Experiences Using GlideinWMS and the Corral Frontend Across Cyberinfrastructures

Mats Rynge, Gideon Juve, Gaurang Mehta, Ewa Deelman
Information Sciences Institute, University of Southern California

Krista Larson, Burt Holzman
Fermi National Accelerator Laboratory

Igor Sfiligoj, Frank Würthwein
Department of Physics, University of California, San Diego

G. Bruce Berriman
Infrared Processing and Analysis Center, California Institute of Technology

Scott Callaghan
Southern California Earthquake Center, University of Southern California

Funded by the NSF OCI STCI program grant #0943725
Outline

Introduction

GlideinWMS

Corral – a new GlideinWMS Frontend

Experiment Setup

SCEC CyberShake – Example application

IPAC Periodogram – Example application

Conclusions
Motivation

Show that a researcher can bring in and combine local resources and national infrastructures to her/his desktop computer

Local Condor pool,
Open Science Grid,
TeraGrid

glideinWMS with the Corral frontend
# Bringing National Cyberinfrastructure Resources to the Scientist’s Desktop

<table>
<thead>
<tr>
<th><strong>Traditional HPC/HTC</strong></th>
<th><strong>Desktop anchored Virtual Resource</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ssh/scp access</td>
<td>Familiar environment</td>
</tr>
<tr>
<td>Grid interfaces?</td>
<td>Access to local data</td>
</tr>
<tr>
<td>Copy data / log in to head node / set up environment / submit jobs</td>
<td>Output location?</td>
</tr>
<tr>
<td>Using more than one resource? Repeat.</td>
<td>Flexibility</td>
</tr>
<tr>
<td></td>
<td>Running across multiple infrastructures protects the scientist from downtimes, technical site problems, allocation issues, and resource availability</td>
</tr>
</tbody>
</table>
## Infrastructure Differences

<table>
<thead>
<tr>
<th>Open Science Grid</th>
<th>TeraGrid</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Throughput Computing</strong>&lt;br&gt;Serial Codes</td>
<td><strong>High Performance Computing</strong>&lt;br&gt;Parallel Codes</td>
</tr>
<tr>
<td>Virtual Organization mapping (many VO users to one local UID)</td>
<td>Automatically mapped (one VO, individual accounts)</td>
</tr>
<tr>
<td>Opportunistic</td>
<td>Allocations</td>
</tr>
</tbody>
</table>
Condor Based Virtual Clusters

GlideinWMS
Pilot Jobs

Overlay a personal cluster on top of grid resources

Condor based pilots: Glideins
GlideinWMS Overview

- Developed to meet the needs for the CMS (Compact Muon Solenoid) experiment at the LHC (Large Hadron Collider)
- **Frontend** watches job queue for demand
- **Factory** uses grid interface to submit jobs (Condor startds)
- >15,000 concurrent jobs in production, 29 million CPU hours over 2 years
A New GlideinWMS Frontend

Corral
Corral

- Developed by Pegasus Workflow Management System team
- Short jobs
- Mixed HPC/HTC workloads
- Repurposed as a glideinWMS frontend
Frontend (Corral in this case, but could also be the VO Frontend)

Corral queries Condor pool for current workload demand

User Desktop (Condor central manger and queue)

Provisioning request

GlideinWMS Factory

Compute Resource

The Factory provisions glideins on remote resources using Globus GRAM jobs

Glidein registering to Condor Pool

Jobs running on the provisioned glideins
This flexibility allows Corral to acquire a mix of resources with different user/group mappings when running across infrastructures.
Multislot Requests

- Mapping demand from user job queue to a factory request to a single grid job requesting N slots
- Efficiency – grow the pool quickly
- Queue limits – only allowed 7 jobs in the queue
Example Applications

Experiment Setup
Desktop Setup

- Condor central manager
  - Collector – for the glideins to register to
  - Schedd – submit jobs
- X.509 security
- 10 sub collectors
- From the users point of view:

  **Standard Condor pool**
**Abstract Workflows** - Pegasus input workflow description

- Workflow “high-level language”
- Only identifies the computation, devoid of resource descriptions, devoid of data locations

**Pegasus**

- Workflow “compiler” (plan/map)
- Target is DAGMan DAGs and Condor submit files
- **Transforms** the workflow for performance and reliability
- Automatically locates physical locations for both workflow components and data
- Provides runtime provenance
CyberShake

Probabilistic seismic hazard analysis workflow

• How hard will the ground shake in the future?
• Considers a set of possible large earthquakes
• 415,000 earthquakes is typical

Uses Pegasus and Condor DAGMan for workflow management

• Hierarchal workflows
• Small set of large parallel jobs
• 840,000 serial jobs, in 78 sub workflows
Probabilistic Seismic Hazard Analysis (PSHA) curve. Estimates the probability that earthquake ground motions will exceed some intensity measure.

Set of PSHA curves interpolated creates hazard map for an area.
CyberShake DAX

PreCVM

VMeshGen

PreSGT

VMeshMerge

SGTGen_x

SGTGen_y

NanTest

SGT Merge

SGT-X

SGT-Y

SGT DAX Template

Runs on Teragrid

extract

seismogram

PeakSA

ZipSeis

Seis.zip

PeakSA.zip

PP DAX Template

Runs on OSG
A mix of MPI and serial jobs
High-Performance Periodogram Calculations in the Search for Exoplanets

Example Application
Periodograms

- Current dataset: ~600,000 stars
- Calculates the significance of different frequencies in time-series data to identify periodic signals.
- Light curve -> Periodogram -> Event -> Event database
- FFT
- Three different algorithms

BLS periodogram for Kepler -4b, the smallest transiting exoplanet discovered by Kepler to date.
Workflow Details

- 11 sub workflows, ~50000 tasks each
- Wall time based job clustering
  - Simple binning
  - Target: 1 hour
- ~800 jobs per sub workflow
- Execute across available resources:
  Local, Open Science Grid, TeraGrid
Size of Condor Pool
Conclusions
The users were shielded from runtime problems such as preemption

Only discovered when examining logs and graphs
Networking
Glideins not being able to register

Backups
Disk and network I/O affected glideins

Disk Space
800 GB output dataset
Thank you!


**Pegasus:** [http://pegasus.isi.edu](http://pegasus.isi.edu)

**SCEC:** [http://www.scec.org/](http://www.scec.org/)

**IPAC:** [http://www.ipac.caltech.edu/](http://www.ipac.caltech.edu/)