

Exploitation of network-segregated CPU resources in CMS

Thursday 20 May 2021 15:13 (13 minutes)

CMS is tackling the exploitation of CPU resources at HPC centers where compute nodes do not have network connectivity to the Internet. Pilot agents and payload jobs need to interact with external services from the compute nodes: access to the application software (cmvfs) and conditions data (Frontier), management of input and output data files (data management services), and job management (HTCondor). Finding an alternative route to these services is challenging. Seamless integration in the CMS production system without causing any operational overhead is a key goal.

The case of the Barcelona Supercomputing Center (BSC), in Spain, is particularly challenging, due to its especially restrictive network setup. We describe in this paper the solutions developed within CMS to overcome these restrictions, and integrate this resource in production. Singularity containers with application software releases are built and pre-placed in the HPC facility shared file system, together with conditions data files. HTCondor has been extended to relay communications between running pilot jobs and HTCondor daemons through the HPC shared file system. This operation mode also allows piping input and output data files through the HPC file system.

Results, issues encountered during the integration process, and remaining concerns are discussed.

Primary authors: DELGADO PERIS, Antonio (Centro de Investigaciones Energéticas Medioambientales y Tecno); PEREZ-CALERO YZQUIERDO, Antonio (Centro de Investigaciones Energéticas Medioambientales y Tecno); ACOSTA SILVA, Carlos (IFAE and PIC); FREY, Jaime (University of Wisconsin Madison (US)); Dr FLIX, Jose (Centro de Investigaciones Energéticas Medioambientales y Tecno); Dr HERNÁNDEZ, José M. (CIEMAT); TANNENBAUM, Todd (University of Wisconsin Madison (US))

Presenter: DELGADO PERIS, Antonio (Centro de Investigaciones Energéticas Medioambientales y Tecno)

Session Classification: Facilities and Networks

Track Classification: Distributed Computing, Data Management and Facilities