Analysis Grand Challenge

Alex Held (NYU)
Oksana Shadura (UNL)

IRIS-HEP Executive Board / Ops Program Grand Challenge Discussion
https://indico.cern.ch/event/1105768/
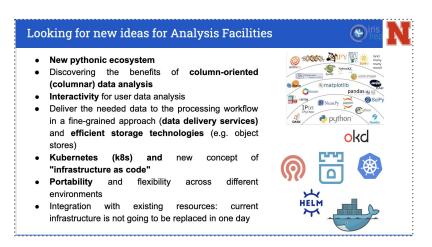
Analysis Grand Challenge

Motivation:
- Allow coping with HL-LHC data sizes by rethinking data pipeline
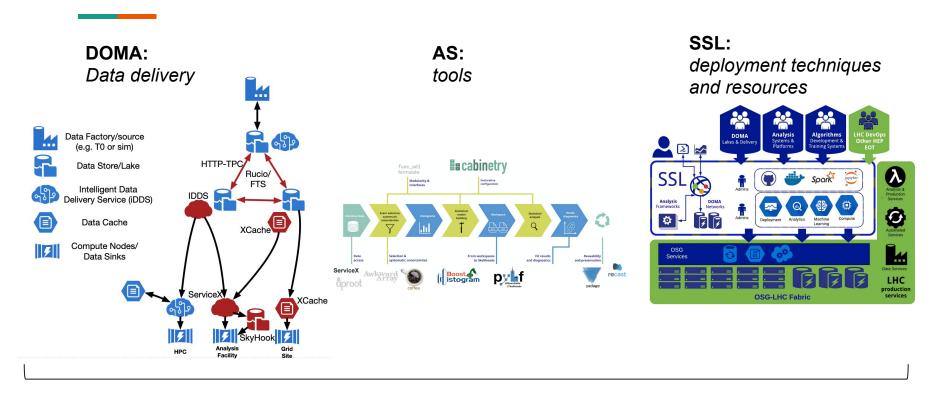- Provide flexible, easy-to-use, low latency <u>analysis facilities</u>

[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.*
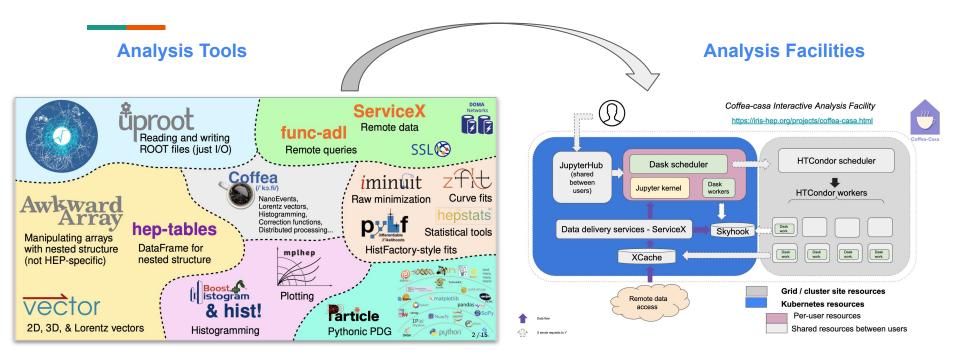
AGC is connecting different IRIS-HEP focus areas

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

Building blocks used for designing AFs

Analysis Tools

Analysis Facilities

Requirements for AFs

Modern authentication (AIM/OIDC), tokens, macaroons, scitokens

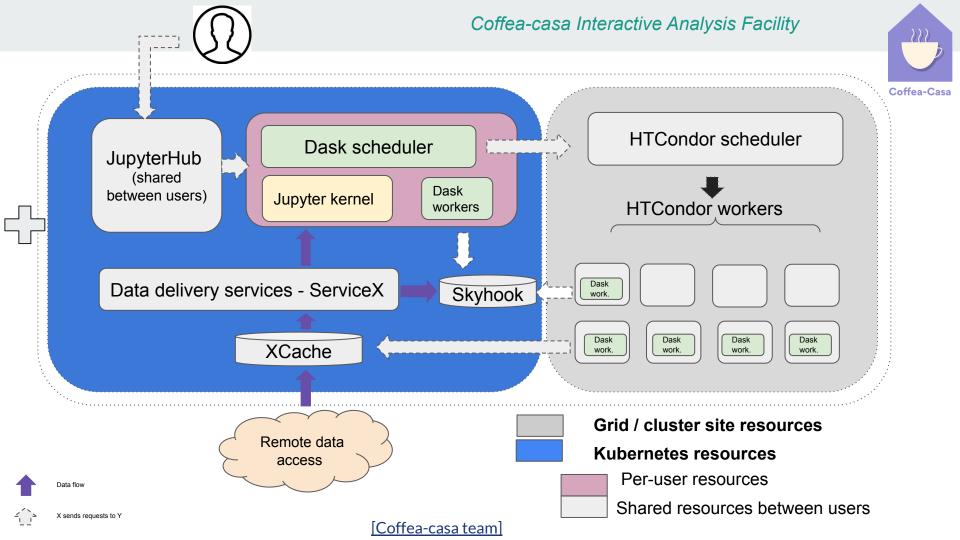Efficient data delivery and data management technologies

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

Coffea-casa Interactive Analysis Facility

Coffea-Casa

JupyterHub (shared between users)

Dask scheduler

Jupyter kernel

Dask workers

HTCondor scheduler

HTCondor workers

Dask work.

Dask work.

Dask work.

Dask work.

Dask work.

Data delivery services - ServiceX

Skyhook

XCache

Remote data access

Grid / cluster site resources

Kubernetes resources

Per-user resources

Shared resources between users

Data flow

X sends requests to Y

[Coffea-casa team]

Designing AF: components of Coffea-casa Analysis Facility

Core

JupyterHub

Parallel processors

Web-based authentication

Dask-scheduler interface

Orchestration

Data access

Scaling

Storage

Coffea-Casa

Overview U.S. CMS Ops WBS4: HL-LHC R&D

U.S. CMS Operations Program

4.1 Analysis Systems

Develop tools and analysis systems for HEP that enable both innovation and the adoption of "industry standard" analytic techniques. Enable rapid interactive analysis of PB datasets.

Grand Challenge: Transform the Scientific Data Analysis Process.

4.1.1 Tools for Advanced Analysis

Provide interfaces and infrastructure to adapt HEP data in order to enable rapid analysis.

- Princeton [Awkward, Uproot, etc.]

Close collaboration with
- CCE / SciDAC / IRIS-HEP, ATLAS
- OSG / WLCG / HSF
- DOE-ASCR, NSF-*

True for entire R&D Program

4.1.2 Analysis Facilities

Prototype and put into production the infrastructure required for a multi-user analysis facility exploiting the tools developed in 4.1.1

- FNAL, UNL*, Purdue*

* Special GPU/hardware investments in 2020 & 2021 to provide heterogeneous resources.

(Note: efforts listed on this and the following slides are only those that receive direct support from the OPS program. Lots of other work is happening!)

Hildreth   USCMS Collaboration Meeting: U.S. CMS HL-LHC Software & Computing R&D

Coffea-casa deployments: existing coffea-casa AF

Coffea-Casa

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.

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 deployments: possible coffea-casa AF

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

We need to understand how we can help with resources (Kubernetes) preparation step

Coffea-casa technical requirements

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

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

OpenData AF @T2 Nebraska
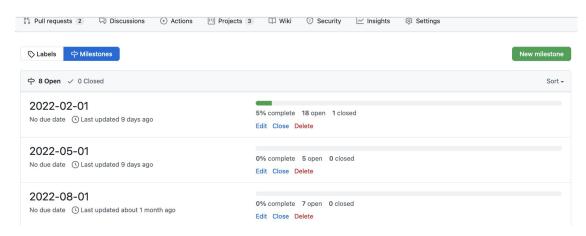**"Coffea-casa"**
https://coffea-opendata.casa

- 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**

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
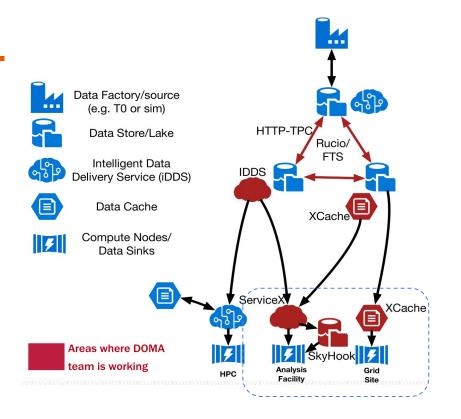- Development documentation and many other items
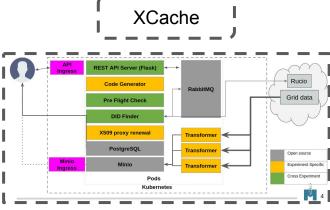
Other significant milestones
- Coffea-casa AF will provide backup for couple of sessions at CMSDAS (first two weeks oj January)
  - We need to provide write capabilities for AF users (testing scitokens)

Data Factory/source (e.g. T0 or sim)

Data Store/Lake

Intelligent Data Delivery Service (iDDS)

Data Cache

Compute Nodes/ Data Sinks

Areas where DOMA team is working

HTTP-TPC

Rucio/ FTS

IDDS

XCache

ServiceX

XCache

HPC

Analysis Facility
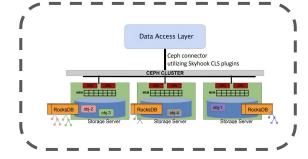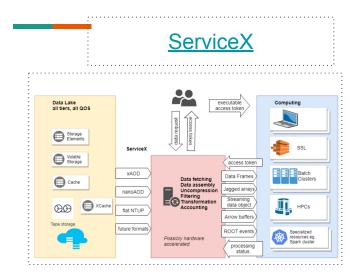
Grid Site

SkyHook

Coffea-casa AF

XCache
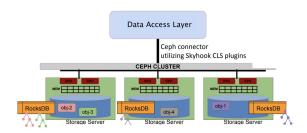
Servicex

Skyhook

14

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

Preparing to use new CMS Open Data for the AGC

- New release of CMS Opendata 2015 imminent (end of this year)

  - AODs + MiniAODs (data) and MiniAODs (simulation)

  - Many TBs per dataset, ideal input for next steps: scaling up and benchmarking

- Currently investigating how to handle MiniAODs:

  - Considering conversion to NanoAOD as fellow project

  - Could also add feature to ServiceX to serve columns directly from MiniAODs

Organizational points

- This was the first AGC meeting split out from the usual EB / Ops Program time slot
  - Idea: more time for discussions without taking over all of the time of the other meeting

- Planning to continue with ~monthly meetings in this new format
  - Will send another poll at the beginning of 2022, and try to find a time slot that works for the majority on a monthly basis

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

AGC engagement and upcoming events

- Identified need for **technical meeting** to **coordinate analysis facility developments**
  - Many similar efforts ongoing at different sites / groups
  - Aiming for meeting in early 2022

- Dedicated **tutorial / training event**: lower entry level, aimed at PhD students new to tools
  - Synergies with HSF training efforts & IRIS-HEP training grand challenge

- Another **showcase event** around spring / summer, possibly aligned with next milestone
  - Could be combined with training event

- Ongoing discussions with ESCAPE project https://projectescape.eu/

- **Effort to engage with more experiments interested in AGC**
  - Organise event(s) to address needs of **experiments beyond ATLAS / CMS**
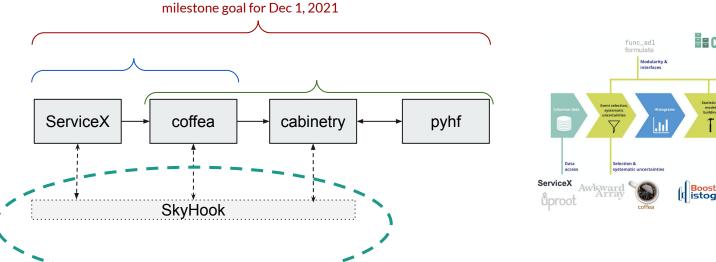 - Synergies with HSF, we are particularly interested in how our tools / workflows may map to other use cases

Towards the next major milestone: June 1, 2022

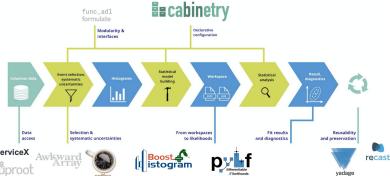- Improve experiment-related coffea-casa setups (e.g. improve experiment specific data access and other features)

  - 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

Backup slides

Expanding 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

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**