



# WBS 2.3.4 Analysis Facilities

Ofer Rind / BNL    Ilija Vukotic / UC    Wei Yang / SLAC



## 2.3.4 FTE Summary

| Institution | Name     | ops FY22 | ops FY23 | Comment  |
|-------------|----------|----------|----------|--|
| SLAC        | Yang     | 0.15     | 0.15     |  |
| SLAC        | Dart     | 0.25     | 0.25     |  |
| BNL         | Poat     | 0.05     | 0.05     | Replaces Strecker-Kellogg who left BNL           |
| BNL         | Huang    | 0.15     | 0.15     | Effort shifted from Liu to Huang                 |
| BNL         | Rind     | 0.55     | 0.30     | Effort shifted to 2.3.5                          |
| BNL         | Benjamin | 0.00     | 0.30     | Technical support and infrastructure development |
| Chicago     | Vukotic  | 0.20     | 0.20     |  |
| Chicago     | Hu       | 0.50     | 0.25     |  |
| Total       |          | 1.85     | 1.65     |  |

Both BNL and Chicago leverage T1/T2 infrastructure effort from 2.3.1 & 2.3.2 respectively, not included here.

note about effort to wbs 5:

BNL: Shuwei Ye 0.5 (AF related)

SLAC: Patrick Pascual 0.2 (sw, k8s, jupyter infrastructure, GPU, etc.)

NIU: Cecilia Duran 0.5



## 2.3.4 Overview of Activities in FY22

- ❖ Operate BNL and SLAC AFs and support users
  - Deployed ATLAS-TALK Discourse server, in collaboration with WBS5
- ❖ UC AF (NSF Shared Tier3) open to users
  - With onboard events
- ❖ GPU deployment at BNL, SLAC.
  - Each lab received RBT and installed a GPU node (details in backup slide)
  - Available for both Jupyter users and batch users
- ❖ Re-organized Tier 3 public documentation at [ReadTheDocs.io](https://ReadTheDocs.io)
  - Incorporated the new AF at UC
  - Rewritten to make more intuitive to users (A. Roepe-Gier)
- ❖ Federated identity at BNL and UC
- ❖ EOS integration for file sharing at BNL and UC
  - Investigation at SLAC
- ❖ Engage IRIS-HEP
  - ServiceX development at UC. Running production, test and integrations instances.
  - Operation, integration testing of Coffea-Casa at UC.
  - OKD cluster supporting REANA and ServiceX at BNL - defining best practices for IRIS-HEP developers to accommodate OKD
  - K8s at SLAC



## 2.3.4 FY22 Milestones Summary

|       |     |   |          |          |  |
|-------|-----|---|----------|----------|--|
| 2.3.4 | 68  | NSF shared T3 in production                                       | Dec 2021 | Nov 2021 | Completed  |
| 2.3.4 | 114 | Jupyterhub at BNL AF opened to collaborators through federated ID | May 2022 | May 2022 | Completed, awaiting final cybersecurity approval |
| 2.3.4 | 116 | Deployment of Lustre file system for Analysis Facility            | Oct 2022 | Oct 2022 | On Schedule                                      |



## 2.3.4: FY22 Milestone Status

### ❖ MS68. NSF shared T3 in production

Completed - Dec 2021 - Nov 2021

- Supported US ATLAS onboard event, IRIS-HEP Analysis workshops
- Developed user database, unix account provisioning, website, user portal, documentation
- Provided federated access to Jupyter via CI-Logon
- Provided access to GPUs from Jupyter
- Migrated features of ML Platform relevant to Tier3/AF
- Provided access to IRIS-HEP SSL (896 cores) opportunistically
- Provided metrics collection scripts and dashboards for all Tier3s
- Continue to improve AF-specific facility monitoring
- Production deployments of IRIS-HEP Coffea-Casa (integrated with ATLAS IAM) and ServiceX
  - For early adopter evaluation & analysis grand challenges
- Risks
  - User support not sufficient to meet user expectations
    - Work with WBS 5 on training events, best practices, documentation
  - Technical manpower not sufficient to implement capabilities available at the labs (data sharing, EOS integration, metrics collection, federation, matching DOE policies)



## 2.3.4: FY22 Milestone Status

- ❖ **MS114. Jupyterhub at BNL AF opened to collaborators through federated ID** (D. Benjamin, C. Caramarcu, K. Casella, R. Hancock, C. Hollowell, S. Kandasamy, J. Lauret, C. Lepore, T. Rao, O. Rind, J. Smith)

**Completed - May 2022 - May 2022**

- Lower barrier to access for ATLAS users without SDCC accounts by providing lightweight jupyterhub accounts keyed to CERN, SLAC or FNAL credentials. Elements of this work included:
  - Implementation of MFA requirement
  - Creation of a new user form and procedures for affiliation confirmation and access approval (2-3 business day goal)
  - Updated EPPN OIDC attribute mapping in LDAP for authorization and local account creation
  - Modifications to `jhub_remote_user_plugin` to support new account mappings
- Procedure has been followed by a number of users with positive feedback
- Risks: Procedure still too unwieldy for users; cybersec approval withheld



## 2.3.4: FY22 Milestone Status

- ❖ MS116. Deployment of Lustre file system for Analysis Facility (D. Benjamin, Z. Liu, T. Rao, O. Rind) On Schedule - Oct 2022 - Oct 2022
  - High performance storage for AF users on Condor pools and IC
  - Currently ~3 PB deployed, including space for HPC globus endpoint.
    - Need to test access patterns with Tier 3 analysis workloads.
    - Need to mount lustre client on shared pool nodes.
    - Explore options for user quota management
  - Risks: Technology mismatch for user access pattern (e.g. small files); lack of personpower for sufficient support



## 2.3.4: Ongoing M&O Activities

- ❖ Coordination, general support and development, user engagement, documentation: Rind (BNL, 0.55), Yang (SLAC, 0.15) **0.7 FTE**
  - a. Rind reducing to 0.3 in FY23; Benjamin being added at 0.3
- ❖ Batch Support: Strecker-Kellogg (BNL, 0.05), Dart (SLAC, 0.15) **0.2 FTE**
  - a. Poat to replace Strecker-Kellogg in FY23
- ❖ Storage (gpfs, quotas, localgroupdisk mgt, users): Liu (BNL, 0.15), Dart (SLAC, 0.1) **0.25 FTE**
  - a. Huang to replace Liu in FY23
- ❖ Interactive/Jupyter development, ML platform, Monitoring: Vukotic (UC, 0.2) **0.2 FTE**
- ❖ UC Tier-3: Hu (UC, 0.25)





## 2.3.4 FY 23 New Activities

- ❖ **Extend common Jupyter environments**
  - Investigate adding support of PHYS and PHYSLITE support in common Jupyter environment?
  - Has a dependency on SW availability
- ❖ **Develop and collect AF usage metrics**
- ❖ **Further exploration of data sharing at all US AFs**
  - In addition to bulk, low sensitivity data sharing
  - explore the possibility of mounting /EOS at AFs
- ❖ **Engage IRIS-HEP, AMG and others user analysis R&D efforts**
  - Continue working with IRIS-HEP on Analysis Grand Challenge planning, readiness, etc.
  - Develop a representative baseline analysis to use for evaluating changes within and across AF sites
  - Connection with ATLAS - Google user analysis project
    - It services as an important connection to industry and knowledge source



## 2.3.4 Risk Assessment

- ❖ Risk 1: AF takes off and demand increases during Run 3
  - We may not have sufficient (unpledged) CPU and Storage resources to meet that demand
  - Mitigation: Identify other available resources or request additional funding
  - Costs: Delays and difficulties for user analysis creating frustration with the tools being provided
  - Probability: Medium
  - Owners: US ATLAS and AFs



## 2.3.4 Risk Assessment

- ❖ Risk 2: Divergence of Analysis Tools and Techniques within ATLAS
  - Development of analysis tools is proceeding under multiple umbrellas, potentially resulting in competing solutions that may not be implementable across all sites. This could create a hurdle for analysis collaboration within ATLAS.
  - Mitigation: Maintain close ties with external efforts (IRIS-HEP, HSF, etc.). Need ATLAS AF task force?
  - Costs: Changing HW and personnel requirements affecting AF deployment timeline (for example, SLAC does not have HW to deploy ServiceX if it is identified as a critical service, would require RBT)
  - Probability: Medium
  - Owners: US ATLAS and AFs



## 2.3.4 Risk Assessment

### ❖ Risk 3: 10 % Funding Cut

- Reduction of personnel effort and hardware, some of which already relies on RBTs at BNL and SLAC
- Mitigation: Distribute reduction of support level and HW resources to minimize impact on users
- Costs: Deterioration of user analysis experience and effectiveness
- Probability: High
- Owners: US ATLAS and AFs



## 2.3.4 FY23 Milestones

- ❖ MS 234-1 Support PHYSLITE based analysis ???
  - Baseline date: FY23QX (depend on work by the AMG group ???)
  - In Jupyter ???
- ❖ MS 234-2 Implement AF usage metrics and collectors:
  - work with WBS 5 to define those metrics
  - baseline date: FY23QX



## 2.3.4: Summary and Conclusion

### ❖ Four areas of work

- Operations
  - **Emphasize stable AF operation during Run 3**
  - Expect much more user activities in Run 3
- Continue to engage users
  - Encourage them to use and explore the AF resources
- Develop and deploy identified new services
  - support PHYS/PHYSLITE based analysis in Jupyter/Python ecosystem environment
  - data sharing
- Engage new analysis models/developers
  - Identify new services, IRIS-HEP, analysis modeling group, etc.

### ❖ Closely work with WBS 5

- tutorial
- user engagement
- testing new services



# Backup slides



## 2.3.4 Summary of Risk Assessment

### ❖ Risk 1

- Do we have sufficient CPU resources at each AF for Run 3 ?

### ❖ Risk 2

- Work related to IRIS-HEP, depend on implementation architecture, and their schedule
- Not clear about timeline, HW requirement and FTE requirement
  - e.g. SLAC does not have HW to deploy ServiceX if it is identified as critical a service - require RBT to help

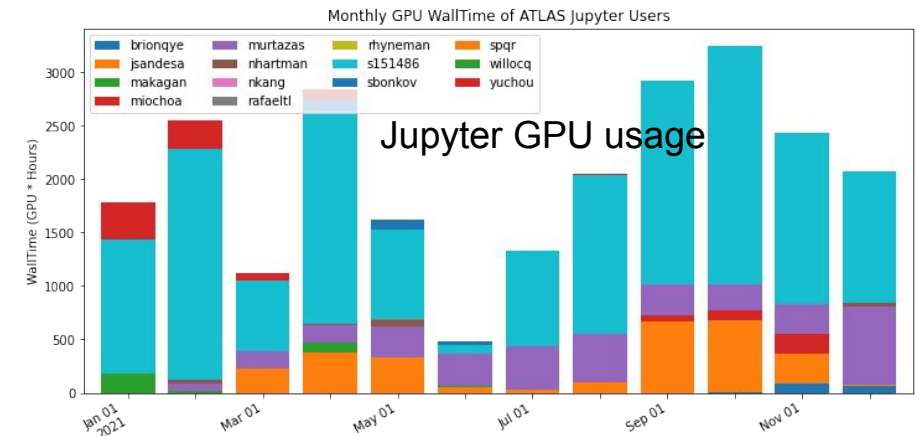
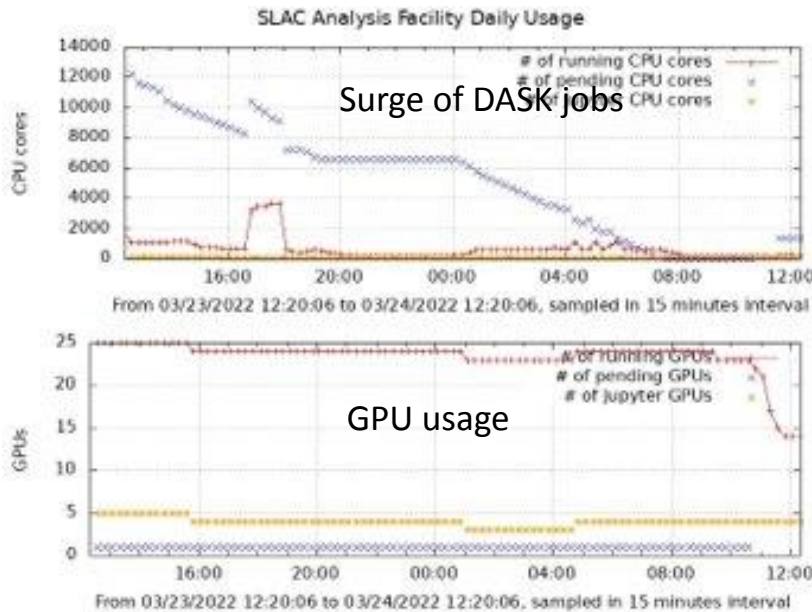
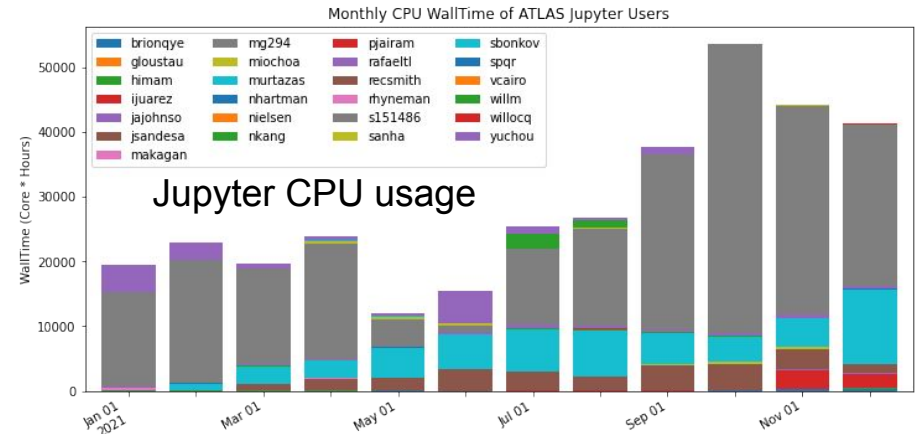
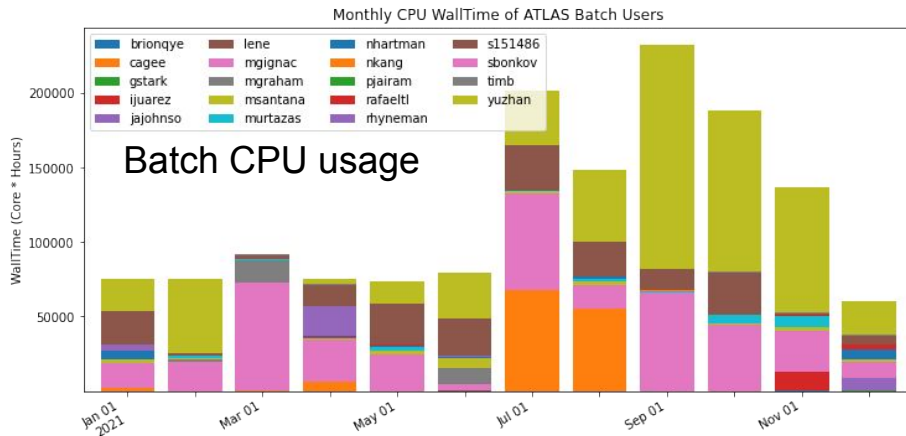
### ❖ Risk 3

- 10% cut: note hardware at BNL and SLAC AF rely on RBTs.
- Evenly cut among at AFs personnel ?





# Highlights From SLAC AF

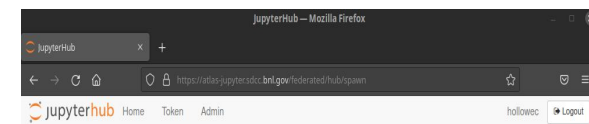
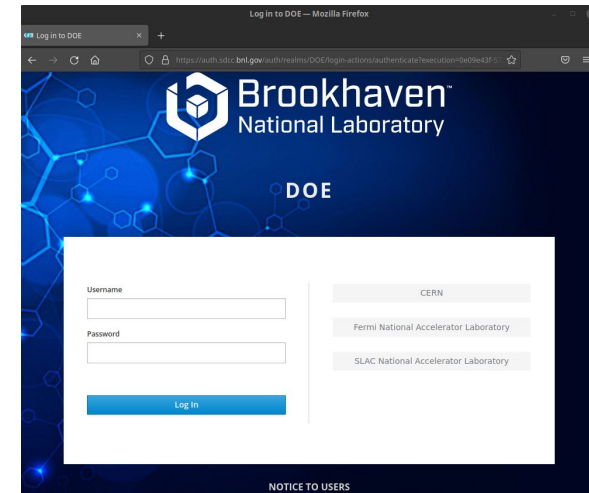


- More Jupyter users than batch users
- Users are utilizing SLAC's large pool of GPUs
- Serious DASK usage.

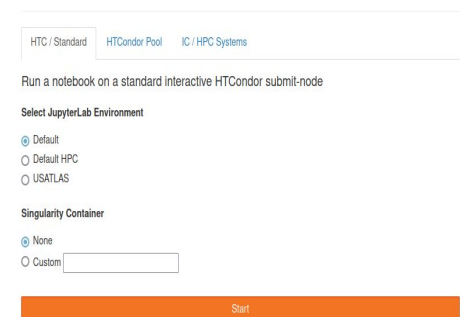
# Federated JupyterHub

- Technical implementation of a federated Jupyter instance for ATLAS complete
  - Allows users to login via BNL/SDCC, CERN, FNAL or SLAC credentials
    - Users without existing SDCC accounts are given “lightweight accounts” where their only access at the facility is Jupyter (i.e. cannot login to SSH gateways, etc.)
  - Implementation requires MFA to satisfy DOE requirements
  - Users must fill out a form and be approved for initial access after their ATLAS affiliation is confirmed
  - Account activation turn-around time goal of 2-3 business days
- EPPN OIDC token attribute mapping used in LDAP for authorization and to tie federated ID to a local UNIX account
  - Reverse proxies doing the authentication
  - Modifications to the Jupyter [jhub\\_remote\\_user\\_plugin](#) to support the mapping
  - Users have access to all SDCC network filesystems
- Being tested by a number of users

Federated JupyterHub Login Screen



SDCC Jupyter Launcher



JupyterHub Web Interface



## 2.3.4 Hardware Purchases

Via RBT

### ❖ At BNL

- 2 new Supermicro SYS-120GQ-TNRT GPU compute nodes added to shared analysis facility
  - Accessible via HTCondor
    - Users added request\_gpus line to their JDFs
  - Per system specs:
    - 2 x Intel Xeon Ice Lake 6336Y CPUs
    - 2 x A100-80 GPUs
    - 256 GB DDR4-3200 MHz RAM
    - 2 x 4 TB NVME drives
    - 10 Gbps NIC

### ❖ At SLAC

- 1. Colfax CX2460s-EK8 (2U) server
  - 2x AMD EPYC (Rome) 7542 CPU. Total 64 cores, 1TB RAM
  - 12TB NVME local scratch, 100 Gbps NIC and IB.
  - 4 Nvidia A100 Redstone GPUs w/ NVLink
  - Integrated into both SLAC Slurm batch (GPU pool) and Jupyter env (which run as Slurm batch jobs)



# Presentations on AF

1. Analysis Facilities activities in DOE multi-purpose computing centers. D. Benjamin, B. Holzman, O. Rind, W. Yang, March 22, 2022, HSF Forum Analysis Kickoff Meeting
2. UC AF presentation???
3. DASK @SLAC and US ATLAS Analysis Facilities, W. Yang, DASK in High Energy Physics, DASK Summit 2021