

The CINES and its archiving platform PAC



The Centre Informatique National de l'Enseignement Supérieur (CINES - www.cines.fr), based at Montpellier (France) is a public establishment under the administrative supervision of the French Ministry of Higher Education and Research. Its main activities rely on HPC (High Performance Computing) and digital preservation.

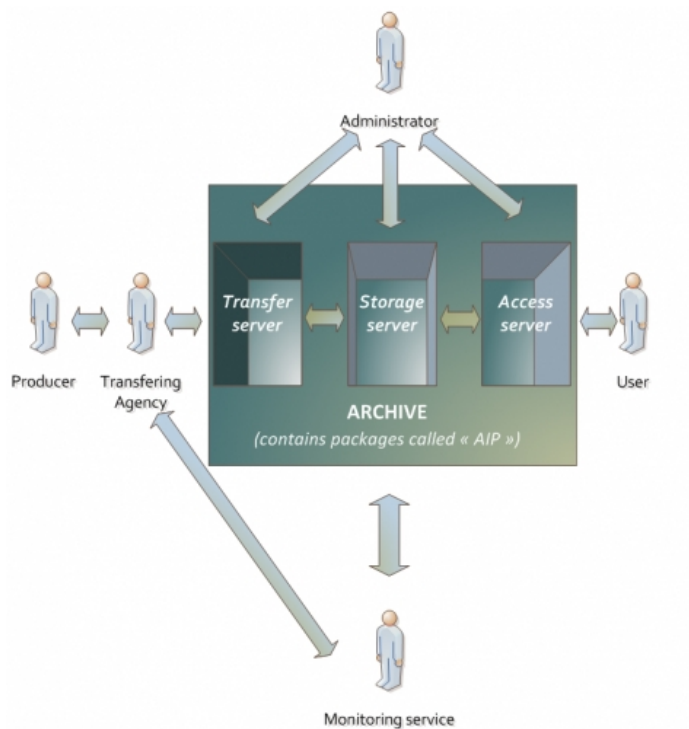
By means of high quality equipments, the CINES offers to its users the opportunity to take up great challenges. These equipments are useful for extreme simulations of infrequent or unrepeatable phenomena. The centre has in its premises the most powerful French supercomputer dedicated to research, named "OCCIGEN", capable of over two million billion operations per second. Various applications from different scientific fields such as climatology, combustion, astrophysics, quantum chemistry, seismic, medicine and biology run through it. The CINES is also part of European projects such as PRACE or HPC-Europa that advance the European HPC.

In addition to this HPC mission, the CINES is directly involved in the development of digital information long-term preservation activities for all research domains including human and social sciences. Since 2006 the CINES' Archiving Platform called "PAC" (Plateforme d'Archivage du CINES) has known two versions. The first version (PAC v1.0) was built in order to be used for digital thesis archiving. In 2008, came the second version (PAC v2.0) with much more infrastructures and means for extensive projects. Thus, the digital preservation activity has evolved and has become so important that a whole department of twelve people has been devoted to work on the development of this area. Among them are, a project manager, archivists, file formats experts, programmers, systems and infrastructures administrators and high qualified technicians.

The PAC system has been conceived as a set of three logical servers, inspired from the model that was proposed by the ISO 14 721 (OAIS ? Open Archival Information System). In order of use, there is:

- The transfer server that receives the documents from the producer
- The storage server that stores the documents
- The access server through which the authorised users can visualize their archives and get a copy.

When the transferring agency sends its files through the network or removable media, the PAC system proceeds with a series of checking. This ensures the validity of files formats, the structure of the deposit. Then, the files can be sent to the storage server with a metadata file which is added to the deposit, containing a unique identifier which will be used to find the archive when needed.



Logical view of the PAC system

One of the great achievements of this project is the creation of the FACILE service. A tool for validating the quality of files formats by ensuring that a file is valid and well formed according to the format selected from a choice list. This check is a required step in the process of long term preservation of digital objects based on format migration strategy. This service is mainly used by the PAC users among which you can find French national libraries, CNRS, French universities, scientific laboratories, etc. FACILE is accessible via internet (www.facile.cines.fr) and there you can also find the list of file formats that can be archived on the PAC platform, with the tools which were used to achieve the validation process.