



Master of Science in Informatics at Grenoble
Parallel, Distributed and Embedded System (PDES)

Multiparameter resource selection for next generation HPC platforms

Dineshkumar RAJAGOPAL

1st September 2015

Research project performed at BULL-SAS

Under the supervision of:
Dr. Yiannis GEORGIOU
BULL-SAS, Echirolles

Defended before a jury composed of:
Prof. Arnaud LEGRAND
Prof. Martin HEUSSE
Prof. Noel DEPALMA
Prof. Olivier GRUBBER
Prof. Olivier RICHARD

Abstract

SLURM(Simple Linux Utility for Resource Management) is a RJMS(Resource and Job Management System) in the supercomputer system software stack. RJMS manages resources and selects the best resources to schedule the user's job request. Resource selection is an internal operation of scheduling, but improper resource selection operation leads to improper resource management(Increases the cost of resource maintenance in HPC cluster). Slurm resource selection policies(select/cons_res or select/linear) are specific implementations in C, rather than using Databases and SQL query to select the resources like Maui, OAR, XtremWeb, etc. Even though specific implementation of select plugin raises performance, evolution of cluster architecture and resource selection criterias loses maintenance of code. Current cons_res plugin consume internal resources(core, memory, disk) of nodes, manage resource information(Resource management) and support topology aware best-fit resource selection policy. Resource management is getting complicated and Resource selection policies are independent from the resource management(cluster topology,resources,resource types and selection criteria), so separating resource management from policy will improve maintenance of the code. To resolve the dynamic nature of resource management by using generic resource management framework called LAYOUTS, to keep all the resources and relationships in global area and giving resource management related specific APIs. In this paper we cover LAYOUTS framework, new consumable resource(cons_res_layout) plugin based on LAYOUTS, performance comparison of cons_res_layout and cons_res plugins, energy based resource selection policy was supported in cons_res_layout and compared energy efficiency of cons_res and cons_res_layout. Using this framework to develop a new select plugin maintain data separately from code(policy), so changes of cluster topology or criterias require changes in data(layouts) only. The new cons_res_layout follows the cons_res operations to support different criterias(power, temperature) and flexibility, with considerable performance overhead. From the experience of developing cons_res_layout, suggested some guidelines about LAYOUTS APIs and internal architectures to SLURM plugin developers.

key words: RJMS, Batch scheduler, Workload manager, SLURM, HPC resource management, Resource selection, Resource allocation, LAYOUTS Framework, Resource allocation policy(RAP)

Contents

Abstract	i
1 Introduction	1
2 State of the art...	3

— 1 —

State of the art...

Some text...

