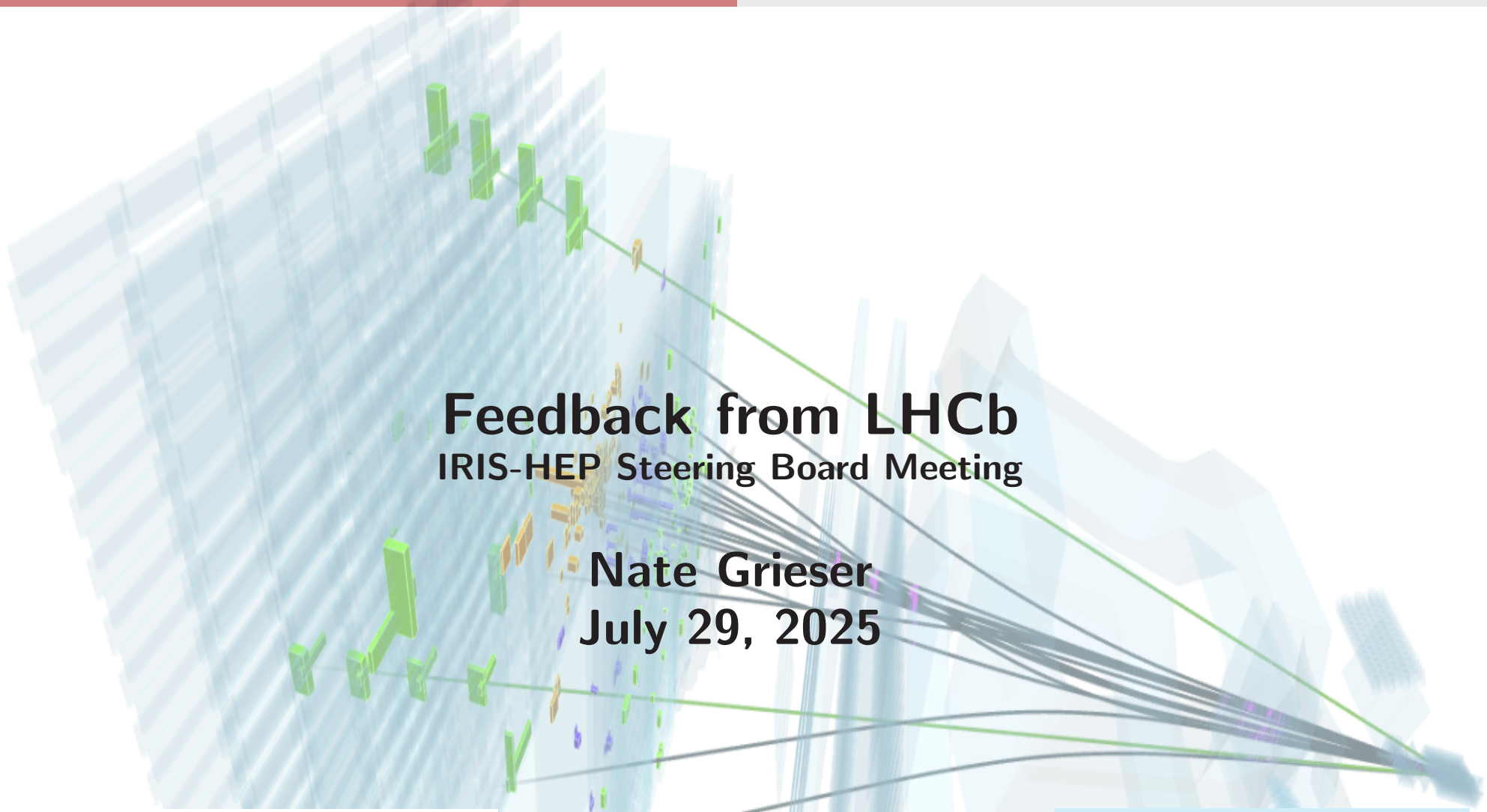


Feedback from LHCb

IRIS-HEP Steering Board Meeting

Nate Grieser

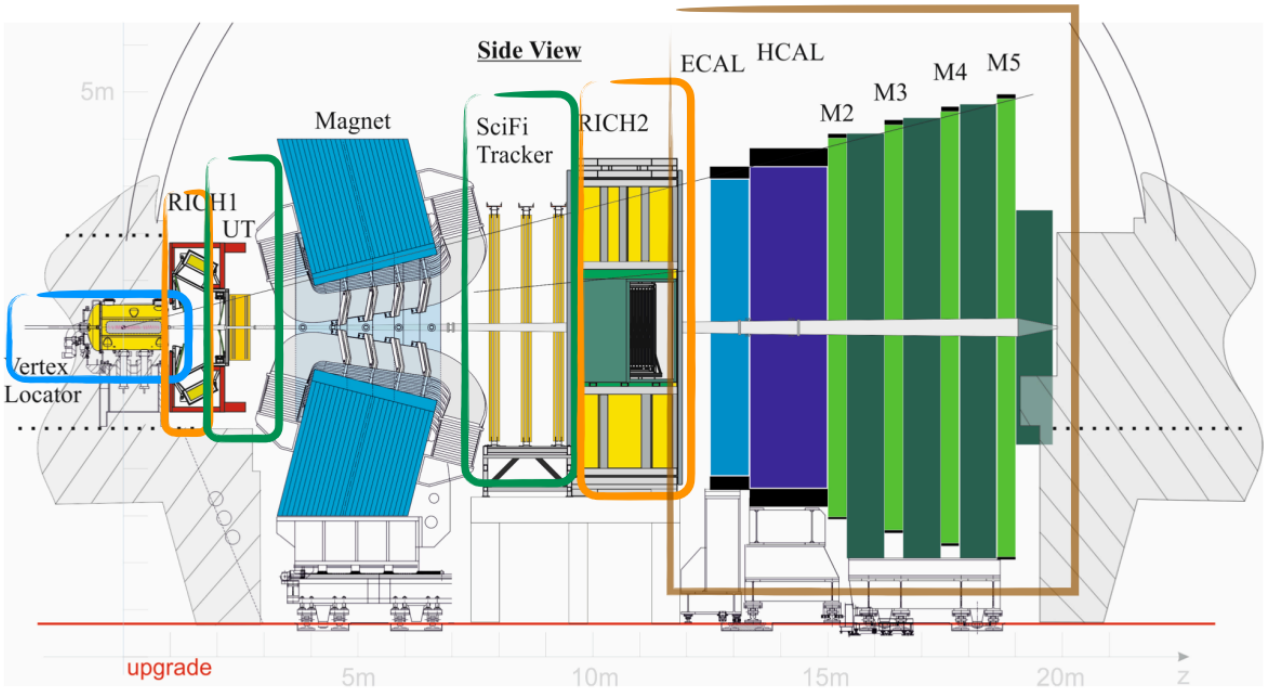
July 29, 2025

University of CINCINNATI

LHCb Upgrade I

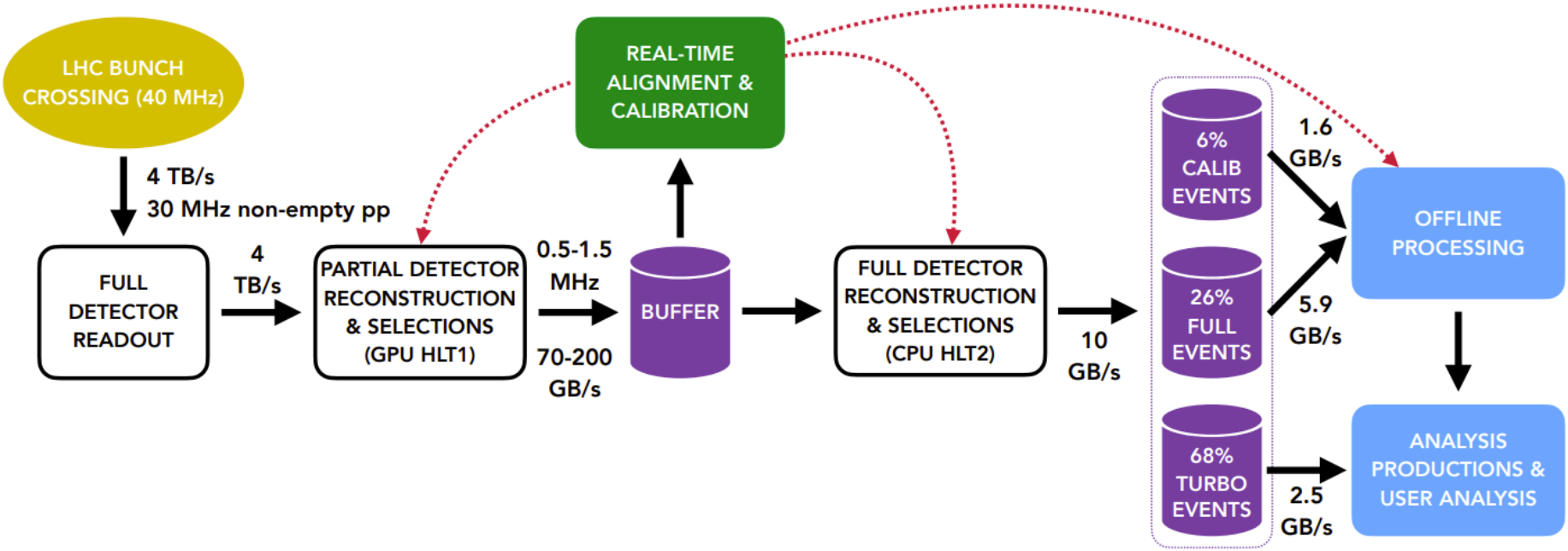
arXiv:2305.10515

Significant upgrades to the LHCb detector were made during the recent LHC LS2



- New Vertexing Detector → Closer to beam; Silicon pixels improve IP resolution
- New PID detector → New photon detectors and improved readout
- New tracking system → Silicon strip UT for improved tracking granularity; Fibre tracking for high particle density momentum resolution
- New Readout Systems
- *BONUS* : New GPU-based trigger

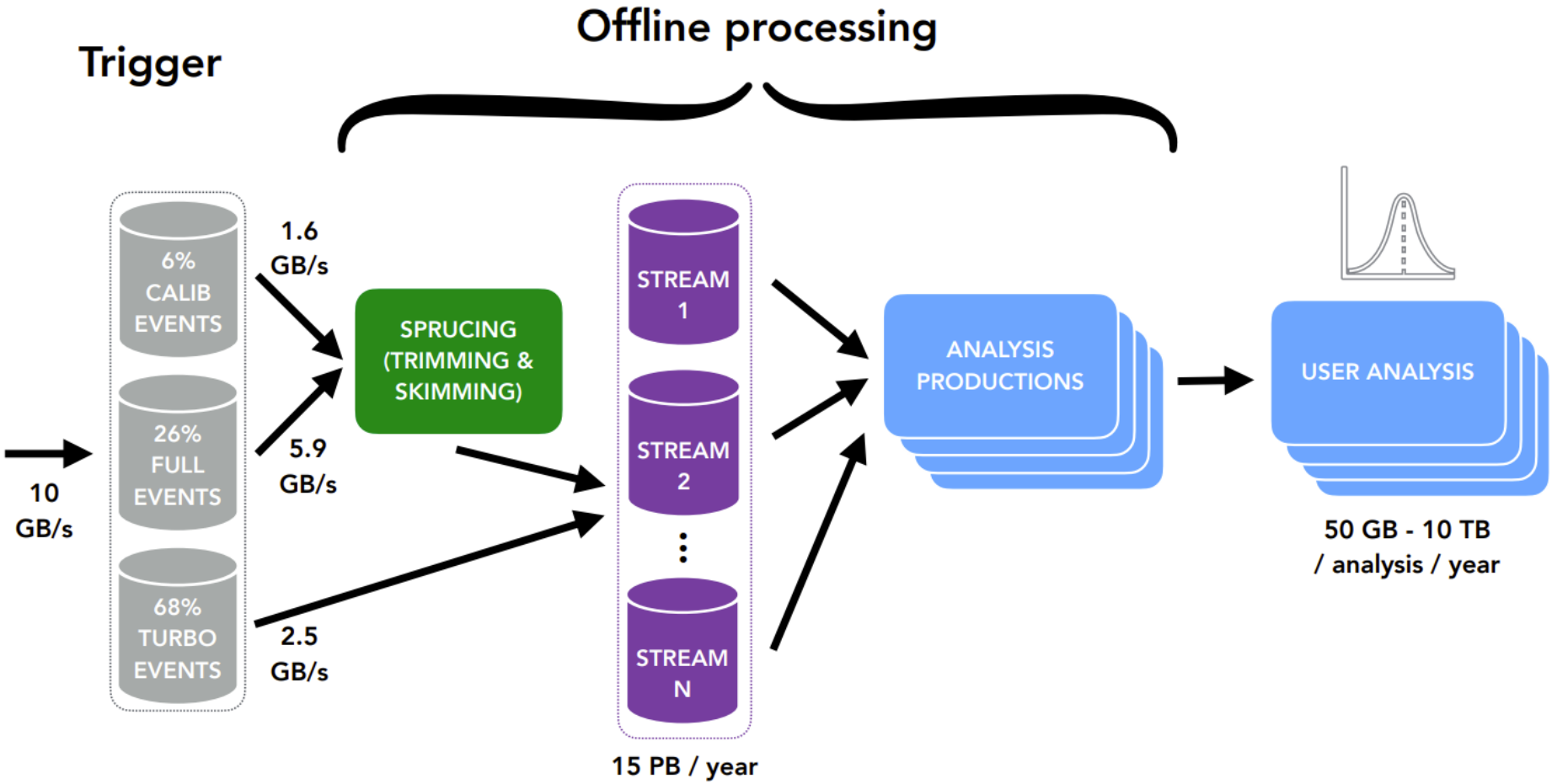
LHCb Run 3 Dataflow



→ LHCb utilizes a purely software-based trigger system

All stages (HLT1, HLT2, Offline) built on top of the same software stack and selections made by analysts

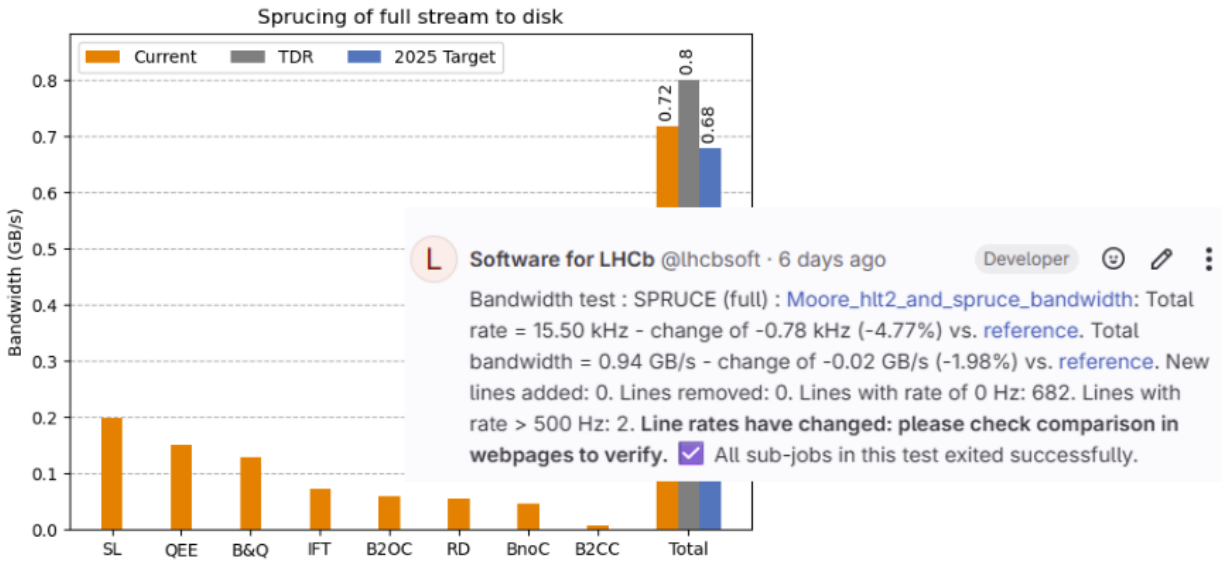
LHCb Run 3 Dataflow – Offline



Huge amounts of data sent offline concurrently in real time

Monitoring Impacts With CI Tests

- Impact of software changes to operational budgets monitored closely
  1. Changing Selections
  2. Changing reconstruction
- WG and line-by-line breakdown provides a transparent understanding of resource use to Physics Coordination



	Stream	Total Retention (%)	Rate (kHz)	Avg Total Event Size (kB)	Total Bandwidth (GB/s)	Avg DstData Size (kB)	DstData Bandwidth (GB/s)
0	sl	9.38	3.4	40.3	0.137	39.8	0.135
1	bandq	4.31	1.56	72.7	0.114	72.3	0.113
2	qee	2.76	1	113	0.113	113	0.113
3	sllow	4.1	1.49	41.6	0.0618	41.2	0.0612
4	b2oc	3.14	1.14	49.5	0.0563	47.1	0.0536
5	bnoc	1.48	0.537	86.9	0.0467	82.3	0.0442
6	rd	2.03	0.738	56.7	0.0419	56.2	0.0415
7	iftlow	5.93	2.15	19.4	0.0417	19.1	0.041
8	qeelow	1.07	0.388	96	0.0372	95.6	0.0371
9	ift	2.69	0.975	30.6	0.0298	28.8	0.0281
10	bandqlow	3	1.09	13.5	0.0146	13	0.0141
11	rdlow	2.52	0.912	14.9	0.0136	14.4	0.0131
12	b2cc	0.0689	0.025	213	0.00532	212	0.00531
13	b2oclow	0.138	0.05	63.5	0.00318	60.6	0.00303
14	b2cclow	0.0345	0.0125	214	0.00268	202	0.00253
15	bnoclow	0	0	0	0	0	0

LHCb Run 3 Dataflow – Analysis Productions

Chris Burr

ew_projects_run3 / mcblock8_magup_w_full_ftuple QEE

Productions / QEE / ew_projects_run3 / mcblock8_magup_w_full_ftuple

State	READY
Size	820 MB
Created	April 11, 2025 at 3:26 PM GMT+2 (2 months ago)
Version	v1r2693
Merge Request	https://gitlab.cern.ch/lhcb-datapkg/AnalysisProductions/-/merge_requests/2693
Task	https://gitlab.cern.ch/lhcb-datapkg/AnalysisProductions/-/issues/1615
Sample ID	44951
Publications	None
Archival	Not flagged
Housekeeping due	July 10, 2025 at 1:27 PM GMT+2 (in 22 days)

DIRAC Production Request 159414

comprises the following transformations:

Transformation 282876

comprises 1 step - output is not kept

Step ID	221565
Application	DaVinci/v65r0

Options

```
{
  "entrypoint": "ew_projects_run3.ftuple:main",
  "extra_args": [
    "--",
    "Full",
    "MagUp",
    "MC"
  ],
  "extra_options": {
    "dddb_tag": "dddb-20240427",
    "data_type": "Upgrade",
    "conddb_tag": "sim10-2024.W40.42-v00.00-mu100",
    "input_type": "ROOT",
    "simulation": "True",
```

Collaborative approach to data processing

→ Centralized productions facilitates analysis preservation

LHCb Run 3 Dataflow – Analysis Productions

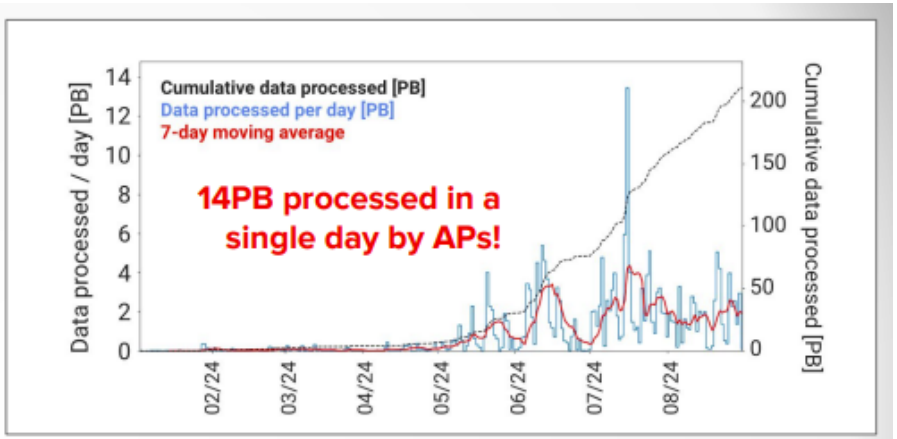
Nicole Skidmore

Analysis productions are an analysis facility in themselves. Ethos is:

- **Declarative** tupling via python options and yamls
  ⇒ Improved analyst experience
- **Centralised** productions run on WLCG
  ⇒ Saves countless analyst-hours
- Exploit **DIRAC transformation system**
  ⇒ Full data provenance - analysis preservation built in
- Full job **testing** on GitLab CI
  ⇒ No buggy jobs on grid

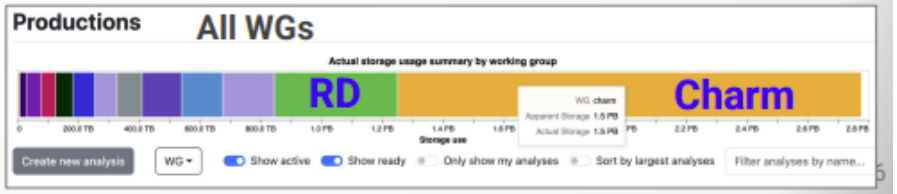
Highly efficient use of the WLCG!

Monitor storage requirements of user ntuples (made with APs) by WG/analysis

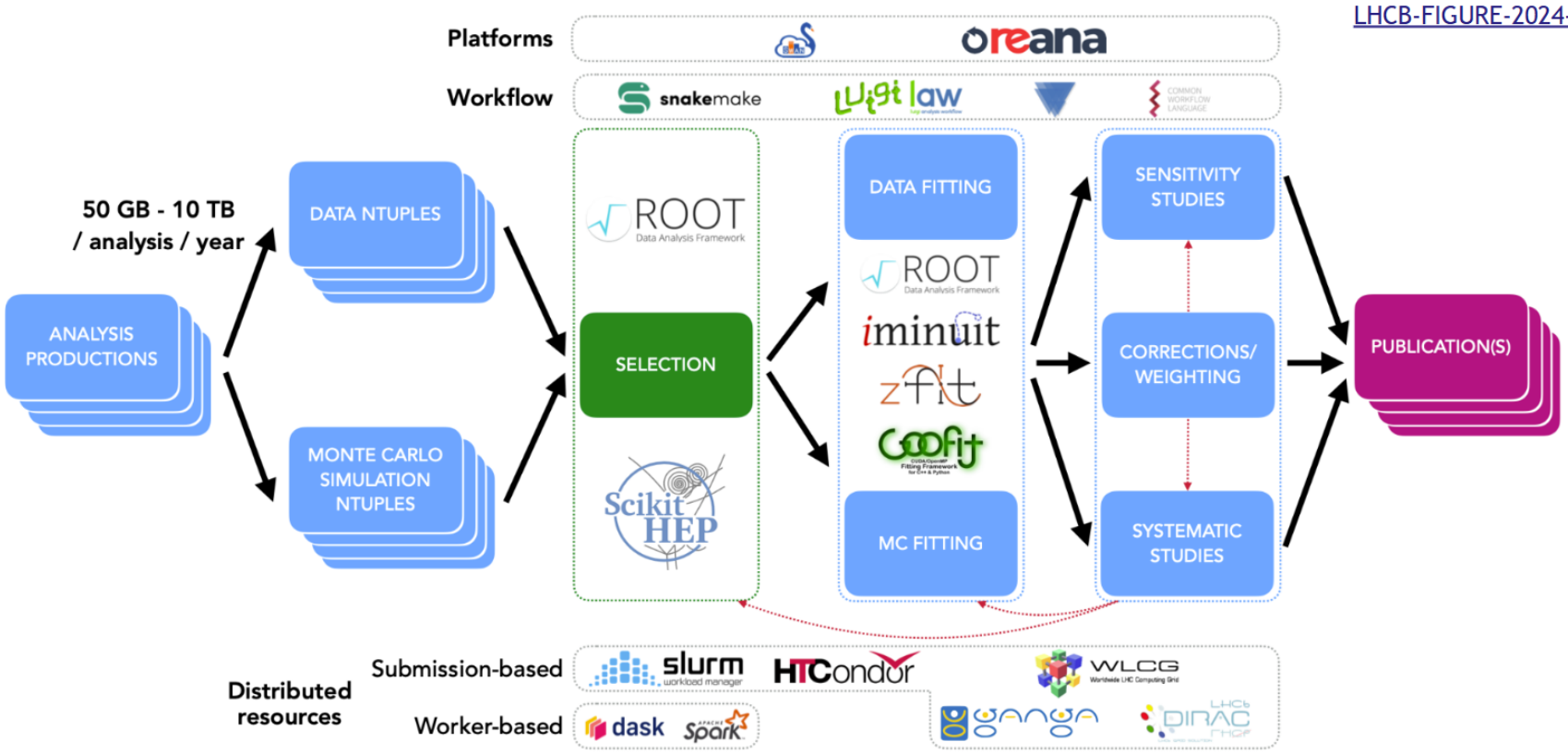


In 2024 "live" APs picked up data as it was Spruced

- **Tuples available within days of data being taken**



LHCb Offline Analysis Structure

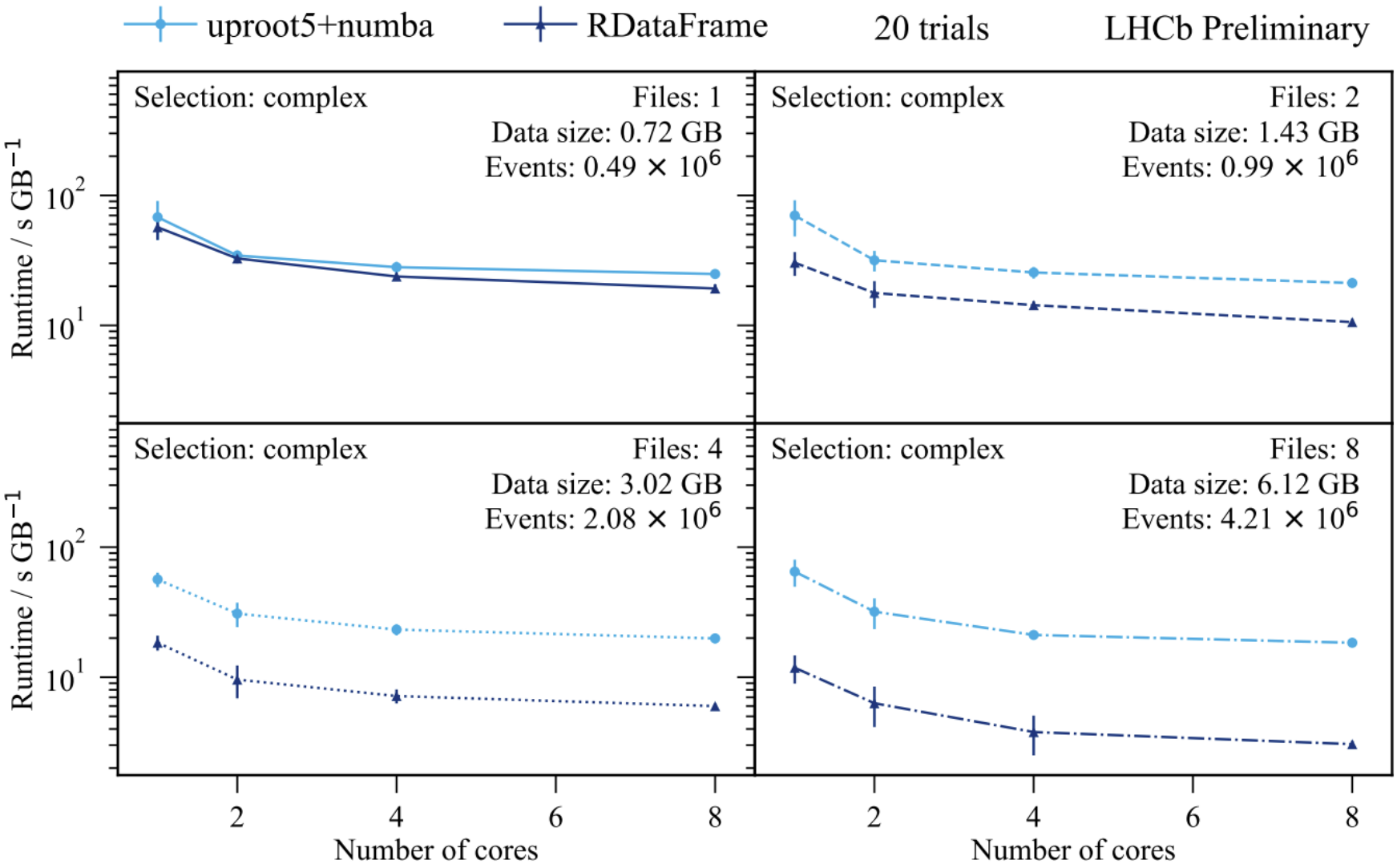


Heavy work on-going for analysis facilities to support this structure

Benchmarking ROOT Tools – RDataframe

Jamie Gooding

Upgrade I gave a huge increase in data sizes → Need efficiency!

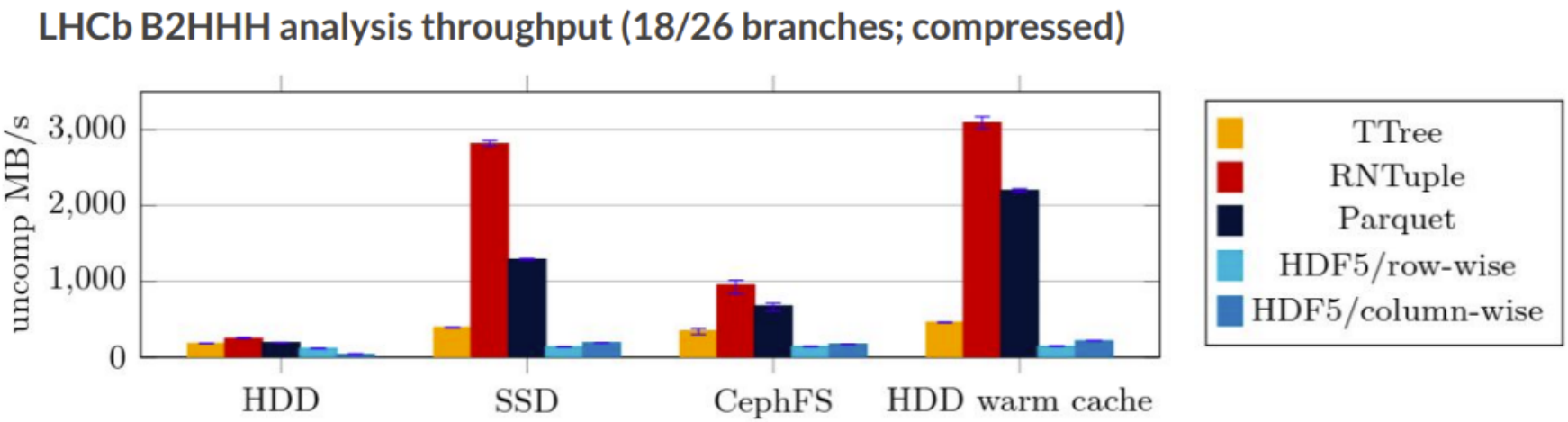


Benchmarking ROOT Tools – RNtuple

Lopez-Gomez, Blomer

Upgrade I gave a huge increase in data sizes → Need efficiency!

→ Moving to RNtuple for output of offline processing stage



→ LHCb added a ROOT liaision (Silia Taider) to provide feedback loop between experiment and software project

LHCb – A Pythonic Experiment

- Python packages often have a simpler implementation and lower learning curve → great for getting students involved
  - About 50% of LHCb analysers using Python-based analysis code!

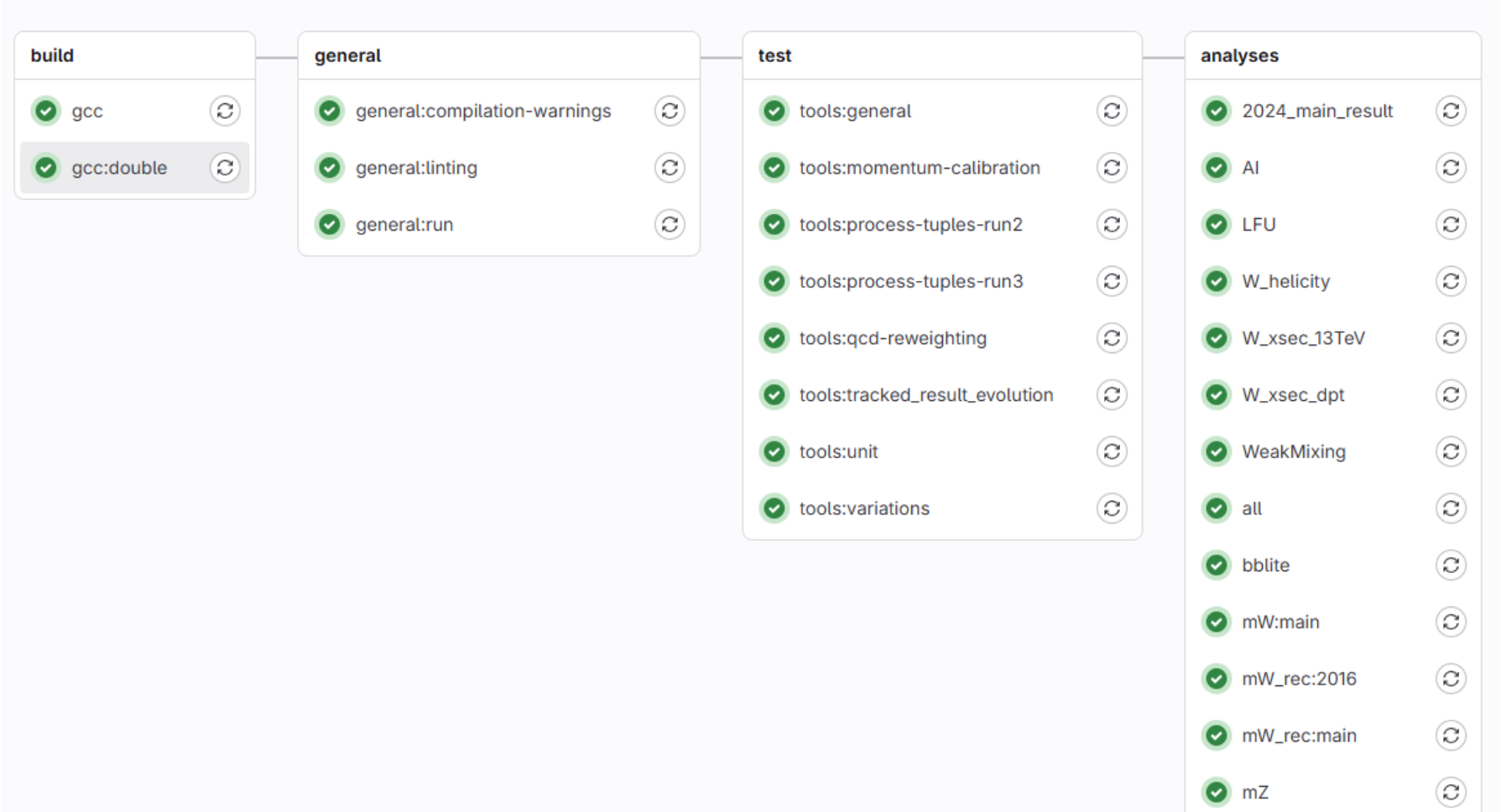
→ Flat data formats for analysts allow for easy access to the full slate of Pythonic tools in the community

- Many collaborators have involvement in development of the SciKit-HEP tools
- Numerous LHCb-specific developments are targeted (*DecayLanguage*, Amplitude Analysis tools, etc.)
- LHCb has many active analysts who are master or bachelors students
  - Python is easier to learn and more applicable to their future

Analysis Preservation with Snakemake and CI

Analysis preservation remains as one of the core principles to analysis development within the collaboration

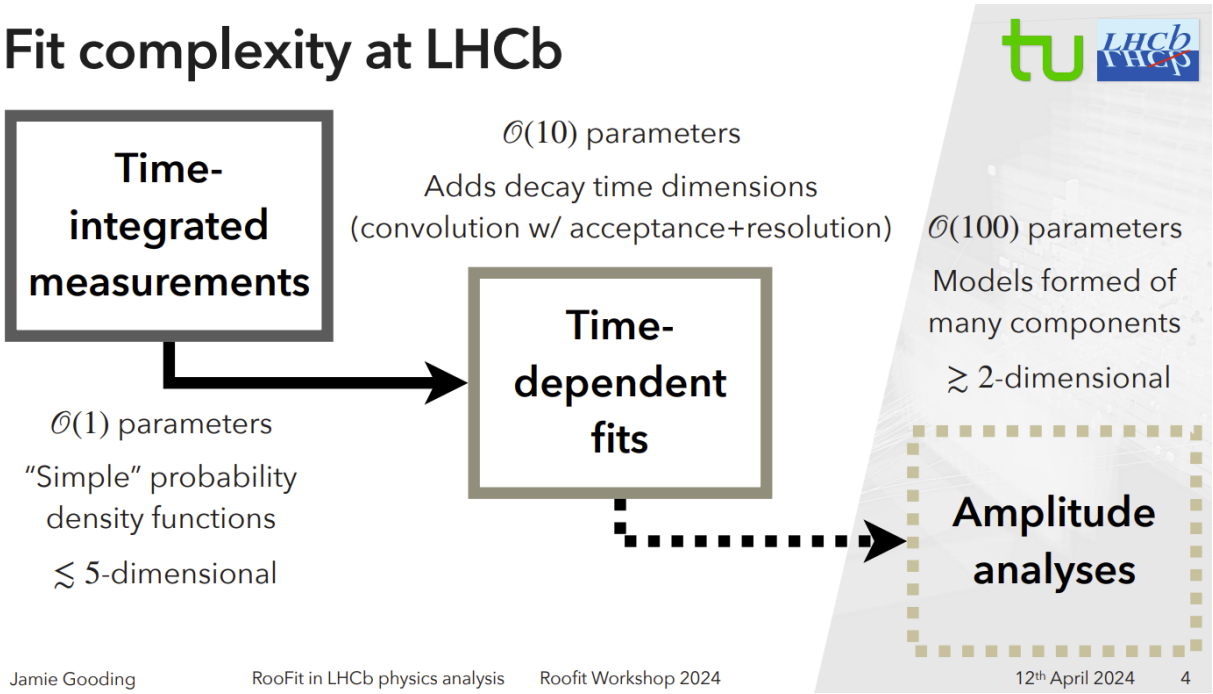
→ Workflow managers (**Snakemake, etc.**) and gitLab CI lead the charge



Fitting Needs of LHCb

Jamie Gooding

Huge variety of fitting needs within the physics program



$\rightarrow$ Often leads to in-house developments

GPU support incredibly import for complex fitting

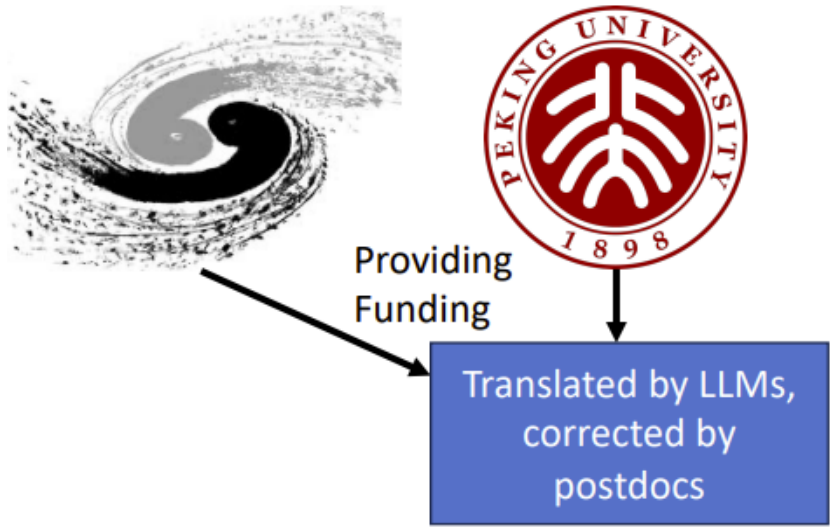
Taking Responsibility for Documentation

Offline Computing project added a work package specifically targeting documentation and training

LHCb

DPA WP7 – Documentation: Visibility and accessibility

- Documentation facilitates the expansion of the physics program efficiently
- Critical that it is **accessible** to the whole scientific community of the experiment



Future Challenges Are the Same

Nicole Skidmore

User requirements for HL-LHC

From Analysis Facilities Community White Paper

• Interactive/fast analysis cycles on "large" datasets for analysis prototyping • Convert interactive to batch-schedulable workloads	Interactive scaling requirements
• Machine learning models • ML Inference within an analysis pipeline	ML support
• Efficient access to collaboration data • Share analysis data artifacts with the collaboration	Storage access and sharing
• Seamlessly share analysis code/workflows with collaborators to run on a chosen infrastructure(s)	Federated Identity
• Instantiate desired software stack • Share environments with colleagues	Analysis portability

20

ATLAS + CMS are solving the hard puzzles now

We need to learn from you for our Upgrade II!

U.S. Activity Across ExHEP

🌎 LHC & DUNE Experiment Collaborating Institutes

Experiment	Total Institutes (Member + Associate)	U.S.-Based Institutes
ATLAS	243	~40
CMS	~240	~45
ALICE	174	11
LHCb	98	~7
DUNE	~175–200 institutions across 32–35 countries	~60% of total → ≈ 105 U.S. institutes
🔢 Total	≈ 930–955	≈ 208 U.S. institutes

→ Small relative contribution of U.S. Institutes

How to increase involvement?

US-LHCb Activities Related to IRIS-HEP Doctrine

Break-down of U.S. institutes and their "tech" activities

1. Cincinnati – Data Processing; Innovative fitting using GPUs; Sim Optimization
2. MIT – Trigger (Alignment + Reco.); Heavy AI innovation activities
3. Syracuse – Primarily Hardware
4. Maryland – Primarily Hardware
5. Los Alamos – Tracking; Primarily hardware
6. Michigan – Analysis Preservation and Open Data

→ Significant growth in last two years of LHCb → EIC targeted groups: Indiana (Trigger), Ohio State (Hardware), Kent State (Tracking)

→ **Overlapping topics already; need to make connections**

BACKUP

Line By Line Comparisons

Changes in rates per line:

	Line	Rate (kHz)			Total Bandwidth (GB/s)		
		Ref	New	Change (%)	Ref	New	Change (%)
0	Hlt2Charm_DpDspToKmEpEpDecision_0	0.0774	0.012	-84.50	0.00451	5.81e-05	-98.71
2	Hlt2BnoC_BdsToEtapKS_etaToGG_LLDecision_0	0	0.012	inf	0	0.00171	inf
3	Hlt2CutBasedInclDimuonFakeMuPlusTrackSSDecision_0	0.0129	0.012	-6.98	0.00113	0.0015	32.58
4	Hlt2B2OC_BuToD0K_D0ToKsLLHHDecision_0	0	0.024	inf	0	0.000236	inf
5	Hlt2Charm_LcpToPpMupMup_RareCharmControlDecision_0	0.0258	0	-100.00	0.0026	0	-100.00
8	Hlt2QEE_JpsiToMuMu_PromptDecision_0	0.271	0.216	-20.27	0.000916	0.000682	-25.55
9	Hlt2RD_BToPpPmEEDecision_0	0.0258	0.012	-53.49	0.00347	0.0018	-48.22
12	Hlt2B2OC_OmbmToOmc0KKPi_Omc0ToPKKPiDecision_0	0.181	0.204	12.96	0.0117	0.077	557.92
13	Hlt2Charm_DpToKpPimPipGDecision_0	0.193	0.108	-44.19	0.00318	0.00114	-64.23
14	Hlt2B2CC_BsToJpsif0wsDecision_0	0	0.012	inf	0	0.003	inf
15	Hlt2Charm_XiccppToLcpMumNu_WSDecision_0	0	0.012	inf	0	0.000424	inf
16	Hlt2RD_XiMinusToLambdaPi_Detached_DD_DDecision_0	0.0129	0.012	-6.98	0.000184	8.87e-05	-51.69
18	Hlt2Charm_LcpXicpToPpKsKs_LLDDDecision_0	0.0258	0	-100.00	0.000158	0	-100.00
19	Hlt2Charm_DstpToD0Pip_D0ToKmPimPipEmNu_SSDecision_0	0.0129	0.012	-6.98	0.000546	0.000115	-78.85
20	Hlt2QEE_SingleHighPtElectronFullDecision_0	0.838	0.732	-12.70	0.0811	0.0865	6.63
21	Hlt2LumiCountersDecision_1	30.8	137	343.63	0.00821	0.0351	327.69
22	Hlt2Charm_LcpToL0PimPipPip_DDDecision_0	0.748	0.972	29.91	0.0151	0.0305	102.61
24	Hlt2IFTTurbo_SMOG2Phi2kkDecision_0	0	0.096	inf	0	0.00184	inf
25	Hlt2RD_BuToKpLLP_LLPToEEDecision_0	0.142	0.06	-57.72	0.00319	0.000374	-88.26
28	Hlt2RD_LbToLMuE_DDDecision_0	0.0129	0.084	551.16	0.00295	0.00208	-29.33
29	Hlt2BandQ_BdForSpectroscopyFullDecision_0	0.0774	0.06	-22.48	0.0225	0.0167	-25.95
30	Hlt2RD_BuToL0PmMuMuSS_DDDecision_0	0	0.024	inf	0	0.00104	inf
32	Hlt2Charm_DpDspToKpMupEmDecision_0	0.0387	0.06	55.04	0.000299	0.00326	990.37
36	Hlt2RD_BdToKstTauTau_KstToKPi_TauToMu_TauToPi_FakePionFromTauDecision_0	0.0258	0.048	86.05	0.000884	0.00561	534.44
38	Hlt2IFTFull_10GeVIsolatedGammaDecision_0	0.0129	0	-100.00	0.00145	0	-100.00
40	Hlt2TrackEff_DiMuon_Downstream_mup_TagDecision_0	0.335	0.192	-42.75	0.0532	0.0299	-43.76
43	Hlt2QEE_DiElectronDisplaced_PersistPhotonsSSDecision_0	0	0.036	inf	0	0.00012	inf