Automatic Certificate Based Account Generation and Secure AJAX calls in a Grid Portal

Mark L. Green, David A. Alexander, Roopa Pundaleekra, James Matykiewicz
Tech-X Corporation (Boulder, CO)
Email: mlgreen@txcorp.com

---------------------
GCE08 Workshop Presentation at SC08
November 17, 2008
• **Small Business Innovative Research Project (SBIR)**
  – Final year of project
  – Main customer DOE-NP, so focus is Open Science Grid (OSG)

• **Collaborating with STAR Computing Group**
  – Use case is STAR nuclear physics experiment
  – C++ analysis framework
    • Handles interface to data (distributed: BNL, LBNL)
    • Has “Logger” to log application-level messages.
  – STAR Unified Meta-Scheduler has high-level expression of analysis task
UCM Primarily Addresses Needs of Solenoidal Tracker at RHIC (STAR)

Experiment....

Computing....

root4star Framework

User Analysis

ROOT

An Object-Oriented Data Analysis Framework

Sub-System Software

Time Projection Chamber

ElectroMagnetic Calorimeter

Triggering

BNL + NERSC + Open Science Grid

(Computing and Storage Resources)
STAR Monitoring Problems Common for Any DOE/OSG Small VO

• **Not enough user feedback to feel comfortable running jobs on the Grid**
  “STAR Experiment Computer Support team say user’s experience too much like throwing jobs over a fence”

• **Varying sources of monitoring information**
  – Tasks submitted with **SUMS** broker (Java)
  – Jobs are **root4star** application (C++)
  – Data fetching wrappers (Python)

• **Varying level of users**
  – Nightly Regression Test & Production Runs
  – Monte Carlo Exploration
  – Individual Analysis
Problem: Aggregating Monitoring Information is Difficult to Get

Single Analysis Task may be many jobs

Grid Scheduler hides Batch Scheduler

Messages not handled directly by Grid

Tasks (many jobs) → Scheduling (site execution) → Application (event messages)

Sometimes user does know number of jobs

Correlating many IDs of jobs and users challenging

Usually don’t get information out until job finishes
Solution: End-To-End Toolset

1. **COLLECTION**
   - Tracking Library

2. **STRUCTURE**
   - Data Store
   - Open-Source Schema for monitoring information
   - Working with CEDPS

3. **PRESENTATION**
   - Web Portal
   - GridSphere Portal
   - Secure with DOE Certificates
   - Uses Google-like Web Application Technology

- Versatile C++ Library
- CCA/Babel for other Languages (F77, Java, Python, etc.)

**Center for Enabling Distributed Petascale Science**
Portal Behaves Like an Application

Idea is to show Task and Job information in dynamic way so user can drill down to job and message details.
Portal Behaves Like an Application

Click on Task
Task Info
Job List
Portal Behaves Like an Application
Portal Behaves Like an Application

Click on Message Link

Message List

Failure Details
Progress: Prototype in place and already of benefit to STAR

- Tracking Library integrated with root4star
- DB in place
- Portal Release to BNL (used with Nightly Tests)
Basic UCM Installation Process

- Make deployment of STAR Portals easy for the administrator
  - Download from https://ice.txcorp.com/trac
  - Install or use existing MySQL server
  - Generate ucm and gridsphere databases with appropriate permissions
  - Edit ucmportal.properties
  - Startup the portal with ucmportal.sh
  - Point your browser to http://<HOST>:<PORT>/gridsphere and complete the online setup for the administrator account
  - Your UCM Portal install is complete
Tech-X Corporation

UCM Database

# ----------------------------------------------
# Portal Server Host Name
#
host.name=cyber.txcorp.com
host.httpport=8080
host.httpsport=8443
host.country=US
host.state=CO
host.city=Boulder
host.organization=Tech-X_Corporation
host.unit=User_Centric_Monitoring

# ----------------------------------------------
# Database Connection Information for UCM DB.
#
ucmdb.name=ucmdb
ucmdb.host=cyber.txcorp.com
ucmdb.port=3306
ucmdb.username=ucmuser
ucmdb.password=strongpassword

# ----------------------------------------------
# Database Connection Information for Gridsphere DB.
#
gsdb.name=gridsphere
gsdb.host=cyber.txcorp.com
gsdb.port=3306
gsdb.username=gsuser
gsdb.password=strongpassword

UCM Tracking Library already required a MySQL database we want to leverage this.
Automate HTTPS Configuration

• What actually happened during the UCM install?
  – From the ucmportal.properties file
    • Using keytool a tomcat.keystore is generated
    • In addition a tomcat.truststore is generated and used to limit only supported Certificate Authorities access
    • The Tomcat conf/server.xml is generated using the ucmportal.properties and stores listed above
    • Files are modified appropriately
  – Both HTTP and HTTPS Portal interfaces are configured
    • HTTP authentication model is username/password
    • HTTPS authentication model is CLIENT-CERT
  – No user accounts or certificates have been generated or stored respectively
Account Generation

• The UCM Portal is configured to access the Virtual Organization Membership Service (VOMS)
  – Accounts are generated when any user accesses the https://<HOST>://<PORT>/gridsphere URI and presents a valid certificate from a trusted Certificate Authority
  – The administrator has the ability to change how frequently the VOMS is queried
  – User accounts will be automatically generated and updated for the UCM Portal

# ----------------------------------------------
# VOMS Connection Information.
# ----------------------------------------------
# TECH-X VO VOMS
voms0.hostcert=/etc/grid-security/hostcert.pem
voms0.hostkey=/etc/grid-security/hostkey.pem
voms0.certtype=PEM
voms0.capath=/etc/grid-security/certificates
voms0.connection=https://cyberrepo.txcorp.com:8443/voms/TECH-X/webui/admin/users/list
Automated Account Generation Control
### Frequency of Account Updates

<table>
<thead>
<tr>
<th>id</th>
<th>last_check</th>
<th>subjectDN</th>
<th>time_increment</th>
<th>next_update</th>
<th>active</th>
<th>stamp</th>
</tr>
</thead>
</table>
VOMS User Mapping

<table>
<thead>
<tr>
<th>id</th>
<th>SubjectDN</th>
<th>Stamp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CN=Alexander Sim 5465822, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>2</td>
<td>CN=Alexandre Suaide 429664, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>3</td>
<td>CN=Dantong Yu 542086, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>4</td>
<td>CN=Douglas L Olson 201098, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>5</td>
<td>CN=Douglas L Olson 916508, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>6</td>
<td>CN=Eric Hjort 89729, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>7</td>
<td>CN=Ian Ballantyne 76180, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>8</td>
<td>CN=Jason A. Smith 236749, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>9</td>
<td>CN=Jerome LAURET 694698, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>10</td>
<td>CN=John Wu 154741, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>11</td>
<td>CN=Lidia Dideniko 827414, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>12</td>
<td>CN=Levente B. Hajdu 105837, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>13</td>
<td>CN=Michael Reuter 965015, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>14</td>
<td>CN=Valeri Fine 224072, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>15</td>
<td>CN=Wayne Betts 602856, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>16</td>
<td>CN=Zhenping (Jane) Liu 607345, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>17</td>
<td>CN=Iwona Sakrejda 40474, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>18</td>
<td>CN=Lee Barnby 268393, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>19</td>
<td>CN=Lidia Dideniko (starreco) 639758, OU=People, DC=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>20</td>
<td>CN=Junmin Gu 111597, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>21</td>
<td>CN=John R. Hover 47116, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>22</td>
<td>CN=Pavel Jakl 81798, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>23</td>
<td>CN=Alexander Withers 351937, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>24</td>
<td>CN=R. Jefferson Porter 227760, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>25</td>
<td>CN=Adam Kokocioski 387979, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>26</td>
<td>CN=Andrew Allen Rose 666335, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>27</td>
<td>CN=Michel Zerola 793743, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>28</td>
<td>CN=Alexandre SHABETA 368093, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>29</td>
<td>CN=Vladimir W. Krukov 457486, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
<tr>
<td>30</td>
<td>CN=Vladimir W. Krukov 457486, OU=People, DC=doegrids, D=...</td>
<td>2008-08-13 15:31:20</td>
</tr>
</tbody>
</table>
Gridsphere User Mapping
Entire UCM Authentication System

1. User Browser Connecting to UCM Portal → HTTPS
2. Obtain Certificate’s Identity → YES
   → Validate Certificate’s Root CA → VALID
   → Update Database Accounts? → YES
   → Generate Member Lists
   → Allowed CA’s
   → VOMS LDAP XML ASCII
3. Validate Certificate’s Identity → NOT VALID
   → Generate Member Lists
   → Not Valid
4. Fail-Over to GridSphere Login Portlet → FAIL
5. Update Database Accounts? → NO
6. Generate Member Lists
7. Manage GridSphere Accounts
8. GridSphere Database
9. Verify Subject DN Mapping → CHECK
10. Credential Database
11. GridSphere Login → PASS
12. Main UCM Portal
DOE Certificate Security in Place

User prompted to choice certificate to present to portal

Portal automatically recognizes user with out additional account password
Portal Behaves Like an Application

Idea is to show Task and Job information in dynamic way so user can drill down to job and message details.
AJAX Communication to UCM Database

Browser

Tomcat Container

Portlet

JSPServlet

shared secret

JSP page

MySQL calls

request for db info

request for page

inside page

JavaScript Code (AJAX)

doView()

returns fragment

references

JavaScript code

database

Portal

Servlet
Servlet Security almost for free!

• **Since we have the automated authentication mechanism in place already**
  – Generate a hash for Servlet security at account generation time
  – Use it for securing Servlet access
  – Only needed to add one database field and algorithm to generate hash
A Portal Solution is Extensible

- Login
- Monitoring
- CEDPS
- Vine Toolkit
- Performance

Client Browser
Portal Tab
  - Portlet Window A
  - Portlet Window B
  - Portlet Window C
  - Portlet Window D

Portal Server
  - Portlet A
  - Portlet B
  - Portlet C
  - Portlet D

Portlet Container

response

vine:toolkit
open-source java grid application framework
More Integration and Scaling to Come

• **More Integration with CEDPS**
  – Connect UCM to Troubleshooting data (job ID)

• **Scaling Tests**
  – Searches are fast enough now, but won’t be with large number of users
  – Refactoring of Tracking Library
  – Redesign of database will be needed

• **More use at BNL**
  – From nightly tests to production runs
  – Look ahead to individual users

• **Integration with SUMS**
  – Already have Java version of library (Babel)
UCM Publications and Wiki

• David Alexander, Chuang Li, Jerome Lauret, Valeri Fine, “User Centric Monitoring (UCM) information service for the next generation of Grid-enabled scientists”, International Conference on Computing in High Energy and Nuclear Physics, 2-7 Sep 2007, Victoria, BC, Canada. Presented as poster and a contributed paper


• Project documentation, releases, issue tracking, & milestones at:

  https://ice.txcorp.com/trac/ucm