AF forum - News

15th September 2022

Mailing list: HSFAFFORUM

MatterMost: hsf-af-forum

Alessandra Forti (WLCG/ATLAS)   Diego Ciangottini (CMS)

Lukas Heinrich (ATLAS)   Nicole Skidmore (LHCb)

Welcome back!

News

- Introduction to AF forum presented at *EOSC-Future ESCAPE Science Projects progress meeting* (July)
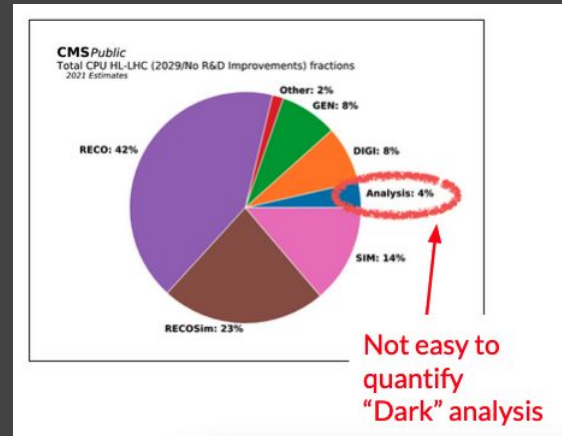 - Well received - possible collaborations in the future?

News

- Ecosystems workshop report is now complete   [Zenodo link](#)
  - Large focus on AFs with proposals for how to move forward

Second Analysis Ecosystem Workshop Report

23-25 May 2022, IJCLab Orsay, Paris

Editors: Peter Elmer (Princeton University), Giulio Eulisse (CERN), Loukas Gouskos (CERN), Stephan Hageboeck (CERN), Allison Reinsvold Hall (US Naval Academy), Lukas Heinrich (Technische Universität München), Alexander Held (University of Wisconsin–Madison), Michel Jouvin (IJCLab), Teng Jian Khoo (Humbold-Universität zu Berlin), Paul Laycock (Brookhaven National Laboratory), Pere Mato Vila (CERN), Jim Pivarski (Princeton University), Jonas Rembser (CERN), Eduardo Rodrigues (University of Liverpool), Jana Schaarschmidt (University of Washington), Elizabeth Sexton-Kennedy (Fermilab), Oksana Shadura (University Nebraska-Lincoln), Nathan Simpson (Lund University), Nicola Skidmore (University of Manchester), Michael Sokoloff (University of Cincinnati), Graeme A. Stewart (CERN, Lead Editor), Gordon Watts (University of Washington)

Facilities publish key metrics related to performance and usage. It is initially envisaged to perform regular user surveys to capture the success of analysis facilities, focusing on user-oriented or user engagement metrics. We can also use this to understand user modalities, preferences and analysis throughput to guide AF development.

CMS Public
Total CPU HL-LHC (2029/No R&D Improvements) fractions
2021 Estimates

- Other: 2%
- GEN: 8%
- RECO: 42%
- DIGI: 8%
- Analysis: 4%
- SIM: 14%
- RECOSim: 23%

Not easy to quantify "Dark" analysis

Next topics

22nd September

　　　Efficiently distributing images via CVMFS with DUCC

29th September (User experience)

　　　NAF @ DESY

　　　3 part talk on user experience from Belle, ATLAS and CMS

On the wish list for next months: design of user-survey, AGC report, Reana, SwiftHEP report (Nov), Caches and federated ID management with both a user and technical POV

Any more to add? :)

AE II topics

- Aim of this forum is to build a community but also produce recommendations
- AE II workshop we agreed to follow up a list of topics to (hopefully) produce this recommendations following up the more **practical** aspects
- Need to discuss today **HOW** to follow this up
- Preference would be to have groups of people willing to put sometime for each topic to follow also in other forum and then create a coherent picture from AF point of view
  - Some topics needs tight cooperation with other projects WLCG/IRIS-HEP/EOSC/… or HSF groups

AE II topics

- Broad topics are in the AE II report
  - Federated identity management
  - DOMA topics
  - Resource sharing
  - Environment sharing
  - Tracking analysis performance
  - AF
- Plus
  - User experience/new analysis techniques (?)
  - IRIS-HEP AGC

Federated identity management

- Federated identity management, identity providers, tokens
- This is followed in two forums: WLCG authz and DOMA BDT
  - Current status is that the infrastructure is still being built: i.e. things are not definitive and not ready yet for AF
  - Storage is the guinea pig (hence DOMA)
  - Infra not ready for AF
- AF are currently being built around technologies that are better suited to tokens than x509
  - Some places like Nebraska already use tokens but not WLCG ones
- Need a couple of presentations with people working on this to start interaction and volunteer sites to follow this up from AF PoV
- Help build the documentation too

DOMA (1) Bulk storage and caches

- The storage part is has several aspects
- Accessing and caching input from the grid bulk storage
  - We have seen some aspects of this with use of XCaches in ATLAS, integration with rucio using Virtual placement and the rucio plugin from ESCAPE
- Range of caches sizes and performance but there is no recommendation on the actual specs
- Xcaches are monitored in the US only if installed with SLATE
  - WLCG xrootd monitoring is a first step towards consolidate also xcaches monitoring but it covers only transfers
- Rucio plugin is an interesting proposal and working with EOSC should be better coordinated

DOMA (2) Shared AF storage

- Recurring topic: Local shared storage for people to seamlessly run from different resources
- Discussed during the SWAN/ScienceBox, EOSC and the INFN multi-site AF presentations.
  - First two use EOS at CERN, the latter has a local cephFS storage
  - As soon as the AF is distributed this becomes problematic
  - If access is from remote or with xrootd gateways it's still not as straightforward and can cause problems to the site
  - For different solutions monitoring of traffic and access needs to be harmonised
- Needs discussion and recommendations

DOMA (3) Object Stores

- Another topic that was covered is the possibility to use object stores
- Several grid sites installed ceph with cephFS for POSIX access
  - Only one or two sites tried to integrate object stores
  - Object stores advantage is scalability because they don't need any metadata db
- AF workflows may have different requirements: which?
- Evaluating object stores without posix access is a large body of work
- Experiments requirements on this need to reviewed and initial recommendations if it is worth pursuing written
- Not top priority but there were people interested in doing this
  - It would be good to collect the outcome of the work

Shared resources and kubernetes

- Recommendation is to colocate AFs with existing sites to maximise resource sharing and k8s is the proposed method of deployment for AF
  - As well as an alternative backend to run users jobs
- It has been embraced by a number of large sites
  - Ricardo gave a comprehensive list of forums people can attend
- In WLCG it is recurringly talked about
  - Last time at June pre-GDB (agreed another pre-GDB in a year)
  - Likely to be a topic at the WLCG workshop
- There aren't recommendations on deployment or for k8s applications developers
  - Short document for app devs to get things to work both on okd and vanilla k8s
  - CERN has a really robust documentation but it looks CERN specific
- May follow up with presentations from beginners sites and from users
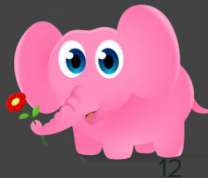
Environment sharing

- Users want to share with colleagues their setup,code, configuration, small amount of input data....
  - Shared storage is the traditional way of doing this but not the only way
- Easy to setup and should help also preservation
- Need recommendations on technologies that help sharing
  - Conda
  - Containers
- How to setup these technologies
  - Particularly containers need discussions on distribution and security
- CVMFS presentation on the 22nd may kickstart further conversation on the topic
- Should also organise a presentation on conda from LHCb

Tracking analysis performance

- Need an agreed upon list of metrics
  - Workflow ID,
  - CPU, RAM, swap,
  - I/O (local storage and network),
  - Software stack,
  - Job failure rate,
  - Time To Completion (TTC),
  - Data source local or cached from a Data Lake,
  - Formats used on input (PHYS, PHYSLITE, AOD, MiniAOD, NanoAOD, NTuple,etc..),
  - Formats written (columns), ratios
  - ......
- Need also the monitoring metrics with no monitoring not helpful
  - Centralised monitoring may take a long time to develop
    - Years, not months!!
  - Other types of site monitoring in the same situation (networking, tape, xrootd, benchmarking...)
    - May draw from that experience
    - Some of the metrics may as well be the same

How to measure user happiness

- Collect user requirements
  Examples:
  - Type of Workloads: classical batch, graph-batch (e.g. workflows/REANA), interactive
  - Scale: expected turnaround time for e.g. full Run-3 MC/Data with systematics
  - How customized of an environment do users need
    (spectrum from: come with own docker images to: use well-curated set of environments
- Define metrics of success for AF
  Examples:
  - Stability / Throughput of Non-interactive/Batch Workloads
  - Time-to-interactive (how quickly can I get a notebook up & running)
  - Time-to-Plot (can I run reliably run distributed analysis in a multi-user setup?)
  - Simulated Stress Workload for N simultaneous users doing a range of tasks

Extra slides