ASGC Site Report

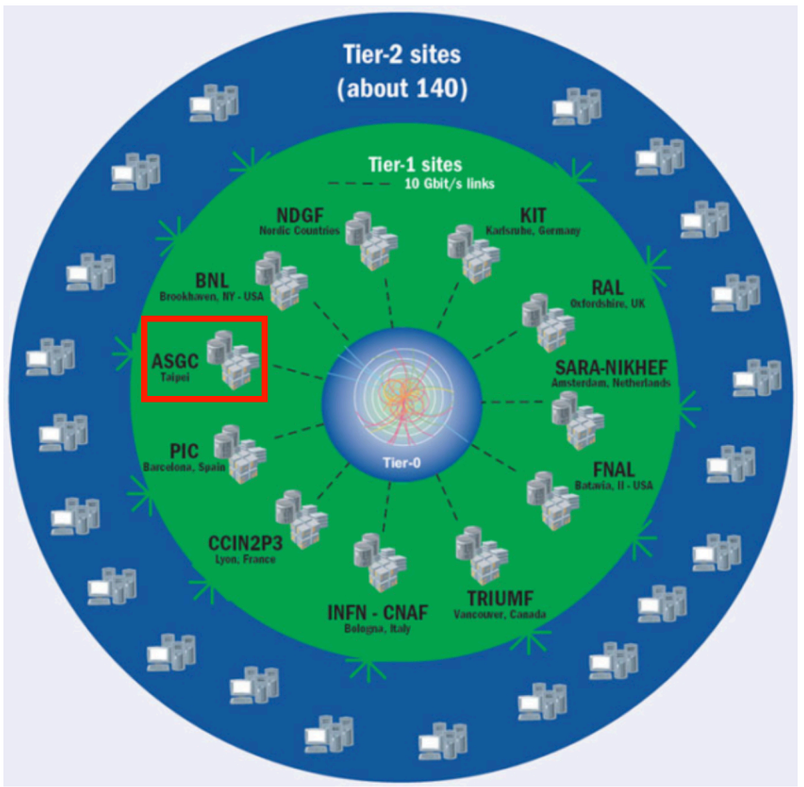
Eric Yen and Felix Lee

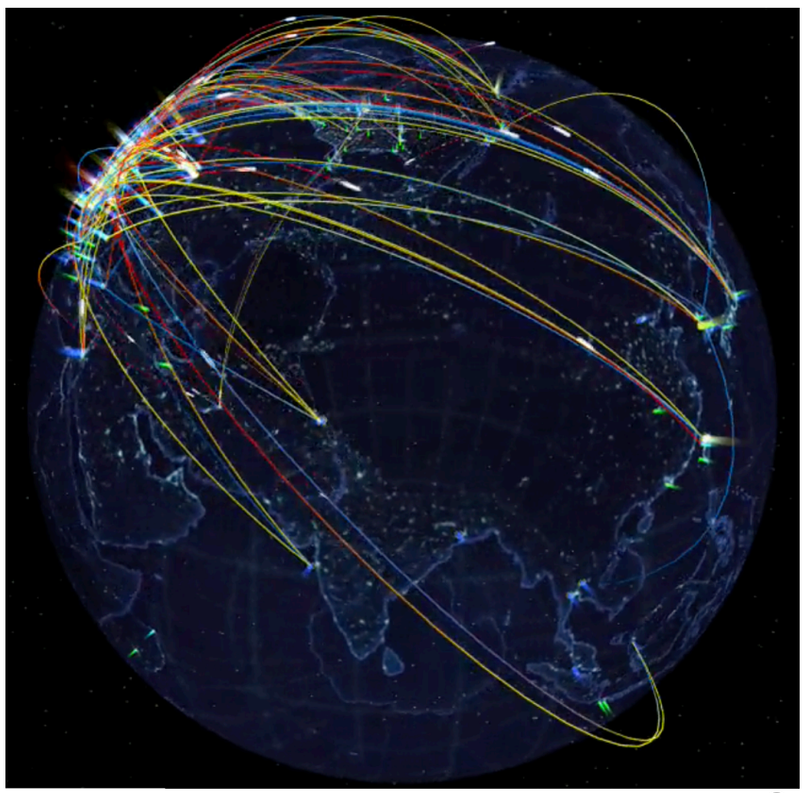
Academia Sinica Grid Computing Centre (ASGC)
Taiwan

Asia Tier Centre Forum (ATCF6) 2022
Krabi, Thailand
22 Nov. 2022

ASGC Introduction

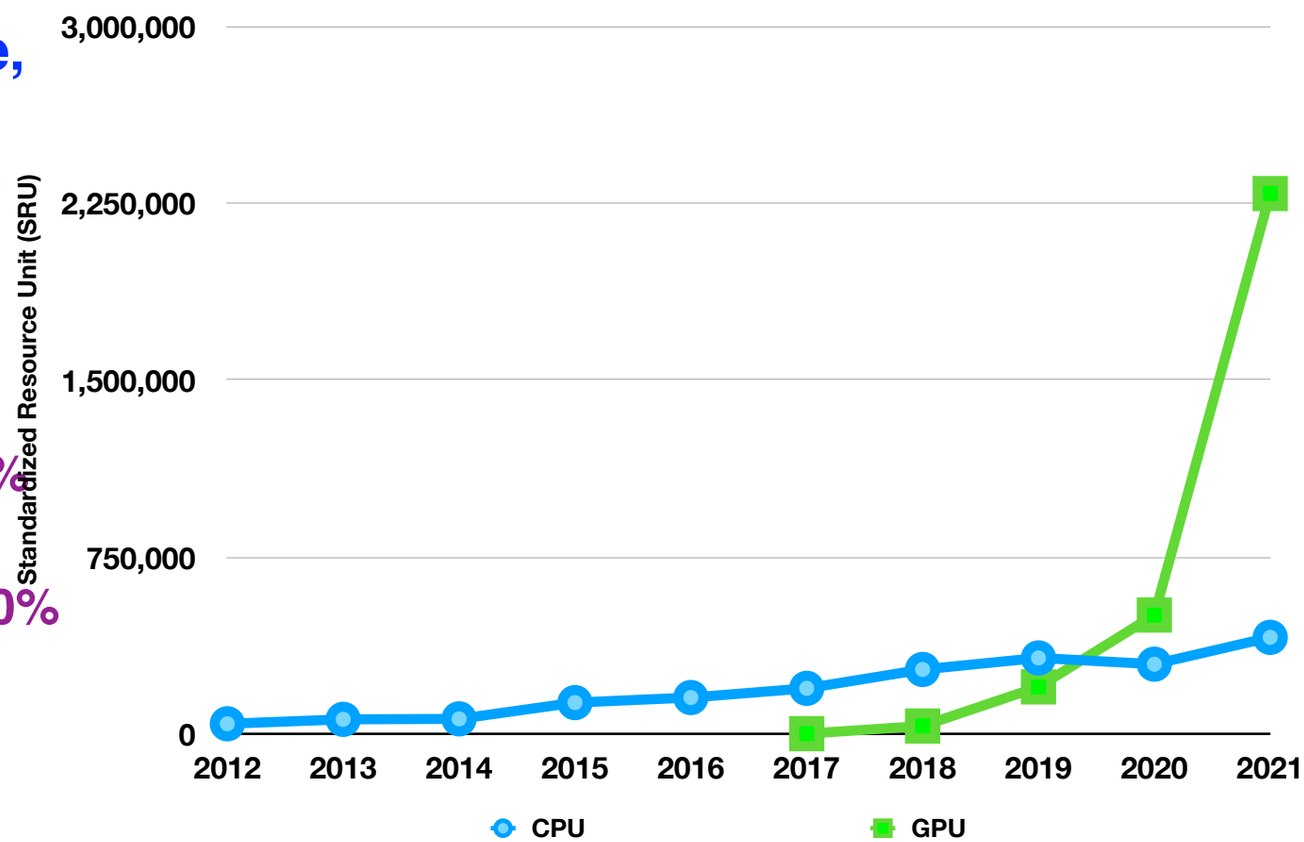
- ASGC joined WLCG development and deployment for the Large Hadron Collider grand challenges since 2001
  - ASGC T1 and WLCG Asian Regional Operation Centre has been operational from 2005
  - Migrating to T2s for ATLAS and CMS (effective from Oct. 2023)
- ASGC has been supporting multi-disciplinary e-Science 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
  - Domains: particle physics (ATLAS, CMS, AMS, KAGRA/LIGO, TEXONO, ICECUBE), physics, Bio-imaging (CryoEM), bioinformatics, next-generation sequencing, earth science, environmental changes, biodiversity and ecology, astronomy & astro-physics, molecular dynamics, remote sensing, information science, humanities and social sciences, etc.
  - 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 - 2021): 15% CAGR
  - Growth of GPU utilization (2017 - 2021): 320% 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)



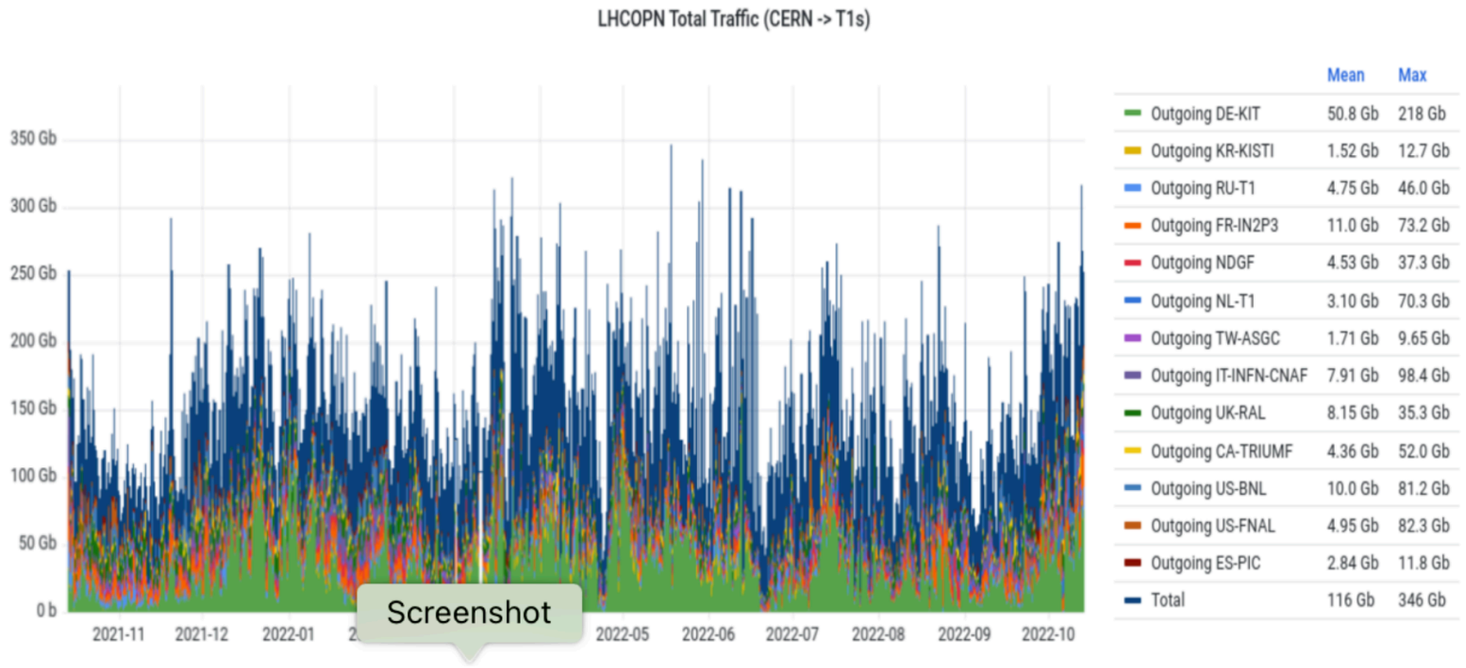
SRU is normalized computing resource unit based on CPU or GPU performance in Linpack.

WLCG Tier-1/2 @ASGC (I)

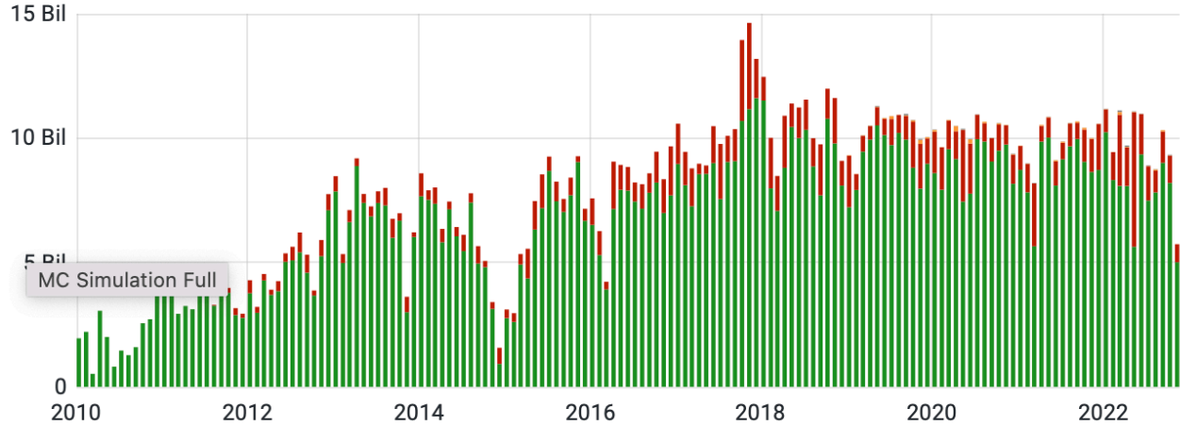
- ATLAS in Taiwan
  - Achievements: Higgs boson; Dark matter; Searches for beyond Standard Model
  - Future plan: H—>bb; Di-Higgs; Dark matter
- Computing Resource - retirement of legacy hardware for energy saving
  - ASGC Tier-1 (2023): 58,760 HEPSpec06 (3,200 CPUCores)
  - Federated Taiwan Tier-2 (2023): 10,896 HEPSpec06 (1,536 CPUCores)
  - 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
- 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

WLCG Tier-1/2 @ASGC (II)

- **CMS**
  - **CMS T3 had been setup @ASGC**
    - **> 20,000 HEPSpec06 (768 AMD 7H12 CPUCores)**
    - **1.7 PB EOS storage: waiting for validation by CMS**
  - **Will be a (federated) CMS T2 of Taiwan**
  - **Should be operational soon (in early 2023 hopefully)**
- **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**

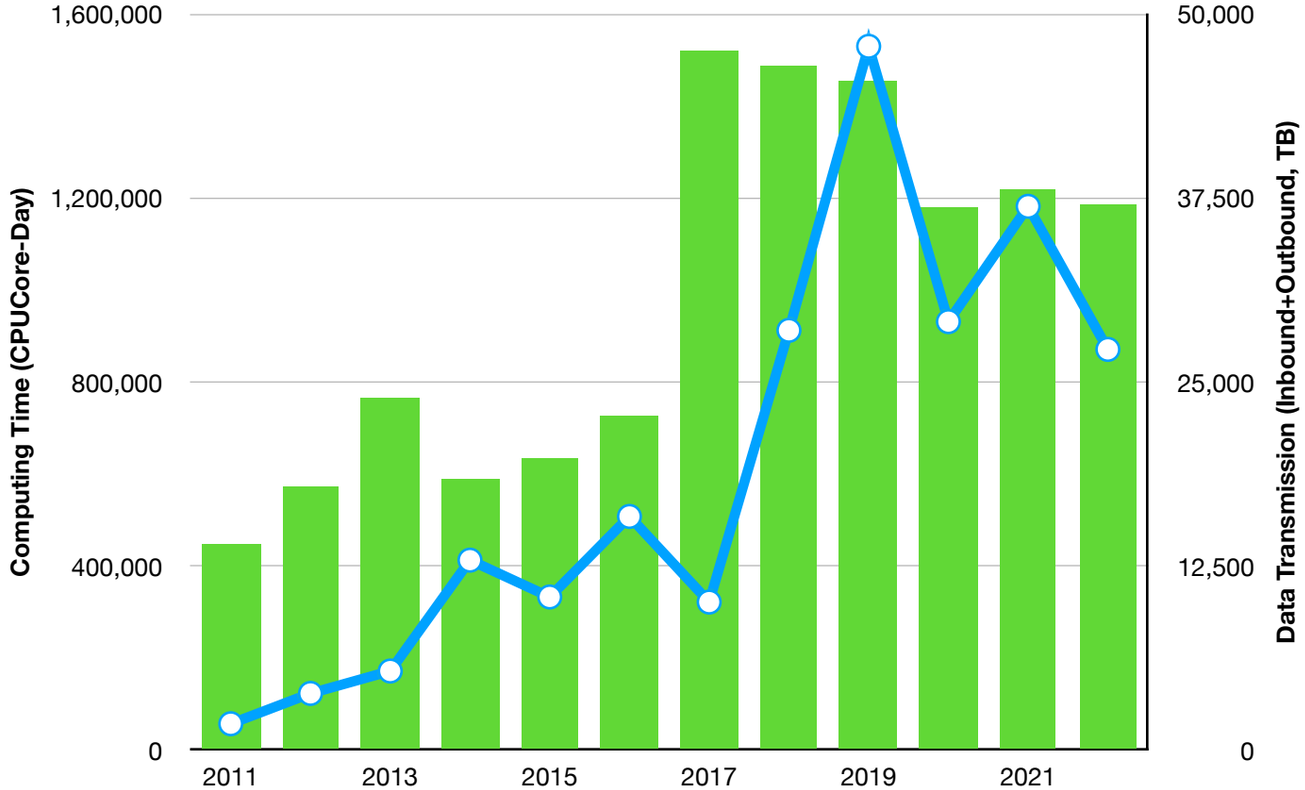


WallClock Consumption of Successful and Failed Jobs - Time Stacked Bar Graph

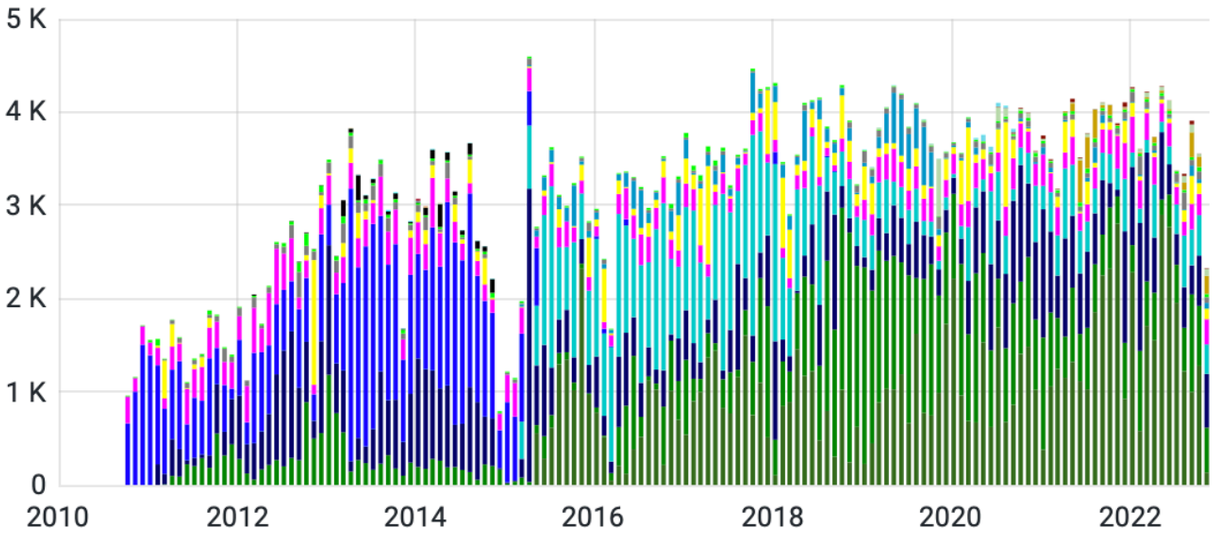


	min	max	avg	total
finished	0	11.6 Bil	6.79 Bil	1.06 Tri
failed	0	5.41 Bil	882 Mil	138 Bil
cancelled	0	211 Mil	12.1 Mil	1.89 Bil
closed	0	59.8 Mil	3.92 Mil	612 Mil

WLCG Activities @ASGC

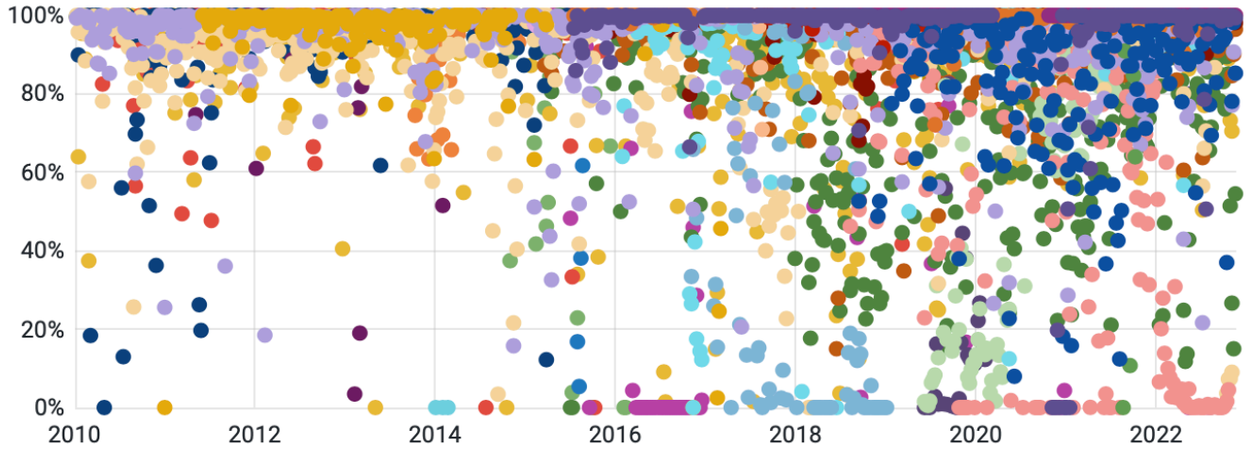


Slots of Running jobs



	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



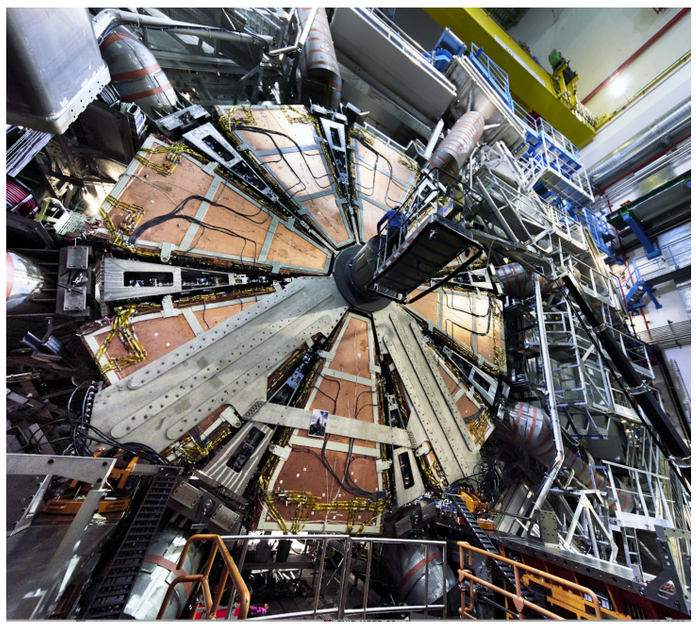
T0 Export	0%	100%	95.8%	100%
Analysis Download	0%	100%	96.2%	99.4%
Data Export Test	97.4%	99.6%	98.5%	98.9%
Production Download	68.1%	100%	99.1%	98.1%

Preparation for Run3 and Run4 (HL-LHC)

- Taiwan ATLAS teams are busy with getting ready for Run3 and preparing for Run4

- **Run3:**
  - Phase-1 upgrade : Trigger router and transceiver for Muon New Small Wheel
  - Monitoring in the ATLAS Control Room

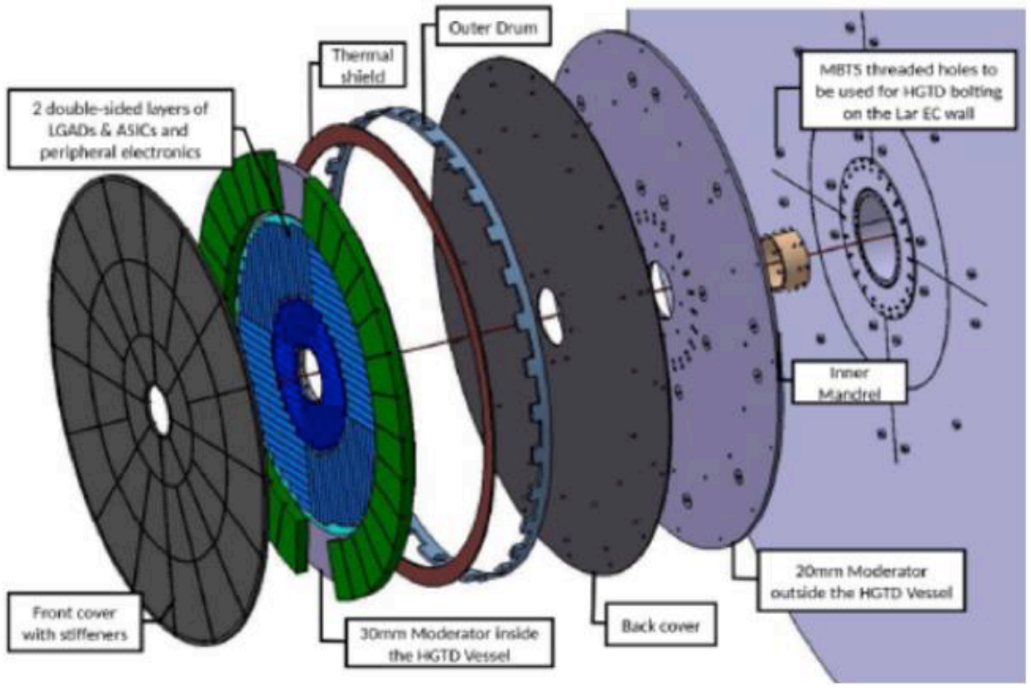
- **Run4:**
  - Phase-2 upgrade :
    - Front end Mezzanine card and fiber cables for Muon Drift Tube
    - High Granularity Timing Detector:
      - Fiber cables, DC/HV supply fan-out
      - Production database and monitoring



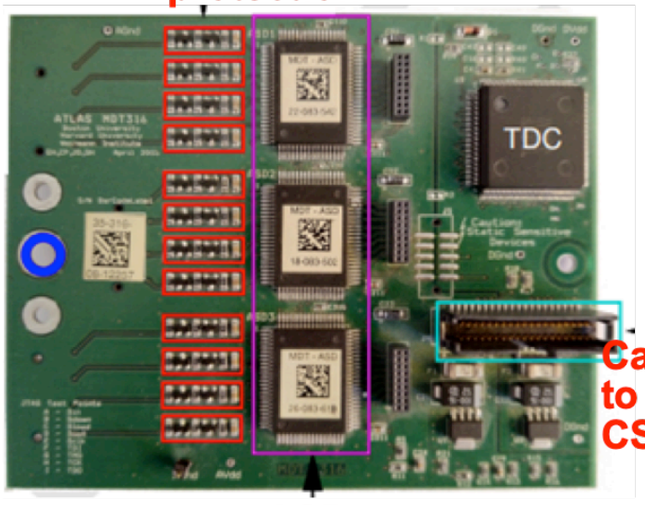
NSW Router module

Trigger Router

New Small Wheel



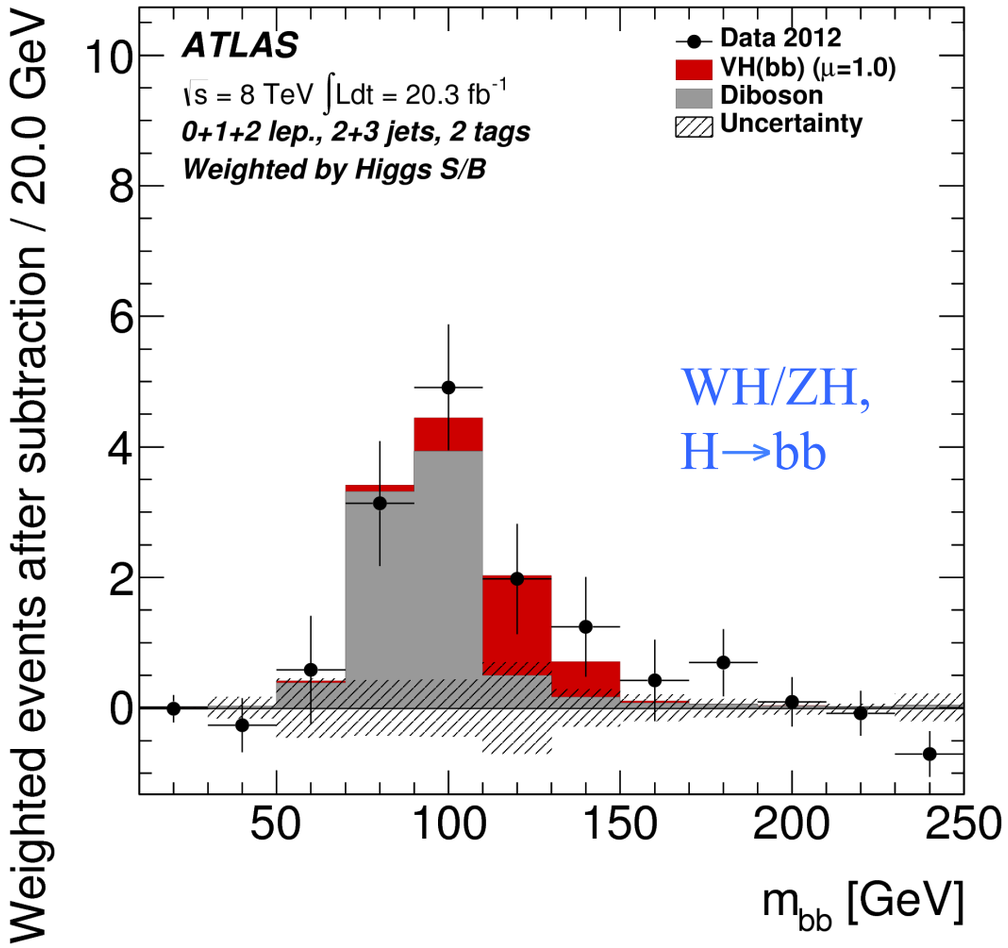
High Granularity Timing Detector



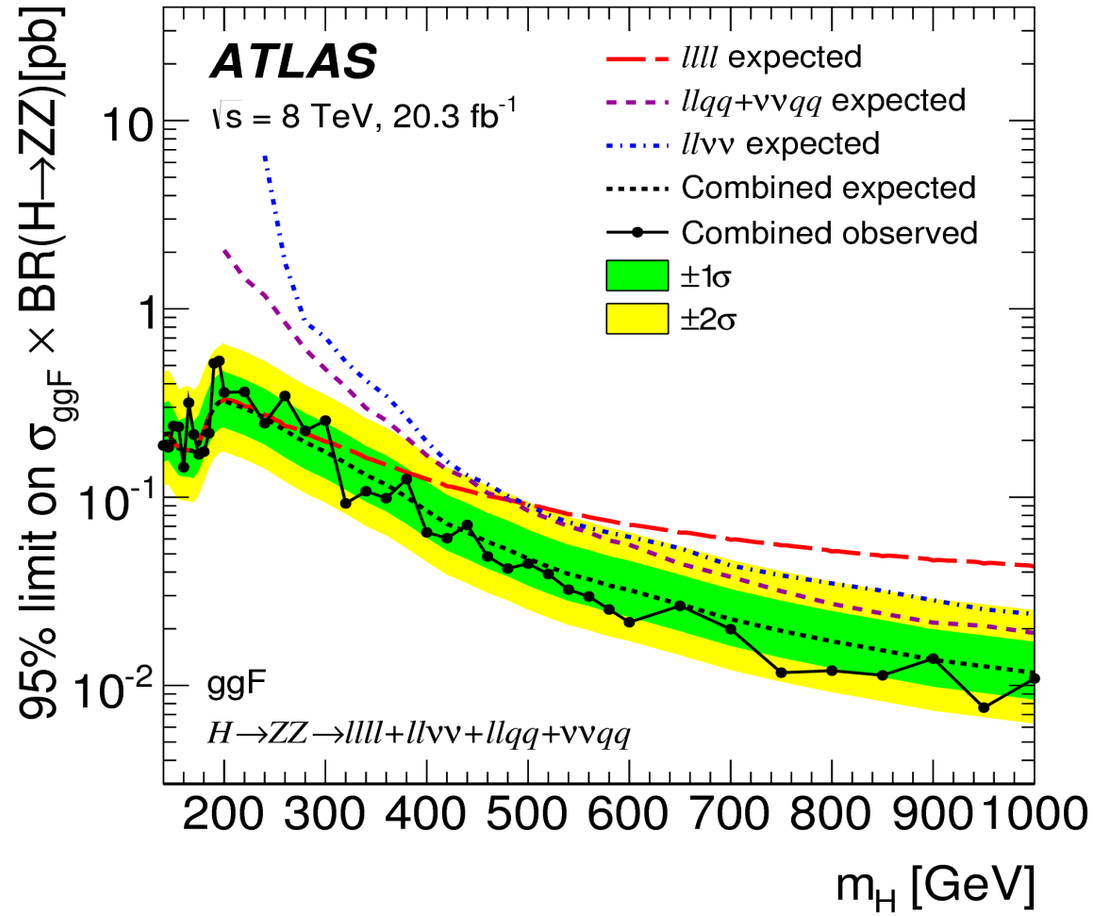
MDT Mezzanine card

Taiwan's Contributions to the Higgs Discovery and Measurements in Run-1 (2011-2012)

- Higgs searches :
  - H→ττ (evidence at 4.5σ, m(ττ) resolution, background)
  - H→WW (obs. at 6.5σ, background estimation)
  - H→bb (main contributor to the ZH→ννbb)
  - Beyond the Standard Model
    - Heavy Higgs, H→ZZ
    - H→invisible particles



Heavy Higgs, H→ZZ

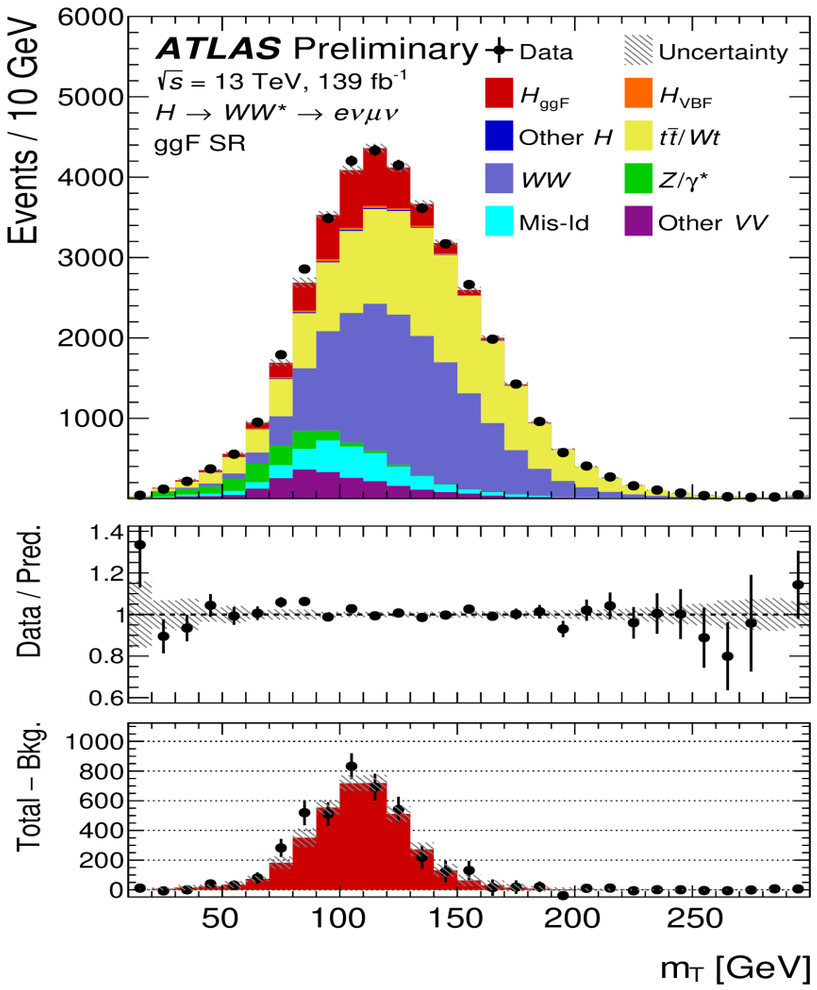


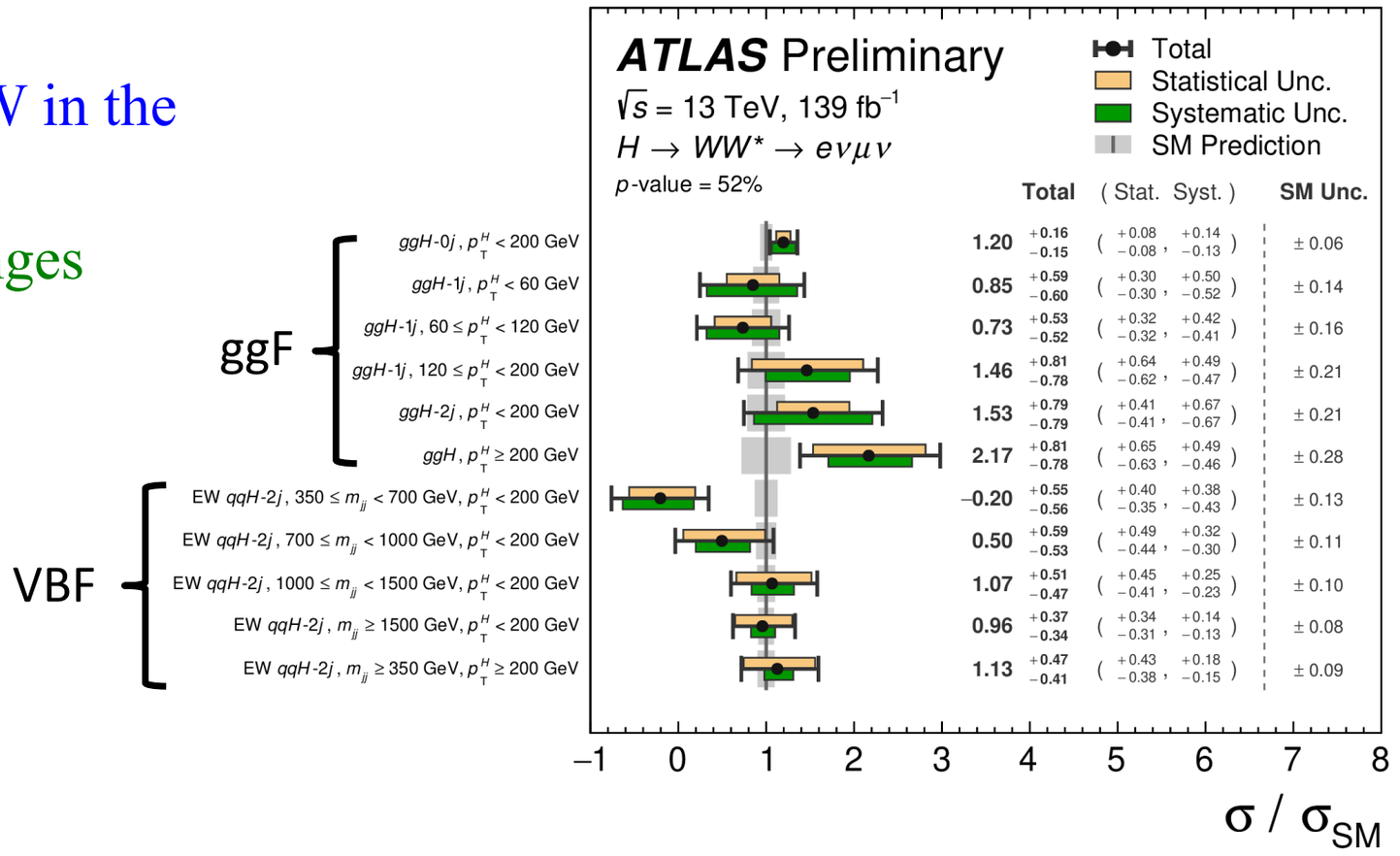
Higgs Boson Measurements

ATLAS-CONF-2021-014

H→WW

- Measure Higgs properties via H→WW in the ggF, VBF production
  - Cross-sections in several pT(H) ranges
- Observed VBF at 6.6σ significance





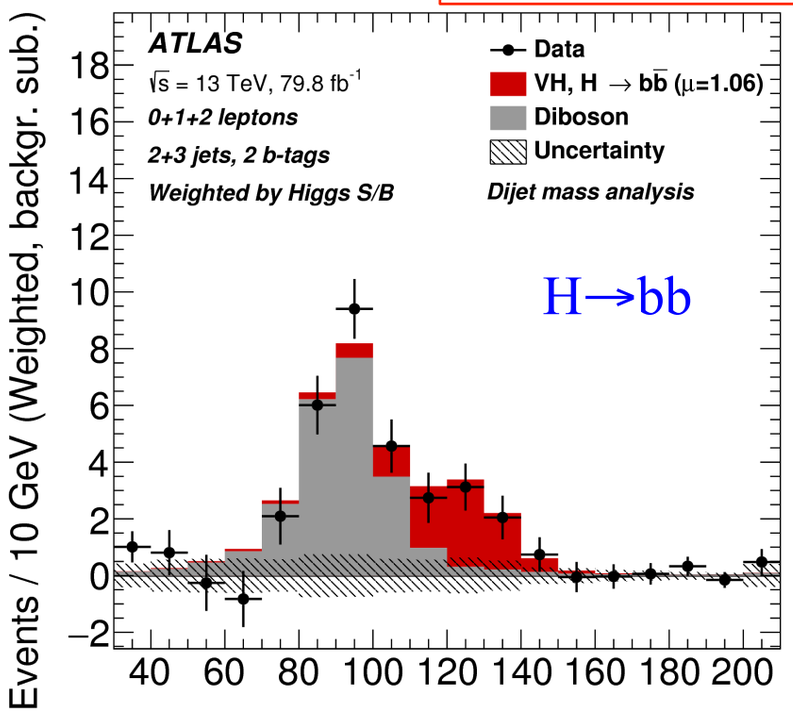
- **NTHU : significant role in this analysis**
- **Yun-Ju Lu** (postdoc) : co-convener of the Higgs HWW working group (2019-2021)
- Results highlighted in ATLAS news briefing :
  - https://atlas.cern/updates/briefing/Higgs-boson-W-boson

Higgs Couplings to 3rd Generation Fermions

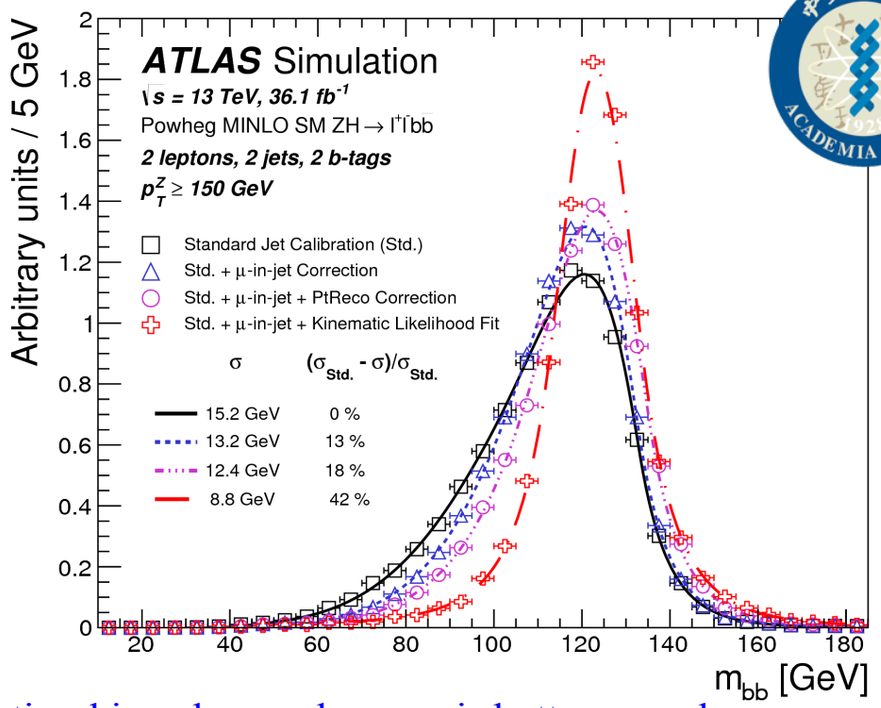
PLB 786 (2018) 59

EPJC 81 (2021)178

- **H→bb :** has largest branching ratio (~58% at m_H=125 GeV)
- ***Considered most difficult decay channel to observe***
- Very large bb pair background ($\sigma_{bb}(QCD) \sim 10^7 \times \sigma(H \to bb)$)
- Concentrate search on WH, ZH production
- Finally observed in 2018 (Run1 + Run2(80 fb^{-1}) data)
  - H→bb at 5.4 σ
  - WH+ZH production at 5.3σ
- Updated analysis with full Run2, also measured σ×BR in several pT(W,Z) bins



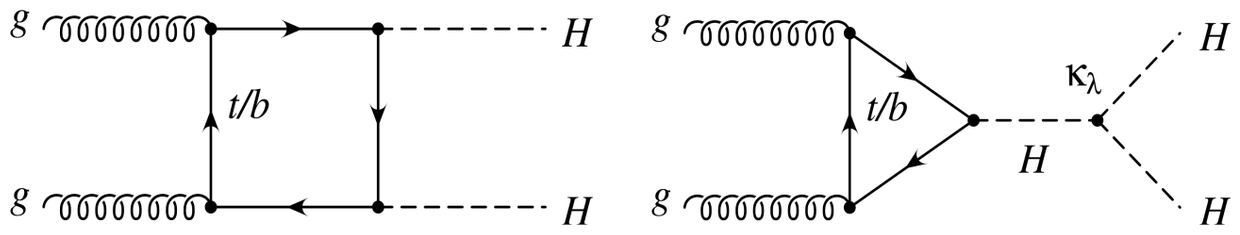
- **Academia Sinica**: major player of this analysis since Run1
- Participated in many aspect of the analysis:
  - Analysis framework development
  - Trigger performance studies
  - Improve m_{bb} mass resolution
  - Main analyzer of the "0-lepton" (ZH→ννbb) channel
- S.M. Wang: co-convener of Higgs Hbb group (2015-2016)
- Observation results highlighted in:
  - CERN press (08/28/18): https://atlas.cern/updates/press-statement/observation-higgs-boson-decay-pair-bottom-quarks
  - CERN COURIER (08/31/18): https://cerncourier.com/a/atlas-observes-higgs-boson-decay-to-b-quarks/



Higgs Boson Pair Production

- Observation of Higgs pair production provides a direct measurement of the Higgs boson self-coupling λ_{HHH} and validate the Higgs mechanism
- HH pair produced predominantly through gluon fusion (ggF)

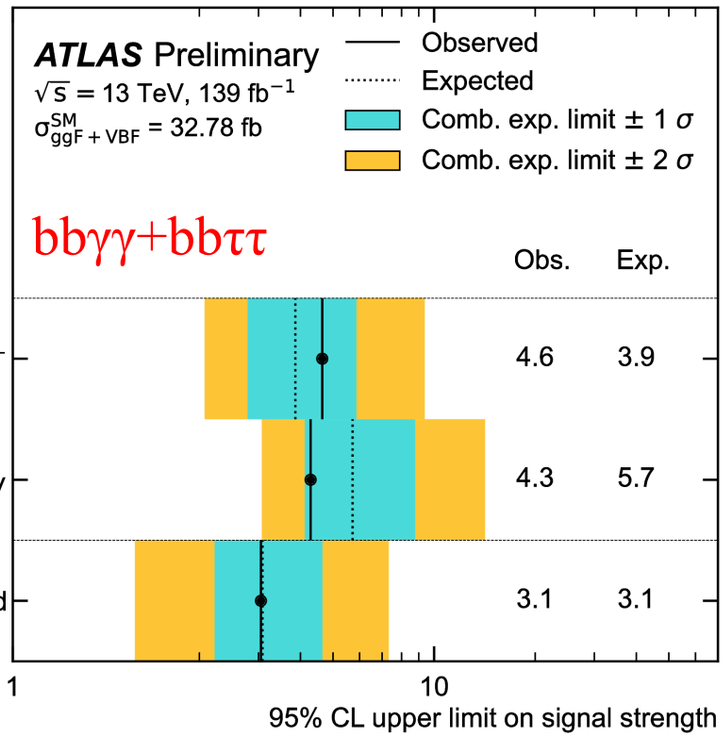
ATLAS-CONF-2021-052



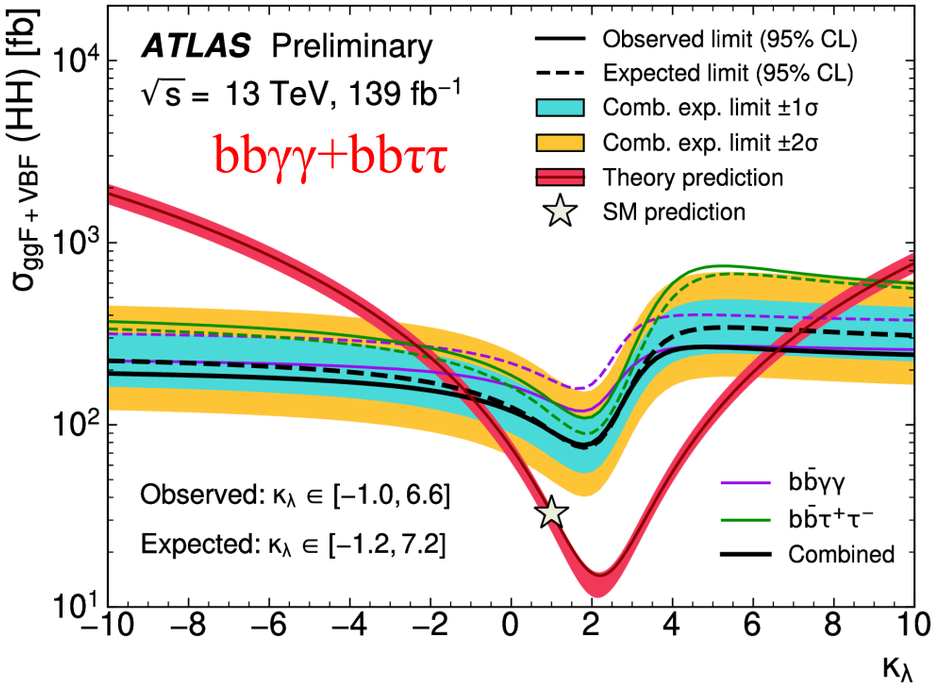
Most sensitive channels :

- HH→bbbb : highest BR, large BG from multi-jets
- HH→bbγγ : clean, but small BR
- HH→bbττ : moderate BG and BR

World best expected limit for single search



- $\sigma(HH)/\sigma^{SM}(HH)<3.1$ (3.1) obs (exp)



$$\kappa_\lambda = \frac{\lambda_{HHH}}{\lambda^{SM}_{HHH}}$$

(Higgs self-coupling modifier)

- Constrained κ_λ : $-1.1<\kappa_\lambda<6.6$

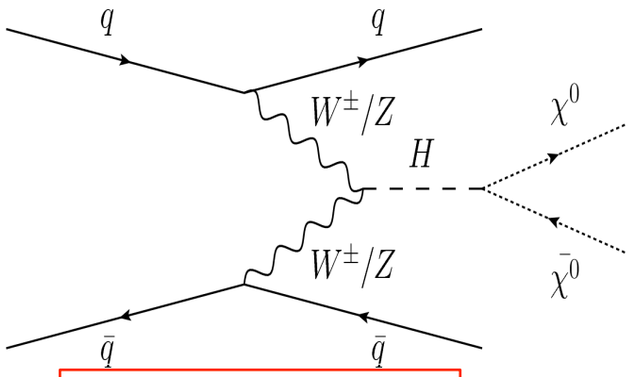
- **Academia Sinica** also contributed significantly in the HH→bbττ search
- Involved in optimization studies, code development, and paper editing
- **Tulin Mete** (postdoc) : co-leading the analysis
- ATLAS briefing : https://atlas.cern/updates/briefing/two-Higgs-better-one

Higgs to Invisible Decays

PRL 122 231801 (2019)

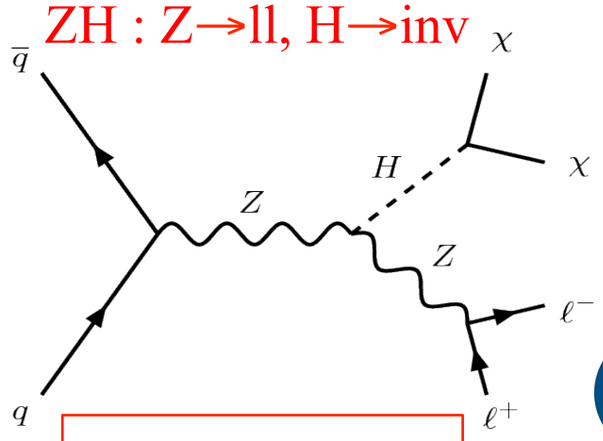
- Some SM extensions predict the Higgs boson acts as a portal between dark sector and the SM sector => Higgs can decay into dark matter particles, if kinematically allowed
- AS searched for H→inv in 2 of 3 searched channels at ATLAS using 36 fb^{-1} Run2 data

VBF : qq→qqH→qq+inv



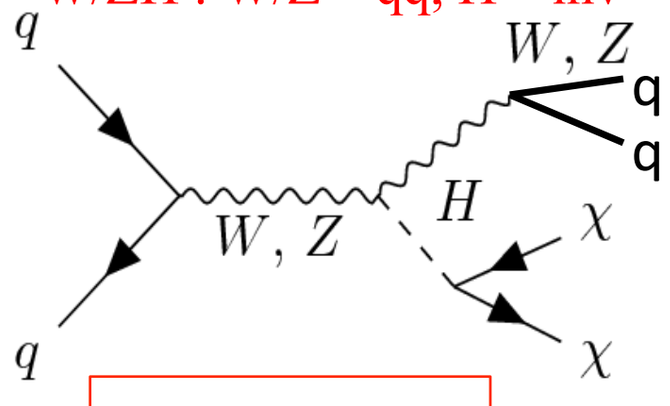
PLB 793 (2019) 499

ZH : Z→ll, H→inv

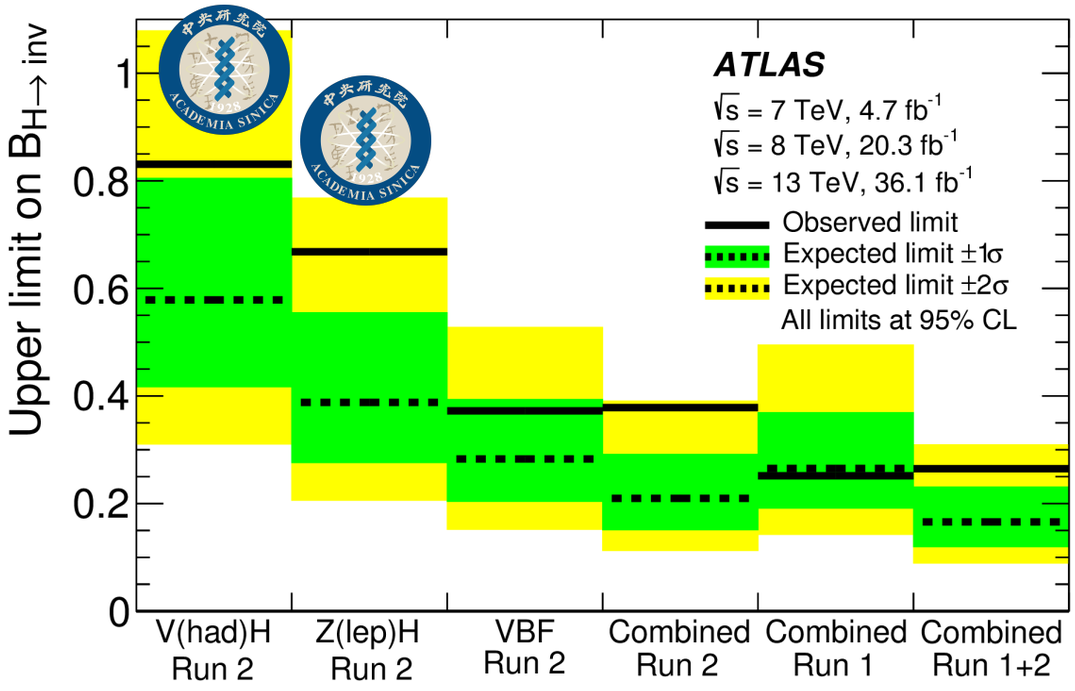


PLB 776 (2018) 318

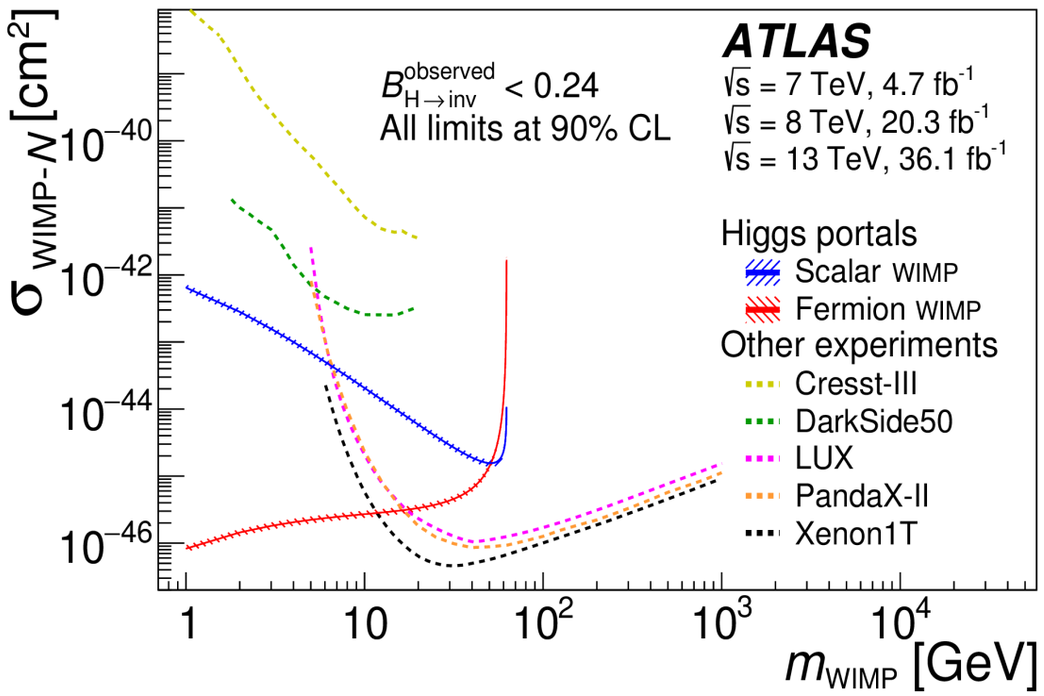
W/ZH : W/Z→qq, H→inv



JHEP 10 (2018) 180



- Combined results from all channels (and with Run1), set limit on BR(H→inv) at 95%CL:
  - BR(H→inv)<0.26 ($0.17^{+0.07}_{0.05}$)

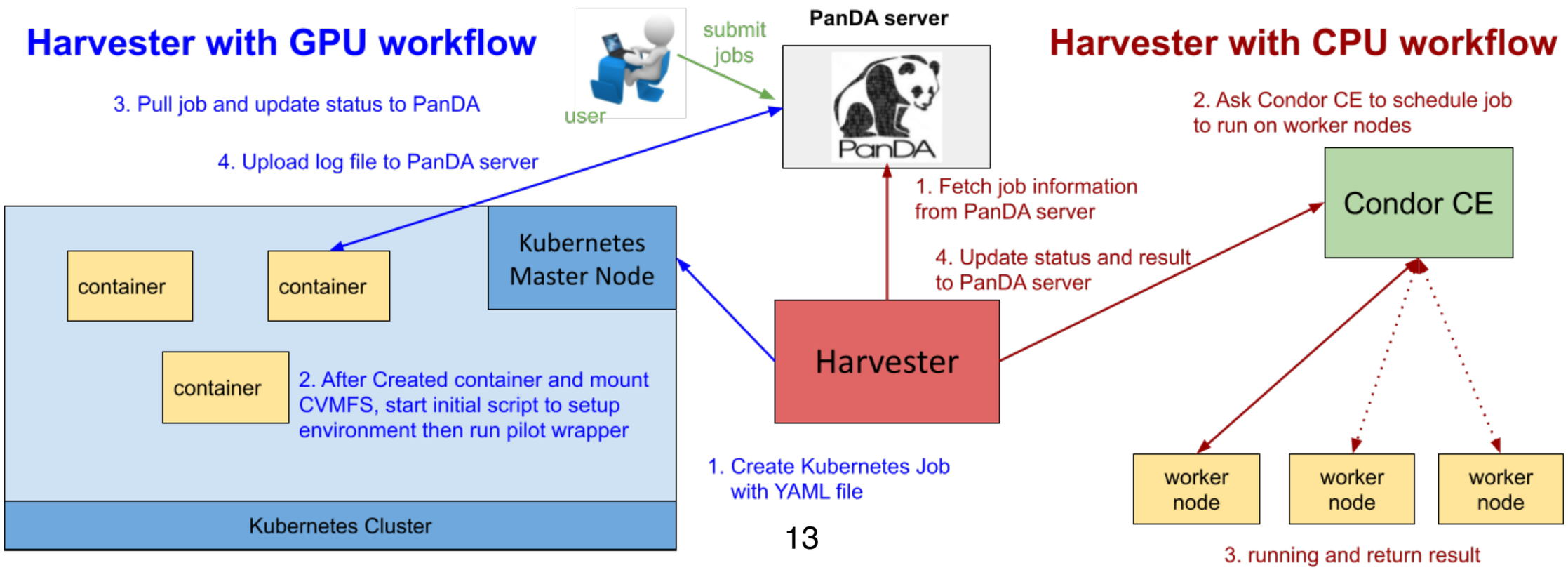


- Set upper limit on WIMP nucleon scattering cross section vs WIMP mass

(WIMP: Weakly Interacting Massive Particle)

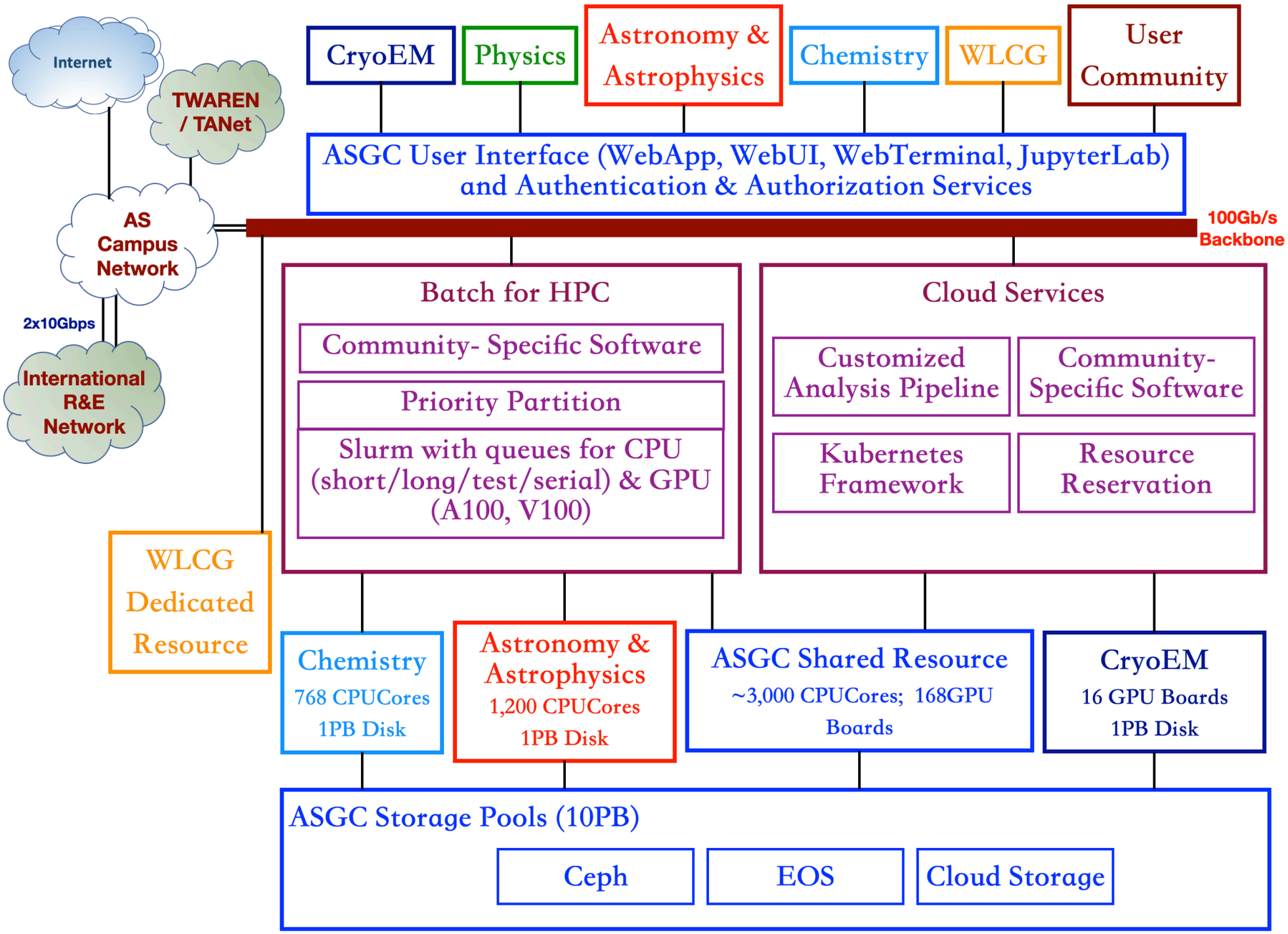
Building Distributed Cloud Infrastructure Supporting Broader Scientific Applications Based on WLCG

- Integrate the whole 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;
- Disciplines: AMS, TEXONO, Gravitational Wave(KAGRA, LIGO, IGWN), NGS, CryoEM, BioSAXS, Drug Discovery, Earth Science, Environmental Changes, Biodiversity and Ecological Monitoring, Lattice Gauge Theory, Condense Matter, proton therapy, and ML/DL applications.



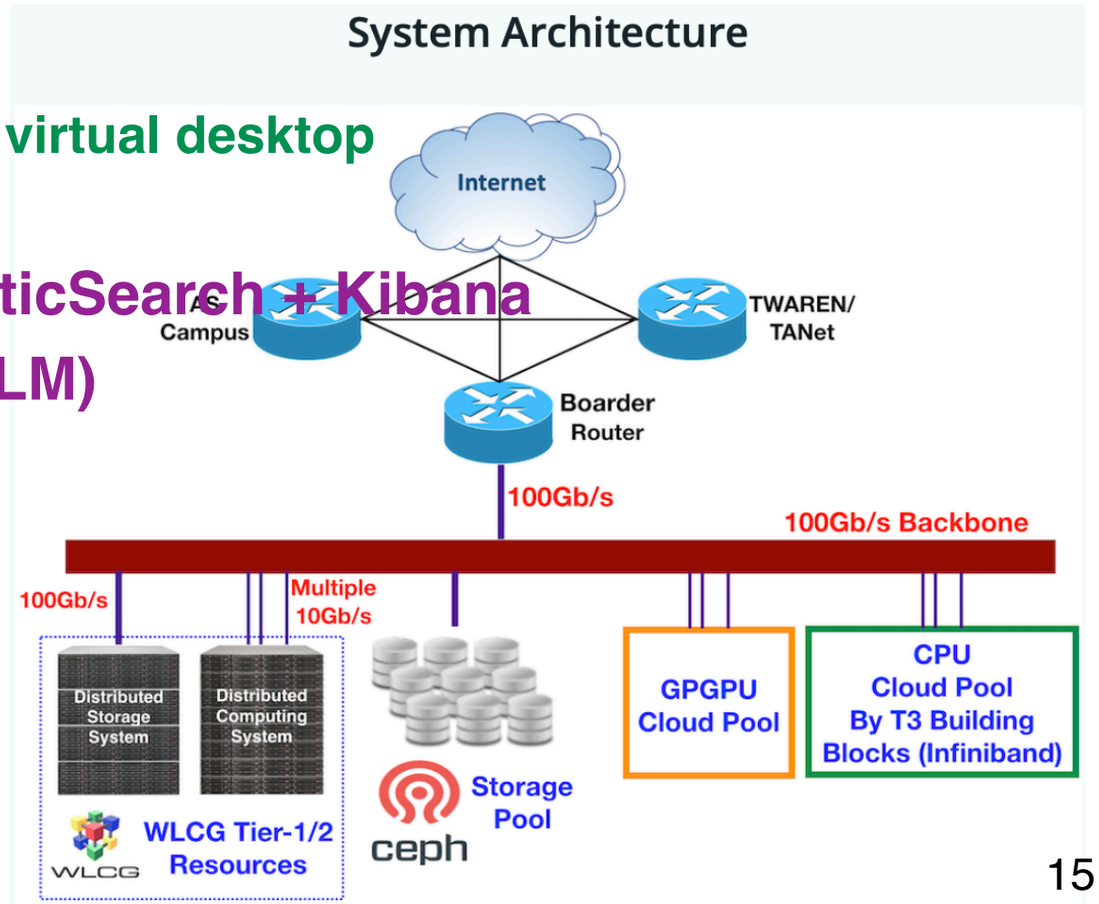
Science Cloud by Federating Cloud and HPC Resources

- **Type of resource**
  - **Core resource: used for providing services**
  - **Shared resource: shared by all users**
  - **Priority resource: invested by user communities. Will be shared to other users whenever facility is free**



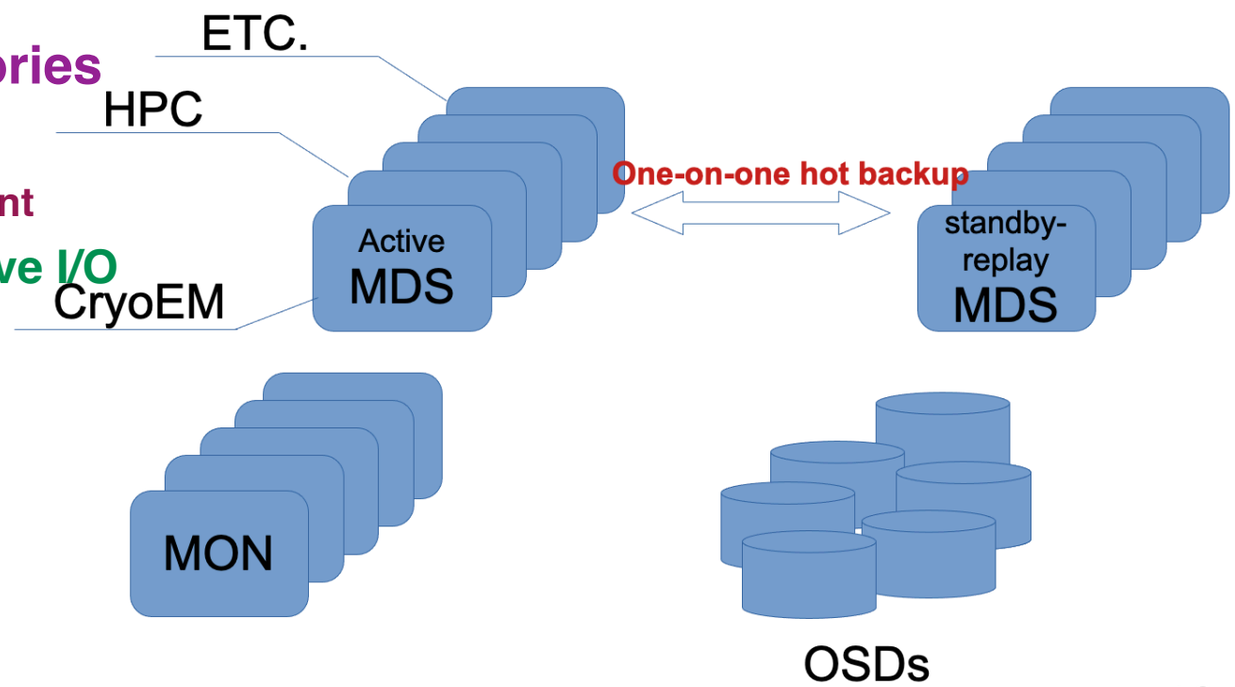
ASGC Science Cloud Infrastructure (DiCOS)

- 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

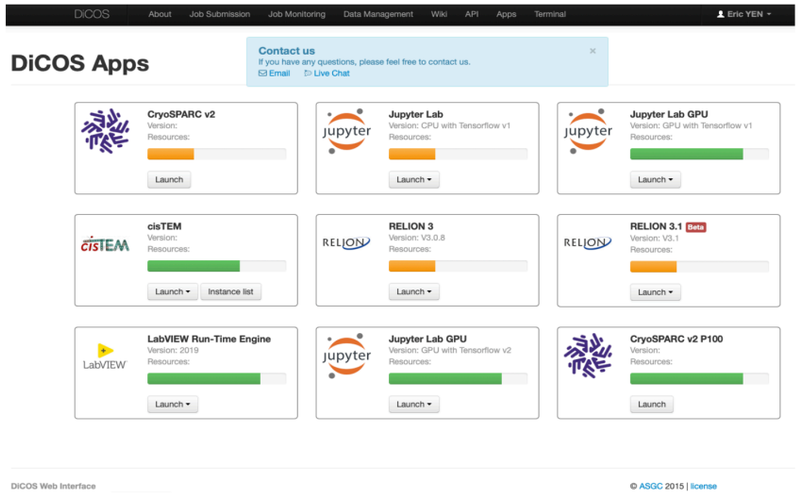


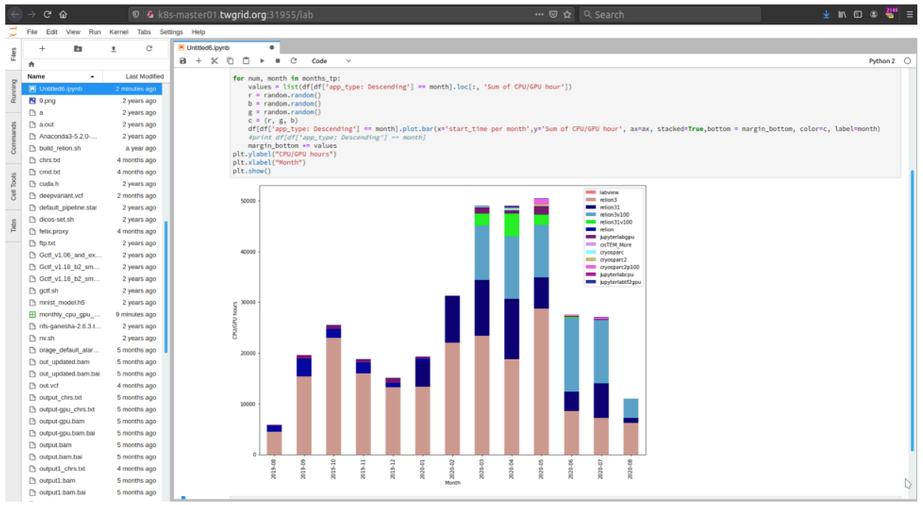
ASGC Science Cloud Storage Architecture

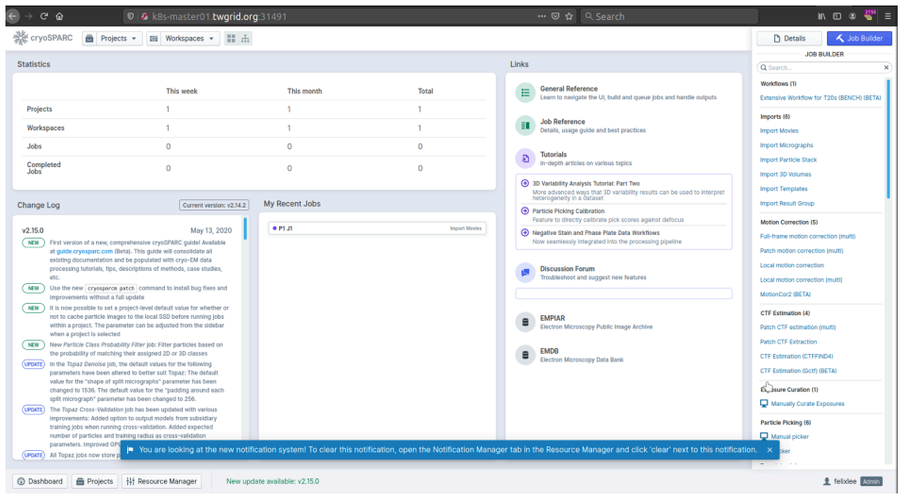
- WLCG Storage System is migrating from DPM to EOS
  - 10PB (2023) after migration
  - Two new EOS system is now under verification by ATLAS and CMS respectively
    - Data has to be moved gradually
- Storage system for other e-Science applications: Ceph-based
  - 8PB (CephFS mainly) in total (1PB used by Openstack Cinder and Glance)
  - Upgrading to AlmaLinux8 and Ceph Pacific by end of 2022
  - Cloud Storage (DropBox-like) services: 2TB/user
    - Accessible from user home directory
- 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
  - 5xMon for continuous growing Ceph cluster
- ML-enabled analysis of Ceph for better reliability and performance is under development

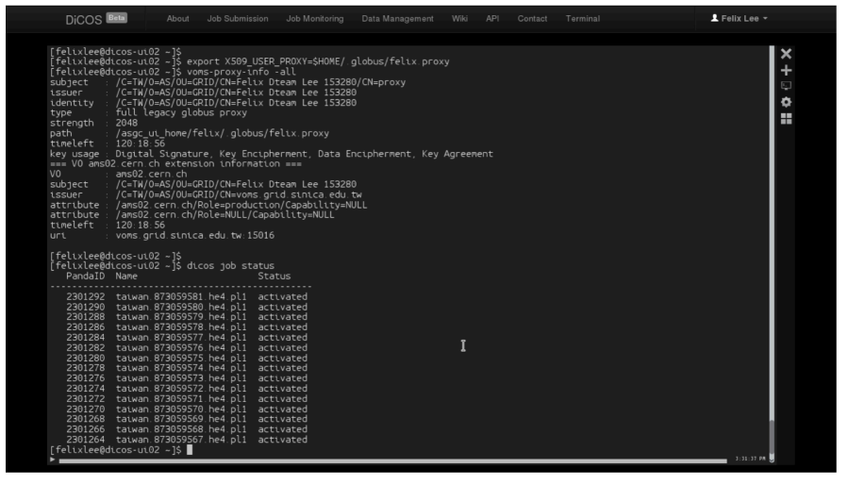


Supporting Big Data & AI in Innovations









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)

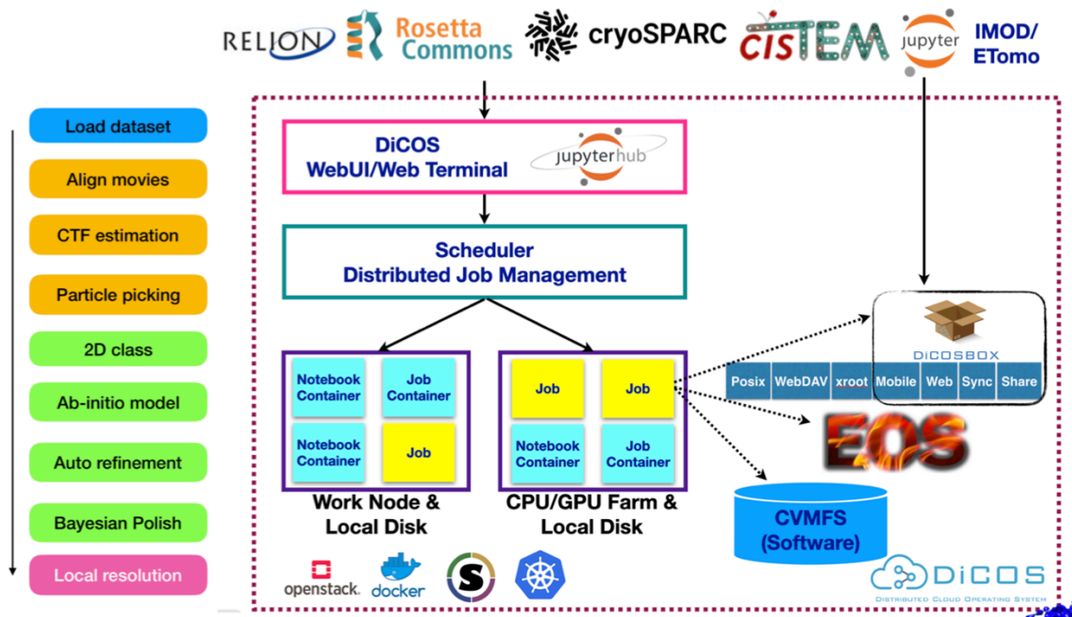
Storage Resource (Ceph/EOS)

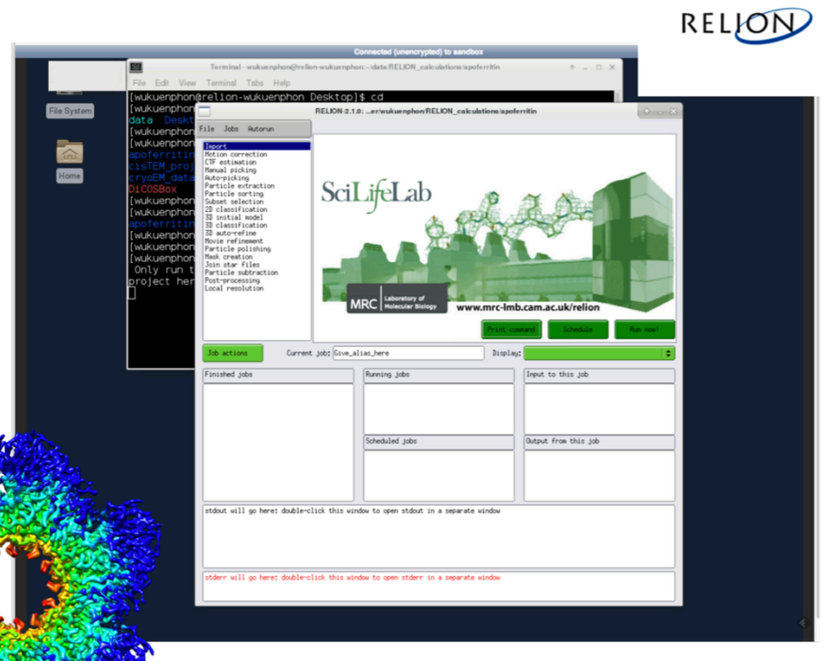
Distributed Data Management & Cloud Storage Services

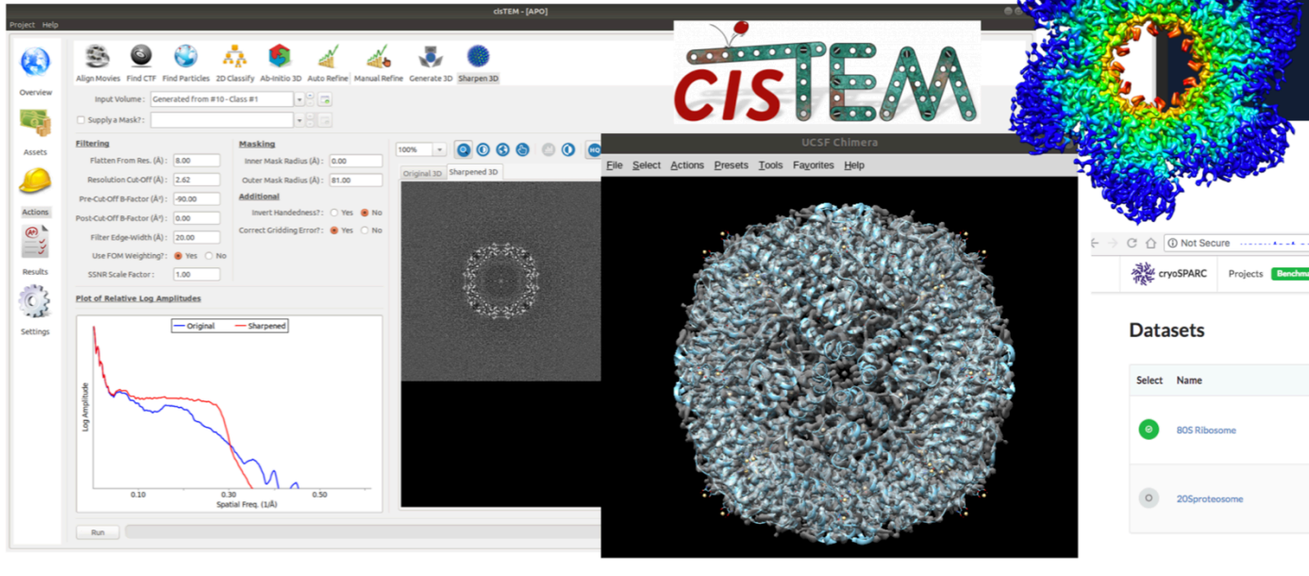
Network & Data Transmission Services

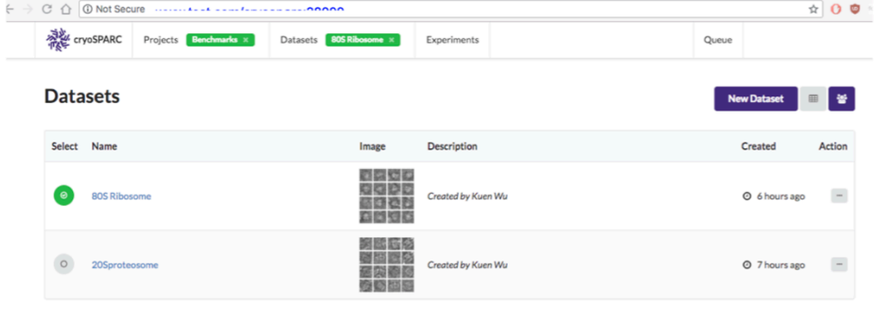
Collaboration With CryoEM Community

- **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**









45 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

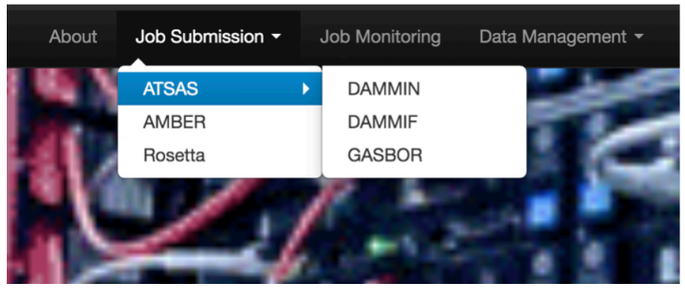
Jupyter Lab gpu 3090
Version: GPU with Tensorflow 3090
Resources: 51%
Launch

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

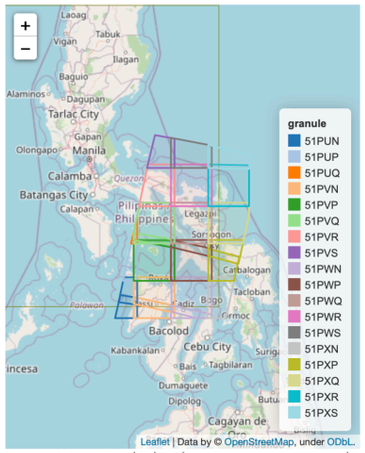
IMOD
Version:
Resources: 66%
Launch

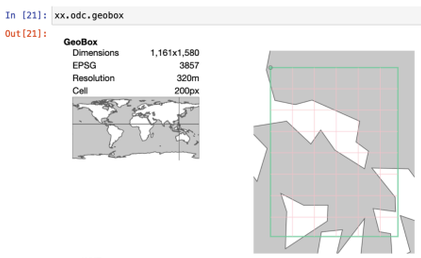
RoseTTAFold
Version:
Resources: 51%
Launch

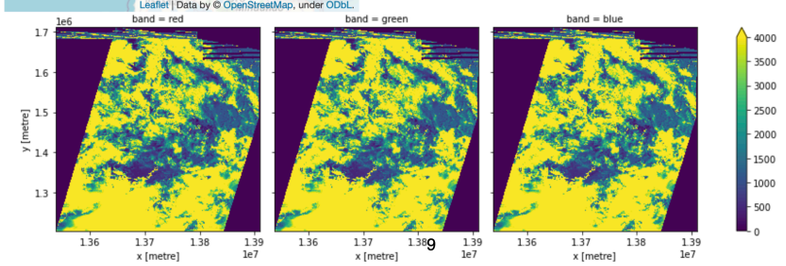
Dynamo
Version:
Resources: 66%
Launch

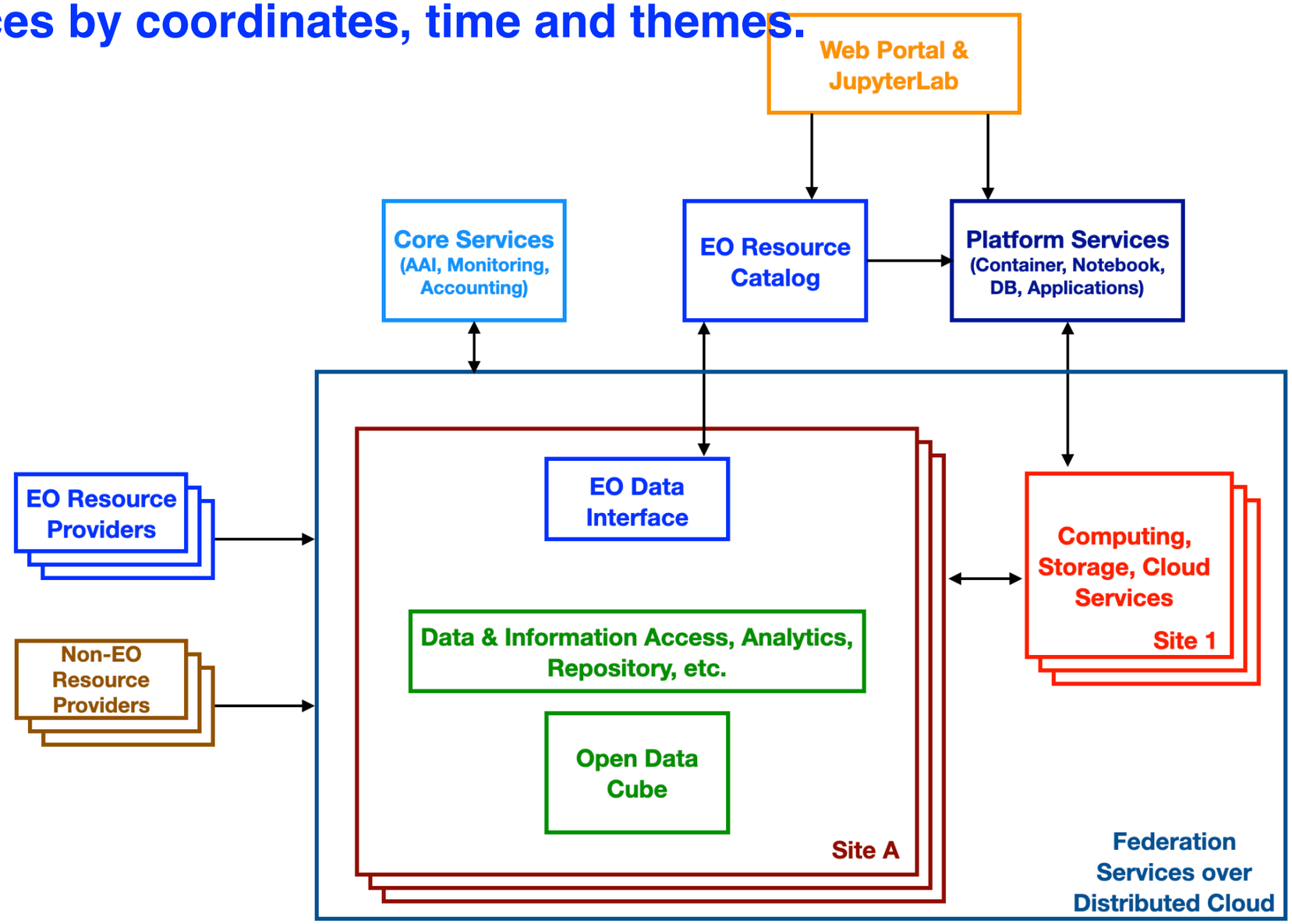
EODC-Enabled Spatiotemporal Information Infrastructure

- **Support gaining insights from big data of heterogeneous resources effectively and leverage the values of big EO data**
  - **Coverage of whole workflow according to applications - including data collection, access, management, simulation, analysis**
  - **Merging insights from ML-enabled BDA and physical intelligence (e.g., domain knowledge)**
- **Earth Observation Data Cube (EODC) serves as an open integration platform of heterogeneous data sources by coordinates, time and themes.**









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

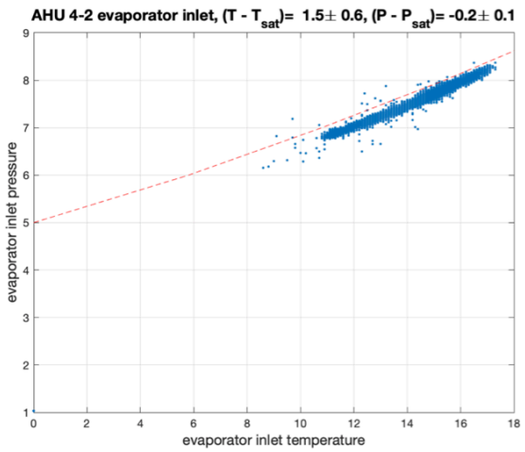
Power efficiency

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

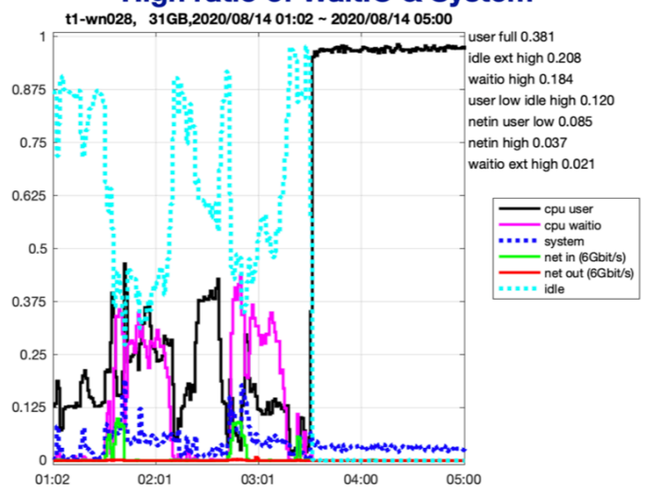
Application and System efficiency

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

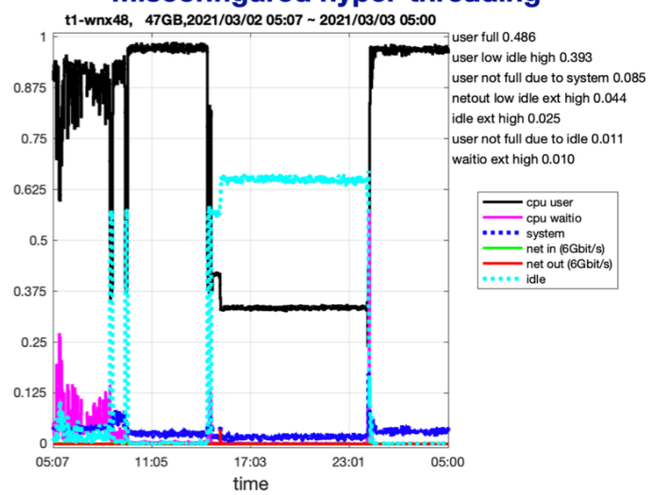
AHU Performance Monitoring



Worknode Monitoring: High ratio of WaitIO & System

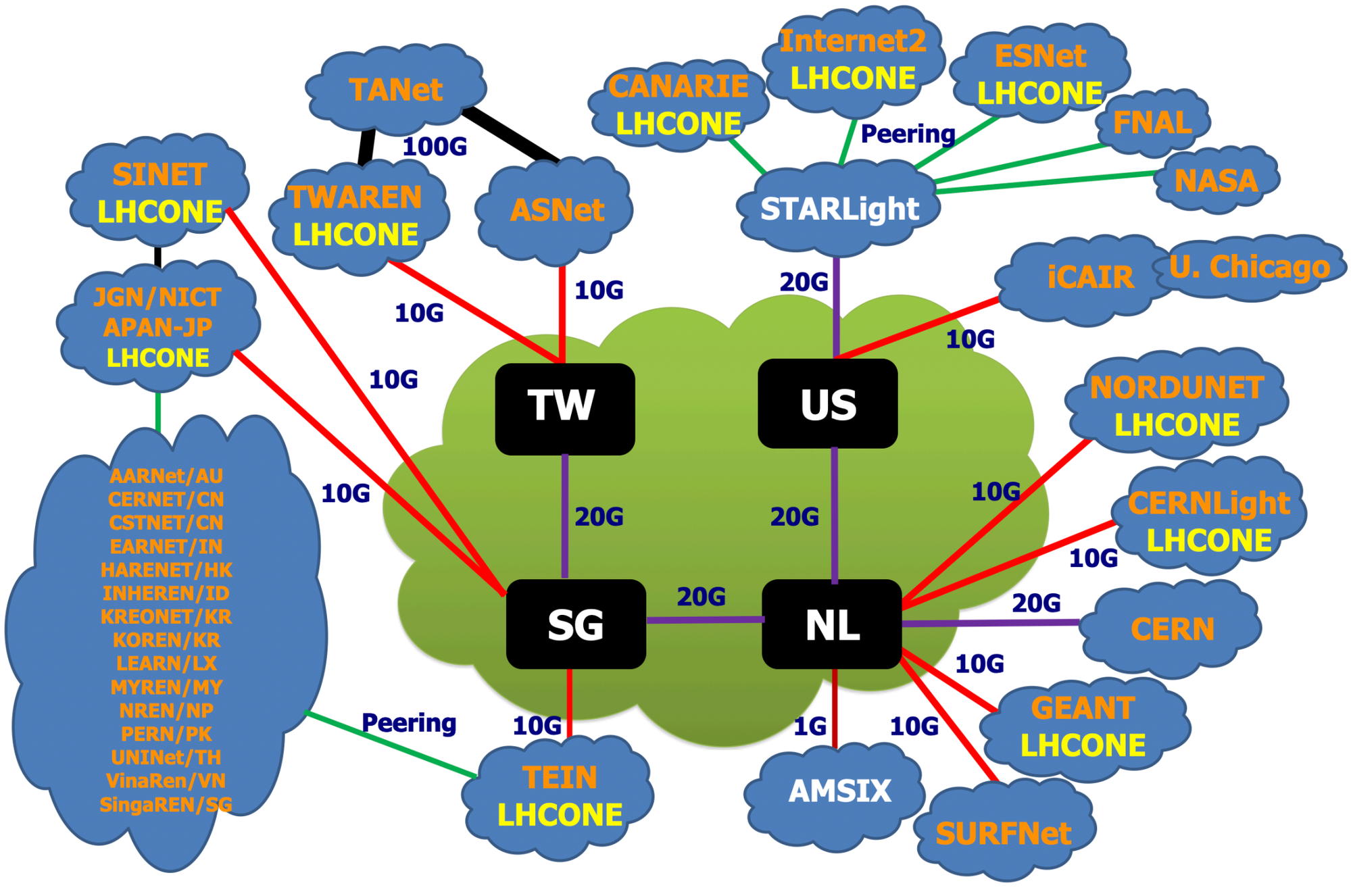


Worknode Monitoring: Misconfigured hyper-threading



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)



Interactions of Science Pulls and Technology Pushes spark innovations and advancement

- CentOS/Linux migration roadmap
- Benchmarking
- IPv6 migration
- Mass storage system for Disk + Tape, e.g., EOS, dCache, etc.
- Distributed computing model
- Federation of distributed storage and data
- Data transmission tools (esp. for long-latency sites)
- Networking: resource virtualization, efficiency optimization, intelligent control, etc.
- ML-enabled data analysis
- Efficiency optimization and intelligent monitoring & control, incl. performance turning, utilization maximization, etc.
- Security Operation Centre (SOC)
- DC & Asset management and monitoring
- Research data management and sharing (FAIR-based)
- Energy saving strategy and evaluation
- …
- All these could be learned and sharing through WLCG Workshop, GDB, LHCOPN/LHCONE and many other events, as well as HEPiX, etc.
- Reorient the Regional Operation Services with support from experiments and WLCG
  - Long latency is not trivial at all and time zone differences is still an issue
  - Apart from collaboration platform as ATCF, effective communication channels have to be in place - Mailing list, Social network tools, information services, etc.

Summary

- WLCG has demonstrated the most viable solutions of global collaboration and e-infrastructure to achieve O(1000)PB data processing and analysis
- WLCG is one of the pillars in supporting the advancement of knowledge frontier, for LHC experiments and many big science projects
- Regional collaboration in Asia is essential and has to evolve accordingly
  - Building capacity together with user communities - ecosystem of technology and collaboration dev.
  - Collaboration platform: ATCF and ISGC, etc.
- 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.
  - New services for M-enabled data analysis and applications are
- Efficiency optimization is not just a pillar of the ASGC Science Cloud, but also the essential contributions to reliability and performance
- Key factors for sustainable operation include HR/Technology, HW, Energy and Networking infrastructure

Welcome To ISGC 2023 & HEPiX Spring 2023 in Taipei

ISGC 2023
Accelerating Time-To-Science Through Computing

- 20th anniversary of ISGC
- International Symposium on Grids and Clouds (ISGC) 2023
  - 19~24 March 2023, Academia Sinica, Taipei, Taiwan
- Call for Abstract
  - On-line Submission: https://indico4.twgrid.org/event/25/abstracts/

- **HEPiX Spring 2023 (27-31 March 2023) will be hosted by ASGC in Academia Sinica, Taipei, Taiwan**

HEPiX Hybrid

- **LHCOPN/ONE**
- **Traffic monitoring of IPv4/v6 respectively - ASGC**
- **multiONE**
- **CNAF-CERN DCI over 400Gbps links (x3)**
  -