CTA status at IHEP

QiuLing Yao
On behalf of Storage group
Computing Center, IHEP
2024-03

Outline

- Infrastructure and Status

- Test and Development

- Next Steps

- Summary

Tape Infrastructure

- 4 Tape libraries shared by various experiments, but tapes are separated
  - IBM TS3500 for BESIII & DYB
    - Frames: 12
    - Drives: 15 LTO7,9 LTO4
    - Tapes: 5k+ LTO7&LTO4
  - IBM TS4500 for LHAASO & YBJ & HERD
    - Frames: 8
    - Drives: 20 LTO7
    - Tapes: 10K+ LTO7
  - IBM TS4500 for JUNO & LHCb
    - Frames: 6
    - Drives: 7 LTO9
    - Tapes: 319 LTO9
  - IBM TS4500 for HEPS (NEW)
    - Frame: 1
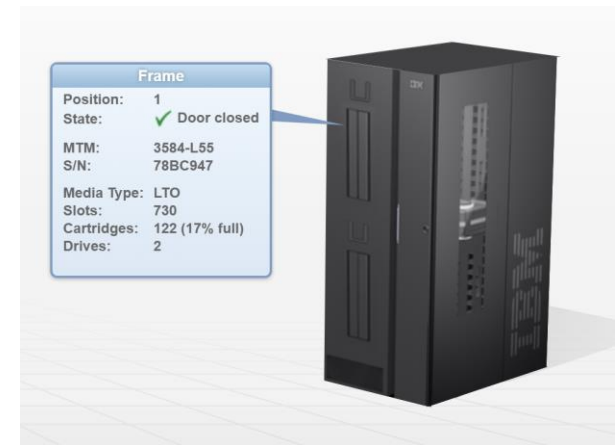   - Drive: 2 LTO9
    - Tape: 120 LTO9

BES lib

LHAASO lib

JUNO lib

HEPS lib

Why EOS+CTA?

- Motivation
  - Massive storage for multiple experiments(PB~EB)
    - BESIII, LHAASO, JUNO, HEPS, LHCb……
  - Long term reservation
    - Important data need to be preserved permanently
    - Can be accessed for several years
  - High performance
    - Tens of thousands of concurrent job access
  - Frequent dynamic data access
- Solution
  - Online: EOS/LUSTRE
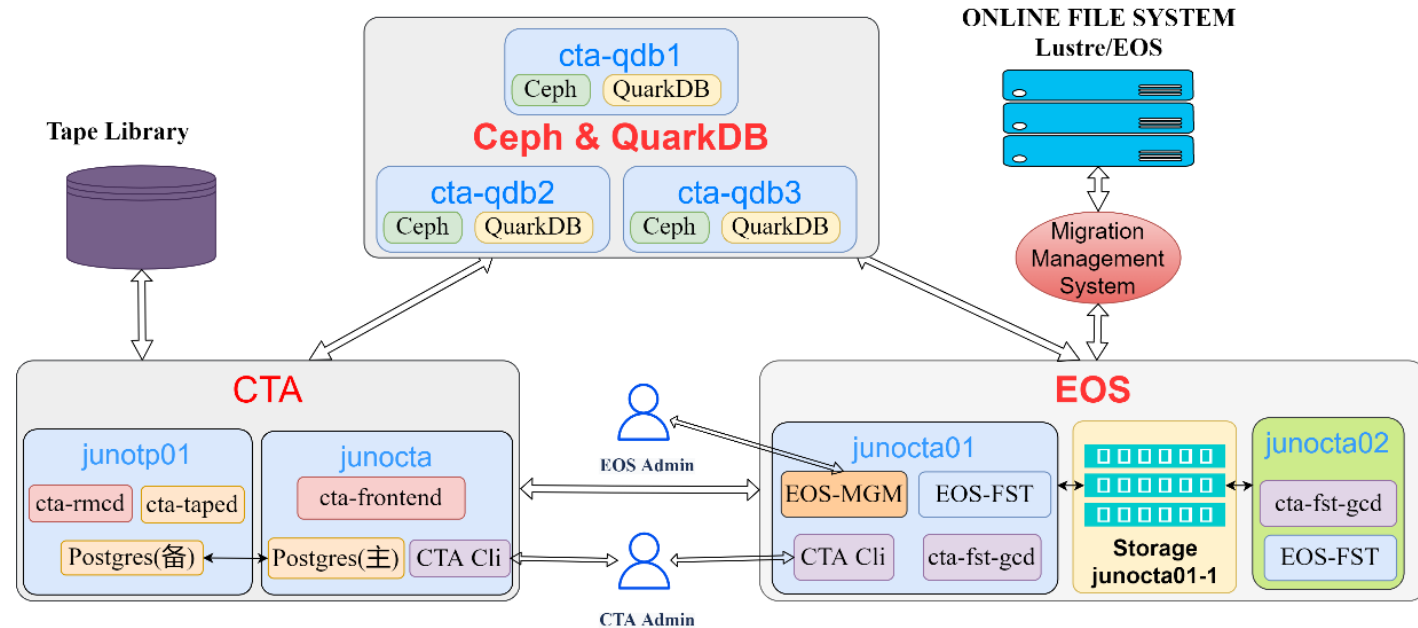   - Find more details in the presentation by Haibo Li: here
  - Offline: CTA

A sample:EOSCTA Architecture for JUNO

- ## Software

  - OS: Alma Linux 9.3

  - EOS: 5.2.8 CTA: 5.10.9

  - PostgreSQL: 13.13
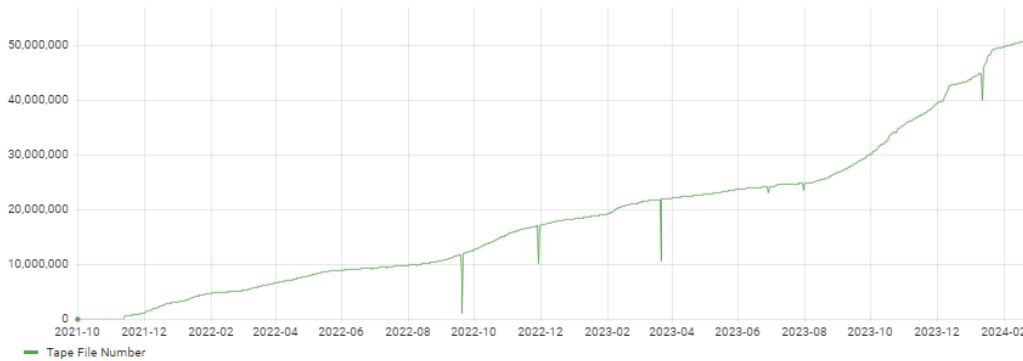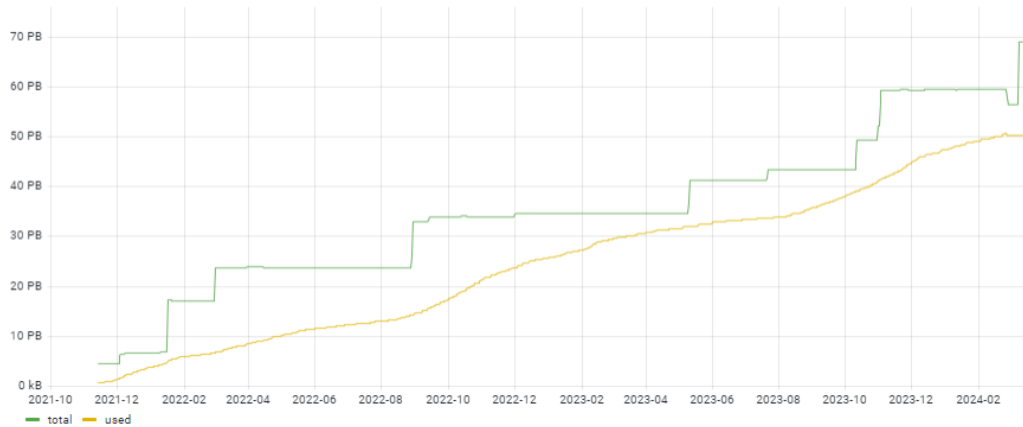
  - Ceph:15.2.15 QuarkDB: 5.2.8

- ## Hardware

  - Little EOS:2 server nodes,1 JBOD Array with 84x20TB

  - 1 CTA frontend and catalogue node

  - 1 tape server node

  - 3 Ceph & QuarkDB nodes
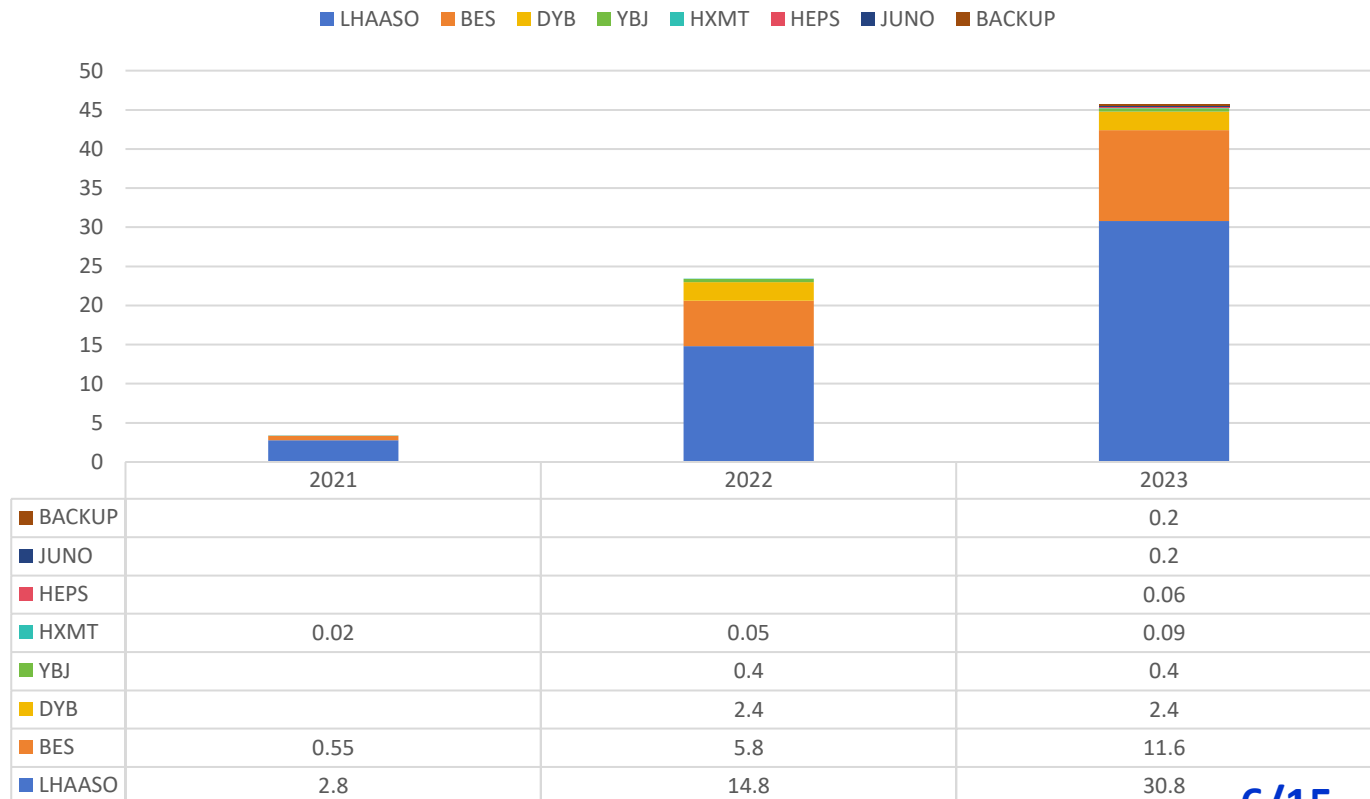
  - Network connection: 25Gb/s FC

Current stats

- 45PiB/46M Data archived to tapes

- All Data except Backup will be saved in two replicas

Storage Distribution by Experiments (PiB)

	2021	2022	2023
BACKUP			0.2
JUNO			0.2
HEPS			0.06
HXMT	0.02	0.05	0.09
YBJ		0.4	0.4
DYB		2.4	2.4
BES	0.55	5.8	11.6
LHAASO	2.8	14.8	30.8

Monitoring

- Ganglia: Server nodes(load, memory, network……)
- Grafana: Statistics of usage recorded in CTA
- Nagios: Send alarms through WeChat App
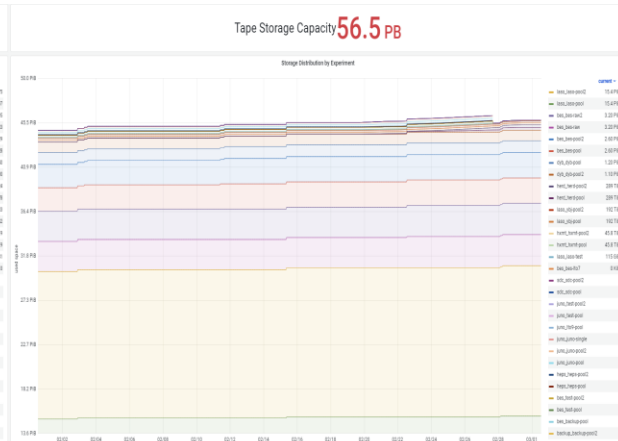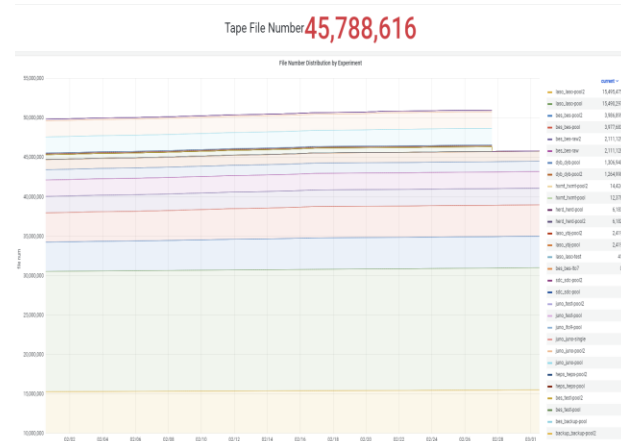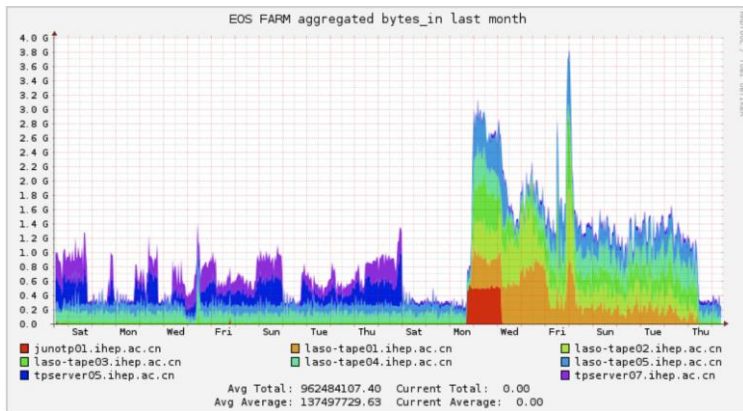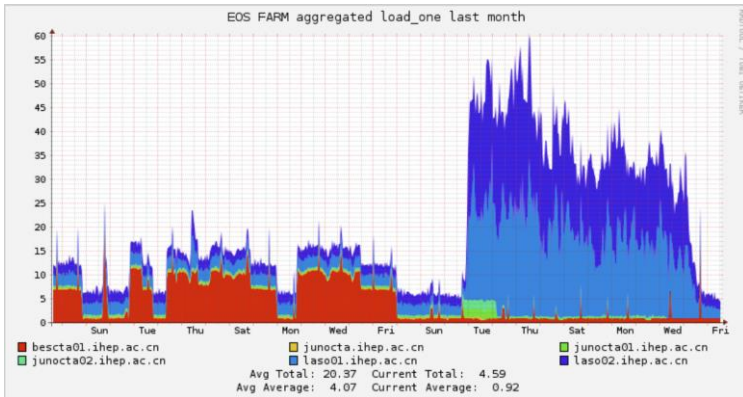
Functional tests of LTO 9

- ## Archive 15k files

  - Passed

  - Write speed ~270MB/s

- ## Retrieve 10k files

  - Failed with RAO error, drives disconnected

  - Modify cta-taped.conf "taped UseRAO"

  - Solution: update the firmware of Drives(P370->Q3F0)

  - Read speed ~ 220MB/s

Mar 21 14:27:39 junotp01.ihep.ac.cn cta-taped[51092]: LVL="WARN" PID="51092" TID="54209" MSG="In RAOManager::queryRAO(), failed to perform the RAO algorithm, will perform a linear RAO." thread="RecallTaskInjector" tapeDrive="JUNO02" transactionId="205310" errorMsg="Failed SG_IO ioctl in DriveG

Mar 21 14:27:40 junotp01.ihep.ac.cn cta-taped[51092]: LVL="ERROR" PID="51092" TID="54200" MSG="Failed SG_IO ioctl in DriveGeneric::positionToLogicalObject Errno=19: No such device" thread="DiskWrite" tapeDrive="JUNO02" tapeVid="L90044" mountId="205310" threadCount="10" threadID="2" fileId="252

library	drive	host	desired	request	status	since	vid	tapepool	vo	files	data	MB/s	session	priority	activity	age	reason
JUNOLIB	JUNO01	junotp01	Up	ArchiveForUser	Transfer	47728	L90024	juno-single	juno	2684	13.4T	280.5	319	0	-	14	-
JUNOLIB	JUNO02	junotp01	Up	ArchiveForUser	Transfer	48470	L90021	juno-single	juno	2366	11.8T	243.6	318	0	-	8	-
JUNOLIB	JUNO03	junotp01	Up	ArchiveForUser	Transfer	48796	L90025	juno-single	juno	2723	13.6T	278.2	316	0	-	4	-
library	drive	host	desired	request	status	since	vid	tapepool	vo	files	data	MB/s	session	priority	activity	age	reason
JUNOLIB	JUNO01	junotp01	Up	Retrieve	Transfer	1645	L90040	juno-single	juno	76	380.2G	229.3	203593	0	-	13	-
JUNOLIB	JUNO02	junotp01	Up	Retrieve	Transfer	1251	L90044	juno-single	juno	56	280.1G	220.4	204384	0	-	9	-
JUNOLIB	TIER01	junotp01	Up	Retrieve	Transfer	1635	L90035	juno-single	juno	77	385.2G	231.7	203609	0	-	3	-

Tape SE Progress

- ## LHCb & JUNO EOSCTA ready for production
  - GSI & Scitokens supported
  - Xrootd and HTTPS supported
  - TPC supported

  Find more details in the report of Yujiang Bi: [here](#)

- ## IAM services for experiments
  - Herd, CEPC: {herd,cepc}-iam.ihep.ac.cn

Welcome to **herd**

Sign in with

eduGAIN

INFN

Ciemat

IFAE Institut de Física d'Altes Energies

EPFL

Info and Privacy Policy

Alma Linux 9 Migration

- Compiled CTA under Alma Linux 9
- Updated OS for EOS&CTA to Alma Linux 9
- Rebulid Ceph cluster from 15 to 17
- Reconstructed EOS & CTA with "latest" version
  - EOS 5.2.8 & CTA 5.10.9

- Updating progress
  - Done for JUNO, LHCb and HEPS experiments
  - BESIII & LHAASO will be done during BEPCII maintenance

CASTOR Migration

- Develop migration scripts

  · Use 'tpread' to retrieve the data from tapes

  · Check the data integrity by 'Alder32 code'

  · Use 'eos cp' to EOSCTA

· Plan to phase out all nodes of CASTOR and LTO4 tapes

	BESIII	DYB	JUNO	LHAASO	YBJ	BACKUP	TOTAL
FILE NUMBER	2,831,504	5,421,476	61,578	6,727,408	603,010	3,500,271,515	3,515,916,491
USAGE(TB)	3402.505	2599.82	29.644	5415.814	525.551	265.845	12,239.179

Problems

- Special characters(*,\,space) lead to failed requests
- Retrieval has more steps than archiving
  - d0:t1 files need to pre-stage first
  - Complicated to user
    - xrdfs stat->xrdfs prepare->xrdfs query prepare->xrdcp
  - A command automatically pre-staging and retrieving?
- Other(set 'undelete', Cancel a retrieval request, Tape Re-write ……)

Next Steps

- Development for experiments
  - JUNO & LHCb Tier1: data challenge is underdoing
  - HEPS: debug with DTS(Data Transfer System) and DOMAS(Data Management System)
- Improvement
  - Data Integrity, monitoring, virtualization
- Maintenance
  - Multiple cta-frontend instance in VMs?
- PostgresSchedDB in production?
  - To retire Ceph clusters

Summary

- CTA is an essential part of storage system at IHEP

- Tests and development are done for various experiments

- We plan to update CTA version on Alma Linux 9

Special thanks to the CTA Team and CTA Community!

Thank you