PROOF on Demand

A. Manafov, P. Malzacher

GSI, Darmstadt, Germany

PROOF on Demand (PoD) [1][2] is a tool-set that automats and dramatically simplifies the process of creating private PROOF [3] clusters.

PoD dynamically sets up a PROOF cluster at a user's request on any resource management system (RMS). It provides a plug-in based system, in order to use different job submission front-ends.

In 2011 the plug-in set has been extended and supports now gLite, LSF, PBS (PBSPro/OpenPBS/Torque), Grid Engine (OGE/SGE), Condor, LoadLeveler, and SSH plug-ins. PoD makes it possible just within a few seconds to get a private PROOF cluster on those systems. If there is no RMS, then the SSH plug-in can be used, which dynamically turns a bunch of machines to PROOF workers.

User community

PoD is being very successfully used by German ALICE collaboration and has been chosen as a standard PROOF tool for FAIR computing [4].

The tool-set is official distributed to all Tier3 ATLAS sites via standard ATLAS software deployment machinery [5].
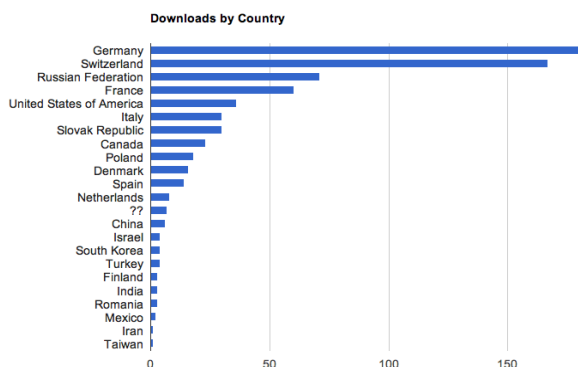

Figure 1: ~700 downloads from 22 countries (since June, 2010), with 180 downloads/year from unique IPs.

A thin client concept

Recently a new major step in PoD development has been made. It now supports not only local PoD servers, but also remote ones.

A newly developed "pod-remote" command made it possible for users to utilize a thin client concept. In order to create dynamic PROOF clusters, users are now able to select a remote computer, even behind a firewall, to control a PoD server on it and to submit PoD jobs. In this case a user interface (UI) machine is just a lightweight control center and could run on different OS types or mobile devices.

All communications are secured and provided via SSH channels. Additionally PoD automatically creates and maintains SSH tunnels for PROOF connections between the UI and a PROOF master.

Dynamic PROOF masters

Future versions of PoD are going to implement a dynamic PROOF masters feature, which is an extension of the recently developed "pod-remote". This feature is going to enable PoD to automatically assign PROOF master processes to worker nodes (WN). Users would need only to run a lightweight PoD/PROOF UI to process their analysis, as the rest will be entirely running on WNs.
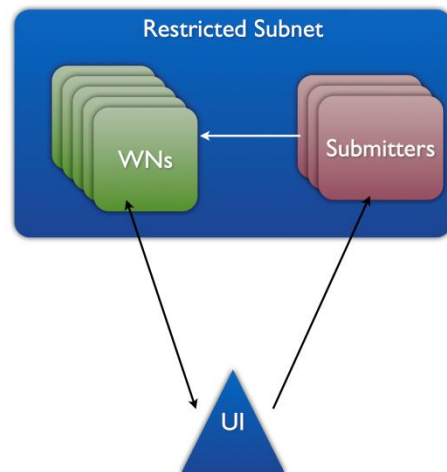

Figure 2: Dynamic PROOF masters.

Summary

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 the HEP community. 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.

References

[1] A. Manafov et al, "PROOF on Demand", PHN-IS-IT-07, GSI Scientific Report 2010.
[2] PROOF on Demand (PoD), http://pod.gsi.de
[3] The Parallel ROOT Facility (PROOF), http://root.cern.ch/drupal/content/proof
[4] M. Al-Turany et al, "PROOF integration in FairRoot, GSI", Scientific Report 2011.
[5] ATLAS Wiki, https://twiki.atlas-canada.ca/bin/view/AtlasCanada/ManageTier3SW