
Fed cloud project status report

Andrey Kiryanov (PNPI)
Dimitrii Krasnopevtsev (MEPhI)
Andrey Zarochentsev (SPbSU)

Requirements for a federated storage

- Single entry point
- Universality: should be usable by at least four major LHC experiments
- Scalability: it should be easy to add new resources
- Data transfer optimality: transfers should be routed directly to the disk servers avoiding intermediate gateways and other bottlenecks
- Fault tolerance: replication/redundancy of core components
- Built-in virtual namespace, no dependency on external catalogues

Options

We had to find a software solution that supports federation of distributed storage resources. This very much depends on a transfer protocol support for redirection. Two protocols that are good at it are xroot and HTTP.

- HTTP-based federation is implemented in DynaFed software developed by IT/SDC group at CERN. This software is highly modular and only provides a federation frontend while the storage backend (s) have to be chosen separately. It would be interesting to try it out but we were looking for more all-in-one solution.
- xroot-based solution is EOS. It's also developed at CERN (we knew where to ask for help), has characteristics closely matching our requirements, and is already used by all major LHC experiments. We decided to give it a try.

Why EOS?

- 1) Complete project
- 2) All WLCG VOs support xrootd

EOS Benefits

1) Requirements:

- Single entry point
- Universality
- Scalability
- Data transfer optimality
- Fault tolerance
- Built-in virtual namespace

2) Extras:

- Optimization of file selection by group or geo-tag
- Automated replication
- Data migration
- Documentation

Initial tests: testbed structure

Initial plan was to test 3 sites (NRC KI, PNPI, SPbSU) with storage servers + “head” at CERN

Due to various reasons most of the presented tests were done with only two sites: PNPI and SPbSU

- NRC KI - MGM (slave/master) + FST + PerfSONAR (T1) + UI
- PNPI - MGM (slave/master) + FST + PerfSONAR + UI
- SPbSU - MGM (slave/master) + FST + PerfSONAR + UI
- CERN - MGM (master) + PerfSONAR + UI

(green means it's already there)

Initial tests: testing plan

- Proof-of-concept test: install and configure distributed EOS, hook up GSI authentication, test basic functionality (file/directory create/delete, FUSE mount, access permissions)
- Reliability test: MGM master-slave migration
- Performance tests: file and metadata I/O, real-life experiment software, network
- Redirection impact test: check if there's performance degradation with remote “head” node
- Data locality test: evaluate EOS geo-tags role in data distribution

Prototype structure: software and tests

- Base OS: SL6/x64
- Storage system: EOS Aquamarine
- Authentication scheme: GSI
- Tests
 - Bonnie++ (file I/O test on FUSE-mounted file system)
 - ATLAS test: standard ATLAS event reconstruction workflow with Athena (thanks to Dmitrii)
 - ALICE test: sequential ROOT event processing (thanks to Peter)
 - PerfSONAR

Problems

- Install script does not work
- GSI authorization is not documented
- MGM master-slave migration does not work as expected (from docs)
 - After a meeting with Andreas we managed to get it to work but still with problems

First results without conclusions

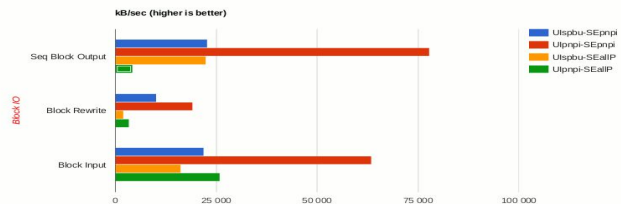
ATLAS tests

bonnie2gchart - Block IO

BNNIE++ tests

<http://alice22.spbu.ru/bonnie/?t=blockio>

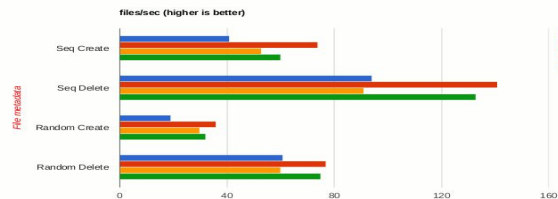
[= index](#)



bonnie2gchart - File metadata

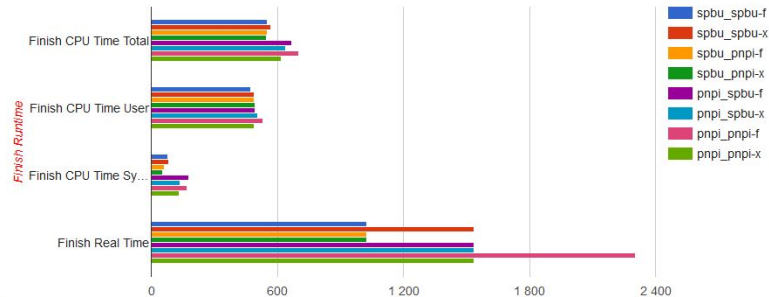
<http://alice22.spbu.ru/bonnie/?t=metadata>

[= index](#)



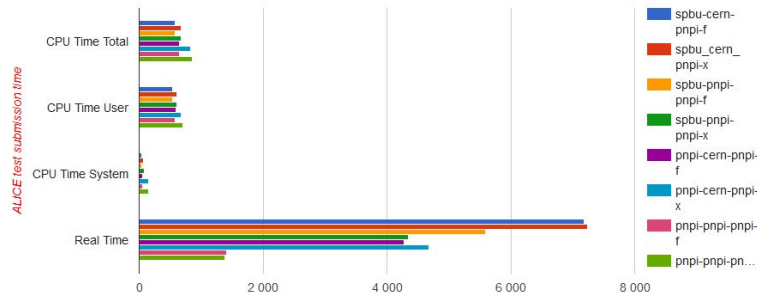
1 of 1

milliseconds (lower is better)



ALICE tests

seconds (lower is better)



Full test plan

Bonnie++

- a. PNPI FST local
- b. SPbSU FST local
- c. UI PNPI to ALL
- d. UI SPbSU to ALL
- e. UI CERN to ALL

ATLAS (2 files on two FST)-----****

- a. UI PNPI to ALL (fuse, xrootd)
- b. UI SPbSU to ALL (fuse, xrootd)
- c. UI CERN to ALL (fuse, xrootd)

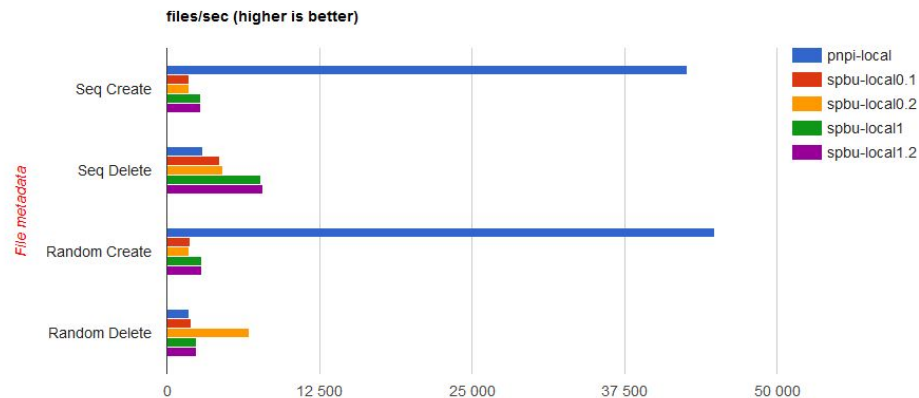
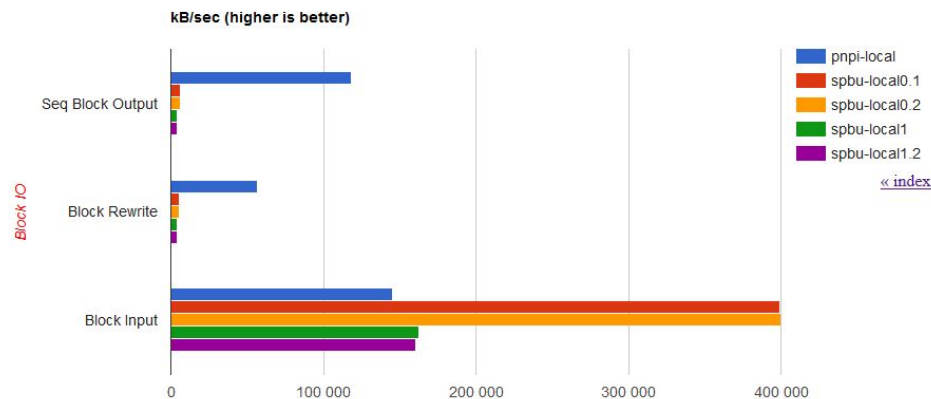
ALICE (322 files on two FST)

- a. UI PNPI to ALL (fuse, xrootd)
- b. UI SPbSU to ALL (fuse, xrootd)
- c. UI CERN to ALL (fuse, xrootd)

PerfSonar

- a. CERN-PNPI
- b. CERN-SPbSU
- c. PNPI-SPbSU

Bonnie local tests - different pools on SPbSU



Problem: need to remove 2nd pool from SPbSU

- 1) EOS supports “fs drain” function.
 - a) Tried by the docs but to no avail
- 2) Consultation with Andreas did not help
 - a) We can use “find”, “fileinfo” and “file move”
 - b) “file move” can not move all files in one turn

Finally after some back-and-forth 2nd pool was removed

Tests in progress