Streaming modes for Run24

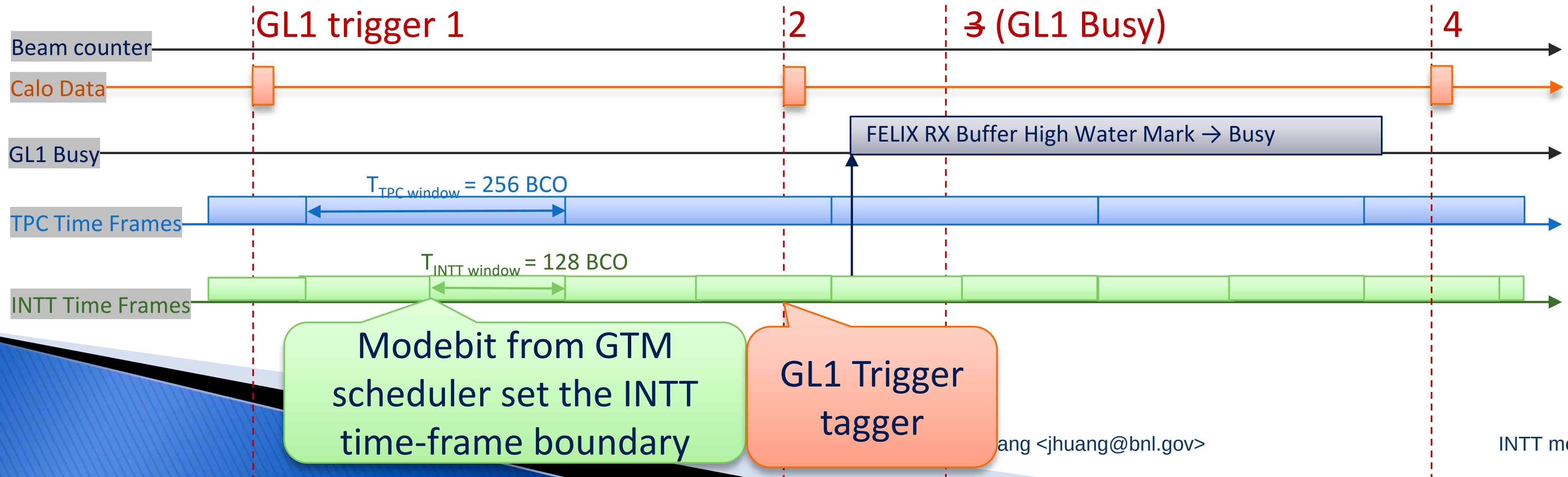
Jin Huang
BNL

Introduction

- As highlighted in beam use proposals (BUP), all low-pT heavy flavor physics in Run24 relies on streaming capability of sPHENIX trackers
  - e.g. low pT D0 that does not have clear signature in calorimeters, nor need them
- Minimally, we will collect 10% of all collision data via a trigger extension trick: record +7us of collision data after each GL1 trigger [see BUPs]
- Depending on the collision rate, there is a chance for TPC to extend the trigger window to record more collision data, up to 100% streaming [ref]
- Regardless TPC operation mode, it is much simpler INTT (and MVTX) record 100% of all hits above chip zero suppression threshold

Discussion

- Chatting with many (including Itaru and Raul), recommend to make 100% streaming for INTT Run24 default :
  - Deliver most physics regardless of TPC streaming modes and triggered program; least error prone (e.g. timing alignment)
- Required INTT work
  - Time-framing data with a periodic mode bit from GTM scheduler, which defines data frames (instead GL1 trigger in Run23);
  - 128BCO time frame width: i.e. within one time frame, FEE clock counter roll over max once
  - Tag GL1 trigger in INTT data stream
  - Contribute to the offline event building to fetch INTT hits for two streams: (1) hits corresponding to GL1 trigger (2) streaming tracker data
  - Active noise channel masking on chips, ensure data rate is small in 100% streaming operation (~1Gbps<<TPC data rate)
- INTT is the ONLY tracking detector to provide per-crossing timing resolution; it is very important to ensure sub-BCO timing alignment of all INTT chips



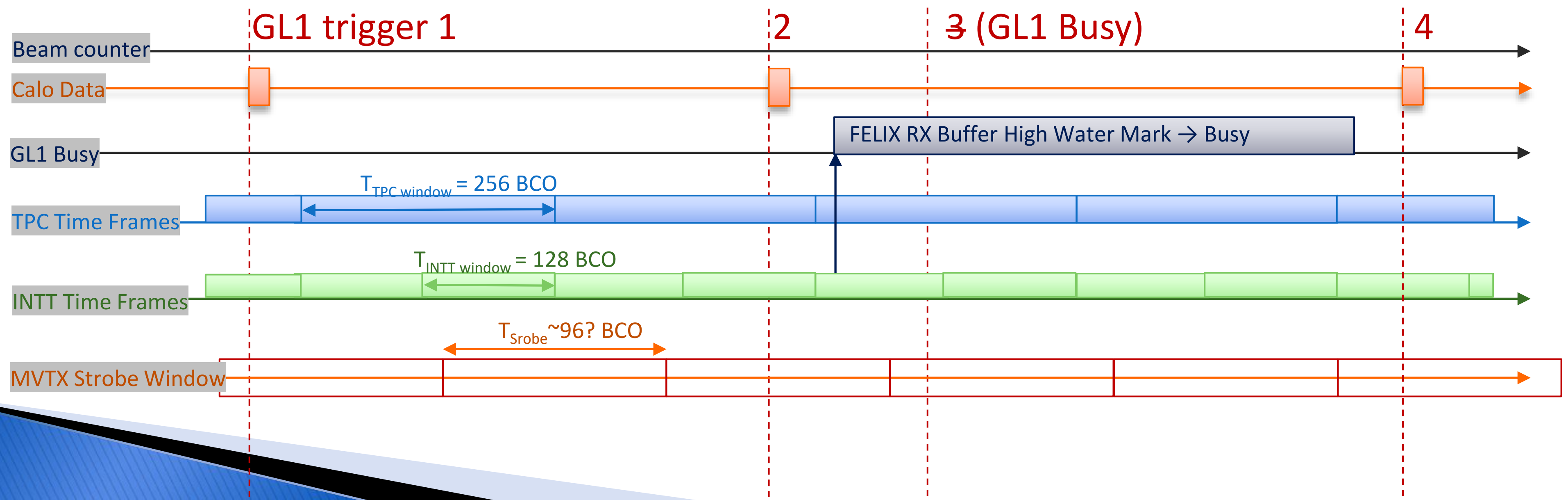
Other ask: add an INTT status badge in Grafana

https://wiki.sphenix.bnl.gov/index.php/Operation_Analytics_Site_(Grafana)

Extra Information

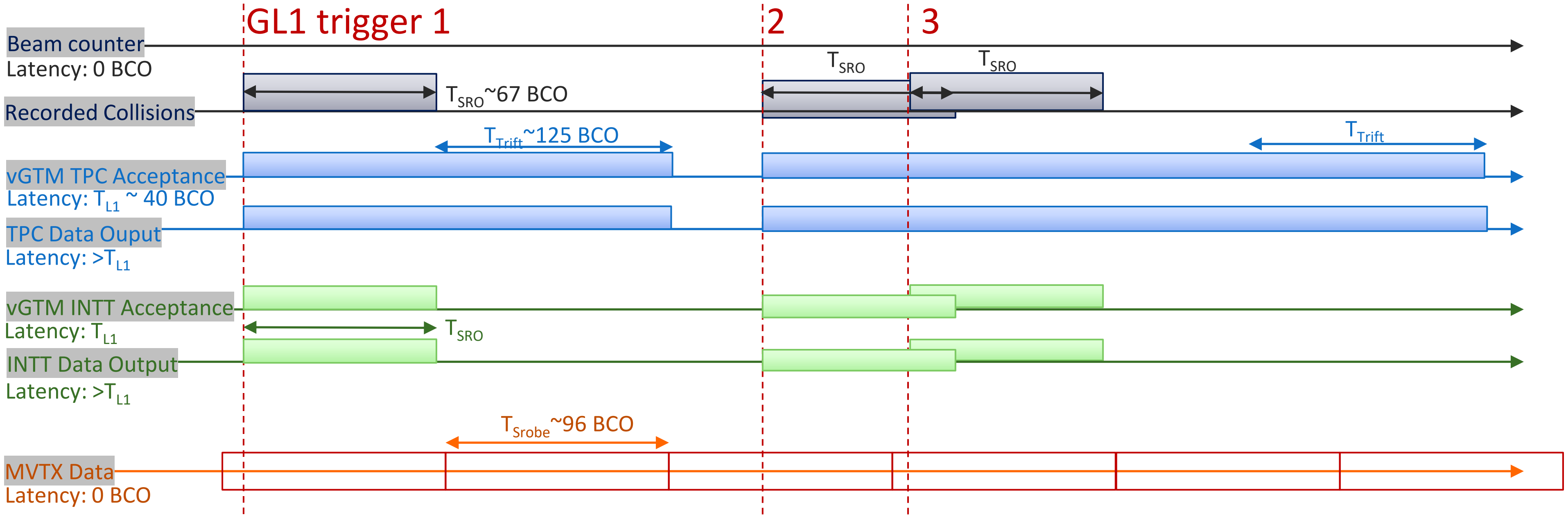
100% streaming operation mode

- GL1 triggered calo readout and tagger on trackers
- Clock triggered vGTM scheduler ModeBit-X to mark time-windows on tracker data
  - 128BCO period for INTT (ASIC BCO Rollover), 256BCO period FA for TPC (>1 drift length)
- Readout Capability Needed:
  - Record taggers of GL1 and ModeBit-X in FELIX data stream (for offline time-window sync)
  - FELIX busy feedback that holds back GL1 trigger if buffer high watermark reached
  - FELIX data stream mark of incomplete-readout for time windows and run stops (scalar sync)



SRO-Mode1-Simple [Recommended]

- vGTM distribute Streaming-Acceptance cross tracking detectors in sync with calo triggers
- Comparing the triggered-mode: simply prolong L1-Acceptance signal to each subsystem, from 1 BCO to T_{SRO}~67 BCOs



CAD production update and implications

https://indico.bnl.gov/event/22290/contributions/87220/attachments/52567/89901/nagle_trigger_update_02-07-2024.pdf

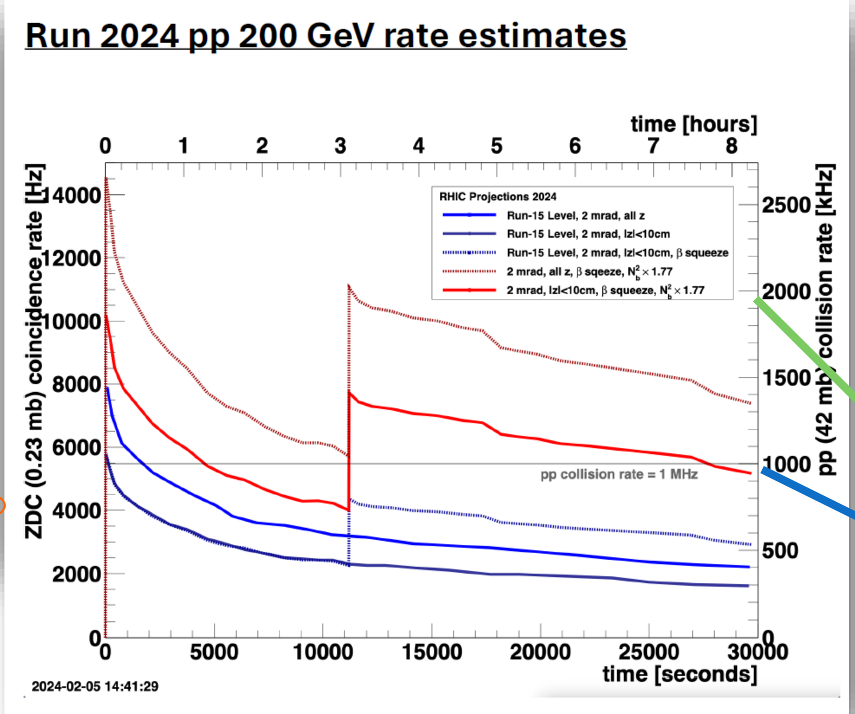
C-AD has some tricks up their sleeves (but no promises).

Increase protons per bunch by 1.33 and hence N_b^2 by 1.77

Beam-beam limit is lower due to crossing angle at sPHENIX, and so also no e-lens needed.

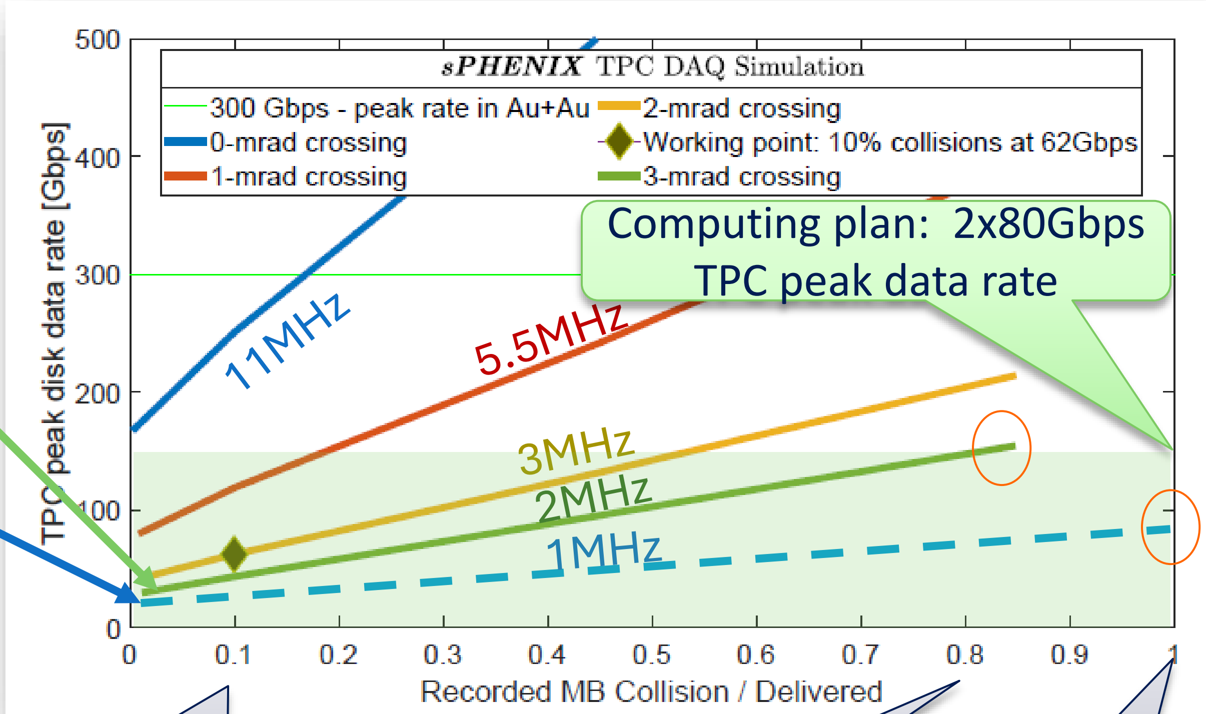
This gives a 77% increase, which on top of the b* squeeze, would give an increase of 247% in delivered luminosity / week.

☹ Again, it has unknown background implications. Needs to be tested.



BUP 22 Fig C.2

+ Additional New 2024 Col.Rate estimations



Computing plan: 2x80Gbps TPC peak data rate

Default 10% trigger-throttled streaming in BUPs

Beyond 50%, throttled streaming is messy, see backup slide 6

100% streaming

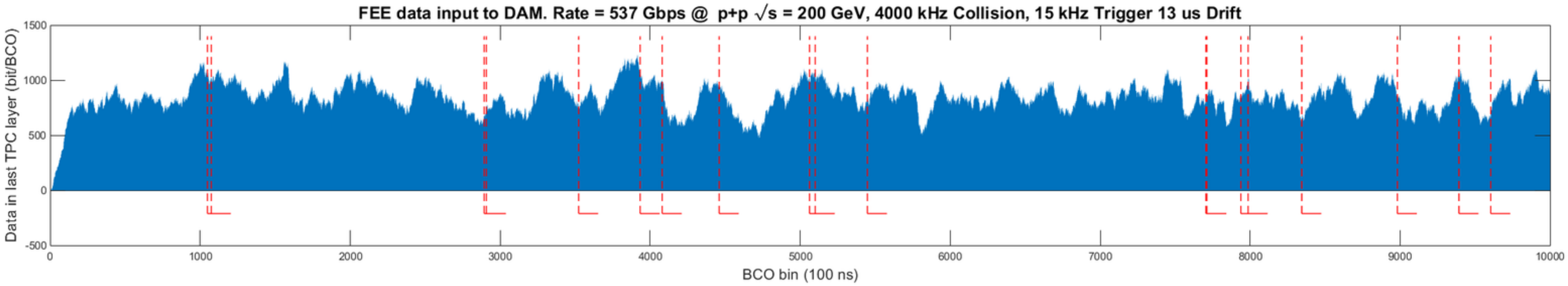
Discussion

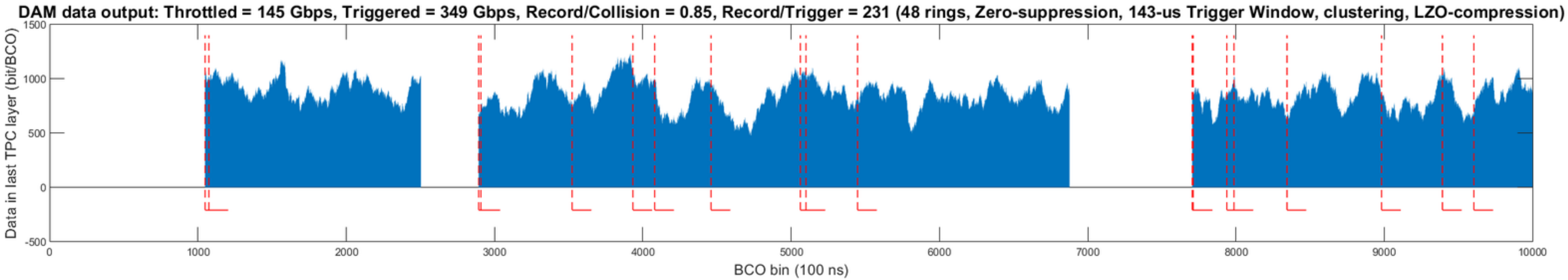
- Many steps to enable TPC streaming readout operation (regardless whether record 10%-100% of data at the backend)
  - Lower detector noise: tremendous progress by Evgeny, Charles, Tom et. al.
  - Study analog behavior: as we speak at https://chat.sdcc.bnl.gov/sphenix/channels/tpc-signal-processing
  - In-ASIC baseline correction: Takao/Hugo tested on TPOT, need firmware update and test in TPC env.
  - In-ASIC zero suppression: Takao tested in early firmware, need firmware update and test in TPC env.
  - FELIX-FEE synchronization: need firmware update, on-going
  - (Trigger throttle mode only) Trigger throttle and digital current calculation: conceptual (see Takao's slides)
  - RCDAQ multi-thread processing : Martin's work on-going, Run23 was limited to ~60Gbps
- Online priority suggestion:
  - Choose whatever operational mode that maximize the potential physics and fit below hardware limitation
  - (If TPC bgd consistent with simulation) that means 100% (<~1MHz collision) or throttled streaming (>1MHz collision) to max DAQ bandwidth
- Offline tracking priority suggestions:
  1. The most important preparation is that triggered event can be tracked, bench marked (majority of sPHENIX physics program). Even in 100% streaming, the triggered TPC data-frame can be extracted as single 13us time-frame regardless
  2. Ensure extended streaming data is good quality (some quality study in collaborate detector check out).
  3. Capability to reconstruct extended trigger frame
  4. Last priority is the capability to reconstruct 100% streaming time-frames

Throttled and 100% streaming modes for Run24

- INTT and MVTX plan aim to use 100% streaming in pp collisions
- For TPC: in BUP, we assume 10% trigger-throttled streaming operation using extended trigger window scheme
  - Responsible for half of HF physics output; and extended trigger window scheme is extremely efficient in cost-benifit at low (e.g. 10%) streaming level.
  - Record 7us extra M.B. crossing data per TPC trigger (on top of 13us drift time)
  - Lead to recording 10% of completed TPC beam crossing at 3MHz collision, and 30% of TPC data ([7+13us] x 15kHz)
- In the past months, it become more likely that collision rate will be low, open opportunity and need of 100% streaming for TPC
  - If collision drop from 3MHz to 1MHz (next page), we could increase fraction TPC data recorded from 30% to 100% at same output rate
  - Benefit: Simplicity online and offline; Avoid any gaps of TPC data (each cost 13us of physics); Best digital current recording ever; max tracks for TPC distortion corr.; max MB physics output

85% SRO @ 143us extended readout window





At 85% streaming, trigger frame is 10x TPC drift, overlap between trigger frame lead to significant overlap between trigger frames (more than x2 increase in processing if unpacked to individual trigger frames)

Why do we care about 100% streaming now?

- In BUP, we assume 10% trigger-throttled streaming operation using extended trigger window scheme
  - Responsible for half of HF physics output; and extended trigger window scheme is extremely efficient in cost-benifit at low (e.g. 10%) streaming level.
  - Record 7us extra M.B. crossing data per TPC trigger (on top of 13us drift time)
  - Lead to recording 10% of completed TPC beam crossing at 3MHz collision, and 30% of TPC data ([7+13us] x 15kHz)
- In the past months, it become more likely that collision rate will be low, open opportunity and need of 100% streaming
  - Collision drop from 3MHz to 1MHz, we could increase fraction TPC data recorded from 30% to 100% at same output rate
  - Benefit: Simplicity online and offline; Avoid any gaps of TPC data (each cost 13us of physics); Best digital current recording ever; max tracks for TPC distortion corr.; max MB physics output
- 100% streaming is only needed if collision rate is low relative to DAQ throughput, and depends on nominal features that is still to be completed (zero suppression in TPC, hit sync, busy feedback)

Possible Tracker Run Modes in Run24

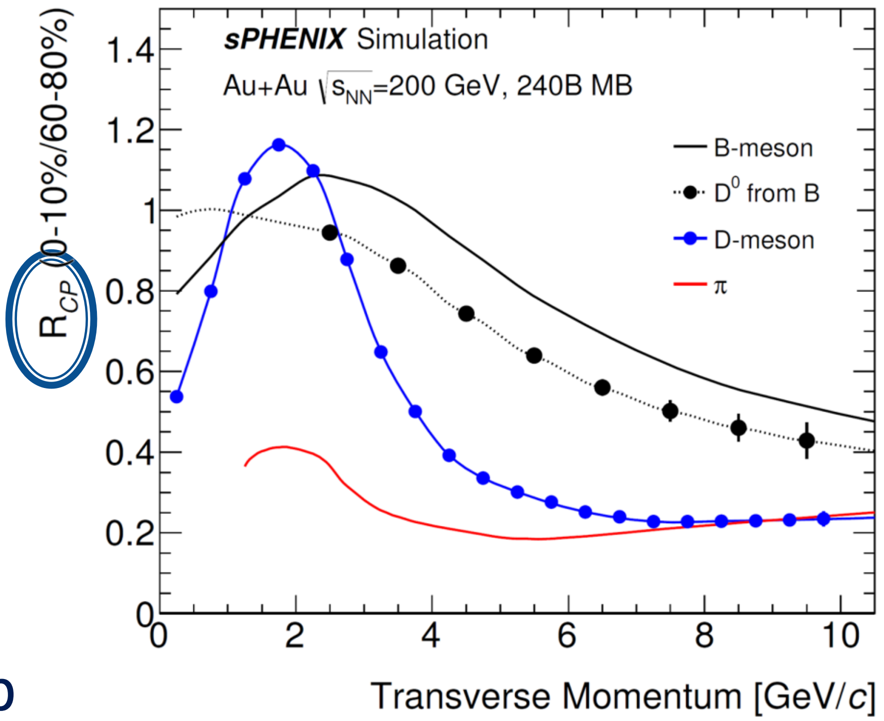
- Both RHIC performance and sPHENIX Beam configuration is still uncertain at this moment. Efforts are made to improve the collision rate
- If collision rate is high:
  - “High” means 100% streaming would **overflow** sPHENIX DAQ recording rate for majority time of a fill
  - We proceed with BUP plan of ~10% trigger-throttled streaming
- If collision is low:
  - “Low” means 100% streaming would **fit within** sPHENIX DAQ recording rate for majority time of a fill
  - Use 100% streaming mode
  - Ask CAD to luminosity level beginning of fill to the max DAQ recording rate
- If collision rate is in-between, we could choose to either
  - Trigger-throttled streaming at beginning of RHIC fill
  - 100% streaming at the later part of the fill

Part I - The physics benefits

- In p+p 200 GeV, recovering the capability of low pT HF physics via vastly increasing the M.B. data sample
- Updated to our write up on overleaf

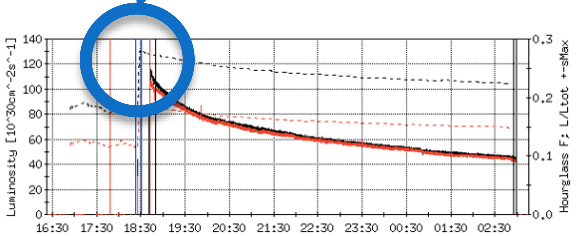
From the last collaboration meeting

- sPHENIX p+p √s=200 GeV at max ~13 MHz collision rate, 200 pb^{-1}
- For analysis using calorimeter signature → usually there is a trigger, e.g. b-jets
- HF meson/baryon: is very hard to trigger in current sPHENIX setup.
- Therefore, the baseline program does not assume p+p data.

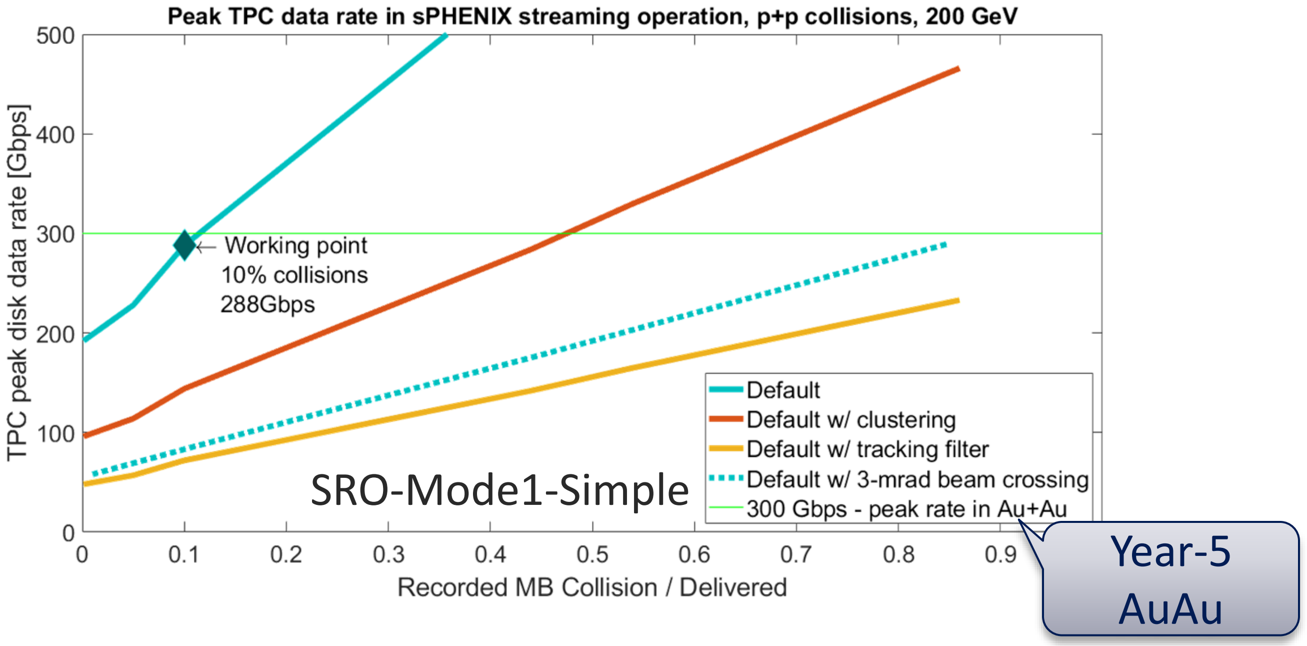


Choice of working point for the dominating data source: TPC @ 10%

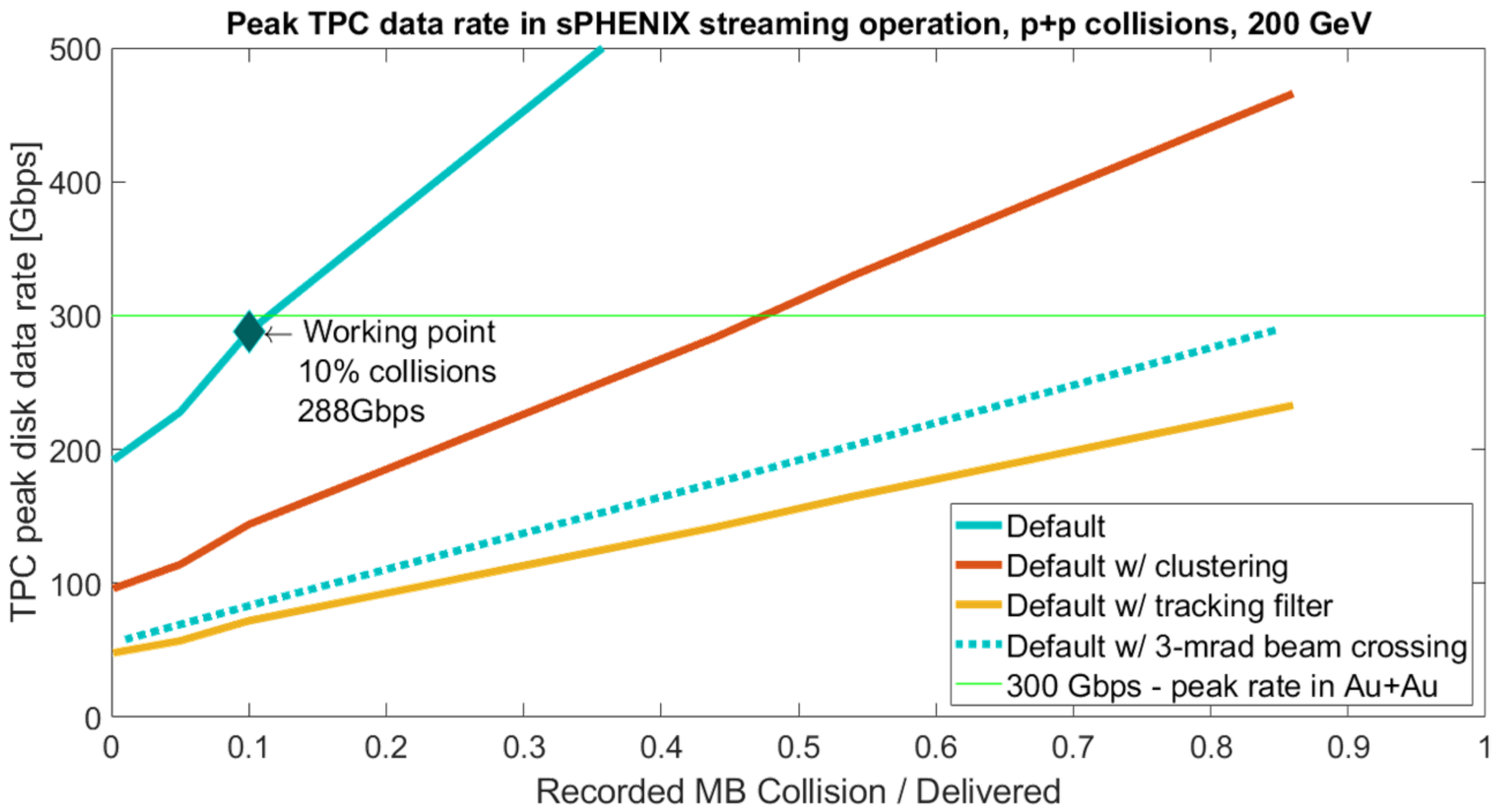
- Consider streaming recording 10% data → peak data rate~ AuAu
- Most our data is out-of-vertex data → narrowing vertex allow collecting dramatically more in-vertex physics
- Higher disk data rate, higher compression → high physics output
- Note this plot is for fill-peak data rate (13MHz collision) used to spec DAMs. It is higher than fill-averaged rate and run-averaged rate in the last talk



Run-15 p+p at 200 GeV store



Rate vs working points



HF Observables in p+p: comparison

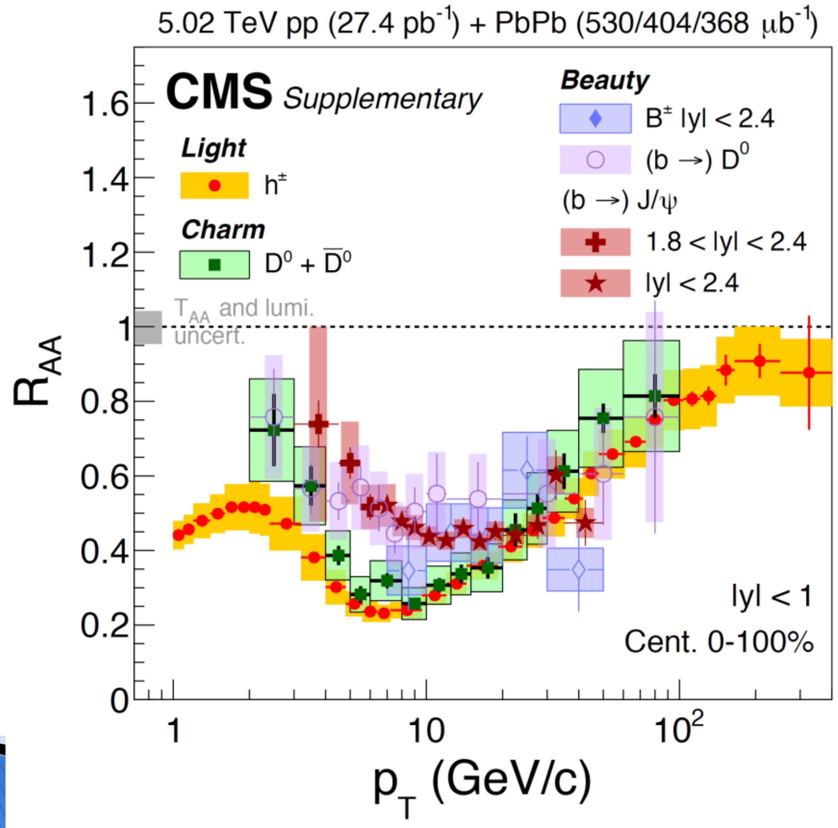
<table>
<tr><th></th><th></th><th>Year-2 , 0-crossing
in current setup
Per-kHz M.B. trigger</th><th>Year-2, 2mrad-crossing
in current setup
Per-5kHz M.B. trigger</th><th>Year -2
w/ Streaming tracker
(in this projection)</th><th>Year -2+4
w/ Streaming tracker</th></tr>
<tr><td rowspan="2">M.B. $p+p$</td><td>Data recorded</td><td>Each 1k Hz M.B. trigger with 2×10^{-4} of M.B. collisions triggered</td><td>Each 5k Hz M.B. trigger with 2×10^{-4} of M.B. collisions triggered</td><td colspan="2">10% M.B. events streaming recorded</td></tr>
<tr><td>Statistics</td><td>0.4 Billion M.B. events
0.01 pb^{-1} recorded</td><td>13 Billion M.B. events
0.15 pb^{-1} recorded</td><td>200 Billion M.B. events
5 pb^{-1} recorded</td><td>800 Billion M.B. events
20 pb^{-1} recorded</td></tr>
<tr><td rowspan="8">Physics reach</td><td rowspan="2">$B \rightarrow D^0 \rightarrow \pi K$</td><td rowspan="2">250 events</td><td rowspan="2">3.8k events</td><td>120k events</td><td>500k events</td></tr>
<tr><td colspan="2">Reference in R_{AA} for B→D^0</td></tr>
<tr><td rowspan="2">$D^0 \rightarrow \pi K$ pair</td><td rowspan="2">250 events</td><td rowspan="2">3.8k events</td><td>120k events.</td><td>500k events</td></tr>
<tr><td colspan="2">Diffusion of c-quarks in angular space</td></tr>
<tr><td rowspan="2">$\Lambda_c \rightarrow \pi K p$</td><td rowspan="2">500 events</td><td rowspan="2">8k events</td><td>250k events.</td><td>1M events</td></tr>
<tr><td colspan="2">Charm hadronization in p+p; reference for $A+A$</td></tr>
<tr><td rowspan="2">Prompt D^0→πK</td><td rowspan="2">75k events</td><td rowspan="2">1.1M events</td><td>40 Million events.</td><td>150 Million events</td></tr>
<tr><td colspan="2">Pinging down tri-gluon correlation via single spin asymmetry</td></tr>
</table>

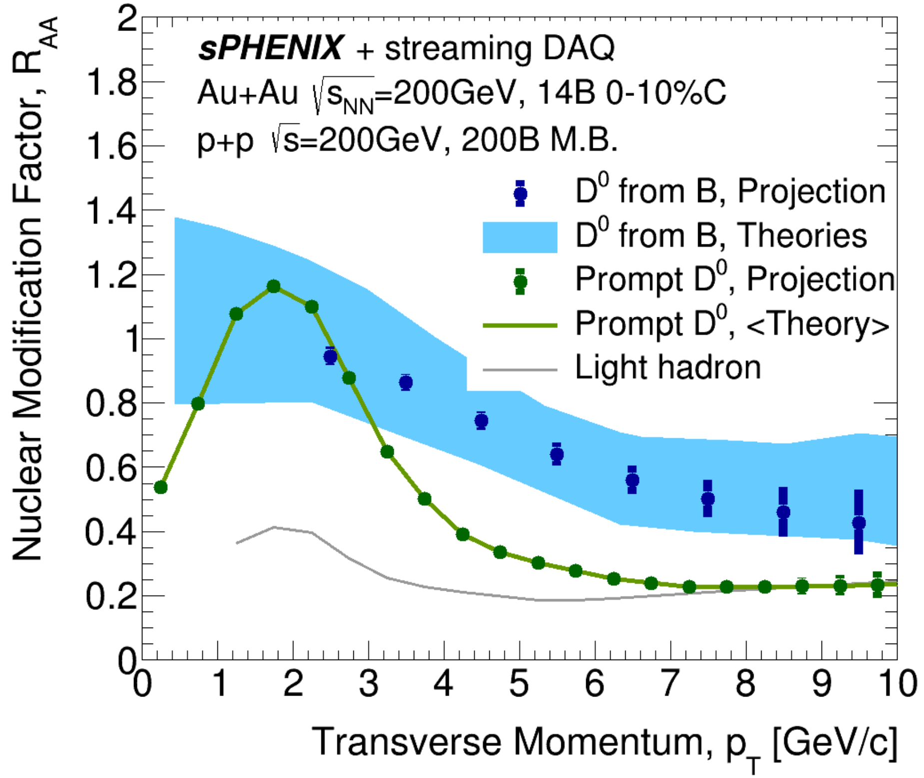
sPHENIX

Recovering HF meson baseline

- Precision b-physics → precision baseline and systematic control

See also: talk X. Dong

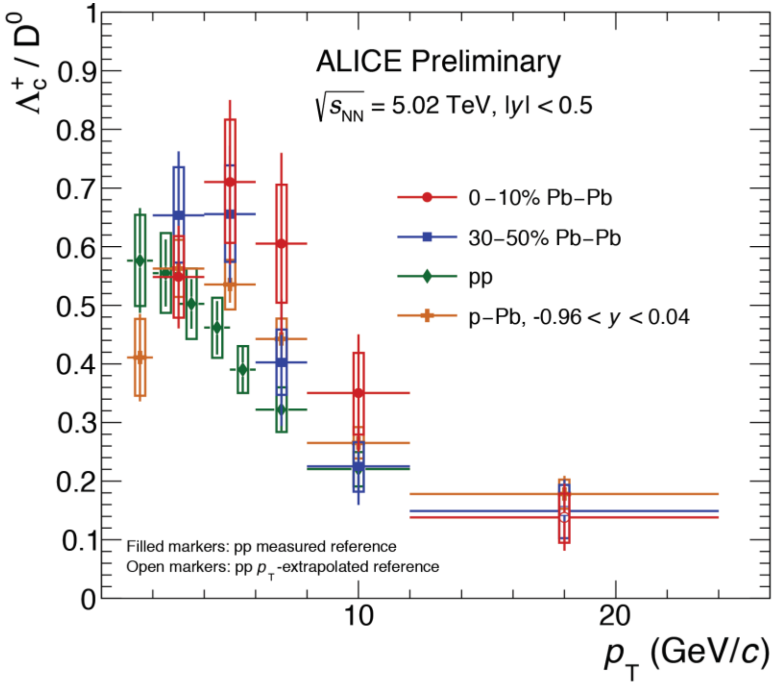


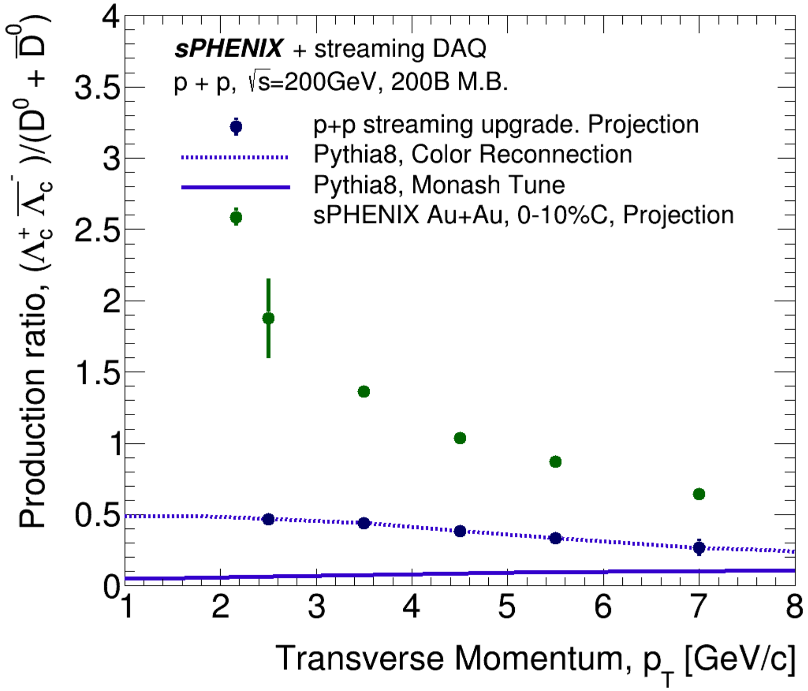


Λ_c and charm hadronization

- Recent data underscores importance of understanding Λ_c in pp and small system in interpretation of AA data

See also: talk Y. Ji

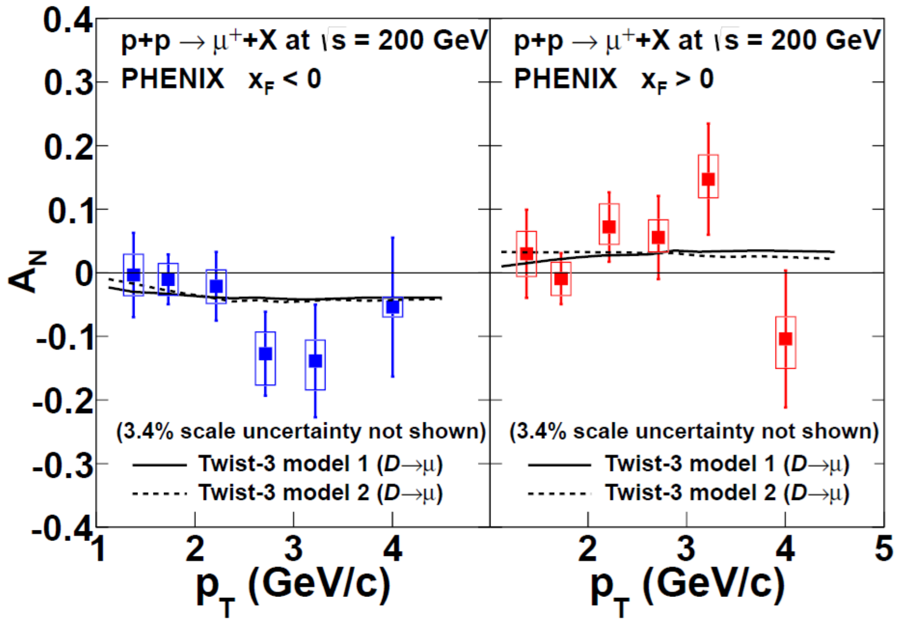




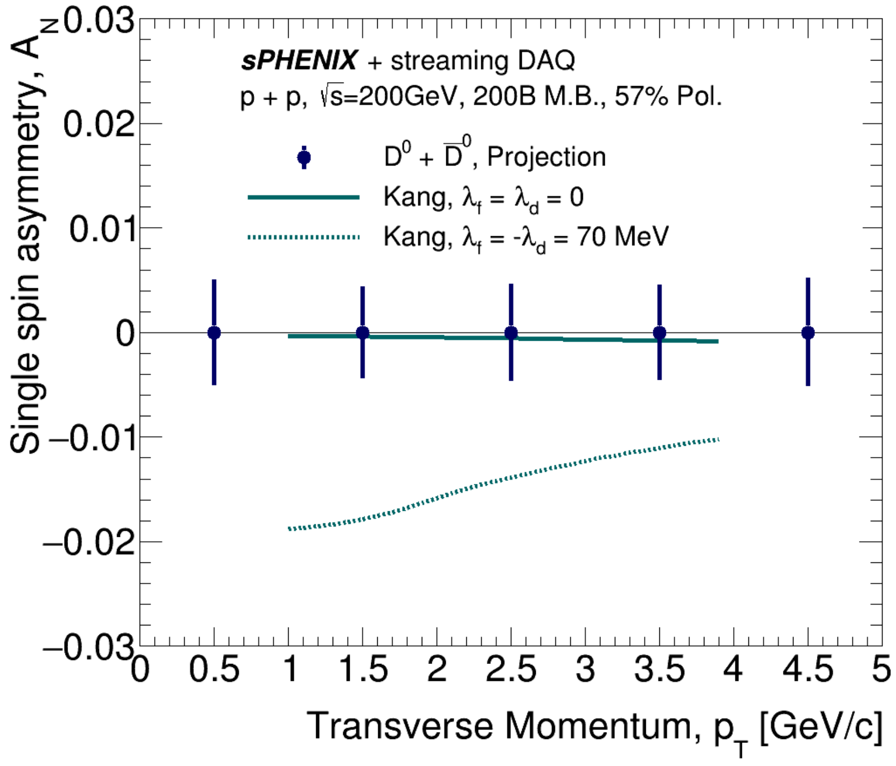
Gluon dynamic in polarized pp col.

- Charm is unique probe of gluon
- D^0 A_N → Tri-gluon correlation

PHENIX, DOI:10.1103/PhysRevD.95.112001

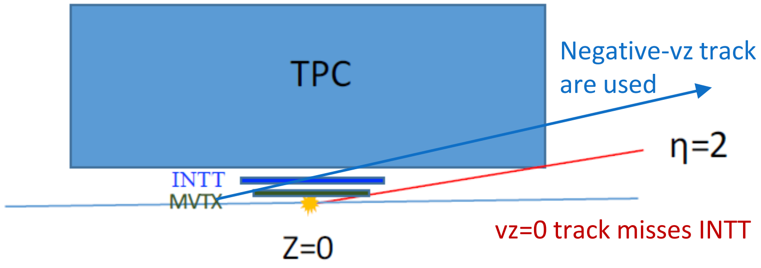


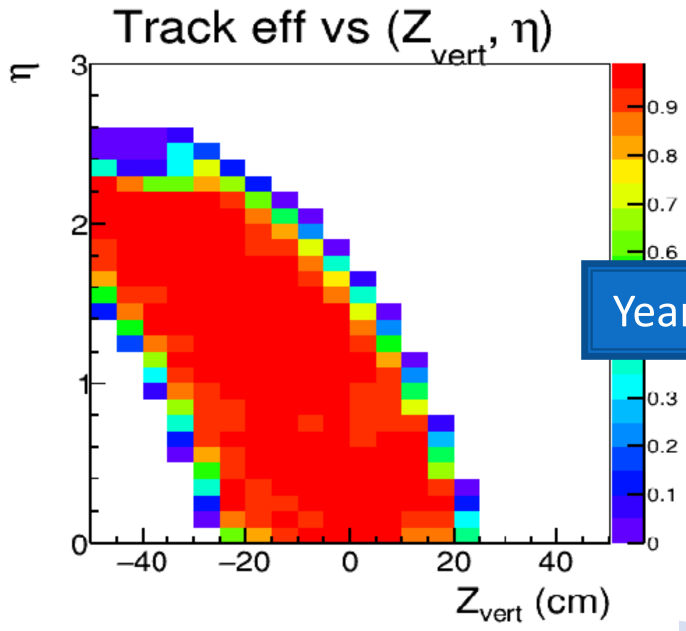
See also: talk Cold QCD TG



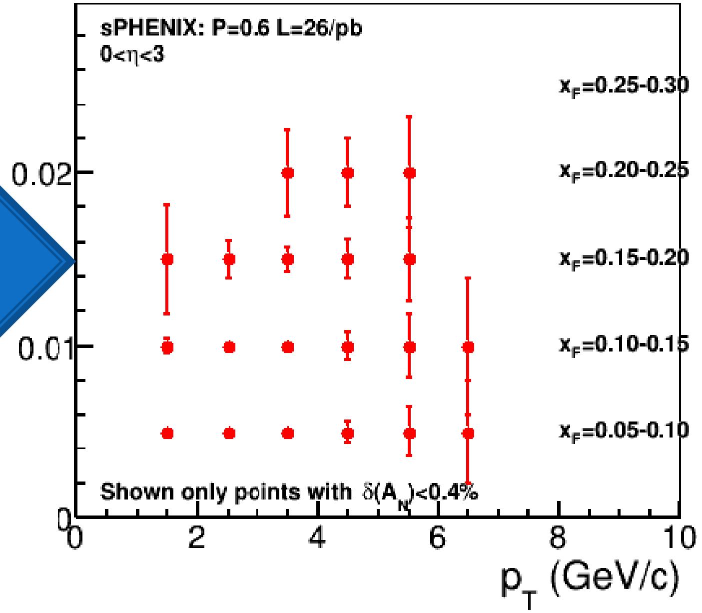
Bridging data gap for the Nuclear A_N mystery

- From Sasha/Christine's talk at Cold QCD TG meeting
- sPHENIX can enormously improve charged hadron A_N
- Note it benefit from a wider vertex distribution, but may tolerate a narrower vertex via beam crossing





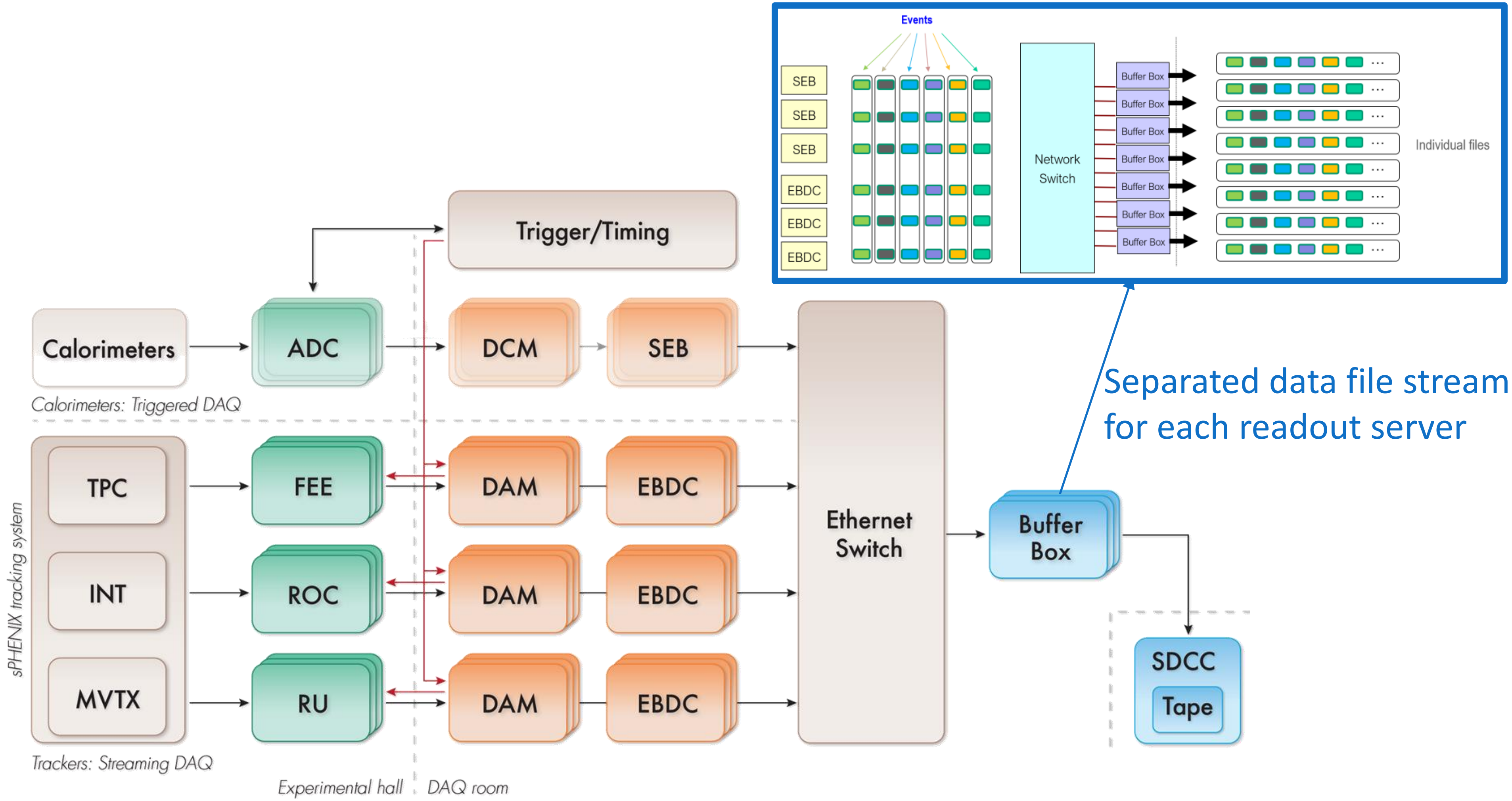
Year2: 10% SRO



Part II – One way it can be done (in my imagination)

- All sPHENIX tracker ASICs supports zero-suppressed streaming readout
- Question here how we introduce minimal change to the triggered DAQ so the DAQ can also partially stream tracker data to tape
- Turns out there is a trick to do so with relatively low cost on total data volume

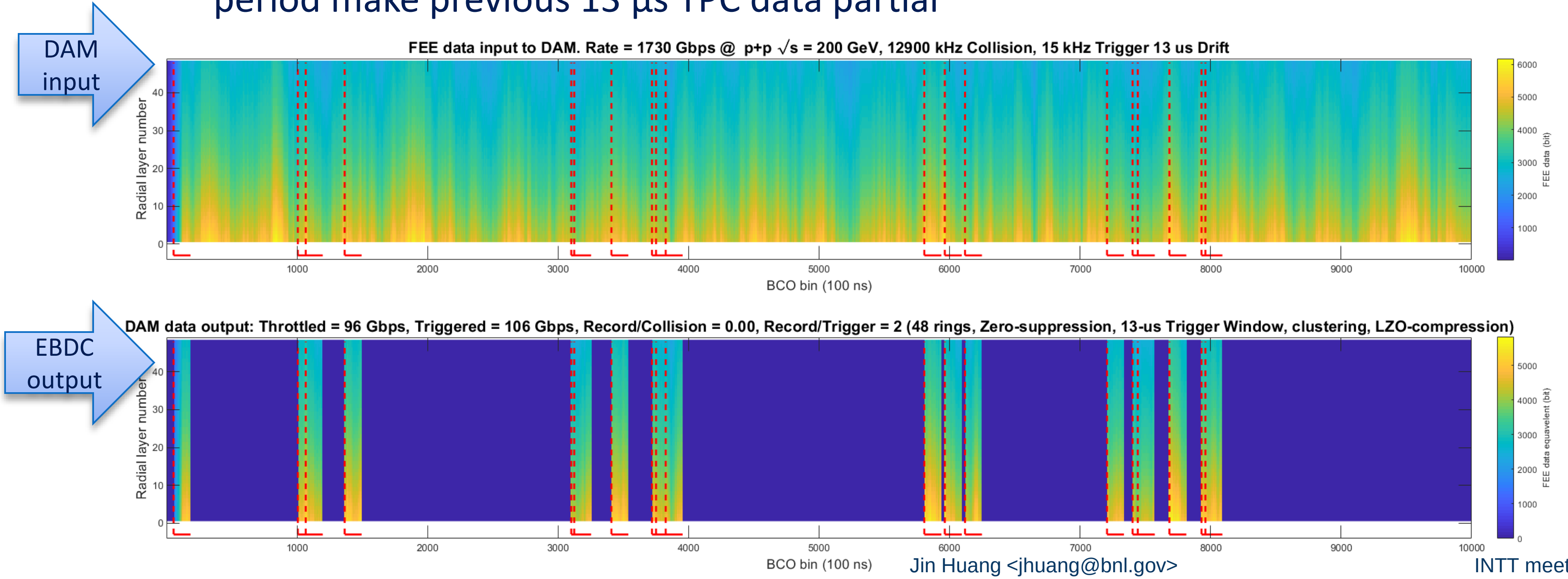
Readout hardware in current plan



See Collaboration meeting DAQ talk by M. Purschke

TPC is loaded with partial pileup events

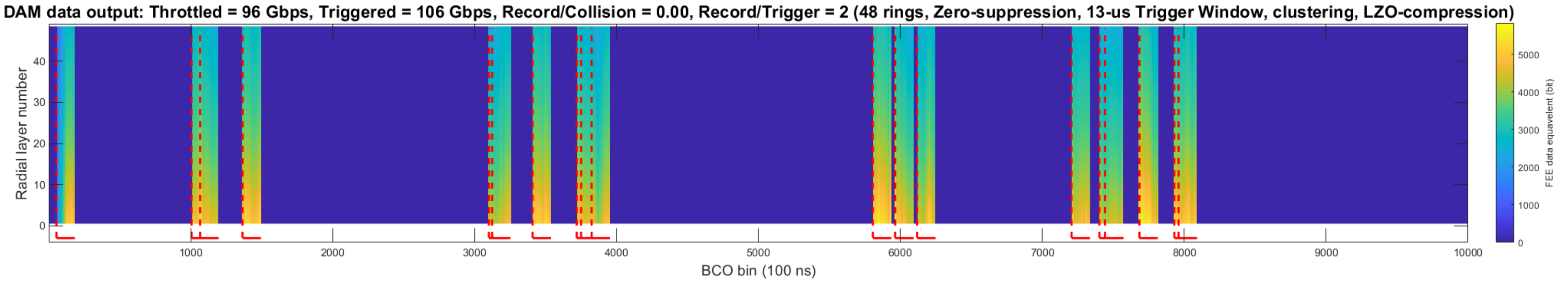
- The majority data are dominated by TPC (longer integration window, may layers, many samples per hit)
- Interestingly, we are already writing out many MB events (partially) to disk.
  - TPC records 13 μs (T_{Trift}) of data for each pp trigger.
  - That include 170 extra M.B. collisions written with the triggered event!
  - *But* they only contain partial event in TPC. i.e. each gap between recording period make previous 13 μs TPC data partial



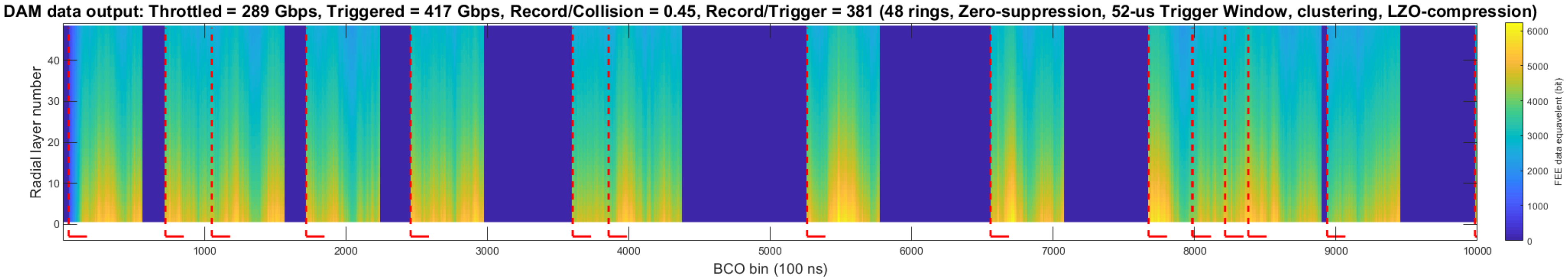
Then, obviously….

- If we allow each TPC recording window at sPHENIX trigger, we allow it to record longer than TPC drift window, we immediate recover large amount of MB data.
- MVTX can extend data with sequential trigger too.
- Tricky part will be INTT, which by reusing FVTX hardware only take 1-cross (0.1us) of data per trigger

TPC record window = drift window (13 μs), almost not recording much M.B. events



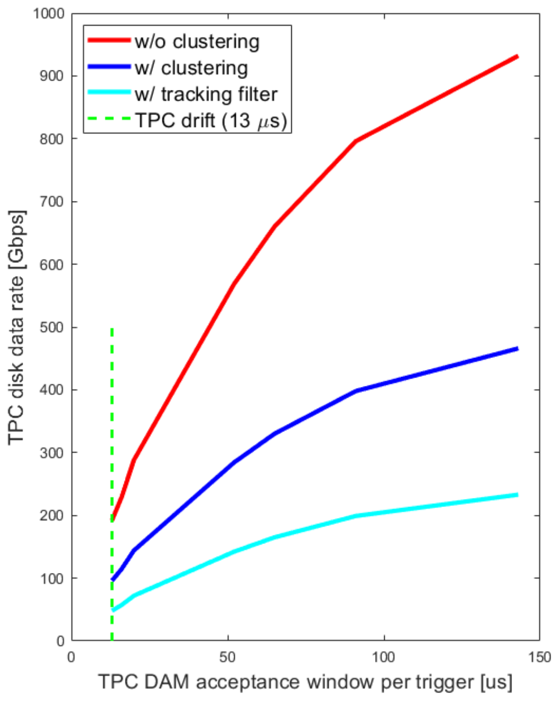
TPC record window = drift window (13 μs) **+ 39 μs**, Recording 45% of M.B. events

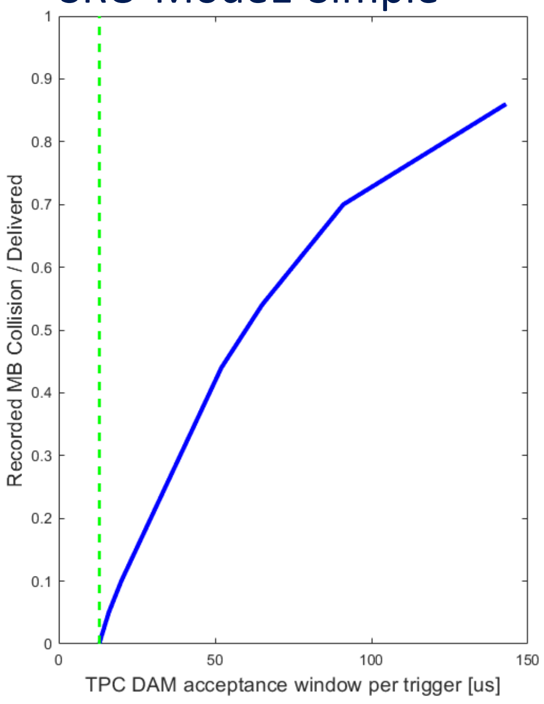


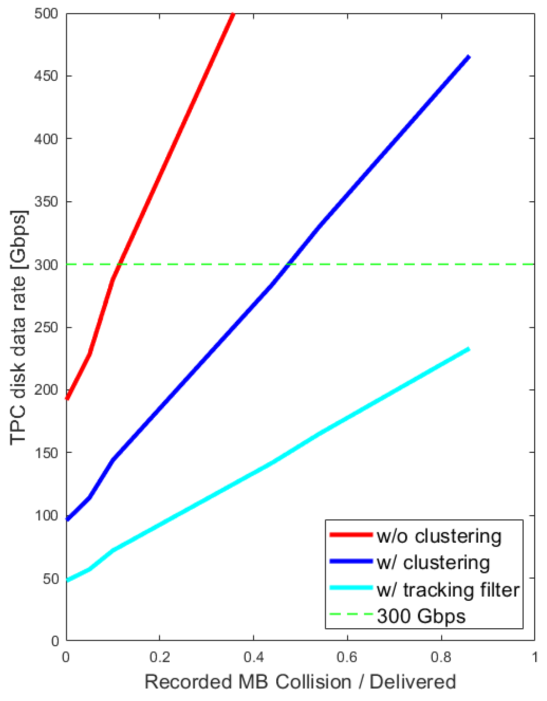
Then what about increased data rate?

- Percentage of M.B. data recorded <-> TPC record window <-> Disk rate
- Also is strongly depending on whether we can do more data reduction than what was assumed on baseline design
- Data rates are TPC EBDC throughput. RCF data transfer rate will further reduced from (60% RHIC uptime x 60-80% sPHENIX uptime)

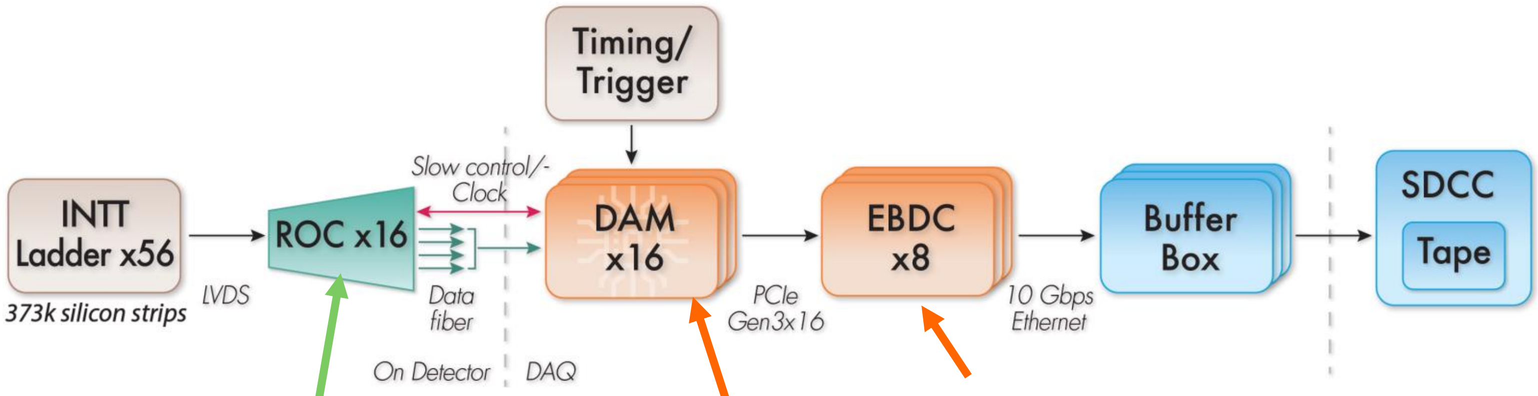
SRO-Mode1-Simple



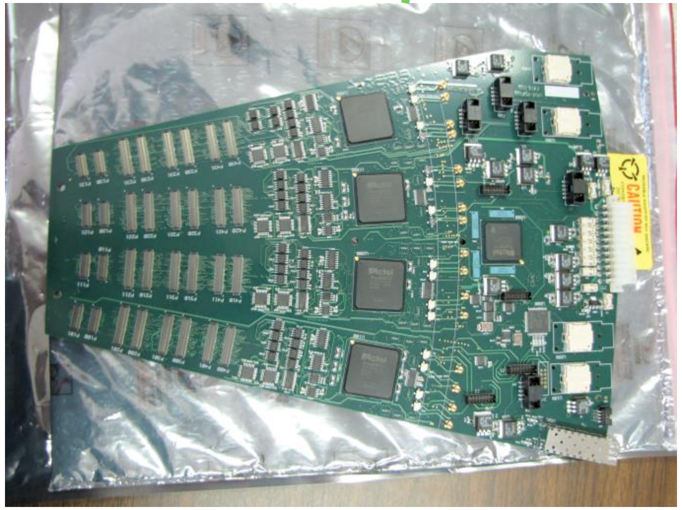


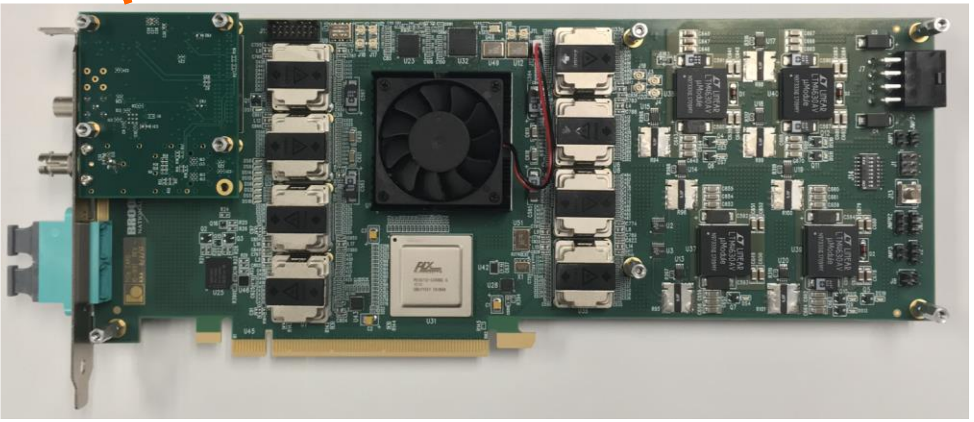


FELIX/DAM adoption for INTT



Relatively low-load commodity server
@ ~0.1 Gbps << TPC

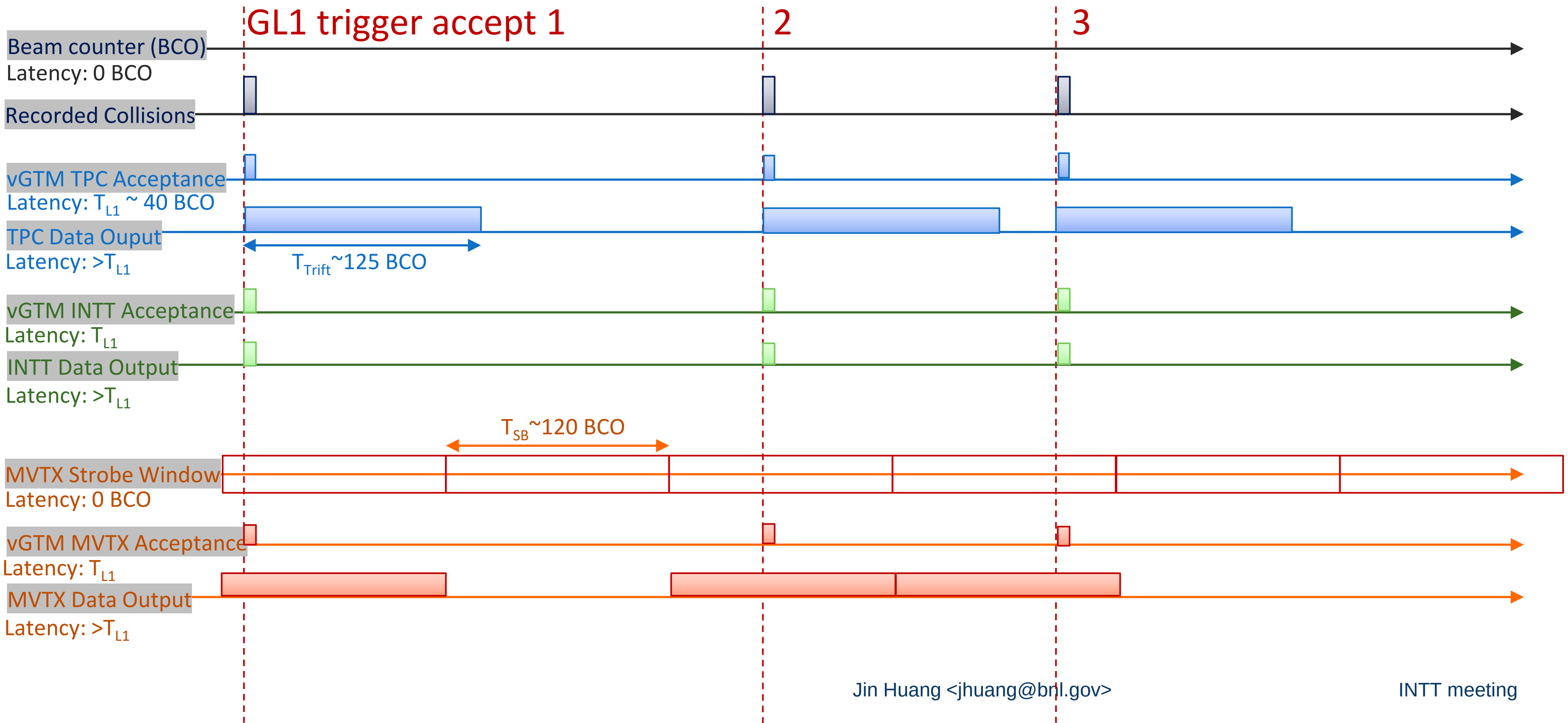




		sPHENIX in Current day-1 setup	sPHENIX w/ Streaming tracker
TPC	DAQ hardware	FEE → DAM → EBDC	Not Changed
	Firmware & Software	Record 13 μs data following a trigger (one TPC drift window), which provide one beam crossing (0.1 μs) of complete collision data	Record 20 μs data following a trigger, providing 7 μs of complete collision data
	Peak data rate	192 Gbps	288 Gbps
INTT	DAQ hardware	ROC → FEM → DCM2 → JSEB2 → Server	ROC → DAM → EBDC New construction of DAM and EBDC following TPC production
	Firmware & Software	Triggered readout of 1 beam crossing (0.1 μs) per trigger	Streaming readout of 7 μs of data following a trigger
	Peak data rate	0.01 Gbps	0.8 Gbps
MVTX	DAQ hardware	FEE → DAM → EBDC	Not Changed
	Firmware & Software	Record one strobe time window of data following a trigger (5-10 μs)	Continue recording strobe time windows until accumulating at least 7 μs of complete collision data
	Peak data rate	3 Gbps	6 Gbps

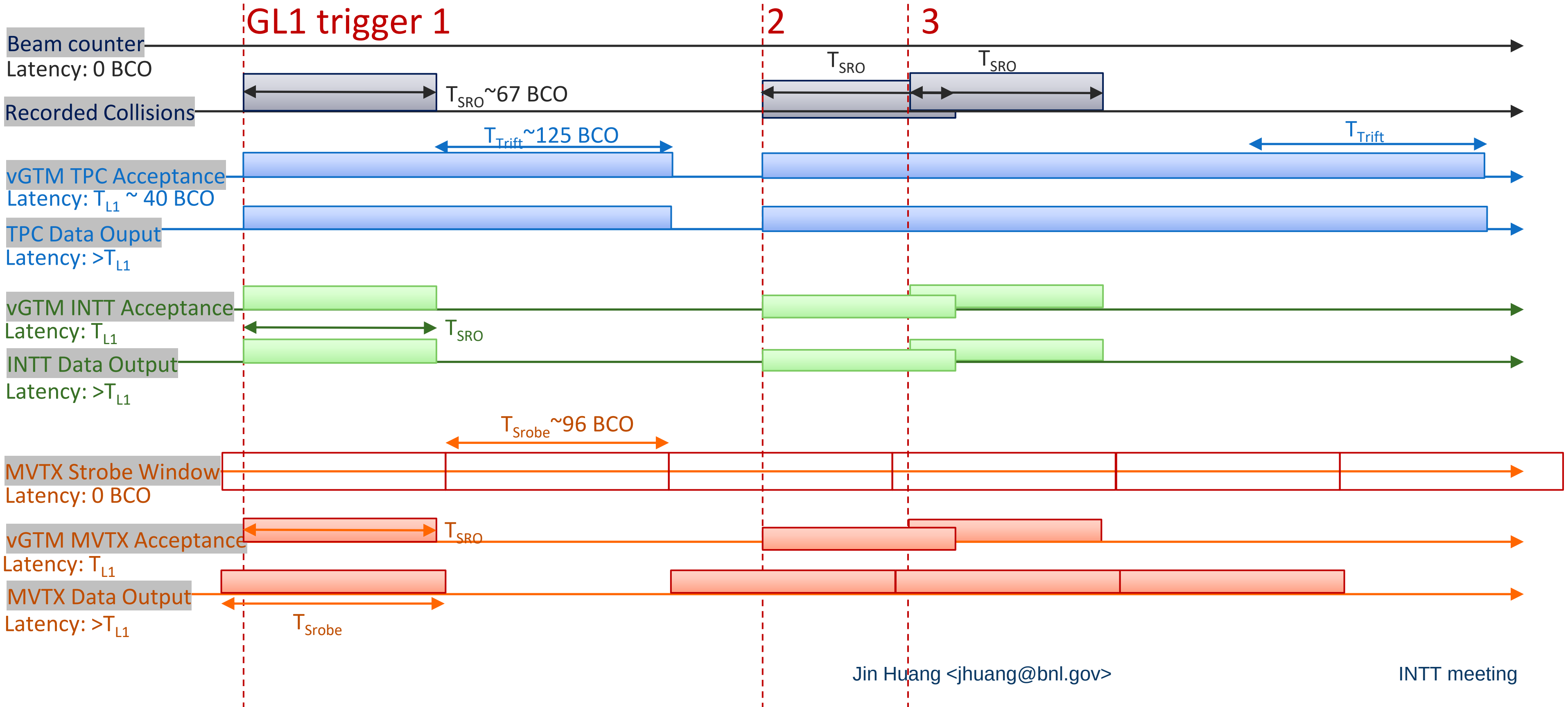
Triggered-Mode:
Current synchronization in beam clock

- vGTM distribute GL1-trigger-accept to each subsystem
- Each subsystem output data gated by trigger + integration time windows
  - TPC=13 us drift, INTT = 1 BCO, MVTX=1 strobe window



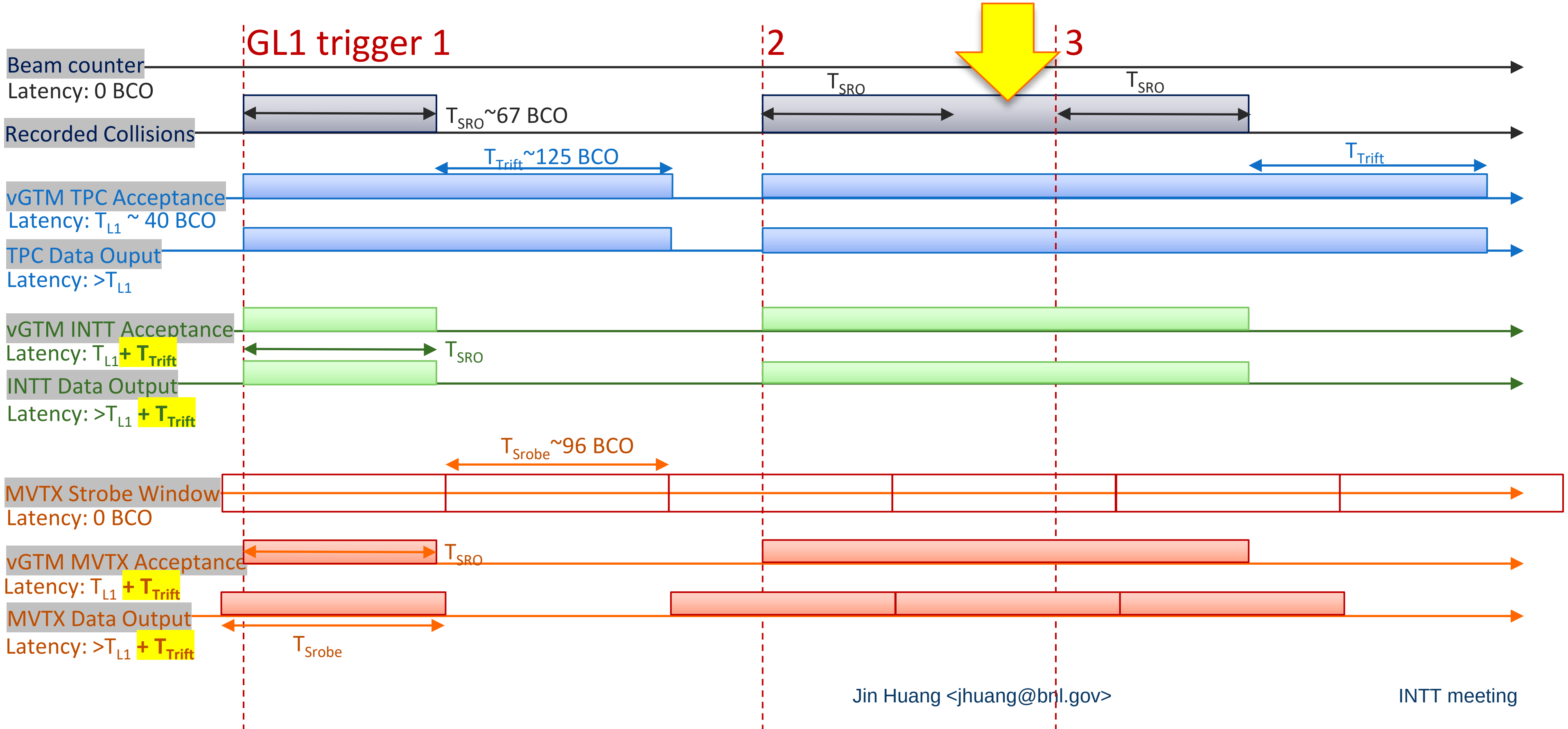
SRO-Mode1-Simple [Recommended]

- vGTM distribute Streaming-Acceptance cross tracking detectors in sync with calo triggers
- Comparing the triggered-mode: simply prolong L1-Acceptance signal to each subsystem, from 1 BCO to T_{SRO}~67 BCOs



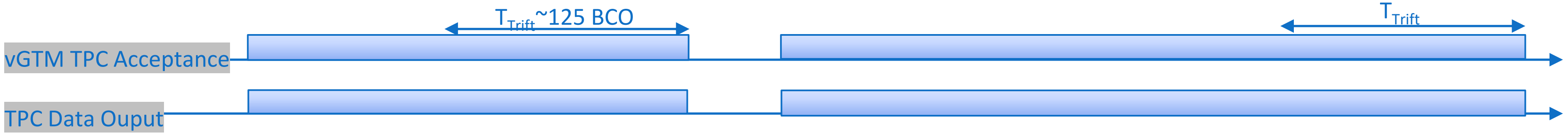
SRO-Mode2-Efficient [Not recommended]

- One more trick to improve efficiency:
  both INTT and MVTX FELIX need to hold data up to additional 13us [T_{Trift}], as GTM to determines whether to produce continued acceptance in case of merged TPC windows
- In the merged window, TPC data is already collected in all three cases. Extended Acceptance-Window allows matching data in MVTX and INTT
- Benefit: roughly doubling SRO recorded events without increasing the TPC (dominant) data rate. Risk: Larger chance for screw-ups and buffer overflows



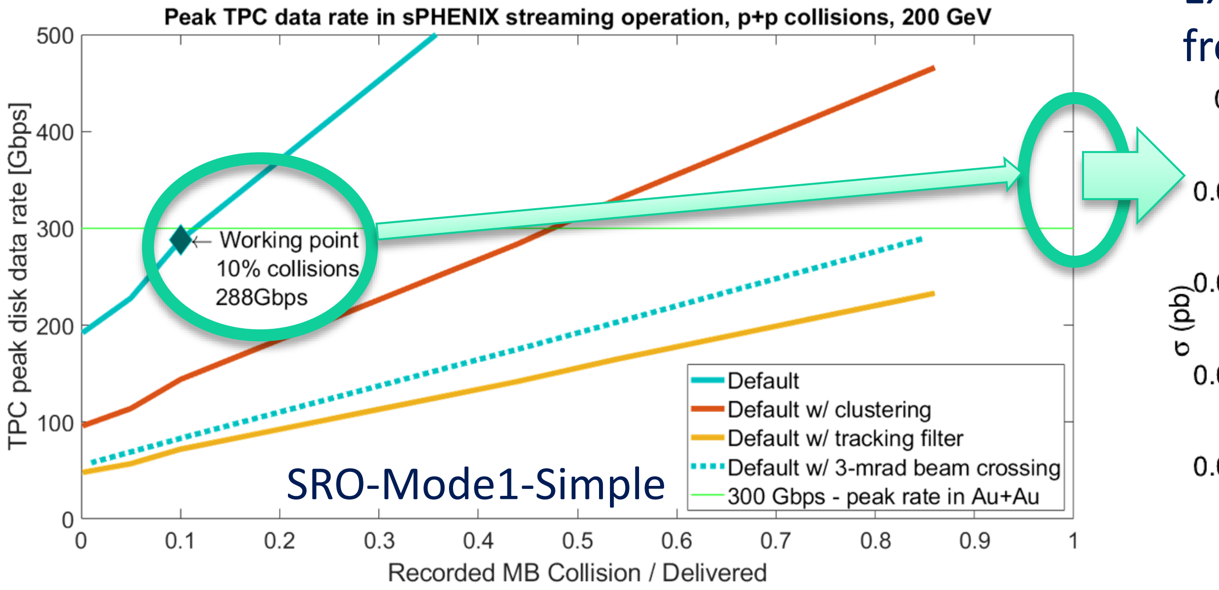
Offline considerations

- Vast majority of pp data in TPC (~99%) and MVTX are from pipe-up collisions
- It would be efficient to reshape analysis to do TPC-calibration, and tracking in continued-time blocks
  - Construct TPC tracklets in [X-Y-Time] space
  - For tracklets in TPC, select one beam crossing determine T0, Translate tracklets into [X-Y-Z] space
  - Continue tracking into INTT. If no matching, ignore this tracklet for this beam crossing (saved to be used in later crossings)
  - For in-time track, continuing tracking into MVTX.
  - Fit in-time track in [X-Y-Z] space

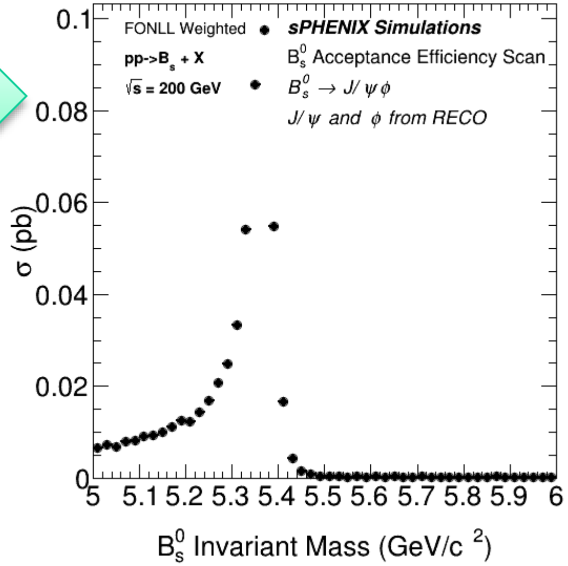


Aggressive streaming in year 4-5?

- Dramatically improved streaming can be a motivation for Run 4 & 5.
- At top luminosity operation in Year 4/5., p+p/Au+Au full streaming FEE->DAM data rates are comparable [Martin S&C review]
    - p+p: with crossing angle we can allow streaming all tracker at ~350Gbps
    - Au+Au: if crossing angle reduce Au+Au collision to 60kHz, TPC 100% streaming data rate ~ 350 Gbps

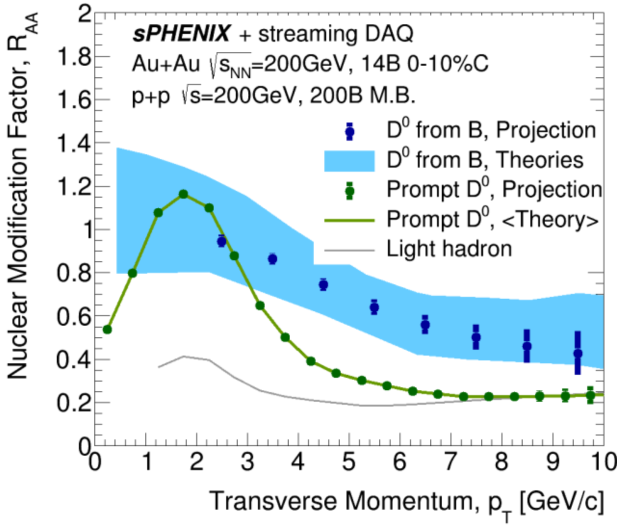


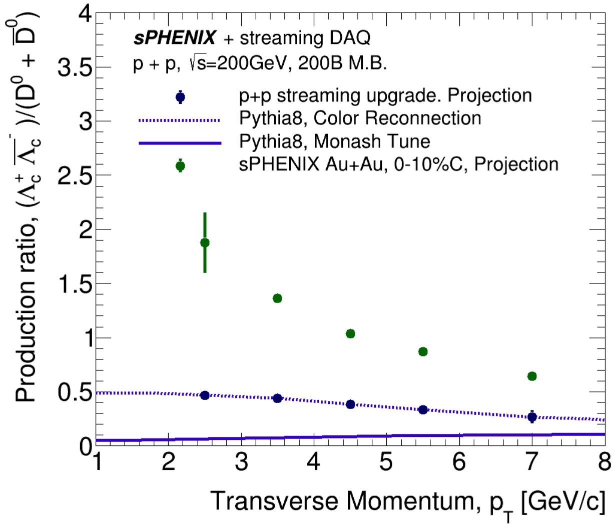
Example channel benefits from larger dataset: B_s [link]

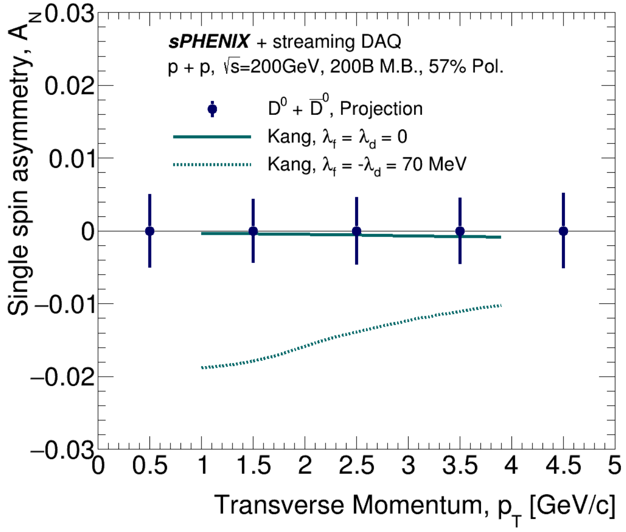


Summary

- A strong physics program within reach with SRO DAQ
  - Covering bottom baseline, charm chemistry, cold QCD
  - Also unexpected physics opportunity with large p+p sample without triggering bias
  - Current projections are scaled from Peripheral Au+Au, need optimization for p+p
- Stream recording tracker data for 10% collisions
  - Modification appears minor for TPC and MVTX
  - INTT updated Readout interface to be same as TPC & MVTX
  - ++ manpower in FPGA, DAQ implementation, and debugging







INTT FEM → FELIX

- Signals between ROC and FELIX/FEM
  - ROC->FELIX data: 4x8 TLK->HFBR-772BEZ MPO fibers
  - FELIX <> ROC slow control: 1 bidirectional SFP
  - FELIX -> ROC timing: 2x fibers with CLK/START signal
- Total 16 ROCs & FELIXs.

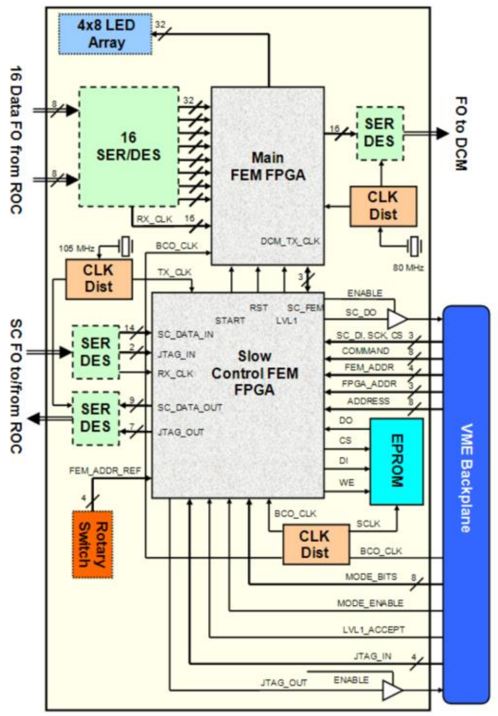
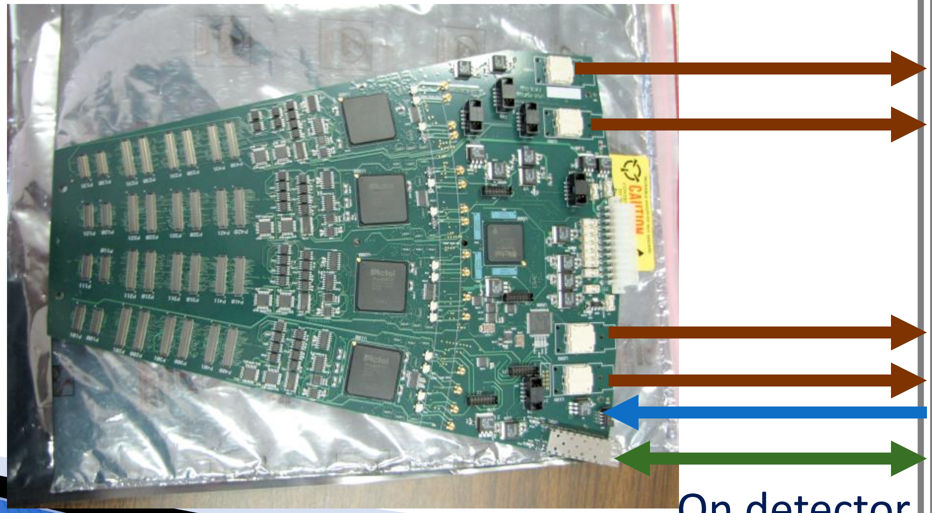
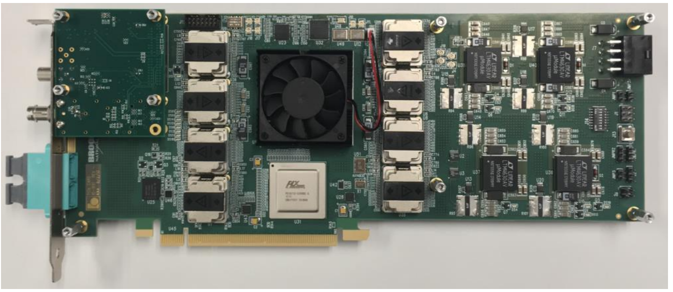


Figure 10: (color online) Block diagram of a FEM board.



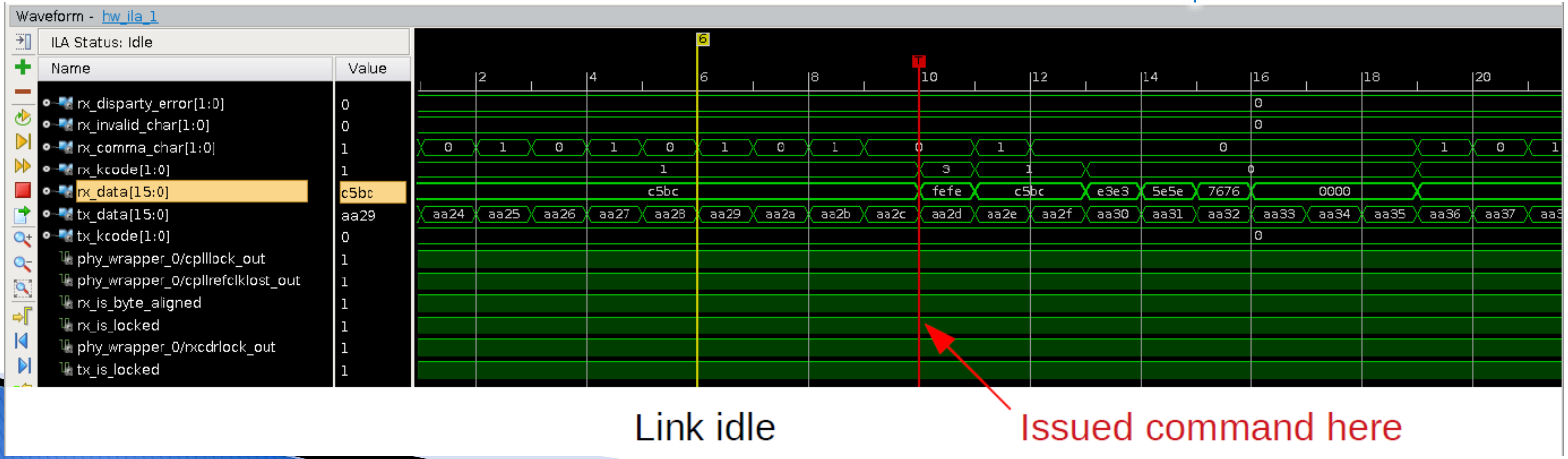


Each FELIX support 48 bi-directional SFP+ link (1 used for timing). Can work with lower speed FVTX SFP too

On detector | Counting house

ROC-FELIX link demo

- After last collaboration meeting
- Demonstrated FVTX/INTT ROC → Kintex-7 SFP link is plausible
- Nonetheless, lots work needed



Can we do it for free? – MB trigger

- D_0, D_{0_bar}->pi/K with pi/K in |eta|<1 is ~ O(10) μb [Pythia8 HardQCD], Population in MB is $2x10^{-4}$
- In baseline trigger bandwidth, we may ask for O(1kHz) MB events
  - Depending on how well vertex trigger may work in high pile up environments
  - Record 0.01% - 0.05% of 8.3 Trillion M.B. collisions within |z|<10 cm
- Recorded D_0->pi/K with pi/K in |eta|<1,
  = 8 Trillion x (0.01-0.05%) x $2x10^{-4}$ = O(1M)
  - Probably can do a neat job for charm D_0 baseline
  - May be enough for Lamda_c (~10% of D, O(100k) events)
  - Difficult for bottom baseline (<1% D_0->pi/K , <8k integrated pT), D-D correlation.

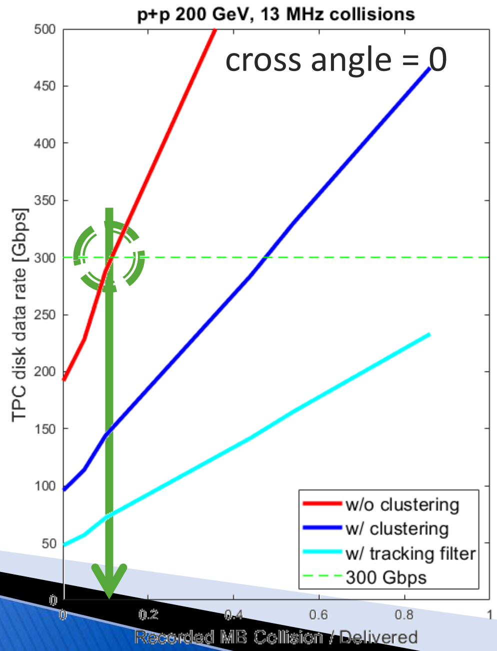
sPH-TRG-2018-001: ~20% pp collision in |z|<10cm

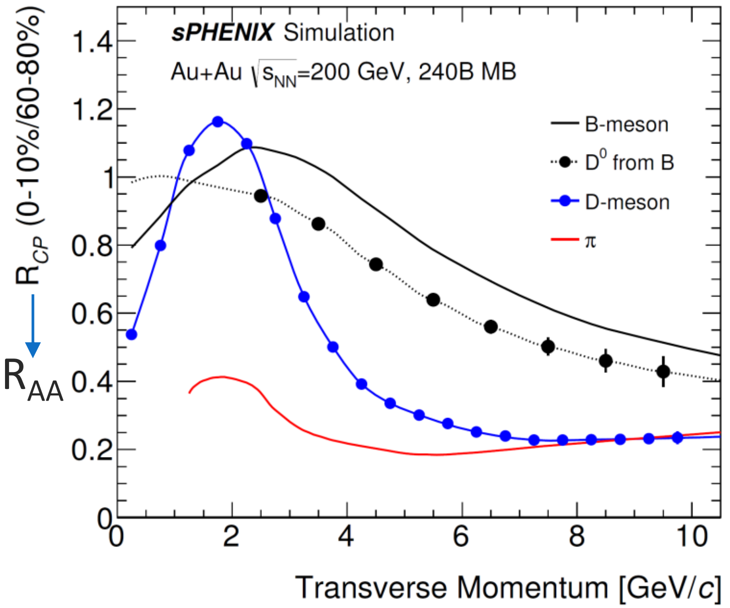
Table 2: Summary of integrated samples summed for the entire five-year scenario.

Species	Energy [GeV]	Rec. Lum.	Samp. Lum.	Samp. Lum. All-Z
Au+Au	200	35 nb^{-1} (239 billion)	80 nb^{-1} (550 billion)	214 nb^{-1} (1.5 trillion)
p+p	200	—	197 pb^{-1} (8.3 trillion)	1.0 fb^{-1} (44 trillion)
p+Au	200	—	0.33 pb^{-1} (0.6 trillion)	1.46 pb^{-1} (2.6 trillion)

HF meson statistics via this upgrade

- Conservatively, if we record **10% of M.B. collision** by increasing TPC recording window to 20 μs…
- 50% increase in overall TPC data rate in pp collision
- x O(100-1000) improve for MB sample,
  Formidable 0.8 Trillion/20pb^{-1} recorded in |z|<10 cm vertex:
  - Recorded D0->pi/K with pi/K in |eta|<1 = 8 Trillion x (10%) x 2x10^{-4} = 160 M
  - Finally ping down Lamda_c (~10% x D0->piK, 16M events)
  - Nice bottom baseline (1M B->D->PiK, integrated p_T), D-D correlation, Exclusive B+
  - Legacy of world last polarized p+p collisions data. Data mining in the future.
- To recover p+p baseline for low-p_T HF hadron program
  - Matching 10% Ncoll in 0-10% AuAu. Nonetheless, much lower combinatorial bgd (pending quantification)





Quantifying in sim.

- p+p projections
  - D_0, B+
  - Lc simulation,
    - Lc reach below 2GeV
- p+A strategy