sPHENIX

Chris Pinkenburg, BNL

322 days till first beam


1st sPHENIX workfest, 2011 in Boulder
Computing corner

sPHENIX Science Mission

- The goals of heavy-ion experiments at RHIC and the LHC as a result of the 2015 Long Range Plan for Nuclear Science are two-fold :
  1. To map the QCD phase diagram with experiments planned at RHIC
  2. To probe the inner workings of quark-gluon plasma (QGP) by resolving its properties at shorter and shorter length scales

sPHENIX Collaboration

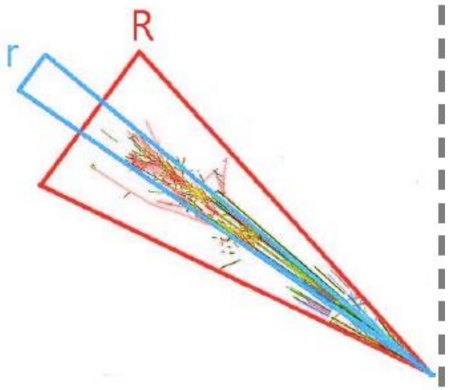
- **Officially formed in 2016**
- **More than 320 members from 84 institutions in 14 countries as of 2021**

Members from around the world gathered around a common science goal

sPHENIX Core Physics Program

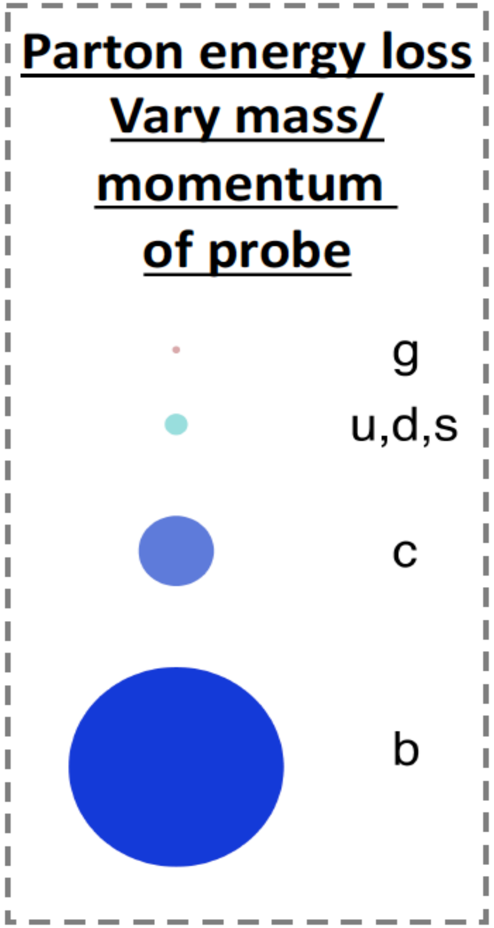
Jet correlation & substructure

Vary momentum/ angular size of probe



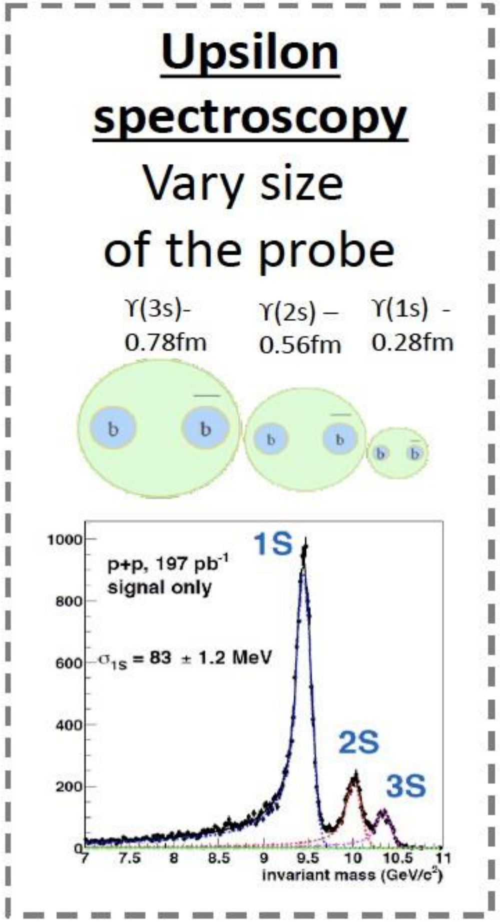
- Guided by the science mission, sPHENIX aims to : probe the QGP in different ways :
  - Vary probe's momentum and angular scale

sPHENIX Core Physics Program



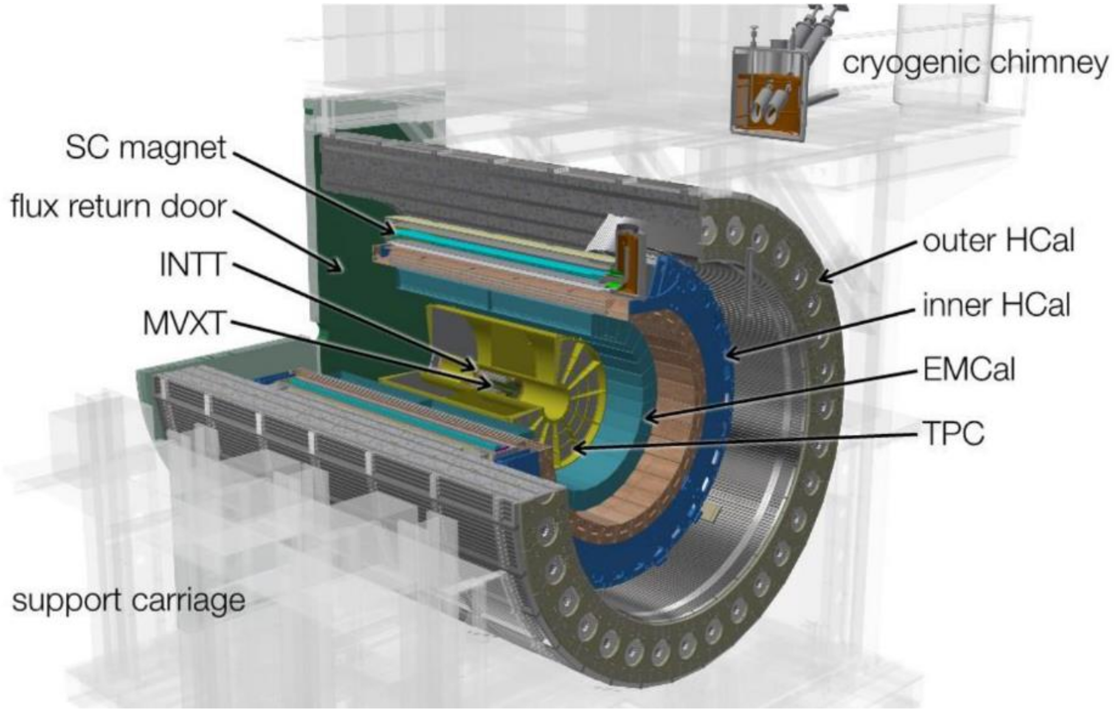
- Guided by the science mission, sPHENIX aims to : probe the QGP in different ways :
  - Vary probe's momentum and angular scale
  - Vary probe's mass and

sPHENIX Core Physics Program



- Guided by the science mission, sPHENIX aims to : probe the QGP in different ways :
  - Vary probe's momentum and angular scale
  - Vary probe's mass and momentum
  - Vary probe's size

sPHENIX Detector Overview



Calorimetry
- Outer Hadronic Calorimeter (oHCAL)
- Inner Hadronic Calorimeter (iHCAL)
- Electromagnetic Calorimeter (EMCAL)

Magnet
- 1.4T superconducting solenoid used by the BaBar experiment

Tracking
- Time Projection Chamber (TPC)
- Intermediate Silicon Tracker (INTT)
- MAPS-based Vertex Tracker (MVTX)

Performance
- **High data rate :** read out rate of 15 kHz for all subdetectors
- **Acceptance :** hermetic coverage over full azimuth & pseudorapidity $|\eta| \leq 1.1$ for the tracking & calorimeter systems

Triggers:
Only lvl1 triggers – no high level triggers
→ No online Event building needed

sPHENIX Detector Overview

Calorimetry

- Outer Hadronic Calorimeter (oHCAL)
- Inner Hadronic Calorimeter (iHCAL)

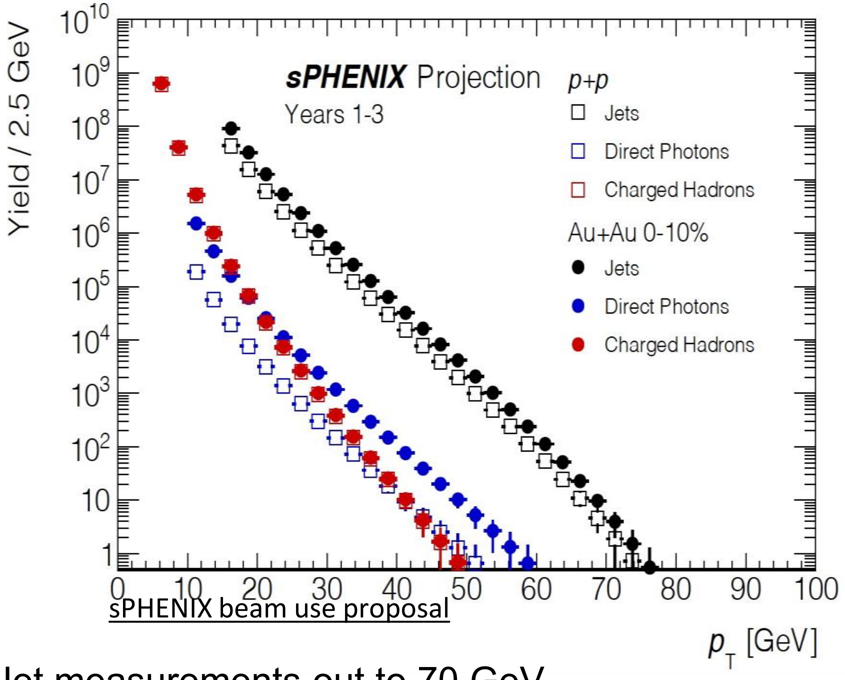
In a nutshell – designed to take a ton of minbias data in a very short period of time

subdetectors

- **Acceptance** : hermetic coverage over full azimuth & pseudorapidity $|\eta| \leq 1.1$ for the tracking & calorimeter systems

Triggers:
Only lvl1 triggers – no high level triggers
→ No online Event building needed

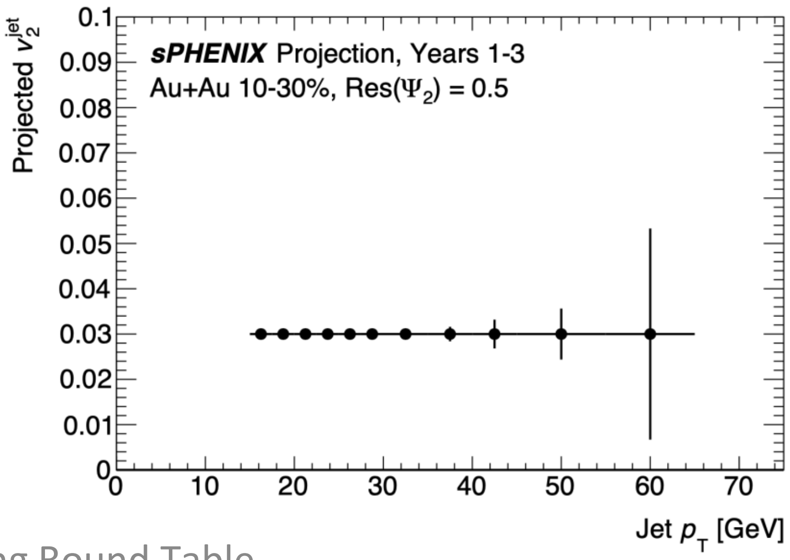
sPHENIX Probes : Jets and Photons



sPHENIX beam use proposal

Signal	Au+Au 0–10% Counts	$p+p$ Counts
Jets $p_T > 20$ GeV	22 000 000	11 000 000
Jets $p_T > 40$ GeV	65 000	31 000
Direct Photons $p_T > 20$ GeV	47 000	5 800
Direct Photons $p_T > 30$ GeV	2 400	290
Charged Hadrons $p_T > 25$ GeV	4 300	4 100

Year	Species	$\sqrt{s_{NN}}$ [GeV]	Cryo Weeks	Physics Weeks	Rec. Lum. $\|z\| <10$ cm	Samp. Lum. $\|z\| <10$ cm
2023	Au+Au	200	24 (28)	9 (13)	3.7 (5.7) nb^{-1}	4.5 (6.9) nb^{-1}



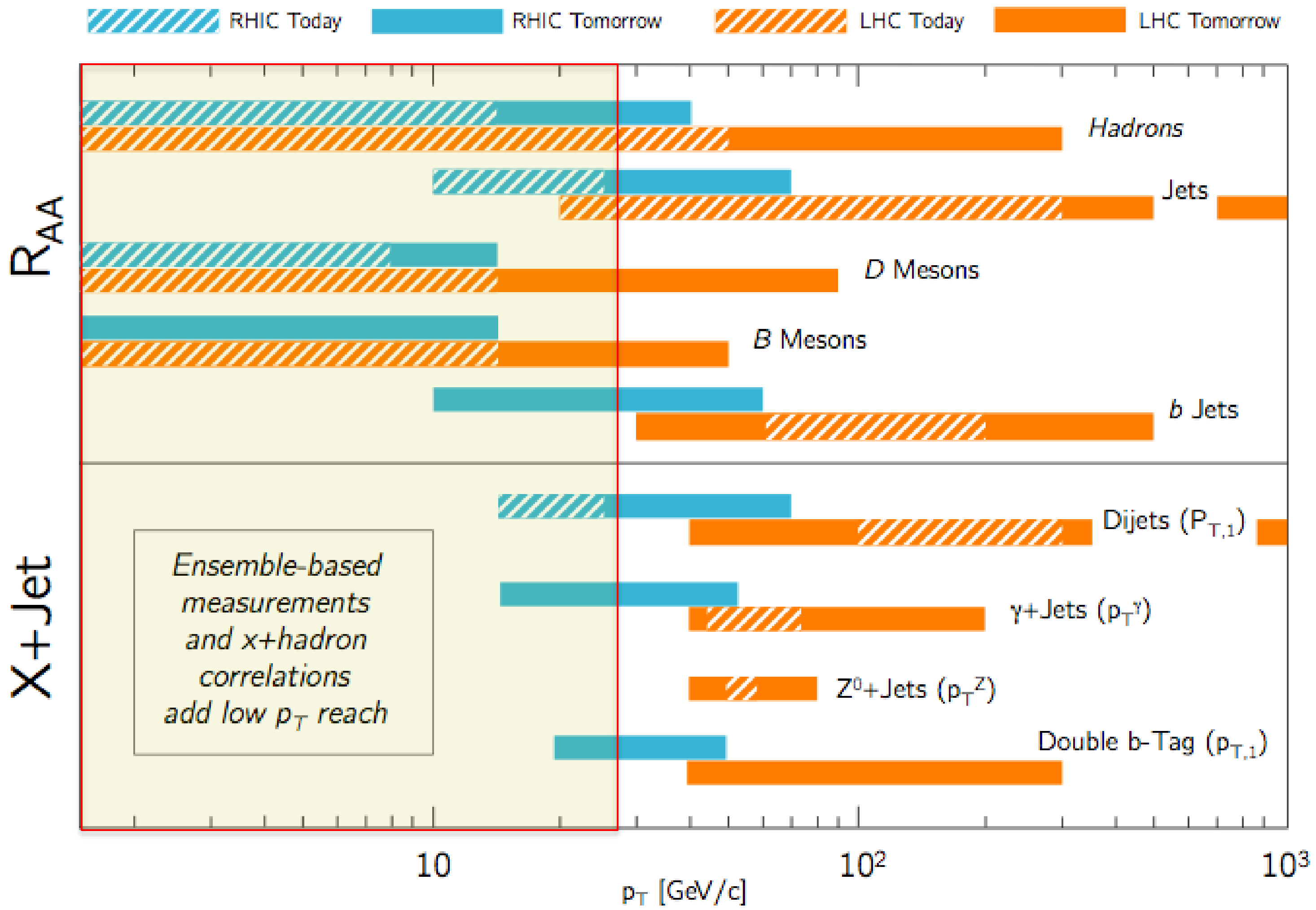
- Jet measurements out to 70 GeV
  - ➡ overlap with LHC measurements
- High stats also for
  - ➡ photons (γ-jet measurements)
  - ➡ charged hadrons (fragmentation functions, substructure)
- Large luminosity Au+Au in first year
  - ➡ dijets, jet v_2

sPHENIX complements LHC expts

Particular strength at low pT

Larger QGP effects but difficult to trigger:

- low pT, highly quenched jets
- low pT charm and beauty



sPHENIX 3-Year Run Plan

sPHENIX Beam Use Proposal (BUP) sPH-TRG-2020-001, August 31, 2020.

Year	Species	$\sqrt{s_{NN}}$ [GeV]	Cryo Weeks	Physics Weeks	Rec. Lum. $\|z\| <10$ cm	Samp. Lum. $\|z\| <10$ cm
2023	Au+Au	200	24 (28)	9 (13)	3.7 (5.7) nb^{-1}	4.5 (6.9) nb^{-1}
2024	$p^{\uparrow}p^{\uparrow}$	200	24 (28)	12 (16)	0.3 (0.4) pb^{-1} [5kHz] 4.5(6.2) pb^{-1} [10%-str]	45 (62) pb^{-1}
2024	$p^{\uparrow}$+Au	200	–	5	0.003 pb^{-1} [5kHz] 0.02 pb^{-1} [10%-str]	0.11 pb^{-1}
2025	Au+Au	200	24 (28)	20.5 (24.5)	13 (15) nb^{-1}	21 (25) nb^{-1}

Year 1 (2023) :

- Commissioning Au+Au
- Measurement of standard Au+Au candles at RHIC

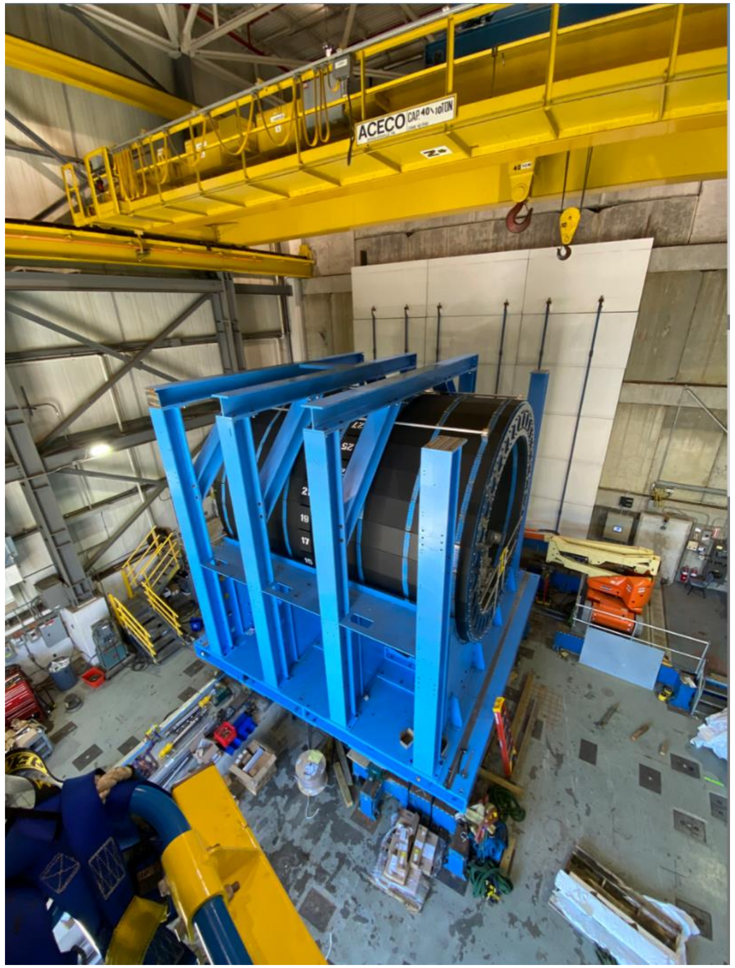
Year 2 (2024) :

- Commissioning p+p
- $p^{\uparrow}+p^{\uparrow}$, $p^{\uparrow}$+Au : HI reference set and cold QCD

Year 3 (2025) :

- Very large Au+Au heavy-ion set for jet and heavy flavor physics
- 141 B events recorded in total

sPHENIX Progress

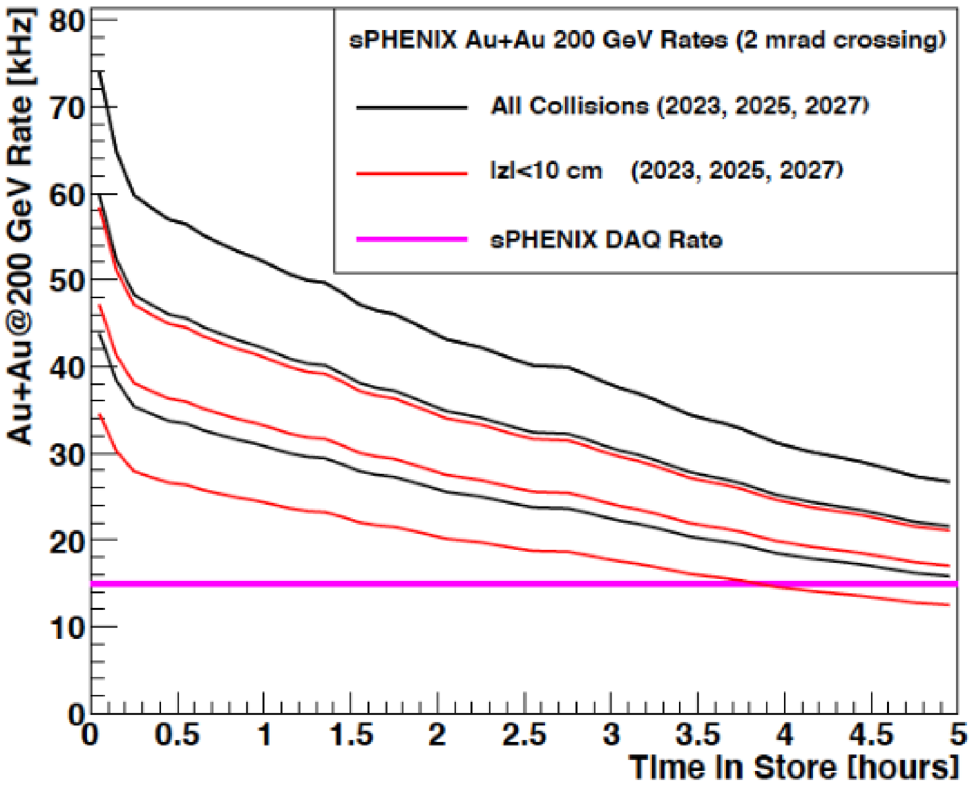


Magnet, outer hcal and rack platform at ip8

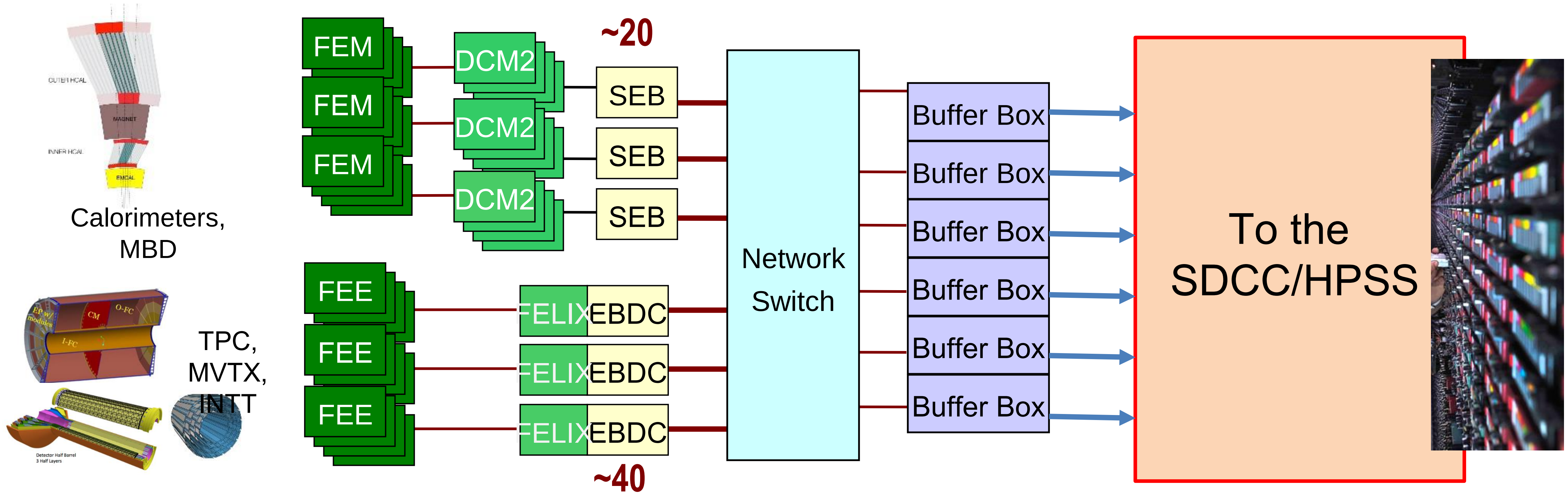
Inner Hcal barrel complete and ready to install EMC

DAQ

- sPHENIX DAQ is rate limited to 15kHz
  - Same rate in p+p and Au+Au
- Beam Crossing Angle reduces rate of collisions outside |z|<10cm by
- The TPC will “see” the full rate – pileup and resulting space charge need to be dealt with
- RHIC has 20 years of experience – collider operations will reach their peak luminosity ~~very quickly~~ on day 1



sPHENIX DAQ



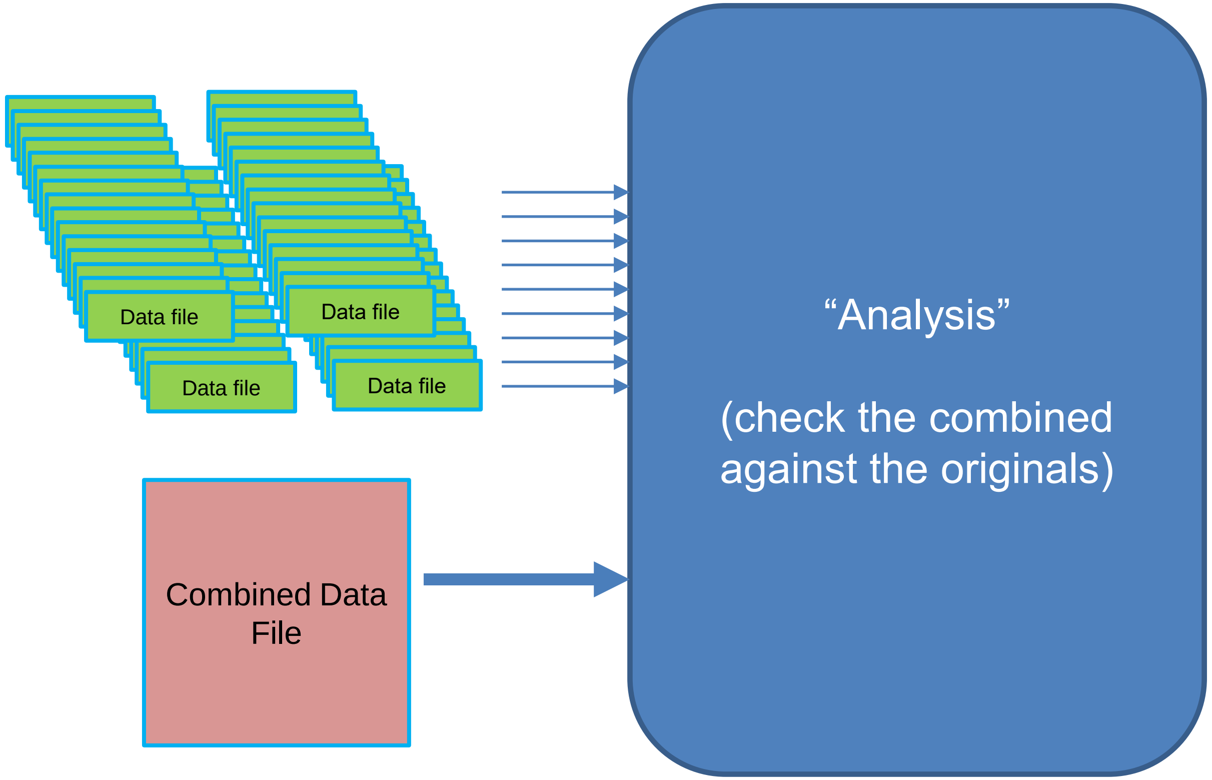
The Buffer Boxes are the only components interfacing with the tape storage system

Event Building Check ✔

We then read back the combined data file and the 60 originals into our analysis framework

Word by word we check that the packets in the combined file are the same as in the originals, and that they are all there and accounted for

500 million events, 200 runs

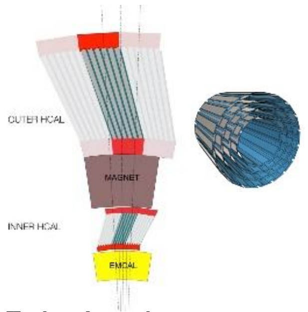


100% success rate!

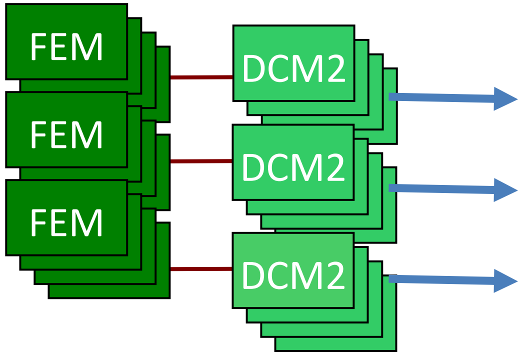
It shows that our framework is able to handle the combined stream, as well as 60 on-the-fly streams

Keep in mind that those files are generated by the real DAQ processes

Two Classes of Front-end Hardware

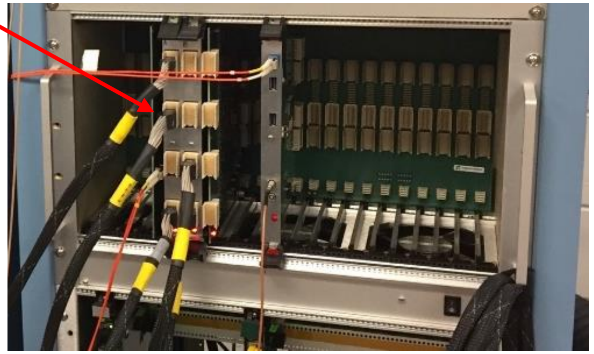


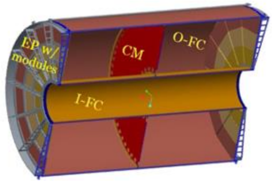
Calorimeters, INTT, MBD



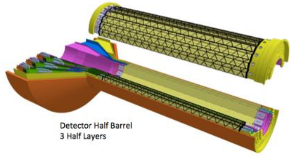
The calorimeters, the INTT, and the MBD re-use the PHENIX "Data Collection Modules" (v2)

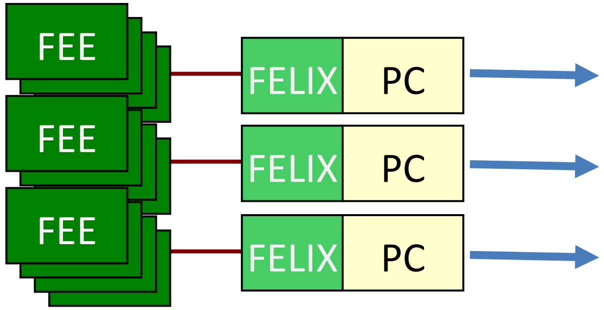
Triggered readout





TPC, MVTX





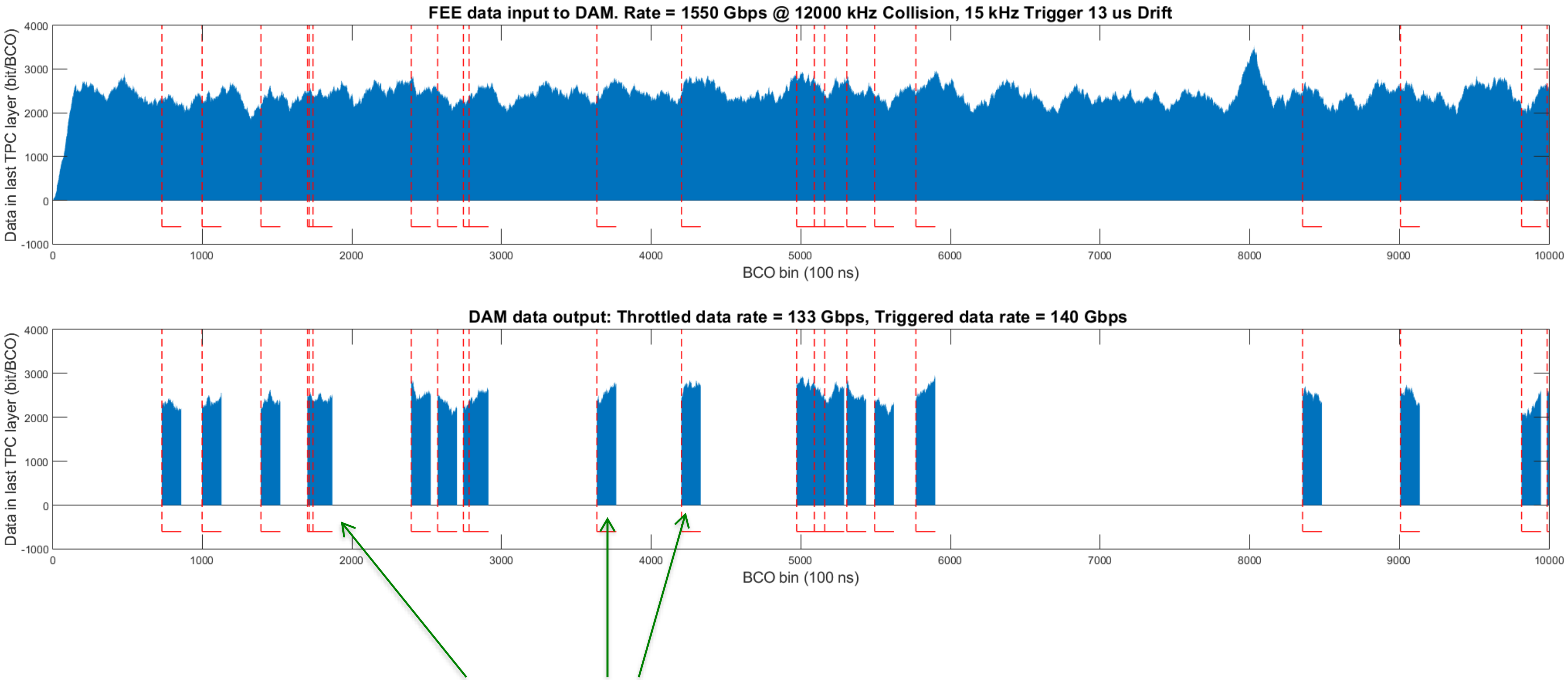
The TPC and the MVTX are read out through the ATLAS "FELIX" card directly into a standard PC

Streaming readout

ATLAS FELIX Card

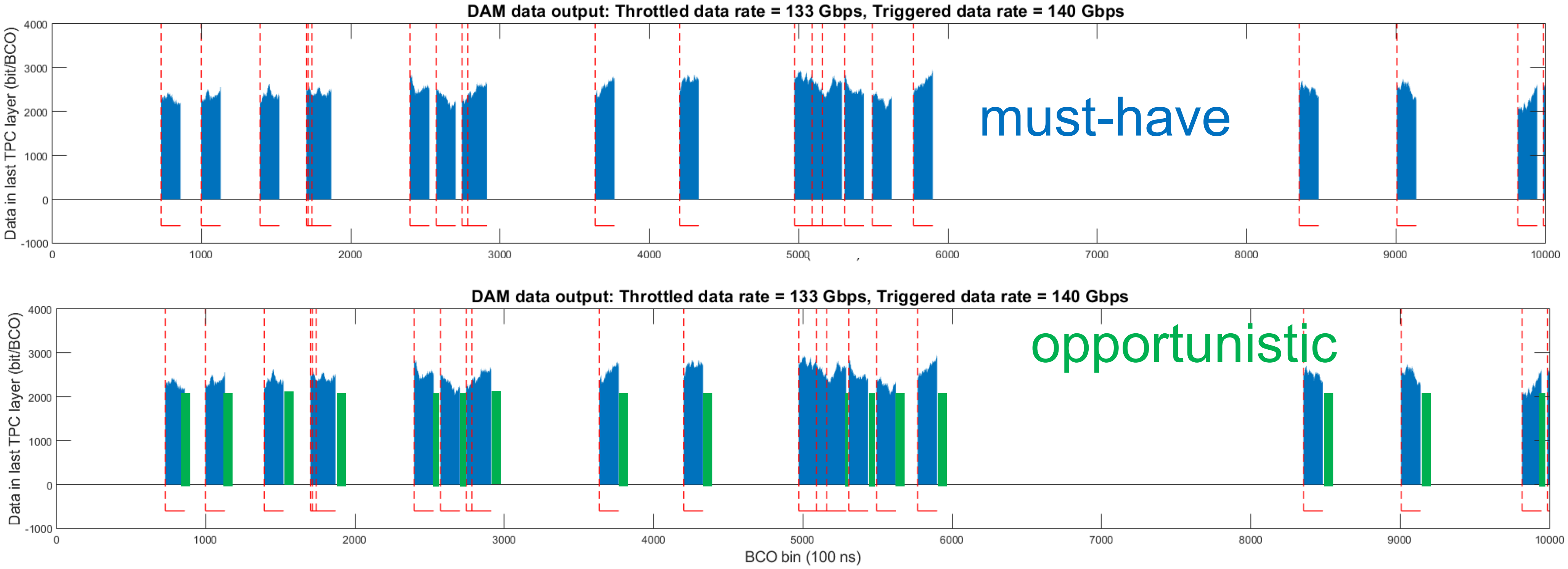
Installed in a PC

Streaming Readout + Triggered Events (concept)



Chunks correlated with triggered events

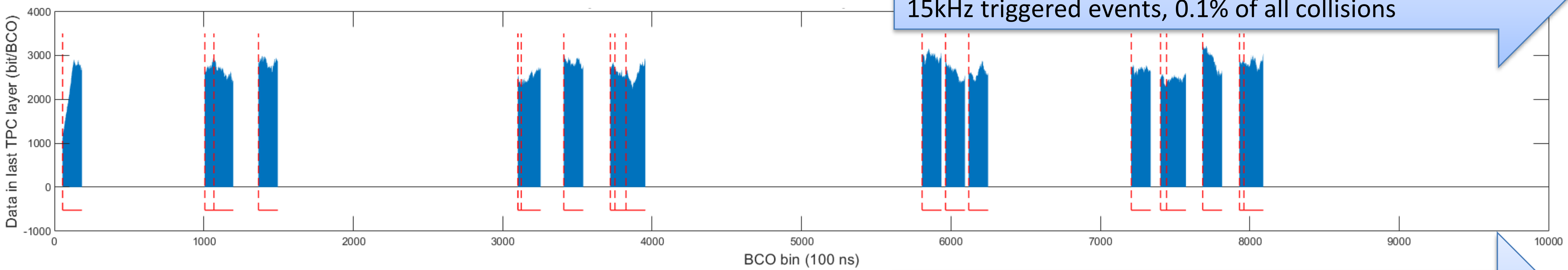
… plus “opportunistic” streaming-only data



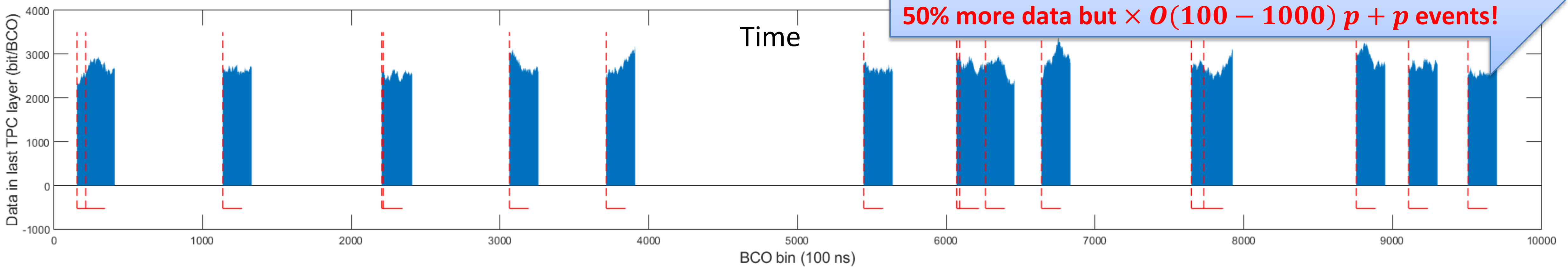
we extend the “stream time” and add tracking-only events without the calorimeters

We can “back-fill” our storage limit with those events

For p+p running, a partial triggerless DAQ?



What we write to disk, 40 Gbps
15kHz triggered events, 0.1% of all collisions



+ Streaming 10% of all collisions, 60Gbps
50% more data but $\times\, O(100-1000)\ p+p$ events!

A treasure chest of truly unbiased p+p events

Event and Data Volume numbers

Numbers taken from the beam-use proposal

Significantly lower data volumes due to the introduction of a beam crossing angle, fewer useless off-vertex collisions (TPC)

Numbers include 30% uncertainty to the high side

Run 1: Au+Au: 13 weeks @ 60% RHIC uptime x 60% sPHENIX uptime ⟶ 43 billion events 72 PB 73Gbit/s

Run 2: p+p, p+A: 21 weeks @ 60% RHIC uptime x 60% sPHENIX uptime ⟶ 69 billion events 78 PB 49Gbit/s

Run 3: Au+Au: 24.5 weeks @ 60% RHIC uptime x 80% sPHENIX uptime ⟶ 107 billion events 180 PB 97Gbit/s

These are conservative uptime figures establishing *minimum* sampled luminosity goals.
We can write up to 230PB/year if needed, so significantly higher uptimes are ok

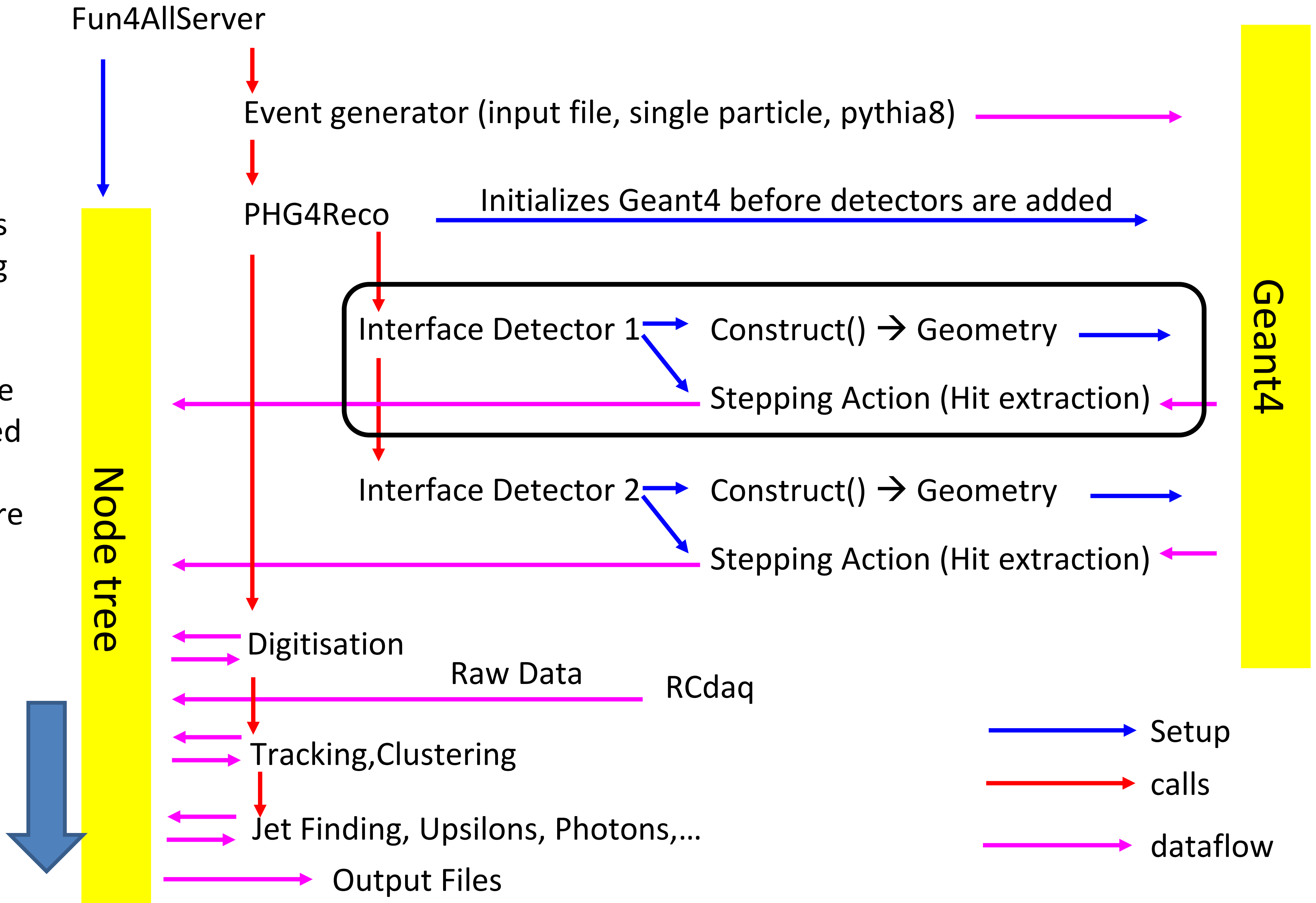
Software

- Very small core group – use what you already have
  - PHENIX took data at 6kHz and was more complicated
  - Students will write most of the code – no fancy “look what I can do with C++”
- Do not re-invent the wheel, use existing modern tools
  - ACTS
  - KFParticle
  - Conditions DB
  - PanDA
- Very short commissioning, flexibility is key
- Get tons of help from other experiments: ALICE, ATLAS, STAR
  - Thank you
- Our events are time ordered which makes reconstruction, calibrations and synchronization of multiple streams a lot easier

Deployment

- Code in github
  - Pull requests trigger extensive CI (including q/a) run in sdcc
  - Addresses the “when did we lose 10% momentum resolution?”
- Singularity/Docker container with sdcc farm image
- Daily and tagged archival builds in cvmfs
- Daily builds with gcc, clang, insure, Coverity, scan-build
  - Keeps our software c++ compliant

Continuous processing chain



The processing is done by chaining up modules. At every step the state of the Node Tree can be saved and the analysis can pick up where it left

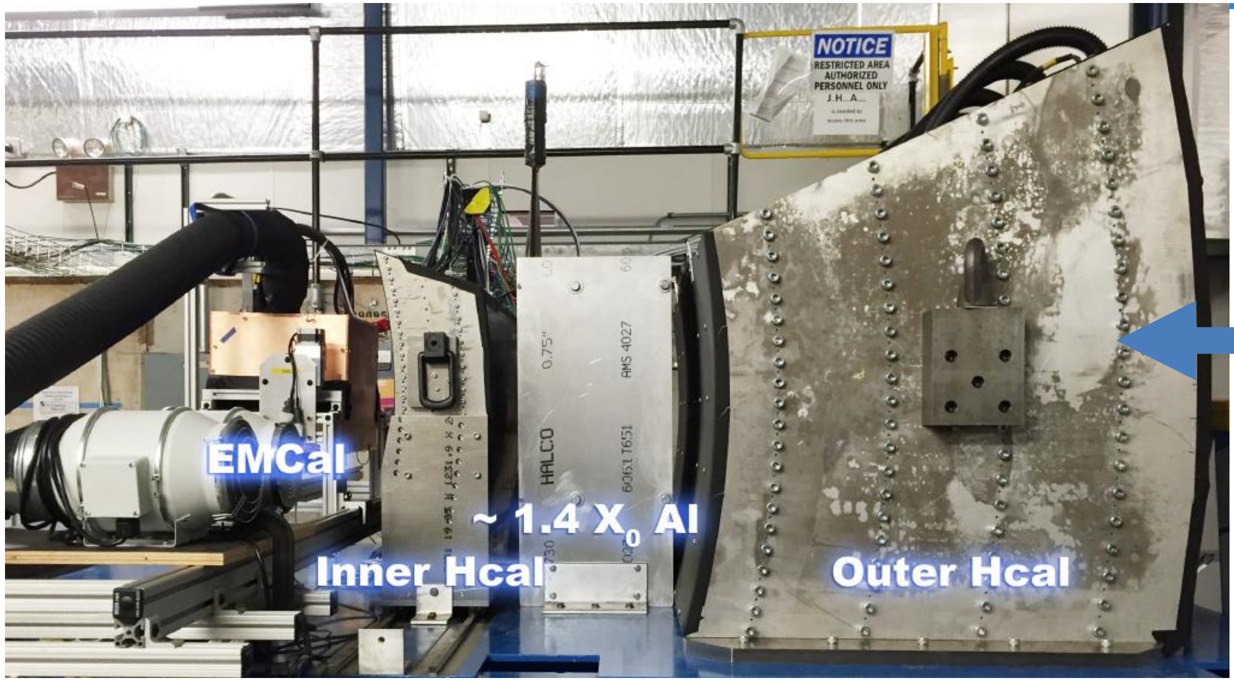
Common reco for raw and simulated data

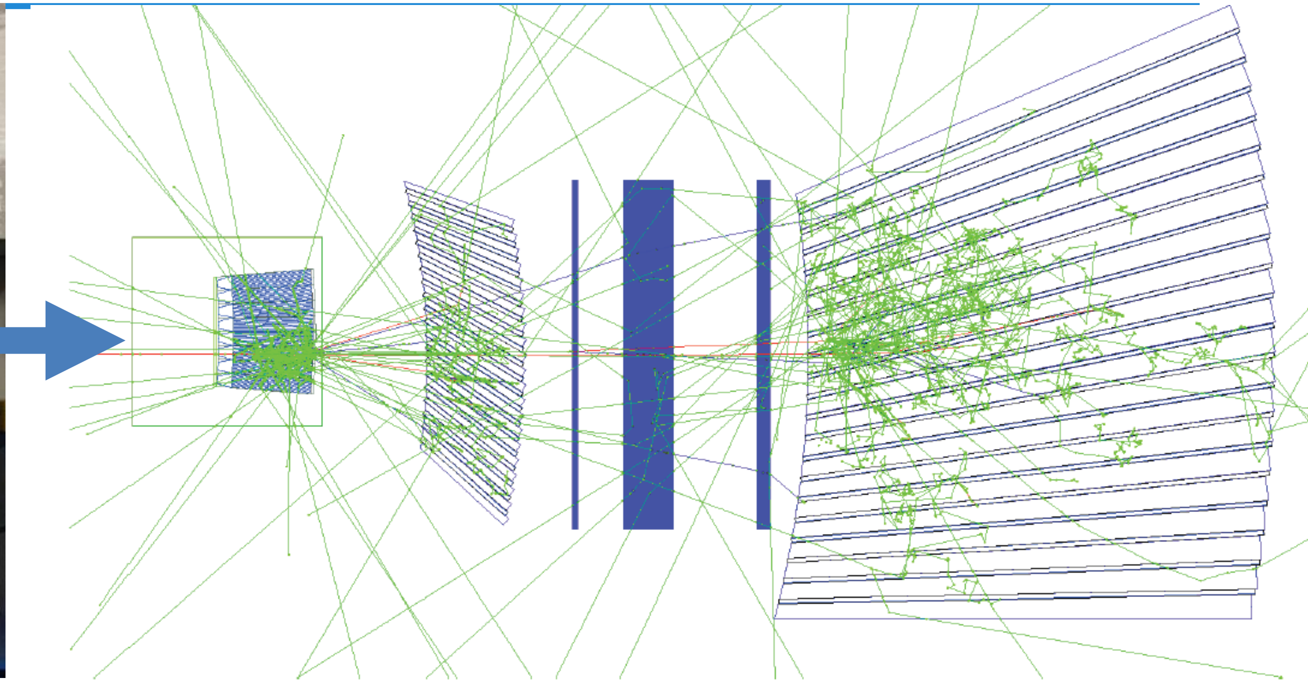
Modular: Each detector is its own entity providing the flexibility needed for complex setups

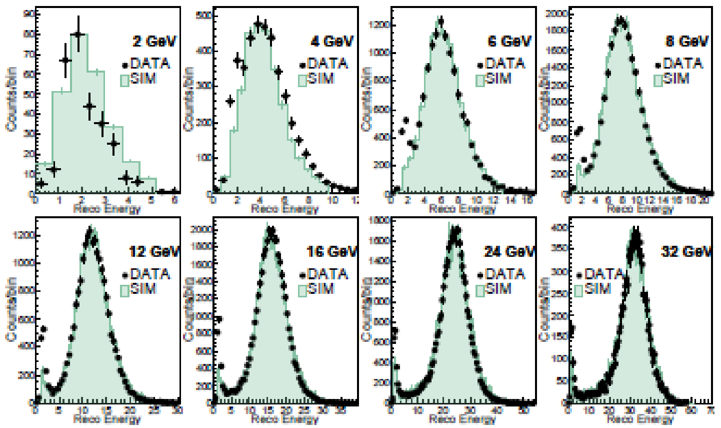
Generic detectors like boxes, cylinders, cones exist and can be configured on a macro level

No volume hierarchy, all detectors are put into world volume

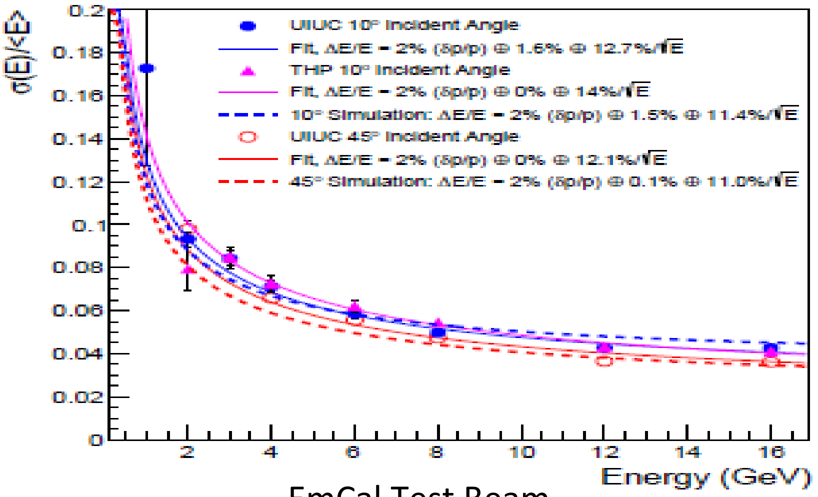
Combining data reconstruction and simulations







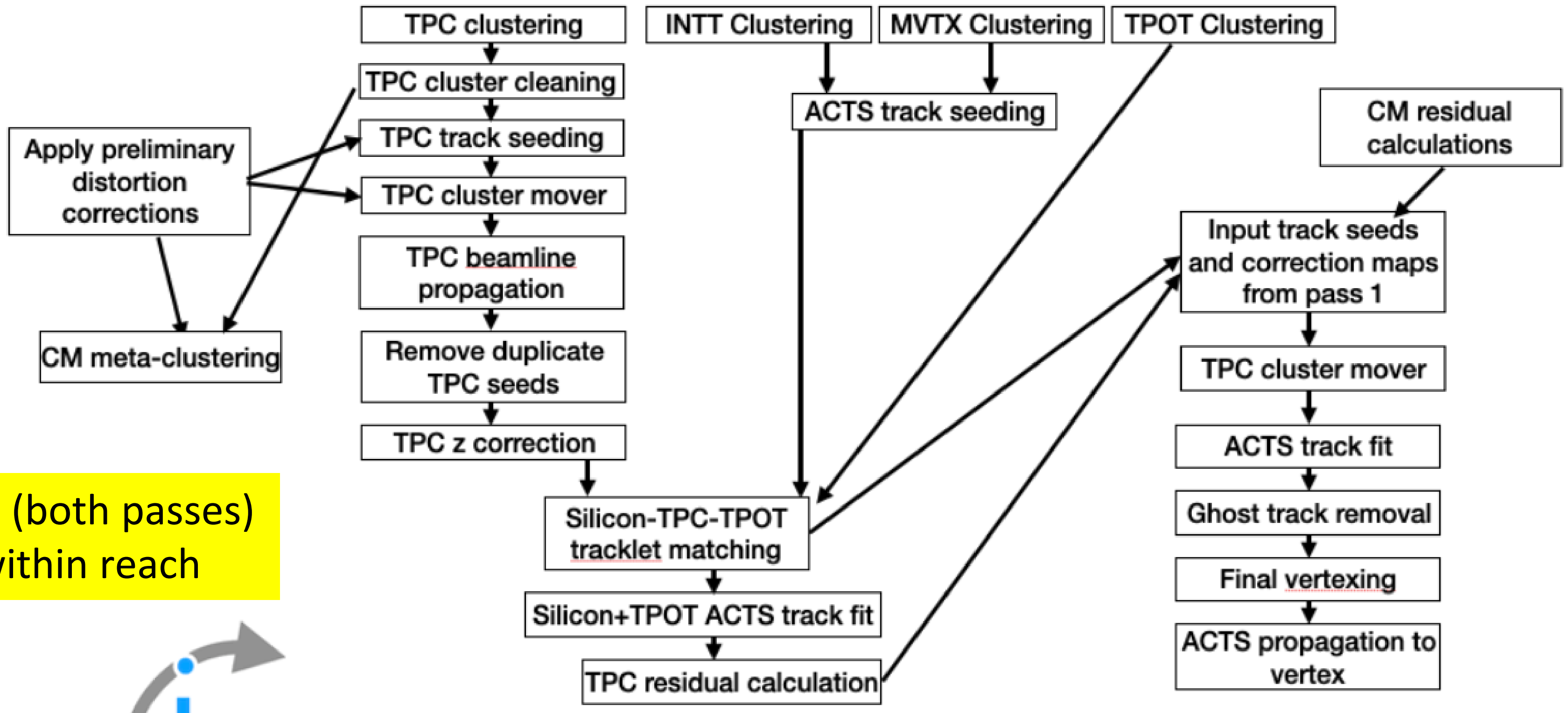
Hadronic Calorimeter Test Beam



EmCal Test Beam

Many simulation needs besides "full detector"

sPHENIX Tracking



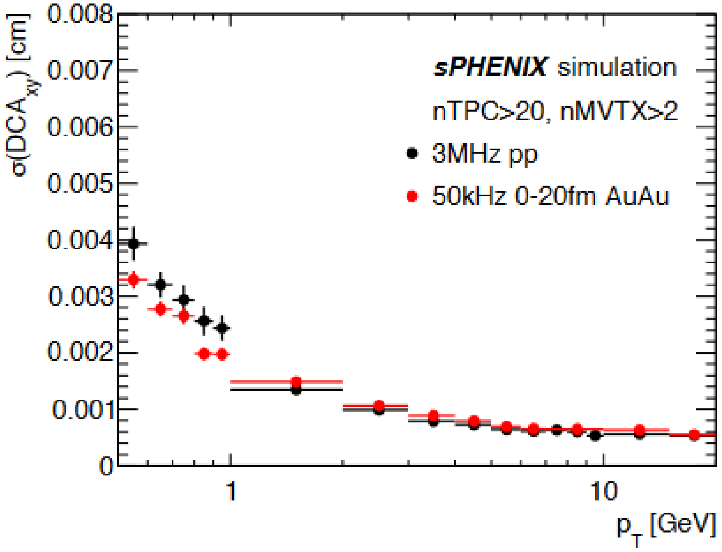
Goal of 10sec/evt (both passes) and 4GB/job within reach

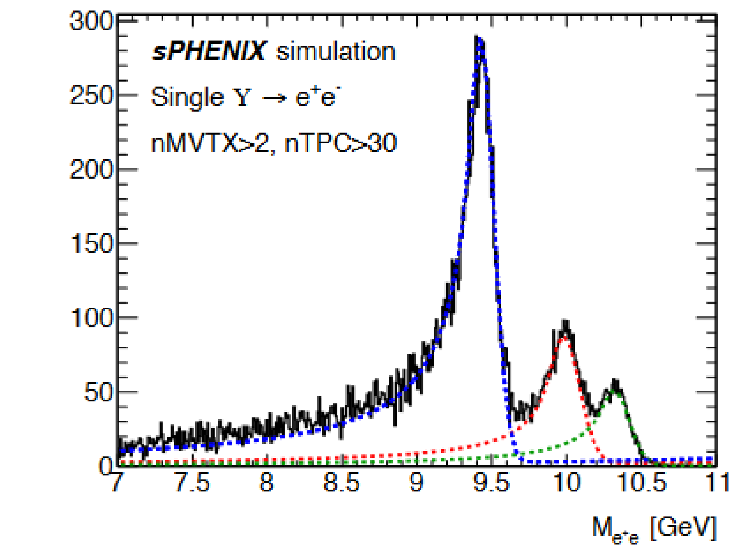
- Distortion correction scheme
  - ➡ Move clusters associated to tracks onto Acts surfaces based on a variety of correction schemes for static, beam induced, or event-by-event fluctuations

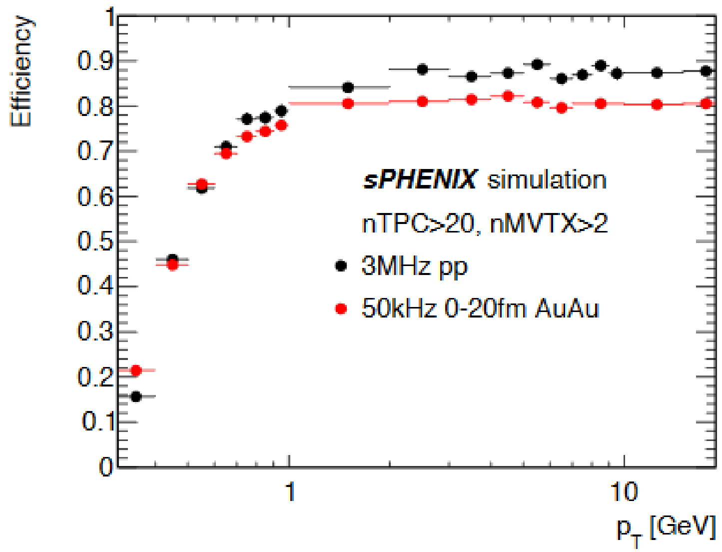
github.com/acts-project/acts

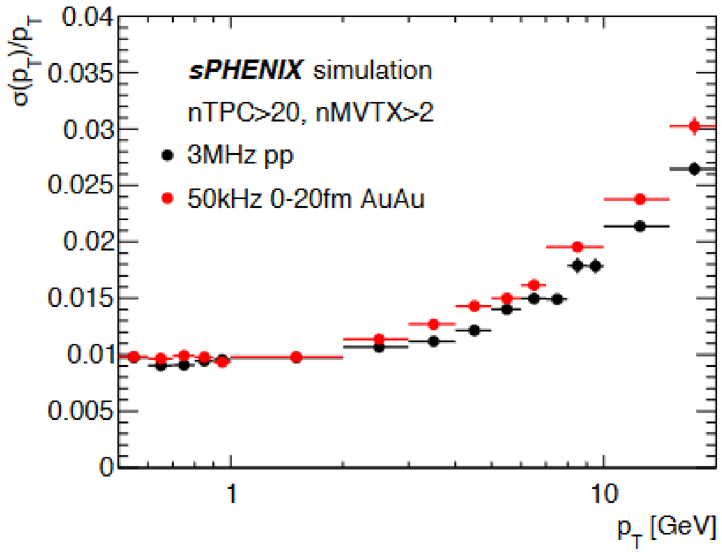
JDO et al., Computing and Software for Big Science 5, 23 (2021)

Tracking Performance





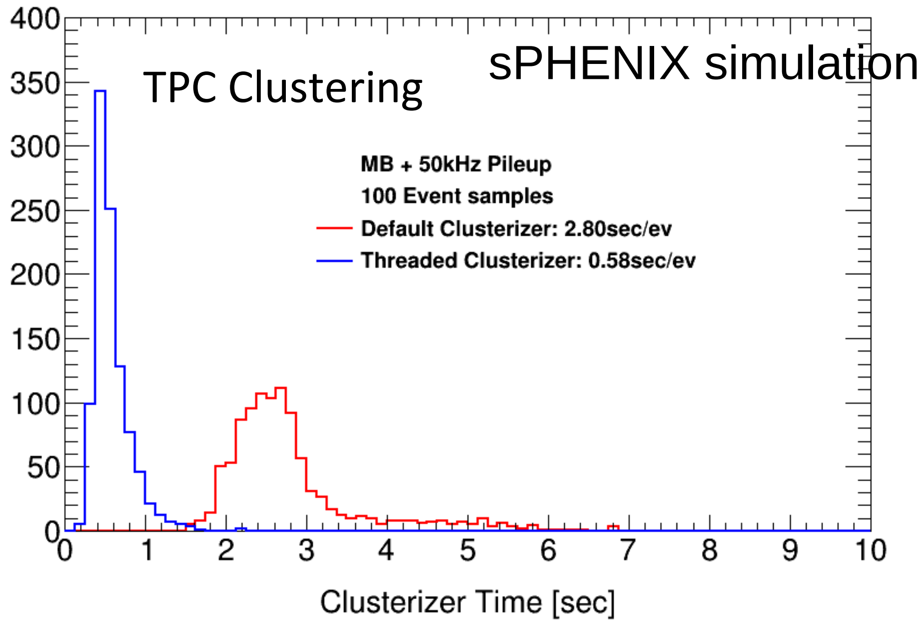




Meets our physics requirements

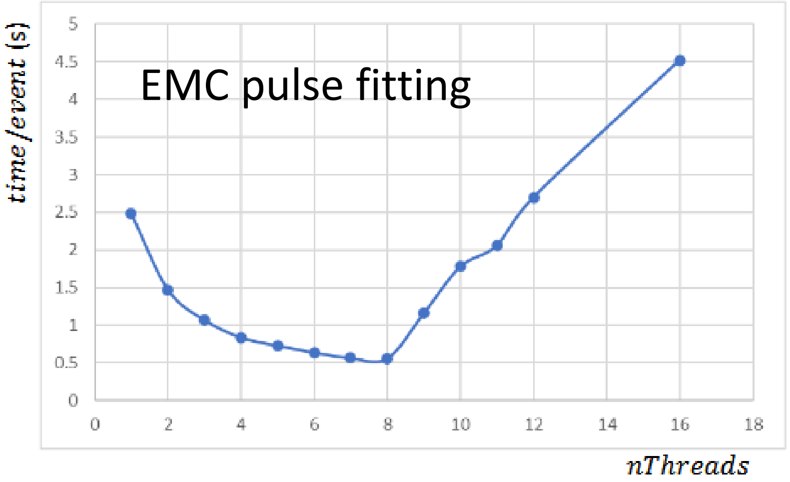
Multithreading

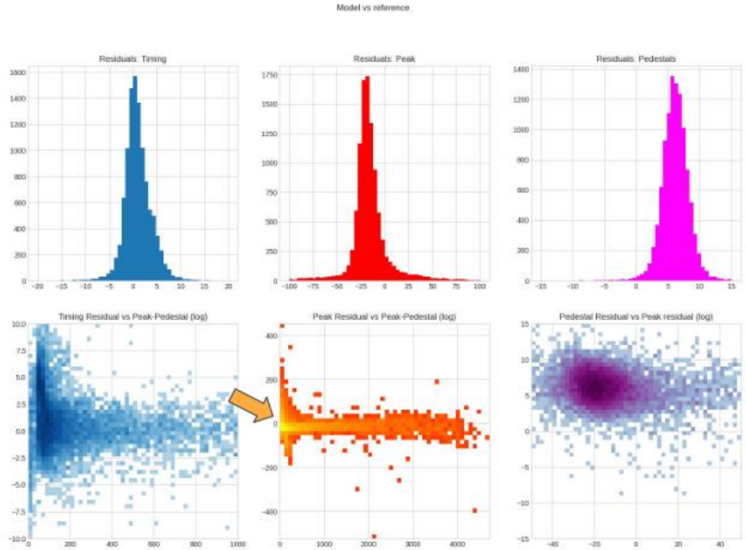
- Our memory is mostly consumed by heavy ion events, multi threading on event by event basis does not help
- Multi threading inside event loop does look promising
- But relies on the node having spare cycles to run those threads



Time per event versus nthreads

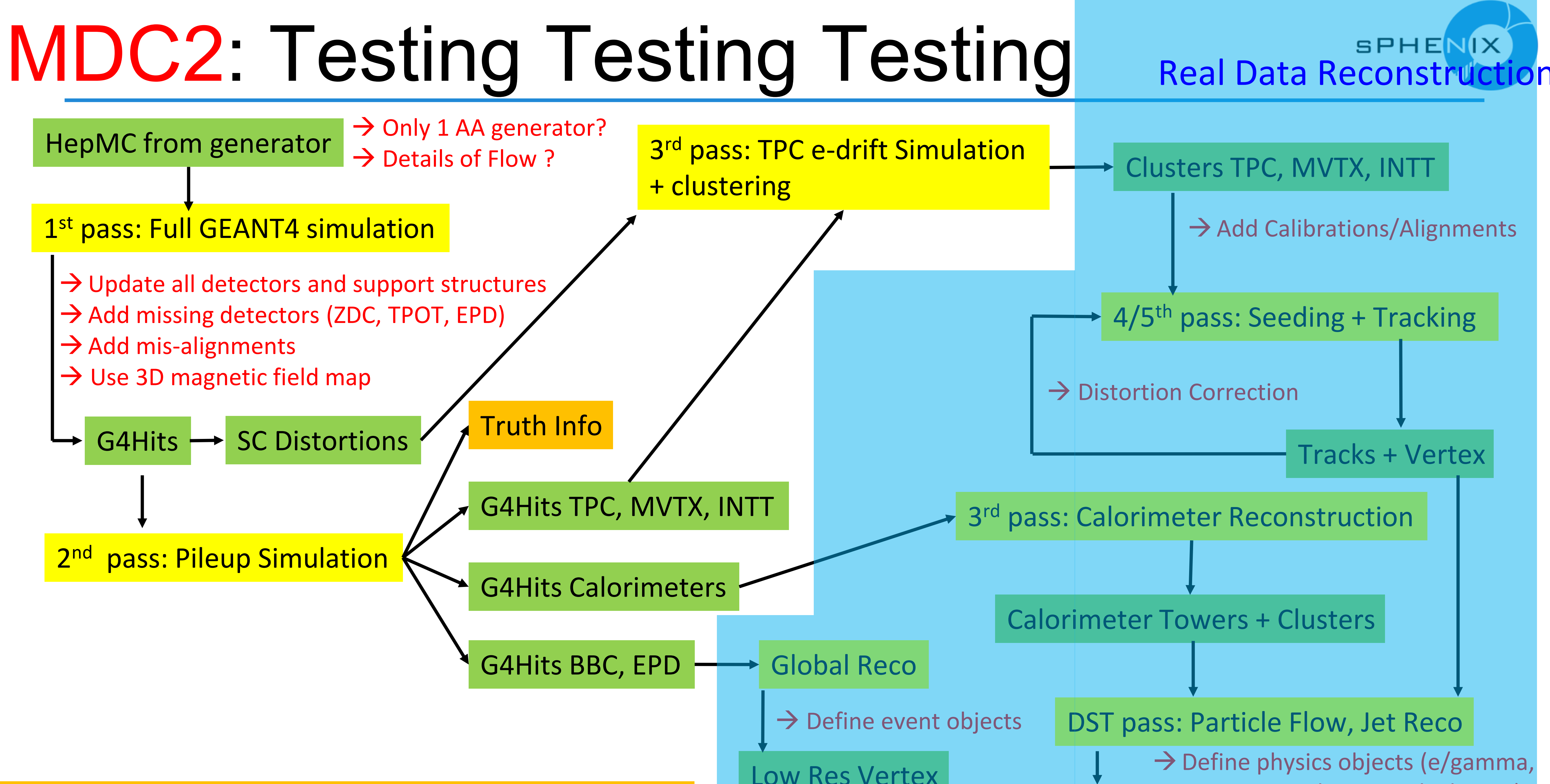
- Averaged over 50 events
- 24,576 channels evaluated per event





EMC Test Beam data pulse fitting with ML

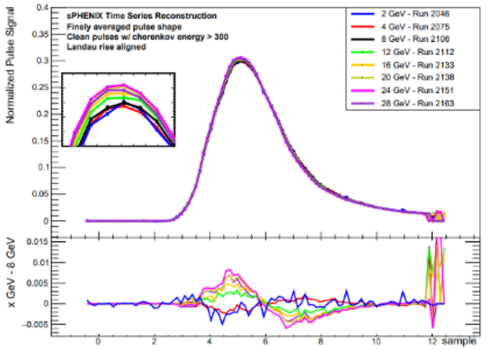
MDC2: Testing Testing Testing

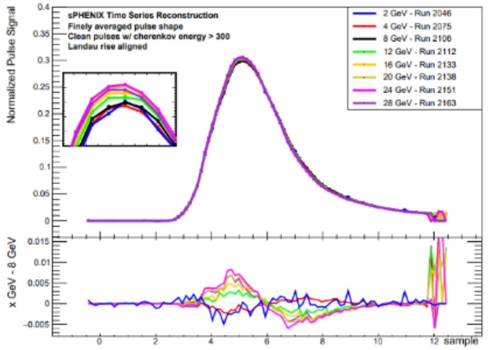


- No more Truth Info used during reconstruction
- Truth Info can be correlated during analysis

Production Workflow

Get a head start by processing calorimeter files separately





…x20

75% data reduction

75% data reduction

Calorimeter Clustering

x40

1st pass tracking

2nd pass tracking

Particle Flow

HPSS

x1

Aggregation

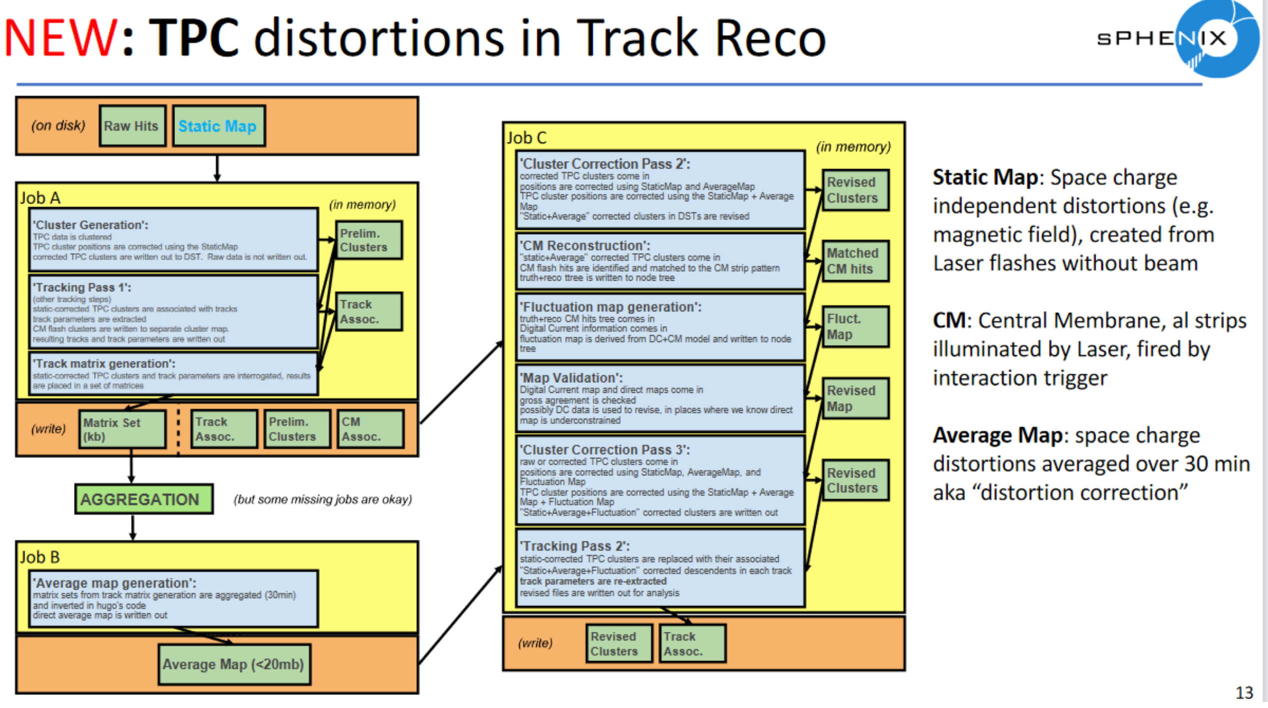
Analysis

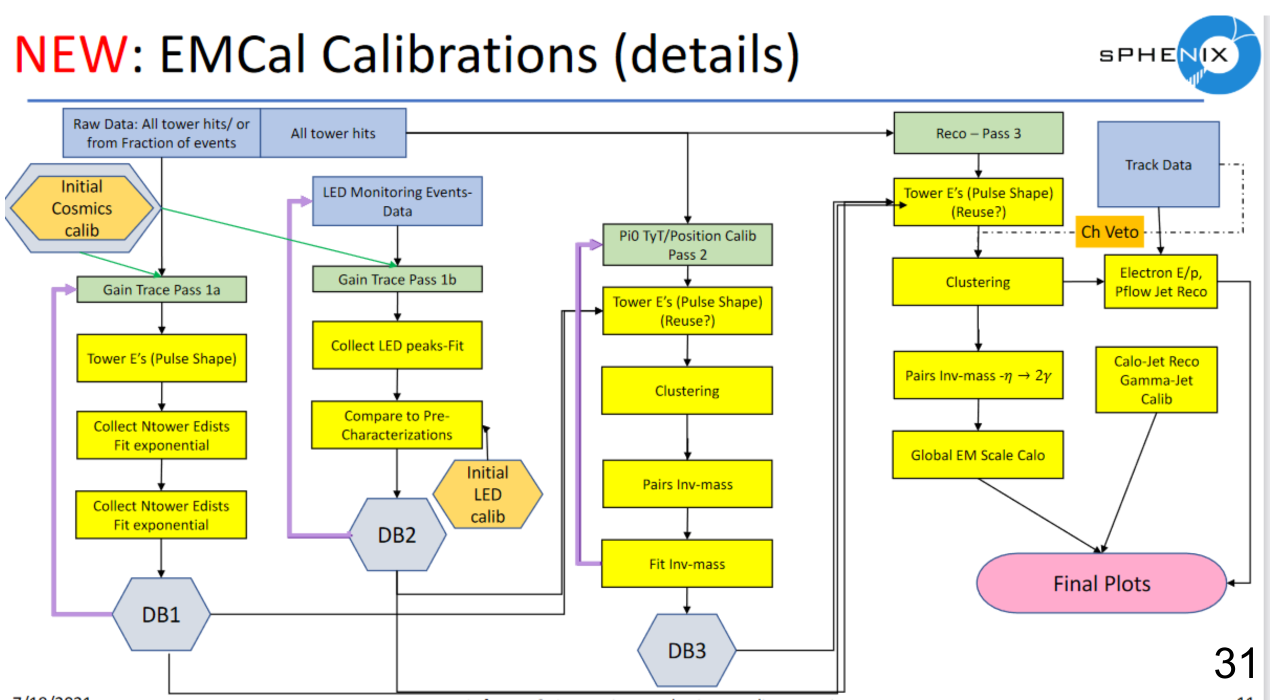
Reduced data format likely ROOT → random access

All DST content in single file to enable analysis from tape

Calibrations

- Two types of calibrations
  1. Distortion corrections (timescale 10ms)
  2. All others (timescale 5mins to years)
- 64 bit (BigInt) beam clock serves as Time stamp
  - event level granularity
  - Our events are time ordered – easy assignment of calibrations to raw data files (looking at first and last event)
  - Gaps in validity (beam off periods) but no overlapping validity ranges
- Distortion corrections
  - No plan to keep calibrations long term (huge data volume)
  - Output of Job A produces distortion calibration for Job C – easy 1:1 match (needs some initial accumulation but then is rolling average)
  - Reprocessing means redoing distortion corrections
  - No need for a conditions DB here – some naming convention will do (and filesystem which can handle this → MDC goal)

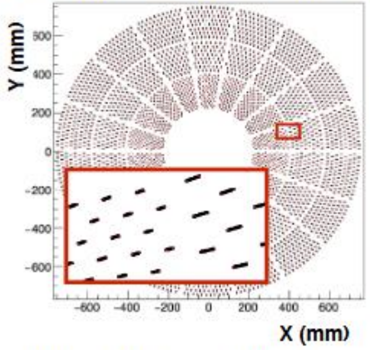




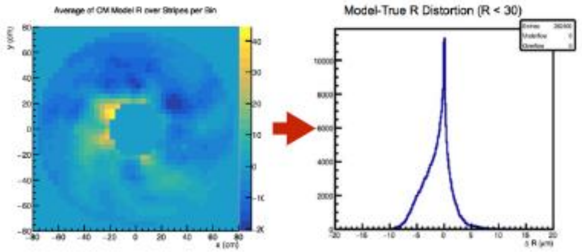
Space charge distortions

Diffuse Laser

- Al stripes on Central Membrane (CM)



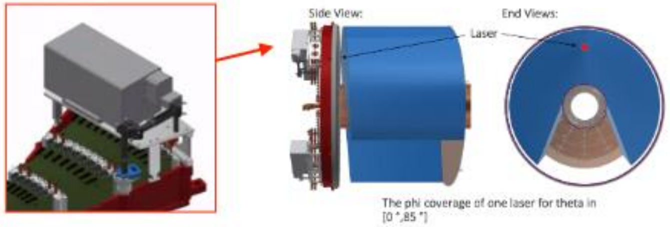
- ~kHz diffuse UV laser releases e-



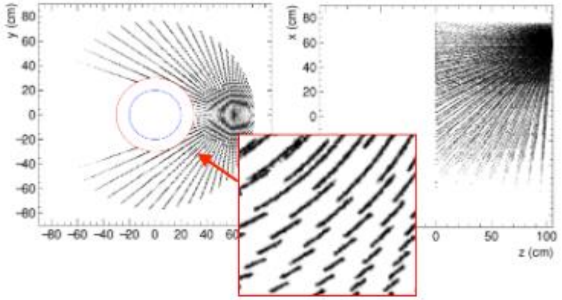
- stripes reco measures distortion at CM
- fast, interleaved with data => monitors *fluctuations*

Direct Laser

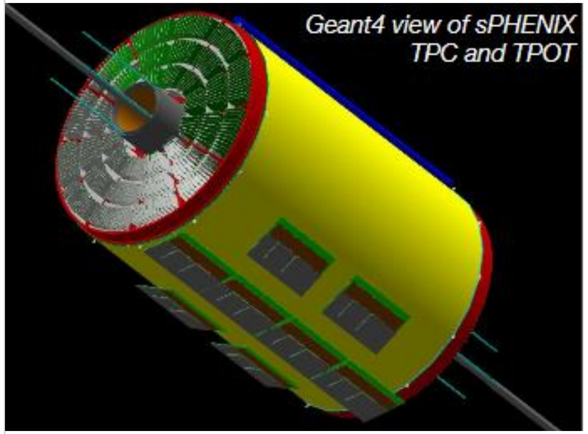
- 4x steerable lasers on each side of TPC

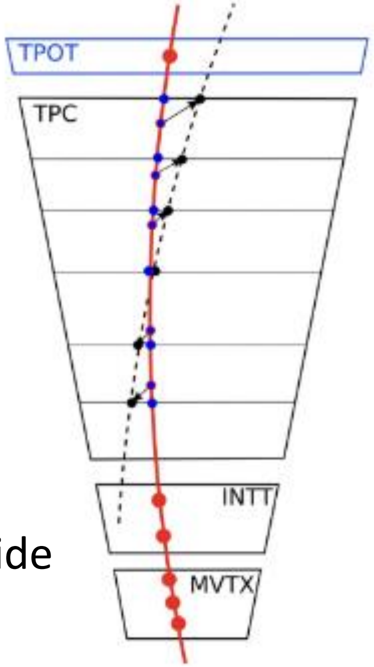


- ~all points can be reached by 2+ lasers
- creates straight-line tracks at ~Hz



- laser reco measures distortion in volume
- slow, not used during run => monitors *static distortion*





Additional Space point outside some areas of the TPC

Last not least: Analysis of the direct currents (continuous readout of the tpc pads)

Diffuse Laser fires in coincidence with an event, e- arrive shortly after e- from event

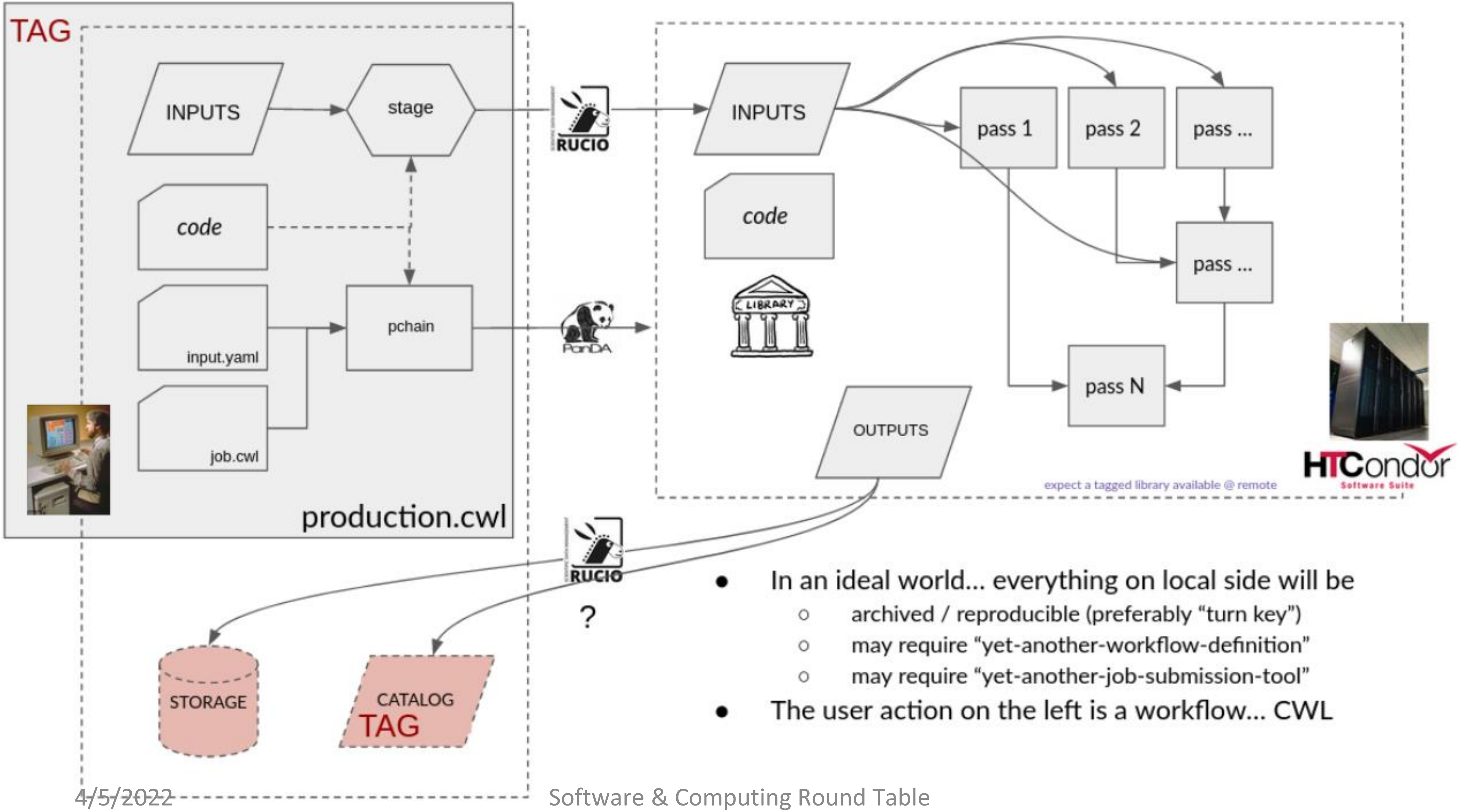
Correction essential for required momentum resolution for Upsilon program

Dataflow is the name of the game

- Our CPU estimate: 25sec/evt reco time
- Two reconstruction passes, near time during data taking and offline between RHIC Runs
  - 107B AuAu events in 24 weeks @ 25sec/evt : 184,290 cores
    - We cannot afford idling cores
  - Speed of the reconstruction is not speed of the job
    - Just run a job and see how long it takes to start up
    - Can be mitigated by longer running jobs (consideration for 24h length)
  - Cores idle during data copying
    - Everything needs to go through 10 Gb/s network
    - Need to work on pre-fetch and post processing copy schemes so cores always have data ready on local disk

sPHENIX Handy Remote Execution Koordinator

Jason Webb (NPPS)
+ PanDA team (NPPS)
+ Rucio team (SDCC)



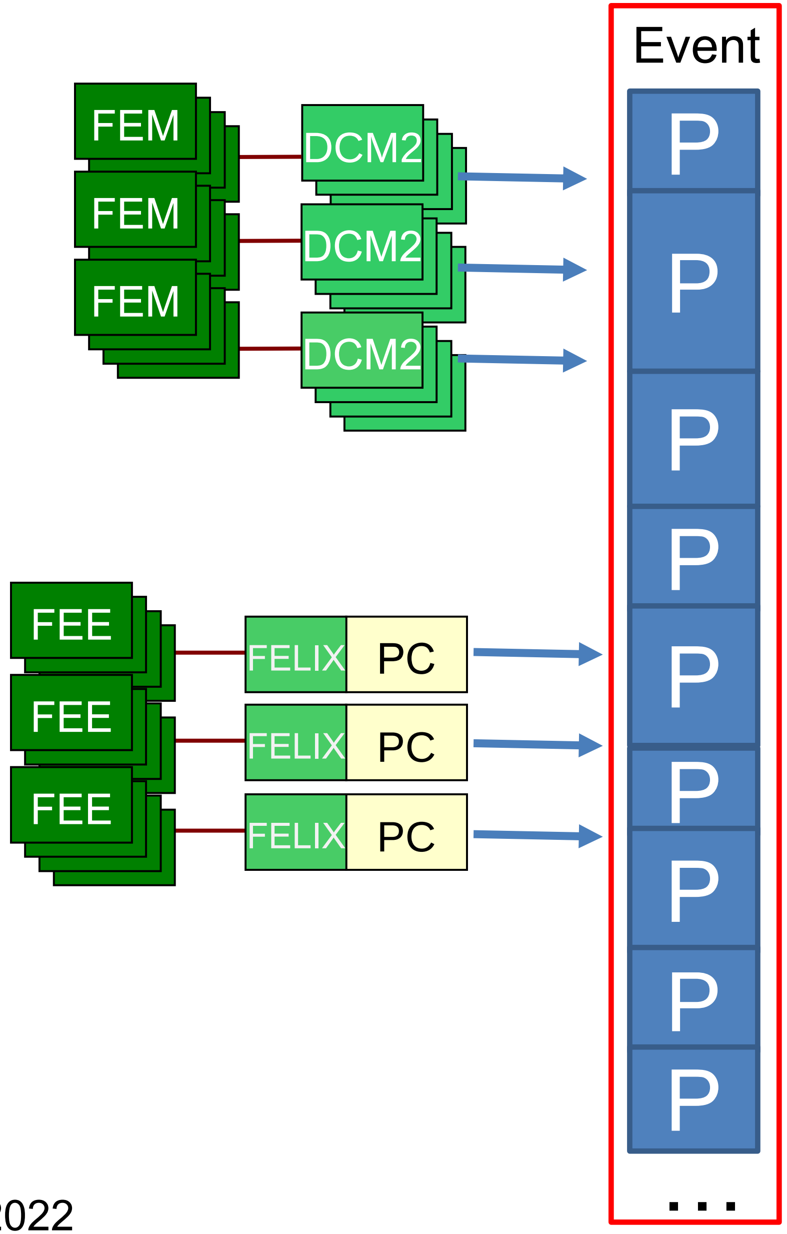
- In an ideal world... everything on local side will be
  - archived / reproducible (preferably "turn key")
  - may require "yet-another-workflow-definition"
  - may require "yet-another-job-submission-tool"
- The user action on the left is a workflow... CWL

Summary

- Full “near time” reconstruction planned followed by second offline pass
- Envisioned processing time per event well within reach
  - Two orders of magnitude speed increase compared to 3 years ago (Thanks: ACTS)
- Tools from the wider community (PanDA, Rucio)
- DB from Belle 2
- Data flow the remaining big challenge
  - Ongoing Mock Data Challenge for testing
- Guinea pig for the EIC

YOU WENT TOO FAR

Data Formats



Each Front-End Card contributes what we call a "Packet" to the overall event structures

A Packet ID uniquely identifies the detector component / front-end card where it comes from

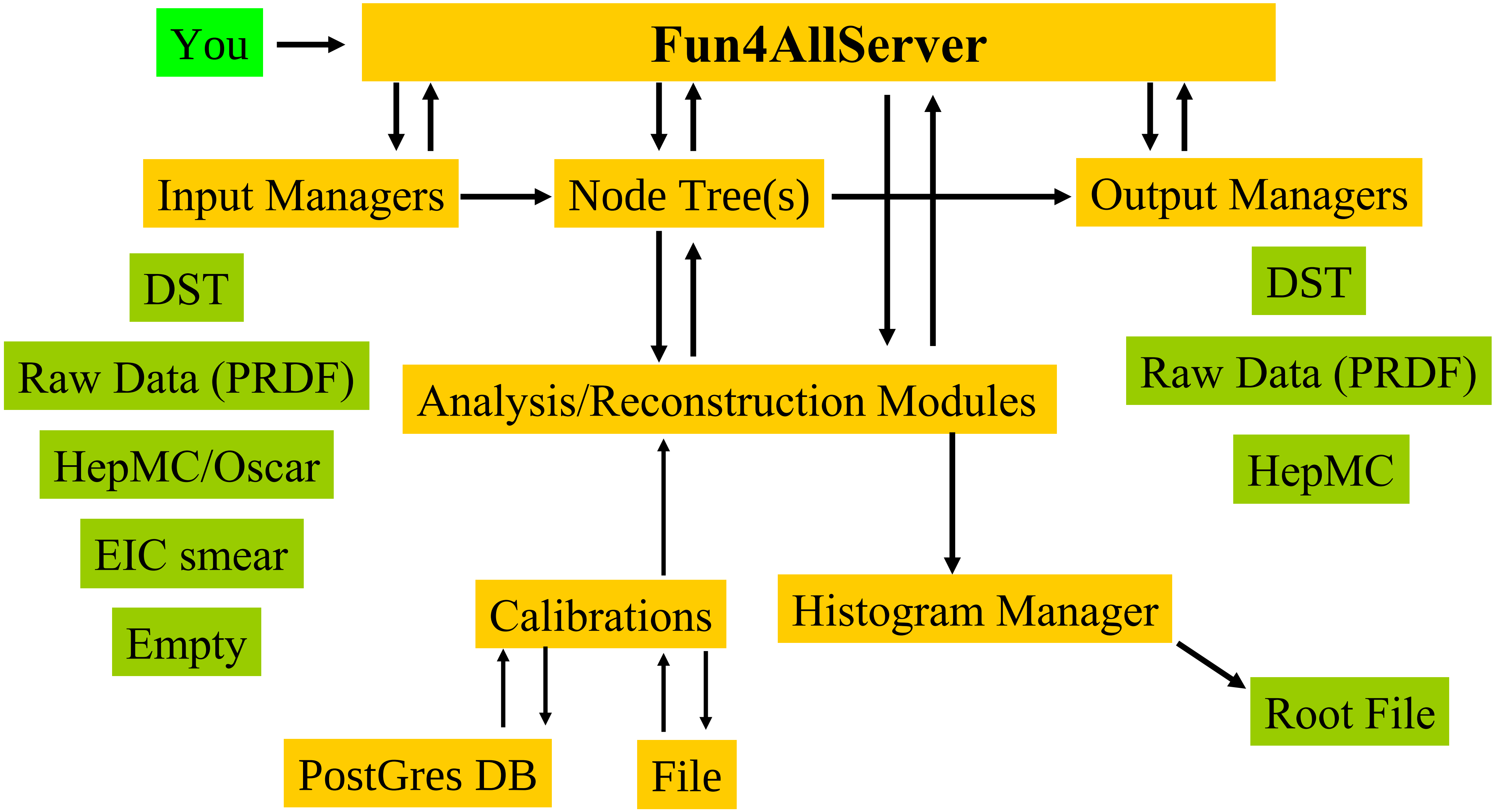
A hitformat field identifies the format of the data, und ultimately selects the decoding algorithm

We can change/improve the binary format and assign a new hitformat for a packet at any time

Insulation of offline software from changes in the online system

About 180-240 such packets in a typical sPHENIX event

Structure of our framework Fun4All



That's all there is to it, no backdoor communications – steered by ROOT macros

MDC2: Testing Testing Testing

HepMC from generator

→ Only 1 AA generator?
→ Details of Flow ?

1st pass: Full GEANT4 simulation

→ Update all detectors and support structures
→ Add missing detectors (ZDC, TPOT, EPD)
→ Add mis-alignments
→ Use 3D magnetic field map

G4Hits

SC Distortions

2nd pass: Pileup Simulation

Truth Info

G4Hits TPC, MVTX, INTT

G4Hits Calorimeters

G4Hits BBC, EPD

Global Reco

→ Define event objects

Low Res Vertex
Centrality
Reaction Plane

3rd pass: TPC e-drift Simulation + clustering

Clusters TPC, MVTX, INTT

→ Add Calibrations/Alignments

4/5th pass: Seeding + Tracking

→ Distortion Correction

Tracks + Vertex

3rd pass: Calorimeter Reconstruction

Calorimeter Towers + Clusters

DST pass: Particle Flow, Jet Reco

→ Define physics objects (e/gamma, jets, KFParticles, ParticleFlow....)

Jets,...

- No more Truth Info used during reconstruction
- Truth Info can be correlated during analysis

Conditions DB coming online

Open ended calibrations (e.g. alignment)

Online_GT (locked)

start=t1	start=t2	start=t3	APPEND IOVs, IOV_END
end=null	end=null	end=null	is not allowed

This is the original model, only start times are allowed, and only appending is allowed
t_insert > t_last

The client will define this as now() + DELTA, where DELTA needs to be defined to be big enough to account for latency etc. but small enough to be useful

Sensible would likely be “next run”, so some work to do on the client side

Thanks to Paul Laycock, Ruslan Mashinistov, Dmitri Smirnov (all NPPS)

Conditions DB

Run wise calibrations

Prompt_GT
(locked)

start=t1
end=t2

start=t2
end=t3

start=t3
end=t4

OVERLAPPING IOVs ARE
NOT ALLOWED!!

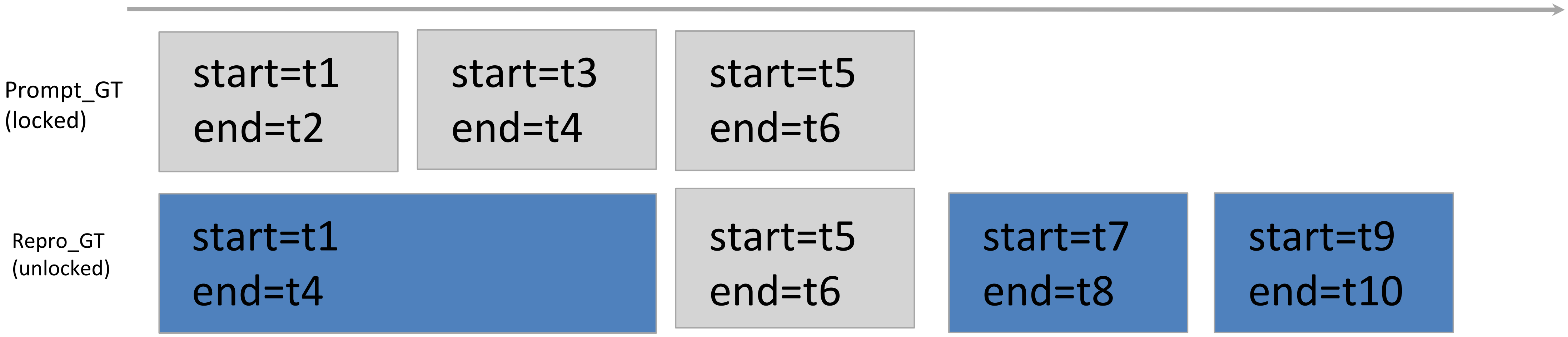
On writing, check that start time comes after the end time of the previous Payload_IOV, AND its end time comes before the start time of the next Payload_IOV - no over-writing so this is a safe operation

Important constraints - the end time of one unit of data will be numerically distinct from the start time of the next unit of data

The service will *always* return the closest PIOV object, it’s up to the client to decide whether to throw an ERROR because of the END of validity time

Conditions DB

Reprocessing preparation



Workflow - Clone the Prompt_GT, this creates an ***unlocked*** GT

You can do what you want with unlocked Global Tags :)

While editing, the service will take care of removing unnecessary PIOVs

Once editing is finished, lock the Repro_GT before use in production

sPHENIX Construction Schedule:

Sequential Build Steps	Location	Start	End	J-22	F-22	M-22	A-22	M-22	J-22	J-22	A-22	S-22	O-22
Complete oHCAL Sector Installation	AH	12/16/21 *Actual*	3/1/22 *Actual*										
Build the Carriage, Platforms, Pole Tip Doors *This CA-D organized team will work in parallel with the sPHENIX techs working on iHCAL*	AH	3/7/22	7/15/22				Pause						
Install the iHCAL Barrel *Install structures, barrel insertion into magnet bore*	AH	6/14/22	7/15/22		iHCAL Barrel Fab in 912			IHCal install prep					
Remove the Shield Wall *sPHENIX moved East→West →East →West two times to get all shield wall parts out*	AH/IR	4/4/22	4/29/22										
On-carriage detector infrastructure AH installs *Cable trays, piping, rack installations*	AH	3/7/22	7/15/22				Pause						
Magnet and Cryo Infrastructure in AH *Valve stem, cold box and other upper platform installations*	AH	6/2/22	7/15/22										
Move sPHENIX To the IR *sPHENIX moves west and north to final IR position*	AH/IR	7/18/22	7/22/22										
Magnet Cryo, Power, Cool Down & Testing in IR *Complete magnet infrastructure pre- and post- sPHENIX roll-in*	IR	4/22/22	9/9/22										
IR Detector Infrastructure *New chiller platform, cables, piping, racks on carriage, etc…*	IR	5/2/22	Through end of project										
EMCAL, TPC install structures, TPOT install in IR *TPOT installed last in order to use TPC installation structures*	IR	7/25/22	10/24/22									Pause	TPOT
Magnet Mapping *Favor magnet mapping over EMCAL install, finish ASAP and then restart EMCAL*	IR	9/6/22	9/27/22										

300k He available in 1008 7/18

Magnet He cool down begins 8/15

Magnet ready full power Ops

You are here

FY22 RHIC Run

Pause EMCAL install for magnet cool down & power testing, and mapping

CERN Team Arrival ~9/6

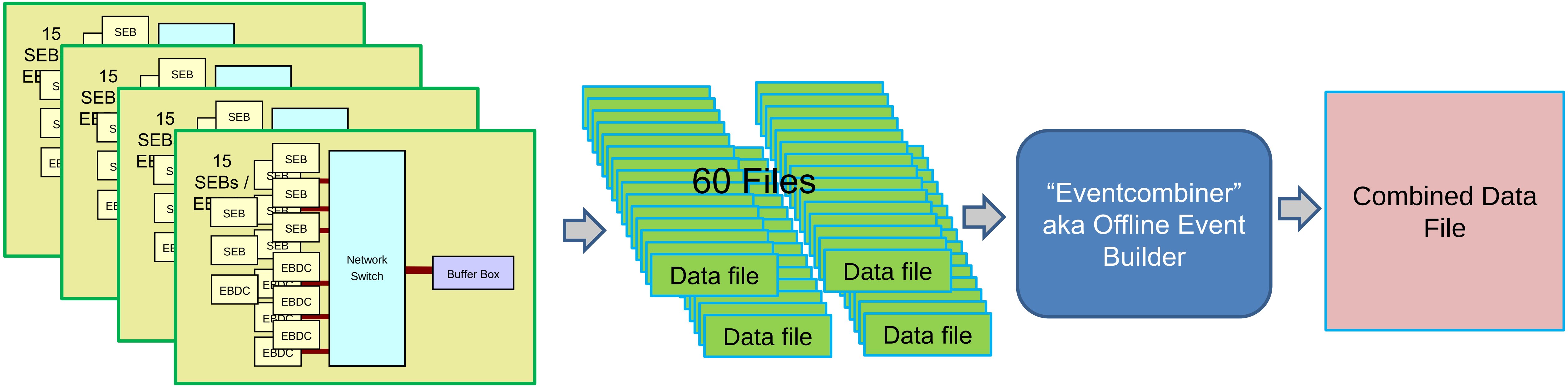
TPOT installed when EMCAL complete

Offline Event-Building from 60 streams

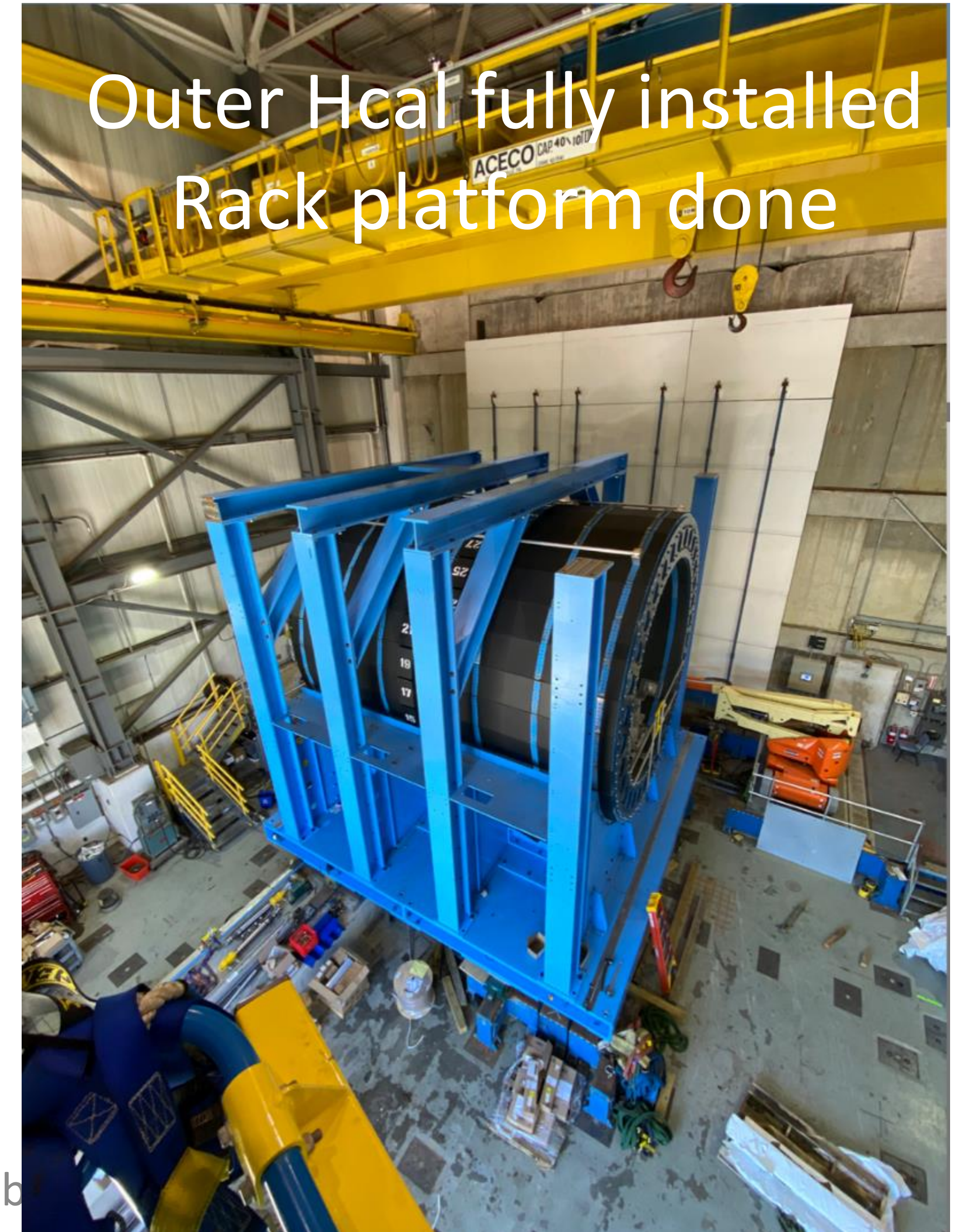
Use 15 SEBs/EBDCs = 15 streams

Configure 4 setups with different packet content that can be combined, mimicking the eventual data makeup

Run 4 "volleys" one after another to get 60 files, 240 packets combined, about the number of packets and event size we expect



sPHENIX Progress



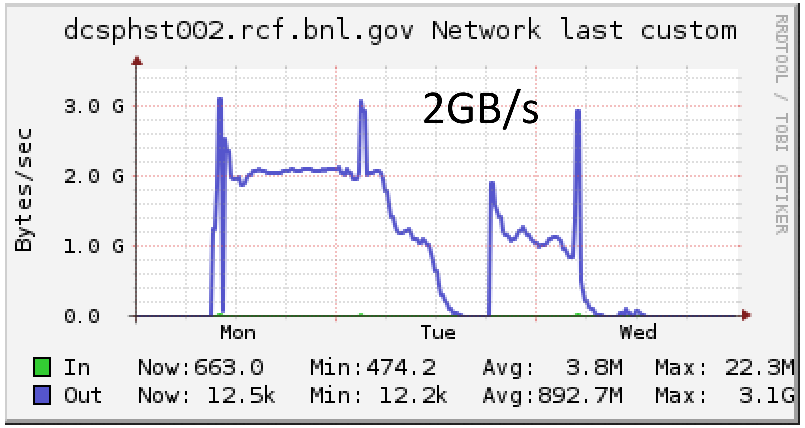
sPHENIX Progress

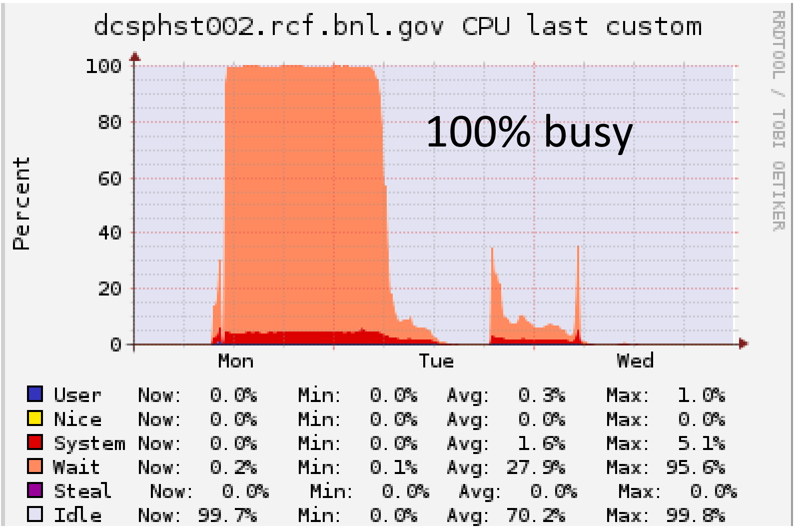
Prepping for EMCal installation into the IHCal

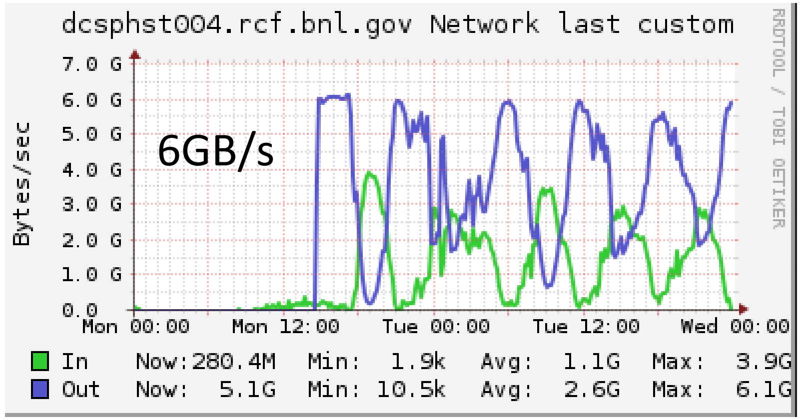
UIUC EMCal Block factory

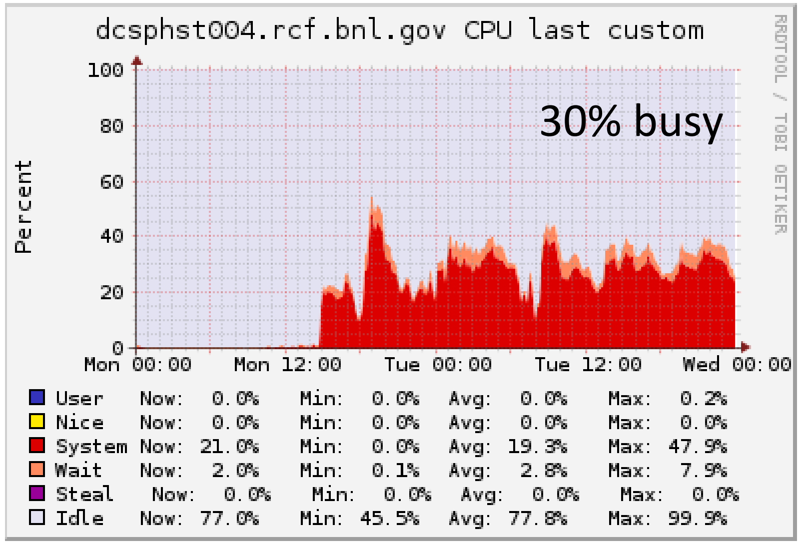
Final EMCal block

First MDC2 result (Help by Tejas Rao+Chris Hollowell)

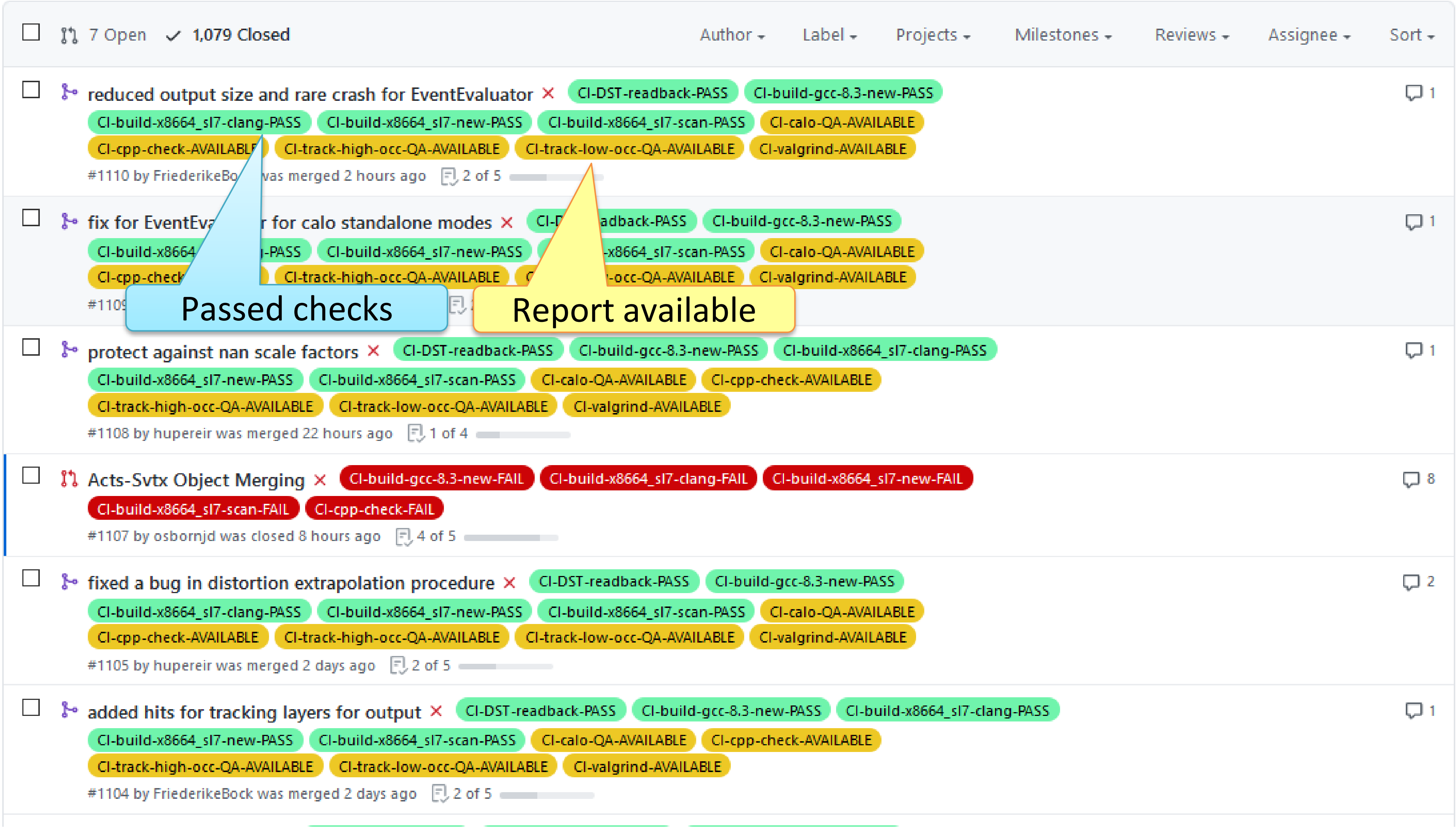








Automatic Pull Request Checks

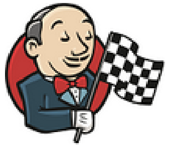


Automatic Pull Request Checks

SPHENIX

Build & test report

Report for commit 6ee3642614c3b27f3b8003e851660714a3044a44:



- build passing builds and tests overall are SUCCESS.
- build passing Build with configuration of gcc-8.3 / new is SUCCESS, Compiler report (full)/(new), build log
  - build passing Generating DST and readback: build is SUCCESS
  - build passing Calorimeter QA: build is SUCCESS
    - QA-calorimeter for e- at p_T=4GeV : combined Chi2/nDoF = -0 / 72, and combined p-Value = 1
    - QA-calorimeter for pi+ at p_T=30GeV : combined Chi2/nDoF = -0 / 72, and combined p-Value = 1
    - QA-calorimetric-jet for e- at p_T=4GeV : combined Chi2/nDoF = -0 / 42, and combined p-Value = 1
    - QA-calorimetric-jet for pi+ at p_T=30GeV : combined Chi2/nDoF = -0 / 42, and combined p-Value = 1
  - build passing Tracking QA at high occupancy: build is SUCCESS
    - QA-Intt : combined Chi2/nDoF = -0 / 72, and combined p-Value = 1
    - QA-Mvtx : combined Chi2/nDoF = -0 / 54, and combined p-Value = 1
    - QA-Tpc : combined Chi2/nDoF = -0 / 56, and combined p-Value = 1
    - QA-tracking : combined Chi2/nDoF = 473.316 / 38, and combined p-Value = 1.52312e-76
    - QA-vertexing : combined Chi2/nDoF = 168.818 / 98, and combined p-Value = 1.16321e-05
  - build passing Tracking QA at low occupancy: build is SUCCESS
    - QA-Intt : combined Chi2/nDoF = -0 / 72, and combined p-Value = 1
    - QA-Mvtx : combined Chi2/nDoF = -0 / 54, and combined p-Value = 1
    - QA-Tpc : combined Chi2/nDoF = -0 / 56, and combined p-Value = 1
    - QA-tracking : combined Chi2/nDoF = 518.111 / 42, and combined p-Value = 2.57133e-83
    - QA-vertexing : combined Chi2/nDoF = 313.342 / 98, and combined p-Value = 2.40642e-24
  - build passing system gcc-8.3 , build new : run the default EICDetector macro: build is SUCCESS, output
  - build passing system gcc-8.3 , build new : run the default fsPHENIX macro: build is SUCCESS, output
  - build passing system gcc-8.3 , build new : run the default sPHENIX macro: build is SUCCESS, output
  - build passing system gcc-8.3 , build new : run the overlap check for sPHENIX macro: build is SUCCESS, output
  - build unstable system gcc-8.3 , build new : Valgrind test: build is UNSTABLE, valgrind report
- build passing Build with configuration of x8664_sl7 / clang is SUCCESS, clang report (full)/(new), build log
- build passing Build with configuration of x8664_sl7 / new is SUCCESS, Compiler report (full)/(new), build log
- build passing Build with configuration of x8664_sl7 / scan is SUCCESS, scan-build report (full)/(new), build log
- build passing cpp-check is SUCCESS, cppcheck report (full)/(new)

Compilation with gcc 4.8, gcc 8.3 and clang

New warnings fail CI

Static code analysis with cppcheck, scan-build

New warnings fail CI

Runtime analysis with valgrind

Full: link to all warnings (legacy, false positives)

New: link to warnings from this PR

Work in progress (new G4 version)