

# Migrating Engineering Windows HPC applications to Linux HTCondor and Slurm Clusters

*Maria Alandes Pradillo*<sup>1</sup>, *Nils Høymyr*<sup>1</sup>, *Pablo Llopis Sanmillan*<sup>1</sup>, and *Markus Tapani Jylhänkangas*<sup>1,\*</sup>

<sup>1</sup>European Organization for Nuclear Research (CERN)

**Abstract.** The CERN IT department has been maintaining different High Performance Computing (HPC) services over the past five years. While the bulk of computing facilities at CERN are running under Linux, a Windows cluster was dedicated for engineering simulations and analysis related to accelerator technology development. The Windows cluster consisted of machines with powerful CPUs, big memory, and a low-latency interconnect. The Linux cluster resources are accessible through HTCondor, and are used for general purpose parallel but single-node type jobs, providing computing power to the CERN experiments and departments for tasks such as physics event reconstruction, data analysis, and simulation. For HPC workloads that require multi-node parallel environments for Message Passing Interface (MPI) based programs, there is another Linux-based HPC service that is comprised of several clusters running under the Slurm batch system, and consist of powerful hardware with low-latency interconnects.

In 2018, it was decided to consolidate compute intensive jobs in Linux to make a better use of the existing resources. Moreover, this was also in line with CERN IT strategy to reduce its dependencies on Microsoft products. This paper focuses on the migration of Ansys [1], COMSOL [2] and CST [3] users from Windows HPC to Linux clusters. Ansys, COMSOL and CST are three engineering applications used at CERN for different domains, like multiphysics simulations and electromagnetic field problems. Users of these applications are in different departments, with different needs and levels of expertise. In most cases, the users have no prior knowledge of Linux. The paper will present the technical strategy to allow the engineering users to submit their simulations to the appropriate Linux cluster, depending on their simulation requirements. We also describe the technical solution to integrate their Windows workstations in order from them to be able to submit to Linux clusters. Finally, we discuss the challenges and lessons learnt during the migration.

## 1 Introduction

For historical reasons, two HPC services used to coexist in the past in the CERN IT department, one based on Windows[4] and the other one based on Linux. In order to make a better use of the resources, the department decided to consolidate all compute intensive tasks in Linux-based clusters. At the same time, the department was also interested in reducing its

---

\*e-mail: [maria.alandes.pradillo@cern.ch](mailto:maria.alandes.pradillo@cern.ch)

dependencies on Microsoft products [5], while the Windows HPC cluster was based on the Microsoft HPC Pack [6]. This meant that all the users of the Windows HPC cluster needed to be migrated to the Linux clusters. The Windows HPC cluster was used to run engineering simulations of Ansys, COMSOL and CST applications. Windows was the platform of choice by the majority of the engineering users who were used to submit their simulations through the Microsoft HPC client. Understanding how to install the engineering applications in Linux and exposing users to Linux were some of the challenges of the migration.

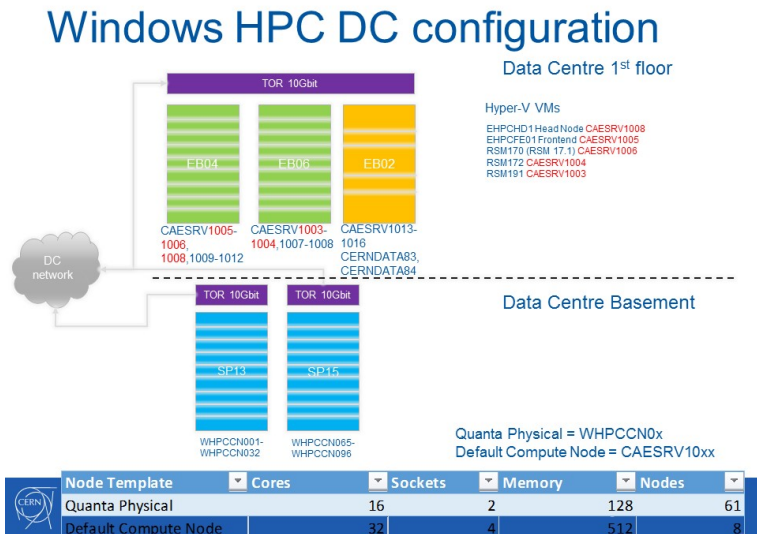
The paper is organised as follows: in Section 2, the Windows and Linux HPC clusters are presented. In Section 3 we detail the engineering use cases. This Section also gives an overview of how the migration process was organised. Section 5 focuses on the communication and training aspects of the migration. Finally, Section 6 summarises the conclusions and lessons learnt.

## 2 HPC Clusters at CERN

At CERN, the main computing challenges are related to High Energy Physics (HEP) computing, but there are also many other computing service needs in the laboratory. For the former, the World LHC Computing Grid [7], as well as local Batch computing service have been set up for efficient High Throughput Computing (HTC), while for some specific application domains, such as accelerator physics, engineering and Theory Lattice-QCD studies, dedicated HPC clusters are used.

### 2.1 Windows HPC Cluster

The CERN Windows HPC service is presented in the figure below.

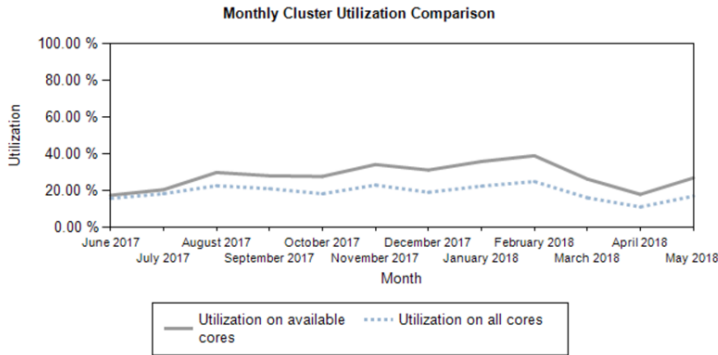


**Figure 1.** Windows HPC cluster at the CERN Data Center

The cluster was comprised of 60 nodes of 16 cores and 128 GB RAM; and 8 nodes of 32 cores and 512 GB RAM. A scratch space was also provided to the users on independent data

servers using the DFS file system. For Ansys simulations, the cluster also provided the Ansys RSM component, that acts as a middle layer between the Ansys user interface and the batch system. The software running on the cluster to manage all these resources was Microsoft HPC Pack 2012 R2.

In order to understand how much the resources were being used, some monitoring statistics were extracted from the cluster head node. Figure 2 shows the Windows HPC resources utilisation in 2018. The Windows HPC cluster was run with full node scheduling.



**Figure 2.** Windows HPC cluster usage in 2018

The Ansys, COMSOL and CST user communities at CERN are very small. Although the simulations run by these users are of critical nature for CERN, the workloads are not big enough to make full use of the available resources. This was an added reason to migrate these workloads to the Linux clusters where the resources are shared with more users and have a higher percentage of utilisation. Moreover, the Windows HPC cluster was managed by one person only. This was also another reason to push for a technology change and move to Linux batch systems where CERN IT has extended experience and a larger team of experts, who could provide better support and long term maintenance of the service.

## 2.2 Linux Clusters

The CERN Batch Service provides computing power to the CERN experiments and departments for tasks such as physics event reconstruction, data analysis, and simulations [8]. It aims to share the resources fairly and as agreed between all users of the system. The service supports both local jobs, and WLCG grid jobs via the HTCondorCE [9]. The current batch computing service is based on HTCondor [10] and currently consists of around 300,000 CPU cores (in some cases, SMT is enabled). This cluster is interconnected using standard ethernet networks, ranging from 1Gbit to 10Gbit. These networks are sufficient for the HTC computing model, where communication between jobs running on different nodes is non-existent. Other than conducting IT operations and providing communication between HTCondor components, the network is mostly used for transferring job inputs and outputs to network filesystems (AFS and EOS).

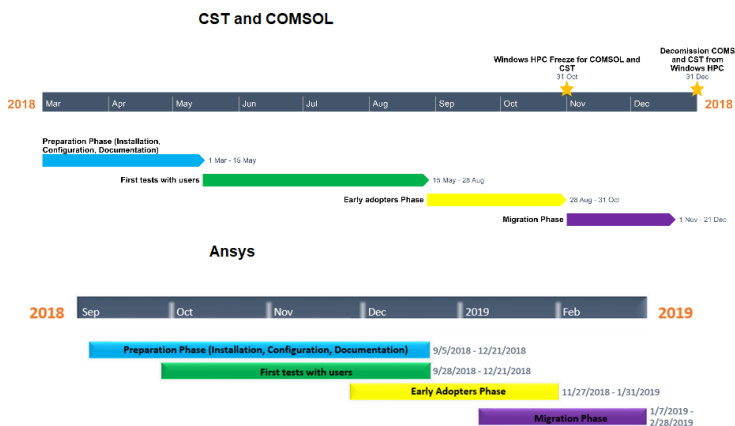
Additionally, there is an HPC service using Slurm [11] for users of MPI [12] applications that cannot fit on a single node on the regular batch service. Access is restricted to approved HPC users, mainly from the Accelerator and Technology sector. The Slurm cluster has about 240 nodes and 4480 CPU cores and consists of several cluster partitions: three 72-node In-

finiband clusters as well as 100 server nodes with low-latency 10Gb Ethernet interconnects. This HPC infrastructure is described in more detail in [13].

Both the HTCondor batch service and the Slurm cluster nodes run CERN CentOS 7 and the same production environment. The Slurm HPC cluster runs physics grid backfill jobs on idle nodes via a dedicated HTCondorCE-Slurm gateway. These backfill jobs are opportunistic and are preempted if an HPC user makes a job submission that needs backfilled resources. Hence, the dedicated HPC resources can also benefit the physics community when not fully utilized by the HPC users.

### 3 Migration

The migration of users was organised in two different stages as presented in the timelines below. First COMSOL and CST users were migrated, then in a second stage, Ansys users were migrated.



**Figure 3.** Migration Timeline

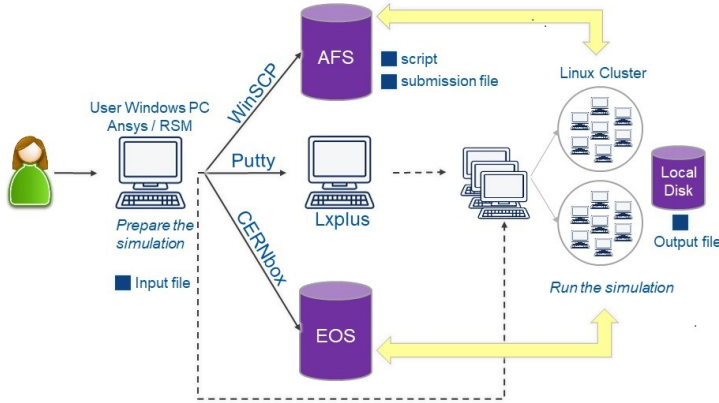
Based on the application used by the engineering simulation, the following Linux clusters were recommended:

- ANSYS Classic/Mechanical: HTCondor
- ANSYS CFX: HTCondor or Slurm, depending on the use case
- ANSYS Fluent: Slurm
- COMSOL: HTCondor
- CST: HTCondor for most solvers, Slurm for wakefield solver.

In the Figure 4 an overview of the Linux infrastructure involved is presented.

In most cases, users prepare their simulation in their Windows PC where Ansys, COMSOL or CST applications are available. The input files are then copied via the CERNbox [14] application to the EOS [15] storage system at CERN, which is our disk based storage system and where we recommend users to store large input and output files. Users then connect via an ssh client (generally Putty) to the LXPLUS interactive Linux cluster where the AFS file system is locally mounted. Users store in AFS the submission file for HTCondor or Slurm batch systems. In the case of Ansys, it is possible to use the Ansys RSM component from

## Infrastructure overview



**Figure 4.** Linux HPC cluster infrastructure

the Windows PC and submit directly to the batch systems. More details are given in the next section for each engineering application.

## 4 Engineering use cases

### 4.1 COMSOL

The COMSOL Multiphysics simulation environment facilitates all the steps in the modeling process – defining the geometry, meshing, specifying the physics, solving, and then visualizing the results. Model set-up is quick, thanks to a number of predefined physics interfaces for applications ranging from fluid flow and heat transfer to structural mechanics and electrostatics. Material properties, source terms, and boundary conditions can all be spatially varying, time-dependent, or functions of the dependent variables.

In 2018, the COMSOL user community at CERN with needs for HPC resources was comprised of 13 users.

COMSOL simulations at CERN are normally memory bound and the recommendation to the users was to target the big memory machines available in the HTCondor cluster. The big memory machines have 1TB of RAM and 24 physical cores. These resources have proven to be more efficient than the ones in the Windows HPC cluster, and have allowed to shorten simulation times considerably due to the reduced communication overhead.

COMSOL users had to learn some basic Linux commands that would allow them to submit jobs to the HTCondor Cluster. Users could connect to the interactive Linux Cluster at CERN, LXPLUS, and submit their jobs from there to the HTCondor batch system. This was in some cases challenging for the users who had never before been exposed to a Linux environment. A step by step documentation and templates were prepared for them to ease the transition.

### 4.2 CST

CST MICROWAVE STUDIO® (CST MWS) is a specialist tool for the 3D Electromagnetic simulation of high frequency components. CST MWS enables the fast and accurate analysis

**Prepare your COMSOL script**

Create a script called, for instance, `my_script.sh` with the following content:

```
#!/bin/bash
export PATH=/afs/cern.ch/project/parc/comsol54/bin/glnx64:$PATH
export LD_LIBRARY_PATH=/afs/cern.ch/project/parc/comsol54/lib/glnx64:$LD_LIBRARY_PATH
export EOS_MGH_URL=root://eosuser.cern.ch

# Copy input file to local condor jobdir:
eos cp /eos/user/m/malandes/inputfile.mph ./inputfile.mph

# Run simulation
comsol batch -usebatchlic -configuration "/tmp/config_@process.id" -tmpdir "/tmp" -prefsdir "/tmp/prefs" -data "/tmp/data_@process.id" -inputfile ./inputfile.mph -outputfile ./outputfile.mph

# Copy output back to EOS
eos cp ./outputfile.mph /eos/user/m/malandes/outputfile.mph

exit 0
```

Please, note that you have to change `/eos/user/m/malandes/` with the EOS path where your input file is stored. Also change the name `inputfile.mph` in the above example with the real file name for your input file. You can also choose another name for your `outputfile.mph` which is meaningful to you i.e. `mixer-out.mph`.

Also make sure you use the right COMSOL paths depending on which version you want to use (See [Supported COMSOL versions](#) section).

Please, note that as of version 5.4, the use of batch licenses can be specified from the command line using the option `-usebatchlic`.

**Figure 5.** Piece of documentation to submit COMSOL jobs to HTCondor

of high frequency (HF) devices such as antennae, filters, couplers, planar and multi-layer structures. CST software makes available Time Domain and Frequency Domain solvers. CST MWS offers further solver modules for specific applications.

In 2018, the CST user community at CERN with needs for HPC resources was comprised of 15 users.

At the moment of the migration, CST simulations at CERN were mostly making use of Eigenmode and Wakefield solvers. For Eigenmode solvers, we recommended the users to submit to the HTCondor cluster as the jobs could benefit from big memory machines, which will allow the simulations to run more efficiently as these workloads are memory bounded. In some cases, in particular for the Wakefield solver, running on Slurm would be more efficient as being able to run on a multi-node environment could speed up the simulation.

CST provides a Linux script collection to help users submit jobs to various batch systems, so that the user can ignore underlying batch system details. In collaboration with CST, we tried to make these scripts work with HTCondor and Slurm, but it didn't work due to some limitations on the scripts. The CST Frontend also provides a macro for job submission to Linux batch systems, but it was not possible to configure it to work with HTCondor and Slurm either. This was due to the authorization model needed by the underlying IT infrastructure that could not be integrated within the CST Frontend. CST users had to therefore connect to the Linux interactive cluster at CERN, LXPLUS, and follow step by step instructions, like with COMSOL.

The exchange and collaboration with the CST users at CERN, together with the help from the official CST support team, were key to be able to get the most of the available resources in our Linux Clusters. A CST HPC webinar was given by CST so the users could better understand the hardware recommendations for each solver, performance aspects and efficiency studies on different hardware configurations.

Thanks to the CST users, some issues were also detected and fixed to be able to better use the resources available in HTCondor and Slurm batch systems, some examples:

- CST was not doing the simulation post-processing stage due to missing HOME environment variable in the submission script
- Some of the nodes in the HTCondor infrastructure were not reporting back properly the number of available cores to CST and thus CST was not making use of all the available cores

### 4.3 Ansys

ANSYS Multiphysics software offers a comprehensive product solution for both multiphysics and single-physics analysis. The product includes structural, thermal, fluid and both high- and low-frequency electromagnetic analysis. The product also contains solutions for both direct and sequentially coupled physics problems including direct coupled-field elements and the ANSYS multi-field solver.

The program is used to find out how a given design (e.g., a machine component) works under operating conditions. The ANSYS program can also be used to calculate the optimal design for given operating conditions using the design optimization feature.

ANSYS comprises many modules, for electromechanical, multiphysics and also Computational Fluid Dynamics (Fluent).

In 2018, the Ansys user community at CERN with needs for HPC resources, was comprised of 80 users.

Ansys has a component called RSM that allows to interact with a batch system through a Graphical User Interface (GUI). HTCondor and Slurm are not supported batch systems by Ansys RSM. In the case of Slurm, this is not a problem as PBS commands are understood by Slurm, and PBS is a supported batch system. In the case of HTCondor, we developed in-house a plugin that could be integrated with Ansys RSM. The plugin was developed by a technical student, Markus Tapani Jylhäkangas, who worked in the project for one year and whose contribution was fundamental for the successful migration to Linux. The plugin was developed in python and XML and it is fully integrated in the Ansys distribution and installation at CERN.

The Ansys use cases at CERN are very diverse and depending on the simulation needs users are advised to use HTCondor or Slurm resources. Documentation for both command line submission and Ansys RSM submission is available. Similarly to COMSOL and CST, there was a good collaboration with Ansys support and the user community, to run simulations in a more efficient way. It is important to help the engineers gain insight on how to analyze how resources are used. In that way they can tune their simulations and scale up better by targeting more nodes or more memory if necessary. One example of collaboration with the users helped to understand performance issues that slowed down significantly the simulations. Thanks to this, we improved the documentation to configure Ansys RSM to store intermediate results on the computing node local file system instead of transferring them each time to the shared AFS file system.

## 5 User communication and training

Several informative meetings were held with the user community throughout the migration process so they had a good understanding of the plans and deadlines to decommission the old infrastructure.

As mentioned before, some technical seminars were organised, in the case of CST and Ansys, to gain expertise on how these applications can make efficient use of HPC resources.

Tutorials were organised with the users to go through the documentation and instructions so they could have hands-on sessions on how to interact with the Linux Clusters.

Since there are several teams involved when submitting simulations to Linux, as not only the users interact with the batch systems but also with the storage systems managed by different teams, proper user support was organised to deliver efficient help to the users. Users were educated to make use of our support system which includes 2nd Line help that was also trained to deal with the most common and basic issues, most of the time related to lack of knowledge of Linux systems.

The Linux HPC clusters are integrated with other IT services like LXPLUS [16], AFS or EOS, which have their own registration procedures, quotas, etc. Documentation was done in a way to simplify users the work to deal with all these extra services. Concrete instructions were given on how to access the relevant resources, increase quota if needed and other relevant operations for engineering simulations in this context.

## 6 Conclusions and lessons learnt

The migration exercise demonstrated that consolidating computing resources in a single infrastructure has proven to be a successful strategy. It allows to share existing resources, increases the resource utilization of costly hardware, and benefits from a larger team of IT experts.

A good collaboration among engineers and IT was key to ensure a smooth migration. The knowledge was very spread as engineers were the ones with application expertise, while the IT team had expertise on the IT infrastructure side. Maintaining a good communication was fundamental to succeed in the migration.

Detailed documentation and clear procedures were critical to let engineers concentrate on their simulations and allow for an easy transition to the new Linux clusters. Moreover, the turn over at CERN is very high. There are many students with short term contracts who have to integrate fast in the engineering teams for a successful contribution. Special care was put on troubleshooting documentation so users can quickly get through the most common configuration mistakes.

Additional work on batch submission from Windows for COMSOL and CST is needed to allow engineers to work from Windows, as Linux is not a well known operating system in this community. It is an extra burden for them to have to deal with Linux and creates a steeper learning curve.

Future work on benchmarking and performance analysis is planned. A thorough benchmarking and analysis procedure could reveal which kind of simulations run most efficiently on each kind of computing resource (HPC cluster, big memory, or regular HTC resources). Consequently, users could make better use of computing resources and reduce simulation times.

## References

- [1] Ansys Inc., *Ansys*, <https://www.ansys.com>
- [2] COMSOL Inc., *Comsol*, <https://www.comsol.com>
- [3] Dassault Systèmes, *CST*, <https://www.cst.com>
- [4] M. Husejko, I. Agtzidis, P. Baehler, T. Dul, J. Evans, N. Himyr, H. Meinhard, *HPC in a HEP lab: lessons learned from setting up cost-effective HPC clusters*, in *Journal of Physics: Conference Series* (IOP Publishing, 2015), Vol. 664, p. 092012
- [5] S.B. Maria Alandes Pradillo, *Experience finding MS Project Alternatives at CERN*, in *EPJ Web of Conferences* (EDP Sciences, 2020), p. To appear
- [6] Microsoft, *Microsoft hpc pack*, <https://docs.microsoft.com/en-us/powershell/high-performance-computing/overview?view=hpc19-ps>
- [7] J. Shiers, *Computer physics communications* **177**, 219 (2007)
- [8] B. Jones, G. McCance, S. Traylen, N.B. Arias, *Scaling Agile Infrastructure to People*, in *Journal of Physics: Conference Series* (IOP Publishing, 2015), Vol. 664, p. 022026



- [9] B. Bockelman, T. Cartwright, J. Frey, E. Fajardo, B. Lin, M. Selmecci, T. Tannenbaum, M. Zvada, *Commissioning the htcondor-ce for the open science grid*, in *Journal of Physics: Conference Series* (IOP Publishing, 2015), Vol. 664, p. 062003
- [10] D. Thain, T. Tannenbaum, M. Livny, Concurrency and computation: practice and experience **17**, 323 (2005)
- [11] A.B. Yoo, M.A. Jette, M. Grondona, *Slurm: Simple linux utility for resource management*, in *Workshop on Job Scheduling Strategies for Parallel Processing* (Springer, 2003), pp. 44–60
- [12] D.W. Walker, J.J. Dongarra, *Supercomputer* **12**, 56 (1996)
- [13] P. Llopis, C. Lindqvist, N. Høimyr, D. van der Ster, P. Ganz, *Integrating HPC into an agile and cloud-focused environment at CERN*, in *EPJ Web of Conferences* (EDP Sciences, 2019), Vol. 214, p. 07025
- [14] Labrador, Hugo González, A. and Bocchi, Enrico, Castro, Diogo, Chan, Belinda, Contescu, Cristian, Lamanna, Massimo, Lo Presti, Giuseppe, Mascetti, Luca, Moscicki, Jakub, Musset, Paul et al., *EPJ Web Conf.* **214**, 04038 (2019)
- [15] X. Espinal, E. Bocchi, B. Chan, A. Fiorot, J. Iven, G.L. Presti, J. Lopez, H. Gonzalez, M. Lamanna, L. Mascetti et al., **898**, 062028 (2017)
- [16] CERN, *Lxplus service*, <http://information-technology.web.cern.ch/services/lxplus-service>