中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

高能所計算中心
IHEP Computing Center

Networking and Computing Status of IHEP

Fazhi QI

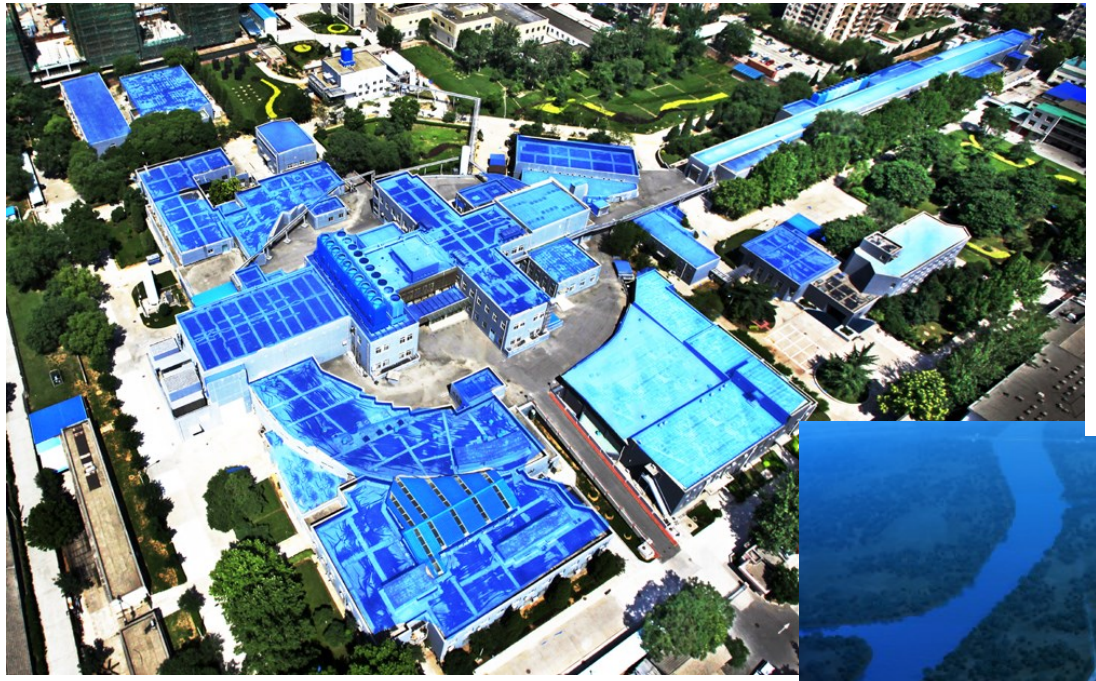
CC, IHEP

Outline

- About IHEP and the data challenge
- Networking Status
- Computing & Storage Status
- Summary

About IHEP

- More than 3 campus, 1450 FTEs, ~1000 graduated students/Postdoc



Yuquanlu,Beijing

Huairou
Beijing

Guangdong

HEP Experiments @IHEP

高能所計算中心
IHEP Computing Center

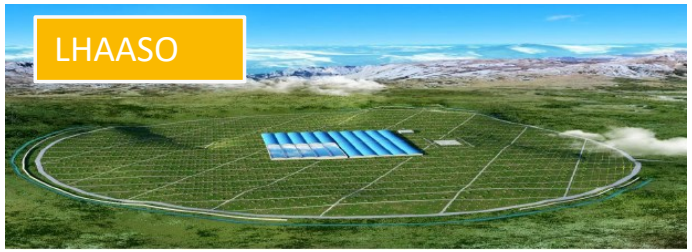
BESIII (Beijing Spectrometer III at BEPCII)

DYB (Daya Bay Reactor Neutrino Experiment)

JUNO (Jiangmen Underground Neutrino Observatory)

LHC

YBJ (Tibet-ASgamma ARGO-YBJ Experiments)



Large High Altitude Air Shower Observatory

AliCPT

China Spectrum Neutron Source

HEPS

Belle II

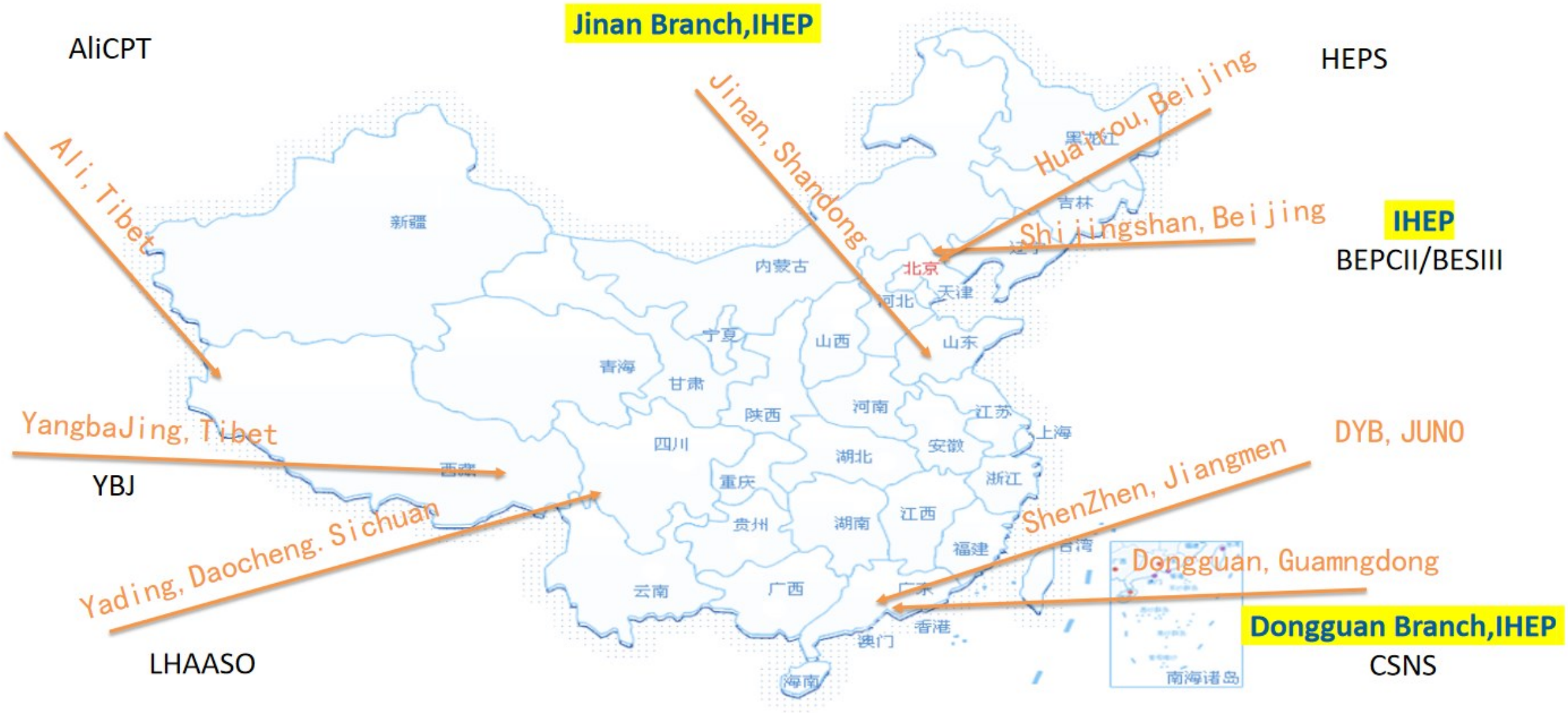
CEPC

PROTO DUNE SP

Distributed Science Facilities & Campus of IHEP

高能所計算中心
IHEP Computing Center

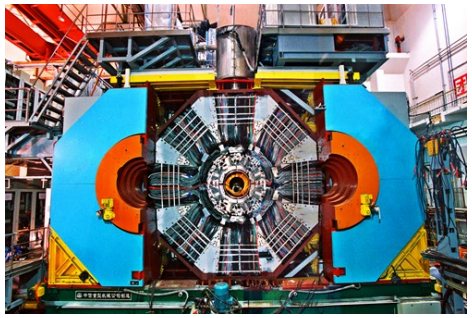
Large Science Facilities of IHEP

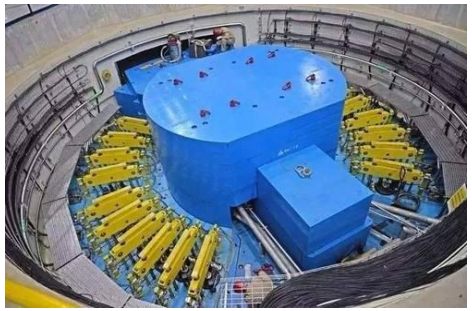


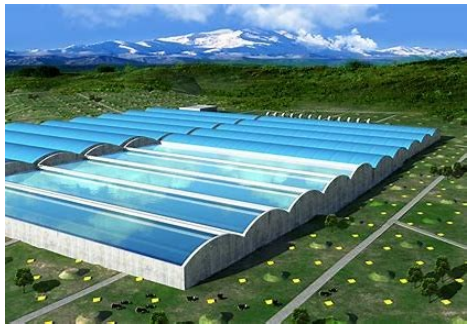
Data Challenges @IHEP

高能所計算中心
IHEP Computing Center

- BECPII/BESIII: Running
  - ~1PB raw data/year, >10PB+ totally now
- DYB: stopped at the end of 2020
  - >2PB raw data totally, data analysis is ongoing
- LHAASO: Running
  - >6PB raw data/year, will run for 20 years
- CSNS: Running
  - 300TB raw data /year for Phase I(3 beamlines, now)
  - > 2PB raw data/year for Phase II (+7 beamlines, 2023)
  - > 5PB raw data/year for Phase II I(+7 beamlines, 2030)
- JUNO: data taking from 2022
  - >2PB raw data/year, will run for 10 years
- HEPS: data taking from 2024
  - >200PB raw data/year for Phase I (15 beamlines)








HEPS Beamlines & Data volume

- More than 90 beamlines volume
- Phase I, 14 ID beamlines+1Bending Magnet beamline selected

Beamlines	Burst output(Byte/day)	Average output(Byte/day)
B1 Engineering Materials Beamline	600TB	200TB
B2 Hard X-ray Multi-analytical Nanoprobe (HXMAN) Beamline	500TB	200TB
B3 Structural Dynamics Beamline (SDB)	8TB	3TB
B4 Hard X-ray Coherent Scattering Beamline	10TB	3TB
B5 Hard X-ray High Energy Resolution Spectroscopy Beamline	10TB	1TB
B6 High Pressure Beamline	2TB	1TB
B7 Hard X-Ray Imaging Beamline	1000TB	250TB
B8 X-ray Absorption Spectroscopy Beamline	80TB	10TB
B9 Low-Dimension Structure Probe (LODISP) Beamline	20TB	5TB
BA Biological Macromolecule Microfocus Beamline	35TB	10TB
BB pink SAXS	400TB	50TB
BC High Res. Nanoscale Electronic Structure Spectroscopy Beamline	1TB	0.2TB
BD Tender X-ray beamline	10TB	1TB
BE Transmission X-ray Microscope Beamline	25TB	11.2TB
BF Test beamline	1000TB	60TB
Total average:		805.4TB/day, 24.16PB/month

Outline

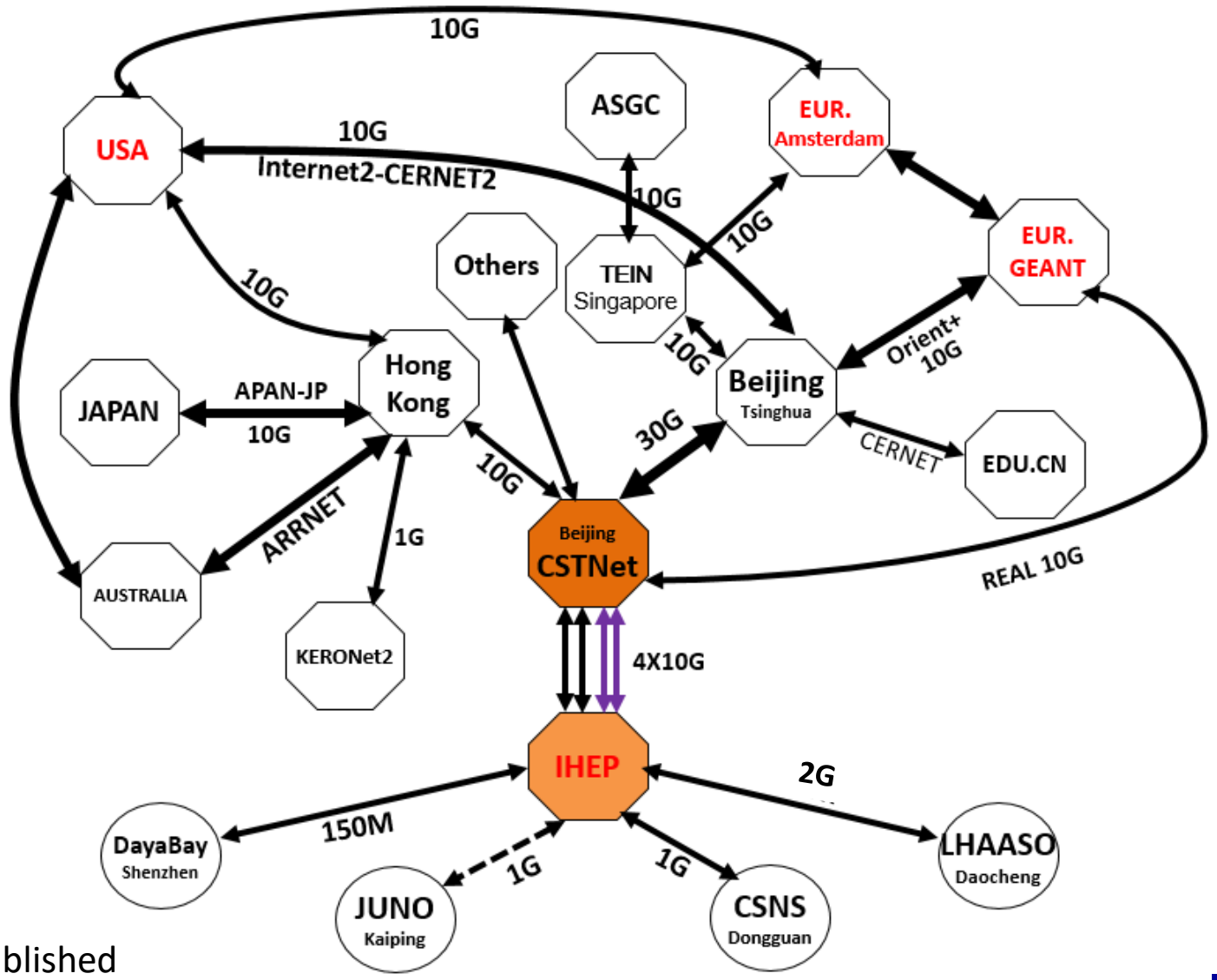
- About IHEP and the data challenge
- Networking Status
- Computing & Storage Status
- Summary

CSTNet International Links

高能所計算中心
IHEP Computing Center

International and domestic links

- Dedicated links for remote sites
  - Shenzhen (Dayabay)
  - Dongguan (CSNS):1G-->5G-->100G
  - Tibet (YBJ/ARGO)
  - Chengdu (LHAASO)
  - Kaiping (JUNO)
- Internet performance
  - IHEP-Europe: 10 Gbps + 10 Gbps
  - IHEP-USA: 10 Gbps + 10 Gbps
  - ~17 PB/year data exchange
- LHCONE
  - Peering to ESNet, Internet2, GEANT were established



Network Traffic Monitoring

高能所計算中心
IHEP Computing Center

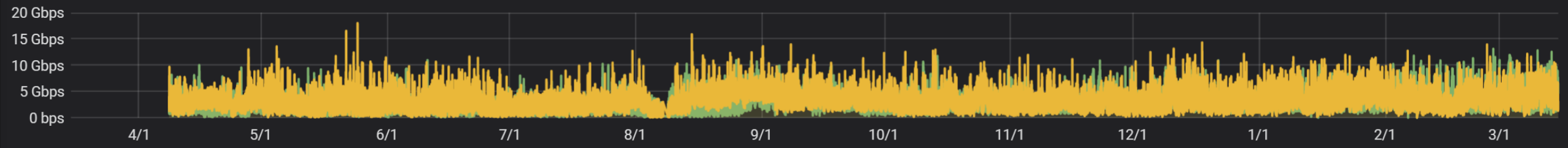
流入数据/Incoming Data

7.796 PB

流出数据/Outgoing Data

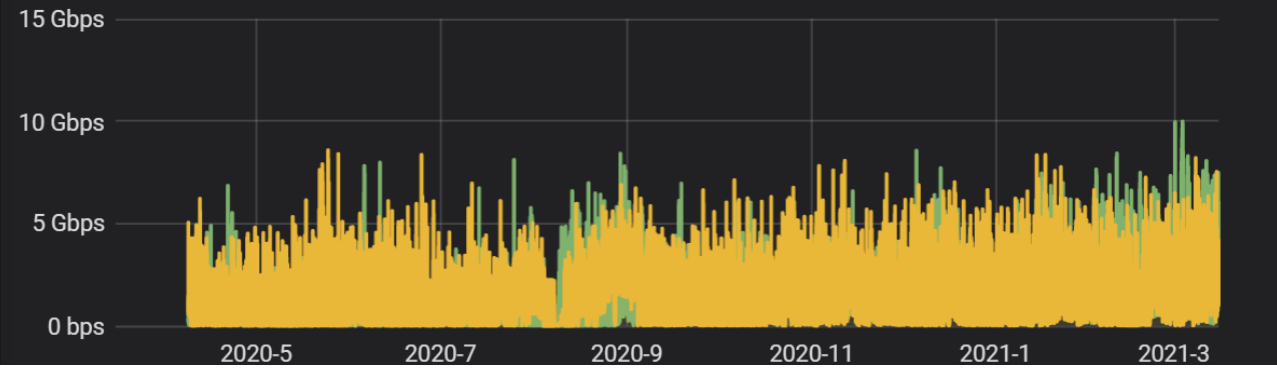
9.888 PB

总体进出流量/Total traffic

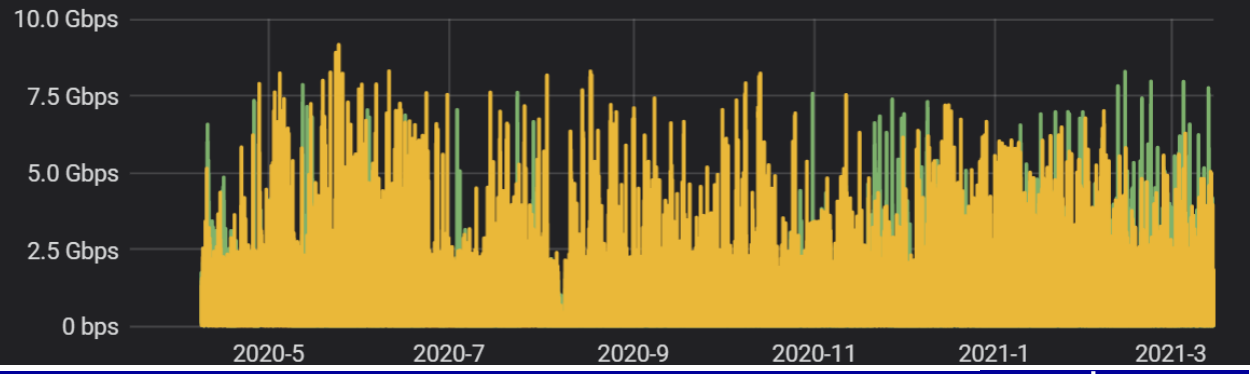


	max	avg	current
Incoming traffic	13.11 Gbps	2.12 Gbps	5.04 Gbps
Outgoing traffic	18.02 Gbps	2.69 Gbps	1.45 Gbps

LHCONE IPv4流量/LHCONE IPv4 Traffic



LHCONE IPv6流量/LHCONE IPv6 Traffic



Outline

- About IHEP and the data challenge
- Networking Status
- Computing & Storage Status
- Summary

Resources @IHEP Computing Center

高能所計算中心
IHEP Computing Center

- 46K CPU cores, 200 GPU cards, locates in Beijing and Dongguan computing center
  - For more than 10 experiments
  - HTCondor cluster runs for HTC jobs
  - Slurm cluster runs for HPC jobs
  - WLCG tier2 sites & DIRAC sites
- About 45PB disk storage, 20PB tape storage
  - Luster and EOS as two main file systems
  - Castor for tape storage， EOS CTA coming soon
- Data Center Network
  - IPv4 / IPv6 dual stack
  - Ethernet / IB protocols supported
  - 8*100Gbps backbone, 25Gbps for nodes

Overview of Computing & Storage

高能所計算中心
IHEP Computing Center

COMPUTE

1,476 Servers | 46,979 Cores | 190 Gpu Cards

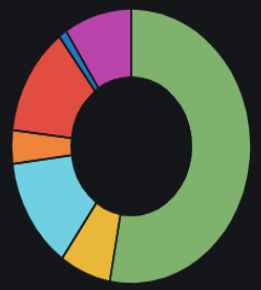
Runing Jobs

23,102 Jobs

Queue Jobs

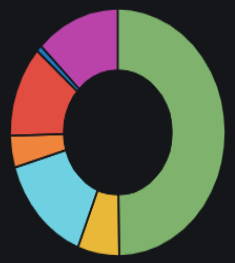
105,201.5 Jobs

Node Distribution



	current
HTC	778
GRID	102
HPC	193
Login_Server	58
Storage_Server	189
Tape_Server	16
OTHER	135

Cpu Core Distribution



	current
HTC	23,820
GRID	2,988
HPC	6,872
Login_Server	2,000
Storage_Server	5,704
Tape_Server	392
OTHER	6,068

STORAGE

Disk Space 46.5 PB | Tape Space 21.0 PB

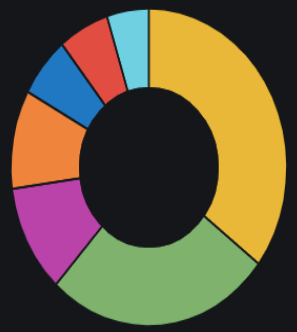
Disk Storage Read Throughput

12.6 GB/s

Disk Storage Write Throughput

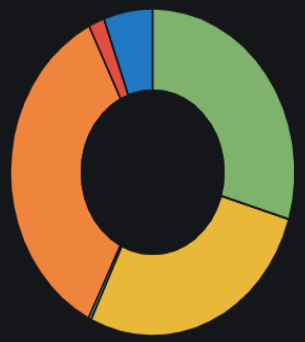
1.10 GB/s

Used Disk Space Distribution



	current
LHAASO	10 PB
BES	7 PB
OTHER	3 PB
DYB	3 PB
JUNO	2 PB
CEPC	2 PB
HXMT	1 PB

Used Tape Space Distribution



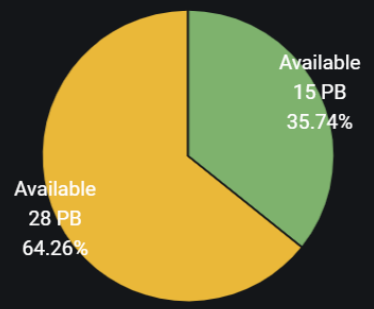
	current
bes_all	3 PB
dyb	3 PB
juno	30 TB
lhaaso	3 PB
user	170 TB
ybj	526 TB

Storage Overview

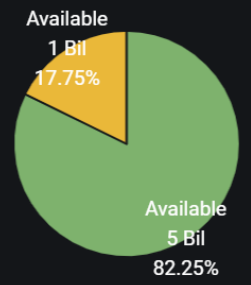
storage overview

Read Throughput	Write Throughput	Current Metadata IOPS	Total Space	Used Space	Total Files
12.9 GB/s	900 MB/s	9.15 K	46.5 PB	28.7 PB	1,090,739,585.66

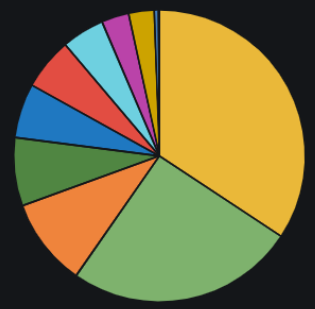
Space Usage



Inode Usage

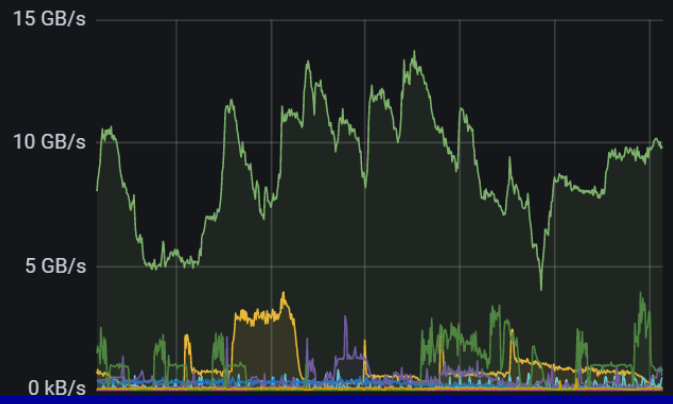


Space Distribution



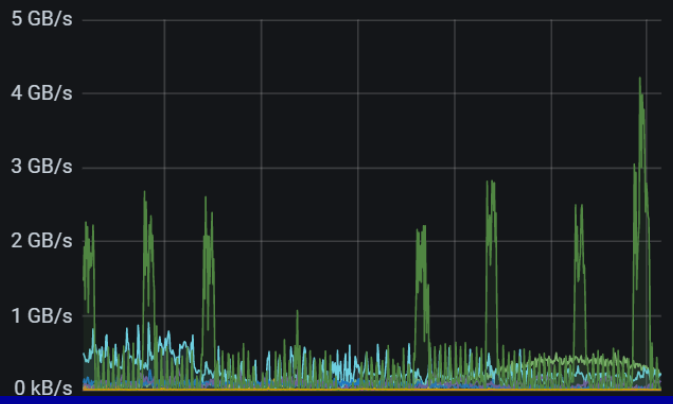
	current
LHAASO	10 PB
BES	7 PB
DYB	3 PB
OTHER	2 PB
JUNO	2 PB
CEPC	2 PB
HXMT	1 PB
HPC	853 TB
GRIDDPM	796 TB

Read Throughput



	max	avg	current
BES	13.7 GB/s	9.03 GB/s	9.80 GB/s
LHAASO-lhmt	3.96 GB/s	791 MB/s	845 MB/s
LHAASO-ihep	2.99 GB/s	415 MB/s	601 MB/s
JUNO	817 MB/s	116 MB/s	473 MB/s
PUBLICEOS	447 MB/s	52.9 MB/s	87.9 MB/s
DYB	3.96 GB/s	864 MB/s	83.3 MB/s
HXMT	574 MB/s	230 MB/s	28.3 MB/s
HEPS	25.1 MB/s	989 kB/s	84 kB/s

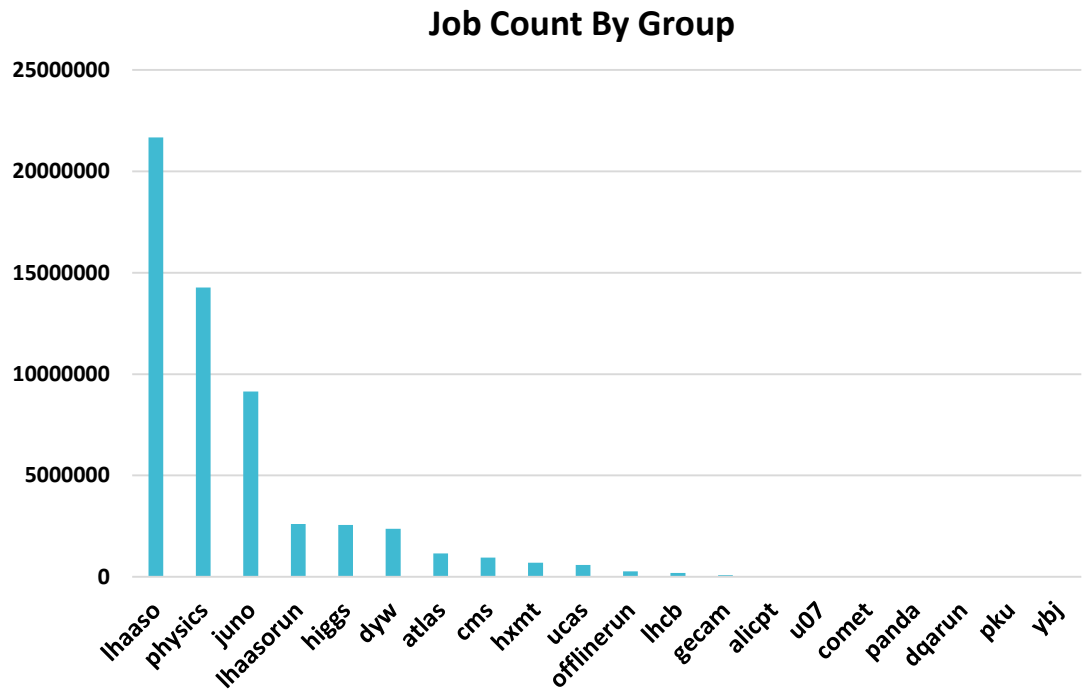
Write Throughput

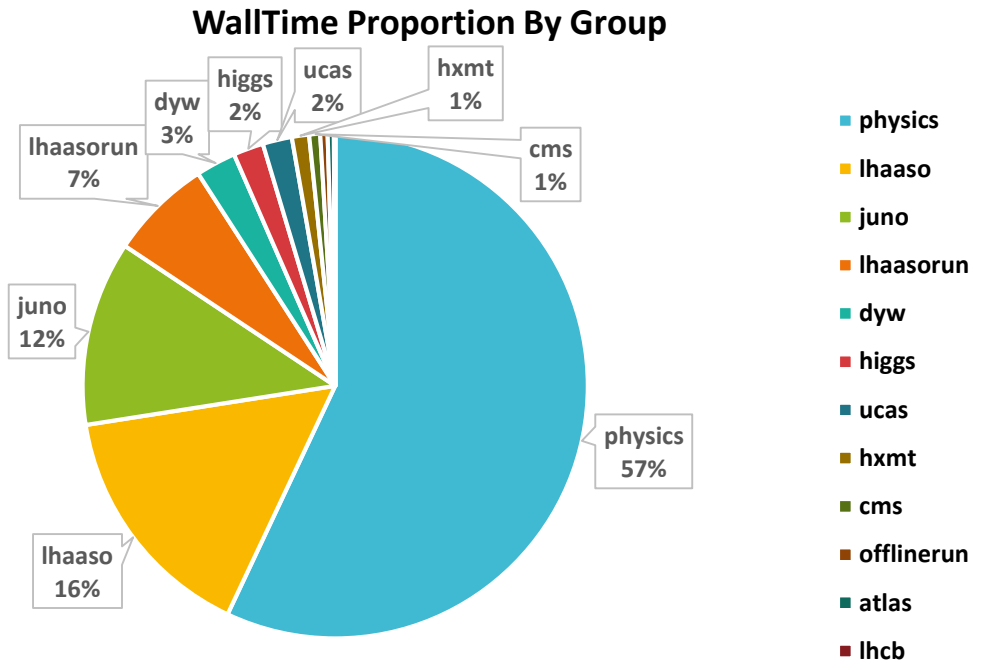


	max	avg	current
BES	500 MB/s	164 MB/s	227 MB/s
JUNO	907 MB/s	255 MB/s	223 MB/s
LHAASO-ihep	314 MB/s	75.0 MB/s	105 MB/s
LHAASO-lhmt	4.22 GB/s	426 MB/s	38.9 MB/s
HXMT	269 MB/s	65.0 MB/s	5.14 MB/s
PUBLICEOS	26.3 MB/s	4.69 MB/s	2.51 MB/s
HEPS	14.3 MB/s	25.7 kB/s	0 kB/s
MBH	6.13 MB/s	11.5 kB/s	0 kB/s

HTCondor Cluster Status

- Jobs : 2020 Oct. – 2021 Mar.
  - Top 3 : LHASSO, BESIII, JUNO
  - 65.4M jobs
  - 81.2M CPU*hours

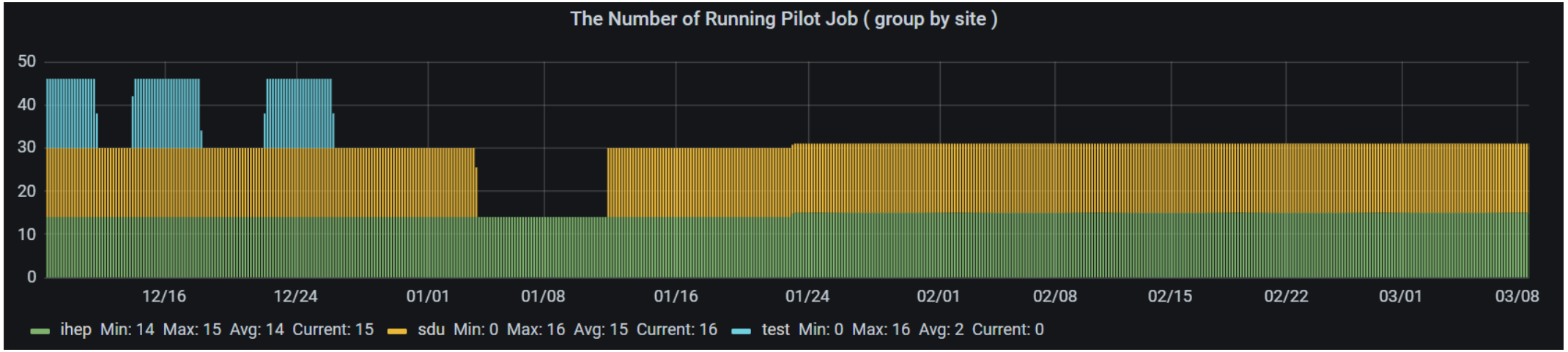




D-HTC : Distributed HTC platform

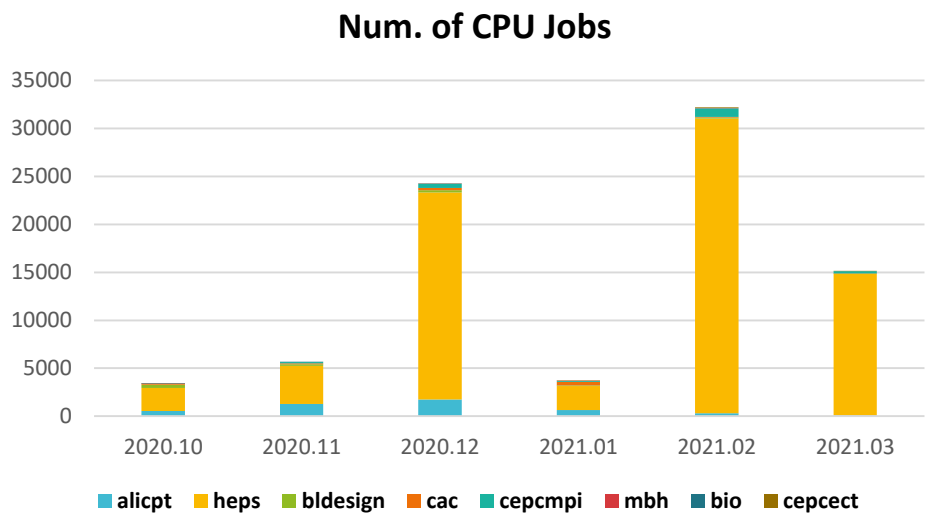
高能所計算中心
IHEP Computing Center

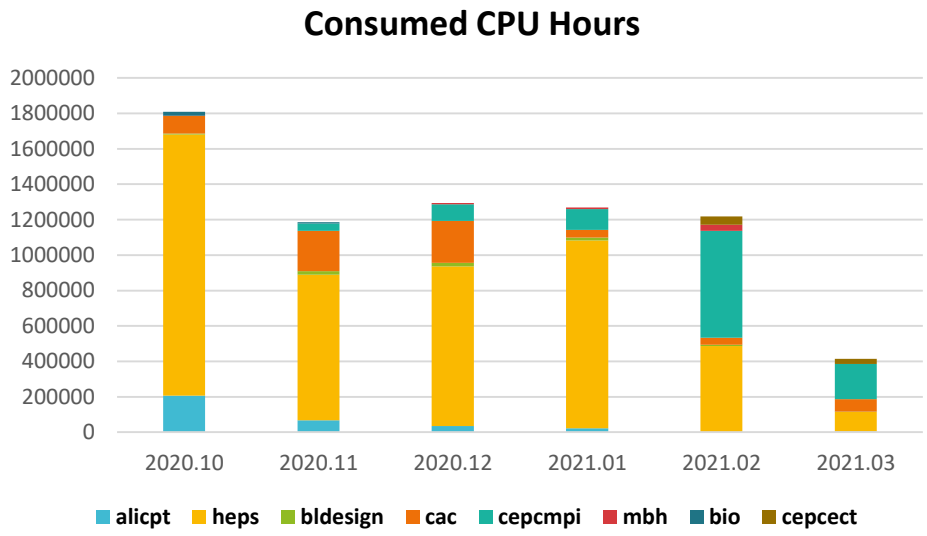
- Distributed Computing Platform for LHAASO Experiment
  - Based on HTCondor Glidein
- Built a test bed to integrate resources from remote sites.
  - HTCondor tokens for daemon authentication
  - Singularity for job running environment
  - Corsika and KM2A simulation job tests done(LHAASO)
- Job tests on for 3+ months



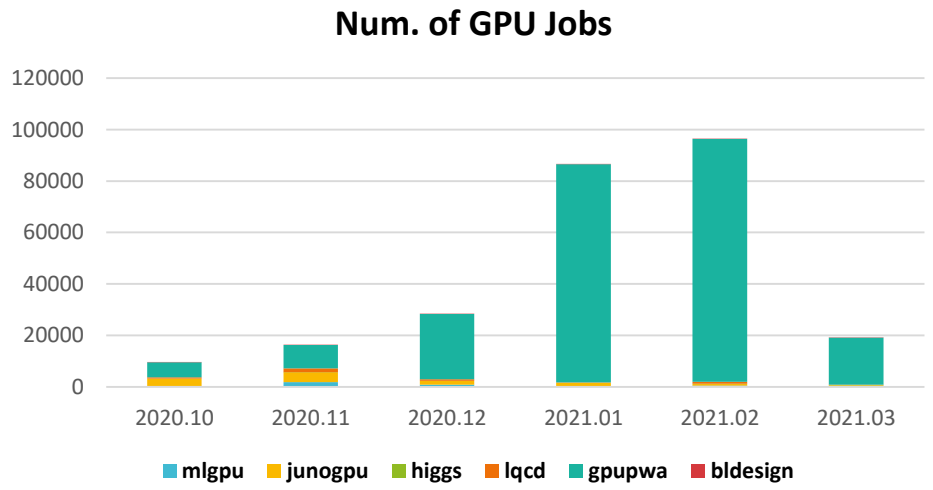
Slurm Cluster Status

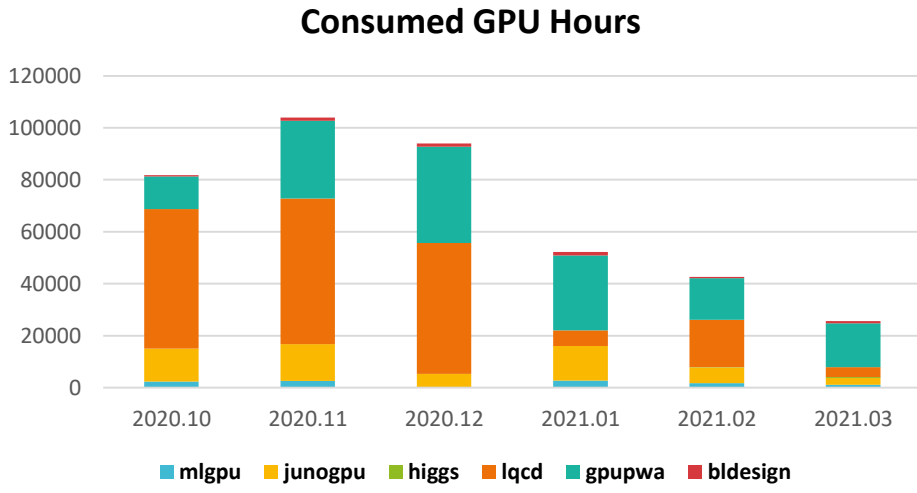
- CPU apps
  - Nine apps in total
  - 84Kjobs
  - 685K CPU*hours





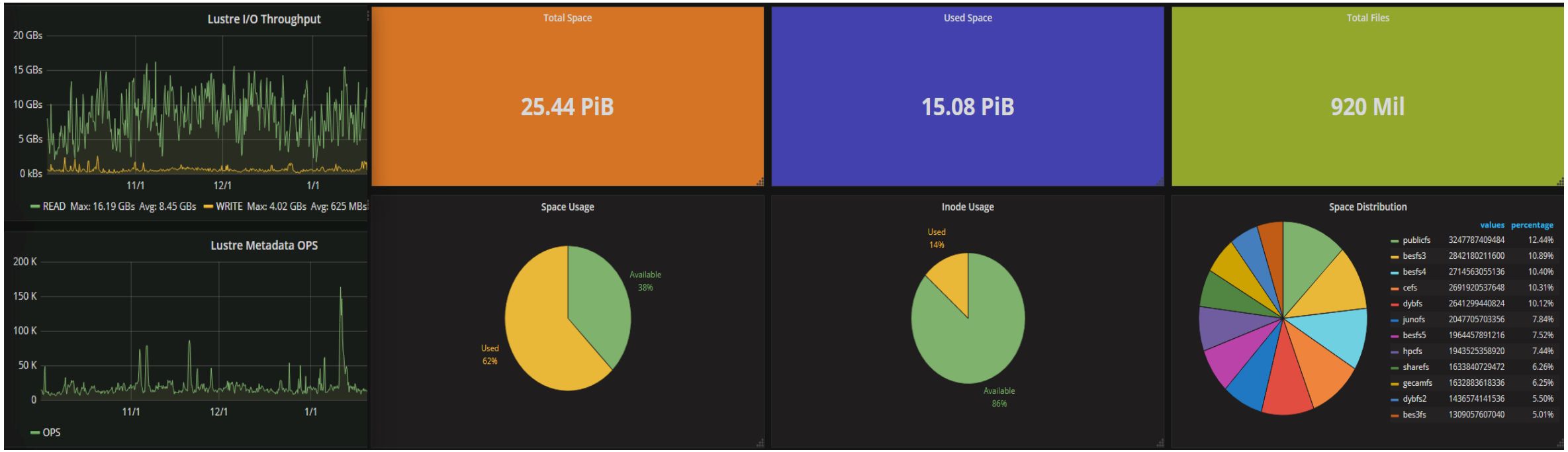
- GPU apps
  - Six apps in total
  - 240K jobs
  - 400K GPU*hours





Luster Statistics

- Many years in production @IHEP
- Total space > 25PB, 15PB used
- For BESIII,DYB,HEPS…



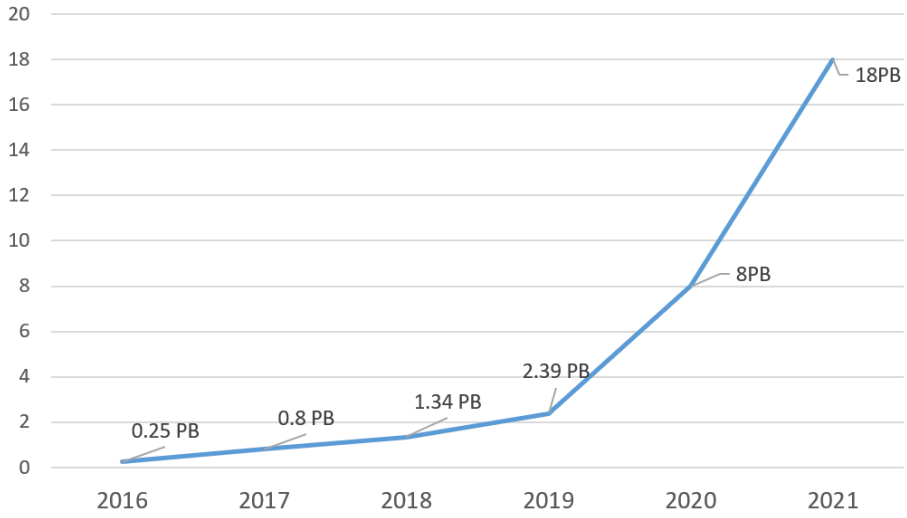
EOS Statistics

- 4 instances， 3 for physics experiments, 1 for IHEP Box

Instance	Capacity	Num. of fs	Num. of Files	Num. of directories
LHAASO-Beijing	14.5PB	1048	155 Mil	20 Mil
LHAASO-Daocheng	2.5PB	78	44 Mil	4 Mil
HXMT	806TB	13	40 Mil	2 Mil
IHEPBox	200TB	5	30 Mil	2.5 Mil

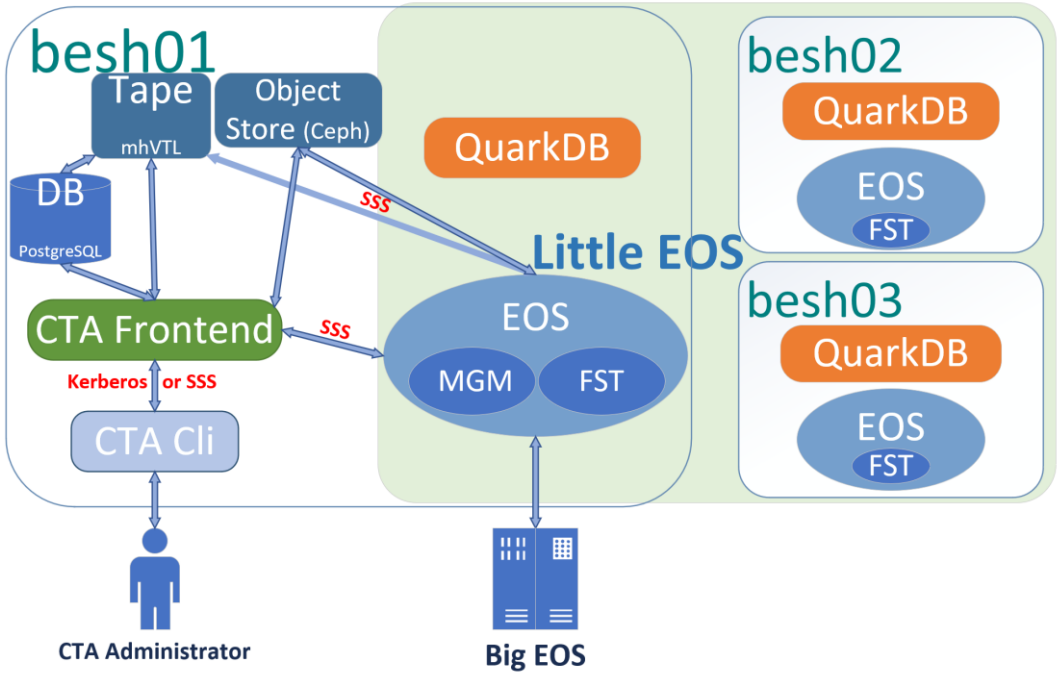
- 18PB capacity with 26PB coming soon

Raw Capacity	~ 18 PB+26 PB coming
Disk server	~49
Number of fs	1144
Number of files	~266 Mil
Number of directories	~10 Mil
Peak Read throughput	>50 GB/s



EOS CTA Status

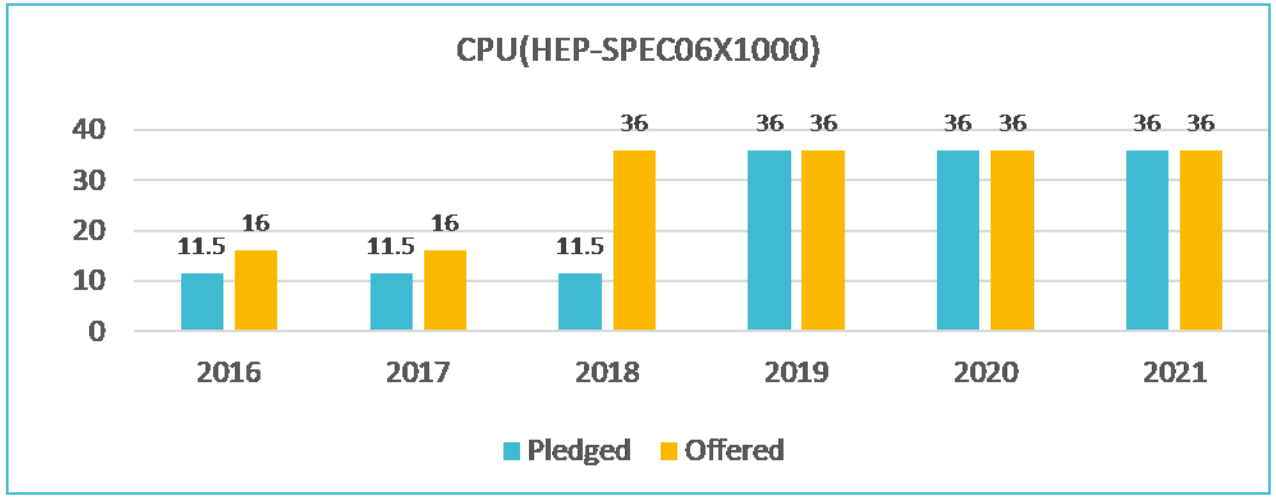
- Motivation
  - Evaluation for JUNO and other possible experiments
- Testbed Setup
  - 3 nodes, ready for function test
  - EOS + CTA + Ceph + mhVTL + PostgreSQL
- Next
  - All tests will be finished in June. 2021

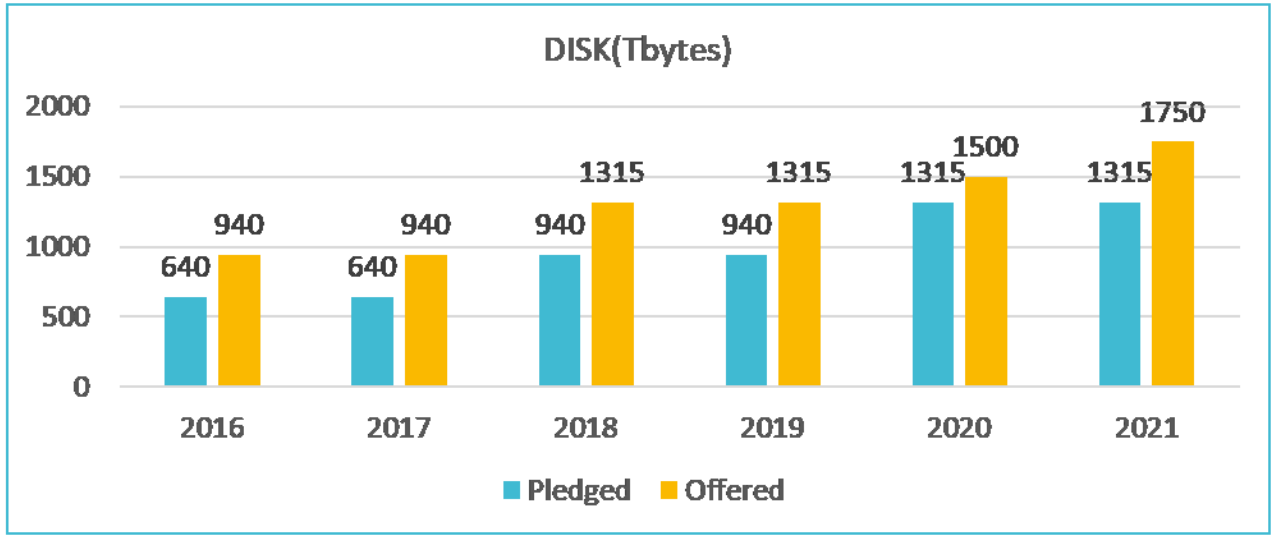


BEIJING-LCG2 Tier2 Resources

高能所計算中心
IHEP Computing Center

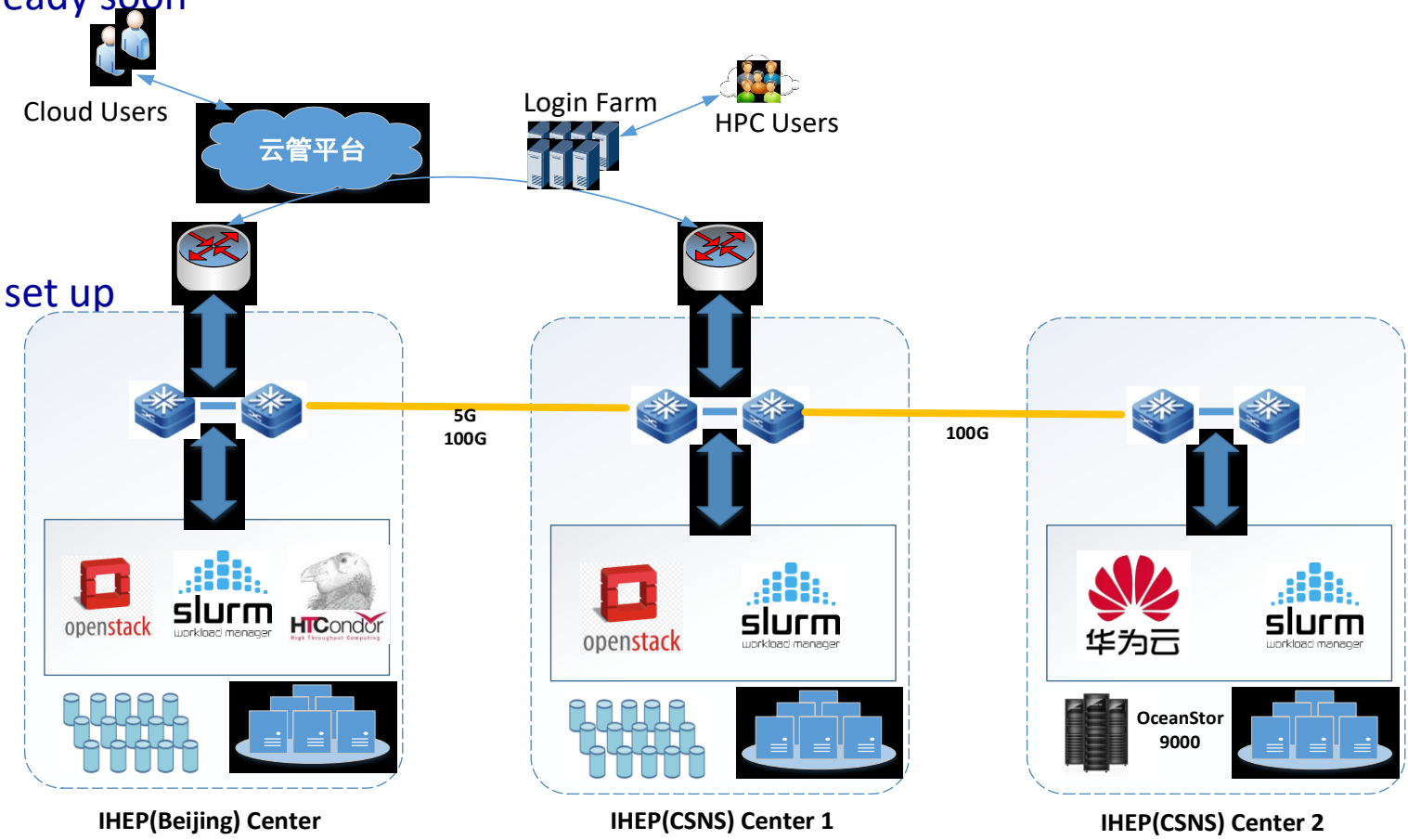
- CPU: 3048 cores
  - Intel Golden 6140 2160 Cores
  - Intel E5-2680V3: 696 Cores
  - Intel X5650 192 Cores
- Batch: HTCondor-CE
- VO: ATLAS/CMS/LHCb/BelleII/CEPC
- DPM: 1055TB
  - 4TB * 24slots with Raid 6, 5 Array boxes
  - DELL MD3860 8TB*60 slots
  - DELL ME4084 10TB*42 slots
  - DELL ME4084 12TB*84 slots





New Computing Center @CSNS in 2021

- Funded by local government (Guangdong), for Scientific Computing
- All the devices are from HUAWEI
- ~30,000 CPU cores & 6PB disk storage will be ready soon
  - 20,000 X86 CPU cores
  - 10,000 ARM CPU cores
  - 100Gbps RoCE network connection
- Deliciated link between IHEP and CSNS will be set up
  - 5G bps + 100Gbps
- A remote site of IHEP computing center
  - The jobs will be scheduled to this site
- 2021.4 in production



Outline

- About IHEP and the data challenge
- Networking Status
- Computing & Storage Status
- Summary

- We are facing the huge data challenges, especially for the photon science experiments
- Computing & storage resources increased according to the requirements from experiments, distributed/remote sites are important
- More and more network bandwidth is needed for data movements and jobs schedule

Thanks