Analysis Grand Challenge

Alex Held (NYU)
Oksana Shadura (UNL)

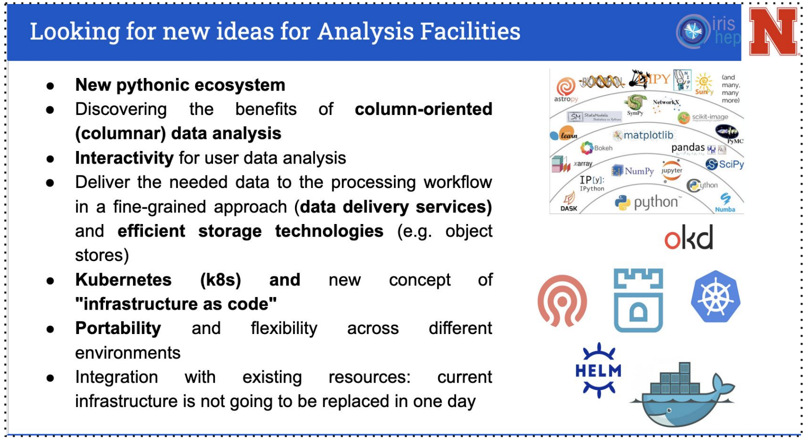
SwiftHep/ExcaliburHep workshop

https://indico.cern.ch/event/1033028/

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
- Provide flexible, easy-to-use, low latency analysis facilities

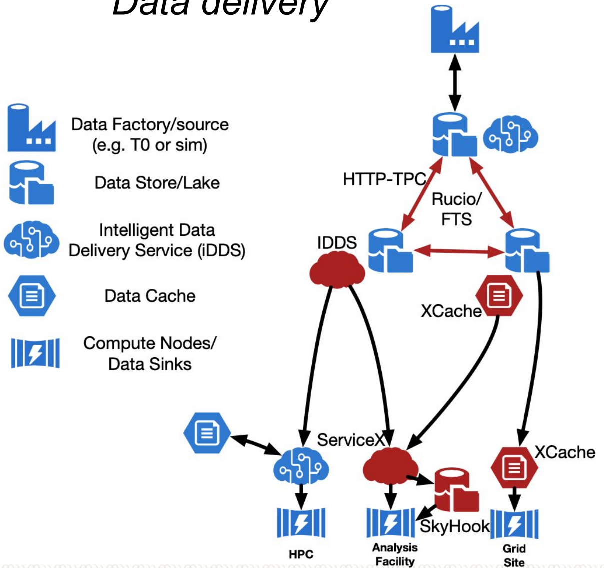


Coffea-casa vCHEP 2021 plenary talk

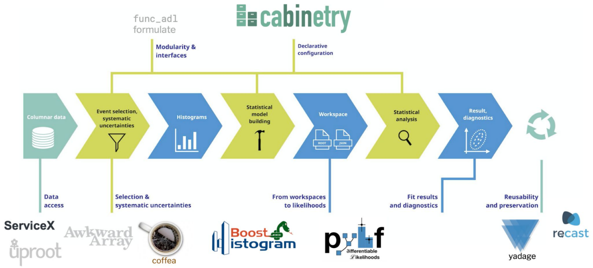
Analysis Grand Challenge will be conducted during **2021–2023**, leaving enough time for tuning software tools and services developed as a part of the IRIS-HEP ecosystem before the start-up of the HL-LHC and *organized together with the US LHC Operations programs, the LHC experiments and other partners.*

AGC is connecting different IRIS-HEP focus areas

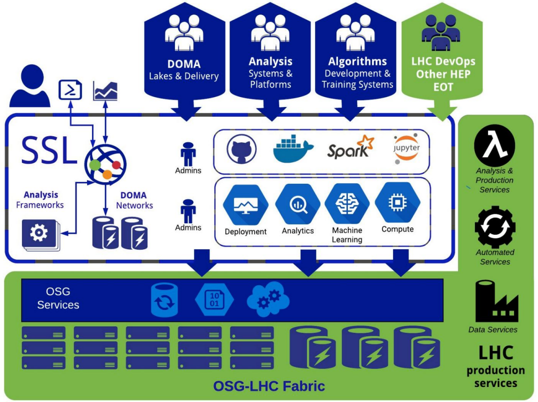
DOMA:
Data delivery



AS:
tools



SSL:
deployment techniques and resources



Activities

- Define target analysis and dataset to be used in the Analysis Grand Challenge
    - Building towards full analysis, now working with smaller prototypes
- Define and investigate baseline 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 to be used for Analysis Grand Challenge and to execute the Analysis Grand Challenge

Analysis Grand Challenge Analysis requirements

- **Analysis needs to include features encountered in realistic ATLAS/CMS analyses**
  - Handling of large data volumes (rough number we have in mind: ~*200 TB*)
  - Handling of different types of systematic uncertainties
  - Ideally use of modern formats *NanoAOD / PHYS / PHYSLITE*, which would make it easier for ATLAS/CMS analyzers to adopt to their use case
- **Not intended to send physics message with analysis**
  - Want to show realistic workflow, not make physics claims
  - No need for real data, simulation fully sufficient (ideally many samples to simulate book-keeping)
- **Want to demonstrate enhanced functionality**
  - Possibility to end-to-end optimize physics analysis, potentially via automatic differentiation
  - Analysis needs to run on analysis facility
- **Analysis needs to be sufficiently specified for others to re-implement**
  - Ideally: data is open and available to everyone (or scheduled to become public in the near future)
  - Hoping to learn from comparing to implementations developed by others outside IRIS-HEP
  - Want to turn parts of analysis into mini-benchmarks for facility and tool benchmarking

Using Open Data for the AGC

- Existing large datasets in Open Data restricted to CMS Run-1
- Would ideally prefer **modern ATLAS PHYS/PHYSLITE or CMS NanoAOD formats**
  - Makes AGC implementation more relevant to current/future analyses & ideally re-usable
- Following up with ATLAS & CMS to understand whether we may be allowed to use (a) new dataset(s) for technical demonstration

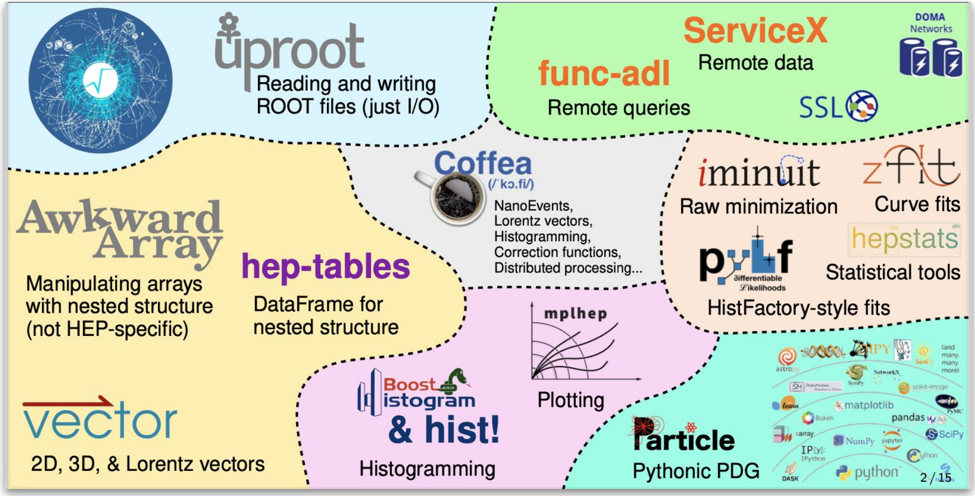
Benchmarks

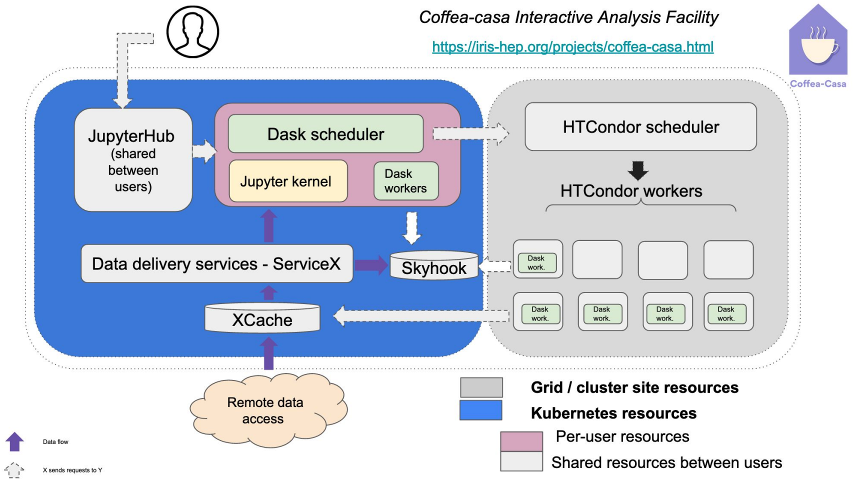
- **HSF DAWG** interested in expanding existing ADL benchmarks
- **HSF DAWG and AGC** identified several potential **directions for extensions**
  - *Testing interfaces between different tools in analysis pipeline*
  - *Handling of systematic uncertainties*
- Idea: to specify **AGC** sufficiently well 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**

Building blocks used for designing AFs

Analysis Tools

Analysis Facilities

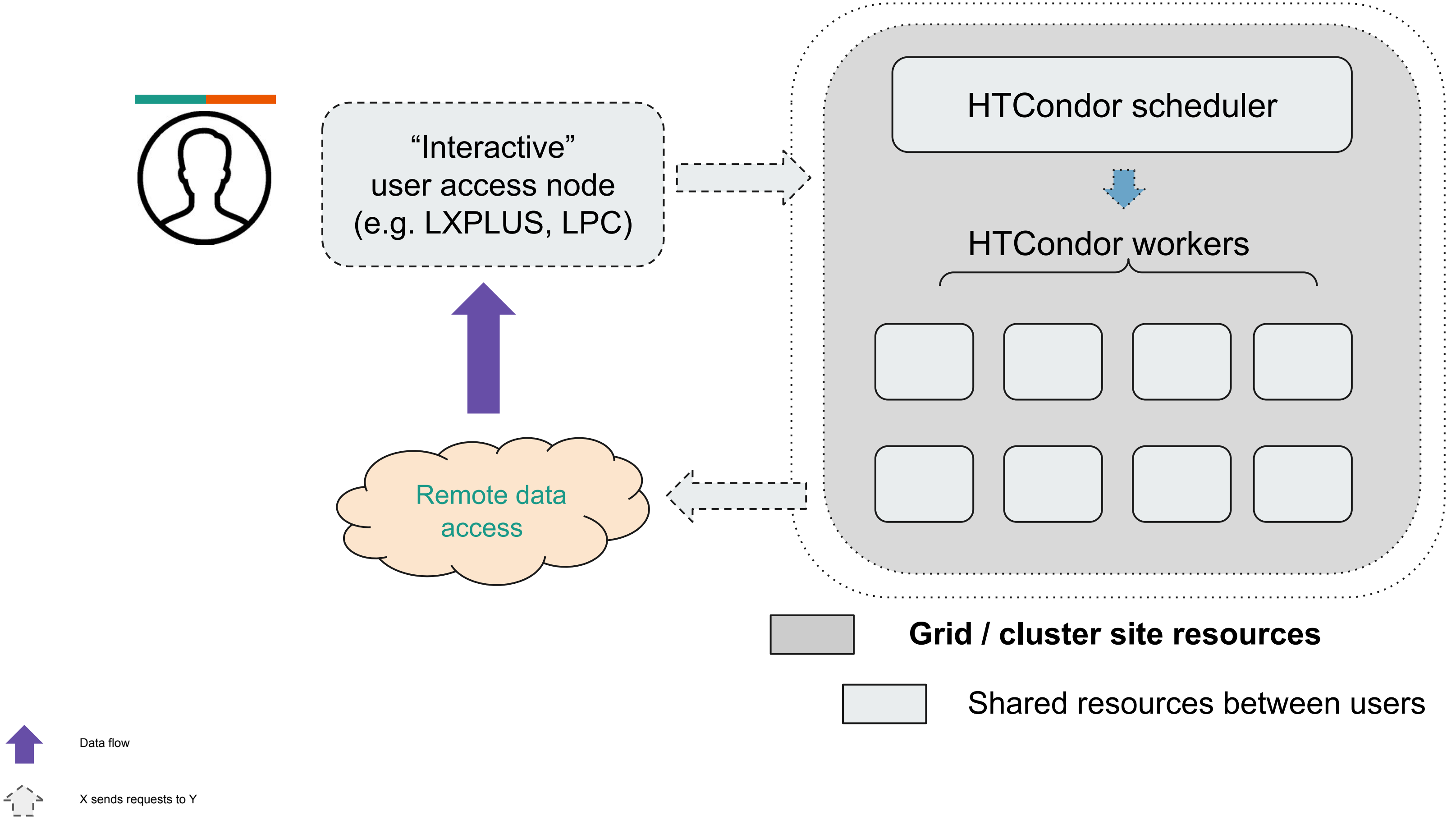


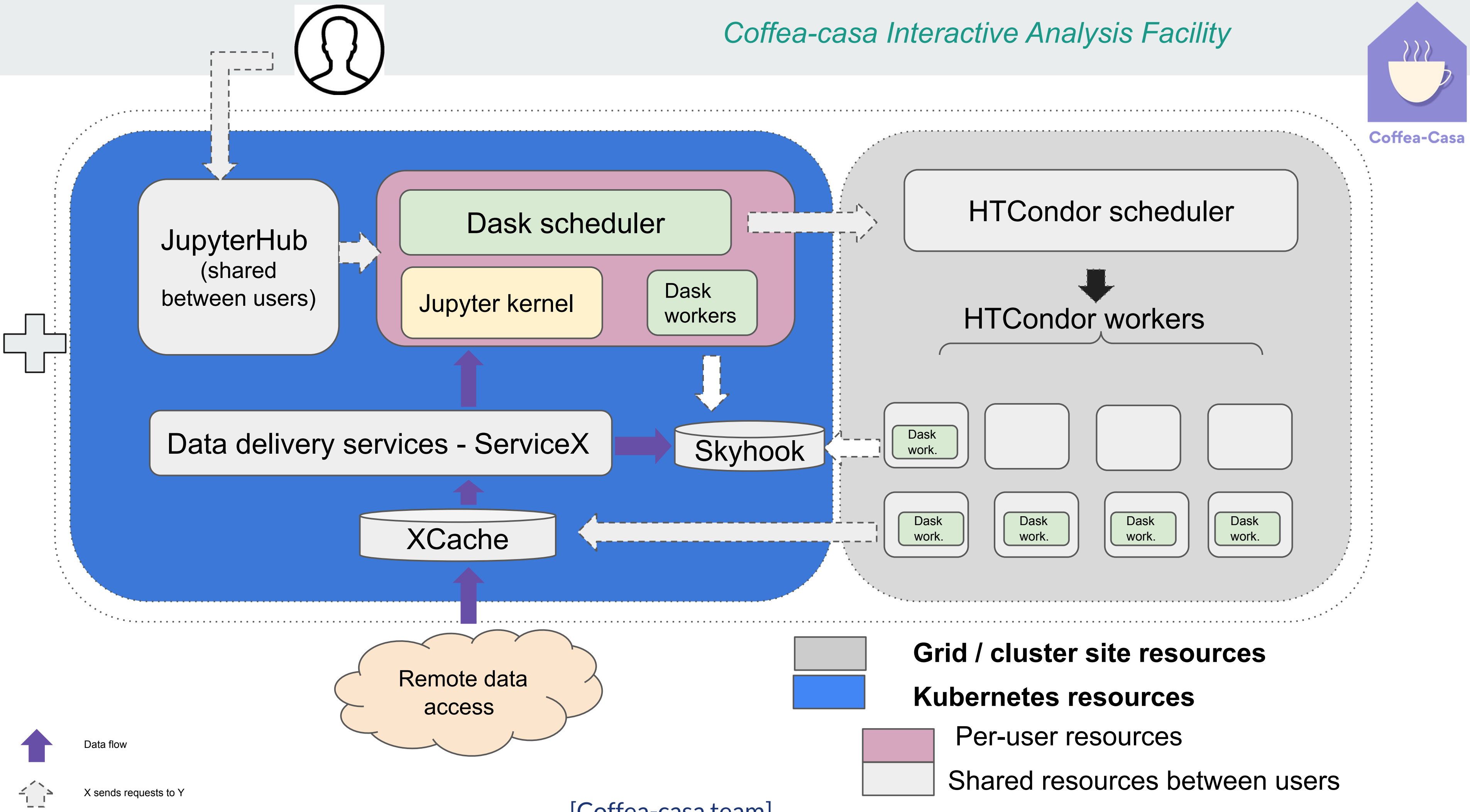


Building blocks used for designing AFs

- Modern authentication (AIM/OIDC), tokens, macaroons
- Efficient data delivery and data management technologies
- Columnar analysis and support new pythonic ecosystem
- Modern deployment and integration techniques
- Support for object storage
- Efficient data caching solutions
- Easy integration with existing HPC resources

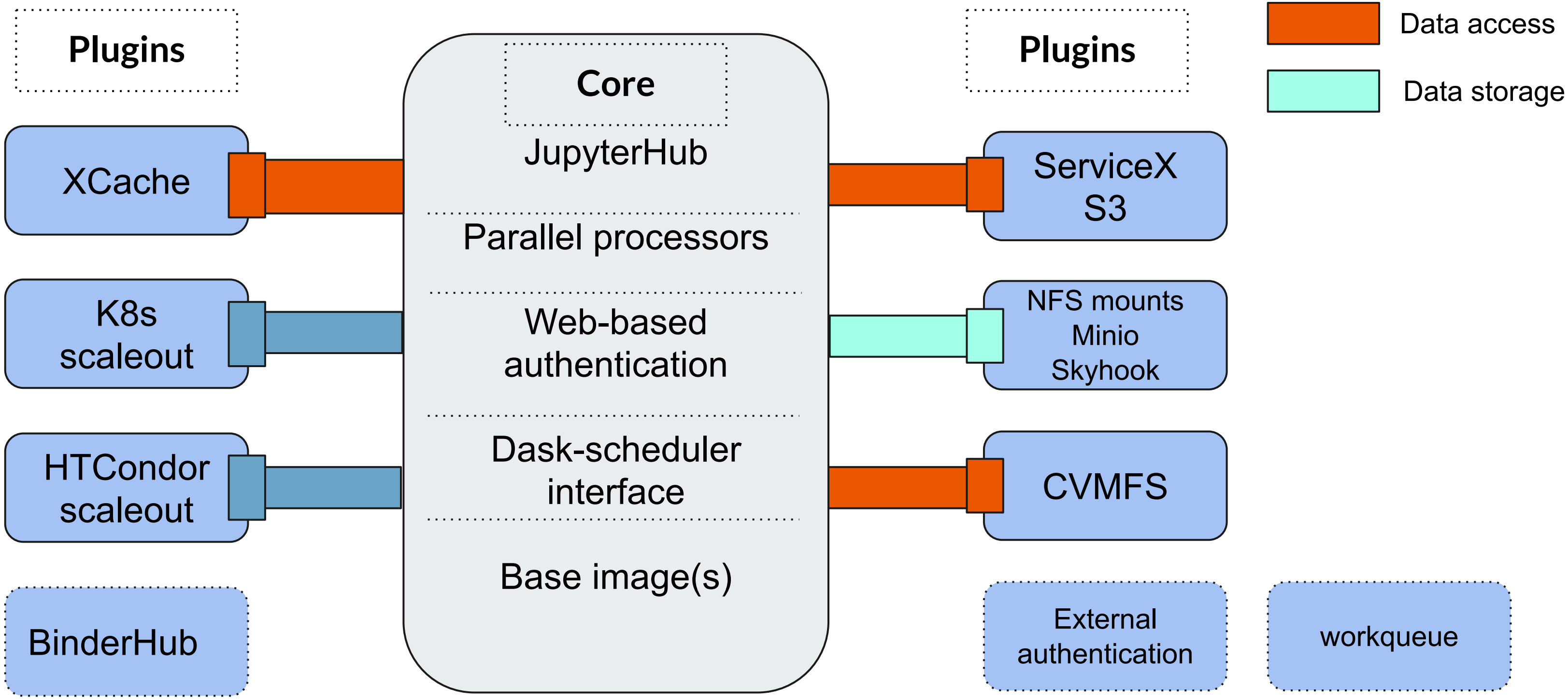
Simplified diagram of hypothetical Analysis Facility currently used by users





[Coffea-casa team]

Coffea-casa components



[Coffea-casa team]

Coffea-casa deployments: existing coffea-casa AF

Coffea-Casa

- Coffea-casa style AF facilities with the possible outcome of adding more sites as soon as we gain experience

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"
(coming very soon)

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

Coffea-casa technical requirements

Coffea-Casa

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

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

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**

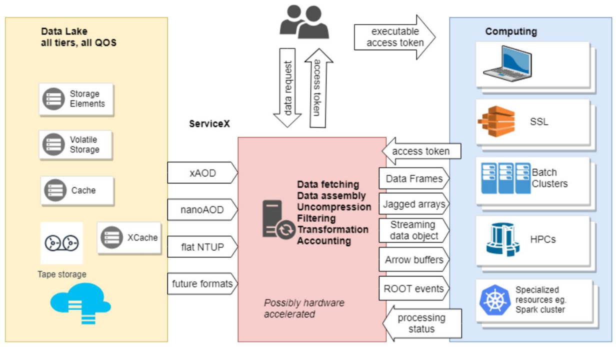
and

- 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**

[Coffea-casa team]

Data delivery services

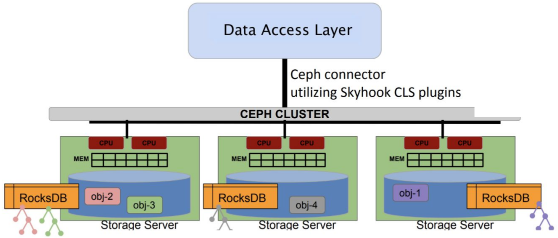
ServiceX



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

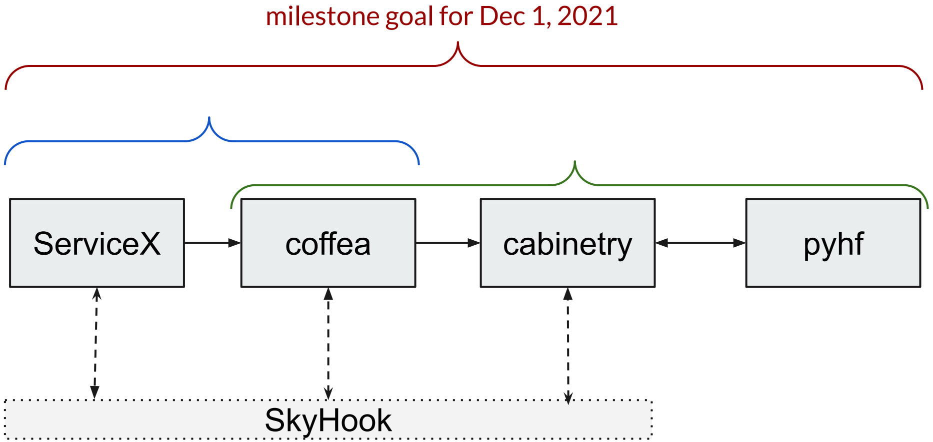


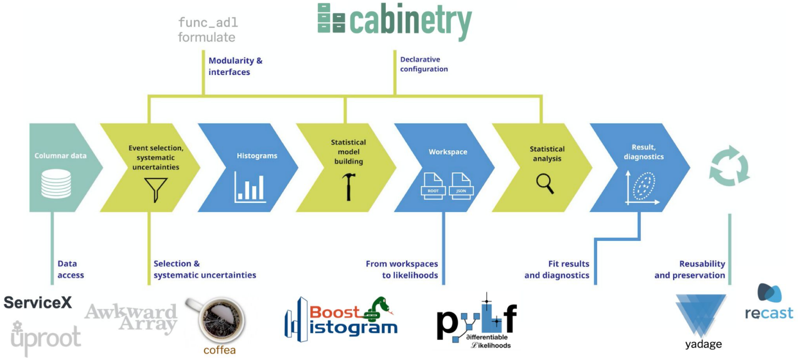
The Skyhook DM is converting event data from ROOT files to the internal object-store format

- 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

Ongoing activities: building analysis pipeline

- Demonstration of **ServiceX -> coffea -> cabinetry -> pyhf pipeline on ATLAS Open Data** (ATLAS H>ZZ* open data example)
  - Testing API developments: its compatibility and user friendliness





IRIS-HEP AGC Tools 2021 Workshop 3–4 Nov 2021

- Training event showing **IRIS-HEP toolchain at *coffea-casa***, aimed at PhD / postdoc level
  - Format: brief introduction to individual packages, notebook talks focusing on interfaces between tools
  - Initially using Open Data examples, then splitting into ATLAS / CMS - specific tracks
  - 2 afternoons CERN time (15:30 - 19:30) on Nov 3/4 (after US ATLAS computing bootcamp scheduled for Oct 18-22)

Check out **IRIS-HEP AGC Tools 2021 Workshop 3–4 Nov 2021**
https://indico.cern.ch/event/1076231/

Backup slides

AGC analysis definition: proposed approach

In 2021

- stick to **ATLAS H>ZZ*** example for demonstrations, allows testing interfaces and is ready
- identify possibility of using datasets with new ATLAS / CMS formats in parallel
    - May involve re-formatting existing Open Data if no new datasets become available

June 2022 milestone (“benchmark performance of prototype system components for AGC”)

- If **modern ATLAS / CMS formats** are available: design analysis around what samples we get
- Otherwise: extend **CMS H>tautau Open Data** analysis (Open Data record) with systematic uncertainties covering all uncertainty types identified in taxonomy (different types requiring different approaches)

March 2023 milestone (execute AGC)

- If **new ATLAS / CMS data** is unavailable: (reformatted) **CMS Run-1 Open Data**-based analysis