The HPC PowerStack: *A Community-wide Collaboration Towards an Energy Efficient Software Stack*

Core committee members: Aniruddha Marathe (LLNL), Barry Rountree (LLNL), Carsten Trinitis (TUM), Christopher Cantalupo (Intel), Jonathan Eastep (Intel), Josef Weidendorfer (TUM), Martin Schulz (LRZ, TUM), Masaaki Kondo (RIKEN, Univ. of Tokyo), Matthias Maiterth (LMU, Intel), Ryuichi Sakamoto (Univ. of Tokyo), Siddhartha Jana (EEHPC-WG/Intel), Stephanie Brink (LLNL), Tapasya Patki (ECP, LLNL)

Scan the OR code for more details about this project

WHAT IS THE HPC POWERSTACK INITIATIVE?

- A community-wide collaboration to incorporate power-awareness across various layers of the HPC *software* ecosystem.
- Since 2016, this collaboration is being supported and driven by a broad spectrum of vendors, labs, and academia that focuses on different layers of the HPC stack.
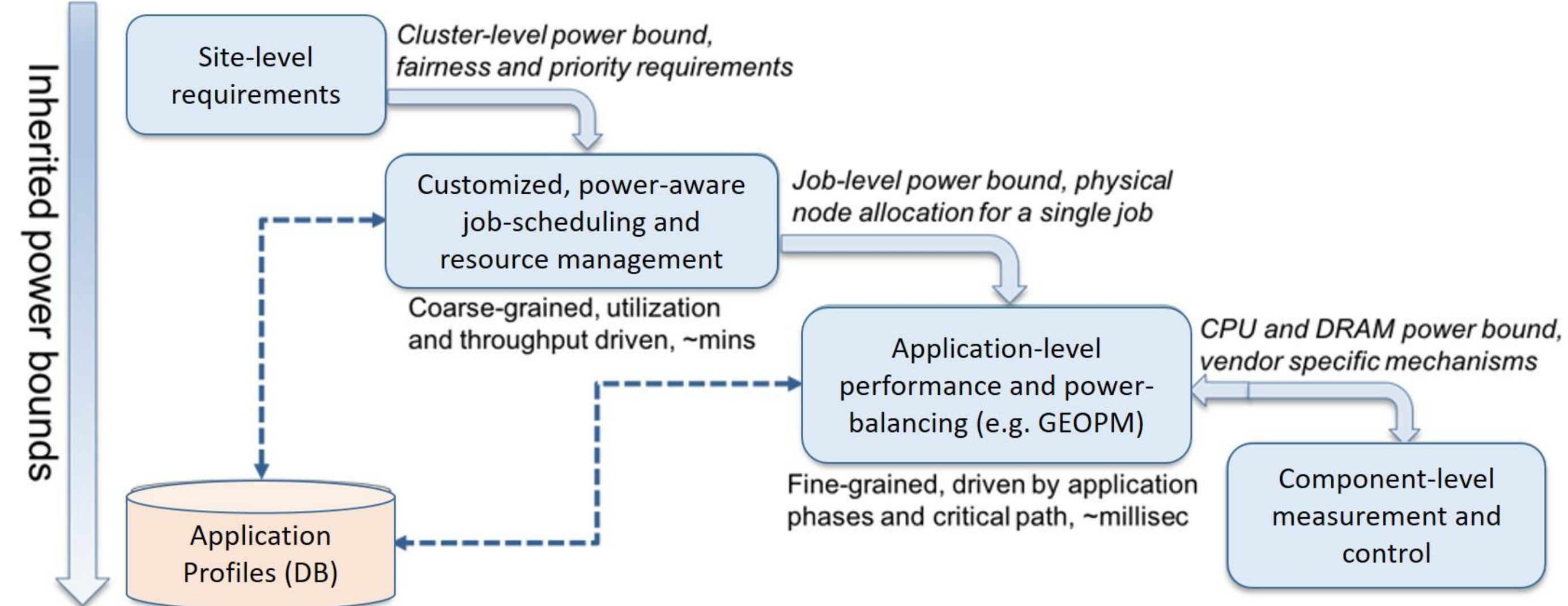
COLLABORATION CHARTER

(A) Identify **key software actors** needed in a system power stack: job-schedulers, application-level runtime, hardware knobs;
(B) Understand the **roles and responsibilities** of the actors;
(C) Reach a **consensus** on their interoperability;
(D) Emphasize on **portability** (thereby keeping the collaboration vendor-neutral);
(E) Combine **existing R&D prototypes and build a community** that actively participates in development and engineering efforts.

## CURRENT LIST OF PARTICIPATING MEMBERS (and growing!)

- *National labs:* LLNL, LANL, Sandia, Argonne, Riken, STFC/Hartree, Cineca, LRZ, Grenoble
- *System Integrators:* Cray, Fujitsu, HPE, ATOS/Bull, IBM
- *Chip Vendors:* x86 (Intel, AMD), ARM, POWER (IBM)
- *Job scheduler / Resource manager vendors:* PBSPro (Altair), ALPS (Cray), Cobalt (Argonne), Flux (LLNL), LSF (IBM)
- *Academia:* TU-Munich, TU-Dresden, UniBo, SDU, Univ of Tokyo, LRZ,
- *Facility and Operations:* EEHPC-WG (Energy Efficient HPC Working Group)

CONTRIBUTING TO THE POWERSTACK COLLABORATION

The HPC community is invited to participate in working groups, contribute to the design of the PowerStack, and subscribe to:

- Mailing lists for announcements: powerstack-announce@googlegroups.com
- Slack channel for discussion: https://powerstack.slack.com
- Git repo for open collaboration: https://gitlab.com/powerstack

PROJECT ROADMAP AND METHODOLOGY
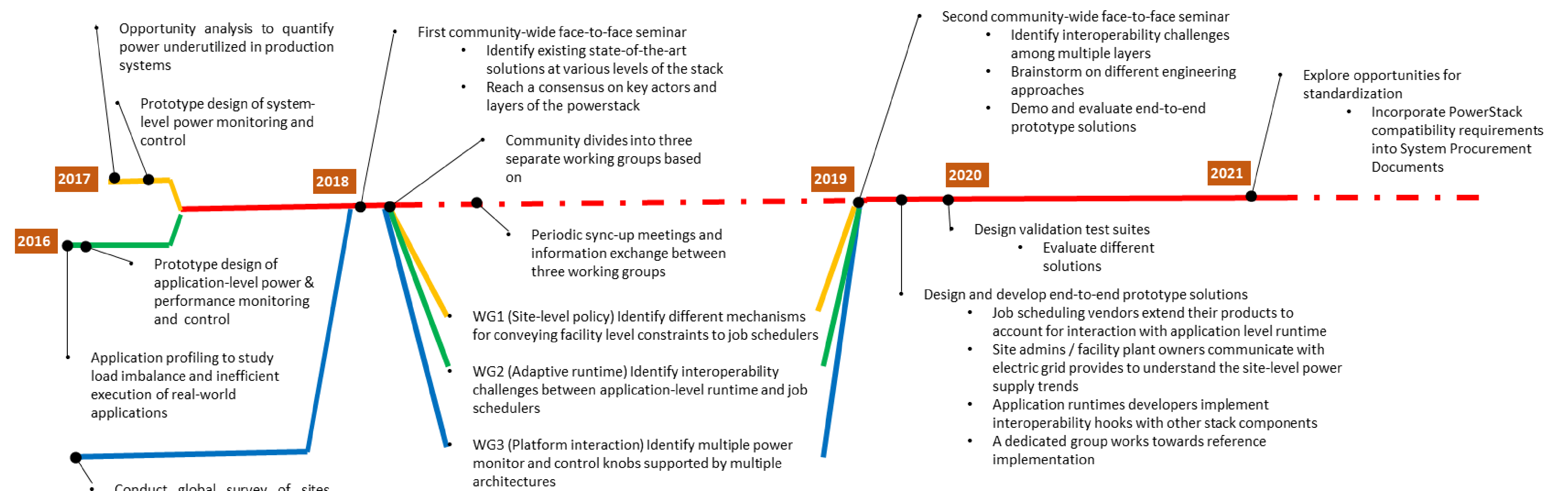
## SNAPSHOT OF RESULTS FROM THE THREE WORKING GROUPS

SYSTEM LEVEL POWER STEERING : WORKING GROUP 2

R&D Goal: Opportunity Analysis of Power-aware Scheduling in Overprovisioned Systems
Active Contributors: Univ of Tokyo/RIKEN, LLNL
Related Publication: refer cited article - [4]
Empirical Analysis: An increase in system-size leads to wasted power that can instead, be redirected towards applications for boosting performance.
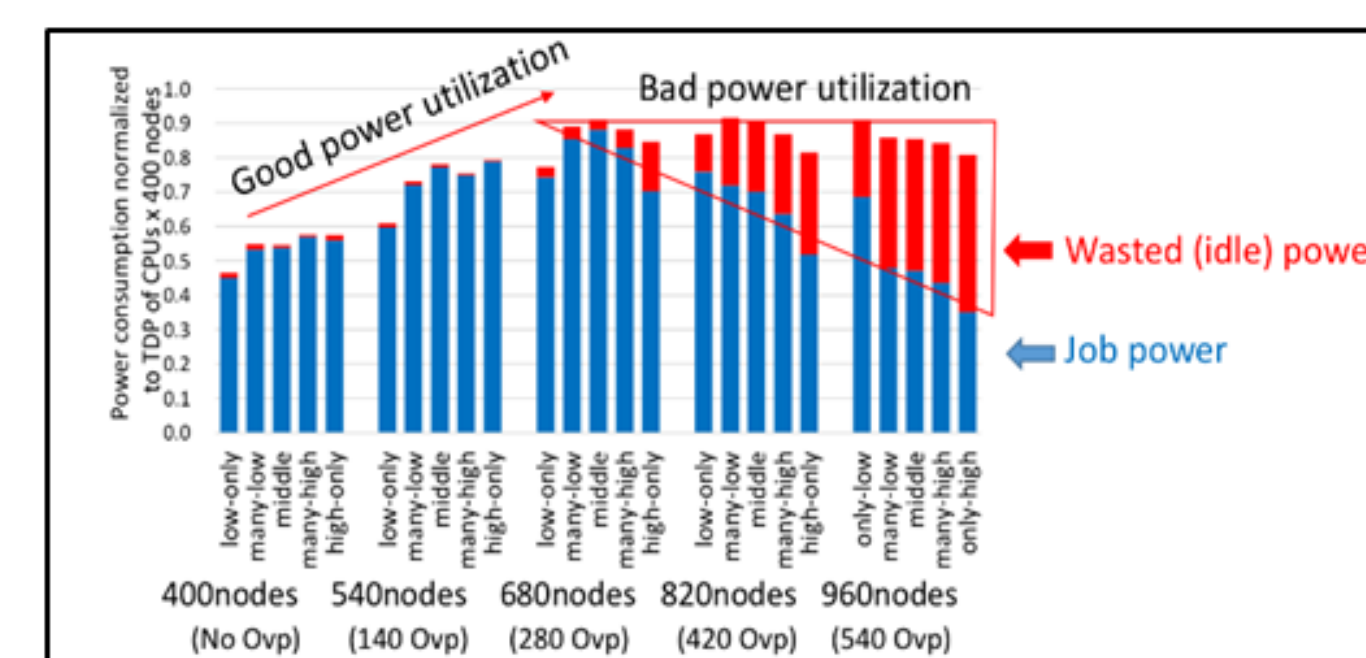

APPLICATION LEVEL PERFORMANCE AND POWER BALANCING IN POWER CONSTRAINED SYSTEMS : WORKING GROUP 2+3

R&D Goal: Design and develop community's first open source application-level power management framework - called GEOPM (https://geopm.github.io), which can interoperate with different layers of the HPC software stack.
Active Contributors: Argonne, LLNL, Hartree, LRZ, CINECA, Sandia, IBM, ARM, Marvell, HPE, Cray, Intel, TUM, UniBo, U. of Tokyo, U. of Arizona, EE HPC WG
Related Publication: refer cited article - [9]
Empirical Analysis: 5% to 30% of performance improvement depending on application design and architecture of power-constrained systems
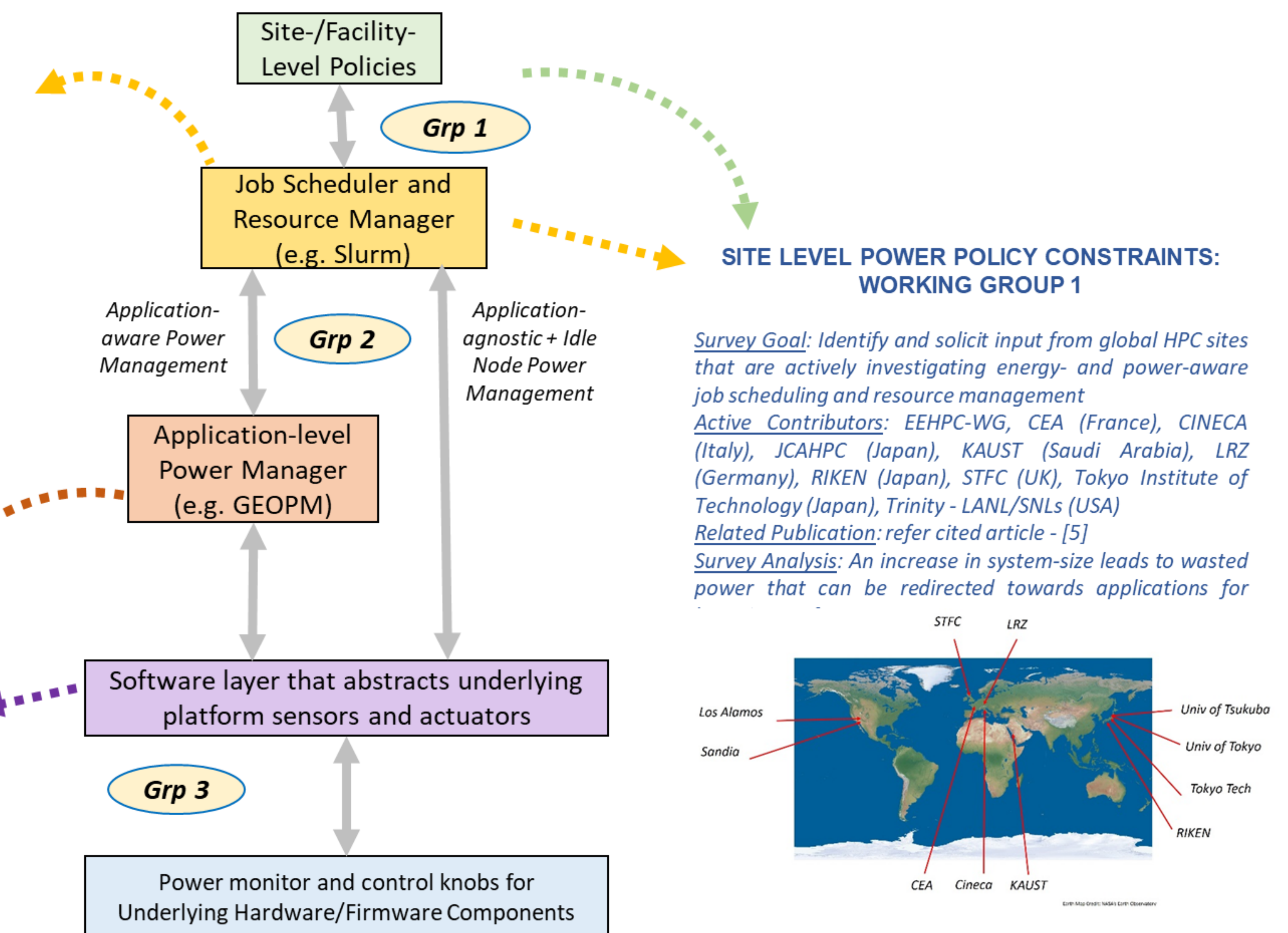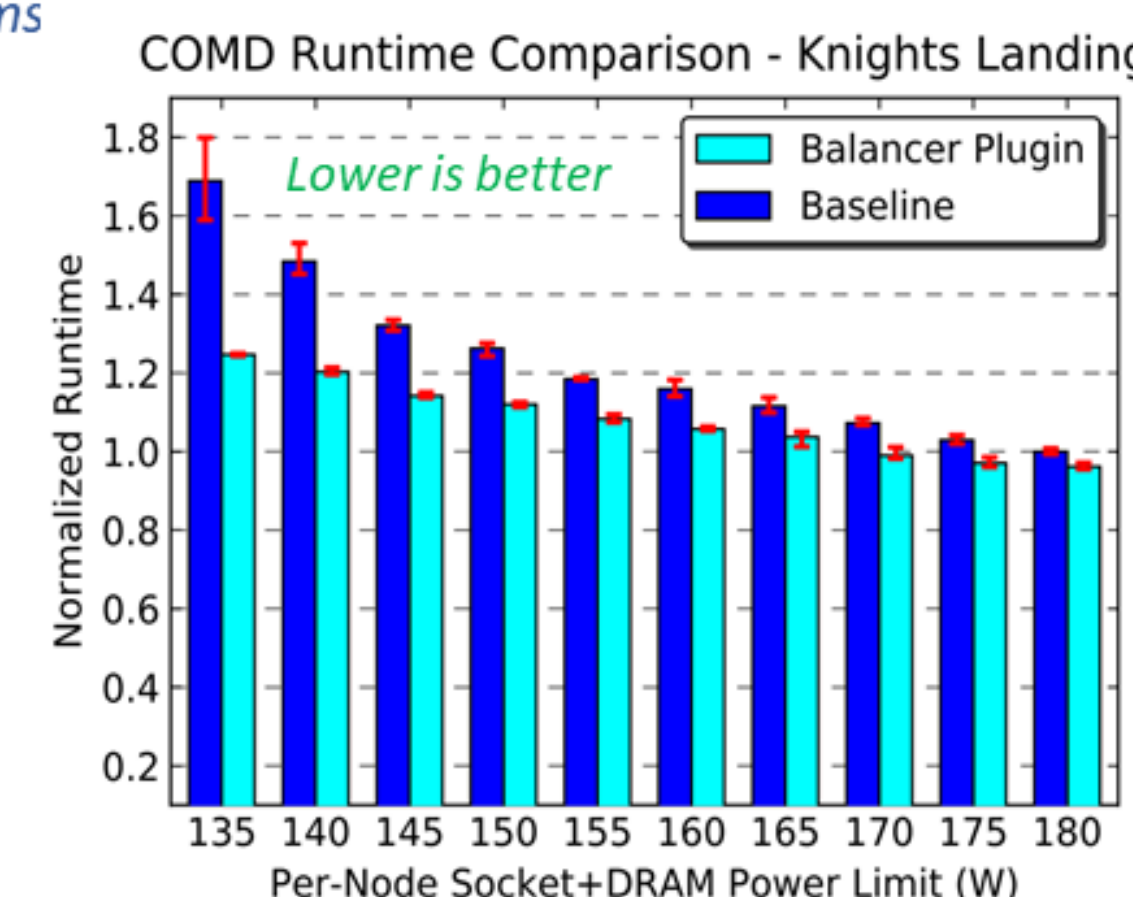
SITE LEVEL POWER POLICY CONSTRAINTS: WORKING GROUP 1

Survey Goal: Identify and solicit input from global HPC sites that are actively investigating energy- and power-aware job scheduling and resource management
Active Contributors: EEHPC-WG, CEA (France), CINECA (Italy), JCAHPC (Japan), KAUST (Saudi Arabia), LRZ (Germany), RIKEN (Japan), STFC (UK), Tokyo Institute of Technology (Japan), Trinity - LANL/SNLs (USA)
Related Publication: refer cited article - [5]
Survey Analysis: An increase in system-size leads to wasted power that can be redirected towards applications for
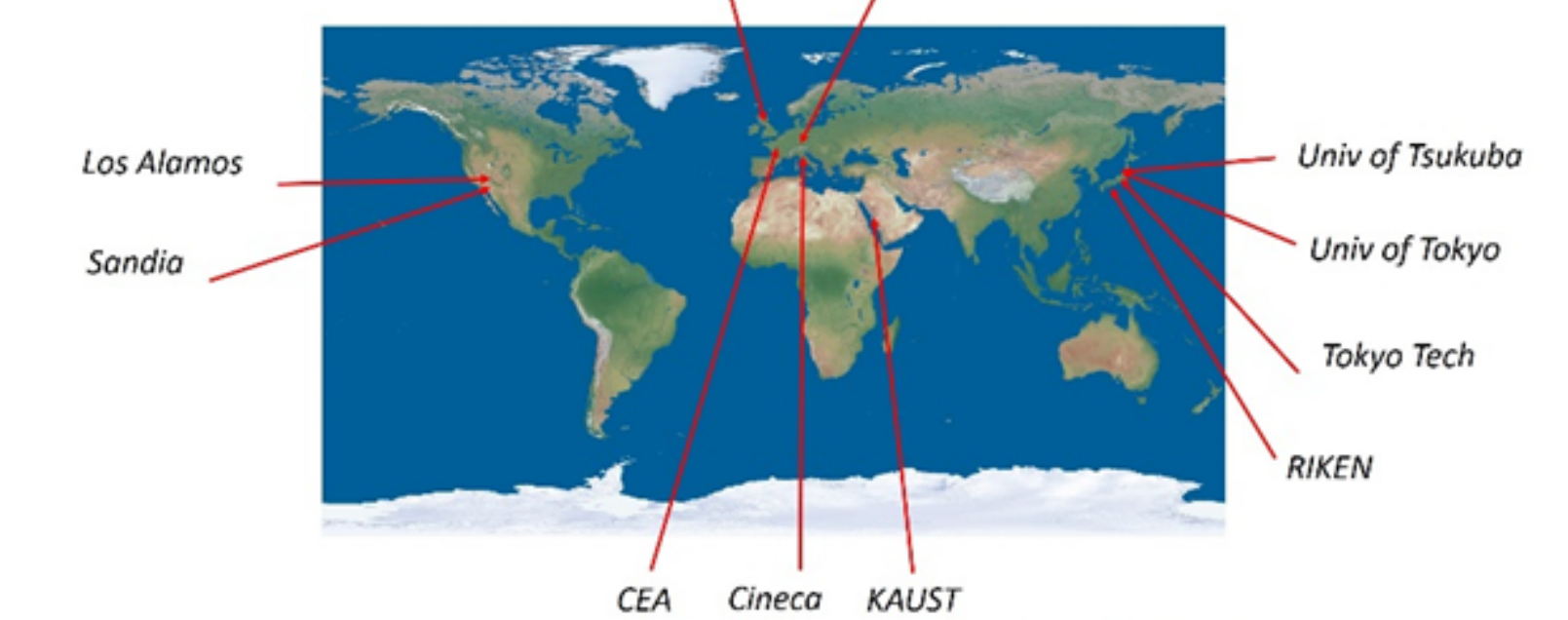
RECENT PUBLISHED LITERATURE

1. HIPEAC 2019 Blog article, "PowerStack: A global response to the power management problem for exascale"
2. The PowerStack Initiative (A Community-driven Effort) - EEHPC-WG Webinar Series, October 2018
3. OSTI Technical report, "A Strawman for an HPC PowerStack", August 2018
4. IPDPS 2018, Proceedings, Sakamoto et al., "Analyzing Resource Trade-offs in Hardware Overprovisioned Supercomputers
5. HPPAC 2018, Maiterth et al., "Energy and Power Aware Job Scheduling and Resource Management: Global Survey - Initial Analysis"
6. Power and Performance Optimization at Exascale - insideHPC, March 2018
7. Energy efficiency and the software stack - InsideHPC, December 2017
8. A global survey of HPC center energy and power-aware job scheduling and resource management, November 2017
9. ISC 2017 proceedings, Eastep et al., "Global Extensible Open Power Manager: A vehicle for HPC Community Collaboration on Co-Designed Energy Management Solutions"