XrootD in cloud environment

Integration of XRootD
into the cloud infrastructure for ALICE data analysis

MIKHAIL KOMPANIETS, OKSANA SHADURA, PAVLO SVIRIN, VOLODYMYR YURCHENKO, ANDREY ZAROCHENTSEV

CERN, Saint Petersburg State University (RU), National Technical Univ. of Ukraine "Kyiv Polytechnic Institute"(UA), Bogolyubov Institute for Theoretical Physics(UA)

Abstract

Cloud technologies allow easy load balancing between different tasks and projects. From the viewpoint of the data analysis in the ALICE experiment, cloud allows to deploy software using Cern Virtual Machine (CernVM) and CernVM File System (CVMFS), to run different (including outdated) versions of software for long term data preservation and to dynamically allocate resources for different computing activities, e.g. grid site, ALICE Analysis Facility (AAF) and possible usage for local projects or other LHC experiments.

We present a cloud solution for Tier-3 sites based on OpenStack and CEPH distributed storage with an integrated XRootD based storage element (SE). One of the key features of the solution is that CEPH has been used as a backend for Cinder Block Storage service for OpenStack and in the same time as a storage backend for XRootD, with redundancy and availability of data preserved by CEPH settings. For faster and easier OpenStack deployment the PackStack solution, which is based on the Puppet configuration management system, is applied. CEPH installation and configuration chains are structured, converted to Puppet manifests and integrated into Packstack. This solution can be easily deployed, maintained and used even in small groups with limited computing resources and small organizations which usually have lack of IT support. The proposed infrastructure has been tested on two different clouds (SPbSU & BITP) and integrates successfully with the ALICE data analysis model. The present work is supported in part by Saint-Petersburg State University research grant 11.38.197.2014.

Idea of distributed cloud storage

- Typical Openstack principle schema: Controller Node, Network Node, Compute Nodes (Figure 1)
- All resources (CPU, RAM, storage) are shared between Openstack with CEPH, RadosFS/CEPHFS & XrootD services (Figure 2)
- Fast deployment using one unified answer file (Packstack developed by RDO community)
- Easy deployment through puppetized solution (you need to have only one node as an entrance point)
- We had tested solution using Bonnie++ benchmarks ,PROOF Benchmarking & MonAlisa tests which proves full compatibility with any types of data analyses (local computing, Proof analyses, grid job processing)
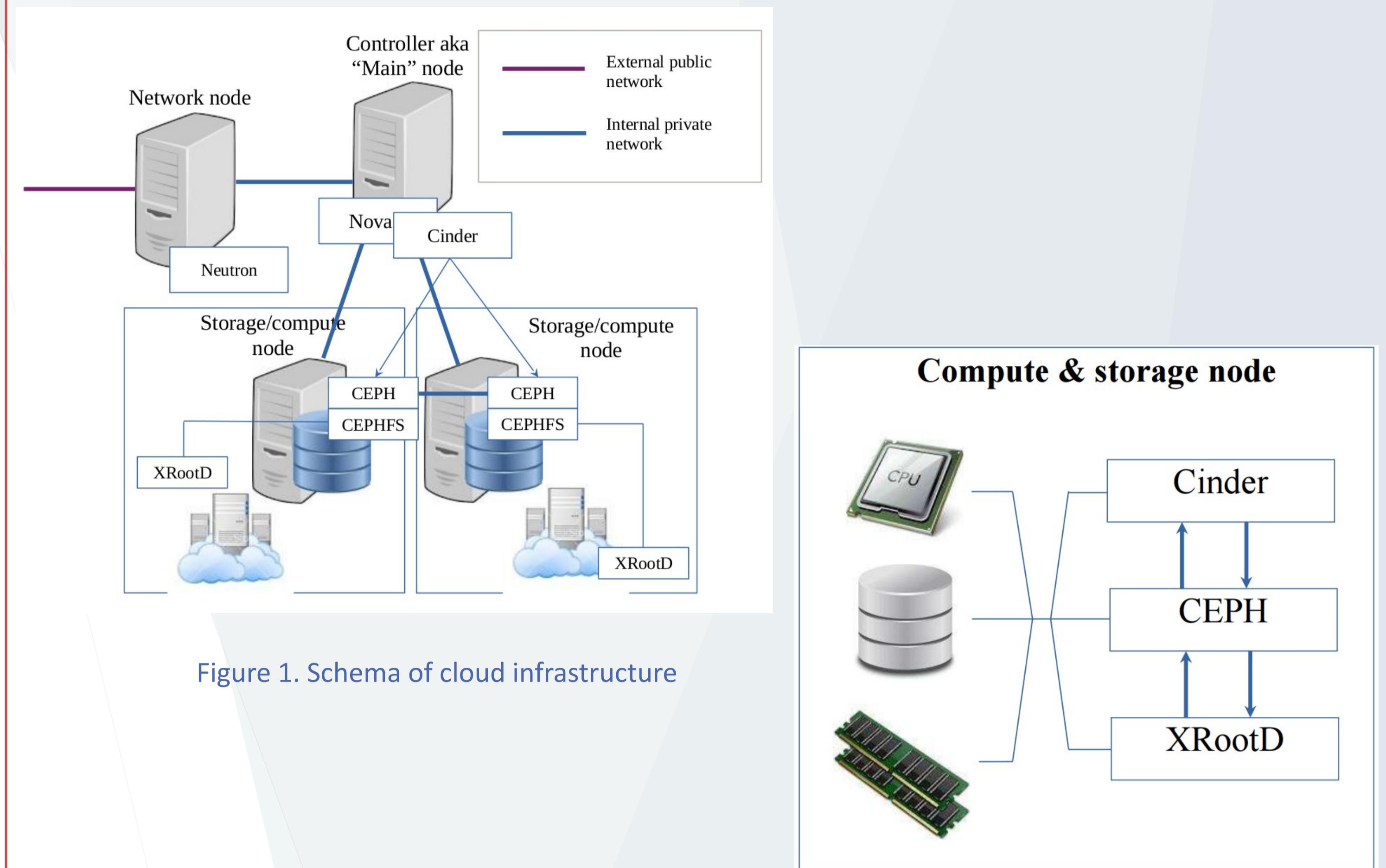
Deployment schema

Figure 1. Schema of cloud infrastructure

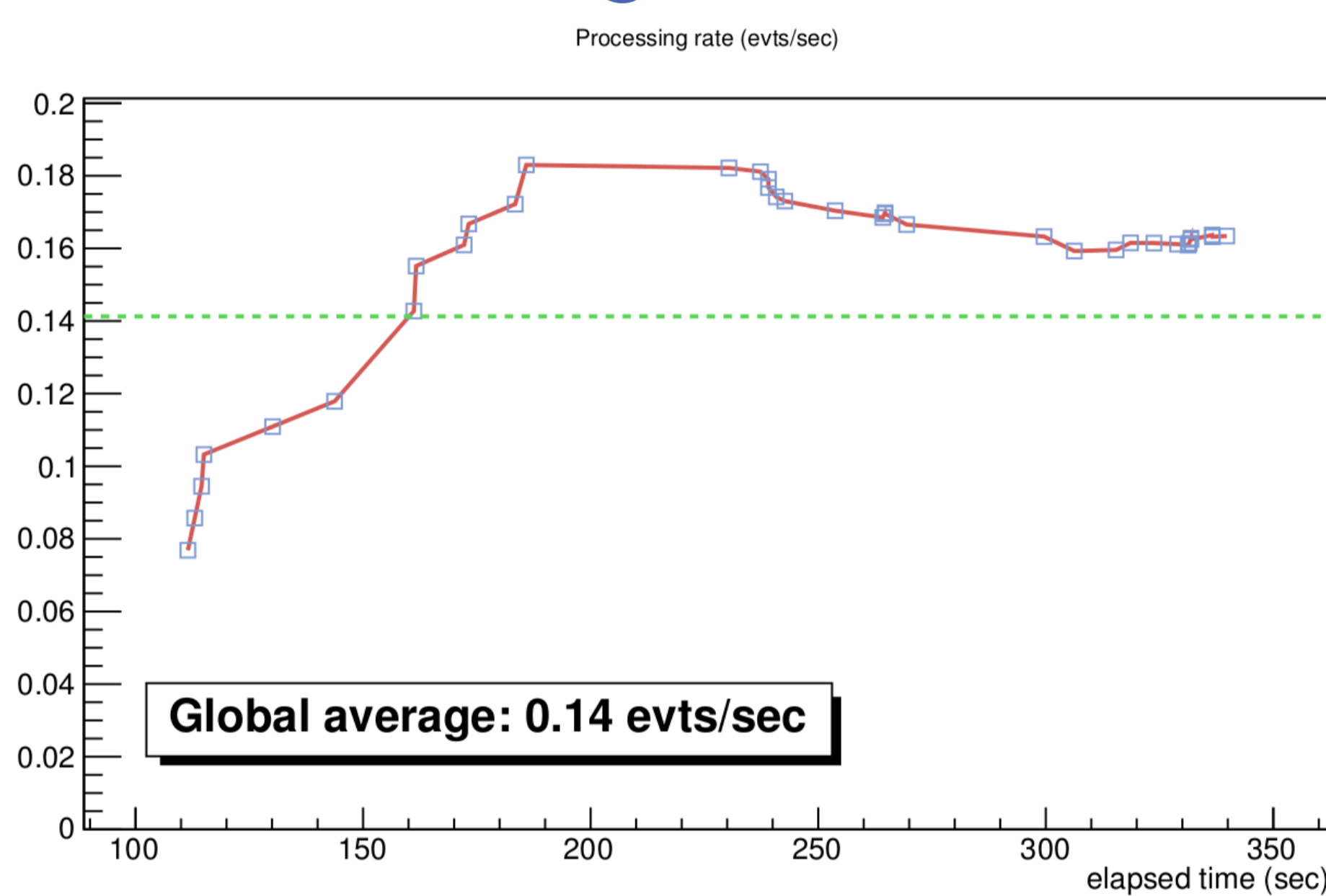Figure 2. Concept of resources usage

Benchmarking
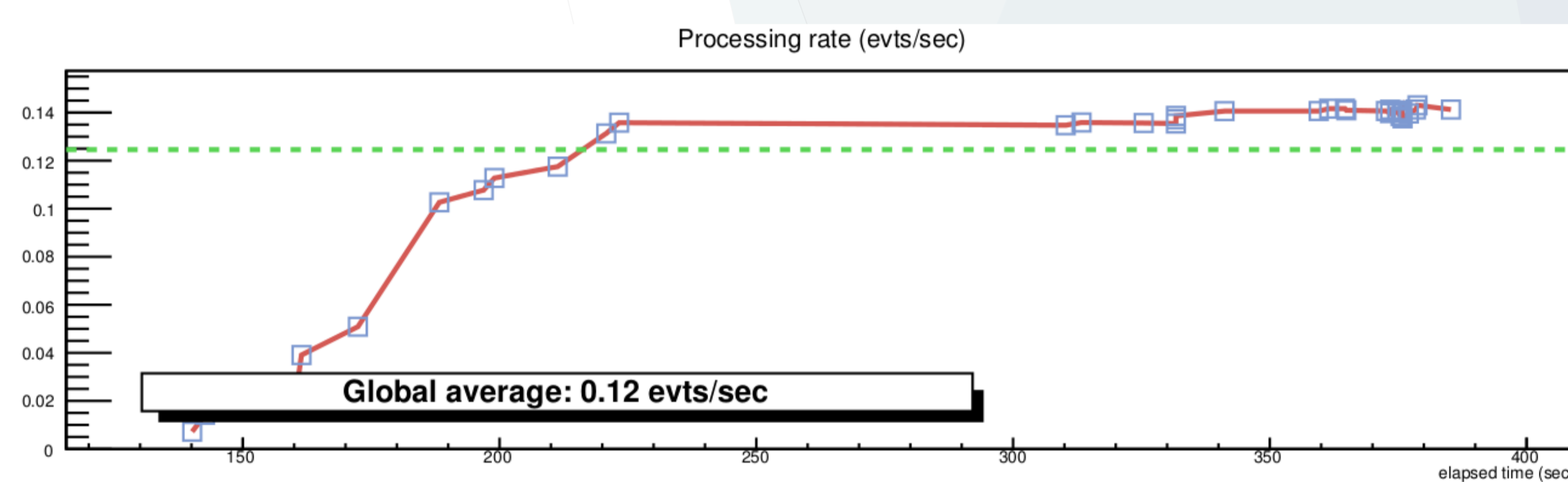

Figure 3. PROOF benchmarking on cloud XRootd

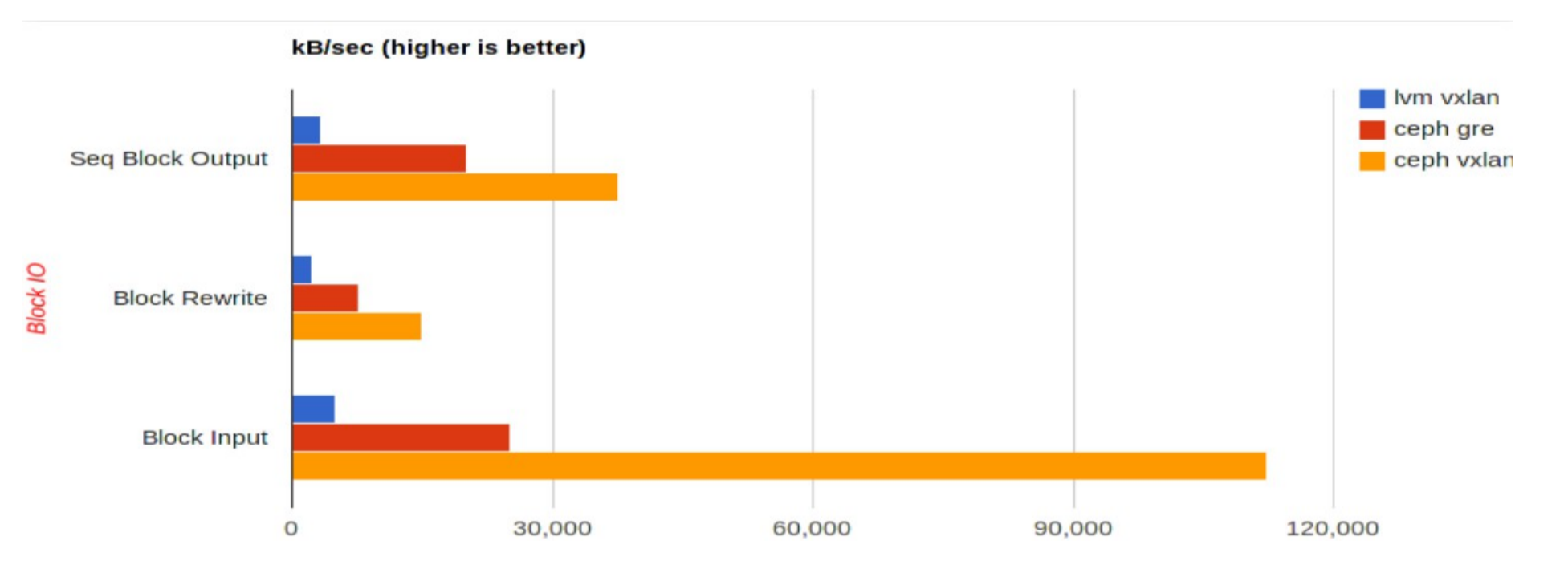Figure 4. PROOF Benchmarking on local XrootD

Figure 5. Bonnie++ benchmarks on LVM & CEPH based VMs

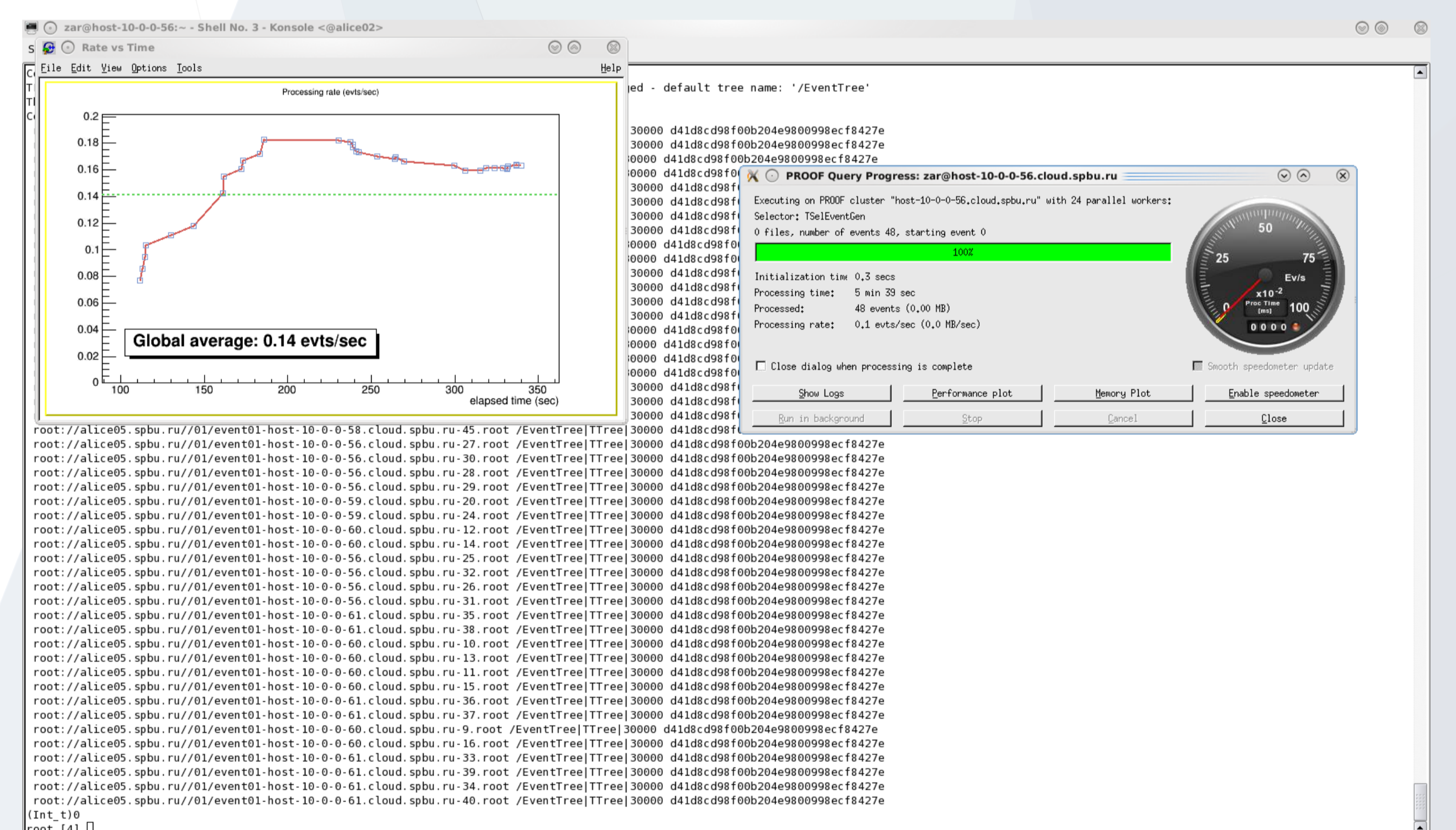Figure 6. MonAlisa testing of CEPH based storage

Figure 5. PROOF Benchmarking snapshot from XrootD on 24 cores (6 VMs)

Problems:

- **RadosFS**
  In first tests we were using RadosFS solution, but all current attempts were unsuccessful because of problems with permissions (chmod) between Xrootd layout and RadoFS. In the end all tests were done using CephFS.
- **Network connection between important entities**
  How to manage fast and efficient management of connections and network traffic between VM and storage? We managed to create special network solution for this reason (check "Current solutions")
- **Packstack (fast integration of cloud)**
  For easy deployment of our cloud computing farm we were searching appropriate cloud platform solution with CEPH support. At the end we stopped on PackStack as most appropriate platform.
- **EC2 integration for our solution**
  Current CernVM solutions are widely using EC2 APIs, but in OpenStack with Neutron network service there is no possibility to use of EC2 service without Neutron API support (EC2 VPC network → should be visible as Neutron network entity & with support of public/private networks)

Current solutions:

- We decided to use CEPHFS as a main layout
- Special networking solution
  1. For more flexible future support we decided to use Neutron and not old Nova-network.
  2. For more efficient management of network traffic between VM and physical node we are creating additional network using Neutron tools, adding on physical node new interface used for this network (find-tag-approach by cron scripts that check VMs -> network tag and create interface with correct tag and IP) and solve network routing to new "xrootd server" through special script on controller answering on requests from virtual machines.
- Development of new modules for Packstack (Ceph anf Xrootd Puppet manifests)
- Testing of new solution for EC2 service integrated with Nova API and Neutron API

Future plans:

- Tests in real scalable environment (data center) with increasing quantity of CPUs
- Benchmarking and testing two solution (with RadosFS or CephFS)
- Comparison of results
- Test Packstack with CEPH based on Centos7
- Test new EC2 API with CernVM products

References
[1] The Grid: A system of tiers. http://home.web.cern.ch/about/computing/grid-system-tiers.
[2] CEPH. http://ceph.com/ .
[3] A. J. Peters. EOS overview. US-ALICE Grid operations review, 2014 .
[4] Packstack RDO. https://openstack.redhat.com/
[5] Puppet. http://puppetlabs.com/