



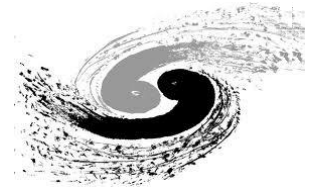
IHEP Site Report

Shi, Jingyan

shiy@ihep.ac.cn

Computing Center, IHEP

Outline



1

Resources at IHEP

2

Current Status and Next Plan

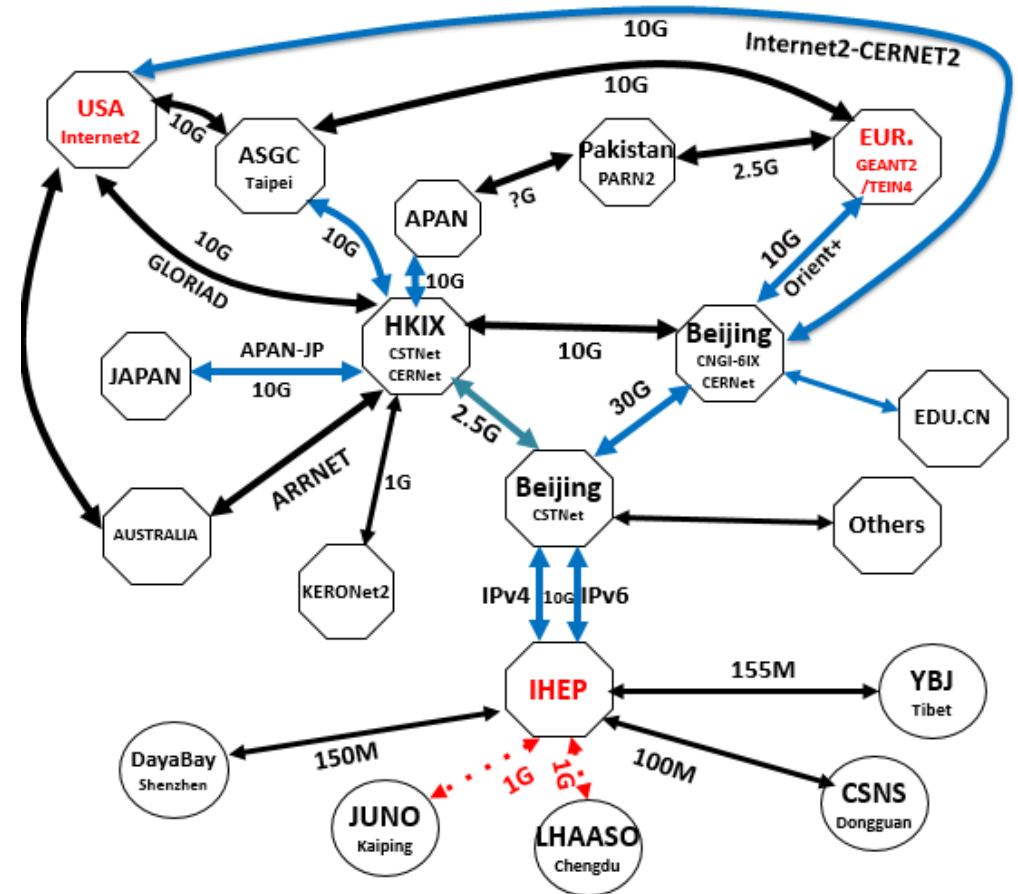
3

Summary

Resources -- network



- Dedicated Links for three other IHEP sites (two in the future)
 - Shenzhen (Dayabay)
 - Dongguan (CSNS)
 - Tibet (YBJ/ARGO)
 - Kaiping (JUNO)
 - Chengdu (LHAASO)
- Good Internet connections
 - IHEP-Europe: 10 Gbps
 - IHEP-USA: 10 Gbps
 - ~5.4 PB/year data exchange



Resources -- Computing



- Local Cluster

- HTCondor Cluster: HPC

- Migrated from PBS by the end of 2016
 - Single core, Series Jobs,
 - ~10,000 CPU cores

- Slurm Cluster: HPC

- Created in Jan, 2017
 - MPI, GPU jobs
 - ~3000 CPU cores;
 - 3 GPU servers with 12 GPU cards

- Cloud: integrated with HTCondor

- ~1000 CPU cores

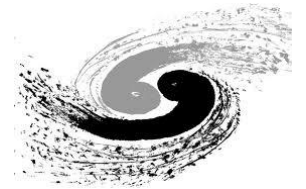
- Grid Site

- Torque Cluster:

- 888 CPU cores



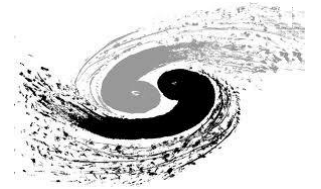
Resources -- Storage



- Local Cluster
 - 9.3 PB+ disk storage
 - Lustre: 8.5 PB
 - EOS: 0.8 PB
 - Other: 1.5 PB
 - 5 PB tape storage: Castor
- Grid Site
 - DPM: 400TB
 - dCache: 540TB



Outline



1

Resources at IHEP

2

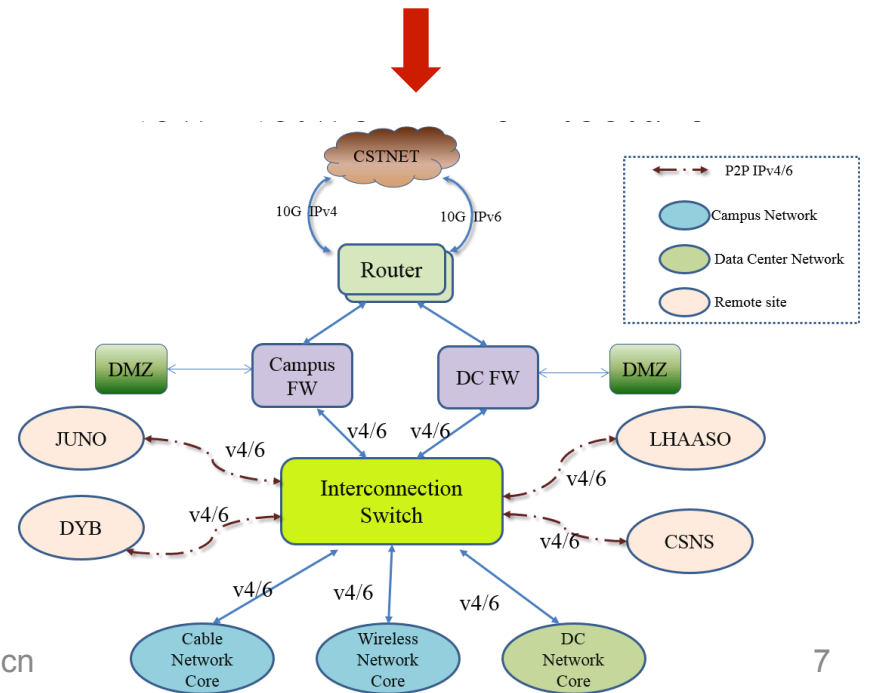
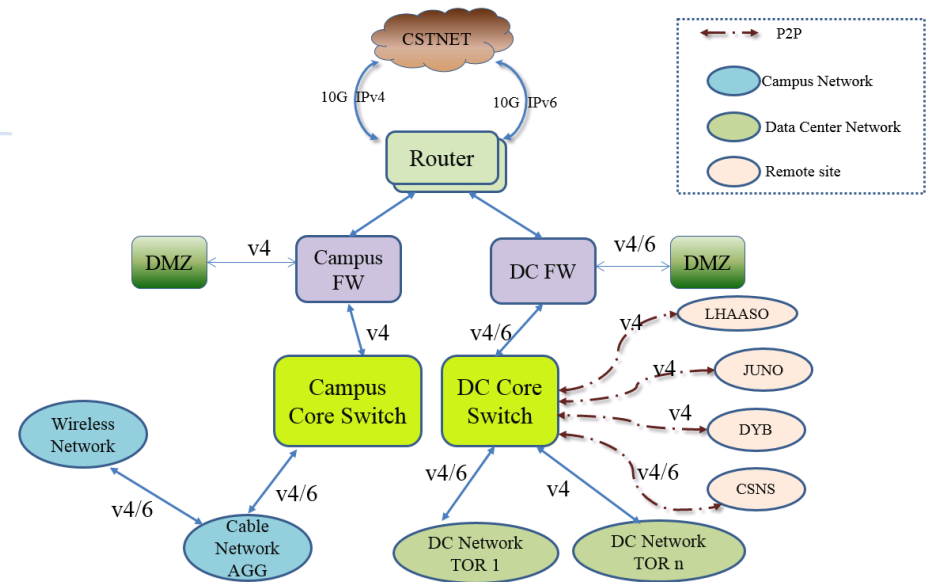
Current Status and Next Plan

3

Summary

Plan for Network

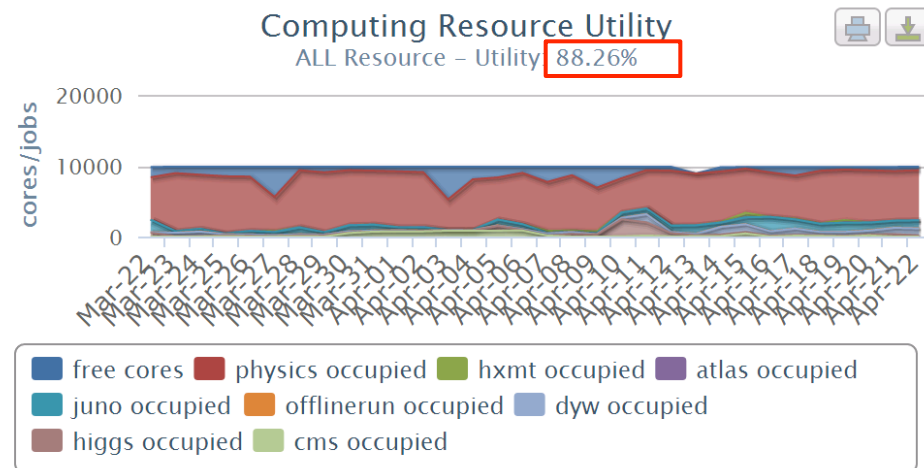
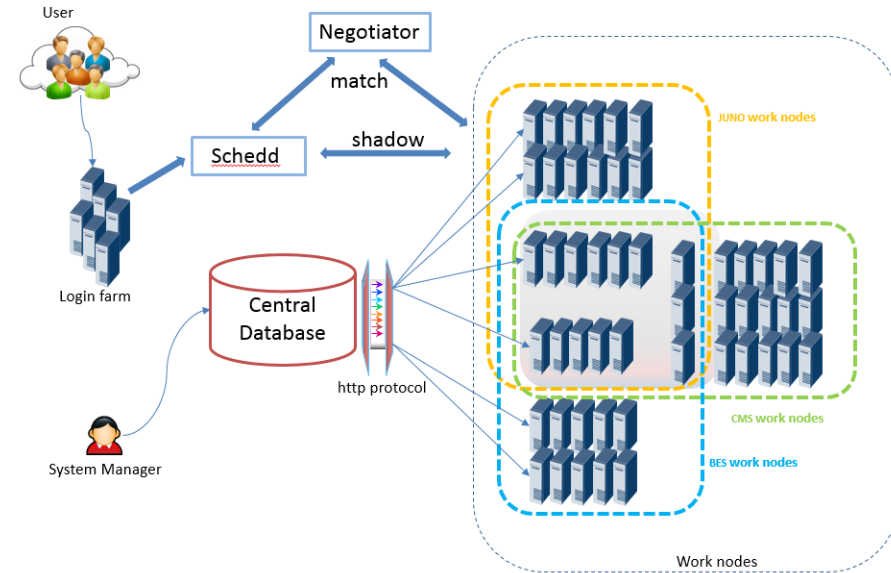
- Physical division by function provides failure isolation.
- Easy management and easy debugging when network failure happens
- Will be deployed in Aug.



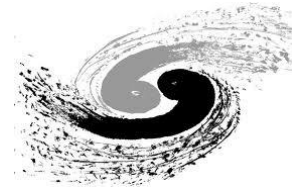
HTCondor Status



- Migration to HTCondor from PBS step by step with risk control
 - Jan, 2015 ~ 1,100 CPU cores
 - May, 2016 ~ 3,500 CPU cores
 - Dec, 2016 ~ 10,000 CPU cores
- Hep_job toolset provided for users
- Shared scheduling strategy provides high resource utilization: **over 80%**

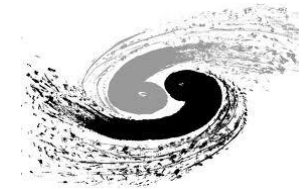


Plan for HTCondor



- More functions to Hep_job tool set
- Accounting system development
- Nodes attributes management based on database
- Virtualization: Virtual machine, Docker

Cloud Computing



- Based on Openstack Kilo
- Two kinds of cloud services provided
- IHEPCloud: Infrastructure As A Service
 - 14 compute nodes – 352 virtual cores
 - 331 cores used: 234 virtual machines running
- User Oriented Self Service
- Virtual Computing Cluster
- Virtual cluster: 1041 cpu cores
- Vcondor:
 - Developed by IHEP
 - <https://github.com/hep-gnu/VCondor>
 - Start virtual machine on demand to fit real computing requirement
- Transparent to users

http://cloud.ihep.ac.cn

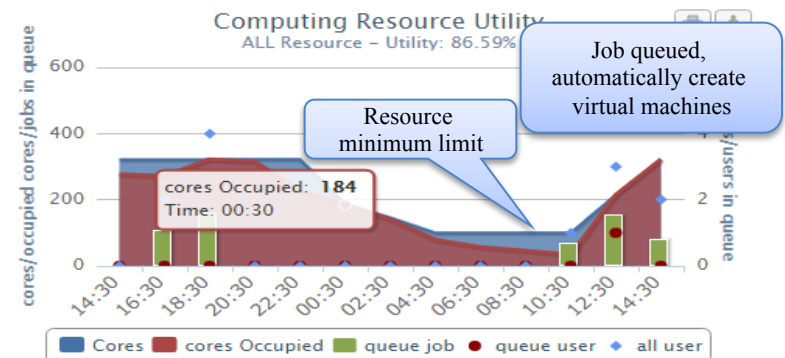
登录

用户名

密码

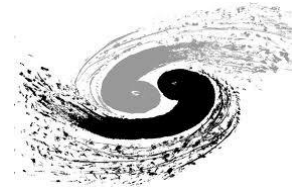
帮助

登入



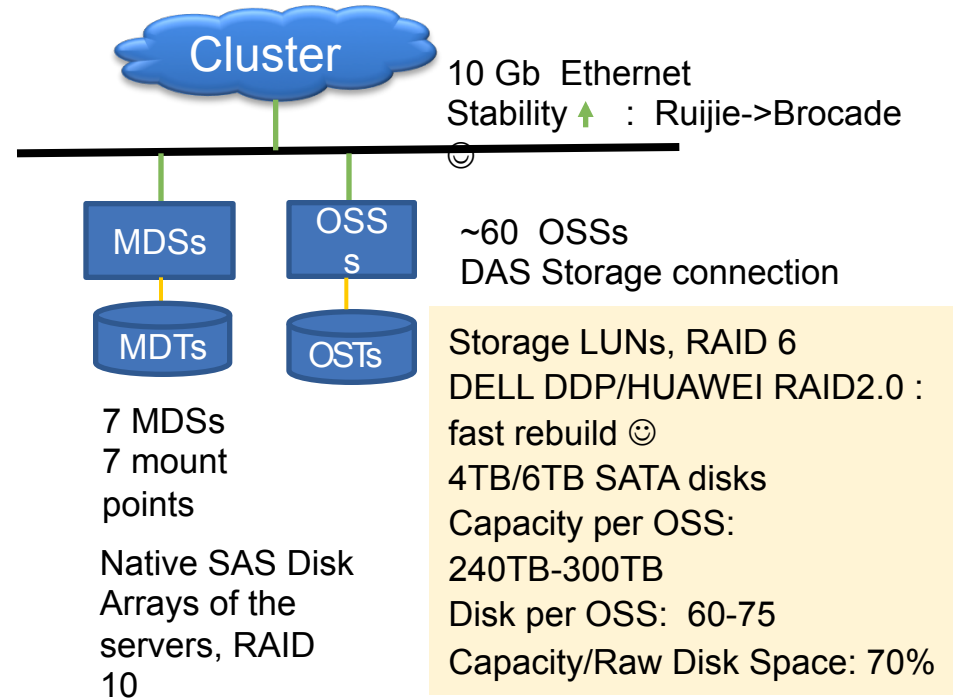
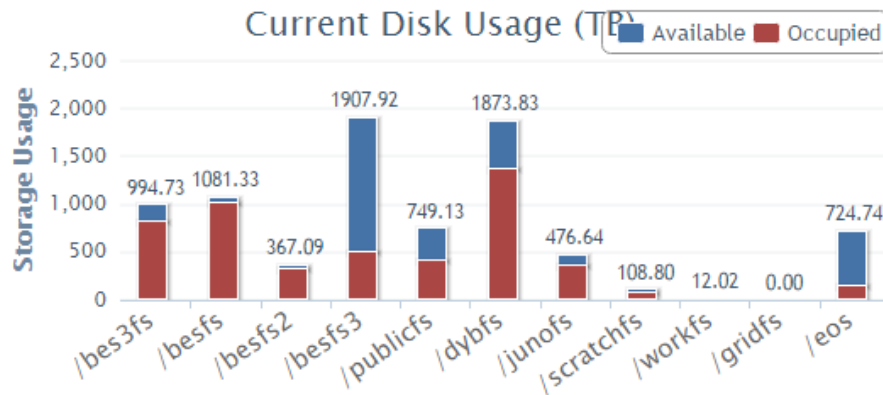
Effect of VCondor dynamic scheduling

Lustre Status

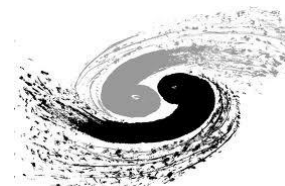


- Capacity: 8.5 PB

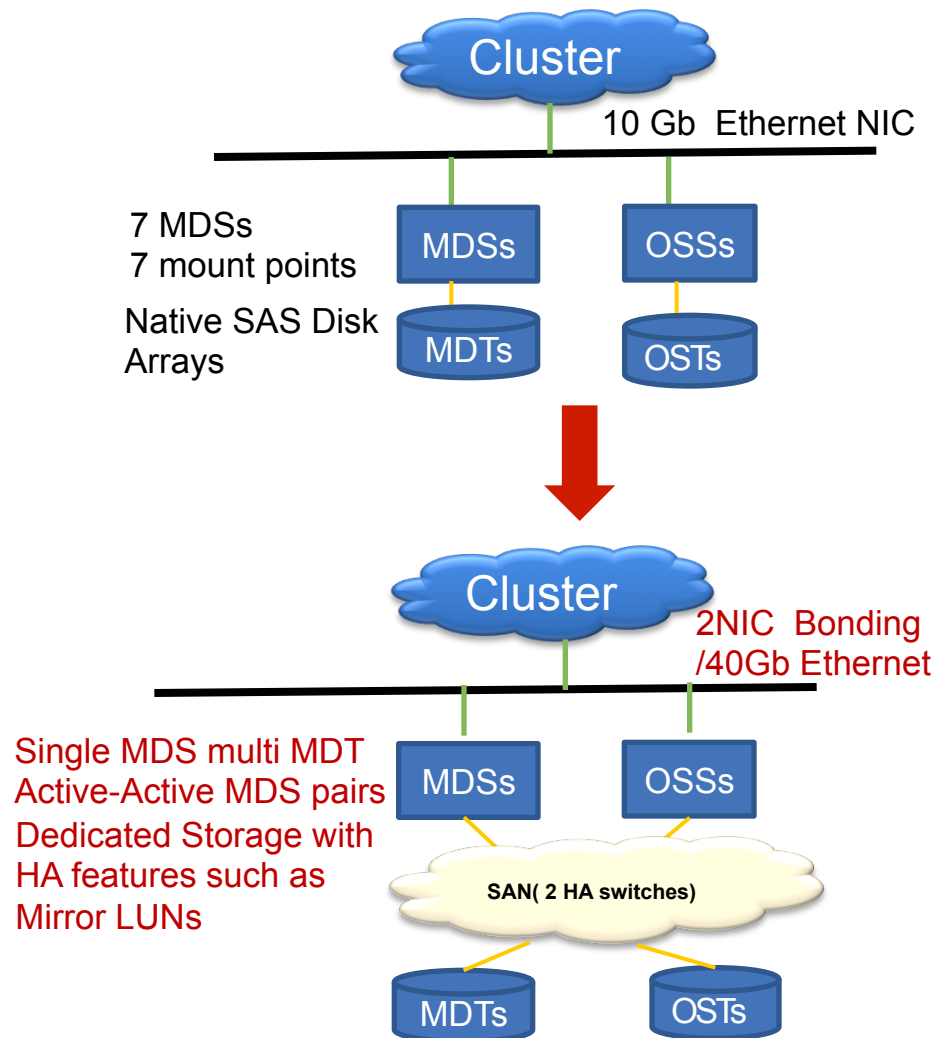
- Version: Lustre 2.5.3
- Average Usage: 60%
- Hardware:
 - Dell MD3860f
 - Huawei OceanStor V5500



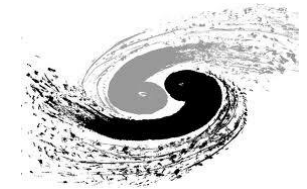
Plan for Lustre



- Upgrade to next stable community version; maybe the Intel open source “commercial version” this year
 - Fix Idlm-client related bugs
 - Reduce mount-points by static MDC
- Improve data reliability
 - reconfigurations of MDT connection
 - Mirrors of MDT LUNs
- Improve data availability
 - SAN
 - MDS active-active HA
 - Mirrors of MDT LUNs
- Improve IO performance
 - Upgrade of network connection to match the backend disk performance

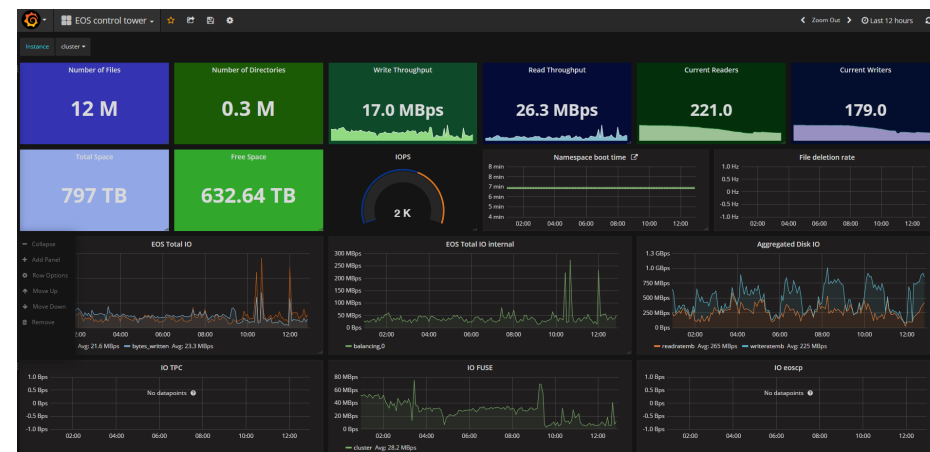
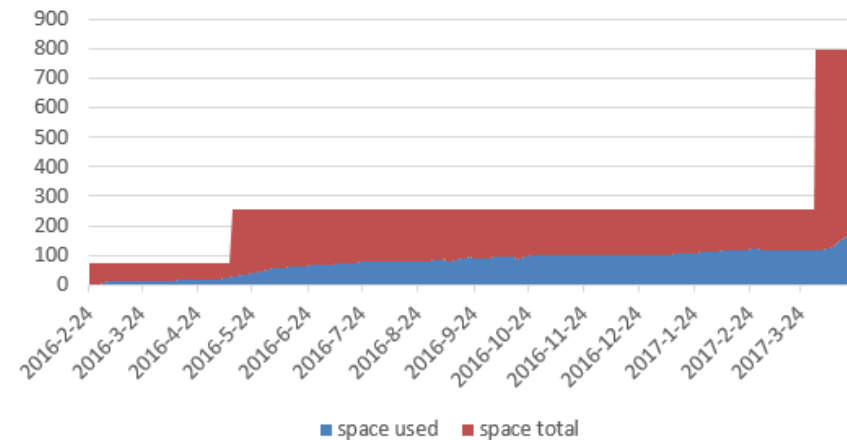


EOS Status

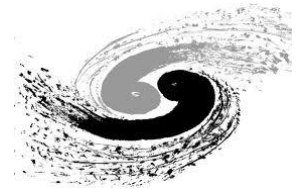


- EOS for batch computing
 - 5 servers with 24 X 4T and 180 X 6T SATA disks
 - 1.1 PB raw capacity provides 0.8 PB available space
 - 10Gb network link to file servers
 - 344,501 directories and 12,139,103 files
 - Now mainly used by LHAASO experiment
- EOS for IHEPBox
 - IHEPBox: A cloud disk system base on OwnCloud
 - 192 TB raw capacity provides with 145TB available space

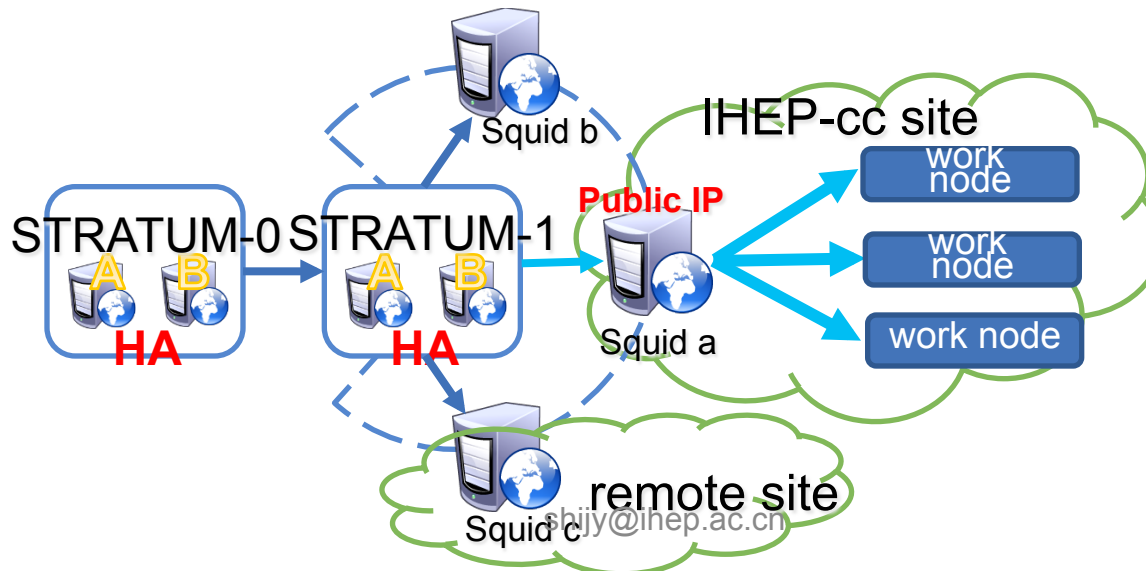
EOS for batch computing Space Usage



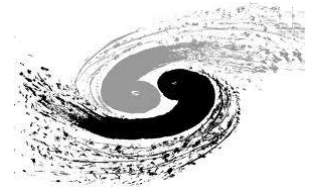
CVMFS



- Cvmfs-stratum-zero.ihep.ac.cn: 19TB Storage
 - For IHEP Experiment Software
- Cvmfs-stratum-one.ihep.ac.cn: 33TB Storage
 - Replica repository of IHEP-startum-zero , cern, openscience grid, egi etc.
- Squid01.ihep.ac.cn: 128G Mem
 - Cvmfs cache to reduce the network latency
- Thank David Dykstra from Fermi Lab for the help of IHEP-CVMFS creation



Outline



1

Resources at IHEP

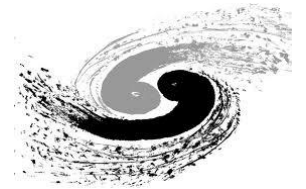
2

Current Status and Next Plan

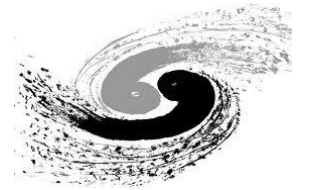
3

Summary

Summary



- New network architecture will be deployed
- Resource utility increased since the migration to HTCondor Cluster
- Plan to Lustre aims to improve stability and IO performance
- EOS scale expanded
- CVMFS service provided



Thank you!
Question?