Overview of KISTI GSDC

ALICE Tier-1 Operation and Data Computing Services

Byungyun Kong
On behalf of KISTI-GSDC



15 October 2026 @ AFAD2026

Introduction

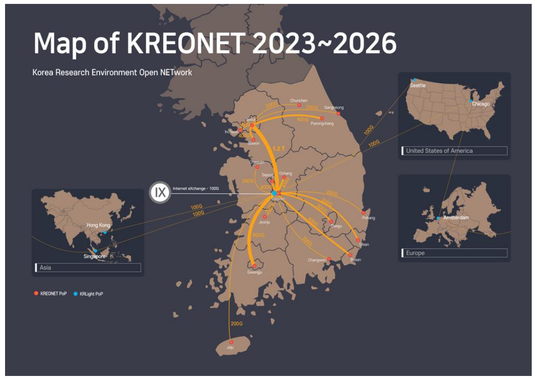
KISTI

Korea Institute of Science and Technology Information

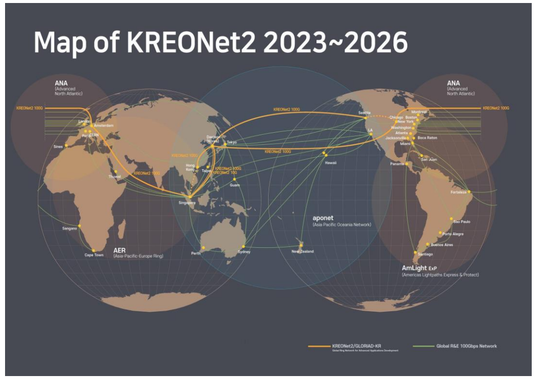


- Government-funded research institute founded in 1961 for national information services and supercomputing
- National Supercomputing Center
  - **5th Supercomputer**
    - **Nurion** - Cray CS500 system
      - 25.7 PFlops at peak, ranked 11th of Top500 (2018) ⇨ 144th (June 2026)
    - **Neuron** - GPU system, 1.24 PFlops
  - **6th Supercomputer** (**HANGANG**)
    - 614 PFlops at peak, 8,496 GPUs
    - Will start end of this year
  - **KREONet/KREONet2** - National/International R&E network









GSDC

Global Science experimental Data hub Center





NETMIKO

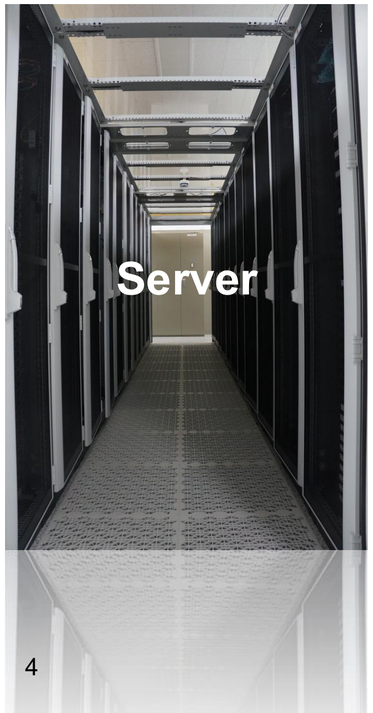


BIRD

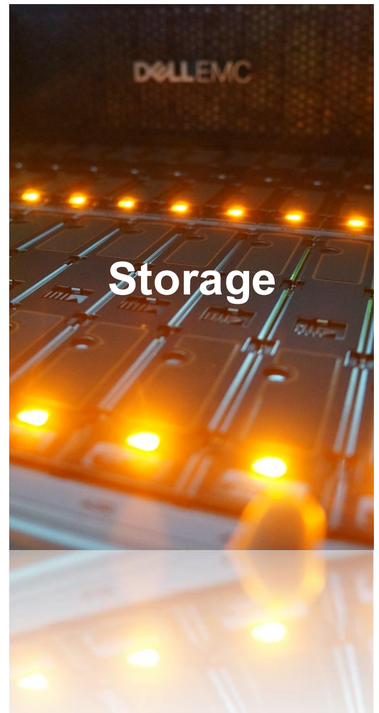
ANSIBLE

puppet

- Government-funded project, started in 2009 to promote Korean fundamental research through providing computing power and data storage
- **Datacenter for data-intensive fundamental research**
  - Preserving data from domestic or overseas large and complex scientific instruments as well as simulation-R&D activities
  - Providing services based on technology development: distributed computing structure, high availability storage system, infra integrated management, disk-based custodial storage







Role of GSDC for Data-intensive Research

KISTI-GSDC

Process

Constructing Facility → Conducting Experiment → Data Comp. Env. → Analyzing Data → Scientific Achievement

Subject

CERN(EU) | LIGO (US) | KEK (JP)
TEM | PAL | RENO
Large-scale Facilities

Users | User Groups
Research Community

Data Computing for Basic Science as a Service

Data Provider



Domestic | Overseas
Large-scale Facility

Requirements by experiments/community →
Data Computing Research Environments

APIs

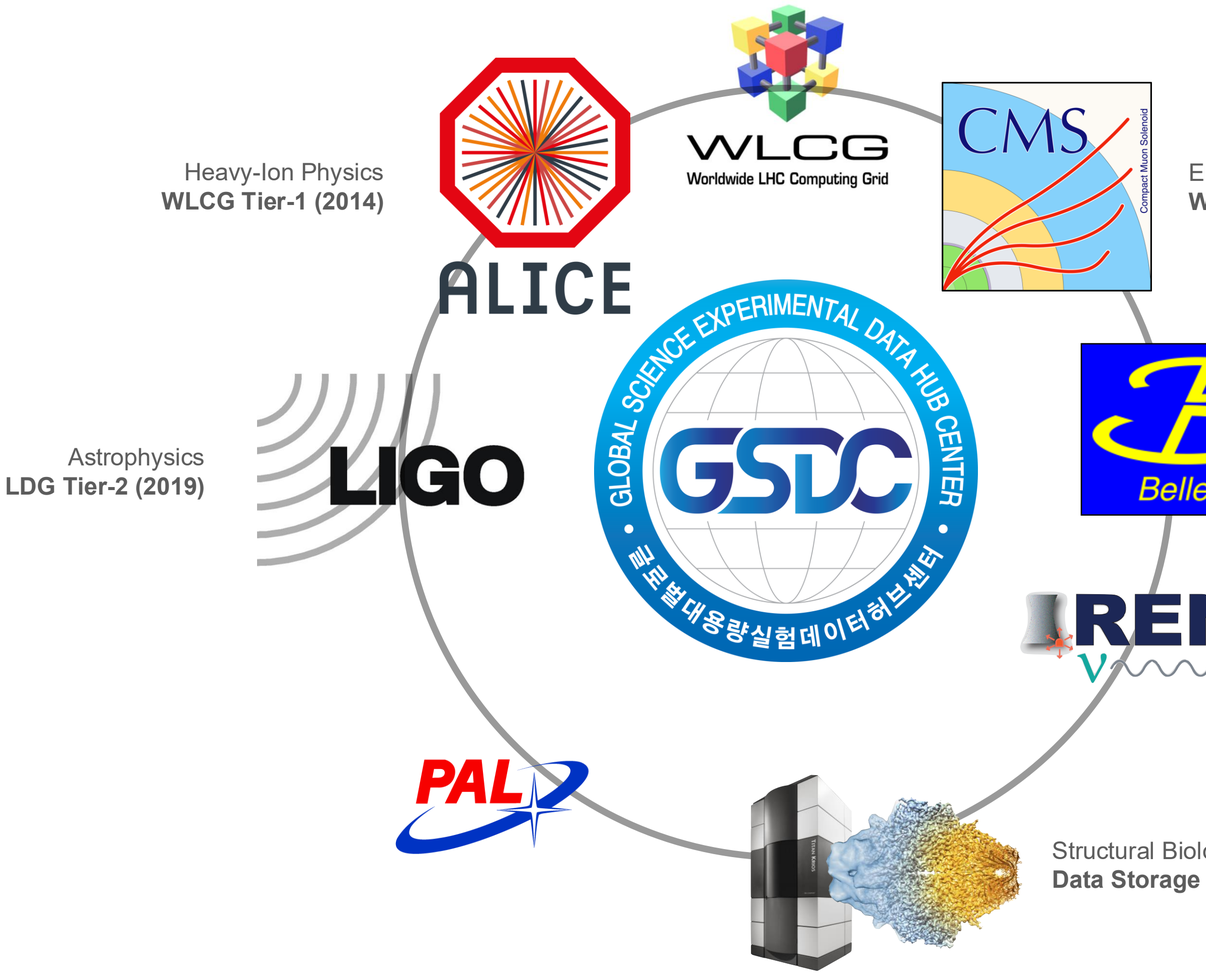
System Auto -mation	Resources Provisioning	Data Management	Data Analysis	Monitor -ing
Identity /Policy Mgmt.	• Server (Container, VM, Baremetal) • Storage (Block, Object, File) • Network	• Transfer • Movement • Metadata • Access ctrl. • Archiving	• Batch • Parallel/Dist. • Streamline • In-memory • DeepLearning	

Data Center Infrastructure

Data-Intensive Research

ALICE CMS LIGO RENO PAL 포항가속기연구소

Supporting Experiments



Heavy-Ion Physics
WLCG Tier-1 (2014)

Elementary Particle Physics
WLCG Tier-2 (2018)

Astrophysics
LDG Tier-2 (2019)

Elementary Particle Physics
Computing cluster

Neutrino Physics
RAW Storage

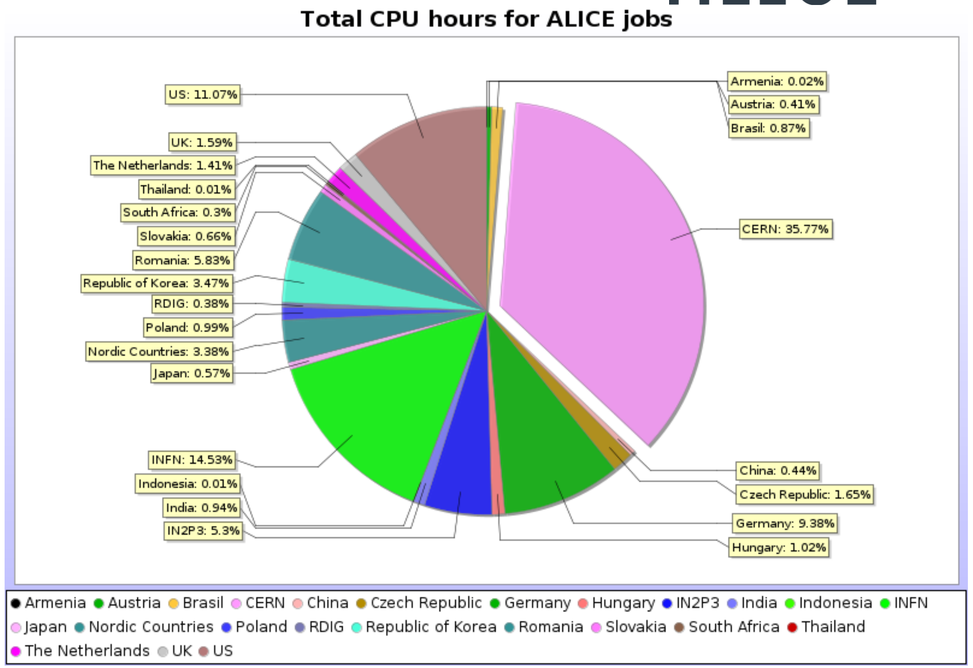
Structural Biology using Cryo-EM/XFEL
Data Storage & Analysis Pipeline



Tier-1 Operation status

WLCG Tier-1 @ KISTI-GSDC

Flagship Service for Data-intensive Computing



- A WLCG Tier-1 in Asia for the ALICE experiment
  - Contributing about 10% of T1 resource requirements of ALICE
    - More than 3% of total (T0+T1+T2+AFs) resource requirements of ALICE
    - Availability/Reliability: 99.34% / 99.34% (2026/01-09)
- CE
  - HTCondor-based, whole-node submission enabled (for N-core jobs)
- SE
  - XRootD/EOS based disk storage
  - Archival SE : CDS, the disk-based one powered by EOS
- Networking
  - LHCOPN : 20G dedicated link between Daejeon (KR) and Geneva (CH)
  - LHCONE : 100G provisioned by KREONet connecting to EU, US and Asia (SG/HK)

KISTI ALICE T1 Structure Overview

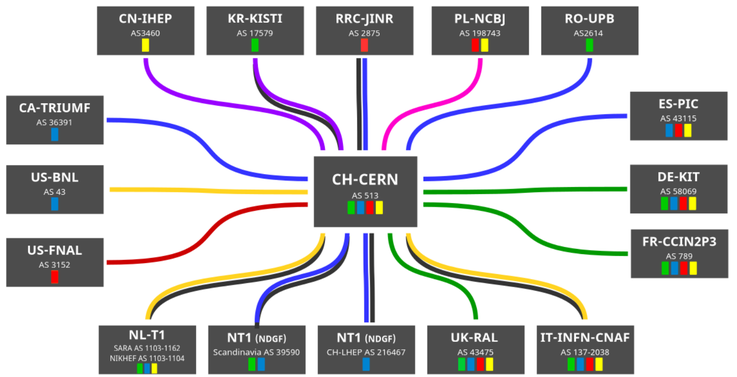


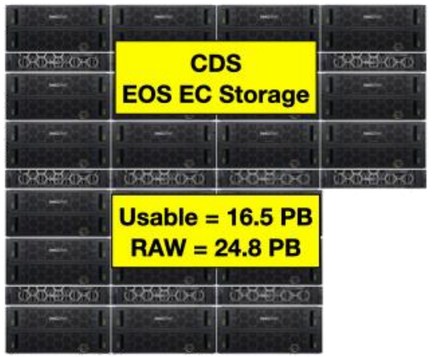
Simulation
Organised Analysis
Reconstruction
User Analysis
Job Queue
Job Broker
allocation
WLCG CRIC
publish
Accounting
APEL
statistics
Job
Data
CERN
HTCondor
6,760 Cores
106.81 kHS
HTCondor-CE
VO-Box
submission
CE
scheduling
Batch Cluster
HTCondor
query
collect
AuthN/AuthZ
BDII/Argus
collect
collect
read only
data i/o
80 PB
Erasure-Coded storage
Detector
ALICE O2
80PB/yr
O2 Buffer
real-time data transfer (> 25Gb/s required)
Archival Storage
raw
Disk Storage
processed
archive
~ 1EB
T0

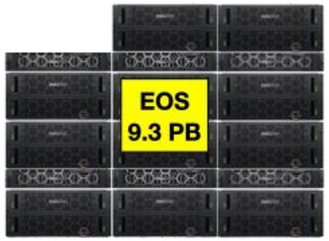




LHCOPN
KISTI-CERN 20Gbps

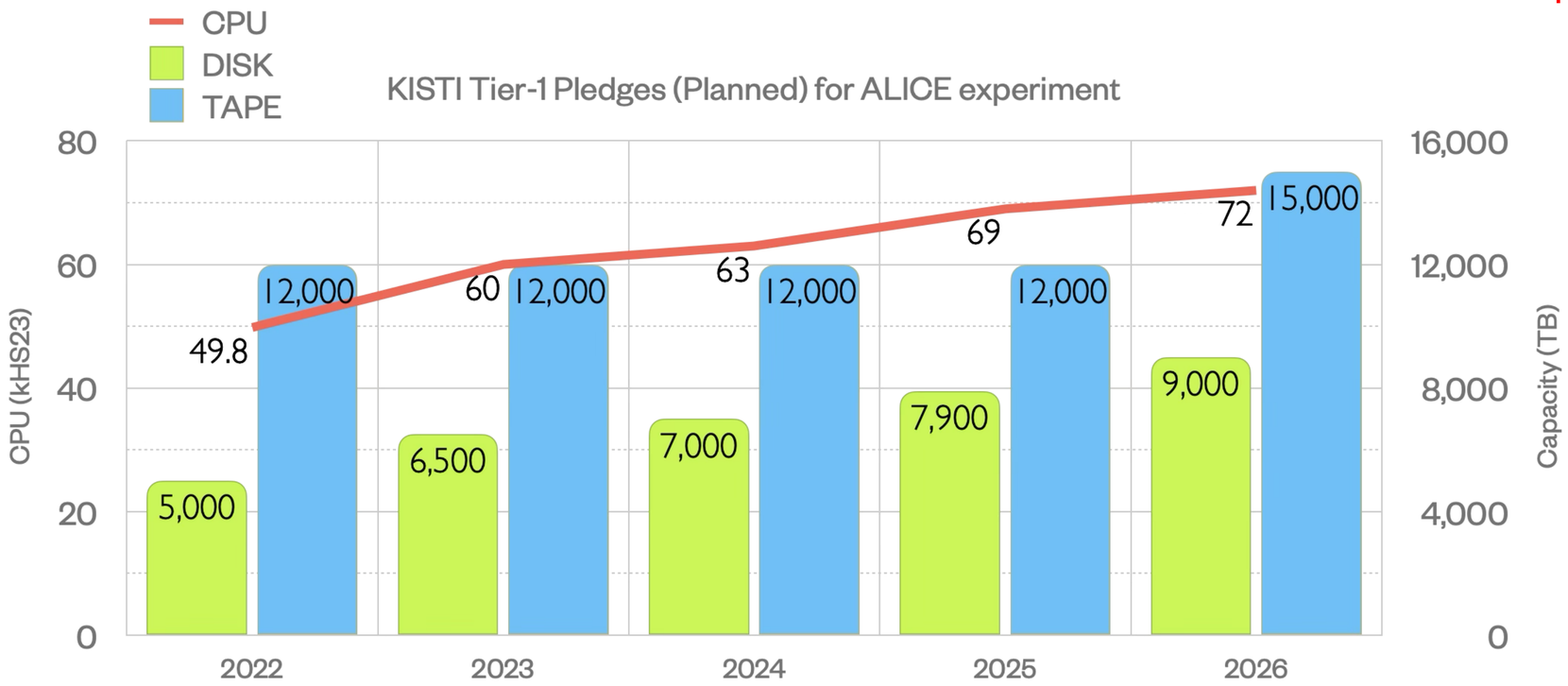






Tier-1 Pledges

10% Contribution to ALICE Tier-1 Computing Requirements



Installed (Planned)	2022	2023	2024	2025	2026
CPU (kHS)	49.8	60.0	63.0	69.0	72.0
DISK (TB)	5,000	6,500	7,000	7,900	9,000
TAPE (TB)	12,000	12,000	12,000	12,000	15,000

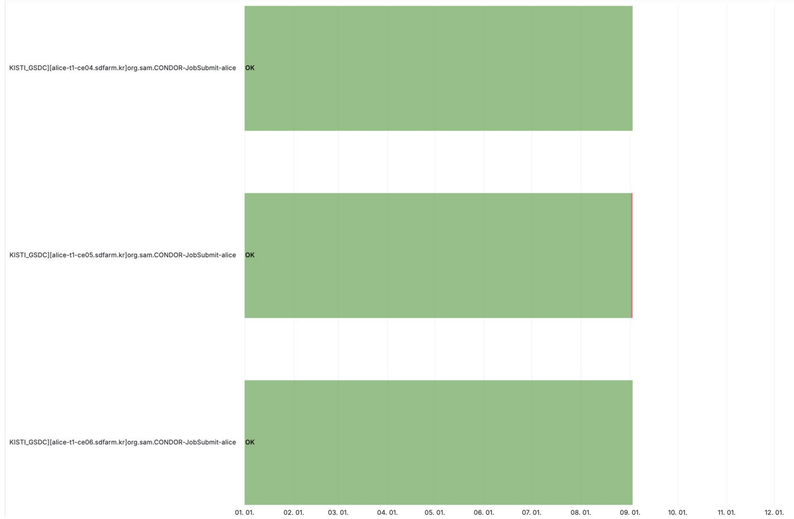
Computing Elements

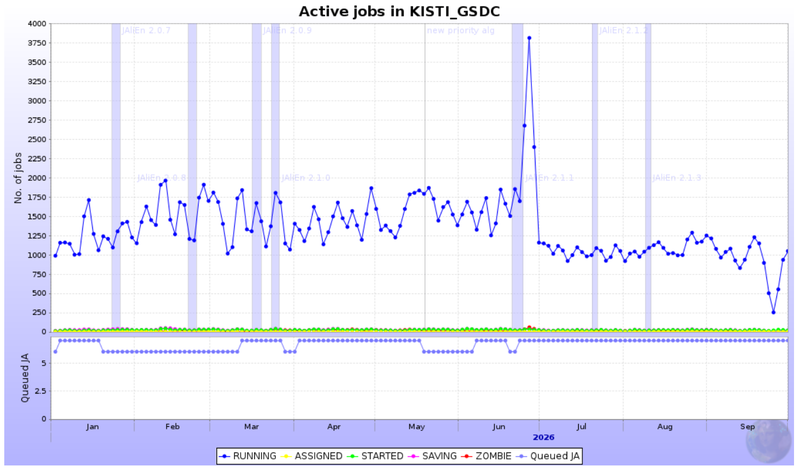
- Computing Resource

HTCondor-CE Hostname	Number of Worker Nodes	HepScore 23 Benchmark score
alice-t1-ce04.sdfarm.kr	12	21,029
alice-t1-ce05.sdfarm.kr	22	31,791
alice-t1-ce06.sdfarm.kr	10	53,988

- Pledge: 72 kHS23 (2026)
- Total: 106.8 kHS23

- Large-Scale Legacy Infrastructure Refresh
  - Decommissioned a substantial portion of aging servers nearing end-of-life (EOL)
- High Compute Density per Node
  - Achieved 5,398 HEPscore per node with the newly deployed hardware





Number of jobs at KISTI_GSDC

	Series	Last value	Min	Avg	Max
1.	1	364	0	415.9	5970
2.	2	101	0	188.3	2542
3.	4	5	0	22.85	1324
4.	8	720	2	623.7	1022
5.	16	1	0	1	1
6.	64	1	0	1	1
7.	96	1	0	0.8	1

Storage Elements

EOS Disk

- Structure
  - Running a container-based EOS server directly with podman
    - Container Image: AlmaLinux 9.6
    - 6 EOS servers (v5.4.8)
      - 1x(MGM+QDB+FST) + 4x(QDB+FST) + 1x (FST)
      - MGM DNS (eos-disk.sdfarm.kr)
- Status
  - Total : 9.30 PB
  - Used : 6.61 PB

CDS Archiving

- Structure
  - Running a container-based EOS server directly with podman
    - Container Image: AlmaLinux 9.6
    - 14 EOS servers (v5.4.8)
      - 1x (MGM+QDB) + 2x (QDB) + 12x (FST)
      - MGM DNS (eos-archive.sdfarm.kr)
- Status (Installed / Usable after EC)
  - Total : 24.80 / 16.53 PB
  - Used : 20.52 / 13.68 PB

Disk storage elements

GSDC

	AliEn SE		Catalogue statistics (1024-base units)						Storage-provided information (1024-base units)						Functional tests							Last day add tests		Demotion
SE Name	AliEn name	Tier	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage	Version	EOS Version	add	get	rm	3rd	IPv6	HTTP	Last OK add	Successful	Failed	factor
1. KISTI_GSDC - EOS	ALICE::KISTI_GSDC::EOS	1	8.263 PB	5.803 PB	2.46 PB	70.23%	100,493,171	FILE	8.263 PB	5.871 PB	2.392 PB	71.05%	Xrootd 5.9.4	5.4.8							02.09.2026 09:41	24	0	0
Total			**8.263 PB**	**5.803 PB**	**2.46 PB**	**70.23%**	**100,493,171**		**8.263 PB**	**5.871 PB**	**2.392 PB**	**71.05%**			**1**	**1**	**1**	**1**	**1**	**1**				

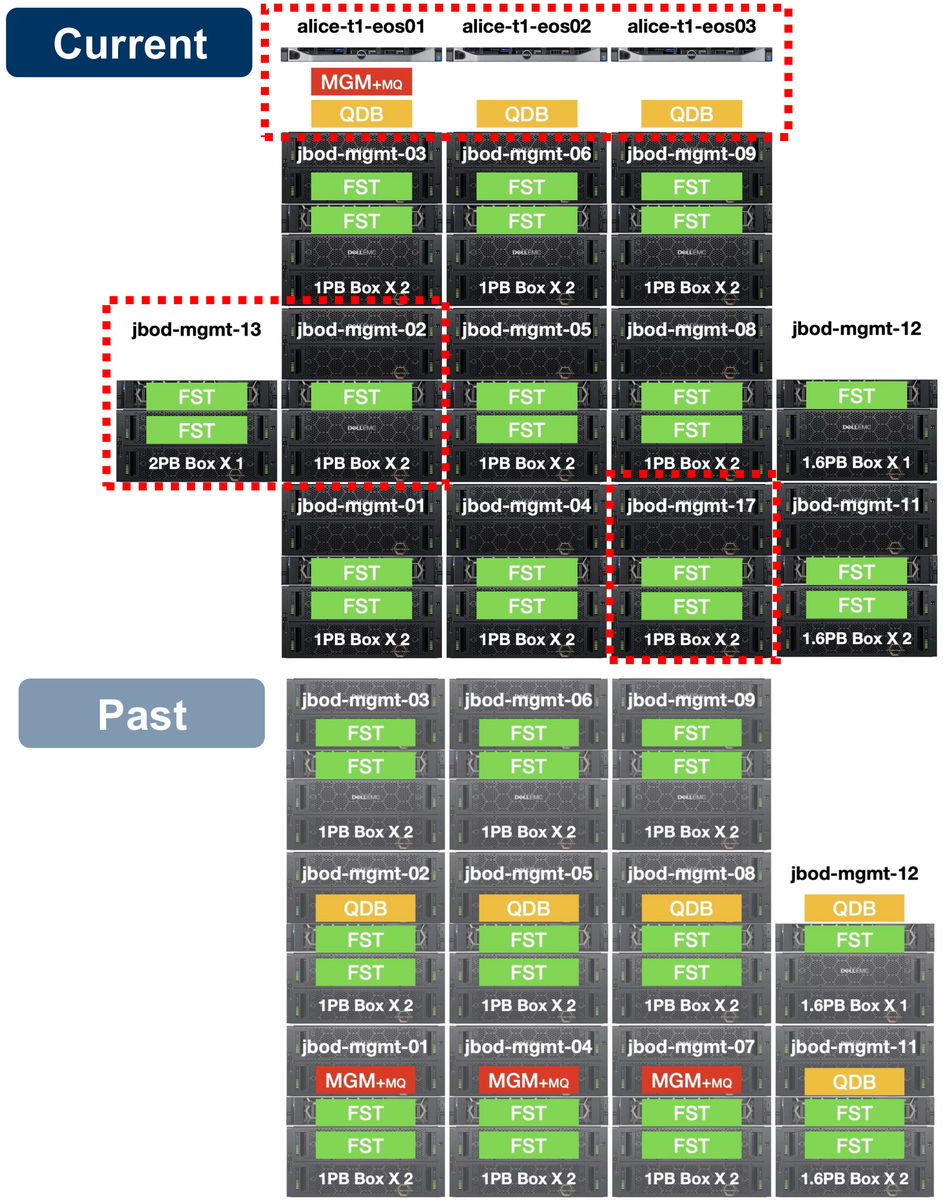
Custodial storage elements

GSDC

	AliEn SE		Catalogue statistics (1024-base units)						Storage-provided information (1024-base units)						Functional tests							Last day add tests		Demotion
SE Name	AliEn name	Tier	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage	Version	EOS Version	add	get	rm	3rd	IPv6	HTTP	Last OK add	Successful	Failed	factor
1. KISTI_GSDC - CDS	ALICE::KISTI_GSDC::CDS	1	12.5 PB	12.5 PB	-	100%	13,809,522	FILE	21.95 PB	18.22 PB	3.723 PB	83.04%	Xrootd 5.9.4					Test...		-	02.09.2026 09:27	23	1	3.267%
Total			**12.5 PB**	**12.5 PB**	**0**	**100%**	**13,809,522**		**21.95 PB**	**18.22 PB**	**3.723 PB**	**83.04%**			**1**	**1**	**1**	**0**	**1**	**0**				

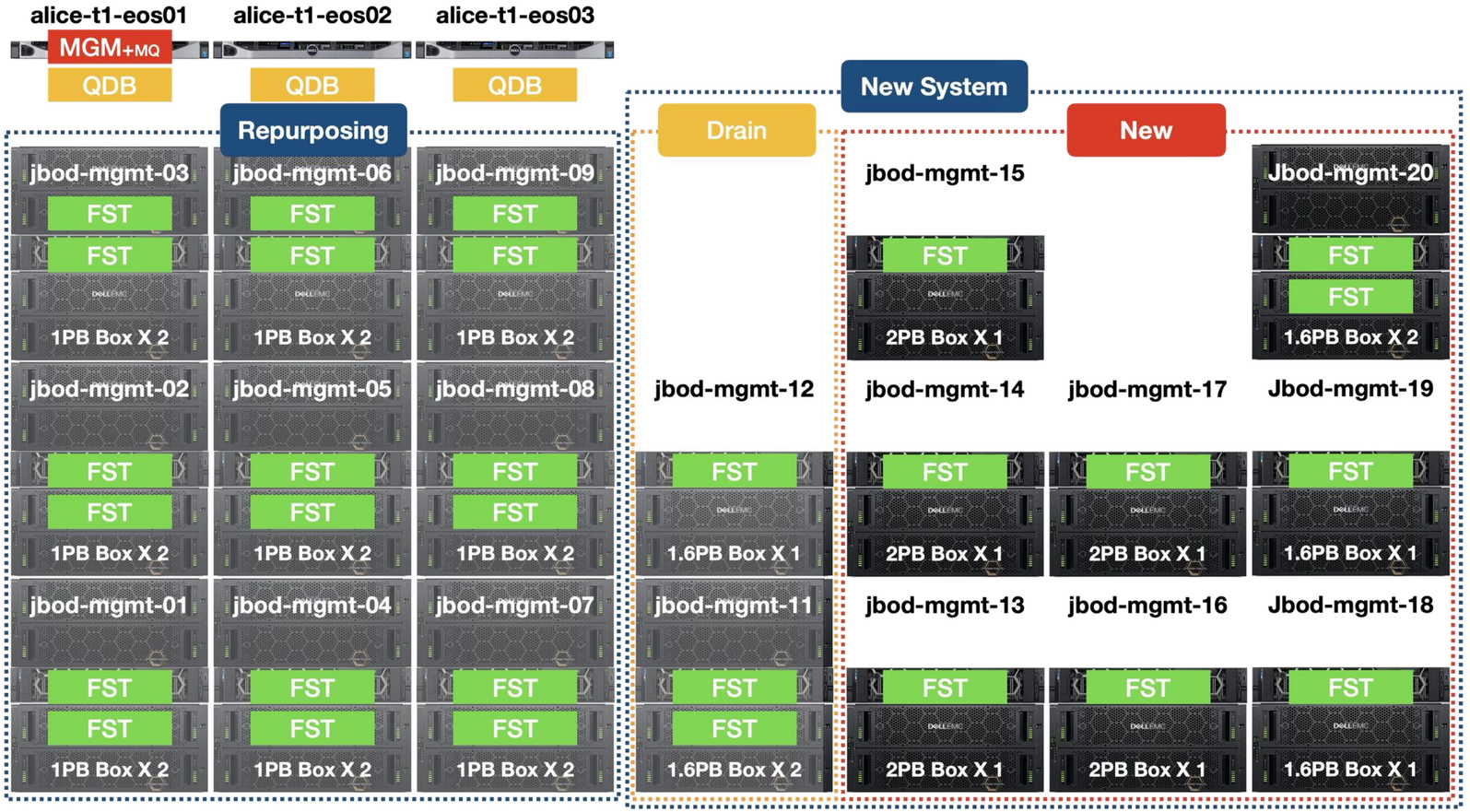
CDS Reconfig. & H/W Refresh

- **Control Plane Decoupling (MGM & QDB Isolation)**
  - Separated MGM and QDB from physical FST servers and migrated to dedicated VMs.
- **High-Load FST Workload Rebalancing**
  - Distributed high-throughput FSTs across balanced host nodes
- **Legacy Hardware Modernization**
  - Decommissioned aging storage nodes and upgraded to high-performance servers
- **Heterogeneous Storage Testing (Pilot)**
  - Attached JBOD disk boxes with different capacities to a single test node
  - Validated EOS/QDB handling and scheduling stability across asymmetric disk sizes



CDS Optimizing Structure

- **Scheme Optimization (12+4+2 → 10+2)**
  - Transitioning the CDS storage architecture from the legacy 12+4 to a 10+2 layout
  - Improves usable capacity efficiency while maintaining high data availability and fault tolerance
- **Disk Repurposing**
  - Integrating viable legacy CDS host chassis and qualified existing hard drives into the new setup
  - Minimizes capital expenditure by retiring only obsolete components while maximizing existing hardware

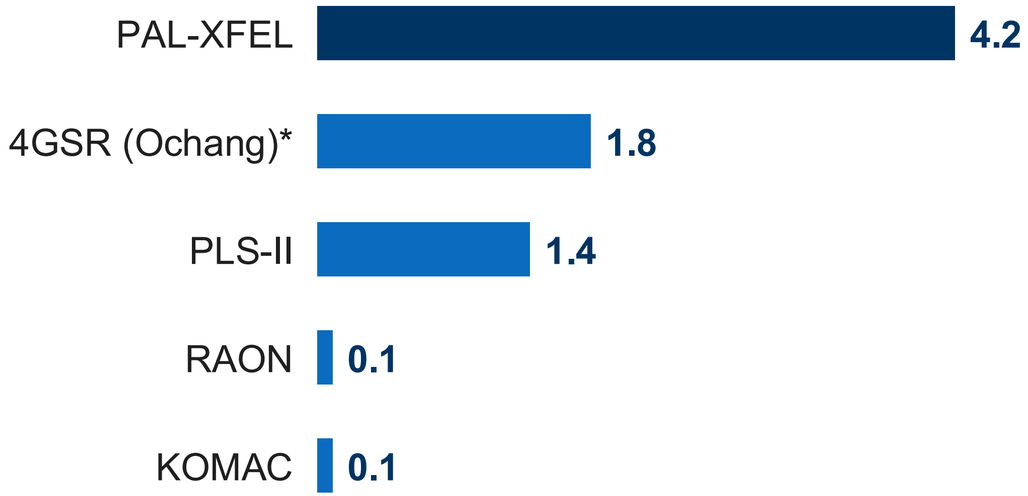


Computing for Accelerator Facilities

Data Challenges at Accelerators

Petabyte-scale data, managed by individual researchers

Annual raw data by facility (PB/yr)

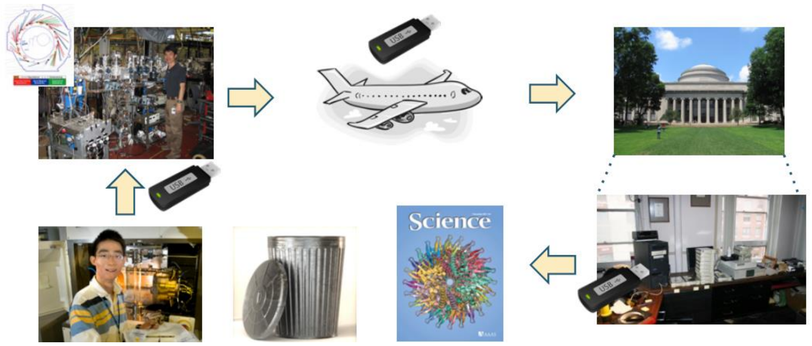


7.6 PB/yr raw → 9.9 PB/yr incl. derived data

Global: APS-U ~100 PB/yr · NSLS-II >80 PB/yr · SSRF ~30 PB/yr

* Under construction (2022–2027). Source: KISTI ISP for the integrated data hub (2026)

2025 user survey of facility researchers



Small-data approach: frequent manual copies, one dataset at a time

85%

manage data individually (personal 70.5%, institution 14.5%)

12.9%

use facility-linked computing (own resources 83.5%)

69%

expected cut in analysis time with HPC

From Planning to Program

A proposed national data hub for accelerators, 2027–2030

1

Jan–Mar 2025

Expert council: facility operators, users and system/IT experts

2

Aug–Nov 2025

ISP: interviews with KOMAC, PAL (PLS-II, XFEL) and RAON operators

3

Jan 2026

Final report: target model, 11 implementation tasks, resource sizing

4

2027–2030 (proposed)

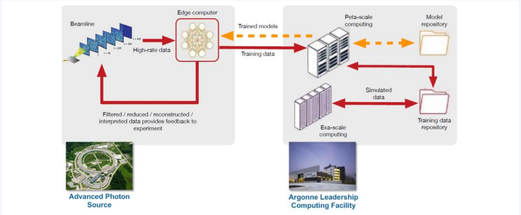
Strategic R&D program (2+2 yrs) led by KISTI with KAERI, KBSI, PAL, IBS

Reference models — integrating facilities with data centers



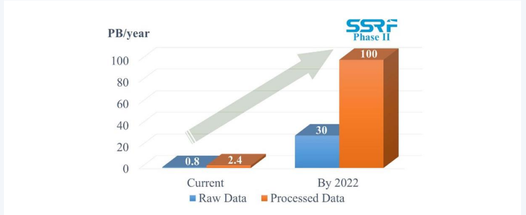
NERSC Superfacility (US)

Facility + data + computing + network + SW/AI; from ALS to fusion, linac, LZ



APS + ALCF (US)

Light source linked to leadership computing; 3.6 PB at APS, 200 PB at ALCF



SSRF + BDSC (China)

Big Data Science Center; ~30 PB/yr raw data expected after 2022



WLCG (HEP)

42 countries, 170 data centers; KISTI serves as a WLCG Tier-1

Target Architecture: Hub & Spoke

Onsite computing at each facility, one central hub

SPOKES · facility sites

PAL-XFEL | PLS-II | 4GSR | KOMAC | RAON

- DAQ integration & auto-metadata
- Onsite high-performance buffer
- Pre-processing & fast feedback
- DTN (data transfer node)

KREONet 100G

Science DMZ + DTN

HUB · KISTI integrated data hub

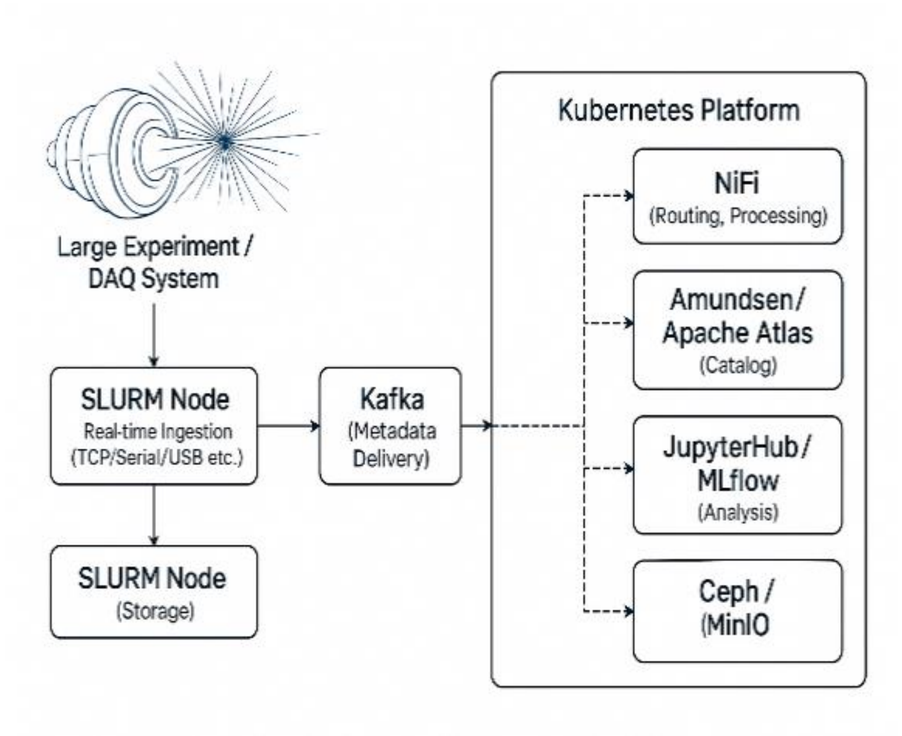
Services	Web portal · Data search & API · Analysis environments · Training
Platform	Kubernetes · Workflow engine · Data catalog (ICAT) · Federated AAI
Infrastructure	CPU / GPU clusters · Tiered storage (buffer → analysis → archive) · Security

To-Be workflow defined in the ISP

1 Plan → 2 Prepare → 3 Generate → 4 Transfer → 5 Store → 6 Analyze → 7 Share → 8 Preserve

Hub Infrastructure Design

On-premise cloud with tiered storage (ISP target model)



ISP reference stack: Kafka, NiFi, Amundsen/Atlas, JupyterHub/MLflow, Ceph/MinIO on Kubernetes

Onsite
High-performance buffer, data transfer servers, InfiniBand interconnect

Network
KREONet 100G between sites and hub, Science DMZ, DTN-based transfer

Compute
Kubernetes for containerized workloads; GPU and high-performance cloud nodes

Storage
Three tiers: high-performance buffer, analysis, archive

Access
Keycloak authentication; portal, catalog API and analysis pods

Security
Intrusion detection/prevention, anti-ransomware, security monitoring

Sizing Method

Compute, GPU and storage derived from beamline workloads

How the ISP sized the hub

Compute (CPU)
Data per research group × cores per TB-day (50; XFEL 10) × 3 concurrent groups × beamlines, + 15% overhead + 10% management

GPU
Average GPUs per beamline (operator interviews) × beamlines; 8 GPUs per node

Buffer
10% of annual raw data, at each site and at the hub

Analysis / archive
(Raw + 30% derived data) × 2 years / × 5 years

Facility	Beamlines	Raw PB/yr	Hub cores	GPUs
PAL-XFEL	3	4.2	10,238	14
PLS-II	36	1.4	4,266	80
KOMAC	4	0.1	325	14
RAON	9	0.1	4,875	32
4GSR	10	1.8	5,484	46
Total	**62**	**7.6**	**25,188**	**186**

Onsite computing adds 5,000 cores. Source: KISTI ISP final report (2026)

Capacity Plan & Rollout

Phased rollout from PAL-XFEL (2027) to 4GSR (2030)

2027	2028	2029	2030	2031~
PAL-XFEL	**PLS-II**	**KOMAC + RAON**	**4GSR**	**Further expansion**
+18.7 PB	+19.9 PB	+14.5 PB	+15.3 PB	priority: life science, astronomy, then fusion, nuclear
10.2k cores	4.3k cores	5.2k cores	5.5k cores	
14 GPUs	80 GPUs	46 GPUs	46 GPUs	

68 PB

storage by 2030 — buffer 1.5 · analysis 17.4 · archive 49.4

~30k

CPU cores — 25.2k hub analysis + 5k onsite

186

GPUs — 24 nodes of 8× H200-class

100G

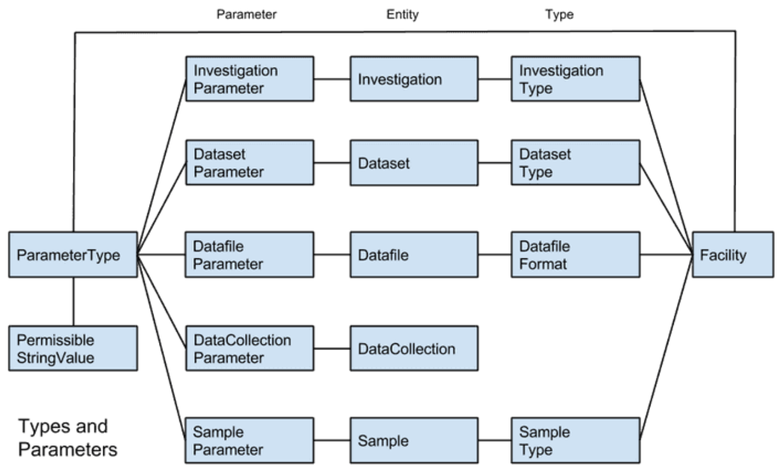
KREONet links between facility sites and the hub

Assumptions: raw data kept 5 yrs, analysis copy 2 yrs, derived data +30%, up to 3 concurrent groups per beamline. Source: KISTI ISP (2026)

Data Management & Access

ICAT-based catalog, DMP lifecycle and tiered access

Metadata model (ICAT)



Investigation → dataset → datafile, extended with typed parameters

Tiered access policy (FAIR)

Tier	Users	Data	Conditions
Operator-only	Facility staff	Raw data, logs, intermediate files	Internal network only
Researchers	Proposal PI and team, internal or external	Approved proposal data incl. raw	Account approval; download and analysis
Public	Researchers, citizens	Datasets converted to open	CC license, DOI, attribution
Restricted	PI and designated staff	Sensitive, pre-publication or patent data	Closed environment, expert review

DMP-based lifecycle: automatic migration and recall between analysis and archive tiers

Summary

- **Reliable Tier-1 Operation for ALICE**
  - A/R 99.34% / 99.34%, stable LHCOPN (20G) & LHCONE (100G); ~10% of ALICE T1 requirements
  - Infrastructure refresh: 106.8 kHS23 installed vs 72 kHS23 pledge; CDS control-plane decoupling, moving to a 10+2 EC layout
- **Computing for Accelerator Facilities**
  - Proposed national data hub (2027–2030): hub & spoke over KREONet 100G
  - Hub: Kubernetes-based platform, three-tier storage, GPU and Science DMZ / DTN transfer
  - Phased rollout PAL-XFEL → PLS-II → KOMAC/RAON → 4GSR: ~68 PB, ~30k cores, 186 GPUs
  - Data managed with an ICAT-based catalog, DMP lifecycle and tiered access (FAIR)

Thank you