

EOS Usage at IHEP

On behalf of Computer Center, IHEP

Haibo Li
2017-02-03

Contents

- About IHEP && IHEPCC
- Why Use EOS?
- EOS Deployment at IHEP
- EOS Experienced
- Summary



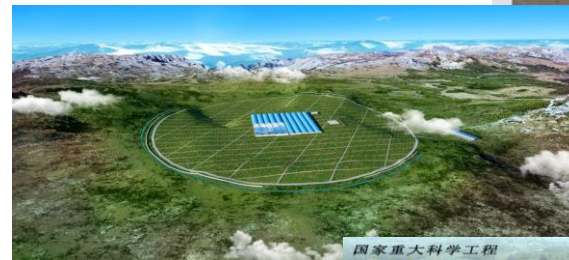
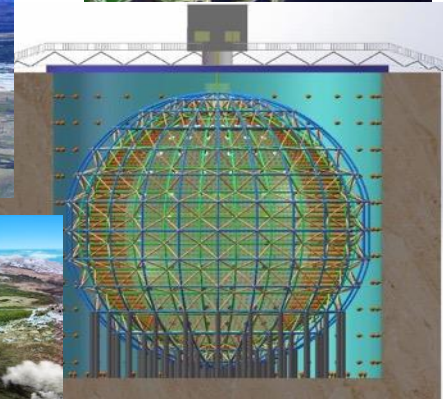
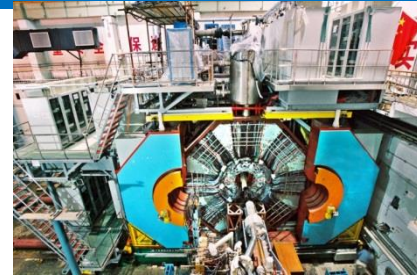
IHEP at a Glance

- Institute of High Energy Physics, Chinese Academy of Sciences
- ~1500 staffs, with ~1200 scientists and engineers
- Four(Six) sites currently
 - Beijing, Dongguang(CSNS), Shenzhen (dayabay), Tibet (Yangbajing), Jiangmen (JUNO), Chengdu (LHASSO)
- The largest fundamental research center in China with following research fields:
 - Experimental Particle Physics
 - Theoretical Particle Physics
 - Astrophysics and cosmic-rays
 - Accelerator Technology and applications
 - Synchrotron radiation and applications
 - Nuclear analysis technique
 - Computing and Network applications
 - ...



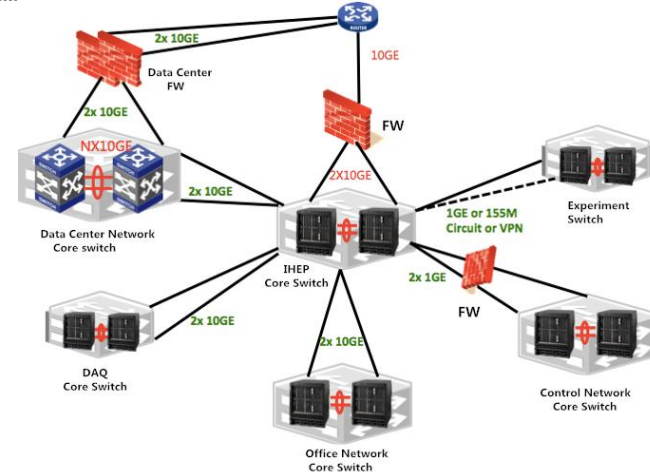
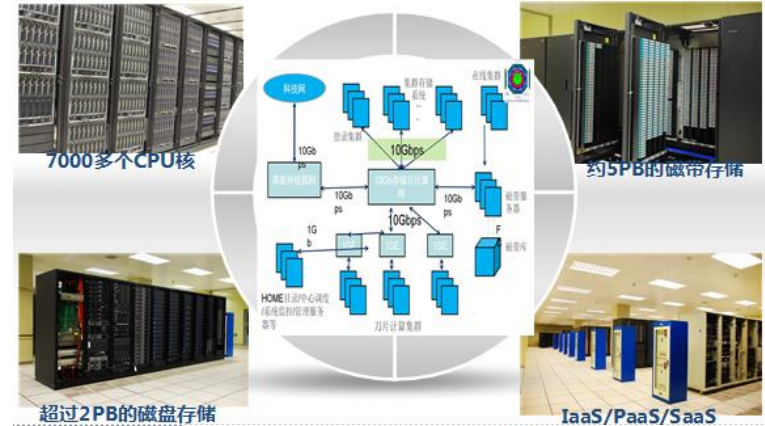
Major Projects

- BEPCII/BESIII
 - 36 Institutes from China, US, Germany, Russian, Japan,...
 - 5PB data in 5 years
- Daya Bay Neutrino experiment
 - 39 Institutes from China, US, ...
 - 400TB/year data collected
- Yangbajing in Tibet
 - Cosmic-ray observatory,
 - Collaborations of China, Italy, Japan
 - ~200TB raw data per year.
- JUNO
 - Jiangmen Underground Neutrino Observatory
 - ~ 1PB Raw data per year
- LHAASO
 - the Large High Altitude Air Shower Observatory
 - ~ 2PB Raw data per year
- CSNS
 - Chinese Spallation Neutrons Source
- LHC
 - Members of ATLAS and CMS
 - WLCG Tire-2 at IHEP
- AMS (Alpha Magnetic Spectrometer)



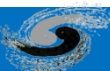
IHEP CC

- Computing Center, IHEP
 - 36 + 5 Staffs , 20 project staffs, 15 Students
- Serve for the HEP Experiments
 - Infrastructure
 - Operation
 - Network and Security
 - Computing & Storage
 - Basic IT services
 - Database
 - Applications Development
 -



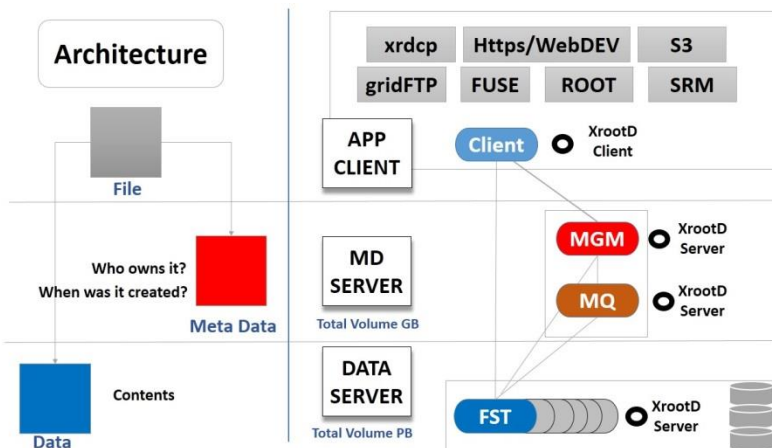
IHEPCC(cont.)

- Computing
 - ~13,500 CPU cores, 300 GPU cards
 - Migrated to HTCondor in 2016
- Storage
 - 5PB LTO4 tapes managed by CASTOR 1
 - 8.2 PB of Lustre
 - 734 TB of gLuster with replica feature
 - 400TB of EOS
 - 1.2 PB of other disk spaces



Why use EOS?

- Current existing storage systems issues
 - metadata is managed statically, which leads to performance bottleneck
 - Metadata and file operations are tightly coupled, difficult to scale for a closed system
 - Local data and remote data are managed separately
 - Traditional RAID technology causes too much time consumption for data recovery and system crash in case of host failure



- EOS has a very comprehensive document management capabilities, including multiple copies of the main switch, load balancing, etc.



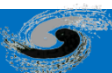
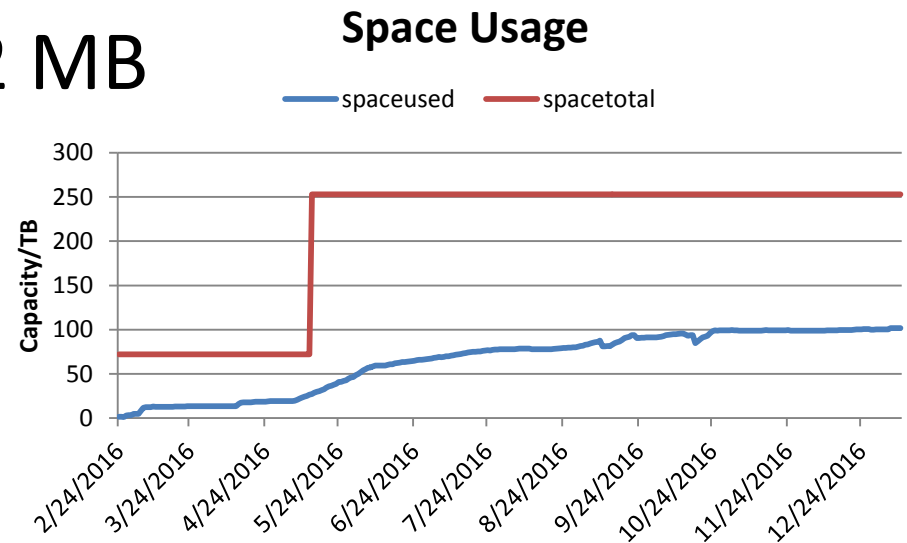
EOS Deployment at IHEP

- Thanks to the support from CERN EOS Team, two instances have been built.
- **LHAASO EOS**
 - Used for LHAASO experiment batch computing
 - 230 TB presently
 - 3 servers
 - 3 dell disk array box (raid6)
 - Each server has 10Gb network link
- **Public EOS**
 - Backend storage for IHEPBox based on Owncloud
 - 145 TB presently
 - 4 servers
 - Each server with 12 disks and 1Gb link
 - Replication mode
- **Future plan**
 - Extend EOS to more experiments



LHAASO EOS status

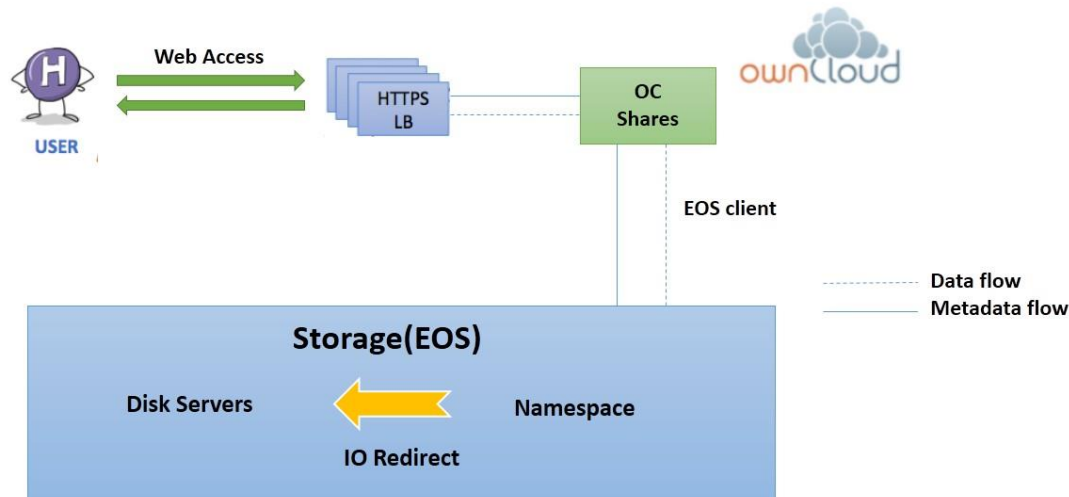
- Space total: 231 TB
- Space used: 102 TB
- # files: 3.19 M
- # dirs: 126 K
- Average file size: ~ 32 MB



Public EOS

IHEPBox = OwnCloud + EOS

IHEPBOX is a cloud disk system made up of OwnCloud and EOS. Cloud disk system is mostly used to share a large number of small files, and IHEPBOX using EOS memory storage metadata and data with multiple copies of the characteristics, making IHEPBOX has a high file read and write performance and data reliability.



IHEPBox



IHEP 文件

LI Haibo

全部文件

收藏

分享给您的文件

您分享的文件

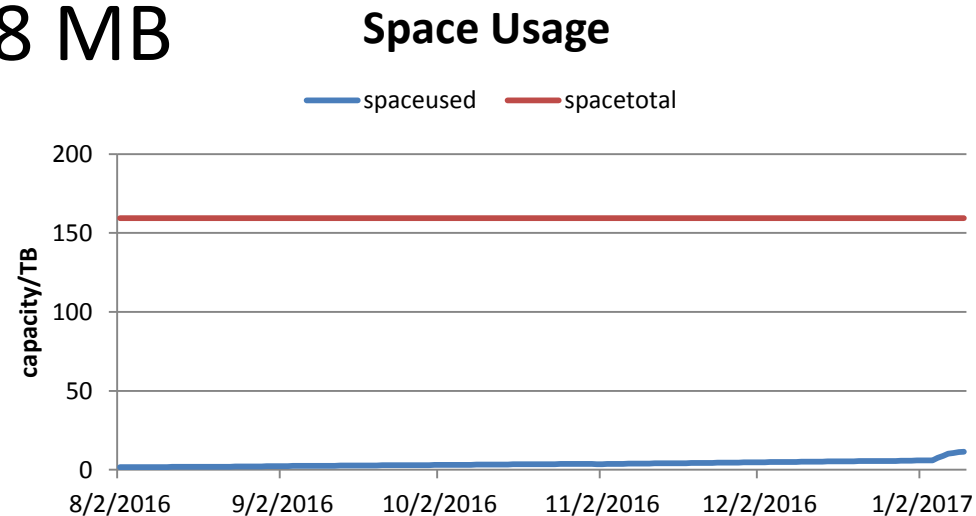
分享链接的文件

标签

名称	大小	修改日期
Documents	35 KB	5个月前
IHEP-CC	1.3 GB	1个月前
Photos	663 KB	5个月前
Public software	370.4 MB	2年前
我的	0 KB	5个月前

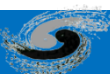
Public EOS status

- Space total: 145 TB
- Space used: 13 TB
- # files: 3.39 M
- # dirs: 395 K
- Average file size: ~ 3.8 MB



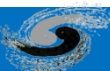
Problems experienced in LHAASO EOS

- Problems mainly on fuse client side
 - Stable after upgraded to 0.3.222
 - eosd consumed a lot of memory
 - HTCondor modified `/proc/sys/kernel/pid_max`
 - Eosd is related to `pid_max`
 - Fixed with the help from CERN EOS Team



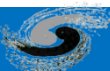
Problems still exist

- Remote site deployment
 - Lack of public IP address
 - Remote site may not have enough replication storage, cache support ?
- Small files performance
- Master-slave switchover sometimes fails.
 - Difficult to reproduce the errors.
- Unstable when switching between groups
 - Such as switching between “cold” and “hot” groups



Summary

- Built two EOS instances, running well
 - LHAASO EOS for batch computing
 - Public EOS provides backend storage support for IHEPBox
- More support from CERN EOS team
 - Willing to help with the majority of EOS



Thank you!

