

PROOF on Demand (PoD): dynamic PROOF clusters on the fly

PROOF is an extension of ROOT enabling interactive analysis of large sets of files in parallel on clusters of computers. Normally users get PROOF provided by administrators as a pre-installed shared cluster. To avoid certain disadvantages of “static” PROOF clusters PoD has been developed. PROOF on Demand is a tool-set, which dynamically sets up a PROOF cluster at a user request, on any resource management system (RMS). PoD is a user oriented product with an easy to use GUI and a command-line interface. It is fully automated, and no administrative privileges, special knowledge or preconfigured nodes are required to use it. It provides a plug-in based system, in order to use different job submission front-ends. PoD is currently shipped with gLite, LSF, PBS (PBSPro/OpenPBS/Torque), Grid Engine (OGE/SGE), Condor, and SSH plug-ins. PoD makes it possible just within a few seconds to get a private PROOF cluster on any RMS. If there is no RMS, then SSH plug-in can be used, which dynamically turns a bunch of machines to PROOF workers. The SSH plug-in is also a perfect solution for a Cloud based PROOF clusters. PoD algorithms efficiently support two types of connections: native PROOF connections and packet-forwarding connections. This helps to properly handle most kinds of workers, with and without firewalls. Upcoming versions of PoD are going to support an out-of-server user interface. Users will be able to select a remote computer acting as PoD server (PROOF master). In this case PoD UI will be just a lightweight control center and could run on different OS types. Also an AliEn plug-in is going to be developed in collaboration with the ALICE Offline team. This cooperation will help PoD to provide a fast interactive PROOF experience on the AliEn Grid. As PoD matures as a product, it is used more and more as a standard for setting up dynamic PROOF clusters in many different institutions in HEP community. Additionally there are already several Cloud based installations, which use PoD as a PROOF cluster solution. With PoD there is no need to maintain a dedicated PROOF analysis facility. PoD users create themselves private dynamic PROOF clusters on general purpose batch farms, Grid or Cloud systems. The concept of PoD, new PoD development as well as user experience will be covered in this presentation.

Primary authors : Mr. MANAFOV, Anar (GSI Helmholtzzentrum für Schwerionenforschung GmbH) Co-authors : Dr. MALZACHER, Peter (GSI Helmholtzzentrum für Schwerionenforschung GmbH) Presenter : Mr. MANAFOV, Anar (GSI Helmholtzzentrum für Schwerionenforschung GmbH)