TIDC Annual Meeting 2024
on Nov/22nd, 2024

COMPUTING REPORT

Kai-Feng Chen
National Taiwan University

INTRODUCTION

- **Computing is an essential service for HEP** –– no matter what kind of experiments or detectors, we always need sufficient computing power to process, to store, and to analyze data!
  - Resources and knowledge required for modern system operations are non-trivial –– need to consider space/power/cooling/knowhow/ ***people***.
  - Optimal is to have a centralized system instead of clusters running individually, not to add lots of management loads to each group.
- Thanks to the supports from ASGC, now our **TIDC cluster** is operating smoothly too!

This talk includes a lot of materials from Eric Yen & Felix Li (ASGC) and Chun-Yu Lin (NCHC). Many thanks for the contributions as well as the support team in the back!

WHAT DO WE HAVE NOW?

- **ASGC** serves as the computing arm of TIDC :
  - ASGC was serving as a WLCG Tier-1 center from Dec 2005 to Oct 2023.
  - Migrated to WLCG Tier-2 center for ATLAS after Q3 2023.
  - CMS Tier-3 initiated in 2022, in collaboration with TIDC, NTU/NCU.
  - Supporting multiple current and future HEP/GW projects.
- **NCHC** supports fundamental science with dedicated or shared resources:
  - 2015 WLCG Tier-2 for CMS .
  - 2022 GW Network (KAGRA/LIGO/Virgo) via OSG.

STATUS OF CMS TIER2 @ NCHC

- **Operations is aligned with WLCG and IT community for continuous software upgrade:**
  - Migration to token-based authentication for both computing&storage (still on-going)
  - Oversea networking thru TWAREN/Internet2 and v4/v6 dual-stack
  - TWAREN has three >10 gbps thru Chicago, LA, and Singapore
  - OS upgrade to Alma Linux 9

CPU models:

10 AMD EYPC 7713 (128c Milan @2023)

24 Intel E5-2670v2 (10c Ivy Bridge @2014)

120 Intel E5-2630v3 (8c Sandy Bridge@2013)

+400 vcore on NCHC's internal cloud

partially shared with OSG

STORAGE

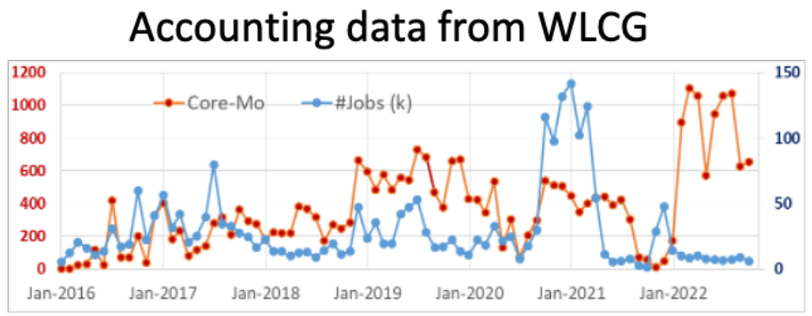
- **2PB capacity based on dCache.v9** and distributed over three branch of NCHC.
  - Support access via WebDAV, XROOT and deprecate SRM.
  - Now (2024 Nov), **~300 TB used by domestic users** (~150TB used by chkuo and zhenggan)
- Utilizing (internal) cloud storage is under evaluation.

	Total (TB)	Free (TB)	Free%
Hsinchu (NTU 2021)	404	169	42%
Taichung (2022)	1,344	847	63%
Tainan (NTU 2015, NCU 2017)	328	85	26%
Total	**2,076**	**1,101**	**53%**

SITE AVAILABILITY

- Operations under regular monitoring with metric based on computing, storage, and networking.
  - On-going token authentication on-going and trigger WARNINGs, still mainly reply on X509.
- **Storage are fairly stable** – have to prepare to move Tainan's storage out.
- **More computing nodes** are gathering.
- Many helps from TWAREN teams for resolving oversea networking.

TIMELINE



Totally ~2k core / 1+ PB, managed by HTCondor/ARC-CE and dCache distributed storage.

- 2023 Jun: Transit from TWAREN/Internet2 (PacWave) to CERN/LHCONE and IPv6 dual-stack start to work.
- 2023 Mar: Join OSG for IGWN computing via NCloud (~500 vcores, thanks to CNSu).
- 2022 Oct: Several old ASUS nodes and 5 EPYC nodes installed in K4. IPv6 deadline for T2.
- 2022 Mar: 1.2PB disk nodes in Taichung (Thanks to Spiraea).
- 2021 Nov: NTU 500TB storage (K4); Taichung 1PB storage; Tainan storage relocation (P4->L04 south).
- 2021 Jun: CMS PhedEX transit to Rucio, ARC accounting migrate from ActiveMQ to EGI ARGO messaging service
- 2020 Feb: **F5 ⇒ Bravo cluster (C7)**; **MoST project(~800k NTD for 2y)**
- 2019 Jan : **IBM1350A ⇒ Formosa 5**; CREAM CE migrate to ARC+Condor
- **2016 Jan: WLCG CMS Tier-2 MoU (WLCG Collaboration Board)**
- 2015 Aug: NTU 512TB MSA2040 installed; **MoST project(~1M NTD for 2y)**
- 2015 Feb-Jul: **Phase-2 project(480k)** w/ 512-core. Join Belle-2 MC campaign.
- 2014 Nov-2015 Jan: **Phase-I init project (50k)** w/ **60-core @ IBM1350A**
- 2013 Mar: Initial request from CMS community

Thanks to Chun-Yu for long term commitment and operation!

ASGC OVERVIEW & STATUS

- **ASGC is providing big data analysis and computing services for the R&E communities in Taiwan as a core facility:**
  - Funded by both Academia Sinica and National Science and Technology Council
  - Primary scientific collaborations: WLCG (ATLAS, CMS), AMS, Gravitational Wave, ICECube/ Neutrino, EIC, QCD, CryoEM, condense matter, etc.
  - Based on the core technologies of WLCG
- **System efficiency as well as AI-enabled analytics are the new focus.**

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

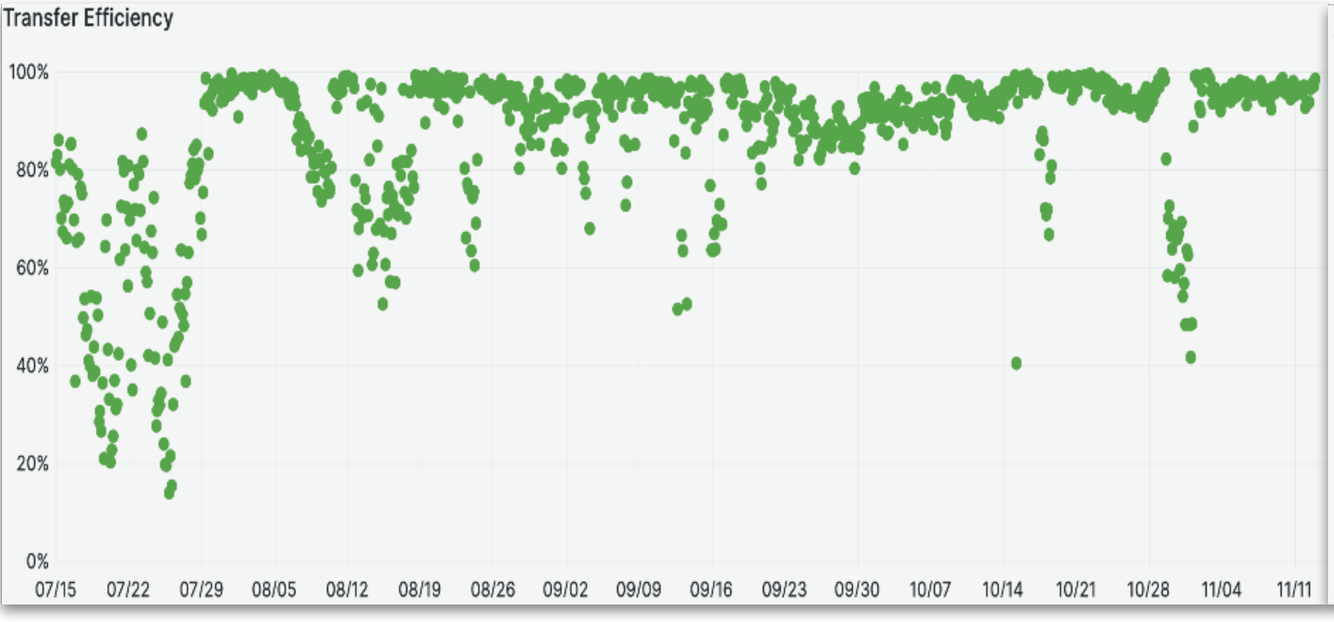
WLCG ATLAS TIER2 IS RUNNING SMOOTHLY

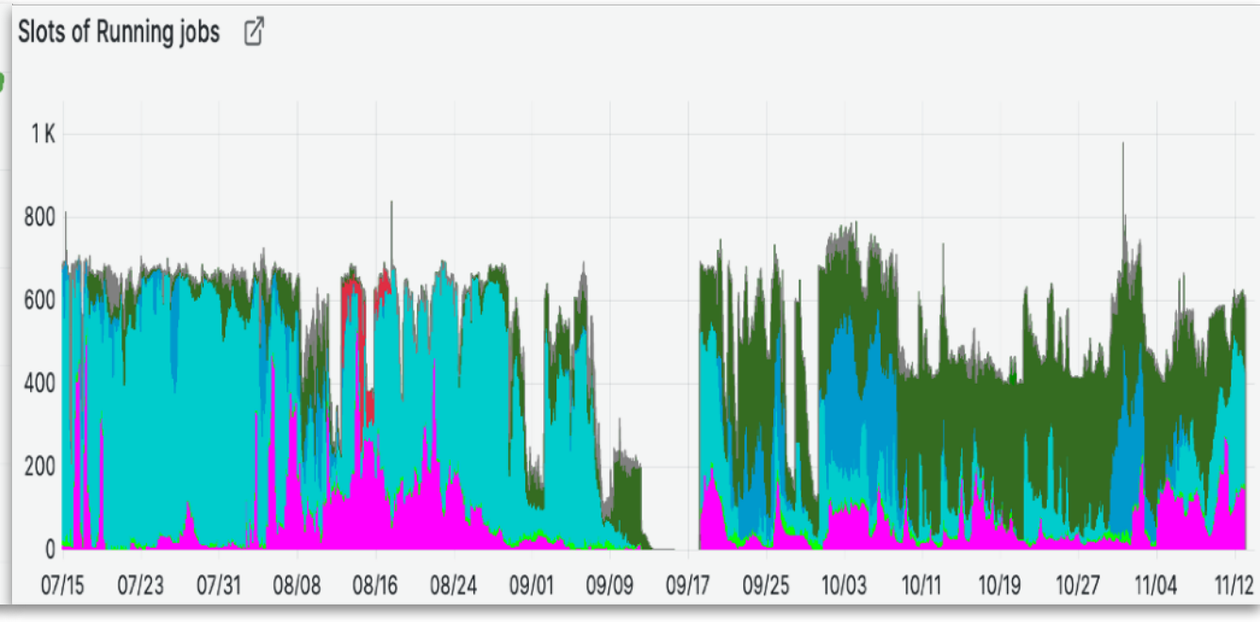
- **Site has been in production since July 2024 with helpful and careful investigation of USATLAS**
  - Connection with LHCONE was reestablished, efficiency was also affirmed
  - Local storage was reconfigured.
  - Passed test jobs on 10 July.
- Supporting MC Simulation jobs and analysis jobs mainly afterwards
- **Pledge in 2024:**
  - **CPU: 30K HEPscore23**
  - **Disk: 5 PB, managed by EOS**
- Job slots: 2,208, ~839 job slots are available now because of OS migration.

Current Status:
>55,219 CPU Core-Days (2024)
Transfer efficiency: > 90%
CPU Efficiency: > 82%
12% User Analysis - 74% eff.
49% MC Simulation - 83% eff.
39% Event Generation - 100% eff.
Storage used: > 1.1PB
Issue: limited network bandwidth (3Gbps) connecting to LHCONE

WLCG ATLAS TIER2 IS RUNNING SMOOTHLY (II)

- Next steps:
  - **Making better use of ASGC resource**, e.g, pre-stage for analysis jobs
  - Will take USATLAS operation shift in Asia time zone in 2025
- Migration to AlmaLinux 9 - will be accomplished before end of 2024
  - 2-stage: ~ 1,370 job slots in AlmaLinux online in Sep.
  - HTCondor and ARC-CE also need to be upgraded

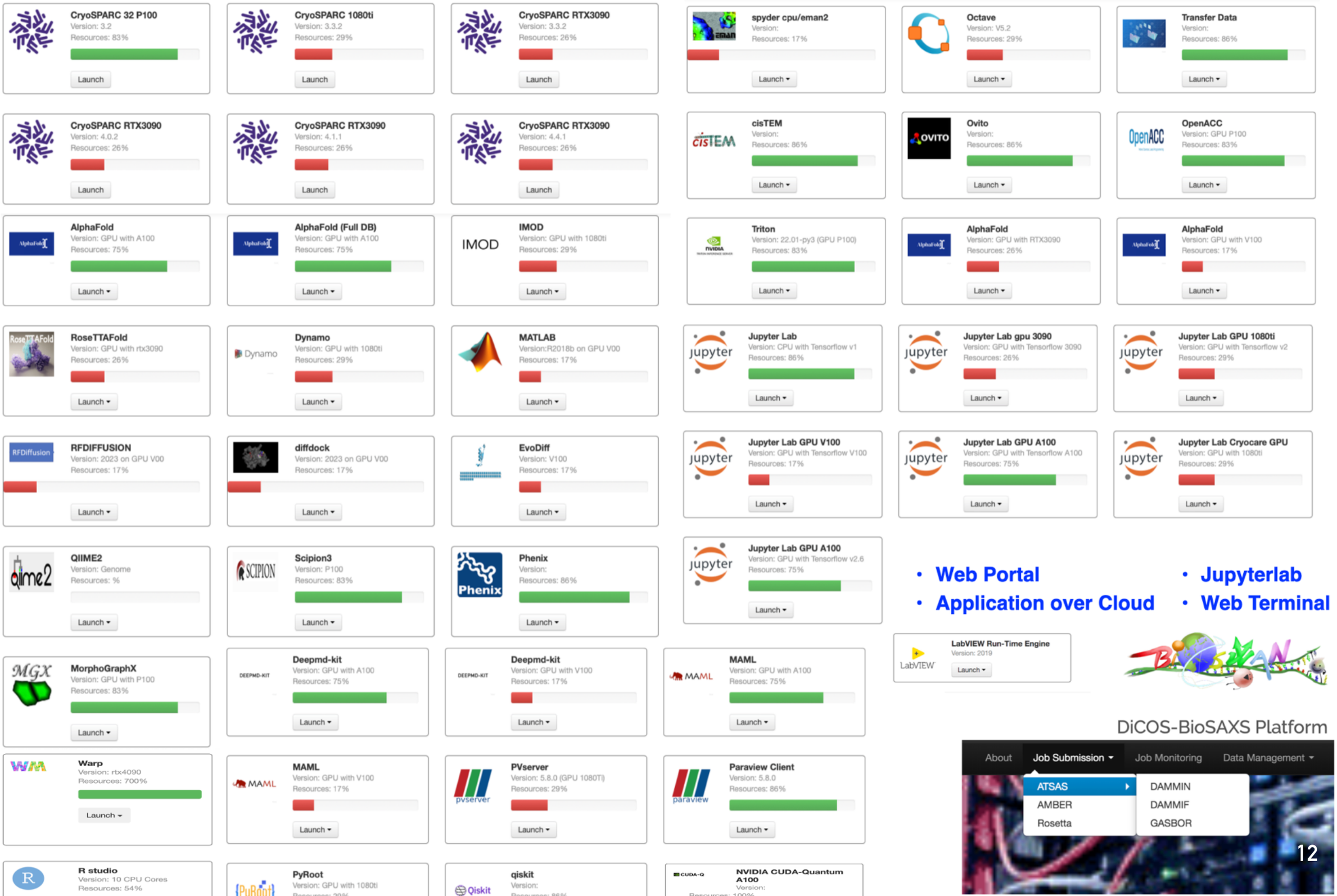




WLCG-BASED COMMON INFRASTRUCTURE

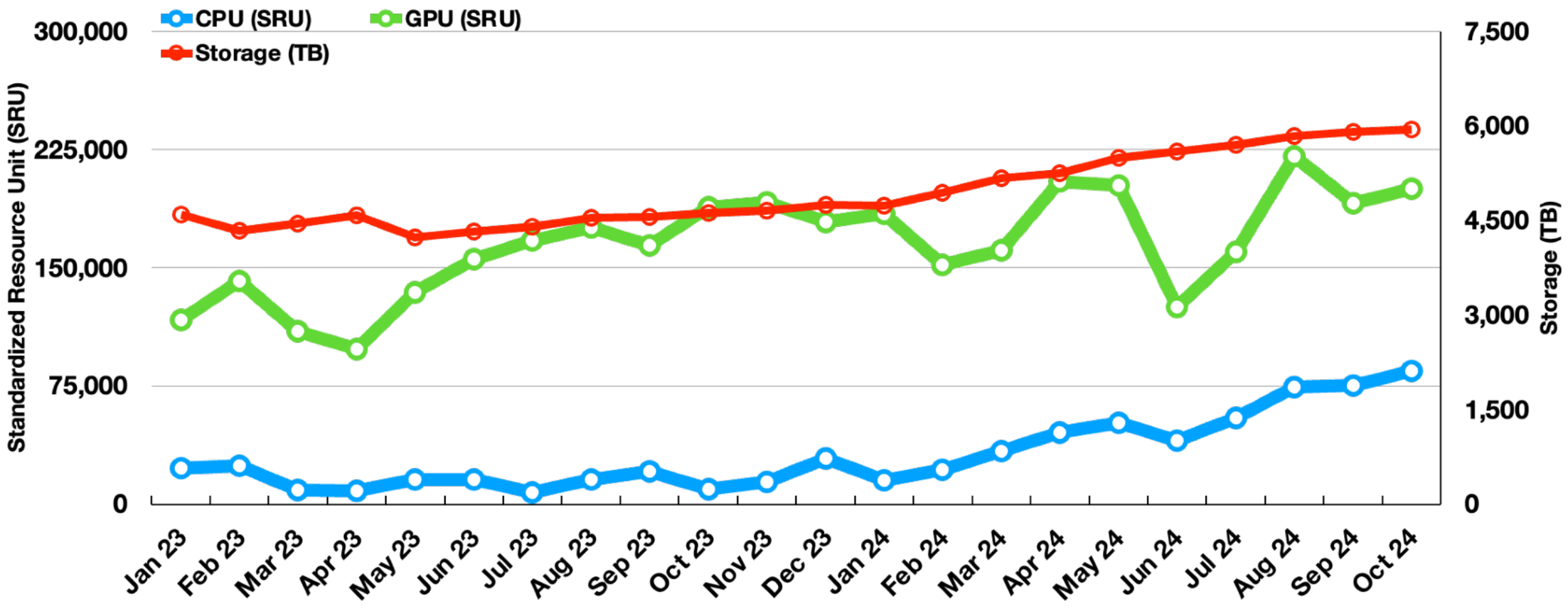
- **PanDA + RUCIO** serves as the core of the common distributed infrastructure
  - Federation of distributed institute resources
  - Federation of core facilities (including CryoEM, NSRRC, NMR, computing centre, etc.)
- **Web-based Cloud services & Slurm clusters** are provided for various computing models
  - **VM for core services and on-demand worker nodes** managed by OpenStack
  - Containerized resources managed by Kubernetes framework for software on-demand services and part of core services
    - ✓ **Batch and interactive GUI jobs: Jupyterlab, virtual desktop**
    - ✓ **GPU Cloud**
    - ✓ **SaaS: web-based application environment with microservices**

73+ WEB APPLICATIONS PROVIDED (SOFTWARE-AS-A-SERVICE)



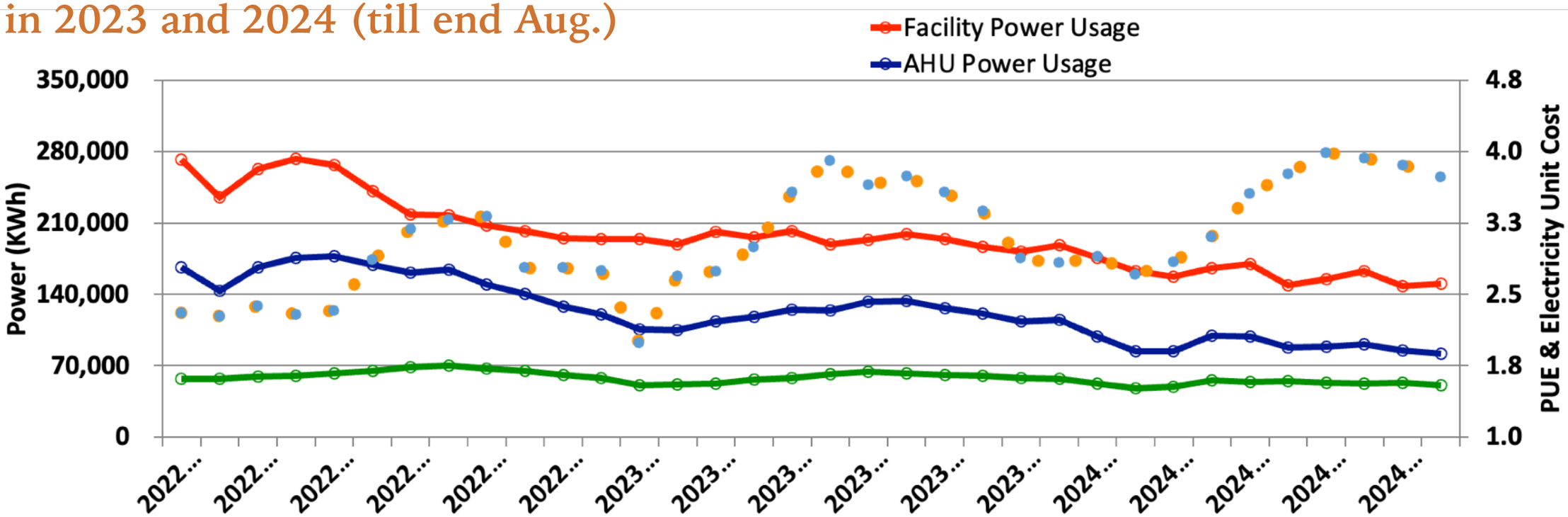
ASGC SCIENCE CLOUD RESOURCE USAGE

- **>98% reliability in 2024**
- Supporting O(1M) CPU jobs from O(100) users a year by +2K CPUCores annually
  - With larger memory/core configuration
- **Demands of high-end GPU keep increasing:**
  - Scientific computing: Lattice QCD, Neutrino, HEP experiment and analysis, Astronomy, Life Science, etc.
  - ML/AI applications including (smaller scale) LLM



ENERGY EFFICIENCY IMPROVEMENT

- **Reliability enhanced by intelligent monitoring and control is the key approach**
- Retirement of legacy hardware
- Improvement of AHU efficiency, including the replacement by top-flow cold air
  - Anomaly detection
  - Well-prepared backup plan
- Energy-sensitive operation:
  - Plan for power efficient hardware: e.g., non-X86 CPUs
  - Power saving - shutdown some idle WNs when the waiting queue is quite short
    - ✓ 20% power usage reduction in 2023 - Effective on 3 CPU clusters (> 3,000 CPUCores) from May 2023
- Overall, DC power usage achieves around 20% reduction rate achieved per annum in 2023 and 2024 (till end Aug.)



FUTURE WORKS

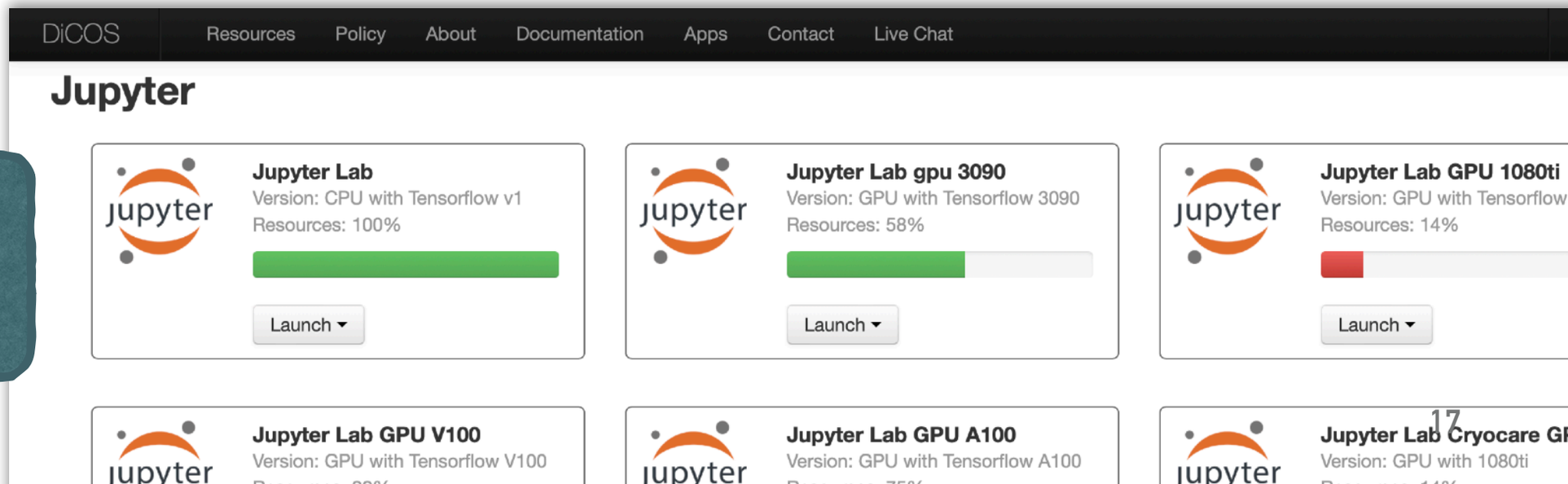
- **Supporting the high-performance data analysis needs of TIDC groups by making better use of ASGC resources.**
- **Both the capacity of users and ASGC are growing according to collaborations**
  - Extended from local ATLAS/CMS groups, to other TIDC experiments and to broader scientific communities
  - Providing ASGC service training and thematic training/workshop
  - **Collaboration platform: ISGC annual event since 2002**
- **Upcoming Services:**
  - Data efficiency and reliability: enrich backup and archive services by integration of disk & tape storages
  - Two Factor Authentication service will be delivered by end of 2024

ACCESS TO USER ANALYSIS RESOURCES

COMPUTING SERVICES @ ASGC

- Currently ASGC is operating the **NSTCCore** services to support general HPC users & applications:
  - 2,944 CPU cores + 2,176 CPU cores (by end 2024)
  - 24x A100 GPU, 16x4090, 48xV100, 56x3090 (+ 12x L40S soon)
  - Working space: Ceph filesystem > 10PB is online (+2.5PB by end 2024)
  - Batch jobs: Slurm management system, with the entry UI: slurm-ui.twgrid.org.
  - Interactive jobs with web-based UI via dicos.grid.sinica.edu.tw

It is possible to start a Jupyter notebook with GPU supports.

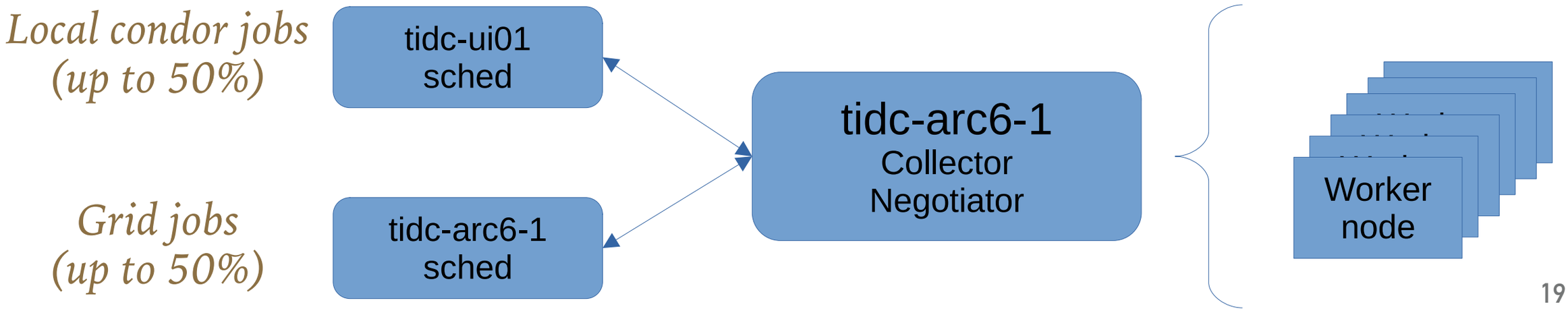


COMPUTING SERVICES @ ASGC (II)

- ASGC is welcoming new users –– in order to use the resources, **group PI has to create a group account first** and then ask the users to register;
  - PI has the privilege to monitor the user activities (*and will receive the cost bill monthly.*)
- Group/user accounts creation:
  https://canew.twgrid.org/ApplyAccount/groupcreate.php
  https://canew.twgrid.org/ApplyAccount/ApplyAccount.php
- TIDC cluster is managed under the same dicos system –– if you would like to use TIDC CPU/storage, you have to apply the same accounts too.
  - At this moment people affiliated to NTU and NCU should fill Kai-Feng Chen and Chia-Ming Kuo as PI.

TIDC CLUSTER OVERVIEW

- Seeding cost allocated under TIDC project with equipments injected from individual grants (George's ASP, budget from Ming & myself, etc).
- Configured as separated **Condor clusters** for grid/local usage:
  - CPU: 768 cores (*AMD EPYC 7713*) + 768 cores (*Intel CPU E5- 2650 v4*)
  - EOS storage: 650 Tb (60% reserved for grid) + 400 Tb
  - Shared user/group working space (the same as NSTCCore services mentioned earlier).
  - UI: `tidc-ui01.grid.sinica.edu.tw`, accessible via SSH. (also `tidc-ui03` now!)
  - CE: `tidc-arc6-1.grid.sinica.edu.tw`, accessible with grid certificate.



TIDC CLUSTER OVERVIEW (II)

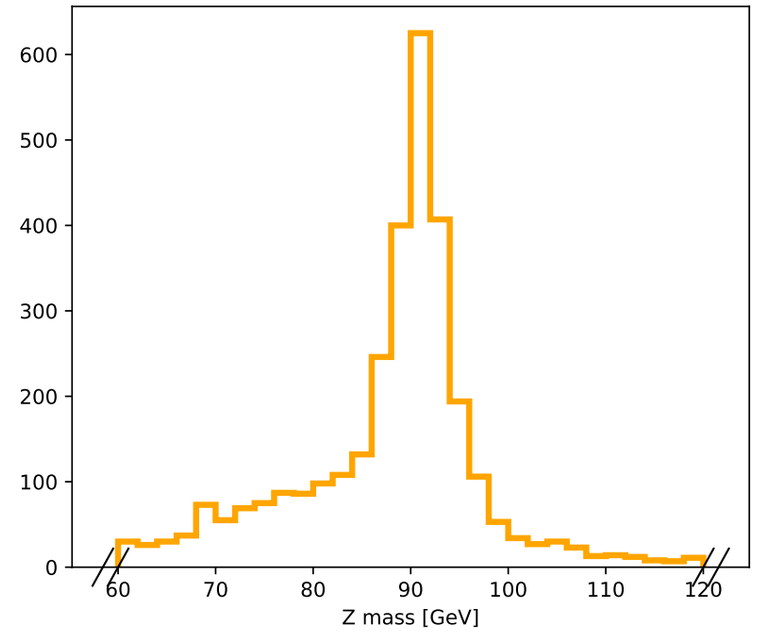
- **Software access:**
  - CVMFS is available –– so basically all the CERN related libraries/tools should be accessible directly.
  - We have only tested CMS software (CMSSW) and analysis frameworks (RDataFrame/Coffea) so far, but the general analysis flow should be similar for other HEP projects too.

- **Storage access:**
  - Grid EOS storage is accessible by xrootd: `root://tidc-smstor1.grid.sinica.edu.tw/eos/`
    - Or via fuse mount under UI: `/eos`
  - Private EOS area: `/eos/tidc/⟨group_name⟩` (50 Tb / group)
  - Working space: `/ceph/work/⟨group_name⟩` (3 Tb / group).

ps. This is basically too small for modern analyses, PI can ask for more when needed.

HOW TO USE?

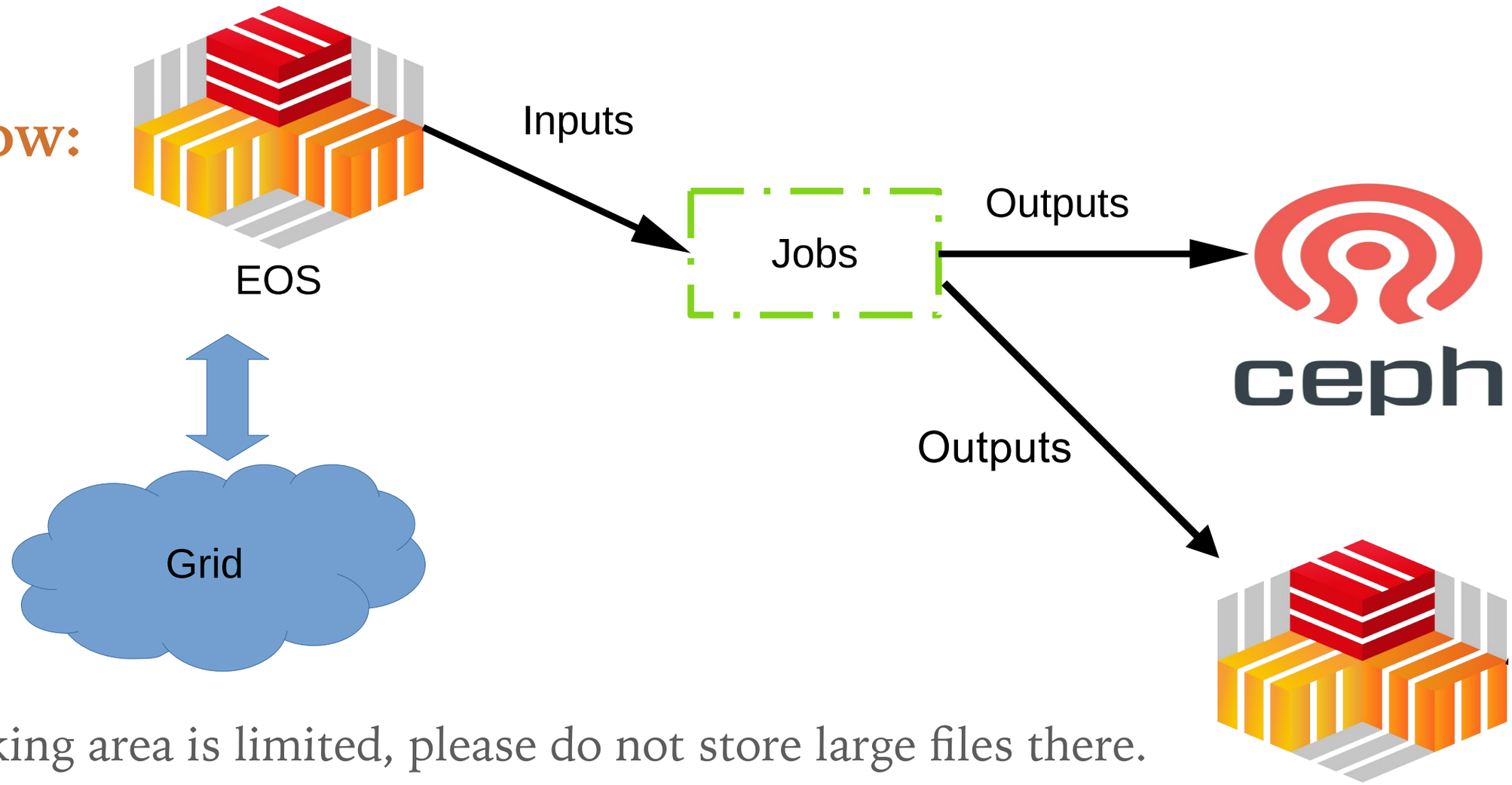
- **ASGC is running hands-on tutorials every 3 months.** A dedicated one held at NTU last year (*many thanks to ASGC & TIDC again!*): https://indico4.twgrid.org/event/35/ with a special session for TIDC cluster.
- Topics of interests:
  - Access to TIDC & Condor schedular (by Felix Lee): https://indico4.twgrid.org/event/35/#b-595-hands-on-computing-servi
  - General CMS software setup & condor jobs (by You-Ying Li): https://indico4.twgrid.org/event/35/#b-597-hands-on-analysis-framew
  - Running analysis with RDataFrame (by Cheng-Han Wu) & Coffea (by Yu-Hsuan Chou): https://indico4.twgrid.org/event/35/#b-587-computing-service-for-he



$Z \to e^+e^-$ peak generated by Coffea running on TIDC system, using CMS open data.

ANALYSIS WORKFLOW

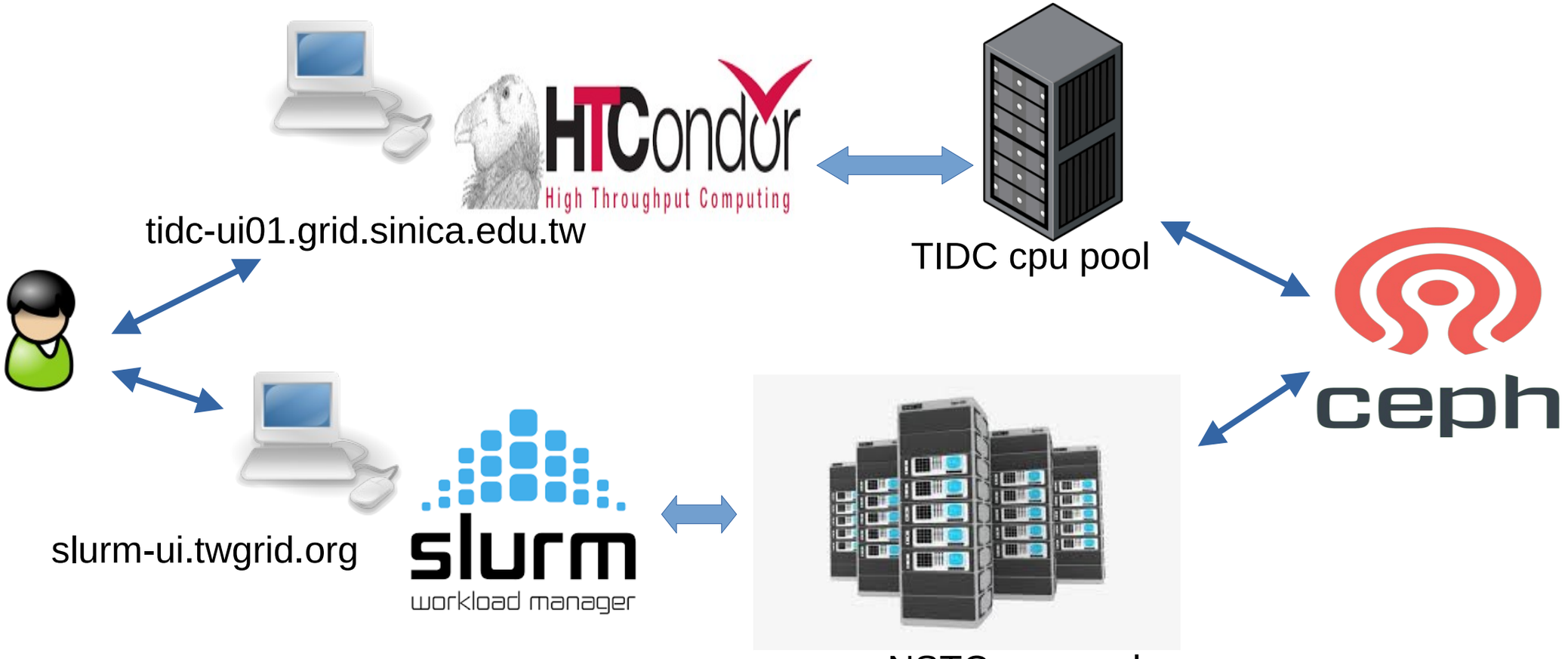
- Typical workflow:



- Ceph working area is limited, please do not store large files there.
- EOS is more optimized for ROOT access, better keep your output ROOT files there if needed.
  - It is a bad idea to write your private files back to grid EOS (might be killed by global grid management), please use the private EOS area instead.

ANALYSIS WORKFLOW

- **Integration with NSTCCore services:**
  - It is possible to use the CPU/GPU allocated under NSTCCore too (as the Dicos user area & Ceph workspace are shared).
  - No direct access to experimental software but possible to run general analysis tasks (e.g. statistical studies, ML training, etc).



SUMMARY & PROSPECTS

- Many thanks to ASGC's support, now we have our **TIDC cluster up and running**!
  - Please apply group/user accounts to access the resources.
  - Now it is CMS compatible, but it should be able to run the analysis jobs for other experiments as well (*CERN projects should be straightforward*) –– further integrations can be discussed!
- **Please be gentle to ui01 / not to overload it!**
  More and stronger UIs are coming, converting from NCU machines! (e.g. ui03)
- Expected to double the specs based on the summit grant early next year.

HAPPY RUNNING ANALYSIS JOBS!