Creating easy to use HPC portals with NICE EnginFrame and SLURM

Andrea.Rodolico@nice-software.com
Alberto.Falzone@nice-software.com
Paolo.Maggi@nice-software.com
Summary

- About NICE
- Introduction to NICE EnginFrame
- Why a SLURM plugin for EnginFrame?
- The SLURM plugin and its current limitations
- Q&A
About NICE

### Company
- Focus on technical computing since 1996
- HQ in Italy (offices in USA, GER, UK)
- Partners all around the world
- Always profitable and self-funded

### Expertise
- Industry veterans around Grid & HPC solutions
- Vertical solutions, Cloud computing, Remote Visualization

### Core business: Access to Grid / HPC / Cloud solutions
- Work, visualize, and collaborate in HPC
- EnginFrame and DCV product families
# NICE Customers and Market Segments

<table>
<thead>
<tr>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadarko, AECL, Hess, Bayerngas, BHP</td>
</tr>
<tr>
<td>Billiton, Beicip, British Gas, Centrica,</td>
</tr>
<tr>
<td>Chevron, Conoco-Phillips, Dong, Dowell,</td>
</tr>
<tr>
<td>DSC-Libya, ENI/Agip, GazPromNeft, GDF,</td>
</tr>
<tr>
<td>Logelco, Maersk Oil, Marathon Oil, Nexen,</td>
</tr>
<tr>
<td>National Oilwell Varco, Novatek, Papuan Oil,</td>
</tr>
<tr>
<td>PetroChina, Rosneft, Schlumberger,</td>
</tr>
<tr>
<td>Sinopec, Sonatrach, Statoil, Talisman</td>
</tr>
<tr>
<td>Energy, TNK-BP, TNNC, TOTAL, WG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aerospace &amp; Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRBUS, Air Products and Chemicals, AVIC,</td>
</tr>
<tr>
<td>Procter&amp;Gamble, SelexGalileo, Goodrich</td>
</tr>
<tr>
<td>Aerospace, Kimberly Clarke, Magellan</td>
</tr>
<tr>
<td>Aerospace, NORDAM, Northrop Grumman,</td>
</tr>
<tr>
<td>Raytheon, Sikorsky, Thales</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Automotive &amp; Industrial Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M, ABB, Altran, Audi, ARRK, BMW,</td>
</tr>
<tr>
<td>Bridgestone, Bosch, Continental, Daimler,</td>
</tr>
<tr>
<td>Delphi, Dow, Faurecia, Ferrari, Hyundai,</td>
</tr>
<tr>
<td>JLR, Lear, Magneti Marelli, McLaren, PSA,</td>
</tr>
<tr>
<td>RedBull, Tata Steel, Toyota, TRW, VW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Life Sciences and Medical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baxter, Bayer, Biolab, DEISA project, HHMI,</td>
</tr>
<tr>
<td>Johnson&amp;Johnson, Novartis, SIB, Partners</td>
</tr>
<tr>
<td>Healthcare, Pharsight</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accent, Samsung SDI, SensorDynamics,</td>
</tr>
<tr>
<td>Bank of Italy, Deutsche Bank</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research &amp; Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beihang U, Birmingham U, Buffalo U,</td>
</tr>
<tr>
<td>CILEA, Georgia State U, INFN, Harvard U,</td>
</tr>
<tr>
<td>Liverpool U, Messina U, Huazhong Normal U,</td>
</tr>
<tr>
<td>TU Ilmenau, Yale U</td>
</tr>
</tbody>
</table>

© Copyright NICE srl, 2013
What is NICE EnginFrame?

Enterprise Grid
Desktop Scavenging
HPC App. Portal
Commercia HPC ASP
HPC SaaS
Collaborative design
Visualization farm
Open Grid ASP
HPC Clusters

© Copyright NICE srl, 2013
EnginFrame Key Features

- User friendly Web based access to technical computing applications (batch and interactive)
- Flexible service offering for end users
  - Application-oriented HPC job submission and monitoring
  - Access to 2D / 3D remote desktops
- Support all major HPC schedulers: LSF, Grid Engine, PBS, Torque, Moab, OpenLava.... and SLURM
- Multiple remote display protocols support: RealVNC, Tiger/Turbo/TightVNC, NICE DCV, HP RGS, ...
- Data management
- Flexible authentication delegation (NIS, LDAP, AD, Kerberos, …)
- Fine grained authorization system
- Accounting and monitoring of resource usage
Batch Job / Workflow Submission

- User friendly, Application-oriented Job submission
- Hide complexity of Underlying scheduler
- Flexible and efficient Input file management
Monitoring

Jobs, Hosts, Queues, Licenses, ...
Web-based Interactive Session Management
Data Transfers & File Management

The file manager component allows to seamless navigate and access server-side files from the web browser.
Job Data Management

Application data can be organized into projects

Application data can be marked as starred

Metadata can be associated to application data
CAE / HPC Workflow Integration

![Diagram of NICE software interface](image)

### CAE Services
- **CFD Services**
  - Fluent
  - PowerFlow
- **Crash services**
  - LS-Dyna
  - PamCrash
  - Radioss
- **FEA services**
  - Abaqus
  - Nastran

### Desktop Sessions
- **Linux Session (2D/VNC)**
- **Windows Session (2D/VNC)**
- **Linux Session 3D**
- **Windows Session (3D/DX)**
- **Interactive Applications**
- **Create Location**

### Post-processing with CEI EnLiten
- **Delete**
- **Properties**
- **Dependencies**
- **Permissions**
- **Design Points**
- **Command Window**
- **Simulation Details**
- **Refresh**

#### Final crash

### Data
- **Upload**
- **Download**
- **Create Folder**
- **Delete**
- **Compress**

#### Location:

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Date Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>binout0000</td>
<td>126 KB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>car.es</td>
<td>5.93 MB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>d3dump01.0000</td>
<td>251 KB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>d3dump01.0001</td>
<td>22 KB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>d3full01</td>
<td>126 KB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>d3plot</td>
<td>126 KB</td>
<td>Today 22:28</td>
</tr>
<tr>
<td>d3plot01</td>
<td>126 KB</td>
<td>Today 22:28</td>
</tr>
</tbody>
</table>

### Jobs
- **Kill**
- **Suspend**
- **Resume**

<table>
<thead>
<tr>
<th>ID</th>
<th>Status</th>
<th>Queue</th>
<th>Submission Time</th>
<th>Running on</th>
<th>Job Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1297</td>
<td>Done</td>
<td>normal</td>
<td>Today 22:28:00</td>
<td>quad</td>
<td>Final crash</td>
</tr>
</tbody>
</table>

*Image and text content from NICE srl, 2013. Copyright belongs to NICE srl, 2013.*
HPC + Visualization Workflows
Multiple Clusters/Schedulers

- Submit/monitor jobs on multiple HPC schedulers through a single EnginFrame instance

Thanks to the EnginFrame modular architecture, support for new HPC schedulers can be added with no changes to the EnginFrame core.
Why a SLURM plugin for EnginFrame?

- In the last year we observed a growing number of requests related to our capability to support SLURM (both from academy and industry)
- SLURM popularity seems to grow more and more
- There is an active and dedicated community
- We are getting very positive feedbacks about SLURM
- We have internal persons that like to play with new technologies
- We found a customer that got us access to its SLURM cluster and help us for testing activities (many thanks to the guys from Buffalo University)
- We all love David ;)

Development started!
Software Stack

- NIS / PAM / LDAP / AD
- NICE EnginFrame (Grid Portal & WS Gateway)
  - Custom plugins
  - ...core plugins...
  - GRID plugins...
    - SLURM
    - LSF
    - ...
- SLURM 2.5.7 or above (Job Scheduler & Resource Manager)
- Batch Applications
  - MPI, ...
- 2D/3D Interactive Applications
  - DCV, RGS, ...

Computing Infrastructure

Storage Infrastructure
The SLURM Plugin

- The SLURM plugin has been developed as a backend for the EnginFrame GRID plugin
- Interaction with SLURM happens at the command line level:
  - SLURM CLI commands are invoked (\texttt{sbatch}, \texttt{scontrol}, ...)
  - The output of the command is parsed and translated into XML documents (GridML)
- Feature parity with all the other major supported schedulers:
  - HPC job submission and monitoring
  - Resource monitoring
  - Submission and managing of interactive sessions
Example: Job Details (scontrol output)

$ scontrol show jobs 351395
JobId=351395 Name=jB11c16
  UserId=ajs42(142404) GroupId=kofke(45550)
  Priority=52967 Account=kofke QOS=normal
  JobState=COMPLETED Reason=None Dependency=(null)
  Requeue=0 Restarts=0 BatchFlag=1 ExitCode=0:0
  RunTime=1-08:27:18 TimeLimit=1-16:00:00 TimeMin=N/A
  StartTime=2013-09-09T05:44:08 EndTime=2013-09-10T14:11:26
  PreemptTime=None SuspendTime=None SecsPreSuspend=0
  Partition=general-compute AllocNode:Sid=f07n05:62117
  ReqNodeList=(null) ExcNodeList=(null)
  NodeList=k07n27
  BatchHost=k07n27
  NumNodes=1 NumCPUs=16 CPUs/Task=1 ReqS:C:T=::*:*:*
  MinCPUsNode=16 MinMemoryCPU=3000M MinTmpDiskNode=0
  Features=(null) Gres=(null) Reservation=(null)
  Shared=OK Contiguous=0 Licenses=(null) Network=(null)
  Command=/ifs/projects/kofke/ajs42/virial/hs/B11/jB11c16
  WorkDir=/ifs/projects/kofke/ajs42/virial/hs/B11
Example: Job Details (GridML)

```xml
<grid:job-list type="slurm" filtered="true" sorted="true" paginated="true" filter="" sort-by="" max-results="1" start-index="1" results="1" total-results="1">
  <grid:job id="351395" type="slurm">
    <grid:name>jB11c16</grid:name>
    <grid:owner>ajs42</grid:owner>
    <grid:account>kofke</grid:account>
    <grid:status ef="Running" grid="RUNNING">RUN</grid:status>
    <grid:total-cpu-usage>1-08:15:06</grid:total-cpu-usage>
    <grid:submission-time>2013-09-01T17:40:20</grid:submission-time>
    <grid:execution-time>2013-09-09T05:44:08</grid:execution-time>
    <grid:queue>general-compute</grid:queue>
    <grid:execution-host>k07n27</grid:execution-host>
    <grid:parallel max="16" min="16"/>
    <grid:swap-usage>0</grid:swap-usage>
  </grid:job>
</grid:job-list>
```
Example: Job Details

Known limitation: some job details not available at writing time

The GridML is rendered
Example: All Jobs
Example: Cluster Load
Example: Host Details

**Known limitations:** free memory value and others not available

<table>
<thead>
<tr>
<th>Status</th>
<th>ok</th>
<th>ok (base)</th>
</tr>
</thead>
</table>

| Job Slots       | 0/1 |           |
| CPU utilization | 0%  |           |
| Physical memory | 0/512 |          |
| Free tmp space  | 2048 MB |         |

<table>
<thead>
<tr>
<th>Hardware</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Host type</td>
<td>x86_64</td>
</tr>
<tr>
<td>Number of processors</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resources</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>vnc</td>
<td>vnc</td>
</tr>
<tr>
<td>dcv</td>
<td>dcv</td>
</tr>
</tbody>
</table>
### Host Details: Limitations

#### Comparison of retrieved informations  - What is missing

**[root@slurm01 ~]# scontrol show node slurm02**

NodeName=slurm02  Arch=x86_64  CoresPerSocket=1  
CPUAlloc=1  CPUErr=0  CPUTot=1  CPULoad=0.09  
Features=vnc,dcv  
Gres=mem:256  
NodeAddr=slurm02  NodeHostName=slurm02  
OS=Linux  RealMemory=512  AllocMem=0  Sockets=1  Boards=1  
State=ALLOCATED  ThreadsPerCore=1  TmpDisk=2048  Weight=1  
BootTime=2013-08-01T23:37:11  
SlurmdStartTime=2013-08-06T13:24:56  
CurrentWatts=0  LowestJoules=0  ConsumedJoules=0  
ExtSensorsJoules=n/s  ExtSensorsWatts=0

---

**[root@slurm01 ~]# scontrol show node slurm02**

NodeName=slurm02  Arch=x86_64  CoresPerSocket=1  
CPUAlloc=1  CPUErr=0  CPUTot=1  CPULoad=0.09  
Features=vnc,dcv  
Gres=mem:256  
NodeAddr=slurm02  NodeHostName=slurm02  
OS=Linux  RealMemory=512  AllocMem=0  Sockets=1  Boards=1  
State=ALLOCATED  ThreadsPerCore=1  TmpDisk=2048  Weight=1  
BootTime=2013-08-01T23:37:11  
SlurmdStartTime=2013-08-06T13:24:56  
CurrentWatts=0  LowestJoules=0  ConsumedJoules=0  
ExtSensorsJoules=n/s  ExtSensorsWatts=0

---

<table>
<thead>
<tr>
<th>Status</th>
<th>Host lsf706linux</th>
<th>Host slurm02</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Load</strong></td>
<td><strong>ok</strong></td>
<td><strong>busy</strong></td>
</tr>
<tr>
<td><strong>ok</strong></td>
<td><strong>ok (base)</strong></td>
<td><strong>busy (base)</strong></td>
</tr>
<tr>
<td><strong>CPU run queue length:</strong></td>
<td>1.6</td>
<td>1/1</td>
</tr>
<tr>
<td>Averaged 15 sec (r15s)</td>
<td>1.6</td>
<td>1/1</td>
</tr>
<tr>
<td>Averaged one minute (r1m)</td>
<td>0.0</td>
<td>1%</td>
</tr>
<tr>
<td>Averaged 15 minutes (r15m)</td>
<td>0.1</td>
<td>1%</td>
</tr>
<tr>
<td>Physical memory</td>
<td>792MB/950MB</td>
<td>0/512MB</td>
</tr>
<tr>
<td>Swap space</td>
<td>68MB/945MB</td>
<td>2048 MB</td>
</tr>
<tr>
<td>Free tmp space</td>
<td>13.97GB</td>
<td></td>
</tr>
<tr>
<td>Paging rate</td>
<td>0.2 pages/sec</td>
<td></td>
</tr>
<tr>
<td>I/O throughput</td>
<td>56 Kb/sec</td>
<td></td>
</tr>
<tr>
<td>Idle time</td>
<td>1035 minutes</td>
<td></td>
</tr>
</tbody>
</table>

---

**[root@slurm01 ~]# scontrol show node slurm02**

NodeName=slurm02  Arch=x86_64  CoresPerSocket=1  
CPUAlloc=1  CPUErr=0  CPUTot=1  CPULoad=0.09  
Features=vnc,dcv  
Gres=mem:256  
NodeAddr=slurm02  NodeHostName=slurm02  
OS=Linux  RealMemory=512  AllocMem=0  Sockets=1  Boards=1  
State=ALLOCATED  ThreadsPerCore=1  TmpDisk=2048  Weight=1  
BootTime=2013-08-01T23:37:11  
SlurmdStartTime=2013-08-06T13:24:56  
CurrentWatts=0  LowestJoules=0  ConsumedJoules=0  
ExtSensorsJoules=n/s  ExtSensorsWatts=0

---

<table>
<thead>
<tr>
<th>Status</th>
<th>Host lsf706linux</th>
<th>Host slurm02</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Load</strong></td>
<td><strong>ok</strong></td>
<td><strong>busy</strong></td>
</tr>
<tr>
<td><strong>ok</strong></td>
<td><strong>ok (base)</strong></td>
<td><strong>busy (base)</strong></td>
</tr>
<tr>
<td><strong>CPU run queue length:</strong></td>
<td>1.6</td>
<td>1/1</td>
</tr>
<tr>
<td>Averaged 15 sec (r15s)</td>
<td>1.6</td>
<td>1/1</td>
</tr>
<tr>
<td>Averaged one minute (r1m)</td>
<td>0.0</td>
<td>1%</td>
</tr>
<tr>
<td>Averaged 15 minutes (r15m)</td>
<td>0.1</td>
<td>1%</td>
</tr>
<tr>
<td>Physical memory</td>
<td>792MB/950MB</td>
<td>0/512MB</td>
</tr>
<tr>
<td>Swap space</td>
<td>68MB/945MB</td>
<td>2048 MB</td>
</tr>
<tr>
<td>Free tmp space</td>
<td>13.97GB</td>
<td></td>
</tr>
<tr>
<td>Paging rate</td>
<td>0.2 pages/sec</td>
<td></td>
</tr>
<tr>
<td>I/O throughput</td>
<td>56 Kb/sec</td>
<td></td>
</tr>
<tr>
<td>Idle time</td>
<td>1035 minutes</td>
<td></td>
</tr>
</tbody>
</table>

---

**[root@slurm01 ~]# scontrol show node slurm02**

NodeName=slurm02  Arch=x86_64  CoresPerSocket=1  
CPUAlloc=1  CPUErr=0  CPUTot=1  CPULoad=0.09  
Features=vnc,dcv  
Gres=mem:256  
NodeAddr=slurm02  NodeHostName=slurm02  
OS=Linux  RealMemory=512  AllocMem=0  Sockets=1  Boards=1  
State=ALLOCATED  ThreadsPerCore=1  TmpDisk=2048  Weight=1  
BootTime=2013-08-01T23:37:11  
SlurmdStartTime=2013-08-06T13:24:56  
CurrentWatts=0  LowestJoules=0  ConsumedJoules=0  
ExtSensorsJoules=n/s  ExtSensorsWatts=0

---

**Hardware**

<table>
<thead>
<tr>
<th>Host type</th>
<th>X86_64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host model</td>
<td>Intel_EM64T</td>
</tr>
<tr>
<td>Number of processors</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Hardware**

<table>
<thead>
<tr>
<th>Host type</th>
<th>x86_64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of processors</td>
<td>1</td>
</tr>
</tbody>
</table>
Example: Interactive Sessions

Connect to interactive session on mouse click

Sample analyses of CO₂ at MLO
Example: Session Details

Test VNC SLURM

Owner
Project
Status

Session
Remote Host
Cluster
Operating System
Creation Time
Size
Color Depth
Remotization Protocol
slurm02
linux
Linux
Aug 02, 2013 17:46:35
1280x800
16M colors
Virtual Network Computing (VNC)

Sharing
Collaborators
Viewers

Job
Job ID
Job Manager
498
slurm
Conclusions

- NICE EnginFrame can now be used to create easy to use technical computing / HPC portals for your SLURM-based computing infrastructure
- SLURM 2.5.7 or higher version is required
- Future work:
  - How to retrieve the missing informations for hosts and jobs is under evaluation
  - MPI support to be tested in deep
  - Support for previous SLURM version under evaluation

Thanks to Doris Sajdak, Martins Innus and the other great guys @ Buffalo University for their active collaboration during the testing activities of the SLURM Plugin