Holistic Approach to Energy Efficient HPC

Vadim Elisseev[1], Robert Tracey[1], Kirk Jordan[2]

[1]IBM Research, UK [2]IBM T.J. Watson Research Center

Project Summary

We present a Holistic Approach to energy and power management, which can be described as Energy Aware Scheduling (EAS). EAS uses performance and power consumption models and software hardware co-design for implementing various energy/power aware scheduling policies at the node, job and cluster levels.

Motivation

Energy efficiency is one of the foremost design goals of the Exascale HPC with power efficiency requirement of around 50GFlops/Watt. As of November 2018, the most power efficient HPC system on the Green500 [1] list is Shoubu System B located at ACCC, RIKEN with 17GFlops/Watt. New approaches to energy efficiency are required, which optimize throughout the whole HPC stack - from firmware and hardware through to the OS, applications and workload managers.

Acknowledgements

This work was supported by the STFC Hartree Centre Innovation Return on Research programme, funded by the Department for Business, Energy and Industrial Strategy.

References

[1] The GREEN 500 : https://www.top500.org/green500.

[2] V. Elisseev et al., Energy Aware Scheduling Study on BlueWonder, E2SC@SC18.

[3] GEOPM: https://geopm.github.io/.

[4] M. Puzovic et al., Improving Performance and Energy Efficiency on OpenPower Systems Using Scalable Hardware-Software Co-Design, IWOPH@ISC18.

[5] S. Plimpton, Fast Parallel Algorithms for Short-range Molecular Dynamics,J. Comput. Phys., 1995.

[6] M. Puzovic et al., A Study on Cross-Architectural Modelling of Power Consumption Using Neural Networks, Supercomputing Frontiers and Innovations, Vol. 5, No. 4 ,2018.

Energy Aware Scheduling

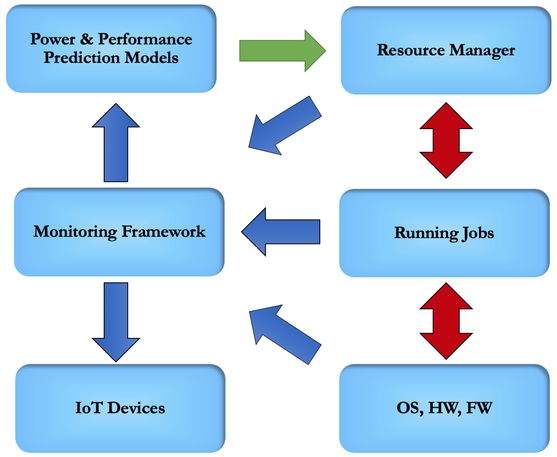

Figure 1: EAS components and interactions

- **Models:** Targeting wide range of workloads, working with various schedulers, cross-architectural, using ML/DL
- **Monitoring Framework:** Hardware counters, server and switches sensors, IoT Devices, historical jobs data, data storage and data access
- **Cluster level policies:** Energy budget, managing idle nodes, energy aware faire share
- **Job level policies:** Min. time to solution, min. energy to solution, power capping, load Balancing
- **Node level controls:** DVFS, power capping, power states

Research and Development Platform

EAS proof of concept platform is currently under development.

- IBM POWER™ (GPU + NVLINK) servers,
- Infiniband and 10GigE networks,
- IBM Spectrum Scale ®storage subsystem,
- Room and rack level IoT devices,
- RHEL 7.x operating system.

Resource Manager

We are using IBM Spectrum LSF as a resource manager, where EAS policies can be implemented as scheduler plug-ins, for example minimum energy to solution or energy budget [2]. We are also considering EAS for Kubernetes.

Runtime optimizations

Our choice for runtime optimizations is the The Global Extensible Open Power Manager (GEOPM), which is an open source framework for power and energy optimizations [3]. We have ported GEOPM to the OpenPOWER architecture [4].

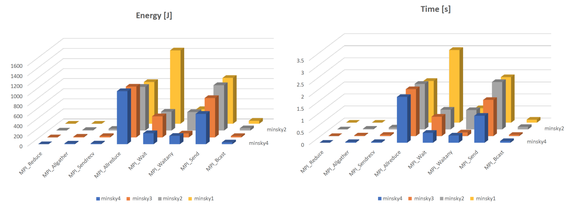

Figure 2: GEOPM: LAMMPS [5] MPI Regions

Monitoring Framework

Monitoring framework is using Elasticsearch® and hardware specific software to collect power and performance metrics from servers, storage, switches and IoT devices as well as workloads stats from a resource manager.

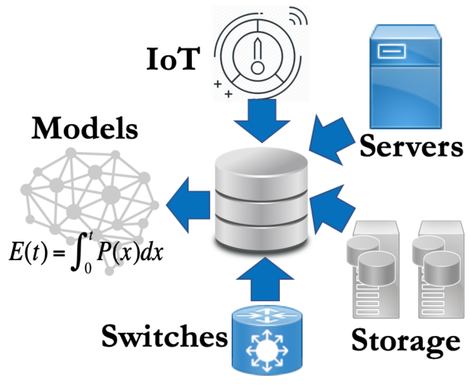

Figure 3: Intelligent monitoring framework for EAS provides data for performance and power predictions models.

Prediction Models

We are exploring different power and performance prediction models using ML and DL algorithms such as Levenberg-Marquardt, Scaled conjugate gradient, Resilient, Bayesian Regularization across different microarchitectures [6].

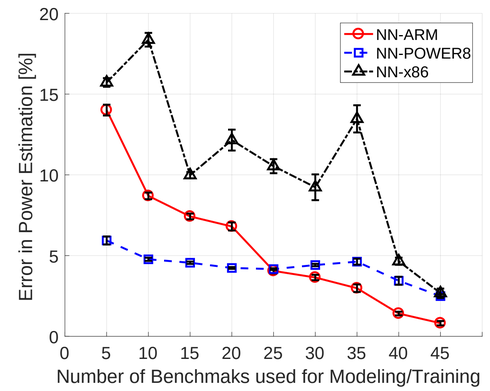

Figure 4: Power estimation accuracy as a function of number of benchmarks.

IoT Devices

Computer room and rack level multi-sensors from iButtonLink, LLC measuring temperature, voltages and humidity with proprietary software for data collection.

Plans and Timelines

EAS is an ongoing project with milestones reflected in multiple publications: [2], [6], [4].

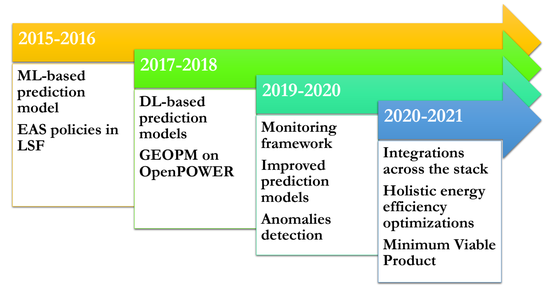

Figure 5: EAS project timelines and milestones.