ALICE - ARC integration

Anderlik, Csaba; Gregersen, Anders Rhod; Kleist, Josva; Peters, Andreas; Siaz, Pablo

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
NDGF – Nordic Data Grid Facility, is the organization responsible for the management of the T1 distributed over the nordic region; Softwarewise, the Nordic T1 is based on ARC, the Grid middleware developed by NorduGrid.

This poster presents our approach to integrate AliEn and ARC, in the sense that ALICE data management and job processing can be carried out on the NDGF infrastructure, using the client tools available in AliEn.

ALICE – A Large Ion Collider Experiment at LHC, involves ~1000 scientists (from ~100 of geographycally distributed institutes) and therefore implies a highly distributed data flow; The produced data consists of a large number of GB size files which needs to be further processed and analyzed.

AliEn – Alice Environment – is a distributed environment to deal with the computing needs of the experiment.

The interoperation has two aspects: a data management part and a job management part. Here we focus on the latter, job management is somewhat “cumbersome” due to the different computing models employed by AliEn and ARC. AliEn uses an Agent based pull model while ARC handles jobs through the more “traditional” push model. The solution comes as a module implementing the functionalities necessary to achieve AliEn job submission and management to ARC enabled sites. This approach is planned to be deployed in two stages.

Stage 1

- Distributed T1 with individual AliEn VO-boxes installed at each member site: Aalborg, DCSC_KU, NSC, LUNARC, Jyvaskyla, CSC, UiB, UiO.
- Each VO-box runs the Following services: CE, ClusterMonitor, PackMan, MonaLisa.
- Several backends: Torque, Loadleveler, SGE, ARC(UiB)
- Accounting using information from MonaLisa combined with SGAS

Stage 2

- Single ALICE VO-box for NDGF T1, submitting JAs to ARC servers (green circles) at each participating node, data management is done using a single dCache door for NDGF; each local storage element will run as a dCache pool (blue circles) for the central dCache server.