

PROOF-CAF meeting

27/01/2009

- New ROOT v5/22-00
- Dynamic PROOF over AliEn
- New PROOF clusters
- Open Issues
- AoB

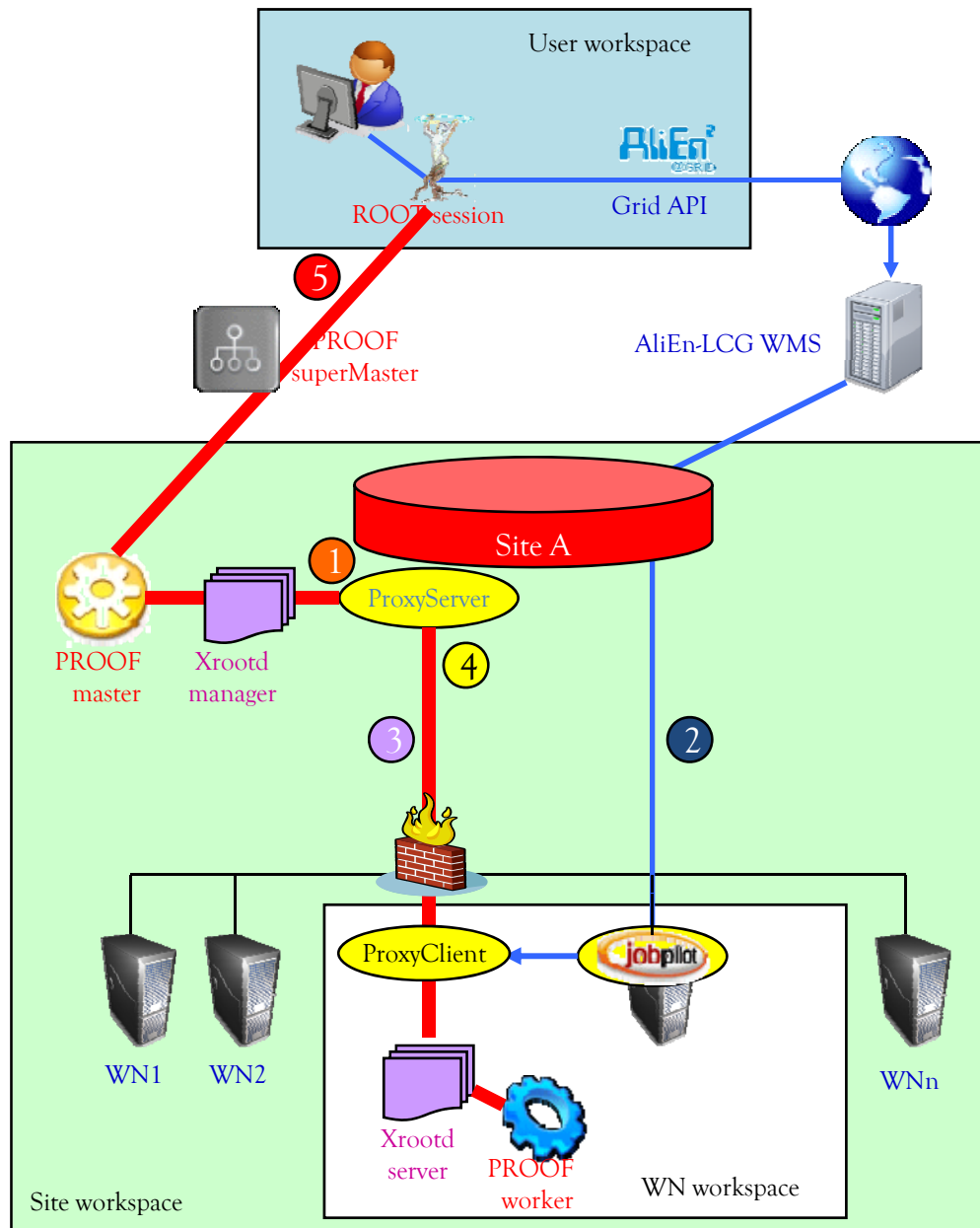
(<http://indico.cern.ch/conferenceDisplay.py?confId=50986>)

New ROOT v5/22-00

- ROOT v5/22-00 deployed to CAF
 - AliRoot v4-16-Release deployed to afs
 - AF package v4-16-Release + user conf script ready on lxplus
 - Tutorial tasks successfully run on the test instance (Port 11093)
- Keep next tutorial session (Thu 29th) with v5/21-01-alice
- Authentication
 - New plug-in for GSI xrootd auth finds a mismatch between the server name and the one in the server certificate

```
server certificate CN 'lxb6041.cern.ch' does not match the expected
format(s): ' */lxfsrd0506.cern.ch' (default); exceptions are
controlled by the env XrdSecGSISRVNAMES
```
 - Provide each machine (or at least the master) with its own certificate? (Latchezar)

Dynamic Distributed PROOF



- 1 A ProxyServer service starts Xrootd and PROOF
- 2 Pilot Grid jobs are submitted to the Grid to start ProxyClients
- 3 A ProxyClient starts an Xrootd server and registers to the ProxyServer
- 4 A ProxyServer keeps the list of all the workers running at the WNs
- 5 A User PROOF session connects to the (super)Master that, in turns, starts PROOF workers

Current Minimal Setup

VoBox

AliEn WMS

User box

```
[voalice06] ~-> netstat -vatn | grep LISTEN
tcp 0 0 0.0.0.0:20000 0.0.0.0:* LISTEN
tcp 0 0 0.0.0.0:1093 0.0.0.0:* LISTEN
tcp 0 0 0.0.0.0:1094 0.0.0.0:* LISTEN

[voalice06] ~-> less proof.conf
master voalice06.cern.ch
```

```
{ Executable=~/.bin/gLitePROOF.sh";
  Packages={"VO_ALICE@ROOT:v5-21-01-alice",...};
  Jobtag = {"comment: Pilot job to start Proof Workers"};
  .....
  InputFile={"LF:~/proofAgent/xpd.cf",
             "LF:~/proofAgent/proofagent.cfg.xml"};
  InputDataCollection="LF:~/LHC08c18_37872,nodownload";
```

```
[aliendb06c.cern.ch:3307] ~/ > ps -s -id 22792005
mmeoni 22792005 RS ~/.bin/gLitePROOF.sh
mmeoni -22792006 R 00:00:00 ~/.bin/gLitePROOF.sh
mmeoni -22792007 R 00:00:00 ~/.bin/gLitePROOF.sh
mmeoni -22792008 R 00:00:01 ~/.bin/gLitePROOF.sh
mmeoni -22792009 R 00:00:01 ~/.bin/gLitePROOF.sh
mmeoni -22792010 R 00:00:01 ~/.bin/gLitePROOF.sh
mmeoni -22792011 R 00:00:00 ~/.bin/gLitePROOF.sh
mmeoni -22792012 R 00:00:01 ~/.bin/gLitePROOF.sh
```

```
[voalice06] ~-> less proof.conf
master voalice06.cern.ch
worker alicesgm@localhost port=20001
worker alicesgm@localhost port=20002
worker alicesgm@localhost port=20003
worker alicesgm@localhost port=20004
worker alicesgm@localhost port=20005
worker alicesgm@localhost port=20006
worker alicesgm@localhost port=20007
worker alicesgm@localhost port=20008
worker alicesgm@localhost port=20009
worker alicesgm@localhost port=20010
```

```
root [0] TProof::Open("alicesgm@voalice06.cern.ch")
Starting master: opening connection ...
Starting master: OK
Opening connections to workers: OK (10 workers)
Setting up worker servers: OK (10 workers)
PROOF set to parallel mode (10 workers)
(class TProof*)0xafa6e0
root [1]
```

Dynamic PROOF: problems...

- Will the VoBox be capable to sustain such additional load?
 - Xrootd (messaging traffic), ProxyServer (flat traffic routing)
 - Local Proof Master
 - Low load during coordination activity among the workers
 - High load during workers' startup and final output merging
 - PROOF file merging feature needs a scratch space on disk, normally on the master
 - PROOF files lookup and validation steps require to open all the files (trees + get #events) in the input datasets
 - **Possible alternative: use the 1st AliEn WN as master?**
- PROOF users and Grid users
 - Use proof setting xpd.multiuser=1
- Send pilot jobs in advance (Pablo) to pre-allocate workers
 - Cannot re-use workers of userA for userB (!)
- Integrate the ProxyServer in MonALISA for automatic deploy/install/monitor

New PROOF clusters at CERN

a) alice-HLT

- Meeting with Cvetan and Stefan Boettger (19/11/08)
 - Pledged ~600 cores, initial test with 1 master + 3 workers
 - Cluster up and running, xrootd and xproofd centrally configured (afs), no local storage, tutorial exercises successfully run on CAF datasets (ROOT v5-21-01-alice)
- 19/01: emailed for additional CPU (and for xrootd storage). No answer yet...

b) CERNMAC

- Xrootd installation required lots of fix (autotools, wget, getopts, hostname, killall, pgrep, pkill)
- Still not working:
 - No file with folder prefix is staged (bug submitted to Andy. Darwin defines the flag to force file path creation to be short instead of int)
 - Staging from Global Redirector (mps scripts)
- Given the new HLT cluster, CERNMAC will be used as AliEn SE

Open Issues (1)

Follow up on Staging

a) Modification of stager daemon to use bulk prepare (Marco)

- DS manager log:

```
TProofAliceManager::ProcessDataSet (2008-11-13 14:26:59.238810000+0100):  
Processing dataset LHC08d_62084 (user mmeoni, group PWG0)  
In parent; waiting for 8569...  
TProofAliceManager::ProcessDataSet: Requesting staging for 94 entries  
Info in <TXNetSystem::Prepare>: buffer ready: issuing prepare ...  
Srv err: session is active  
TProofAliceManager::ProcessDataSet: Bulk Stage returned 1.  
Info in <TProofDataSetManagerFile::ScanDataSet>: processing a maximum of 500  
files
```
- Bug submitted to Andy, cc'ing Andreas, hypothesis of slow reactivity by the stager script (but the error pops up after few seconds)

b) Replace of touch file by bulk prepare: not yet tested (Marco)

c) IsStaged not working with cmsd (Andy)

- After being staged, a file is not automatically seen by requests such 'stat'
- Nevertheless, opening a file works fine
- Bug submitted to Andy...

Open Issues (2)

DS improvement: co-locate (Marco)

- Staging related files on the same disk (ESD, Kinematics..)
- DS staging: current user use-case:

```
gridColl = gGrid->OpenCollection(alien_collection_name)
proofColl = gridColl->GetFileCollection();
gProof->RegisterDataSet("dsname", proofColl);
```

- Presently in AliEn we can only create a collection of collections...
 - ...but not a collection of related files
 - Cardinality of the individual collections would not be guaranteed
- DS manager needs to be heavily modified
 - Co-locate only possible at 1st registration, not afterwards

Open Issues (3)...

Timeout for file merging (Gerri)

- (in case xrootd is unavailable)

Local usage of par files (Fons)

Bug in package rebuilding (Fons)

- Verify if cp -a is used in par file target

Garbage collector (Marco)

- No necessary at the moment (2.5/30TB used)