NEXT MEETING:
August 13th, 2024
09:00 AM Pacific Time
https://lbnl.zoom.us/j/97161857437

Energy Efficient High Performance Computing Working Group
06/11/2024 Meeting Report

INTRODUCTION: The Energy Efficient High Performance Computing Working Group (EE HPC WG) held a meeting on 06/11/24. This Working Group is composed of members representing major governmental departments and independent agencies, private sector representatives, and members of the academic community. More information can be found at the working group's website, https://eehpcwg.llnl.gov/ .

MINUTES:

OperationalData Analytics (ODA) Team, presented by Michael Ott:

We held a Birds of Feather (BoF) at ISC and I think it was one of the better BoFs. We've received pretty good feedback from at least 6 attendees and especially good ratings in terms of interactivity. Once again we used Mentimeter for audience participation. It really pays off that we went through all this effort to create this Mentimeter slide deck. It helps us think about what we want to discuss during the meeting. We also signed up a couple of new members for the EE HPC Working Group in general, and for the ODA Team.

At ISC, there was also the Monitoring & Operational Data Analytics (MODA) workshop. There were a couple of very interesting lightning talks. We suspect that the people working on ODA don't have a strong publication focus. So, offering the possibility to share hands-on experience in a format that doesn't require one to write a full paper increases participation.

We have started preparations for a BoF for SC24. The deadline is the 5th of July. We are thinking that we would like to stress a bit more the things that you can do that might seem less sophisticated, but still deliver strong results. That was some feedback we had during the ISC BoF.

The Team is still working on establishing a generic dashboard that would work for multiple sites. Tim Osborne and Rachel Palumbo from Oak Ridge initiated this idea. They have tried several different approaches but haven't yet settled on one that works well.

Liquid Cooling How-to and TUE Teams, presented by David Sickinger:

The Liquid Cooling How-to Team released a whitepaper in October of last year and created a webpage with the document. The intent was that it would be a living document, and that we could update it over time. If people haven't had a chance to review it and want to offer up any feedback or suggestions for improvement, we're open to that. Here is a link to that webpage:

https://sites.google.com/lbl.gov/ee-hpc-wg-liquid-cooling/home. We are discussing putting in a submission for an SC24 BoF on liquid cooling with a focus on next generation technology.

The other thing that we're starting to talk about is revisiting TUE, the total usage effectiveness metric. It can reflect what choices are made when it comes to deploying liquid cooled systems. We're re-evaluating the energy efficiency of our flagship systems. TUE helps with extracting out the power that is more accurately characterized as belonging to the facility; things like CDUs, rear door heat exchangers. We're documenting what that does to our overall PUE. I think there are 2 levels to look at. One is from the ARPA Cooler Chips program, where they have some very aggressive goals for measuring server fan energy. The other is from an operational perspective for different data centers. What can we do to make TUE useful and relevant, but not a big burden. Maybe we can estimate some things? Maybe some use case examples that are showing the differences. So, there is renewed interest in that TUE metric. The EE HPC WG has a webpage dedicated to TUE, and there are a lot of great papers from the 2013-14 timeframe.

Comments from Norm Bourassa: I am working at NERSC to get an iTUE and TUE. It's the server fans that are causing the problem because there's a bunch of them. Each individual fan is not a lot of energy, but they total up to what I'm estimating is about 28% of our air cooled rack power. I'd like to find a method to calculate that. I'm eagerly awaiting this group to start up again. The level of accuracy that we need for server fans is not that high. If we're calculating an iTUEand a TUE to 2 significant figures, then we only need to be able to reliably estimate the fan power in air cooled servers to 3 significant figures. I would argue that we could do that by just understanding what the free flow power curve is of that fan and a typical system impedance, meaning the air resistance through an air-cooled server. We could find the operating point based on the RPM speeds. That's telemetry that we can easily get from these server fans, and I could see us easily, being able to calculate a very reasonable fan power estimate for the air-cooled servers. If we had a library of air resistance curves for typical air-cooled servers, we could very quickly move to a fan power estimate, using the RPM telemetry from the existing servers. I don't see liquid cooling as being that difficult in terms of the metrics, because the CDUs are separate, and they're often powered separately, and they're well metered. It's very easy to separate out the overhead cooling power for liquid cooling.

Power Measurement Methodology Team presented by Thomas Ilsche

The ISC BoF was well attended and there were presentations on the new number one and number two systems. There was quite some focus on the actual results of the list. We also had a good discussion on the methodology update, particularly about the introduction of a level 0. It is controversial in the community. We think we have a good plan that involves reviewing the submissions, especially the level 0 submissions. We are also discussing with the Top 500 and the Green500 to update the rules with respect to what you can do in the submission. We are somewhat optimistic about making progress there, but it is maybe not the fastest moving ship on the planet.

Question from Brandon Biggs - INL: Do you have any resources where I can read more about your power collection work?

Thomas answer: Here's the link to the latest methodology team update ISC24 BoF slides: https://docs.google.com/presentation/d/1RkM79NpGDfGNZwxFJUKwnI3dgPq7k5-0/edit#slide=id.p1

Power API Team presented by Natalie Bates

The Power API and Redfish presented at a Webinar on expansion efforts to include liquid cooling and a recording is available on the EE HPC WG Website.

Conferences Update presented by Torsten Wilde, Sid Jana and Natalie Bates

We have two upcoming workshops. One is the virtual EE HPC WG Workshop on June 25th and 26th. This workshop has invited speakers only. The other is the in-person Sustainable HPC State of the Practice Workshop which will be held as part of the Cluster 2024 Conference in Kobe, Japan on September 24th. This second workshop is a paper workshop with submitted papers, peer reviews and accepted papers presented and published. Links to websites for both workshops can be found at this link: https://eehpcwg.lbl.gov/upcoming-events.

We have 2 invited speakers for the virtual EE HPC WG Workshop on June 25th and 26th. Johannes Kirnberger from the AI and sustainability team at OECD. Johannes focuses on the organization's policy guidance on green and digital twin transitions and the environmental impact of generative AI. The other invited speaker is Andrew Chien, from University of Chicago, and an affiliated member of the National Science Foundation. Andrew will focus on sustainability and HPC. We also have 4 sessions planned. One session is on motivating end users for energy efficiency which will be led by Matthias Maiterth. We have a second session that's on sustainability metrics for impact led by Rob Bunger. Another on advanced facility cooling controls and Vali Sorell is leading that. The fourth session. Session 4 is on Green500 measurements run by Thomas Ilsche.

For SC24 in November (Atlanta, Georgia USA), we already have a booth on the exhibitor floor. We also have a half-day for a symposium on sites with net-zero goals by 2035 at the Sustainable Supercomputing Workshop. The other half of that workshop will be for papers. Birds of Feather submissions are due July 5th. We have also already made a submission for a panel on Sustainability, but the results of that submission are not known yet.

Beyond SC24, later this year, we are hoping to have another showcase event like the ones that we have hosted in prior years for RIKEN/Fugaku, ORNL/Frontier and CSC/Lumi.

Other announcements:

There is a call for journal papers on Recent Trends and Advances for Energy Efficient HPC Systems. The deadline is end of October. See the following links for more information

https://www.frontiersin.org/research-topics/61501/recent-trends-and-advances-for-energy-efficient-hpc-systems

https://www.frontiersin.org/journals/energy-efficiency

Participants included:

Brandon Biggs, INL

Christian Wassermann, RWCH Aachen

David Sickinger, NREL

Francis Belot, CEA

Michael Ott, LRZ

Natalie Bates, EE HPC WG

Norm Bourassa, LBNL

Philipp Hematty, University of Heidelberg

Russ Pearson, Independent

Torsten Wilde, HPE

Sid Jana, Intel

Thomas Ilsche, TU Dresden