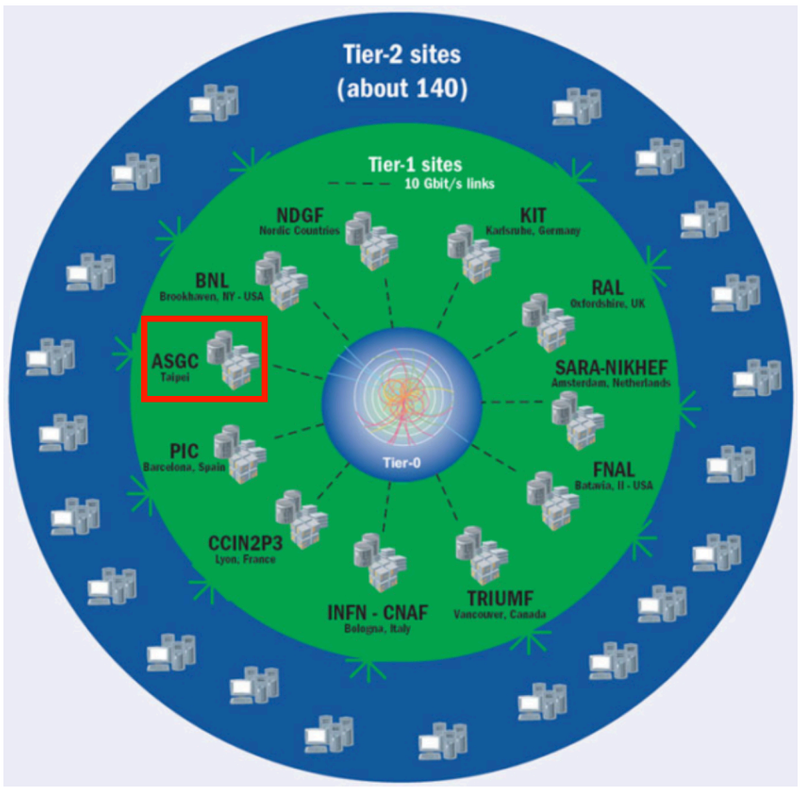
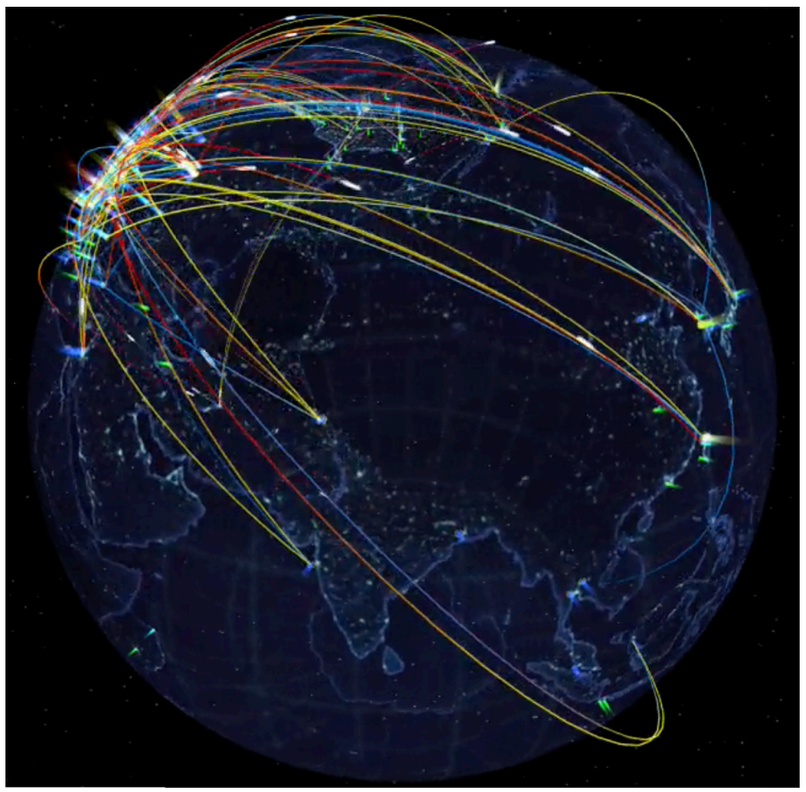
TW Report @HEPiX

ASGC Introduction

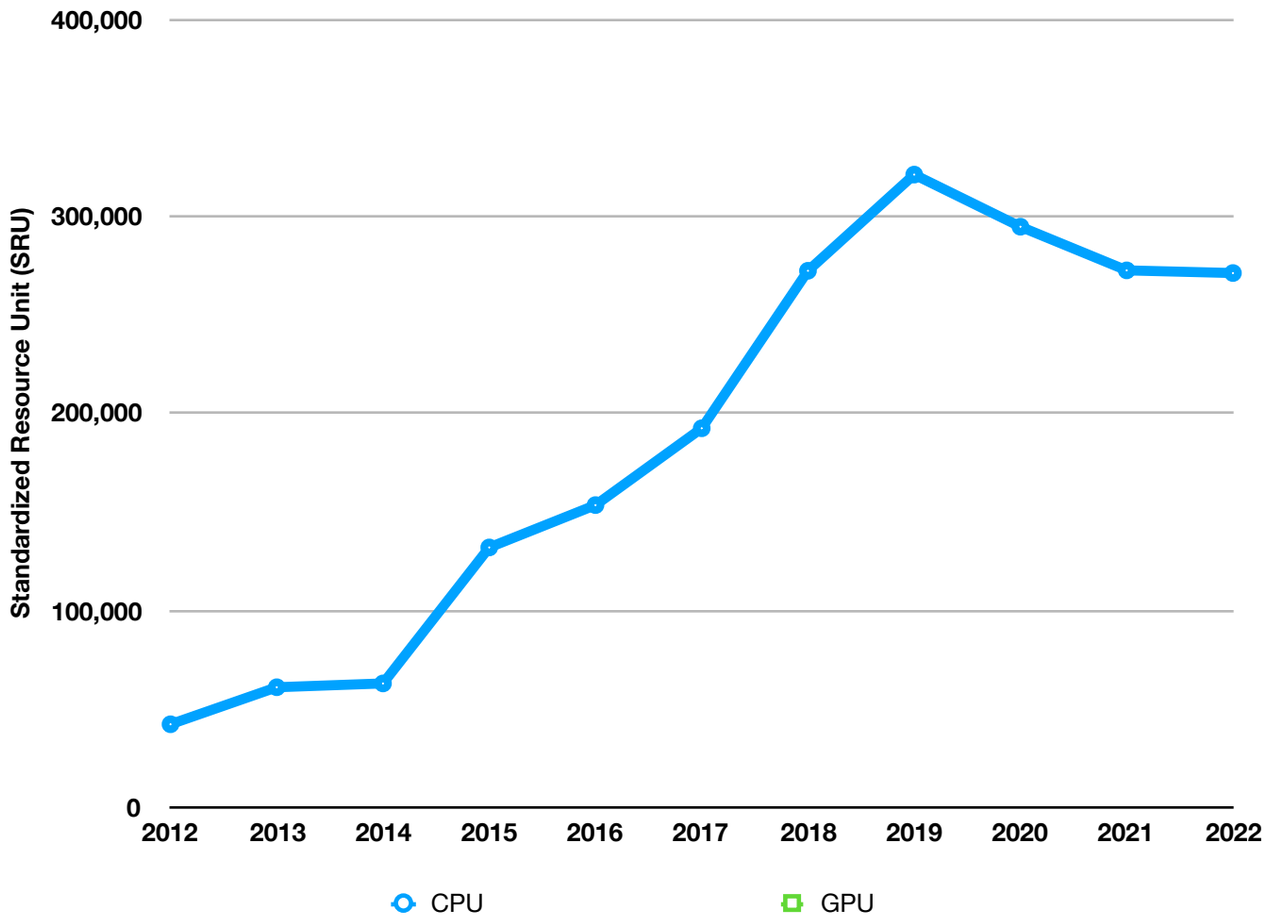
- ASGC joined WLCG development and deployment for the Large Hadron Collider grand challenges since 2001
  - ASGC T1 and WLCG Asian Regional Operation Centre has been operational from 2005
  - Migrating to T2s for ATLAS and CMS (effective from Oct. 2023)
- ASGC has been supporting multi-disciplinary e-Science applications of Academia Sinica from 2006, based on WLCG core technologies
  - The research infrastructure, platform and services are improved progressively along with growing scientific applications of various disciplines
- System efficiency optimization (including power, thermal, system and applications, etc.) is also a strategic goal of ASGC aided by machine learning technologies
- ASGC becomes the Core Facility for big data and scientific computing of AS from 2023





Scientific Collaborations and Resource Status

- Supporting e-Science and big data analysis based on WLCG core technologies in Taiwan and Asia
- Research collaborations since 2006: Workflow customization and Efficiency optimization
  - 50+ research groups, 150+ users
  - Flexible collaboration model
- Asia regional collaborations on e-Science, especially for hazard risk analysis have been conducted as part of a series of EU-funded/EGI-lead project from 2008
- Resource Usage
  - Availability: 99%
  - Growth of CPU utilization (2012 - 2022): 20 % CAGR
  - Growth of GPU utilization (2017 - 2022): 170% CAGR
- Data centre operation in 24x7 since 2001
  - 2MW, 400 tons AHUs, 112 racks in 800 m^2
  - 10,976 CPU Cores, 204 GPU (36x A100), 30PB disk storage (Oct. 2022)



SRU is normalized computing resource unit based on CPU or GPU performance in Linpack.

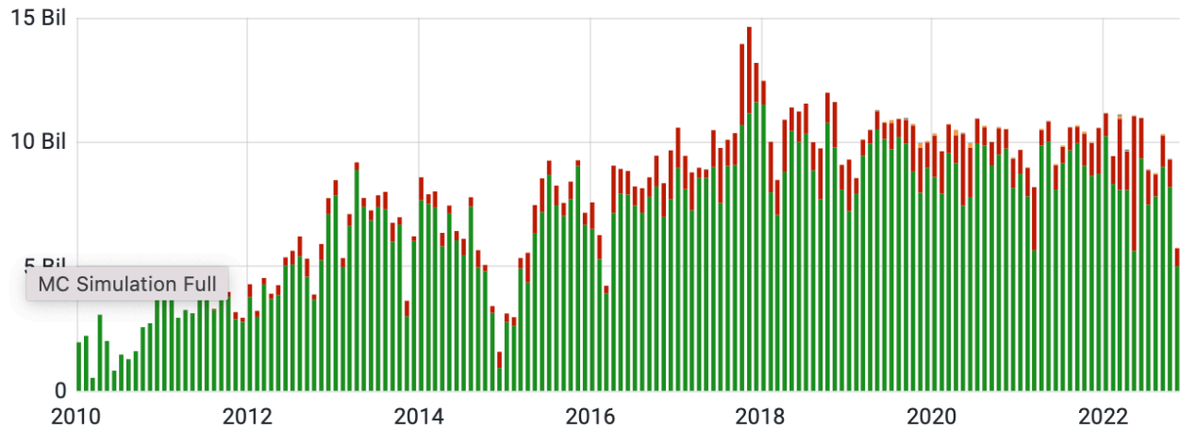
Scientific Collaborations

- Particle physics: ATLAS, CMS, AMS, KAGRA/LIGO/IGWN, ICECube, EIC
- Physics: Surface, Material, Quantum Field, Biophysics
- Astronomy and Astrophysics
- Structural Biology, Drug Discovery, NGS, Bioinformatics, CryEM
- Ecology and Biodiversity informatics
- Computational Chemistry, Biophysical Chemistry, Chemoinfomatics
- Seismology and earth science
- Environmental changes and hazard mitigation
- ML-enabled data analysis
- Research infrastructure and e-Science: including open data, research data management

WLCG Activities @ASGC

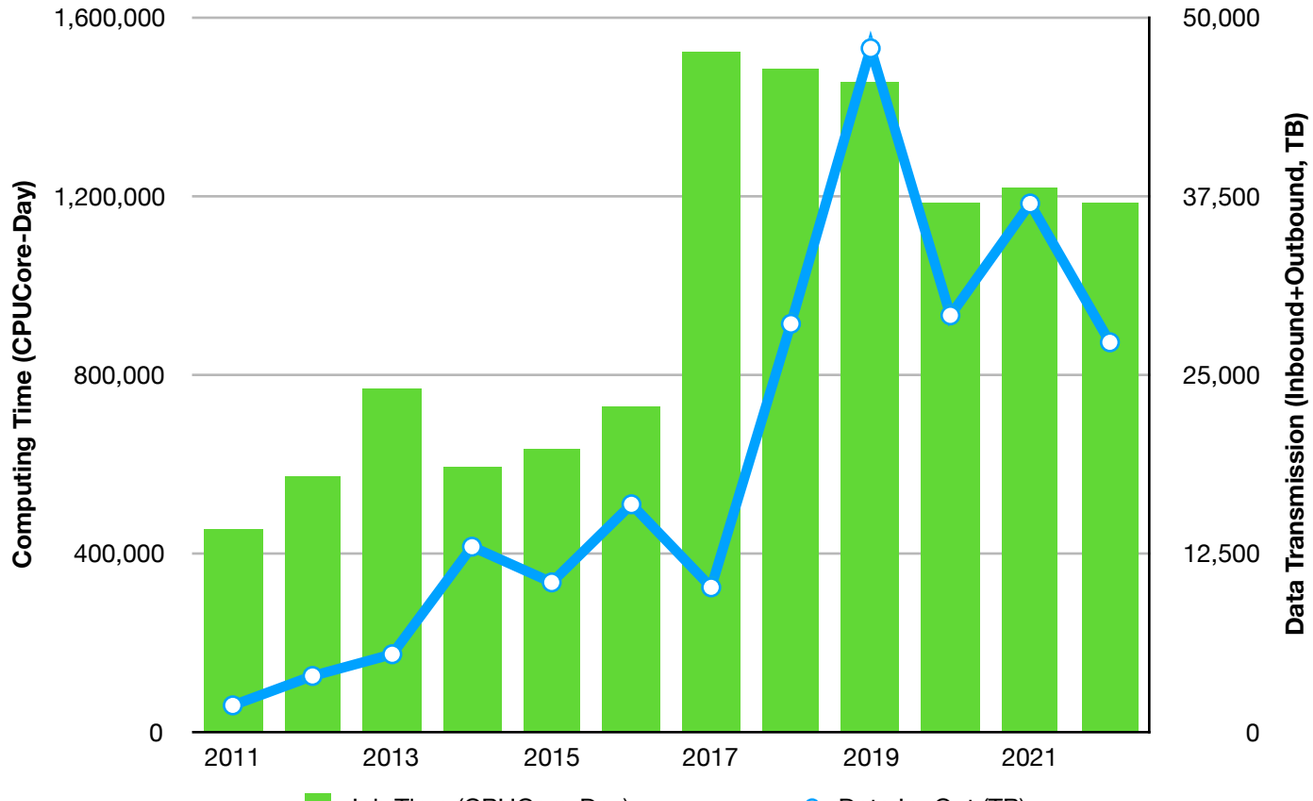
- ATLAS in Taiwan
  - Achievements: Higgs boson; Dark matter; Searches for beyond Standard Model
  - Future plan: H—>bb; Di-Higgs; Dark matter
- Computing Resource - retirement of legacy hardware for energy saving
  - ASGC Tier-1 (2023): 58,760 HEPSpec06 (3,200 CPUCores)
  - ATLAS Federated Taiwan Tier-2 (2023): 10,896 HEPSpec06 (1,536 CPUCores)
  - CMS T3: 20,000+ HEPSpec06 (768 CPUCores), 1.7 PB EOS storage
  - GPU would be available after validation of new computing models (ATLAS, CMS)
- Storage Resource (2023) of ASGC T1 and FTT T2: 9.6PB + 1.1 PB
  - Migration from DPM to EOS is under validation by ATLAS
- Data Networking
  - 30+ PB data (Inbound + Outbound) transferred in 2022
  - Able to fully utilize the 2x10Gbps links between TW and CERN - reached 19.8bps at peak
- Activities for ATLAS
  - Finished 1,200 billion events, 300 PB in 2010 - Nov. 2022 (#processed data and MC events)
  - Development of High Granularity Timing Detector (HGTD) DB and backup support
  - Support Folding@Home for COVID-19 studies
- Contributions to ATLAS Software and Computing
  - Participating development of ATLAS Harvester/Panda and RUCIO
  - Deeply involved with ATLAS data preparation activities

WallClock Consumption of Successful and Failed Jobs - Time Stacked Bar Graph

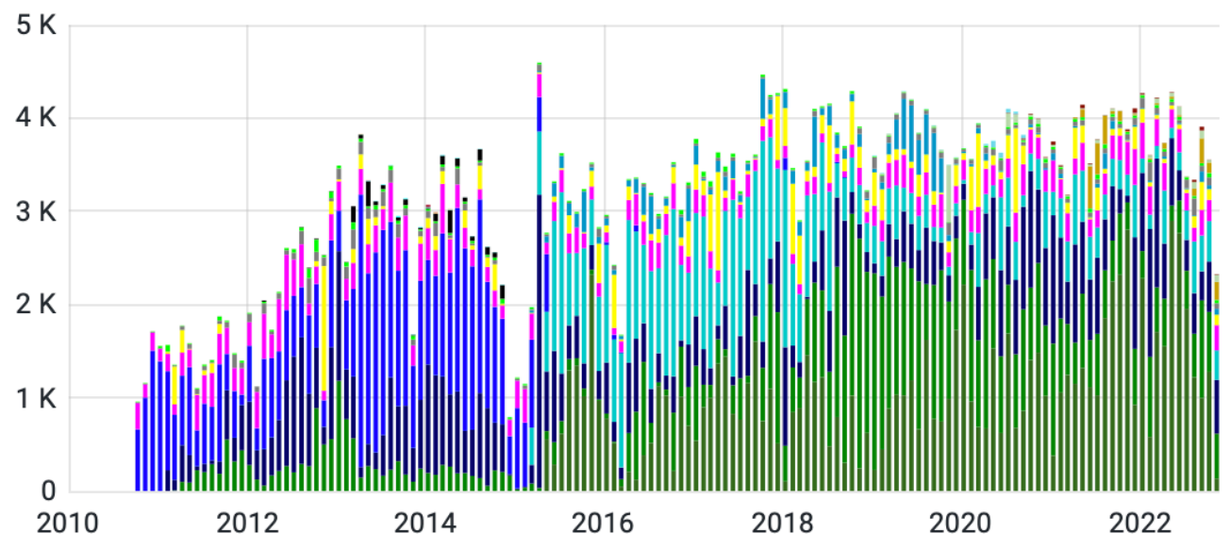


	min	max	avg	total
finished	0	11.6 Bil	6.79 Bil	1.06 Tri
failed	0	5.41 Bil	882 Mil	138 Bil
cancelled	0	211 Mil	12.1 Mil	1.89 Bil
closed	0	59.8 Mil	3.92 Mil	612 Mil

WLCG Activities @ASGC

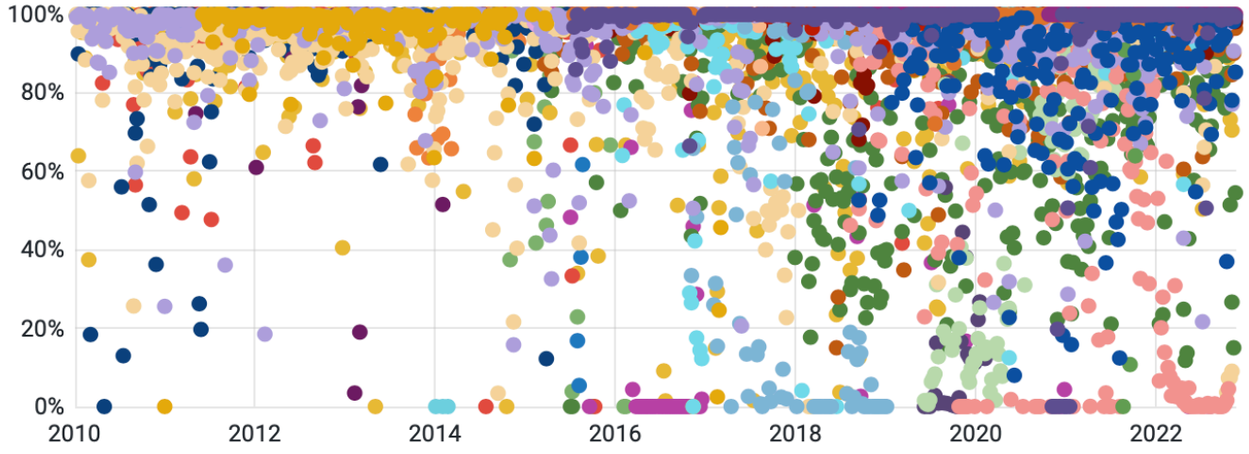


Slots of Running jobs



	min	max	avg	current
MC Event Generation	0	2.90 K	615	134
Group Production	0	2.46 K	505	481
MC Reconstruction	0	3.15 K	499	575
MC Simulation Full	0	2.15 K	402	318

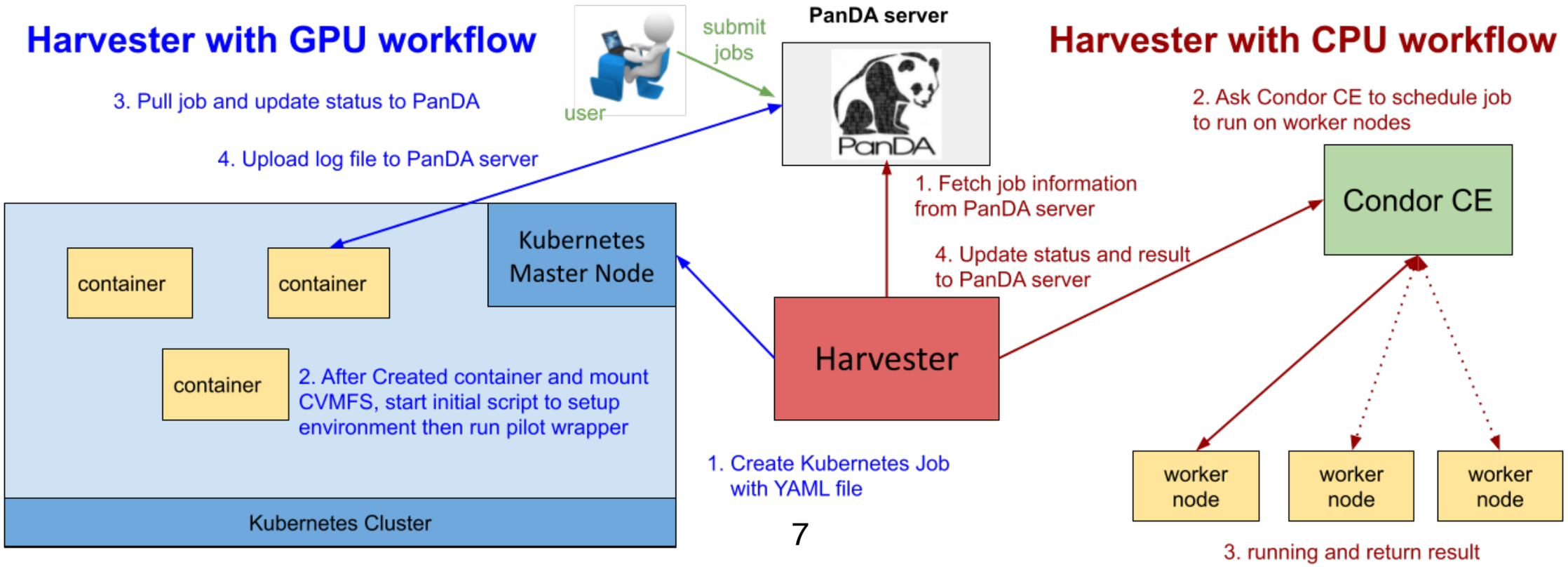
Transfer Efficiency



T0 Export	0%	100%	95.8%	100%
Analysis Download	0%	100%	96.2%	99.4%
Data Export Test	97.4%	99.6%	98.5%	98.9%
Production Download	68.1%	100%	99.1%	98.1%

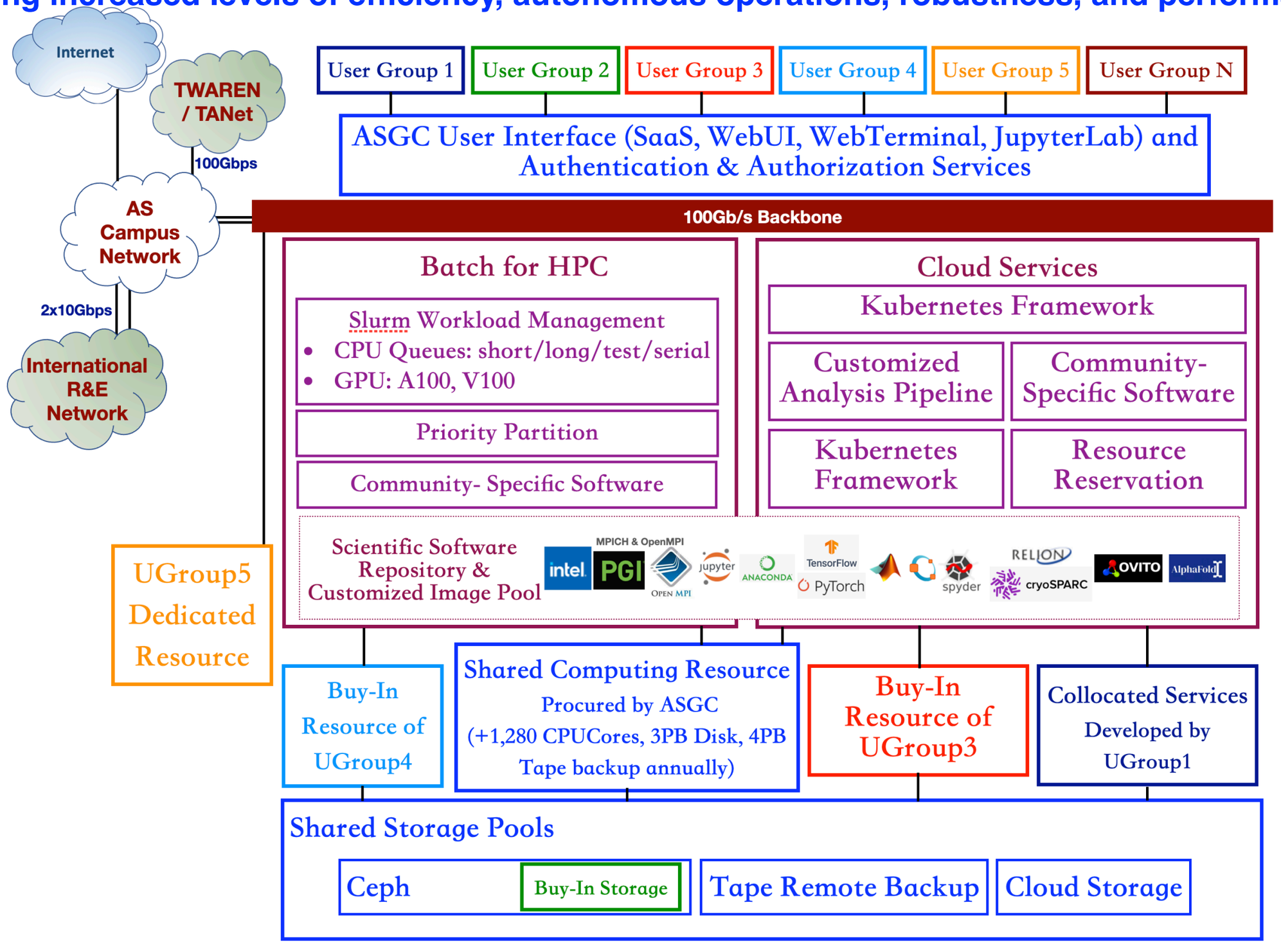
Building Distributed Cloud Infrastructure Supporting broader Scientific Applications based on WLCG

- Integrate & optimize data analysis pipeline, develop web portal/science gateway, and optimize system efficiency
- Facilitate GPU computing for big data analytics through DiCOS
- Computing model, system architecture and services, solution and technology are continuously improved by user experiences and advanced ICT
- Flexible virtual cluster over distributed heterogeneous resources
  - GPU, CPU with/without infiniband
  - Shared filesystem/storage by Ceph
  - Job scheduler through HTCondor and Kubernetes (with containers)
  - Containerization of DiCOS core components: analysis pipeline robustness; portability; maintainability;
- Federated research infrastructure achieved by distributed cloud based on WLCG core technologies



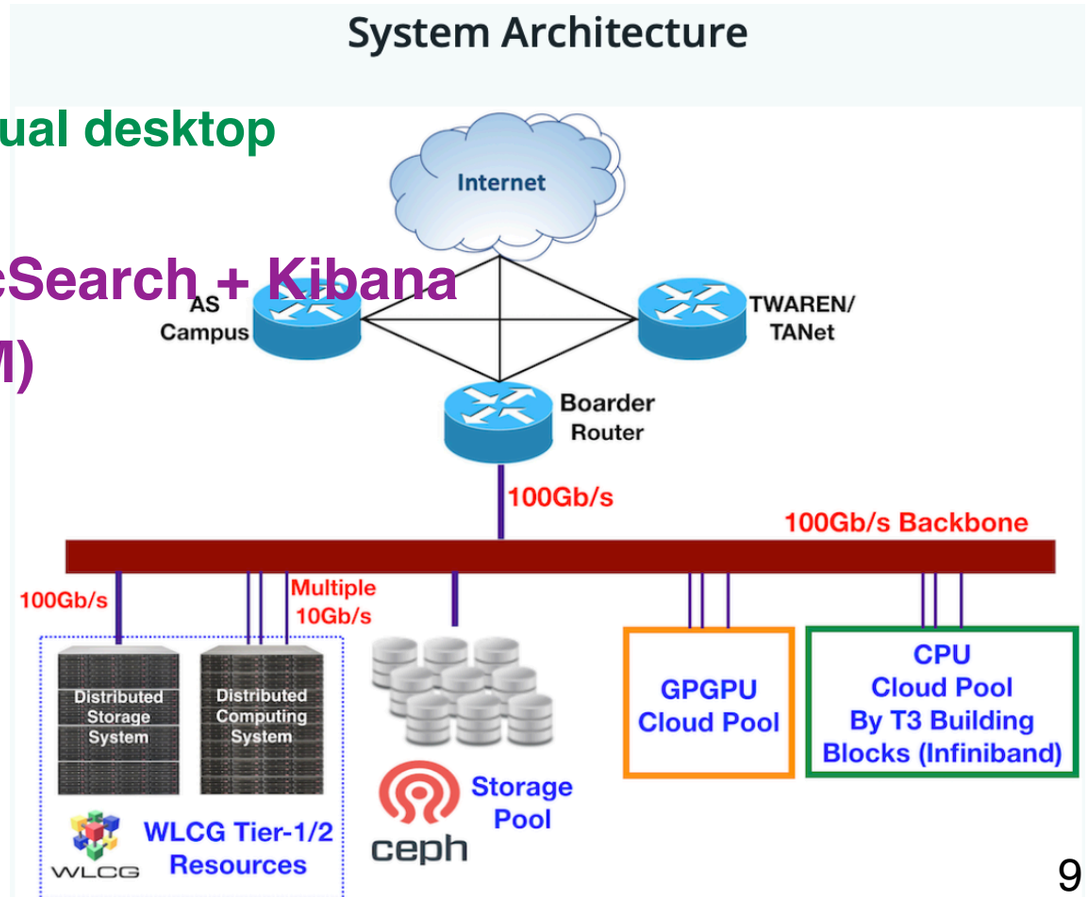
Science Cloud with Flexible Collaboration Models

- Collaboration models are defined based on requirements and computing models of user communities
- Responding to new emerging technologies, resources classes and usage modes, diverse application classes, etc.
- Achieving increased levels of efficiency, autonomous operations, robustness, and performance



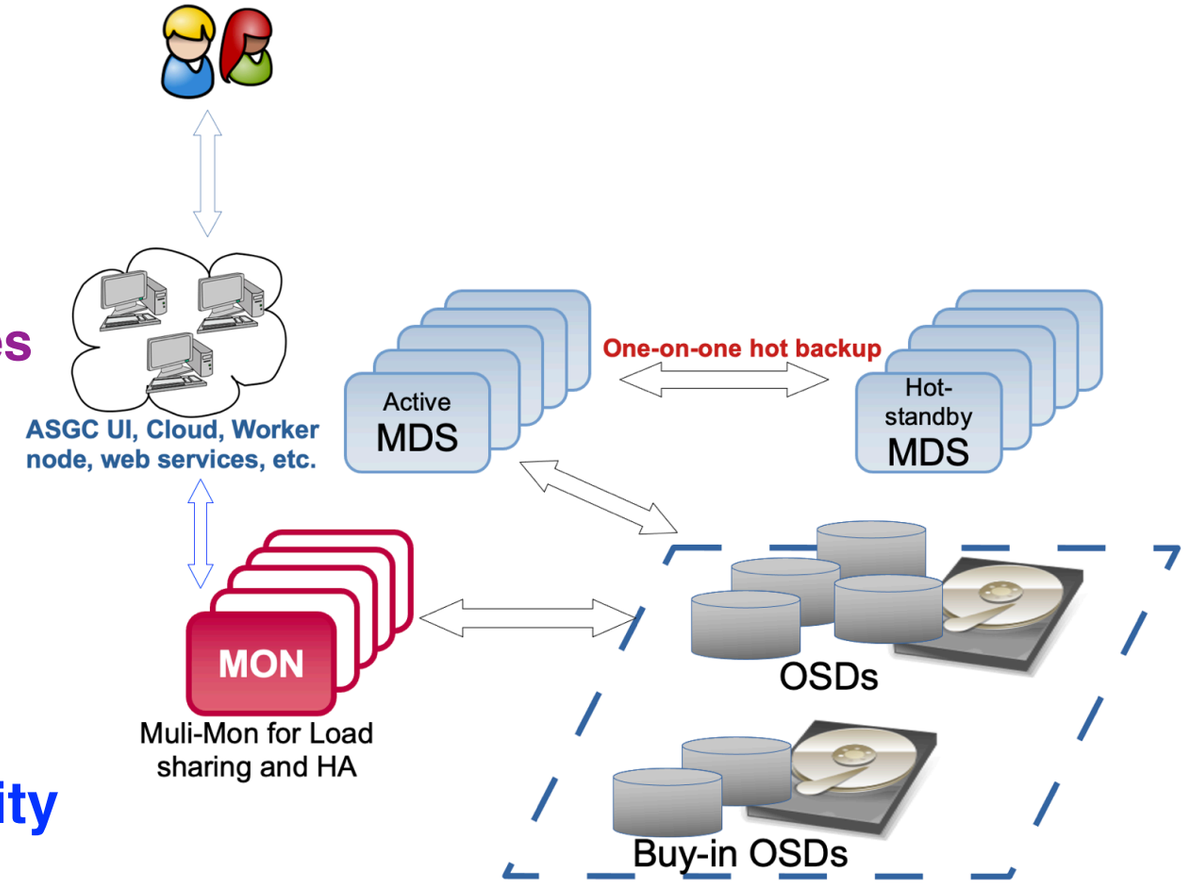
ASGC Science Cloud Infrastructure (DiCOS)

- OpenStack Cloud: for core services and on-demand worker nodes maintained by Openstack-ansible
  - Multiple cells/Region for various configurations and capabilities
    - e.g. GPU, 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
- Operation and management
  - Source control: Gitlab
  - Puppet-based deployment of components
  - HELM



ASGC Science Cloud Storage Architecture

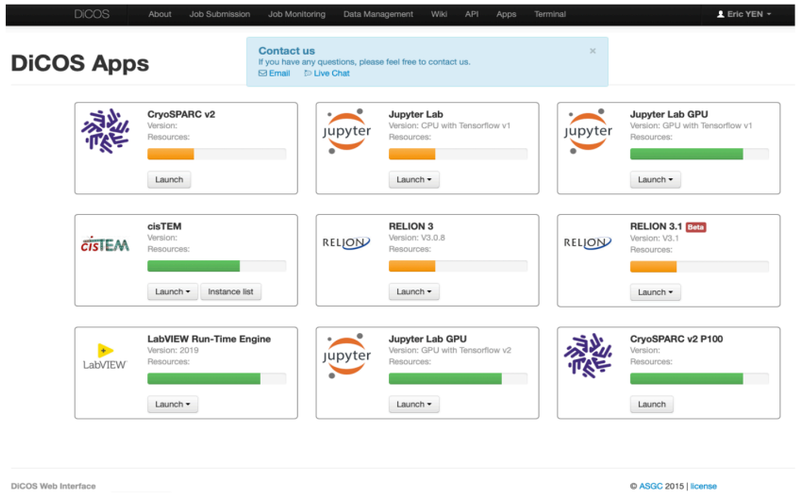
- WLCG Storage System is migrating from DPM to EOS
  - 10PB (2023) after migration
  - Two new EOS system is now under verification by ATLAS and CMS respectively
    - Data has to be moved gradually
- Storage system for other e-Science applications: Ceph-based
  - 7PB (CephFS mainly) in total (1PB used by Openstack Cinder and Glance)
  - Upgrading to AlmaLinux8 and Ceph Pacific from 2022
  - Cloud Storage (DropBox-like) services: 2TB/user
    - Accessible from user home directory
- Tape-based remote backup system (4PB) will be established in 2023
- Scaling Ceph and HA
  - 5-MDS w/ one-on-one hot backup
    - Enhanced availability
  - Pin MDS for specific application/group/directories
    - To avoid split/merge subtree across the MDS
      - Which would somehow cause slow response to client
    - Also, could allocate slightly powerful H/W to serve I/O intensive services
      - e.g. much bigger MDS memory against others
  - 5-Mon for continuous growing Ceph cluster
- ML-enabled analysis of Ceph for better reliability and performance is under development

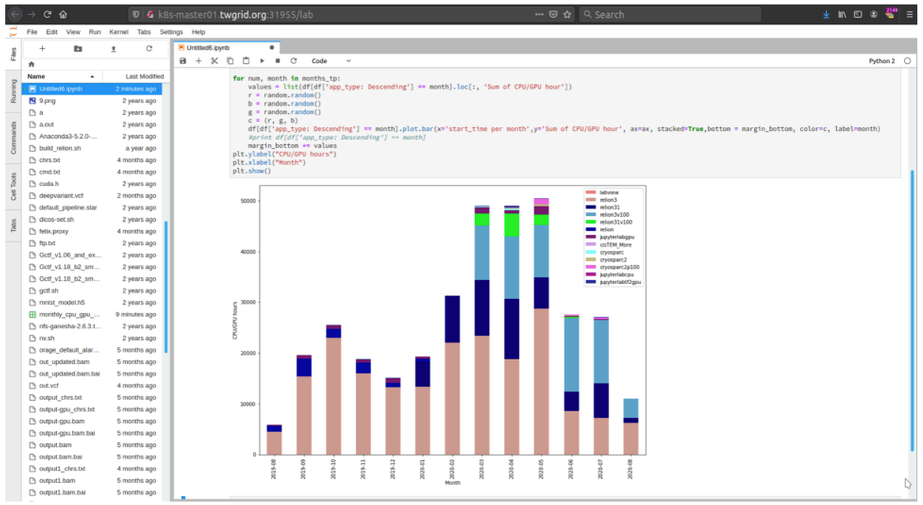


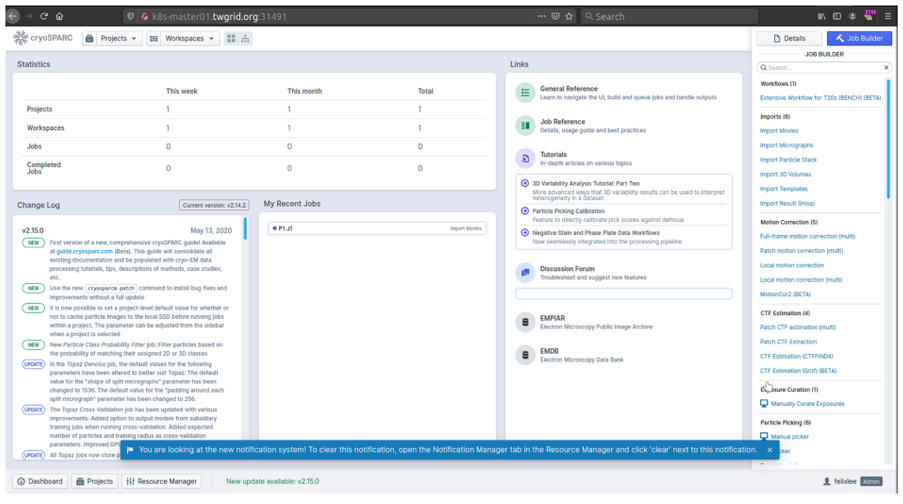
EOS Migration from DPM

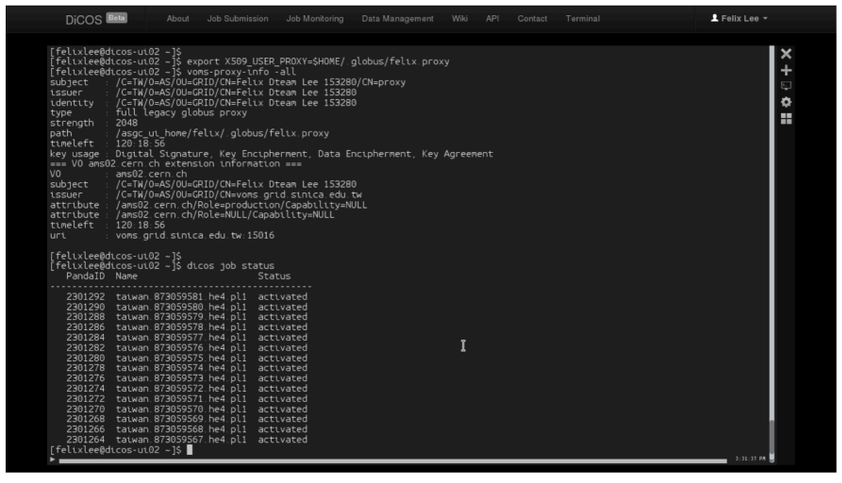
- Background: No script or tool available for direct DPM2EOS migration.
- Straightforward and safe approach - moving data from DPM to EOS.
  - Moving 9PB data from DPM to EOS is not trivial at all
  - Not possible to prepare another 9PB for EOS for the migration
- Current procedures: moving each space pool one by one, starting from data disk space (6.4PB)
  - Efficient data migration processes suggested by ATLAS
- Issues: Troubleshooting and validation take some efforts
  - EOS is easy to setup and maintain in general but document about VO support is insufficient
    - Configuring VOMS mapping via EOS VID mechanism is barely mentioned.
  - Data migration progress is slow, due to:
    - We need to bother Atlas DDM team to do high-level data moving.
  - While there are also large number of DPM sites waiting for migration and re-validation …

Supporting Big Data & AI in Innovations









CLI | **Web Portal** | **DiCOS APP** | **Jupyter Notebook** | **Science Portal** | **Web Browser/ Terminal**

Application-specific/ Generic Learning Engines

scikit learn

ROOT Data Analysis Framework

Deep Learning Engines/Frameworks

NVIDIA TRITON INFERENCE SERVER

PyTorch

TensorFlow

Keras

Computing Resource (Cloud/Grid/Slurm)

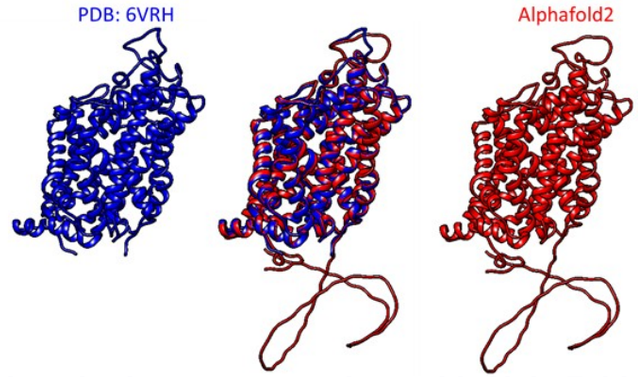
Storage Resource (Ceph/EOS)

Distributed Data Management & Cloud Storage Services

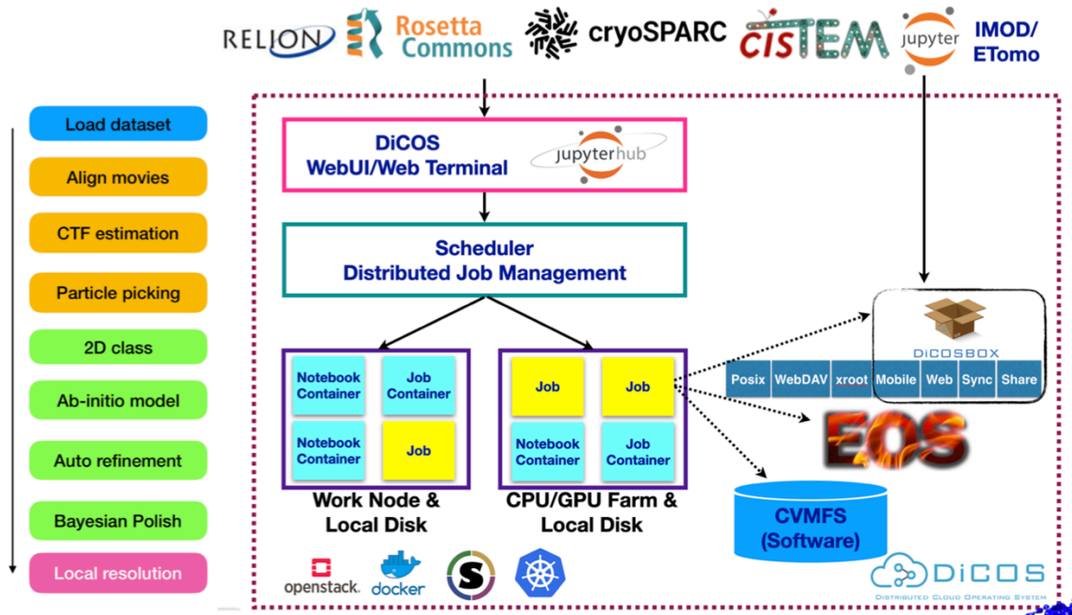
Network & Data Transmission Services

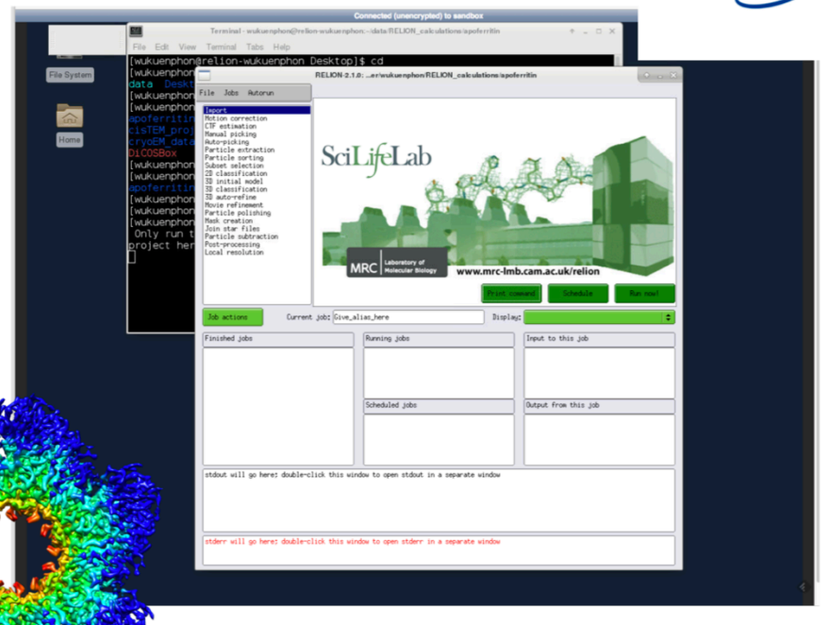
Collaboration with CryoEM Community

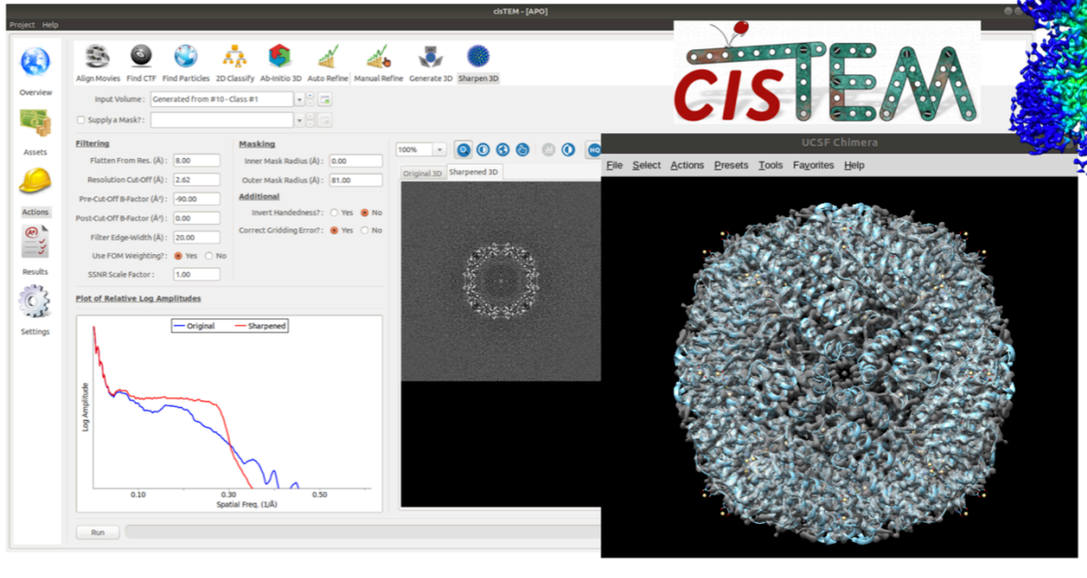
- **ASGC provides**
  - Development of web applications, portals and JupyterLab interfaces according to research workflows
  - Software package as container services
  - Data flow and application performance optimization
  - Reduced latency between CryoEM facility and big data analysis facility
- **CryoEM user community uses 43% computing resources (GPU > 90%) and 1PB storage**
- **ML-enabled functions: particle picking; ab initio 3D classification; unexpected structure discovery with minimum bias; structure determination**
- **Tools and database from AlphaFold2 and RosettaFold are also supported**

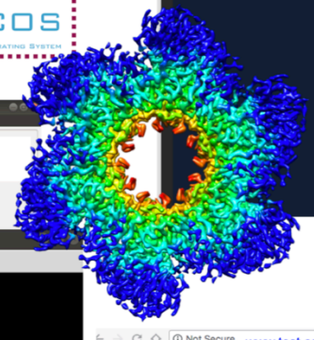


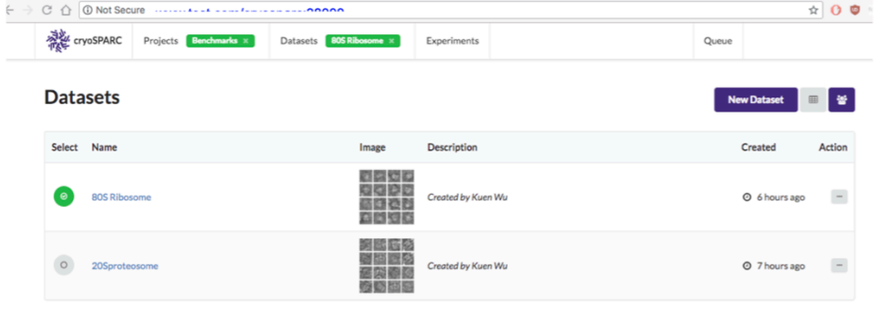
Legobody case comparison with AlphaFold











50+ Web Applications Provided

- Web Portal
- Application over Cloud
- Jupyterlab
- Web Terminal

PHYS

Deepmd-kit
Version: GPU with A100
Resources: 12%
Launch

Deepmd-kit
Version: GPU with V100
Resources: 80%
Launch

MAML
Version: GPU with A100
Resources: 12%
Launch

MAML
Version: GPU with V100
Resources: 80%
Launch

PVserver
Version: 5.8.0 (GPU 1080Ti)
Resources: 66%
Launch

Paraview Client
Version: 5.8.0
Resources: 97%
Launch

PyRoot
Version: GPU with 1080ti
Resources: 66%
Launch

Jupyter

Jupyter Lab
Version: CPU with Tensorflow v1
Resources: 97%
Launch

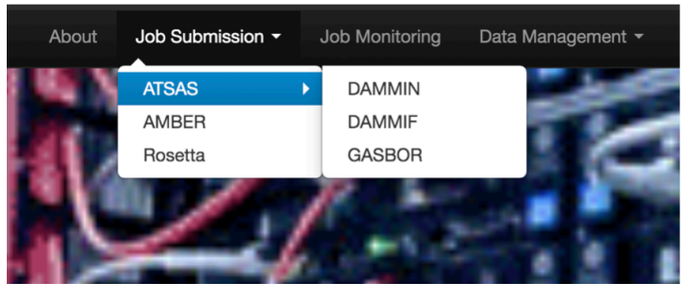
Jupyter Lab gpu 3090
Version: GPU with Tensorflow 3090
Resources: 51%
Launch

Jupyter Lab GPU V100
Version: GPU with Tensorflow V100
Resources: 80%
Launch

Jupyter Lab GPU A100
Version: GPU with Tensorflow A100
Resources: 12%
Launch

LabVIEW Run-Time Engine
Version: 2019
Launch

DiCOS-BioSAXS Platform



Other

spyder cpu/eman2
Version:
Resources: 97%
Launch

Octave
Version: V5.2
Resources: 66%
Launch

Transfer Data
Version:
Resources: 97%
Launch

cisTEM
Version:
Resources: 100.0%
Launch

Ovito
Version:
Resources: 97%
Launch

OpenACC
Version: GPU P100
Resources: 50%
Launch

Triton
Version: 22.01-py3 (GPU P100)
Resources: 50%
Launch

AlphaFold
Version: GPU with V100
Resources: 80%
Launch

AlphaFold
Version: GPU with A100
Resources: 12%
Launch

IMOD
Version:
Resources: 66%
Launch

RoseTTAFold
Version:
Resources: 51%
Launch

Dynamo
Version:
Resources: 66%
Launch

Research Data Management - Depositar

- **Deposit, Discover and Reuse - An extensible and programmable research data repository (based on CKAN)**
  - **Features: file preview (CSV, WMTS, Shapefile, ...), Wikidata as control vocabularies, Archival Resource Keys (ARKs) for PIDs, data catalogs in DCAT, etc.**
- **Sample datasets:**
  - **ark:37281/k5d515442 (Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan)**
  - **ark:37281/k507h112r (Cultural Landscapes in Emerging Digital Scholarship)**

https://data.depositar.io

Advocating Good Data Practices: From Research Data Repository to Research Data Management

Ming-Syuan Ho, Cheng-Jen Lee, Chia-Hsun Ally Wang, Tyng-Ruey Chuang | Contact: data.contact@depositar.io
Institute of Information Science | The Research Center for Information Technology Innovation, Academia Sinica, TAIWAN

Background

- Lacking national-level funder policies about Research Data Management (RDM).
- Ministry of Science and Technology (MOST) of Taiwan started to require Data Management Plans (DMP) for some research grants on sustainability in 2020.
- The *depositar* is a general-purpose research data repository developed at Academia Sinica; it is open to all to use.

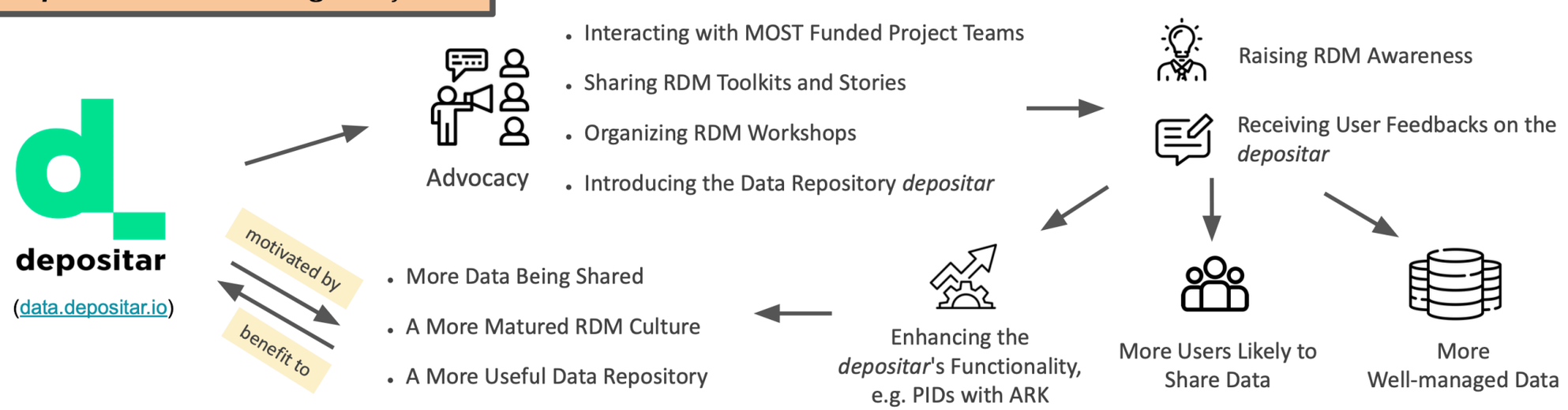
What the *depositar* Team is Doing

- Interact with some MOST funded projects on implmenting research data management.
- Develop "RDM Hub" (rdm.depositar.io) to introduce RDM toolkits and stories.
- Organize RDM workshops for researchers to share their experience.
- Operate and enhance the data repository *depositar* according to user needs.

Lessons Learned

- Funder policies are critical in driving good RDM practices; it is a gradual process.
- A feasible DMP is better than an ideal DMP. Experience sharing is essential in cultivating a culture of research data management.
- Frequent communication between project teams and data repositories is helpful in making data FAIR (Findable, Accessible, Interoperable, and Reusable).

The *depositar* in RDM Ecological Systems



System Efficiency Optimization

- Goals: maximize power, thermal and system (Comp, Storage, Network, application) efficiency
- Strategy: intelligent monitory and control assisted by ML technologies
- Example: Thermal management, Compute/storage/network anomaly detection, Power saving of work nodes
- AHU monitoring and control: ~500 warnings and 192 overheat alarms issued in 2022 (by end Oct)
  - Detection of refrigerant operating issues and abnormal components; Efficiency optimization
  - 13 sensors x 16AHU; 18K data points/day;
  - Realtime monitoring, adjustment and diagnostics: refrigerant operating issue; abnormal components detection; efficiency tuning; ML-based automatic detection of critical problems;
- System Anomaly Detection
  - Classify machine status into 5 clusters daily: based on CPU-user, CPU-wio, CPU-system, CPU-idle, Network In/Out
  - >30M records/day from all systems of ASGC are covered
  - 146 events in 14 types identified during March 2020 - March 2021

Cooling efficiency

- Water-cooled condenser
- Detection of refrigerant operating issues and abnormal components
- Realtime monitoring, adjustment and diagnostics (18K data points everyday/ refrigerator): ML-based automatic detection of critical problems;.
- PUE ~ 1.5 (1:2)
- R&D conditional cooling architecture for innovative green single rack cloud center based on space technology

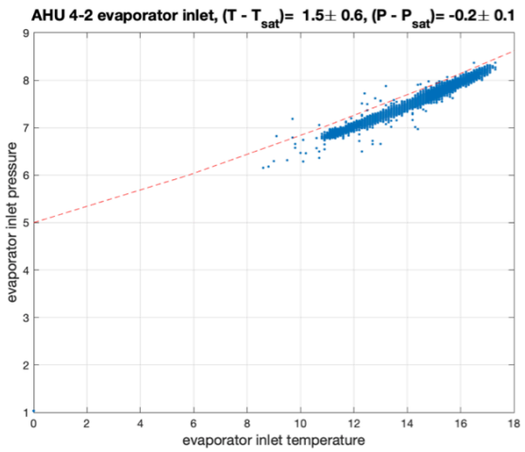
Power efficiency

- Real time monitoring
- UPS only for critical services (10% saving)
- Power saving algorithm based on job status (25% power saving)

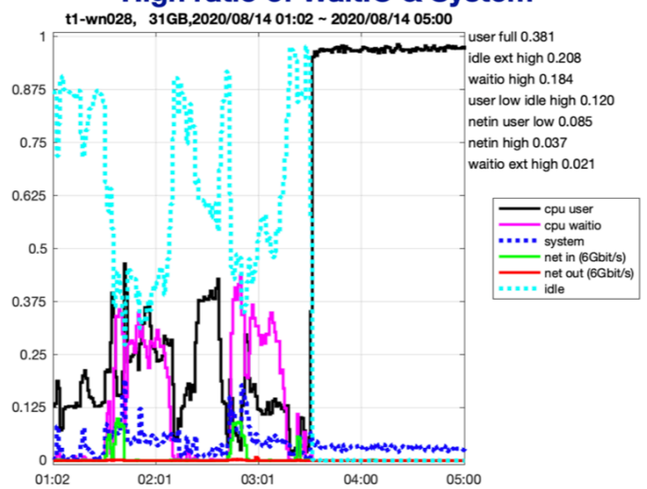
Application and System efficiency

- Monitor all nodes every 10 seconds.
  - Node efficiency
  - Service efficiency
  - Job efficiency
  - Data efficient

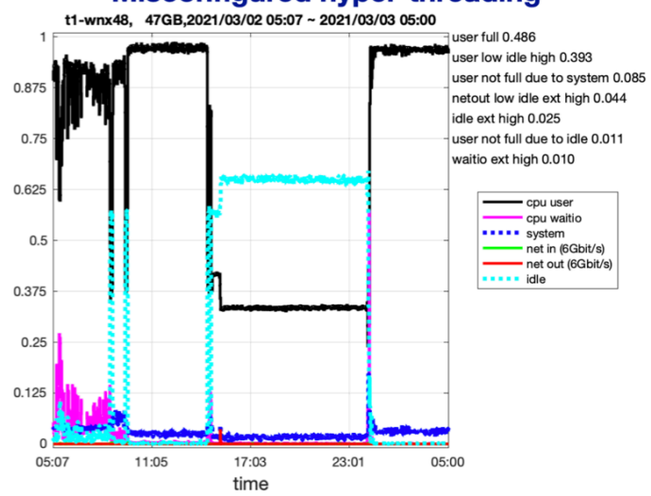
AHU Performance Monitoring



Worknode Monitoring: High ratio of WaitIO & System

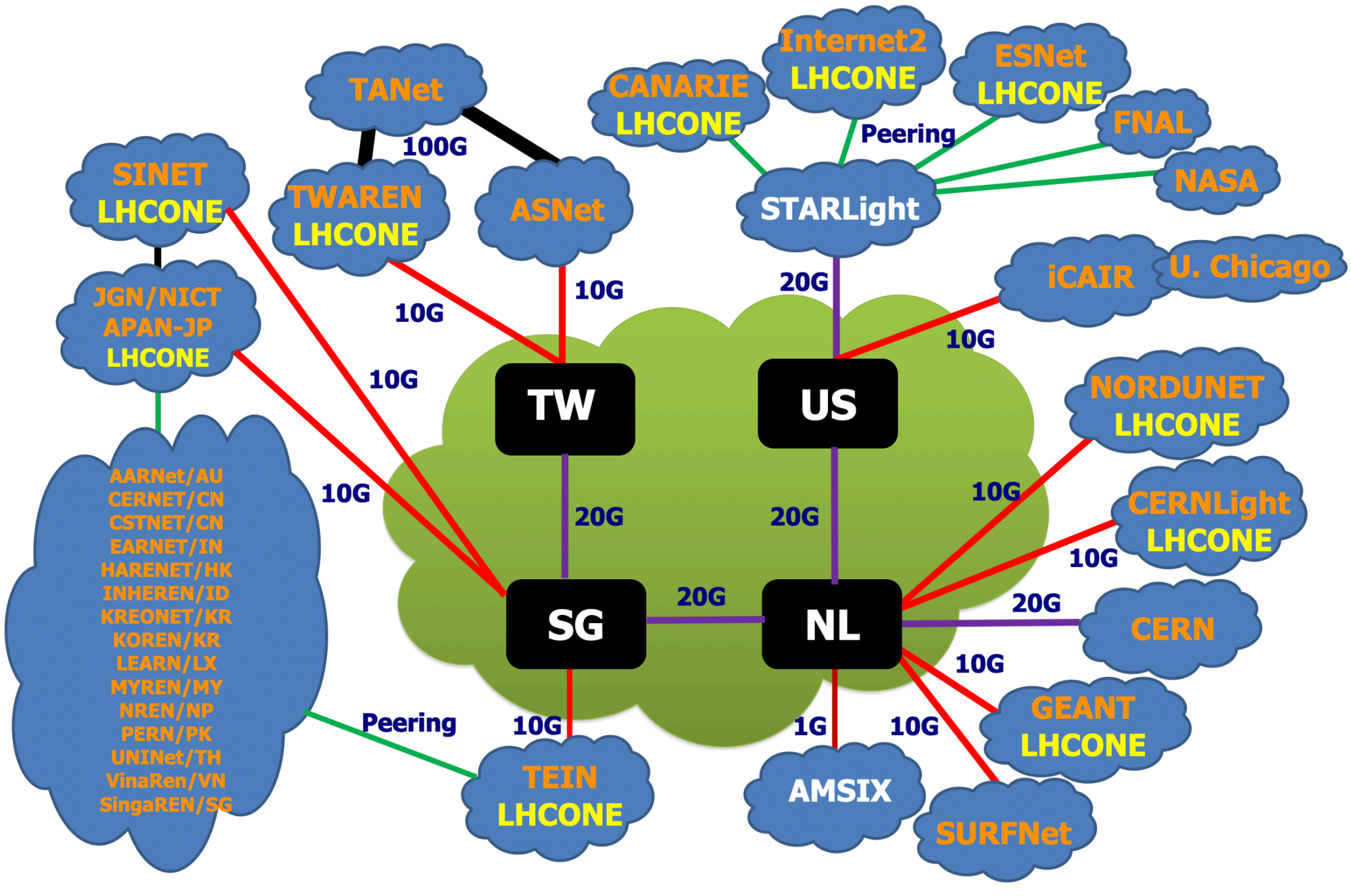


Worknode Monitoring: Misconfigured hyper-threading



ASGCNet is responsible for providing high-throughput research network from Taiwan to Europe and Asia

- Achieved 19.5Gb/s and 19.6Gb/s In/Out performance respectively at the same time over 2x10Gb/s international backbone between TW-SG-NL-CERN
- Automatic backup with JGN and TEIN for LHCONE/LHCOPN traffic
- All WLCG services are in IPv4/6 dual-stack
- Inside ASGC DC: 100Gb/s backbone operational since early 2020
- SDN (experiment) and VRF (in production)



Summary

- Based on WLCG core technologies, ASGC is supporting big data analysis and AI in innovations for broader disciplines
  - Flexible Collaboration models - turning research needs into services
  - Upkeep of scientific computing and big data analysis systems
  - Workflow integration, customization and efficiency improvement
  - Resource federation for extension of the research infrastructure
    - Facility owner has priority usage but resource has to share with AS users
    - Bring your own hardware or investment, etc.
    - Distributed cloud development and operation
    - Research data management framework is available for FAIR-enabled open data
    - Capacity building: consulting, training, workshop, hackathon, etc.
- Demands of ML-Cloud services are booming, which will be a core of the Science Cloud in Academia Sinica
- Efficiency optimization is not just a pillar of the ASGC Science Cloud, but also the essential contributions to reliability and performance