

New and Noteworthy JuFo and PTP 6.0

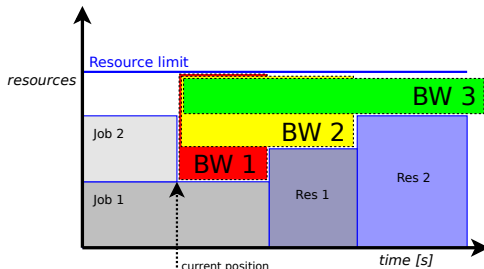
October 26, 2012 | Carsten Karbach

Part I: JuFo

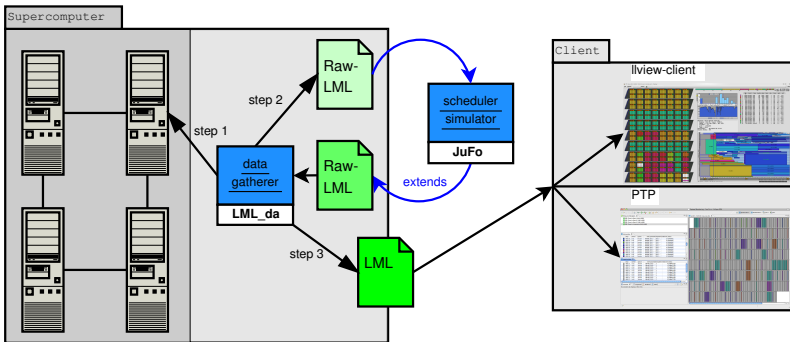
October 26, 2012 | Carsten Karbach

Overview

- **Configurable** simulator for global job schedulers for **on-line prediction** of job dispatch dates
- Based on analysis of JSC batch systems **Moab** and **Loadleveler**
- **Integrated** with monitoring system **LLView**
- **LML** as configuration and communication data format
- **Use-cases:**
 - **User** predicts start dates of submitted jobs
 - **Administrator** simulates job scheduler performance with various input parameters



Architecture



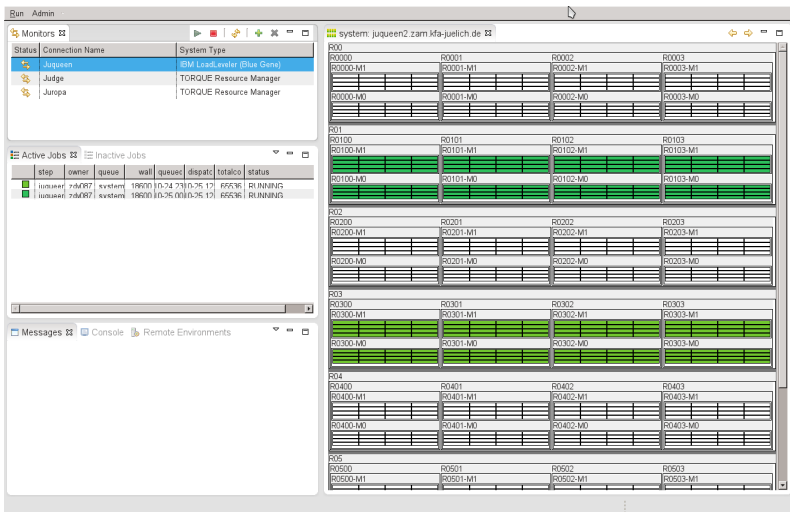
Features

- Supported **scheduling algorithms**
 - First-Come-First-Served
 - List-Scheduling
 - Backfilling
- Available **simulation parameters**
 - Generic job **prioritization**
 - Advanced **reservations**
 - Jobs can request CPUs, GPUs, memory
 - **Nodesharing**
 - **Queue** constraints
- Test framework for evaluating JuFo's accuracy

Part II: PTP Monitoring Updates

October 26, 2012 | Carsten Karbach

Standalone Monitoring Client



The screenshot displays the Standalone Monitoring Client interface, which is divided into several panels. The top-left panel, titled "Monitors", shows a list of monitoring connections:

Status	Connection Name	System Type
🔌	Jaqueen	RM Loadleveler (Blue Gene)
🔌	Judge	TORQUE Resource Manager
🔌	Juropa	TORQUE Resource Manager

The bottom-left panel, titled "Active Jobs", shows a table of active jobs:

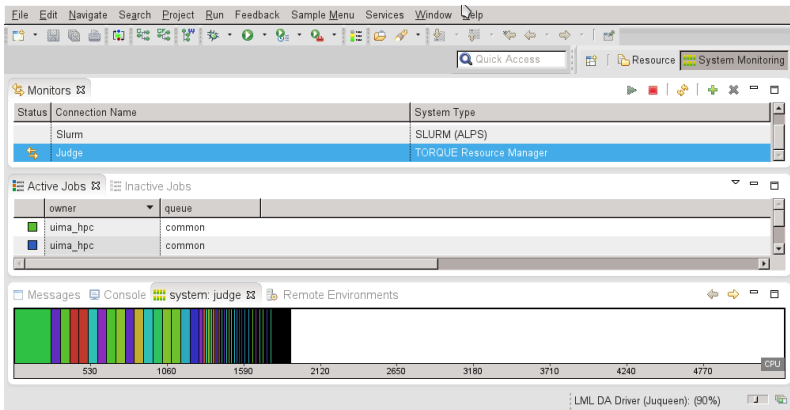
	step	owner	queue	wall	queued	dispatc	totalco	status
🟢	juqueen	zhd087	svstem	18900	0-24 23	0-25 12	65536	RUNNING
🟢	juqueen	zhd087	svstem	18900	0-25 00	0-25 12	65536	RUNNING

The main right-hand panel displays a grid of resource monitors for systems R00 through R05. Each system (R000, R010, R020, R030, R040, R050) is represented by a 4x4 grid of sub-monitors. The sub-monitors are labeled as follows:

- R000: R0001, R0002, R0003, R0003-M1, R000-M0, R0001-M0, R0002-M0, R0003-M0
- R010: R0101, R0102, R0103, R0103-M1, R0100-M1, R0101-M1, R0102-M1, R0103-M1, R0100-M0, R0101-M0, R0102-M0, R0103-M0
- R020: R0201, R0202, R0203, R0203-M1, R0200-M1, R0201-M1, R0202-M1, R0203-M1, R0200-M0, R0201-M0, R0202-M0, R0203-M0
- R030: R0301, R0302, R0303, R0303-M1, R0300-M1, R0301-M1, R0302-M1, R0303-M1, R0300-M0, R0301-M0, R0302-M0, R0303-M0
- R040: R0401, R0402, R0403, R0403-M1, R0400-M1, R0401-M1, R0402-M1, R0403-M1, R0400-M0, R0401-M0, R0402-M0, R0403-M0
- R050: R0501, R0502, R0503, R0503-M1, R0500-M1, R0501-M1, R0502-M1, R0503-M1

The grid shows various colored bars (green, black, white) representing the status of each resource monitor.

Usage bars on top level



Usage bars can now summarize the **entire system load**

Further updates

Recent enhancements

- Jobs are searched in table and Nodes View when selected
- LML_da adapter for supporting *Monte Rosa* Cray system at CSCS
→ **SLURM ALPS** combination
- **Level-of-detail** can be chosen **separately** for each connection
- **Double-buffering** enabled for Windows

Future work

- **Layout configuration** via Eclipse client
- Support **multiple node displays** for each connection
(e.g. to visualize power usage, node states, I/O activity)
- **GPU** monitoring