

The JINR - Russia Analysis Facility

L. Valova^{1,2}, G. Shabratova³, G. Stiforov³, R. Semenov¹

e-mail: jancurova@jinr.ru, ¹Laboratory of Information Technologies, JINR, Dubna

²TUKE, FEI, Kosice, Slovakia

³Veksler and Baldin Laboratory of High Energy Physics, JINR, Dubna

Abstract

The JRAF (JINR - Russia Analysis Facility) runs PROOF (Parallel ROOT Facility) cluster software for ALICE experiment and it is part of AAF (ALICE Analysis Facility). Simulated, reconstructed and analysis data are located on local disks of workers. It is used to perform a fast interactive analysis and calibration. The aim of the JRAF is conceptually different from analysis on the Grid. The JRAF does not make it possible, to analyse all data taken by ALICE at once, because its space is limited. But storage on workers runs in cache mode, so when all space is used, old data are removed and new are staged. The JRAF is a PROOF cluster having 48 cores and 24 TB of disk space are ready for ALICE analysis today. Due to fast staging (speed of network is then limit), it is possible to run an analysis and see results after a few minutes or even seconds, thus allowing very fast development cycles.

Keywords PROOF, ROOT, ALICE, JRAF, Grid, datasets

The JINR Analysis Facility (JRAF) is a cluster at Dubna running PROOF cluster software supported by AAF. It can be used for prompt analysis of pp data, as well as selected PbPb data. Furthermore calibration programs can be run on the JRAF. PROOF allows interactive parallel analysis on a local cluster. Interactive means that you see the results right away (contrary to a batch job where you have to wait for the job to finish before you see the results). Parallel means that several nodes execute sub sets of your data at the same time. Jobs (packets) are distributed by PROOF master using TPacketizer class. You can connect to a PROOF system from your usual ROOT prompt using TProof class. Using PROOF is aimed to be transparent, that means you can execute the same analysis code locally and on a PROOF system without changing code in user's task. Of course, certain rules have to be observed. A schema of a PROOF system can be seen in figure 1.

The PROOF based analysis and calibration, main tasks of PROOF facility at ALICE, are conceptually different from the analysis and calibration on Grid. The latter is based at largest available amount of processing data. So it gives the most statistically well-to-do results. However, these results will not

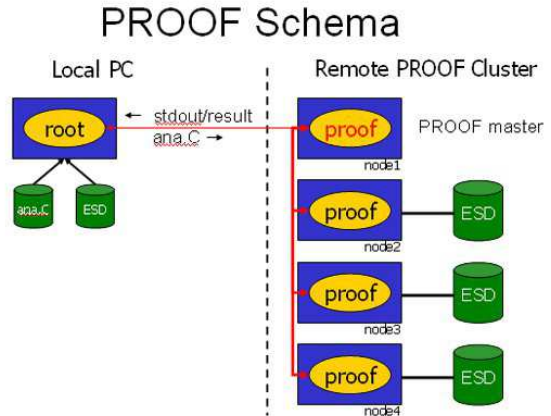


Figure 1: PROOF schema

be available in a few minutes. PROOF facility lets you get these results very fast at restricted statistics located at disks of workers, thus allowing very fast development cycles. The new type of AAF running PROOF clusters is used today for:

- Prompt analysis of proton-proton data
- Pilot analysis of Pb-Pb data
- Calibration & Alignment

The JRAF Analysis Facility (JRAF)

The JRAF will run PROOF for ALICE. Simulated data with different LHC periods starting from PDC06 and measured reconstructed data in both formats ESD(Event Summary data) and AOD (Analysis Object Data), where AOD is approximately 10 times smaller and it is more used by analysis. Once ALICE starts data taking and reconstruction is done, data can be accessible within few minutes, hours (this depends on size of data and speed of network) and they are available on the JRAF when automatic staging copy files from ALICE storage elements. Users can register datasets they want, and data are automatically staged to the JRAF PROOF cluster. Status of data availability is checked every five minutes so user is able to analyse part of dataset already staged. It can be used to perform analysis and calibration within minutes. The aim of the JRAF is conceptually different from the analysis on the Grid, due to

