

Flexi-Cluster: A Simulator for a Single Compute Cluster

Flexi-Cluster is a flexible, discrete-event simulation model for a single compute cluster, such as might be deployed within a compute grid. The model, written in SLX¹, facilitates investigation of alternate algorithms for admission control, queue management and scheduling within a single compute cluster. The key innovation in the model is to permit users to develop their own functions for making admission and scheduling decisions and for managing a job queue and to insert those functions within a model framework that provides simulated jobs, a simulated compute cluster and an instrumented measurement collection system.

Flexi-Cluster consists of three main elements: a job generator, a compute cluster and jobs. The job generator constructs a stream of jobs with user-specified characteristics, including: duration over which jobs are generated; load rate; the number of processors required by three jobs widths (narrow, typical and wide); proportion of cluster load absorbed by each job width; mean, minimum and standard deviation in job lengths; base value for jobs and value skew for high-value jobs; delay before a job begins to decay in value; base decay rate for late jobs and decay skew for high-urgency jobs; proportion of jobs that have normal value and urgency; penalty rate for expired jobs. The combination of job values and urgencies result in a job mix with up to four classes. The compute cluster is defined by the number of available processors, processor speed and user-designated functions to decide job admission and to schedule job execution. The cluster may also be designated to use first-fit backfilling or not. Each job provides four methods (reject, queue, defer and run) that allow the cluster to change job state. The current version of Flexi-Cluster comes with three admission control modules (admit all jobs, admit jobs randomly, admit jobs with optimum probability), three scheduling modules (first-come-first-served, priority based on decay over job length and random) and associated queue management modules (first-in-first-out, random, first-fit and decay-over-job-length).

The measurement framework captures both job-related events and summary statistics in aggregate, as well as by job width and job class. Captured events include job arrivals, rejections, admissions, deferrals, executions, completions and expirations. The framework also monitors changes in processor usage and queue size over time. Summary statistics include: cluster utilization; total revenue, revenue per second and revenue per job; decay penalty; count of jobs arrived, accepted, rejected, completed, deferrals, late and expired; average queue length, queued work, earliness, lateness and deferred time.

Flexi-Cluster executes quite quickly, completing 250 simulated years of cluster operations at 100 % load in as little as 10 minutes, depending on the processor used. The next version of Flexi-Cluster will include the EASY backfilling algorithm, job pricing determined by the compute cluster, and job arrivals conditioned on asking price.

¹ SLX is a commercial simulation system available from Wolverine Software, see: <http://www.wolverinesoftware.com>.