

Below is a report summarizing the Energy Efficient HPC Working Group SC18 presence.

## **REPORT: EE HPC WG and SC18**

The EE HPC WG has ~800 members with ~50% from supercomputing centers, 30% from the vendor community and 20% from academe. The membership is mostly from the United States, but Europe is also well represented. Asia – and especially Japan – is a growing part of the membership. The EE HPC WG provides a unique forum for bringing together people with expertise in supercomputing centers facilities and those with expertise in HPC systems.

The EE HPC WG has many active teams that work on promoting energy efficient HPC with whitepapers, research papers, methodologies, case studies and best practices. Some of the topical areas include liquid cooling, operational data analytics, metrics, power/energy measurement, procurement, electric grid integration and power/energy management and control, including job scheduling and resource management. The overall purpose of the SC18 activities is to solicit feedback and participation from the broader community as well as serve as a face-to-face collaborative forum. This report shall document the results of these activities.

SCxx has had a supercomputing data center focus since at least SC09, when a "Data Center of the Future" Competition was part of that year's Sustainability Thrust. The Energy Efficient HPC Working Group has taken on the mantle of continuing to promote participation at SCxx by people with an operational and/or facilities focus. The annual EE HPC WG workshop is a cornerstone for that outreach. It is buttressed by EE HPC WG organized Birds of Feather (BoF), Panels, Research Papers and Exhibitor Booth. There is not a comparable organized effort for this target audience at SCxx.

The EE HPC WG presence at SC18 included a workshop, 6 BoF sessions, a panel, a workshop research paper and an exhibitor booth. Participation was strong in all of these sessions and the booth. Thanks to all who participated. For those who couldn't make it, presentations are posted on the EE HPC WG website.

All of the presentations and some supporting material can be found on the EE HPC WG website ( [https://eehpcwg.llnl.gov/conf\\_sc18.html](https://eehpcwg.llnl.gov/conf_sc18.html) ) under Conferences and then SC18 Conference.

InsideHPC published a video recording of the [SC18 Panel](#) we organized on "*If you can't measure it, you can't improve it*", *Software Improvements from Power/Energy Measurement Capabilities*

## Table of Contents

|  |          |
|--|----------|
| <b>WORKSHOP</b> .....  | <b>2</b> |
| KEYNOTE: GREEN COMPUTING INVOLVES A LOT OF INNOVATION, BUT ALSO A LOT OF SWEAT.....          | 3        |
| STATE OF THE WG TEAMS.....   | 4        |
| POWER GRID .....   | 5        |
| THERMOSYPHON- IMPROVED WUE .....   | 6        |
| MACHINE INSTALLATIONS.....   | 6        |
| DWAVE INSTALLATION.....  | 7        |
| CHALLENGES IN HOLISTIC MONITORING AND DATA INTEGRATION.....                                  | 7        |
| PANEL DISCUSSION – SOFTWARE FOR ENERGY EFFICIENCY .....                                      | 8        |
| <b>EE HPC WG BOOTH</b> .....   | <b>8</b> |
| <b>BIRDS OF FEATHER (BoF)</b> .....  | <b>9</b> |
| ENERGY EFFICIENCY CONSIDERATIONS FOR HPC PROCUREMENTS .....                                  | 9        |
| THE FACILITY PERSPECTIVE ON LIQUID COOLING: EXPERIENCES AND PROPOSED OPEN SPECIFICATION..... | 10       |
| POWERAPI AND REDFISH: STANDARDIZING POWER MEASUREMENT AND CONTROL FOR HPC.....               | 10       |
| A LOOK AHEAD: ENERGY AND POWER AWARE JOB SCHEDULING AND RESOURCE MANAGEMENT (EPA JSRM) ..... | 10       |
| THE GREEN 500: TRENDS IN ENERGY EFFICIENT SUPERCOMPUTING.....                                | 11       |
| DATA ANALYTICS FOR SYSTEM AND FACILITY ENERGY MANAGEMENT .....                               | 12       |

## WORKSHOP

The 9<sup>th</sup> Annual EE HPC WG Workshop was held on Monday, November 12<sup>th</sup> from 9:00 to 5:30. Attendance was comparable to recent prior years. There were ~100 participants in the room at any point in time with ~150 different individuals attending at some point during the day. There were mostly all-day participants, but some participants were attending multiple events and targeted specific sessions.

Everyone who completed a survey thought the workshop was valuable (10 total respondents) with 80% rating it ‘very valuable’, 20% ‘somewhat valuable’ and nobody who found it ‘only a little valuable’. The survey data is summarized in rank order for each session in the table below.

| Session               | Very valuable | Somewhat valuable | Only a little valuable | Other |
|-----------------------|---------------|-------------------|------------------------|-------|
| Keynote               | 90%           | 10%               | 0%                     | 0%    |
| Thermosyphon          | 70%           | 10%               | 20%                    | 0%    |
| Power Grid            | 60%           | 40%               | 0%                     | 0%    |
| Machine Installations | 60%           | 40%               | 0%                     | 0%    |

|                            |     |     |     |     |
|----------------------------|-----|-----|-----|-----|
| State of the WG            | 60% | 30% | 10% | 0%  |
| Operational Data Analytics | 60% | 10% | 20% | 10% |
| Software Stack             | 50% | 50% | 0%  | 0%  |
| Quantum                    | 50% | 30% | 10% | 10% |

#### WORKSHOP AGENDA:

- Opening Remarks, Introductions, Agenda
- Keynote Satoshi Matsuoka, RIKEN
  - “Green Computing Involves a lot of Innovation, but also a lot of Sweat”
- State of the WG Team Reports
- The Power Grid
  - HPC Power Increasingly Challenges the Power Grid: Josip Loncaric, (LANL)
  - Forecasting Extreme Site Power Fluctuations Using Fast Fourier Transformation: Ghaleb Abdulla, (LLNL)
  - Electrical Utility Interface Protections and Challenges: Gary New (NCAR)
- Thermosyphon- Improved WUE: David Sickinger (NREL), David Smith (SNL)
- Machine Installations
  - Astra Computer New HPC Data Center: David J. Martinez (SNL)
  - SuperMUC-NG: Dr. Herbert Huber (LRZ)
  - Cooling System Overview: Summit Supercomputer: David Grant (ORNL)
- D-WAVE Installation: Josip Loncaric; (LANL)
- Challenges in Holistic Monitoring and Data Integration
  - Challenges in Holistic Monitoring and Data Integration: Andrea Bartolini (University of Bologna)
  - Monitoring and Data Integration at LLNL: Kathleen Shoga (LLNL)
  - What’s new at NERSC? High Resolution Data Monitoring: Challenges in Providing EE; Norm Bourassa (NERSC)
- Panel Discussion – Software for Energy Efficiency (PowerStack, PowerAPI, READEX, GEOPM)
- Closing Remarks

#### **KEYNOTE: GREEN COMPUTING INVOLVES A LOT OF INNOVATION, BUT ALSO A LOT OF SWEAT**

Satoshi Matsuoka was the keynote speaker. He is the Director of RIKEN Center for Computational Science, a Professor at the Tokyo Institute of Technology and a Fellow at the AI Research Center which is part of the National Institute of Science and Technology (AIST), the largest national lab in Japan. Satoshi has been a leader in green computing for almost two decades, since power and energy consumption first emerged as an issue for HPC. His presentation described energy efficiency strategies, technologies and approaches at Tokyo Tech, AIST and RIKEN. Satoshi’s presentation was dense and filled with a lot of very interesting content. Below are major takeaways:

- HPC best achieves major power savings from new devices, new architectures and novel cooling techniques. Furthermore, unlike Enterprise and Business Cloud Computing, HPC cannot benefit at all from server consolidation and not much from DVFS.
- The Tsubame series of supercomputers at Tokyo Tech showed x100,000 speed-up in 17 years while demonstrating energy efficiency and power savings from massively parallel many core technology, efficient power control and novel cooling techniques. Tsubame supercomputers took Green500 List awards in 2010, 2013 and 2017.
- AI Bridging Cloud Infrastructure (ABCI) is a supercomputer at AIST that was in the Top10 of both the Top500 and Green500 Lists in 2018. It is housed in a very efficient HPC data center that uses free cooling all year and cooling pods.
- RIKEN has been operating the K-computer, which draws 12k-15MW and uses hybrid cooling with a water/air ratio of 70:30. Improvements in power management and facility energy efficiency have resulted in significant operational cost reduction and PUE improvements over the past 6 years.
- RIKEN has announced that it will replace the K-computer with the post K-computer in 2020. The peak power will more than double as will voltage fluctuations, but idle power remains the same and energy efficiency is projected to be world-class. The system will also have fine-grained power management and control capabilities.

## STATE OF THE WG TEAMS

The next session was an update on the State of the EE HPC WG Team. Each report included information about the impacts of the work, deliverables, current activities, next steps and help needed. Below were the presenters, their affiliation and the names of the teams for which they presented a report.

|                  |   |   |
|------------------|---|---|
| Josip Loncaric   | Los Alamos National Laboratory            | Electric Grid Integration                                 |
| Ghaleb Abdulla   | Lawrence Livermore National Laboratory    | Operational Data Analytics                                |
| Sid Jana         | Intel Corporation                         | Energy & Power Aware Job Scheduling and Resource Mgmt     |
| Wu Feng          | Virginia Tech and Green500                | HPC System Power Measurement Methodology                  |
| Gert Svensson    | KTH Swedish Royal Institute of Technology | Energy Efficient Procurement Considerations               |
| Dale Sartor      | Lawrence Berkeley National Laboratory     | Liquid Cooled Server Rack Specification                   |
| David Grant      | Oak Ridge National Laboratory             | Liquid Cooling Controls                                   |
| Anna Maria Baily | Lawrence Livermore National Laboratory    | Reliability, Availability, Serviceability & Manageability |
| Ryan Grant       | Sandia National                           | Power API   |

|               |            |  |
|---------------|------------|--|
|               | Laboratory |  |
| Natalie Bates | EE HPC WG  | Other Teams: Dashboards, Liquid Cooling Commissioning, TUE |

## POWER GRID

There were three presentations on the power grid; the first one focused on HPC system power fluctuations, the second one on modeling and predicting site-wide power fluctuations and the last one on electrical utility interface protections and challenges.

**HPC Power Increasingly Challenges the Power Grid**, presented by Josip Loncaric (LANL). A 20-40 MW computer could impose fast 15-30 MW power transients. What is the impact that this could have on the electrical grid? Research at LANL has shown that transient voltage disturbances from the HPC system do propagate on their distribution system with up to 5% sag at full load. These voltage disturbances are occurring because of a confluence of three major trends: 1. CMOS power demands are growing, 2. power management techniques are reducing idle power demands (i.e., wider gaps between peak and idle demands), and 3. bulk synchronous applications have uneven and peaky loads. Josip urges the community to understand the impact that their sites could have on the distribution grid.

**Forecasting Extreme Site Power Fluctuations Using Fast Fourier Transformation**, presented by Ghaleb Abdulla (LLNL). LLNL's site-wide power shows periodic and large amplitude power spikes and WAPA (LLNL's energy provider) asked for notification when the site's total energy consumption goes up/down by 750 KW over a 15 minutes window. LLNL then generated email alerts using real-time data that met this request. This prompted WAPA to ask for predictive information, since they schedule LLNL's load on a 2-6 day ahead basis. Using the same data used for the email alerts and analyzing three years of alerts, LLNL determined that over 50% of large power spikes occur between two hours in the morning and one hour in the evening. These are attributable to employee schedules. The remaining spikes were not periodic and can be attributed to both scheduled and random events.

**Electrical Utility Interface Protections and Challenges**, presented by Gary New (NCAR). What level of service is the electrical utility mandated to provide? Some things to consider are aging infrastructure, weather, evolving technology and variable loads. The site can employ technologies to mitigate risk of power quality issues; these include uninterruptible power supplies and surge protection devices. Transformer and distribution gear must also be considered. Monitoring is also critical, both information provided by the utility and to the utility.

## THERMOSYPHON- IMPROVED WUE

David Sickinger (NREL) and David Smith (Sandia NL) made a presentation on experiences with thermosyphon technology at NREL and Sandia NL. In both cases, a thermosyphon cooler is deployed along with a cooling tower in a hybrid cooling system. The thermosyphon cooler saves water relative to the cooling tower under the right conditions. The thermosyphon is best used when the temperature is colder. For example, Sandia will use the thermosyphon all the time from October through April and only during the nights from May through September.

## MACHINE INSTALLATIONS

There were three presentations on recent HPC system installations. All three of the presentations described the cooling features of the HPC systems installed. Some mention was also made about peculiarities of the facility cooling system.

**Astra Computer New HPC Data Center**, presented by David J. Martinez (Sandia NL). The Astra Computer uses HPE's Modular Cooling System (MCS). Astra's thirty-six compute racks (1.2MW load) are cooled by 12 fan coils. Heat leaving the compute racks is minimal; 99% of the heat created by the compute racks stays in the cabinet. The system doesn't require the internal plumbing of liquid disconnects and cold plates running across the CPUs and memory. With the new HPC data center, these coils are cooled without compressors year round and no evaporative water at all almost 6000 hours per year.

**SuperMUC-NG**, presented by Dr. Herbert Huber (LRZ). LRZ has been a pioneer of direct warm water cooling in order to maximize the use of free-cooling through-out the year. SuperMUC-NG uses Lenovo's direct warm water cooling that allows for 50 degrees C inlet water temperature and 60 degrees C outlet water temperature at the node level. Conventional chillers are still needed to cool the system waste heat of air-cooled components such as network switches, power supplies and storage. They are also needed to cool the heat flux of the system into the room air (temperature of cooling devices in the nodes is more than 40°C which means high surface temperature of nodes and compute node racks). LRZ uses adsorption chillers to generate chilled water cooling capacity with heat re-use from the HPC system.

**Cooling System Overview- Summit Supercomputer**, presented by David Grant (ORNL). ORNL's IBM Summit supercomputer uses direct liquid cooling and rear-door heat exchangers with over 100,000 liquid connections. Summit shares a chilled water loop with two other ORNL supercomputers, but is able to use cooling towers only for most of the load. During the months of March through October 2018, ORNL's ratio of cooling tower to chiller cooling was 80/20.

## DWAVE INSTALLATION

Prepared by Loren Serna and Ron Velarde and presented by Josip Loncaric (LANL) was a presentation about the facility requirements for Ising, LANL's D-Wave quantum annealing computer. There are very tight facility requirements for low vibration & other disturbances, highly reliable power and cryogenic cooling. It is a physics experiment that operates at 10 milli-kelvin, which is 300 times cooler than interstellar space.

Furthermore, the cryogenic cooling system must be ultra-reliable because any failure costs months of recovery. Beyond the extensive cryogenic cooling system, facility requirements included a vibration isolated concrete pad, additional PDU and UPS systems, a generator with failover controls and liquid nitrogen dewars.

Prepared by Loren Serna and Ron Velarde and presented by Josip Loncaric (LANL) was a presentation about the facility requirements for Ising, LANL's D-Wave quantum annealing computer. There are very tight facility requirements for low vibration & other disturbances, highly reliable power and cryogenic cooling. It is a physics experiment that operates at 10 milli-kelvin, which is 300 times cooler than interstellar space.

Furthermore, the cryogenic cooling system must be ultra-reliable because any failure costs months of recovery. Beyond the extensive cryogenic cooling system, facility requirements included a vibration isolated concrete pad, additional PDU and UPS systems, a generator with failover controls and liquid nitrogen dewars.

## CHALLENGES IN HOLISTIC MONITORING AND DATA INTEGRATION

There were three presentations on Challenges in Holistic Monitoring and Data Integration each reporting on capabilities, technologies and experiences at their sites.

**Challenges in Holistic Monitoring and Data Integration**, presented by Andrea Bartolini (University of Bologna – DEI, Italy). The presentation introduces the ExaMon monitoring framework. It then presents how AI can be coupled with ExaMon to deliver datacentre automation. Finally, it introduces edge computing as a way to overpass technological challenges in dealing with fine-grain data and real-time analysis and artificial intelligence.

**Monitoring and Data Integration at LLNL**, presented by Kathleen Shoga (LLNL). The presentation describes the many data sources – including not only HPC system data, but application data- that LLNL is attempting to integrate. It then describes the tools used to aggregate, process and visualize that data. Finally, it shows a usage case that correlates rack temperature with application job data.

**What's new at NERSC? High Resolution Data Monitoring: Challenges in Providing EE**, presented by Norm Bourassa (NERSC). After a short overview of the high resolution data monitoring system architecture, the presentation described usage cases

that demonstrate the value of their system. These included real-time PUE measurement and improving energy efficiency. The energy efficiency examples showed improvement with facility pumps and heat exchangers, fans within the HPC system, and hot-aisle air systems. The presentation concluded with challenges ahead.

## **PANEL DISCUSSION – SOFTWARE FOR ENERGY EFFICIENCY**

Steve Martin was the moderator of this panel. There were four panelists, each of whom made a short presentation on community wide open initiatives on software for energy efficiency.

**GEOPM**, presented by Jonathan Eastep (Intel). GEOPM is a runtime that runs asynchronously on compute nodes, monitors the application's behavior through lightweight profiling, then leverages learning and control system techniques to discover runtime patterns in the application and tune hardware knobs in the underlying hardware platform to exploit those patterns. It is open source software and supports extensible tuning strategies through plugins. Examples of pre-bundled plugins include: one that improves application time-to-solution subject to a bound on job power and another that improves energy-to-solution subject to a bound on impact to time-to-solution.

**READEX – Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing**, presented by Andreas Gocht (TU Dresden). READEX is a tool-aided methodology for automatic tuning of parallel applications that works by dynamically adjusting system parameters to actual resource requirements. It is a European Commission funded research project with participation from multiple European sites. Applications are first instrumented to provide information and then they are assessed for their dynamism potential (run-time tuning). Both design-time and run-time optimization occurs using power measurement infrastructure and tuning plug-ins.

**PowerAPI**, presented by Ryan Grant (Sandia NL). The PowerAPI is a portable API for power measurement and control. This Power API provides multiple levels of abstractions to satisfy the requirements of multiple types of users. It drives interface standardization to enable interoperability of different implementations of power components within the HPC system.

**PowerStack**, presented by Siddhartha Jana (Intel). The PowerStack is a global collaboration of engineering efforts towards a well-defined, community-wide stack that accounts for power-awareness across various layers of the HPC software ecosystem. The PowerStack explores hierarchical interfaces for power management at three specific levels: batch job schedulers, job-level runtime systems, and node-level managers.

## **EE HPC WG BOOTH**

The EE HPC WG booth opened with the Exhibitor Floor opening gala on Monday and was staffed during the day Tuesday through Thursday for the rest of the week. The EE HPC WG booth looked inviting and volunteer staffing was excellent. It was very



effective at recruiting new members and providing a presence for the EE HPC WG. We have already signed up for a booth next year.

## **BIRDS OF FEATHER (BoF)**

There were six Birds of Feather sessions organized by the EE HPC WG. Four of them were sixty minute sessions and two were ninety minute sessions.

### **ENERGY EFFICIENCY CONSIDERATIONS FOR HPC PROCUREMENTS**

This BoF was moderated by Jason Hick (LANL). There were speakers from four different sites. The presentations were concise and there was time for questions from the audience. The room was more than half full. There were some questions from the audience. The presentations were all interesting and fairly diverse in content as well as level of detail.

**CORAL-2 Facility Integration and Energy Efficiency**, presented by Anna Maria Bailey (LLNL). This was a high-level presentation on the facility integration and energy efficiency implications for LLNL during the CORAL-2 procurement. Three vendors were compared and evaluated with a normalized and scaled analysis to a common performance. Questions to consider include efficiency, cooling requirements, electrical requirements. Facility requirements grow with each generation of HPC so limitations and expectations should be included in the scope of work. Once a selection is made, continue to work with the vendor.

**Power Management Oakforest-PACS (JCAHPC)**, presented by Toshihiro Hanawa (University of Tokyo). The Oakforest PACS procurement document was ~80 pages long with ~5 pages related to power and energy efficiency. There were requisites for facility installation, cooling system, job management and operation management. One of the more unique requisites was that the HPC system name-plate power requirement (6MVA) would exceed the available power to the system (4.6MVA) and that there would be power capping and power saving capabilities to manage the delta.

**A Look into the Future of Energy Efficiencies**, presented by David Martinez (Sandia National Laboratory). Looking beyond today's procurements, there will be new technologies for electrical and cooling systems. This presentation speculates on some of those technologies and suggests the implications this might have for the infrastructure.

**CoolMUC-3 A Direct Warm Water Cooled HPC System**, presented by Herbert Huber (Leibniz Supercomputing Center). This procurement stipulated that at least 97% of the system waste heat must be discharged by direct warm water cooling (inlet temperature of 55 degrees C). The solution required direct cooling of switches and power supplies as well as provisions for thermal isolation.

## **THE FACILITY PERSPECTIVE ON LIQUID COOLING: EXPERIENCES AND PROPOSED OPEN SPECIFICATION**

This BoF was organized as a panel session. It was one of two BoFs that were scheduled for 90 minutes. Dale Sartor (LBNL) was the moderator and the panelists were Anna Maria Bailey (LLNL), Herbert Huber (LRZ), David Grant (ORNL) and Dave Martinez (Sandia NL). There was good attendance and audience participation. Dale opened the BoF with an overview of warm liquid cooling, then allowed for introductory remarks from panelists on their strategies and lessons learned with liquid cooling. He then gave an overview of specifications for a liquid cooled rack. The floor was then open for panel and audience discussion.

## **POWERAPI AND REDFISH: STANDARDIZING POWER MEASUREMENT AND CONTROL FOR HPC.**

This BoF was organized as a panel discussion. It was moderated by Torsten Wilde (HPE) and panelists included Ryan Grant (Sandia NL and lead on PowerAPI), Jeff Autor (HPE and lead on Redfish), Steve Martin (Cray) and Todd Rosedahl (IBM). There was a brief presentation to start the discussion. The presentation compared PowerAPI and Redfish and then gave an update on the initiative's recent activities. It was held at the same time as the EE HPC WG BoF on Energy and Power Aware Job Scheduling and Resource Management, so the two may have contended for participants. There were not a lot of attendees at this BoF, but there was very lively discussion. Some of the key topics discussed were 1) in verses out of band measurement requirements and 2) measurement frequency requirements.

## **A LOOK AHEAD: ENERGY AND POWER AWARE JOB SCHEDULING AND RESOURCE MANAGEMENT (EPA JSRM)**

Siddhartha Jana (Intel) and Kevin Pedretti (Sandia NL) co-moderated this BoF. There were three other speakers from three different sites. The BoF was well attended, but there was very little time left after the presentations for audience participation.

**EPA JSRM Overview**, presented by Siddhartha Jana (Intel). This presentation summarized the work done by the EPA JSRM Team; a global survey investigating motivations for deploying EPA JSRM and solutions adopted.

**Energy-Aware Scheduling at LRZ**, presented by Michael Ott (Leibniz Supercomputing Center). LRZ's motivation for energy-aware scheduling is to save energy without impacting performance. They have deployed an EPA JSRM solution evaluates codes as to whether or not they are compute or memory bound and then sets the frequency with which the job is run based on this information. Their future plans are to enhance this capability for changing frequency within the application run (micro-tuning). They also plan to use historical application runs to influence job scheduling; e.g., schedule "hot" jobs on adsorption chiller island and balance applications temporally and spatially.

**Exascale Computing Project (ECP) PowerSteering Project**, presented by Tapasya Patke (LLNL). This presentation gave a birds-eye view of a DOE project with goals to implement GEOPM across ECP-enabled applications and demonstrate safe-usage under power or energy constraints and, ultimately, showing runtime performance improvement while maintaining power at or under the system-specified bound. Initial results were presented for several research efforts that are under this project.

**Trinity Advanced Power Management**, presented by Kevin Pedretti (Sandia NL). Although LANL's Trinity system is not power constrained, it is anticipated that future systems will be and so Trinity has capabilities that are helping LANL/Sandia to evaluate how to best use and operate future DOE platforms under a constrained power budget. Several potential usage cases and the power management architecture was reviewed. Some initial results were presented that demonstrated power band and ramp rate management.

### **THE GREEN 500: TRENDS IN ENERGY EFFICIENT SUPERCOMPUTING**

This was the second 90 minute BoF. It was organized as a series of presentations with Q&A following each presentation, followed at the end by a ceremony for giving awards. The first presentations were given by the organizers of the Green500 and Top500. These were followed by presentations from sites that were recipients of the Green500 award and/or had used a high quality power measurement methodology (initiative of the EE HPC WG). Wu Feng (Virginia Tech), Erich Strohmaier (LBNL) and Tom Scogland (LLNL) led the BoF as Green500 and Top500 organizers.

**Measuring Energy Consumption on OLCF's Summit Supercomputer**, presented by Jim Rogers (ORNL). The ORNL facility in which Summit resides has high resolution power and energy instrumentation that allows for the highest quality Level 3 (L3) measurement. Seven industrial meters were used reporting power each second (151,284 measurements). The Summit measurement captured total mechanical load including pumps, fans, lights and control systems. It also captured Infiniband switch loads, management system loads, and file system loads. The rest of the presentation described efficiency trends based on Titan and Summit, projecting to Exascale. From Titan to this first installation of Summit, there has been a 7x increase in energy efficiency. Summit was as 14.66GFlops/Watt, but 50GFlops/Watt is required for the 20MW target for sustained exascale system. The Summit ranked #3 in the Green500 for November 2018.

**GREEN500 L2 Measurement on AI Bridging Cloud Infrastructure (ABCI)**, presented by Hirotaka Ogawa (National Institute of Advanced Industrial Science and Technology (AIST)). The ABCI system has very high resolution power and energy instrumentation on ¼ of the nodes, allowing for a Level 2 (L2) measurement. AIST and Fujitsu reported that power monitoring makes it easy to find problems. The ABCI system ranked #4 in the Green500 for November 2018.

**Level3 Measurement of "Shoubu System B"**, presented by Sunao Torii (ExaScaler Inc.) and Ryo Sakamoto (PEZY Computing K.K.). Neither the facility in which Shoubu resides nor the Shoubu system itself have the high resolution power and energy

instrumentation required for a L2 or L3 measurement. Instead, the submitters put in place external instrumentation (clamp sensors) in order to make the L3 measurement. While preparing to do the measurement, the submitters realized an opportunity for optimizing their cooling sub-system with variable frequency pump controllers. The feedback given by the submitters was that the L3 measurement is not too difficult for under 100KW scale machines, but that understanding the methodology is not so easy. They also reported that the L3 measurement was great in helping to find out hidden overhead. Shoubu System B ranked #1 on the Green500 List for November 2018.

## **DATA ANALYTICS FOR SYSTEM AND FACILITY ENERGY MANAGEMENT**

This BoF was moderated by Torsten Wilde (HPE). There were three speakers from three different sites. The three presentations were dense and interesting. The room was full. There were some questions from the audience, but not a lot of discussion in part due to the amount of time taken by the presenters.

**Opportunities of ML-based Data Analytics in ABCI**, presented by Ryousei Takano (AIST). AI Bridging Cloud Infrastructure (ABCI) is an open innovation platform for advancing AI research and development. It is housed in a new datacenter that has some chillers, but mostly uses passive cooling towers. The facility and HPC system are highly instrumented, the data is collected, aggregated and available for analysis. Two usage cases were described, one for cooling system optimization and the other for improving HPC system utilization. Improvements for both cases described are in process.

### **Combine Out-of-Band Monitoring with AI and Big Data for Datacenter**

**Automation**, presented by Andrea Bartolini (University of Bologna). Building on the presentation made at the EE HPC WG Workshop, Andrea described edge computing as a way to overpass technological challenges in dealing with huge data rates. The rest of the presentation focused on a high-resolution, out-of-band power and performance monitoring solution that has been deployed on one of CINECA's HPC systems (1PF D.A.V.I.D.E).

**Operational Data Analytics in support of Facility Management**, presented by Ghaleb Abdulla (LLNL). LLNL has built and deployed an operational data analytics system that has been in use for more than five years. The presentation describes several use cases where LLNL benefited from having this capability. The first case describes cooling system optimization that resulted in a pump system redesign. Another case described correlating wind gust data with voltage sags, suggesting cause and effect. LLNL's instrumentation for power includes the entire site, not just the HPC facilities.