Migrating Engineering Windows HPC applications to Linux HTCondor and SLURM Clusters

CHEP 2019
Maria Alandes Pradillo, Nils Høimyr, Pablo Llopis Sanmillan, Markus Tapani Jylhankangas – CERN IT Department
Engineering Use Cases
Background

• Two HPC services used to coexist in the past in CERN IT
  • Windows HPC service managed by IT-CDA
  • Linux clusters in HTCondor and SLURM managed by IT-CM
• IT Strategy is to consolidate compute intensive jobs in Linux
  • Making a better use of resources
  • Moreover…
    • This is also in line with CERN IT strategy to reduce its dependencies on MS products
Windows HPC Utilisation

Monthly Cluster Utilization Comparison

Utilization on available cores vs. Utilization on all cores

Month:
- June 2017
- July 2017
- August 2017
- September 2017
- October 2017
- November 2017
- December 2017
- January 2018
- February 2018
- March 2018
- April 2018
- May 2018
Some other constraints

- End of warranty period for Windows HPC HW in 2019
  - Legal warranty period for systems in the CERN Data Center is 3 years; extendable to 5 years
  - Windows HPC machines were around 5 years
  - Many nodes were already giving memory errors since August 2018
- Departure of Windows HPC service manager
Migration Plan COMSOL and CST

- 13 COMSOL users
- 15 CST users
Migration Plan Ansys

2018
- Sep
- Oct
- Nov
- Dec
2019
- 2019

<table>
<thead>
<tr>
<th>Phase</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation Phase (Installation, Configuration, Documentation)</td>
<td>9/5/2018 - 12/21/2018</td>
</tr>
<tr>
<td>First tests with users</td>
<td>9/28/2018 - 12/21/2018</td>
</tr>
<tr>
<td>Early Adopters Phase</td>
<td>11/27/2018 - 1/31/2019</td>
</tr>
<tr>
<td>Migration Phase</td>
<td>1/7/2019 - 2/28/2019</td>
</tr>
</tbody>
</table>

Ansys Windows HPC users per department (according to e-group ansys-rsm-users)

Active Ansys Windows HP users per department in 2018
Challenges 1

- Lack of Linux expertise among engineers
  - Steep learning curve
    - A lot of effort in user support going to solve “easy” Linux issues
  - Big change for Windows HPC users
  - Big effort to prepare tutorials and detailed documentation

KB0005869: Preliminary steps to run Engineering Application on Linux Clusters

The following instructions explain how to get access to the Linux clusters to run simulations for the following engineering applications:

- Ansys
- COMSOL
- CST
- LS-DYNA

IMPORTANT: Note that basic Linux knowledge is required except for those Ansys users who will work from RSM. All the other users, please, make sure you are able to do basic file management operations via the command line, like creating a new folder, copying files, etc. Also, editing files is needed, so please, learn some basics for tools like vi or any other Linux editor.

Register to LXPLUS
In order to work with Linux clusters, you need to have access to LXPLUS. Please, follow the instructions in this KB to know how to activate this.
You can log in to LXPLUS from a Windows machine using a program like Putty (available in CMF).

Register to EOS
When working with Linux, your simulation and results are stored in EOS. Note that for the time being, Ansys RSM users can’t work with EOS, so you can skip this step.
Infrastructure overview

User Windows PC
Ansys / RSM

WinSCP
Putty

Prepare the simulation

Input file

AFS

script
submission file

Lxplus

Linux Cluster

EOS

Local Disk
Output file

Run the simulation
Challenges 2

- Strong collaboration among heterogeneous groups to succeed
  - Knowledge and expertise wide spread across departments and groups
  - IT support involves storage, computing and application
  - Application expertise is sitting with the engineers
  - Performance tuning requires several iterations among IT experts and engineers
Challenges 3

• Lack of HTCondor and SLURM support in CST, COMSOL and Ansys
  • Difficult to debug technical issues with vendor support teams at the beginning
  • Not possible yet to submit jobs from Windows GUI for COMSOL and CST
• For Ansys
  • HTCondor plugin developed in-house for RSM submission
    • Still gaining experience and understanding some performance issues
  • Interaction with SLURM via PBS commands
Ansys plugin for HTCondor

Project ID: 52873

HTCondor plugin for Ansys RSM.

Auto DevOps
It will automatically build, test, and deploy your application based on a predefined CI/CD configuration.

Learn more in the Auto DevOps documentation

Enable in settings

- master / ansys-rsm-htcondor

fix method call
Pablo Llopis Sanmillan authored 1 month ago

Name | Last commit | Last update
--- | --- | ---
scripts | fix method call | 1 month ago
xml | Use +WCKey=Ansys | 4 months ago
Conclusions

• Consolidating computing resources in a single infrastructure has proven to be a successful strategy:
  • Allows to share existing resources
  • Uses powerful hardware in an efficient manner
  • Benefits from a team of IT experts

• A good collaboration among engineers and IT was key to ensure a smooth migration

• Detailed documentation and clear procedures are critical to let engineers concentrate in their simulations

• Further work on batch submission from Windows for COMSOL and CST, and better performance for Ansys RSM are still needed
  • To allow engineers work from Windows
  • To benefit the most from the existing computing capacity
Acknowledgements

• Markus Tapani Jylhankangas, our technical student, who did a great job and wrote the HTCondor plugin for Ansys RSM

• CERN engineers for their patience and cooperation as early adopters:
  • Rui Vizinho
  • Agnieszka Chmielinska
  • Ricardo Filipe Pereira
  • Eduardo Cano
  • Alessandro Cattabiani
  • Alexandre Amorim
Extra slides
## Linux HPC

<table>
<thead>
<tr>
<th></th>
<th>Cores</th>
<th>RAM</th>
<th>OS</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTCondor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 cores</td>
<td>16 GB</td>
<td>CC7</td>
<td>Open access</td>
</tr>
<tr>
<td></td>
<td>16/32 cores</td>
<td>500 GB</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24 cores</td>
<td>1 TB</td>
<td></td>
<td>Big Memory nodes. Access needs to be granted</td>
</tr>
<tr>
<td></td>
<td>48 cores</td>
<td>500 GB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLURM</td>
<td>16 cores with HT</td>
<td>128 GB</td>
<td></td>
<td>Low-latency 10Gbit Ethernet</td>
</tr>
<tr>
<td></td>
<td>20 cores</td>
<td>128 GB</td>
<td></td>
<td>Infiniband. Access needs to be granted</td>
</tr>
</tbody>
</table>

**Cores**
- HTCondor: 8, 16, 24, 48
- SLURM: 16, 20

**RAM**
- HTCondor: 16 GB, 500 GB, 1 TB
- SLURM: 128 GB

**OS**
- HTCondor: CC7
- SLURM: HT

**Comments**
- HTCondor: Open access
- SLURM: Big Memory nodes. Access needs to be granted
- SLURM: Low-latency 10Gbit Ethernet
- SLURM: Infiniband. Access needs to be granted
Stats in 2019

Total number of HTCondor jobs

- CST: 3000
- Ansys: 2000
- NA: 1000
- Comsol: 500

Values:
- Sum of CPU
- Sum of WallTime

1E+09
900000000
800000000
700000000
600000000
500000000
400000000
300000000
200000000
100000000
0

- Ansys
- Comsol
- CST
- NA

CERN

CHEP - November 2019
Windows to Linux HPC migration
Windows HPC

Data Centre 1st floor

Hyper-V VMs
- EHPCHD1 Head Node CAESRV1008
- EHPCFE01 Frontend CAESRV1005
- RSM170 (RSM 17.1) CAESRV1006
- RSM172 CAESRV1004
- RSM191 CAESRV1003

Data Centre Basement

Quanta Physical = WHPCCN0x
Default Compute Node = CAESRV10xx

<table>
<thead>
<tr>
<th>Node Template</th>
<th>Cores</th>
<th>Sockets</th>
<th>Memory</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quanta Physical</td>
<td>16</td>
<td>2</td>
<td>128</td>
<td>61</td>
</tr>
<tr>
<td>Default Compute Node</td>
<td>32</td>
<td>4</td>
<td>512</td>
<td>8</td>
</tr>
</tbody>
</table>
Windows HPC Statistics per job template 2018