ePIC Streaming Computing Model Working Group
Meeting - Aug 22, 2023

Short recap of what has been discussed so far

I) WG presentation: people, work plans, organization …

ePIC Streaming Computing Model Working Group
Kick-off meeting - July 11, 2023

Software and Computing Coordinator (Markus)
+ Deputy Coordinator **Operations (Wouter)**
+ Deputy Coordinator **Development (Sylvester)**
+ Deputy Coordinator **Infrastructure (Torre)**
Guiding Principles: DE&I, Software Principles, Sustainability

Operation WGs:
- **Production (CD)**
- **User Learning**
- **Validation (CD)**

Development WGs (CI):
- **Physics and Detector Simulation**
- **Reconstruction**
- Analysis Tools

Infrastructure WGs:
- **Streaming Computing Model**
- Multi-Architecture Computing
- Distributed Computing

Cross-cutting WG:
- Data and Analysis Preservation

- Structure of Software and Computing in ePICS
- Presented by Markus on May 2nd 2023 kickoff-meeting

- ePIC SRO Computing Model WG belongs to 'infrastructure'
- Co-conveners: M.Battaglieri (INFN), J.Huang (BNL) + J.Landgraf (BNL)

Marco Battaglieri
Co-convener

Jin Huang
Co-convener

Jeffery Landgraf
Co-convener for electronics & DAQ WG
Kindly helping SRO group organization during next months when Jin focuses on ePHENIX commissioning

Infrastructure

- **Streaming Computing Model (**Marco Battaglieri, Jin Huang, Jeff Landgraf (interim)**)**:
  - **Charge**: Development of the computing model for the compute-detector integration using streaming readout, AI/ML, and multi-architecture computing (CPU, GPU, …) with a specific focus on the data flows after the FEE layer.
  - **Priorities for 2023**:
    - Establish a collaborative dialogue with the Electronics and DAQ WG.
    - Define the requirements and high-level design for a computing model that enables rapid processing for the data for physics analyses, while leveraging external compute resources, including those provided by international partners.
    - Coordinate activities on prototyping streaming computing systems, together with the Physics and Detector Simulation and Reconstruction Framework and Algorithms WGs.
    - Document a streaming computing model that can be redefined further with international partners.

Work plan:
- Define a draft of SRO Computing Model by end of September
- Discuss it during the fall
- Present the ePIC Computing Model at the EIC Resource Review Board meeting scheduled for December 7 and 8 2023

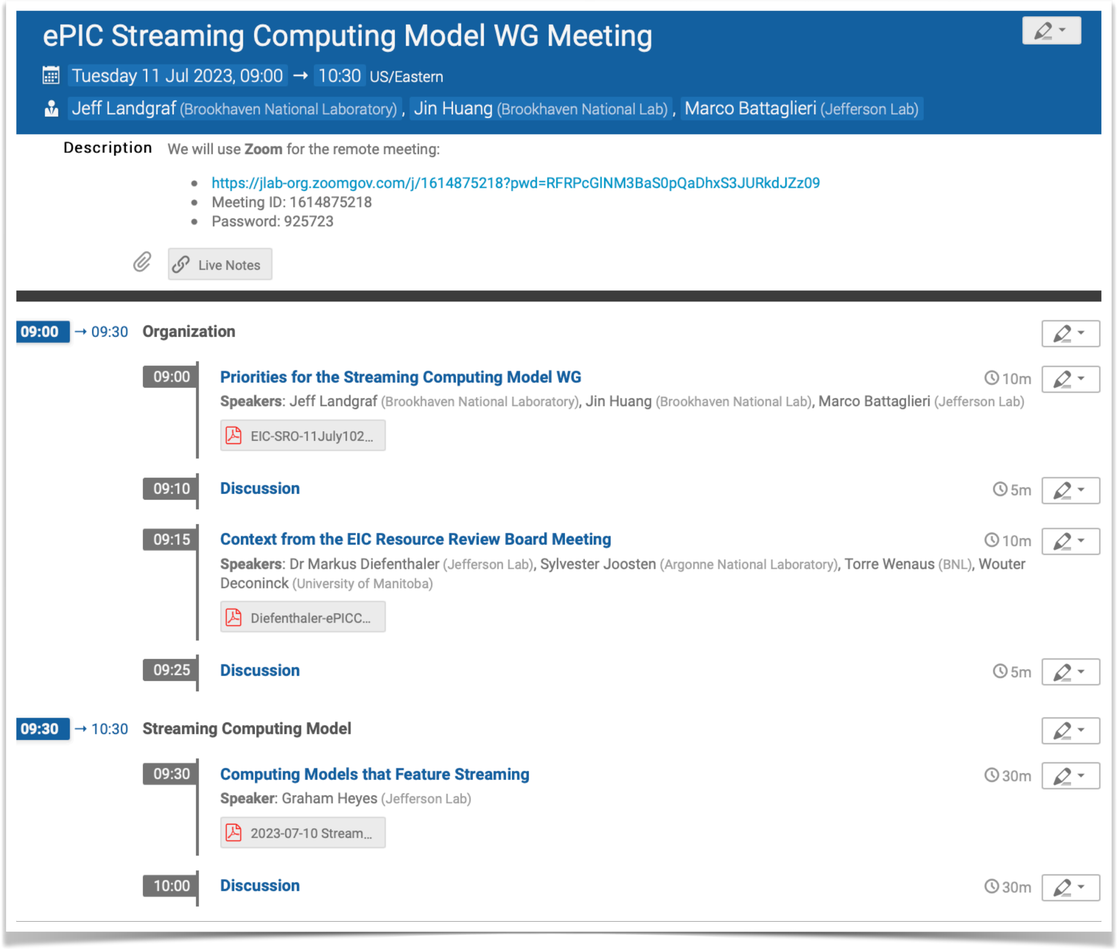
Work Group organization:
- Two meetings before EICUG/ePIC meeting in Warsaw meeting (July 11 and 18)
- Bi-weekly meetings starting from August with contributions from experts (general, EIC, BNL, JLab and foreign institutions CERN, FAIR, KEK, …) and other ePICs WG (DAQ, SW, Physics, …)
- From september focus the discussion on ePIC SERO computing model

Regular weekly meeting on Tuesday at 9:00 AM EDT

ePIC Streaming Computing Model Working Group

Meeting - Aug 22, 2023

II) Computing model for HEP/NP physics



ePIC Streaming Computing Model WG Meeting

Tuesday 11 Jul 2023, 09:00 → 10:30 US/Eastern

Jeff Landgraf (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Marco Battaglieri (Jefferson Lab)

Description We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

Live Notes

09:00 → 09:30 **Organization**

09:00 **Priorities for the Streaming Computing Model WG** 10m
Speakers: Jeff Landgraf (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Marco Battaglieri (Jefferson Lab)
EIC-SRO-11July102...

09:10 **Discussion** 5m

09:15 **Context from the EIC Resource Review Board Meeting** 10m
Speakers: Dr Markus Diefenthaler (Jefferson Lab), Sylvester Joosten (Argonne National Laboratory), Torre Wenaus (BNL), Wouter Deconinck (University of Manitoba)
Diefenthaler-ePICC...

09:25 **Discussion** 5m

09:30 → 10:30 **Streaming Computing Model**

09:30 **Computing Models that Feature Streaming** 30m
Speaker: Graham Heyes (Jefferson Lab)
2023-07-10 Stream...

10:00 **Discussion** 30m

- Real-time data processing is a general issue/opportunity (with commercial/ custom solutions already available!)
- SRO is an opportunity to simplify data acquisition
- Need of flexibility to cope with technology evolution
- Every DAQ is a streaming system!
- Should SRO include high-level analysis too or just stop at data collection?
- Row data and 'events': where to place the separation line? physics?

What is this talk about?

- This is not a proposal for a computing model for ePIC. It is a short look at:
  - What streaming data is in the broadest sense
  - The opportunities and challenges for processing such data
- Outline
  - Definition of streaming data
  - The current dominant data processing model
  - The need for a new approach
  - Data Streaming as a disruptive model
  - Science use cases
  - Computing models
  - Hardware architectures
  - Data steering
  - Summary

2 Jefferson Lab

G.Heyes: Computing models for future streaming

Need to know what we (EIC) already did and what others are doing ….

ePIC Streaming Computing Model Working Group

Meeting - Aug 22, 2023

III) Computing model for ePIC

ePIC Streaming Computing Model WG Meeting

Tuesday 18 Jul 2023, 09:00 → 10:40 US/Eastern

Jeff Landgraf (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Marco Battaglieri (Jefferson Lab)

Description We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

Live Notes

09:00 → 09:10 Kick-Off Meeting Summary

- 09:00 **Key Points and Action Items from July 11** — 5m
  Speaker: Marco Battaglieri (Jefferson Lab)
  EIC-SRO-18July203...
- 09:05 **Discussion** — 5m

09:10 → 10:40 Streaming Computing Model

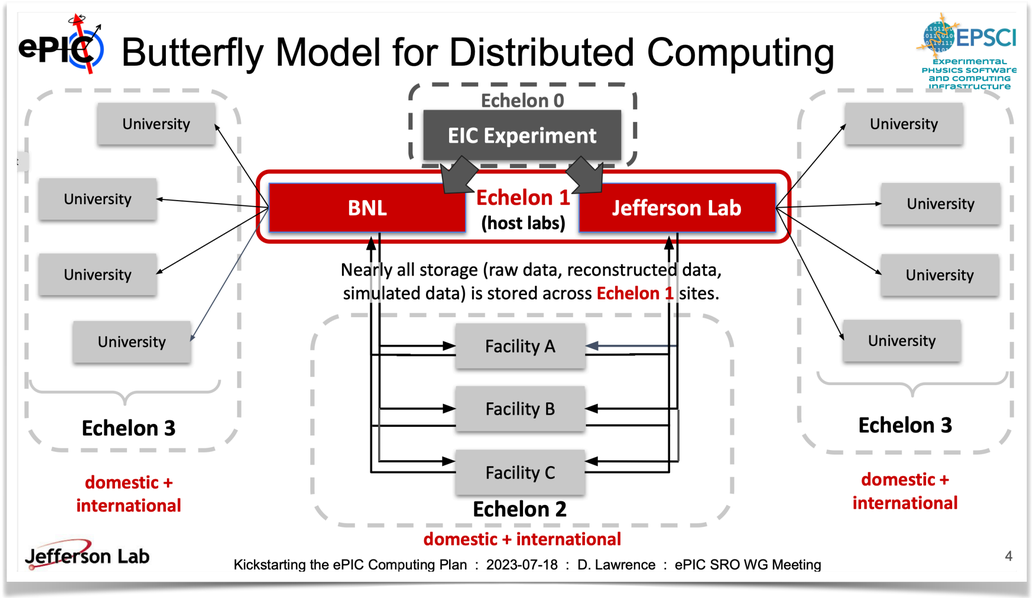
- 09:10 **International Perspective** — 15m
  Speaker: Andrea Bressan (University of Trieste and INFN)
  INFN_Computing_I...
- 09:25 **Discussion** — 15m
- 09:40 **Computing Models for the EIC** — 30m
  Speaker: David Lawrence (Jefferson Lab)
  2023.07.18.ePIC_C... Google Slides
- 10:10 **Discussion** — 30m

INFN (@CERN) Perspective

- Structured in TIERS (TIER0@CERN)
- Several Tier2 distributed in various locations + TIER1(ICSC)
- LHC-HL specs: 350 PB/y (raw data); ~400+100 Gbps to Tier1s (quasi-real time)
- T1 to T2 probably by tapes

Computing model for EIC

- Experience gained with ECCE, ATHENA and the EIC YR
- Butterfly model



D.Lawrence: Computing model for EIC

ePIC Streaming Computing Model Working Group

Meeting - Aug 22, 2023

III) Computing model for ePIC

ePIC Streaming Computing Model WG Meeting

Tuesday 18 Jul 2023, 09:00 → 10:40 US/Eastern

Jeff Landgraf (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Marco Battaglieri (Jefferson Lab)

Description We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

Live Notes

09:00 → 09:10 Kick-Off Meeting Summary

- 09:00 **Key Points and Action Items from July 11** — 5m
  Speaker: Marco Battaglieri (Jefferson Lab)
  EIC-SRO-18July203...
- 09:05 **Discussion** — 5m

09:10 → 10:40 Streaming Computing Model

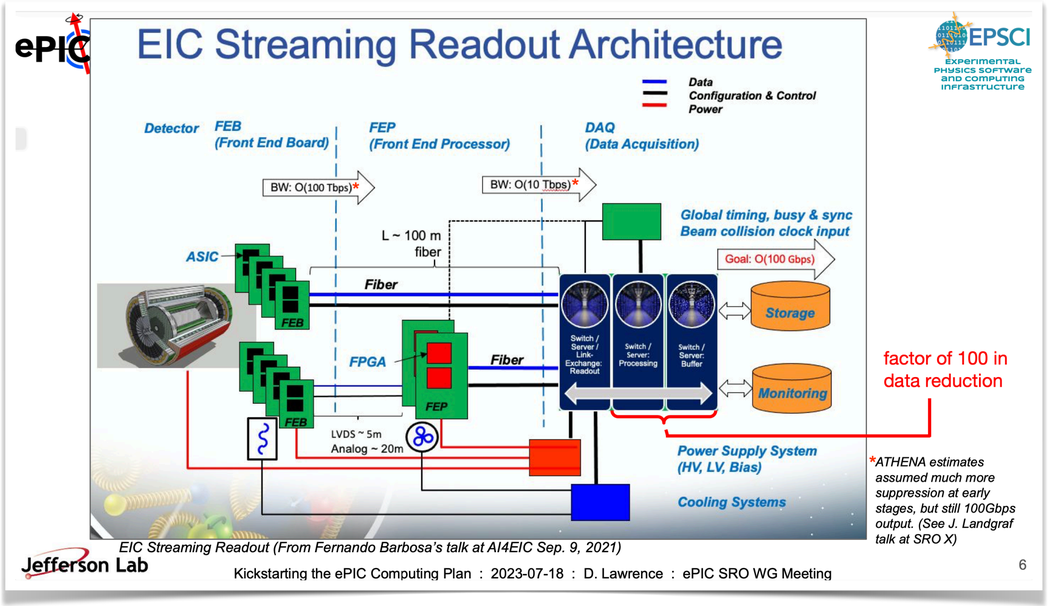
- 09:10 **International Perspective** — 15m
  Speaker: Andrea Bressan (University of Trieste and INFN)
  INFN_Computing_I...
- 09:25 **Discussion** — 15m
- 09:40 **Computing Models for the EIC** — 30m
  Speaker: David Lawrence (Jefferson Lab)
  2023.07.18.ePIC_C... Google Slides
- 10:10 **Discussion** — 30m

INFN (@CERN) Perspective

- Structured in TIERS (TIER0@CERN)
- Several Tier2 distributed in various locations + TIER1(ICSC)
- LHC-HL specs: 350 PB/y (raw data); ~400+100 Gbps to Tier1s (quasi-real time)
- T1 to T2 probably by tapes

Computing model for EIC

- Experience gained with ECCE, ATHENA and the EIC YR
- Butterfly model
- ePIC Streaming readout architecture
- estimates for row data size needed to size: CPUs, storage, networking
- Challenges/opportunities: calibrations, AI, ...



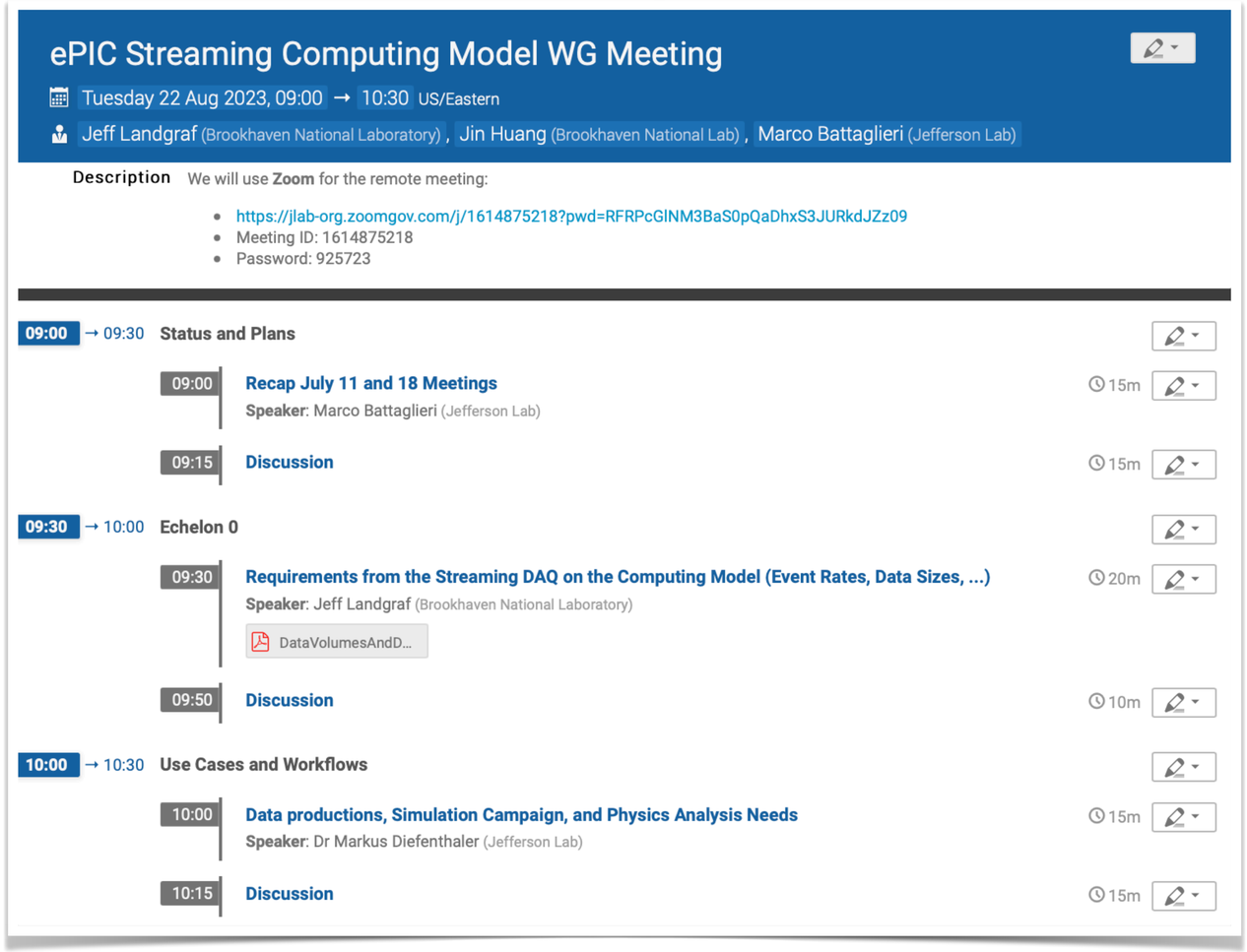
D.Lawrence: Computing model for EIC

Need to collect in formation about needs and specs!

ePIC Streaming Computing Model Working Group

Meeting - Aug 22, 2023

Today: ePIC requirements (aka Echelon-0 rates)



ePIC Streaming Computing Model WG Meeting

Tuesday 22 Aug 2023, 09:00 → 10:30 US/Eastern

Jeff Landgraf (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Marco Battaglieri (Jefferson Lab)

Description We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

09:00 → 09:30 Status and Plans

- 09:00 **Recap July 11 and 18 Meetings** — 15m
  Speaker: Marco Battaglieri (Jefferson Lab)
- 09:15 **Discussion** — 15m

09:30 → 10:00 Echelon 0

- 09:30 **Requirements from the Streaming DAQ on the Computing Model (Event Rates, Data Sizes, ...)** — 20m
  Speaker: Jeff Landgraf (Brookhaven National Laboratory)
  DataVolumesAndD...
- 09:50 **Discussion** — 10m

10:00 → 10:30 Use Cases and Workflows

- 10:00 **Data productions, Simulation Campaign, and Physics Analysis Needs** — 15m
  Speaker: Dr Markus Diefenthaler (Jefferson Lab)
- 10:15 **Discussion** — 15m

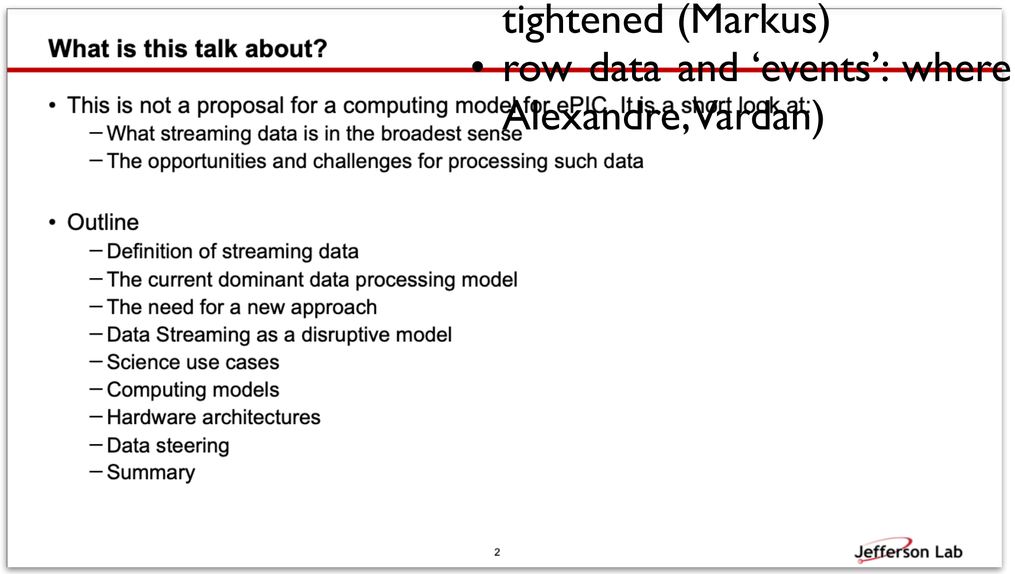
Before October 19-20 review:

- Use cases and workflows (reconstruction, storage, distribution, simulations, high-level analysis and role of Echelon-X): requirements from Physics WGs
- Alignment and calibrations and near real-time data processing

… more in next Tuesday's meetings!

Short recap of what has been disc used so far

- General talk about computing models (not only for hep/np physics) (Graham)
- Real-time data processing is a general issue (with commercial/custom solutions already available!) (Graham)
- SRO is an opportunity to simplify data acquisition (Graham)
- Flexibility to cope with technology evolution (Graham)
- Every DAQ is a streaming system! (Jeff)
- Streaming as a model for the whole analysis pipeline? (Pietro)
- The connection between Physics WGs and Computing Models needs to be tightened (Markus)
- row data and 'events': where to place the separation line? physics? (Jan, Markus, Alexandre,Vardan)
- Butterfly model: detector computing hand in hand (Markus D)
- The computing model affects the role of international partners (Pietro A)
- SRO to speed up physics analysis (Markus D)



What is this talk about?

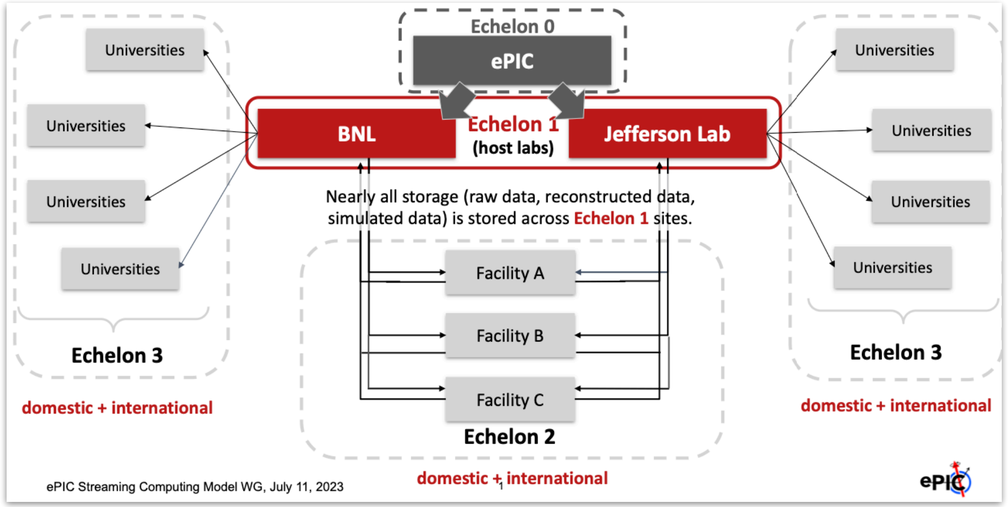
- This is not a proposal for a computing model for ePIC. It is a short look at:
  - What streaming data is in the broadest sense
  - The opportunities and challenges for processing such data
- Outline
  - Definition of streaming data
  - The current dominant data processing model
  - The need for a new approach
  - Data Streaming as a disruptive model
  - Science use cases
  - Computing models
  - Hardware architectures
  - Data steering
  - Summary

2 Jefferson Lab

- General talk about computing models (not only for hep/np physics) (Graham)
- Real-time data processing is a general issue (with commercial/custom solutions already available!) (Graham)
- SRO is an opportunity to simplify data acquisition (Graham)
- Flexibility to cope with technology evolution (Graham)
- Every DAQ is a streaming system! (Jeff)
- Streaming as a model for the whole analysis pipeline? (Pietro)
- The connection between Physics WGs and Computing Models needs to be tightened (Markus)
- row data and 'events': where to place the separation line? physics? (Jan, Markus, Alexandre,Vardan)

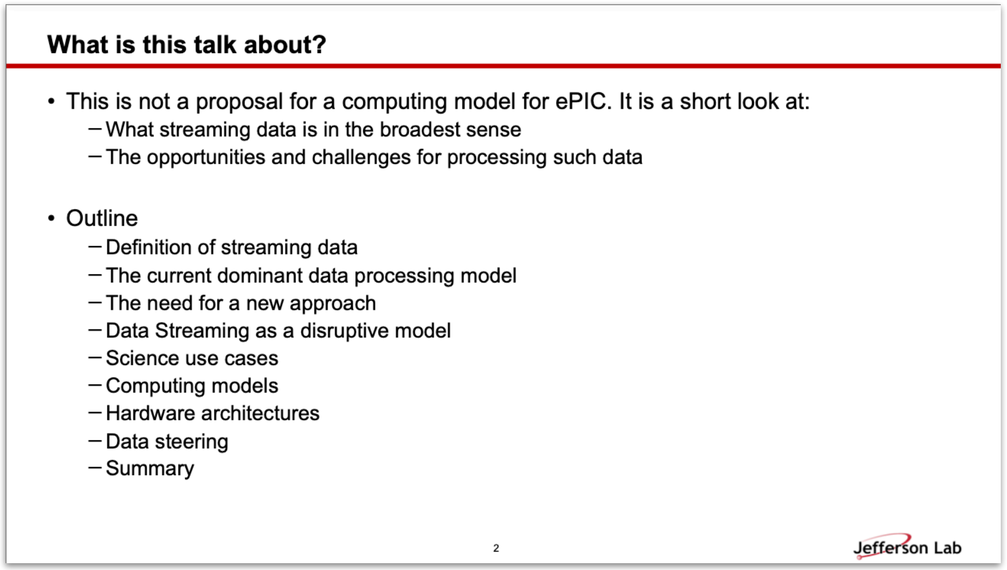
ePIC Streaming Computing Model Working Group

Meeting - Aug 22, 2023



- Presentation of the ePIC SCM WG

- Butterfly model: detector computing hand in hand (Markus D)
- The computing model affects the role of international partners (Pietro A)
- SRO to speed up physics analysis (Markus D)



- General talk about computing models (not only for hep/np physics) (Graham)
- Real-time data processing is a general issue (with commercial/custom solutions already available!) (Graham)
- SRO is an opportunity to simplify data acquisition (Graham)
- Flexibility to cope with technology evolution (Graham)
- Every DAQ is a streaming system! (Jeff)
- Streaming as a model for the whole analysis pipeline? (Pietro)
- The connection between Physics WGs and Computing Models needs to be tightened (Markus)
- row data and 'events': where to place the separation line? physics? (Jan, Markus, Alexandre, Vardan)