



UChicago

PI Chats

Rob Gardner
April 23, 2021



Questions

1. Who is actively working on the project from the group and what are they doing?
2. What are contributions of those individuals in your group to the various area(s), i.e. AS, IA, DOMA, SSL, SSC, OSG-LHC, MGMT
3. What specific collaborations has the group made with other groups within IRIS-HEP?
4. What specific collaborations has the group made with other groups or individuals external to IRIS-HEP?
5. Which contributions has the group made to the "Intellectual Hub" aspect of IRIS-HEP? (Including community building activities, training, outreach and broader impact.)
6. Please list any papers/presentations/software-repositories connected to the above.
7. How is your project managed? How is progress measured? How are risks identified and mitigated?
8. Are metrics aligned with the project? Do metrics accurately reflect the progress, success or impact of the project?



Who is actively working on the project from the group and what are they doing

- Rob Gardner (PI)
- Andrew Chien (CS faculty)
- Andrew Eckart (CS MSci)
- Suchandra Thapa (Software Professional)
- Lincoln Bryant (DevOps Engineer)
- Ilija Vukotic (Physics Software Professional)



What are contributions of those individuals in your group to the various area(s), i.e. AS, IA, DOMA, SSL, SSC, OSG-LHC, MGMT

- | | |
|---|----------------|
| ■ Rob Gardner (PI) | SSL, DOMA, MGT |
| ■ Andrew Chien (CS faculty) | SSL |
| ■ Andrew Eckart (CS MSci) | DOMA-SX |
| ■ Suchandra Thapa (Software Professional) | SSL |
| ■ Lincoln Bryant (DevOps Engineer) | SSL |
| ■ Ilija Vukotic (Physics Software Professional) | DOMA-SX |



Accelerated Data Delivery (Note)

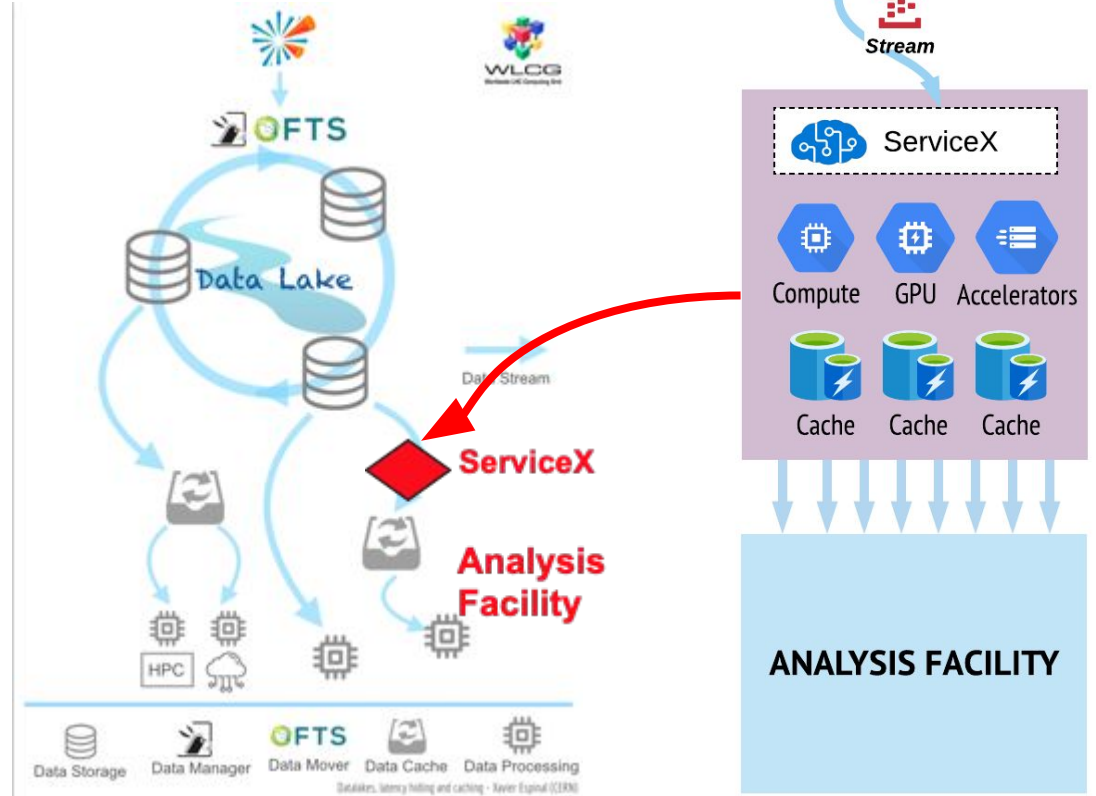
High level goal is to explore (hardware) accelerating data delivery from ROOT format to columnar formats

Two milestones

- **Opportunity assessment** which profiles baseline transformer performance, evaluates technology options (Dec 2020)
- If cost/benefit relative to ServiceX on [cluster baseline](#) indicates, build **prototype system and benchmark its performance** (July 2021)

Andrew Chien has signaled his intention to step away from IRIS-HEP as his other research has become more of a focus (sustainability, book)

Unfortunately our graduate student decided to explore other CS research areas. So this activity is delayed until we find a new CS student.





What specific collaborations has the group made with other groups within IRIS-HEP?

- DOMA ServiceX, Skyhook
 - *Illinois, Washington, Santa Cruz*
- Analysis Services
 - *NYU, Illinois*
- OSG-LHC
 - *Wisconsin, Nebraska, UCSD*
- SSC
 - *Princeton*
- SSL
 - *Wisconsin, UCSD*



What specific collaborations has the group made with other groups or individuals external to IRIS-HEP?

- DOMA ServiceX
 - *Texas, Fermilab*
- Analysis Services
 - *Fermilab*
- Pacific Research Platform
- Chameleon Project
- FuncX
- REANA team
- HTCondor team
- ATLAS Distributed Computing Analytics
 - *Frontier-Squid*
 - *PerfSONAR analytics*
- PanDA, Harvester team
- Open OSG for COVID-19
- FAB (FABRIC Across Borders)



Which contributions has the group made to the "Intellectual Hub" aspect of IRIS-HEP? (Including community building activities, training, outreach and broader impact.)

- Presentations at ATLAS Software and Computing meetings
- Myraid other meetings (Ops, IRIS-HEP)
- Monthly SSL call on Kubernetes topics, open to community (these have dropped off as others have emerged - OSG-LHC k8s hackathons, CNCF on k8s and research computing)
- WLCG-k8s dropped off too
 - *some retrenching, disillusionment phase?*
 - demonstrators, solving specific problems
- Analysis Facility
 - *working on 'shared Tier3' with US ATLAS*

COLLABORATIONS

- WLCG Kubernetes Working Group (L. Bryant co-chair)
- OSG-LHC, SLATE (containerization, deployments)
- DOMA (deployments, functional testing, scale testing - ServiceX, Kafka)
- CoDaS-HEP instructors, US ATLAS Ops, Pacific Research Platform
- REANA development team (sharing unpriv deployments)
- ATLAS ADC (Analytics deployments on SSL), ATLAS PanDA (Harvester containerization)



Please list any papers/presentations/software-repositories connected to the above.

1. Presentations
 - a. *Facility R&D with the Scalable Systems Laboratory, ATLAS ADC Technical Coordination Board, 3/1/21 ([slides](#))*
 - b. *SSL project website lists presentations (<https://iris-hep.org/ssl.html>)*
 - needs updating
2. Publications
 - a. *The Scalable Systems Laboratory: a Platform for Software Innovation for HEP (CHEP2019)*
 - b. *ServiceX - A Distributed, Caching, Columnar Data Delivery Service (CHEP2019)*
 - c. *Towards Real-World Applications of ServiceX, an Analysis Transformation System (Kyungeon Choi, UT Austin), vCHEP 2021 (supporting role)*
3. Software repositories
 - a. <https://github.com/ssl-hep/ServiceX>
 - b. <https://github.com/ssl-hep>



Are metrics aligned with the project? Do metrics accurately reflect the progress, success or impact of the project?

Metrics are well aligned

Metric	Target	Current Value
M.6.1: Number of deployed and operated platform services	Design phase targets taken from initial DOMA and AS estimates. A suitable metric for the execution phase will be designed.	12
M.6.2: Number of SSL user groups	Determined by number of developers & other devops requiring resources	>9
M.6.3: Number of engaged institute areas per quarter	5 Institute areas were targeted for engagement	4 (DOMA, AS, Training, OSG-LHC)
M.6.4: Number of integrated SSL cluster resources	Determined by scale demands requirements of development teams, both within the Institute and outside	2 (1 dedicated, one shared)



How is your project managed? How is progress measured? How are risks identified and mitigated?

1. Management
 - a. *Daily contact through Slack*
 - b. *Brainstorming design sessions*
 - c. *Bi-Weekly meetings (DOMA-ServiceX) (DOMA :()*
 - d. *GitHub issue tracking, PRs and releases managed by Ben*
 - e. *Weekly group meetings*
 - f. *Many other meetings - regular and out of band we're asked to join*
2. Progress
 - a. *SSL - c.f. Metrics (app deployments, IRIS-HEP groups engaged, etc.)*
 - b. *ServiceX - towards 1.0 release (much help from Ben & Gordon)*
3. Risks - examples
 - a. *SSL - technology risk (Kubernetes)*
 - i. *LHC community adoption, pushback by Facility incumbents*
 - ii. *Lack of contributions, expertise, effort*
 - iii. *Culture of isolation, island projects*
 - b. *ServiceX*
 - i. *scalability - testing & measurement*
 - ii. *community adoption, physicist interface*
 - iii. *integration with PaNDA/iDDS (as backend "executor") - has been delayed due to lack of expertise & effort*