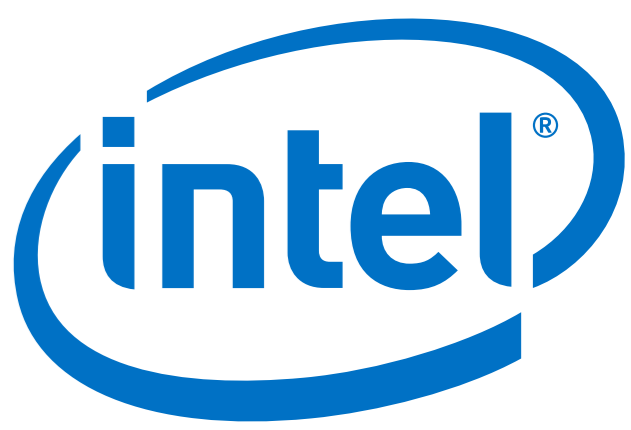


Global Extensible Open Power Manager

For open source code, documentation, and publication list, see:

<https://geopm.github.io/geopm>



A runtime framework for energy management that is open source, extensible, and scalable. Contributed to accelerate community collaboration toward SW/HW co-designed energy management.

Siddhartha Jana, Asma Al-Rawi, Steve Sylvester, Christopher Cantalupo, Brad Geltz, Brandon Baker, and Jonathan Eastep

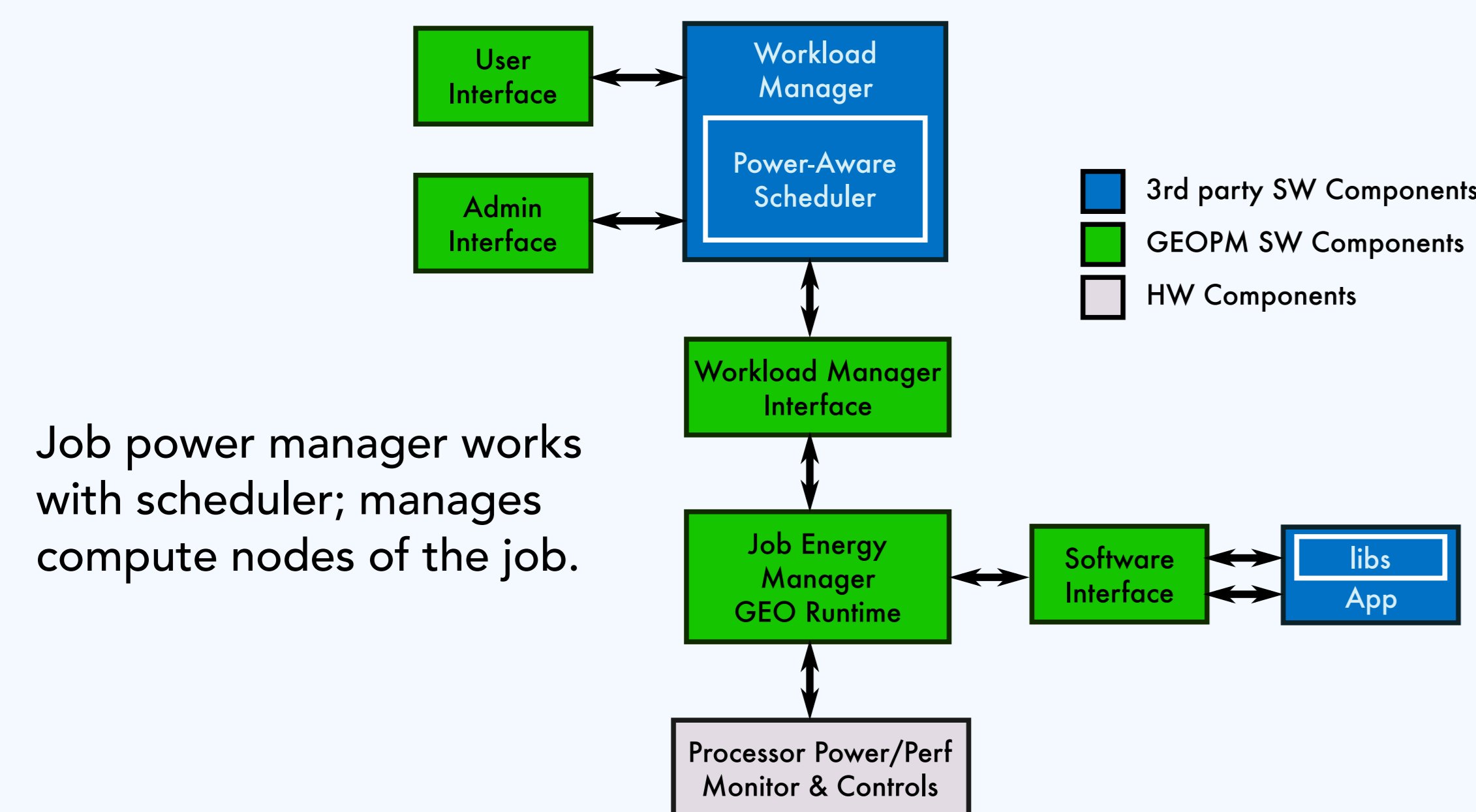
siddhartha.jana@intel.com, asma.h.al-rawi@intel.com, steve.s.sylvester@intel.com, christopher.m.cantalupo@intel.com, brad.geltz@intel.com, brandon.baker@intel.com, jonathan.m.eastep@intel.com

Performance of Future HPC systems will be Limited by Costs of Scaling Power

Overcoming the gap requires a breakthrough in energy management and increased collaborative research between HW vendors and the HPC SW community.

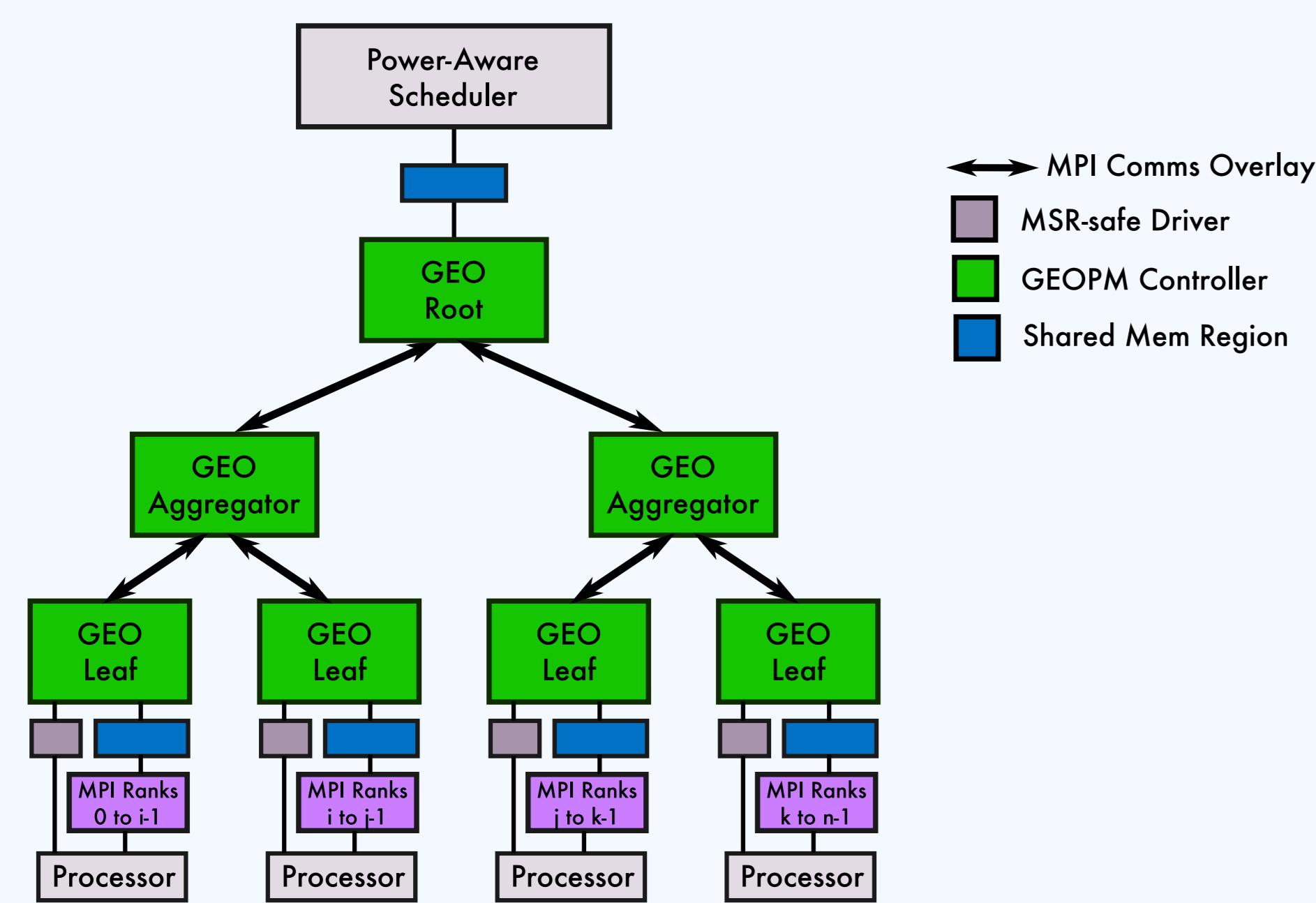
HPC Community may fall short of 2018-2020 Exascale performance goals by an estimated 2-3x margin due to prohibitive costs of scaling power.

GEOPM Interfaces & HPC System Stack Integration



Feedback-guided control system collects feedback from app/libs via GEOPM APIs and tunes power caps, DVFS settings, etc. It can operate with a configurable objective function.

Hierarchical Design & Communication Mechanisms

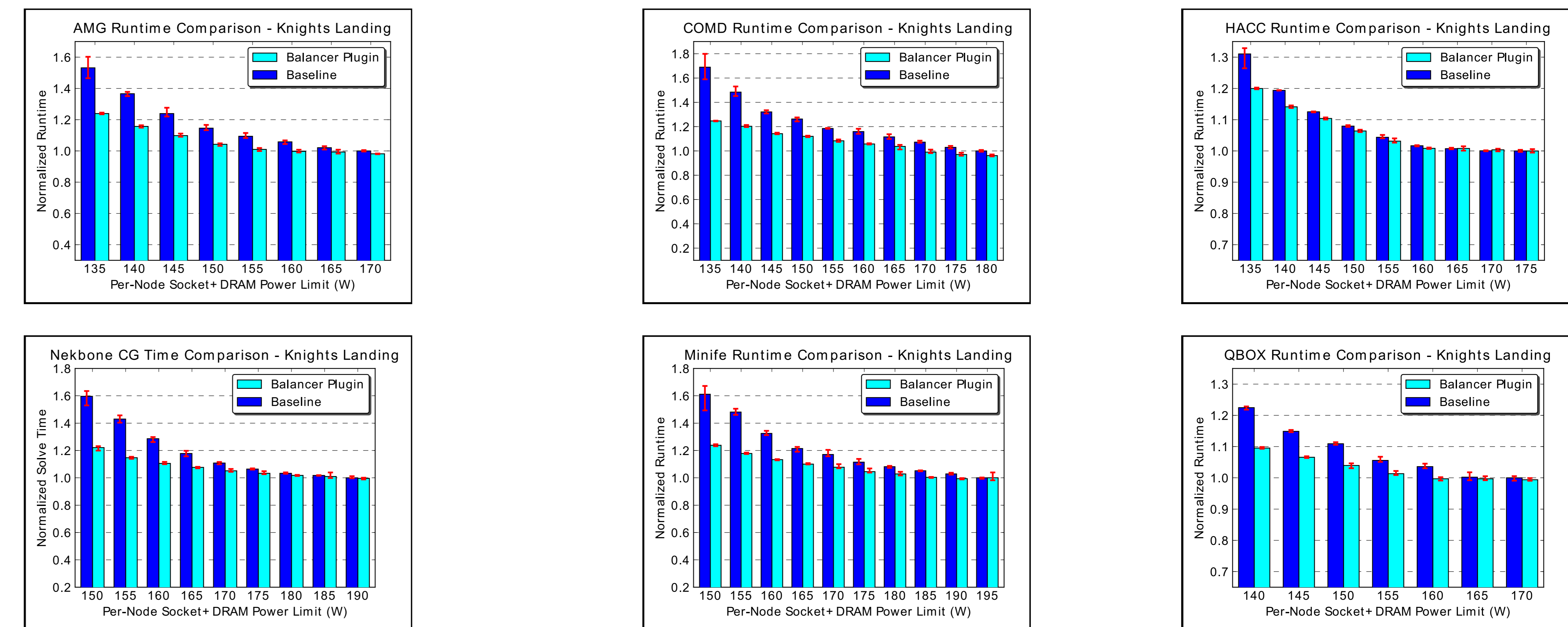


- Scalable tree-hierarchy of controller agents
- Recursive control and feedback approach
- Tree topology and placement optimized via MPI Cartesian grid
- Extensible design via plugins

Primary Contributions

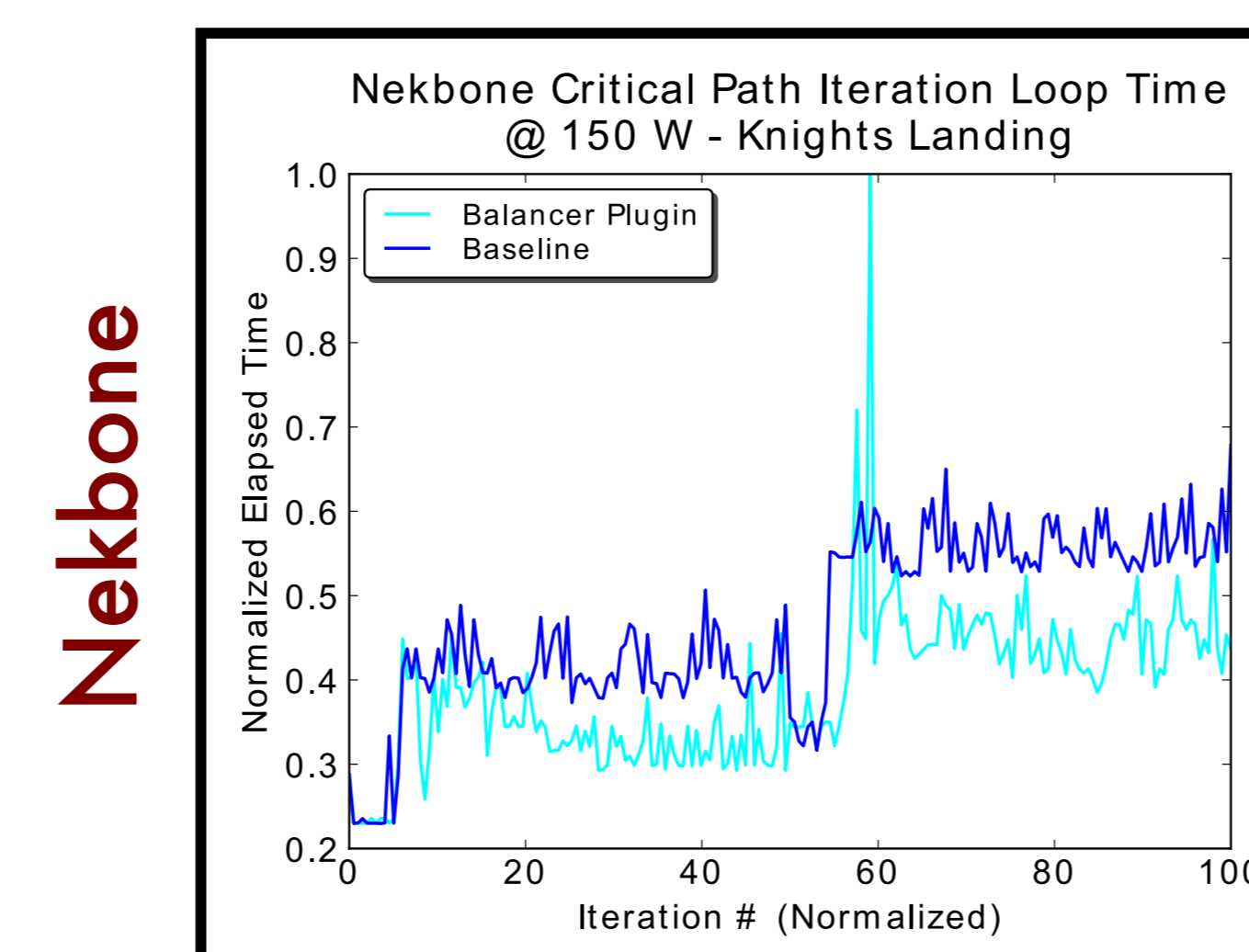
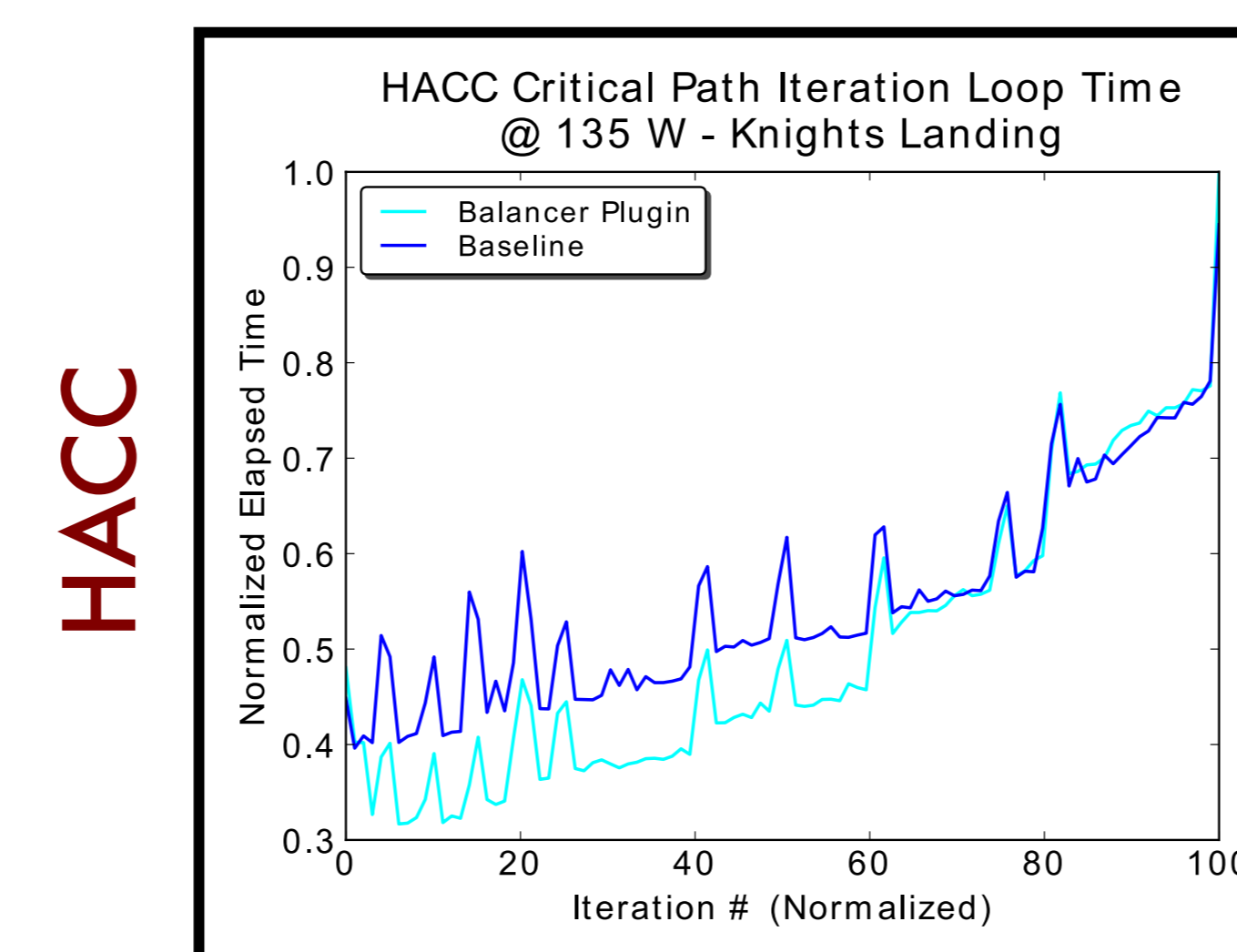
- Hierarchical scalable runtime for globally optimizing HW power and performance control knobs
- Open source extensible runtime is provided as an infrastructure for collaboration and accelerated research
- Includes example plug-ins which guide HW to better use of limited power for significant performance benefits
- **Demonstrate up to 30% runtime improvement**

Evaluation of GEOPM Power Balancing Plugin on a 12-Node Knights Landing Cluster

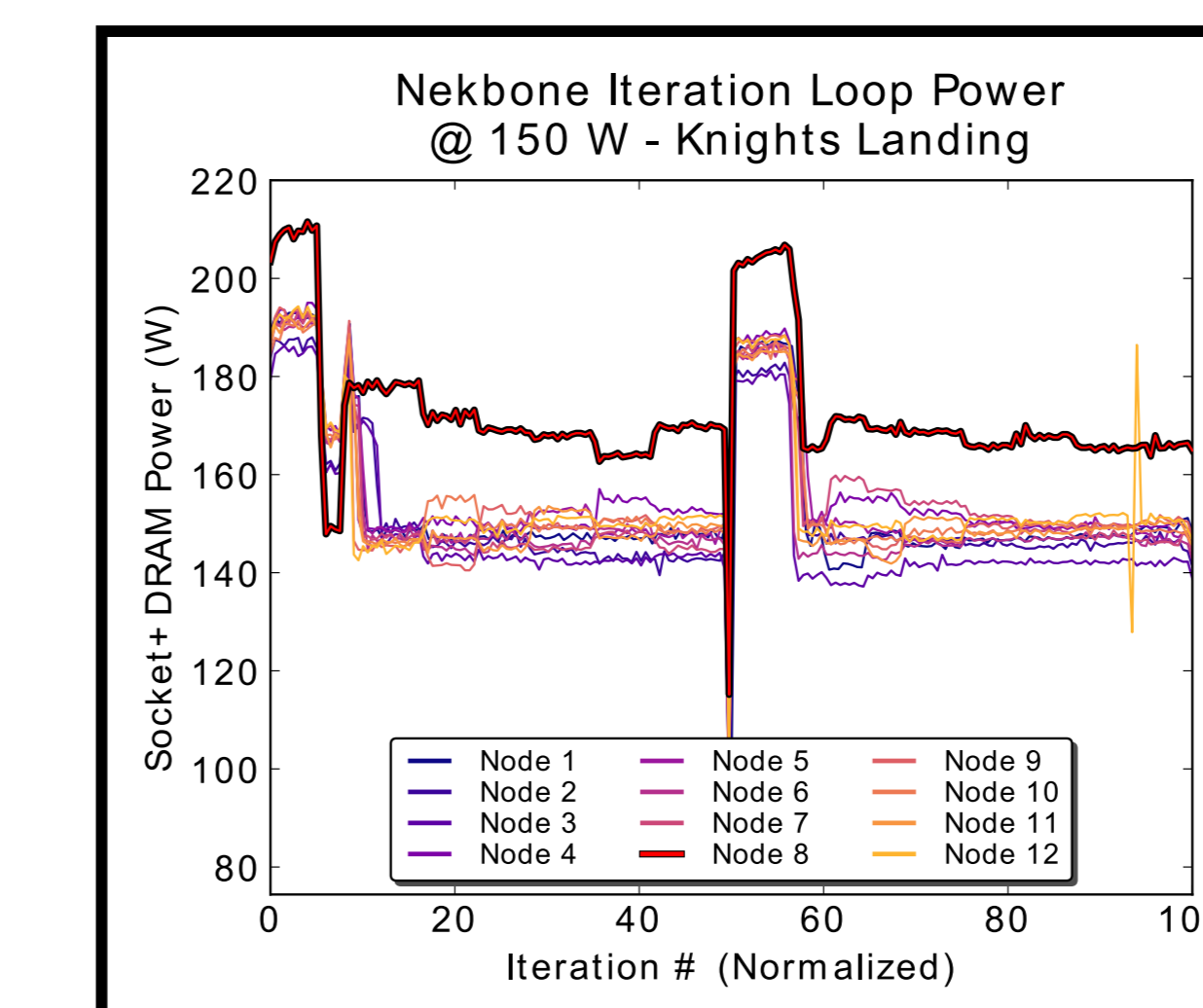
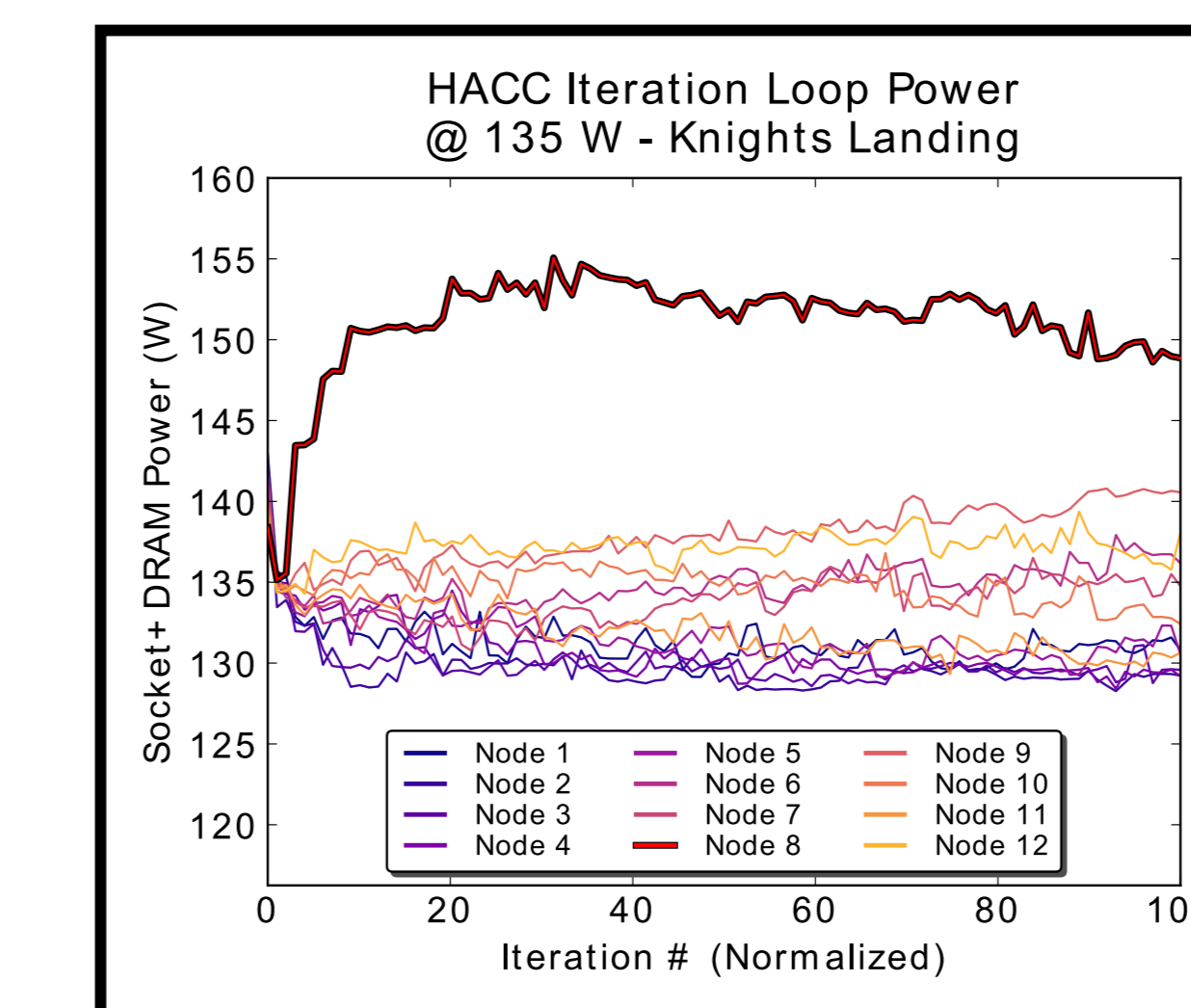


Up to 30% runtime improvement using GEOPM Power Balancing Plug-In

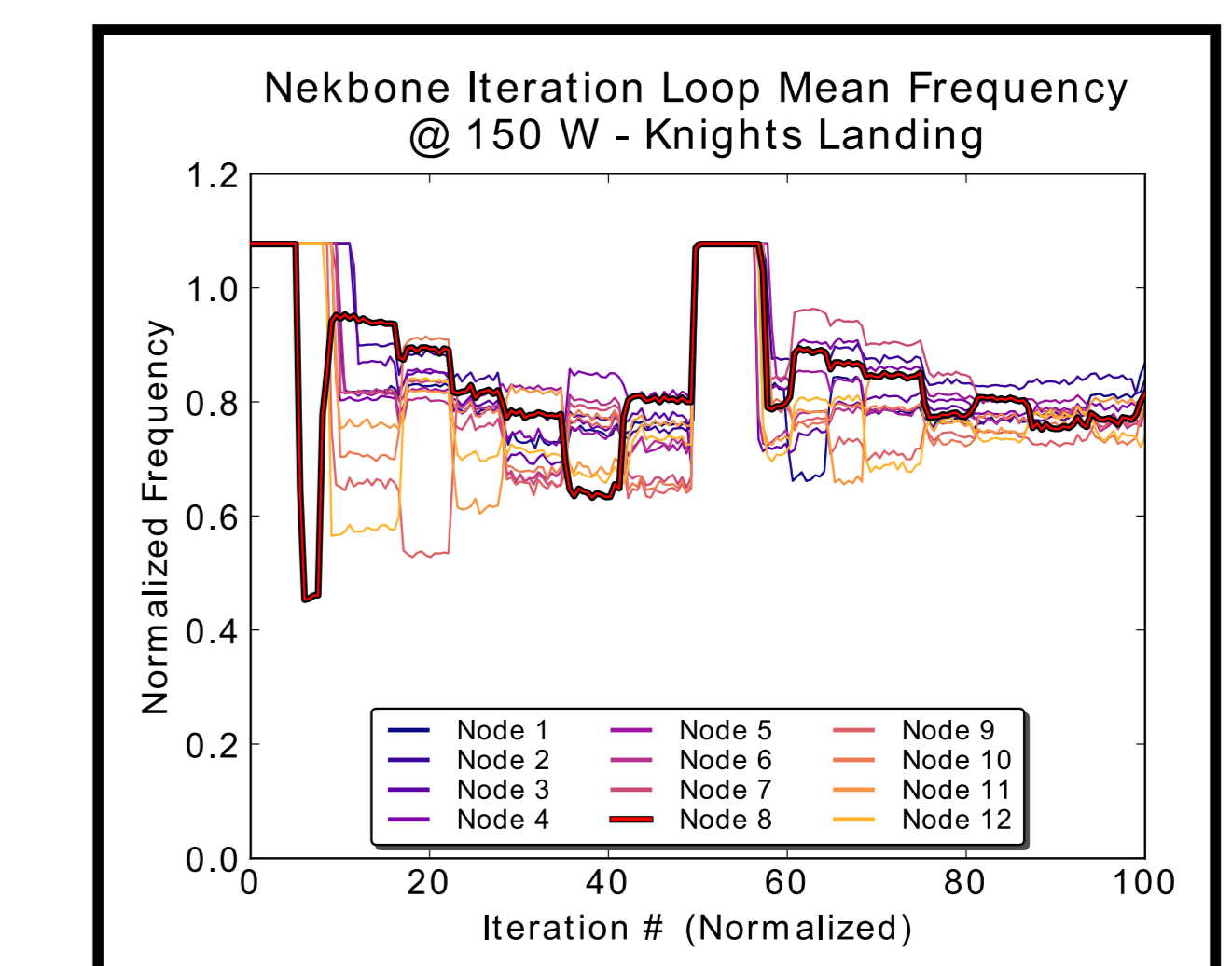
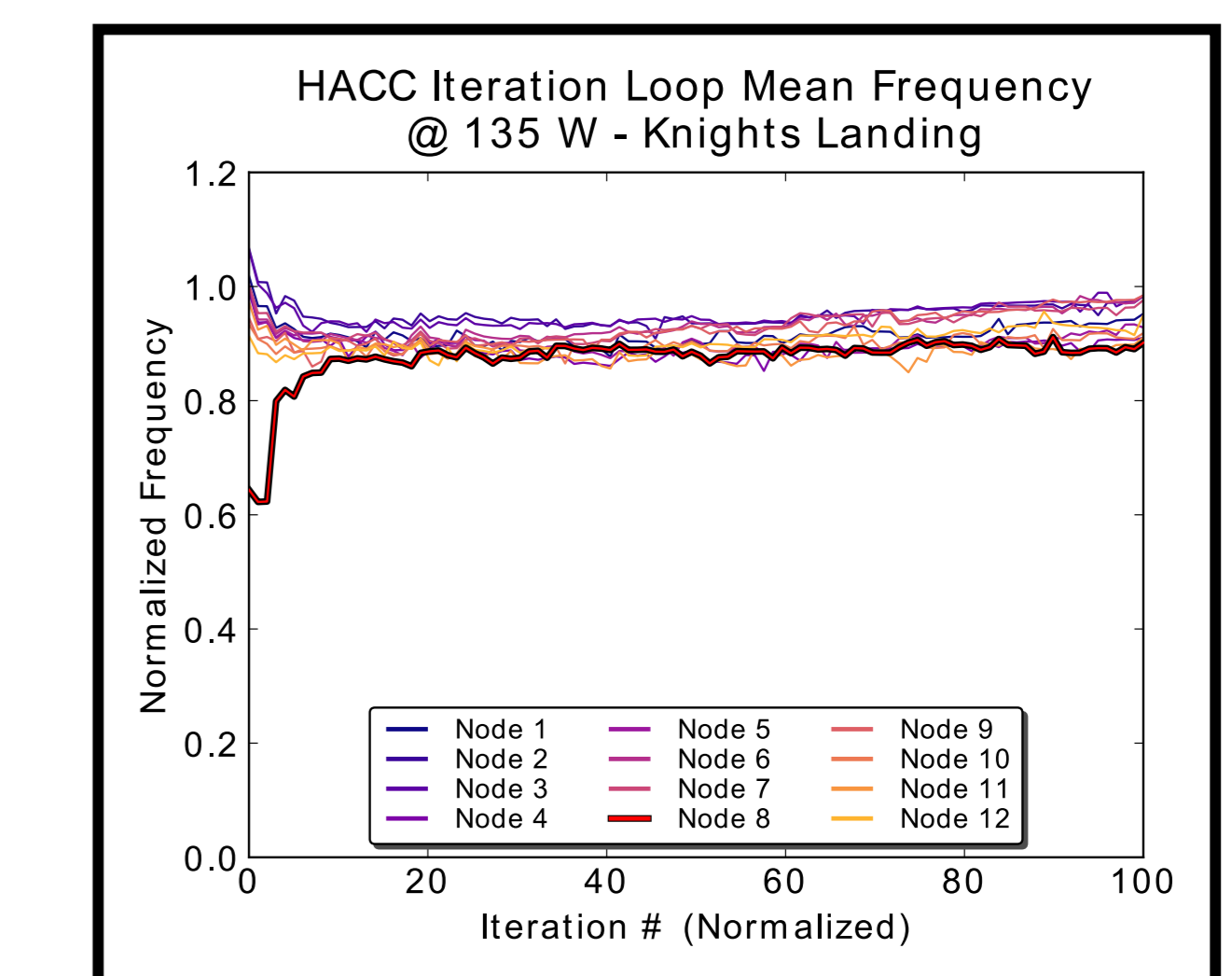
Time for the critical path node to complete each iteration



Power allocated in each node in each iteration



Mean Frequency in each node in each iteration



Future Work

- Spin up further collaborations with the community
- Research new energy optimization strategies via GEOPM's plugin framework
- Integrate with emerging power-aware scheduling functions in workload managers (e.g. SLURM)
- Explore tuning power-performance knobs in SW libraries/runtimes on application layer of HPC stack

First production deployments anticipated on: Theta(Argonne) • Quartz (LLNL) • SuperMUC-NG (LRZ)

Call for Collaboration

The authors are seeking collaboration to explore tuning power-performance knobs in software libraries/runtimes like MPI or OpenMP as well as knobs in the library application layer of the HPC stack.

