EIC Project Detector-1

"We need a real name"

EIC ~~ECCE~~ Computing Model

or

"Seamless Data Processing from DAQ to Analysis"

David Lawrence - JLab

(and a bunch of other people)

Workshop X on
Streaming Readout
May 17-19, 2022

EIC Project Detector-1
"We need a real name"

ECCE Computing Plan Document

- Outlines preliminary computing plan
- Many model concepts aimed at EIC and are not necessarily ECCE specific
  - *similis* WLCG for all LHC experiments
- Review period complete for submission to NIM
  - https://github.com/ecce-notes/ecce-note-comp-2021-01/releases/tag/v2.1

Scientific Computing Plan for the ECCE Detector at the Electron Ion Collider

J. C. Bernauer[55,56,57], C. T. Dean[32], C. Fanelli[37], J. Huang[4], K. Kauder[4], D. Lawrence[28], J.D. Osborn[49], C. Paus[37], J. K. Adkins[36], Y. Akiba[53], A. Albataineh[69], M. Amaryan[47], I. C. Arsene[73], C. Ayerbe Gayoso[38], J. Bae[61], X. Bai[78], M.D. Baker[4,28], M. Bashkanov[87], R. Bellwied[67], F. Benmokhtar[14], V. Berdnikov[12], F. Bock[49], W. Boeglin[16], M. Borysova[83], E. Brash[10], P. Brindza[28], W. J. Briscoe[20], M. Brooks[32], S. Bueltmann[47], M. H. S. Bukhari[27], A. Bylinkin[69], R. Capobianco[66], W.-C. Chang[2], Y. Cheon[59], K. Chen[7], K.-F. Chen[46], K.-Y. Cheng[40], M. Chiu[4], T. Chujo[76], Z. Citron[1], E. Cline[55,56], E. Cohen[44], T. Cormier[49], Y. Corrales Morales[32], C. Cotton[78], J. Crafts[12], C. Crawford[70], S. Creekmore[49], C.Cuevas[28], J. Cunningham[49], G. David[4], M. Demarteau[49], S. Diehl[66], N. Doshita[85], R. Dupré[23], J. M. Durham[32], R. Dzhygadlo[19], R. Ehlers[49], L. El Fassi[38], A. Emmert[78], R. Ent[28], R. Fatemi[70], S. Fegan[87], M. Finger[8], M. Finger Jr.[8], J. Frantz[48], M. Friedman[22], I. Friscic[88], D. Gangadharan[67], S. Gardner[18], K. Gates[18], F. Geurts[52], R. Gilman[54], D. Glazier[18], E. Glimos[49], Y. Goto[53], N. Grau[3], S. V. Greene[79], A. Q. Guo[25], L. Guo[16], S. K. Ha[86], J. Haggerty[4], T. Hayward[66], X. He[17], O. Hen[37], D. W. Higinbotham[28], M. Hoballah[23], T. Horn[12], A. Hoghmrtsyan[1], P.-h. J. Hsu[45], G. Huber[74], A. Hutson[67], K. Y. Hwang[87], C. Hyde[47], M. Inaba[64], T. Iwata[85], H.S. Jo[31], K. Joo[66], N. Kalantarians[81], G. Kalicy[12], K. Kawade[60], S. J. D. Kay[74], A. Kim[66], B. Kim[61], C. Kim[51], M. Kim[53], Y. Kim[51], Y. Kim[59], E. Kistenev[4], V. Klimenko[66], S. H. Ko[58], I. Korover[37], W. Korsch[70], G. Krintiras[69], S. Kuhn[47], C.-M. Kuo[40], T. Kutz[37], J. Lajoie[26], S. Lebedev[26], H. Lee[61], J. S. H. Lee[75], S. W. Lee[31], Y.-J. Lee[37], W. Li[52], W. Li[55,56,84], X. Li[9], X. Li[32], Y. T. Liang[25], S. Lim[51], C.-h. Lin[2], D. X. Lin[25], K. Liu[32], M. X. Liu[32], K. Livingston[18], N. Liyanage[78], W.J. Llope[82], C. Loizides[49], E. Long[72], R.-S. Lu[46], Z. Lu[9], W. Lynch[87], D. Marchand[23], M. Marcisovsky[13], P. Markowitz[16], H. Marukyan[1], P. McGaughey[32], M. Mihovilovic[71], R. G. Milner[37], A. Milov[83], Y. Miyachi[85], A. Mkrtchyan[1], P. Monaghan[10], R. Montgomery[18], D. Morrison[4], A. Movsisyan[1], H. Mkrtchyan[1], A. Mkrtchyan[1], C. Munoz Camacho[23], M. Murray[69], K. Nagai[32], J. Nagle[65], I. Nakagawa[53], C. Nattrass[77], D. Nguyen[28], S. Niccolai[23], R. Nouicer[4], G. Nukazuka[53], M. Nycz[78], V. A. Okorokov[43], S. Orešič[74], C. O'Shaughnessy[32], S. Paganis[46], Z Papandreou[74], S. F. Pate[42], M. Patel[26], G. Penman[18], M. G. Perdekamp[68], D. V. Perepelitsa[65], H. Periera da Costa[32], K. Peters[19], W. Phelps[10], E. Piasetzky[62], C. Pinkenburg[4], I. Prochazka[8], T. Protzman[34], M. L. Purschke[4], J. Putschke[82], J. R. Pybus[37], R. Rajput-Ghoshal[28], J. Rasson[49], B. Raue[16], K. Read[49], K. Reed[73], R. Reed[34], J. Reinhold[16], E. L. Renner[33], J. Richards[66], C. Riedl[68], T. Rinn[4], J. Roche[48], G. M. Roland[37], G. Ron[22], M. Rosati[26], C. Royon[69], J. Ryu[51], S. Salur[54], N. Santiesteban[37], R. Santos[66], M. Sarsour[17], J. Schambach[49], A. Schmidt[20], N. Schmidt[49], C. Schwarz[19], J. Schwiening[19], R. Seidl[53], A. Sickles[68], P. Simmerling[66], S. Sirca[71], D. Sharma[17], Z. Shi[32], T.-A. Shibata[41], C.-W. Shih[40], S. Shimizu[53], U. Shrestha[66], K. Slifer[72], K. Smith[32], D. Sokhan[18,24], R. Soltz[35], W. Sondheim[32], J. Song[9], J. Song[51], I. I. Strakovsky[20], P. Steinberg[4], P. Stepanov[12], J. Stevens[84], J. Strube[50], P. Sun[9], X. Sun[7], K. Suresh[74], V. Tadevosyan[1], W.-C. Tang[40], S. Tapia Araya[26], S. Tarafdar[79], L. Teodorescu[5], A. Timmins[67], L. Tomasek[13], N. Trotta[66], R. Trotta[12], T. S. Tveter[73], E. Umaka[26], A. Usman[74], H. W. van Hecke[32], C. Van Hulse[23], J. Velkovska[79], E. Voutier[23], P.K. Wang[23], Q. Wang[69], Y. Wang[7], Y. Wang[63], D. P. Watts[87], N. Wickramaarachchi[12], L. Weinstein[47], M. Williams[37], C.-P. Wong[32], L. Wood[50], M. H. Wood[6], C. Woody[4], B. Wysiouch[37], Z. Xiao[63], Y. Yamazaki[30], Y. Yang[39], Z. Ye[63], H. D. Yoo[87], M. Yurov[32], N. Zachariou[87], W.A. Zajc[11], J. Zhang[78], Y. Zhang[63], Y. X. Zhao[25], X. Zheng[78], P. Zhuang[63]

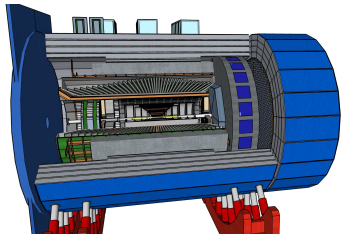
[1] A. Alikhanyan National Laboratory, Yerevan, Armenia
[2] Institute of Physics, Academia Sinica, Taipei, Taiwan
[3] Augustana University, Sioux Falls, , SD, USA
[4] Brookhaven National Laboratory, Upton, 11973, NY, USA
[5] Brunel University London, Uxbridge, , UK
[6] Canisius College, 2001 Main St., Buffalo, 14208, NY, USA
[7] Central China Normal University, Wuhan, China
[8] Charles University, Prague, Czech Republic
[9] China Institute of Atomic Energy, Fangshan, Beijing, China
[10] Christopher Newport University, Newport News, , VA, USA
[11] Columbia University, New York, , NY, USA
[12] Catholic University of America, 620 Michigan Ave. NE, Washington DC, 20064, USA
[13] Czech Technical University, Prague, Czech Republic
[14] Duquesne University, Pittsburgh, , PA, USA
[15] Duke University, , , NC, USA
[16] Florida International University, Miami, , FL, USA
[17] Georgia State University, Atlanta, , GA, USA
[18] University of Glasgow, Glasgow, , UK
[19] GSI Helmholtzzentrum fuer Schwerionenforschung, Planckstrasse 1, Darmstadt, 64291, Germany
[20] The George Washington University, Washington, DC, 20052, USA
[21] Hampton University, Hampton, , VA, USA
[22] Hebrew University, Jerusalem, , Isreal
[23] Universite Paris-Saclay, CNRS/IN2P3, IJCLab, Orsay, France
[24] IRFU, CEA, Universite Paris-Saclay, Gif-sur-Yvette France
[25] Chinese Academy of Sciences, Lanzhou, , , China

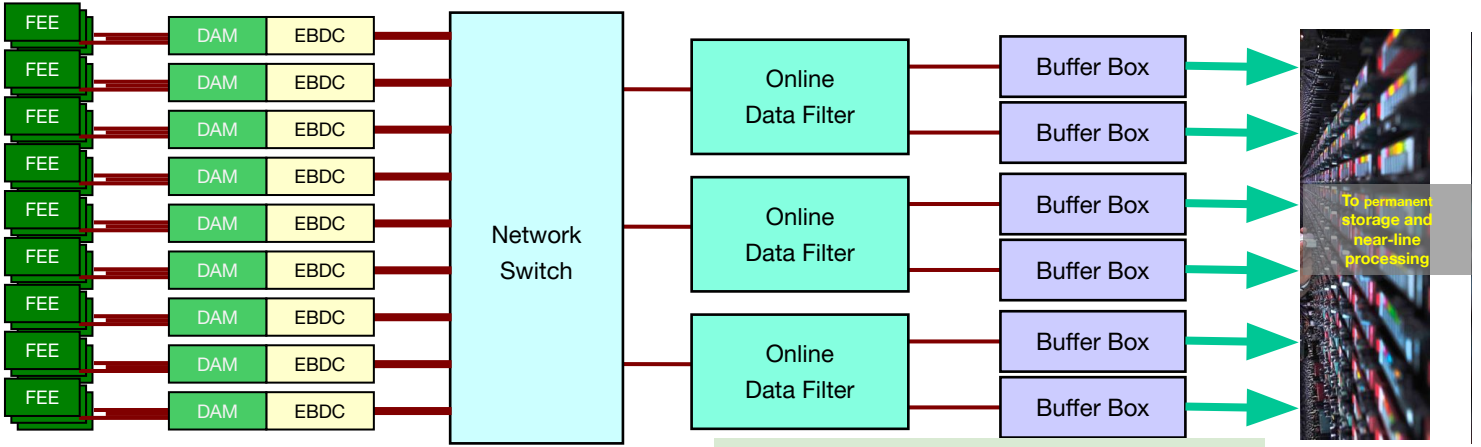
Preprint submitted to Nuclear Instruments and Methods A May 17, 2022

DAQ: Overview

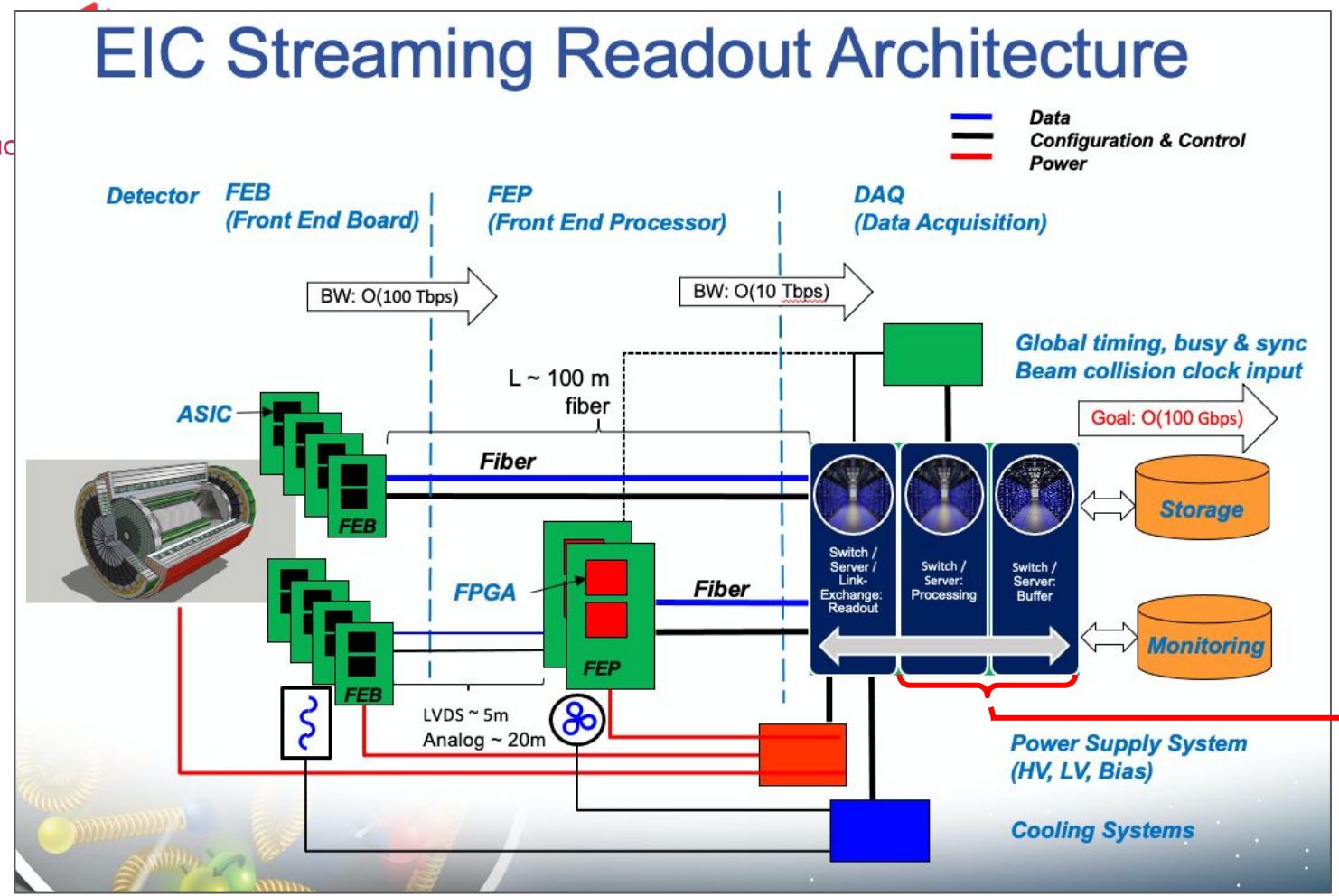
EIC Project Detector-1
"We need a real name"

- **Community solidly behind Streaming Data Acquisition System (SRO)**
  - Widely recommended by experts: EIC Computing Consortium, EIC Yellow Report
  - No need to wait for all signals from single crossing to read out data
  - Removes nearly all deadtime
  - Less restrictions for filter criteria and potentially less bias





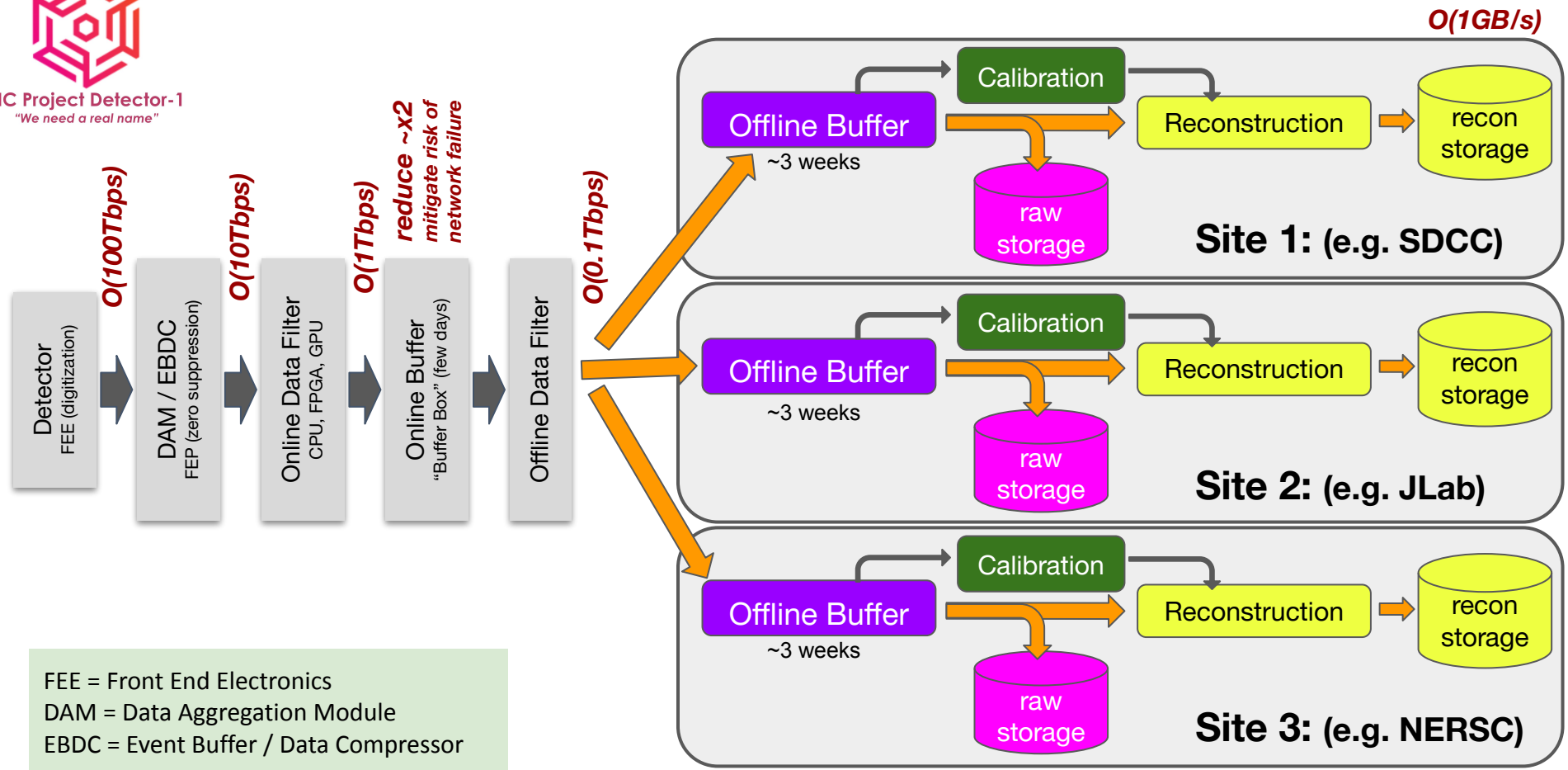
FEE = Front End Electronics
DAM = Data Aggregation Module
EBDC = Event Buffer / Data Compressor



factor of 100 in data reduction

EIC Streaming Readout (From Fernando Barbosa's talk at AI4EIC Sep. 9, 2021)

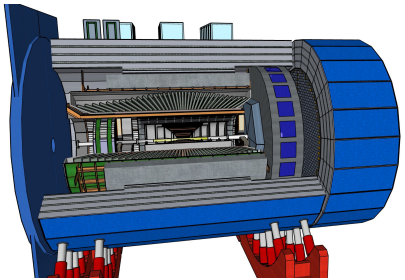
EIC Project Detector-1
"We need a real name"



FEE = Front End Electronics
DAM = Data Aggregation Module
EBDC = Event Buffer / Data Compressor

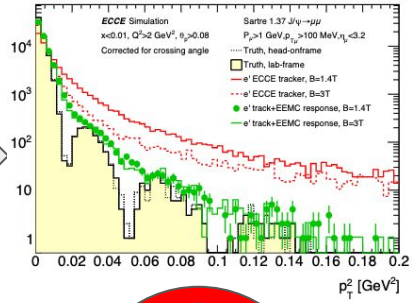
EIC Project Detector-1
"We need a real name"

Pipeline from detector to publication does not include data from tape!



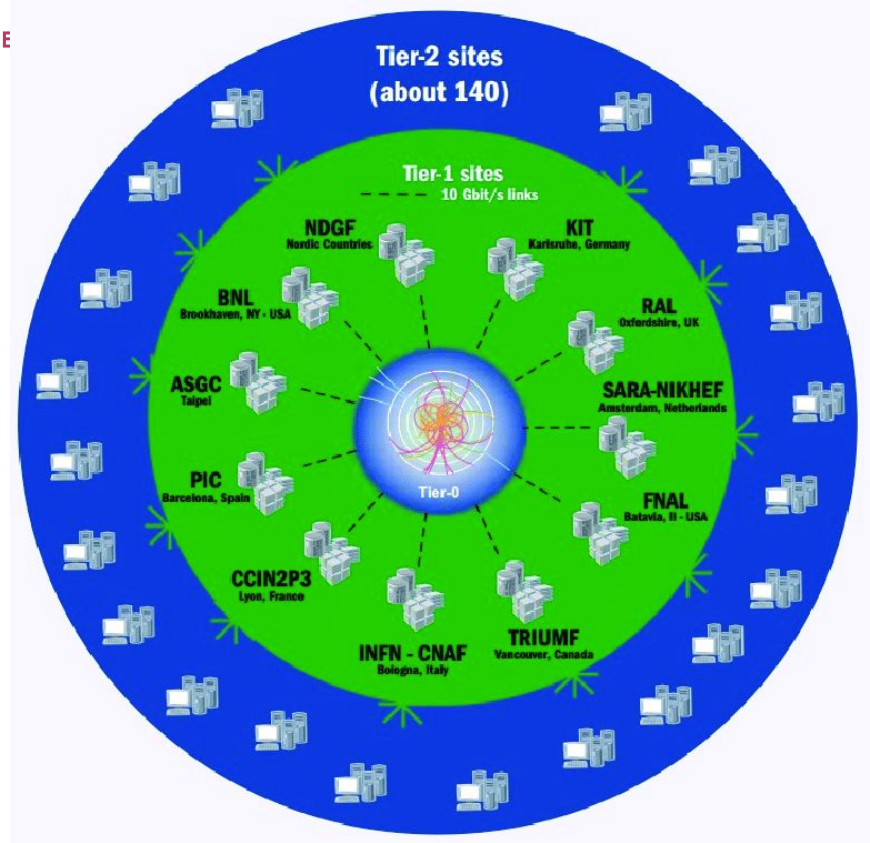
DAQ, BUFFERING, FILTERING, CALIBRATION, RECONSTRUCTION, ANALYSIS

=



Tape is a safety feature that
you hope to never use!

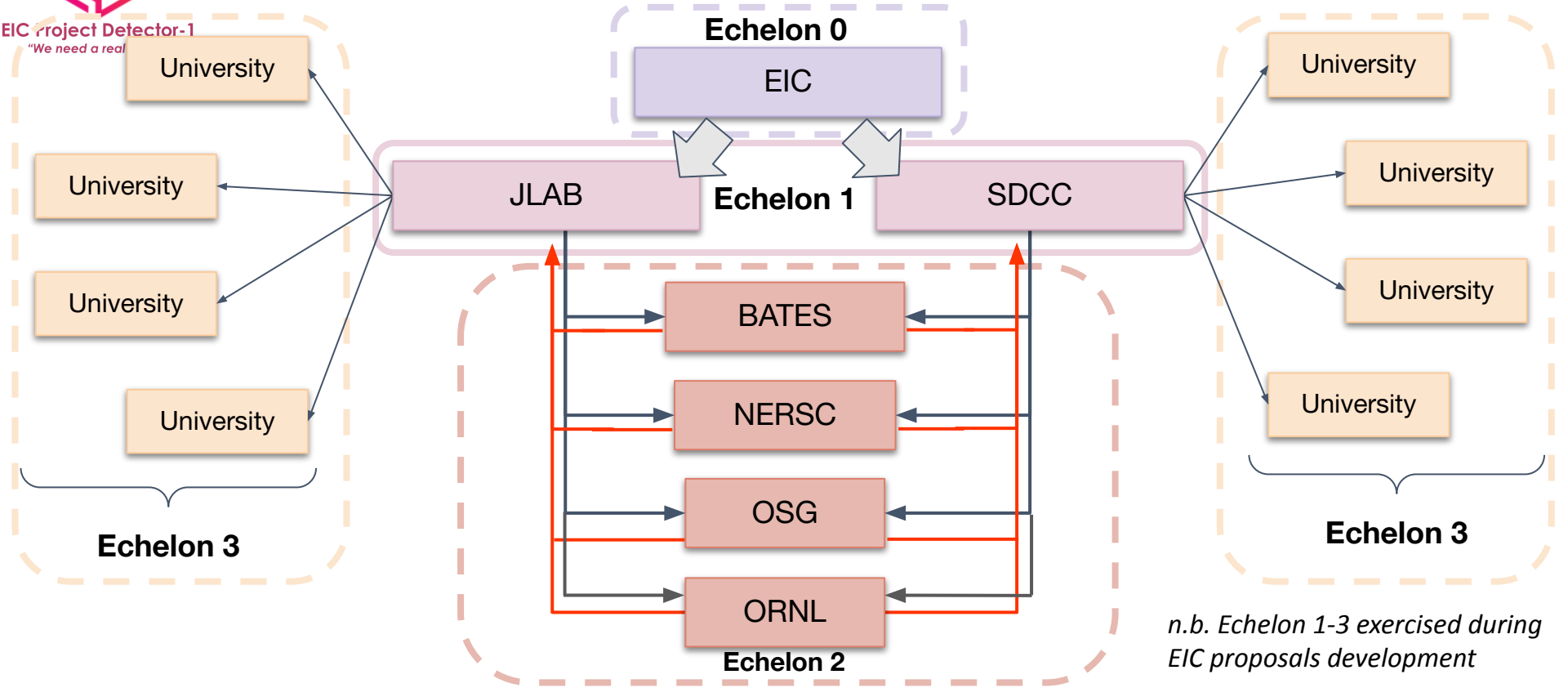
Worldwide LHC Computing Grid (WLCG)



- **Tier 0 - LHC**
  - Store all raw data
  - 20% of LHC computing
  - Initial reconstruction
  - Distribute raw data and initial recon to all Tier 1 sites
- **Tier 1 - multiple sites(13)**
  - Combined capacity to store all data from Tier 0
  - Compute for large scale reprocessing
  - Distribute data to Tier 2 sites
  - Store simulated data from Tier 2 sites
- T**ier 2 - Universities and Scientific sites (155)**
  - Compute for specific analysis tasks
  - Some reconstruction
  - Simulation
- **Tier 3 - End users**
  - small clusters or personal computers

https://home.cern/science/computing/grid-system-tiers

EIC Butterfly Model



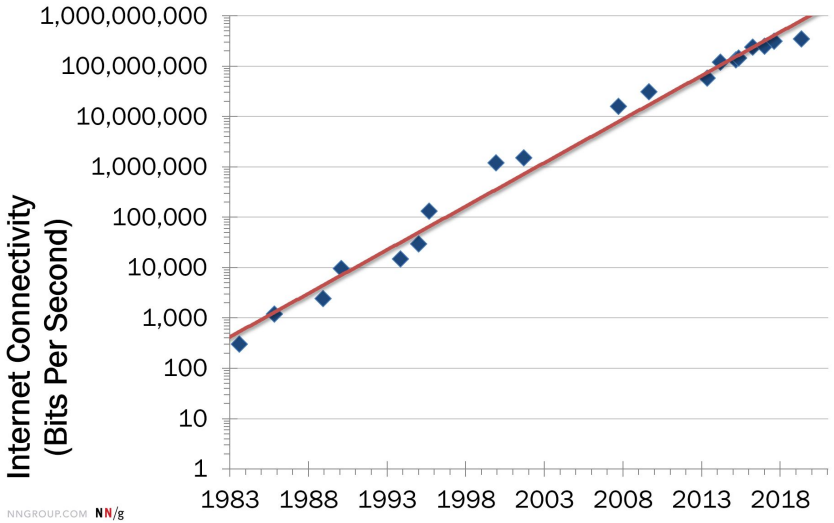
n.b. Echelon 1-3 exercised during EIC proposals development

*Nearly all storage (raw data, reconstructed data, simulated data) is stored across **Echelon 1** sites*

EIC Project Detector-1
"We need a real name"

Network Bandwidth for EIC

- BNL has 400Gbps connection to ESNet today (2021)
- Nielsen's law claims 50% increase on average every year
- 8 years from now will see x25 increase
- If BNL only increases x10, then streaming entire EIC-Detector-1 raw data set will take only **2.5%** of total bandwidth



https://www.nngroup.com/articles/law-of-bandwidth/

https://science.osti.gov/Funding-Opportunities/Digital-Data-Management

DOE Requirements

To integrate data management planning into the overall research plan, the following requirements will apply to all Office of Science research solicitations and invitations for new, renewal, and some supplemental funding issued on or after October 1, 2014. **These requirements apply to** proposals from all organizations including academic institutions, **DOE National Laboratories,** and others. These requirements do *not* apply to applications to use Office of Science user facilities.

All proposals submitted to the Office of Science for research funding must include a Data Management Plan (DMP) that addresses the following requirements:

1. DMPs should **describe whether and how data** generated in the course of the proposed research **will be shared and preserved**. If the plan is not to share and/or preserve certain data, then the plan must explain the basis of the decision (for example, cost/benefit considerations, other parameters of feasibility, scientific appropriateness, or limitations discussed in #4). At a minimum, DMPs must describe how data sharing and preservation will enable validation of results, or how results could be validated if data are not shared or preserved.
2. DMPs should provide a **plan for making all research data displayed in publications** resulting from the proposed research **open, machine-readable, and digitally accessible to the public** at the time of publication. This includes data that are displayed in charts, figures, images, etc. In addition, the underlying digital research data used to generate the displayed data should be made as accessible as possible to the public in accordance with the principles stated above. This requirement could be met by including the data as supplementary information to the published article, or through other means. The published article should indicate how these data can be accessed.
3. DMPs should **consult and reference** available information about **data management resources** to be used in the course of the proposed research. In particular, DMPs that explicitly or implicitly commit data management resources at a facility beyond what is conventionally made available to approved users should be accompanied by **written approval from that facility**. In determining the resources available for data management at Office of Science User Facilities, researchers should consult the published description of data management resources and practices at that facility and reference it in the DMP. Information about other Office of Science facilities can be found in the additional guidance from the sponsoring program.

4. DMPs must **protect confidentiality, personal privacy**, Personally Identifiable Information, and U.S. national, homeland, and economic security; recognize proprietary interests, business confidential information, and intellectual property rights; avoid significant negative impact on innovation, and U.S. competitiveness; and otherwise be consistent with all applicable laws, regulations, and DOE orders and policies. There is no requirement to share proprietary data.

DMPs will be reviewed as part of the overall Office of Science research proposal merit review process. Additional requirements and review criteria for the DMP may be identified by the sponsoring program or sub-program, or in the solicitation.

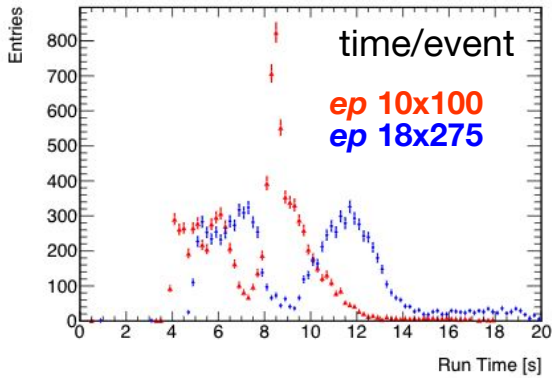
EIC Project Detector-1
"We need a real name"

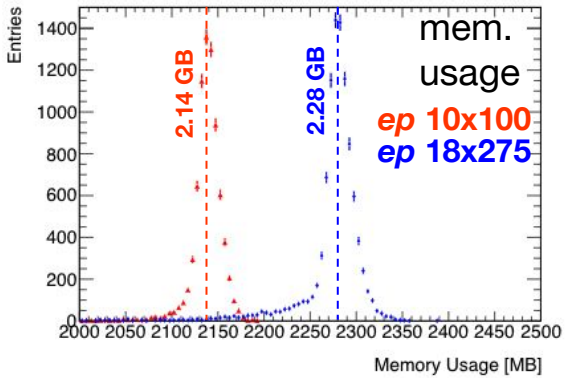
CPU

Simulation focused on next four years where detector refinement will be key

Year	Number of Events [$\times 10^6$]	Storage [TB]	CPU hours [kHours]
2022	200	50	450
2023	100	25	225
2024	100	25	225
2025	50	12.5	110
Total	450	112.5	1010

n.b. not all cores are equal. These reflect rough average of cores currently in use (i.e. some modern and some several years old)





Raw Data Requirements *(estimated)*

EIC ~~ECCE~~ Runs	year-1	year-2	year-3
Luminosity	$10^{33}\text{cm}^{-2}\text{s}^{-1}$	$2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$	$10^{34}\text{cm}^{-2}\text{s}^{-1}$
Weeks of Running	10	20	30
Operational efficiency	40%	50%	60%
Disk (temporary)	1.2PB	3.0PB	18.1PB
Disk (permanent)	0.4PB	2.4PB	20.6PB
Data Rate to Storage	6.7Gbps	16.7Gbps	100Gbps
Raw Data Storage (no duplicates)	4PB	20PB	181PB
Recon process time/core	5.4s/ev	5.4s/ev	5.4s/ev
Streaming-unpacked event size	33kB	33kB	33kB
Number of events produced	121 billion	605 billion	5,443 billion
Recon Storage	0.4PB	2PB	18PB
CPU-core hours (recon+calib)	191Mcore-hrs	953Mcore-hrs	8,573Mcore-hrs
2020-cores needed to process in 30 weeks	38k	189k	1,701k

EIC Project Detector-1
"We need a real name"

Heterogeneous Hardware

- Anticipate supporting heterogeneous hardware at all stages of pipeline
  - GPU, TPU, CPU, FPGA, …
- AI/ML support
  - Expected at all stages (design, simulation, filtering, calibration, reconstruction, analysis, …)
  - Type of hardware depends on AI model (some models run faster on CPU)
- Custom hardware-specific algorithms
  - Only if necessary
  - Easier for front end hardware-wise, but limits diagnostics/emulation offline

EIC Project Detector-1
"We need a real name"

Summary

- Data pipeline starts with SRO DAQ, and couples through to reconstruction
  - Use buffering to allow 2-3 weeks for calibration
  - Tape is for safety only and is not part of primary pipeline
- Expect to support AI/ML and Heterogeneous Hardware at all stages
- Butterfly model focuses most storage on a pool of Echelon 1 sites
  - Reduces the number of single points of failure
  - Dynamically adjust resource location depending on demand and availability
  - EIC level (not Detector-1 specific)
- Expect data rate to storage to be O(100) Gbps
  - Reconstructed data rate will be ~1/10 and kept live on disk
- Details on data preservation and sharing have yet to be worked out
  - should be done at EIC level

EIC Project Detector-1
"We need a real name"

Backups

EIC Project Detector-1
"We need a real name"

Tape Storage

Numbers derived based on steady-state production estimate of:

O(100)Gbps (signal data only, after filtering)
10^{34} $cm^{-2}s^{-1}$ luminosity
30 weeks running / yr
60% operational efficiency

from Yellow Report

n.b. data stream from detector O(100)Tbps

New Storage	year-1	year-2	year-3
Luminosity	$10^{33}cm^{-2}s-1$	$2x10^{33}cm^{-2}s-1$	$10^{34}cm^{-2}s-1$
Weeks of Running	10	20	30
Operational efficiency	40%	50%	60%
Data Rate to Storage	6.7Gbps	16.7Gbps	100Gbps
Raw Data Storage (no duplicates)	4PB	20PB	181PB
Recon Storage	0.4PB	2PB	18PB
Total Storage (no duplicates)	4.4PB	22PB	200PB

EIC Project Detector-1
"We need a real name"

Disk Storage

Numbers derived based on raw data produced in given year

Temporary stores raw data for 3 weeks for calibration and reproduction to complete

Permanent stores reconstructed data (cumulative)

Total Disk	year-1	year-2	year-3
Disk (temporary)	1.2PB	3.0PB	18.1PB
Disk (permanent)	0.4PB	2.4PB	20.6PB
TOTAL	1.6PB	5.4PB	38.7PB

Stage	Input/Output	Reduction Factor	Technology options
Compute Interface (e.g. FELIX)	100Tbps/10Tbps	$\times 10^{-1}$	FPGA
Online Event Filter	10Tbps/1Tbps	$\times 10^{-1}$	FPGA, (GPU), CPU
Online Buffer	1Tbps/0.5Tbps	$\times 5x10^{-1}$	$< disk >$
Offline Event Filter	0.5Tbps/100Gbps	$\times 2x10^{-1}$	FPGA, GPU, CPU
Reconstruction	100Gbps/10Gbps	$\times 10^{-1}$	(FPGA), GPU,CPU
Total	**100Tbps/10Gbps**	$\times 10^{-4}$	