Distributed Cloud Research Infrastructure & Applications in Taiwan

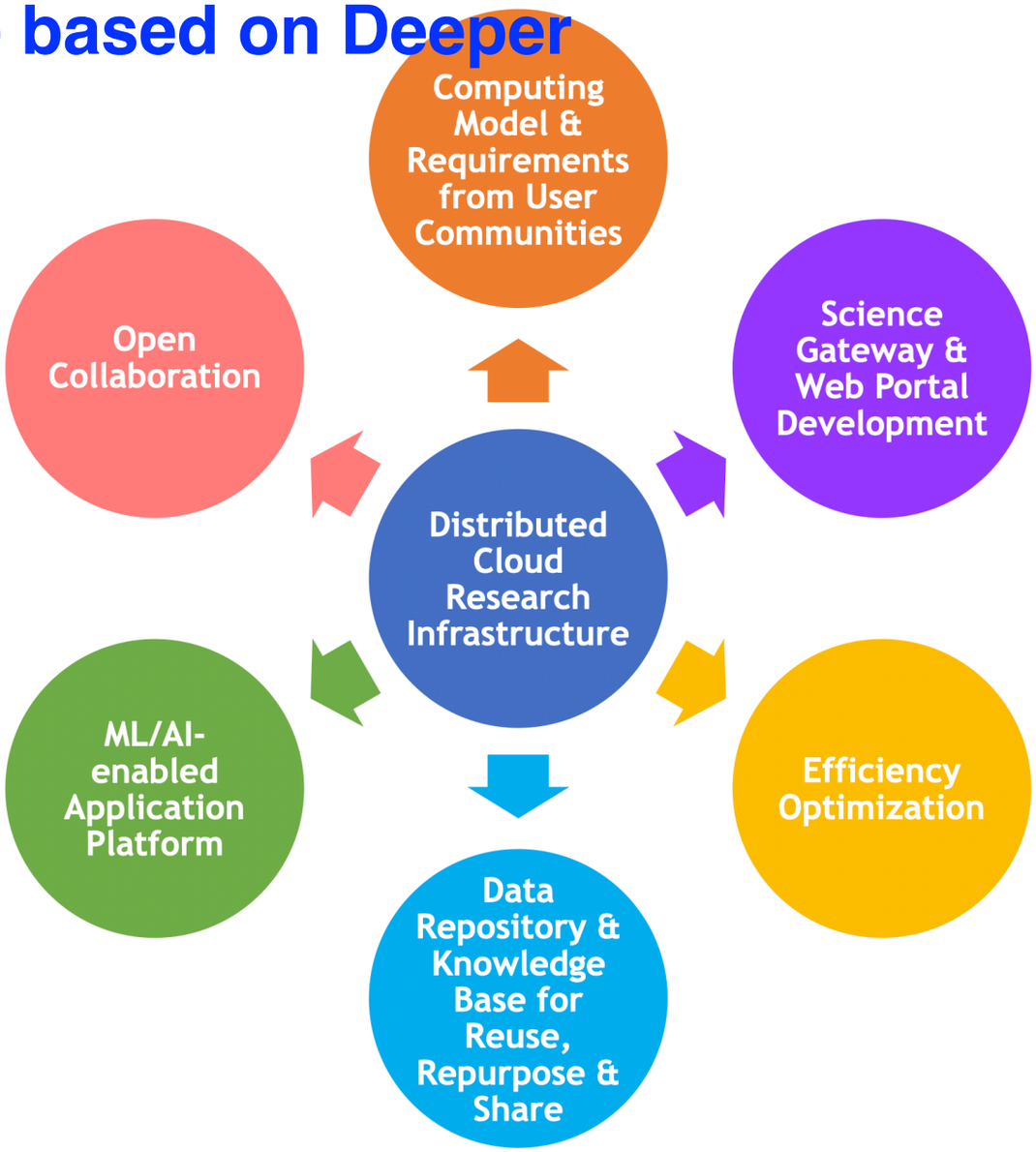
Felix Lee

Academia Sinica Grid Computing Centre (ASGC)
Taiwan

ISGC 2021
23 March 2021

e-Science for the Masses

- In addition to support big data analytics efficiently, all types of research applications of various disciplines could also benefits from e-Science
- Instead of cycle provision, the whole analysis pipeline is integrated and optimized over e-Infrastructure, with close collaboration with user communities
- Open collaboration model for e-Science based on Deeper Understanding approach
- Driving Advancement of Science, Technology, Collaboration & Services by More Case Studies
- Progressive Enrichment and Enhancement by more Case Studies and Support Innovative Knowledge Discovery
- Customized Science Gateway & API for flexible Integration
- Open Science, Reproducibility, Repurposing, Reuse and Share



Primary User Communities

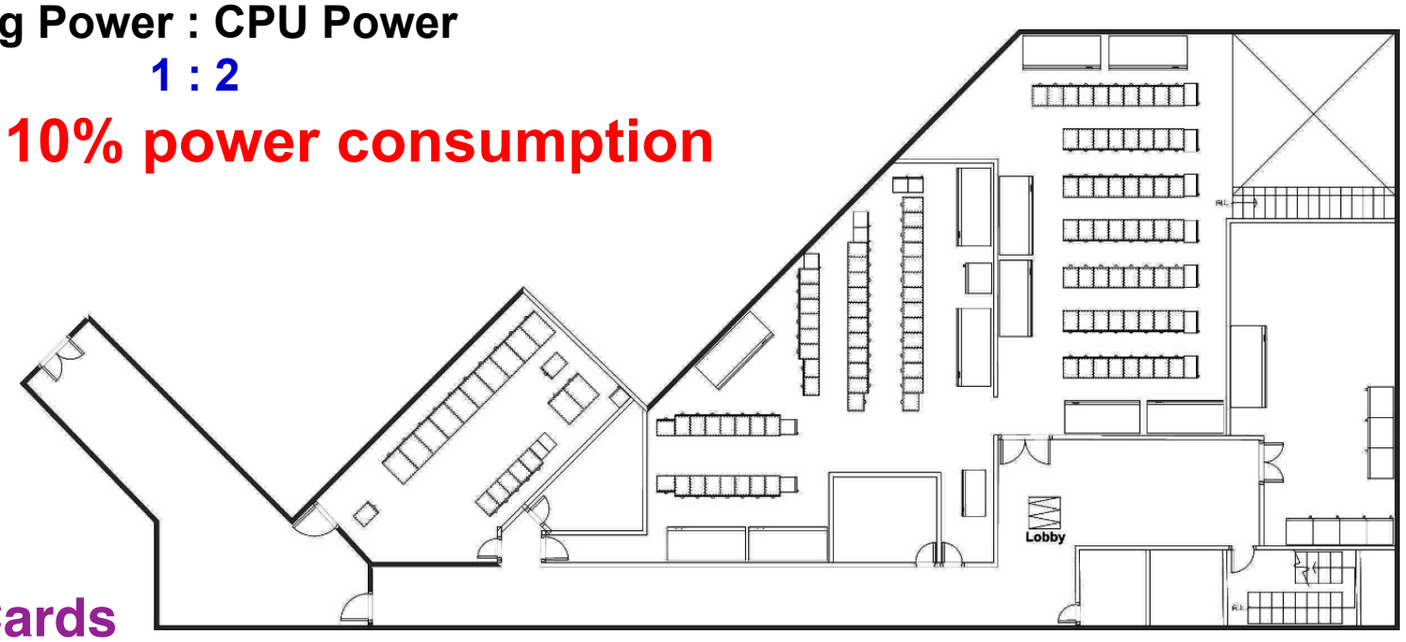
- User Communities
  - ATLAS/LHC: search for the origin of matter
  - AMS: explore the universe with Space-borne detector
    - Launched to ISS in 2011; ~160BN cosmic ray events recorded
  - TEXONO/Neutrino: search for the origin of left-right asymmetry;
  - Gravitational wave: KAGRA and LIGO
  - CryoEM: Assist analyzing protein structure and cellular ultrastructure for further solutions of bio-technology issues and serious disease problems.
  - BioSAXS: advanced bio-structural researches by biological small-angle X-ray scattering beamlines
  - NGS: precision medicine;
  - Drug discovery; Proton therapy; Lattice gauge theory; Condense Matter; Earth science; Environmental changes; Hazard risk analysis; Biodiversity and ecology monitoring; etc.
- Types of services: HPC, HTC, data backup and archive, data transmission, analysis pipeline integration, web portal development and customization, as well as ML-enabled application platform, etc.

ASGC Resources

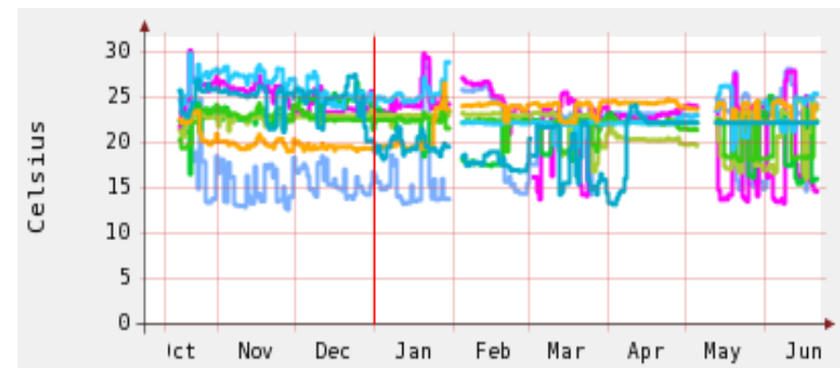
Cooling Power : CPU Power

1 : 2

No UPS to save 10% power consumption



- **Total Capacity**
  - 2MW, 400 tons AHUs
  - 112 racks in ~ 800 m²
- **Resources (2021)**
  - 20,922 CPU Cores
  - 868,352 GPU CUDA Cores/160 GPU Cards
    - A100 Server *3 (8 GPU Card for each) will be deployed in 2021
  - 20 PB Disk Storage
  - 2x10Gb links to CERN and primary NRENs worldwide
- **WLCG Tier-1 Center since 2005**
- **Supporting HPC & HTC in Academia Sinica by distributed cloud operating system (DiCOS)**
  - Usage > 1M CPUCore-Days in 2015 (ATLAS used 61%)
  - Usage > 2M CPUCore-Days in 2019 (ATLAS used 68%)
- **R&D on system efficiency optimization by intelligent monitoring & control**



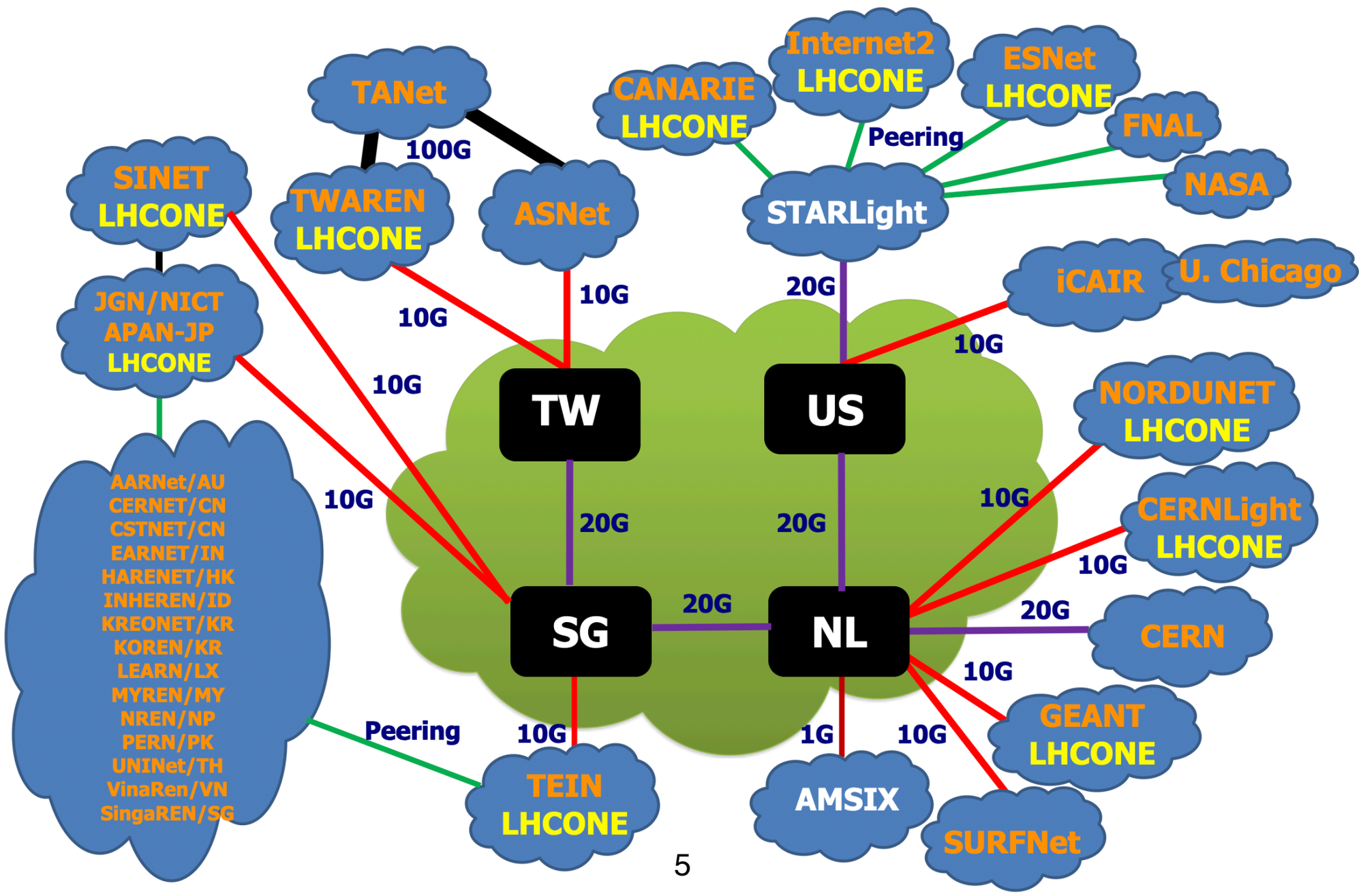
Monitoring the power consumption and temperature of every piece of equipment every 10 seconds.

All software used are open-source codes developed by ASGC and an international collaboration led by CERN

4

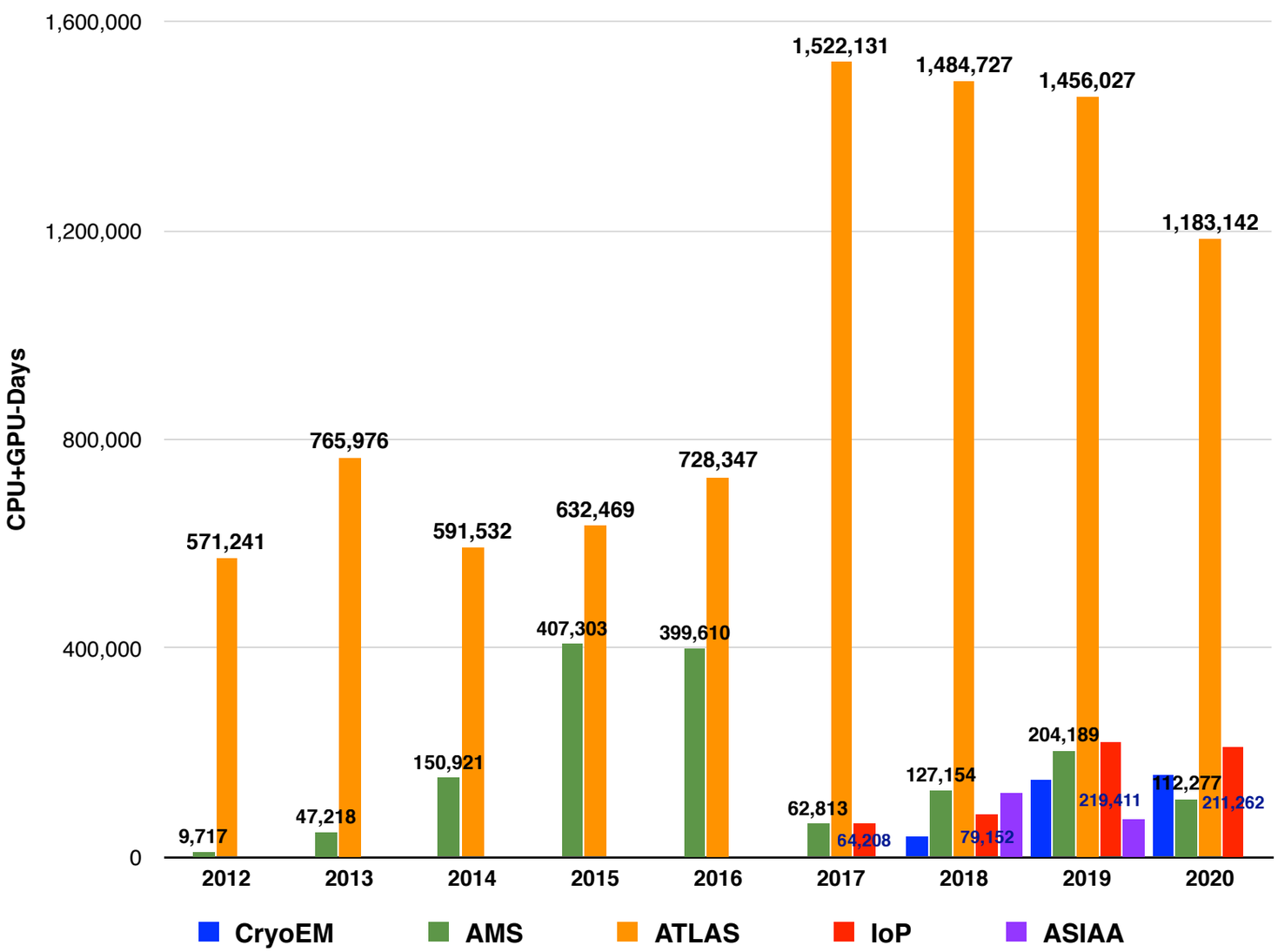
ASGCNet is responsible for providing high-throughput research network from Taiwan to Europe and Asia

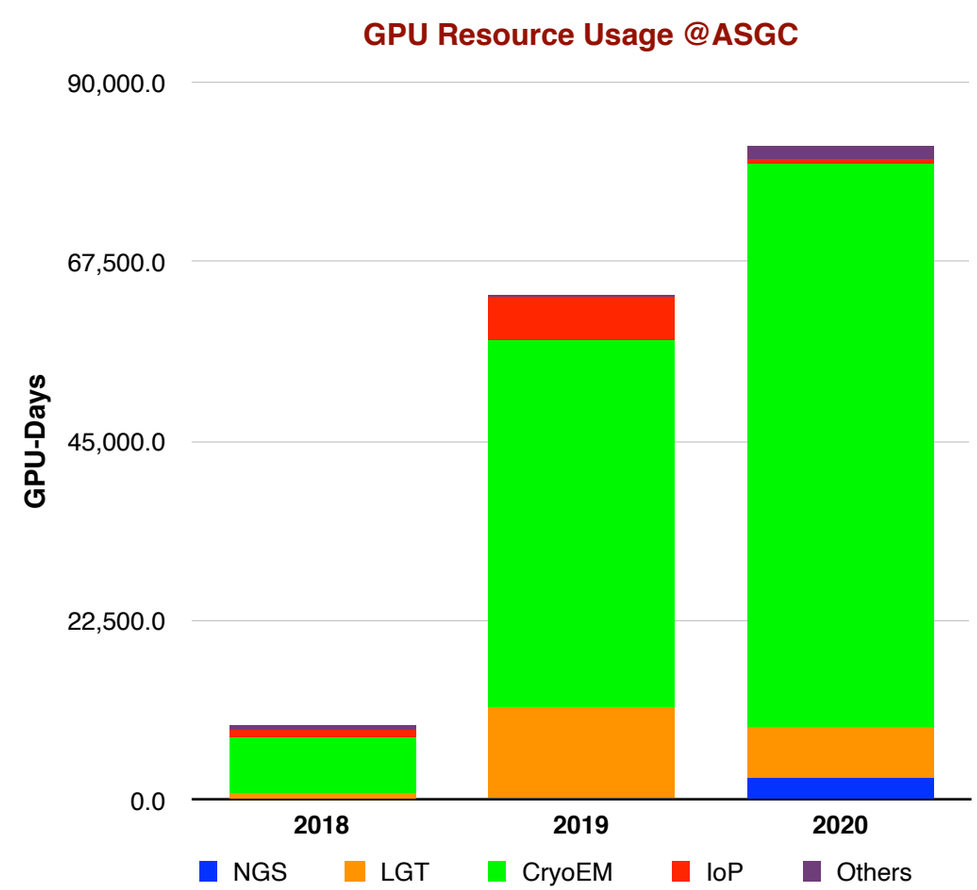
- Achieved 19.5Gb/s and 19.6Gb/s In/Out performance respectively at the same time over 2x10Gb/s international backbone between TW-SG-NL-CERN
- Automatic backup with JGN and TEIN for LHCONE/LHCOPN traffic
- All WLCG services are in IPv4/6 dual-stack
- Inside ASGC DC: 100Gb/s backbone operational since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



ASGC Resource Usage Statistics

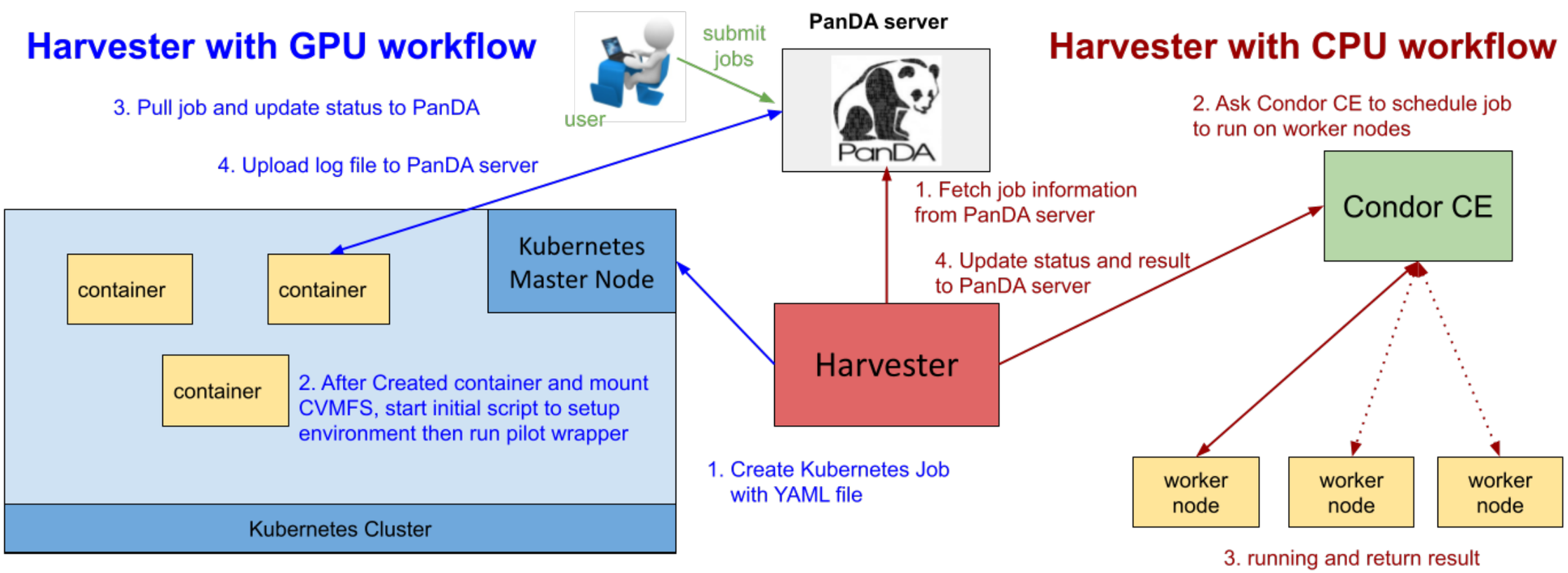
- **ATLAS and AMS are the largest users, but others are growing fast.**
- **GPU usage increases exponentially since 2018. Main applications include: Bio-imaging, NGS, ATLAS, Gravitational Wave, Bioinformatics, Lattice QCD, Soundscape, Efficiency, and ML/AI etc.**
- **Integrated by DiCOS with efficient file system/storage.**





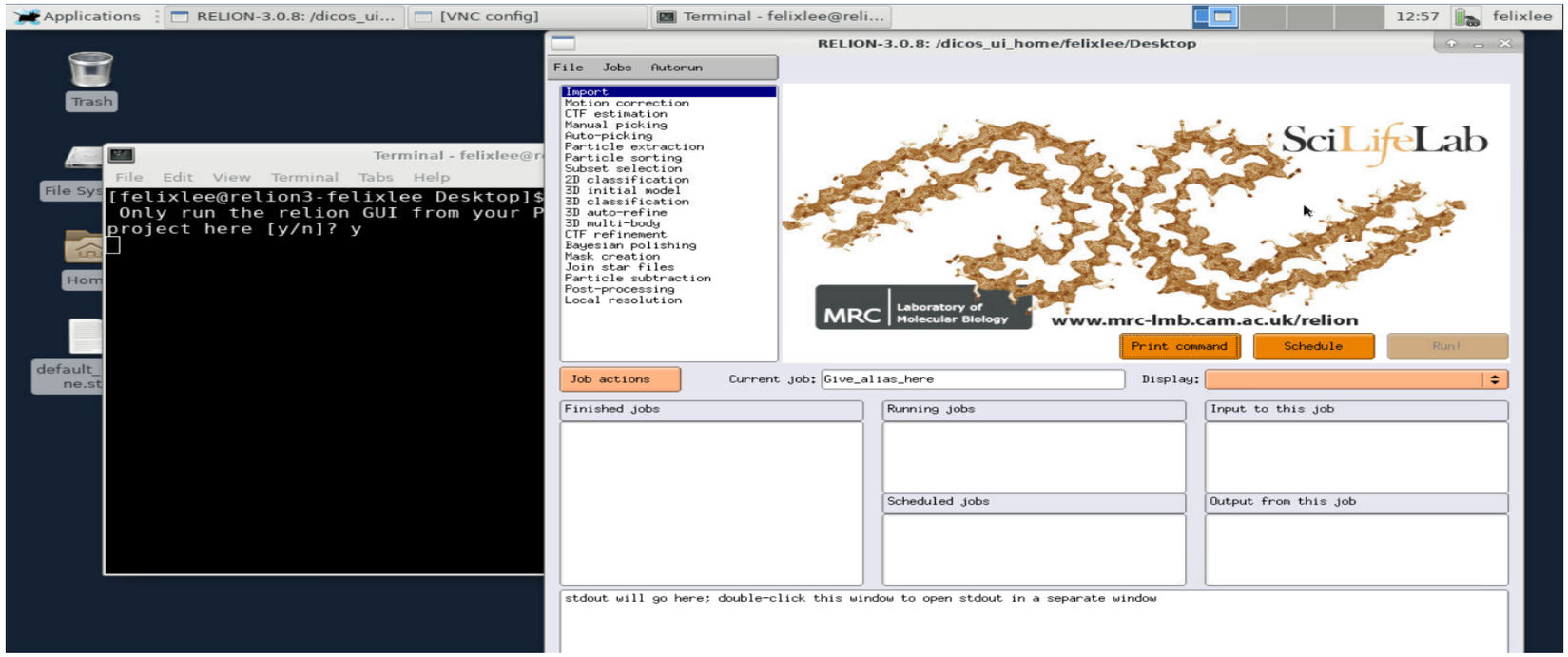
Building Distributed Cloud Infrastructure Supporting Broader Scientific Applications Based on WLCG Core Technologies

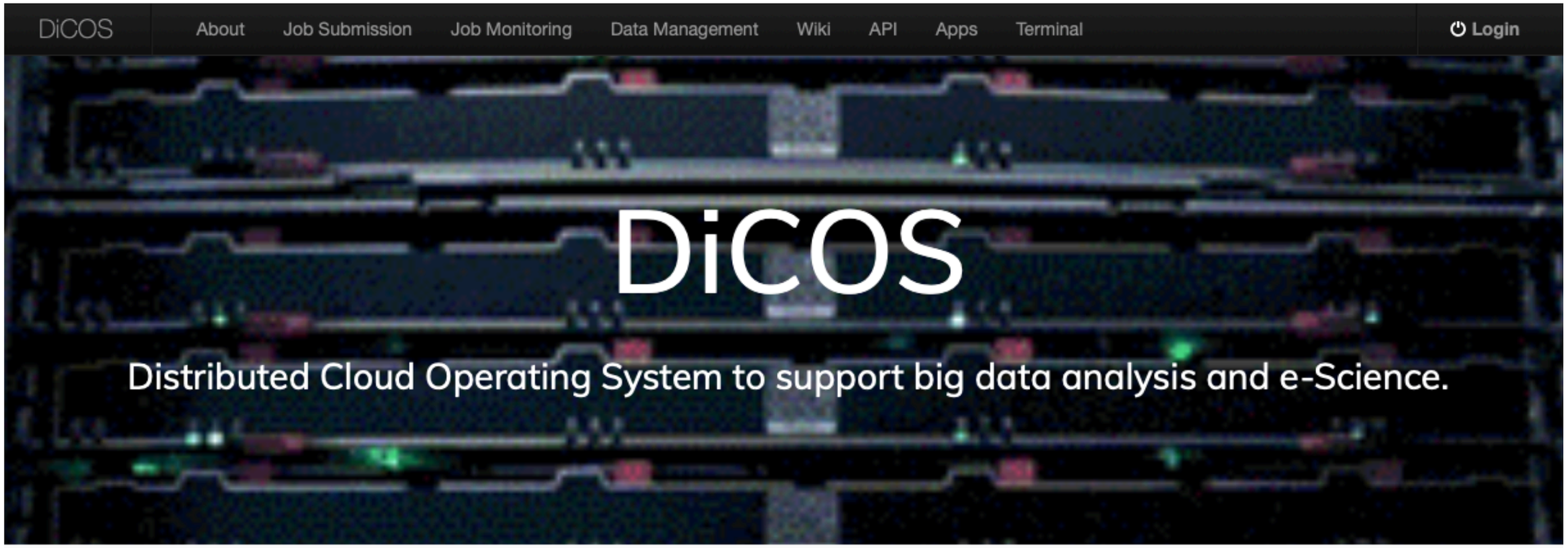
- Integrate the whole data analysis pipeline, develop web portal/science gateway, and optimize system efficiency
- Facilitate GPU computing for big data analytics through DiCOS: >100K GPUCard-Day used in 2020
- Computing model, system architecture and services, solution and technology are continuously 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;
- Support interactive UI for user analysis: user application, JupyterLab, Web UI, etc.
- Provide infrastructure, platform and application services by federating resources across institutes



Cloud and Ceph Storage Systems

- Cloud Services: both core services and applications are containerized
  - Openstack
    - For Generic services and on-demand worker nodes
    - OS Upgrade: CentOS7 → CentOS8
    - Upgrade Migrate from Newton to Ussuri
  - Kubernetes
    - For GPU cloud and other core services
    - Offering batch, interactive GUI jobs and services
      - Such as remote Jupyterlab and virtual desktop
- Ceph Storage
  - RBD for Openstack (1.4PB): Upgraded from Hammer to Nautilus
  - Cephfs for local users (4.4PB): Upgrade to Nautilus is under way
- DiCOSBox over EOS for cloud storage services (1TB/user)

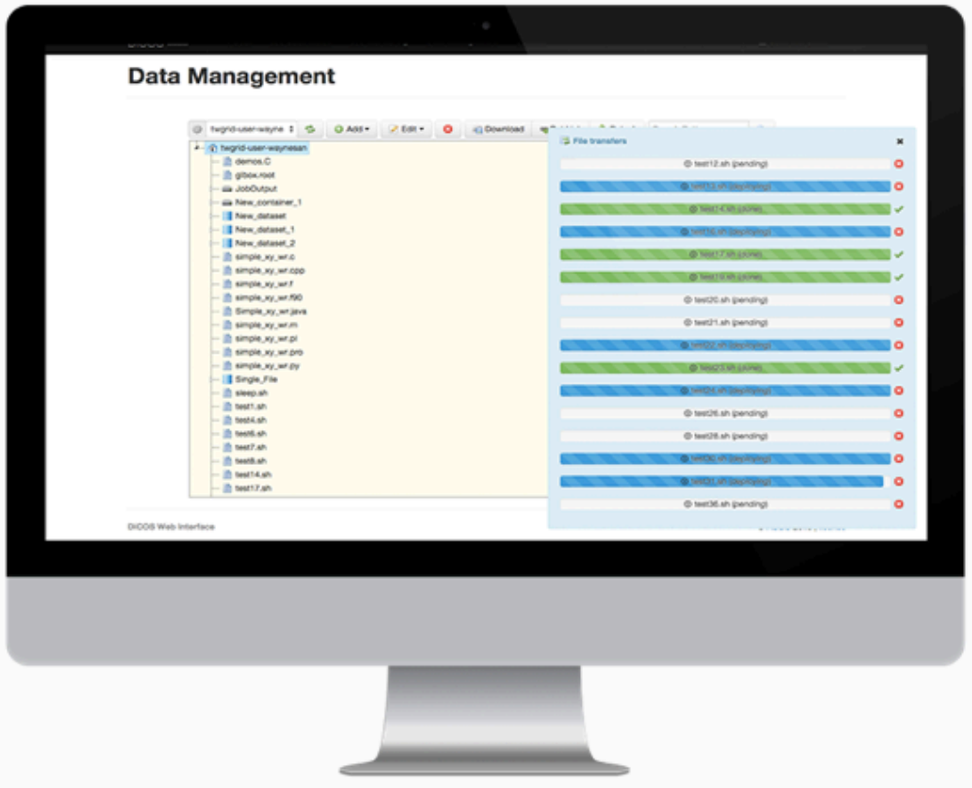




DiCOS

Distributed Cloud Operating System

https://dicos.twgrid.org



About DiCOS

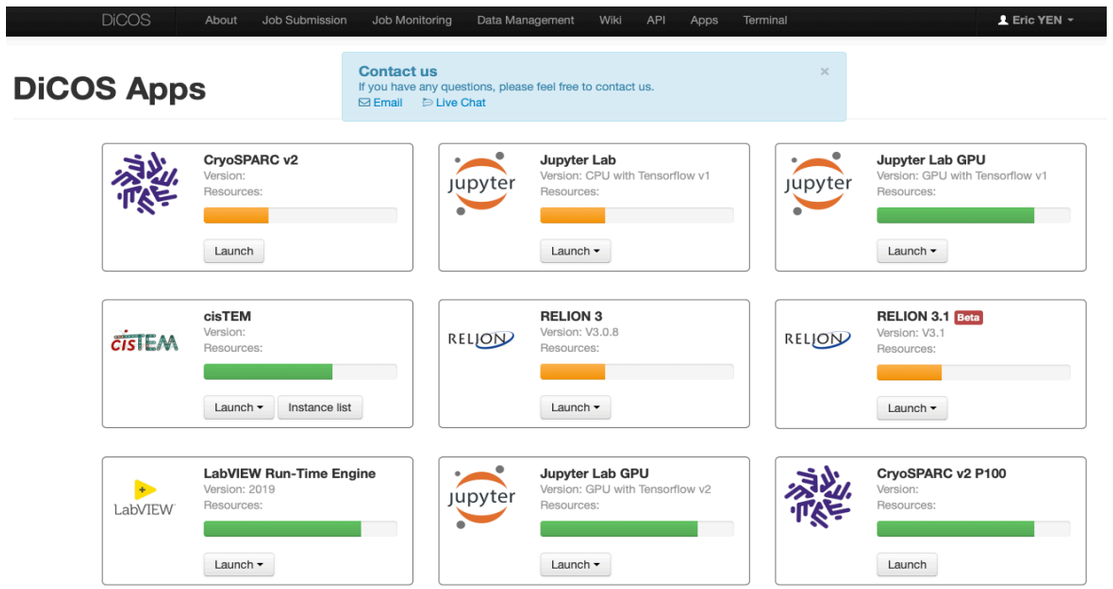
DiCOS is the operating system for the distributed Cloud environment, to provide scalable infrastructure, flexible platform as well as intelligent data and job tools for high performance scientific and generic computing purposes.

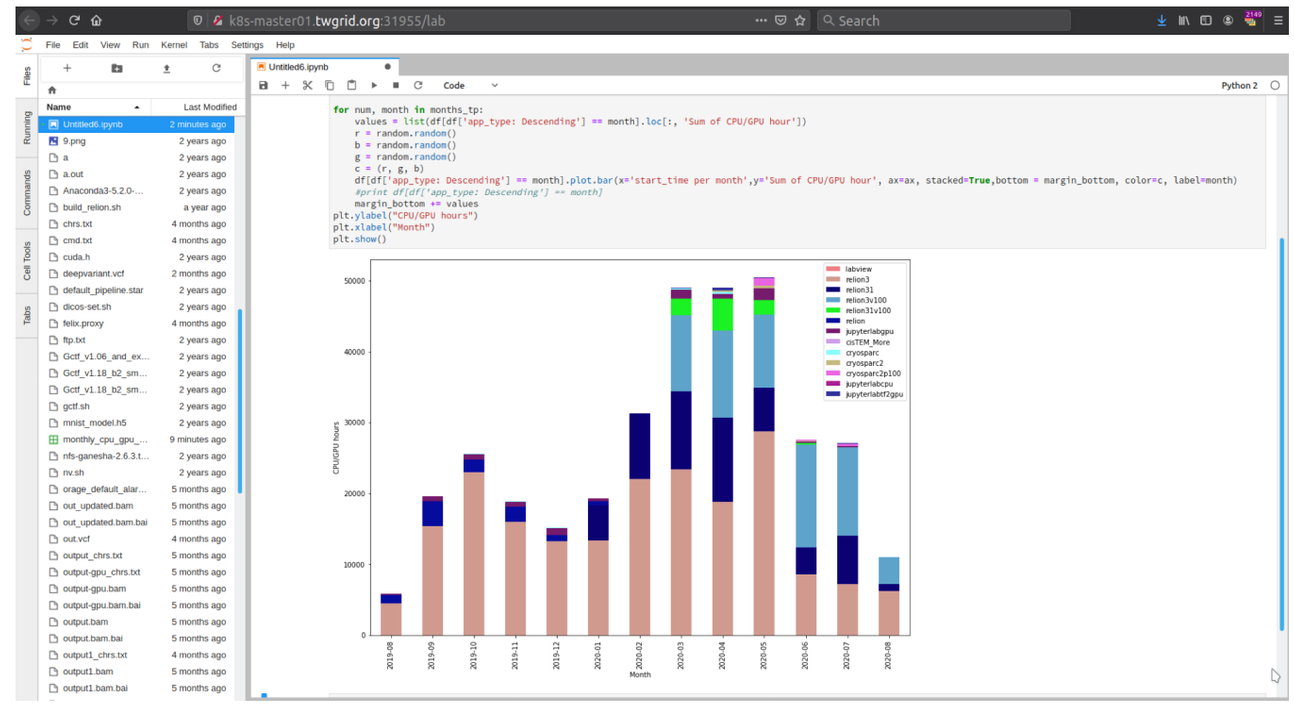
In addition, together with the fanless rack system, DiCOS also serve as the building block of various levels of resource centers for the distributed Cloud.

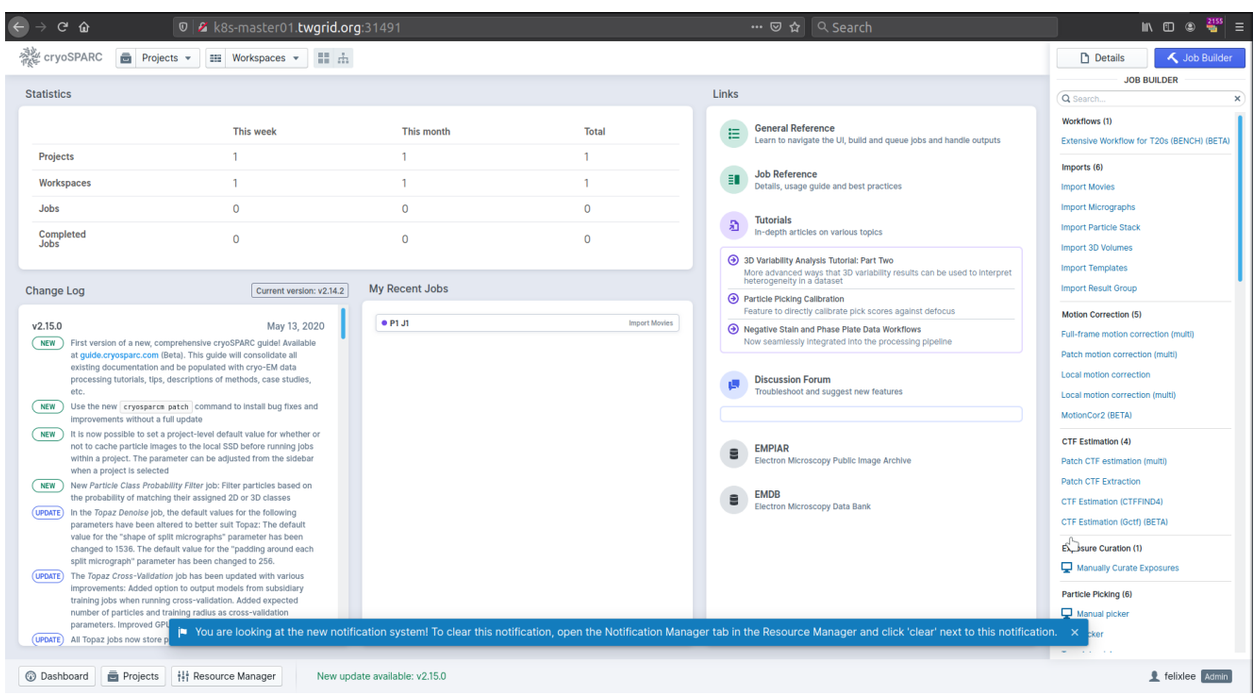
DiCOS-UserManual (English) DiCOS-UserManual (Chinese)

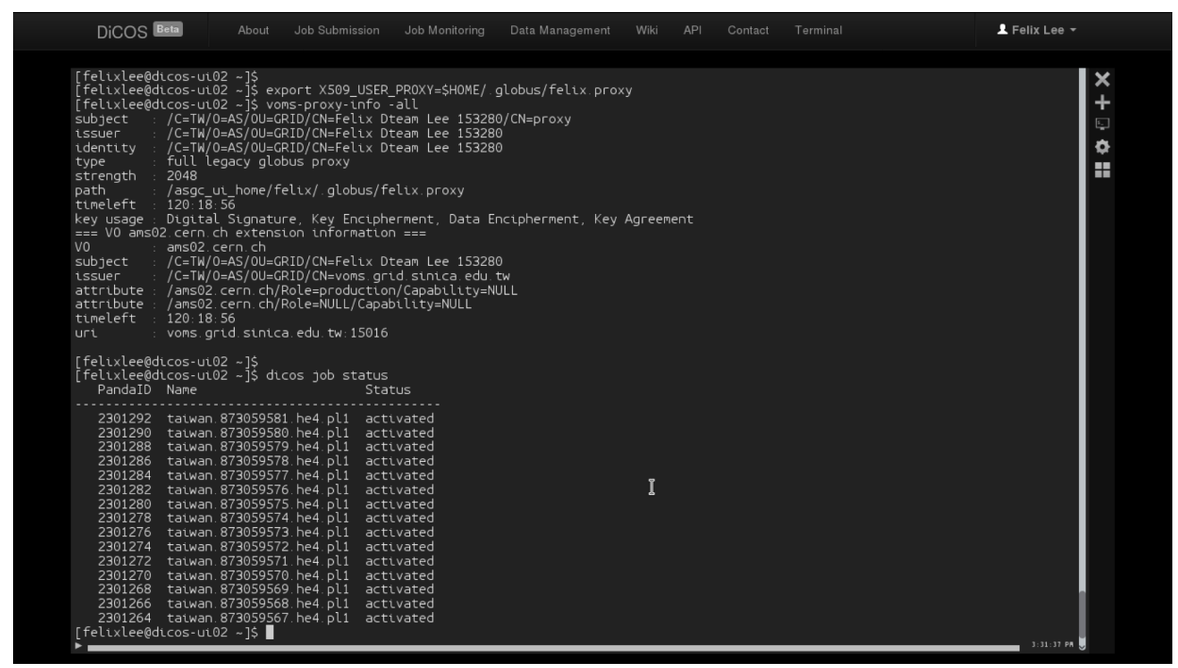
Interactive UI

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



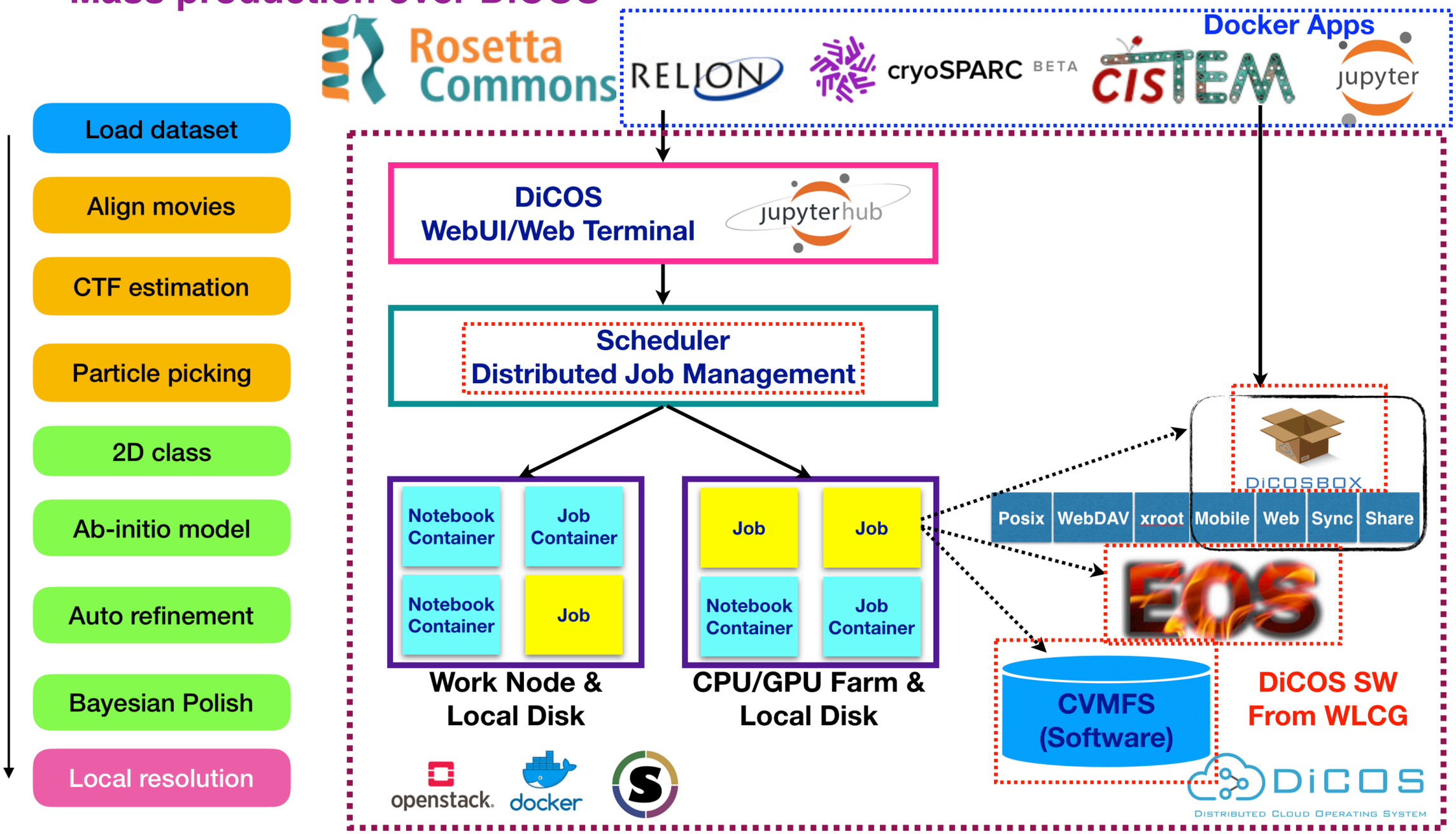






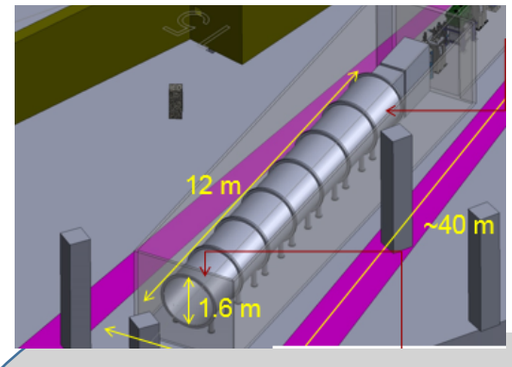
Supporting Cryo-EM Applications by DiCOS

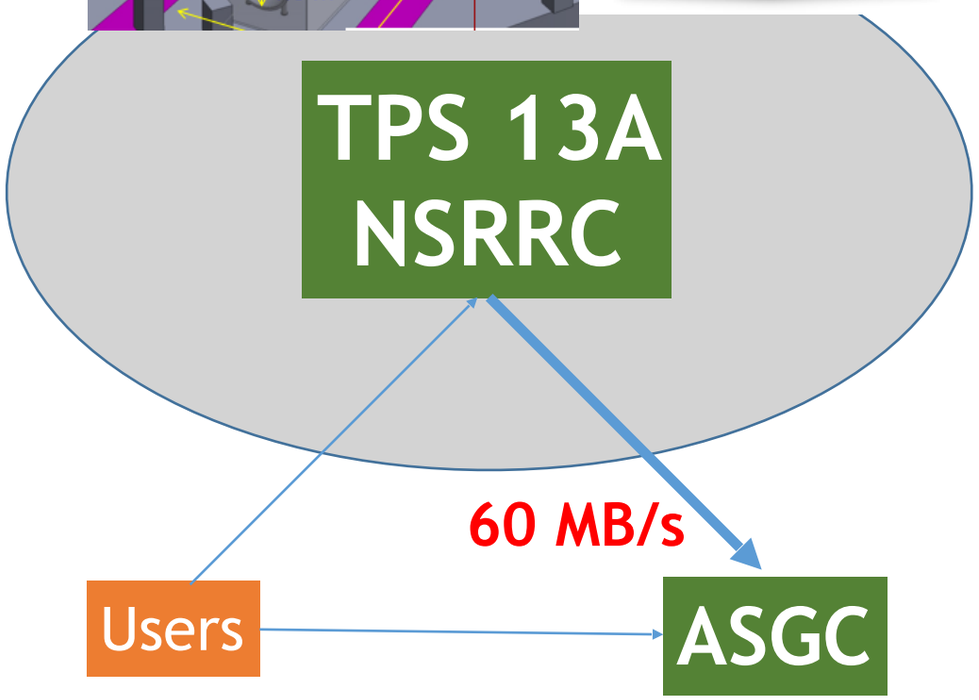
- Primary GPU (single precision) and CPU (300-600 threads) users, O(TB) input/job
- Customization and development
  - Web UI
  - SW package as container
  - Data flow and performance optimization
  - Mass production over DiCOS

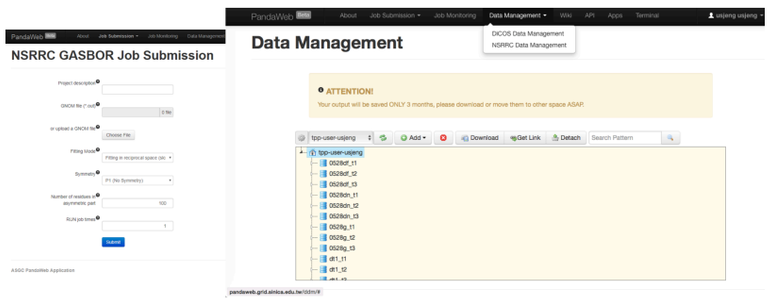


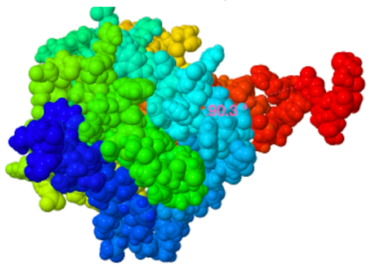
Scalable and Efficient BioSAXS Data Analysis

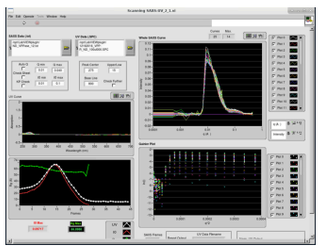
Extending DiCOS capability for protein data analysis and collaborations between NSRRC and AS











BioSAXS: Biological Small Angle X-ray Scattering

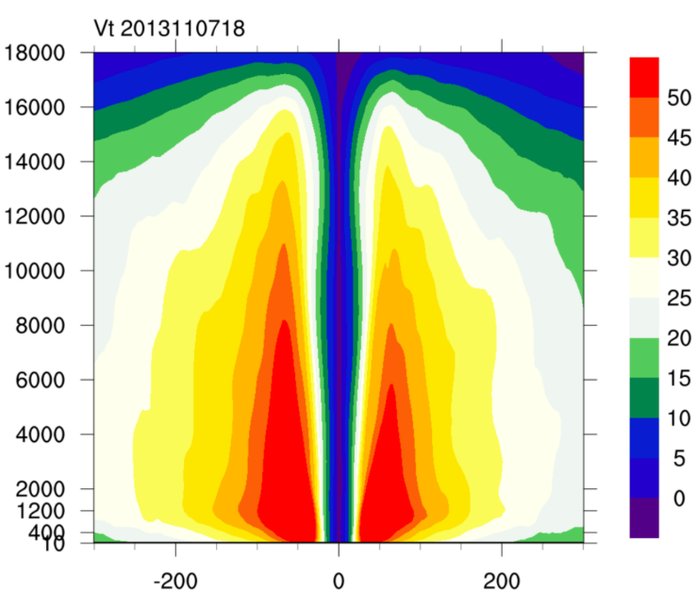
Goal: Biological structures and structural kinetics in atomic-to-micrometer length scales and in μs – min time resolution

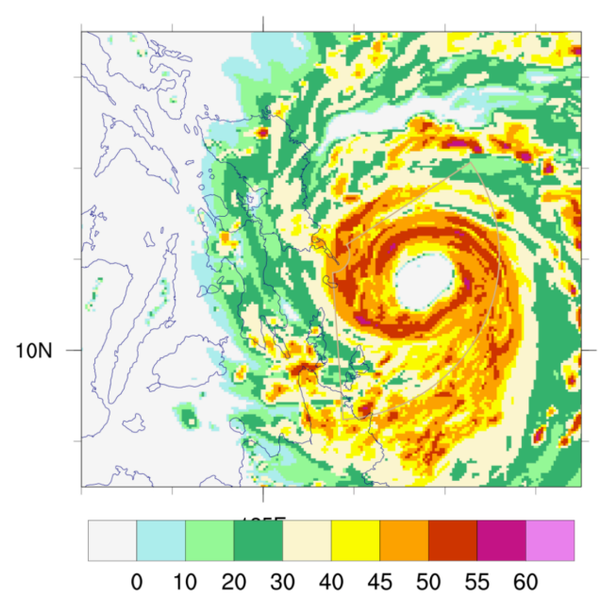
- **Main tasks: data synchronization, computing model implementation, software package integration, workflow validation, efficiency optimization, Web Portal**
- **100+ domestic users**
- **Daily data production: ~240 GB (6 MB/s)**
- **1000 CPU-cores, 100TB storage are available**
- **Single Sign-On enabled for TPS 13A users**
- **Users can login to ASGC to finish analysis by DiCOS or download data to their own PC**
- **Both WebUI (porting ATSAS software), Command Line UI (Direct access to ATSAS/AMBER/Rosetta, etc), Labview Web UI and JupyterLab UI were implemented in DiCOS by ASGC**
- **Data transmission performance needs to be further optimized**

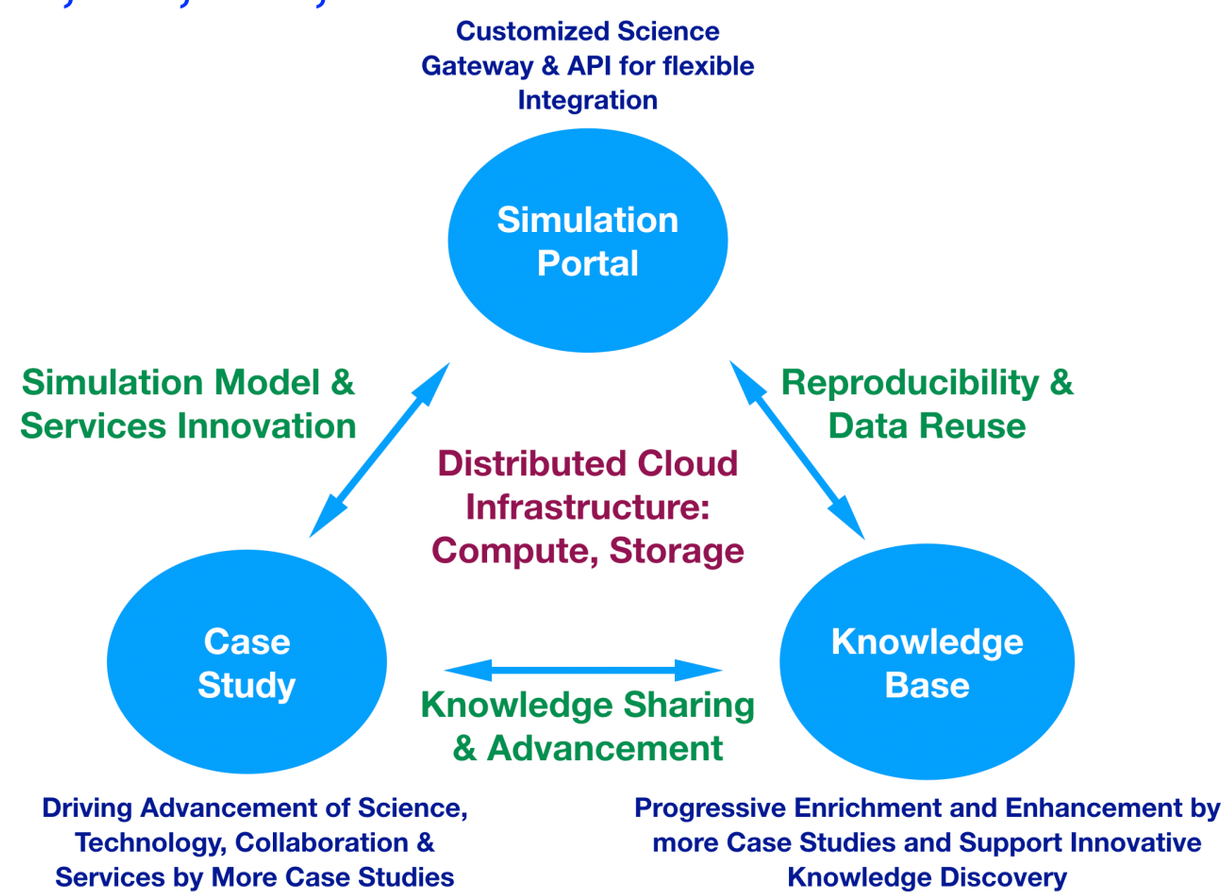
Disaster Mitigation Capacity Building by Deeper Understanding Approach in Asia

- Quantify risks and reduce risk determinants (Vulnerability x Exposure x Hazards) by evolved knowledge of physical processes behind the disaster events
- Improve efficiency of numerical simulations according to the knowledge of root cause and drivers of target hazards
- Case study approach: develop best practices; find out gaps to get insights; discover challenges and further requirements
- Moving towards open science model
- Partners: BD, ID, IN, JP, MM, MY, PH, TH, TW, VN, EGI, APAN

Improved Vertical Wind Field Structure and Eyewall Contraction for Typhoon Haiyan

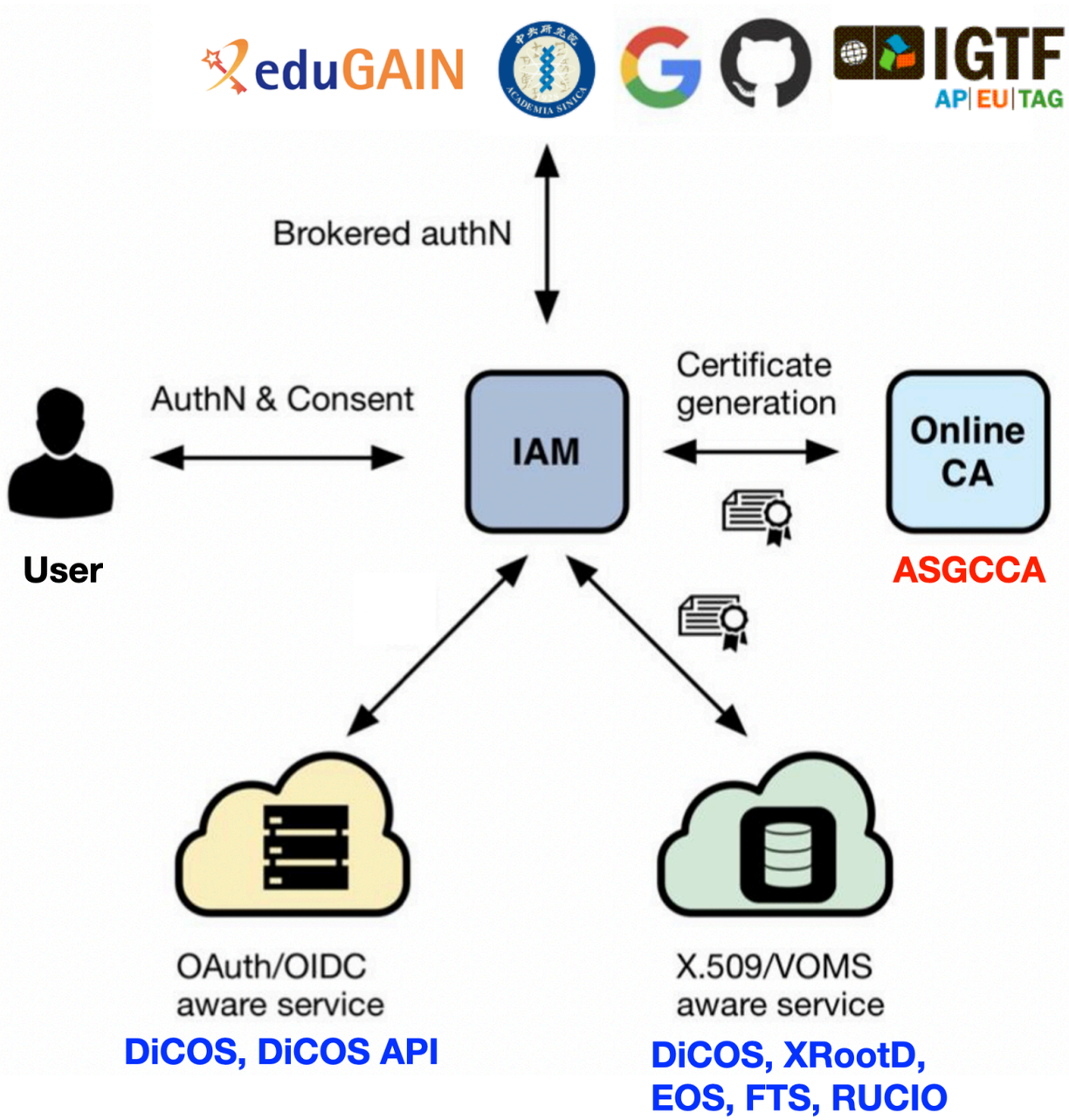






Token-Based AAI

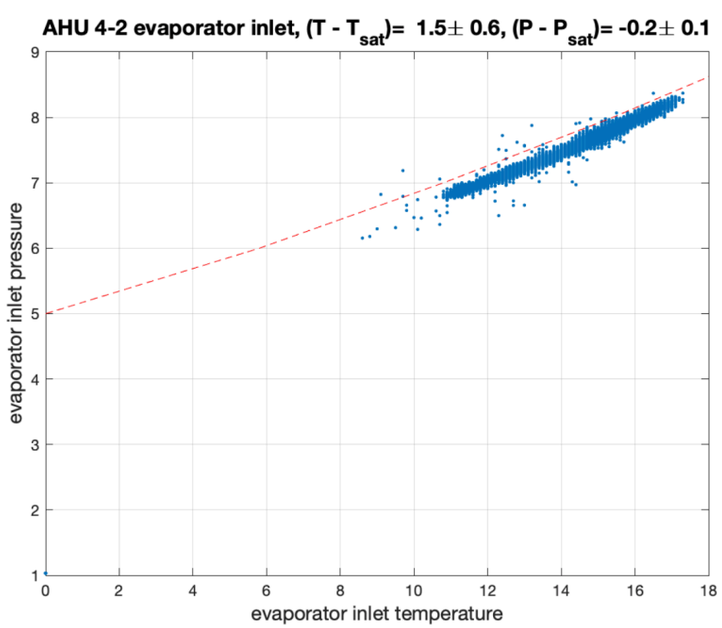
- Basic Services according to community needs
  - User Identity, Community Attribute Services, Access Protocol Translation, Authorisation, and End Services
- User's home institute credential based authentication
  - Academia Sinica SSO, SAML, OpenID Connect (Google, Microsoft, Github) and X.509 certificates
- Supporting community AAI
  - Either local community or international community
  - Trust building between users and service providers
- Supporting federation of infrastructure services: identity management, access, resource, etc.
- Improving AAI services based on user experiences over IAM



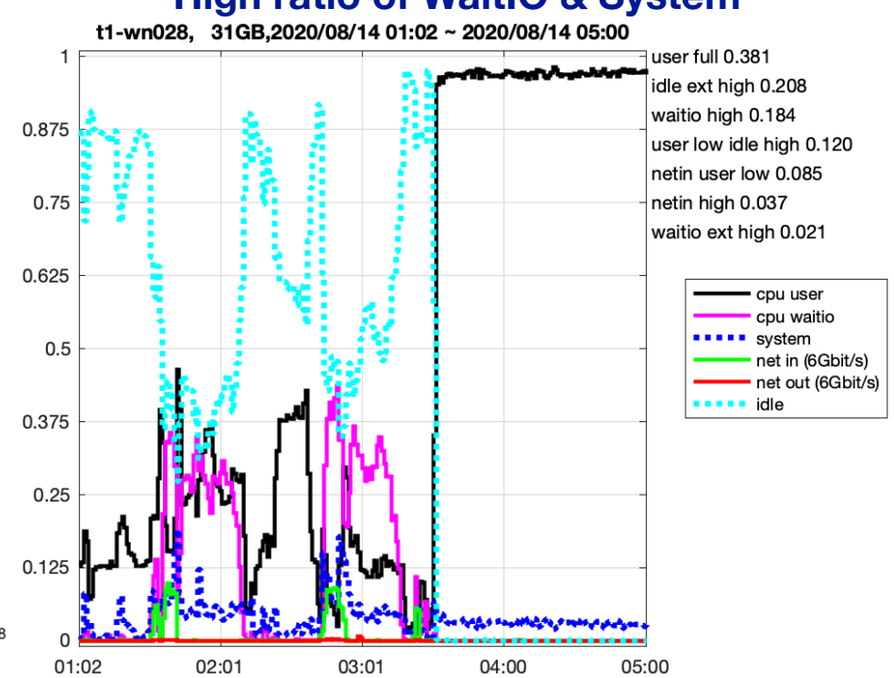
System Efficiency Optimization

- Goals: maximize application performance by available resources dynamically, in terms of power, thermal and system (Comp, Storage, Network, application) efficiency
- Scope: Power, Thermal and Distributed Cloud System management
- Strategy: intelligent monitory and control through ML
- Example: Thermal management, Compute/storage/network anomaly detection, Power saving of worknodes
- AHU monitoring and control
  - Detection of refrigerant operating issues and abnormal components; Efficiency optimization
  - 13 sensors; 18K data points/day;
  - Realtime monitoring, adjustment and diagnostics
- 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

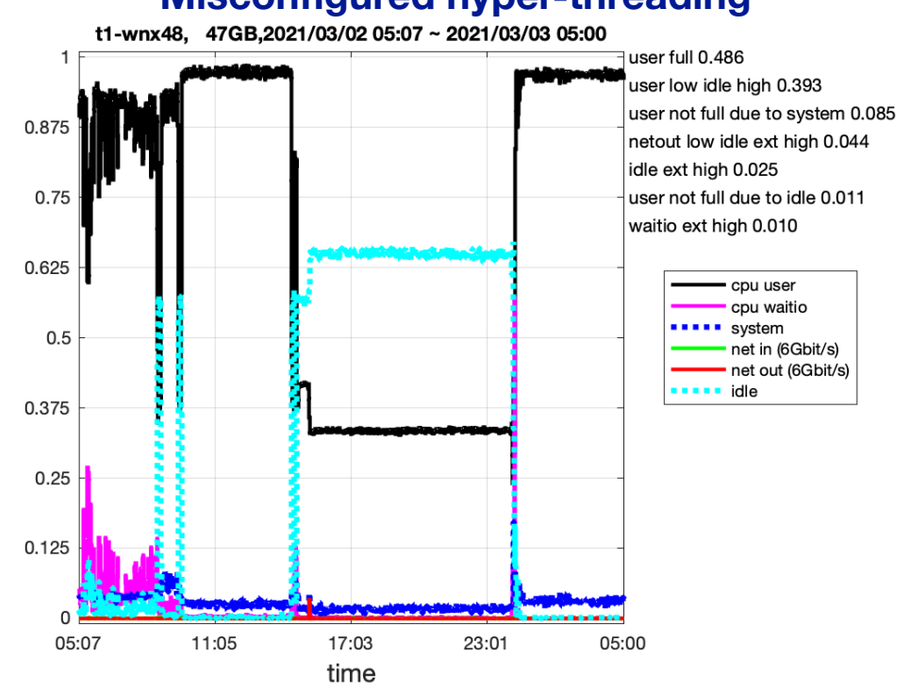
AHU Performance Monitoring



Worknode Monitoring: High ratio of WaitIO & System



Worknode Monitoring: Misconfigured hyper-threading



Summary

- ASGC is supporting multiple O(PB) scale research applications by the common distributed cloud research infrastructure based on WLCG technologies and collaborations
- The scale, functionality and technology of RI would be advanced along with the extension of scientific applications, based on user community requirements mainly
- The RI is also supporting regional collaborations and extending e-Science for the masses
- System efficiency based on intelligent monitoring and control is one of our strategic focus.
- Moving towards open science: Open source, reproducibility, FAIR-based open sharing and open collaboration