PROOF on Demand (PoD)

Peter Malzacher, Anar Manafov (GSI Darmstadt, Germany)

PoD is a set of utilities developed at GSI, which allows starting a PROOF cluster at user request, on any resource management system. Installation is simple and doesn't require administrator privileges, and all the processes run in user space. PoD gives users, who don't have a centrally-administrated static PROOF cluster at their institute, the possibility to enjoy the full power of interactive analysis with PROOF.

PoD is a specially designed solution to provide a PROOF cluster on the fly.

Native PROOF connections

Whenever possible, PoD setups direct PROOF connections between nodes. It results in a full functional PROOF cluster. Users get native speed and the whole range of PROOF features. To use native connections an incoming traffic must be allowed on PoD workers for a defined port. Otherwise PoD uses packet-forwarding algorithms.

Multiuser/-core environment

PoD implements automatic port mapping algorithms to properly handle cases when several users start PoD instances (servers/ workers) on the same machine. PoD also automatically manages situations when multiple PoD workers are started on the same node. Private PoD instances can't disturb each other.

Easy to use

The process of installation is very simple and fully automated. PoD works out of the box. Its distribution contains precofigured modules and everything users need to just immediately start to work with it right after the installation.

Packet-forwarding

When worker nodes are behind a firewall then PoD uses its packet-forwarding algorithms to maintain the PROOF traffic. The algorithms are very efficient, there will be no speed penalty, but some PROOF functions are limited.

GUI & Command-line

PoD provides a simple and intuitive graphics user interface in order to simplify access to its functionality. For user's convenience there is also a command line interface, it helps to manage a PoD cluster remotely or use it in a batch mode.

Different job managers

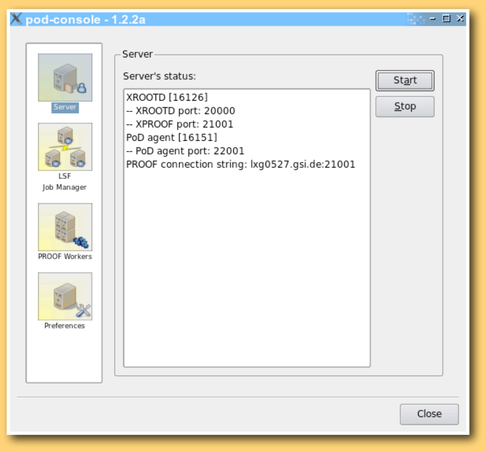
PoD supports different job managers via a plug-in system. It is a very easy to extend system. PoD currently supports gLite and LSF. PoD v2.1.4 will be shipped with an SSH plug-in. In future versions more plug-ins will be implemented (PBS, Condor and SGE).

3 steps to start your private PROOF cluster

Start

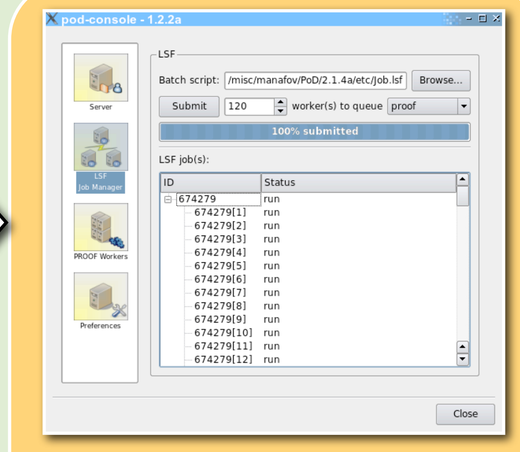
GUI

Cmd Line



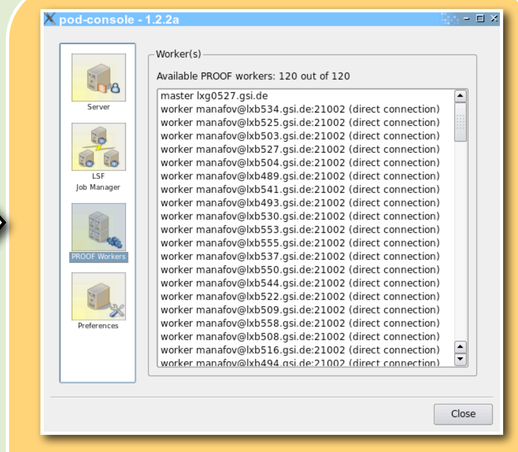
pod-server start

The machine, where a user starts the PoD server will be the PROOF master. Stopping of the PoD server will trigger automatic shutdown of PoD workers.



pod-submit ...

The user chooses a job manager plug-in to distribute PoD workers to an resource. It's also possible to accumulate workers from different resources in one PoD session, i.e. submit from different plug-ins.



pod-info ...

PROOF workers appearing online. The start-up time for jobs (grid/local) ranges from a few seconds to hours, depending on the chosen resource. When workers are online, they can be reused for many PROOF sessions.

Your dynamic PROOF cluster is ready!

You can now connect to your cluster to process a PROOF analysis as if it would be a regular static PROOF cluster.

If you close your ROOT session or an analysis crashed there is no need to resubmit PoD workers. PoD automatically reuses workers.

At any time you can add more PROOF workers (step #2).

PoD will automatically shutdown the cluster or a particular worker when idle for a pre-defined amount of time.

PoD: http://pod.gsi.de **PoD Weblog:** http://proof-on-demand.blogspot.com **PROOF:**http://root.cern.ch/drupal/content/proof **D-Grid Initiative:** http://www.d-grid.de [Work supported by BMBF, FK 01AK802G]