PoD

PROOF on Demand

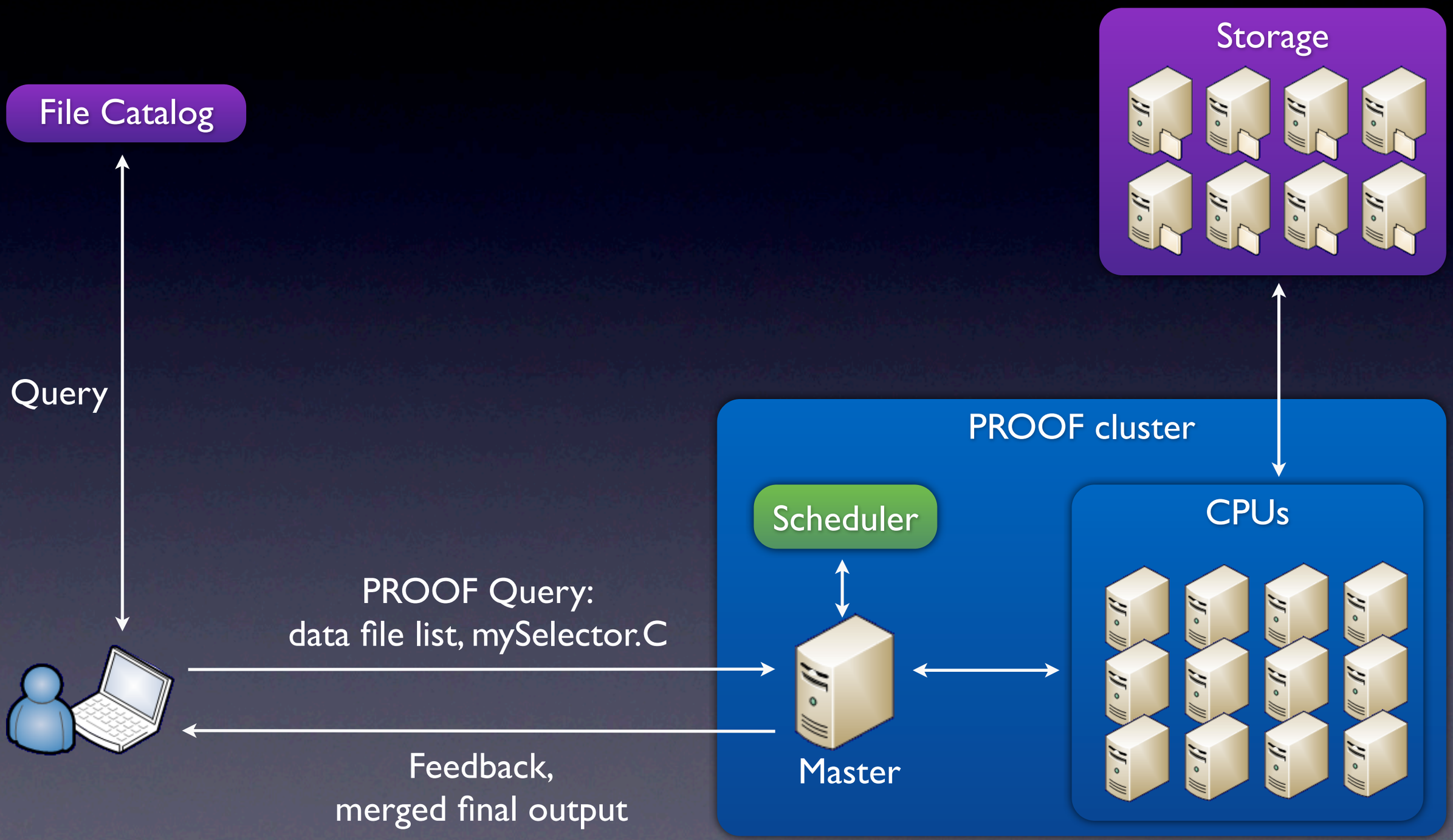
Anar Manafov, GSI Darmstadt

5th EGEE UF, 2010-04-15

HEP Data Analysis

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

Dynamic cluster

User

- can entirely control it,
- can setup and use it on demand,
- can reserve a desired amount of workers,
- can select a preferable master host,
- doesn’t need admins to take an action,
- doesn’t disturb other users.

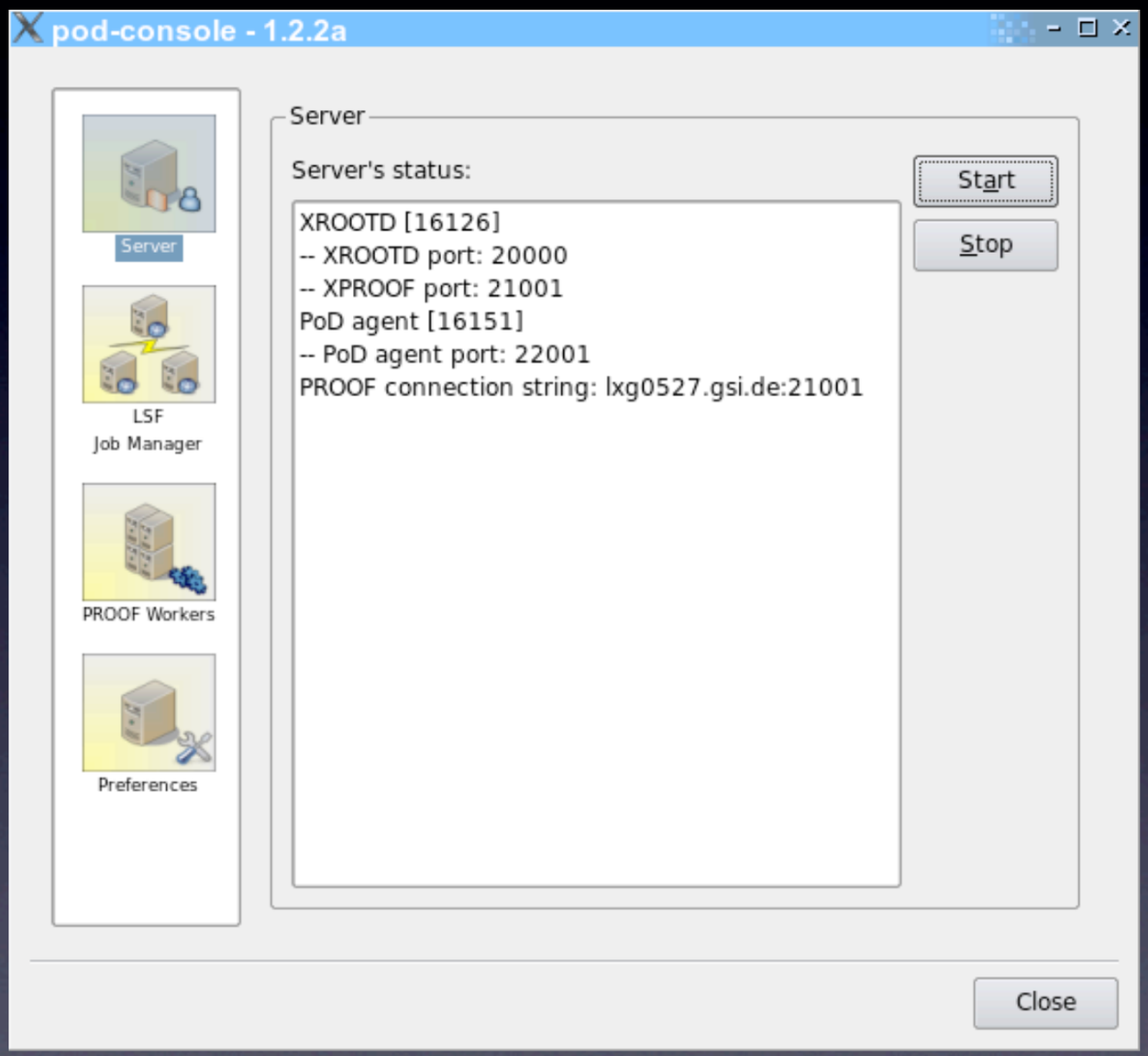
PoD

PROOF on Demand

developed at GSI by the Scientific Computing group

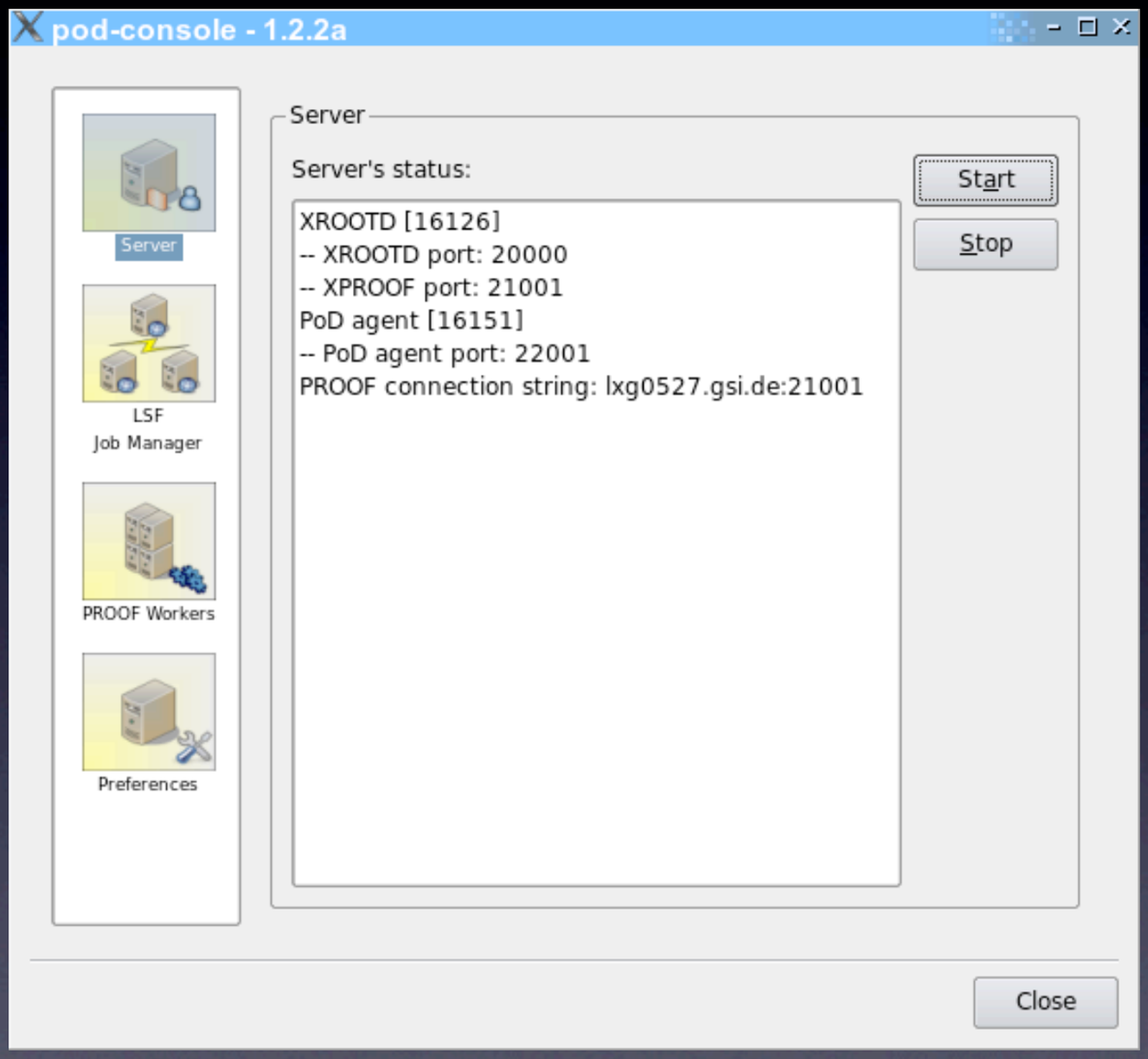
3 steps to set your private PROOF cluster up

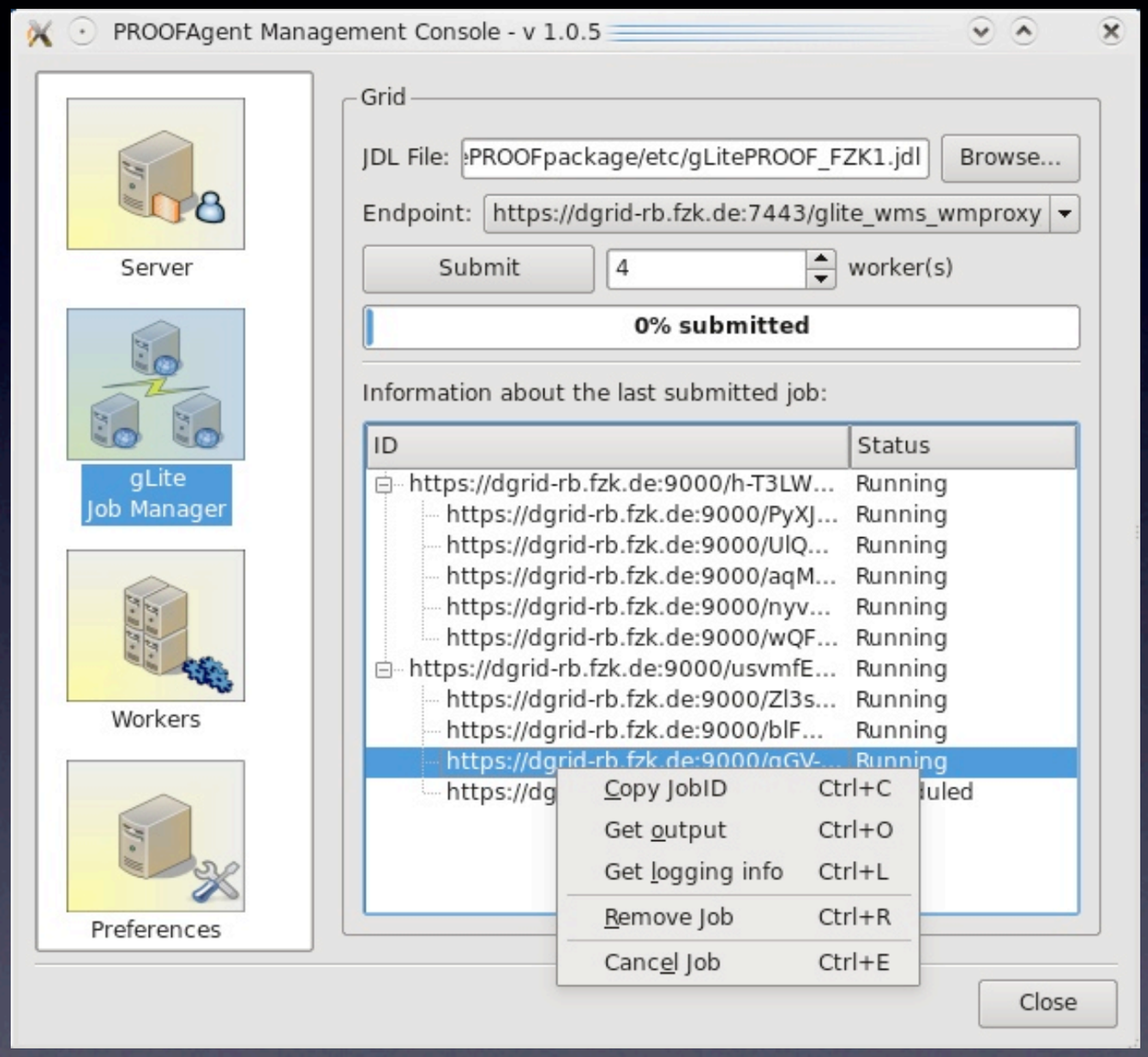
PoD server



3 steps to set your private PROOF cluster up

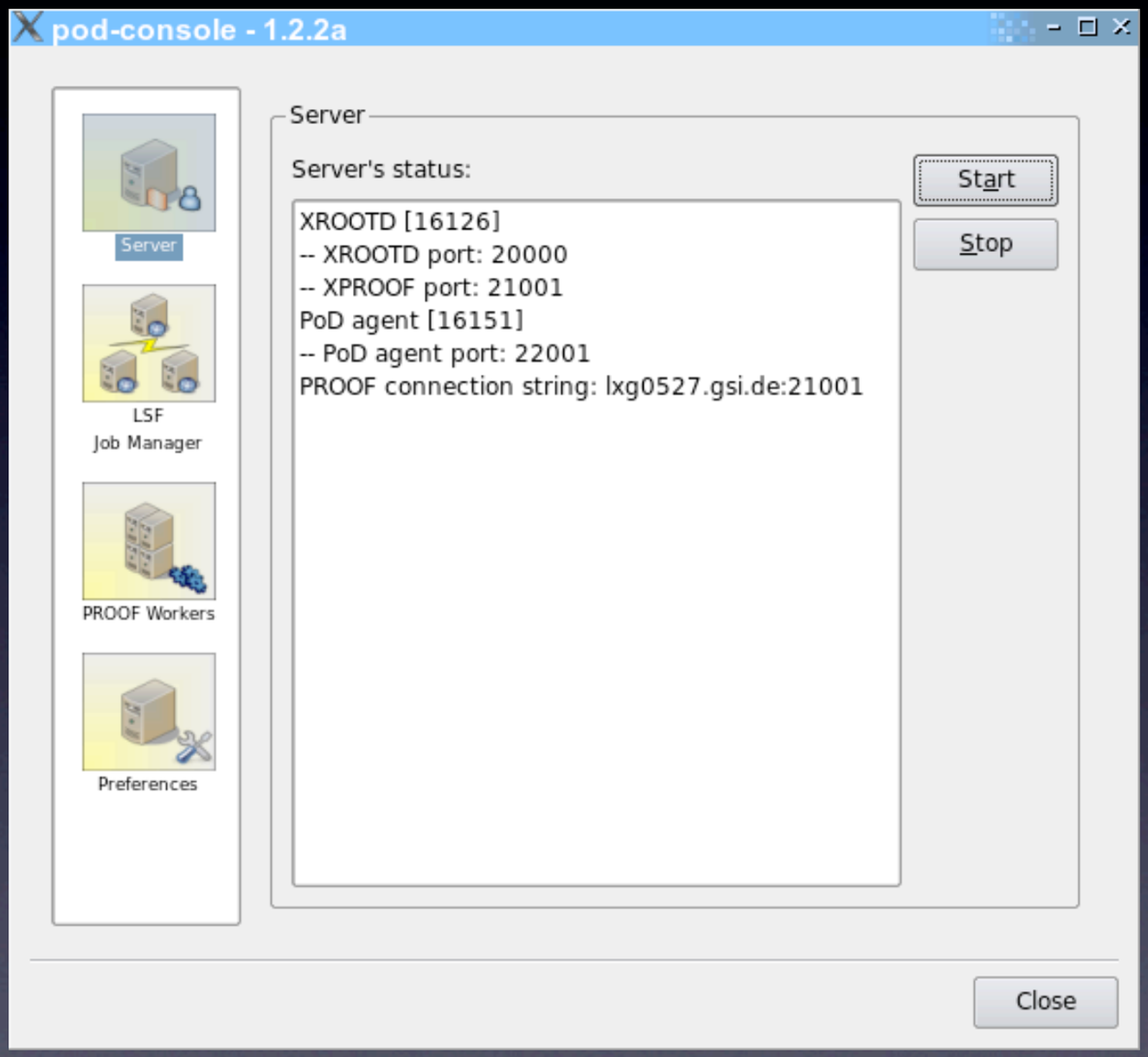
PoD server → Job Manager (gLite, PBS, LSF)

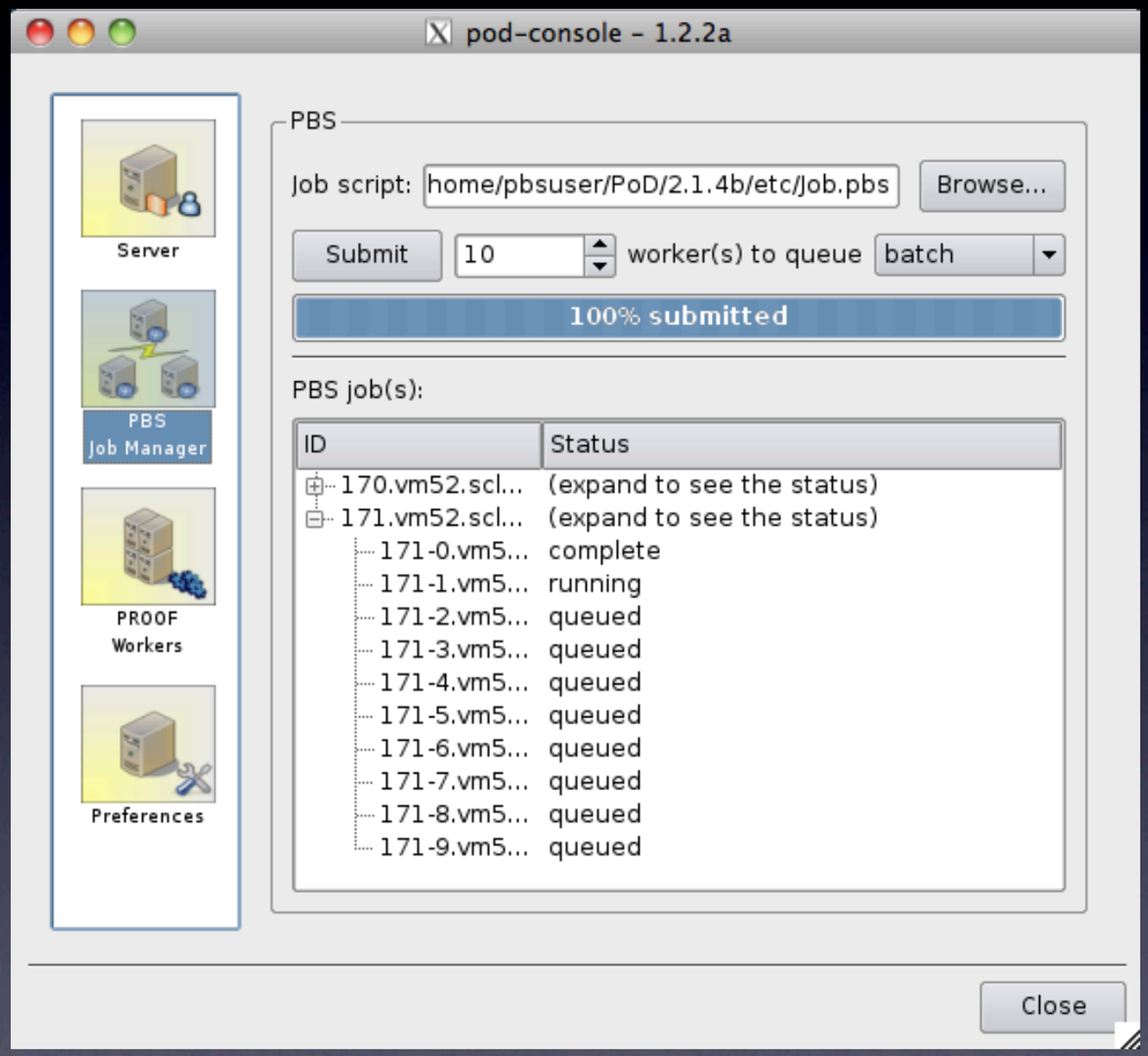




3 steps to set your private PROOF cluster up

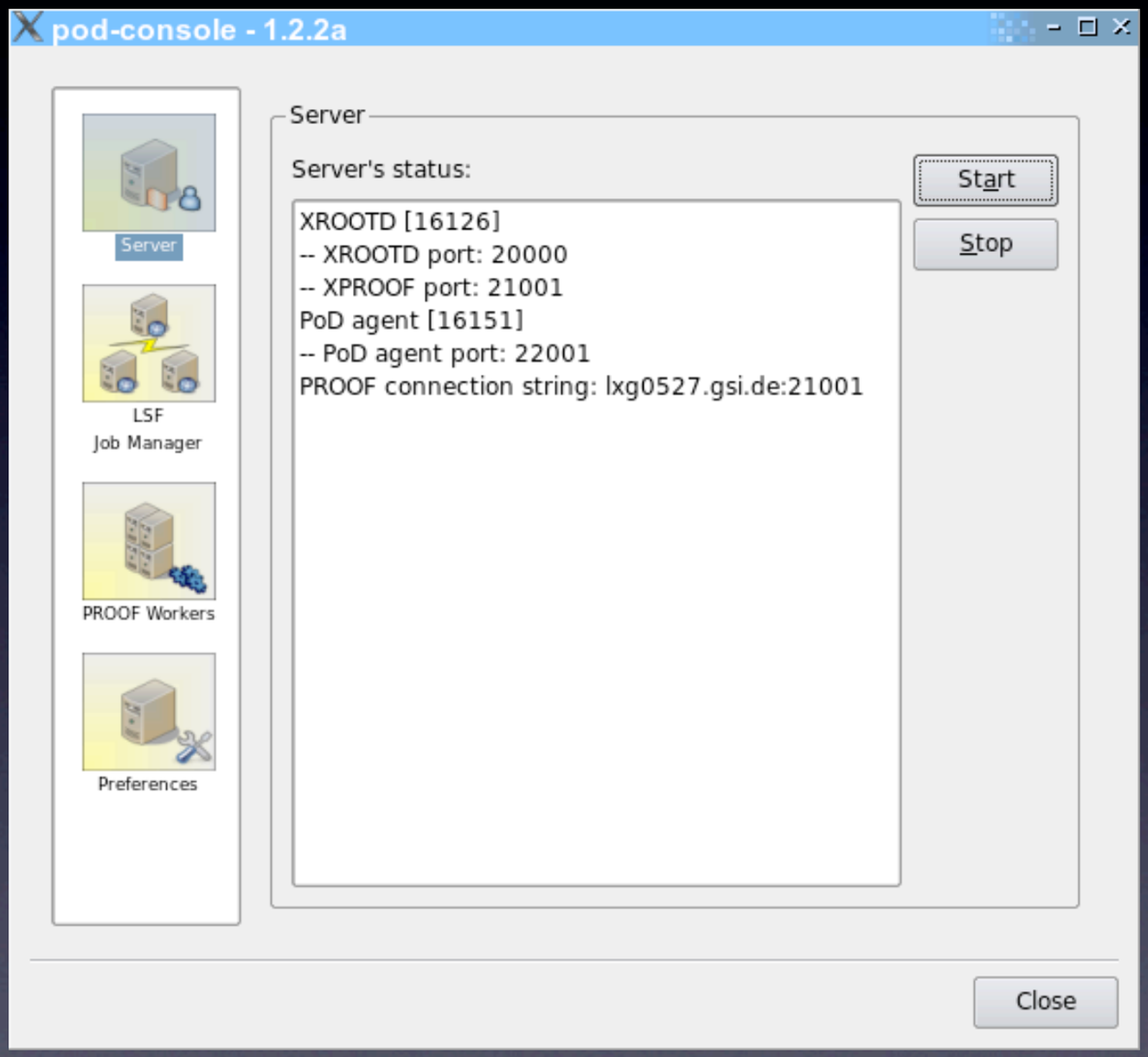
PoD server → Job Manager (gLite, PBS, LSF)

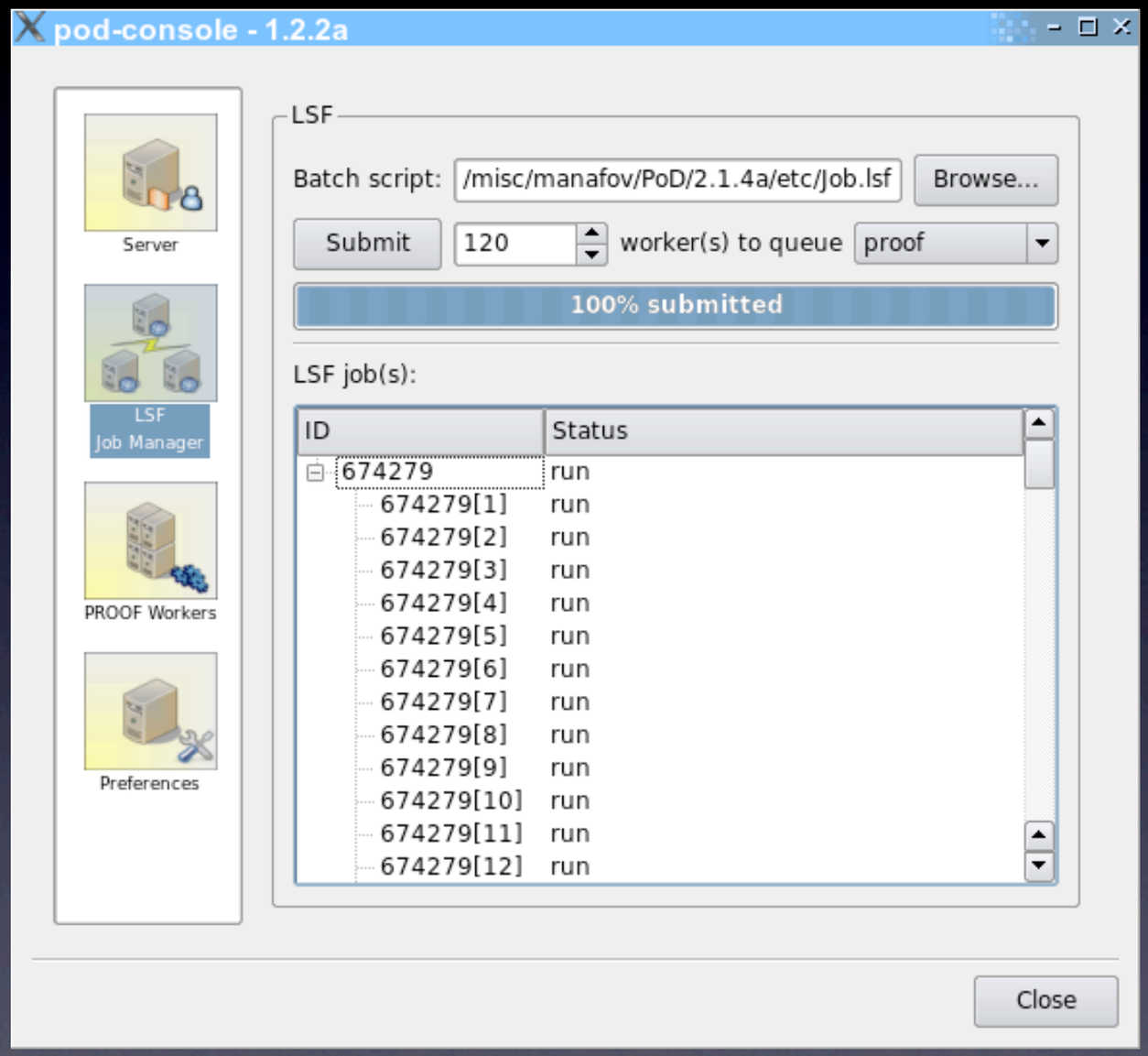




3 steps to set your private PROOF cluster up

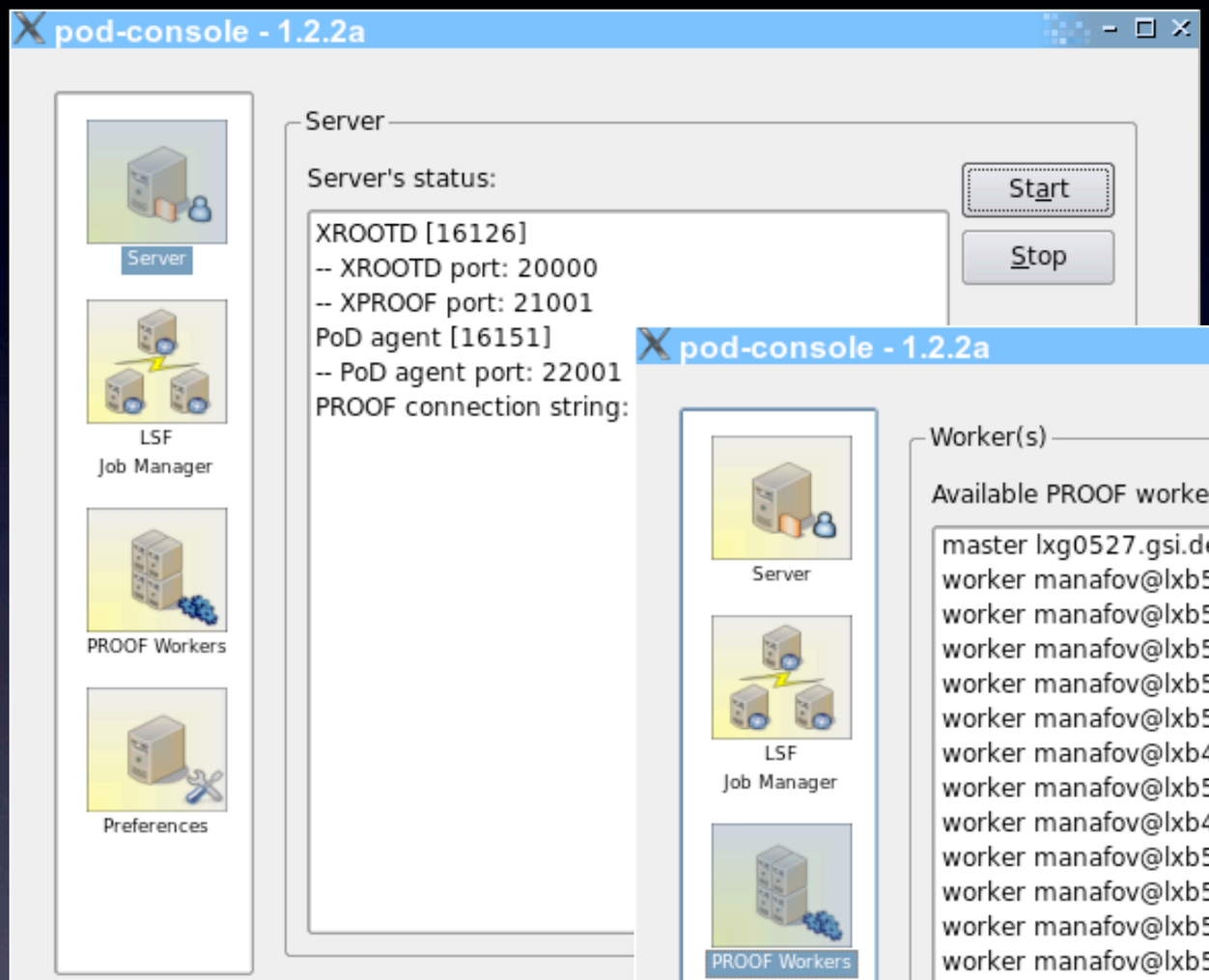
PoD server → Job Manager (gLite, PBS, LSF)

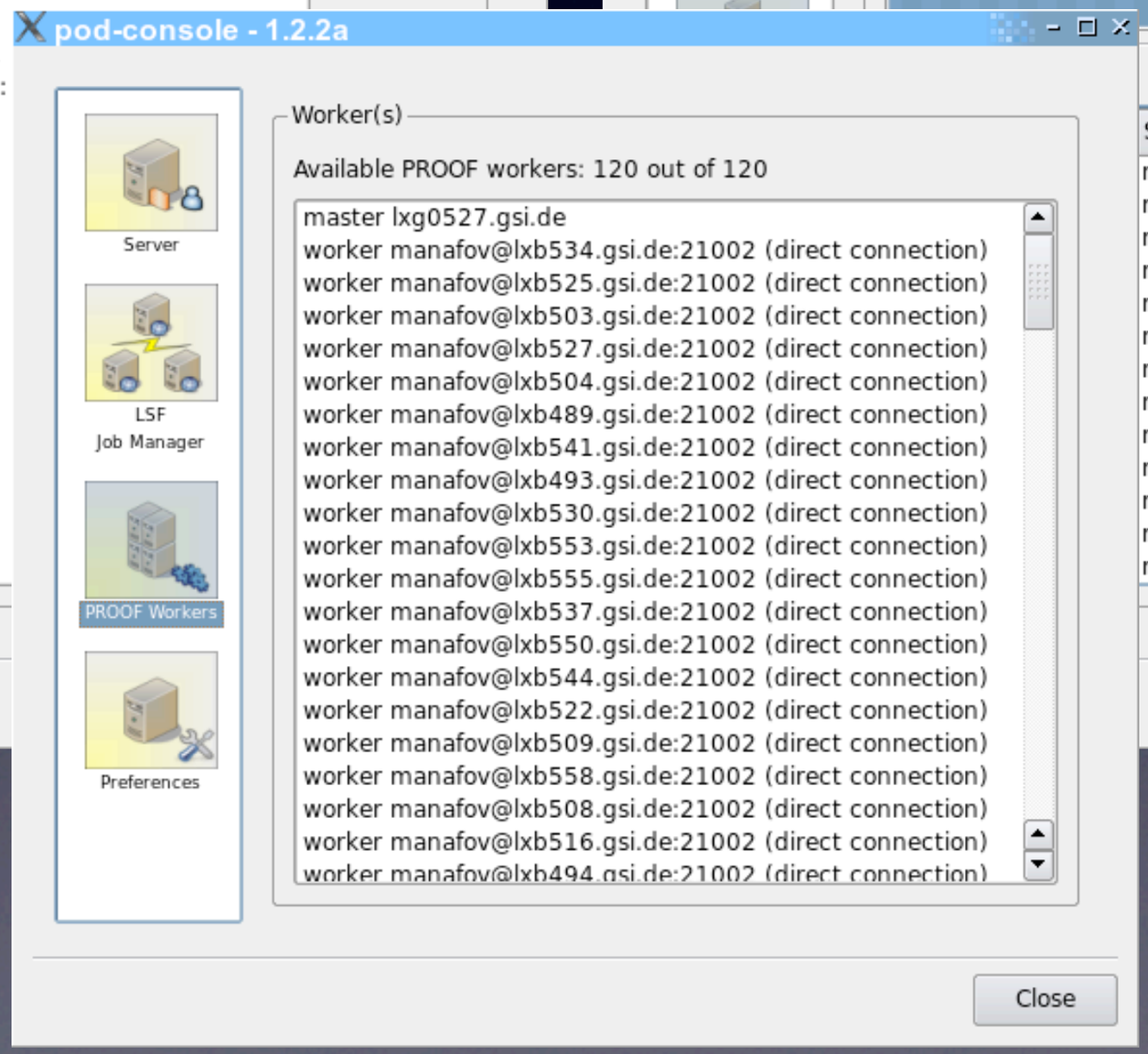


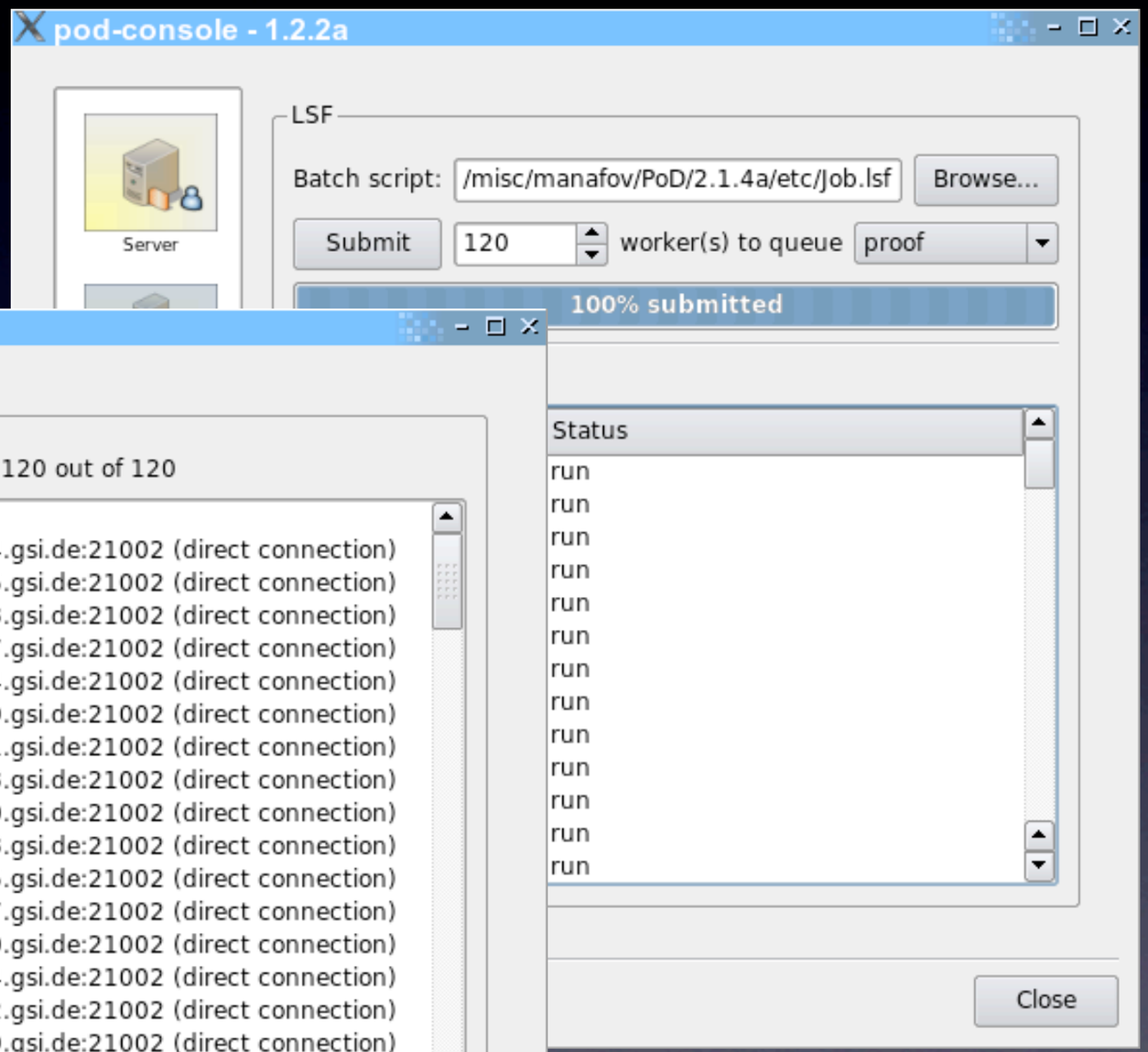


3 steps to set your private PROOF cluster up

PoD server → Job Manager (gLite, PBS, LSF) → your PROOF cluster







PoD v2.1.X

pod-console

PoD server management

PROOF workers monitor

Job Manager plug-in system

gLite

LSF

PBS

SSH

...

glite-api-wrapper (GAW)

PoDWorker

PoD utilities and command line tools

Configuration files (PoD, xrootd, PROOF)

pod-user-defaults

pod-agent

gLite Grid API

LSF API

PBS (torque) API

ssh/scp

...

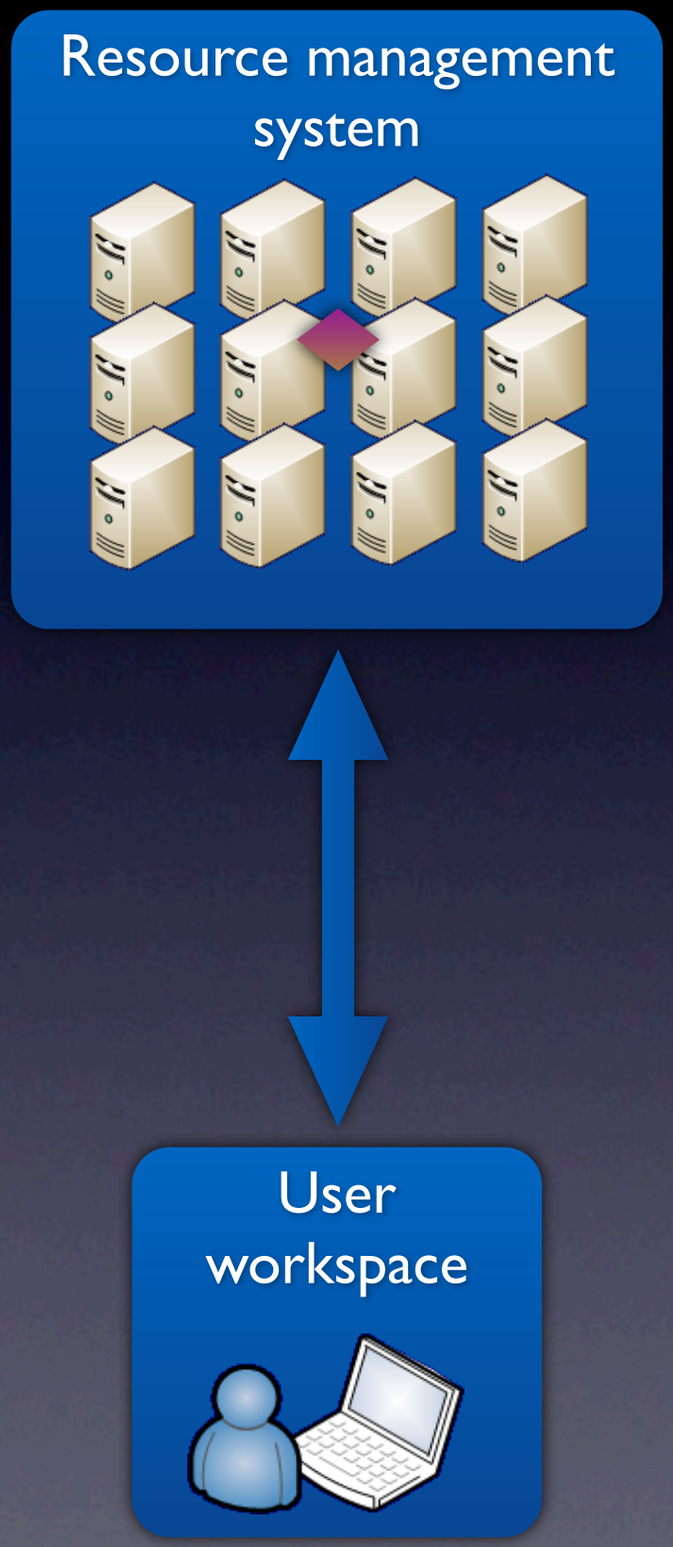
xrootd/xproof plug-in

Resource management system

User workspace

Resource management system

User workspace


Resource management system
User workspace

Resource management system

Worker node workspace

Firewall

User workspace

User workspace

Resource management system

Worker node workspace

Firewall

pod-agent server

xrootd

User workspace

User workspace

Resource management system

User workspace

Worker node workspace

PoDWorker job #1

xrootd

pod-agent worker #1

Firewall

pod-agent server

xrootd

User workspace

Resource management system

User workspace

Worker node workspace

PoDWorker job #1

xrootd

PoDWorker job #2

pod-agent worker #1

pod-agent worker #2

Firewall

pod-agent server

xrootd

User workspace

Resource management system

User workspace

Worker node workspace

PROOF worker #1

PROOF worker #2

PoDWorker job #1

xrootd

PoDWorker job #2

pod-agent worker #1

pod-agent worker #2

Firewall

pod-agent server

xrootd

PROOF master

User workspace

Resource management system

User workspace

Worker node workspace

PROOF worker #1

PROOF worker #2

PoDWorker job #1

xrootd

PoDWorker job #2

pod-agent

pod-agent

xrootd

PROOF master

pod-agent server

User workspace

Key features

- Easy to use
- GUI & Command-line
- Different job managers
- Multiuser/-core environment
- Native PROOF connections
- Packet-forwarding
- User defaults - configuration

PoD at GSI

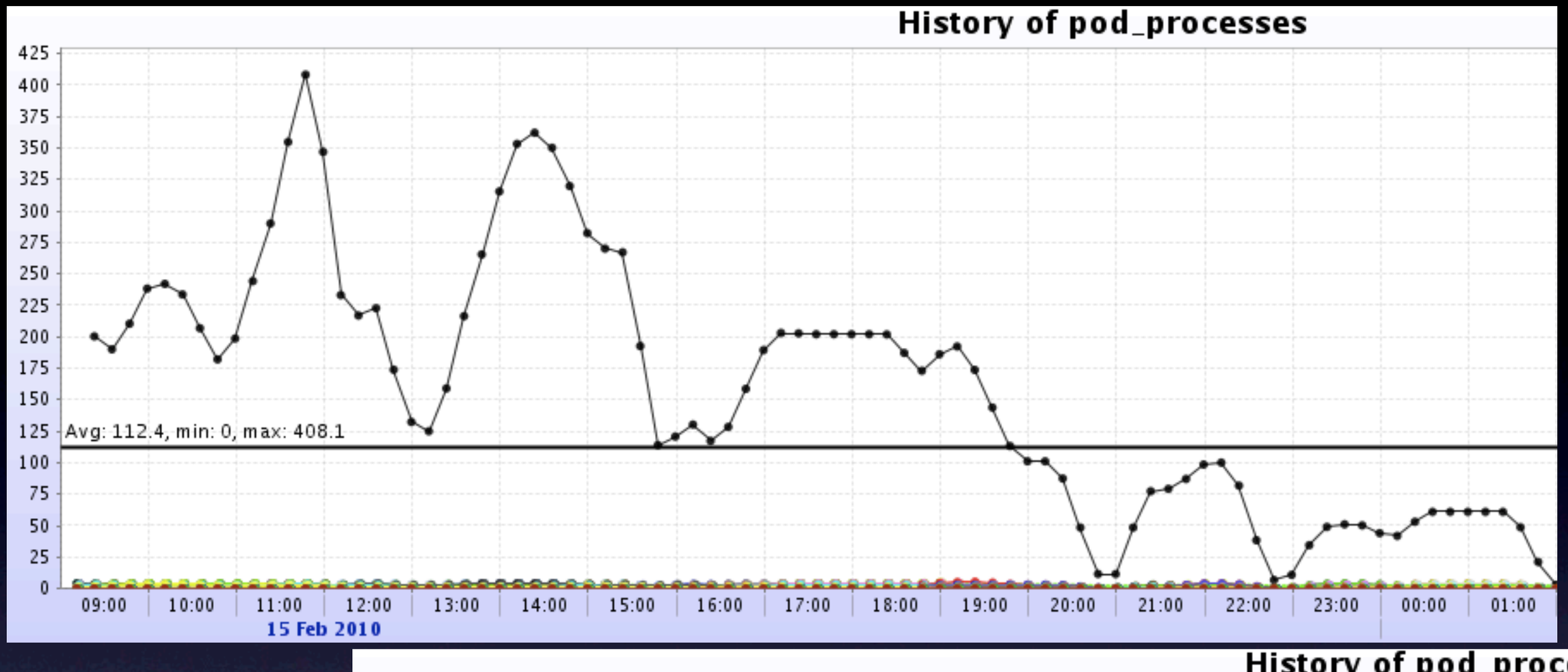
Dedicated LSF queue

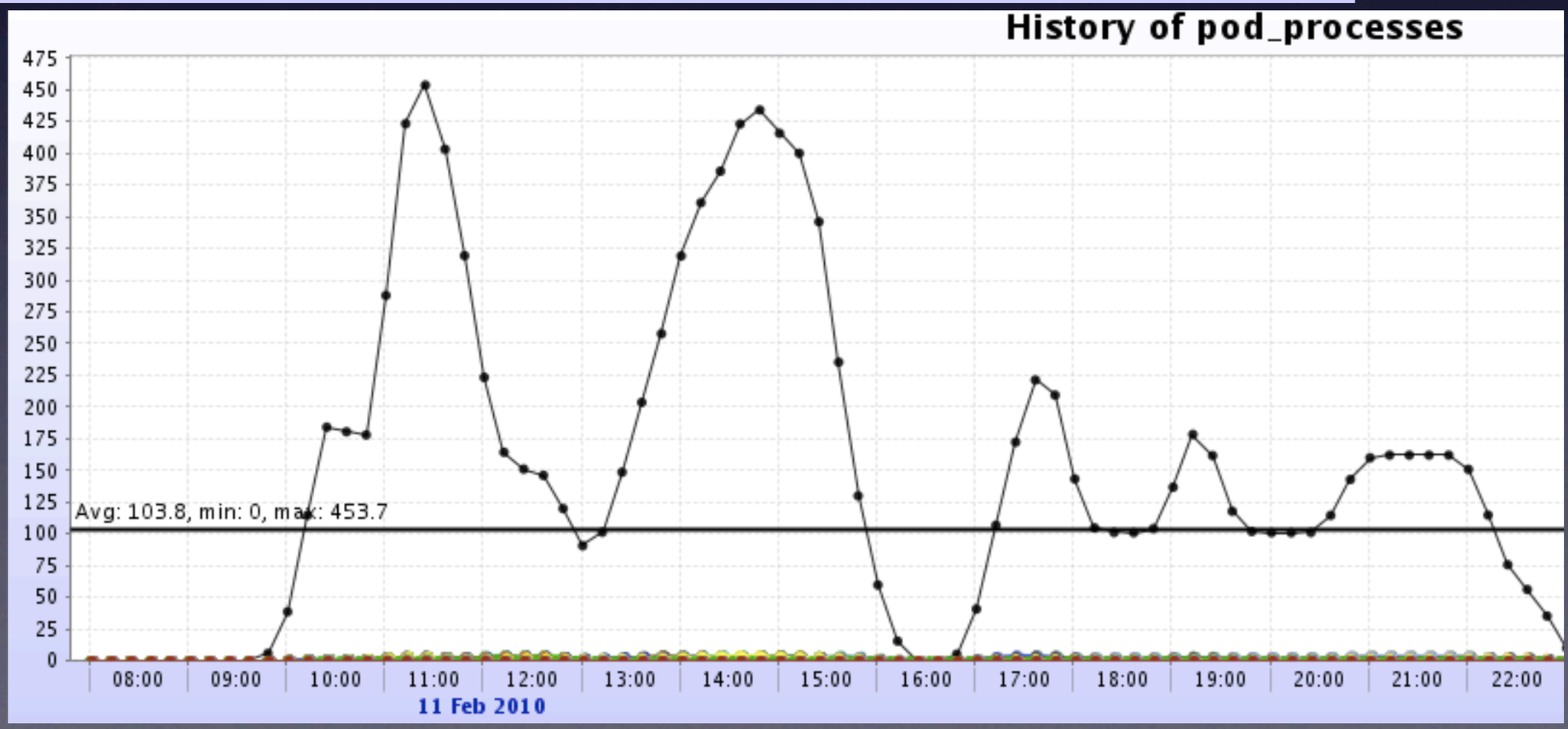
- preemptive, max. 120 jobs per user and max. 4 hours run-time per job.

Data located on the lustre FS.

Mainly used by ALICE group (GSI, Heidelberg, Münster).

In average we have 2-5 concurrent users with 20-120 workers each.


History of pod_processes
425
400
375
350
325
300
275
250
225
200
175
150
125
100
75
50
25
0
Avg: 112.4, min: 0, max: 408.1
09:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
22:00
23:00
00:00
01:00
15 Feb 2010


History of pod_processes
475
450
425
400
375
350
325
300
275
250
225
200
175
150
125
100
75
50
25
0
Avg: 103.8, min: 0, max: 453.7
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
22:00
11 Feb 2010

ToDo

- an SSH plug-in,
- “out of server” UI,
- a native Mac OS X implementation of UI,
- an AliEn plug-in.

http://pod.gsi.de

BackUp slides

pod-console - 1.2.1 <@lxb333>

Server

LSF Job Manager

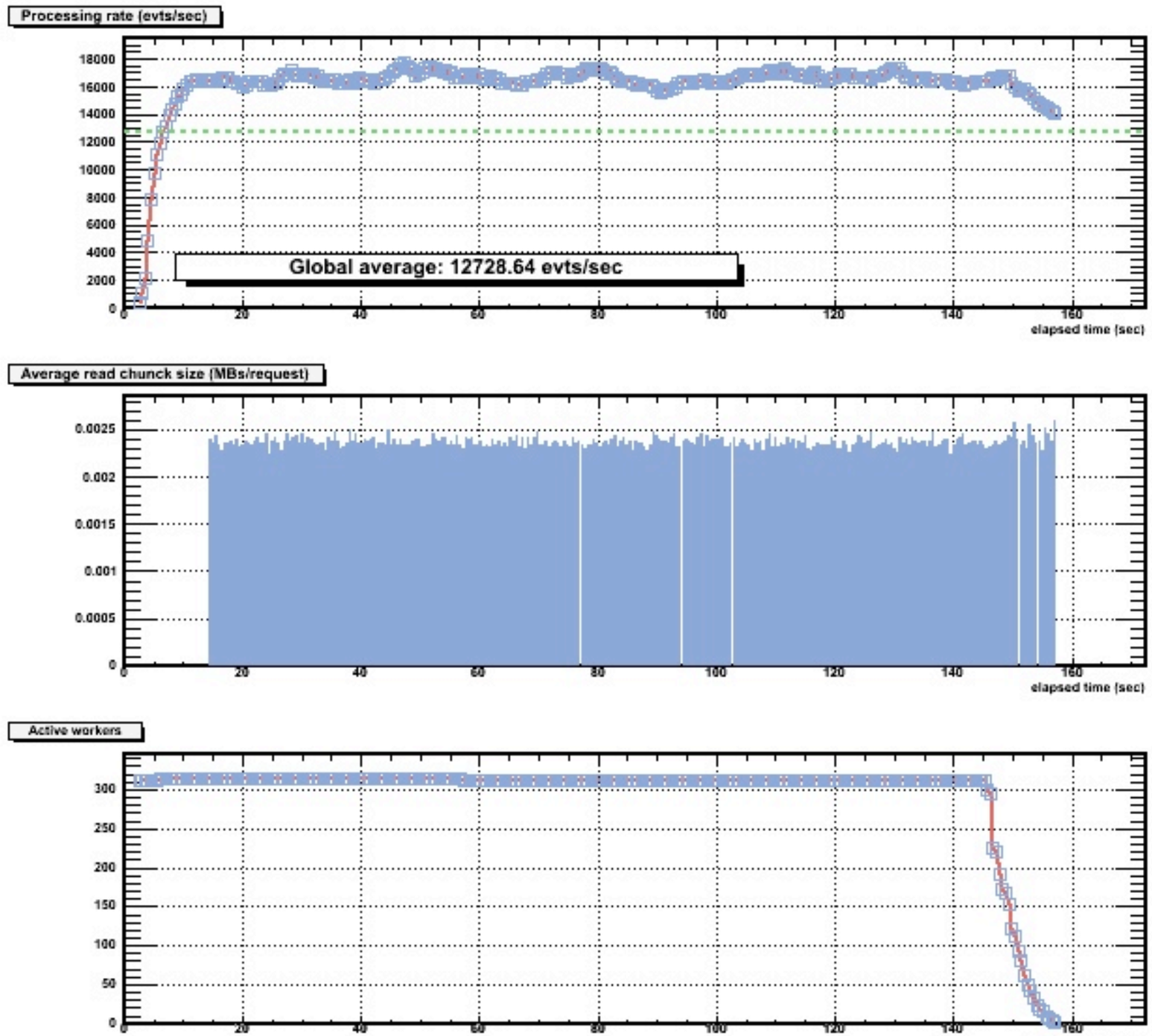
PROOF Workers

Preferences

Worker(s)

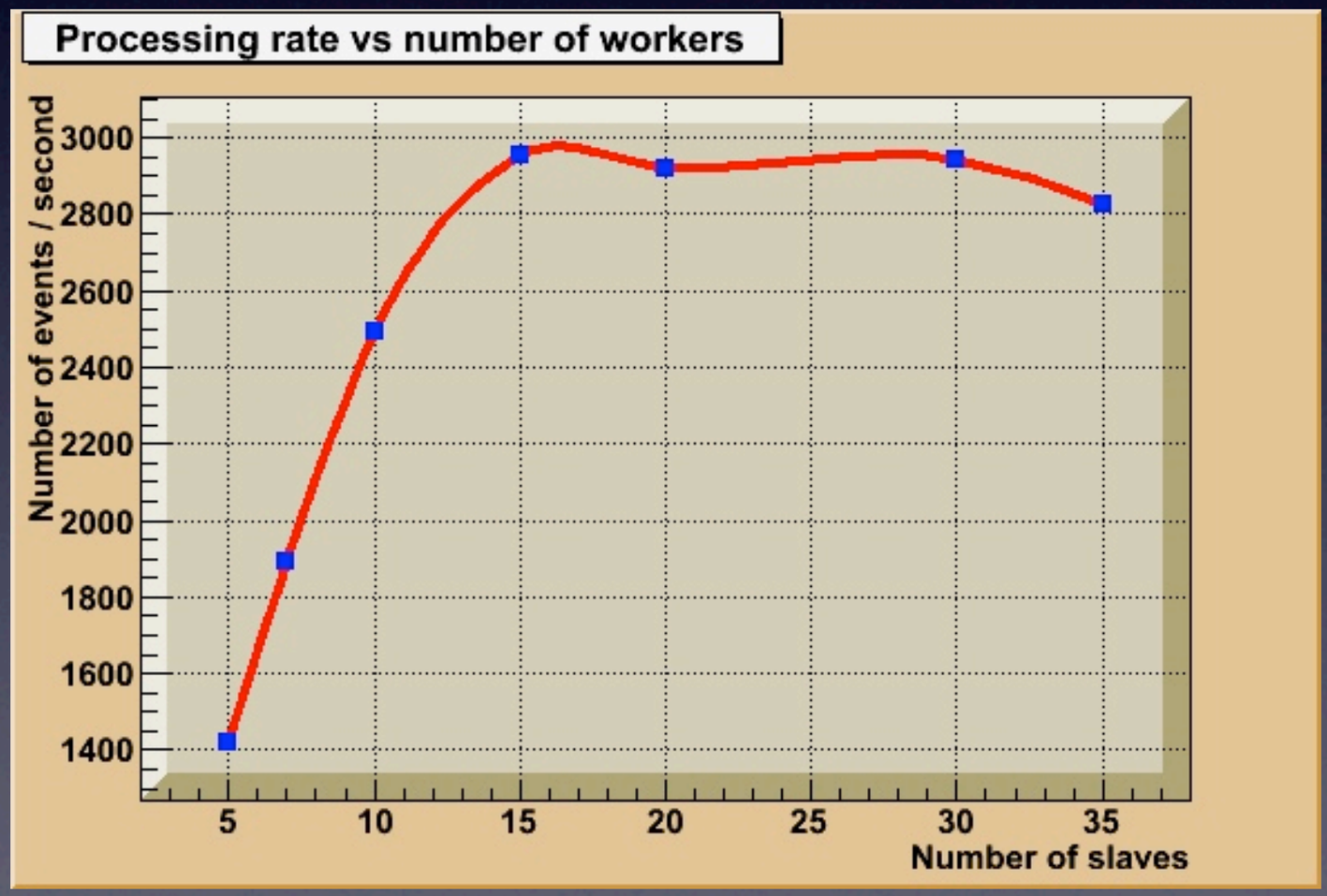
Available PROOF workers: 317 out of 319

master lxb333.gsi.de
worker mfasel@lxb511.gsi.de:21001 (direct connection)
worker mfasel@lxb544.gsi.de:21001 (direct connection)
worker mfasel@lxb541.gsi.de:21001 (direct connection)
worker mfasel@lxb521.gsi.de:21001 (direct connection)
worker mfasel@lxb529.gsi.de:21001 (direct connection)
worker mfasel@lxb556.gsi.d
worker mfasel@lxb400.gsi.d
worker mfasel@lxb505.gsi.d
worker mfasel@lxb449.gsi.d
worker mfasel@lxb407.gsi.d
worker mfasel@lxb396.gsi.d
worker mfasel@lxb519.gsi.d
worker mfasel@lxb462.gsi.d
worker mfasel@lxb495.gsi.d
worker mfasel@lxb386.gsi.d
worker mfasel@lxb466.gsi.d
worker mfasel@lxb429.gsi.d
worker mfasel@lxb477.gsi.d
worker mfasel@lxb384.gsi.d
worker mfasel@lxb524.gsi.d



User experience
PoD & gLite

T-3 for ATLAS,
the gLite site is IN2P3-CPPM
DPM + xrootd



the diagram from Karim Bernardet <bernardet@cppm.in2p3.fr>
(Marseille)