Phase 3 processing scheme

full dress rehearsal and phase III

Stefano Lacaprara

stefano.lacaprara@pd.infn.it

INFN Padova

Data Production meeting,
KEK, 7 February 2019

Outline

- Phase II reprocessing
- GCR5 processing
- toward phase III processing

Phase2: Proc7

pre-proc7 done (14/12/2018)

- prerelease-03-00-00b, GT data_reprocessing_validation_release-03-00-00
  - improvement in tracking code, no new payload for CDC
  - produced cdst from HLT skims (prod6) hlt_bhabha, hlt_gamma_gamma, hlt_mumu_2trk
  - First exercise for new Data production (SL) and calibration manager (Umberto)
- Issues:
  - Most of raw data on tape on hsm, need to prestage them (tools available hstage)
  - setup of script, GT, etc via Jira ticket and PullRequest (good)
  - Processing was fast (once data on disk) about 1 day: fully at KEKCC (as expected)

proc7 done (8/1/2019) Experiment3-proc7 confluence page

- release 03-00-00 (2/1/2019 as scheduled)
- GT: data_reprocessing_proc7 using pre-proc7 and release (4/1)
  - proc7 started 5/1 ended 8/1 at KEKCC
  - input RAW 613 runs, output: mDST, cDST and DST
    - ★ only customization: Set isMC: 0 in metadata
  - also started production proc7b on grid (next slide)
  - feedback very fast. Tracking degradation was found BII-4359
    - ★ intense debugging and immediately start preparation for (pre-)proc8 with fixed code/payload

Grid processing proc7b

- week after proc7, we started a reprocessing of exp3, phase2 on the grid proc7b identical to proc7
- need some time to setup processing correctly, need to learn `gb2_prod_tools` (help from Ueda-san)
- two test pre-processing (3 runs each) submitted on 16/1
- two batch of jobs because some input file are labelled “Beam” and some “Physics”

6798 Done: all looked fine, so I started the full processing

6799 Running: Transaction are Done= 6/6 but

- ▸ (24/1): TransId:21689 Registered Done:2 Total:2 100.0% 1/1 ??

- Full production (started 17/1):

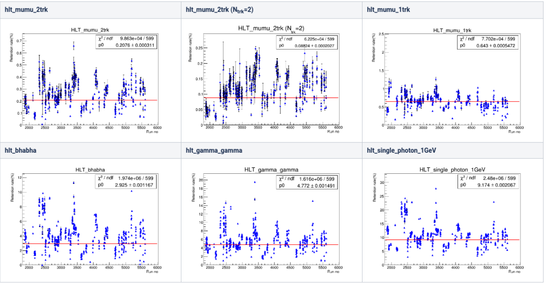
6857 Running: submitted on 2019-01-17 15:28:18, Done=400/402 (was 182/402 on 24/1)

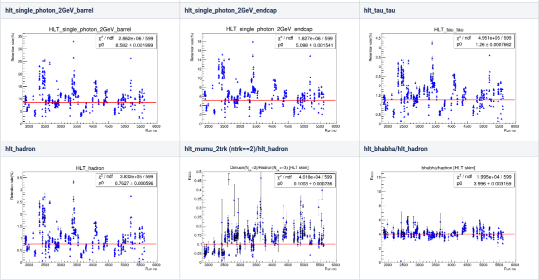
6858 Running: submitted on 2019-01-17 15:42:48 Done=872/874 (was 187/874 on 24/1)

- NB: `proc7` took about 2.5 days at KEKCC
- Investigating with computing experts: apparently the last remaining jobs are done, but are in a strange state, and they are reported as not done

HLT skims (and offline)

- preproc7 was done starting from prod6 HLT-skims (raw)
- for proc7 we produce new HLT skims, based on code by Karim
- input **mdst/cdst** (no dst)
- output: skim **mumu_2trk mumu_1trk hadron bhabha gamma_gamma single_photon_1GeV single_photon_2GeV_barrel single_photon_2GeV_endcap tau_tau** exp3 skims
- later Karim produced offline skims **offskim_mumu, offskim_dstar**
  - do we still need all of them? Do we need more?
- retention rate stable wrt prod6 (Karim)

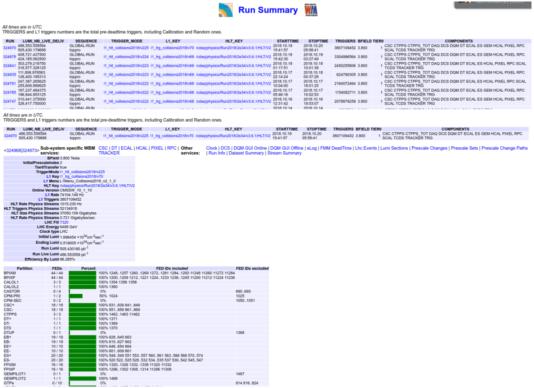




Luminosity

- Provided by Xing-Yu, based on ECL information with the Bhabha and Di-γ events [BELLE2-NOTE-PH-2018-027]
- Proc7: $(503.5 \pm 0.2)\,\mathrm{pb}^{-1}$ and $(499.0 \pm 0.6)\,\mathrm{pb}^{-1}$, respectively
  - was 479.8 pb^{-1} and 494.9 pb^{-1} for prod6 (correction suggested by N.Gabyshev to account for phiCMS efficiency dependence for Bhabha gives: 491.5 pb^{-1})
- also run-by-run lumi available link
- We need an offline run summary page to collect these kind of information
  - see Run+Summary confl page
  - now a txt file (extracted from DAQ run DB) link
  - Good starting point, but we need something more code and user friendly: Run+Summary
  - possibly with access via software and maybe also web interface?

An example: CMS run summary page



Run Summary

All times are in UTC.
TRIGGERS and L1 triggers numbers are the total pre-deadtime triggers, including Calibration and Random ones.

RUN	LUMI_NB_LIVE_DELIV	SEQUENCE	TRIGGER_MODE	L1_KEY	HLT_KEY	STARTTIME	STOPTIME	TRIGGERS	BFIELD	TIER0	COMPONENTS
324970	486,553.558594 505,430.179688	GLOBAL-RUN toppro	l1_hlt_collisions2018/v225	l1_trg_collisions2018/v70	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.19 15:41:57	2018.10.20 05:58:41	3807109452	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324878	408,721.437500 424,185.062500	GLOBAL-RUN toppro	l1_hlt_collisions2018/v224	l1_trg_collisions2018/v69	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.18 15:42:30	2018.10.19 03:27:49	3304996564	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324841	303,279.218750 316,577.328125	GLOBAL-RUN toppro	l1_hlt_collisions2018/v223	l1_trg_collisions2018/v69	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.18 01:17:51	2018.10.18 10:01:39	2450255806	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324835	111,806.976563 126,400.195313	GLOBAL-RUN toppro	l1_hlt_collisions2018/v223	l1_trg_collisions2018/v69	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.17 22:14:24	2018.10.18 00:37:28	624790305	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324791	247,387.265625 255,609.890625	GLOBAL-RUN toppro	l1_hlt_collisions2018/v222	l1_trg_collisions2018/v68	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.17 10:04:00	2018.10.17 18:01:25	2164072464	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324785	187,237.484375 196,644.953125	GLOBAL-RUN toppro	l1_hlt_collisions2018/v222	l1_trg_collisions2018/v68	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.17 05:46:16	2018.10.17 10:02:53	1154082711	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG
324747	310,441.375000 326,417.750000	GLOBAL-RUN toppro	l1_hlt_collisions2018/v222	l1_trg_collisions2018/v68	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.16 12:31:42	2018.10.16 19:53:07	2078879259	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG

All times are in UTC.
TRIGGERS and L1 triggers numbers are the total pre-deadtime triggers, including Calibration and Random ones.

RUN	LUMI_NB_LIVE_DELIV	SEQUENCE	TRIGGER_MODE	L1_KEY	HLT_KEY	STARTTIME	STOPTIME	TRIGGERS	BFIELD	TIER0	COMPONENTS
324970	486,553.558594 505,430.179688	GLOBAL-RUN toppro	l1_hlt_collisions2018/v225	l1_trg_collisions2018/v70	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2	2018.10.19 15:41:57	2018.10.20 05:58:41	3807109452	3.800	1	CSC CTPPS CTPPS_TOT DAQ DCS DQM DT ECAL ES GEM HCAL PIXEL RPC SCAL TCDS TRACKER TRG

<324968|324973> **Sub-system specific WBM services:** CSC | DT | ECAL | HCAL | PIXEL | RPC | TRACKER **Other services:** Clock | DCS | DQM GUI Online | DQM GUI Offline | eLog | FMM DeadTime | Lhc Events | Lumi Sections | Prescale Changes | Prescale Sets | Prescale Change Paths | Run Info | Dataset Summary | Stream Summary

BField	3.800 Tesla
InitialPrescaleIndex	2
Tier0Transfer	true
TriggerMode	l1_hlt_collisions2018/v225
L1 Key	l1_trg_collisions2018/v70
L1 Menu	L1Menu_Collisions2018_v2_1_0
HLT Key	/cdaq/physics/Run2018/2e34/v3.6.1/HLT/V2
Online Version	CMSSW_10_1_10
L1 Rate	74104.148 Hz
L1 Triggers	3807109452
HLT Rate Physics Streams	1015.230 Hz
HLT Triggers Physics Streams	52134918
HLT Size Physics Streams	37050.109 Gigabytes
HLT Rate Physics Streams	0.721 Gigabytes/sec
LHC Fill	7320
LHC Energy	6499 GeV
Clock type	LHC
Initial Lumi	1.696454 $\times 10^{34}\mathrm{cm}^{-2}\mathrm{sec}^{-1}$
Ending Lumi	0.519005 $\times 10^{34}\mathrm{cm}^{-2}\mathrm{sec}^{-1}$
Run Lumi	505.430180 pb^{-1}
Run Live Lumi	486.553559 pb^{-1}
Efficiency By Lumi	96.265%

Partition	FEDs	Percent	FED IDs *included*	FED IDs *excluded*
BPIXM	44 / 44	100%	1248..1257 1260..1269 1272..1281 1284..1293 11248 11260 11272 11284	
BPIXP	44 / 44	100%	1200..1209 1212..1221 1224..1233 1236..1245 11200 11212 11224 11236	
CALOL1	3 / 3	100%	1354 1356 1358	
CALOL2	1 / 1	100%	1360	
CASTOR	0 / 4	0%		690..693
CPM-PRI	1 / 2	50%	1024	1025
CPM-SEC	0 / 2	0%		1050..1051
CSC+	18 / 18	100%	831..839 841..849	
CSC-	18 / 18	100%	851..859 861..869	
CTPPS	3 / 3	100%	1462..1463 11462	
DT+	1 / 1	100%	1371	
DT-	1 / 1	100%	1369	
DT0	1 / 1	100%	1370	
DTUP	0 / 1	0%		1368
EB+	19 / 19	100%	628..645 663	
EB-	19 / 19	100%	610..627 662	
EE+	10 / 10	100%	646..654 664	
EE-	10 / 10	100%	601..609 661	
ES+	20 / 20	100%	548..549 551 553..557 560..561 563..566 568 570..574	
ES-	20 / 20	100%	520 522..525 528..532 534..535 537 539..542 545..547	
FPIXM	16 / 16	100%	1320..1326 1332..1338 11320 11332	
FPIXP	16 / 16	100%	1296..1302 1308..1314 11296 11308	
GEMPILOT1	0 / 1	0%		1467
GEMPILOT2	1 / 1	100%	1468	
GTPe	0 / 10	0%		814 816..824

Experience gained (kekcc)

- Still learning the job, great help from Jake, Karim, Umberto, and many other

✓ Very good communication: JIRA, PullRequest, mails, chat, skype

🔧 Need to document in detail all phases of process, and keep up to date the documentation;

✓ Very good interaction with calibration team, clear definition of GT creation process helped a lot.

- some “ad hoc” tuning of standard reconstruction needed in pre-proc, which is fine(-ish)
- in real processing we need to use only standard unpacking/reconstruction/etc from release
  - ✗ what if we need a quick change? Still acceptable to have non-standard (but documented) mod in steering file or need for a fast patch-release?
  - not an issue for phaseII re-processing, but is for exp5 cosmic run, and likely (?) for phaseIII
- issue of tape vs disk storage at KEKCC:
  - hstage tool is working
  - for future (starting from proc8): keep last two dataset on disk (`gpfs` with copy to tape `ghi`)

Experience gained (grid)

✗ processing on grid was done (is being done) with the left hand

- ✓ just setup (need learn about `gb2_prod_tools`), and then fire (and almost forget)
- ▸ need more babysitting and more careful monitoring from my side: need to interact more with computing group, starting now
- ✓ first experience: setup and submission are well structured, and can be automatized easily for phaseIII processing
- ▸ large task can be problematic
- 🔧 coordination with Racha (skim) and Ale (MC) to set ProdID in filename: need to think a good solution

Phase2: toward Proc8

- **pre-proc8 (done)**
  - script to run `pre-proc8`
  - release-03-00-01
  - include some modifications to tracking: Change the CDC ADC threshold from Sasha Glazov
  - EKLM track matching BIIDP-1120 and BKLM additional branches BII-4458
  - GT from Umberto: `data_reprocessing_preproc8`
  - input: HLT skims `["hlt_bhabha", "hlt_gamma_gamma", "hlt_mumu_2trk", "hlt_hadron" (new)]`
  - special sub production with scan on CDC ADC threshold (0-2-4-6-8e^{-7}) Done
- **pre-proc8b (starting)**
  - Same as pre-proc but with release-03-00-02 and updated GT `data_reprocessing_preproc8b` (new payload for SVD)
  - setting up script, will start hopefully today or tomorrow
- **proc8**
  - Most likely will use release-03-00-02
  - aggressive schedule:
    - ✓ [Jan 21] First pre-proc8 with modifications to tracking settings with release-03-00-00 and data_reprocessing_proc7 Done (with delay 26/1)
    - ✓ [Feb 1] Deadline for software modifications to be used for proc8 release-03-00-02 is out
    - 🔧 [Feb 15] Updated calibration constants provided for proc8
    - 🔧 [Feb 18] Reprocessing begins

Experiment 5 Global Cosmic Run

✓ Started on 15/1

✓ good interaction with computing group (Hara-san)

1 Hara-san copy (by hand?) file for offline usage (sroot→root conversion). Only "Cosmic trigger data for detector performance" to be processed

1.1 notify that new runs are available BIIDP-1097

✓ Register the same files (root) on grid as well, for grid processing Done

2.1 DP (me) will process the runs and produce cdst/mdst/(dst)

2.2 once done (about 1 hour) they new runs are moved to final location

2.3 update confluence page https://confluence.desy.de/display/BI/Experiment+5+-+full+dress+rehearsal

- As for exp3, need to define release, GT, input, scripts, output
- Two processing so far (there will be more) **GCR5a** and **GCR5b** (there will be a **GCR5c** in 1-2 weeks)

Exp5 GCR5

First processing GCR5a (stopped)

- **release-03-00-00** and “wrong” GT `data_reprocessing_proc7` (done for phase II, so no SVD-PXD)
- setup script and automation, process new runs as soon as available
- processing stopped as soon as a better GT (and patch release) available
- Eventually will delete these obsolete processing

Second processing GCR5b (running)

✓ using **release-03-00-01** and snapshot of running global tag `data_reprocessing_prompt`

✗ Include better steering file to use PXD and SVD data, non standard cosmic processing (not good!)

  ✗ still no tracking using PXD and SVD (only CDC). Would have required complex mod to steering file

✗ found a (serious) problem in SVD geometry (Giulia). Need a new payload and reprocessing

- two options:
  - A wait for new release (possibly with SVD+PXD in tracking) and new GT (with correct Payload for SVD)
  - B start immediately GCR5c with custom payload (localDB) (custom steering **and** payload? eech!)
  - ▸ Some discussion with tracking and SVD people: we will go for [A], Nils will try to include SVD+PXD in tracking in 1/2 weeks

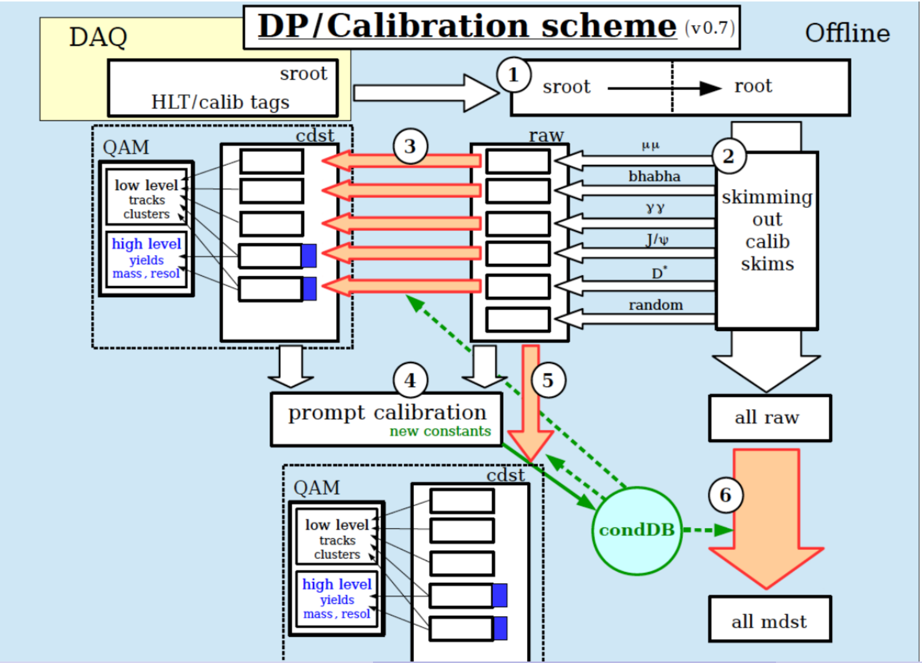
Experience from GCR5

✓ Together with PhaseII reprocessing, a ideal playground to gain experience toward phaseIII

✓ Communication with computing group is good (Mail, Jira ticket, PullRequest)

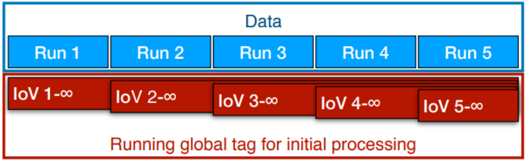
- need to understand how to automatize in a robust way the submission of jobs as soon as new runs are available
  - ▶ so far I'm running on a dedicated LSF queue from a standard KEKCC node
  - ▶ still too much manual work, but scripts are in place and we can automatize most of 2.x
  - 🔧 watchdog for `RunListCosmic` and submit automatically as soon as new runs are listed under test
  - ▶ eventually we will need to do this on the grid (data is already copied)
  - ▶ need to play a bit to gain some direct experience (initial thoughts later)
- as for exp3, calibration group is providing GT in a timely manner
  - ▶ issue: running GT or snapshot?
  - ▶ after this morning GT tutorial, the answer is running GT, when is implemented as described
  - ▶ in the meanwhile, I think snapshot is the way to go for reproducibility

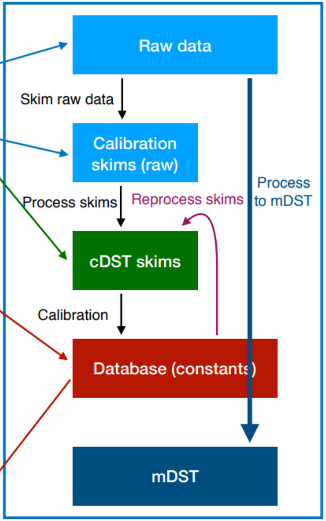
DP/Calibration scheme (v0.7)



Fast reprocessing scheme

- ROOT formatted raw data on offline system, registered, and replicated to raw data processing centers
- Calibration skims from raw data processed to cDST at “calibration center”
- Prompt calibration and QAM run at “calibration center” (includes multiple reprocessing to cDST with updated tracking for dE/dx and TOP calibration)
- Constants added to running offline global tag for initial reprocessing
- Latest runs reprocessed to mDST
- When offline calibrations and software updates are complete, reprocess full data to mDST





What we are doing now

Phase II - reprocessing

- HLT-skim for RAW produced (reuse previous production)
- cdst of selected HLT-skim processed
- validation and calibration from cdst
- production of PayLoad (to condDB)
- bug fix need new patch release
- reprocessing of cdst with better code/payload
- validation and further improvement
- re-processing of full phase II data with state-of-the-art calib and code (until next iteration)
- iterate

GCR Exp5 - prompt processing

- produce cdst/mdst from RAW data (no HLT skim)
- validation and calibration from cdst
- new PayLoad (to condDB) and patch
- include new GT (snapshot from calib manager) and patch into steering script (if possible - no patch release)
- re-process all runs and produce cdst/mdst. Stop previous processing
- (removal of obsolete cdst/mdst not done yet)
- Kind of ok since the data collected is not too much, but we need to move closer to processing schema (or re-discuss it)

Toward Phase III

1. Data collected by DAQ **sroot**

- **sroot** → **root** by computing group (CG aka Hara-san)
- **root** copied and registered to grid (CG)

2.0 DataProduction notified

2.1 DP produces skims (HLT and/or offline?) (RAW) on calibration center (eg KEKCC)

3 DP produces cdst with running (or snapshot) GT and pass to calibration

4 CalibTeam does its magic and produce improved payload

5 (optional) DP produces new cdst with improved payload

- iterate until everybody is happy (or not too sad, at least)

6 DP process RAW into mdst with happy GT

  - initially run at calibration center (eg KEKCC) and produce mdst locally
  - publish them on the grid? How, who?
  - run on grid in parallel (eventually only) and produce mdst directly on grid

7. ...

8. Profit (Announce to collaboration)

Toward Phase III (2)

what do we need for step 2.

- Setup of steering script
  - DP responsibilities: need to be standard reconstruction as defined in release
    - what if we need small/quick change?
    - (it has already happened for CGR5)
    - for cdst processing I think that this is kind of ok (not ideal, but we need some flexibility)
    - otherwise we might need fast patch release for data processing (better but slower and possible latency)
  - GT provided by Calibration manager (prompt_data_processing)
    - see next slide: running GT vs snapshot
- **step 2.1** HLT Skim
  - In the schema, it is the first step and is to be done locally
  - current steering is mostly I/O and fast. Steering script mostly ready from phaseII (Karim), likely to need update (eg monopole skim?)
  - skim also for mdst not in the scheme: we had them for phase2 and widely used.
  - when: before or after mdst processing? For phase II now are done after but we are using the processing-1 ones for cdst . . .
  - where: mdst will be eventually produced on grid, so need to produce skim there as well, but:
    - yet an other step before profit
    - most of raw skim already done at calibration center for cdst: duplication

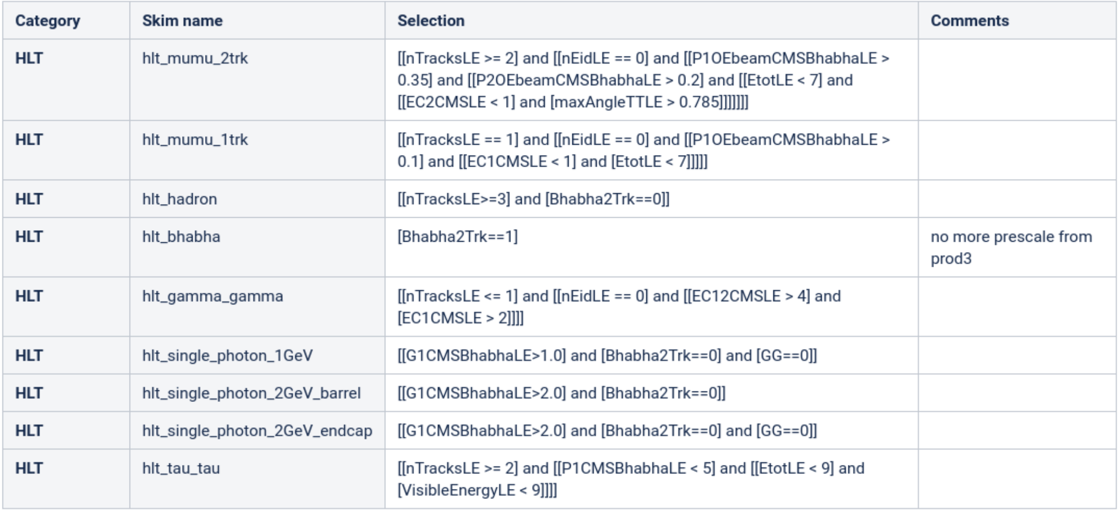
Grid processing

- Not plan A for initial phase III data taking.
- exploit limited luminosity to use local (kekcc) fast processing and re-processing to achieve a reasonably stable operation (unpacker/reconstruction/calibration)
- in parallel run on grid to gain experience for a smooth transition
  - Start already with GCR5
- some issue from my limited experience so far:
  - watchdog for new runs to appears (which is some time after the copy/replication has started)
  - need to develop tool to create `json` with runs to be processed when they are available (should be easy)
  - ProdID for each run or set of runs? Bigger is not better!
  - automation might have some issue (eg to submit I need a voms proxy, which expires in 24h)
  - submission is done via personal grid certificate, which last 24 hours: not possible to setup a fully automatic process (need to renew the certificate every day) certificate renewal service is possible (up to 1 week), still . . .
  - other solution/idea?
  - need to use the available monitoring tool to find problems asap (eventually an important task for Data Production shift)

Additional stuff

Additional or backup slides

Skim definition



Category	Skim name	Selection	Comments
HLT	hlt_mumu_2trk	[[nTracksLE >= 2] and [[nEidLE == 0] and [[P1OEbeamCMSBhabhaLE > 0.35] and [[P2OEbeamCMSBhabhaLE > 0.2] and [[EtotLE < 7] and [[EC2CMSLE < 1] and [maxAngleTTLE > 0.785]]]]]]	
HLT	hlt_mumu_1trk	[[nTracksLE == 1] and [[nEidLE == 0] and [[P1OEbeamCMSBhabhaLE > 0.1] and [[EC1CMSLE < 1] and [EtotLE < 7]]]]]	
HLT	hlt_hadron	[[nTracksLE>=3] and [Bhabha2Trk==0]]	
HLT	hlt_bhabha	[Bhabha2Trk==1]	no more prescale from prod3
HLT	hlt_gamma_gamma	[[nTracksLE <= 1] and [[nEidLE == 0] and [[EC12CMSLE > 4] and [EC1CMSLE > 2]]]]	
HLT	hlt_single_photon_1GeV	[[G1CMSBhabhaLE>1.0] and [Bhabha2Trk==0] and [GG==0]]	
HLT	hlt_single_photon_2GeV_barrel	[[G1CMSBhabhaLE>2.0] and [Bhabha2Trk==0]]	
HLT	hlt_single_photon_2GeV_endcap	[[G1CMSBhabhaLE>2.0] and [Bhabha2Trk==0] and [GG==0]]	
HLT	hlt_tau_tau	[[nTracksLE >= 2] and [[P1CMSBhabhaLE < 5] and [[EtotLE < 9] and [VisibleEnergyLE < 9]]]]	

Running global tag

- running GT is, by construction, open (NEW)
- namely can change after being used to process a given run range.
- update can be forward update, namely valid from a given run to infinity (by policy, by design, or by gentle-person agreement?)
- scenario 1.
  - we produce cdst from run X to run X+10
  - calibration team analyze them, and come up with update payload, IoV $X - \infty$
  - we wait news from calibration team for a possible updated payload before producing mdst for physics
- scenario 1.2
  - calibration a week later: no wait, we do have even better payload for IoV $X - \infty$
  - calib upload payload to running GT with IoV $X - \infty$
  - we have mdst produced with not-up-to-date for $X - X + n$ runs.
  - we don't care (but it would be hard for analysis to understand what happened)
  - we do care, and reprocess run $X - X + n$ runs (and DP goes crazy pretty fast)
- processing with snapshot of runnig GT would guaranteed to know precisely what have been used for processing that run
- would need to be updated regularly by calib coord (which might go crazy ...)
- not clear to me.

Final slide

I need a deputy
(we all neeed)

Bibliography I