An Open Job Scheduling Service for Large-Scale Data Processing

Zhenghua Xue
zhxue@cnic.cn
Computer Network Information Center (CNIC)
Chinese Academy of Sciences (CAS)
Outline

- Scientific Data Challenge
- Scientific Data Grid (SDG), CAS
- Job Scheduler of SDG (SDGJS)
- Critical Enabling Technologies
- Application Cases
- Done and Doing about Big Data
Data Driven Scientific Discovery

- Data is regarded as the most valuable thing.
Data Driven Scientific Discovery

- Data Intensive Computing (DIC) in China

2011.7 Communications of China Computer Federation
2011.8 Programmer
Scientific Data Challenge

- Big
- Distributed
- Heterogeneous
- Mul-Sources

Share: Data’s shameful neglect

- Difficult to discover
- Difficult to access
- Difficult to deal with
- Being lost
Scientific Data Grid (SDG), CAS

- **Goal of SDG**
  - Integrating isolated scientific data
  - Aggregating distributed computing and storage resources
  - Developing or leveraging simple-using grid middleware
  - Providing rich scientific data service: computing, storage, analyzing and visualization
Organizational Architecture of SDG

A cross-cutting disciplines team above 70 developers on data, computing, storage, and applications
System Architecture of SDG
On-line Sites of SDG
On-line Grid Application Service

Auto-discovery of Sun Activity Data Computing and 3D Visualization on Earth's Atmosphere 3D Visualization of Molecular Structure Integrated Data Access Geographically Distribute
Computing Service of SDG

Computing Service Process

1. User Login
2. User Authentication
3. Query Data
4. Submit Jobs
5. Decompose Tasks
6. Find Computing Resources
7. Transfer Data Queried to Computing Resources
8. Monitor Running Status of Jobs
9. Visualize Computing Results
10. Return Visualized Results
Job Scheduler of SDG

Architecture

User Authentication

Job Control
- start
- suspend
- ... resume
- kill

Job Monitor
- cpu time
- wall time
- ... memory
- sub process

Resource Monitor
- cpu usage
- memory usage
- ... average load
- net traffic

Job Schedule Policy
- FIFO
- HPF
- ... LJF

Web Service

Job Assign

Job && Resource Statistics && Analysis

User Authorization

User Mapping

jobs

results
SDGJS-User Management

● Unified user management: Single Sign on

● Separating authentication from Authorization

● Mapping SDG user to the account of data resources or computing resource

● LDAP based user management in internal computing cluster.
SDGJS-Resource Monitoring

Communication Business Logic

Data Access Tier
- Site Info
- Service Des
- CPU Usage
- Bandwidth
- Run Job Amount

SOA-based Business Logic Tier
- Site Info
- Service Des
- CPU Usage
- Bandwidth
- Run Job Amount

Presentation Tier
- Site Info
- Service Des
- CPU Usage
- Bandwidth
- Run Job Amount

Resource Monitoring
SDGJS-Job Monitoring

Monitoring and Statistic

- Job waiting time
- Job running time
- Consumed CPU
- Consumed Mem
- Accessed data resources

Job Trace
SDGJS-Job Scheduler

- For specific application
  - Goal function
  - $T = T(C) + T(DT)$

- For whole jobs
  - NP hard
  - Simulation by genetic algorithm

Filtering-Evaluation-Determination
SDGJS-Data Transport

Hessian2: Web Services protocol designed for object-oriented transmission

Performance Comparison for Small/Large Object Lists
SDGJS-Data Transport

Globus GridFTP: Large Data Transport

Transport Performance with Various Parallel Streams
Transport Performance with Various TCP Buffer Sizes
Application Case: High Resolution Surface Display

1. Model parameter input
2. Land use data request
3. Dynamic data retrieval for model results
4. Notice of submitted calculation task initiation
5. 90-meter digital elevation data

<table>
<thead>
<tr>
<th>Data Grid Resource Node Cluster</th>
<th>Grid Resource Main Node</th>
<th>Loess Plateau Area</th>
<th>Northeast Black Soil Area</th>
<th>Southwest Mountains Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Grid Calculation Node Cluster</td>
<td>Calculation Node 1</td>
<td></td>
<td>Calculating Node 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Calculating Node 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Calculating Node 4</td>
<td></td>
</tr>
</tbody>
</table>

- China Academy of Sciences
- Data Grid Center
Application Case: Biological Computation

- **Gene Alignment Project**

- **Data:**
  - Microbiology Institute
  - World Data Center for Microorganisms
  - Wuhan Virus Institute

- **Computing:**
  - CNIC
  - Microbiology Institute

- **Adopted Middleware:**
  - Data search
  - Data transport
  - Job scheduler
  - User authentication
Application Case: Biological Computation
Thanks!
Any Questions?