Analysis Grand Challenge

Alex Held (NYU)
Oksana Shadura (UNL)

IRIS-HEP / Ops Program Analysis Grand Challenge Planning

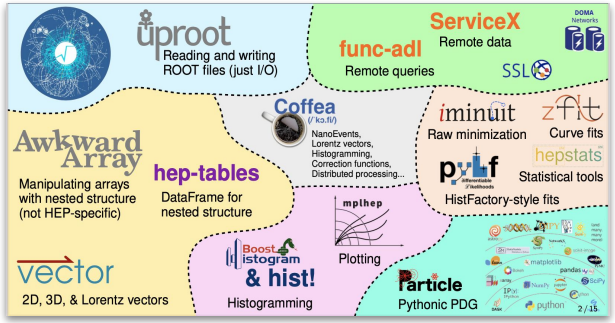
Jan 28, 2022: https://indico.cern.ch/event/1120687/

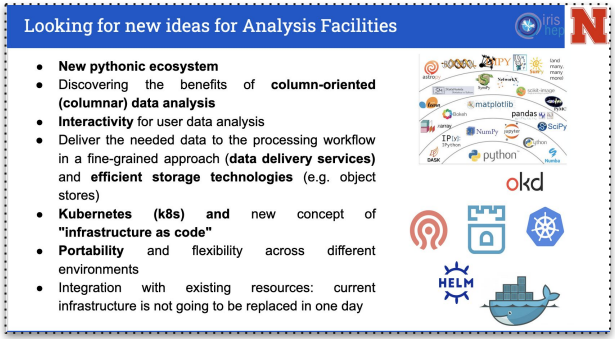
This work was supported by the U.S. National Science Foundation (NSF) Cooperative Agreement OAC-1836650 (IRIS-HEP).

Analysis Grand Challenge

Motivation:

- Allow coping with HL-LHC data sizes by rethinking data pipeline
  - Evaluating the new Python analysis ecosystem and integrating a differentiable analysis pipeline
- Provide flexible, easy-to-use, low latency analysis facilities



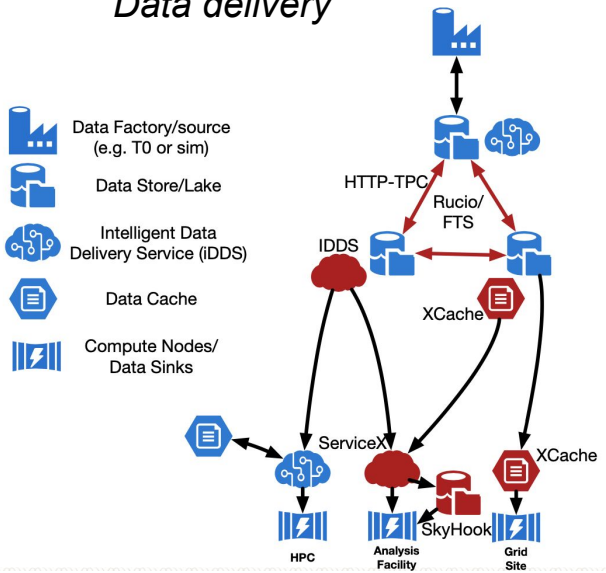


Coffea-casa vCHEP 2021 plenary talk

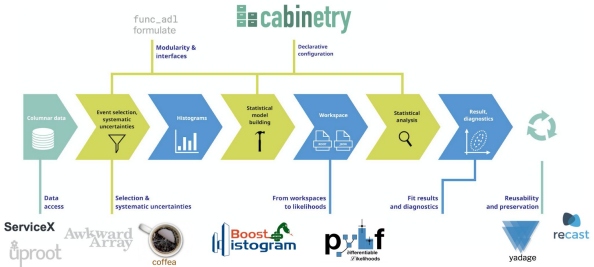
Analysis Grand Challenge will be conducted during **2021–2023**, leaving enough time for tuning software tools and services developed as a part of the IRIS-HEP ecosystem before the start-up of the HL-LHC and *organized together with the US LHC Operations programs, the LHC experiments and other partners.*

AGC is connecting different IRIS-HEP focus areas

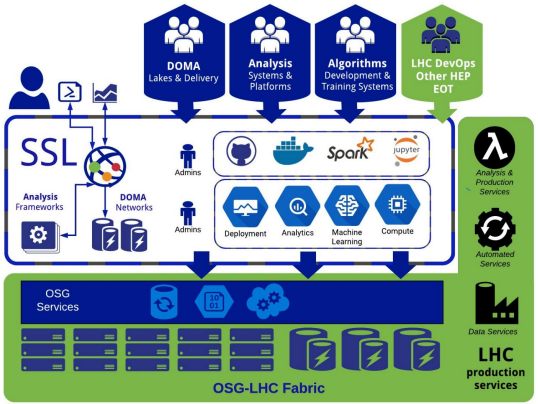
DOMA:
Data delivery



AS:
tools



SSL:
deployment techniques and resources



Activities

- Define target analysis to be used in the Analysis Grand Challenge
  - Working with smaller prototypes while building up to this
- Define and investigate programming interfaces between components
  - Investigate differentiable pipeline as a part of AGC
- Prototyping and deploying Analysis Facilities for executing Analysis Grand Challenge
- Coordinate with AS, DOMA, SSL, and operations programs to benchmark performance of prototype system components and to execute the Analysis Grand Challenge

Target analysis for AGC

Define target analysis to be used in the Analysis Grand Challenge

Towards a benchmark analysis

- Main AGC analysis example will be based on Run-2 CMS Open Data
  - CMS released ~400 TB of 2015 miniAOD at the end of 2021
  - Want to work with nanoAOD inputs
    - Results in realistic workflows as envisioned in ATLAS / CMS with PHYSLITE / nanoAOD
  - Expect to also be able to eventually include upcoming 2016 Open Data (will be miniAOD + nanoAOD)

- **Current efforts**
  - Converting miniAOD to nanoAOD: need to find suitable software versions, in contact with CMS experts
  - Searching for alternative to EOS for hosting (known bottleneck): SSL, other alternatives?

Analysis selection

- Previously we thought about extending the Run-1 Higgs->tautau example analysis
- New Open Data release provides a lot of flexibility
  - All major MC samples available for many analyses
  - Would like to do a generic search in a ttbar phase space
    - More familiarity with relevant objects / phase space / systematics / techniques
    - Possible synergies and collaboration with Swift-HEP / University of Manchester
- **Considerations**
  - Implications of developing a new analysis example? Possible issues with Open Data policies?
  - Would presumably model example analysis after a published CMS analysis

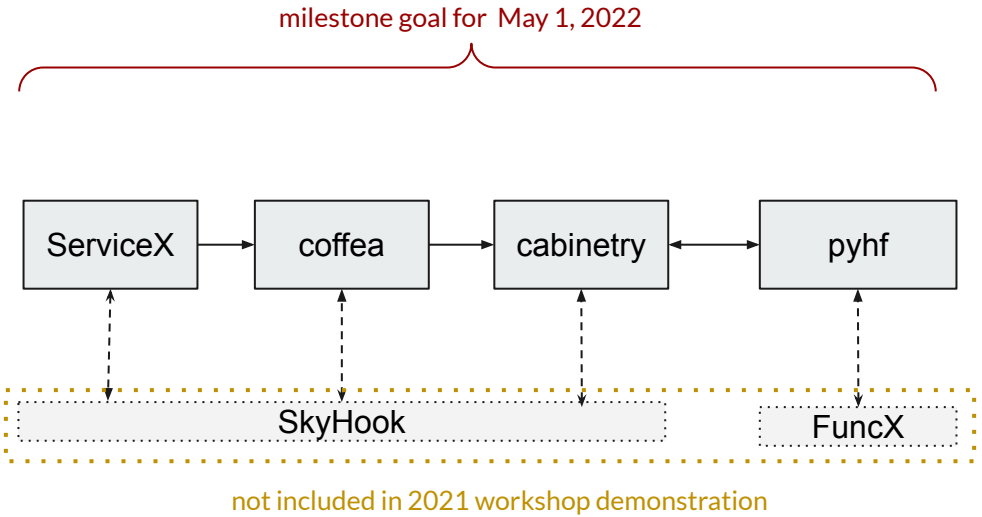
Software components for AGC

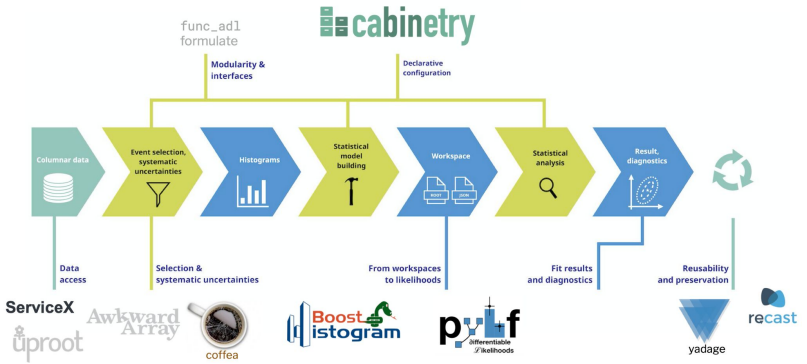
Define and investigate baseline programming interfaces between components

Expanding analysis pipeline: software components

Expanding existing analysis pipeline

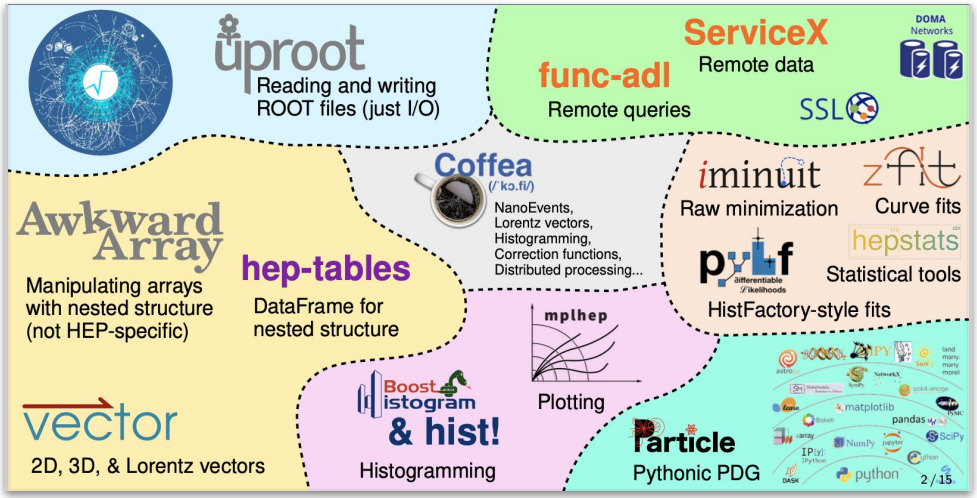
- Demonstration of **ServiceX ->Skyhook-> coffea -> cabinetry -> pyhf pipeline on Open Data**



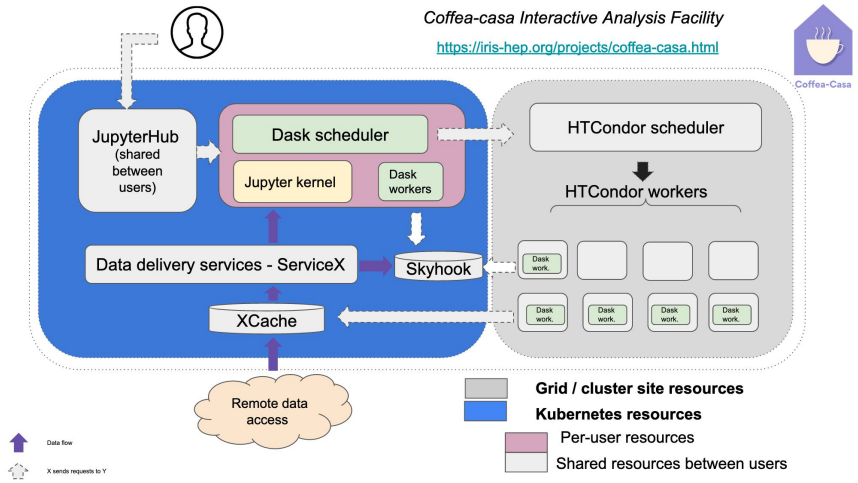


Building blocks used for designing AFs

Analysis Tools



Analysis Facilities



Future activities in this area

- Test existing interfaces and check if we need any cleanup
  - Simplify boilerplate code for *ServiceX-coffea* and *Skyhook-coffea* interaction
  - Brainstorm about possible *Cabinetry-coffea* integration
- Collect more feedback from users
- Preliminary benchmarks

Considerations

- Please let us know if there's some software or service you'd expect to see here that's missing?

Analysis Facilities for AGC

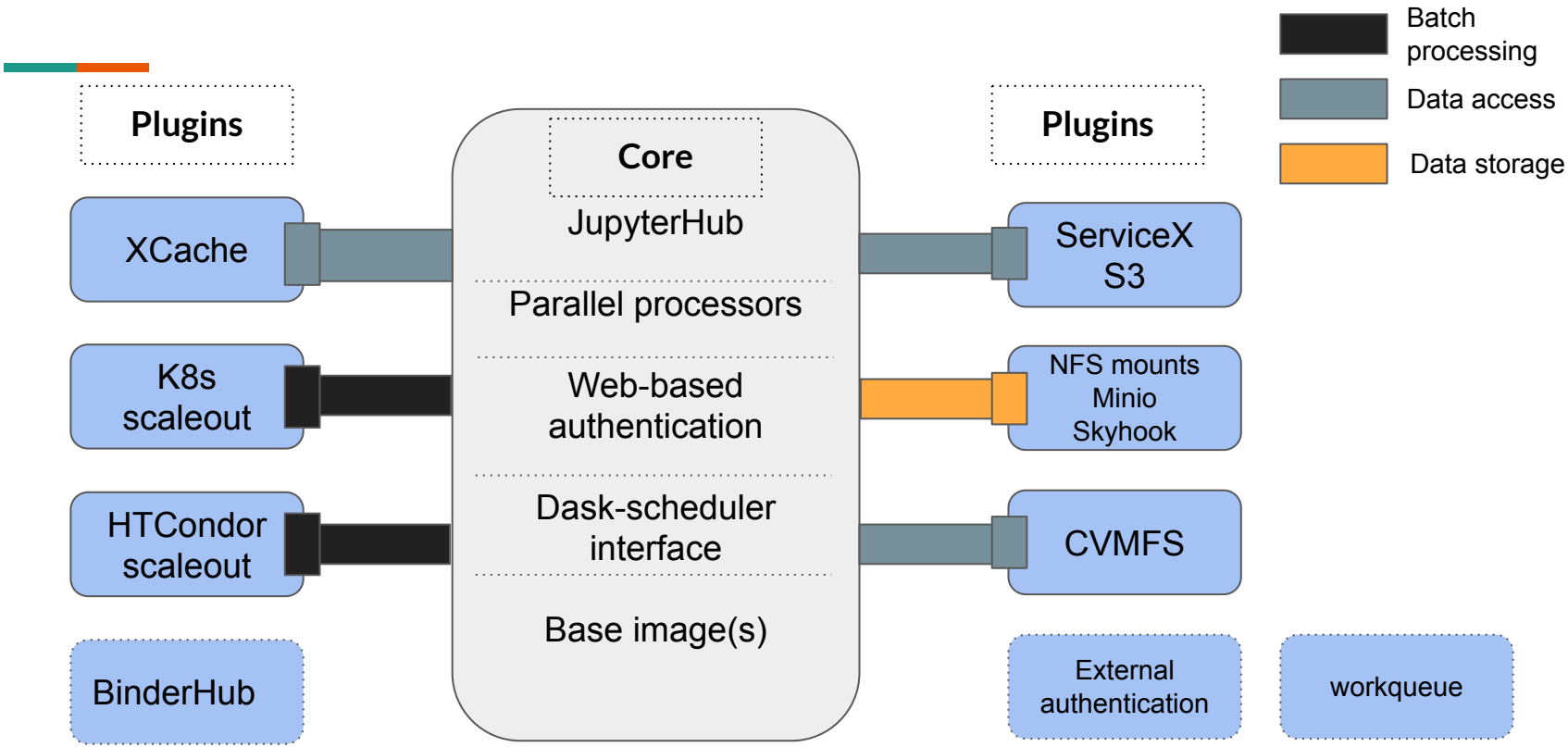
Prototyping and deploying Analysis Facilities for executing Analysis Grand Challenge

Requirements for AFs

Modern authentication (AIM/OIDC), tokens, macaroons, scitokens	Ongoing R&D on moving to use scitokens natively for AF (write/read)
Efficient data delivery and data management technologies	Ongoing work on integration ServiceX/Skyhook data delivery services
Columnar analysis and support new pythonic ecosystem	
Modern deployment and integration techniques	
Support for object storage	
Efficient data caching solutions	Integrating XCache in analysis pipeline
Easy integration with existing HPC resources	Looking to add support for other batch systems and task managements frameworks

[Coffea-casa team]

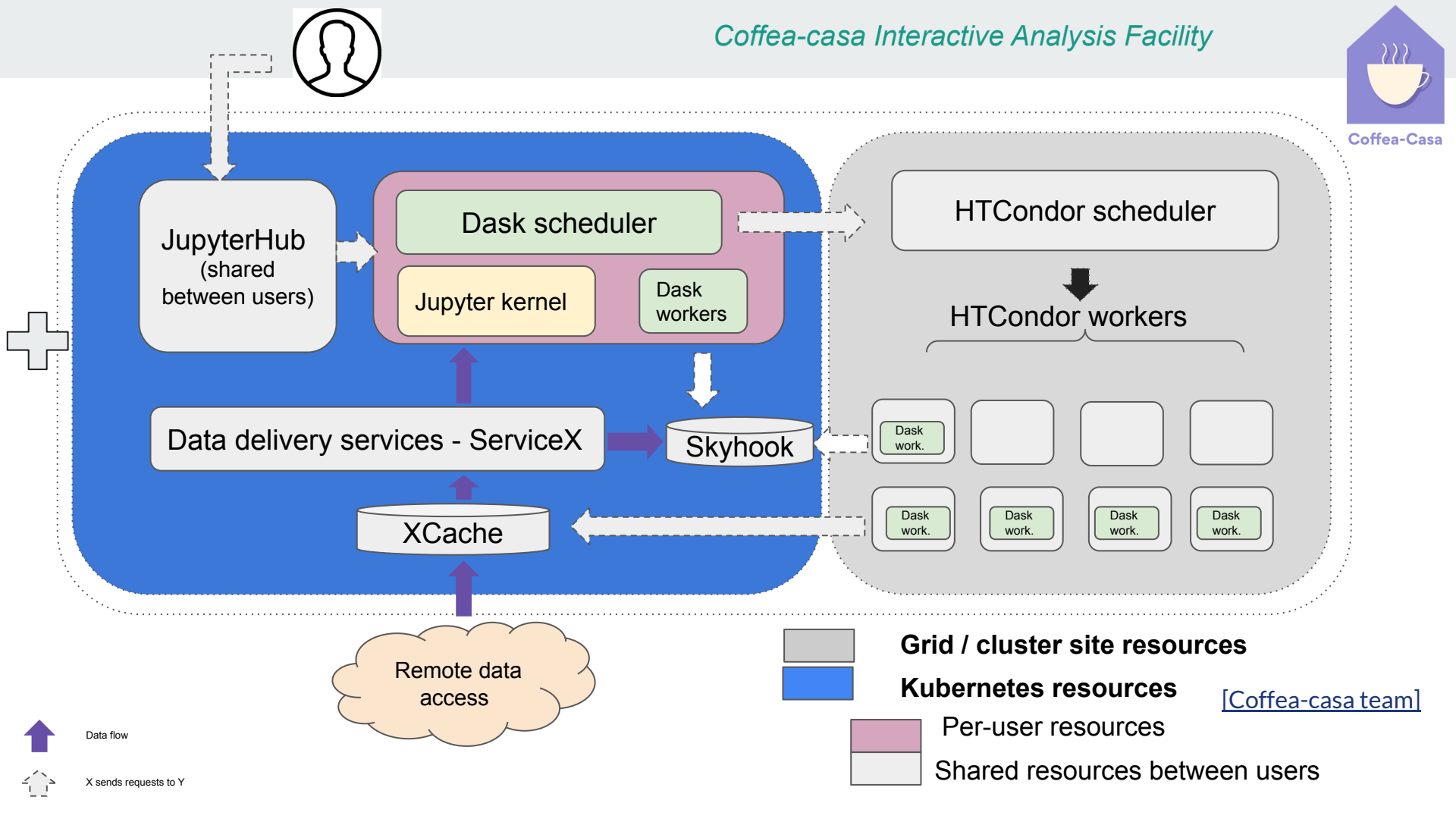
Designing AF: components of Coffea-casa Analysis Facility



[Coffea-casa team]

Coffea-casa Interactive Analysis Facility

Coffea-Casa



[Coffea-casa team]

Coffea-casa technical requirements

(Coffea-casa@ UNL is given here as an example)

CMSAF @T2 Nebraska
“Coffea-casa”
https://coffea.casa

OpenData AF @T2 Nebraska
“Coffea-casa”
https://coffea-opendata.casa

- Storage (CEPH via Rook.io) is on 4x Dell R710 nodes with 2x Xeon X5650 2.67GHz procs, 96GB RAM, 10Gb networking and 3x 1.92TB SSD each
- Old CPU nodes consist of various dual socket 4 and 8 core Opteron and Xeon CPUs with 2-4GB/core RAM and 1GbE networking
- 3x Modern CPU nodes are Dell R440 with Xeon Gold 6126 2.6GHz procs, 192GB RAM, and 10Gb networking.
- **Total “old” is ~256 cores of various ages and ~7TB triply replicated SSD CEPH storage**

  and
- 12x Dell R750 each with dual Xeon Gold 6348 28C/56T CPUs, 512GB RAM, 200Gb networking and 10x 3.2TB NVMe
- **Total “new” is 672 cores / 1344 threads and ~100TB triply replicated NVMe CEPH storage**

News:
- We deployed new hardware at UNL
  - Getting ready to move to new custer
- Deployed Skyhook on the old AF setup @ UNL for further functionality tests
  - Next step is a performance test (to be done on a new hardware)

[Coffea-casa team]

Coffea-casa deployments: existing coffea-casa AF

CMSAF @T2 Nebraska
“Coffea-casa”
https://cmsaf-jh.unl.edu

OpenData AF @T2 Nebraska
“Coffea-casa”
https://coffea-opendata.casa

ATLAS AF @Scalable System Lab (UChicago)
“Coffea-casa”

New facility with ATLAS IAM, setting this up generated valuable feedback for future coffea-casa developments.

Partners

Elastic AF @ Fermilab

Developed by: Burt Holzman, Maria Acosta (FNAL)

We are also in contact with BNL team to evaluate possibility to use coffea-casa experience at BNL facility

How we can involve more partners interested in AGC?

Coffea-casa deployments: possible coffea-casa AF

Got question about possible installation in 2022:
NYU and TU Munich
(in 2021 we got also request for testing coffea-casa AF from York University/DUNE)

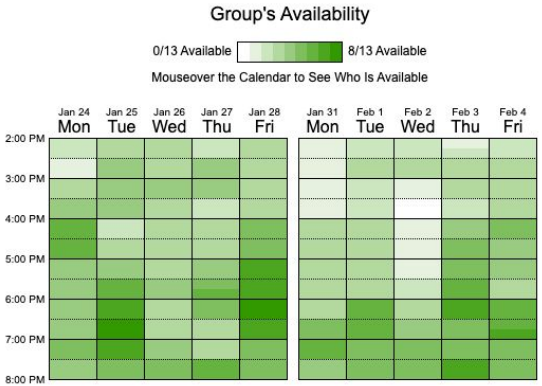
Collecting requirements since each setup has something unique :)

(we also need to understand if we can help with resources (Kubernetes) preparation step)

Future activities in this area

- Deploy Skyhook at Coffea-casa @ UNL and Coffea-casa @ SSL, test its integration with ServiceX through enabled POSIX persistent volume backend
- Test ServiceX performance with Open Data used in combination with XCache and Skyhook at SSL and UNL coffea-casa AF deployments
    - Brainstorm how to figure out how we can use efficiently shared cache for users

Next steps



- Had hoped to establish fixed **time slot for monthly meetings**
  - No ideal time found, maybe Tuesdays 18:00 CET / Friday 17:00 CET next time
  - Any other ideas for finding a suitable time?

- Planning **next AGC workshop** for April 25/26 or April 27/28
  - Format similar to previous workshop (https://indico.cern.ch/event/1076231/): 2 afternoons CERN time

AGC engagement and upcoming events

- Identified a need for **technical forum** for **analysis facility developments**
  - Mark (together with us) proposed a new HSF activity: HSF AF Forum
  - Iterating based on feedback about this idea from HSF coordination
  - Expect a blueprint meeting as kick-off
- Ongoing discussions with ESCAPE project https://projectescape.eu/
- **Effort to engage with more experiments interested in AGC**
  - Idea to organise event(s) to address needs of **experiments beyond ATLAS / CMS**
    - **Possibly together with AGCTools workshop**

Related IRIS-HEP Fellow proposals (see more here)

- **Enabling support for MiniAOD Transformer for ServiceX Data Delivery Service**: ServiceX is a distributed, cloud-native application that extracts columnar data from HEP event data and delivers it to an analyst. The func_adl data query language is used to tell ServiceX how to extract the data (the columns, simple cuts, etc.). The func_adl data query language has two backends that are currently part of ServiceX - one based on C++ for ATLAS data and CMS data, and one based on columnar processing using uproot and awkward arrays. The C++ backend currently runs only on the ATLAS binary format (xAOD) and CMS binary format (CMS AOD). This project will modify the C++ backend to also run on CMS MiniAOD binary files (available publicly as a part of Run 2 CMS Opendata release). The MiniAOD transformer is an important ingredient for a physics analysis workflow envisioned in the Analysis Grand Challenge. (Contact(s): Gordon Watts Ben Galewsky Oksana Shadura Alexander Held)

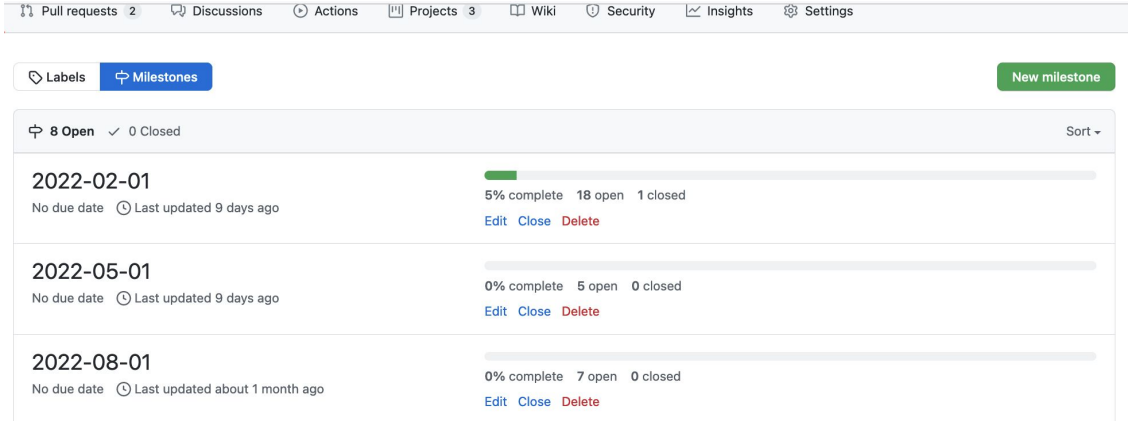
- **Benchmarking of prototype analysis system components**: The Analysis Grand Challenge of IRIS-HEP focuses on performing a high energy physics analysis at scale, including all relevant features encountered by analyzers in this context. It is performed using tools and technologies developed within both IRIS-HEP and the broader community, making use of the Python ecosystem and the required cyberinfrastructure to run at scale. This project will happen after a first preliminary benchmarking has been performed, and it will build on that: the prospective fellow will use pieces of an example physics analysis to study the performance of different system components in more detail. Fellows are expected to have prior Python experience and interest in working with a diverse stack of analysis tools available in the ecosystem. (Contact(s): Oksana Shadura Alexander Held)

- **Metrics to define user activities and engagement on the various coffea-casa Analysis Facility deployments**: coffea-casa is a prototype of analysis facility (AF), which provides services for "low latency columnar analysis", enabling rapid processing of data in a column-wise fashion. These services, based on Dask and Jupyter notebooks, aim to dramatically lower time for analysis and provide an easily-scalable and user-friendly computational environment that will simplify, facilitate, and accelerate the delivery of HEP results. The goal of the project is to define a set of various user engagement metrics, collected from Jupytehub and other AF services, as well from underlying infrastructure (e.g. Kubernetes) and available through Elasticsearch. Expected results are the development of the various metrics, a data collection infrastructure for them, and possibly visualization dashboards. (Contact(s): Brian Bockelman Ken Bloom Oksana Shadura)

Backup slides

Planned milestones for deliverables of coffea-casa AF

(outcome of coffea-casa dev meeting in November)



https://github.com/CoffeaTeam/coffea-casa/milestones

2022-02-01

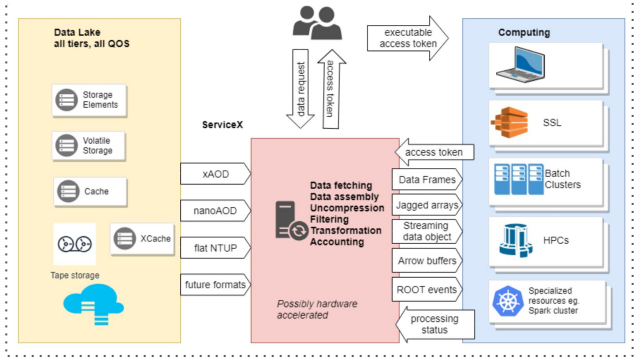
- New hardware deployed in UNL
- Skyhook deployment
- Production ready Helm charts
- Better integration testing
- Development documentation and many other items

Other significant milestones

- Coffea-casa AF will provide backup for couple of sessions at CMSDAS (first two weeks oj January)
  - We need to provide write capabilities for AF users (testing scitokens)

Data delivery services

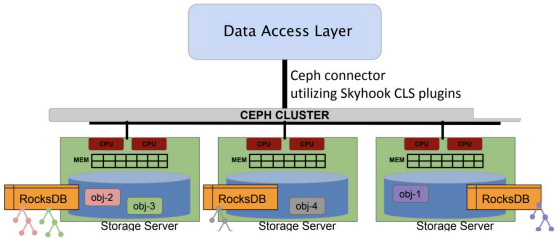
ServiceX



ServiceX provides user level ntuple production

- Converts experiment-specific datasets to columns (e.g. NanoAOD, DAOD)
- Enable simple cuts or simple derived columns and fields (*heavy-weight analysis will still happen via some separate processing toolchain*

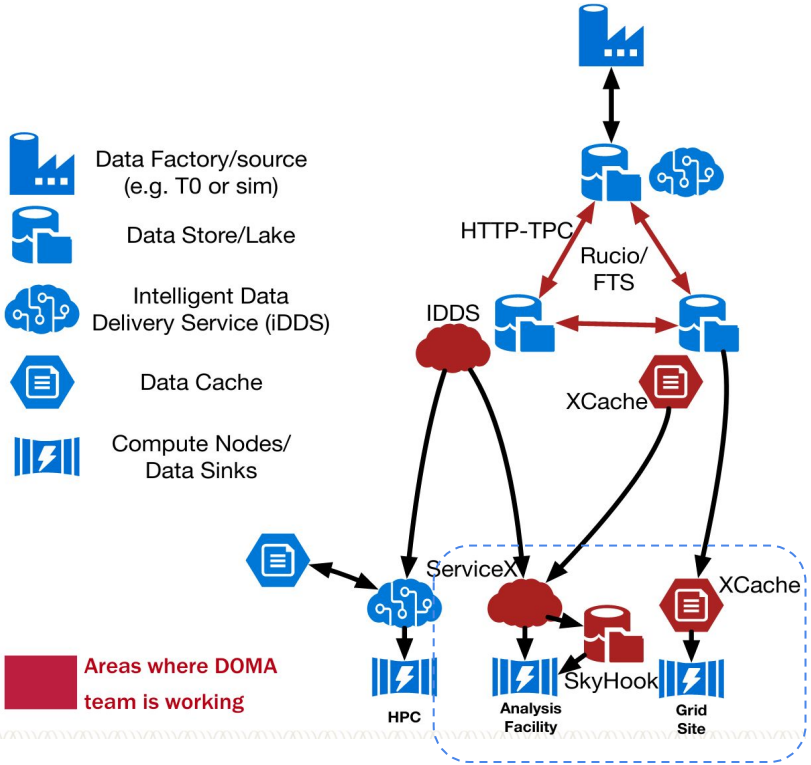
Skyhook DM



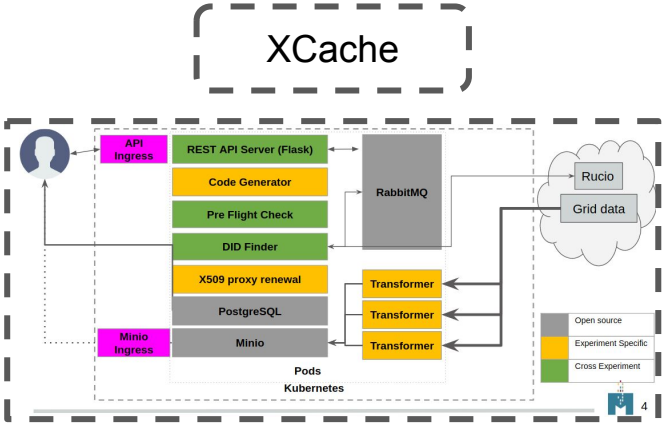
The Skyhook DM is converting event data from ROOT files to the internal object-store format

- Mechanism to access data kept in CephFS through the popular Arrow libraries
- Enables pushing down filters, projections, compute operations directly to the storage backend to minimise network overhead
- Allows writing files to a POSIX filesystem

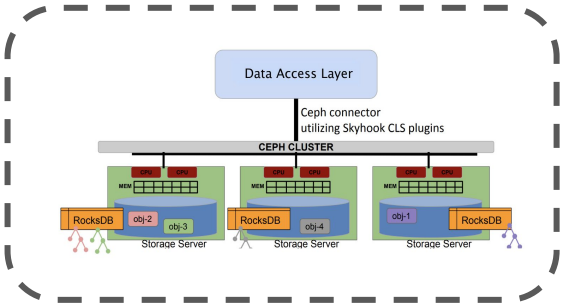
Analysis Facility and Distributed Ecosystem (Data Lakes)



Coffea-casa AF



Servicex



Skyhook

Current AGC milestones

- **Dec 1, 2021:** Demonstrate ServiceX -> coffea -> cabinetry -> pyhf pipeline (analysis-grand-challenge/issues/1)
  - Demonstrated at AGC workshop: https://indico.cern.ch/event/1076231/contributions/4560405/ ☑
- **Dec 1, 2021:** Execute IRIS-HEP AGC tools soft-launch event (analysis-grand-challenge/issues/2)
  - Done on Nov 3/4: https://indico.cern.ch/event/1076231/ ☑
- **June 1, 2022:** Coordinate with AS, DOMA, SSL, and operations programs to benchmark performance of prototype system components to be used for Analysis Grand Challenge (analysis-grand-challenge/issues/5)
  - In progress

Benchmarks

- **HSF DAWG** interested in expanding existing ADL benchmarks
- **HSF DAWG and AGC** identified several potential **directions for extensions**
  - *Testing interfaces between different tools in analysis pipeline*
  - *Handling of systematic uncertainties*
- **Idea:** to specify **AGC** sufficiently well so that it can be used as **very large benchmark**
  - Also want to split into **sub-tasks** that can be used for benchmarks
  - Detailed specification may attract other users to write new implementations
- Specification of *N* new benchmarks potential new **milestone for AGC**

Towards the next major milestone: June 1, 2022

- Improve experiment-related coffea-casa setups (e.g. improve experiment specific data access and other features)
    - Test integration of **SkyHook in *coffea-casa@UNL and SSL (as a testbed)***:
- Deploy and test all packages and services (e.g. related to AGC) at various analysis facilities
- Benchmark performance of prototype system components for AGC
- Work with HSF DAWG group about specification of new sub-benchmarks as a potential new milestone for AGC
- Develop analysis example used for next round of demonstration (based on new CMS Open Data)
    - Considering extending Run-1 Higgs->tautau example analysis