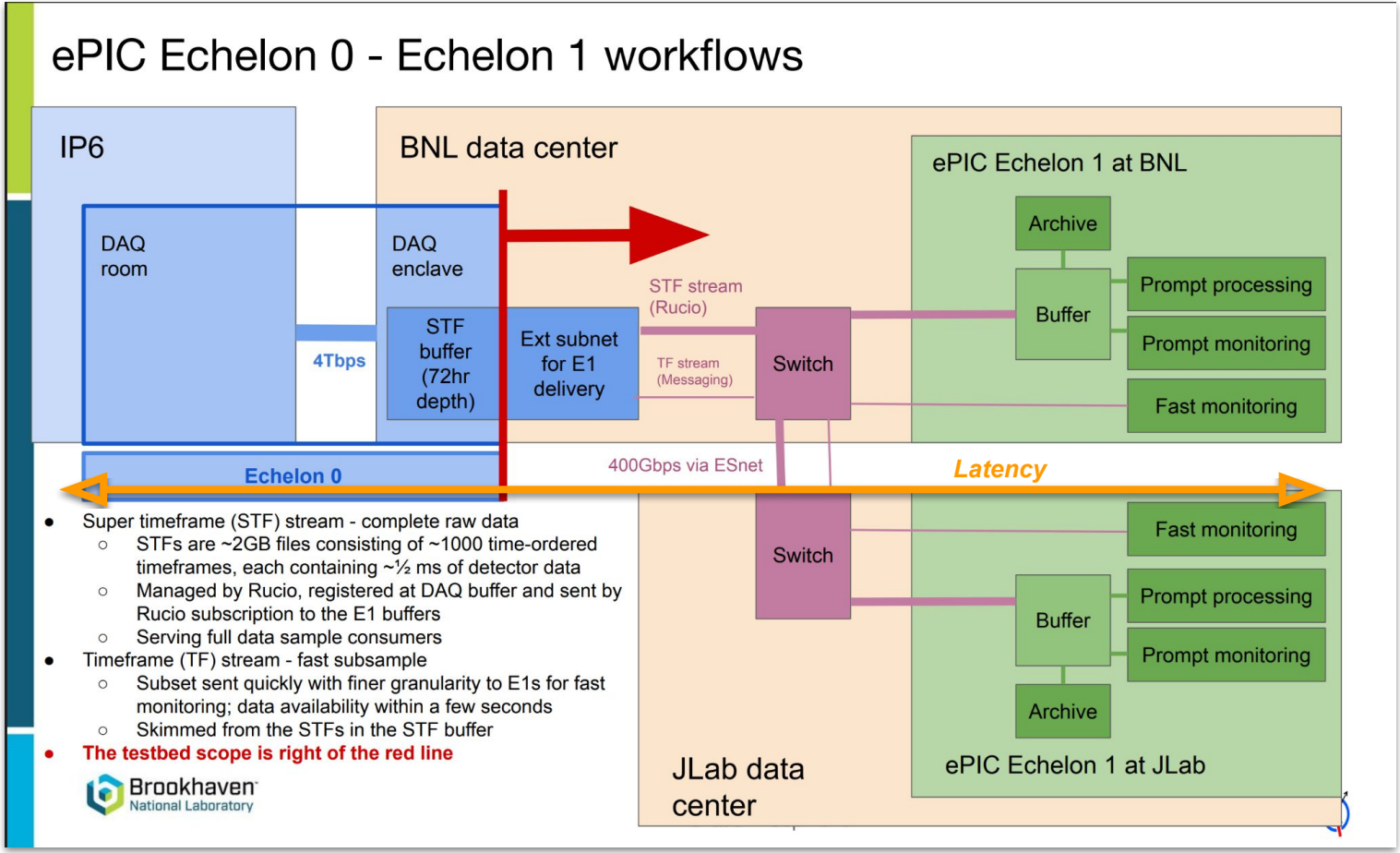
ePIC Latency Requirements for Data Transfer and Processing

Oct. 8, 2025

Latency of interest: Time between interaction in experiment reconstructed event at E1 sites.

Assume the time it takes to get into STF buffer is << than total latency for *Fast Monitoring*



This slide shown without permission from Torre's slide deck here:
https://indico.cern.ch/event/1587990/contributions/6692003/attachments/3137737/5567974/Slides.pdf

The ePIC Streaming Computing Model Version 2, Fall 2024

https://zenodo.org/records/14675920

4.4 Prompt Reconstruction

A defining characteristic of ePIC's streaming data model is the events are reconstructed in near real time from the streaming data, modulo time varying calibrations that will require later reprocessing for a final fully calibrated reconstruction. The prompt availability of reconstructed data, and concurrent calibration cycle consuming it, is a crucial element of ePIC's objective to have a rapid, near real time turnaround of the raw data to production, as expressed in the software principles[17]. The stringent low latency and high availability requirements of prompt reconstruction, together with the locality of its inputs at the Echelon 1 sites, makes this a processing activity limited to Echelon 1. Prompt reconstruction uses streaming based processing described in Section 6.1 below, taking time frames as produced by the DAQ as input and producing event (single interaction) based data as output, for processing by analysis software.

Prompt/Fast monitoring is the more stringent requirement

4.5 First Pass Reconstruction

It is expected that the Echelon 1 facilities will have insufficient compute resources to perform the complete first pass reconstruction for incoming data. The prompt reconstruction workflow at Echelon 1 will process, at a minimum, the sample necessary for monitoring, diagnostics, quick-turnaround calibration and so on. The remaining first pass reconstruction processing will be shared with Echelon 2 facilities. The maximum acceptable completion time is about 2-3 weeks. This timescale is driven by calibrations. Given the expectation of relatively low data rates during commissioning and early running, and the need to commission, validate and stabilize the use of Echelon 2s for first pass reconstruction, it is likely that Echelon 2s will be integrated after the first pass reconstruction workflow at Echelon 1 is operating smoothly and Echelon 2s are validated as ready.

Full recon is planned on the calibration timescale of weeks

Torre's ePIC Prompt Processing Orchestration Spreadsheet:

https://docs.google.com/spreadsheets/d/1H7jtsPd0YDldR10u-SALgu-hTuEsbCSNYz60Qs4YMHs/edit?gid=0#gid=0

30	**Time to the first plots based on a full processed event sample corresponding to one STF file. This number is key to setting the scale**	

Prompt processing scenario		
Acceptable latency from run start to first E1 full-STF plots (s)	**30**	**Time to the first plots based on a full processed event sample corresponding to one STF file. This number is key to setting the scale**
Working backwards from that...		
Acceptable latency from run start to first STFs arriving at E1 (s)	5	Not consistent with relying on Rucio. At run start, need to send STFs express to E1s, like fast monitoring
Express data rate out of DAQ buffer to the two E1s (MB/s)	900	This is the express fraction only.
Target STF processing time during prompt processing (s)	**25**	**Wall clock time to process a full STF equivalent of data**
Cores per STF to attain target	2702	Number of concurrent reco processing threads needed to process a full STF in desired time
Threads per reco instance (EICrecon level parallelism)	10	The event level parallelism of an EICrecon job
Workflow orchestration level parallelism	270	The remaining parallelism comes from parallel workflow. This number (or half the number across 2 E1s) is too big to hit one STF
STF arrivals at DAQ during STF processing time target	15	A new STF arrives at DAQ every half second, can spread the prompt processing across them if we get them to the E1
STFs expressed to each E1 per processing time window	**2**	**Processing spread across this number of STFs at each E1 to attain STF processing time target**
Orchestration fan-out per STF arrival	135	Concurrency level on STF processing to attain STF processing time target
Orchestration fan-out per E1, ie divide by 2	**68**	**This is the parallelization for each STF that the orchestration must deliver if both E1s contribute equally to prompt monitoring**
Rucio STF transport latency to E1 (min)	5	Meanwhile, Rucio is doing the bulk STF transfers, with a large latency
STF arrivals at DAQ during that latency	180	With many new STFs arriving during that latency
STFs needed to maintain target STF processing time during that latency	24	At run start, we must express transfer STFs to E1s to achieve the latency target
Fraction of STFs needing fast delivery, at least early in run	**13%**	Amounting to this fraction of the STFs. If we want to maintain a low-latency view of the last few min, this express fraction has to be sustained, if we are using a full-STF transfer scenario. (See the STF sampling scenario below.)

Torre's ePIC Prompt Processing Orchestration Spreadsheet:

https://docs.google.com/spreadsheets/d/1H7jtsPd0YDldR10u-SALgu-hTuEsbCSNYz60Qs4YMHs/edit?gid=0#gid=0

STF sampling scenario			
Assuming still that what we get from DAQ is STFs, send the express data by sampling the STFs, e.g. to an object store			
STF-equivalent processing time target (s)	30	as above	
STF-sample transfer time (s)	2		

The value of the *processing time target* affects:

1. E0 to E1 Network latency requirement
2. Amount of compute needed for Fast Monitoring

Echelon 0 -> 1

Considerations:

1. What is the timescale for gleaning insights from the data stream?
    a. How many events are needed to monitor a property?
    b. How much beam time is needed to produce that many events?
    c. Are we monitoring 100% of events, or some fraction?

2. What actions can be taken on that timescale?
    a. Adjustments to beam(s) (e.g. steering, luminosity)?
    b. Adjustments to DAQ (e.g. thresholds, time offsets)?
    c. Adjustments to Detector (e.g. LV,HV)?

3. Who implements the action?
    a. Human?
    b. Computer?

E0 to E1 Latency is driven by the time it takes to produce actionable information plus the time to implement the action.

Backups

Timing

- Timing system jitter: 5ps
- Best detector timing resolution: 20-30ps
- Bunch crossing rate: 98.5MHz
- Interaction rate: <=500kHz
- DAM Boards (streams): 95
- Time frame(TF): 0.6ms
- Super Time Frame(STF): 1000 TFs = 0.6s
  - ~45k events of interest/STF
  - ~19 hrs/core/STF to process

Resource	Type	Amount
Outgoing bandwidth	raw data	200Gbps
	Monitoring, slow controls, misc. meta data	$\leq$ 1Gbps
	Capacity headroom	$\approx$ 200Gbps
	TOTAL	**400Gbps**
Incoming bandwidth	monitoring, calibration	$\leq$ 1Gbps
Storage	Disk (outgoing data buffer w/ 24hr)	1PB

Table 2 Echelon 0 networking and storage requirements.

Resource	Type	Amount
Outgoing bandwidth	Raw data -*immediate* ($\frac{1}{6}$ of total)	17Gbps
	Raw data -*replay* (contingency)	50Gbps
	monitoring, slow controls, misc. meta data	1Gbps
	TOTAL	**68Gbps**
Incoming bandwidth	monitoring, calibration, slow controls *(from E0, E1, and Echelon 2)*	1Gbps

Table 3 Echelon 1 networking requirements. Values shown are for a single E1 site.