Scalable Systems Laboratory (SSL)

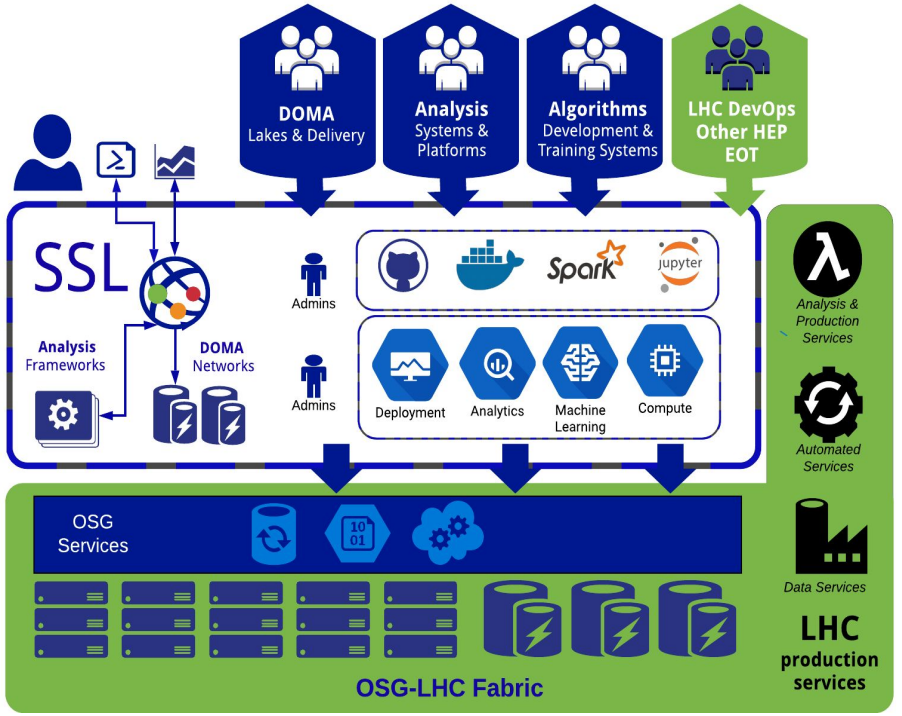
Input to Year 3

Rob Gardner

May 26, 2020

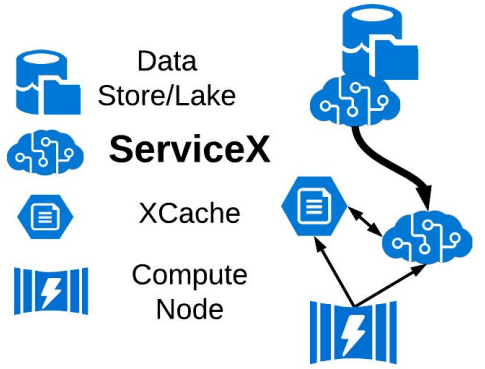
Focus Area: Scalable Systems Laboratory (SSL)

- Vision: Community innovation platform & facility R&D, path to production
- Work with IRIS-HEP research areas, Blueprint, LHC software and computing teams, OSG-LHC
- Utilize industry stand tooling (cloud native Linux container packaging and orchestration: Docker, Kubernetes, Helm)
- Leverage related NSF CI projects SLATE, Pacific Research Platform and others

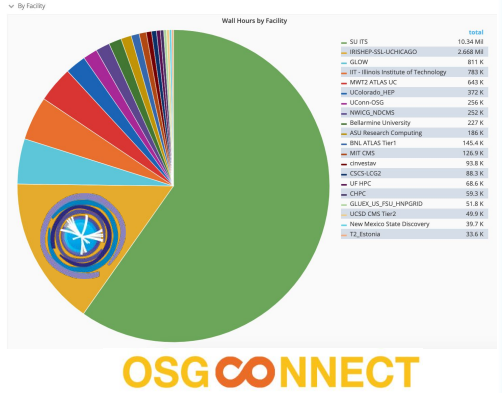


Major Activities

- Re-purposed existing cluster to form a 'kernel' of a distributed Kubernetes service
- Provided testbed for DOMA ServiceX development and scalability testing
- Deployed of CoDaS-HEP training platform
- Supporting ATLAS Harvester edge containerization
- Deployed REANA service
- Additional activities:
  - *Provided various ATLAS and WLCG analytics services*
  - *Configured the SSL to 'backfill' with OSG Connect or ATLAS PanDA*
- Partnership established with Pacific Research Platform to share expertise and access to CPU and GPU resources



reana



SSL Highlight

- Provisioned Jupyter machine learning environment for 55 students attending the CoDaS-HEP school
- Scheduling backend to GPU resources from the NSF Pacific Research Platform and CHASE-CI
  - *More in the following presentation on Training*

Purpose

A computational platform optimized for machine learning applications, supporting the second school on tools, techniques and methods for Computational and Data Science for High Energy Physics (CoDaS-HEP), 22-26 July, 2019, at Princeton University.

Links

CODAS-HEP.org

2019 School Program

HEP Software Foundation

Current Organization

This was the WBS for the SSL area in years 1, 2:

- WBS 6.1 - creating and operating scalable cyberinfrastructure
- WBS 6.2 - establish devOps patterns through blueprint meetings & workshops
- WBS 6.3 - participate in functional testing
- WBS 6.4 - provide database services for metrics aggregation and dashboards
- WBS 6.5 - as needed, provide backend cyberinfrastructure for training

The structure carries forward but aligned with institute-wide challenge goals and priorities

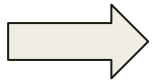
Summary of SSL Deployments

DOMA::ServiceX	Data transformation and delivery service for LHC analyses
Frontier Analytics	Analyze and improve data access patterns for ATLAS Conditions Data
perfSONAR Analytics	Network route visualization based on perfSONAR traces
Parsl / FuncX	Parallel programming in Python, serverless computing with supercomputers
Large-Scale Systems Group @ UChicago	Serverless computing with Kubernetes
DOMA::Skyhook	Programmable storage for databases, scaling Postgres with Ceph object store
REANA	Reusable Analysis Service
CODAS Platform	JupyterLab notebooks, access to GPU resources on the Pacific Research Platform for annual summer CoDaS-HEP training event
SLATE & OSG	Backfilling otherwise unused cycles on SSL with work from the Open Science Grid using the SLATE tools

Activities and Collaborations

ACTIVITY

- Re-purposed existing cluster to form a 'kernel' of a distributed Kubernetes service
- Provided testbed for DOMA **ServiceX**, **Skyhook**
- Deployed of **CoDaS-HEP** training platform; **FuncX**
- Supporting ATLAS **Harvester** edge containerization
- **REANA** Helm dev & deploy
- Additional activities:
  - *ATLAS and WLCG analytics services (**perfSONAR viz, Frontier-Squid** analytics)*
  - *Configured the SSL to 'backfill' with **OSG Connect** or ATLAS **PanDA***
  - *SSL big contributor to **COVID-19***
- Partnership established with Pacific Research Platform to share expertise and access to CPU and GPU resources



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)

Activities and Challenges

- All areas are making progress.

- The SSL has aggressively identified a course for flexible, declarative and dynamic facility infrastructure based on new, fast-moving open source technology, not precisely aligned with scientific computing infrastructure

- So this has brought a number of challenges:
  - *Expanding the Kubernetes substrate (involving more resource providers)*
    - Challenge of lack of Kubernetes expertise in our community
    - Effort available to train, on-board, and collaborate
  - *Support and management of "applications" and "users" of the SSL*
    - Lack of expertise of IRIS-HEP developers with service deployment
    - Integration with external, dependent services (e.g. storage, caches)

Delivering to Experiments, Blueprint, Other meetings

- The delivery to the experiments happens in a number of ways
  - *Helping IRIS-HEP research areas with deployments and testing*
  - *Delivering to LHC computing resource providers*
  - *Helping experiments with containerized service deployment*
  - *Providing education and training infrastructure*
- Blueprint meeting with Analysis Systems for scalable platforms
- SSL-TEAM monthly telecons
- Chicagoland k8s-HEP meetup
- Leadership in WLCG working groups
  - *WLCG Kubernetes WG (co-chair)*
  - *WLCG Federated Operations Security WG (co-chair)*

Year3 Milestone Ideas

1. Realize a **lightweight**, three-site "k8s substrate" which has the following ingredients and capabilities:
   a. *SSL single sign-on and minimal user/group management*
   b. *Distributed storage solution based on Rook (Ceph)*
   c. *HTCondor, Spark and Ray task spawners from interactive notebooks*
   d. *Deliver a **k8s deployment pattern** that can reproduced at a single site (e.g. a Tier3, or a local network of 'gaming' -Nvidia cards- laptops in an LHC analysis group)*
      - ==> National Scale SSL (c.f. slides following)
2. **Analysis Challenge**, and 10 PB day long Grand Challenge (next slide)
3. Create a **scalable federated notebook service** to support interactive machine learning education and analysis
   a. *Based on recent discussions w/ AS, CERN IT*

Grand Challenges

As discussed at the NSF Review ("**analysis challenge**", c.f. response to review questions), and a grand challenge currently being discussed by CMS (**10 PB One Day Challenge**, cf here)

Inform SSL priorities

National Scale SSL - declarative, dynamic, flexible

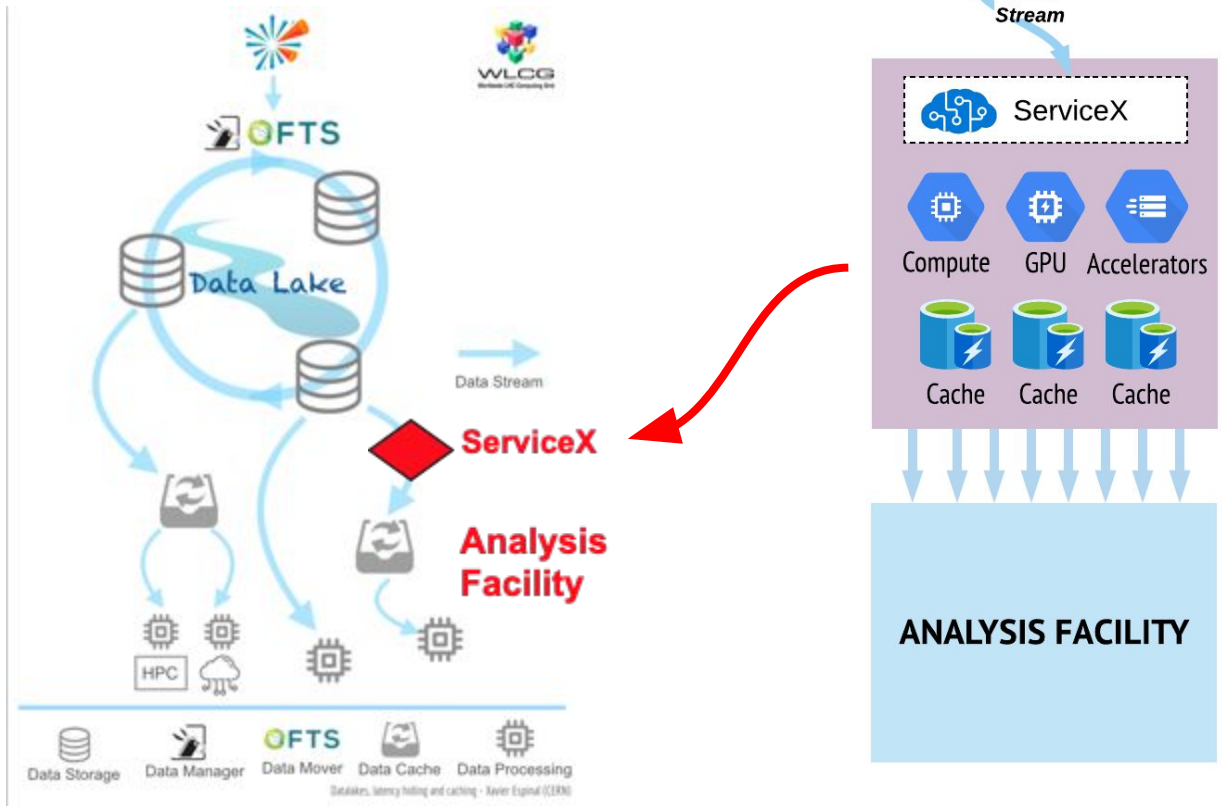
- **Hyperconverged k8s**

National Scale SSL

Capabilities:

- Platform for service deployment teams
- Primary focus on data delivery and analysis

Accelerator Testbeds



Apply hardware acceleration here

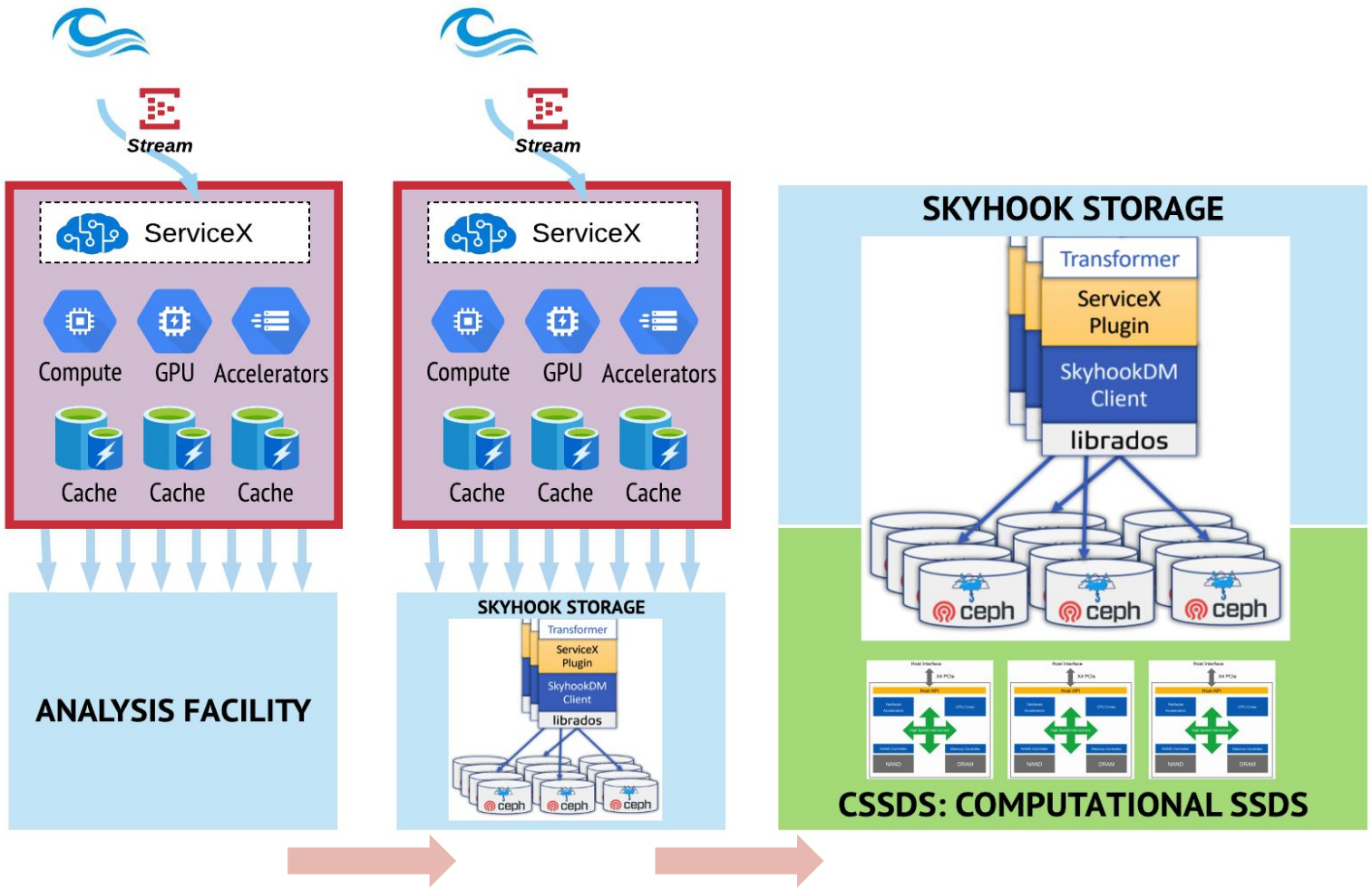
- Accept streamed ROOT data from WAN
- Accelerate decompression
- Branch extraction to columnar format
- Optionally apply user filters
- Box or Boxes

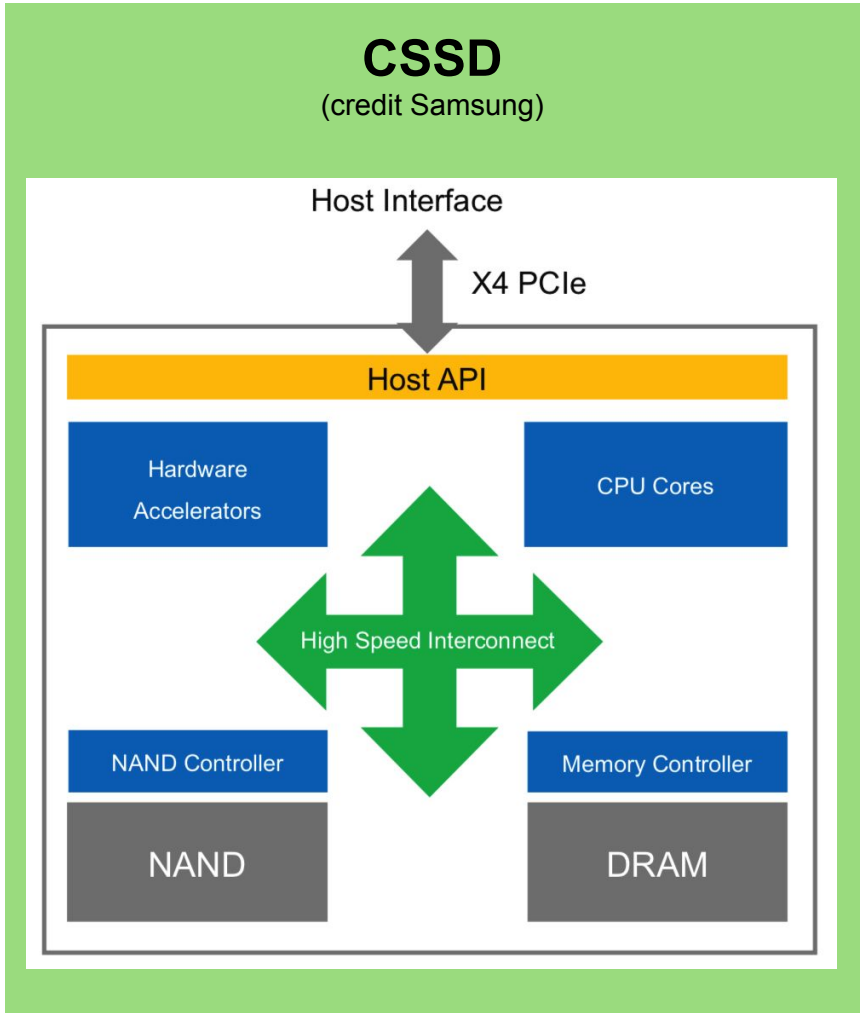
A possible 2 year goal

Demonstrate ingestion via ServiceX on a single host that achieves 50-100Gbps sustained ingest rate while doing the following:

- Uncompress root baskets and reformat into columnar data format to support much higher speed on reads later. Maybe add some indexing for fast retrieval if useful?
- Apply an event filter on keeping only events for which at least one object within a container of objects within a basket exceeds some user defined threshold.
  - Assume filtering such that O(10)% of events survive on output (5-50% should all work well enough).
- Apply an object list filter
  - User defined set of baskets are to be kept while all else is dropped.
  - This can be done even without uncompressing.
- Apply a simple algorythm like 4-vector addition or alike, possibly also more complex thing like JEC (apply a scalefactor to all objects in a specific container according to their place in a 2D map of eta-pT, i.e. object characteristic)
- Combine all three of the above.

Computational Storage with CSSDs

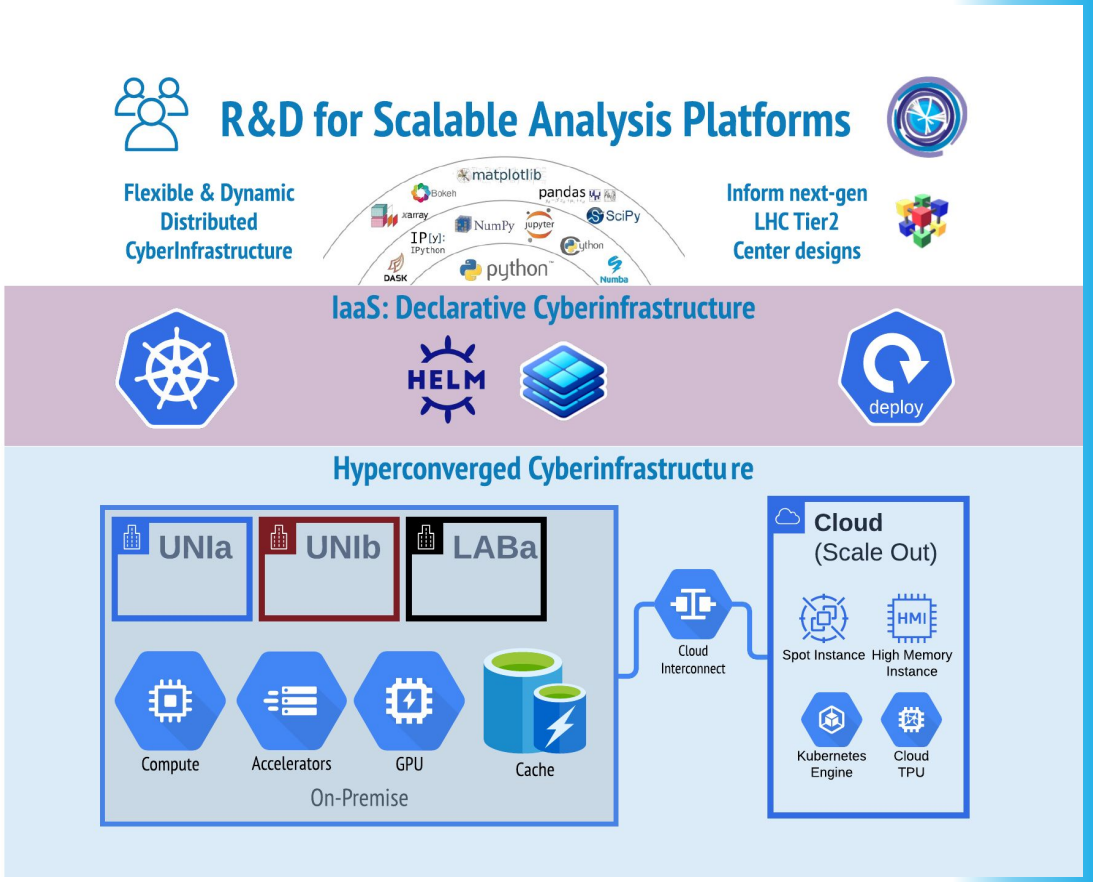




Are there new opportunities where effort from IRIS-HEP can make an impact? Is the alignment of the focus areas in IRIS-HEP appropriate?

Yes. Opportunities we're looking at:

1. R&D to evolve the LHC computing facilities to be more flexible, declarative, dynamic, and responsive to LHC physicists for scalable analysis platforms.
   - *inform by grand challenges*
2. Supporting the Snowmass effort with an analysis platform (JupyterLab access to GPUs, parallel frameworks: HTCondor/Ray/Spark).
   - *in progress*



Summary

This week we'd like to brainstorm during the parallel session about these potential Year 3,4 plans

Importantly we need to find R&D partners:

- provide k8s infrastructure
- support AS and DOMA development teams
- support early adopters (real users)
- move infrastructure patterns into production facilities