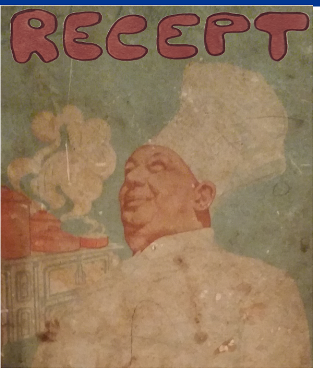
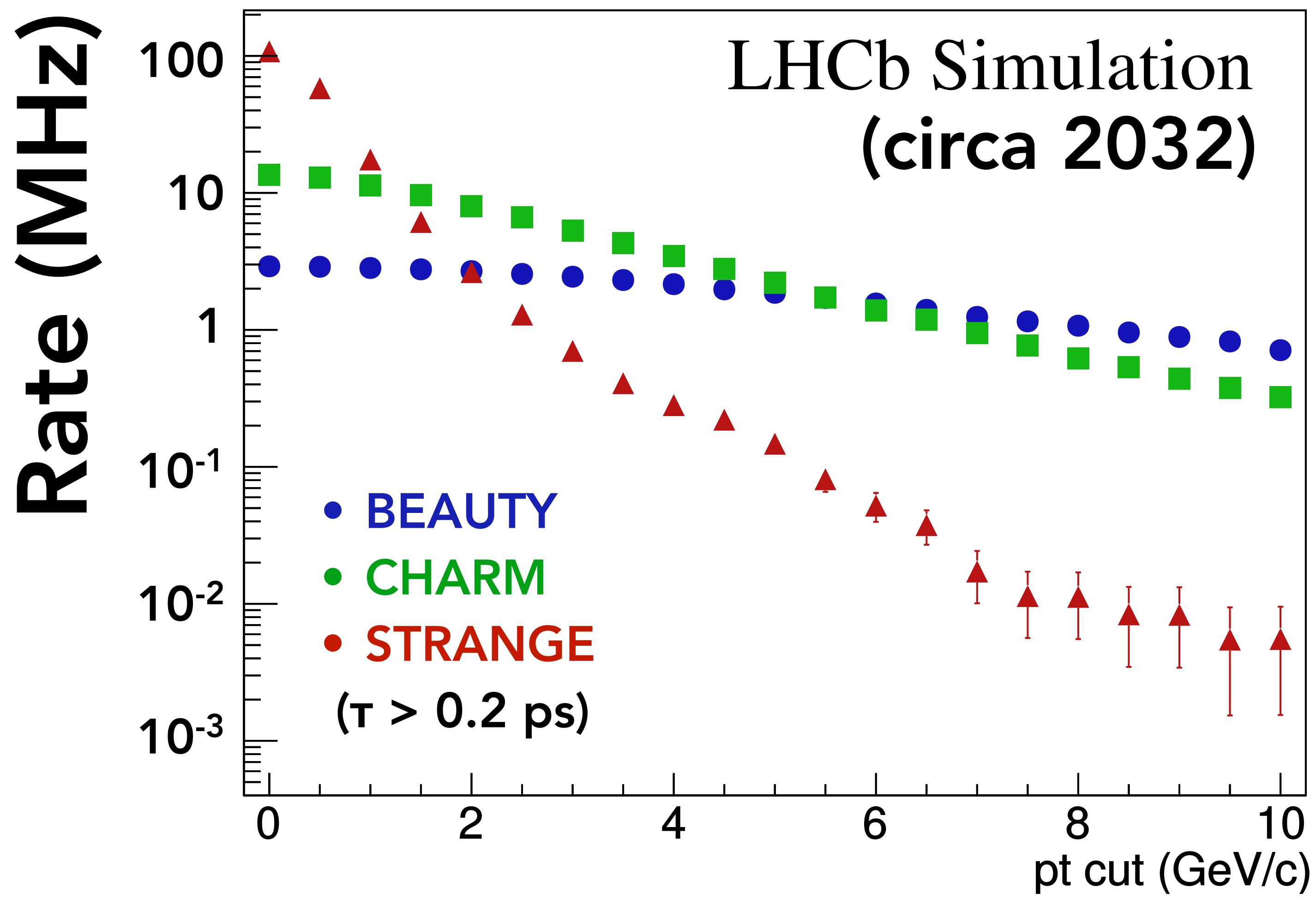
Data Challenges
(on the road towards ultimate precision)



Partially reconstructed signals



CNRS

LPNHE PARIS

European Research Council
Established by the European Commission

Vladimir V. Gligorov, CNRS/LPNHE

With many thanks to the LHCb, ATLAS, CMS, and Belle II collaborations for input and comments!

TUPIFP Workshop, Warwick, 16.04.2018

Data @ ultimate precision experiments

	LHCb	ATLAS/CMS	Belle II
Today			
2021			
2027			
2032			

Data rate read out from detectors by the DAQ system for further online or offline processing

Data written to permanent storage for long-term analysis

See first backup slide for disclaimers, simulation is discussed later on in the talk

Data @ ultimate precision experiments

	LHCb		ATLAS/CMS		Belle II
Today		5.6 Gb/s	150 Gb/s	~10 Gb/s	
2021					
2027					
2032					

Data rate read out from detectors by the DAQ system for further online or offline processing

Data written to permanent storage for long-term analysis

See first backup slide for disclaimers, simulation is discussed later on in the talk

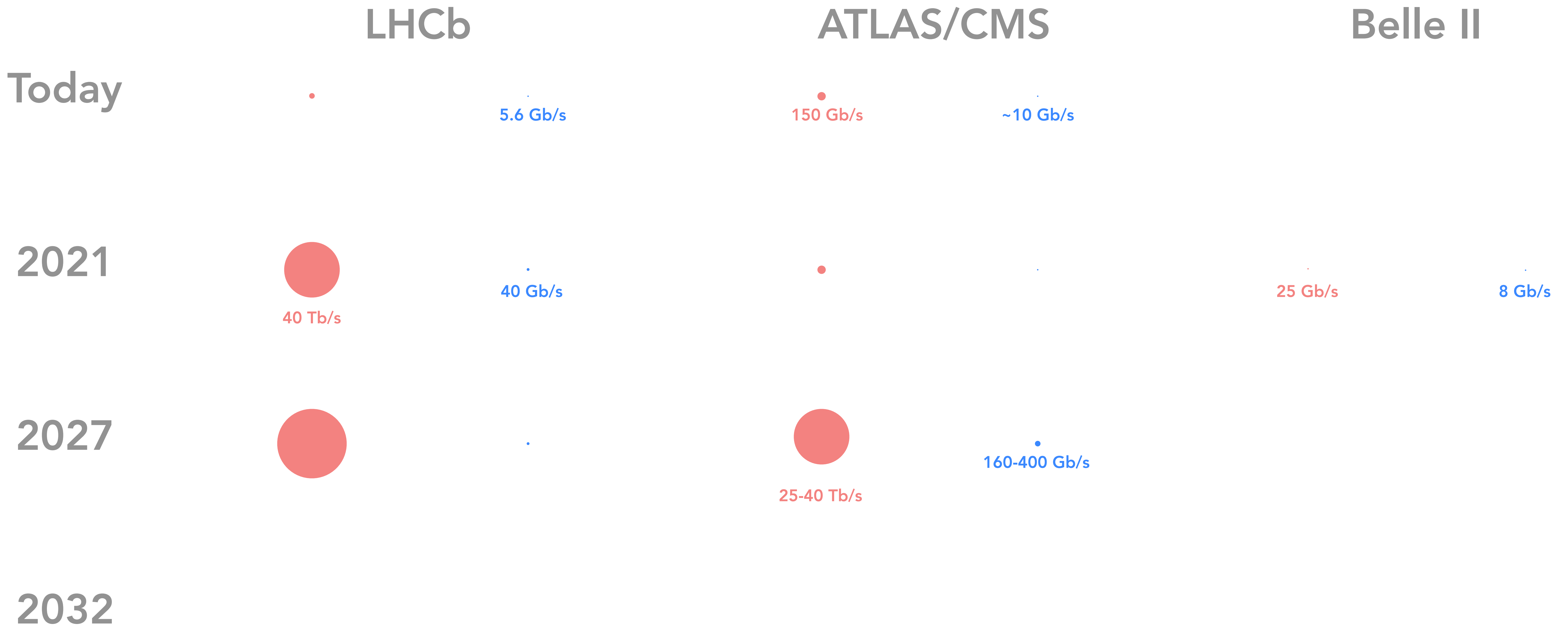
Data @ ultimate precision experiments

Data rate read out from detectors by the DAQ system for further online or offline processing

Data written to permanent storage for long-term analysis

See first backup slide for disclaimers, simulation is discussed later on in the talk

Data @ ultimate precision experiments

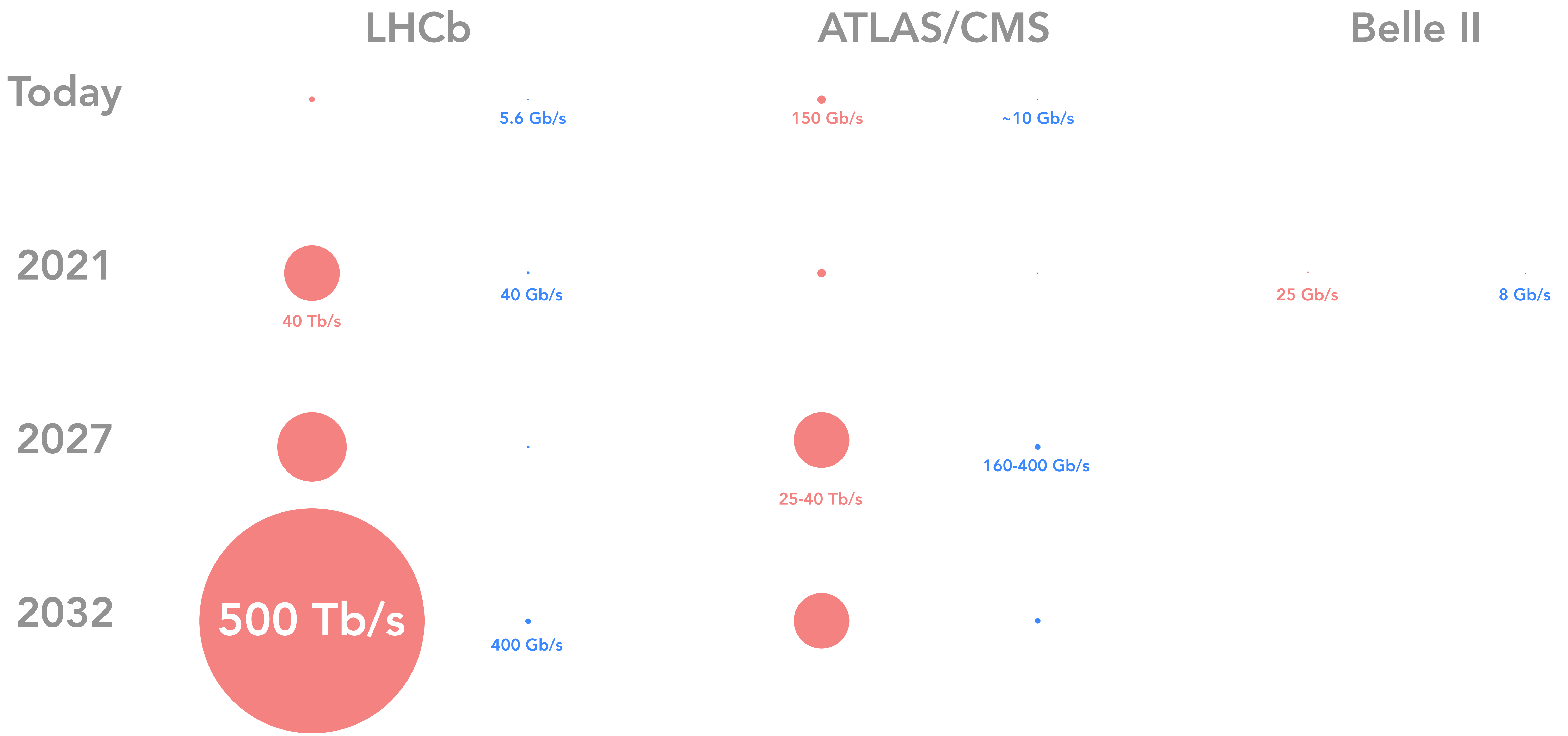


Data rate read out from detectors by the DAQ system for further online or offline processing

Data written to permanent storage for long-term analysis

See first backup slide for disclaimers, simulation is discussed later on in the talk

Data @ ultimate precision experiments



Data rate read out from detectors by the DAQ system for further online or offline processing

Data written to permanent storage for long-term analysis

See first backup slide for disclaimers, simulation is discussed later on in the talk

Context : other large data sources

LHCb 2032

2000

Eb/year

Context : other large data sources

Context : other large data sources

LHCb 2032

2000
Eb/year

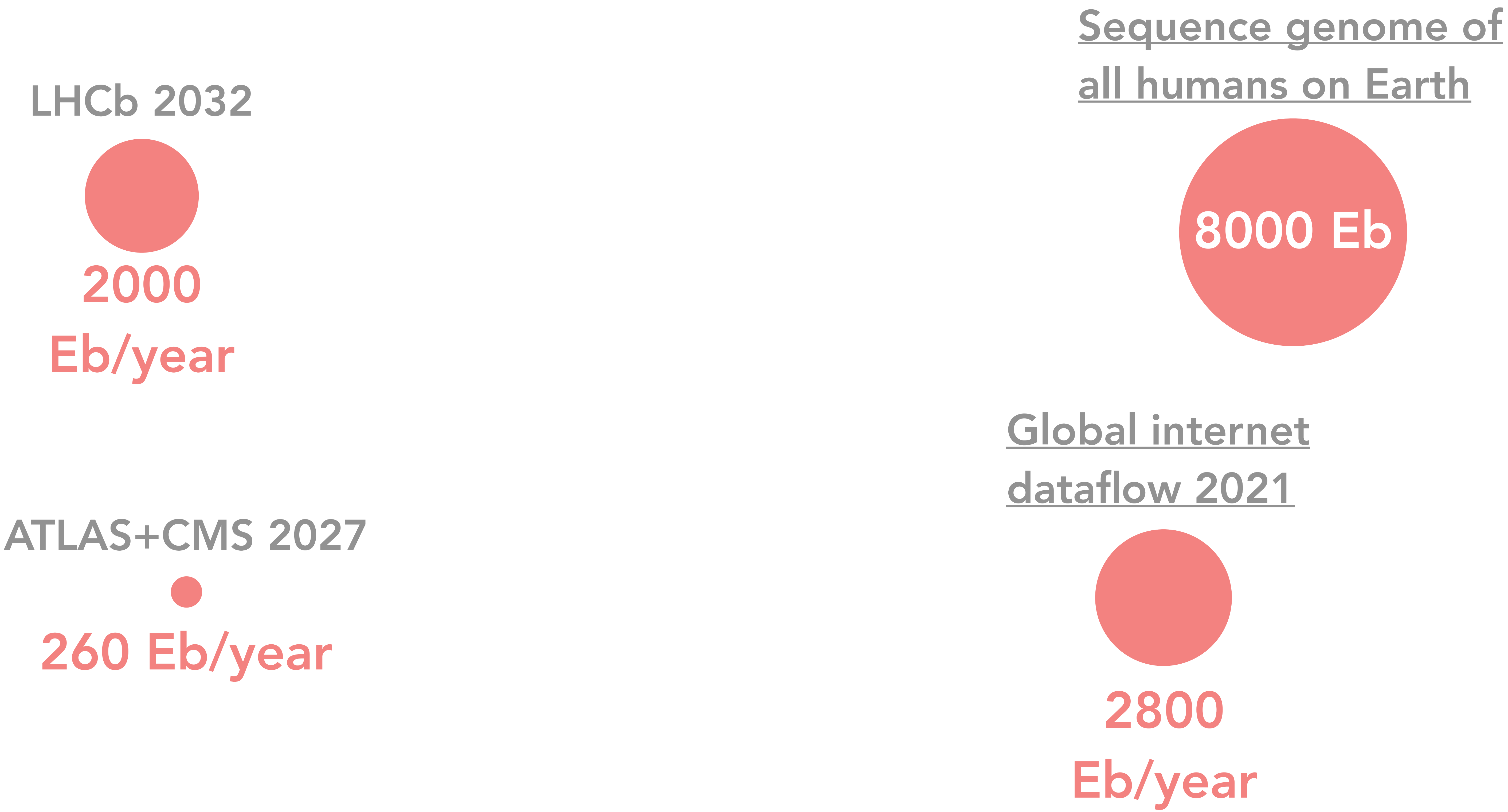
ATLAS+CMS 2027

260 Eb/year

Global internet dataflow 2021

2800
Eb/year

Context : other large data sources



Context : other large data sources

LHCb 2032

2000 Eb/year

Square Kilometre Array (2030s)

~30000 Eb/year

Sequence genome of all humans on Earth

8000 Eb

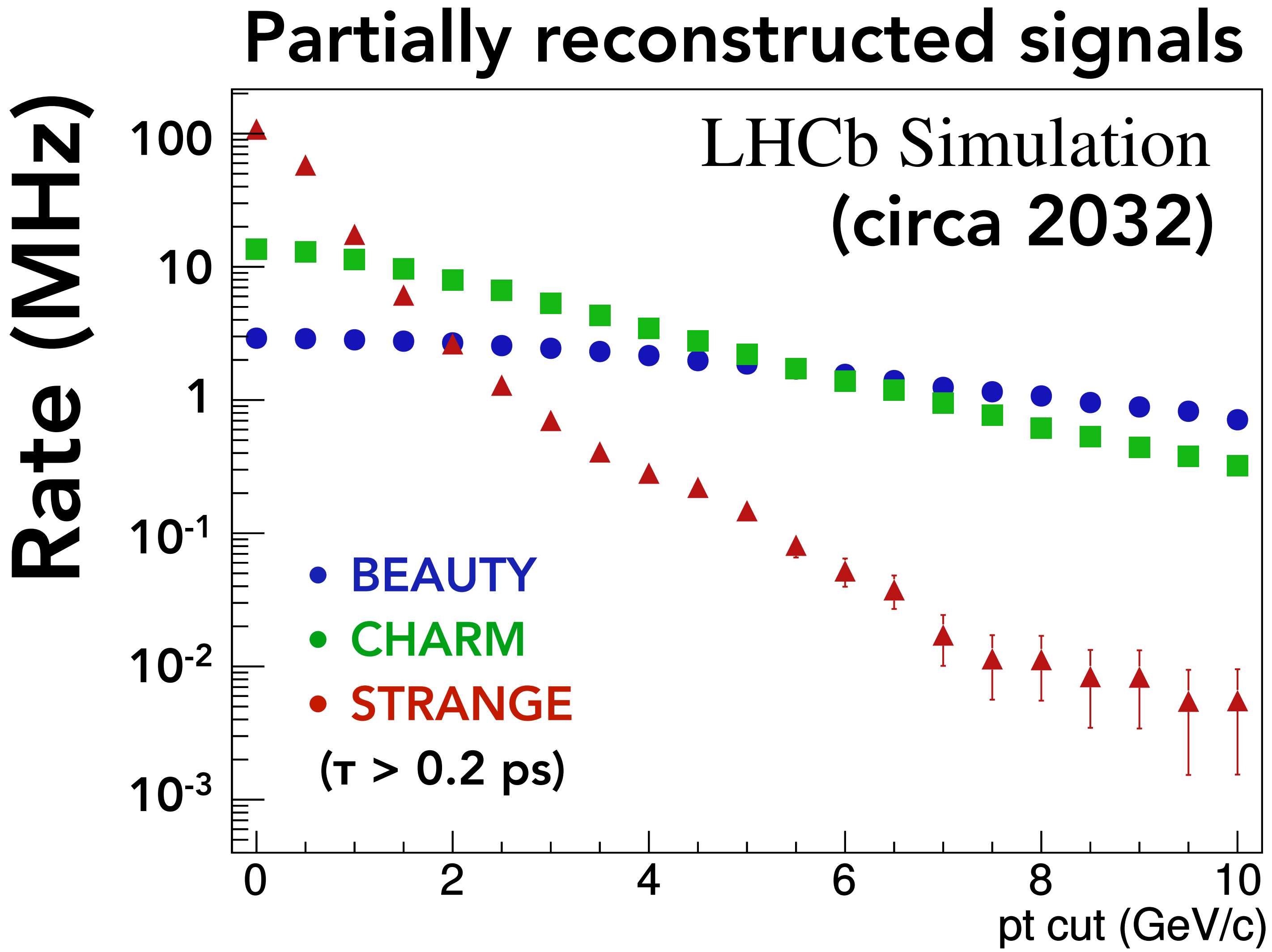
ATLAS+CMS 2027

260 Eb/year

Global internet dataflow 2021

2800 Eb/year

Context : signal density evolution

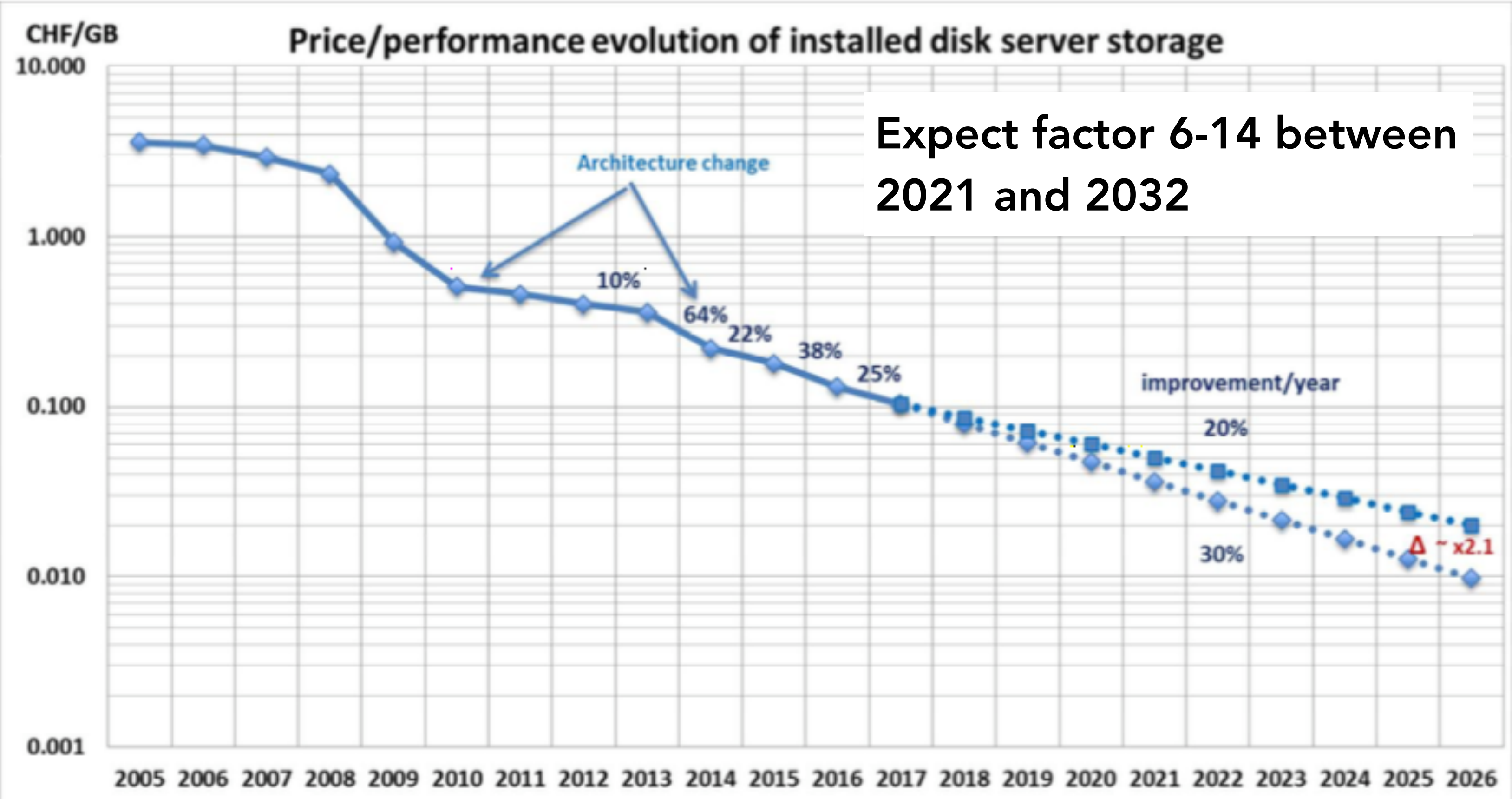


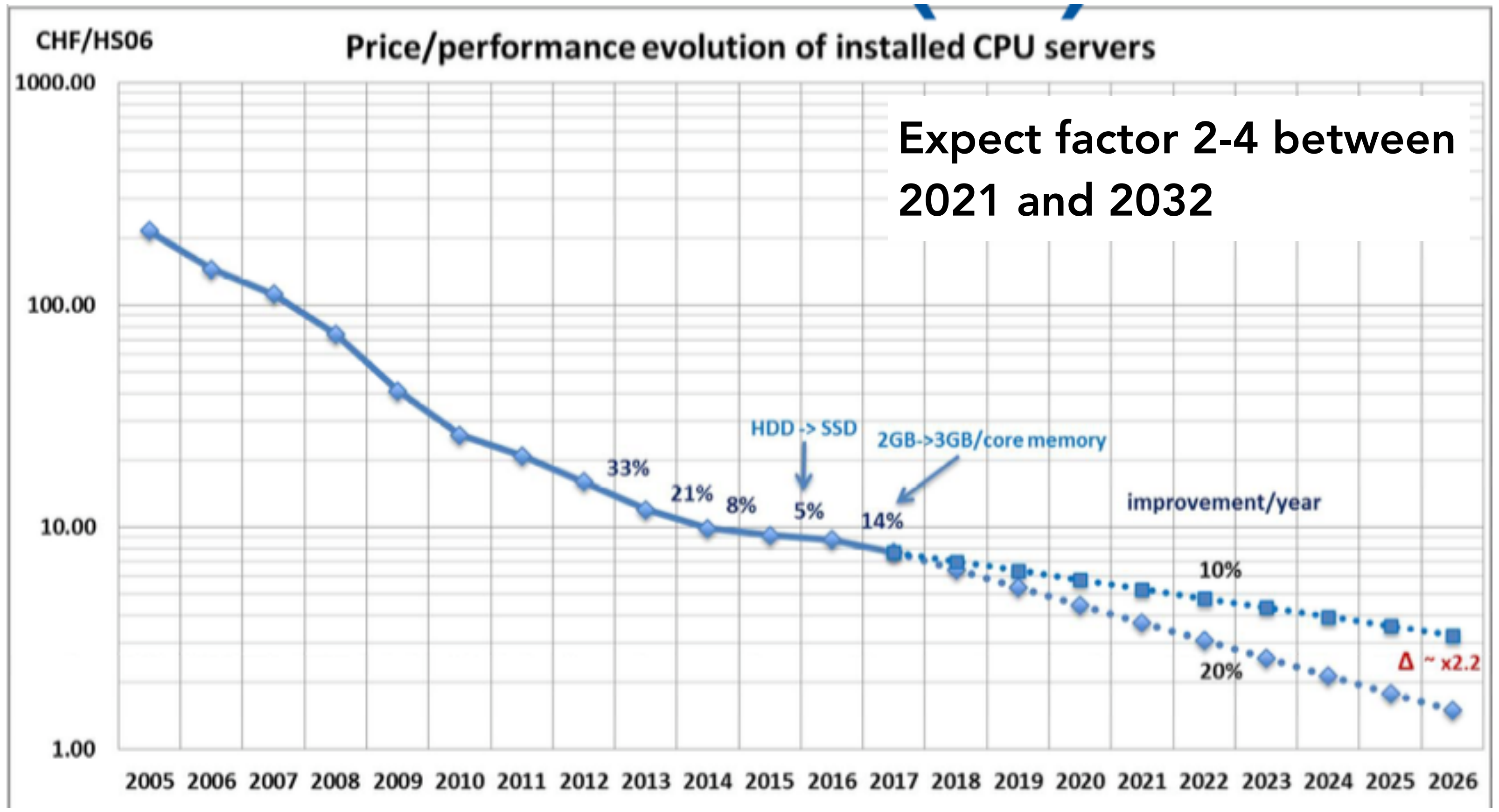
Almost all bunch crossings at our accelerators will contain interesting signal

Triggers are a limitation to be overcome : all this data should be analyzed!

Note that in HL-LHC period the ATLAS/CMS effective luminosity entering the software trigger is roughly the same as LHCb Upgrade I (1/30th the rate, 30x the pileup)!

Context : expected technology evolution

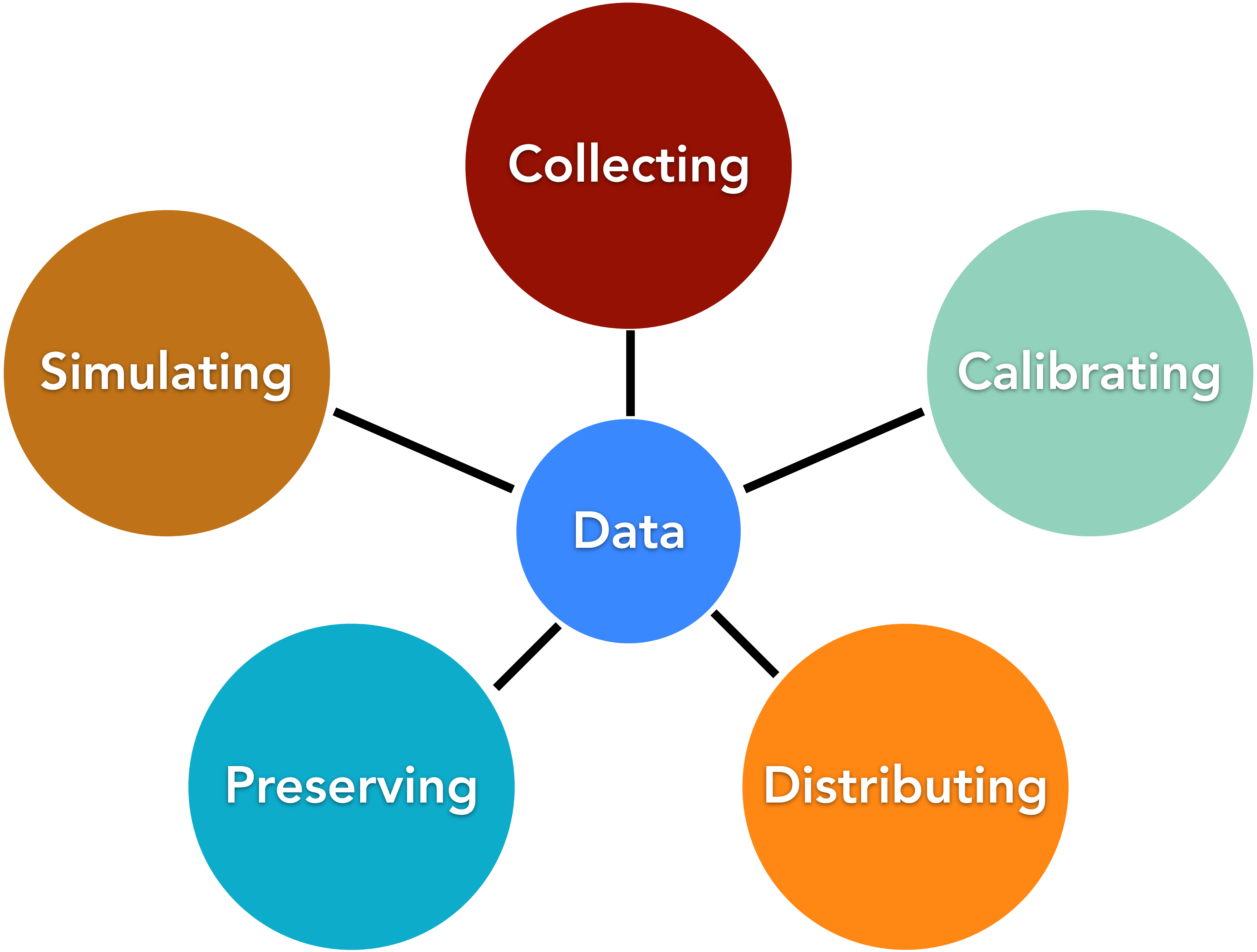


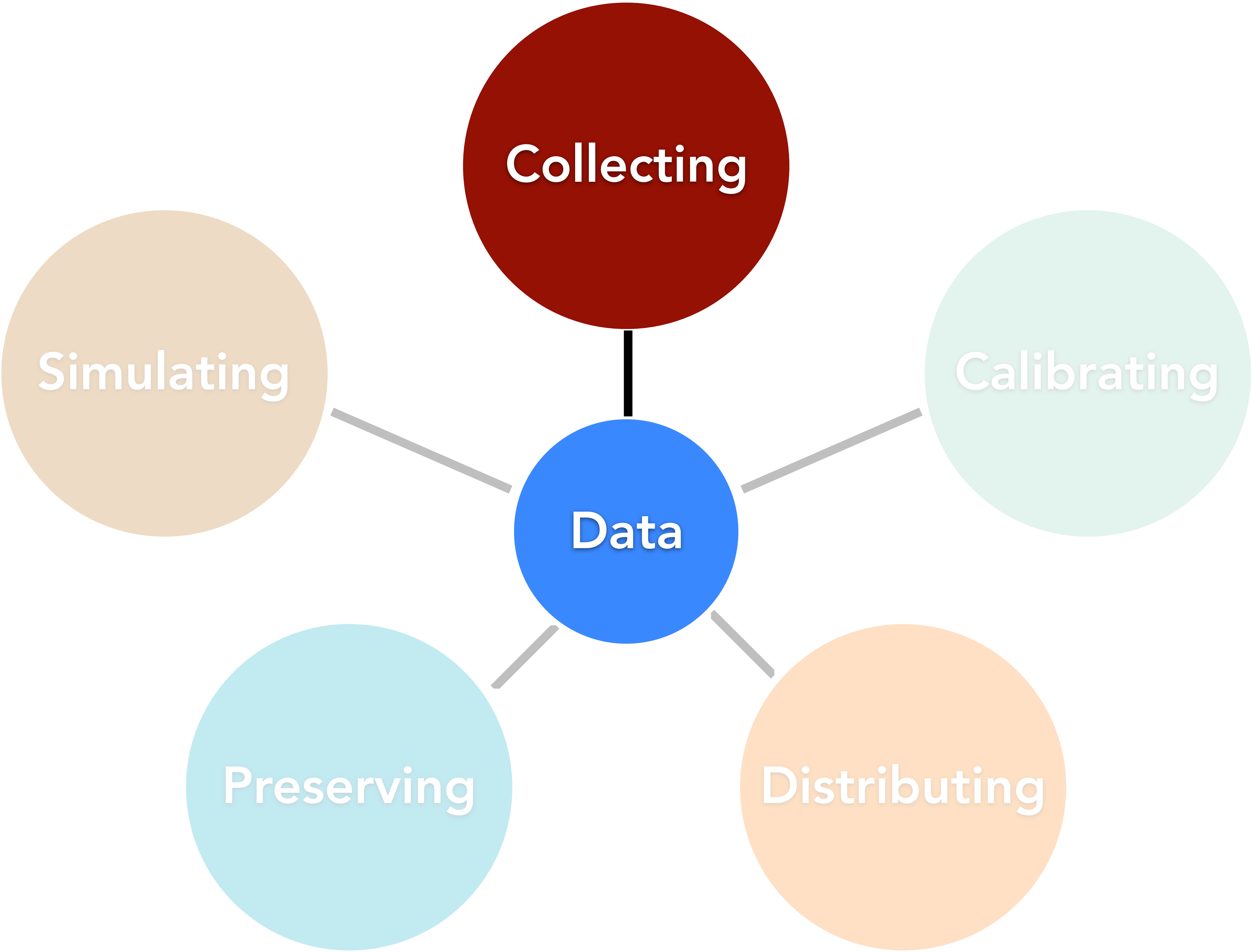


Price/performance gains in CPU & disk servers slowing down

Difficult to predict price/performance evolution of coprocessors & hybrid architectures, market driven.

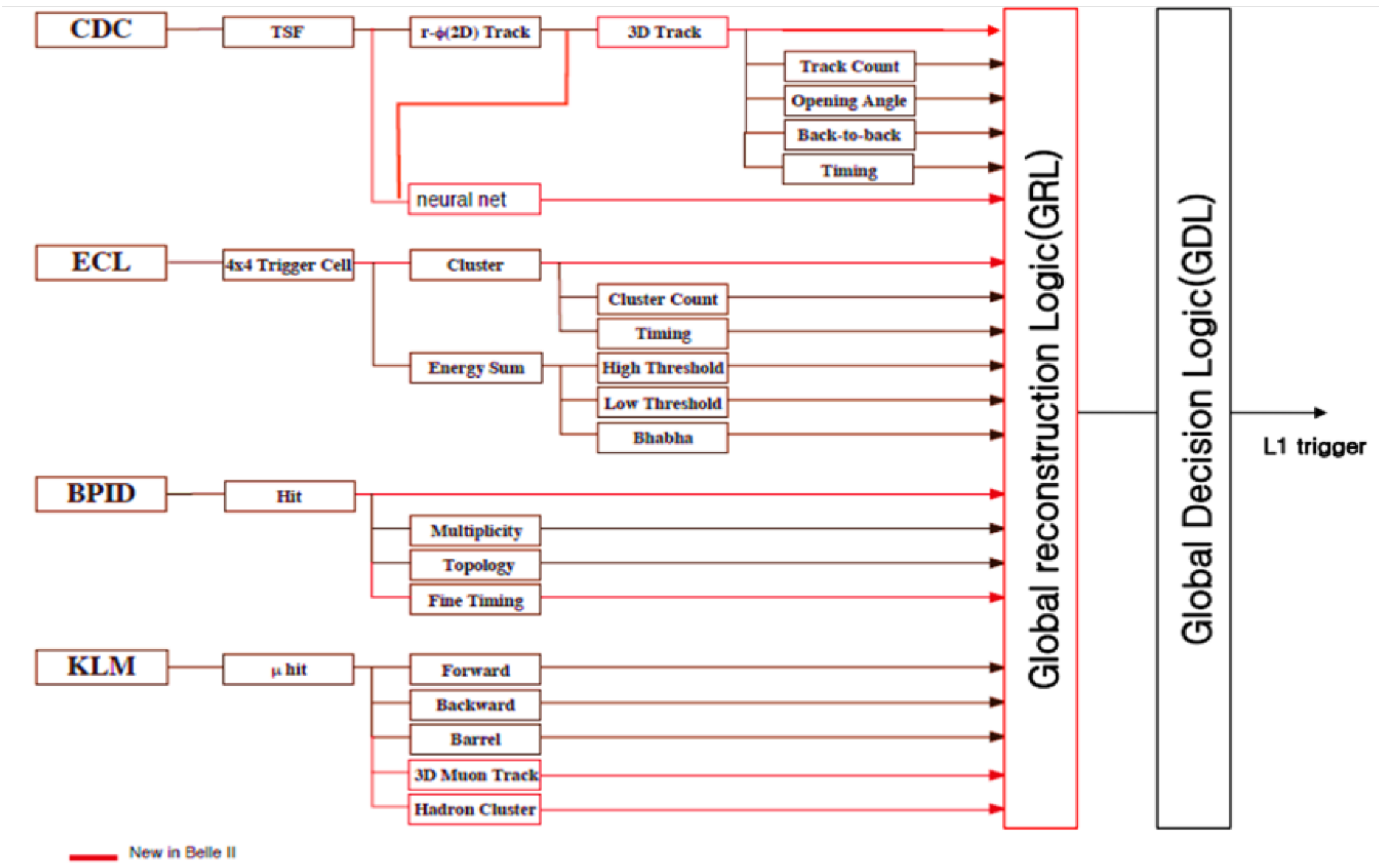
See Helge Meinhard’s talk at LHCb Elba workshop for details


Collecting
Simulating
Calibrating
Data
Preserving
Distributing


Collecting
Simulating
Calibrating
Data
Preserving
Distributing

Belle–II DAQ & trigger

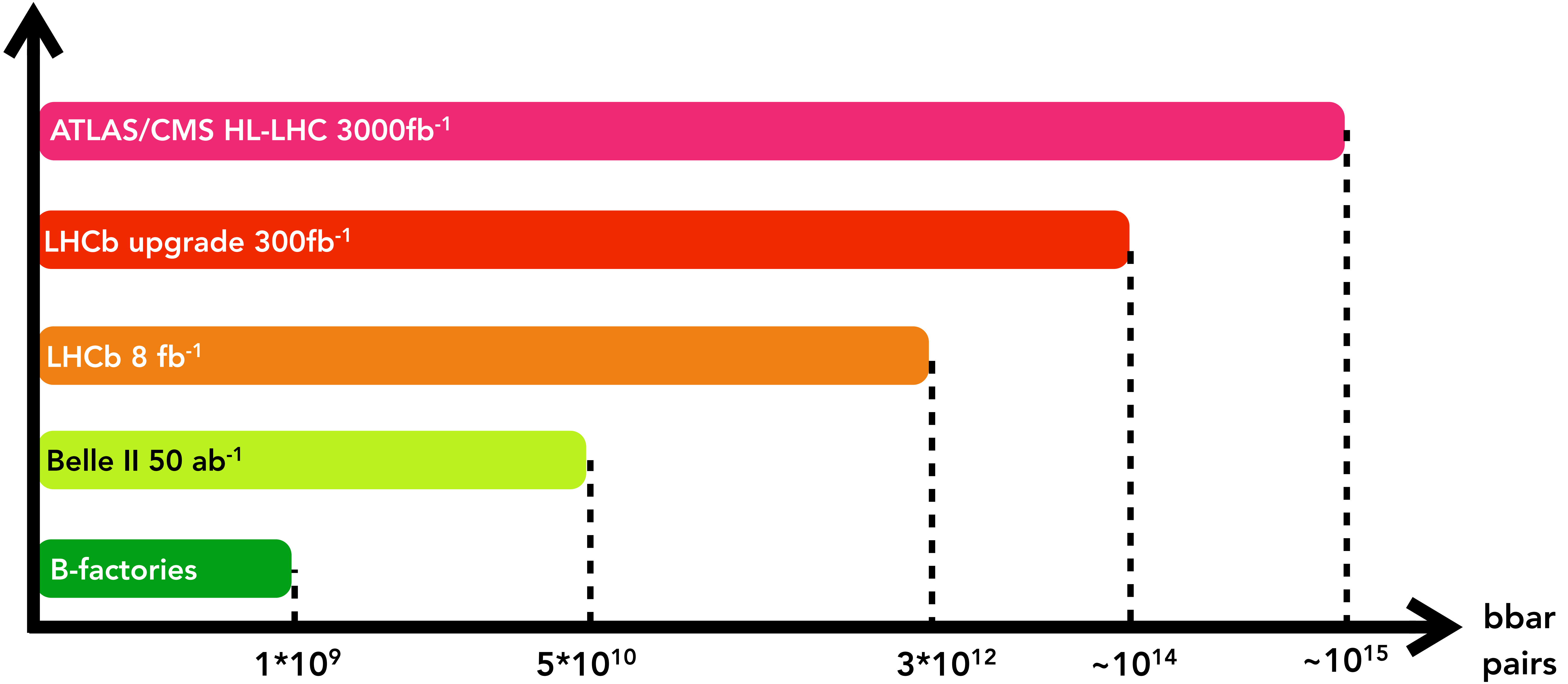
More details



Based around fixed-latency trigger which is ~100% efficient for core beauty physics

Major challenges are reconstruction, analysis, and simulation. Note in particular that the Belle II user-level data volume (ntuples & MC) comparable in size to other experiments.

HL-LHC, the biggest flavour factory of all



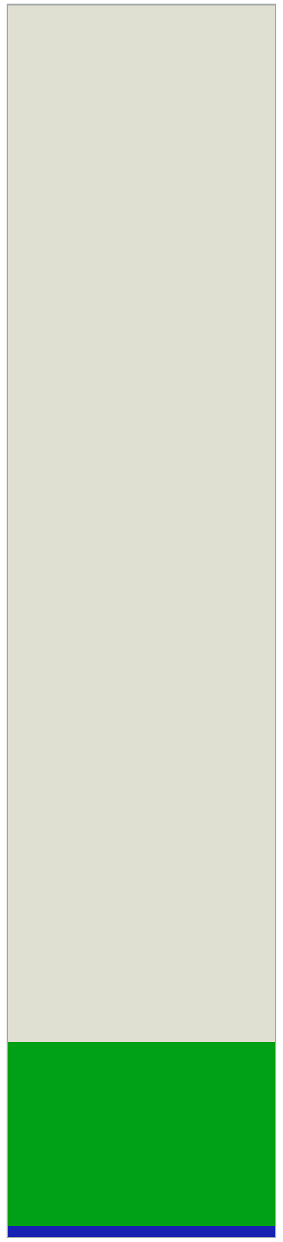
Islands of signal in a sea of QCD

One pp collision

Islands of signal in a sea of QCD

One pp collision

0.6% beauty
15% charm
85% uds



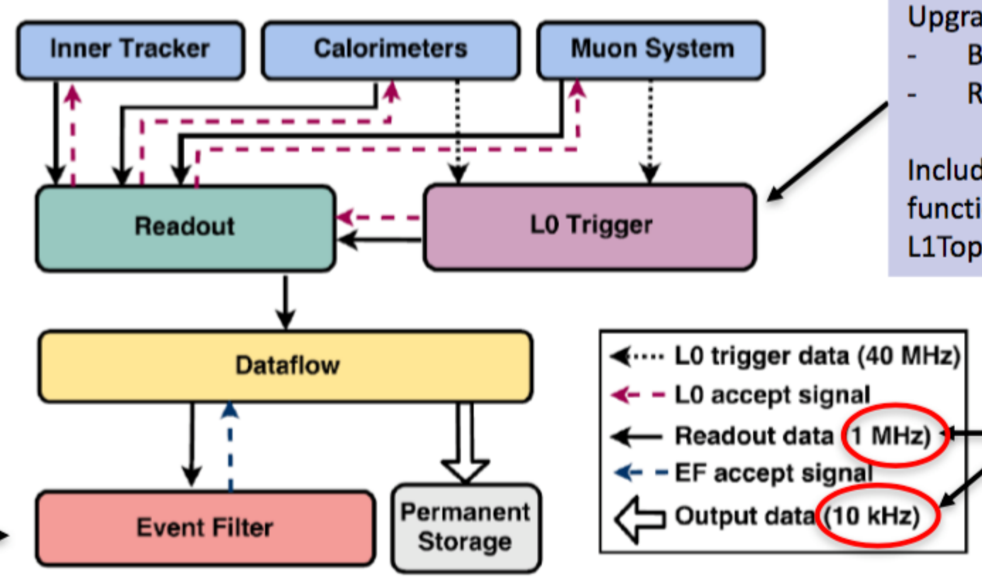
One bunch crossing @ $7.5 \cdot 10^{34}$ cm^{-2}s^{-1}

120% beauty
3000% charm
0% uds

Most pp collisions do not contain heavy flavour, but all bunch crossings do! Above percentages roughly based on measured cross-sections extrapolated to 4π, just to give an idea…

The ATLAS & CMS DAQ and triggers

ATLAS



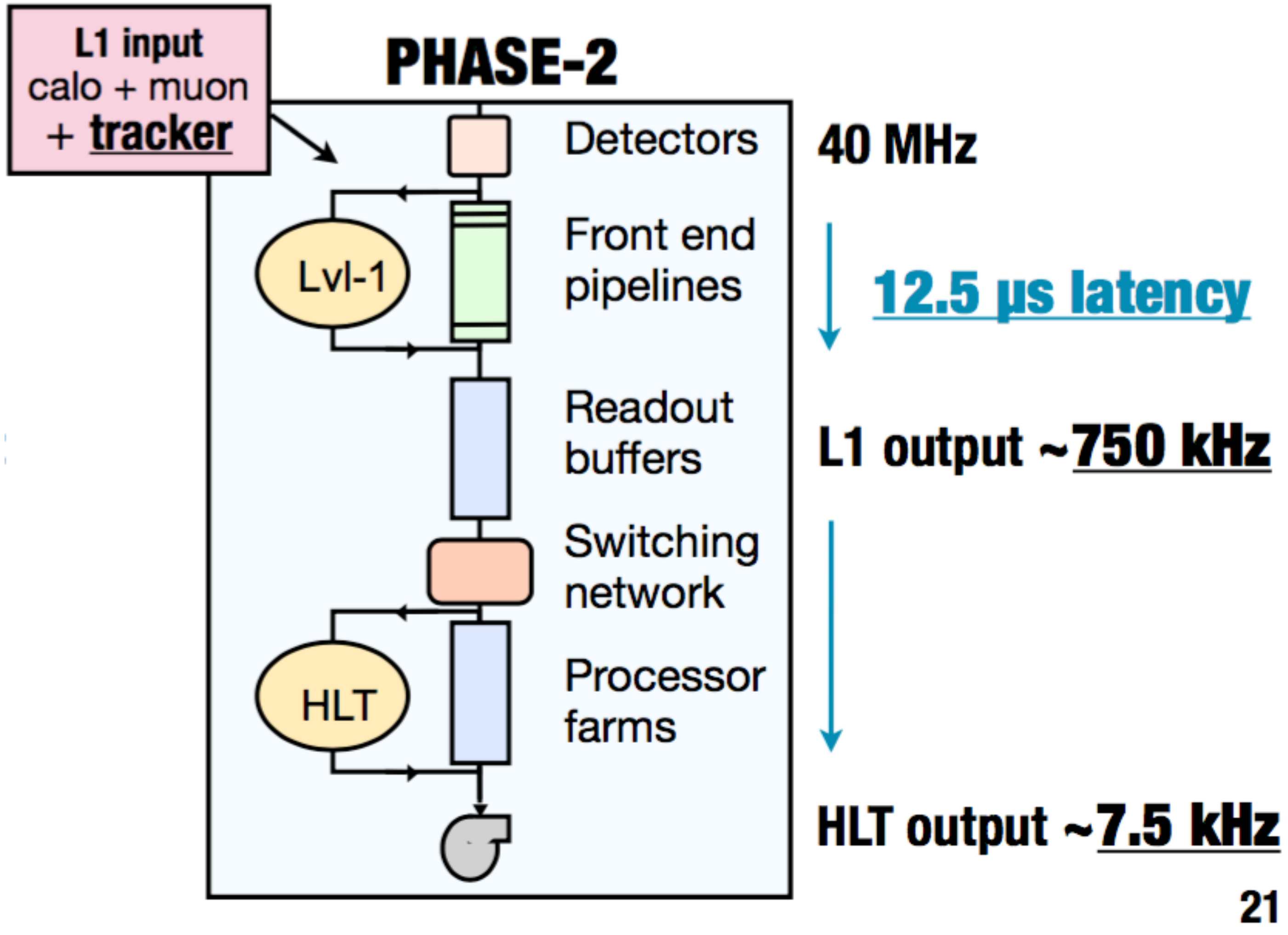
Upgrades to muon trigger
- Better efficiency
- Reduced rate

Includes 'Global" processor with functionality akin to current L1Topo

10x more than now

Including Hardware Track Trigger

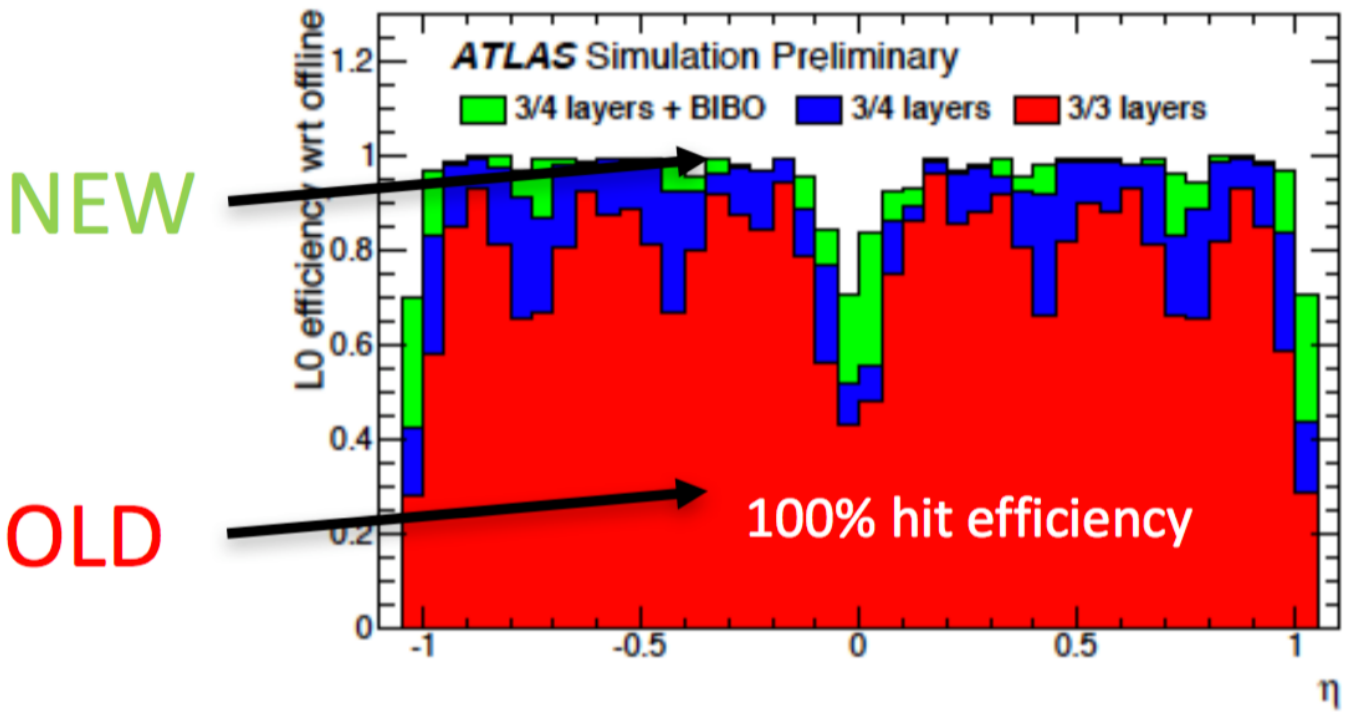
CMS



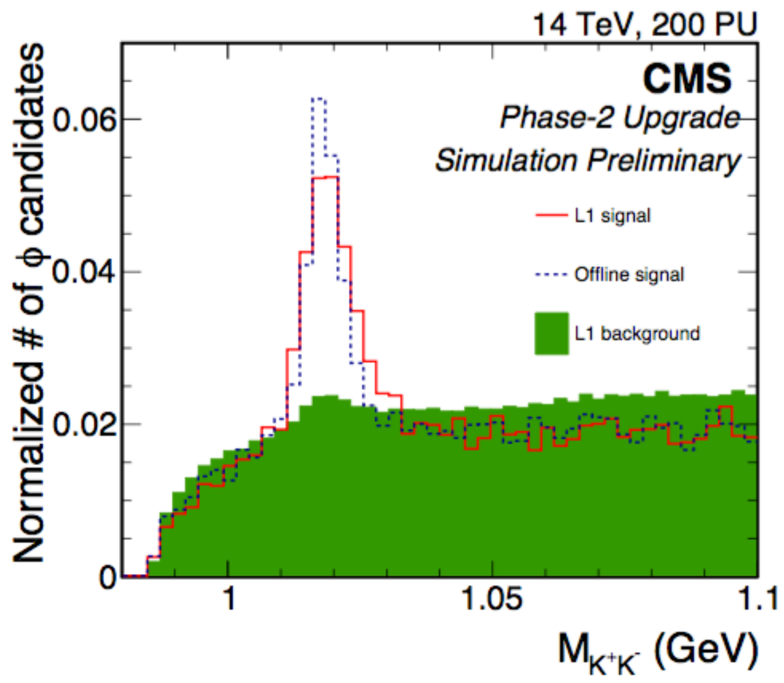
Similar architectures : ATLAS track trigger comes in at 1 MHz, CMS at 40 MHz but with an efficiency which drops rapidly for tracks not coming from the beamline.

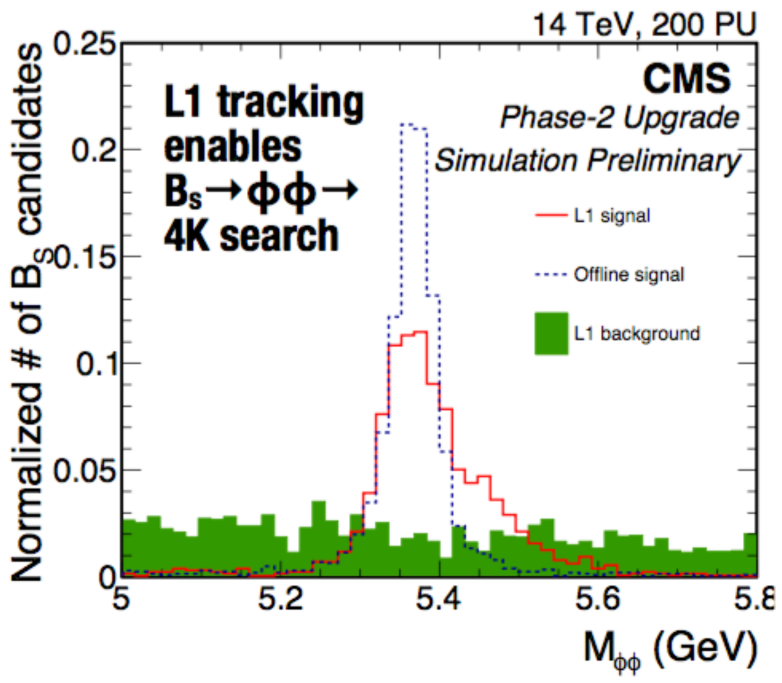
Gains from HL-LHC GPD triggers

ATLAS



CMS



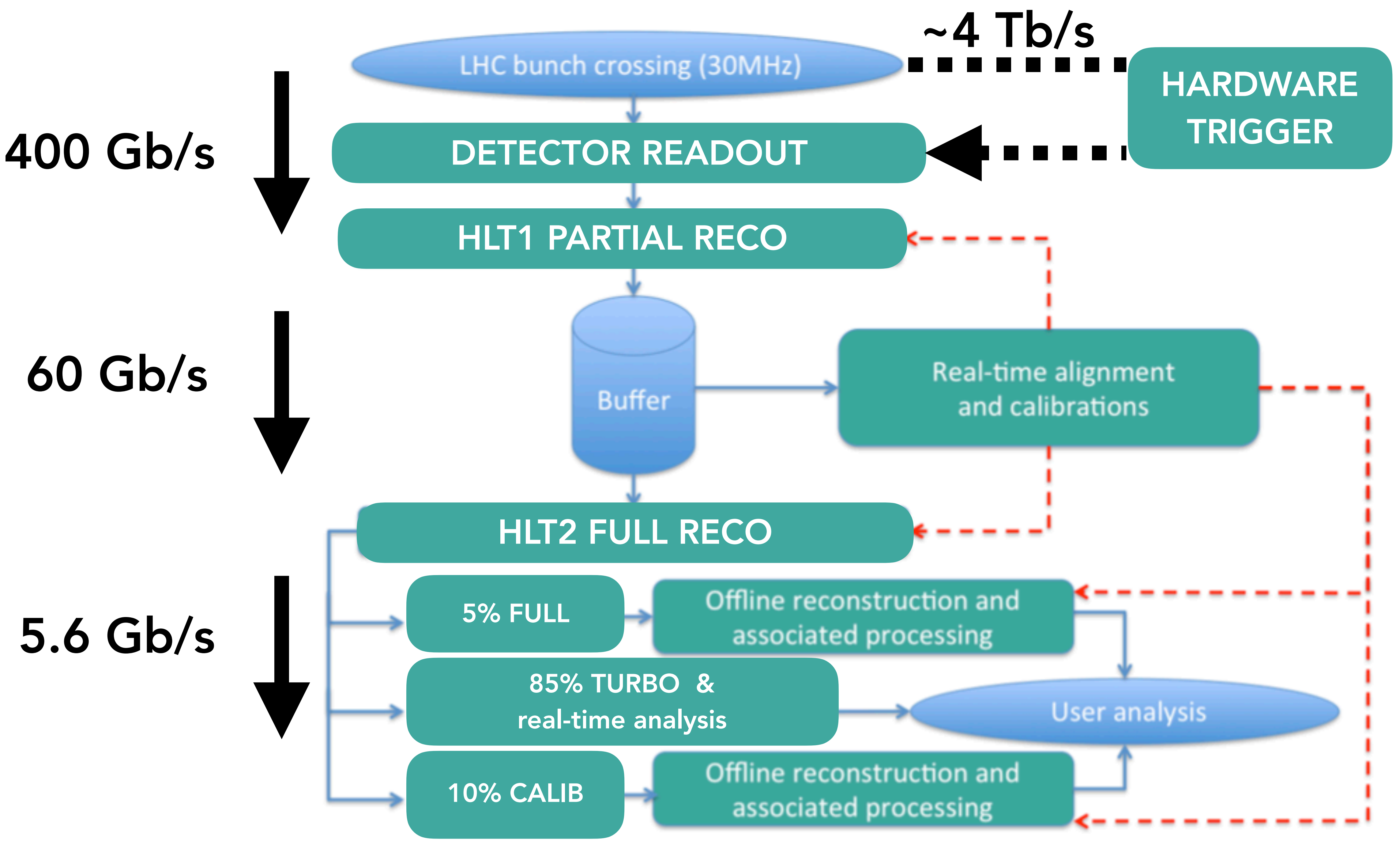


Major improvements to ATLAS muon system may enable HL-LHC muon trigger thresholds to remain the same as in Run 2 — under investigation!
CMS track trigger may be transformative, e.g. allowing CMS to study $B_s \to \varphi\varphi$ with an interesting precision. Personal remark : why not $B_s \to \varphi\gamma$ as well?!

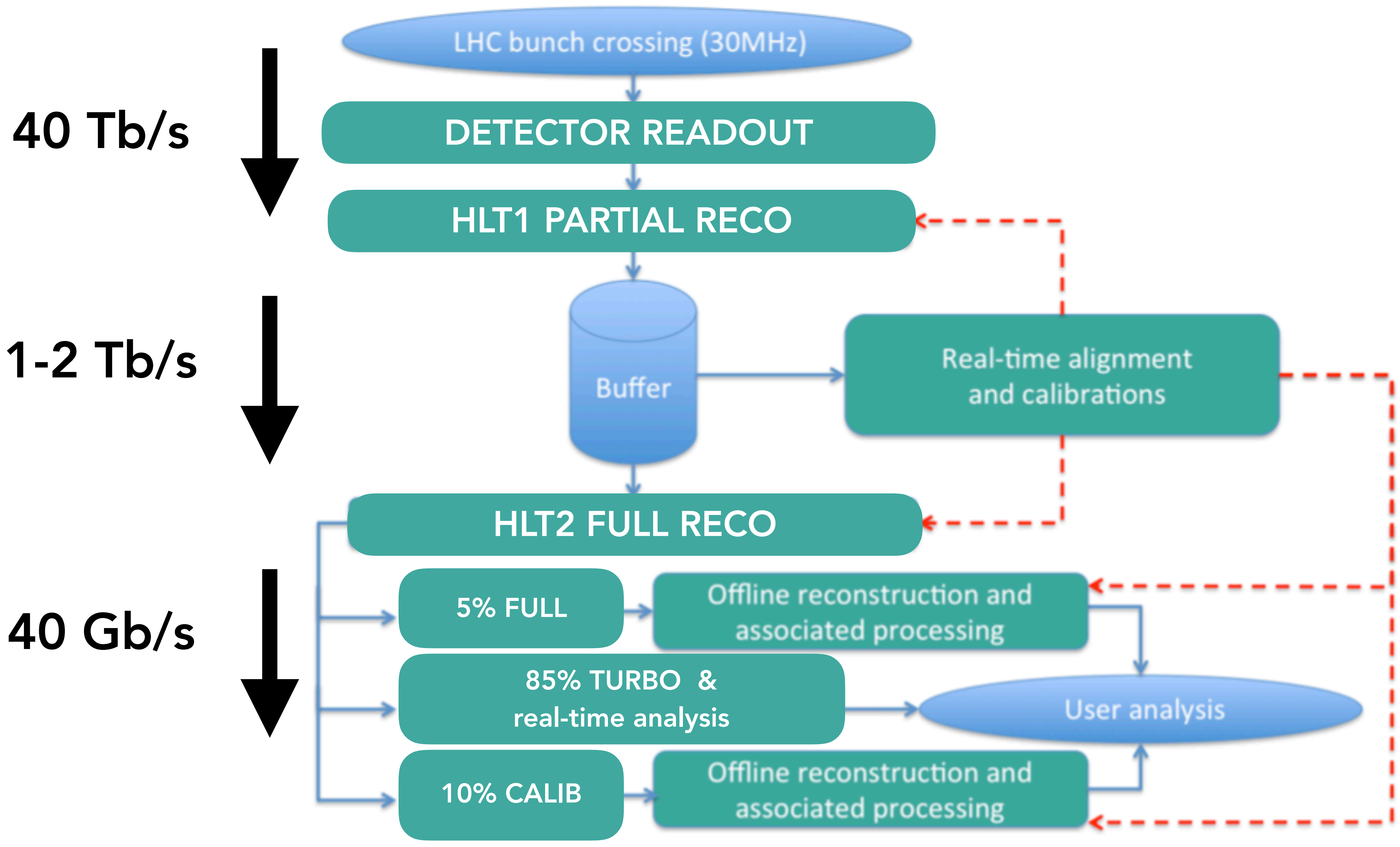
The real-time analysis paradigm

Fully reconstruct, align, and calibrate the detector in real time, and extract the signal with analysis-level precision in real time, then compress by discarding all detector information.
LHCb, ATLAS, CMS and ALICE pursuing similar strategies!
HLT resource constraints critical for full exploitation of this approach.

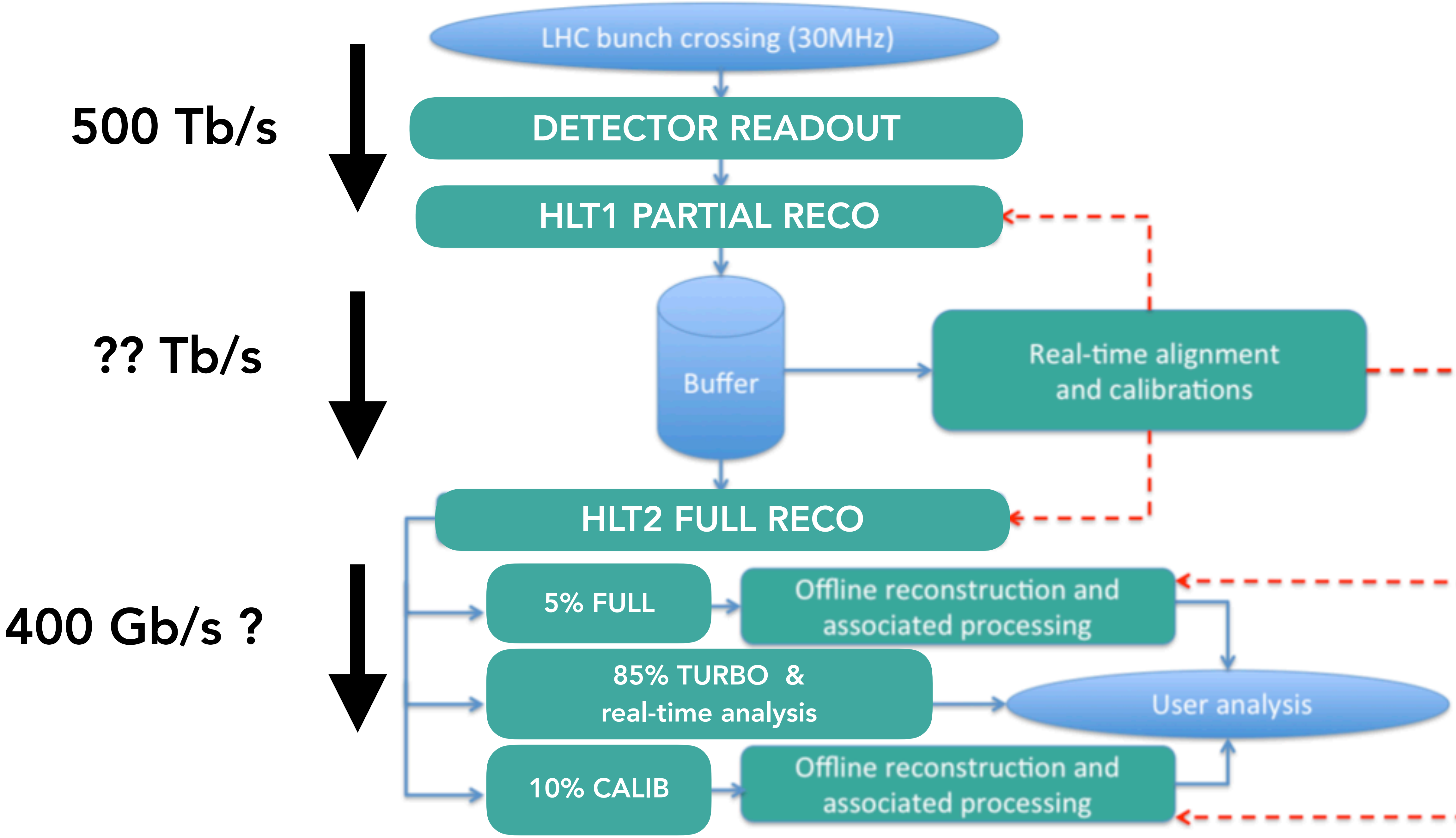
LHCb real-time analysis, present



LHCb real-time analysis, upgrade I

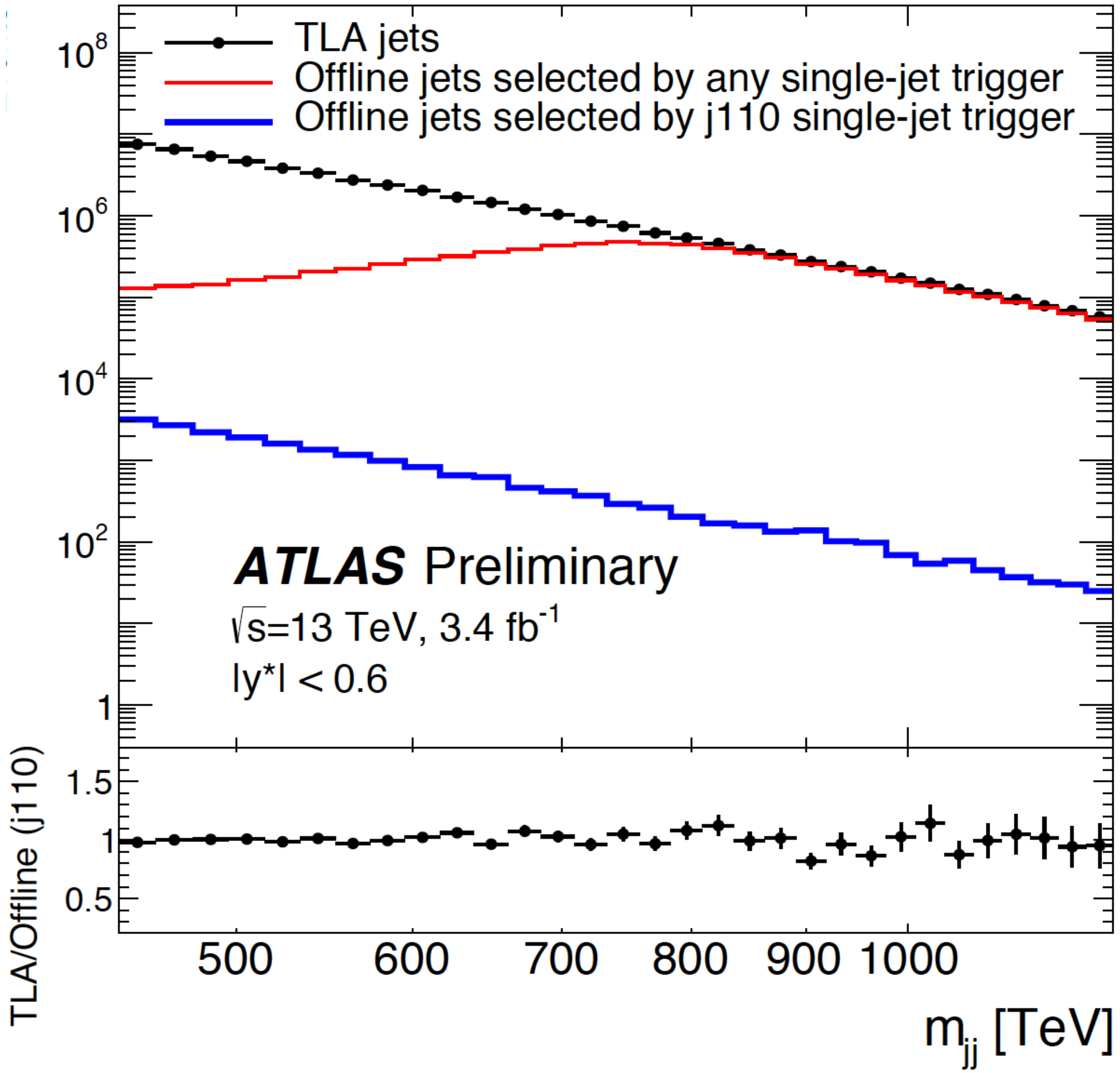


LHCb real-time analysis, upgrade II

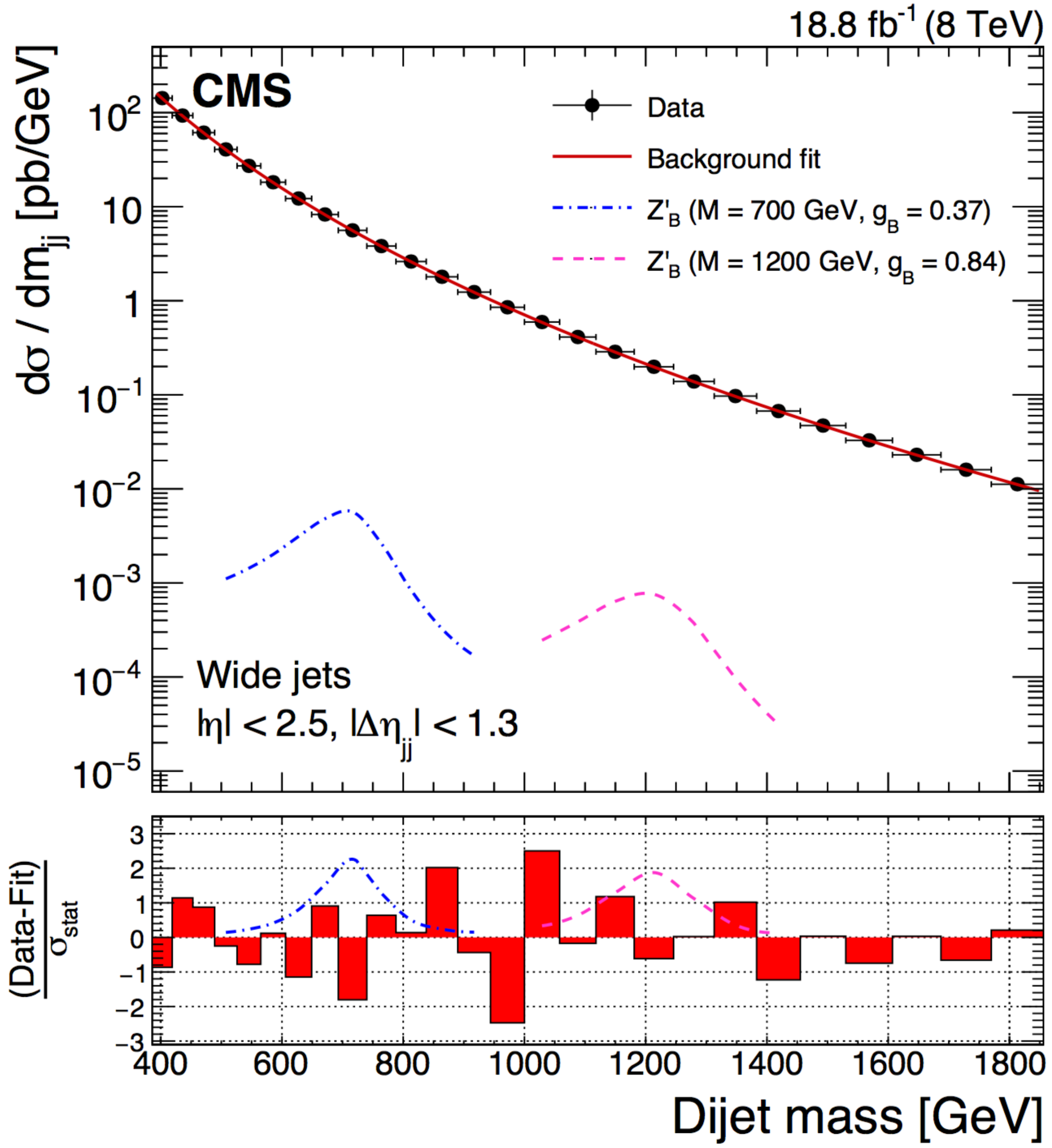


ATLAS & CMS real-time analysis

ATLAS dijet analysis

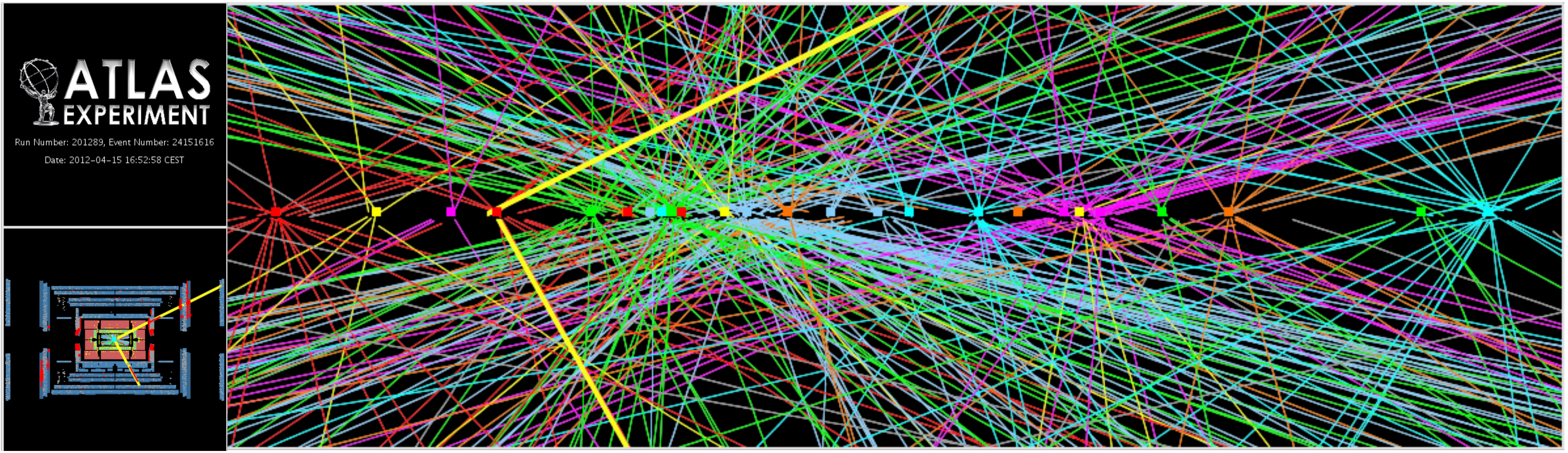


CMS dijet analysis



At ATLAS & CMS so far used for dijet analyses, expand to flavour physics in HL-LHC?

Towards a solution : pileup suppression



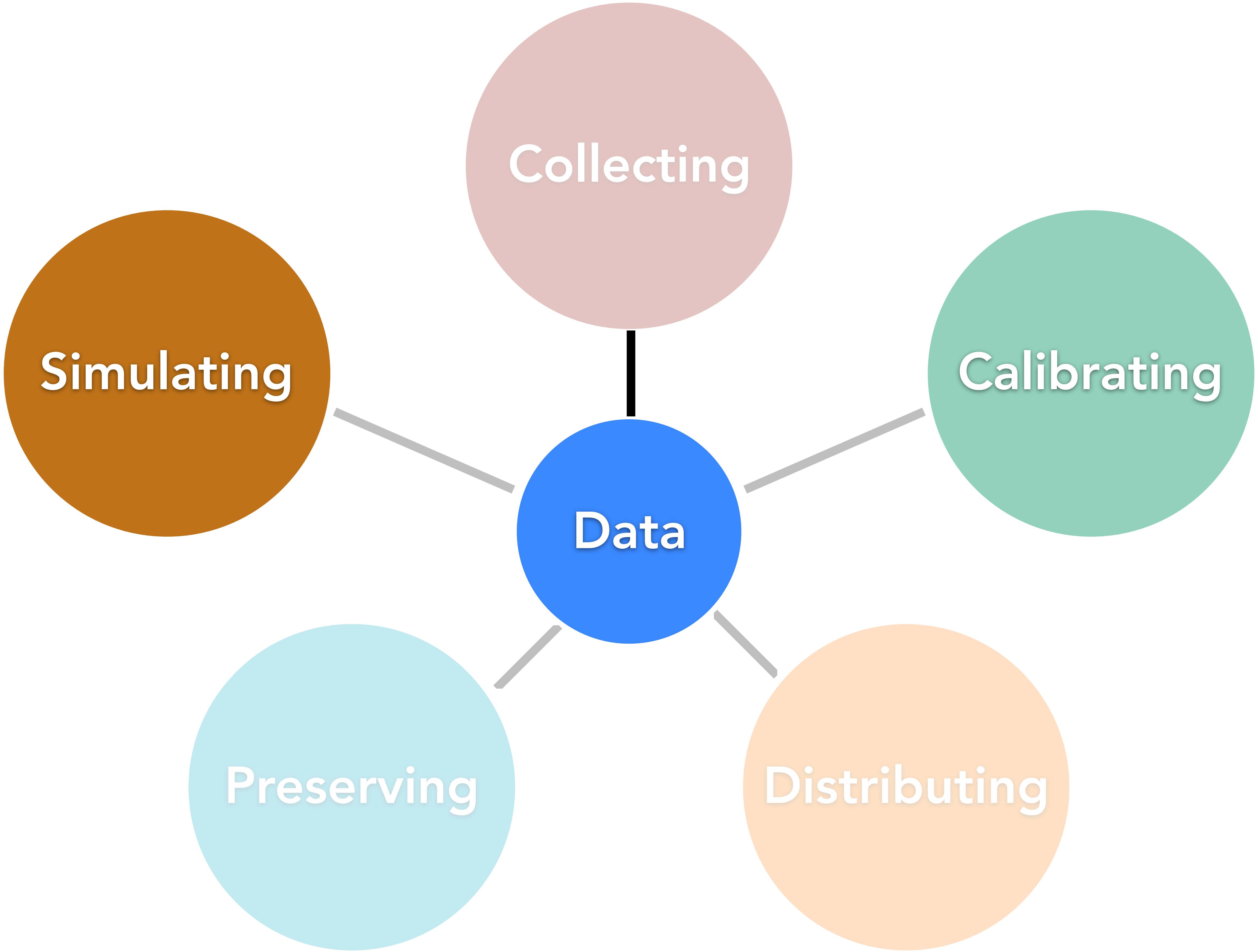
Almost all bunch crossings will contain interesting signal, most proton-proton collisions will not

➡ Our triggers should select collisions, not bunch crossings

Requires ~offline-quality real-time reconstruction, detector alignment&calibration

Requires access to “rest of event” information (tagging, isolation…) in real-time

Could be applied to flavour physics at the LHC for all detectors in HL-LHC era!


Collecting
Simulating
Calibrating
Data
Preserving
Distributing

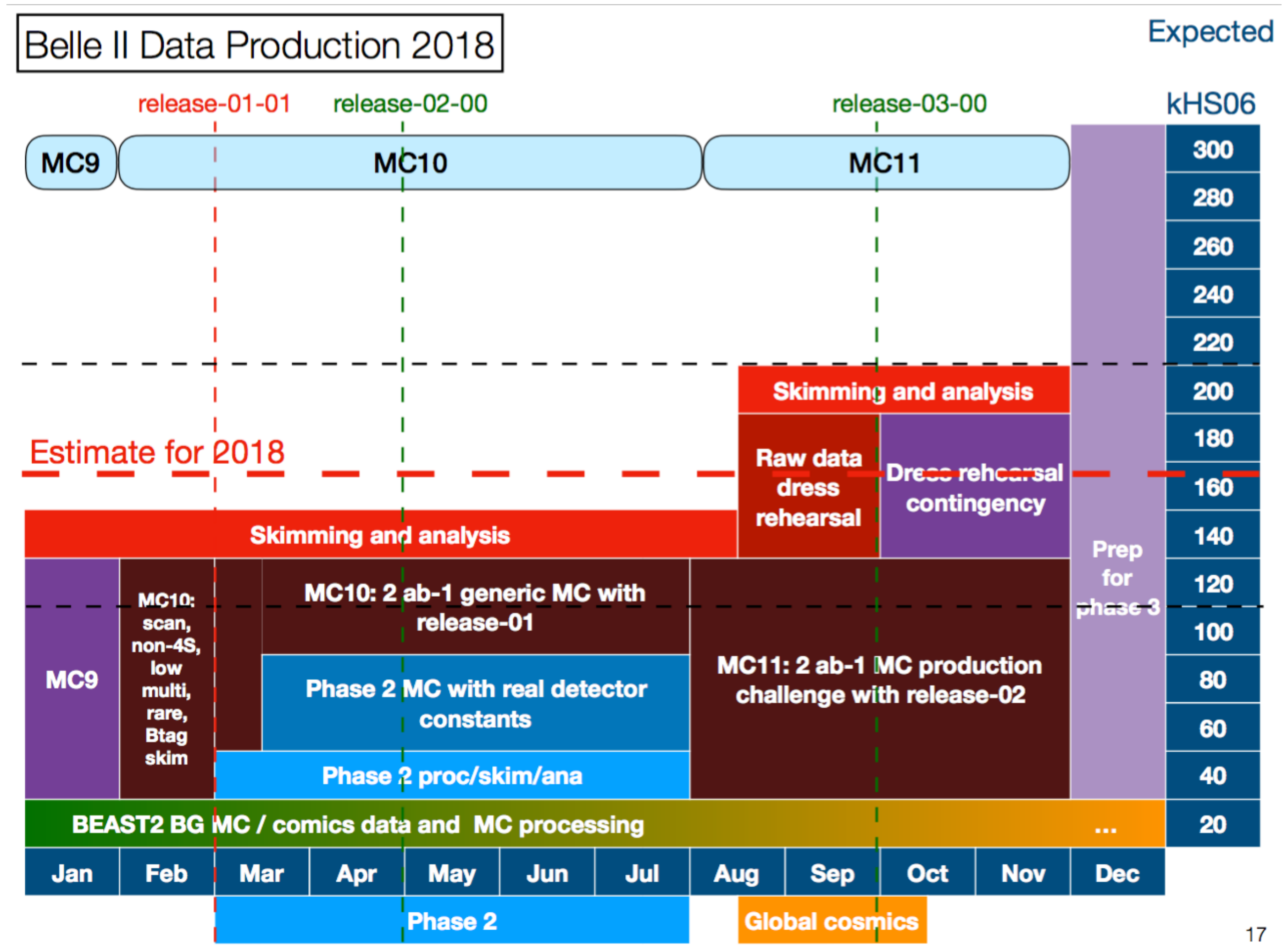
Belle-II calibration and simulation

W/o signal MC

Year	2018 Jan- Dec	Jan 2019 – Mar2020	Apr 2020 – Mar 2021	Apr 2021 – Mar 2022	Apr 2022 – Mar 2023
Total tape (PB)	1.6	10.7	24.4	46.2	73.3
Total Disk (PB)	3.5	12.1	21.7	25.9	45.0
Total CPU (kHEPSpec)	175	400	503	590	753

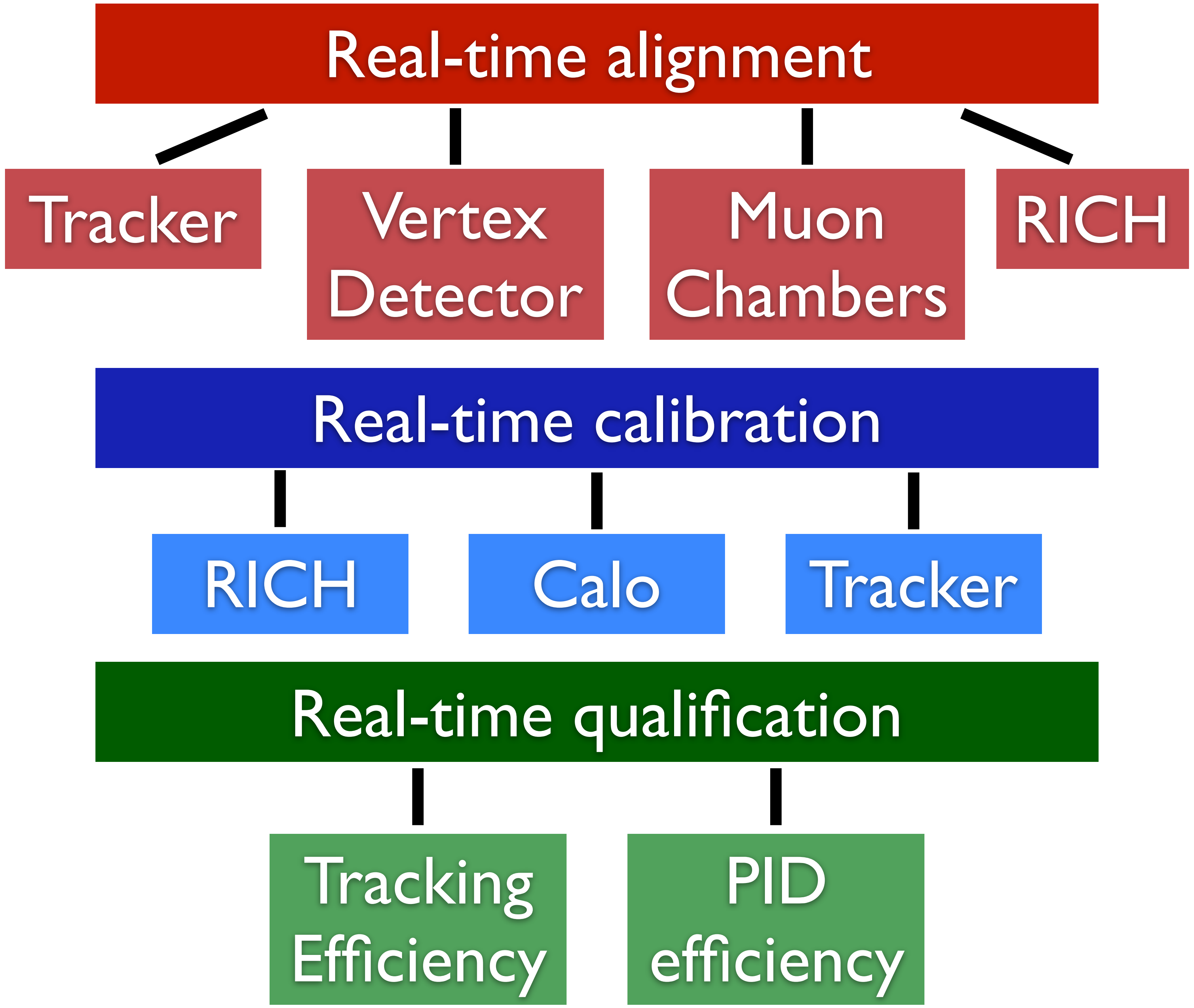
With signal MC

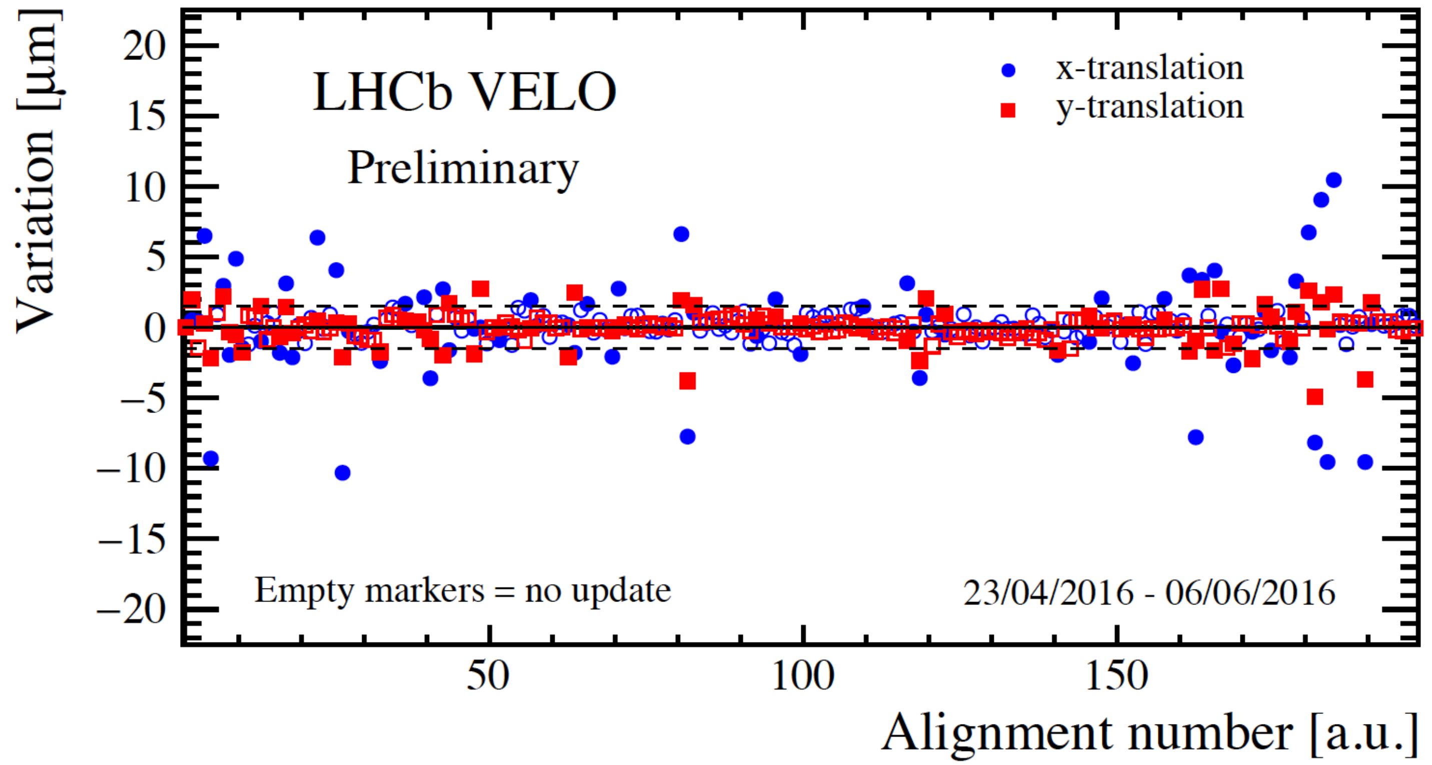
Year	2018 Jan- Dec	Jan 2019 – Mar2020	Apr 2020 – Mar 2021	Apr 2021 – Mar 2022	Apr 2022 – Mar 2023
Total tape (PB)	1.6	10.7	24.4	46.2	73.3
Total Disk (PB)	3.5	12.2	21.9	26.4	45.7
Total CPU (kHEPSpec)	175	403	510	599	766

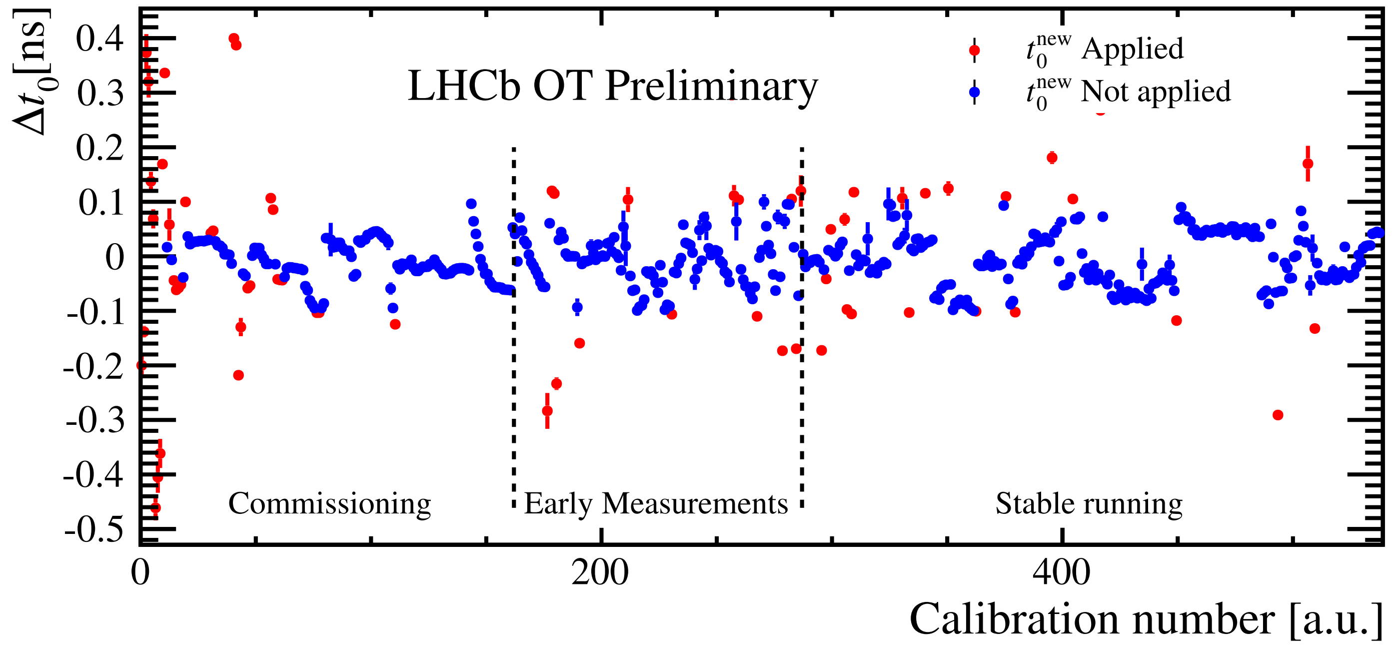


As trigger is fully efficient, based on calibrating and reprocessing the detector after data is taken. Simulation requirements driven by background processes, not signal.
Preparations for datataking are well underway, CPU resources driven by simulation.

LHCb real-time alignment & calibration





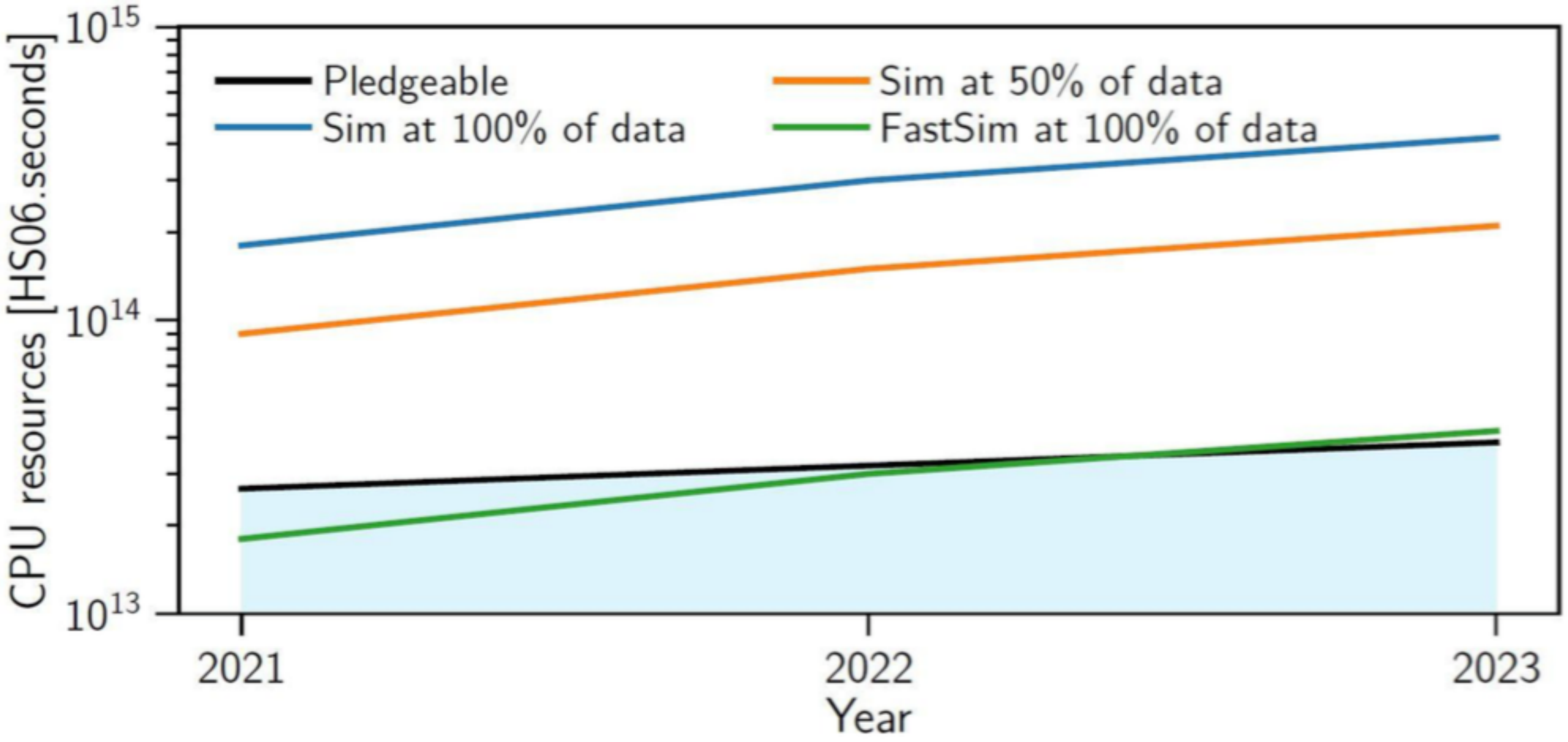


Deployed in 2015 and continuously developed and improved since. Forcing ourselves to perform alignment & calibration in real-time led to many improvements to algorithms.

Simulating @ ultimate precision

More details

Legend: "Sim at 50% of data" = FullSim sample is 50% the datasize, FastSim sample is 50% the datasize



FastSim assumed to be 1/10 faster than FullSim

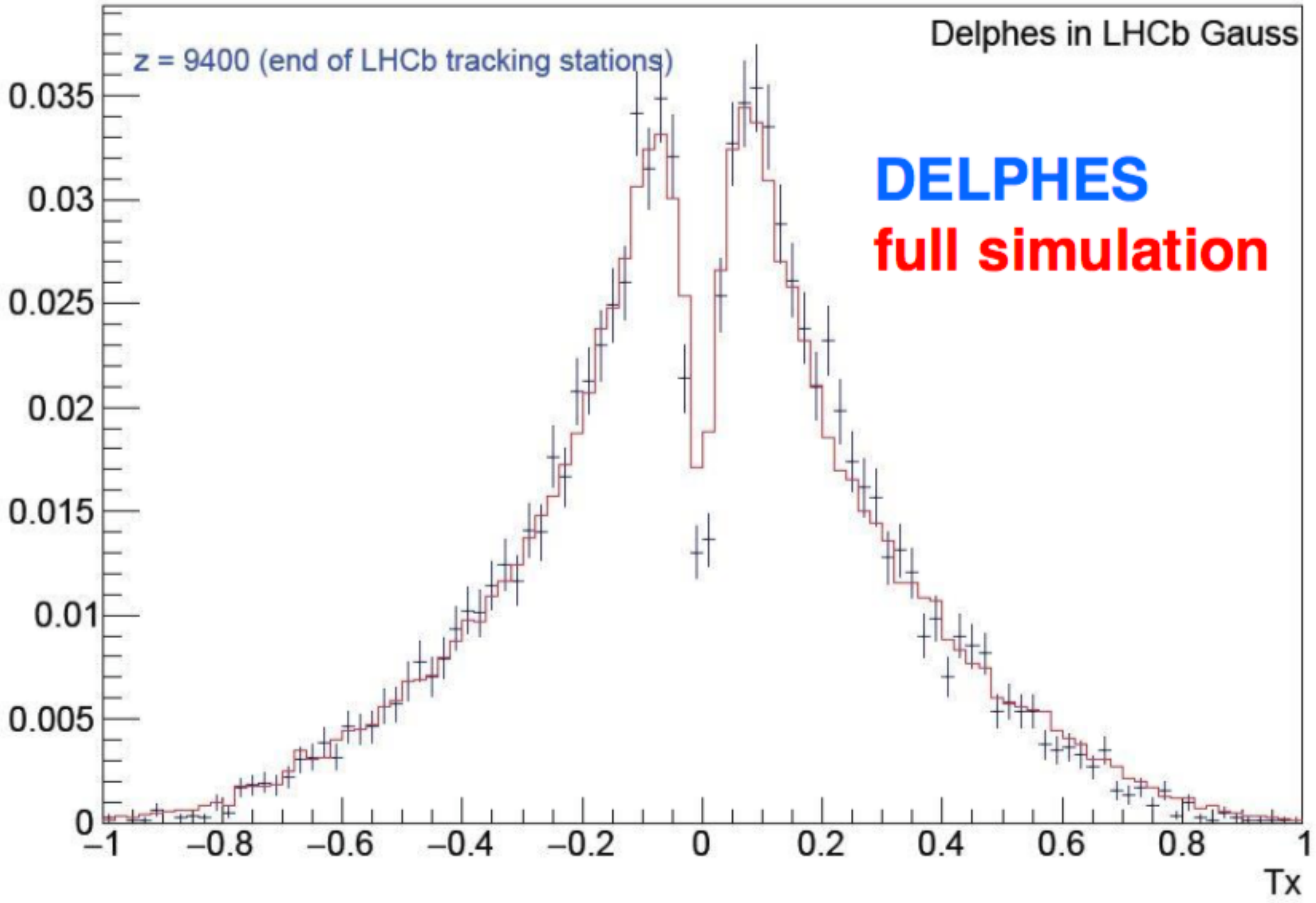
Example of the problem for LHCb, but similar issues exist for other collaborations
A huge computing challenge, may get even more complex if processing hardware at GRID sites diverges from real-time centres. Is WLCG model tenable in real-time analysis era?

Fast or partial simulations

	Generator	Digitization	Reconstruction
Purpose	"True" picture of bunch crossing	Simulate passage of particles through detector material	Apply same real-time and offline processing as done to data
Cost	Depends on process, to some extent irreducible	1-100x cost of reconstructing data	1x cost of reconstructing data w/software, more if emulating hardware triggers
Solution	Generate signal & pileup separately, reuse one pileup for many signal events	Parametric or ML based simulation trained on data & full GEANT simulation	Reconstruct only part of detector, e.g. tracker

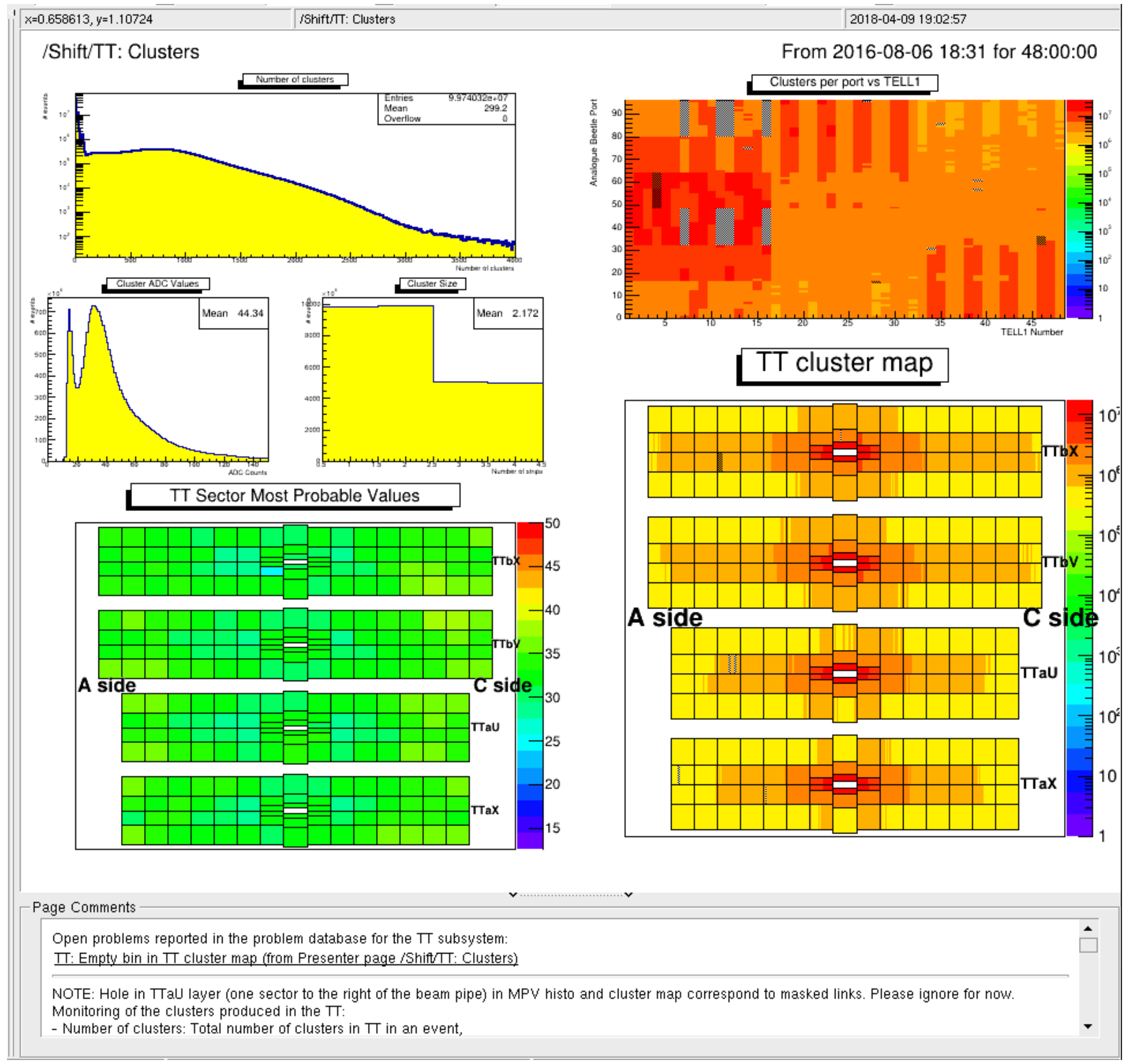
Many different approaches on the market, best one likely to be analysis dependent. Key is to maintain a flexible framework which allows to deploy best one for each problem.

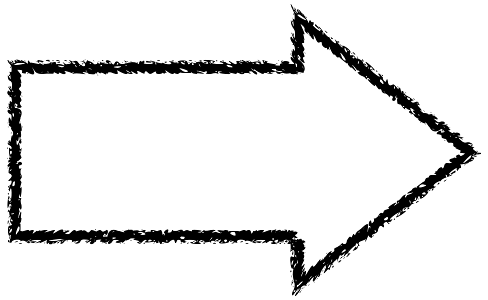
Calibrating parametric/NN simulations

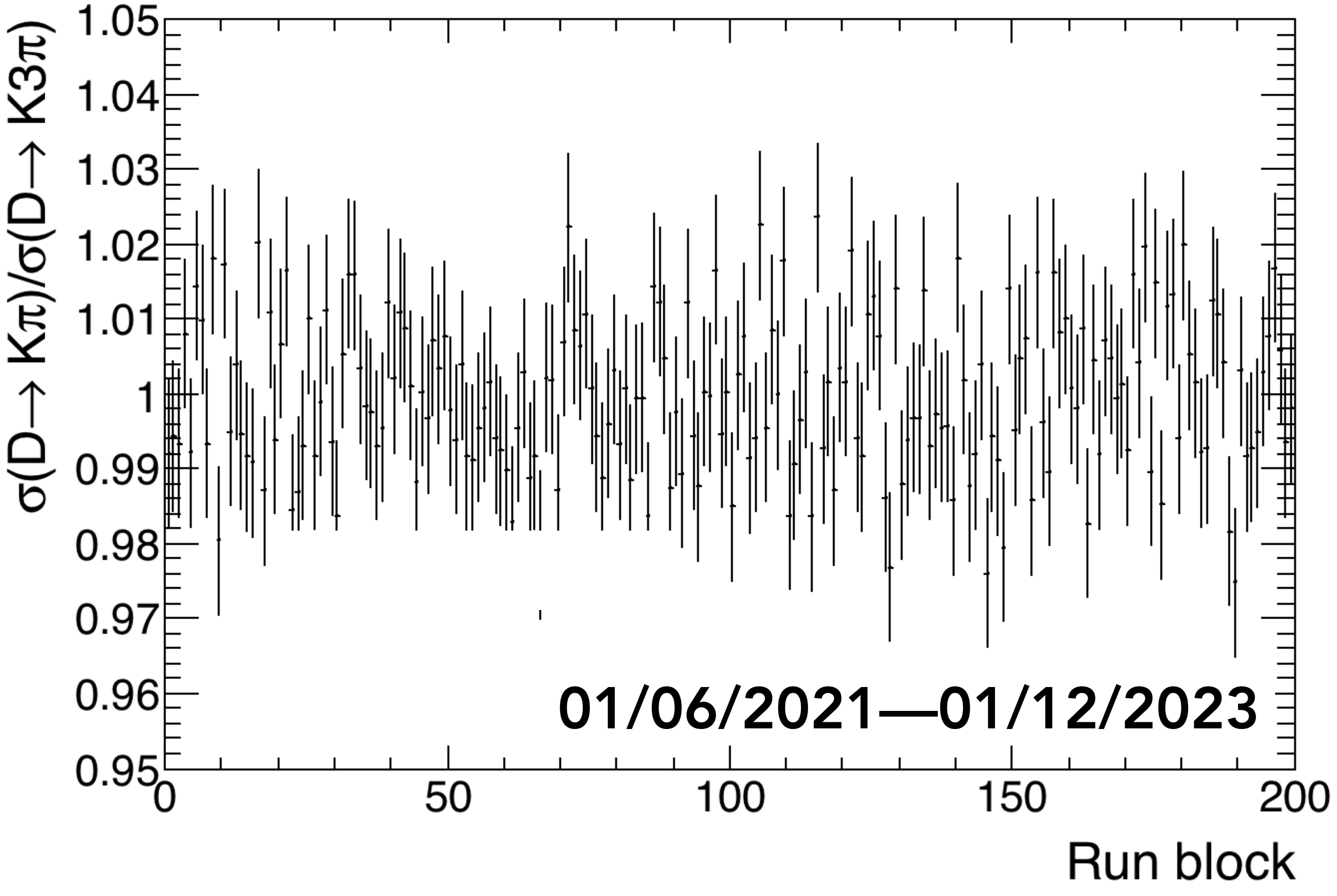


GEANT is not the best simulation by accident : tuned using decades of fixed-target data
Fast simulations can surely get the core of the distributions right, but tails&correlations?
Fast simulations rely on precise data-driven measurements of detector performance!

Towards self-calibrating data





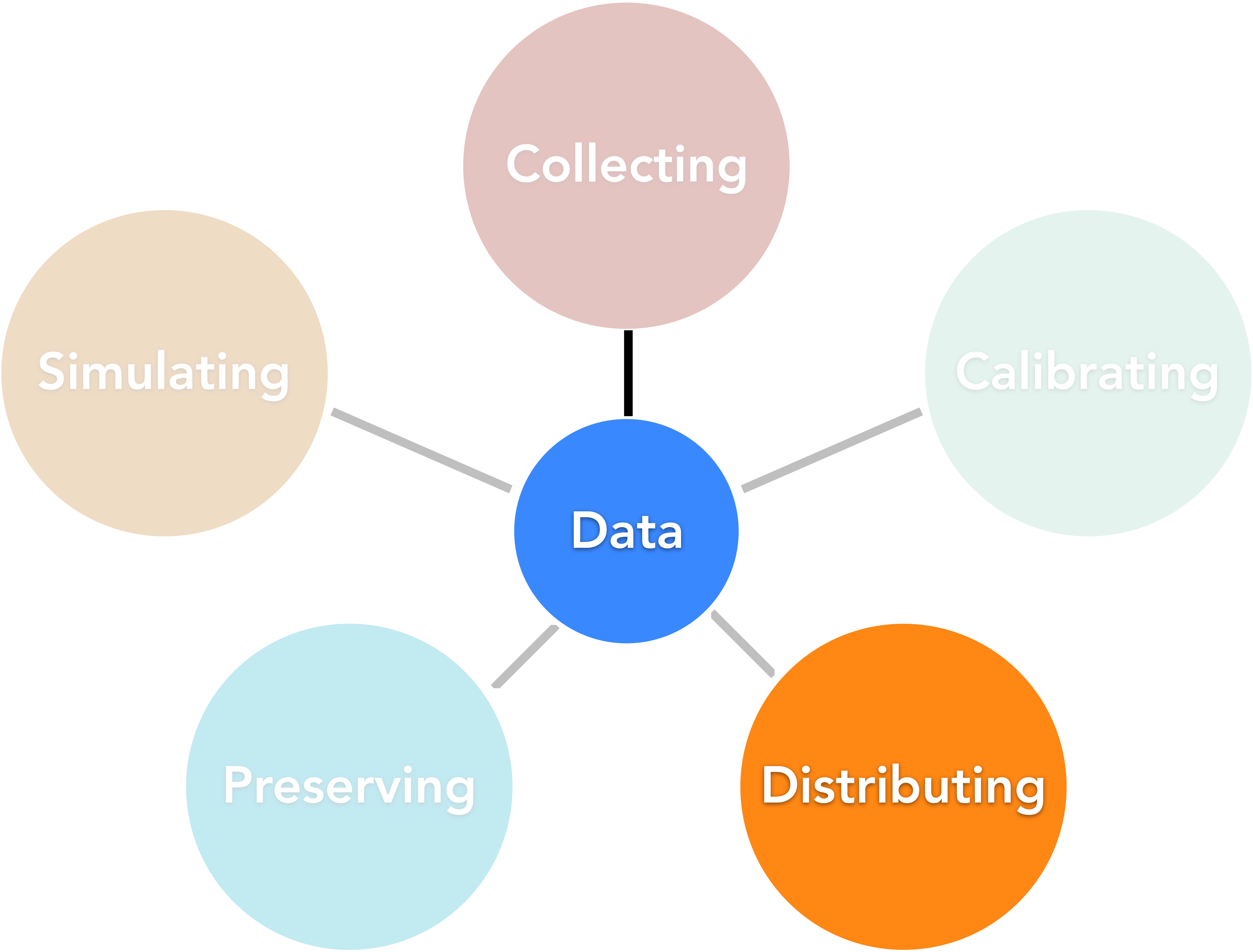


For ultimate precision LHC detectors will analyze data taken three decades apart : a challenge in itself

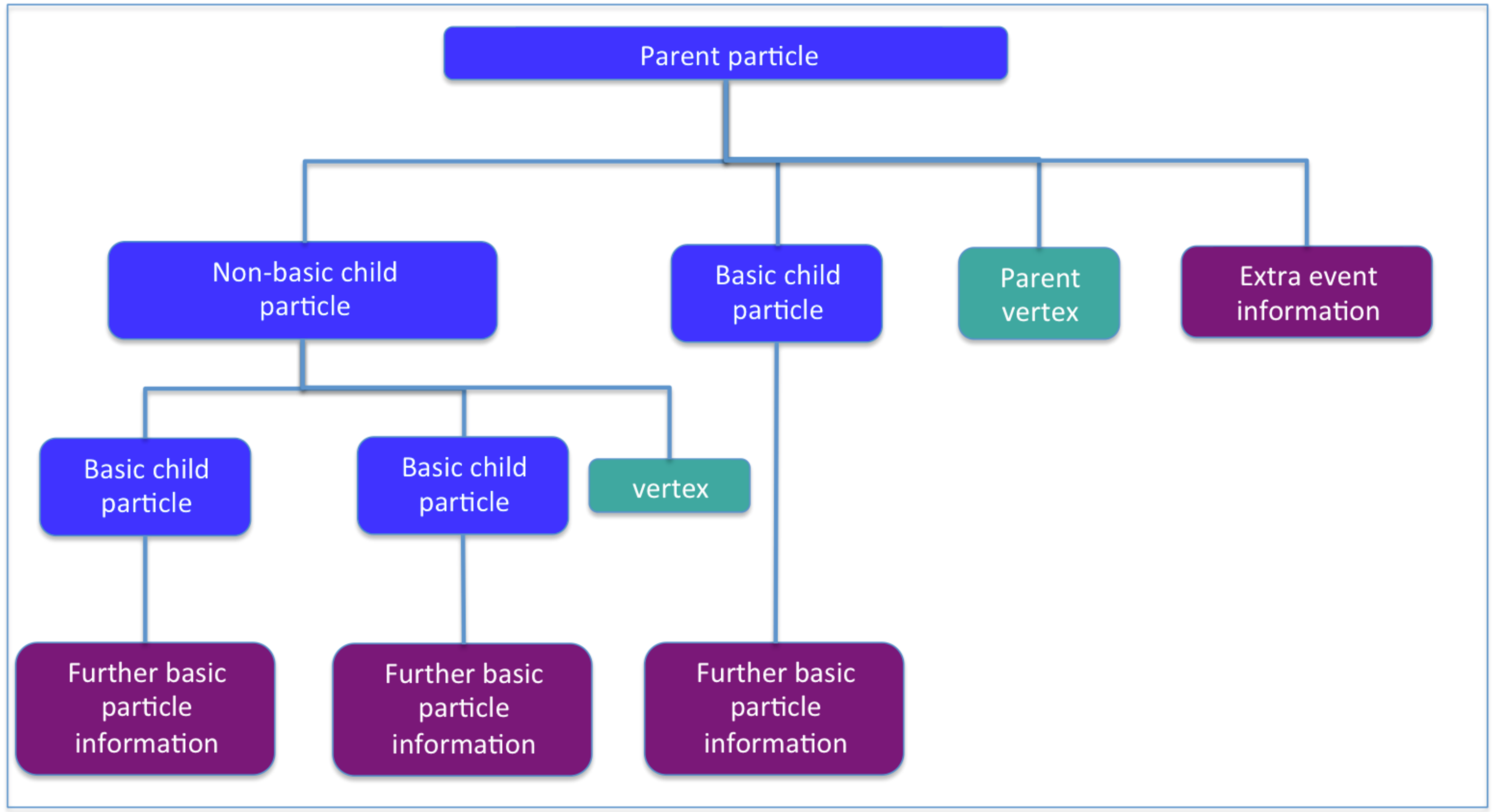
Rough detector monitoring and data quality not sufficient

Measure standard candles in real-time, automatically tuning fast/full simulation with the output?

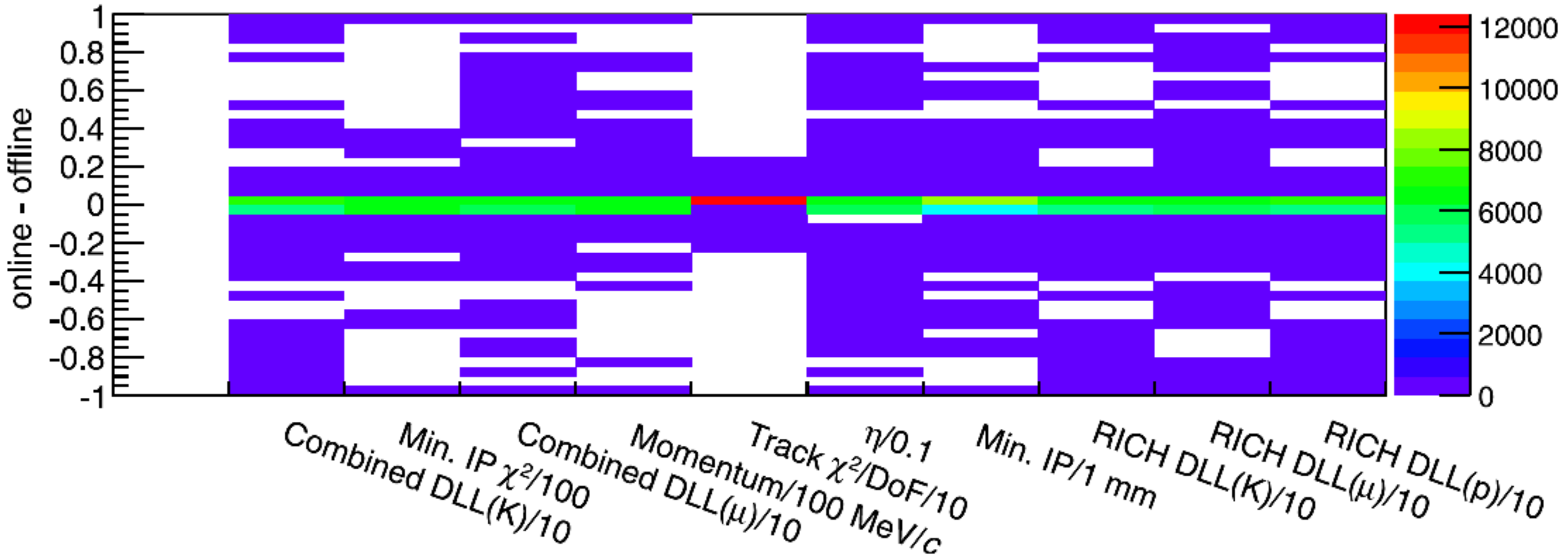
Very helpful if Belle II could measure some absolute branching fractions below the permille level!


Collecting
Simulating
Calibrating
Data
Preserving
Distributing

Before distributing we must compress



Same data, different compression



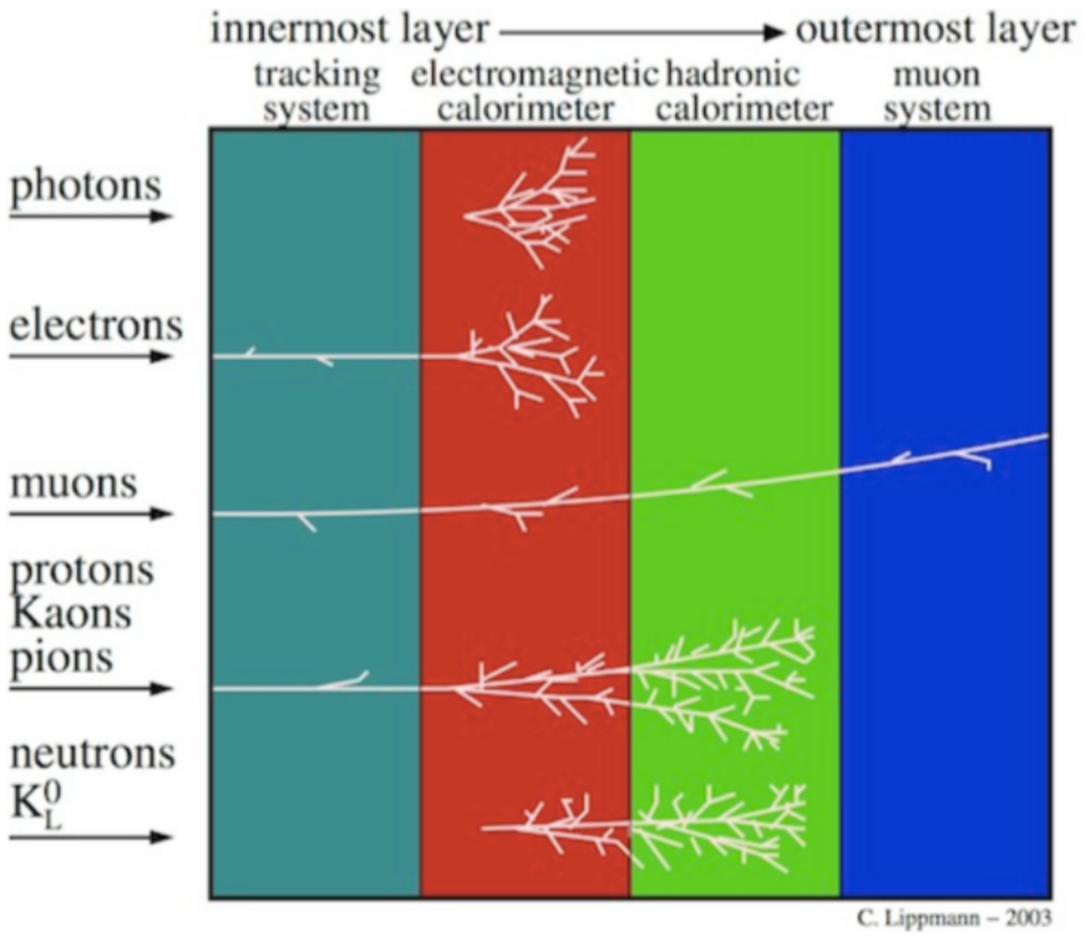
Over the past ten years, have continuously increased levels of data compression

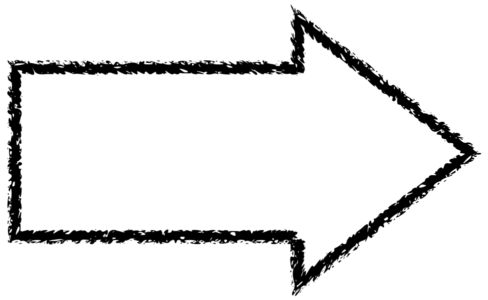
Compression happens at two levels : selectively persist only information really needed for an analysis (for example about the signal candidate) and pack this information into the minimum possible space.

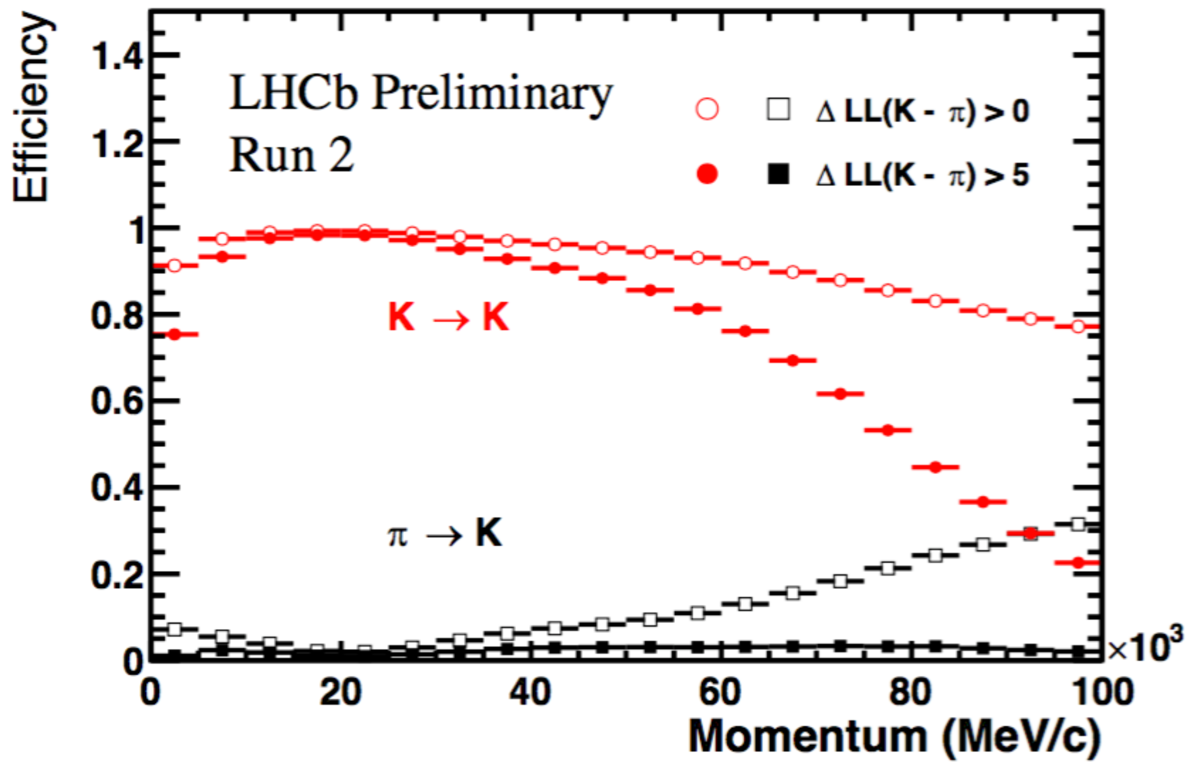
Changing compression can lead to visible changes in the data, must be carefully maintained!

Finding the optimal representation

More details







Example : LHCb particle identification relies on tens of variables per reconstructed track

If we cannot save all reconstructed data for the signal particles, need to compress it!
What is the minimal representation of the data which preserves sensitivity to the physics observables and allows unforeseen systematics to be controlled?
Is representation analysis dependent or can we make generally useful compressions?

Streaming & data access

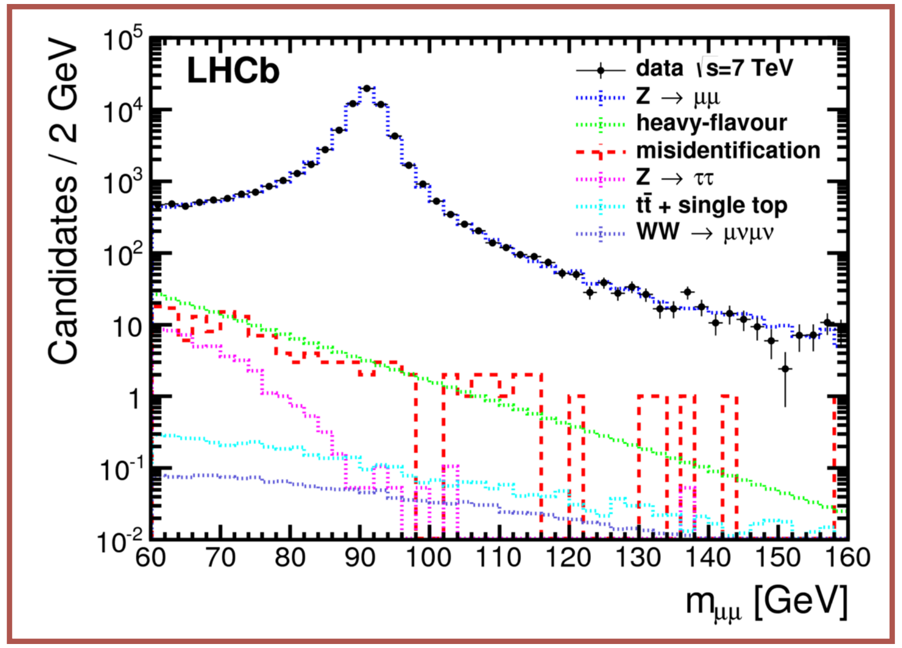
Analysis streams

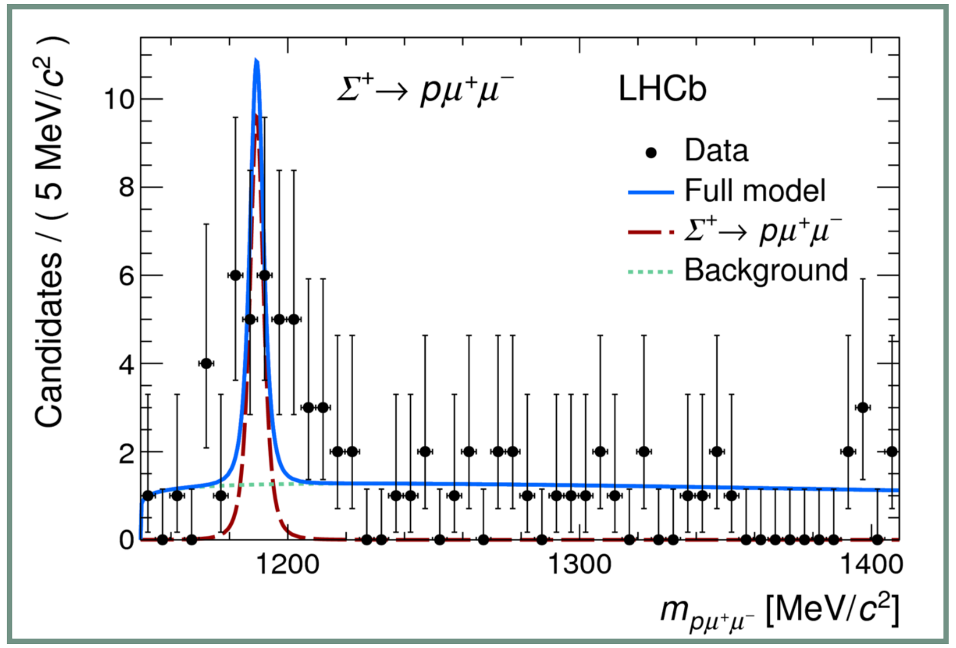
...

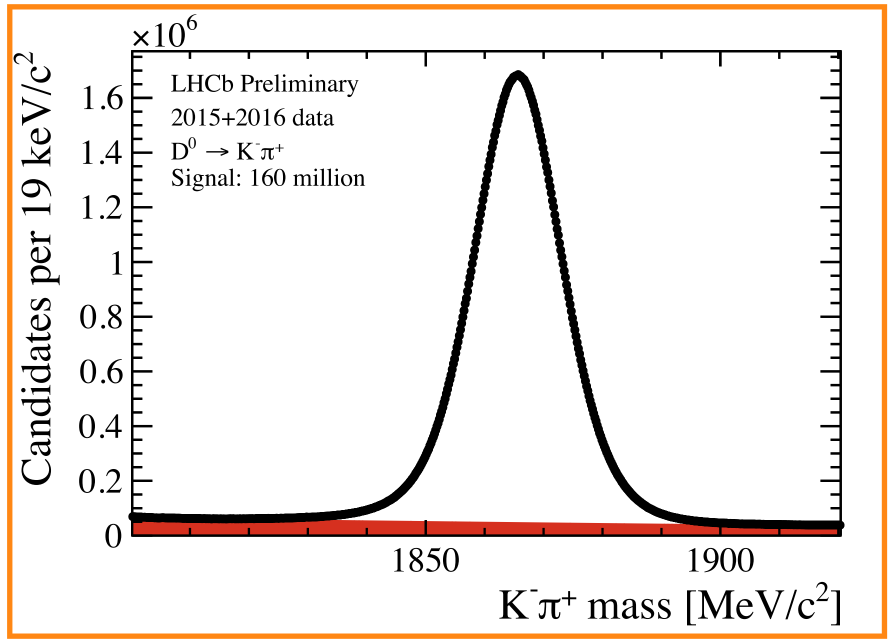
EW

Rare decays

Charm

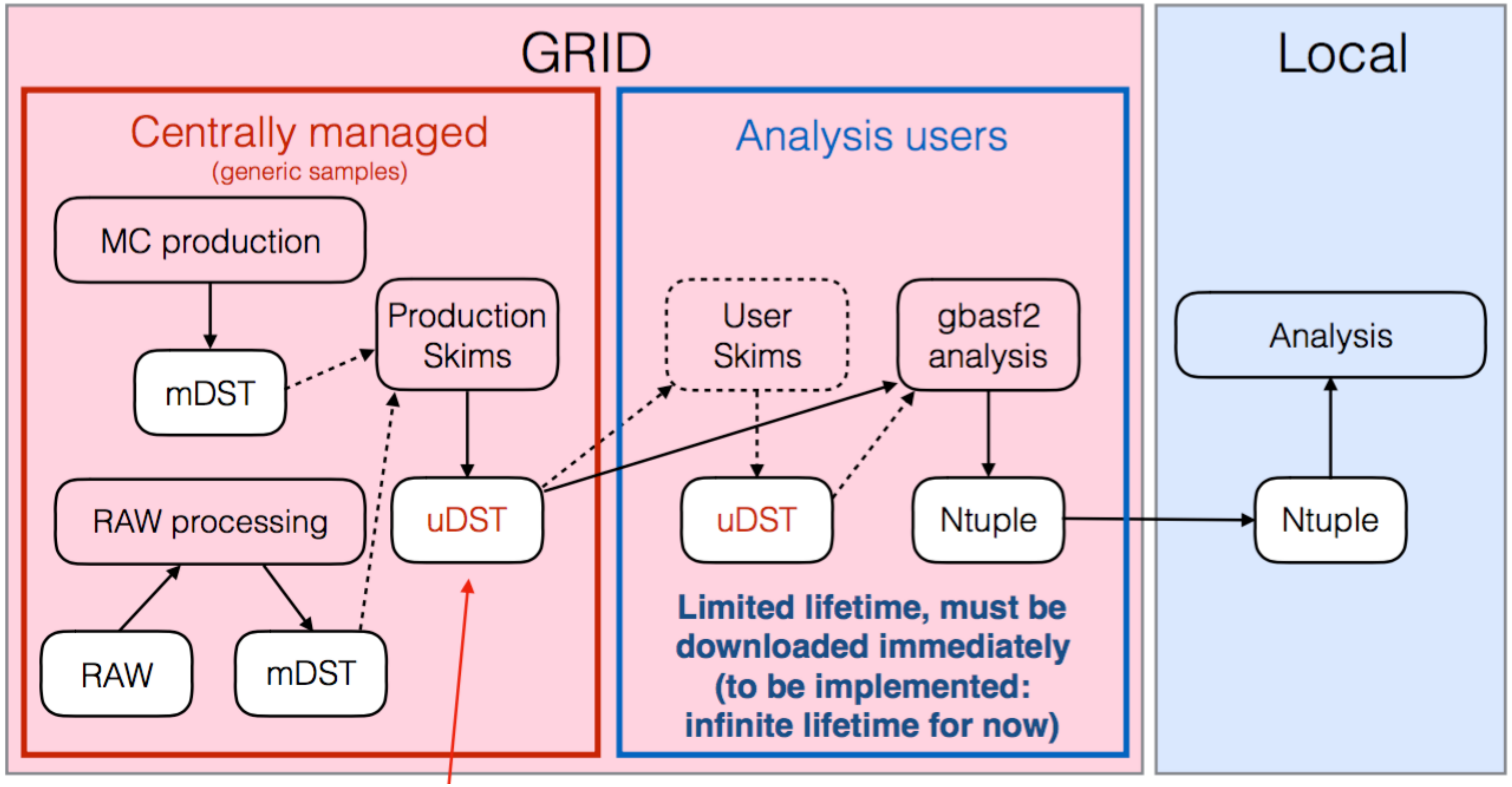






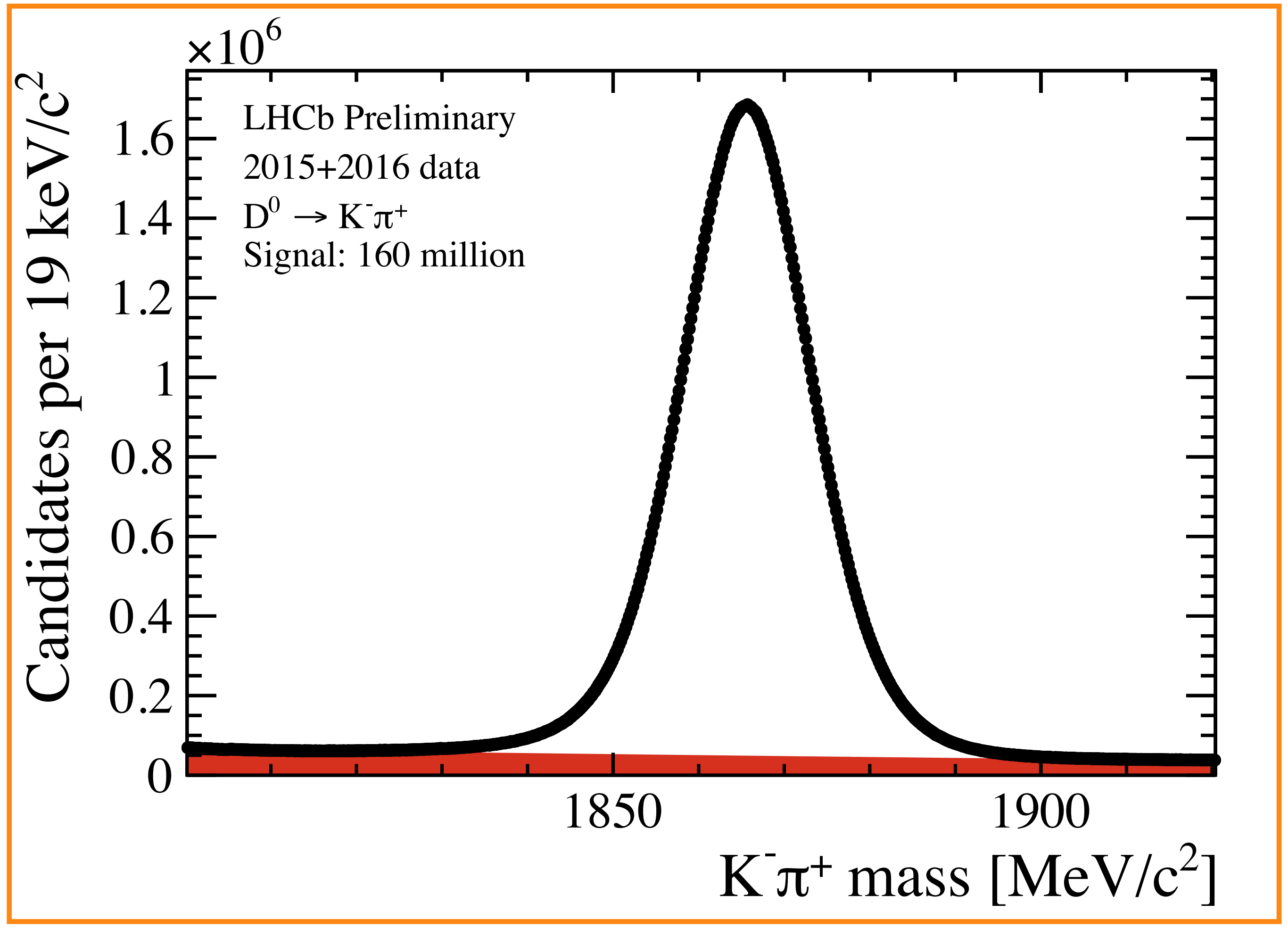
Even after compression and real-time analysis, implausible for each analyst to access all the data
Segregate high yield modes from rarer, easier to process signals, but must also group all control modes in the same stream as the signal → tough optimization problem with O(1000) signal selections

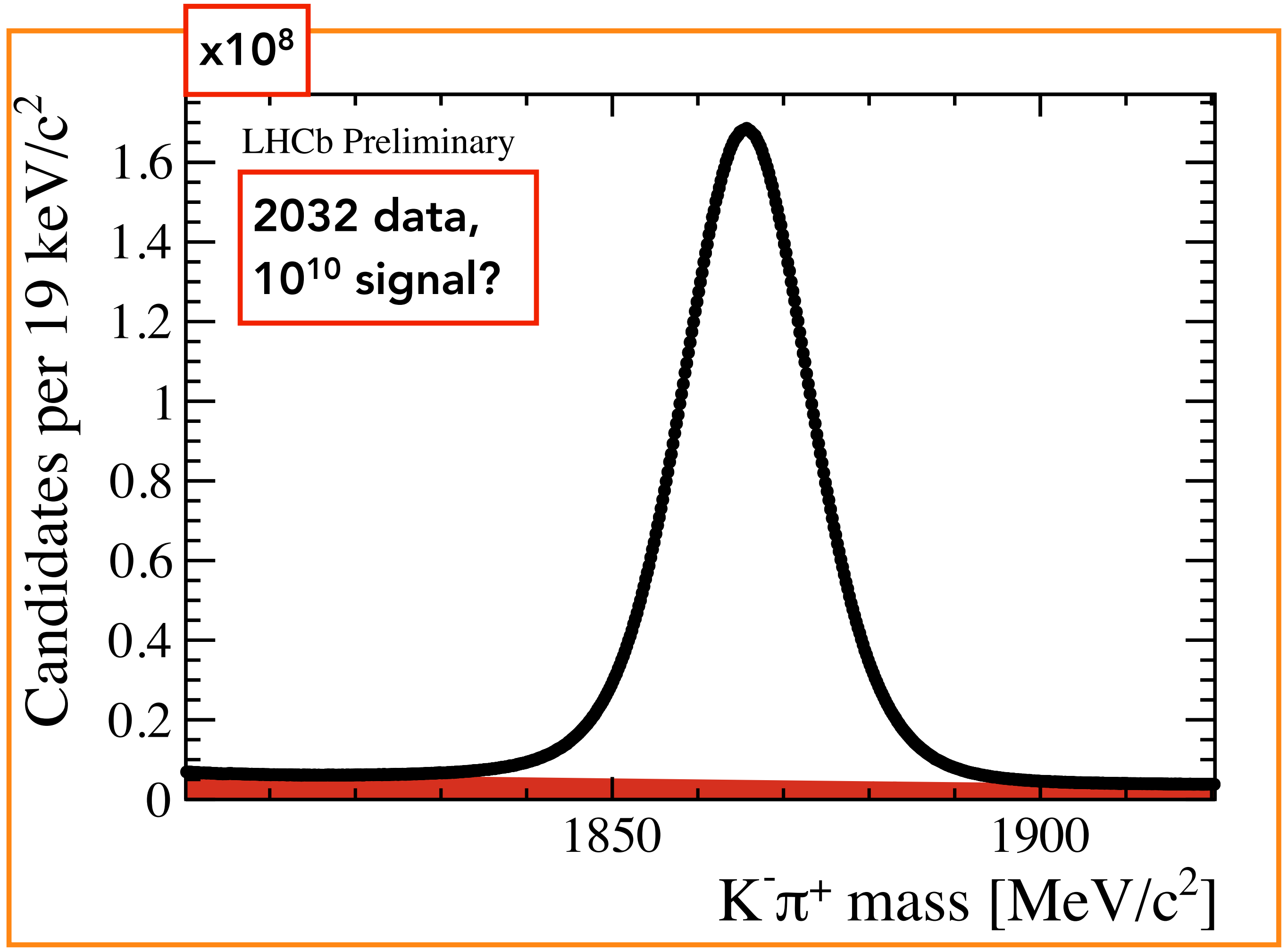
Belle II skimming strategy



The Belle II data analysis challenge is of a similar scale to the other ultimate precision experiments
Similar solutions being pursued in respect of “skimming”/streaming the data to the users for analysis

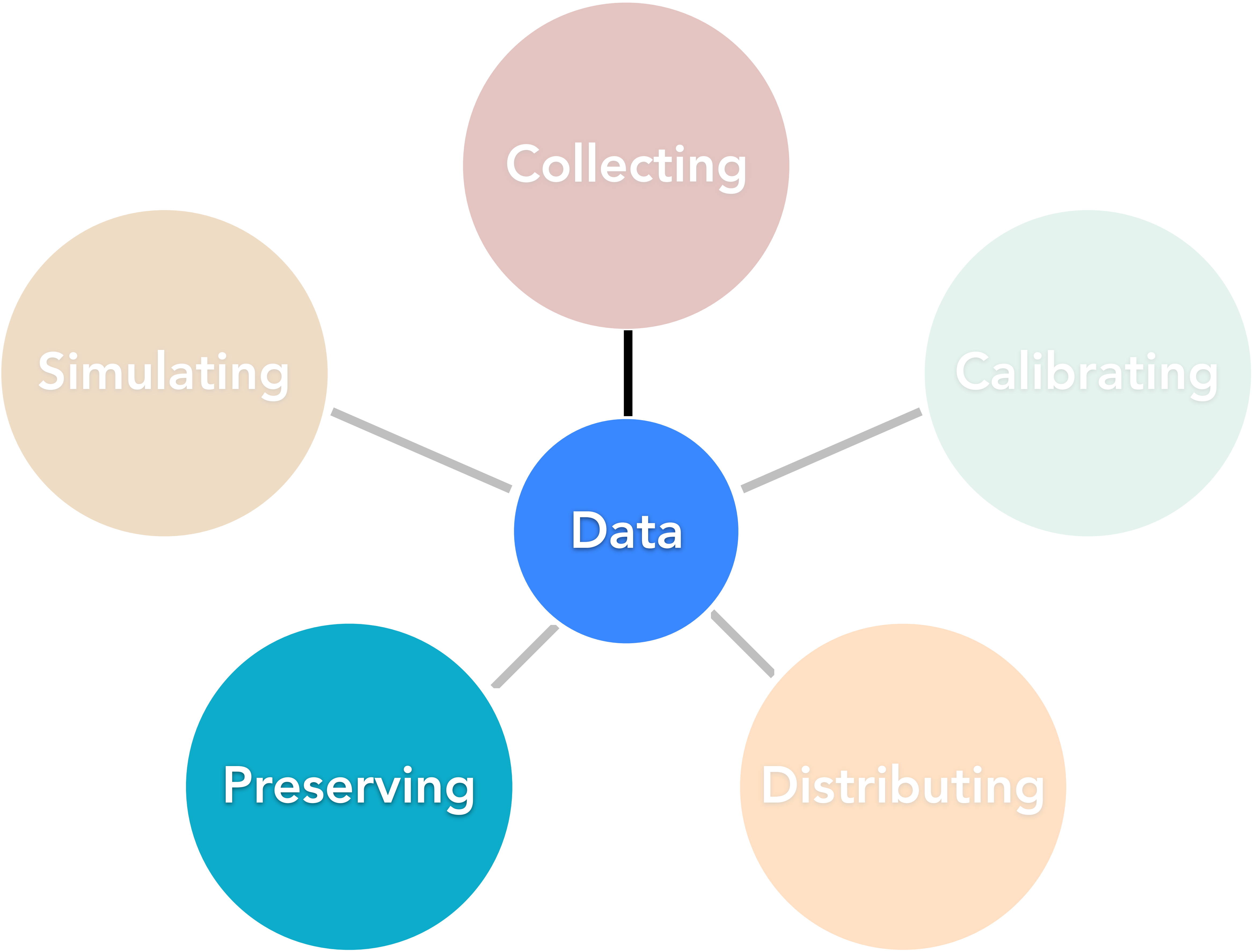
Analysis trains or analysis factories?





Over ~10^7 signals/mode/year, maybe not plausible for users to run directly over even streamed data. Prepare final analysis steps on small data subsets, then schedule rest for centralised production. Requires a huge cultural, not only technical shift, in particular with respect to PhD thesis supervision.

And for ultimate precision, will need the ultimate data-driven measurements of detector performance using tag-and-probe and standard candle measurements, and then this 10^{10} is not extreme!


Collecting
Simulating
Calibrating
Data
Preserving
Distributing

Analyzing decades old data

Old data analysis checklist

Can I read it with the new software?

If not, will old SW compile on the new machines?

Will floating point operations give same answer on the new machines?

If not and I have to emulate, will this be fast enough?

If I run into problems, do people still remember what the centralized or real-time processing did 10 years ago so I can understand it?

Can I generate simulations for old data with the new Pythia tunings?

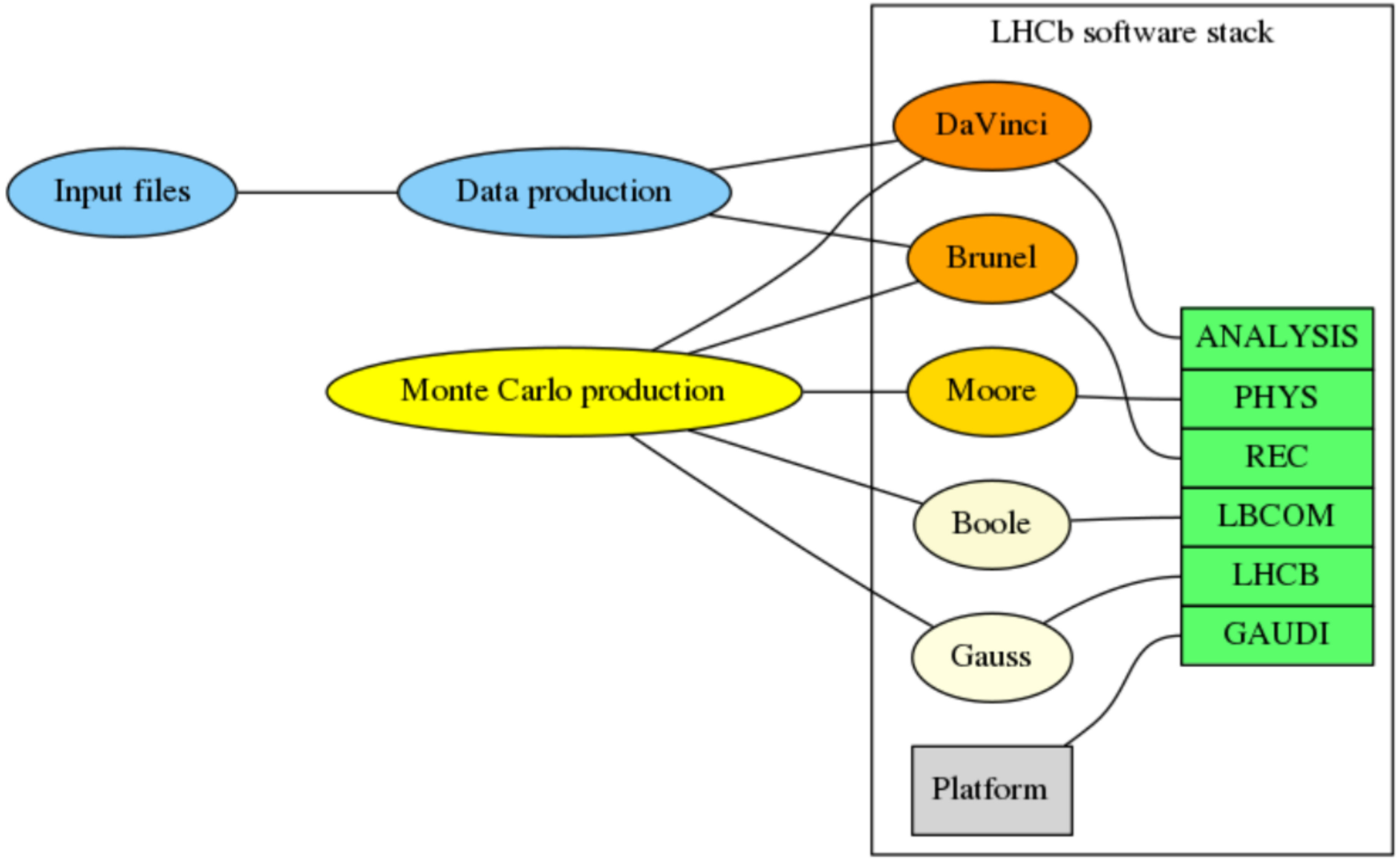
…

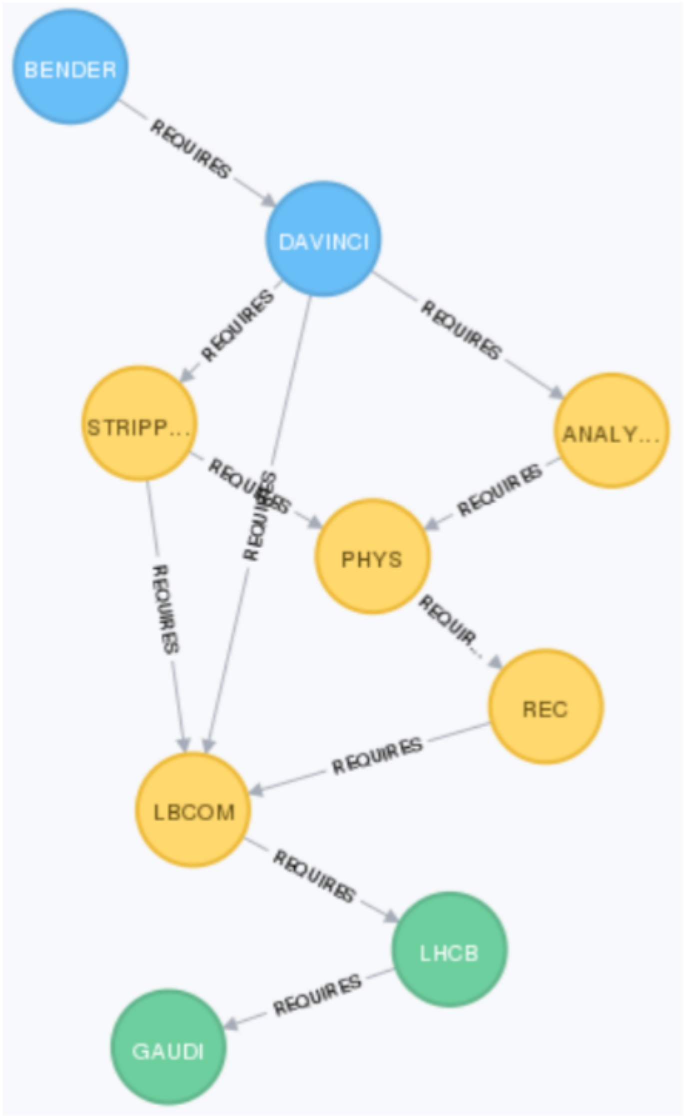
…

Simply being able to reliably keep analyzing this data is a major challenge!
Compatibility between old and new datasets will be particularly psychologically important if we start seeing BSM signals. Requires major efforts on data&analysis preservation.

Encapsulating the workflows

More details

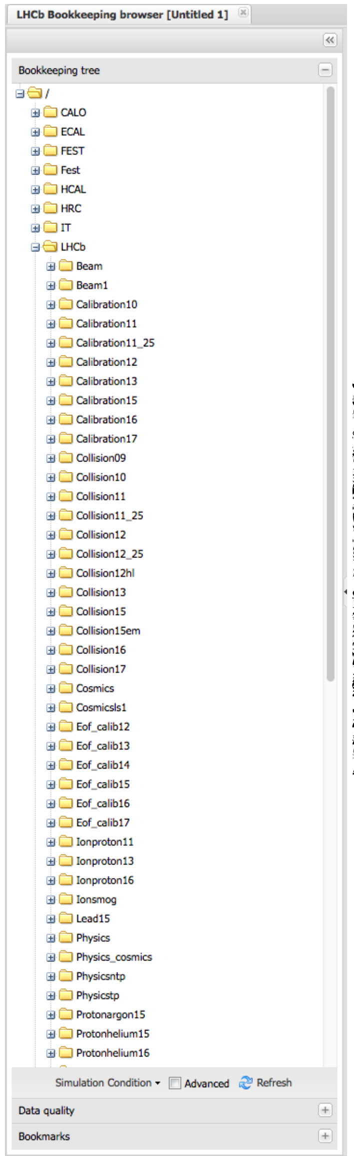




Vital to automatically encapsulate old workflows and watermark data provenance

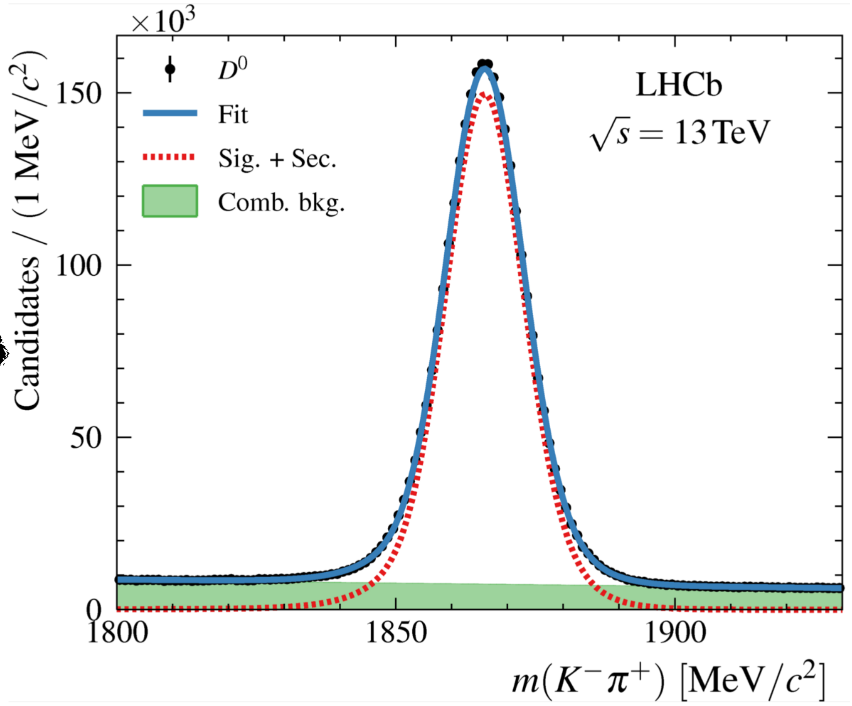
Code preservation and documentation

Streamed data



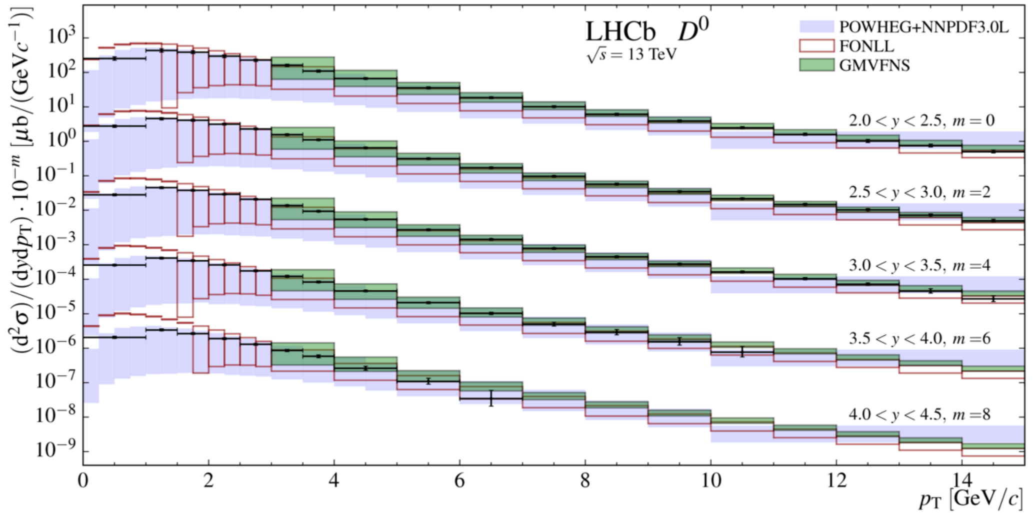
Make ntuples
Train selections
Build PDFs and fit
…

Intermediate analysis data



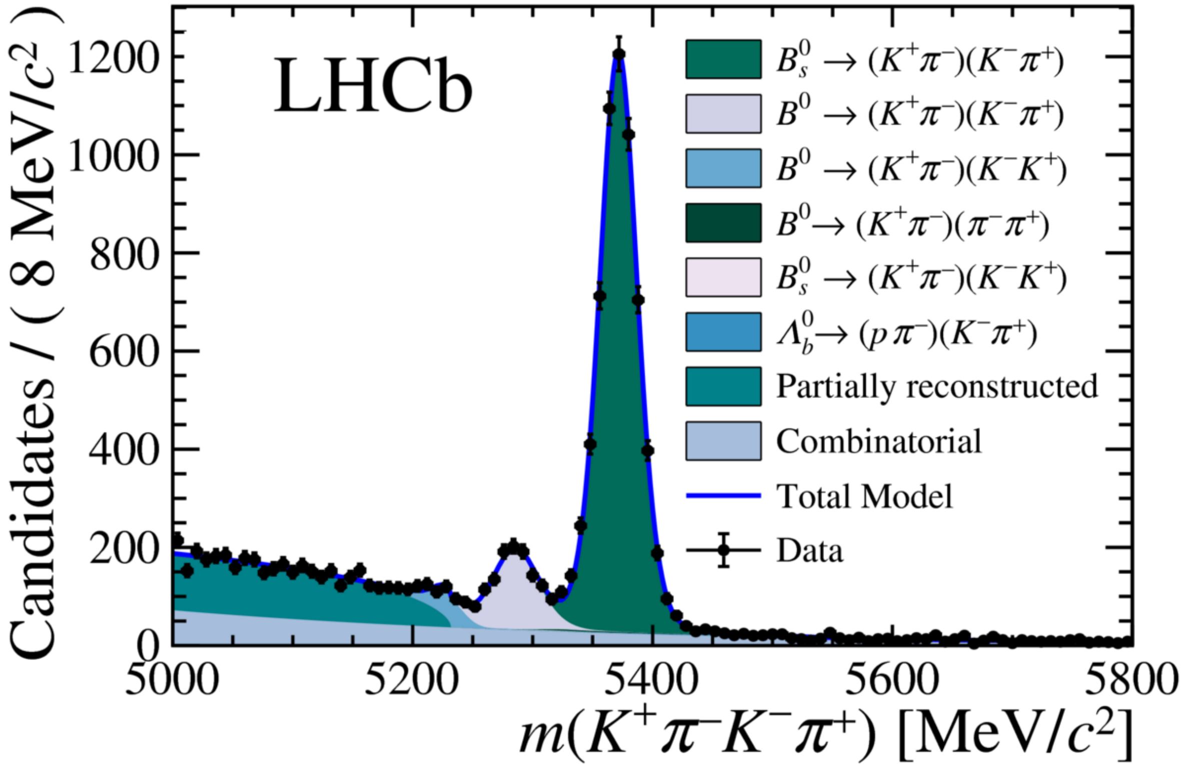
Apply efficiency corrections
Run toys
Evaluate systematics
…

Final physics output



Equally vital to preserve analyses and the ability to rerun the full analysis chain without intervention
Complicated by the fact that rerunning e.g. the ntuple making may not be realistic in analyses with very high signal yields → must combine data & analysis preservation!
Some test cases already exist in LHCb, but a long way from a generally applicable framework.

Large data, common data, open data

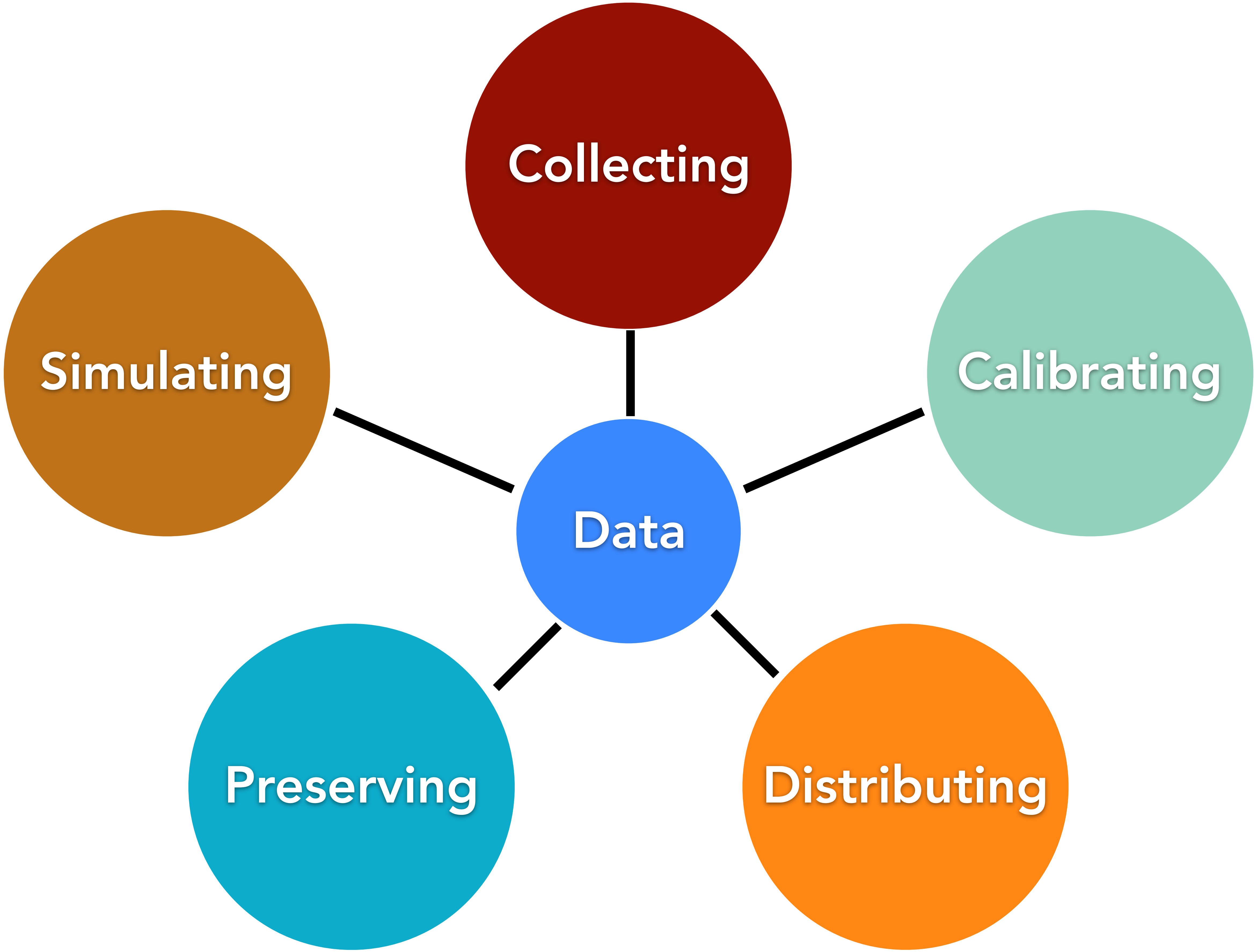


Parameter	$\phi_s^{d\bar{d}}$ [rad]	$\lvert\lambda\rvert$	f^{VV}	f_L^{VV}	$f_\parallel^{VV}$	$\delta_\parallel^{VV}$	$\delta_\perp^{VV}$	f^{SV}	f^{VS}	δ^{SV}	δ^{VS}	f^{SS}	δ^{SS}
Yield and shape of mass model	0.012	0.001	0.001	0.004	0.004	0.011	0.020	0.002	0.003	0.023	0.023	0.004	0.012
Signal weights of mass model	0.012	0.007	0.002	0.006	0.005	0.024	0.112	0.004	0.005	0.049	0.022	0.005	0.047
Decay-time-dependent fit procedure	0.006	0.002	0.001	0.006	0.002	0.007	0.017	0.003	0.002	0.007	0.027	0.001	0.009
Decay-time-dependent fit parameterisation	0.049	0.013	0.021	0.025	0.026	0.187	0.202	0.042	0.029	0.159	0.234	0.064	0.227
Acceptance weights (simulated sample size)	0.106	0.078	0.004	0.031	0.029	0.236	0.564	0.037	0.039	0.250	0.290	0.015	0.256
Other acceptance and resolution effects	0.063	0.008	0.005	0.018	0.005	0.136	0.149	0.006	0.004	0.167	0.124	0.017	0.194
Production asymmetry	0.002	0.002	0.000	0.000	0.000	0.001	0.017	0.002	0.002	0.002	0.008	0.000	0.002
Total	0.141	0.089	0.024	0.046	0.042	0.333	0.641	0.071	0.065	0.346	0.405	0.069	0.399

Parameter	f^{ST}	f^{TS}	δ^{ST}	δ^{TS}	f^{VT}	f_L^{VT}	$f_\parallel^{VT}$	f^{TV}	f_L^{TV}	$f_\parallel^{TV}$	δ_0^{VT}	$\delta_\parallel^{VT}$	$\delta_\perp^{VT}$	δ_0^{TV}	$\delta_\parallel^{TV}$	$\delta_\perp^{TV}$
Yield and shape of mass model	0.002	0.004	0.111	0.023	0.001	0.003	0.001	0.001	0.043	0.025	0.023	0.055	0.110	0.053	0.018	0.065
Signal weights of mass model	0.004	0.006	0.151	0.105	0.002	0.003	0.001	0.001	0.043	0.029	0.025	0.131	0.126	0.080	0.073	0.150
Decay-time-dependent fit procedure	0.001	0.002	0.248	0.017	0.002	0.004	0.002	0.002	0.008	0.005	0.012	0.069	0.025	0.062	0.017	0.030
Decay-time-dependent fit parameterisation	0.006	0.017	0.736	0.247	0.011	0.053	0.019	0.008	0.080	0.048	0.286	0.308	0.260	0.260	0.228	0.405
Acceptance weights (simulated sample size)	0.014	0.015	1.463	0.719	0.026	0.145	0.054	0.027	0.199	0.102	1.117	1.080	0.888	0.712	0.417	0.947
Other acceptance and resolution effects	0.002	0.003	0.184	0.226	0.015	0.024	0.004	0.005	0.045	0.017	0.163	0.168	0.191	0.229	0.246	0.171
Production asymmetry	0.001	0.001	0.037	0.026	0.001	0.003	0.001	0.002	0.012	0.006	0.015	0.030	0.018	0.003	0.007	0.041
Total	0.031	0.033	1.688	0.817	0.049	0.165	0.063	0.048	0.252	0.143	1.171	1.159	0.970	0.802	0.546	1.076

Parameter	f^{TT}	f_L^{TT}	$f_{\parallel_1}^{TT}$	$f_{\perp_1}^{TT}$	$f_{\parallel_2}^{TT}$	δ_0^{TT}	$\delta_{\parallel_1}^{TT}$	$\delta_{\perp_1}^{TT}$	$\delta_{\parallel_2}^{TT}$	$\delta_{\perp_2}^{TT}$
Yield and shape of mass model	0.000	0.045	0.019	0.037	0.002	0.038	0.027	0.009	0.079	0.114
Signal weights of mass model	0.000	0.066	0.025	0.024	0.002	0.147	0.046	0.112	0.123	0.215
Decay-time-dependent fit procedure	0.001	0.022	0.022	0.014	0.004	0.127	0.036	0.068	0.058	0.040
Decay-time-dependent fit parameterisation	0.005	0.051	0.071	0.113	0.038	1.213	0.199	0.685	0.820	0.476
Acceptance weights (simulated sample size)	0.003	0.135	0.110	0.127	0.077	1.328	0.454	1.348	1.443	1.161
Other acceptance and resolution effects	0.002	0.031	0.028	0.056	0.024	0.226	0.275	0.156	0.343	0.301
Production asymmetry	0.000	0.002	0.001	0.008	0.003	0.005	0.002	0.062	0.015	0.043
Total	0.007	0.176	0.142	0.205	0.107	1.825	0.573	1.546	1.706	1.330

Ultimate precision analyses will increasingly involve measuring multiple correlated observables
As we get closer to the systematics limit, systematics correlations will become more important
Some form of open data probably necessary to enable combination of knowledge between collaborations, at least at the level of publishing fit results or likelihoods as ROOT files.


Collecting
Simulating
Calibrating
Data
Preserving
Distributing

We are all software developers now…

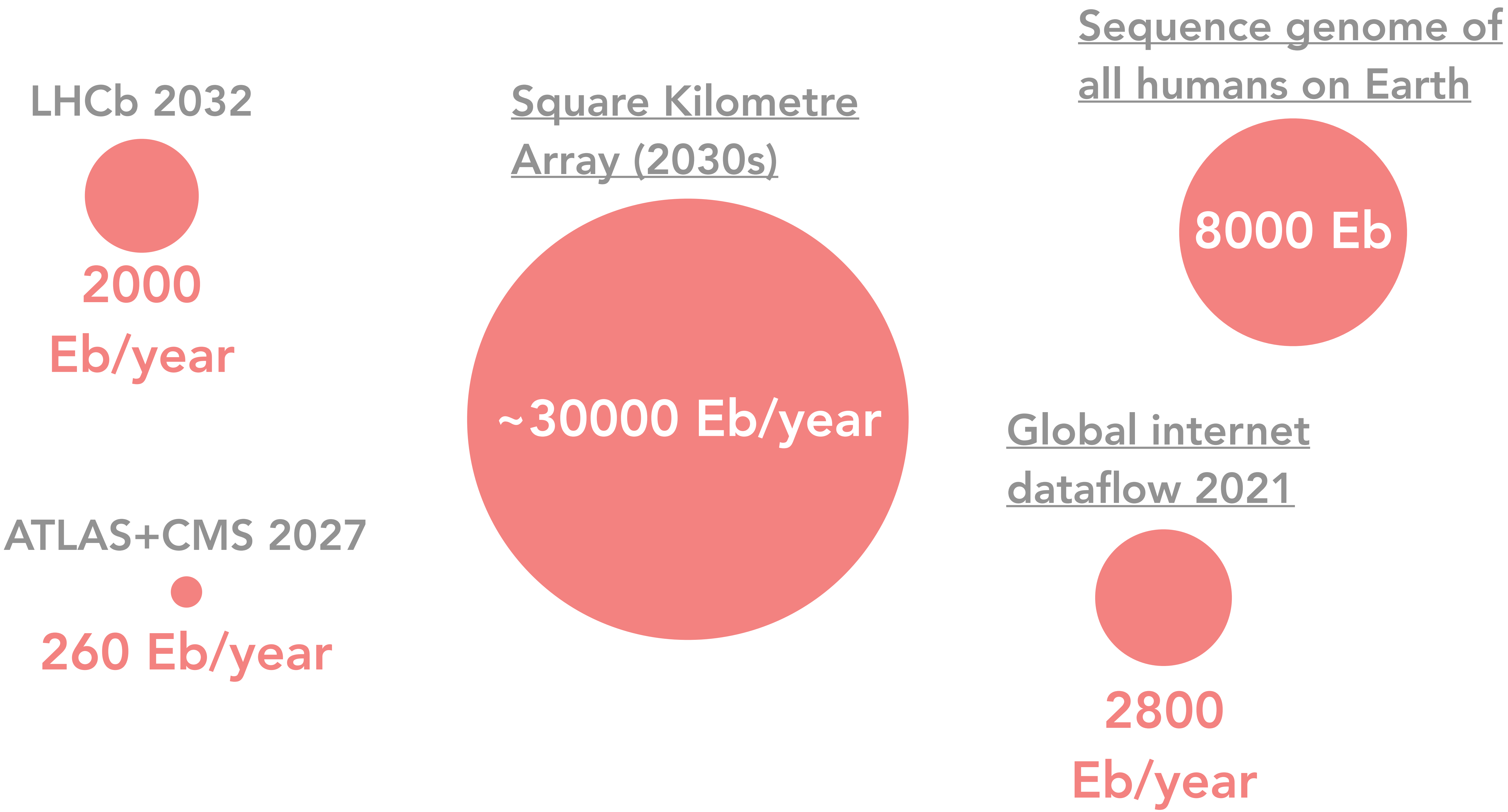
If we want to solve the data challenges which lie on the road towards ultimate precision, we will need to get comfortable with our tools

This does not mean blindly following what is done in industry, any more than working on silicon detectors or calorimeters would

It does mean making it possible to provide long-term support for physicists whose core “technical” skill is software in the same way as exists for physicists who work on hardware detectors

Critical to understand that our financial and resource constraints are orders of magnitude from what private sector would require to do the same job. So will always need custom hybrid processing architectures tuned to our specific problems, and the physicists to build&tune them.

…lucky to have big problems to solve!



More data

Disclaimers

User-level analysis has been completely ignored in this talk; it could deserve a talk on its own, particularly e.g. advanced GPU fitting methods for complex analysis like amplitude fits.

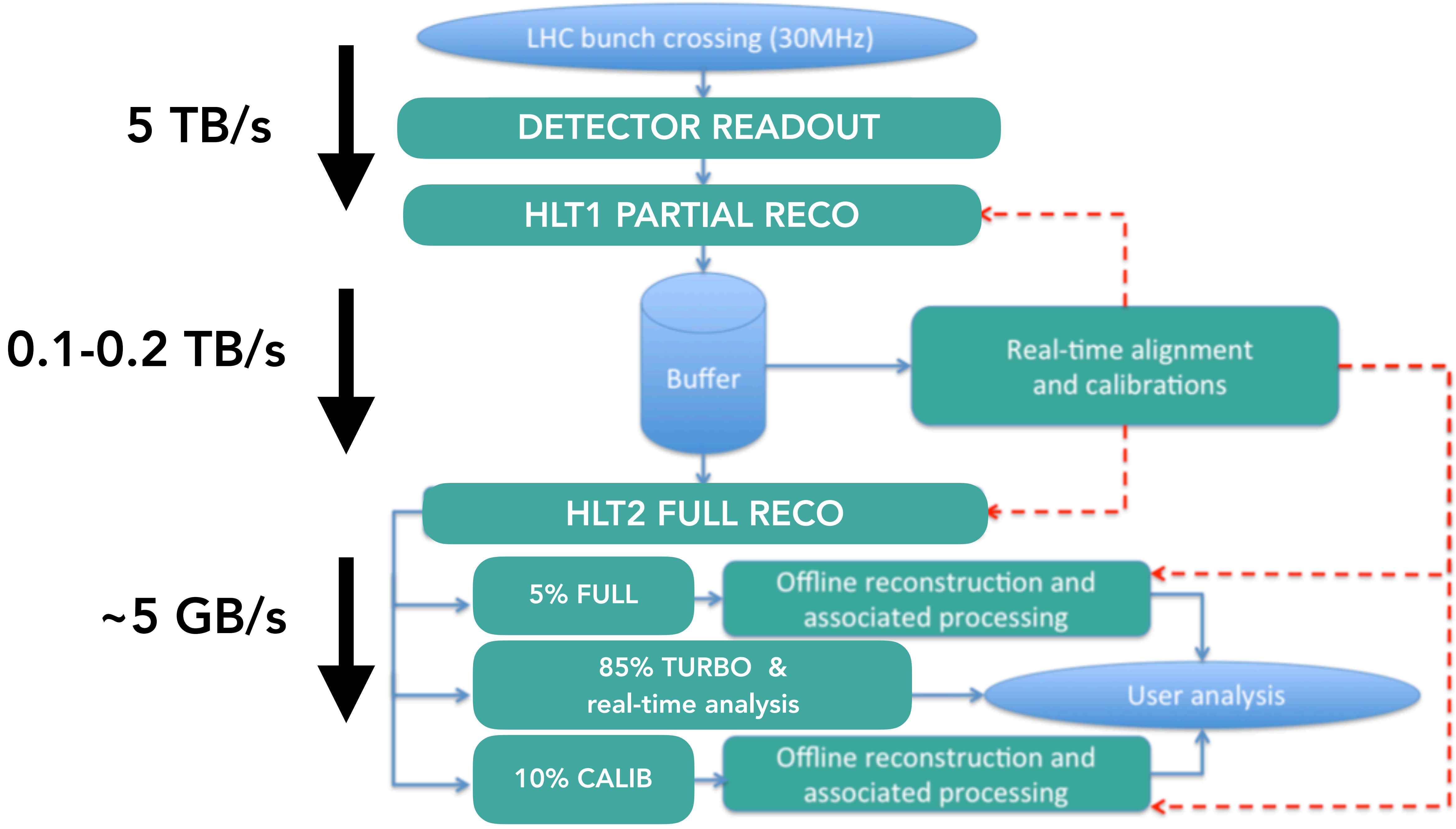
The data rates given on slide 2 are for what comes off the detector front-end electronics, after zero-suppression but before any processing. This is contrary to the usual HEP approach of quoting only the data persisted offline, but is coherent with the point that the biggest challenge(s) relate to real-time data processing and analysis.

Upgrade I data processing chain

Need to use full tracker information in order to preselect interesting events $\Longrightarrow$ triggerless readout.

HLT2 roughly 10x slower than HLT1 per event in Run I/II $\Longrightarrow$ processing model relies on a 25-50 reduction in the event rate at the HLT1 stage.

Almost all analyses will go to TURBO, but maintain raw data for some calibrations and for a limited number of well-motivated analysis use cases (e.g. EW). Real-time analysis & selective persistence critical to maintain the broadest possible physics programme.



GPD vs. LHCb UII data rate comparison

CMS detector	LHC Run-2	HL-LHC Phase-2		LHCb Upgrade II
Peak ⟨PU⟩	60	140	200	
L1 accept rate (maximum)	100 kHz	500 kHz	750 kHz	30 MHz
Event Size	2.0 MB [a]	5.7 MB [b]	7.4 MB	1.5 MB ?
Event Network throughput	1.6 Tb/s	23 Tb/s	44 Tb/s	~500 Tb/s
Event Network buffer (60 seconds)	12 TB	171 TB	333 TB	??
HLT accept rate	1 kHz	5 kHz	7.5 kHz	??
HLT computing power [c]	0.5 MHS06	4.5 MHS06	9.2 MHS06	??
Storage throughput	2.5 GB/s	31 GB/s	61 GB/s	50 GB/s ?
Storage capacity needed (1 day)	0.2 PB	2.7 PB	5.3 PB	??

ATLAS globally similar but TDR is still under review so numbers not public

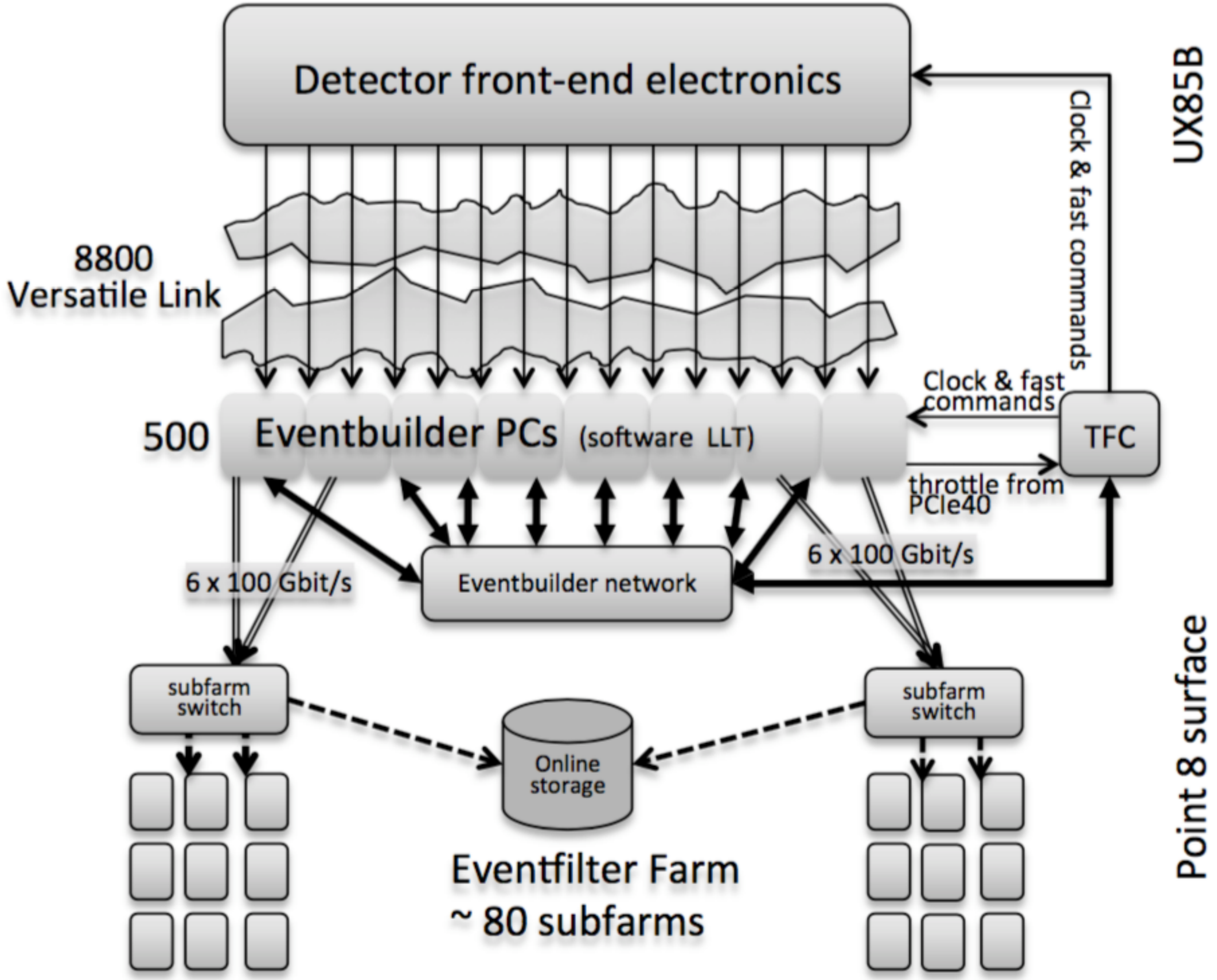
Upgrade II DAQ must process 10x the HL-LHC GPD data rate
Upgrade II Offline must process same data volume as GPDs

Challenges & evolution of LHCb DAQ

GBT link : 4.8 Gb/s Upgrade I
Assume evolution to 10 Gb/s for HL-LHC using aggressive error handling : missing factor 5 compared to data rate growth.

Event-building : current network is 500 servers with 100 Gb/s links. 200 Gb/s readily available, keep an eye on price/ performance scaling beyond this?

Farm : carry out R&D in next years on optimal use of hybrid architectures (GPU/ CPU/FPGA), remain flexible



Partnership with industry seems key to ensuring cost-effective solutions to DAQ & processing