Multiparameter resources selection for next generation HPC platforms

Yiannis Georgiou, David Glesser,
Dineshkumar Rajagopal
BULL S.A.S
Email: {yiannis.georgiou, david.glesser,
dineshkumar.rajagopal}@bull.net

Matthieu Hautreux
CEA DAM
Email: {matthieu.hautreux}@cea.fr

***Abstract*—SLURM(Simple Linux Utility for Resource Management) is a Resource and Job Management System(RJMS) in the High Performance Computing(HPC) System software stack. It is a middleware system software between User applications and Operating systems to distribute the HPC Computing resources(Nodes) to the user's Job requirements(Constraints) effectively. Computing nodes internal architecture is evolving and having many-socket,many-core and multi-threaded features to increase the HPC and user applications performance and throughput. SLURM select/cons_res resource selection plugin consuming nodes internal resources of cores,sockets,memory,GPUs to satisfy the users requirement. This will increase the performance and throughput of HPC environment. Current resource selection(cons_res) plugin in the SLURM is slow and not easy to adapt for the new HPC and Compute nodes architecture, Because of using linear data structure(Bitstring) and having limited data and relationships between HPC resources and sub-resources. If we represent the internal nodes architecture informations and relationships between sub-resources for current Intel based cluster, than the resource selection will not be maintainable for the future architectural evolution and new type of resources(GPUs,MICs,Intel Xeon Phi,Quantum Computers). To Resolve that by using general purpose resource management framework called LAYOUTS in the slurm to represent complete informations and relationships between resources to manipulate physical(e.x Node, Switch, etc.) and virtual(Simply anything e.x Rack, Room, etc.) resources in the HPC for different architectures easily. LAYOUTS framework used tree data structure to represent relationships between resources and sub-resources, So it improved the performance of basic operations to sub-linear and improved the performance of cons_res plugin performance drastically/considerably.**

I. Introduction

RJMS is complicated and changed by the changes in the HPC hardware and application requirements. RJMS has to support dynamical changes of the resources and easy to maintain. Current resource management does not support that flexibility and resource selection is not getting complicated for the future HPC. Resource selection plugin managing HPC resources and allocating best resources to the user's job based on the different criteria(number-of-tasks, core, memory, disk-space, GPU, power and temperature). Parameter(criteria) for resource selection is changing over the decades. For an example power and temperature is not a criteria in the Peta-scale HPC system, But it is an important factor in the future Exa-scale systems. So our resource manager should support the changes. Consumable resource(cons_res) selection plugin is developed for Peta-Scale systems to consume resources within the single node to improve the performance and throughput in the HPC system, But cons_res plugin does not manage resources information best manner. Lack of resources information management complicated resource selection, So it is not easy to add a new criteria to the resource selection with the current plugin only flexibility mechanism. Previous cons_res plugin limitations will be resolved By using LAYOUTS framework. Current LAYOUTS framework implemented in the SLURM to support generic resources(physical,virtual,simply anything) and relations to support future HPC resources and architectural changes. Selection plugin left the resources information to the LAYOUTS framework and having only resource selection algorithm to select the best resources based on the different parameters(criteria). LAYOUTS managing resources informations and relationships between the resources in the central place to avoid resources information in the different plugins locally. Currently it supports tree relationship between resources,But multi-tree and graph relationships to make the LAYOUTS framework to represent relationship between resources without any restrictions. LAYOUTS framework internal details will be explained in more details in **Section III**.