



**NSTC** 國家科學及技術委員會  
National Science and Technology Council



# ASGC Open Science Cloud

Felix Lee, Eric Yen  
Academia Sinica Grid Computing Centre (ASGC)

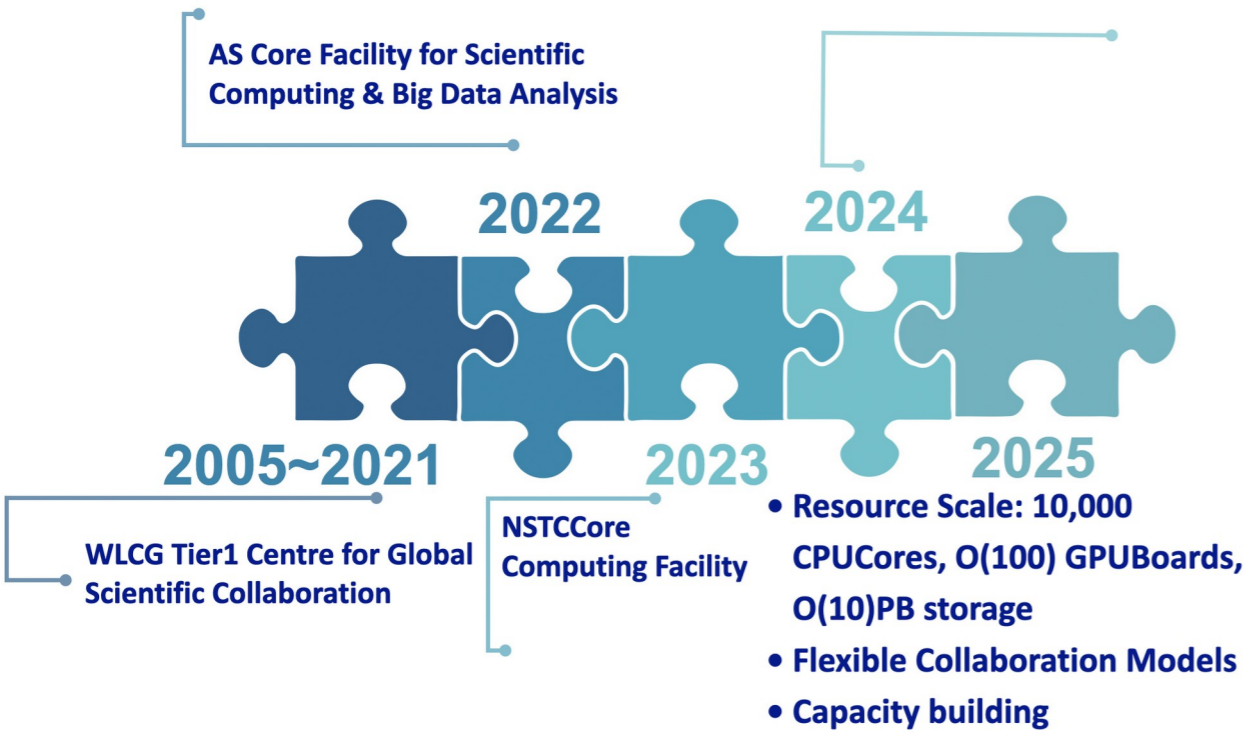
**ISGC 2026**  
**18 March 2026**

# ASGC Open Science Cloud - Accelerating Discovery & Innovation

- ASGC was founded for the deployment and development of WLCG system in 2005.
- Vision: accelerating scientific discovery and strengthening scientific computing ecosystem
  - User-oriented, service-based approach
    - Flexible collaboration models and customized services are strategic focus
    - Reliability and efficiency are primary objectives
  - Integrated platform of data, algorithm and computing
  - Federating community, application/service, RI & supporting team (service R&D)

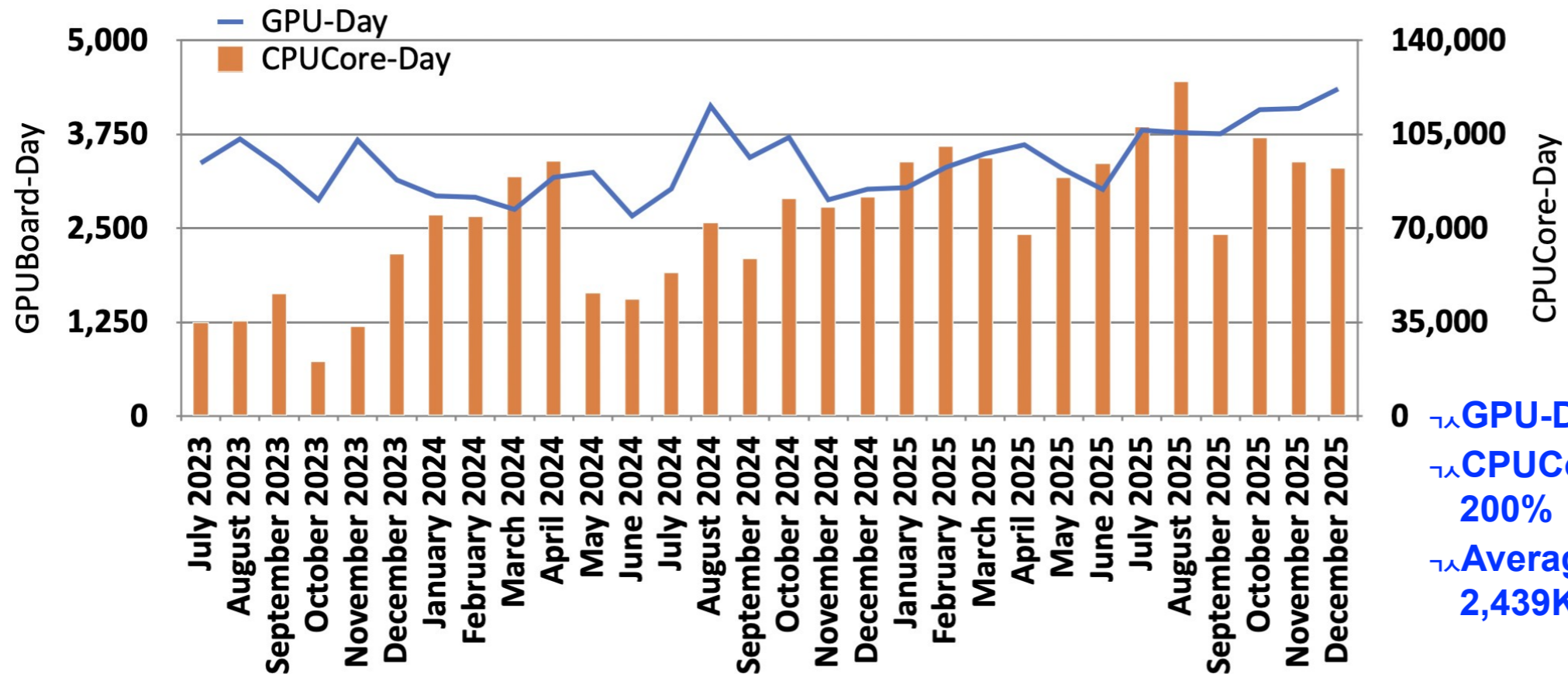


- |                                   |                         |
|-----------------------------------|-------------------------|
| Neutrino/MHEP                     | CryoEM                  |
| Quantum Materials Physics         | Bioimaging              |
| Physics of Active & Living Matter | Drug Discovery          |
| Astrophysics                      | Computational Chemistry |
| Earth Science                     | Biodiversity & Ecology  |



# ASGC Core Facility is an integral part of the research infrastructure

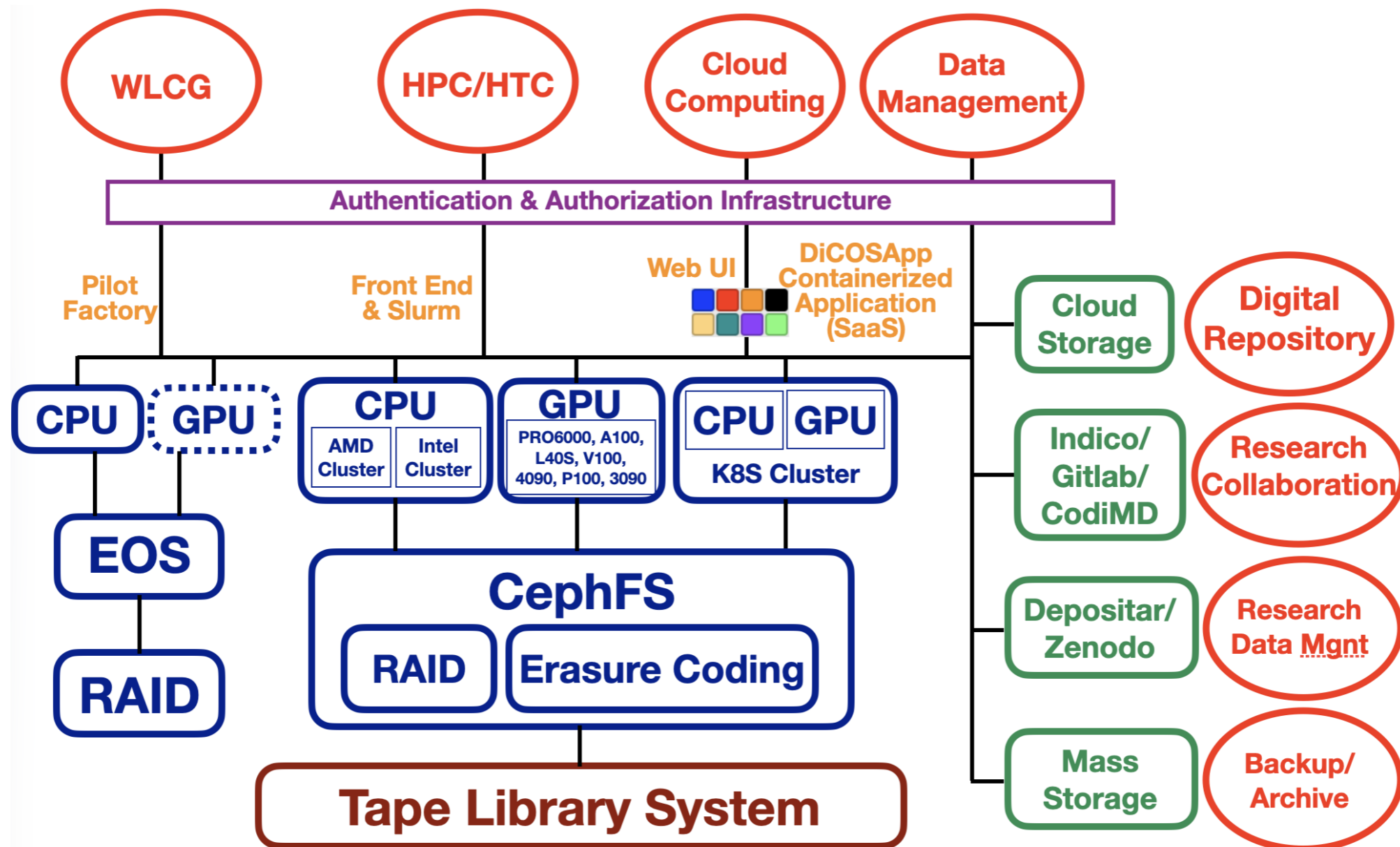
- ASGC has been the Core facility for scientific computing and big data analysis, funded by AS and NSTC from 2023
- Supporting particle physics communities together with TIDC, serving as the computing arms
- Physics and Life Sciences are the primary user communities of ASGC
  - MHEP keeps driving the advancement of scientific computing system and technology
  - ASGC is extending services beyond the domain boundaries, based on the collaborations with broader user communities
- 645 Users in 203 PI Groups, from 30 Institutes (Dec. 2025)



- GPU-Day growth in 2025: 30%
- CPUCore-Day growth in 2025: 200%
- Average annual finished jobs: 2,439K

# ASGC Service Architecture - Scalable Science Cloud

- Integrating data, algorithm and computing
- Scientific workflow implementation with integrated compute, storage and network resources
- Persistent services: Analysis facility, Web portals
- New resource and new services online yearly, based on users requirements



# Computing & Storage Resources

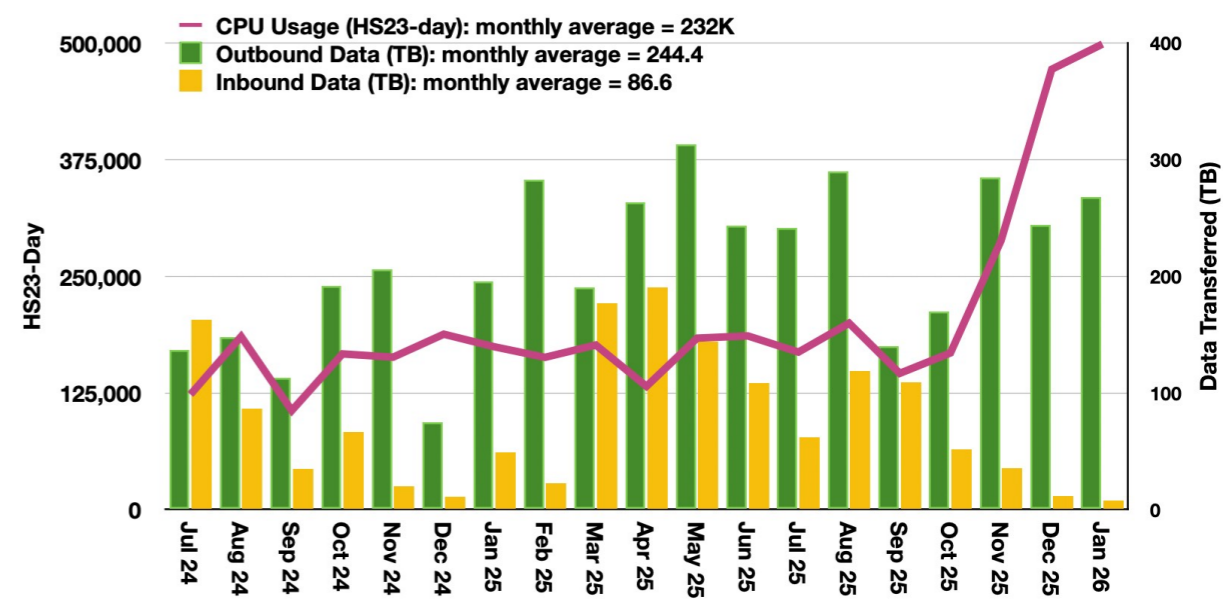
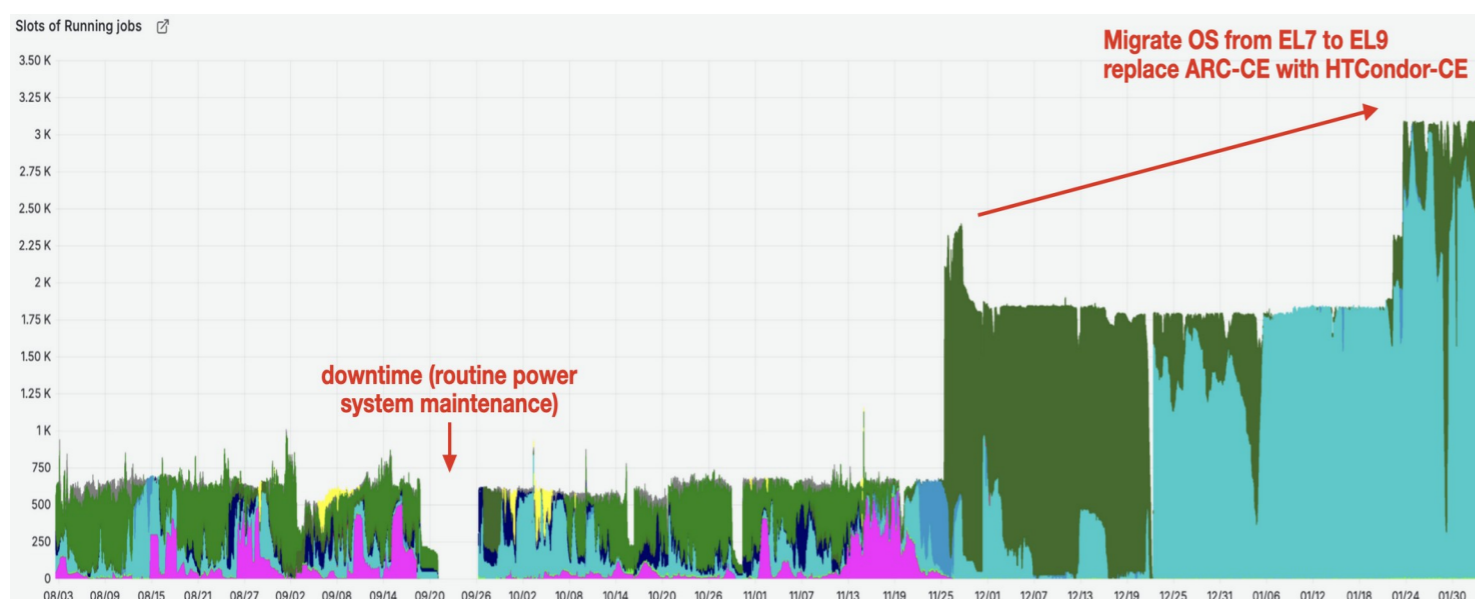
CPU Cluster	Spec	CPU Cores	GB RAM/ Core	Nodes	Performance (TFLOP/Core)	Network	Onboard Date
EDR2	AMD Turin 9645	1,920	8	10		100Gb/HDR	Mar 2026
EDR1	AMD Genoa 9654	3,840	8	20	1.1	100Gb	Mar 2024, Dec 2024
Intel-G4	Intel Xeon 6448H	512	8	4	1	100Gb	Sep 2024

GPU Cluster	Spec	GPU	Architecture	Nodes	CPU Cores	Server RAM (GB)	CPU	Network	Onboard Date
B6000	RTX Pro 6000 96GB	16	Blackwell	2	32	1,536	Intel 6517Px2	200Gb	Mar 2026
L40S	L40S 48GB	4	Ada Lovelace	1	32	768	AMD 9374x1	10Gb	Dec 2024
4090	RTX 4090 24GB	16	Ada Lovelace	2	40	1536	Intel 4416x2	10Gb	Jul 2024
3090	RTX 3090 24GB	56	Ampere	7	32	1024	AMD 7302x2	100Gb	Dec 2020
A100	A100 NVL 80GB	24	Ampere	3	64	512	AMD 7542x2	100Gb, NVL	Jun 2020

Storage	Capacity (TB)	System	Price	Remarks
Ceph File System	11,000	Online storage managed by Ceph distributed file system, with (8, 3) erasure coding	1,000 per TB-year	+3PB by end of Mar 2026
Tape Lib System	10,000	LTO9 (18TB/Tape)	300 per TB-year	Service from 2025

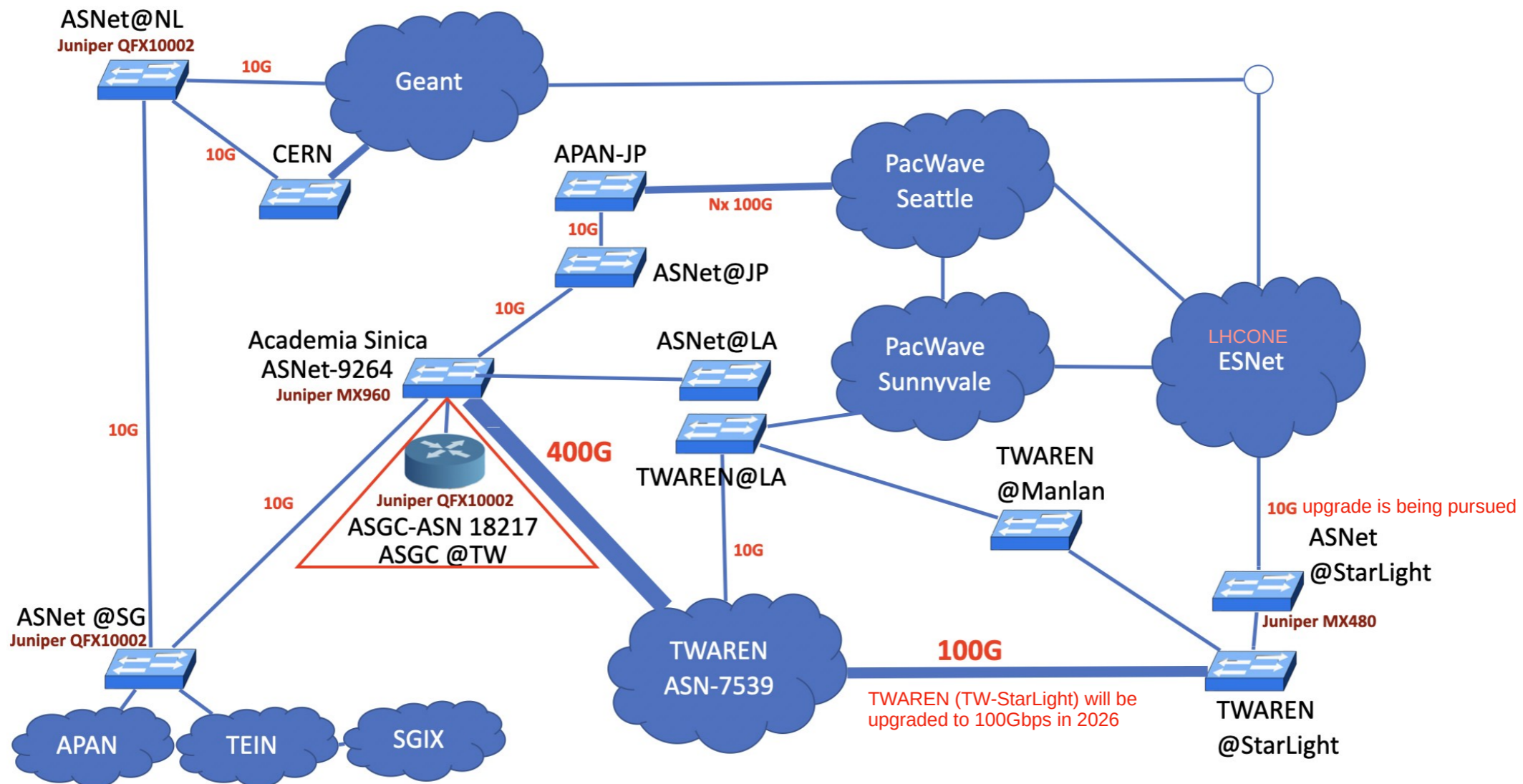
# WLCG Tier-2 Site for ATLAS

- Site has been in production since July 2024 in association with USATLAS
  - Connection with LHCONE upgraded from 3Gbps to 5Gbps on 1 May 2025
  - Local storage RSE reconfigured and CRIC updated accordingly
- Stably supporting MC Simulation jobs and analysis jobs mainly
- Current Status (Feb 2025 - Jan 2026):
  - 292.5K CPUCore-Days (2.8M HS23-day), 2.5PB storage (EOS, growing from 0.93PB in Aug. 2024)
  - Data Transmission (inbound + Outbound): 1.1 + 2.9 = 4 PB with efficiency > 95%
- CPU Efficiency: > 88%
  - 77% MC Simulation & Event Generation
  - 16% User Analysis
  - 6% Group Production & Analysis
- Pledge in 2025 & 2026
  - CPU: 30K HEPscore23
  - Disk: 5 PB, managed by EOS
- Job slots - 3,200
  - ~ 3,000 job slots are available now
- Next Step:
  - Making better use of ASGC resource, e.g, pre-stage for analysis jobs
  - Efficiency improvement



# Networking

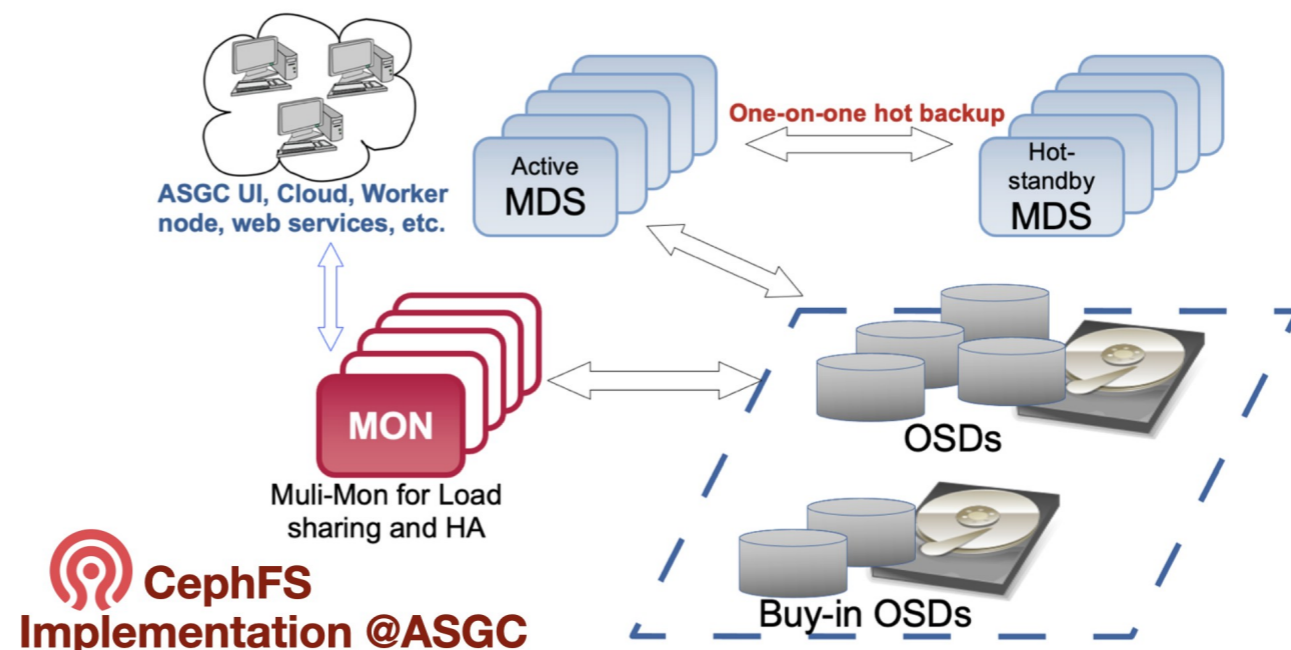
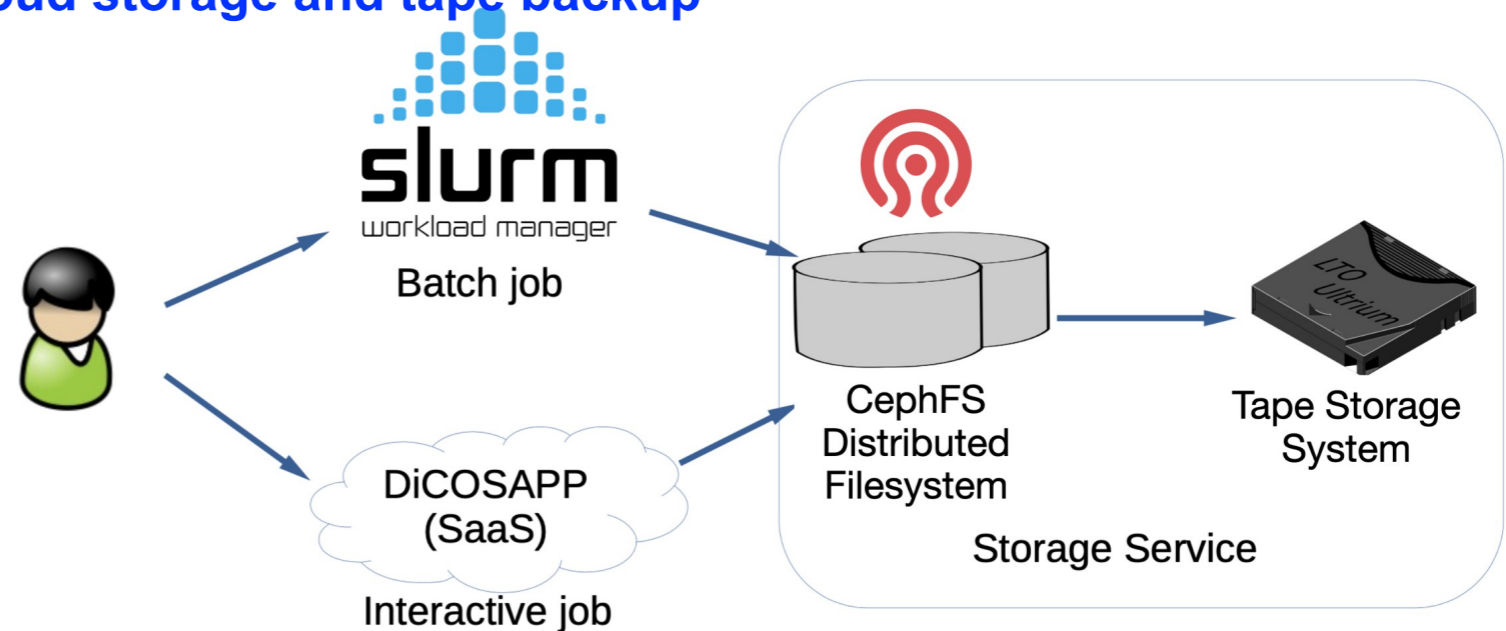
- **Status:**
  - **Connecting with LHCONE through ESnet**
    - 3Gb bandwidth upgraded to 5 Gb from May 2025
  - **Internal : 10 Gbps**
    - **backbane: 100 Gbps**
    - **to router : 10 Gbps**
- **TWAREN connecting with StarLight will be upgraded to 100Gb in March 2026**
  - **Bandwidth between TW-FTT and LHCONE through ESNet could be upgraded afterwards**
- **Asymmetric routing issue**
  - **No jumbo frame and IPv6 support for traffic coming into ASGC through general Internet routes**



TWAREN : TaiWan Advanced Research and Education Network

# Data & Storage Services - CephFS

- Ceph distributed storage system - O(10)PB scalable data pool
- High performance - 4.2GB/s throughput, 1.5K op/s rd, 2K op/s wr
- Reliability - Erasure coding (8, 3), with 462 OSDs, 51 hosts and multiple MDS & MONs
- Scalability - from O(10)TB in 2012 to 15PB in 2025
- Data transmission and local backup supported
- New Services in 2026: data transfer, cloud storage and tape backup

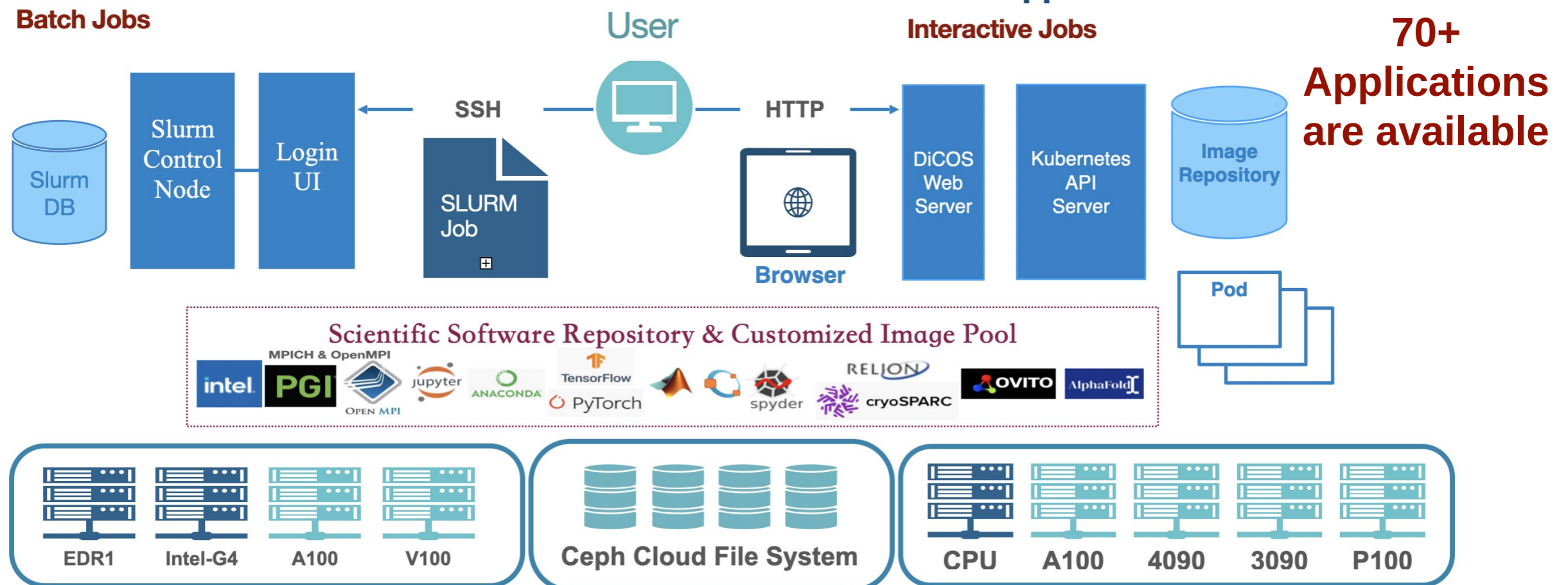


# Computing Services: Optimization of Application and System Efficiency

- HTC/HPC & Cloud Computing Platform
- Integration of computing model and software
- Research data workflow and management
- Information security
- Application efficiency
- Customized IaaS, PaaS, and SaaS

## HPC platform Slurm System Architecture Batch Jobs

## Software-as-a-Service Web-based Cloud Platform DiCOSApp SaaS Architecture Interactive Jobs



# 73+ Web Applications Provided (Software-as-a-Service)

<b>CryoSPARC 32 P100</b> Version: 3.2 Resources: 83%	<b>CryoSPARC 1080ti</b> Version: 3.3.2 Resources: 29%	<b>CryoSPARC RTX3090</b> Version: 3.3.2 Resources: 26%	<b>spyder cpu/eman2</b> Version: Resources: 17%	<b>Octave</b> Version: V5.2 Resources: 29%	<b>Transfer Data</b> Version: Resources: 86%
<b>CryoSPARC RTX3090</b> Version: 4.0.2 Resources: 26%	<b>CryoSPARC RTX3090</b> Version: 4.1.1 Resources: 26%	<b>CryoSPARC RTX3090</b> Version: 4.4.1 Resources: 26%	<b>cisTEM</b> Version: Resources: 86%	<b>Ovito</b> Version: Resources: 86%	<b>OpenACC</b> Version: GPU P100 Resources: 83%
<b>AlphaFold</b> Version: GPU with A100 Resources: 75%	<b>AlphaFold (Full DB)</b> Version: GPU with A100 Resources: 75%	<b>IMOD</b> Version: GPU with 1080ti Resources: 29%	<b>Triton</b> Version: 22.01-py3 (GPU P100) Resources: 83%	<b>AlphaFold</b> Version: GPU with RTX3090 Resources: 26%	<b>AlphaFold</b> Version: GPU with V100 Resources: 17%
<b>RoseTTAFold</b> Version: GPU with rtx3090 Resources: 26%	<b>Dynamo</b> Version: GPU with 1080ti Resources: 29%	<b>MATLAB</b> Version: R2018b on GPU V00 Resources: 17%	<b>Jupyter Lab</b> Version: CPU with Tensorflow v1 Resources: 86%	<b>Jupyter Lab gpu 3090</b> Version: GPU with Tensorflow 3090 Resources: 26%	<b>Jupyter Lab GPU 1080ti</b> Version: GPU with Tensorflow v2 Resources: 29%
<b>RFDIFFUSION</b> Version: 2023 on GPU V00 Resources: 17%	<b>diffdock</b> Version: 2023 on GPU V00 Resources: 17%	<b>EvoDiff</b> Version: V100 Resources: 17%	<b>Jupyter Lab GPU V100</b> Version: GPU with Tensorflow V100 Resources: 17%	<b>Jupyter Lab GPU A100</b> Version: GPU with Tensorflow A100 Resources: 75%	<b>Jupyter Lab Cryocare GPU</b> Version: GPU with 1080ti Resources: 29%
<b>QIIME2</b> Version: Genome Resources: %	<b>Scipion3</b> Version: P100 Resources: 83%	<b>Phenix</b> Version: Resources: 86%	<b>Jupyter Lab GPU A100</b> Version: GPU with Tensorflow v2.6 Resources: 75%	<ul style="list-style-type: none"> <li>• Web Portal</li> <li>• Application over Cloud</li> </ul>	
<b>MorphoGraphX</b> Version: GPU with P100 Resources: 83%	<b>Deepmd-kit</b> Version: GPU with A100 Resources: 75%	<b>Deepmd-kit</b> Version: GPU with V100 Resources: 17%	<b>MAML</b> Version: GPU with A100 Resources: 75%	<ul style="list-style-type: none"> <li>• Jupyterlab</li> <li>• Web Terminal</li> </ul>	
<b>Warp</b> Version: rtx4090 Resources: 700%	<b>MAML</b> Version: GPU with V100 Resources: 17%	<b>PVserver</b> Version: 5.8.0 (GPU 1080T) Resources: 29%	<b>Paraview Client</b> Version: 5.8.0 Resources: 86%	<b>LabVIEW Run-Time Engine</b> Version: 2019	
<b>R studio</b> Version: 10 CPU Cores Resources: 54%	<b>PyRoot</b> Version: GPU with 1080ti Resources: 29%	<b>qiskit</b> Version: Resources: 86%	<b>NVIDIA CUDA-Quantum A100</b> Version: Resources: 100%		

LabVIEW Run-Time Engine  
Version: 2019  
Launch



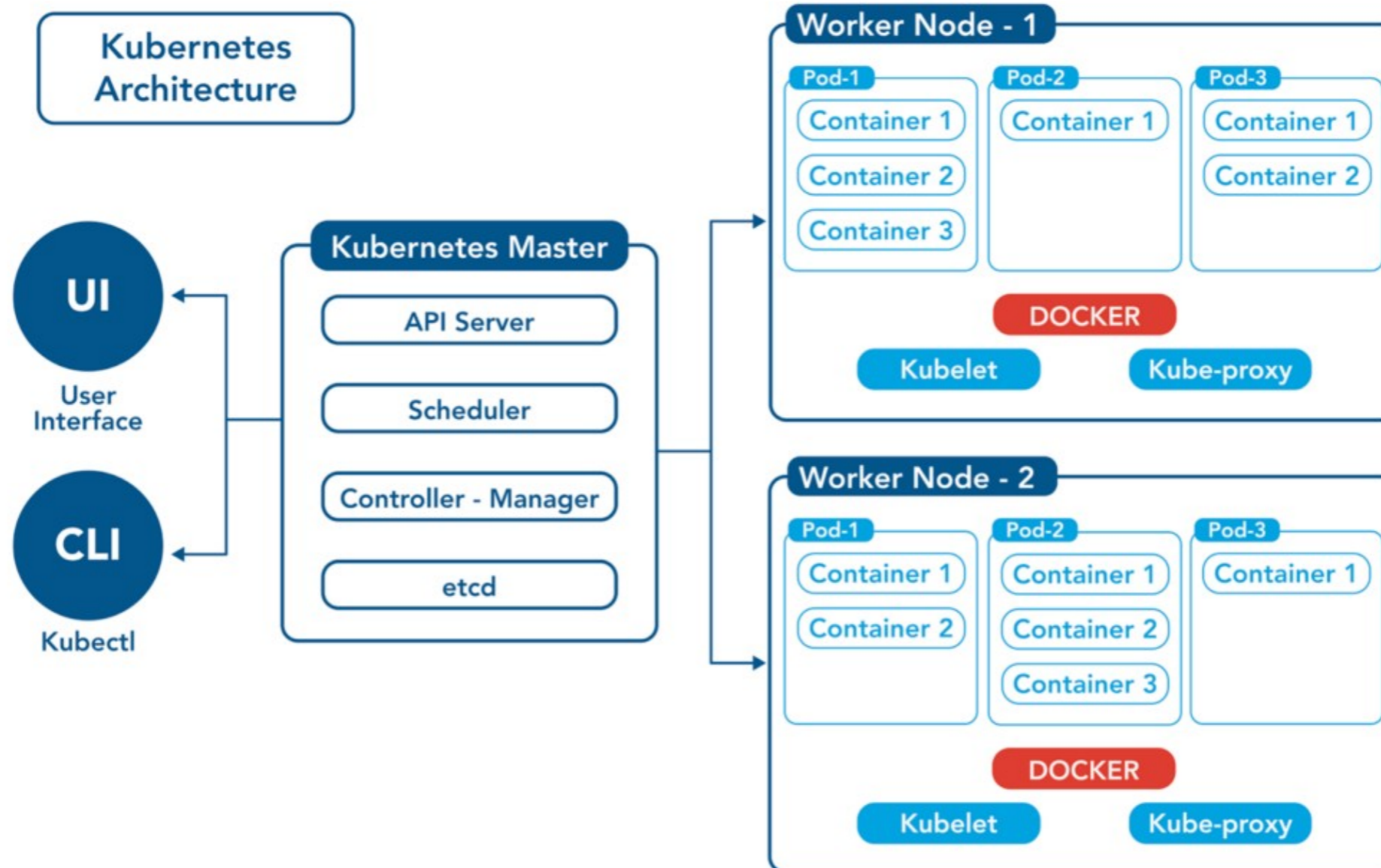
## DiCOS-BioSAXS Platform

About Job Submission Job Monitoring Data Management

- ATLAS
- AMBER
- Rosetta
- DAMMIN
- DAMMIF
- GASBOR

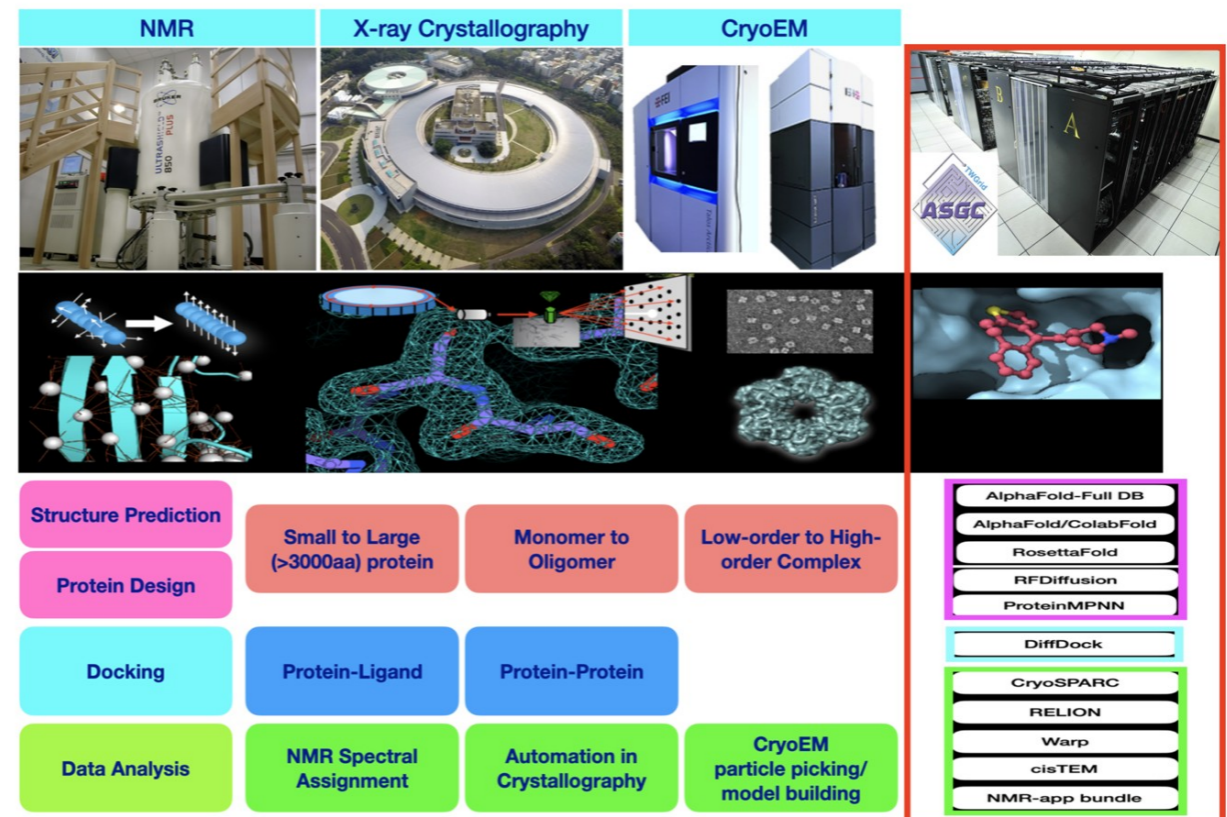
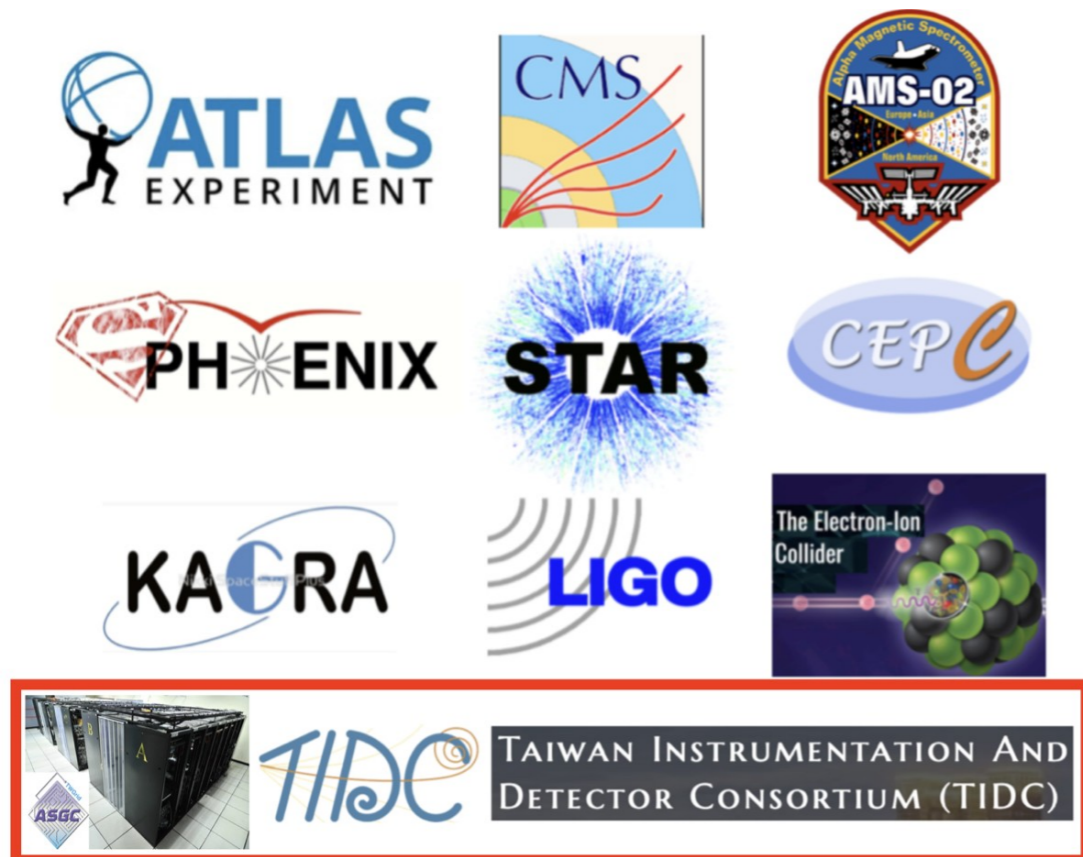
# Resilient Cloud Research Platform

- Faster Research: HPC/HTC using CPU & GPU cluster
- User application and workflow Orchestration
- High Availability & Elastic Scaling
- Resource Efficiency
- Hybrid & distributed infrastructure
- Software-as-a-Service (SaaS)
- Batch, Interactive (Jupyter) and Web



# Flexible Collaboration Model: User-Driven, Service Oriented

- Pay-as-you-go
- Resource Integration: Collocation & Buy-In
- Service Hosting
- Reservation
- Experiment & Collaboration support
- Advanced Service: Web Application and portal, efficiency tuning, workflow optimization, customization, hands-on training
- User Engagement : weekly User Meeting, User Committee, dissemination & outreach



# Customized Services for Structural Biology Community

- Serving as the computing arm of AS CryoEM Core Facility (ASCEM) from 2018: computing model implementation, efficiency tuning, capacity building
- ColabFold for protein structure predictions in batches
- Molecular Docking: DiffDock
  - Able to finish 140 DiffDock predictions in < 60min, using DiCOS@RTX3090
- Protein Design Challenges: AlphaFold, RosettaFold, RFDiffusion
- DiCOS-BioSAWN Platform supports NSRRC BioSAXS beam line data analysis using ATSAS/Rosetta/Amber/ Gromacs-SWAXS
- The primary GPU computing users in ASGC: 55 PI Groups (32 in AS), 150 users

The image displays a grid of software service tiles, a workflow diagram, and a platform interface. The tiles include:

- CryoSPARC RTX3090**: Version: 4.7.0, Resources: 50%. Launch button.
- RELION 5 beta RTX3090**: Version: V5 beta, Resources: 50%. Launch button.
- Warp**: Version: rtx4090, Resources: 100%. Launch button.
- ColabFold**: Version: GPU with rtx3090, Resources: 36%. Open and Delete buttons. End Time: Jan. 28, 2026, 12:04 p.m.
- NAMD3**: Version: rtx4090, Resources: 100%. Launch button.
- Gromacs**: Version: GPU RTX4090, Resources: 100%. Launch button.
- MZMINE**: Version: CPU, Resources: 54%. Launch button.
- AlphaFold**: Version: GPU with RTX3090, Resources: 47%. Launch button.
- RFDiffusion**: Version: 2023 on GPU RTX4090, Resources: 18%. Open and Delete buttons. End Time: Jan. 30, 2026, 10:20 a.m.

The workflow diagram shows the process from **Cryo-EM Facility** to **Data Store**, then to **Preprocessing (On-the fly processing, Motion Correction)**, and finally to **ASGC**. The ASGC section is highlighted with a dashed box and includes:

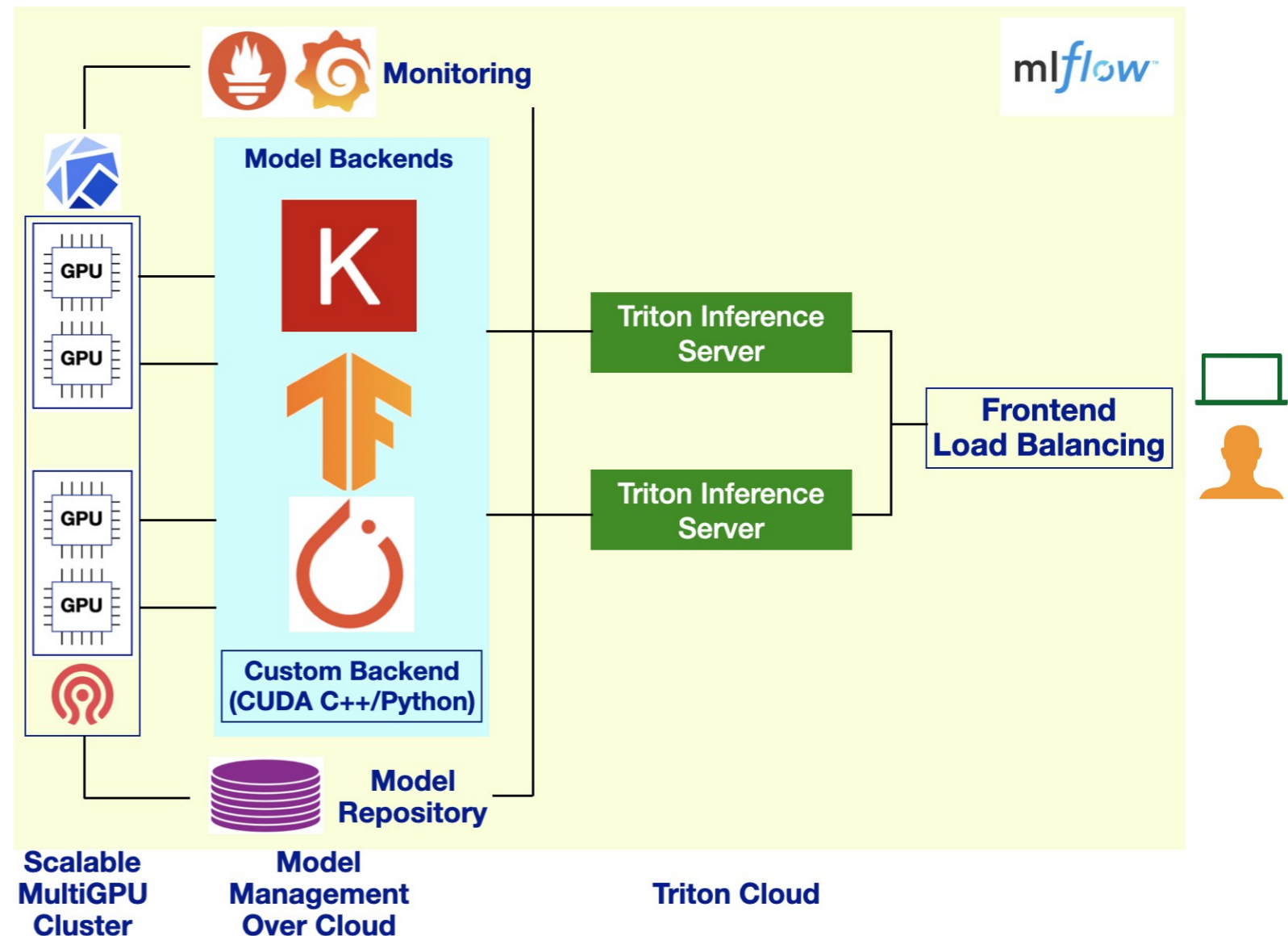
- Scientific Software Platform with High Performance Computing Resources
- Data Store

The **DiCOS-BioSAWN Platform** interface is also shown, featuring a navigation menu with options like **Job Submission**, **Job Monitoring**, and **Data Management**, and a dropdown menu listing **ATSAS**, **AMBER**, **Rosetta**, **DAMMIN**, **DAMMIF**, and **GASBOR**.

AI-enabled Protein-related Tools

# Services of ML-enabled data analytics

- **ML/AI application platform service - SW library, HW, integration and application development management (e.g, MLflow, Triton)**
  - Build up customized ML platforms for user specified projects - Deploy ML packages ready environment in order to help ML development smoothly and provide on-demand computing power
  - Upkeep of the application framework
  - Workflow and data pipeline integration: tracking, model serving, packaging, model registry, evaluation & tracing
  - Efficiency Improvement
  - Inference-as-a-Service (incl. Model repository): e.g., Triton-based services
- **Potential use cases**
  - Users who bring existing source code - ASGC could help to setup a virtual environment and confirm source code running normally
  - Share and reuse of trained/verified models
  - MLOps over the cloud
- **Approaches**
  - Supporting Kubernetes/Jupyter lab for development purpose
    - Create Kubernetes/Jupyter lab environment with user specified ML packages ready.
    - Support on-demand scalable CPU/GPU computing power.
  - Supporting containerized environment (e.g, Docker image) for deployment purpose
    - Create takeout images in Docker format as an option for user who wants to train/predict model
    - Docker images could be downloaded from ASGC server and deployed on users' Docker Desktop on Windows/Linux.



# Deep-learning-based Speech Enhancement (with Its Application to Assistive Oral Communication Technologies)

**You-Jin Li, Rong Chao, PI: Yu Tsao**

**Research Center for Information Technology Innovation, Academia Sinica**

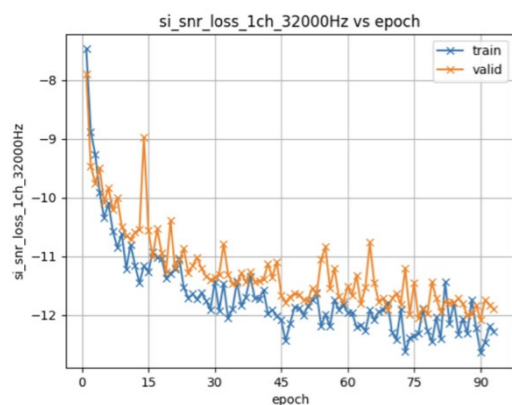
## Universal Speech Enhancement

MODEL 2: SFI\_SEMAMBA (10 Layers)

- devices: **NVIDIA A100 GPU x 1**
- batch size: **2**

**Epoch 251:**

PESQ ↑ : 2.5904  
 ESTOI ↑ : 0.8374  
 SDR ↑ : 14.4522  
 LSD ↓ : 2.5686  
 MCD ↓ : 3.6416  
 DNSMOS ↑ : 3.0515  
 NISQA\_MOS ↑ : 3.3971  
 SpeakerSimilarity ↑ : 0.7261

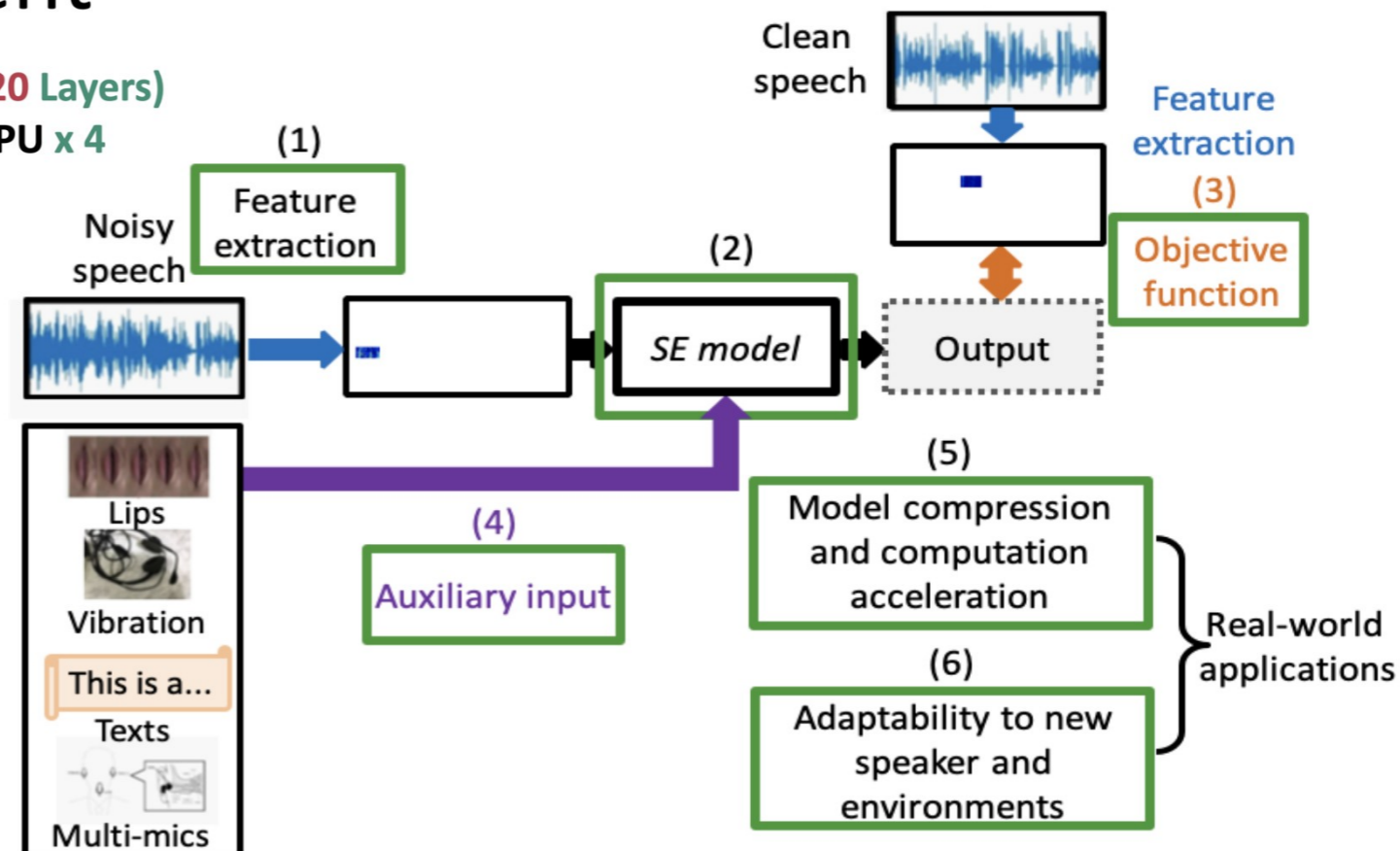
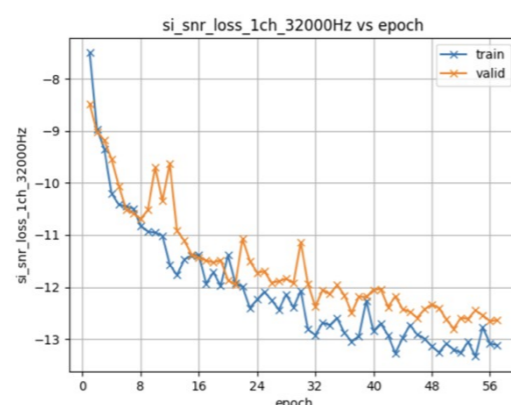


MODEL 4: SFI\_SEMAMBA (20 Layers)

- devices: **NVIDIA A100 GPU x 4**
- batch size: **4**

**Epoch 168:**

PESQ ↑ : 2.8140  
 ESTOI ↑ : 0.8667  
 SDR ↑ : 15.9479  
 LSD ↓ : 2.5437  
 MCD ↓ : 3.2805  
 DNSMOS\_OVRL ↑ : 3.0645  
 NISQA\_MOS ↑ : 3.6297  
 SpeakerSimilarity ↑ : 0.7698



# Potential Generative AI Applications for AMS/HEP

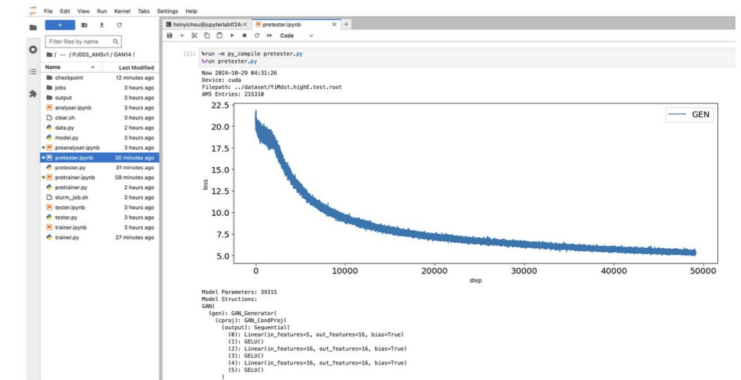
## Computing Resource

[Remote Terminal]  
 slurm-ui-asiop.twgrid.org  
 slurm-ui03.twgrid.org

```

1 import torch
2 import torch.nn as nn
3 import torch.nn.functional as F
4
5 device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
6
7 class PixelShuffleID(torch.nn.Module):
8     def __init__(self, upscale_factor):
9         super(PixelShuffleID, self).__init__()
10        self.upscale_factor = upscale_factor
11
12    def forward(self, x):
13        batch_size, channels, length = x.size()
14
15        # Check if channels can be evenly divided by upscale_factor
16        if channels % self.upscale_factor != 0:
17            raise ValueError(f"Channels must be divisible by (self.upscale_factor)")
18
19        # Perform pixel shuffle
20        channels //= self.upscale_factor
21        x = x.view(batch_size, channels, self.upscale_factor, length)
22        x = x.permute(0, 1, 3, 2) # Rearrange dimensions
23        x = x.contiguous().view(batch_size, channels, length * self.upscale_factor)
24        return x
25
26 class PixelUnshuffleID(torch.nn.Module):
27     def __init__(self, downscale_factor):
28         super(PixelUnshuffleID, self).__init__()
29         self.downscale_factor = downscale_factor
30
31     def forward(self, x):
32        batch_size, channels, length = x.size()
33
34        # Check if the length is divisible by downscale_factor
35        if length % self.downscale_factor != 0:
36            raise ValueError(f"Length must be divisible by (self.downscale_factor)")
37
38        # Perform pixel unshuffle
39        out_channels = channels * self.downscale_factor
40        out_length = length // self.downscale_factor
41
42        # Reshape and permute the tensor
43        x = x.view(batch_size, channels, out_length, self.downscale_factor)
  
```

[DiCOS Apps]  
 Jupyter Lab GPU

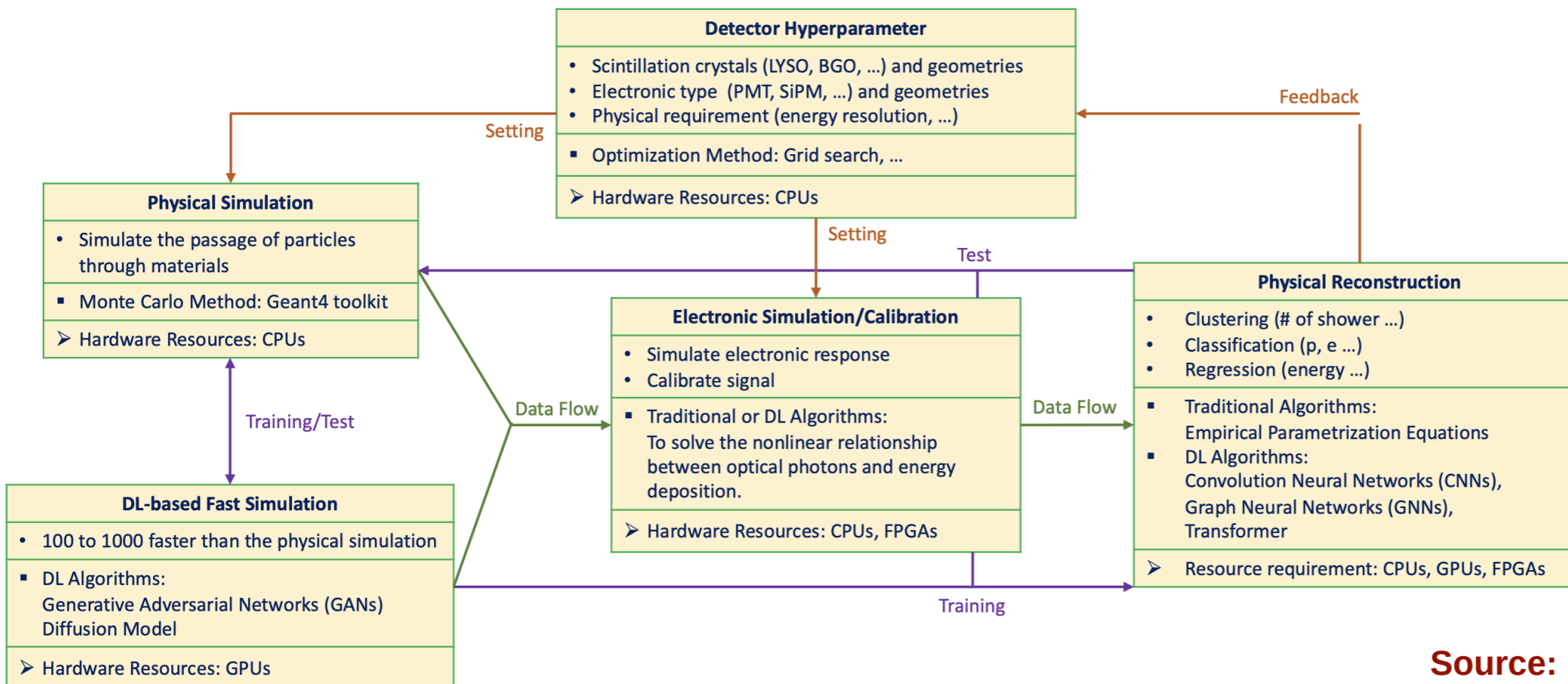


[Slurm]  
 CPUs, GPUs A100, V100, ...

```

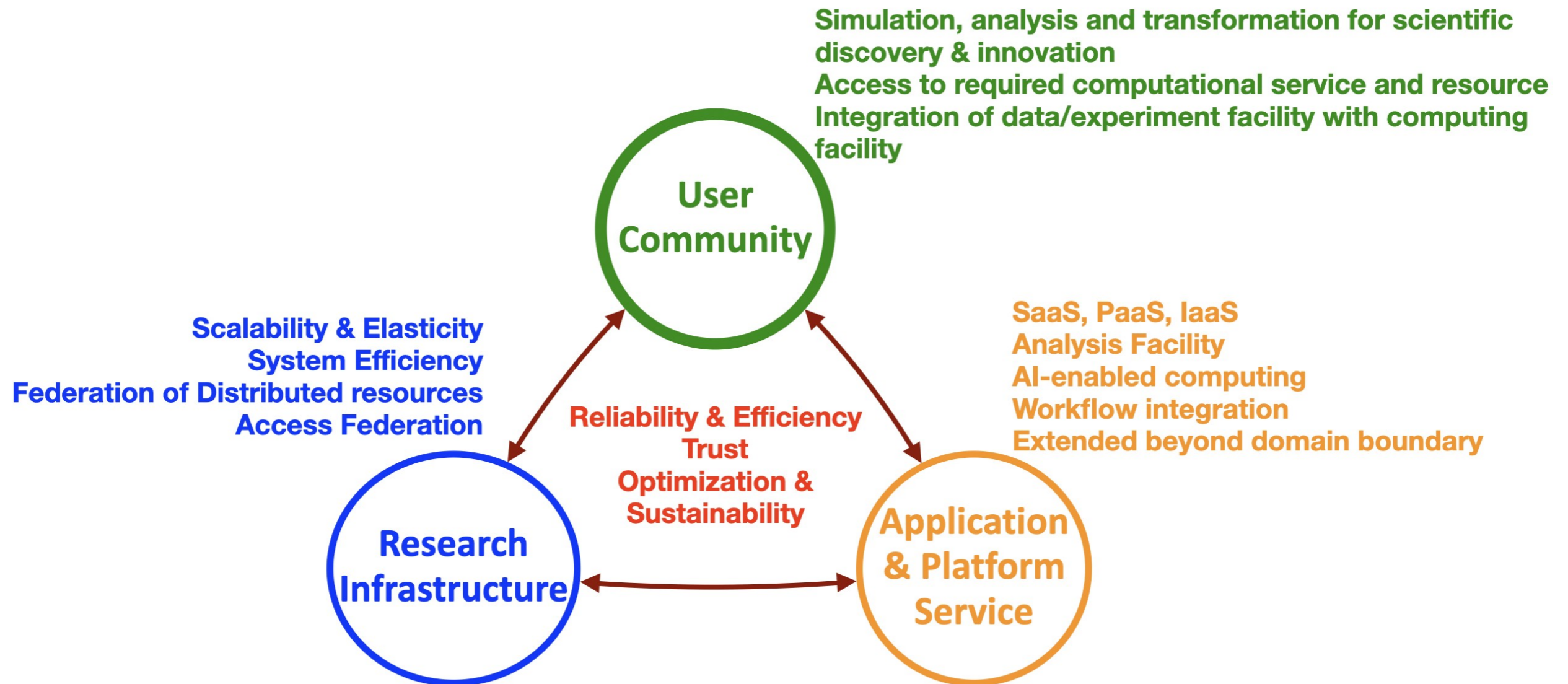
(ml) [hsinyichou@slurm-ui-asiop GAN14]$ squeue | grep a100
230030      a100      HMC0034      twchiu PD      0:00      1 (Nodes required for
231049      a100      HMC0036      twchiu PD      0:00      1 (Priority)
232091      a100      HMC0035      twchiu PD      0:00      1 (Priority)
232098      a100      HMC0033      twchiu PD      0:00      1 (Priority)
232667      a100      handbook     noke0081 R       18:00     1 hp-teslaa03
232635      a100      handbook     noke0081 R       1:01:53   1 hp-teslaa03
194890      a100      long urg_seja rc PD      0:00      1 (Dependency)
194889      a100      long urg_seja rc PD      0:00      1 (Dependency)
194888      a100      long urg_seja rc PD      0:00      1 (Dependency)
194887      a100      long urg_seja rc PD      0:00      1 (Dependency)
194886      a100      long urg_seja rc PD      0:00      1 (Dependency)
176131      a100      long urg_seja rc PD      0:00      1 (Dependency)
176130      a100      long urg_seja rc PD      0:00      1 (Dependency)
176129      a100      long urg_seja rc PD      0:00      1 (Dependency)
176128      a100      long urg_seja rc PD      0:00      1 (Dependency)
  
```

## Calorimeter Design Tool in High Energy Physics



Source: Dr. HY Chou, IoP, AS

# Capacity Building



# Enhancing User Community Engagement

- Training, dissemination and outreach
- Improving communication and response
- Supporting user access and application efficiency
- Understanding user requirements and experiences
- Facilitating capacity building of users and NSTCCore
- Training and workshop/conference: 250+ participants in 2025
- Weekly user meeting: 2:30pm, Wednesday

## Basic

- 30min – full day
- Get access, details of using compute, data and SaaS services, w/o hands-on, User Forum

## Thematic

- 1 hrs – full day
- Discipline-based Data analysis
- ICT topics: GPU Computing, AI, CPU Trends

## Domain-Specific Workshop

- 30min – 3hrs
- Supporting hands-on, w/o ASGC service introduction

## Conference

- e-Science, community-specific workshop, core technology
- ISGC (International Symposium on Grids & Clouds, since 2002)
- HEPiX, CHEP (Computing in High Energy Physics)
- APAN (Asia Pacific Advanced Network)



# Collaboration with NCHC

- **NSTCCore (operating by ASGC) is part of the national scientific computing infrastructure**
- **Collaboration Model**
  - **Workload distribution Principle: large-scale vs high-throughput**
  - **Quarterly meeting for milestones**
- **Starting from**
  - **Federated IdM - easy access to both sites using the same account (better to be user's home institute SSO account)**
  - **Home directory and data consistency**
  - **Faster access to required services and finish processing efficiently**
- **Challenge**
  - **Security governance - requirements and policy**
  - **Unique UID for each user**
  - **Data & storage consistency**
  - **Compatible Software and environment**
  - **Services**

# Summary

- **Strengthening the Scientific Computing Ecosystem**
  - Supporting Simulation, big data analysis and AI applications at the same time
  - User Community x Resource x Services (incl. technical R&D team)
  - Capacity building: making better use of ASGC resources, and being able to provide advanced services
- **Enforcing AI tools and AI-enabled computing**
  - Innovative implementations of computing models are expected
- **Open data and open science**
- **Efficiency - optimization for application, system and energy**
  - Intelligent monitoring and control
- **Networking: driving the infrastructure upgrade by sciences**
  - Supporting larger data throughput and resource federation for multiple local/regional/international collaborations
  - Long-distance data efficiency
  - Security