15th JLESC Workshop



Contribution ID: 31

Type: Short talk

Controlling the Energy Efficiency of HPC Nodes - A Reinforcement Learning Based Approach

Wednesday 22 March 2023 14:20 (10 minutes)

Exascale systems draw a significant amount of power. As each application deployed map to the various heterogeneous computing elements of these platforms, managing how power is distributed across components becomes a

priority. The ECP Argo project is developing an infrastructure for node-local control loops that can observe application behavior and adjust resources dynamically, power included, for better performance. We have recently developed a control loop using reinforcement learning, with a proximal policy optimization algorithm, trained on an existing mathematical model of application progress response to power capping. This dependency on the mathematical model is a hindrance: progress/instantaneous performance is stochastic (noisy) under a dynamic workload and therefore a good approximation model demands more data, and lengthy characterization studies. Therefore, we are exploring methods for bypassing this mathematical model, like actorcritic methods, and are looking for collaborations with know-how on other options, for example: real-time training, existing fully characterized applications, alternative control loop designs.

JLESC topic

Primary authors: RAJ, Akhilesh (Student Researcher at Argonne National Lab); PERARNAU, Swann (Argonne National Laboratory)

Presenter: RAJ, Akhilesh (Student Researcher at Argonne National Lab)

Session Classification: Short Talks on Advanced Architectures

Track Classification: Advanced architectures