Analysis Facilities

Summary and next steps

Jamie Gooding *on behalf of the Analysis Facilities Plenary organisers*

HSF/WLCG Workshop 2024
Closing plenary, 17th May 2024

HSF HEP Software Foundation

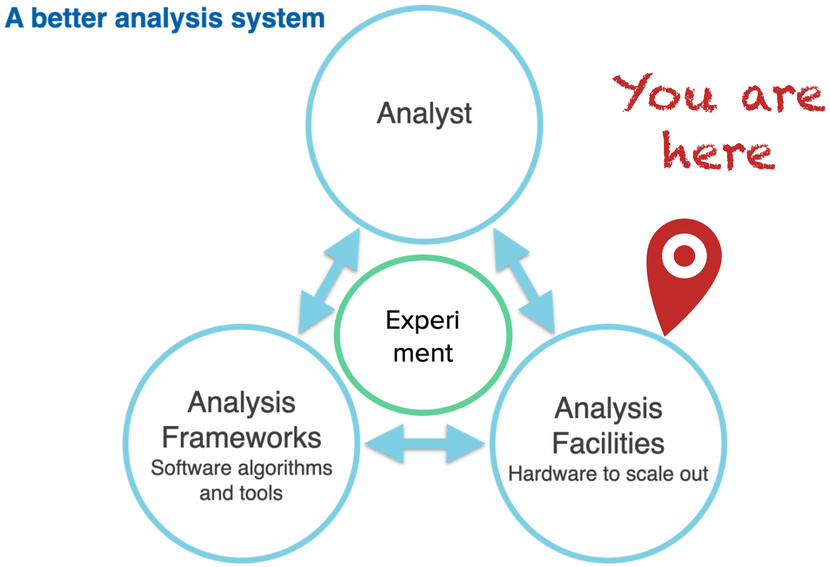
technische universität dortmund

SMARTHEP REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

MARIE CURIE ACTIONS

Introduction

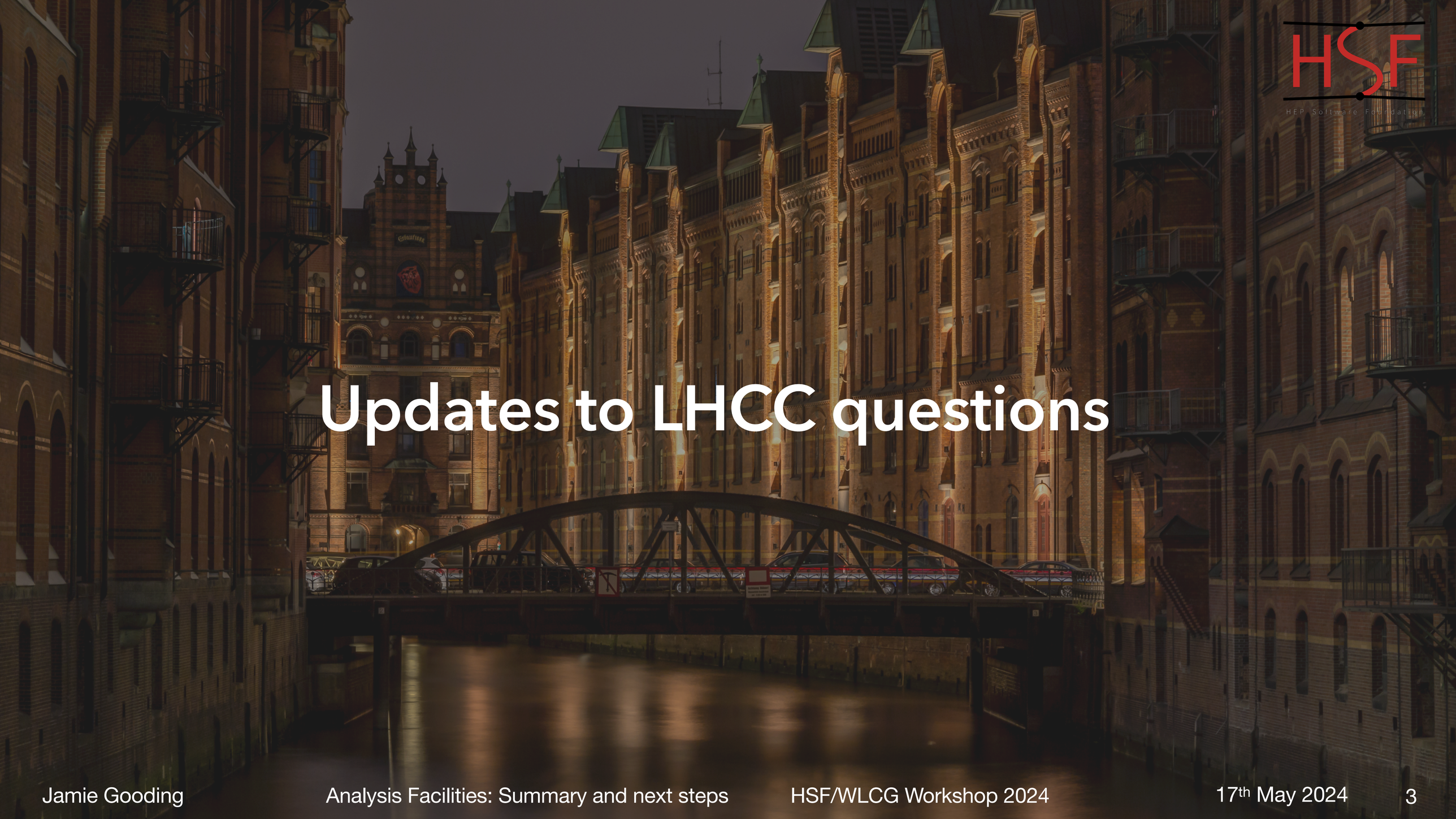
- Why discuss AFs at WLCG/HSF 2024?
  - Broad expertise in running existing AFs
  - Analysis in the HL-LHC era requires computing scaling
    - AFs provide a user-friendly route to deliver this
    - Need to define how this should look before moving forward
- LHCC charge: requested list of questions for experiments
  - Lively discussion yesterday, comments now implemented
- AF White Paper recently released (arXiv:2404.02100)
  - Submitted to Computing and Software for Big Science



Questions for the LHCC on Analysis Facilities, *Alessandra Forti*

Analysis Facilities White Paper

D. Ciangottini[1,b], A. Forti[2,b], L. Heinrich[3,b], N. Skidmore[4,b], C. Alpigiani[5], M. Aly,[2], D. Benjamin[6], B. Bockelman[7], L. Bryant[8], J. Catmore[9], M. D'Alfonso[35], A. Delgado Peris[17], C. Doglioni[2], G. Duckeck[10], P. Elmer[11], J. Eschle[12], M. Feickert[13], J. Frost[14], R. Gardner[8], V. Garonne[6], M. Giffels[36], J. Gooding[15], E. Gramstad[9], L. Gray[16], B. Hegner[31], A. Held[13], J. Hernández[17], B. Holzman[16], F. Hu[8], B. K. Jashal[18,19], D. Kondratyev[20], E. Kourlitis[3], L. Kreczko[21], I. Krommydas[22], T. Kuhr[10], E. Lancon[6], C. Lange[23], D. Lange[11], J. Lange[34], P. Lenzi[1], T. Linden[24], V. Martinez Outschoorn[27], S. McKee[25], J. F. Molina[17], M. Neubauer[26], A. Novak[35], I. Osborne[11], F. Ould-Saada[9], A. P. Pages[28], K. Pedro[16], A. Perez-Calero Yzquierdo[17], S. Piperov[20], J. Pivarski[11], E. Rodrigues[29], N. Sahoo[30], A. Sciaba[31], M. Schulz[31], L. Sexton-Kennedy[16], O. Shadura[32], T. Šimko[31], N. Smith[16], D. Spiga[1], G. Stark[33], G. Stewart[31], I. Vukotic[8], G. Watts[5],



Updates to LHCC questions

LHCC questions 1-3

1. How does the experiment expect the Analysis Model will evolve for Run4 (and Run5) compared to Run3, considering both evolutionary advancements and potentially disruptive revolutions?
    a. Please briefly describe the Run3 Analysis Model and highlight the most important expected changes foreseen/planned for Run4 (and Run5)
    b. Which main analysis workflows run in Run3, and which types of workflows do you foresee being needed in Run4 (and Run5).
    c. How many data reduction steps? How tightly chained are they?
    d. If reduced data formats exist (.e.g. PHYSLITE and nanoAOD): How many analysis will be covered by them?
    e. ATLAS-1, ATLAS-2, ATLAS-4, CMS-1, CMS-3, CMS-4, CMS-5, LHCB-1
2. Please describe the data formats used today and in 2030 for analysis?
    a. Data volumes (per year and total per Run, data and MC)
    b. The number of versions and the number of replicas?
    c. How many will be centrally managed?
    d. Will analysis need to access extra information from other resources?
    e. ATLAS-1, ATLAS-5, CMS-1
3. How much compute power is used today for analysis?
    a. How much is coming from pledged resources and how much from unpledged resources?
    b. Do you have any estimate of how much from local interactive and how much from local batch?
    c. ATLAS-1, ATLAS-5

Analysis Facilities: Questions for the WLCG Experiments

New sub-question

Rewording

LHCC questions 4-5

Analysis Facilities: Questions for the WLCG Experiments

4. What are the main pain points that users experience today in analysis and how does the experiment plan to improve them in the coming years?
    a. What features are currently found to be most beneficial by users? What is missing for a more effective analysis?
    b. Do you foresee any technological or infrastructure evolution/revolution that would help in improving the analysis experience?
    c. ATLAS-6, ATLAS-7, ATLAS-8
5. What is an Analysis Facility from the experiment point of view?
    a. Please briefly describe the present status for what concerns AF in the experiment.
    b. Which functionalities refer, or could refer, to an AF, and which would not make sense to include there (please refer also to the HSF AF)?
    c. Is an AF using resources local or is it distributed or a mix?
    d. For the local resources, would AF cover interactive, batch or both?
    e. Using unpledged vs pledged compute power?
    f. Tailored for certain analysis workflows and/or specific working groups?
    g. Should this support all the users of an experiment, or only a part of it?
    h. ALICE-1, ALICE-7, ALICE-8, ATLAS-1, ATLAS-3, CMS-2, LHCB-1, LHCB-2

Rewording

LHCC questions 6-9

Analysis Facilities: Questions for the WLCG Experiments

6. What are the capabilities and support model that you would need from the AF for you to have them effective?
    a. What technical capabilities?
    b. Which specialised HW?
        i. GPU
        ii. High Memory: do you need special nodes?
        iii. Disk with high IOPS? NVME?
        iv. Caches?
        v. Specific type of storage?
        vi. What bandwidth?
    c. Personpower support
        i. installation, quick answering to users needs, documentation
    d. ALICE-4, ALICE-5, ATLAS-9, LHCB-4, LHCB-5
7. What is the motivation for deploying AF?
    a. Based on the above answers, please describe your strategy towards deployment and integration of AF.
    b. Please detail also whether analyses would be centrally organised on the AF, and if so how.
    c. Please estimate how many AF would be needed for each type defined in Q5.
    d. ALICE-1, ALICE-6, ALICE-3
8. Are you already able to identify a few use cases of analysis which could be useful to benchmark an AF of the type(s) defined in Q5?
9. List R&D that is being done to help address your concerns in questions Q1 and Q4?
    a. CMS-7

LHCC questions: last call!

Last call for comments!

- Document **will be frozen to major comments after today**
- Minor comments (small language/grammar changes) still welcome

Analysis Facilities: Questions for the WLCG Experiments

Analysis Facilities: Questions for the WLCG Experiments

Analysis Facilities Plenary talks

200 Gb/s analysis demonstrator

- Challenge of demonstrating analysis at 25% HL-LHC scale
  - ATLAS data implementation in ServiceX running at UChicago
  - CMS data implementation in Uproot running at Nebraska T2
- Many lessons learned:
  - Understanding memory usage in Python vital
  - Bugs caught (and fixed!), e.g., new Uproot version with XRootD fixes
  - Role of network imbalances
  - Best practices in XCache tuning

Derived Values – Example CMS 'napkin math'

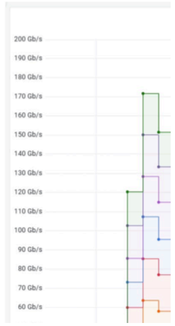
- Start with 200TB read in 30 minutes. => ~900Gbps sustained.
- 25% scale => 200Gbps sustained. Hence, **200Gbps challenge**.

Uproot Results

- Highest data-rate configuration (TaskVine):
  - Data read (compressed): 58.33TB
  - Average data rate: 221Gbps
  - Peak data rate: 240Gbps
  - Files processed: 63,762 (17 failed)
- Highest event-rate configuration (Dask):
  - total event rate : 32,256 kHz
  - Processed 40,276,003,047 events total
  - Per-core event rate : 27.66 kHz

ServiceX Results

- To reduce the overhead of small datasets, we ran on a subset that consisted of the bulk of the data.
- Highlight run:
  - 4 Datasets
  - 146TB total
  - 19,074,862,754 Events
  - 170Gbps
  - Limited to 1,000 pods.
  - Time: 32:28
  - Event Rate: 9,787 kHz



200 Gb/s analysis demonstrator, *Brian Bockelman*

User experience discussion

- Discussion around requested (expected) UX at AFs
- A sample of the discussion:
  - Where to draw line between “installed” tools and what users bring with them?
    - How can users be supported without prescribing tools?
  - Where exactly do our current pain points lie?
  - How do ML tools fit into AF ecosystem?
  - Tools in AFs to ease use of Grid by beginners?

Questions…

- What is impossible?
  - When should a job be done on the GRID rather than an AF?
- Can we (the experiments) collaborate on building these?
- What work gives us the most users?
- Are there technical decision we make that the user will care about?
  - But are unaware of?

I am going to go very fast through this talk!

Tools & Environments

Installed Tools: notebooks, ssh, etc.?

What should the user bring with them?

Method to share container environments in an analysis group
(e.g. docker, dev-containers, binderhub)

Scaling Impact?

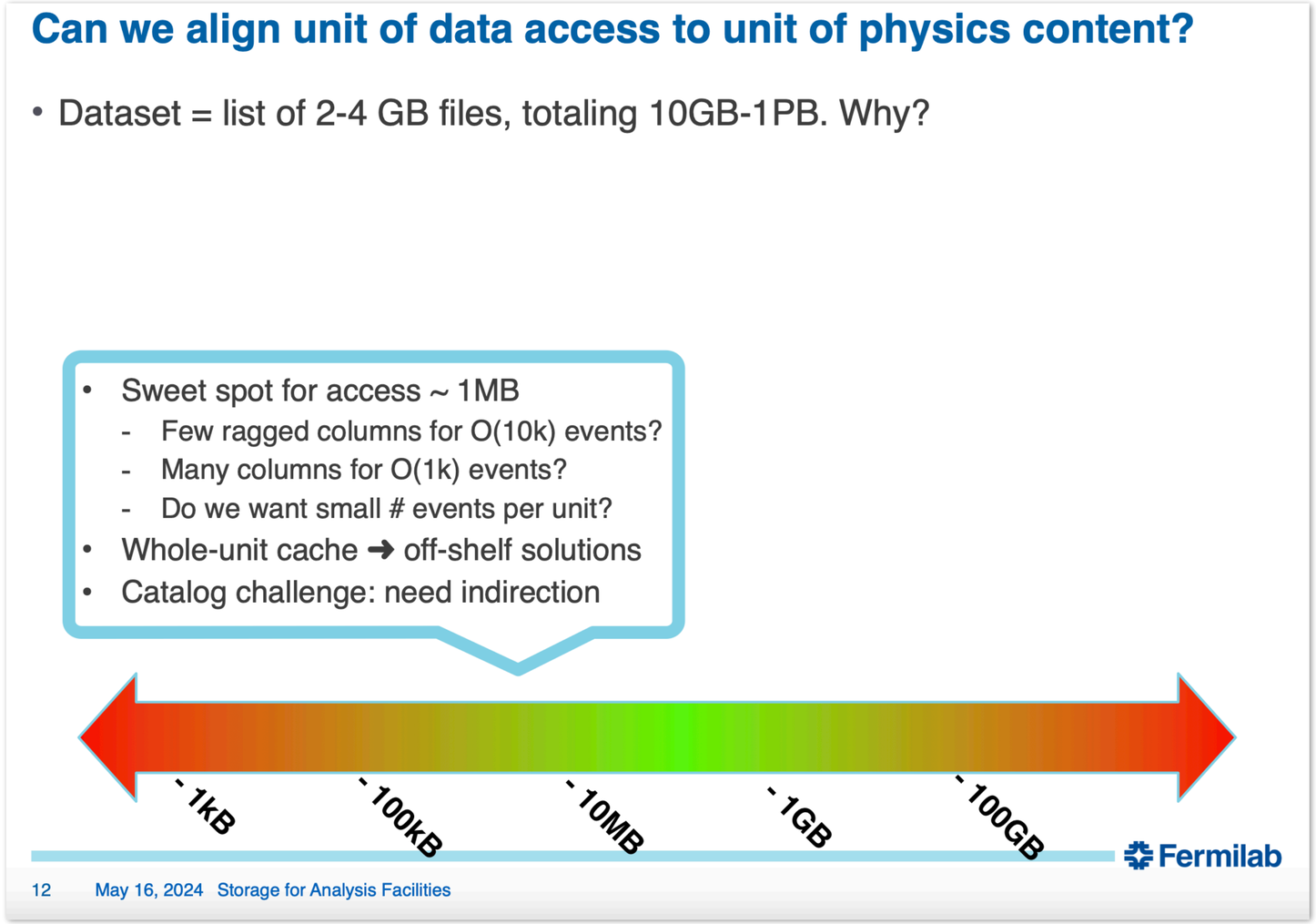
Current Analysis Facilities tend to specialize – is this the right way forward?
- Services provided – ServiceX, REANA, etc.
- Locally installed tools like snakemake…
- Should coffea be there?

Machine Learning → Toolset and workflows – plethora of workflows
- Access to GPU’s (efficient!)
- Workflow support: don’t make a choice?

User Experience at AF, *Gordon Watts*

Data access discussion

- Discussion of need to align data access with physics:
  - ~1 MB objects align well to physics content
- A sample of the discussion:
  - HEP applications generally built for POSIX
    - Community experience with POSIX → how does the path look for user adoption of object store?
  - How can/should token-based auth. be implemented at an AF?



Storage for Analysis Facilities, *Dirk Hufnagel & Nick Smith*

User monitoring discussion

- Discussion of how best to monitor use of AFs
- Already many examples of monitoring best practice
- A sample of the discussion:
  - What is most useful for users/for AFs to see?
  - How to communicate to users key information (e.g., wait/fail reasons)
  - How to tell if users are suffering in silence?

What are the questions, and from whom?

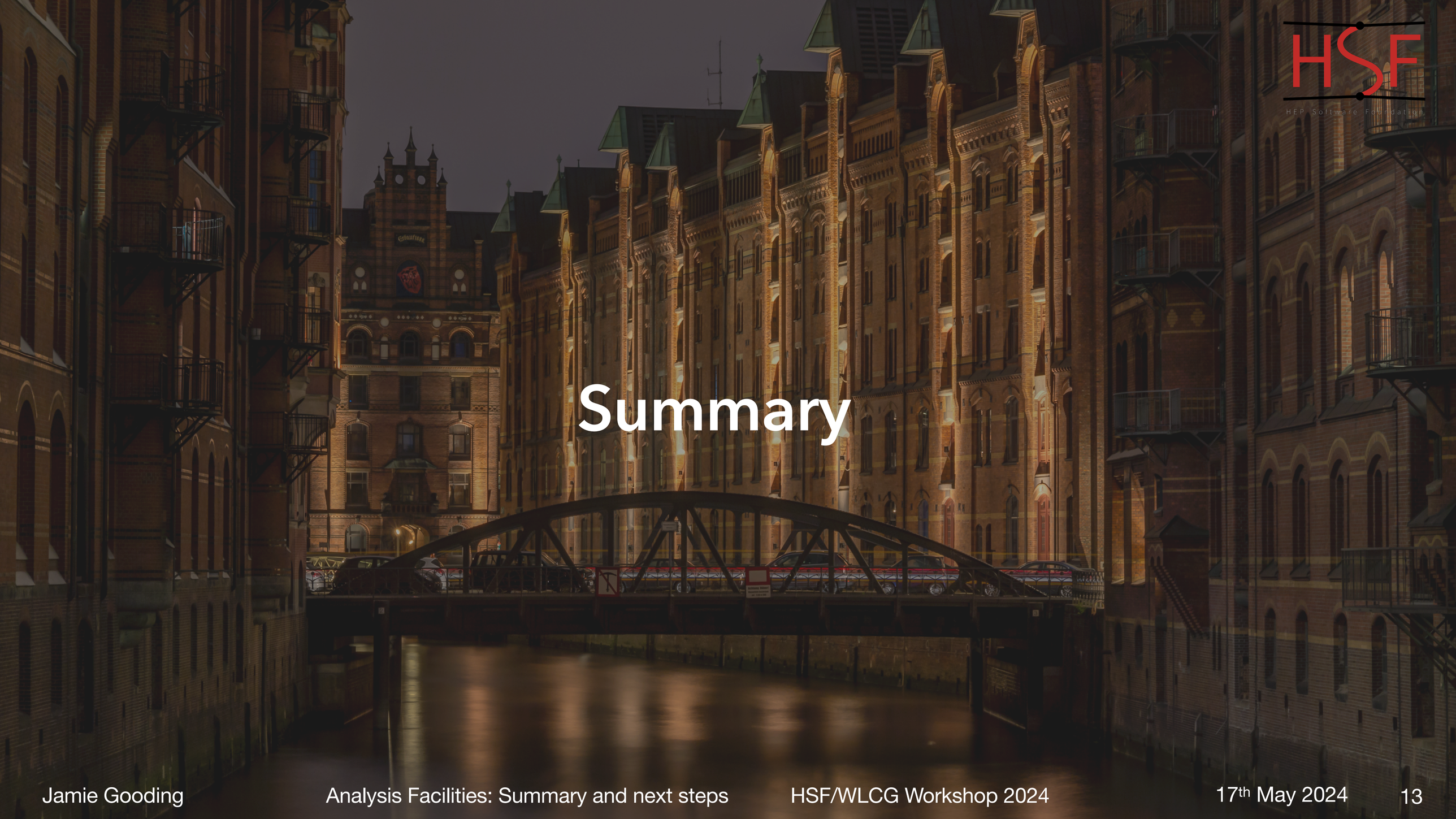
- Users (experience) (1)
  - What resources are available?
  - How long will my jobs wait in queue? Why do they run so slow? Why is my notebook hanging? Why did my last few jobs not finish? Why are my jobs held? Why did my jobs fail? Why are they being held?
  - How do I access my data? Is it local? How do I get X software installed? How do I run with my container?
- Resource providers (trends, performance, facility metrics) (2) (3) (4)
  - What resources (cpu, disk-capacity, disk-fast, network, gpu) are under-provisioned?
  - What are the performance bottlenecks?
  - What are the (unexpressed) requirements?
  - Managing the storage - scratch, precious, freeing up space, group storage
  - Scheduling bursty workflows & precious resources (GPUs, fast storage)
- A fith category: metrics for **framework & platform developers**
  - Which data formats are physicists most often using and by which frameworks?
  - Are performance targets met? (e.g. X TB / Y minutes)
  - Where are the inefficiencies and user pain points?
  - What capabilities are missing?

link

submit your own questions

3

Analysis Facilities Monitoring Discussion, *Rob Gardner et al*



Summary

Summary and next steps

- Questions for experiments to be presented to LHCC by Alessandra in June 2024:
  - Comments and discussion yesterday greatly appreciated
  - Any major comments should be made today
  - Minor comments still welcome
- Plenary talks covered many aspects of AFs:
  - Discussion covered many new points, provided useful insight on current best-practice
  - Outcomes will prove valuable in reaching practical solutions to problems highlighted in AF White Paper
    - *See notes taken live*

Thank you, any questions?