Dynamic PROOF clusters on the fly

Anar Manafov, GSI Darmstadt

HEP Data Analysis

Implement algorithm → Run over data set   Make improvements
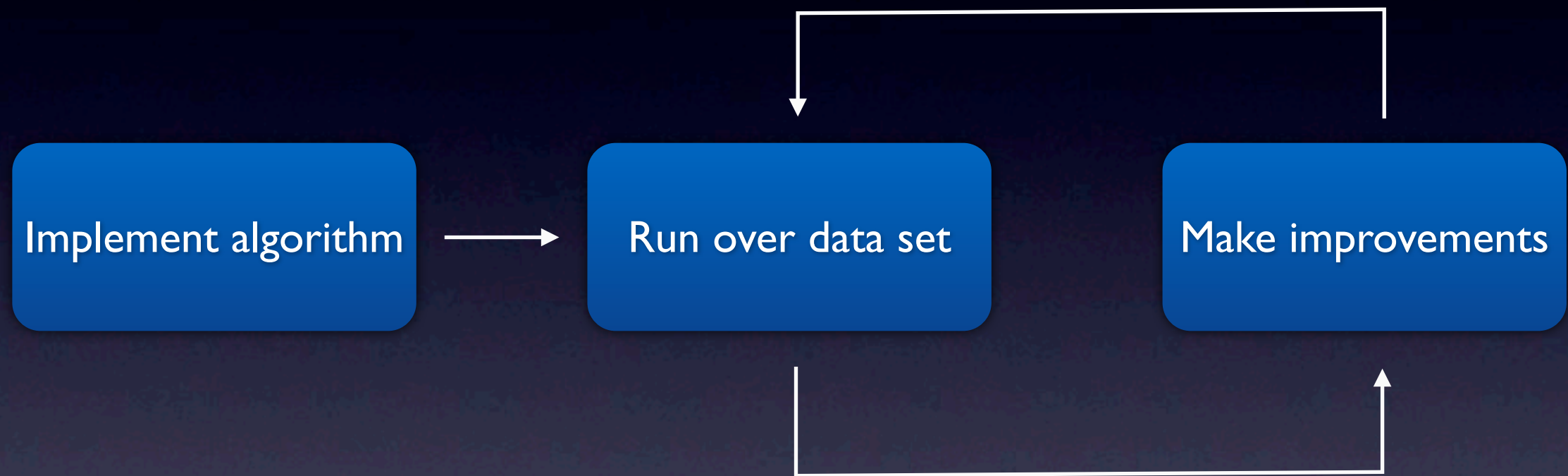
Typical HEP analysis needs a continuous algorithm refinement cycle

PROOF

File Catalog

Storage

Query

PROOF cluster

Scheduler

CPUs

PROOF Query:
data file list, mySelector.C

Master

Feedback,
merged final output
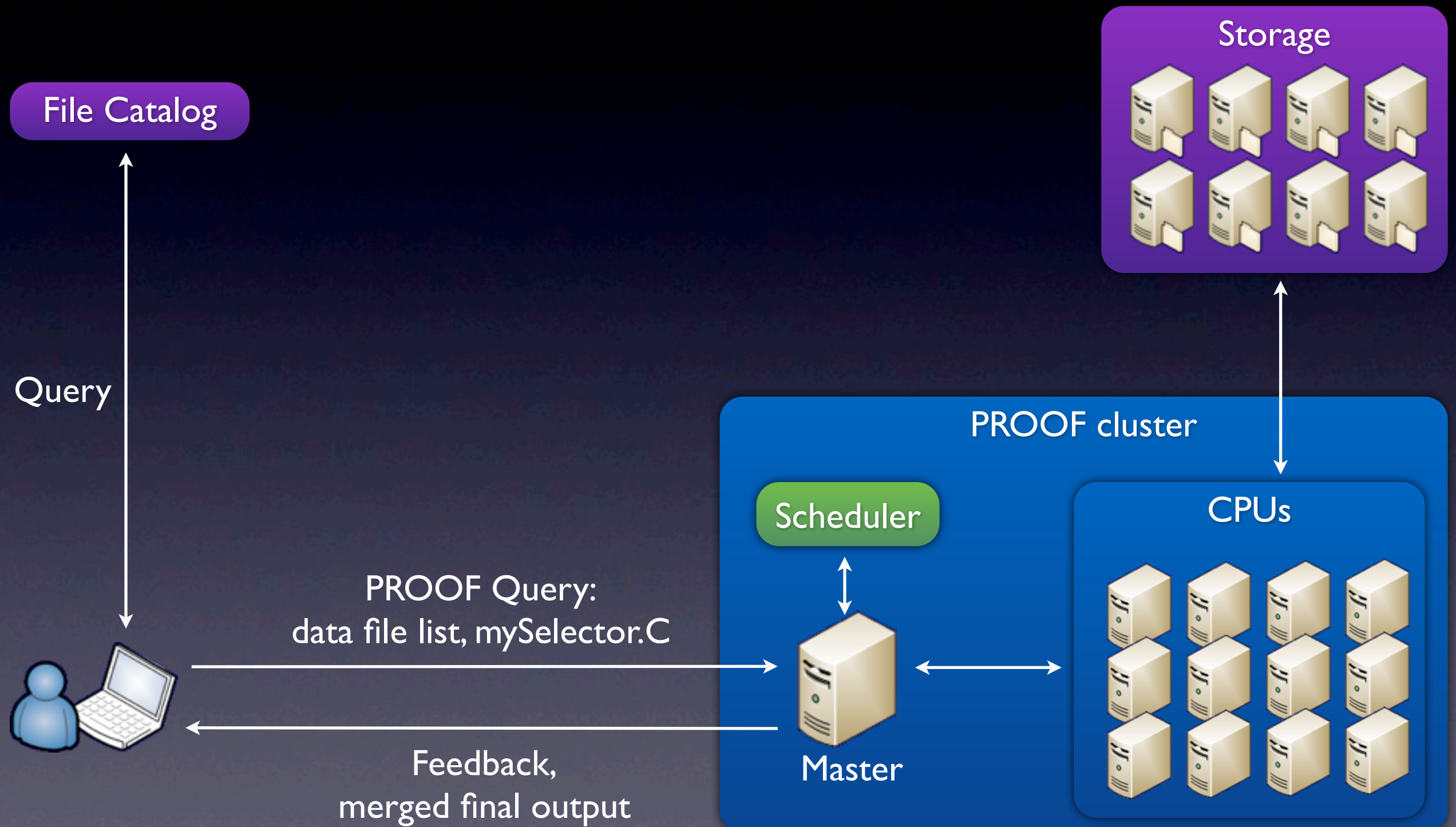
"static"/pre-installed PROOF cluster

- One user can disturb other users.

- From time to time admin. interventions are needed.

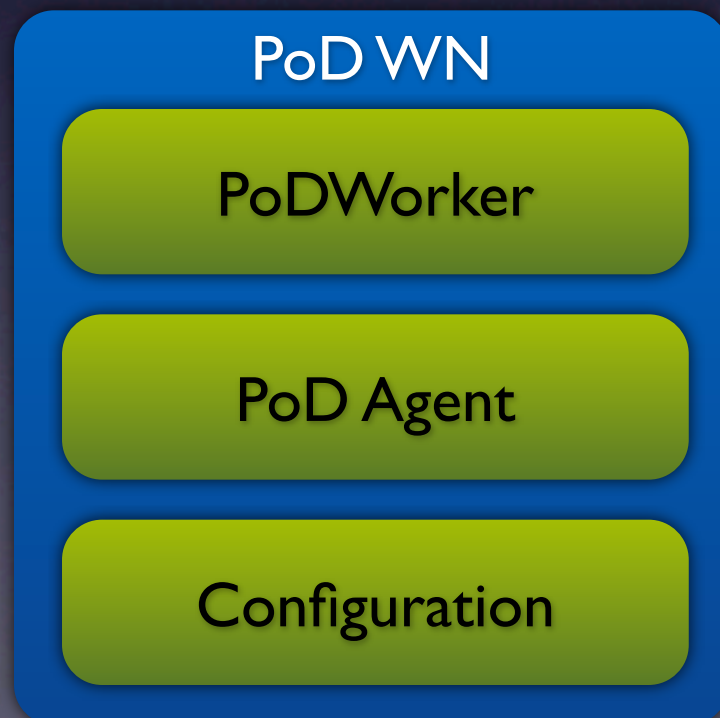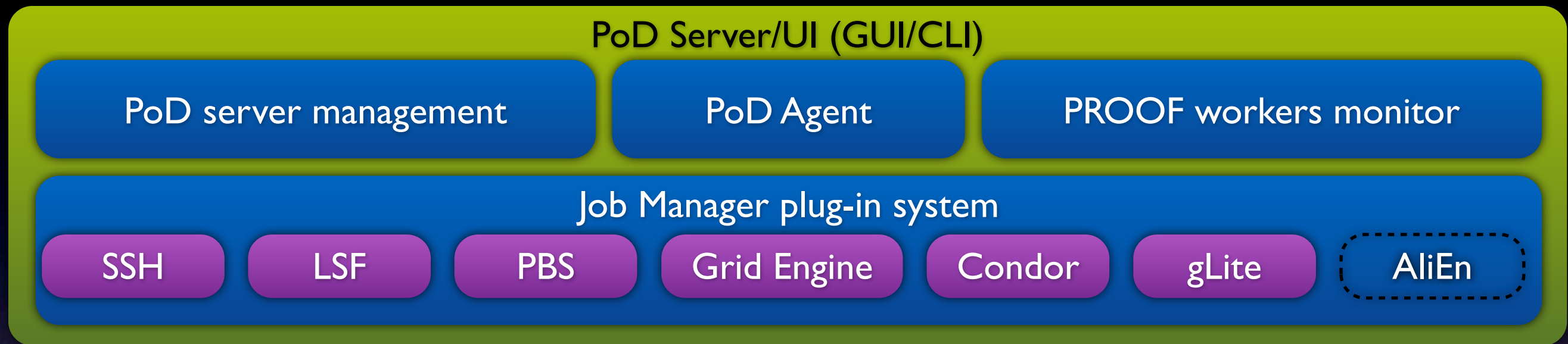- There is only one ROOT/xrootd version for PROOF services.

- There is a master node limitation.

PoD

PROOF on Demand

PoD v3.x

PoD Server/UI (GUI/CLI)

| PoD server management | PoD Agent | PROOF workers monitor |

Job Manager plug-in system

SSH   LSF   PBS   Grid Engine   Condor   gLite   AliEn

PoD WN

PoDWorker

PoD Agent

Configuration

Server and UI can run on the same or different hosts.

PoD supports Linux and Mac OS X.

Live Demo

User Friendliness

- low system/software requirements,

- simple and intuitive installation procedure,

- administrative privileges are not required,

- ready to be used right after the installation,

- provides GUI and CLI,

- very well documented and supported.

3 steps to set your private PROOF cluster up

Start PoD Server

Step #1
pod-server start

Step #2
pod-submit -r [lsf | ge | pbs | condor] -q my_queue -n 100
or
pod-ssh -c pod_ssh.cfg --submit

Step #3
pod-info -n

3 steps to set your private PROOF cluster up

| Start PoD Server | → | Submit PoD Jobs using PoD Job Manager | → | Check for Your PROOF Cluster |

Step #1

```
pod-server start
```
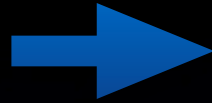
Step #2

```
pod-submit -r [lsf | ge | pbs | condor ] -q my_queue -n 100
or
pod-ssh -c pod_ssh.cfg --submit
```

Step #3

```
pod-info -n
```

3 steps to set your private PROOF cluster up

| Start PoD Server | → | Submit PoD Jobs using PoD Job Manager | → | Check for Your PROOF Cluster |

Step #1

pod-server start

Step #2

pod-submit -r [lsf | ge | pbs | condor] -q my_queue -n 100
or
pod-ssh -c pod_ssh.cfg --submit

Step #3

pod-info -n

3 steps to set your private PROOF cluster up

Start
PoD Server

→

Submit PoD Jobs
using PoD Job Manager
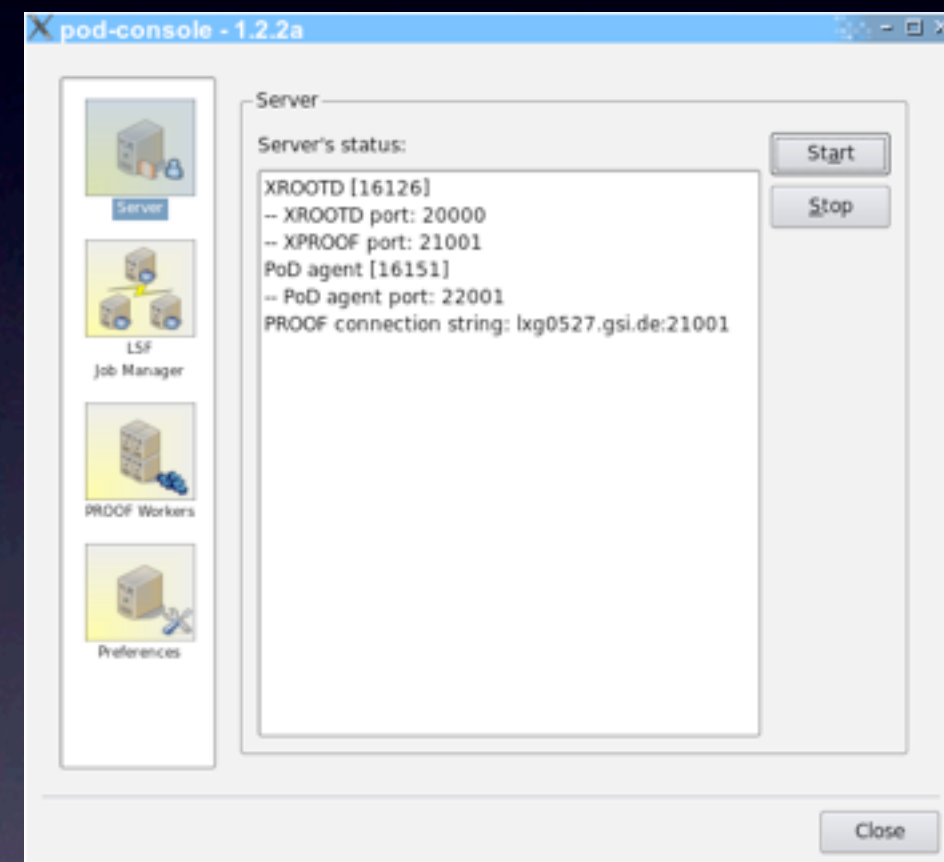
→

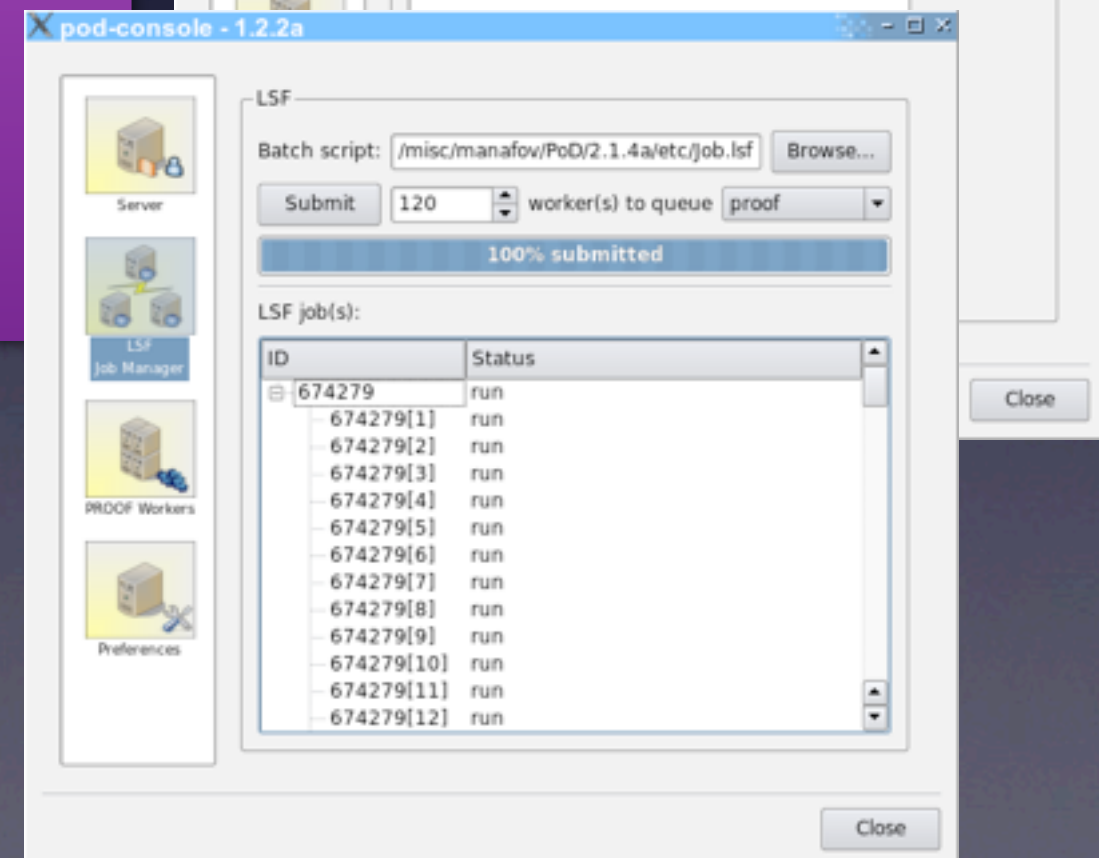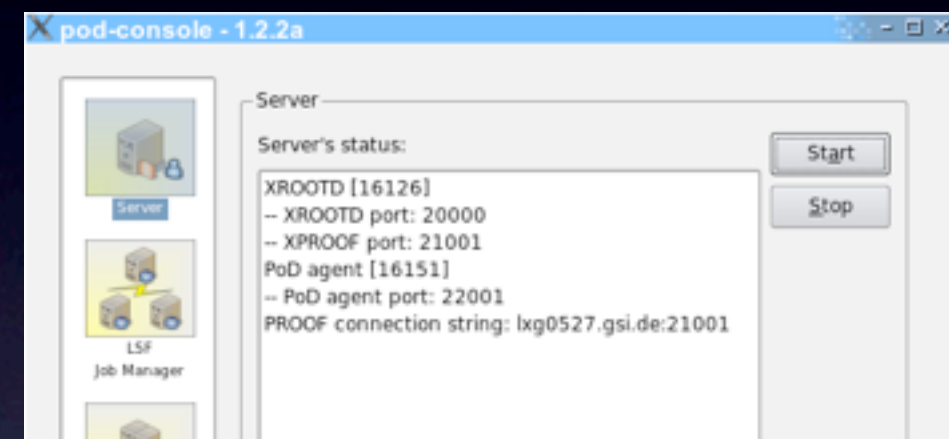Check for Your
PROOF Cluster

Step #1

pod-server start

Step #2

pod-submit -r [lsf | ge | pbs | condor] -q my_queue -n 100
or
pod-ssh -c pod_ssh.cfg --submit

Step #3

pod-info -n

3 steps to set your private PROOF cluster up

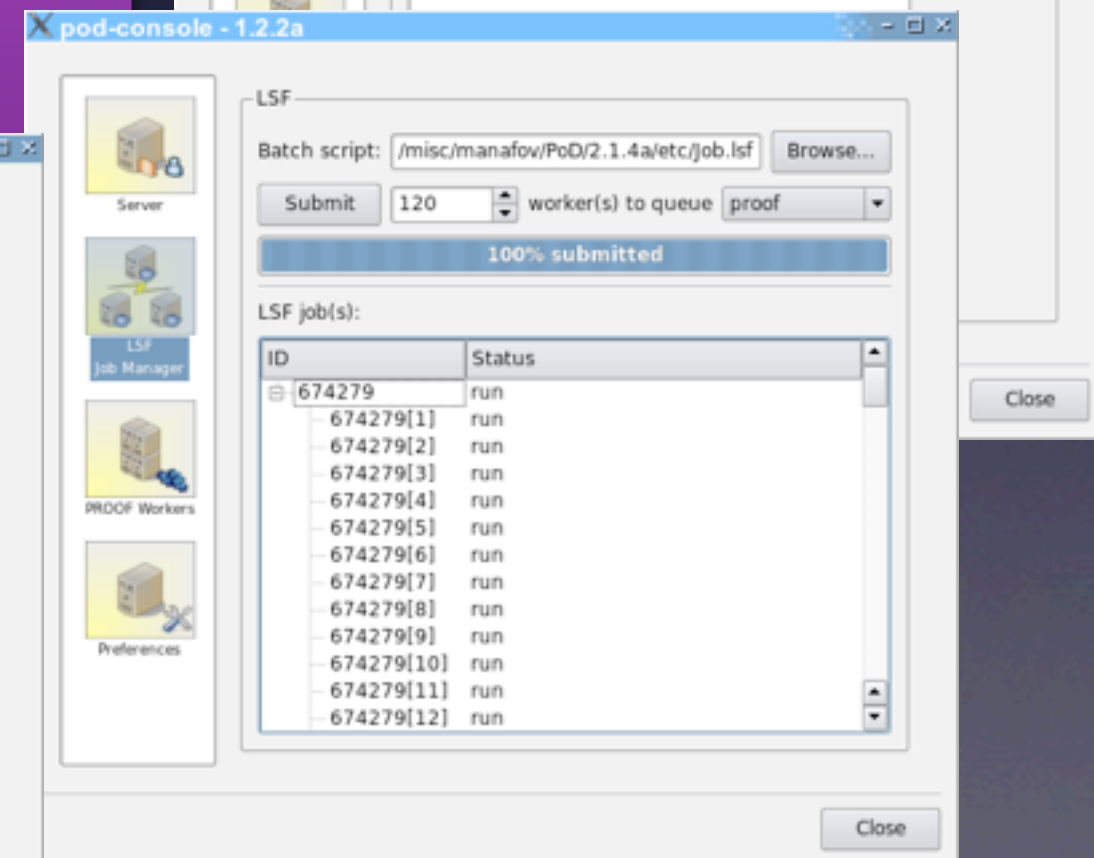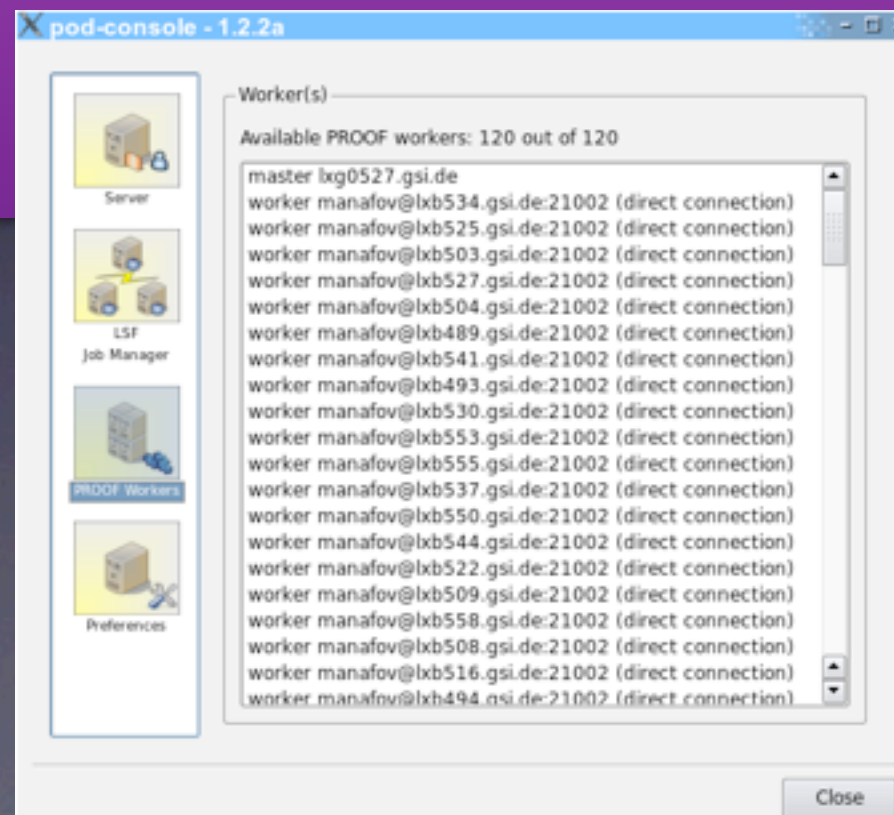| Start PoD Server	→	Submit PoD Jobs using PoD Job Manager	→	Check for Your PROOF Cluster

Step #1

pod-server start

Step #2

pod-submit -r [lsf | ge | pbs | condor] -q my_queue -n 100
or
pod-ssh -c pod_ssh.cfg --submit

pod-info -n

Submit PoD jobs to a selected RMS

A Grid Site

Firewall / Sub-net

Workers

User workspace

PoD UI

PoD Server

Workers

A Local cluster

Whenever possible, PoD setups **direct PROOF connections** between nodes. To use native connections an incoming traffic must be allowed on PoD workers for a defined port. Otherwise PoD uses **packet-forwarding algorithms**.
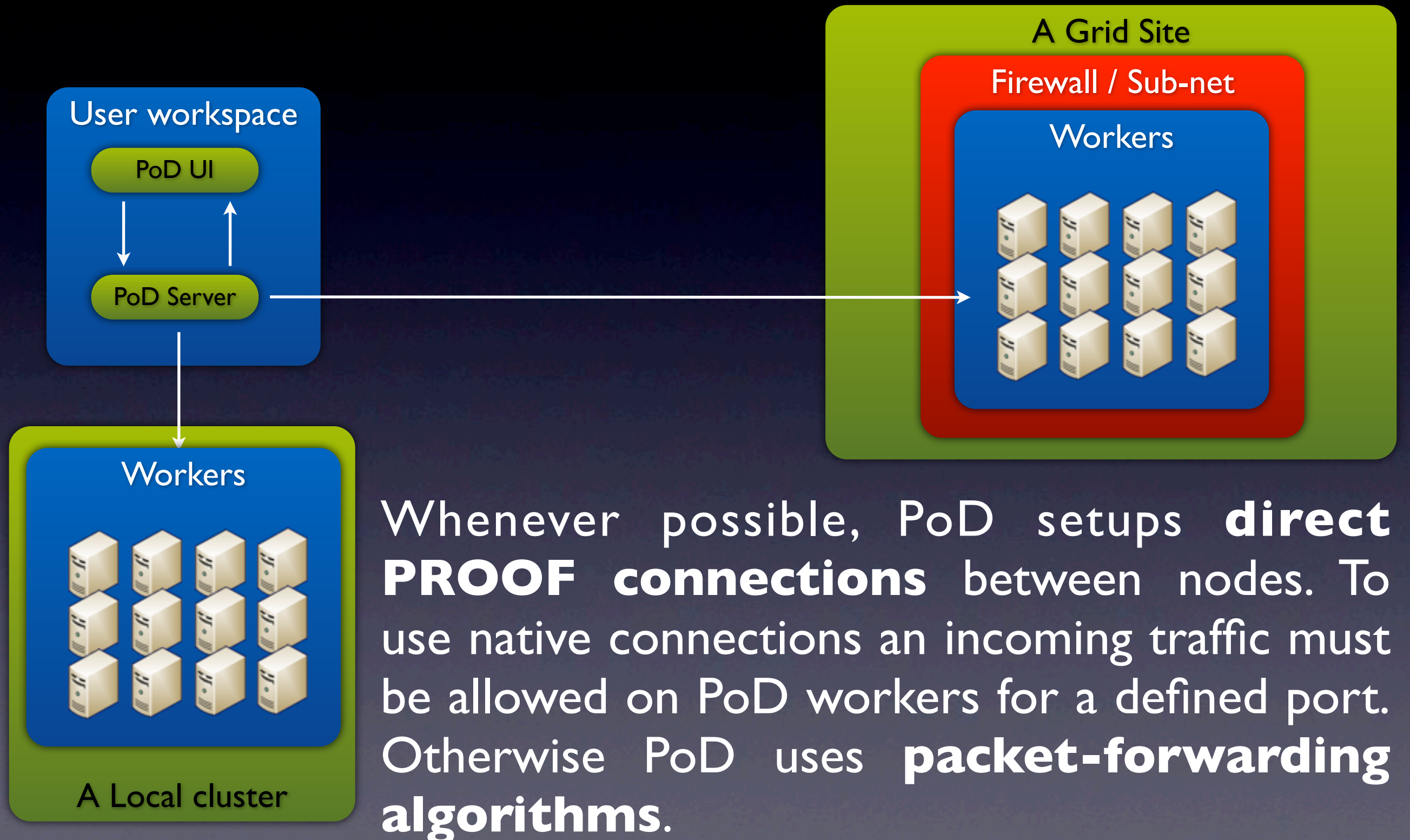
PoD WNs reporting back to PoD server

Server validates PoD WNs, which if validated will become your PROOF WNs.
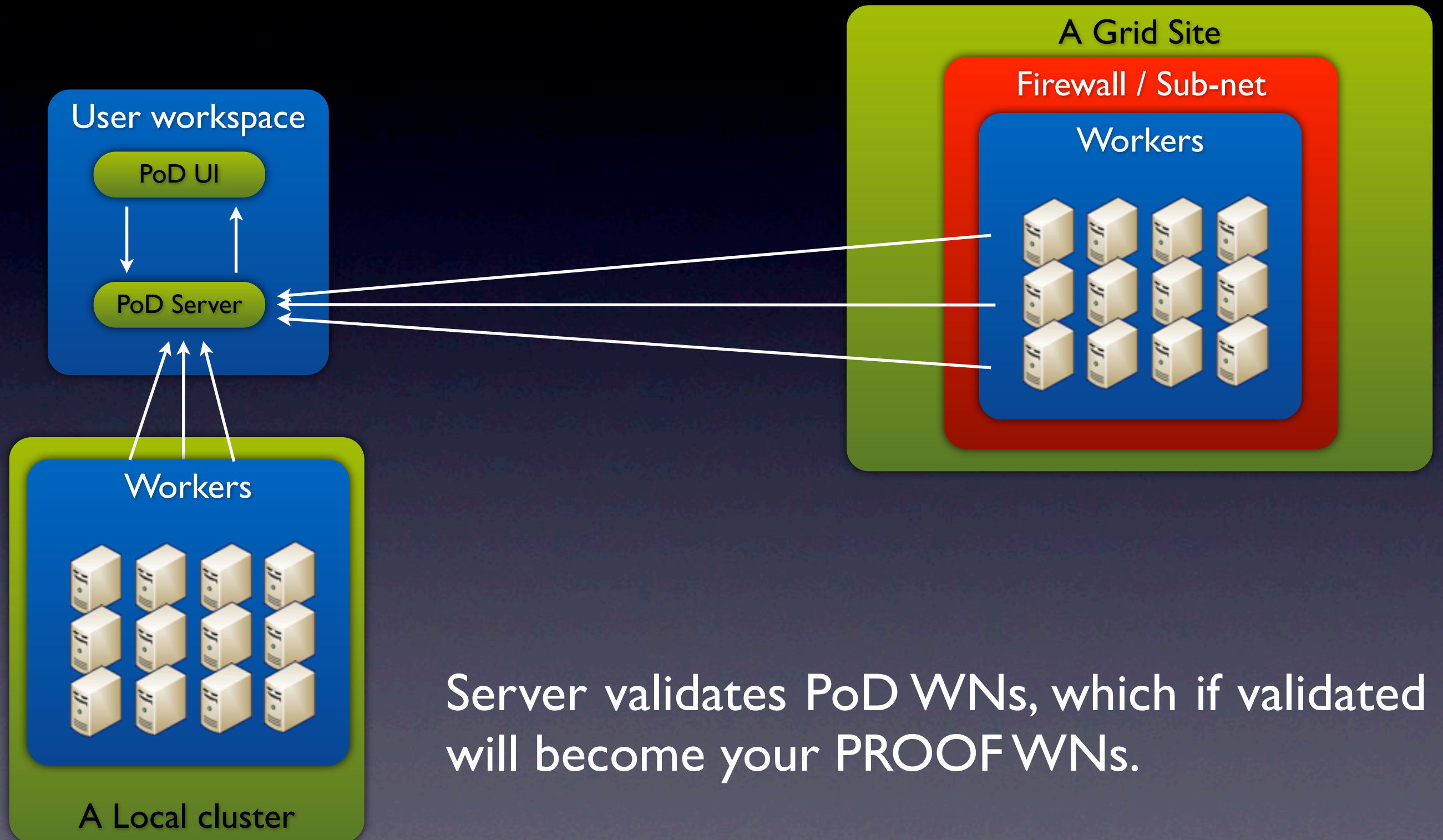
Different job managers

Job Manager plug-in system

| SSH | LSF | PBS | Grid Engine | Condor | gLite | AliEn |

PoD is shipped with a number of plug-ins, which cover all major RMSs, such as local cluster systems, Grid RMSs.

If you don't have any RMS, then the SSH plug-in can be used.
The SSH plug-in is also used to setup PROOF clusters on Clouds.

12

SSH plug-in

a simple CSV file as an input to the plug-in:

```
id  , login@host        , ssh_params , wn_dir     , num_of_workers

r1  , anar@lxg27.gsi.de , -p24       , /tmp/test , 10
a2  , user@lxi001       ,            , ~/pod_wn  ,
125 , doom@host.my      , -p22       , /opt/pod  ,
```

the plug-in is in PoD
CLI only, so-far

PoD UI

Firewall

send PoD worker (ssh/scp)

WN #1

send PoD worker (ssh/scp)

WN #X

PROOF via PoD

User

- can entirely control his/her dynamic cluster,

- can setup and use it on demand,

- can dynamically change an amount of workers,

- can select a preferable master host,

- doesn't need admins to take an action,

- doesn't disturb other users,

- is free to choose a ROOT version for services.

Key features

- Easy to install and use

- Supports private or shared installations

- GUI & Command-line

- Different job managers (SSH, LSF, PBS, GridEngine, Condor, gLite...)

- Multiuser/-core environment

- Native PROOF or Packet-forwarding connections

- User defaults - configuration

Short-term ToDo

- finish the PoD UI and Server disentangling,

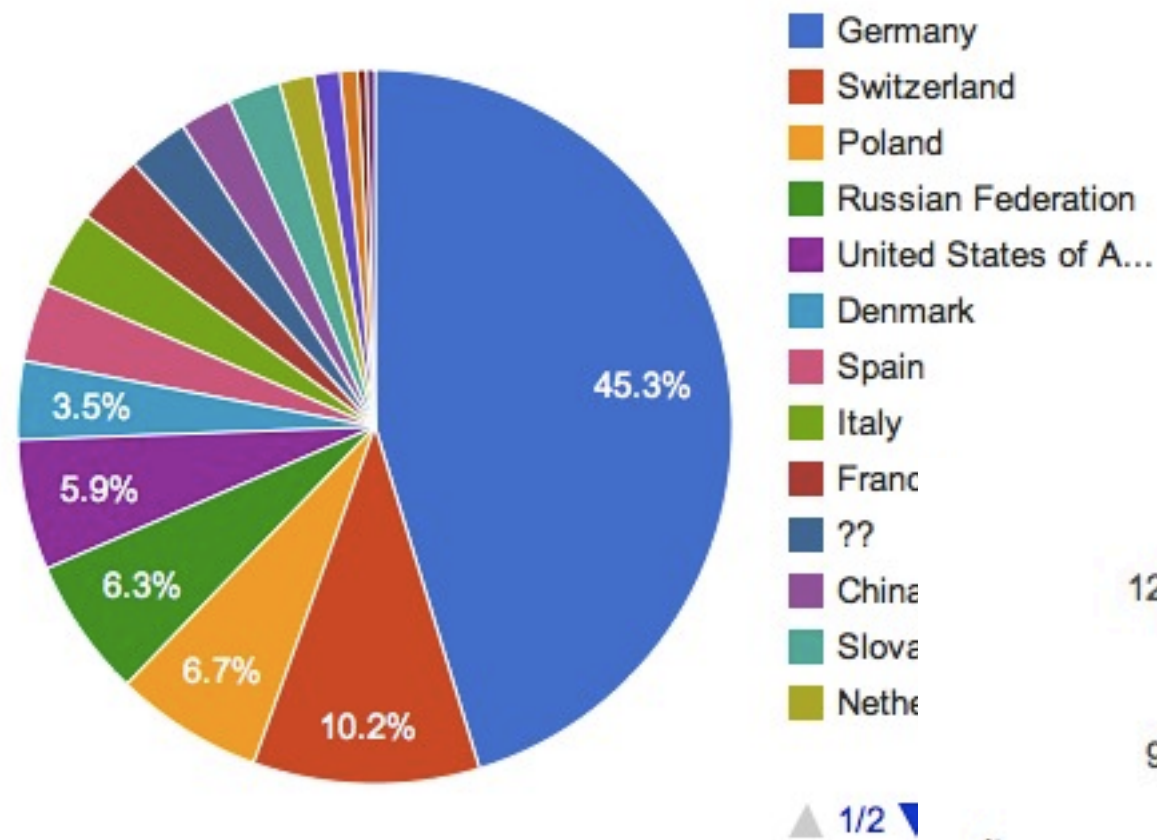- finish the AliEn integration,

- revamp the GUI of PoD.

http://pod.gsi.de

Download Statistics

254 Downloads from 16 countries

(since June, 2010)

Downloads by Country

■	Germany
■	Switzerland
■	Poland
■	Russian Federation
■	United States of A…
■	Denmark
■	Spain
■	Italy
■	Franc
■	??
■	China
■	Slova
■	Nethe

▲ 1/2 ▼

Downloads by Country

Monthly Download Statistics

Downloads by Versions