ASGC Site Report

Eric Yen, Felix Lee

Academia Sinica Grid Computing Centre (ASGC)
Taiwan

HEPiX Spring 2024
Paris, France
15 April 2024

ASGC Overview

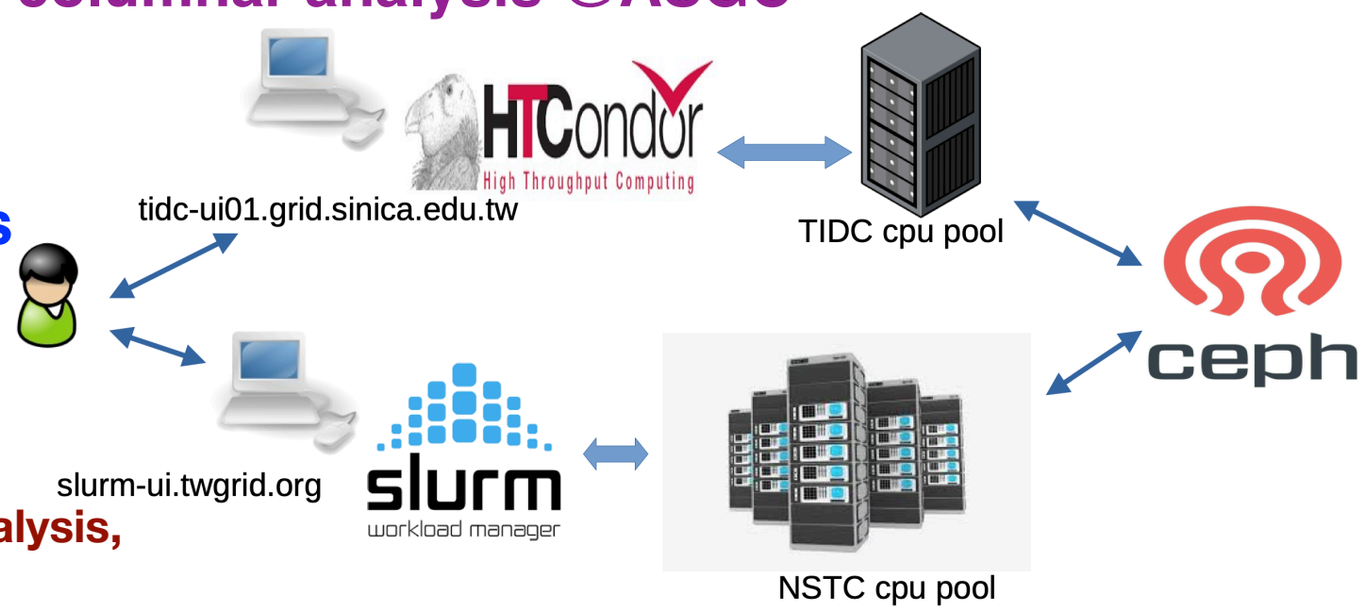
- Founded for participating WLCG and supporting the research collaborations
  - Being/serving WLCG Tier-1 center from Dec 2005 to Oct 2023
  - Migrating to WLCG Tier-2 center for ATLAS after Q3 2023
- ASGC is providing big data analysis and computing services for the R&E communities in Taiwan as a core facility
  - Funded by both Academia Sinica and National Science and Technology Council
  - Primary scientific collaborations: WLCG (ATLAS, CMS), AMS, Gravitational Wave, ICECube/ Neutrino, EIC, QCD, CryoEM, condense matter, etc.
  - Based on the core technologies of WLCG
- System efficiency as well as AI workload are the 3rd focus

ATLAS T2 Site Status

- Migration to ATLAS T2 site after Oct. 2023
- Successfully migrated to EOS from DPM with help from ATLAS in 2023
- Computing Resource of Federated Taiwan Tier-2 (2024):
  - 17,500 HEPSpec06 (1,864 CPUCores)
  - GPU would be available after validation of new computing models (ATLAS, CMS)
- Storage Resource (2024) of ASGC T2: 5 PB
  - EOS has been available for ATLAS
- Connecting to LHCONE by new ASN 18217 through ESNet at Chicago (StarLight) since Jan 2024
  - More connections with LHCONE at SG, JP, AMS or LA are under construction
  - Daily outgoing data volume could reach to 12.5TB (48TB/week) in March 2024
- Ready for production if no further issue arises
  - Site functionality has been validated by daily Hammercloud jobs
- Data efficiency is the only issue now ! Esp. between ASGC and US T2 sites
  - IPv6 + MTU issue had been identified in March
    - IPv4 has kernel level MTU probing supported
- ASGC will keep up the HGTD (High Granularity Timing Detector) DB services together with the ATLAS Taiwan team

CMS T3

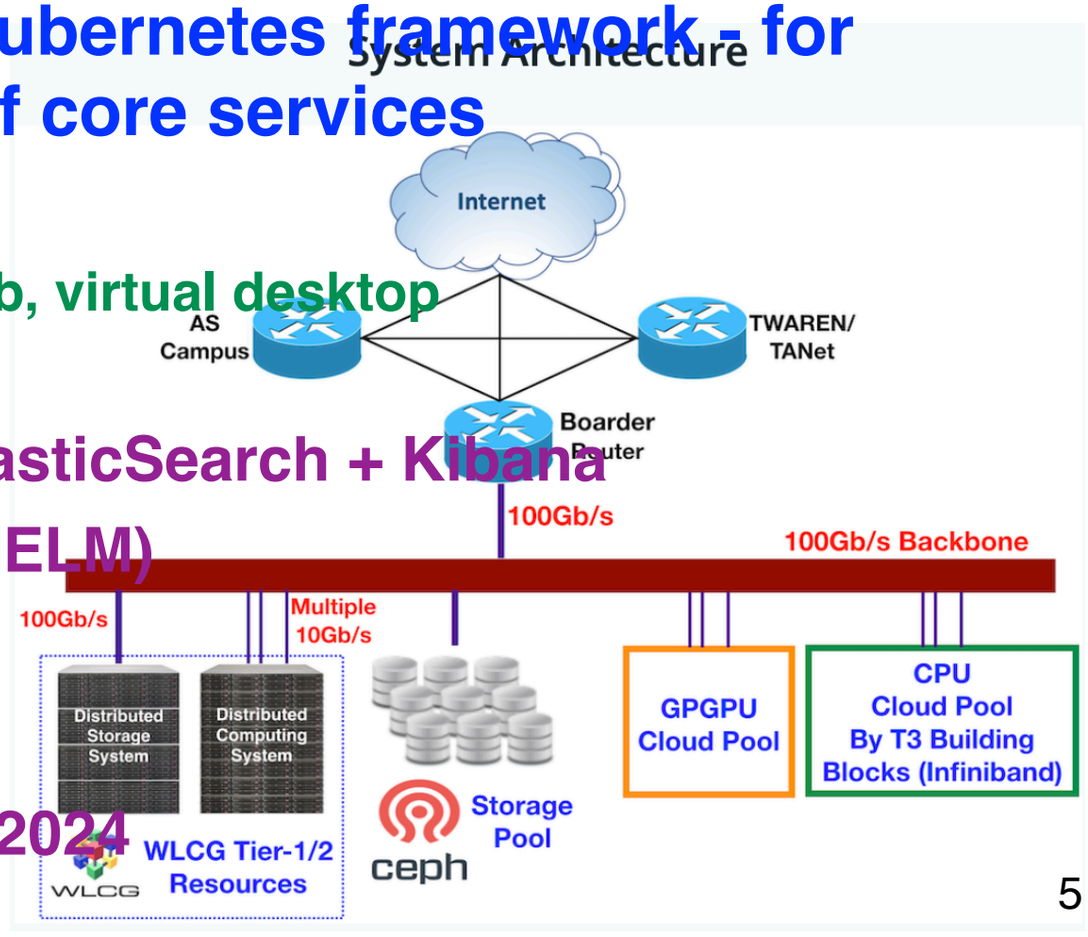
- Operated by ASGC from 2022
- In collaboration with TIDC (Taiwan Instrumentation and Detector Consortium) and local CMS groups (NTU and NCU)
- Analysis facility (HW is provisioned by TIDC)
  - Both Condor/UI and Crab/ARC-CE are available
    - CephFS shared filesystem: 3TB/group by default
    - EOS by xrootd and fuse: 1PB
  - Condor cluster
    - 768 cores(AMD EPYC 7713) + 768 cores (Intel CPU E5- 2650 v4)
    - Supporting user's access to CMS data or CRAB jobs submission
    - JupyterLab
- User training and support are also provisioned
  - e.g., using Coffea with template for columnar analysis @ASGC
- TIDC supports collaborations such as ATLAS, CMS, STAR, sPHENIX, EIC



Coffea = Column Object Framework For Effective Analysis,
https://github.com/CoffeaTeam/coffea

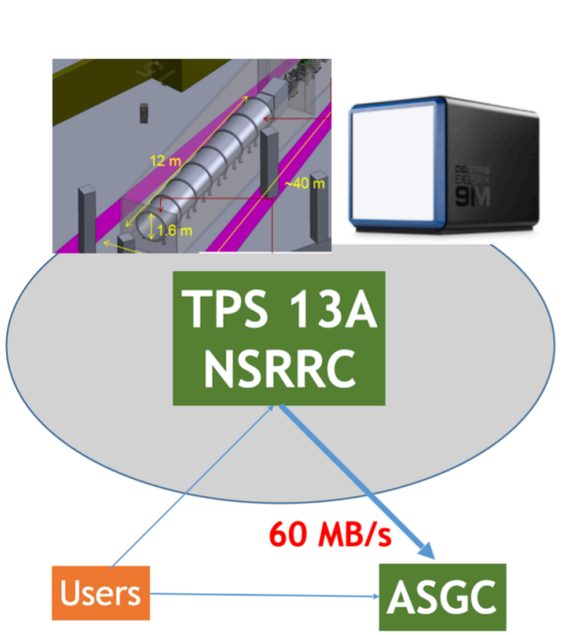
ASGC Science Cloud Infrastructure (DiCOS)

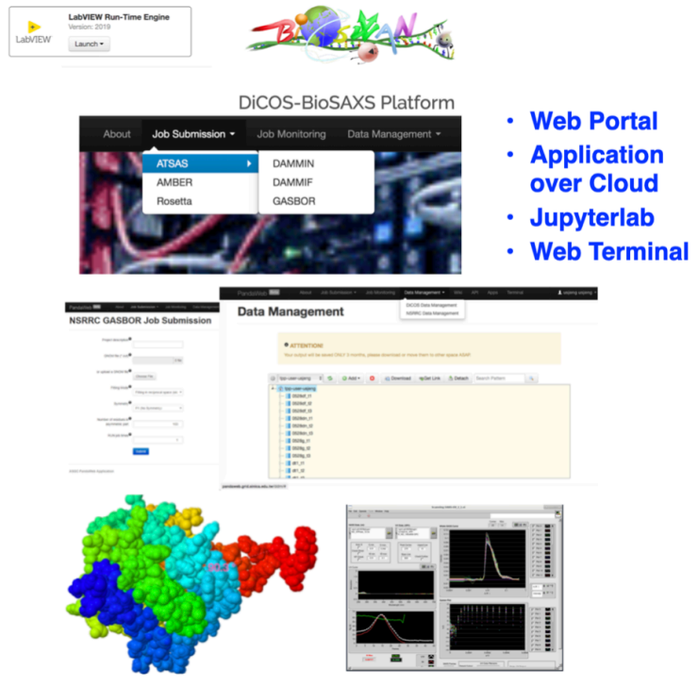
- Resource: 10,000 CPUCores, GPU Boards- A100x52, V100x48, 3090x56
- OpenStack Cloud: for core services and on-demand worker nodes maintained by Openstack-ansible
  - Multiple cells/Region for various configurations and capabilities
    - e.g. Neutron Compute, Nova Compute, …
  - Single hypervisor type: KVM
    - #hypervisors: 100+
    - #VMs: 500+, dynamic provisioning
  - Networking: flat and segmented
- Containerized Resources managed by Kubernetes framework - for software on-demand services and part of core services
  - User cluster:
    - batch, interactive GUI jobs: remote Jupyterlab, virtual desktop
    - GPU Cloud
  - Core Services: distributed cloud cores; ElasticSearch + Kibana
  - High availability is enabled (managed by HELM)
- UI: Web UI/Terminal; JupyterLab
- Linux Strategy
  - Migrating to AlmaLinux9 from CentOS7 in 2024

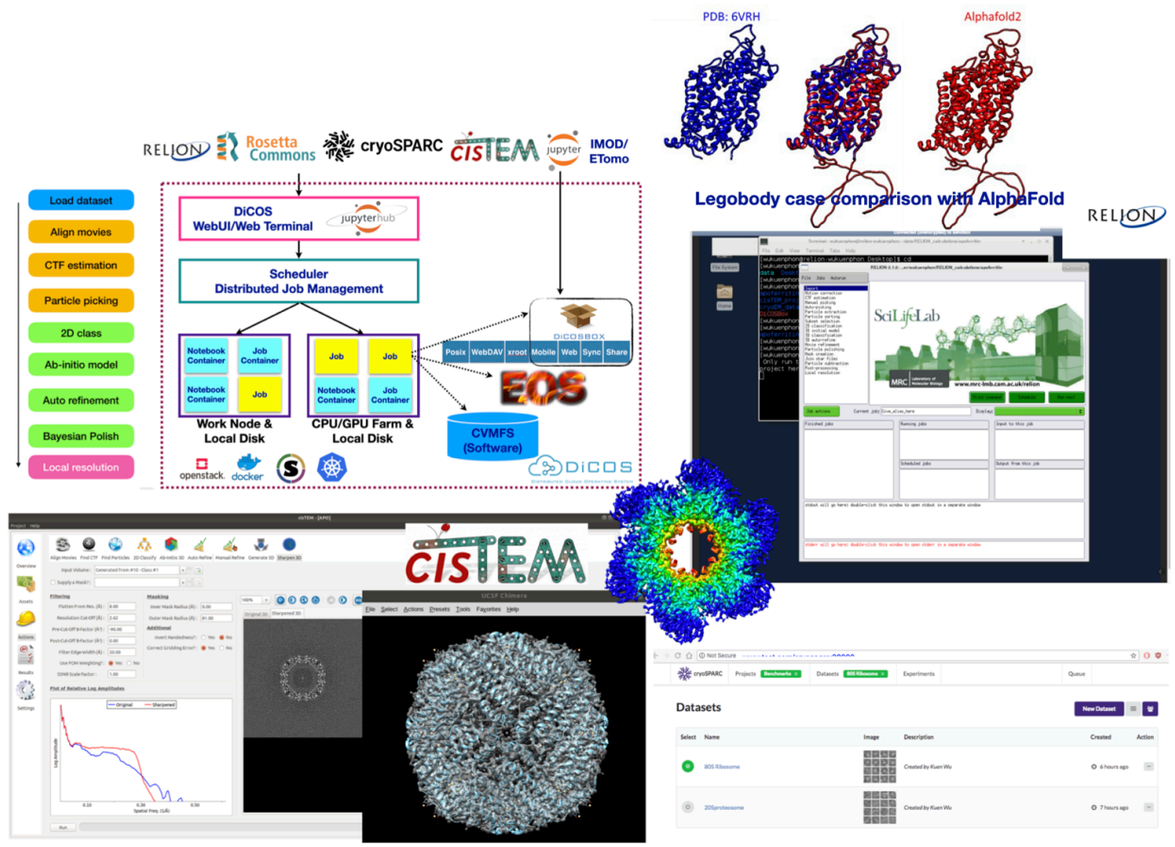


Cross-Facility Data Processing for Structural Biology (CryoEM & X-Ray)

- **Customized automated data analysis pipeline**
  - Cloud Service：better scalability, efficiency, performance
  - Integration & Optimization: data pipeline and workflow (computing model)
    - Reduced latency between experiment facility (data source) and data analysis facility
    - Integrating required software, application framework, storage and analysis workflow
    - Reduced latency between data and training model
  - Web Service: developing Web App, Web Portal or Science Gateway
  - Generalization and new service creation
  - On-the-fly data transmission
- **Example：Structural Biology: NSRRC (BioSAXS) and Taiwan Protein Project (CryoEM Core Facility)**
- **ASGC Science Cloud**
  - Application platform + Computing Infrastructure + Workflow Integration + Efficiency Optimization
  - AlphaFold, RosettaFold, RosettaFold Diffusion, DiffDock

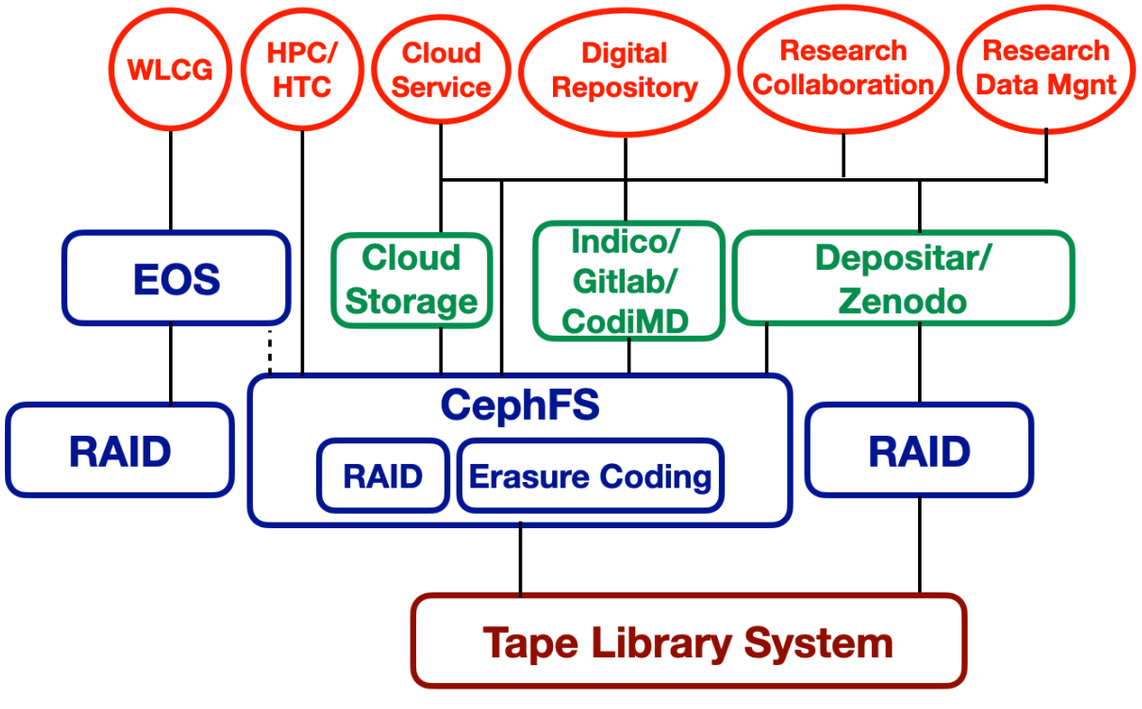


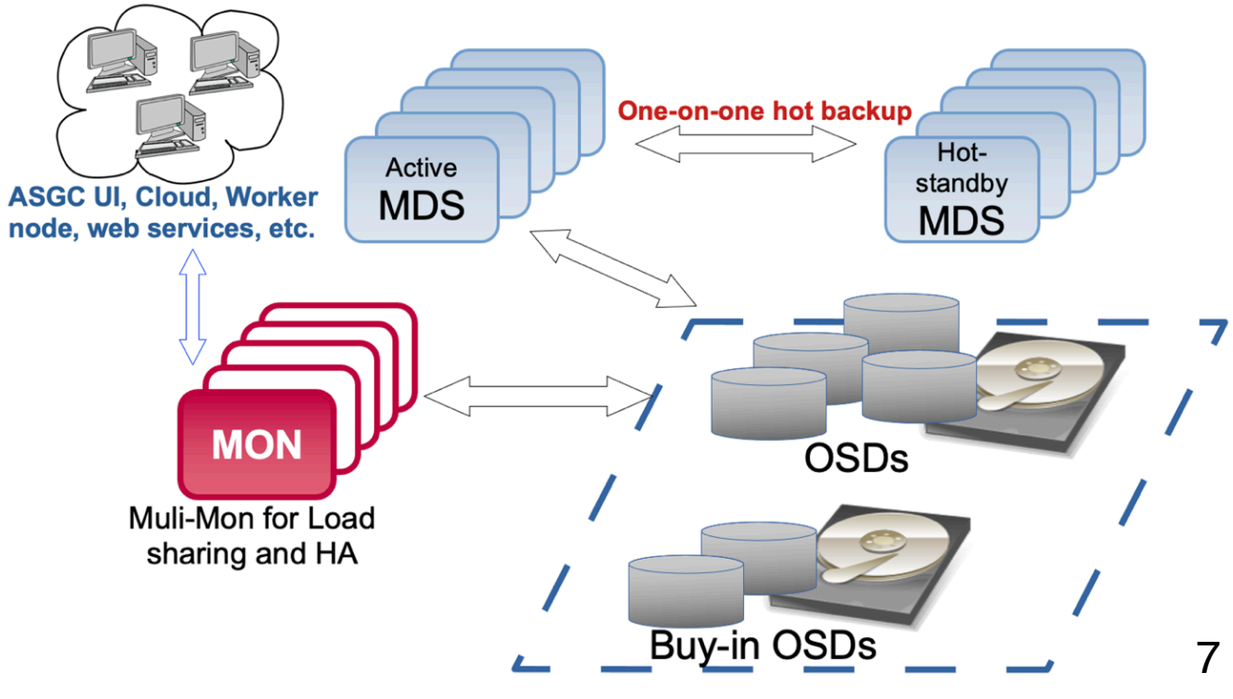




ASGC Science Cloud Storage Architecture

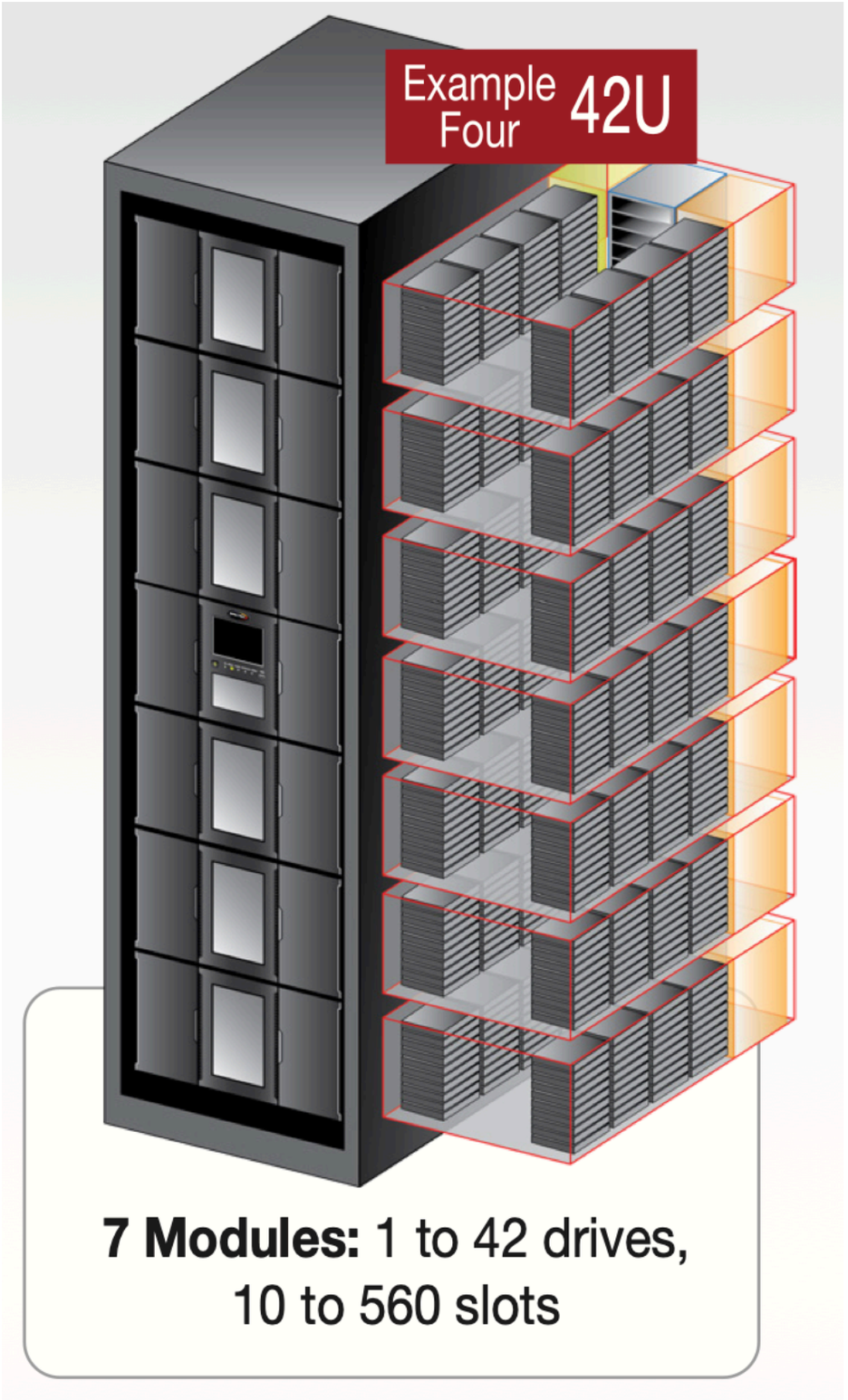
- Scalable and reliable online storage system based on Ceph mainly
- Ceph Configurations: ~10PB
  - 6 MDS + 6 hot-standby (one-on-one backup); 7 MONs
  - 462 OSDs, 51 hosts.
- Services
  - 3 TB/PI Group setup by default; PI could extend the space through management UI flexibly
- Reached 2GB/s R/W throughput so far
- Tape-based remote backup system (12PB) will be established and integrated in 2Q24
- Providing big pool for HPC, HTC, AI and various applications concurrently
- Ceph capacity will be growing to 13PB by end of 2024
  - Plan to procure new 4PB disk servers for Ceph System in 2024 and 2025 respectively





Tape Library System Will be Online by end of June 2024

- Serving as 2nd layer remote backup system
  - For cold data, or 2nd-copy backup
  - For backup of users' core data on Ceph
  - Long-term storage services
  - Reliability of ASGC services will be increased
- Scalability: capacity on demand
  - Max 7 modules x 6u, 42 drives, 560 tape slots
  - LTO-9 tape: 1.44PB (native)/3.6PB (compressed) per module (80x 18TB/tapes)
  - Installed capacity: 12PB
- Tape drive performance: 300 - 400 MB/s using fibre network
- Integration and services: based on EOS and CTA
- Tape-related data services should be operational by end Q2 2024.



Energy Saving

- Reliability enhanced by intelligent monitoring and control is the key approach
- Retirement of legacy hardware
- Improvement of AHU efficiency, including the top-flow cold air
  - Anomaly detection
  - Well-prepared backup plan
- Energy-sensitive operation
  - Plan for power efficiency hardware: e.g., ARM CPU
  - Power saving
    - 20% power usage reduction in 2023 - Effective on 3 CPU clusters (> 3,000 CPUCores) from May 2023
- Overall, 20% DC power usage decreased in 2023

OS Migration Plan

- **e-Science Local clusters & resources:**
  - **Making AlmaLinux9 computing resource available by end of April 2024**
    - **Part of worker nodes will be upgraded from CentOS first**
  - **User can test their code on AlmaLinux9 worker nodes.**
  - **After the CentOS7 EOL, start migrating all the computing resources to AlmaLinux9.**
    - **Depending on user application migration status**
    - **May also help user to migrate their code to Singularity images, in case some users are not able to migrate their codes to AlmaLinux9 in time or some incompatibilities remain.**
  - **Few application servers are using Ubuntu LTS**
- **WLCG resources: subject to the availability of Grid middleware and the WLCG strategy**

Future Plan

- Resource plan in 2024
  - New Intel computing server x1 with 256 Cores
  - New NVIDIA GPU: 4090 x1
  - More storage for Ceph: +4PB
  - New WN: +2,000 CPUCores
- Tape Library: 12PB tape space managed by CTA/EOS will be available in Q3 2024
- Facilitating AI applications on Science
  - Physics-based, data-oriented analytics
  - Application & development environment
- Efficiency optimization enabled by AI will be improved
  - On Thermal, power, application, system (computing, data, networking, security), user support, etc.
- 2FA and SOC are ongoing