Opal-Sigiri: Software as a Service on PRAGMA Testbed

Yuan Luo and Beth Plale, PhD
School of Informatics and Computing, Indiana University
Data To Insight Center, Indiana University

Presentation/Demo at PRAGMA 20th Workshop, University of Hong Kong, March 3rd 2011
D2I PRAGMA Testbed

• Physical Cluster:
  • A 4-node quad-core Rocks Cluster as Virtual Cluster Host Server
  • Master node: pragma.cs.indiana.edu

• Virtual Clusters:
  – Rocks Virtual Cluster:
    • Front node + 3 compute nodes with SGE installation
  – PRAGMA Virtual Cluster:
    • Front node + 3 compute nodes
    • Front node: pragma-f1.cs.indiana.edu
    • Globus + SGE installation
Problem Addressed

• Legacy codes may be frequently wrapped as Web services to provide remote access to application services that are available across the Internet.
• Complexities in interacting with wide variety of computational resource become a challenge due to non-standard job managers
• Scalability and reliability issue of grid job manager
• Access non-grid resources, e.g., departmental and community clusters
• Access to new platform for scientific job executions, e.g., Windows HPC Clusters
Opal Web Service Toolkit

- Submits jobs to back end resources
- Emphasis on ease of deployment of existing applications as web services.
Sigiri

• A uniform abstraction layer over heterogeneous compute resources.
• Supports multiple Job Specifications:
  – Job Submission Description Language (JSDL)
  – Globus Resource Specification Language (RSL)
• Supports multiple compute resources:
  – Local Linux Machine
  – SGE, PBS, Loadleveler
  – Windows HPC Cluster
  – Amazon EC2, Microsoft Azure
Sigiri Architecture
Sigiri Daemon
Sigiri Job States

1. Client Submits the Job
2. Job Received by Tomcat+Axis2
3. Job Received in to the System
4. Queued in the System
5. Picked up by System Daemon
6. Submitted to Resource Manager
7. First Status Update from RM
8. Job Active
9. Job Done
Opal Architecture (revisit)
Opal-Sigiri Integration
Experiences

• Sigiri is more scalable and reliable than WS-GRAM
• Flexible resource selection

• Too many layers in the framework
  – Standalone Sigiri service co-exist with web service

• Redundancy
  – job status representation
Opal-Sigiri Demo

Opal Service setup at D2I PRAGMA Testbed

http://pragma.cs.indiana.edu:8080/opal2

with SGE on 4-node Rocks Virtual Cluster
• Ongoing Work
  – Work with Jim Williams of IU to connect the D2I PRAGMA nodes to IU's fast switches
  – Extend hierarchical MapReduce* work to multiple PRAGMA clusters; Experiment evaluation of AutoDock based virtual screening.

• Future Work (still in idea stage)
  – Trident for deploying into PRAGMA testbed
  – Distribute LEAD II data catalog as resource on PRAGMA testbed
    • data catalog is web service crawler and indexer of meteorology data
  – Provenance collection of MapReduce on Rocks clusters

Thanks!
Questions?

Yuan Luo yuanluo@indiana.edu
Beth Plale, PhD plale@indiana.edu