
Improving Resiliency through Dynamic Resource Management and Scheduling

Felix Wolf*¹

¹Darmstadt University of Technology [Darmstadt] – Karolinenplatz 5 Karolinenplatz 5, Germany

Abstract

Batch systems traditionally support only static resource management wherein a job's resource set is unchanged throughout execution. Node failures force the batch systems to restart affected jobs on a fresh allocation (typically from a checkpoint) or replace failed nodes with statically allocated spare nodes. As future large-scale systems are expected to have high failure rates, this solution leads to increased job restart overhead, additional long waiting times before job restart and excessive resource wastage. In this talk, we present a dynamic resource management mechanism that enables on-the-fly replacement of failed nodes to affected jobs without requiring a job restart. We present a scheduling algorithm for the combined scheduling of various job types and show how the unique features of these jobs and the scheduling algorithm can expedite node replacements. Experimental results derived through simulation show better job throughput even under high failure rates when compared to static resource management, thus improving the overall resiliency.

(Joint work with Suraj Prabhakaran)

*Speaker