A Novel Grid Resource Broker Cum Meta Scheduler

- Asvija B
System Software Development Group,
Centre For Development of Advanced Computing,
Bangalore,
India
Basic Expectations

- To aggregate the heterogeneous resources and manage the workload at a global level.

- To provide unattended, reliable, and efficient execution of jobs on heterogeneous and dynamic resources in a grid.

- To offer a unified interface for managing these loosely coupled resources, maintaining scalability and autonomy.
The Top level Architecture

Web Portal

CLI

Applications

Web Service Implementations for Brokering and Scheduling

Middleware

LRMs

Resources
A Unified Suite

- Brokering Services
- Scheduling Framework
- Policy Engine
- Work Flow Engine
- Job Description Language Engine
Security Features

- Integrates with the GSI
- Fine grained authorization scheme to define the privileges to Users and Administrators at these levels:
  - Grid
  - Site
  - Resource
- Uses the DN for authorization at the Grid Level
- Employs XACML for role based authorization at individual site and resource levels.
Policy Management Engine

- Valid users / user groups
- Resource priorities
- Scheduling classes
- Job-preemption policies
- Usage Quotas on number of Jobs, Nodes, Processors, Wall clock time
- Limits on CPU, Physical Memory, Virtual Memory, Disk utilizations
Job Description Language Engine

- Operating System, Processor Architecture, Speed, Memory
- Temporal, Data dependencies...
- Job Priorities, Job Policy Classes
- State notifications, Actions and Triggers ...
- Job dependencies, Interface definitions ...

Job Description

- Resource Requirements
- Scheduling Requirements
- Job Policy Requirements
- Job Lifetime Management
- Workflow
Job Submission and Status Tracking

- Job Submission to Middleware (Globus Toolkit) - Supporting both WS and Pre – WS components
- Notifications on Status changes with WSN support
- Transparent Co-allocation and Co-reservation of resources through a layered, flexible architecture.
- A wide range of application profiles – Serial, Batch, Array, Parallel, Multi-Jobs and Workflows
Application / Domain Specific Features

- Canned, Extensible Workflows for Application Specific Scheduling.
- Plugin architecture for building customized domain specific workflows.
- Optimizations for known, pre-defined Communication and Storage Patterns.
Scheduling Features

• Best Available Heuristic Match of Heterogeneous resources can be seen as

0-1 Knapsack Problem in Combinatorial Optimization

**Input description:** A set of items $S = \{1, \ldots, n\}$ where item $i$ has size $s_i$ and value $v_i$.

A knapsack capacity $C$.

**Problem description:** Find the subset $S' \subseteq S$ that maximizes the value of $\sum_{i \in S'} v_i$ given that $\sum_{i \in S'} s_i \leq C$; i.e. all the items fit in a knapsack of size $C$. 
Effect of including Network parameter

Turnaround time of Integer Sort Runs

- Run Scheduled by taking network latencies into account
- Run Scheduled without taking network latencies into account

Data Size Set

Time (in secs)

S, W, A, B
Standards Compliance

- Caters to most of the GGF use cases for meta-scheduling
- Security Framework Integrates with the GSI
- Advance Reservation Services as per the GGF Recommendations
- Job description Language in line with the JSDL v1.0 specification
Other popular products in the market

- **Moab** – From Cluster Resources Inc.
  - Feature rich, Separate Policy Engine, Advance Reservations
  - Supports only compute resources – no network and data resources involved in making scheduling decisions

- **GridWay**
  - Seamless Integration with Globus, Adaptive scheduling
  - Comes only with simple greedy round robin scheduling, Trivial support for accounting/logging, Lack of Fine grained policy settings
References

• Grid Brokers and Metaschedulers Market Overview – Gridwisetech.com
  http://www.gridwisetech.com/content/view/110/88/lang,en/

• Gridway – Metascheduling solutions to the grid
  http://www.gridway.org

• Moab – Metascheduler
  www.clusterresources.com/products/mwm
Thank you!

Advanced Computing For Human Advancement