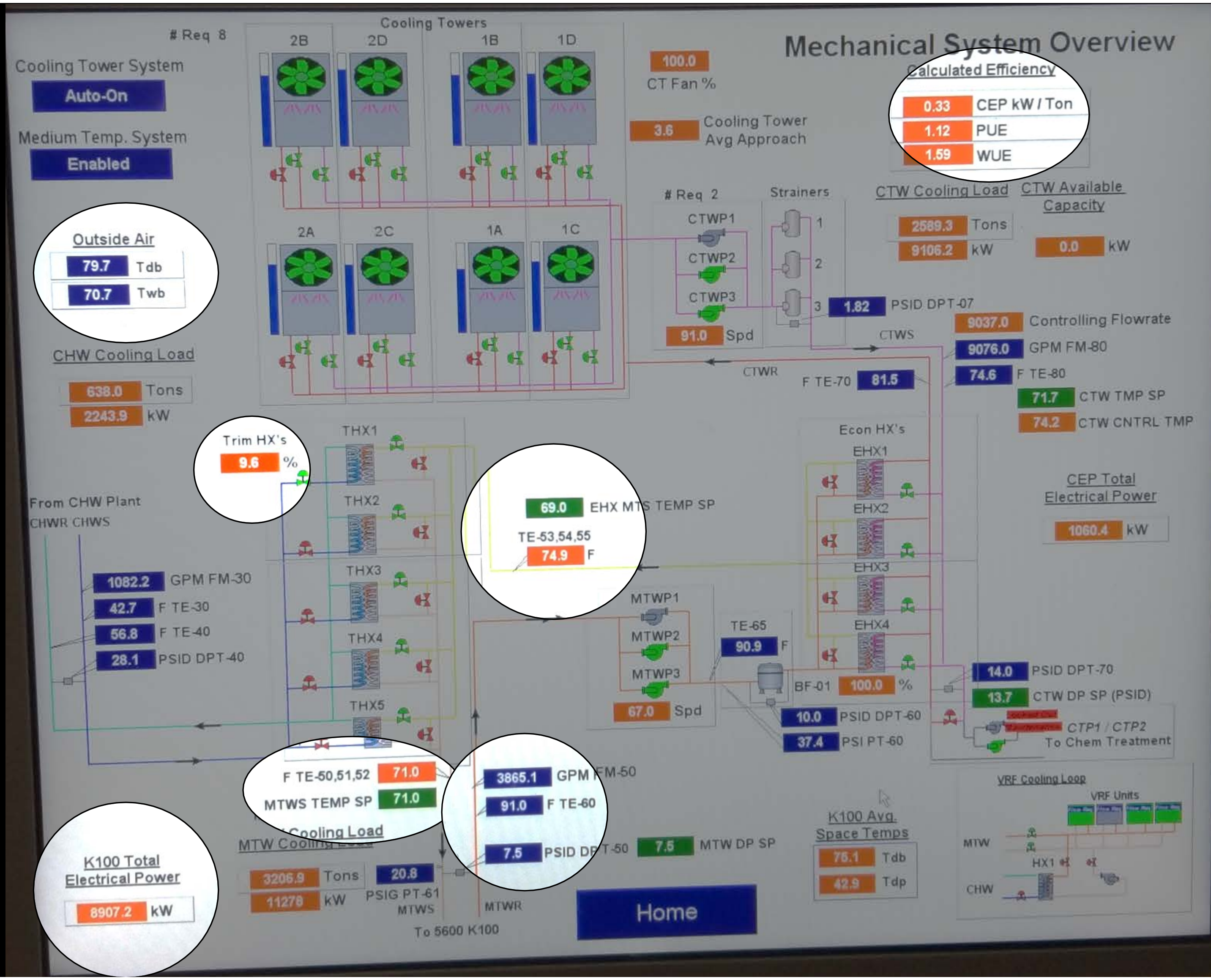
8.16x performance increase ***17.9% avg. power increase***

Titan: ~2.1 GFLOPs/Watt

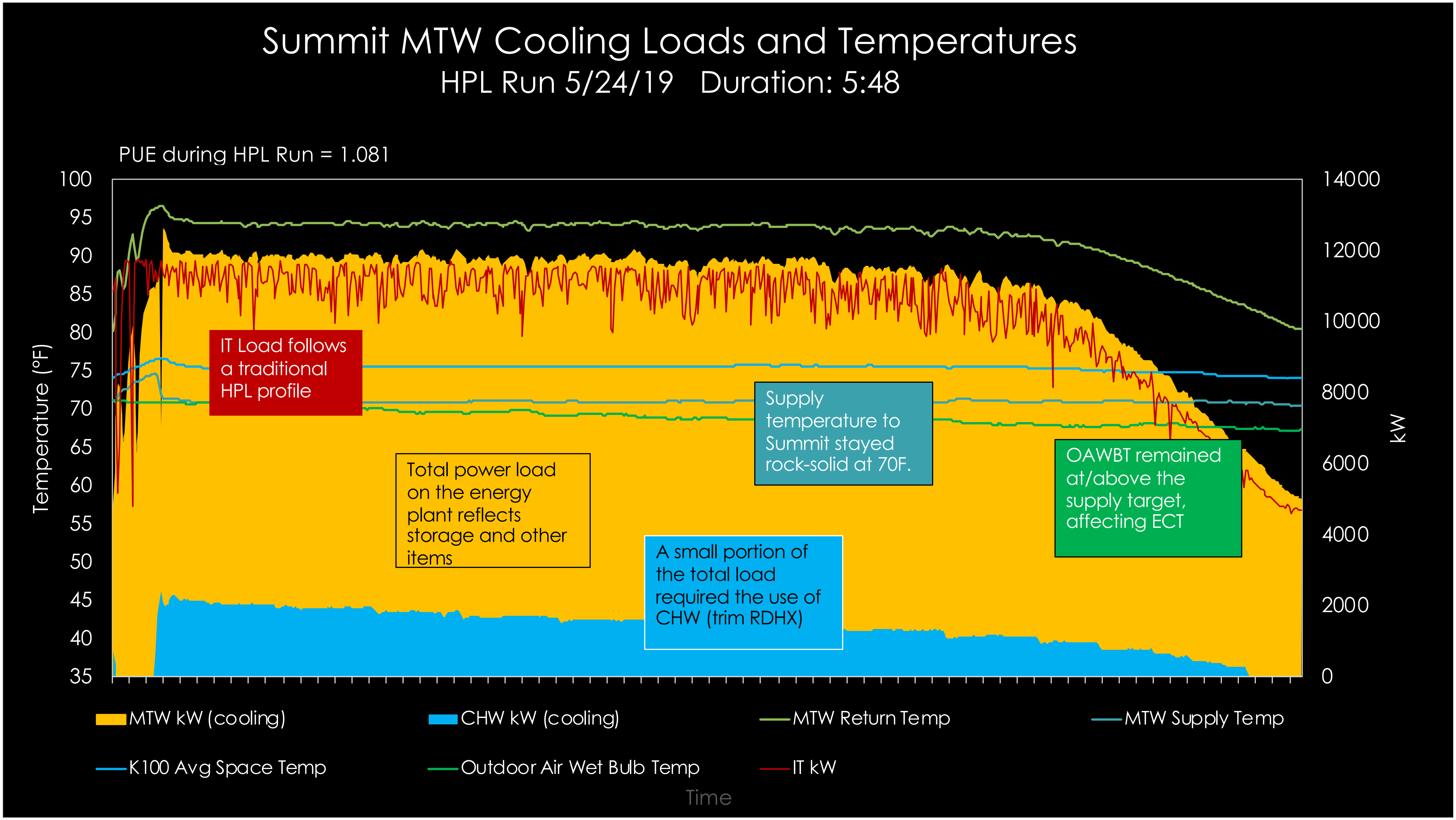
Summit: 14.66 GFLOPs/Watt

>7x increase in energy efficiency

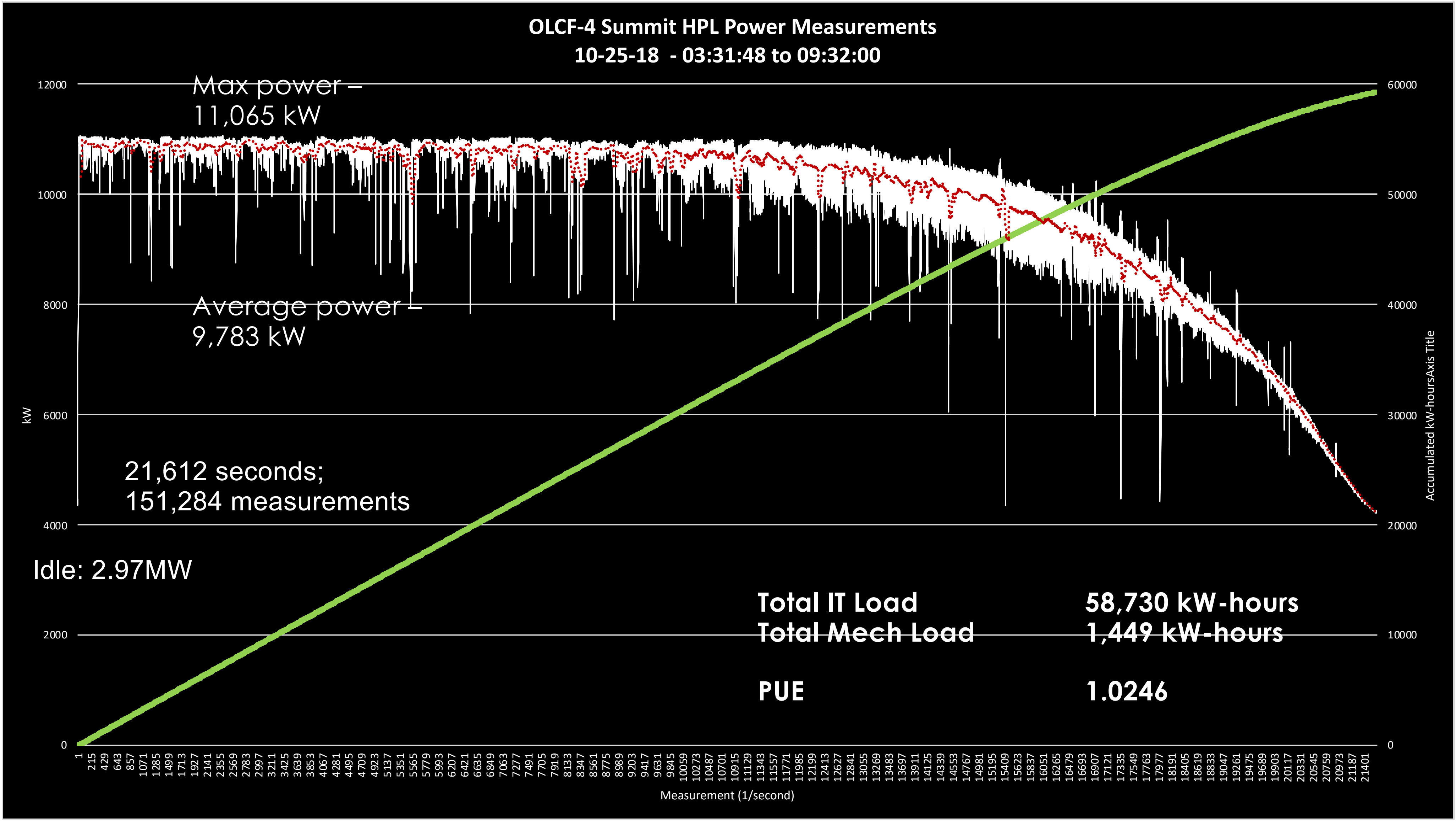

Mechanical System Overview
Calculated Efficiency
0.33 CEP kW / Ton
1.12 PUE
1.59 WUE
Req 8
Cooling Tower System
Auto-On
Medium Temp. System
Enabled
Cooling Towers
2B 2D 1B 1D
2A 2C 1A 1C
100.0
CT Fan %
3.6 Cooling Tower Avg Approach
Req 2
CTWP1
CTWP2
CTWP3
91.0 Spd
Strainers
1.82 PSID DPT-07
CTW Cooling Load
2589.3 Tons
9106.2 kW
CTW Available Capacity
0.0 kW
9037.0 Controlling Flowrate
9076.0 GPM FM-80
74.6 F TE-80
71.7 CTW TMP SP
74.2 CTW CNTRL TMP
CTWS
CTWR
F TE-70 81.5
Outside Air
79.7 Tdb
70.7 Twb
CHW Cooling Load
638.0 Tons
2243.9 kW
Trim HX's
9.6 %
THX1
THX2
THX3
THX4
THX5
From CHW Plant
CHWR CHWS
1082.2 GPM FM-30
42.7 F TE-30
56.8 F TE-40
28.1 PSID DPT-40
69.0 EHX MTWS TEMP SP
TE-53,54,55
74.9 F
Econ HX's
EHX1
EHX2
EHX3
EHX4
CEP Total Electrical Power
1060.4 kW
MTWP1
MTWP2
MTWP3
67.0 Spd
TE-65
90.9 F
BF-01 100.0 %
10.0 PSID DPT-60
37.4 PSI PT-60
14.0 PSID DPT-70
13.7 CTW DP SP (PSID)
CTP1 / CTP2
To Chem Treatment
F TE-50,51,52 71.0
MTWS TEMP SP 71.0
3865.1 GPM FM-50
91.0 F TE-60
7.5 PSID DPT-50
7.5 MTW DP SP
K100 Total Electrical Power
8907.2 kW
MTW Cooling Load
3206.9 Tons
11278 kW
20.8
PSIG PT-61
MTWS
MTWR
To 5600 K100
Home
K100 Avg. Space Temps
75.1 Tdb
42.9 Tdp
VRF Cooling Loop
VRF Units
MTW
HX1
CHW


Mechanical System Overview
Calculated Efficiency
0.33 CEP kW / Ton
1.12 PUE
1.59 WUE
Cooling Towers
Cooling Tower System
Auto-On
Medium Temp. System
Enabled
100.0 CT Fan %
3.6 Cooling Tower Avg Approach
Outside Air
79.7 Tdb
70.7 Twb
CHW Cooling Load
638.0 Tons
2243.9 kW
Trim HX's
9.6 %
CTW Cooling Load
2589.3 Tons
9106.2 kW
CTW Available Capacity
0.0 kW
1.82 PSID DPT-07
9037.0 Controlling Flowrate
9076.0 GPM FM-80
74.6 F TE-80
71.7 CTW TMP SP
74.2 CTW CNTRL TMP
F TE-70 81.5
69.0 EHX MTS TEMP SP
TE-53,54,55 74.9 F
CEP Total Electrical Power
1060.4 kW
From CHW Plant
1082.2 GPM FM-30
42.7 F TE-30
56.8 F TE-40
28.1 PSID DPT-40
TE-65 90.9 F
67.0 Spd
91.0 Spd
BF-01 100.0 %
14.0 PSID DPT-70
13.7 CTW DP SP (PSID)
10.0 PSID DPT-60
37.4 PSI PT-60
CTP1 / CTP2 To Chem Treatment
F TE-50,51,52 71.0
MTWS TEMP SP 71.0
3865.1 GPM FM-50
91.0 F TE-60
7.5 PSID DPT-50
7.5 MTW DP SP
MTW Cooling Load
3206.9 Tons
11278 kW
20.8 PSIG PT-61
K100 Total Electrical Power
8907.2 kW
K100 Avg. Space Temps
76.1 Tdb
42.9 Tdp
VRF Cooling Loop
Home
To 5600 K100

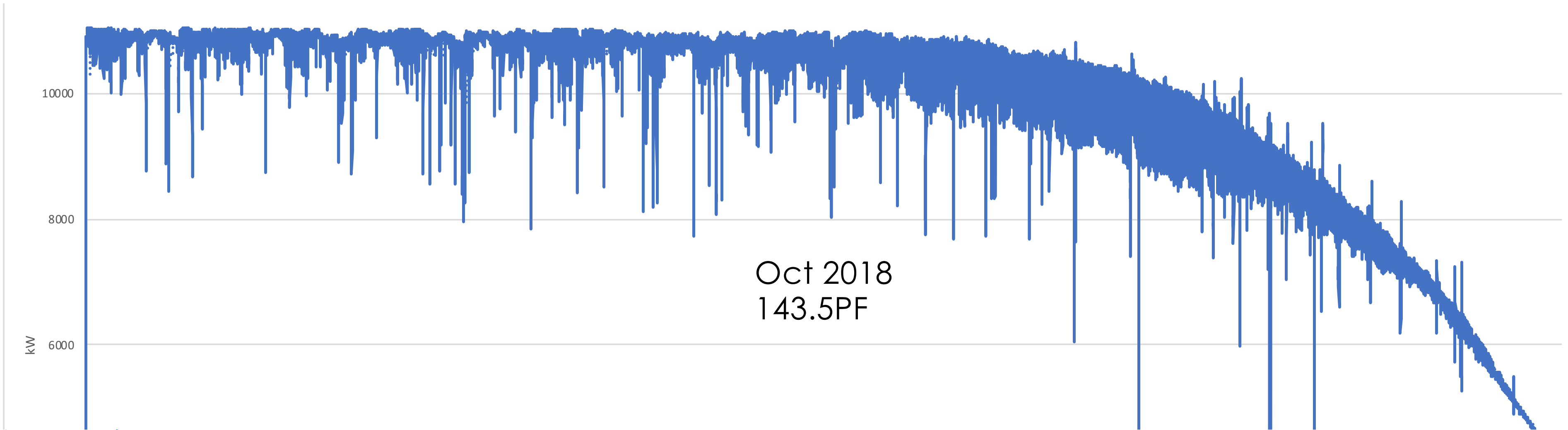
Live Summit Cooling Metrics
Live Summit OpenBMC data stream scraped every 10 sec - All 4606 nodes
Active OpenBMC Datastreams
4606/4606
Active OpenBMC Datastreams
Active Data Streams Current: 4.606 K
Realtime Summit Power
3.36 MW
Supply Temperature
69.7 °F
Flow
3381 gpm
K100 PUE
1.057
Wetbulb (+0h)
15.1 °C
Wetbulb (+12h)
17.3 °C
Summit Total Power
Summit Total Power
GPU Power Distribution (4606 x 6 : 5min)
CPU Power Distribution (4606 x 2 : 5min)
Summit Power Breakdown
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

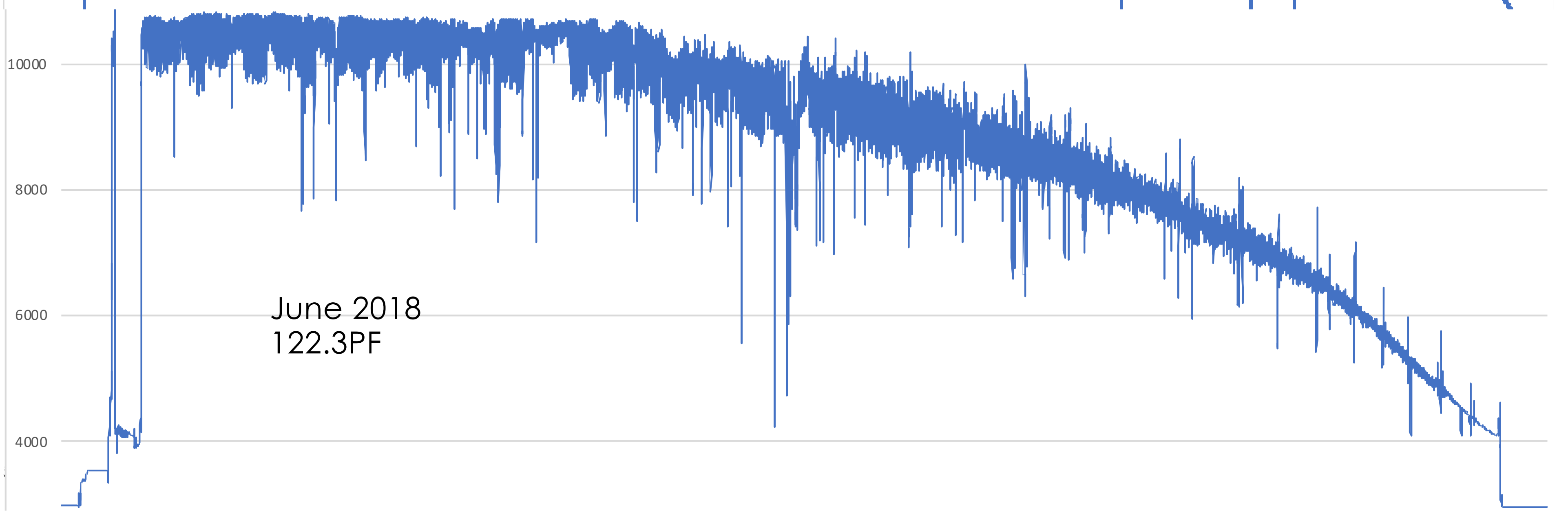


Summit MTW Cooling Loads and Temperatures
HPL Run 5/24/19 Duration: 5:48
PUE during HPL Run = 1.081
Temperature (°F)
100
95
90
85
80
75
70
65
60
55
50
45
40
35
14000
12000
10000
8000
6000
4000
2000
0
kW
IT Load follows a traditional HPL profile
Total power load on the energy plant reflects storage and 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

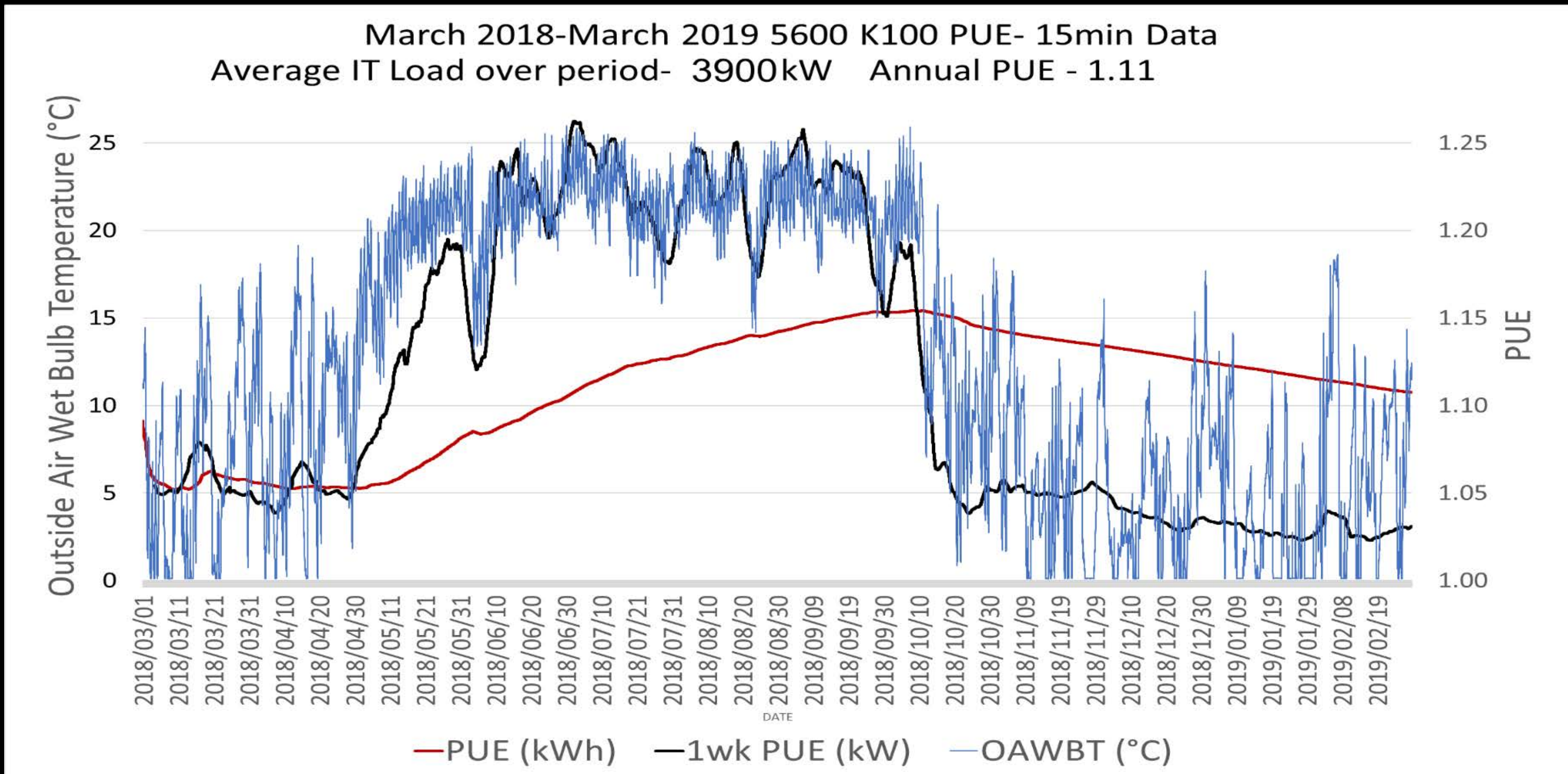


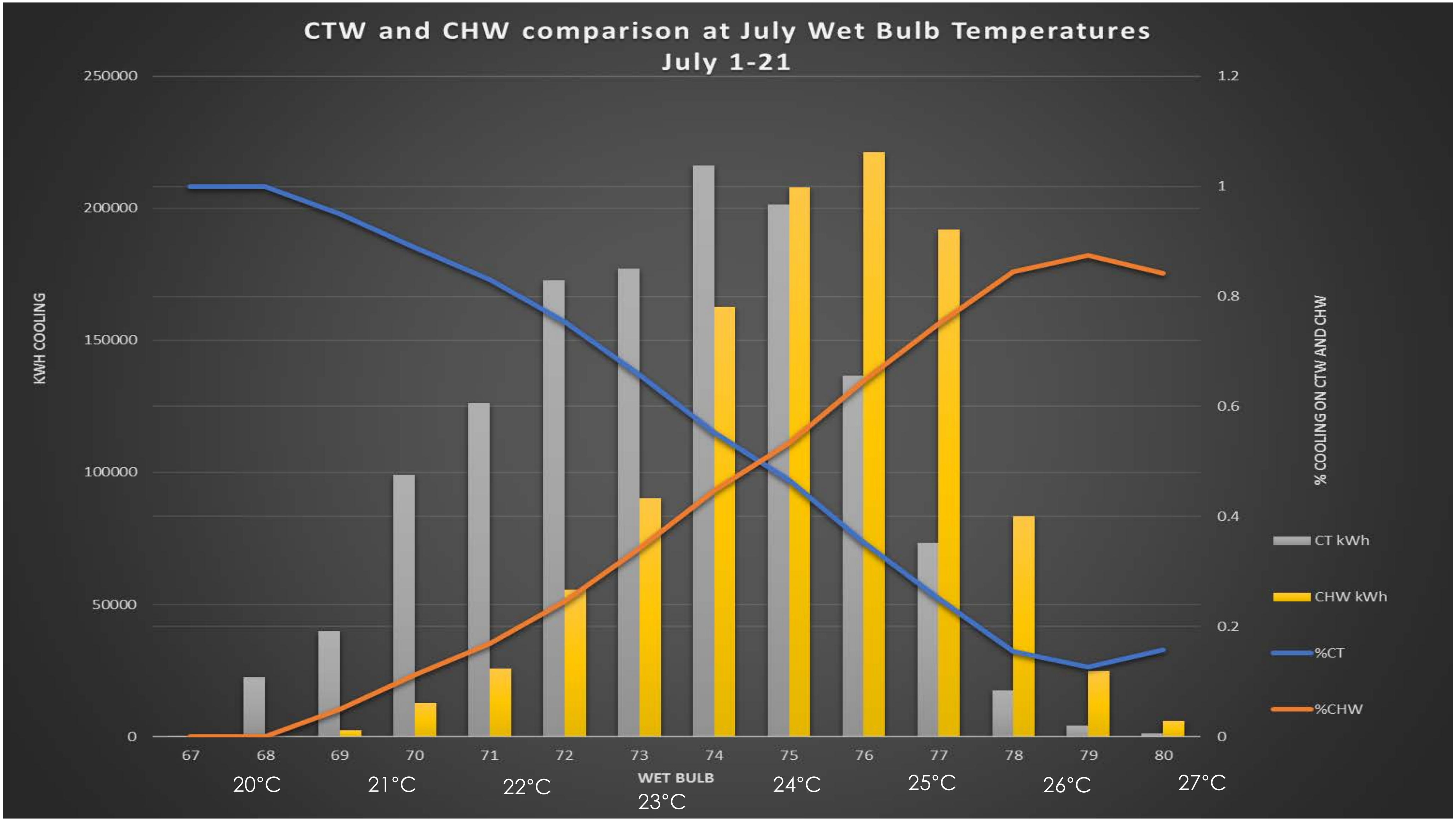
OLCF-4 Summit HPL Power Measurements
10-25-18 - 03:31:48 to 09:32:00
Max power –
11,065 kW
Average power –
9,783 kW
21,612 seconds;
151,284 measurements
Idle: 2.97MW
Total IT Load
58,730 kW-hours
Total Mech Load
1,449 kW-hours
PUE
1.0246
kW
12000
10000
8000
6000
4000
2000
0
Accumulated kW-hoursAxis Title
60000
50000
40000
30000
20000
10000
0
Measurement (1/second)

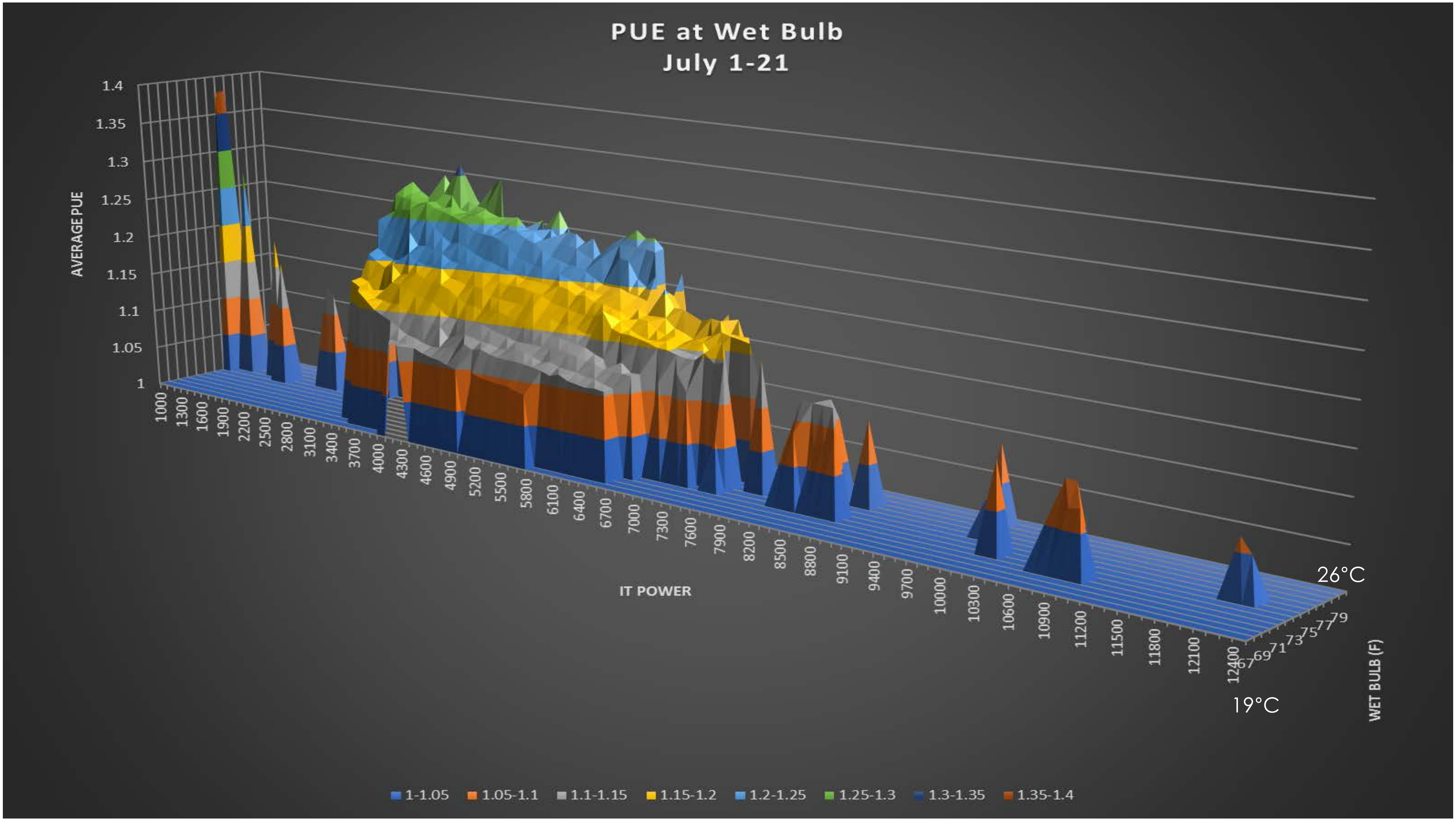

10000
8000
6000
kW
Oct 2018
143.5PF


10000
8000
6000
4000
June 2018
122.3PF

Cooling System Performance - PUE




CTW and CHW comparison at July Wet Bulb Temperatures
July 1-21
KWH COOLING
250000
200000
150000
100000
50000
0
1.2
1
0.8
0.6
0.4
0.2
0
% COOLING ON CTW AND CHW
CT kWh
CHW kWh
%CT
%CHW
67
68
69
70
71
72
73
74
75
76
77
78
79
80
WET BULB
20°C
21°C
22°C
23°C
24°C
25°C
26°C
27°C


PUE at Wet Bulb
July 1-21
AVERAGE PUE
1.4
1.35
1.3
1.25
1.2
1.15
1.1
1.05
1
IT POWER
1000
1300
1600
1900
2200
2500
2800
3100
3400
3700
4000
4300
4600
4900
5200
5500
5800
6100
6400
6700
7000
7300
7600
7900
8200
8500
8800
9100
9400
9700
10000
10300
10600
10900
11200
11500
11800
12100
12400
WET BULB (F)
67
69
71
73
75
77
79
26°C
19°C
1-1.05
1.05-1.1
1.1-1.15
1.15-1.2
1.2-1.25
1.25-1.3
1.3-1.35
1.35-1.4

Caution – Secondary (Closed) Loop Concerns Worsen as Temperatures Rise…

Allowable Wetted Materials
- Lead-free copper alloys with less than 15% zinc.
- Stainless steels
- EPDM
- Polypropylene

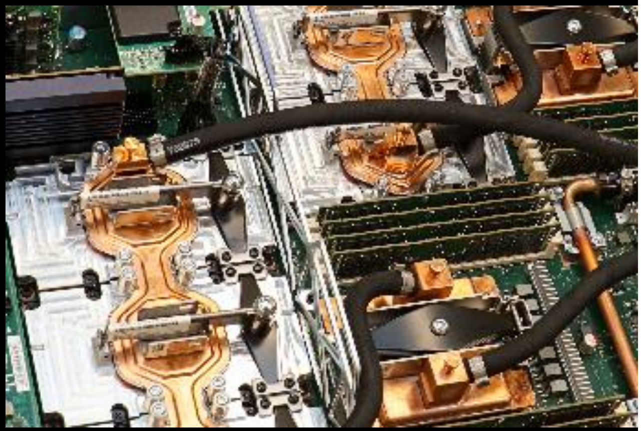
IBM's Water Quality Requirements
- All metals less than or equal to 0.10 ppm
- Calcium less than or equal to 1.0 ppm
- Magnesium less than or equal to 1.0 ppm
- Manganese less than or equal to 0.10 ppm
- Phosphorus less than or equal to 0.50 ppm
- Silica less than or equal to 1.0 ppm
- Sodium less than or equal to 0.10 ppm
- Bromide less than or equal to 0.10 ppm
- Nitrite less than or equal to 0.50 ppm
- Chloride less than or equal to 0.50 ppm
- Nitrate less than or equal to 0.50 ppm
- Sulfate less than or equal to 0.50 ppm
- Conductivity less than or equal to 10.0 μS/cm.
- pH 6.5 – 8.0
- Turbidity (NTU) less than or equal to 1

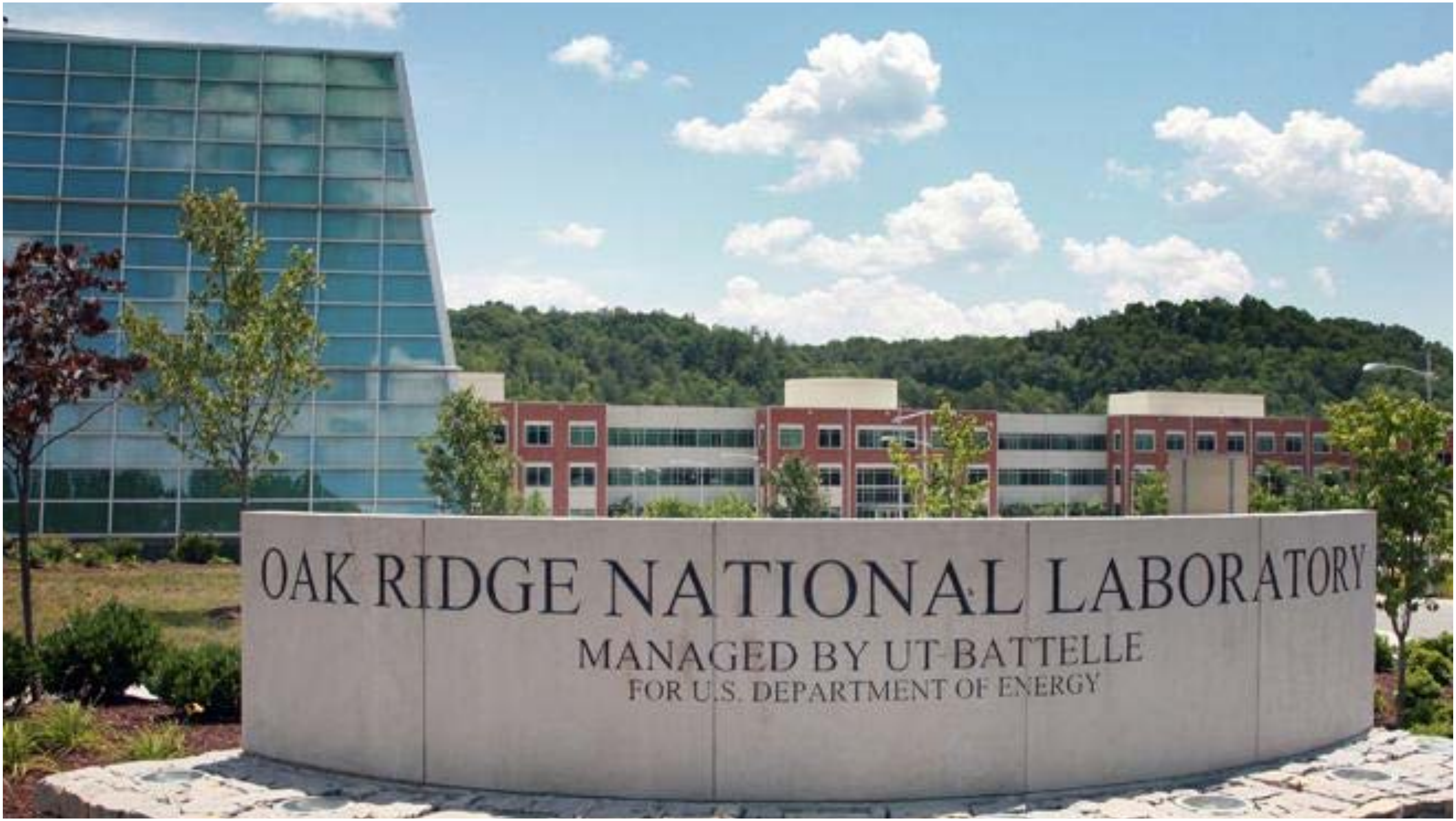
EEHPCWG is looking for common ground among the HPC suppliers for water quality

Accelerated Corrosion
System Blockages
Reduced System Efficiency

Water Chemistry - Biocides
- The choice of biocide depends on whether you are chasing anaerobic bacteria, aerobic bacteria, fungi, and/or algae

Fungi might not be detected in the water, even though it can grow and cause blockage of cooling channels in cold plates




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