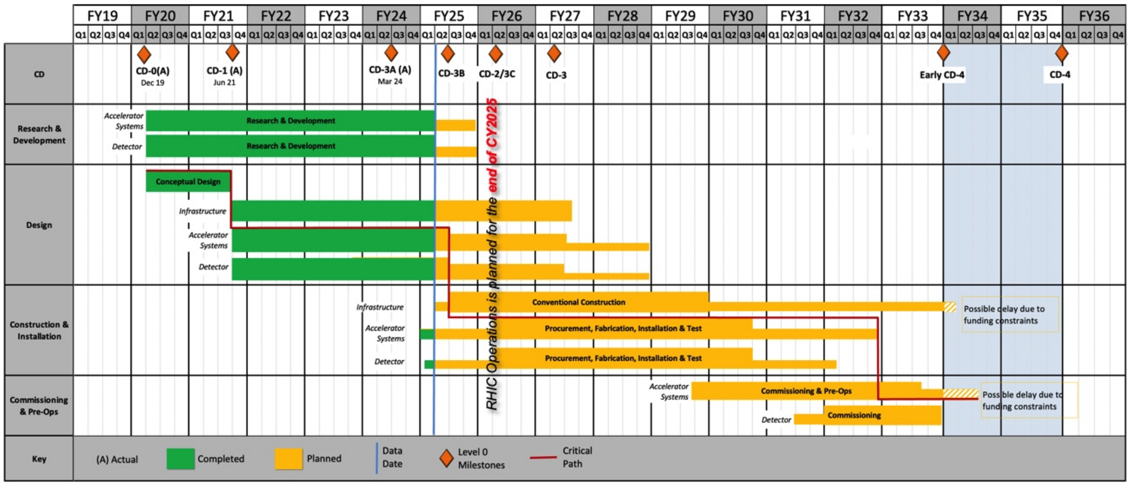

Figure 1: Reference schedule for the EIC Project. The critical path is the accelerator systems. Science operations are expected to begin in approximately a decade.

High-Level Questions

1. **It would be useful to have a brief overview of the EIC program schedule, for reference. Start date, planned extended shutdown periods, etc ...**

   The anticipated operational schedule for the EIC program consists of approximately 6 months of running and 6 months of downtime each year. This yearly cycle is expected to begin in the mid-2030s. A reference EIC Project schedule is shown in Fig. 1, which reflects the construction and commissioning timeline through the start of operations.

   > ***Note from the TIM:*** *Cosmic data taking is planned for 2031 for the commissioning of the detector, a couple of years before the first beam.*

2. **The document describes the online and computing aspects of ePIC. How much of the online part is in scope of the Advisory Committee we are part of? Could be "all", "none" or "from this point". It would be useful to know, as the part not in scope we should treat it as "for information" while the one in scope we should treat it as "for discussion".**

   We aim to build an integrated system in which the Streaming DAQ (E0) and the Streaming Computing (E1–3) are developed in close coordination, with ongoing detailed discussions at their interface. Accordingly, the document covers both components.

   The Streaming DAQ is within the scope of the EIC Project and is reviewed as part of the EIC Project's formal review processes. In contrast, the Streaming Computing is not part of the EIC Project. Instead, it falls under the oversight of ECSJI, with ECSAC serving in an advisory capacity to both ECSJI and the Streaming Computing effort.

   For ECSAC, the scope of discussion starts with the availability of data at the E0 exit buffer, ready for E1 transfer and processing. Content about the Streaming DAQ before this point is included for information only.

   > ***Note from the TIM:*** *For ECSAC, the scope of discussion begins with data at the output buffer of E0.*

3. **We still have difficulties understanding what does what in the DAQ. In particular, we do not understand in Figure 3 which elements 1) buffer the information from the different subsystems in timestamped**

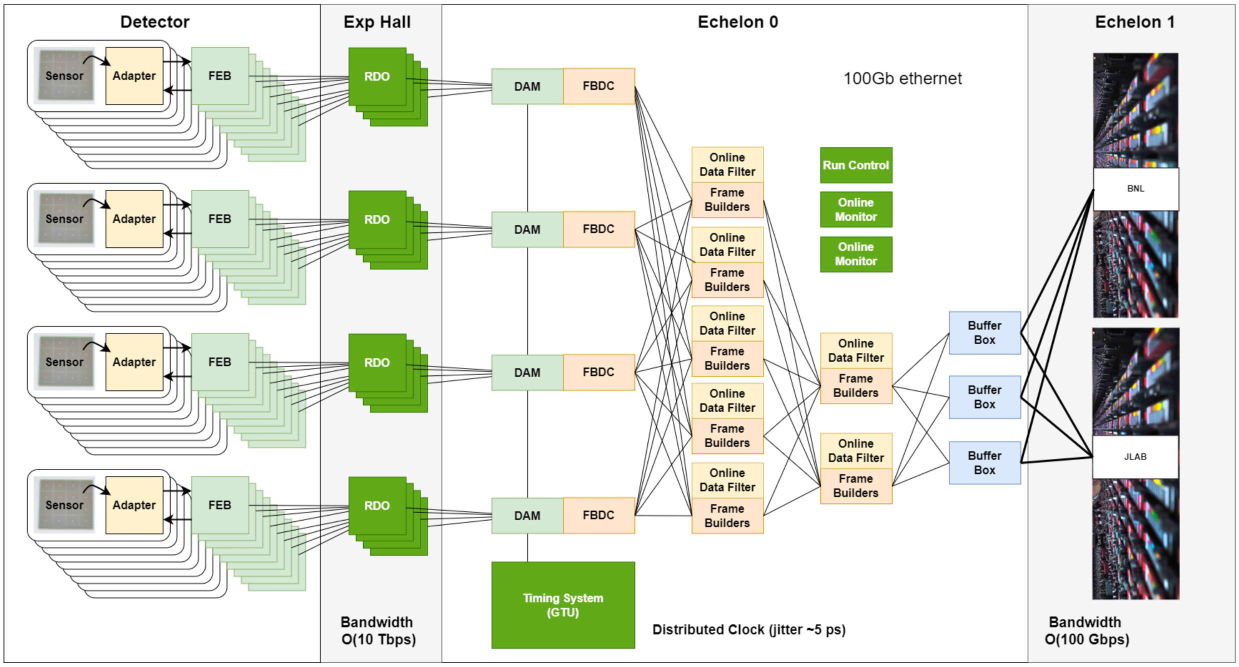

Figure 2: Schematic of the streaming DAQ. The diagram illustrates four logical blocks corresponding to buffering in detector sections, TF assembly, data handling and processing strategy, and frame builder processing and transfer.

datafiles, 2) recombine those timestamped datafiles into a timeframe, 3) compress the timeframe (or the previous datafiles) 4) push the timeframes into the offline.

In Fig. 2, we provide an updated diagram to clarify the data flow and buffering architecture referenced in the original Figure 3. The figure presents the system in four logical blocks, representing the stages of buffering and data handling:

1. **Buffering in detector sections**: Buffering is minimal, but necessary to build detector-section-specific timeframe (TF) buffers. Data from the detector sections is buffered in the DAM/FBDC (Data Aggregation Module / Front-End Board (FEB) Data Concentrator). The available buffer is on the order of $1\,\mathrm{s}$, while the expected required buffer is approximately $5\,\mathrm{ms}$.
2. **TF assembly**: TF builders assemble TFs from multiple detectors. Each frame builder has an available buffer of approximately $10\,\mathrm{s}$, and the system is designed to be scalable. The expected required buffer per frame builder is on the order of $1\,\mathrm{s}$.
3. **Data handling and processing strategy**: The default approach is to read in full detector data, including noise and backgrounds, when data rates are low enough to allow it. The DAM boards are capable of performing additional processing, such as generating selection criteria to eliminate noise hits, performing feature finding, and applying higher-level zero suppression than what is available at the front-end ASIC level (which are hosted on the Front-End Boards (FEBs).
4. **Frame Builder processing and transfer**: Frame Builders gather sub-detector TF contributions into a single buffer for each complete TF. They also have processing capabilities that range from quality assurance and sanity checking to feature finding, data compression, and potentially even High-Level Trigger (HLT) processing. Completed TFs are pushed to E1 via Buffer Boxes.

Note from the TIM: *The timing information is injected into the frontend board, and framing is performed immediately afterward based on that information. The experimental hall bandwidth is quoted between* $2\,\mathrm{Tbps}$ *and* $10\,\mathrm{Tbps}$*, with* $2\,\mathrm{Tbps}$ *being the more accurate estimate. The* $400\,\mathrm{Gbps}$ *output rate accounts for two* $100\,\mathrm{Gbps}$ *streams (one copy to JLab and one to BNL), along with a contingency factor of two. A single stream operates at* $100\,\mathrm{Gbps}$*.*

Note from the TIM: *A Timeframe is* $0.6\,\mathrm{ms}$ *in duration and approximately* $2.5\,\mathrm{MB}$ *in size during early years of running. Timeframes are grouped into SuperTimeFrames consisting of 1000 Timeframes (approximately* $0.6\,\mathrm{s}$ *and* $2.5\,\mathrm{GB}$*). The number 1000 is arbitrary and may change. This grouping takes place in the online farm, which is therefore designed with two levels for Timeframe construction and a third level for ordering and aggregation of complete Timeframes. There is substantial buffering at the DAM level to ensure that asynchronous data from various subsystems can be collected into the same frame once it arrives at the Frame Builders. The Frame Builders themselves are equipped with sufficient buffering to ensure that SuperTimeFrames contain contiguous Timeframes.*

Note from the TIM: *The E0–E1 schematic cites a bandwidth of* $4\,\mathrm{Tbps}$ *between the DAQ at IP6 and the enclave at SDCC. This number should be reviewed and corrected if necessary. It refers to the total bandwidth between the readout computers and the first layer of frame building.*

Note from the TIM: *Failure modes and recovery: A failure of a DAM board would likely lead to a stoppage of the detector to allow for recovery. With approximately 100 DAM boards in the system, the overall probability of failures may be non-negligible.*

4. **We also do not understand (related to the above) if the timeframes are processed in any way online (except for being compressed). What is the role of the online data filters? Do they filter out timeframes and based on which criteria? We are trying to understand how much intelligence is in place to make sure that timeframes contain useful information, as we understand they are kept forever.**

   Yes, TFs may be processed in the Streaming DAQ, but only when necessary to stay within the $100\,\mathrm{Gbps}$ bandwidth budget from E0 to E1. The example is the dRICH, where the gradual increase in luminosity during the early years of operation will provide an opportunity to evaluate and cross-check the data reduction methods.

   The raw digitized data from the detectors is organized into data banks. Any processed data is written into additional data banks within the same file, ensuring that the original raw data banks are not modified.

   To reduce the data volume, some raw data banks may be omitted from certain TFs. However, a fraction of TFs will always include all raw data banks to ensure comprehensive sampling. A potential exception to this would be if raw banks are compressed using lossless compression, which would allow full data preservation at reduced volume.

   Importantly, no TFs are entirely discarded. Even TFs that do not contain identifiable events will still be written out. In those cases, the processed data banks will typically be very small or empty.

   Note from the TIM: *Ideally, filtering should be avoided in E0. However, some detectors produce considerable noise, so a degree of filtering is necessary. Triggering is also technically feasible, as sufficient compute capacity is available, but the baseline approach is to avoid it and perform the processing at E1 and E2.*

5. **A key aspect of the computing model is the need to run first pass reconstruction at Echelon-2. This is because JLAB and BNL will not have enough resources to do prompt reconstruction. While Echelon-2s will provide more resources, the computing model will also be more complicated and that will probably imply further costs for JLAB and BNL. Has the cost been looked at in a holistic way?**

The E1 sites at BNL and Jefferson Lab will have the necessary resources to support the basic operational needs of the ePIC experiment. The E2 sites are intended to accelerate the computing use cases.

We have not performed a holistic cost analysis for our computing model.

> ***Note from the TIM:*** *The execution of some workflows at lower Echelon levels is driven by both technical and sociological factors, which will be further balanced as the collaboration and the computing model reach greater maturity.*

6. **Could you provide an explanation of the services and capabilities you expect to find at Echelon-0, 1, 2, and 3? In several parts of the document one finds "this workflow fits the capabilities of Echelon-X but for the other one we need X-1. It is not however clear what you expect to find in X-1 that you do not find in X. An example is section [4.8] on analysis: what do you need for quark-gluon structure analyses that you find in Echelon-2 and not in Echelon-3? E.g. high memory nodes, high IOPS storage, ... ?**

   E0 is distinct; it represents the Streaming DAQ of the ePIC experiment and is not shared with the other echelons.

   E1 and E2 share roles in data processing. Low-latency use cases run at E1 sites solely, while E2 provides additional capacity to accelerate all other computing use cases.

   E3 corresponds to home institute computing, ranging from individual laptops to small local clusters. These resources are insufficient for many QCD analyses, which typically require:

   - Multidimensional or multichannel analyses
   - Large-scale simulations for estimating systematic uncertainties
   - Inverse problem-solving methods to extract QCD information

   Such tasks exceed the computing and storage capabilities available at E3.

   > ***Note from the TIM:*** *E1 and E2 sites are very similar in terms of the services they provide, with the exception that E1 sites include archiving of raw data as a service. E1 sites are primarily intended for low-latency use cases, whereas E2 sites are designed to accelerate science delivery. E3s, by contrast, are fundamentally different. They represent home institute computing and are not suitable for the most compute-intensive analyses.*

7. **It would be useful to have a description of the data model, with formats of each data tier, a short description of what they contain and the expected event sizes (apart from the timeframes which contain many events). This will help understanding some of the numbers. For example in section [5.3.2] it is very hard to follow the logic of why a complete simulated data sample is 100PB.**

   Our data model is based on EDM4hep, with well-established EDM4eic extensions. It used for simulation and reconstruction workflows. Output data is currently written to ROOT TTree files. The RNTuple format is currently under evaluation and has shown a reduction by a factor of two in file size for our data.

   However, the data formats for TFs and for ePIC data during operations are not yet defined. It is also not yet clear whether the EDM4eic data model will apply to data prior to the E0 exit buffer.

   For our estimate of simulated data, we assumed that the output will mimic raw data, not just physics objects. This approach allows the simulated data to be reprocessed using the same reconstruction workflows as real data.

   Specifically regarding the $100\,\mathrm{PB}$ simulation dataset size, this estimate is based on our resource estimation parameters: a reconstructed event size of $12.5\,\mathrm{kB}$, a multiplier of 1.2 to account for the inclusion of MC truth, and an annual number of simulated events equal to ten times the number of real events.

> ***Note from the TIM:*** *The additional 20% event size attributed to 'truth' information should be verified, as it depends strongly on the type of truth data being included.*

8. Similarly to the data model, it would be useful to have a description of the workflows transforming from one data tier to another, and a rough understanding of the CPU time/event for each transformation. This will help understand the assessment of the CPU needs.

We provided the relevant information during the last review:

Reconstruction and Simulation Tasks	Times
Reconstruction event processing time with background [s]	2
Reconstruction algorithmic speedup factor 10 yrs out	1.5
Simulation event processing time with background [s]	15
Full simulation speedup factor 10 yrs out	1.5
Combined time with background, with speedup [s]	11

The estimates presented are based on our current understanding of the software and performance on modern cores. They will evolve as the software stack matures and additional performance data becomes available.

The ePIC Experiment

9. The text mentions 1.5 fb-1 per week with 60% efficiency. How many weeks of run are foreseen in a year (the question is related with the high-level one about the schedule)

The expected operational schedule for the EIC program consists of approximately six months of running and six months of downtime each year. Please see also our response for Question 1.

The Streaming Data Acquisition System

10. Section [3.1] Different channels will have different readout times. So in order to build the frames you will need some buffering. Where do you buffer?

The maximum readout time from detector channels is on the order of $5\,\mu s$.

To accommodate this and support data assembly, buffering is provided at several levels:

- The DAMs have buffer capabilities on the order of $1\,s$, which is much larger than the typical TF length of $0.6\,ms$.
- The TF buffers provide capacity on the order of $10\,s\, \times$ # TF builders, again much larger than the TF duration.
- The Super Timefram (STF) buffers offer capacity on the order of $10\,s \times$ # STF builders, with a typical STF duration of $0.6\,s$.

These buffer depths ensure that data can be held long enough to accommodate varying readout latencies across subsystems while assembling consistent time-aligned data frames.

11. Section [3.1]. What happens if in a frame you do not find at least one event and how likely is that to happen? Do you throw away the frame? We guess no, as you assume improved SW triggers could find an event

Yes, it is very likely that some TFs will not contain identifiable events, particularly during early running or at lower collision energies. While the maximum collision rate is expected to be around $500\,\mathrm{kHz}$, corresponding to more than 300 hits per TF, the actual rates will be much lower during initial operation and at other energy settings.

All TFs are retained, regardless of event content. The data volume per TF scales with the EIC settings, particularly luminosity and electron and ion beam energies.

Please see also our response to Question 4.

12. **Time-based frames. For our own understanding: do the time-based "frames" overlap with the preceding and following ones? We are asking because, if they are strictly time-sliced, there's a (very small) chance—on the order of 1 in 60,000—that an event could be split between two frames. Has this scenario been considered?**

    No, the TFs do not overlap. This design choice has been explicitly discussed and considered. First, timeframes are grouped into STFs of approximately $1\,\mathrm{s}$. As a result, adjacent (pre- and post-) TFs are stored near each other in the data files and can be combined during processing. Second, the start of a TF is not tied to a fixed bunch crossing. This avoids introducing any potential bias related to spin states in the event that TF joining is not applied. This design ensures a statistically neutral treatment of spin-dependent data.

    > ***Note from the TIM:*** *Events could be split across timeframes, but this will be rare and even rarer with super timeframes. The initial and final parts of a timeframe may be excluded, but care must be taken to avoid introducing bias into the sample.*

13. **Section [3.1]. ATLAS does not use data streaming. We guess the text intends ALICE. We challenge that LHCb thanks to streaming now can publish in weeks, while before it was years**

    The reference to ATLAS is incorrect and should be ALICE, which is an example of a streaming readout experiment. We also acknowledge the point regarding LHCb. The text was written from the perspective of Nuclear Physics, where streaming readout, autonomous experimentation, and heterogeneous computing are expected to enable a turnaround time for physics analysis on the order of weeks, rather than a year. This is a key goal of the ePIC computing model.

14. **Figure 8 is very hard to read, so we can not compare it with Fig 6. But from the text in [3.7] it seems the data into DAM is 10 Tbps while from Fig 6 it reads 1.7 Tbps. May be we misunderstand Figure 6. I assume the RDO data rate is for the input, otherwise it is not coherent with the tape data volume column (input to tape)**

    We will update the diagram to reflect the correct value. The actual number is $2\,\mathrm{TB/s}$, based on simulated data volume tests, and is therefore more reliable than earlier schematic estimates.

    That said, the apparent inconsistency remains. The $100\,\mathrm{Tb/s} \rightarrow 10\,\mathrm{Tb/s} \rightarrow 100\,\mathrm{Gb/s}$ progression was a schematic concept from the Yellow Report era that continues to influence our diagrams, even though it no longer reflects detailed estimates.

    It is important to note that some input parameters are still evolving. For example, the noise levels and ASIC data sizes are not yet fully specified, and synchrotron radiation backgrounds have not yet been included. Considering these uncertainties, a rate of $2\,\mathrm{TB/s}$ is not fundamentally inconsistent with an order-of-magnitude estimate of $10\,\mathrm{Tb/s}$.

15. **The text in [3.7] says most bunch crossing will not result in interesting physics. But they could result in noise that gets read out and timeframed. You will not find interesting events in there, but will you keep those TFs? (same question as above, but now after some more information)**

    Yes.

16. **When referring to the DAQ computing farm in [3.8] do you refer to all the elements on the right hand side of the network switch in Fig 3. (apart from the Echelon-1 of course) ?**

    Yes, the elements referred to are those on the right-hand side of the network switch. This is more clearly illustrated in the updated version of the schematic provided in Fig. 2

17. **In [3.8] there is a broken link to Figure ?? or there is a missing figure. We do not find a figure that matches the text.**

    We overlooked this and will correct the document.

18. **In [3.8] we understand that in both scenarios (DAQ in IP6 and DAQ in B725 enclave) the architecture of the online-offline interface will be the same (Echelon-1 in BNL will be a distinct object from Echelon-0 even if they are co-hosted). Did you consider a hybrid where you leverage the benefits of the co-location?**

    No, a hybrid model was not seriously considered. TF assembly is a significantly different task from data processing, and it requires tight integration with the DAM board hardware. This level of hardware-specific interaction makes co-locating DAQ and data processing tasks impractical within the same echelon.

19. **Do you have enough floor space/power in IP6 for the full online system?**

    There is barely enough space available at IP6 to accommodate the full E0 system. However, considerations such as cooling requirements, future expansion, and the infrastructure needed to support a computing farm are significant factors. These considerations support the idea of splitting the E0 systems, rather than concentrating all components in a single location.

20. **We guess the IP6 option closer to the detector is presumably preferred—assuming space constraints are met. Is there an intermediate solution being considered, such as a surface-level room (like some LHC experiments use) to house online computing equipment?**

    All intermediate solutions under consideration are surface-level options. The guiding idea is to leverage existing physical infrastructure wherever possible, rather than constructing new underground or custom-built spaces.

21. **In [3.9] you mention the condition and calibration databases. Where do you intend to deploy those databases and will you have online databases and offline databases ? If the condition and calibration databases have to be accessed for reconstruction and calibration (which is an offline task) they will need to be accessible by the online resources (at least Echelon-1s).**

    While the deployment strategy for the conditions and calibration database is still to be determined, we can anticipate what it will involve, particularly given our aim for the EIC to be a tightly integrated facility from machine to detector/DAQ to processing and analysis. We will have online data sources (and consumers) for conditions/calibration information, so we must have an online database. We must have an offline database as you note. And we will need mechanisms to selectively migrate/transform information from online to offline in order to provide the data needed offline.

    > ***Note from the TIM:*** *Both online and offline databases are planned, along with a replication mechanism between them.*

22. **In [3.11] you foresee having a buffer depth of 1 week. This is about 7.5 PB of storage. It seems large as at least 1 echelon-1 is located in the same lab as the echelon-0. What are the arguments?**

    This is a known inconsistency that we have accepted up to this point. The concept of requiring $1\,\mathrm{s}$ (or even $1\,\mathrm{min}$) latency to the E1 sites is indeed inconsistent with the need for a large buffer at E0.

The underlying reason is that E0 was planned early, as part of the EIC Project scope, before the full concept for E1 data transfers was in place. As a result, the buffer size at E-0 was specified conservatively.

In addition, we currently do not have firm guarantees on the uptime or reliability of the E1 facilities and the intermediate network connections. Therefore, the buffer size at E0 is likely larger than necessary, but it provides a safety margin. We expect that the buffer size can be substantially reduced as the system architecture and service level expectations are better defined. We will qualify this statement in our revised schematics and document.

> ***Note from the TIM:*** *E0 has a buffer depth of 72 hours, not one week. This will be corrected.*

23. **Data reduction for the RICH (PID) subsystem seems to be crucial to stay within the bandwidth budget. We couldn't find in the document where this reduction will be performed—presumably in the Readout Computer. Will the available processing resources be sufficient to keep up with the acquisition rate? Has the reduction algorithm been developed and its computational cost estimated?**

Current R&D on data reduction for dual-radiator ring imaging Cherenkov detector (dRICH) subsystem is being conducted by INFN. The effort focuses on an Apieron/FPGA-based ML application designed to run directly on DAM boards dedicated to ML-based processing, using the local information from the dRICH only.

As a fallback, there is also a backup plan based on more conventional trigger concepts, which would combine data from other detectors, e.g., the LFHCAL, to determine whether there are likely hits in dRICH that warrant further processing.

> ***Note from the TIM:*** *"Data reduction for the RICH has been studied in detail. It will take place on the DAM board, where sufficient computational capacity can be allocated to handle the task.*

The Computing Use Cases

24. **In [4.1] (and after) we think it would be good to have a diagram that explains for different use cases tha data flow with data rates, etc.. Can we get something along these lines?**

We are not sure what you have in mind. Could you please provide an example diagram, and we will include it in the next review?

25. **Section [4.1] and below. When does the Echelon-0 have the green light to flush data on the buffer? E.g. if there is a disk copy at Echelon 1, if there is a tape copy at Echelon-1, if there are two tape copies at both Echelon-1 ... This has an impact on the buffer sizes and the level of sophistication of the agents responsible for transfers and archive.**

The timing of when E0 is allowed to flush data from its buffer depends somewhat on the specific transfer mechanism used. However, in general, data will likely be released once a successful status (i.e., a "good" error code) is returned from the transfer to E1. Whether a single E1 or both remains to be decided. Current discussion is to use Rucio to manage the E0-E1 transfers, giving us prompt knowledge of replica availability at E1s, so the E0 deletion policy can be applied and easily adjusted.

26. **Section [4.4] If we understand correctly, the prompt reconstruction produces events from timeframes. Are the events still "RAW" after prompt reconstruction, in the sense that they are not in the form of physics objects? Maybe another way to phrase this: are "prompt reconstruction" and "first pass reconstruction" the same things in terms of input and output data formats? And, following from that, is the output format of the two ready for analysis?**

The input and output formats of prompt reconstruction and first pass reconstruction are the same, and the output can be used directly for physics analysis. We aim to provide fast feedback on the status of the experiment and data taking, including physics observables such as DIS kinematics.

> ***Note from the TIM:*** *Prompt reconstruction is primarily for data quality monitoring and is performed on a subset of the timeframes.*

27. **Section [4.6]. Reprocessing seems to be a more relaxed activity wrt first pass reconstruction and in fact can run everywhere, including opportunistic resources even for full reprocessing. Why is that the case? Analyses for publications will run very likely on reprocessed data, particularly for precision measurements.**

 We aim to make the first pass reconstruction as accurate as possible, supported by an autonomous alignment, calibration, and validation scheme. The output of the first pass can be used for physics analysis and paper preparation.

 In particular, during the early years of operation, we expect to learn and improve our approaches to alignment, calibration, reconstruction, and validation. As a result, reprocessing will be necessary, but it will occur on a more relaxed timescale compared to the urgency of delivering the first pass results to analyzers.

> ***Note from the TIM:*** *Note that there is a difference between the initial years of data taking, when detector commissioning and early physics will be of primary importance, and the later years.*

28. **Section [4.7]. we do not fully understand the last part of the section (technical and sociological considerations). If you apply new sw algorithms in reconstruction of real data, I guess you want to apply the same algorithms to reconstruct the simulated data, no?**

 We agree that the same reconstruction algorithms should be applied to both simulation and real data to ensure consistency in analysis.

 However, during commissioning and the early stages of operation, there may be periods of rapid development in reconstruction and related algorithms. In such cases, updates may be applied immediately to prompt and first pass reconstruction of real data, but not immediately reflected in the bulk simulations, as we would not regenerate large simulation datasets for every intermediate update.

> ***Note from the TIM:*** *Note that there is a difference between the initial years of data taking, when detector commissioning and early physics will be of primary importance, and the later years.*

29. **Frame-based Simulation vs Event-based [4.7]. You propose simulating “frames” instead of discrete events to match the real-data reconstruction workflow. While this offers consistency, could the complexity outweigh the benefits? Generating frames requires overlaying many events (long processing time in Geant4) and including various machine backgrounds, which may be hard to model realistically. If the splitting of frames into events happens early in reconstruction anyway, might it be more practical to simulate individual events directly and avoid this overhead?**

 Most simulations will be event-based, for the reasons outlined in the question. To accommodate this, we are developing our simulation and reconstruction software to support switching between event-level and TF-level modes as needed.

 However, for the development of TF-based reconstruction and certain detector studies, we will also require TF-based simulation to closely match the real data flow, to evaluate reconstruction performance, and to determine related systematic uncertainties under realistic conditions. TF-based simulations are expected to account for a smaller, targeted fraction of the total compute resources allocated for simulations.

> ***Note from the TIM:*** *Eventually, both timeframe-based and event-based simulation will be needed. Timeframe-based simulation will be required in smaller volume compared to event-based simulation.*

30. **Use Echelon2 for Analysis [4.8]. What is the motivation for using Echelon 2 resources for analysis as well? Wouldn't this complicate the infrastructure and software environment, since analysis has different characteristics: it is more interactive, less predictable, may require more varied and complex software stacks, and often benefits from fast cache storage. Would it make more sense to build dedicated facilities optimized for analysis instead of trying to generalize Echelon 2 for all use cases?**

    Analysis prototyping and the final stages of analysis, such as studying histograms and other data representations, will take place at E3. However, time-consuming parts of analysis such as iterative procedures, fitting of high-dimensional datasets, or training and validation of ML models will be executed at E1 and E2 sites.

    Many of these tasks are conceptually similar to reconstruction, as they involve applying algorithms across large sets of physics events. At this stage, we do not plan to build dedicated analysis facilities. Instead, our plan is to ensure that the E1 and E2 sites support the required analysis infrastructure.

    > ***Note from the TIM:*** *There is still limited understanding of the analysis requirements. At present, the plan is for E1 and E2 facilities to also support user analysis. This is feasible if analysis is considered from the outset as a key use case in the design of E1 and E2. Regarding input/output formats used by different processing workflows, there is not yet a mature understanding of the analysis needs to determine whether 'reduced' analysis formats are a viable solution, at least for the early years of data taking.*

31. **Digital twin development [4.9]. Has there been an evaluation of the potential benefits of building a digital twin, relative to the significant effort required to develop and maintain it?**

    No. We list the digital twin as a potential use case to optimize the operation of the ePIC experiment. At this stage, our goal is simply to ensure that we retain the option to deploy a digital twin in the future. No formal evaluation or development has been undertaken beyond that point.

Computing Resources

32. **In section [5.3.1] it mentions that Echelon-2 should also do prompt processing. This is a bit in contrast with section [4]. Also, this will mean Echelon-1 will need an output bandwidth of 1/6th of the RAW stream. Where does the number 6 come from?**

    We acknowledge that this point needs to be clarified in the document. Prompt processing will be limited to the E1 sites. The extent to which raw data will be processed at E2 sites remains an open question.

    In the initial version of the document, we assumed that the E1 sites would each process approximately one-third of the total raw data, for a total of two-thirds, while the E2 sites would process the remaining one-third. This model would imply that the two E1 sites would need to distribute one-third of the raw data, or one-sixth each, to the E2 sites.

33. **RAW data archival [5.3.2]. The handling of RAW data is also a bit unclear. The document mentions that RAW data will only be deleted once the reconstruction artifacts are complete and stored, but it's not explicitly stated whether a full set of RAW data is archived permanently. It may help to clearly lay out the storage strategy: what data products are stored, for how long, and where (e.g., temporary, disk, tape, etc.).**

    We plan to permanently store two copies of the raw data, i.e. the TF-based data from E0. However, in the later stages of the experiment, once we have developed a deep understanding of the data, including its background characteristics, we may consider discarding portions of the data that are not associated with physics events.

34. **Echelon-2 processing. It's not clear whether both Echelon 2 sites will process all frames, or if they will share the load. The document suggests the goal is to maintain two complete data copies, but if both sites are fully reconstructing the same data, the computing cost is effectively doubled for the first reconstruction pass. Could this be clarified?**

    We plan to maintain two full copies of the raw data at the E1 sites. Data processing will be distributed between the E1 and E2 sites. The E1 sites will share the load (not duplicate the processing). It will be up to the E2 sites to decide whether to store raw data beyond the fraction needed for their portion of the processing.

    From preliminary discussions, we understand that some countries intend to store a full copy of the raw data as part of their local policy requirements for contributed infrastructure. This copy will not necessarily be streamed continuously to those countries but may instead be batched and retrieved from E1 sites outside of data-taking periods.

    In contrast, physics objects will be distributed globally to support efficient access for analysis across the collaboration.

    > ***Note from the TIM:*** *This point is related to the item above, but also applies more broadly. The estimates are based on a conservative computing model for 2034. 'Conservative' in this context means that more information and data are retained until there is sufficient confidence in the computing model and understanding of the detector. At later stages, some optimizations will be possible, for example by reducing or eliminating timeframes that contain no useful information.*

35. **Section [5.3.1]. Why do you need monitoring, and slow control data at Echelon-2? Even in the case of running reconstruction, you need conditions data, not the full slow control time series. Also, why in Table-3 monitoring, calibration and slow controls are incoming from the other Echelon-1 and Echelon-2's? The slow control information is collected at the detector ..**

    Indeed, we need to clarify that it is conditions data, not slow control data, that is required at E2 for reconstruction. Calibration and conditions information must be shared between E1 and E2 sites for distributed data processing.

36. **Fig.10. Are the numbers for disk or tape? Streaming data (timeframes) are 70PB/year and they go to each Echelon-1, but they remain on disk for 3 weeks (the time to reconstruct with up-to-date calibrations). I guess you do not need to keep the full 70 PB of disk. On the same figure, why does 35PB of streaming data also need to be at Echelon-2? I guess it is to support first pass processing at Echelon-2s, but for that you would need only a small buffer of the data you want to process at a given point in time. 35PB is roughly the volume that each of the two Echelon-1s will first pass process in one year (70PB divided by two Echelon-1s)**

    First, please refer to our answer for Question 34. We leave it to the E2 sites to decide how much of the raw data they will store. The E2 sites could store only a sliding window of the data served from E1s, reducing their storage need. Distributed among all E2 sites, they will need to store at least one full copy of the raw data, which corresponds to approximately $35\,\mathrm{PB}$. You make a good point that the E2s could store only a sliding window of the data served from E1s, reducing their storage need.

    The disk buffer is intended to support autonomous alignment, calibration, and validation of the reconstructed physics objects. This buffer must cover at least three weeks worth of raw data, but the exact requirements still need to be evaluated.

37. **Fig.11 (and related text in [5.3.3]. we do not understand why alignment and calibration need to happen also at Echelon-2s. They are complex workflows, they likely require dedicated input streams, they need to upload back constant at Echelon-1s condition databases, etc .. They require 2k cores out of**

150k, which is on the order of 1%. We guess BNL and JLAB can provide these 2k cores. What are we missing?

We link the roles of E2 resources to the responsibilities of participating countries within the collaboration. For example, Italy is responsible for the dual-radiator Ring Imaging Cherenkov (dRICH) detector for ePIC and will play a significant role in studies of transverse-momentum dependent (TMD) PDFs for quark-gluon imaging.

As part of this role, Italian groups will be involved in validating and approving calibrations for the dRICH, as well as running simulations and analyses related to TMD PDFs. These activities will be supported by and reflected in their E2 contributions to the computing infrastructure.

Distributed Computing

Software

38. **Section 7.1. There is large repetition in the section (bad cut and paste we guess). The Round Tables and the Workshop are repeated twice. BTW, a pity that the S&C Round Tables are not continued.**

    We apologize for the low quality of the text in this section and will work to improve it in future versions of the document. We also regret the discontinuation of the Software & Computing Round Table. At this stage, it is unclear when or if it will be revived.

39. **Programming Languages. There's no mention of programming languages in the document. We assume C++ is the primary language for simulation, reconstruction, and analysis, but it would be good to explicitly state this. Will you continue with the C++/Python model, or are you considering evaluating other languages such as Julia or some level of language interoperability?**

    We are planning to continue with the C++ and Python model. While we follow developments in JuliaHEP with interest, our current focus is on training the community in Python, as it is widely used in the data science community, where many relevant developments in data analysis and statistical methods are occurring. Python also benefits from strong support in university curricula.

    While modern C++ remains challenging for much of the collaboration, our developer base is trained in it and is growing slowly.

    > ***Note from the TIM:*** *At the moment, there is no concrete need to consider more heterogeneous architectures such as GPUs. However, there is some interest in using GPUs for simulation, particularly for optical photons.*

Serving Users

40. **JANA2 Agnostic algorithms [8.1]. We were surprised to read that developers can write reconstruction algorithms completely agnostic of the JANA2 framework. While we understand that the main purpose of these algorithms is to transform data objects in EDM4hep format into other objects, thus independent of JANA2, it seems unlikely that they could be fully decoupled from the framework. In practice, algorithms often need access to configuration parameters, detector conditions, geometry, and other shared services. All of this typically requires interaction with the framework. Could you clarify how this separation is achieved in JANA2, and whether there are limitations or trade-offs in this approach?**

    We have adopted the JANA2 framework to manage our reconstruction algorithms, and it has since evolved substantially in response to our needs. The design has been shaped by three main drivers: the need to integrate cleanly with PODIO-based data models and other components of the key4hep stack, the ability to externally configure existing components, and support for TF splitting required by the streaming readout model.

The algorithms that our developers write must satisfy an abstract algorithm interface model which, in 1400 lines of code, only depends on select external components: DD4hep, podio, EDM4hep, EDM4eic. The underlying philosophy behind this choice is to take a declarative approach to algorithm specification.

The algorithm interface requires algorithms to specify a specific signature for the constructor, and for the `init` and `process` functions. The algorithm interface also defines how algorithms can declare which properties they contain and which services they use. The algorithm interface itself does not implement any service itself, and JANA2-specific code is responsible for providing the correct data collections to the `process` calls, and for connecting command-line or steering-file specified properties and actual JANA2 services to the interface handles.

In addition to removing the need for users to become familiar with JANA2 when not needed, an added benefit is that the algorithms interface enforces modularity and imposes good practices (const-correctness) which aid in keeping the algorithms suitable for concurrent running.

General Note from the TIM: *What happens in E0 when data is not being collected has not yet been discussed. However, E0 is relatively small, consisting of approximately 100 computers directly connected to the boards and another 100 as part of the DAQ farm. These machines are specifically configured for DAQ purposes. There are currently no plans to use them for offline use cases, but this may be considered as an option in the future.*