Nice detector you got there, whatcha gonna do with it?
(Towards an Upgrade II data processing model)

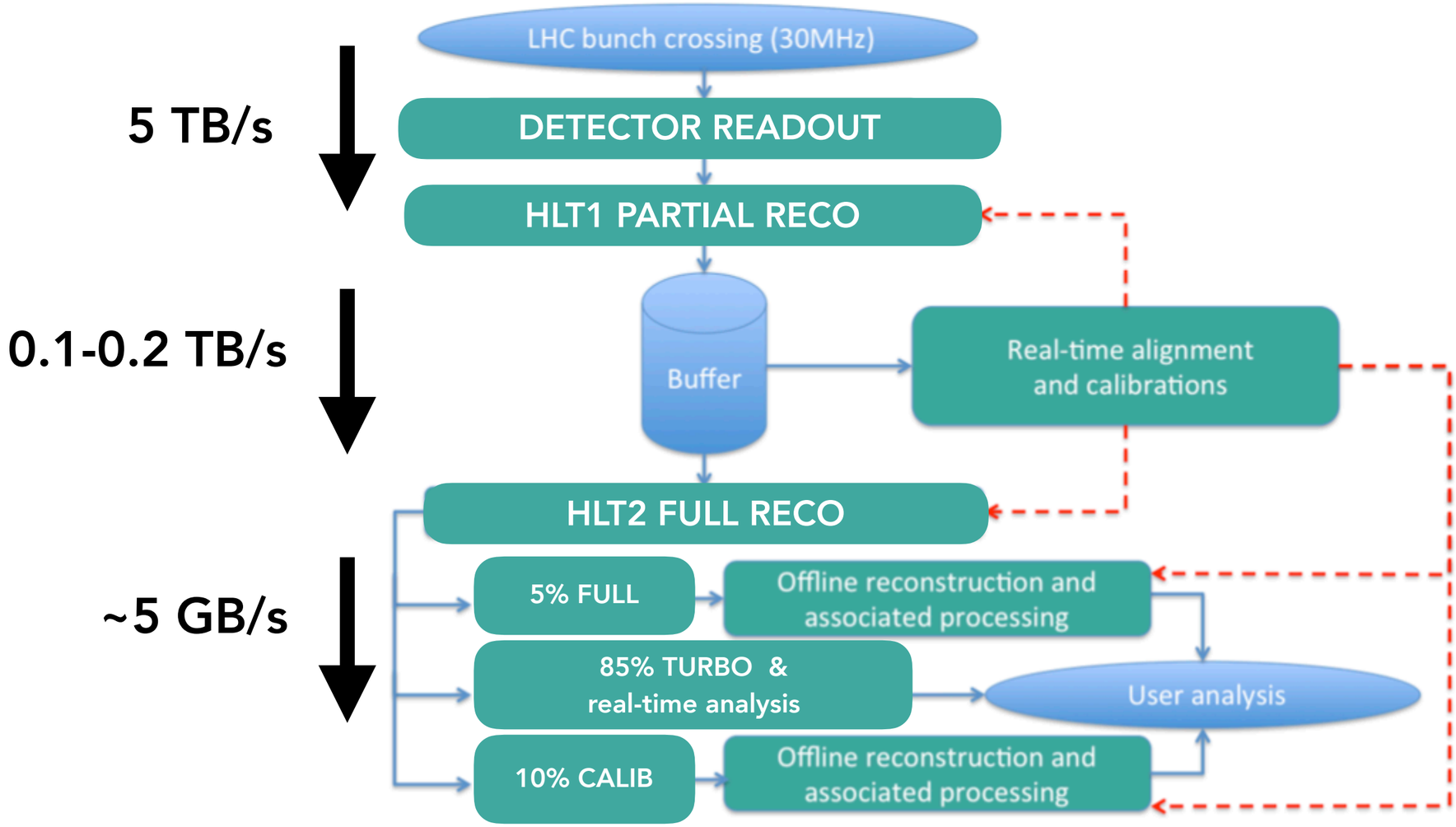
Vladimir V. Gligorov, CNRS/LPNHE

With input from Renaud, Conor, Niko, Marco Cattaneo, Stefan, Concezio, Renato

Upgrade II Workshop, Annecy, 21.03.2018

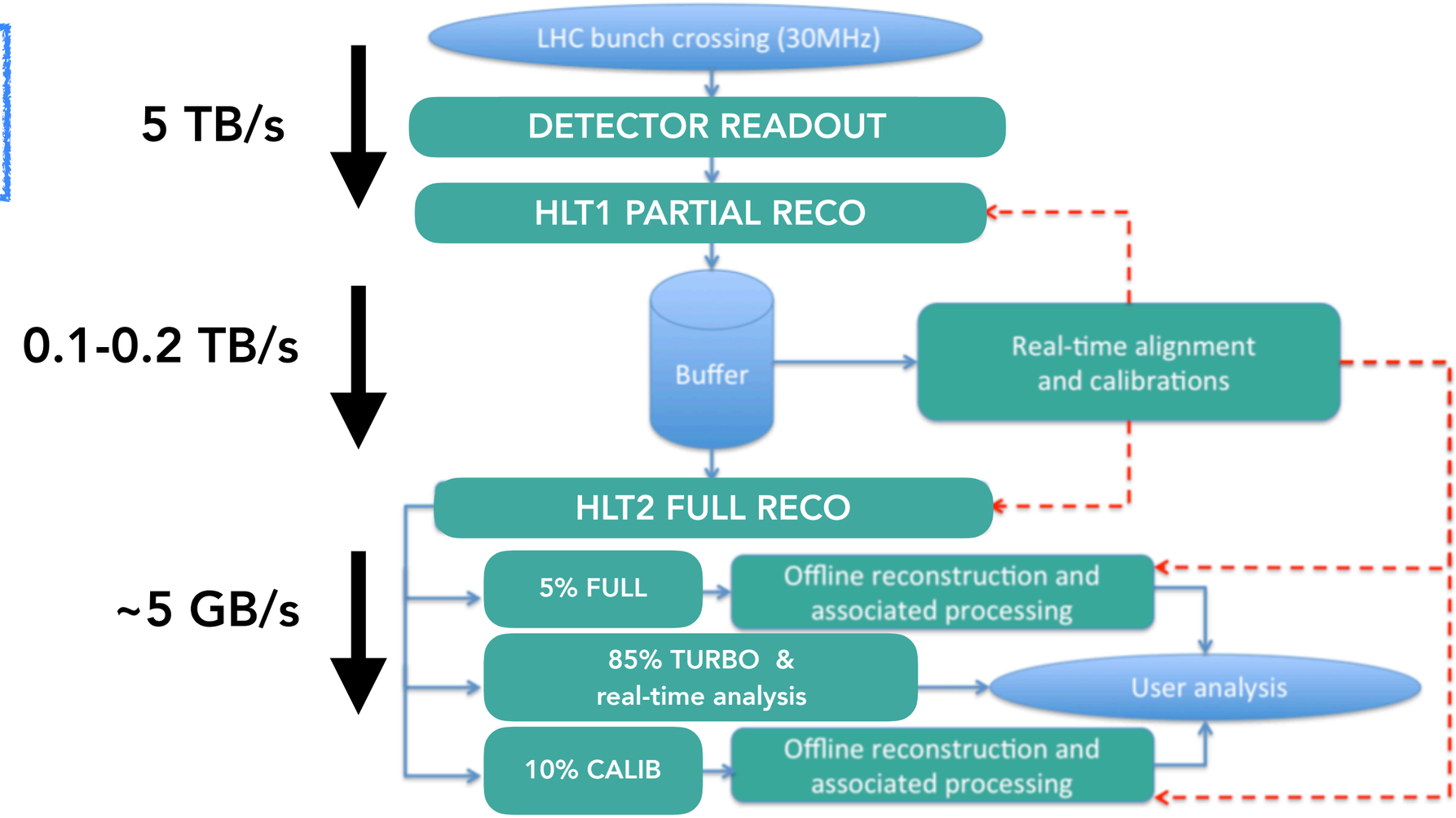
WHAT PROBLEM DO WE NEED TO SOLVE?

Upgrade I data processing chain



Upgrade I data processing chain

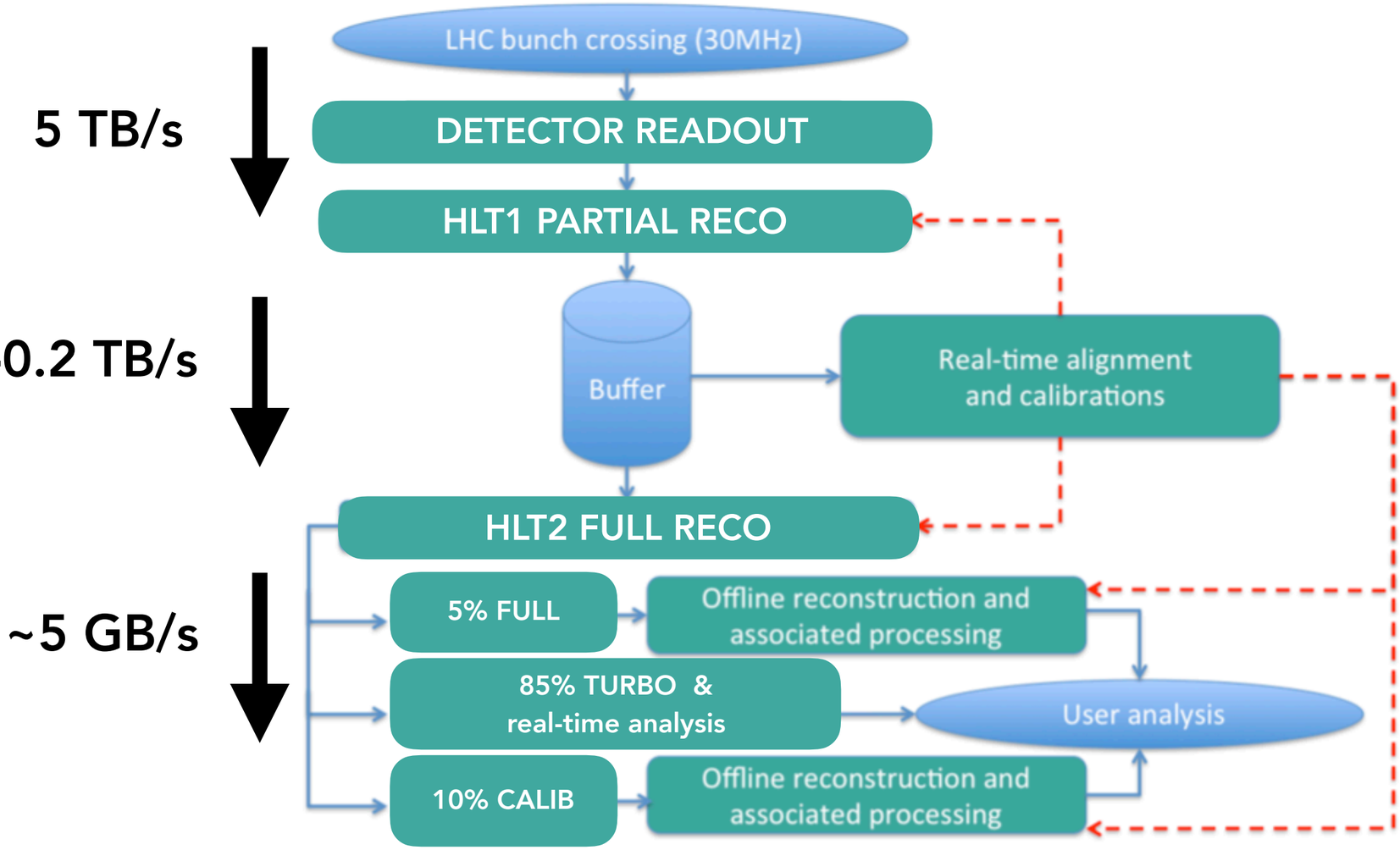
Need to use full tracker information in order to preselect interesting events ⟹ triggerless readout.



Upgrade I data processing chain

Need to use full tracker information in order to preselect interesting events ⟹ triggerless readout.

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

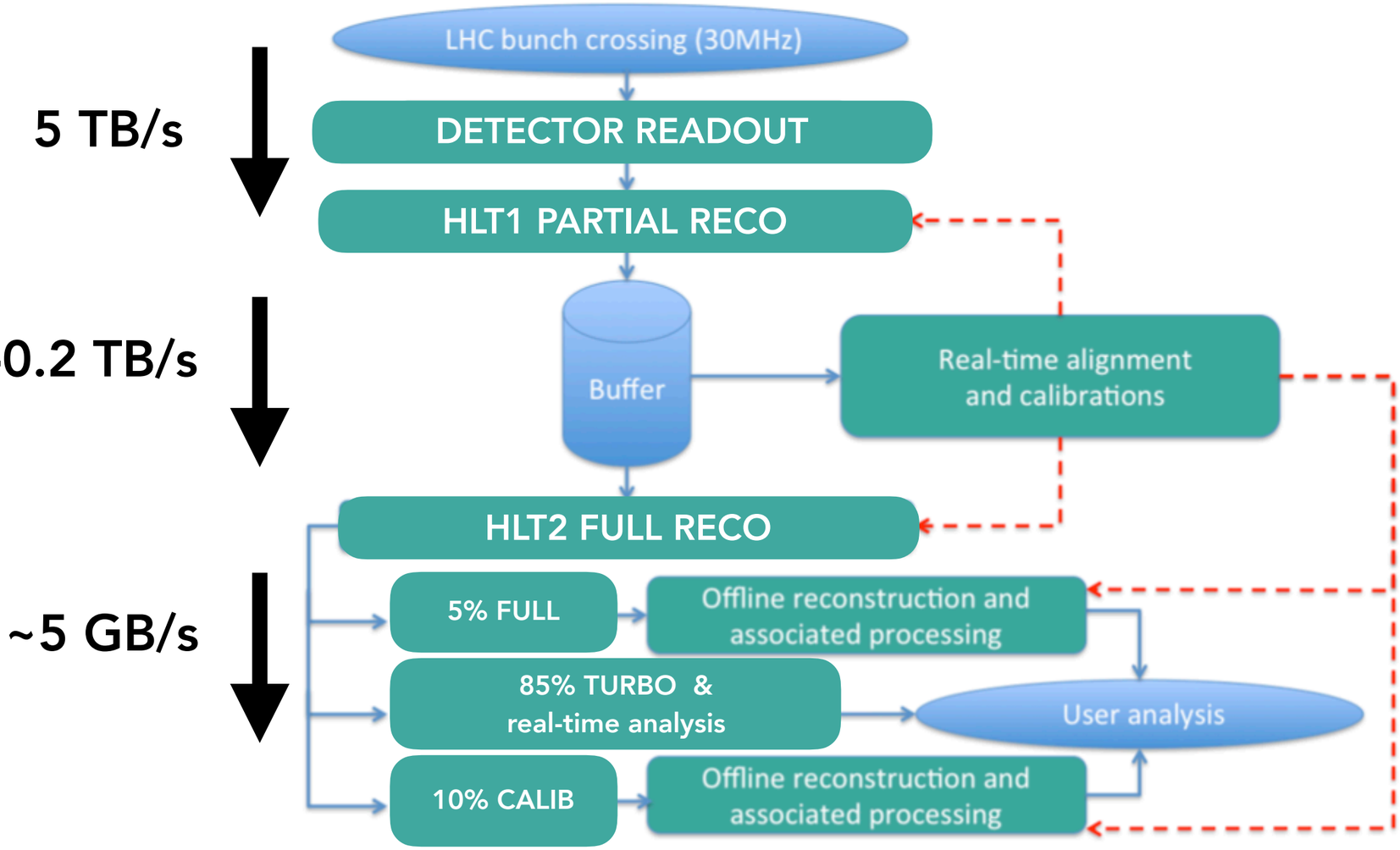


Upgrade I data processing chain

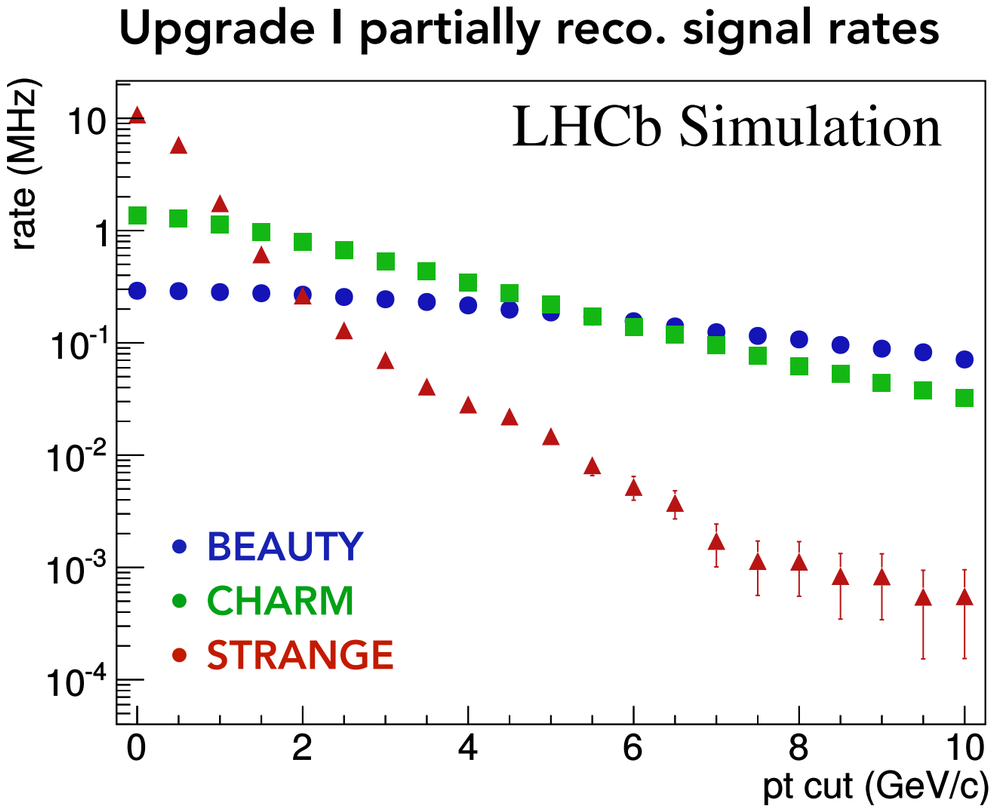
Need to use full tracker information in order to preselect interesting events ⟹ triggerless readout.

HLT2 roughly 10x slower than HLT1 per event in Run I/II ⟹ 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.

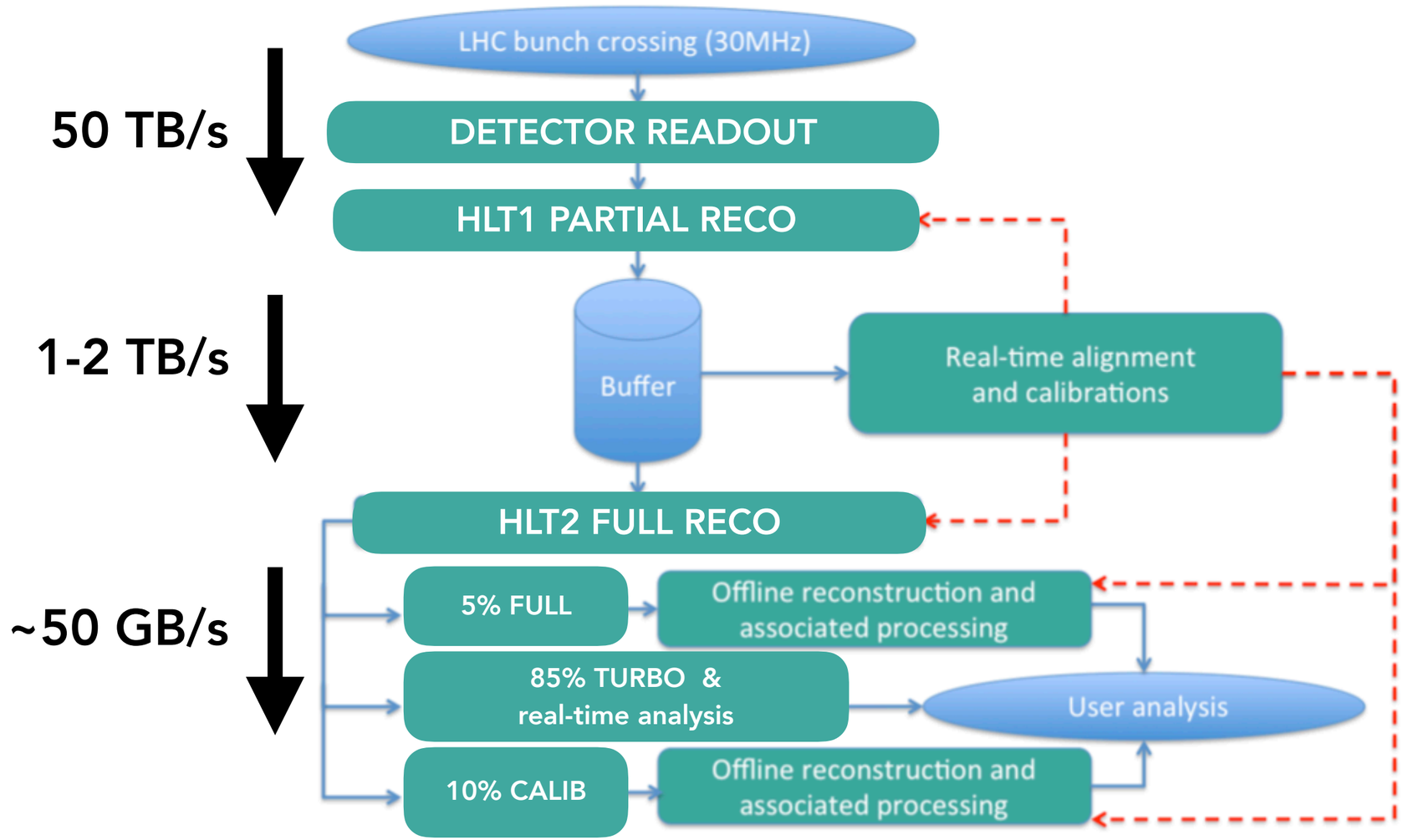


Implications of Upgrade I/II physics case



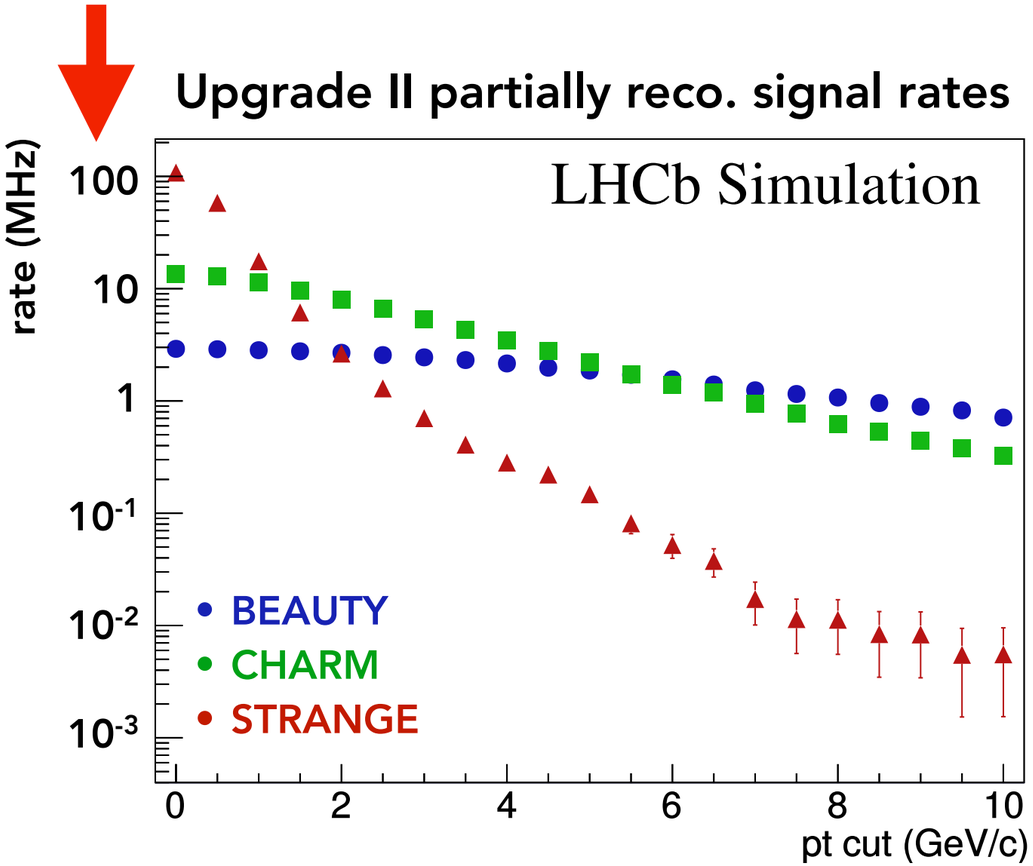
1. Almost all bunch crossings contain interesting signal
2. The job of HLT1 is to preselect signals most likely to be usable for analysis
3. Signal peaks at low P_T, no efficiency plateau to work on
4. Upgrade II must maintain full Upgrade I physics to make x6 worthwhile (?)

Data and signal rates at $2 \cdot 10^{34}$



Naively : it just scales by a factor ~10. However...

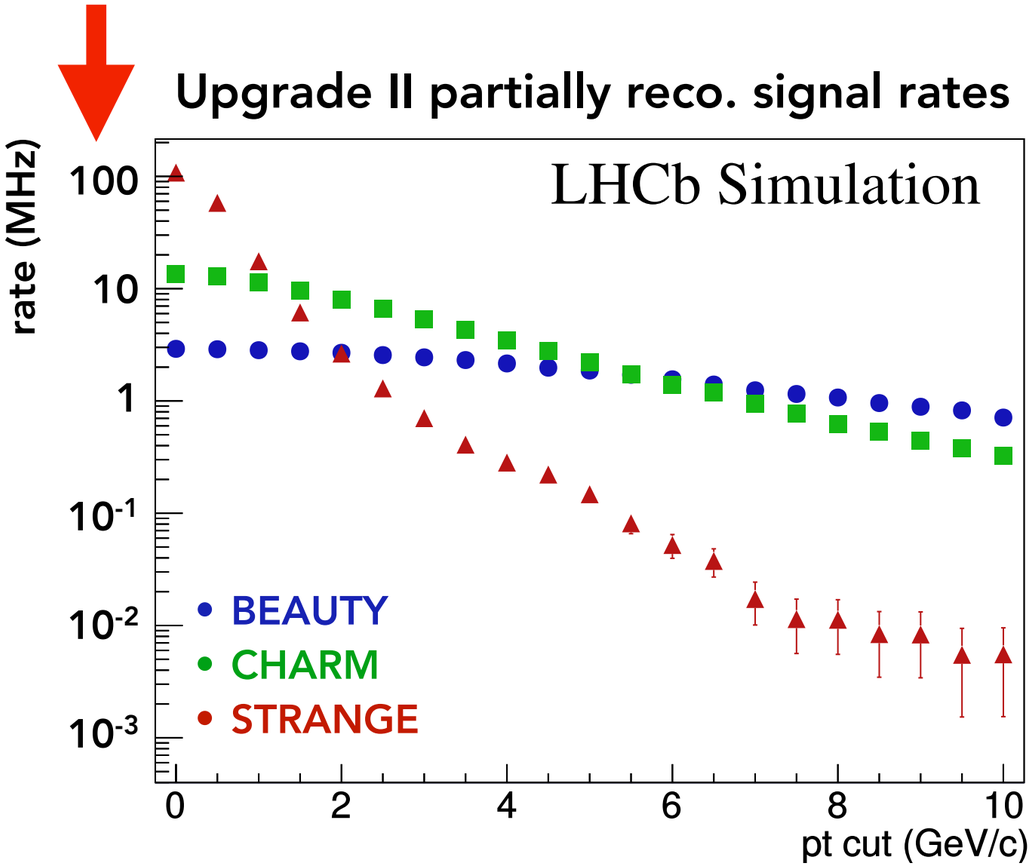
Data and signal rates at $2 \cdot 10^{34}$



Multiply everything by 10

At a pileup of 60, expect 8 MHz (!) of charm hadrons with PT > 2 GeV and τ > 0.2 ps which will leave a two-track vertex reconstructible inside the LHCb detector.

Data and signal rates at $2 \cdot 10^{34}$



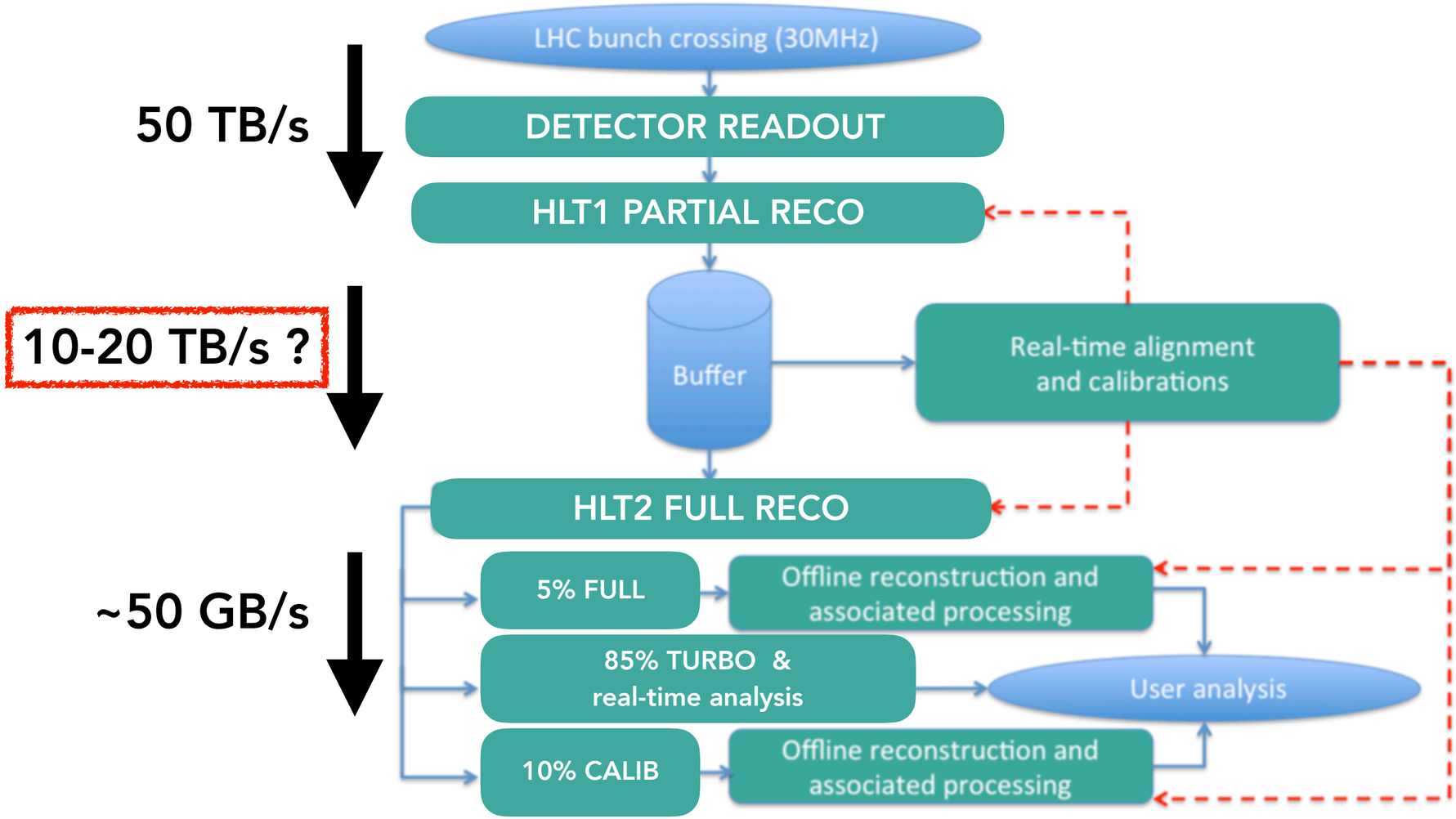
Multiply everything by 10

At a pileup of 60, expect 8 MHz (!) of charm hadrons with PT > 2 GeV and τ > 0.2 ps which will leave a two-track vertex reconstructible inside the LHCb detector.

Our data processing model for Upgrade I which relies on being able to perform a factor 25-50 reduction in *the number of events* at HLT1 breaks down due to *signal* saturation.

Processing complexity the increases quadratically even if the cost of reconstruction algorithms scales linearly. Run HLT2 at ~5-10 MHz or perform a pileup removal (timing?) already at HLT1.

Data and signal rates at $2 \cdot 10^{34}$

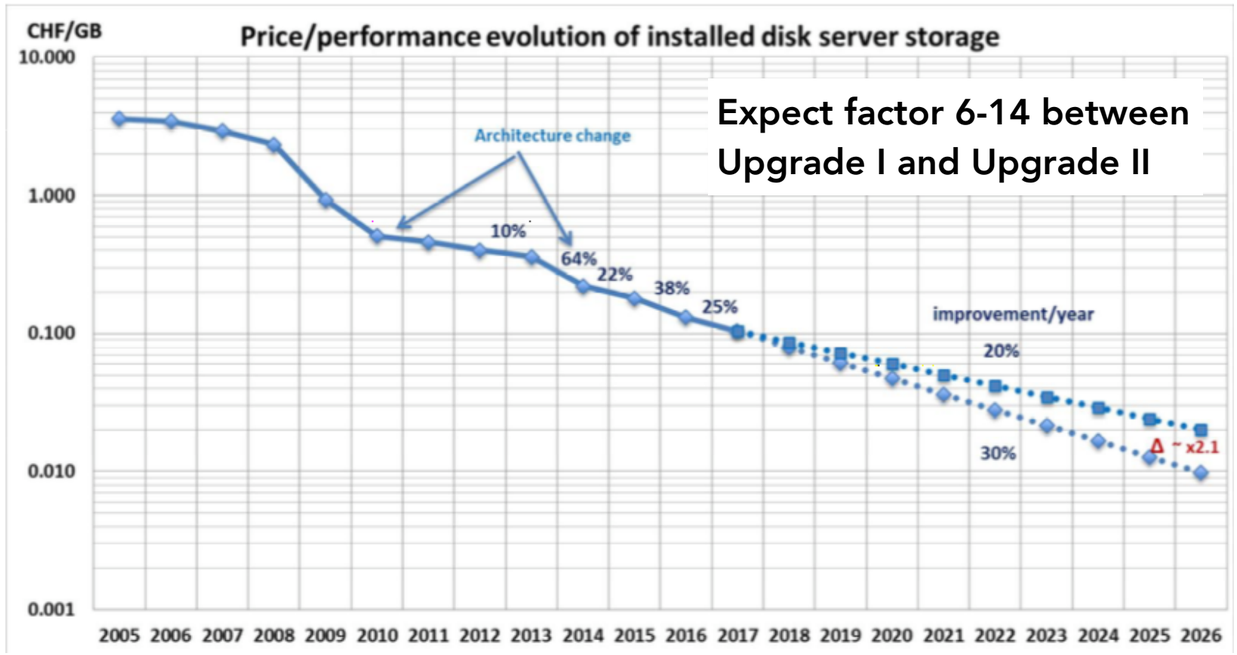


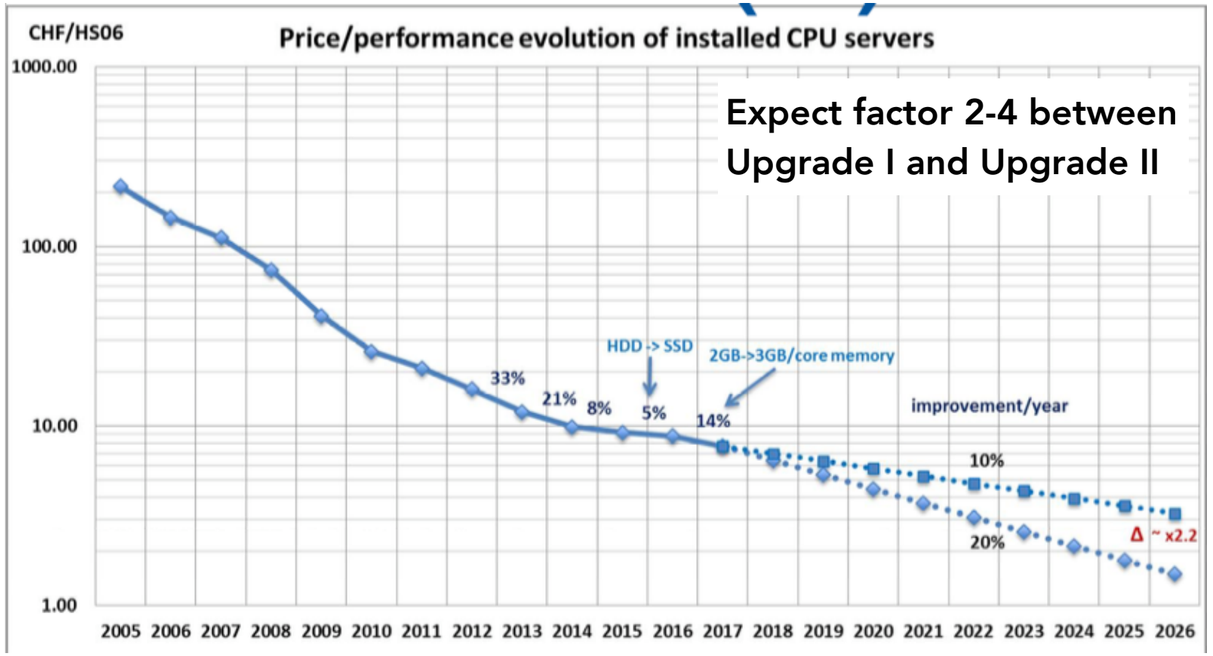
…under these conditions, can HLT1 still select *events*?

SPECIFIC CHALLENGES & POSSIBLE WAYS FORWARD

*Disclaimer : no attempt to be comprehensive, just highlight a few key points

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 Elba talk for many more details

GPD vs. LHCb UII data rate comparison

CMS detector	LHC Run-2	HL-LHC Phase-2	
Peak ⟨PU⟩	60	140	200
L1 accept rate (maximum)	100 kHz	500 kHz	750 kHz
Event Size	2.0 MB [a]	5.7 MB [b]	7.4 MB
Event Network throughput	1.6 Tb/s	23 Tb/s	44 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
Storage capacity needed (1 day)	0.2 PB	2.7 PB	5.3 PB

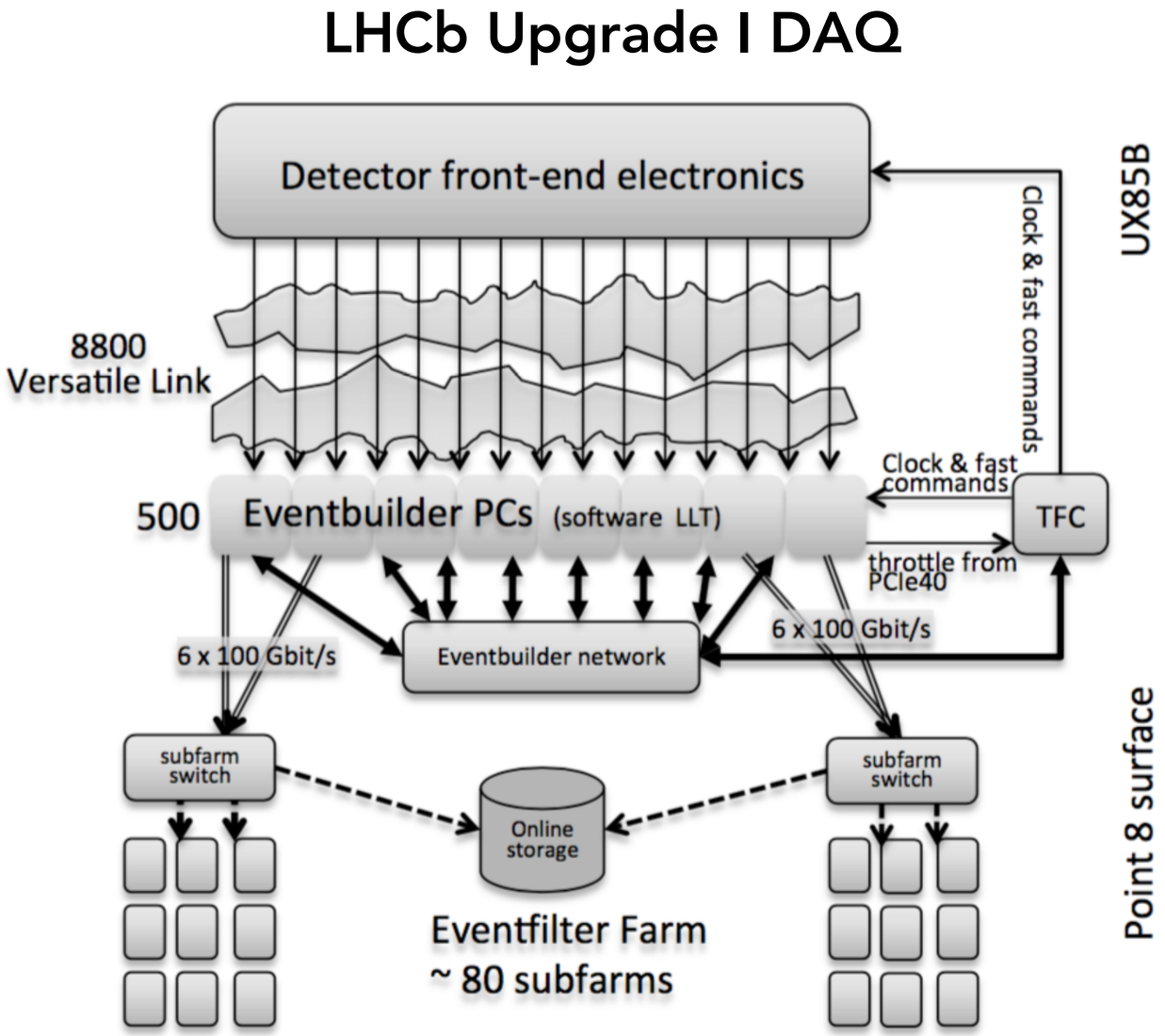
LHCb Upgrade II

30 MHz
1.5 MB ?
~500 Tb/s
??
??
??
50 GB/s ?
??

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

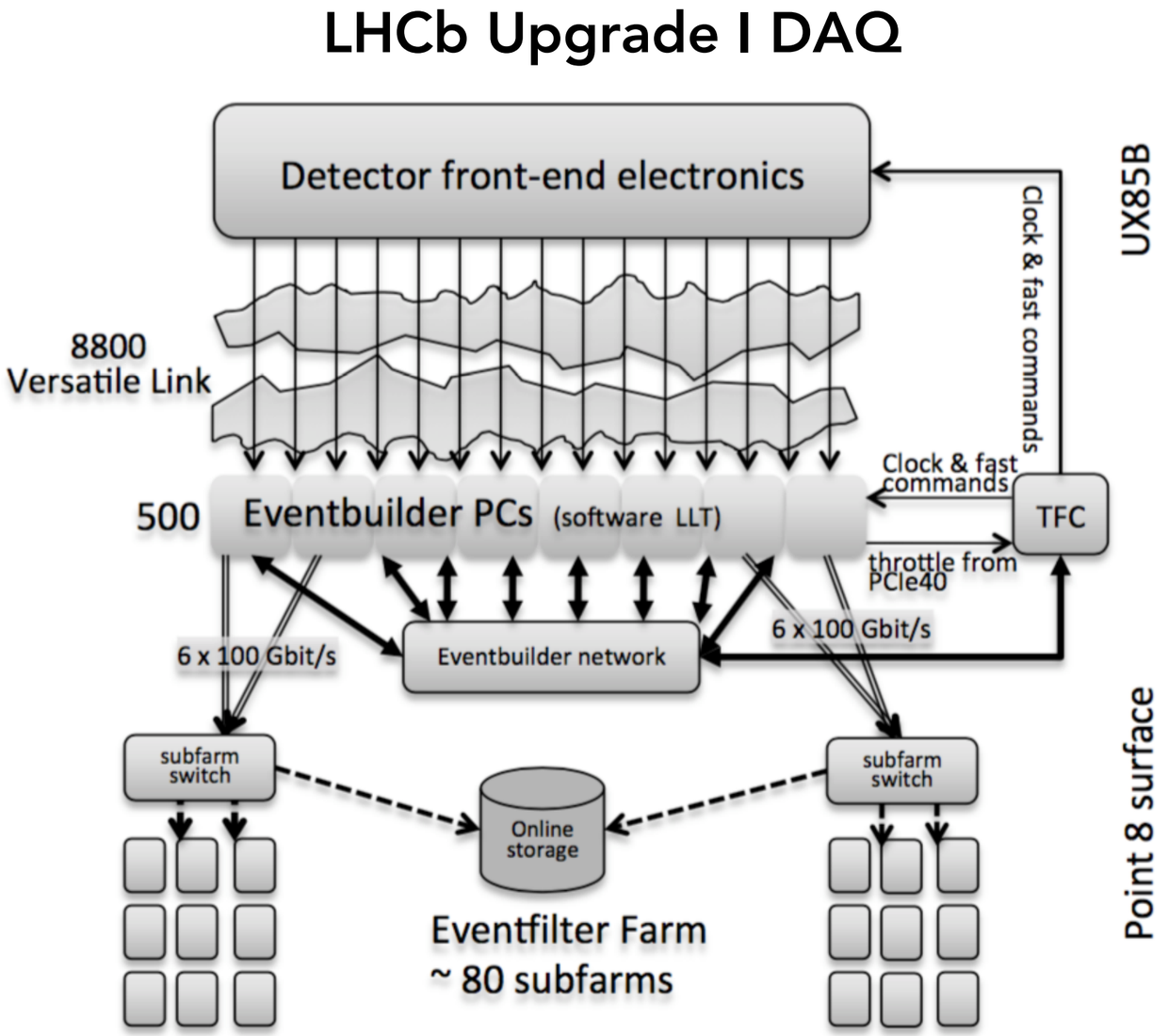
Upgrade II DAQ must process 10x the HL-LHC GPD data rate
Upgrade II Offline must process same data volume as GPDs

Challenges & evolution of DAQ



Challenges & evolution of DAQ

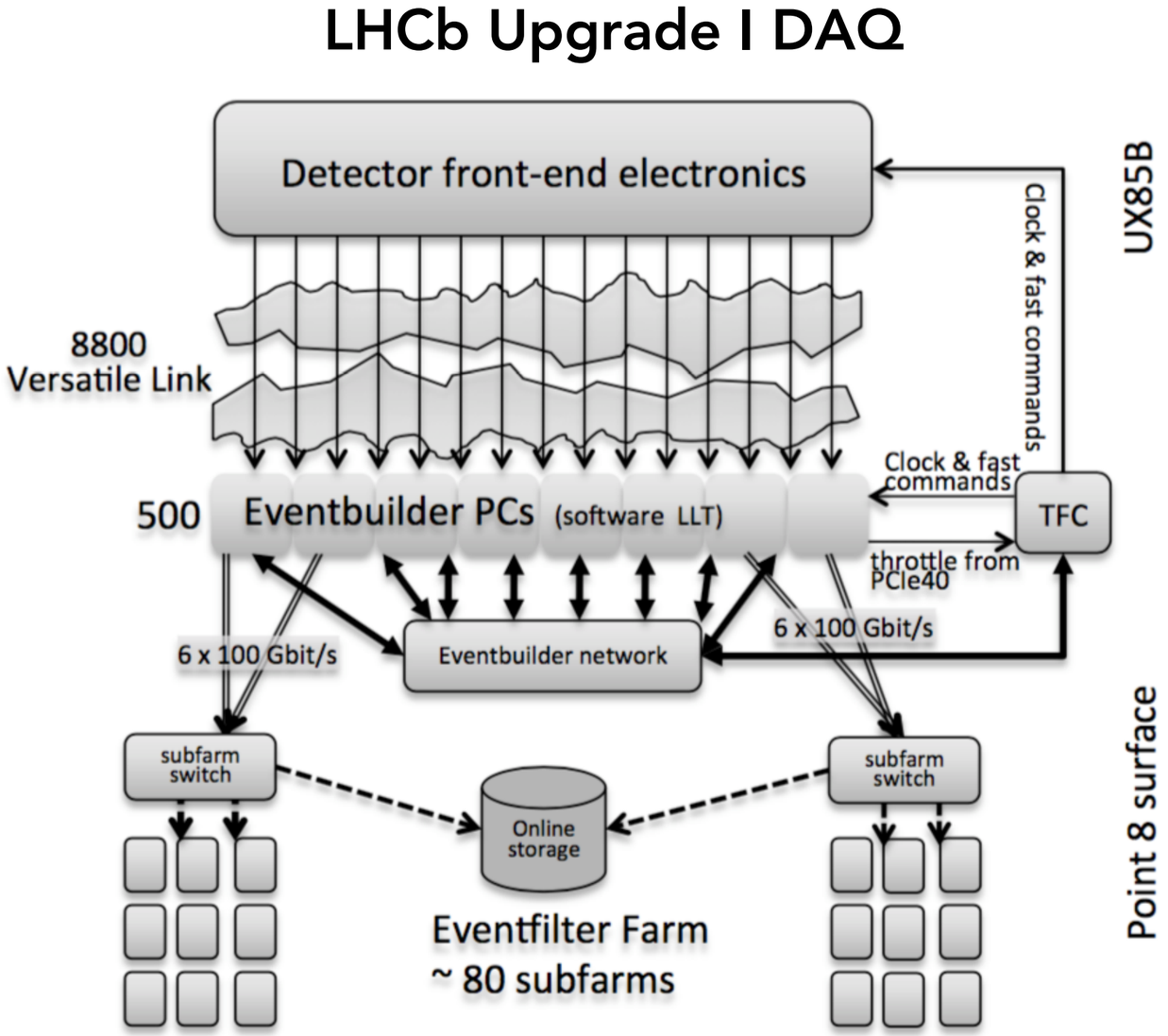
✗ **No possibility of a hardware trigger based on tracking because of the breadth of the physics case.**



Challenges & evolution of DAQ

✗ No possibility of a hardware trigger based on tracking because of the breadth of the physics case.

✗ Cannot save significant bandwidth with reconstruction in front-end as cannot send a subset of interesting objects for each individual detector (see backups for details why)

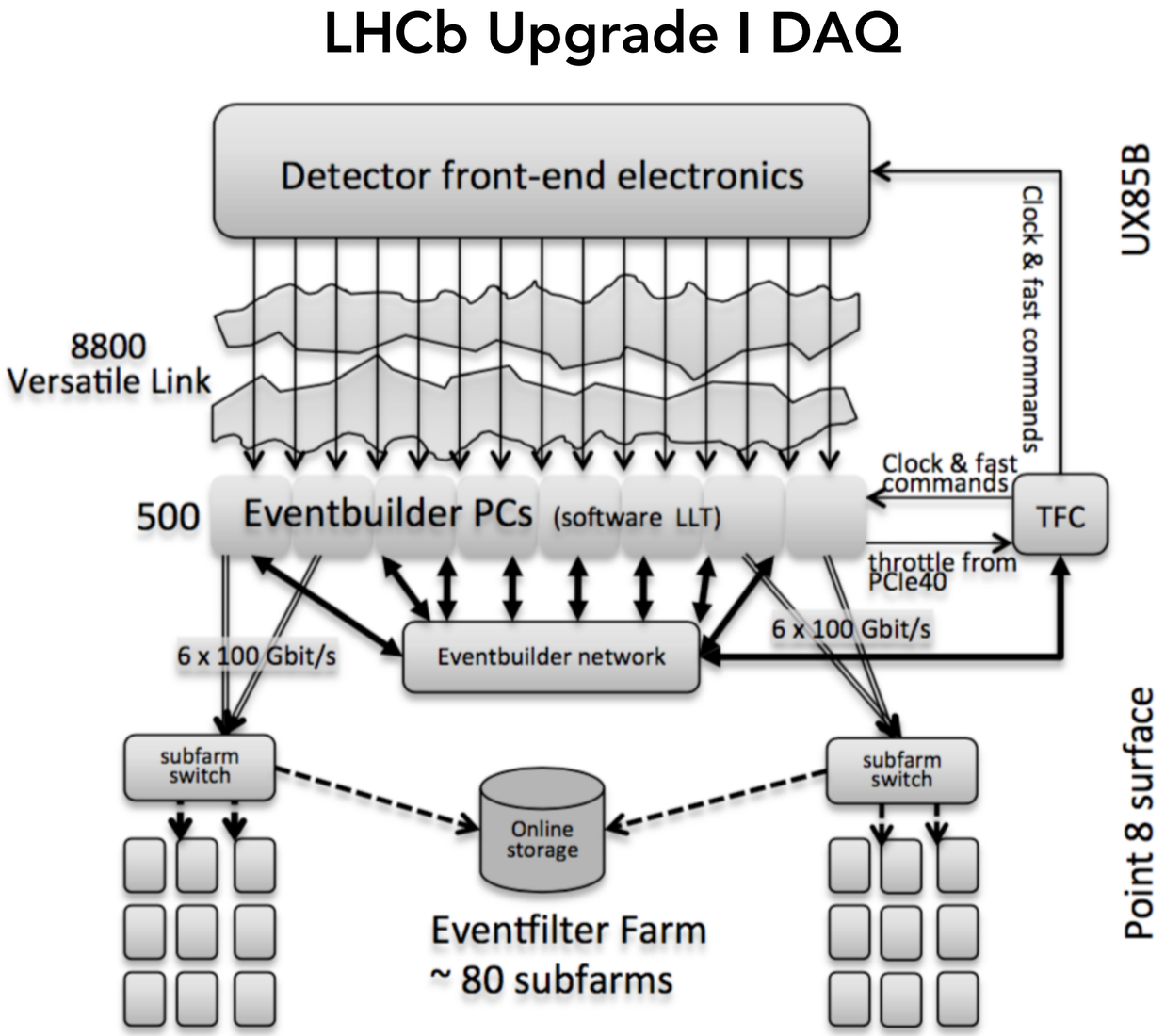


Challenges & evolution of DAQ

✗ No possibility of a hardware trigger based on tracking because of the breadth of the physics case.

✗ Cannot save significant bandwidth with reconstruction in front-end as cannot send a subset of interesting objects for each individual detector (see backups for details why)

✔ Therefore DAQ architecture should stay the same as in Upgrade I. Implement zero-suppression and clustering in front-end electronics, sort & transform to global LHCb coordinates (?) in back-end.



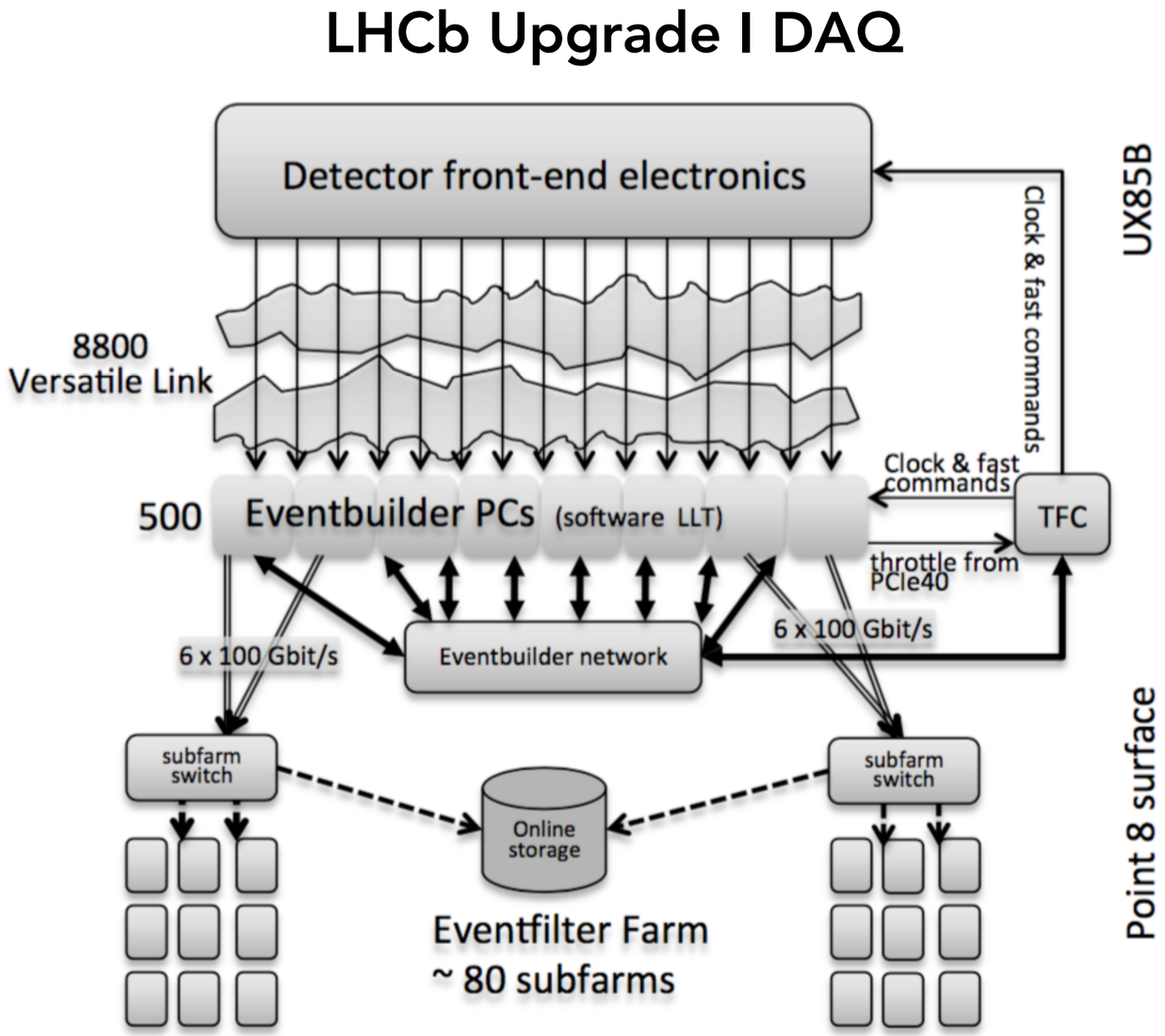
For alternative view see Giovanni's talk to TTFU

Challenges & evolution of 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?

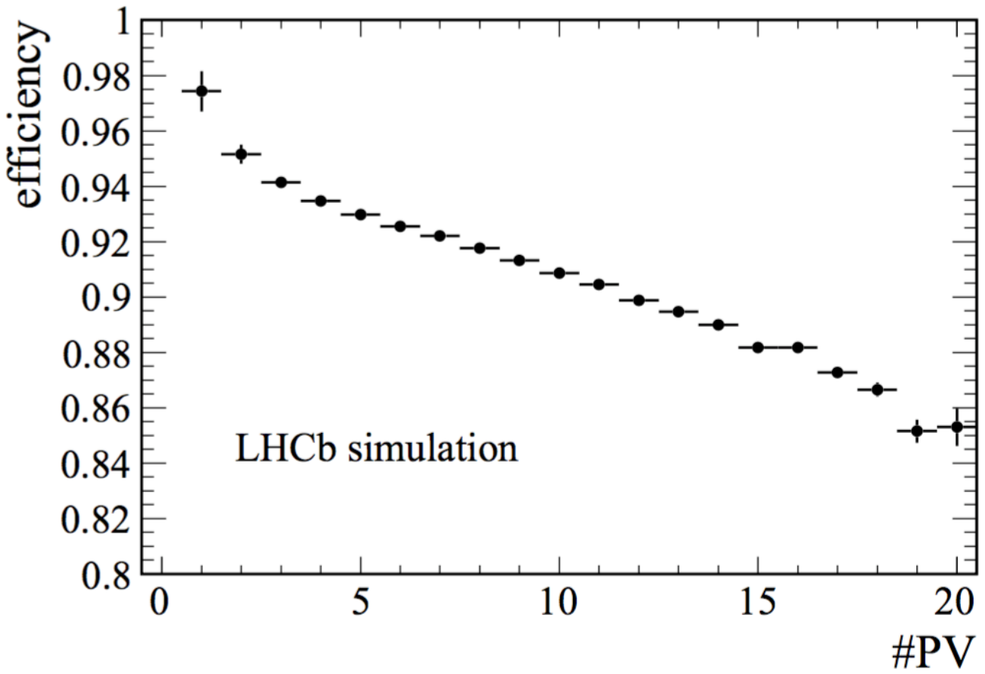
Farm : carry out R&D in next years on optimal use of hybrid architectures (GPU/ CPU/FPGA), remain flexible



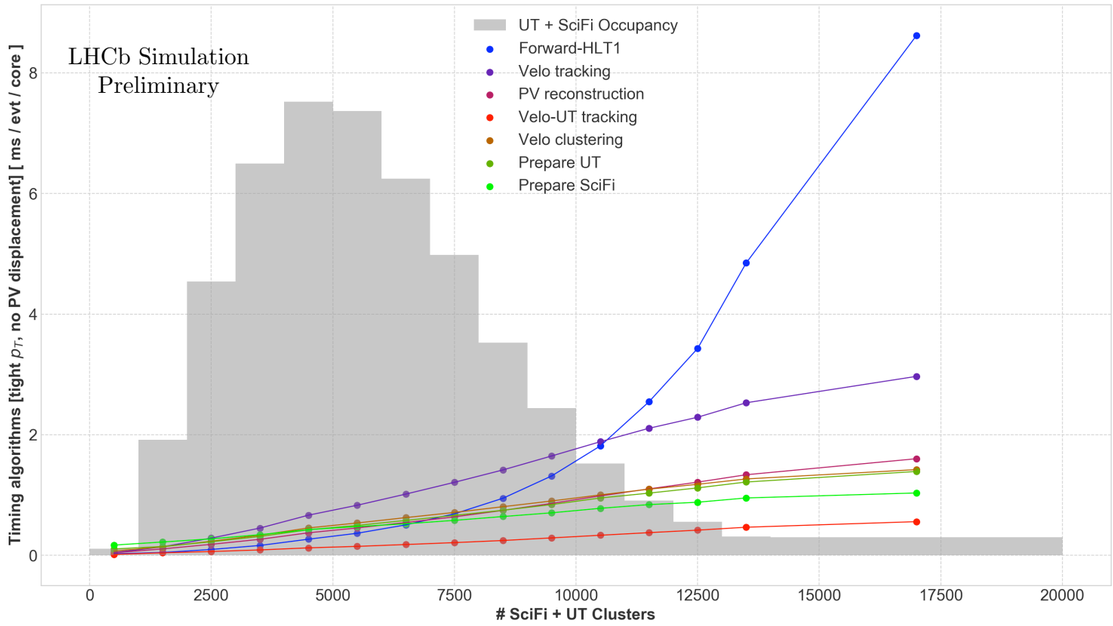
Partnership with industry seems key to ensuring cost-effective solutions to DAQ & processing

Challenges & evolution of reconstruction

LHCb Tracker TDR, forward tracking



Current best Upgrade I "fast" sequence



Challenge : greatly improve efficiency & fake rate while keeping cost of processing scaling linearly with the detector occupancy. Same logic applies to all other parts of the reconstruction.

Interplay with detector evolution

Efficiency/fake rate driven by tracker occupancy & granularity

Intuition : fakes relatively “easy” to remove, especially if we have timing in part of the tracker. Keeping high efficiency and low resource usage significantly harder. Interplay between spatial granularity of tracker & availability of timing seems important.

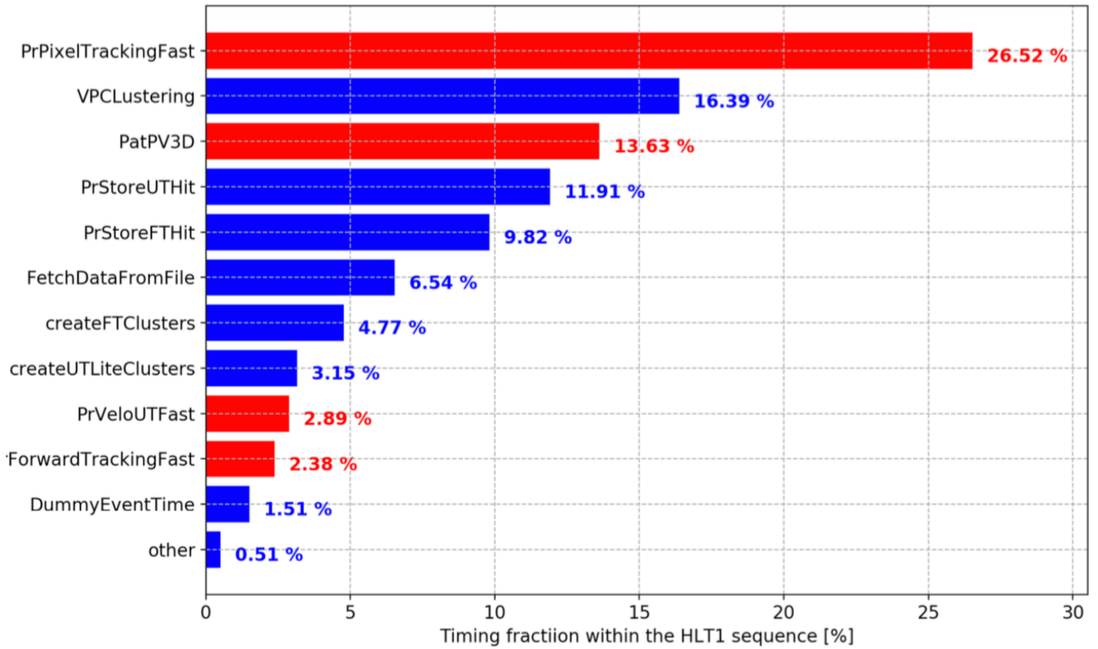
Forward tracking cost driven by long lever arm & poor momentum estimate (large search window) at UT. Ideas to improve this for Upgrade I by matching to single SciFi layer. In Upgrade II define tight search window using timing in UT + one tracker layer?

For RICH, CALO, Muon follow evolution of Upgrade I. All depend on tracking, RICH in particular relies on a track state with a proper covariance matrix between VELO and tracker. Will CALO rely on tracker for electron/photon separation or have a preshower?

Note : see Marco Petruzzo’s dedicated talk tomorrow on ideas for using timing to improve VELO tracking performance.

Challenges & evolution of HLT1

Current best Upgrade I "fast" sequence for displaced vertices



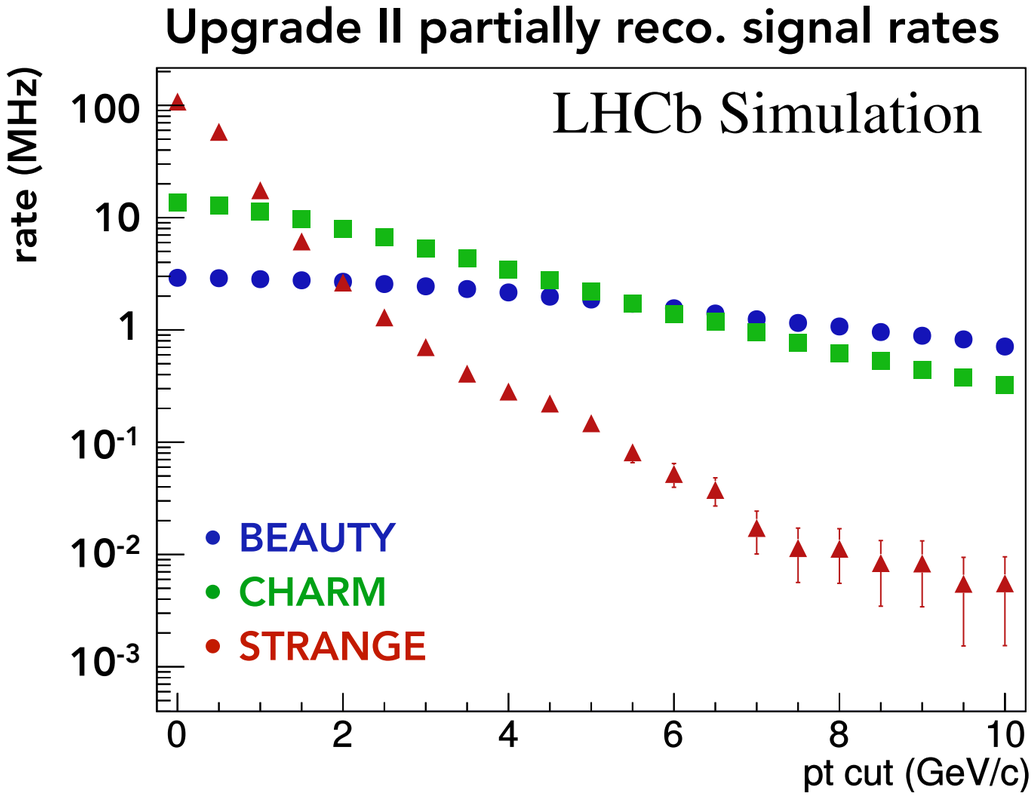
In Upgrade I current cost of HLT1 is 50% data preparation & 50% pattern recognition. Follow evolution of Upgrade I data preparation cost carefully, are CPUs best tool for clustering?

Challenges & evolution of HLT1

Because of Upgrade II *signal* rates, cannot use the current partial HLT1 reconstruction to simply select entire events anymore.

Try to present a 0th draft of alternative strategy, based on having limited timing information in part of the tracker.

➡ **HLT1 finds an interesting signal based on a high-p_T subset of decay products, then uses timing to suppress pileup in full reconstruction. See backups for details.**



Boundary between HLT1 & HLT2 essential for disk buffer cost.

Options for timing & pileup suppression

1. Fast reconstruction of high-P_T tracks, use candidate vertex to define timing window of interest in all relevant subdetectors.

2. Complete track reconstruction within defined timing window, add particle identification information and compatible neutral objects for these tracks. Select events containing exclusive fully-reconstructed signals of interest.

HLT1

BUFFER AT 0.5–1 TB/s

3. Perform a full event reconstruction for the subset of selected events to add e.g. isolation and FT information, refine track properties.

HLT2

May not address all our physics, depending on which parts of tracker can make timing information available. Understanding efficiency of step 2 may be non-trivial. Even with pileup suppression processing cost likely to be significantly greater than for Upgrade I. In particular, would imply significant combinatorics burden in HLT1 for the first time, may be hard.

Options for timing & pileup suppression

1. Fast reconstruction of high-P_T tracks, select events containing candidate vertices which also define timing window of interest for later processing.

HLT1

BUFFER AT 10-20 TB/s

2. Complete track reconstruction within defined timing window, add particle identification information and compatible neutral objects for these tracks. Select events containing exclusive fully-reconstructed signals of interest.

3. Perform a full event reconstruction for the subset of selected events to add e.g. isolation and FT information, refine track properties.

HLT2

Would require an enormous disk buffer : around 500 PB to buffer one fill. However if disk really evolves faster than CPU in the coming years, may not be totally out of the question. Lower HLT1 processing cost, and no combinatorics, thus much more maintainable and benefits from out-of-fill processing. Understanding efficiency of step 2 may be non-trivial.

Pros & cons & interplay with HLT2

Currently HLT2 takes around 10x the processing cost of HLT1, and is largely performed out-of-fill. With both ATLAS/CMS and us luminosity-levelling in Upgrade II, will fill length still accommodate significant out-of-fill processing?

The evolution of this relative cost is hard to estimate, will probably depend a lot on specifics of detector design. Important to have a coherent design of the different subdetectors around viable candidate HLT1/HLT2 sequences.

If timing windows lead to significant hit inefficiency and degradation in the track resolutions, would also need to redo or improve the reconstruction at the end of the HLT2 processing. Cost affordable by definition but significant departure from current model of a single unified reconstruction, may have a non-negligible impact on understanding systematics.

Challenges & evolution offline

The benefit of real-time analysis is that once out of the trigger, scaling is naively linear with the integrated luminosity. Will be easier to extrapolate once Upgrade I computing model is defined later this year.

Exception is cost of simulation, especially non-parametric (“full”) simulation in the case where a significant part of the trigger is performed using co-processors. Will those be natively available on the grid or will we have to emulate them?

This is a key difference between ATLAS/CMS and LHCb : we do not work on an efficiency plateau by definition and hence we have to simulate our trigger for many (most?) analyses. Requires a coherent data processing model from the readout to the ntuple making.

Compressing the TURBO stream?

What if we cannot afford to persist even TURBO data for analysis?

Going beyond TURBO means reducing information about the signal

- More sophisticated packers, e.g. work by Manuel
- Use autoencoders to compress signal information to minimum needed by the analysis ⟹ promising preliminary studies for PID
- Automated analysis tasks to measure per-run likelihoods for key standard candles (lifetimes, masses, Δm_s...) ⟹ promising preliminary studies for tracking efficiencies

Personal view : analyses tasks are going to be crucial for real-time analysis data quality and monitoring. For precision measurements, would think long and hard before persisting reduced signal information, as new sources of systematics tend to arise as the precision increases. Naively more viable for bump hunts.

Should not forget or underestimate the challenge of preserving our code and analyses over the next 15-20 years.

Calibrating the detector & simulation

Real-time alignment and detector calibration will likely scale if reconstruction can be made to scale; if anything will improve by having samples of e.g. Z or other high-P_T signals available more frequently to help align the tracker.

Simulation costs and resources are already increasingly mismatched, will require move to parametric (“fast”) simulations already for upgrade.

For precision measurements, especially those which are not ratios, this actually means shifting the burden onto data-driven methods for understanding the detector performance.

Must tag-and-probe every step of the reconstruction for every particle species and ensure that our data processing strategy enables the collection of the relevant calibration samples across full kinematic range of interest.

CONCLUSION & FUTURE

Conclusions, DAQ, HLT & Reconstruction

The Upgrade II TDAQ has to process 100 times the data of Upgrade I, and 10 times more data than the ATLAS or CMS HL-LHC systems

Unlike ATLAS or CMS our physics case imposes a triggerless readout, and requires that we bring track & neutral reconstruction together with particle identification from practically the start of the processing chain.

Additional or higher granularity detectors means more data to transfer and process, which must be accounted for from the start in the design.

By far the biggest data processing challenge in the history of HEP, huge potential for industrial partnership in developing a cost-effective solution

Conclusions, data access and simulation

Moving more and more to fast or parametric simulations means that we have to have better data-driven ways to understand the detector and calibrate this simulation. Need per-particle reconstruction and identification efficiencies in a data-driven way for all particle species.

Mandatory to collect all the onia and Cabibbo-favoured charm decays possible in a tag-and-probe approach. Important design requirement when preselecting events or parts of an event during the processing.

Streaming and data access particularly challenging, flat cash approach to GRID resources further and further from our real needs. Likely to impose particular difficulties for smaller institutes which cannot afford large local clusters which can run our software and analysis framework.

Overall conclusion

If we run LHCb Upgrade II at $2 \cdot 10^{34}$, the detector readout and reconstruction will be one of the most challenging problems

Current processing model likely scales in terms of technology, but far from clear it scales in terms of cost, in particular DAQ.

Must coherently design subdetectors and data processing model. Early pileup suppression (based on timing?) crucial.

Must draw on lessons of both Upgrade I and on evolution of systematic uncertainties across our whole physics programme.

BACKUP

Pileup suppression model

The essential problem for DAQ in Upgrade II is that almost every bunch crossing contains interesting signal, while almost all the particles produced in those bunch crossings are not interesting.

In an ideal world where every subdetector had high-precision timing information, we could identify displaced high p_T vertices and dileptons, and use their timing to delete most of the pileup in all subdetectors before processing only the collision of interest.

In the real world where we will most likely only have limited-precision timing information for a subset of the tracker, can nevertheless benefit from this approach to reduce

1) The cost of the RICH&Muon reconstruction, which need only be done for the subset of interesting tracks from the signal pp collision
2) The cost of Kalman fitting the tracks, which for now costs more than finding the tracks in Upgrade I.
3) The cost of particle combinatorics, likely to be a big fraction of the time spent in Upgrade I HLT2 and which by definition scales non-linearly with luminosity.

If the CALO also has timing information, can of course benefit much more from this, especially since the majority of our analyses which rely on CALO information can add it after finding a two-track vertex.

These benefits are on top of what can be gained by performing a 4D track finding, and address the problem of having to run HLT2 at ~30 MHz, which may be prohibitively expensive given expected technology evolution.

Why not reconstruction on front-end?

Reconstruction on the front-end electronics has been proposed as a way of saving DAQ bandwidth between the front-end and the back-end. This is however not possible because :

- LHCb reconstructed objects are roughly same size as the raw event (hits/clusters), so simply transmitting them and suppressing the raw information does not save space
- Therefore to save space have to select a subset of interesting reconstructed objects
- But you cannot do this before bringing together the data from all subdetectors because of signal rates

This argument is independent of technology and of whether you can actually perform the reconstruction on the front-ends. Even if you could perform it perfectly, you would not save any bandwidth, and since a triggerless readout is imposed by the physics case (see slide 10 of this talk) you cannot save anything else either.

More detailed arguments can be made for each individual subdetector, e.g. for the VELO vast majority of clusters are associated to tracks, and if you did want to suppress clusters at the front-end (forgetting impact of this on physics) need to transmit at least two states per track. Compared to the average VELO track size (~6.5 clusters) two states is not that much less information to send. Similar arguments apply elsewhere.

Strategies like this work best in detectors like ALICE where the event size is dominated by a single subdetector (the TPC) and where that subdetector's data can be very significantly compressed independently of what happens elsewhere (e.g. by deleting low-momentum curlers), but neither is the case in LHCb.