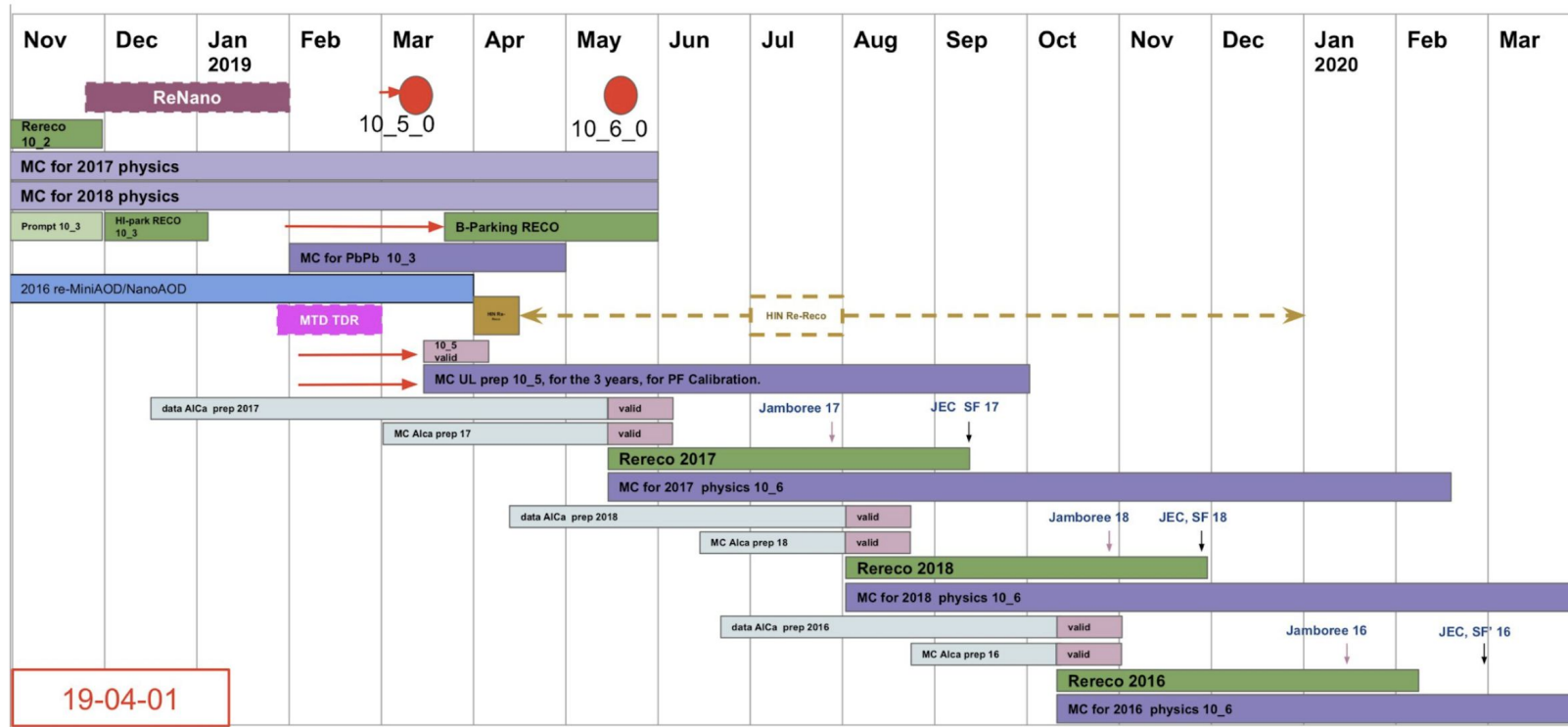


Legacy production and development plans PPD workshop, 16th March 2020

Chayanit Asawatangtrakuldee, Wajid Ali Khan,
Paolo Gunnellini
for the PdMV group

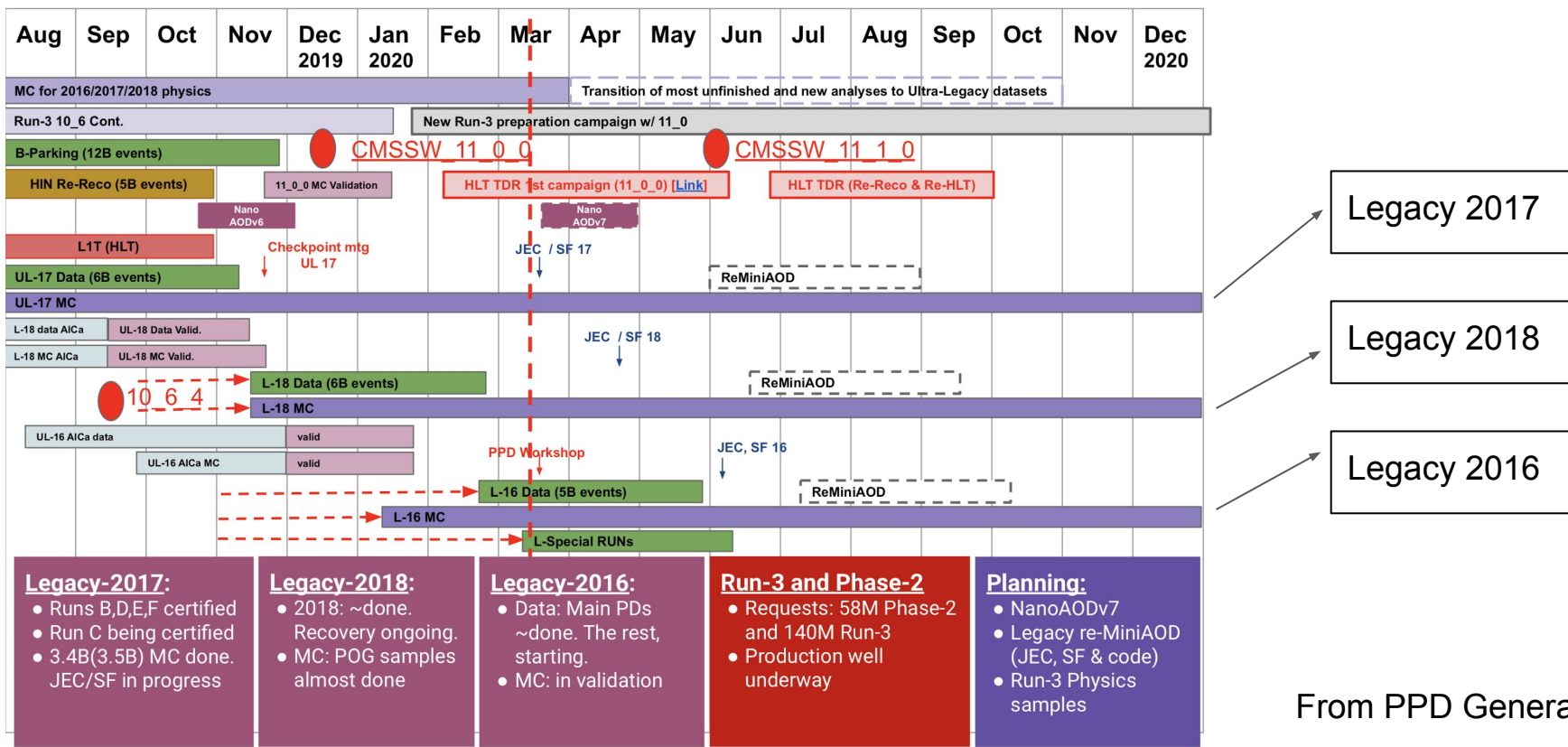
*Chulalongkorn University, National center of Physics (PK),
University of Hamburg*

Tentative time schedule (last year)



From April CMS week, time schedule for UL: [Link](#)

Current time schedule



Highway to Legacy validation

-> **Three major validation campaigns (with various successive iterations)**

Standard (low-stats) validation: reports in ValDB, green-light from validators

- No PU, classical mixing and premix PU simulation, selected runs in data

High-stats validation for POGs (to be compared with RunII MC samples and data)

High-stats validation for trigger group (trigger run in native release - see later)

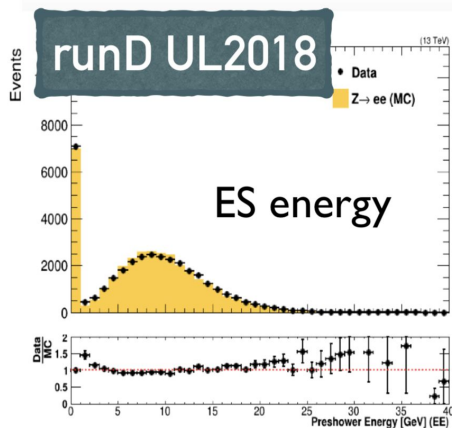
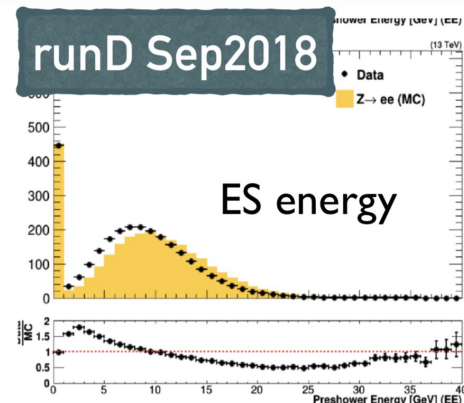
-> Discussion within groups, report at PPD general -> **GREEN LIGHT**

-> **MC and data production starts**

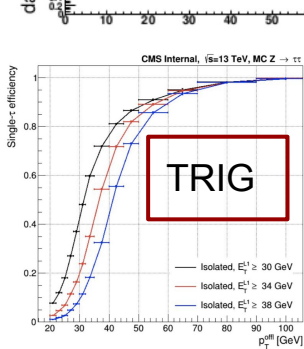
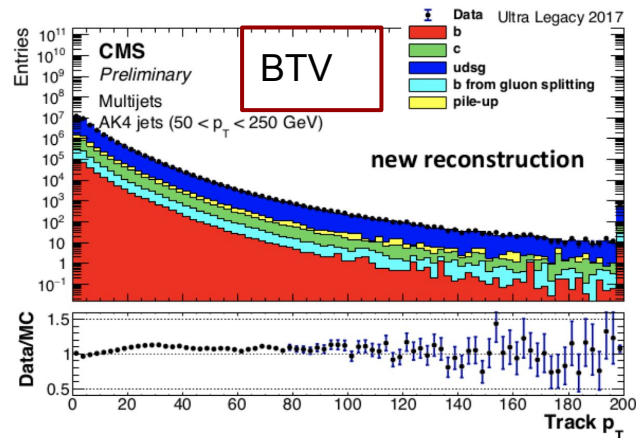
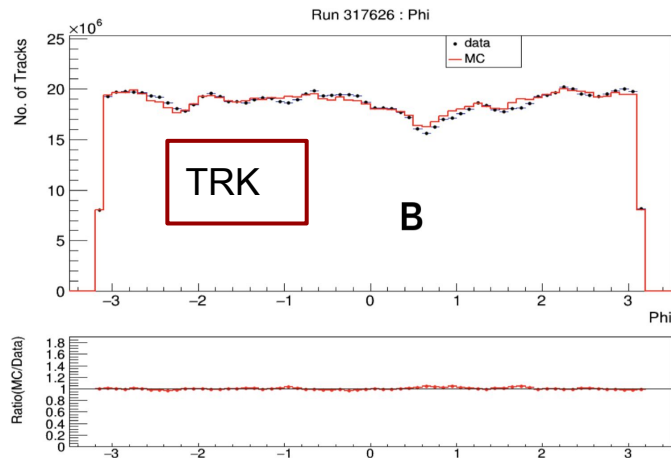
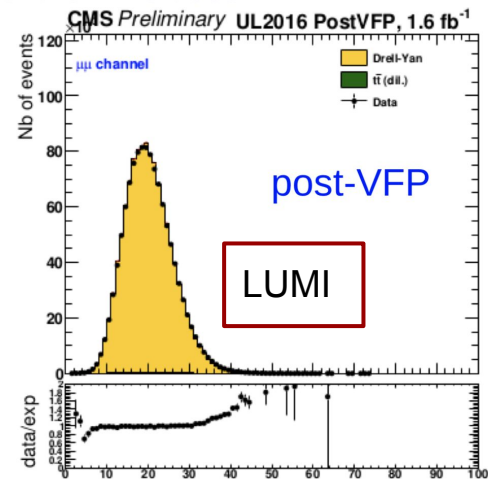
Legacy 2017: final validation on [25 Jul 2019](#)

Legacy 2018: final validation on [24 Oct 2019](#)

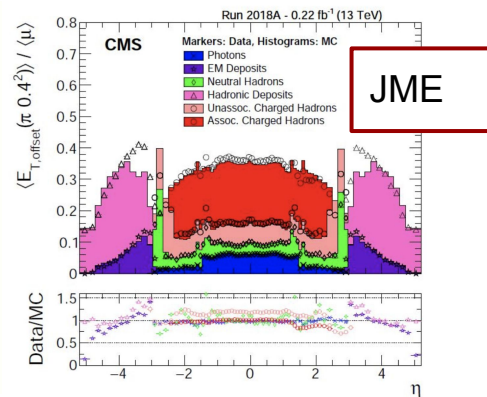
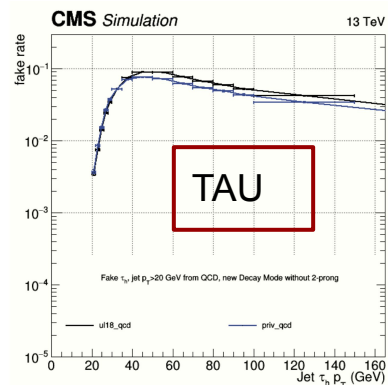
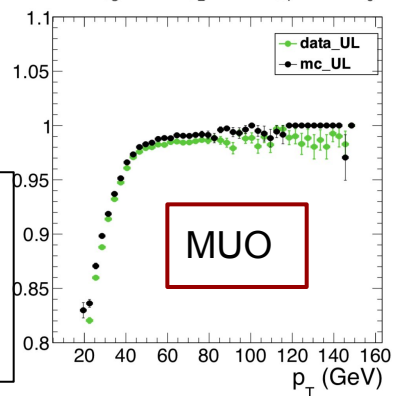
Legacy 2016: final validation (except trigger part) on [30 Jan 2020](#)



Highway to Legacy validation



A lot of
delivered
work from
all groups!



Legacy data production

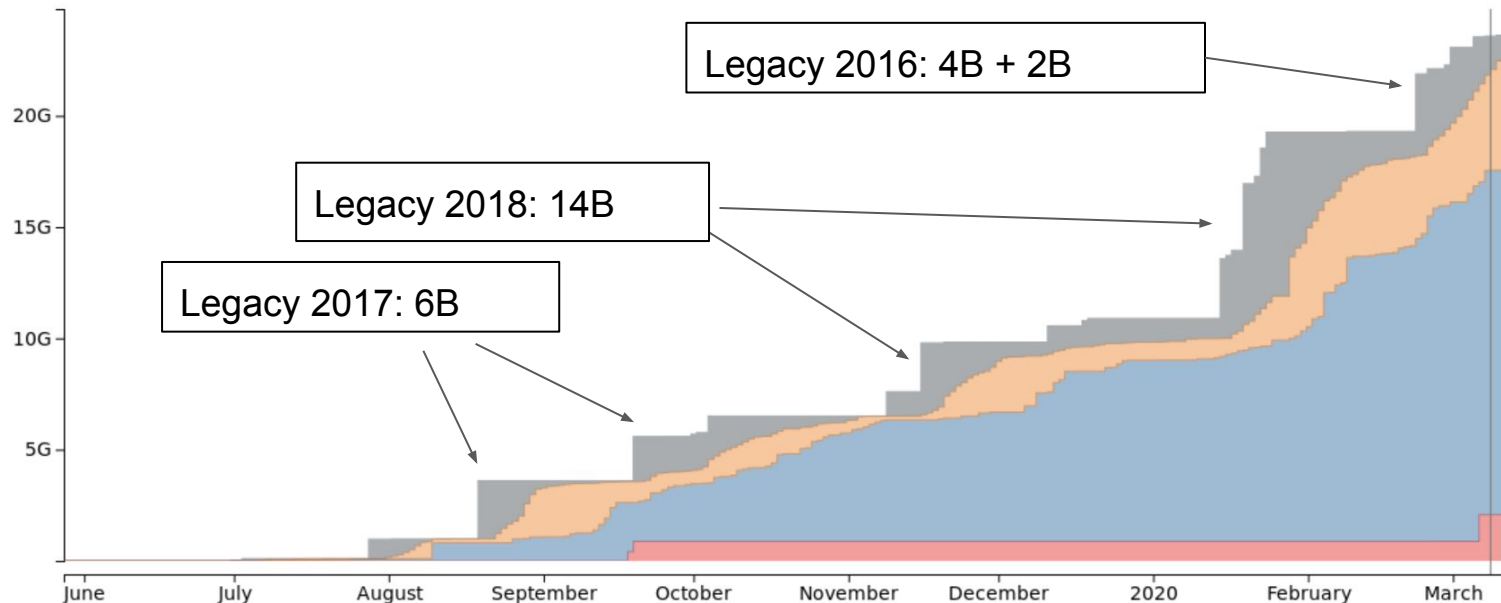
Time: Sun, Mar 8th, 2020, 12:57

Expected events: 23.58G

Events in DAS: 22.09G

Done events in DAS: 15.47G

Deleted/Invalid events: 2.07G



Generated: Tuesday, March 10th, 2020, 20:36

Last update: Tuesday, March 10th, 2020, 20:17

For input: UltraLegacy2017, UltraLegacy2016, UltraLegacy2018

All years injected at high priority and processed in 2-3 months

Legacy data production

Legacy 2016: two different versions for preVFP and postVFP

- Run B-E: preVFP with HIP mitigation
- Run F: part1 in preVFP and part 2 in postVFP
- Run G-H: postVFP without HIP mitigation

No difference/confusion for users: whatever is centrally produced should be used for analyses

- **Dataset name indicates whether HIP mitigation was included or not**

Special runs:

- UL Cosmics: all injected ([Link](#))
- 2017 ZeroBias for LUMI group injected ([Link](#))
- 2018 ZeroBias for LUMI group: waiting for GT, including new Reco beamspot
- 2016 ZeroBias for LUMI group injected ([Link](#))
- ZeroBias for scouting in preparation ([Link](#))
- LowPU for all years in preparation
- 5 TeV run in 2017 in preparation

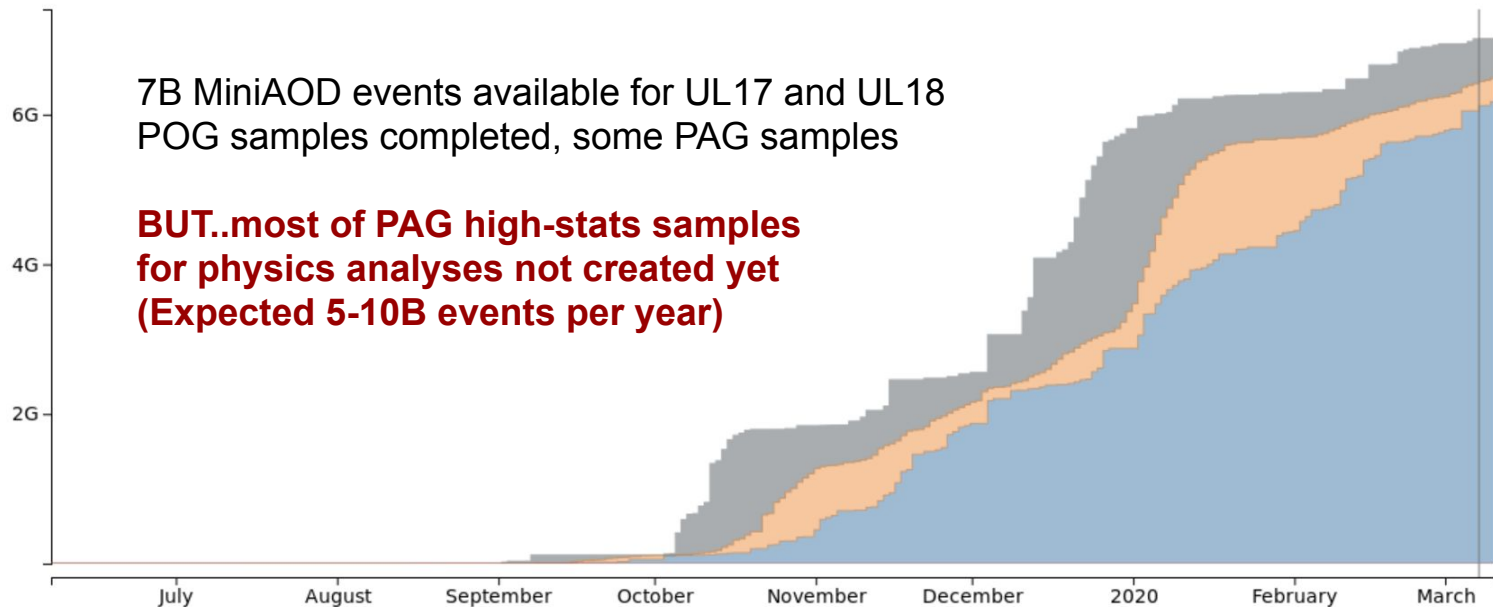
Legacy Monte Carlo production

Time: Sat, Mar 7th, 2020, 12:11

Expected events: 7.01G

Events in DAS: 6.44G

Done events in DAS: 6.12G



Generated: Tuesday, March 10th, 2020, 20:37

Last update: Tuesday, March 10th, 2020, 20:17

For input: RunIISummer19UL18MiniAOD, RunIISummer19UL17MiniAOD

Legacy 2016 MC production

STATUS: Campaign configuration not fully green-lighted:

- > Little differences in trigger menus in native release
- > Discussion ongoing with HLT and L1T groups
- > New relvals submitted with new GT ([Link pre](#) and [post](#))

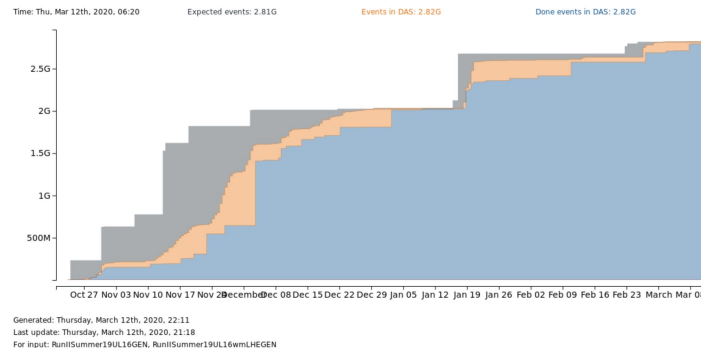
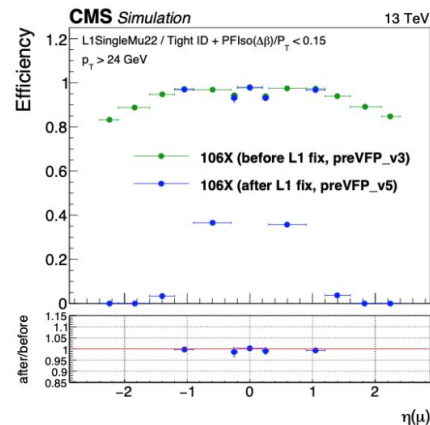
[Link](#) to presentation in PPD general

(FUTURE) CAMPAIGN ORGANIZATION:

- As for data, there will be two different eras:
 - With and without HIP mitigation
- Currently, priority on campaign without HIP mit.
- When campaign has started, production of similar physics samples with HIP mitigation

N.B. Need for new GEN requests for HIP campaign!
Groups and MC contacts will be asked to create them

Separate all campaigns w/ and w/o mitigation from GEN to Nano



Analysis table for Legacy in preparation

	Legacy 2016	Legacy 2017	Legacy 2018
CMSSW (Production)	CMSSW_10_6_8_patch1	CMSSW_10_6_2	CMSSW_10_6_4_patch1
CMSSW (Analysis)	CMSSW_10_6_10 (WILL BE UPDATED)		
Data	AOD: 21Feb2020_UL16 MiniAOD: 21Feb2020_UL16 NanoAOD: -----	AOD: 09Aug2019_UL2017 MiniAOD: 09Aug2019_UL2017 NanoAOD: -----	AOD: 12Nov2019_UL2018 MiniAOD: 12Nov2019_UL2018 NanoAOD: -----
Int. lumi.	---- /fb	---- /fb	---- /fb
GT for data Analysis (AOD, mini, nano)	106X_dataRun2_v27	106X_dataRun2_v20	106X_dataRun2_v24
MC Digi-Reco	RunII Summer19UL16RECO	RunII Summer19UL17RECO	RunII Summer19UL18RECO
MC MiniAOD	RunII Summer19UL16MiniAOD	RunII Summer19UL17MiniAOD	RunII Summer19UL18MiniAOD
MC NanoAOD	RunII Summer19UL16NanoAOD	RunII Summer19UL17NanoAOD	RunII Summer19UL18NanoAOD
GT for MC Analysis (AOD, mini, nano)	106X_mcRun2_asymptotic_v10	106X_mc2017_realistic_v6	106X_mc2018_realistic_v11_L1v1

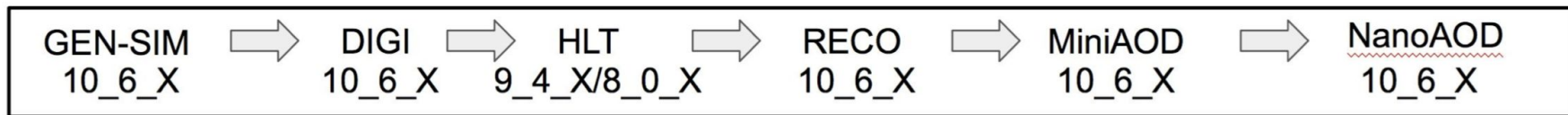
Analysis table for Legacy in preparation

	Legacy 2016	Legacy 2017	Legacy 2018
CMSSW (Production)	CMSSW_10_6_8_patch1	CMSSW_10_6_2	CMSSW_10_6_4_patch1
CMSSW (Analysis)	CMSSW_10_6_10 (WILL BE UPDATED)		
Data	AOD: 21 Feb 2020_UL16 MiniAOD: 25 Jul 2020_UL16 NanoAOD: -----	AOD: 3 Aug 2019_UL2017 MiniAOD: 6 Aug 2019_UL2017 NanoAOD: -----	AOD: 12 Nov 2019_UL2018 MiniAOD: 19 Nov 2019_UL2018 NanoAOD: -----
Int. lumi.	-----/fb	-----/fb	-----/fb
GT for data Analysis (AOD , mini , nano)	106X_dataRun2_v27	106X_dataRun2_v20	106X_dataRun2_v24
MC Digi-Reco	RunII Summer19UL16RECO	RunII Summer19UL17RECO	RunII Summer19UL18RECO
MC MiniAOD	RunII Summer19UL16MiniAOD	RunII Summer19UL17MiniAOD	RunII Summer19UL18MiniAOD
MC NanoAOD	RunII Summer19UL16NanoAOD	RunII Summer19UL17NanoAOD	RunII Summer19UL18NanoAOD
GT for MC Analysis (AOD , mini , nano)	106X_mcRun2_asymptotic_v10	106X_mc2017_realistic_v6	106X_mc2018_realistic_v11_L1v1

DATASETS READY TO BE USED FOR ANALYSIS

Workflow Monte Carlo organization

Legacy campaigns run HLT trigger step in the native release

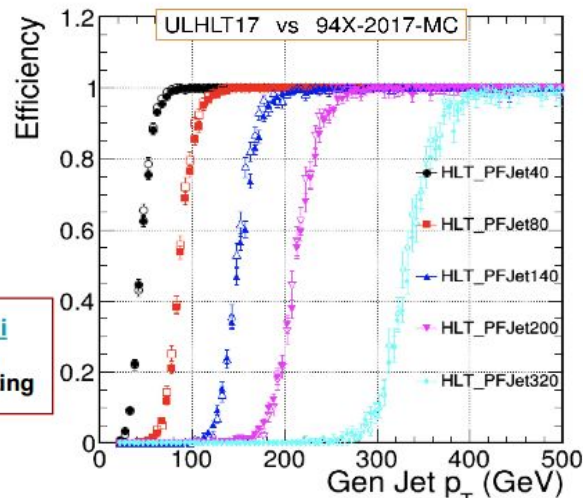


Legacy 2016	CMSSW_8_0_33_UL
Legacy 2017	CMSSW_9_4_14_UL_p1
Legacy 2018	CMSSW_10_2_16_UL

Thanks to all validators who worked on this!

[Marino Missiroli](#)

PPD general meeting



Run 3 preparation (but testing phase now)

Tool and service development*

- PdmV home page (backup)
- GEN validation in McM
- Data injection and monitoring tool
- User-friendly collection of samples per group and campaign



* If interested, please contact us:

cms-ppd-conveners@cern.ch

A physics GEN validation in McM

Requests in McM are “validated” only to test computing parameters (time/event. size/event. code crashes..) so far

IDEA:

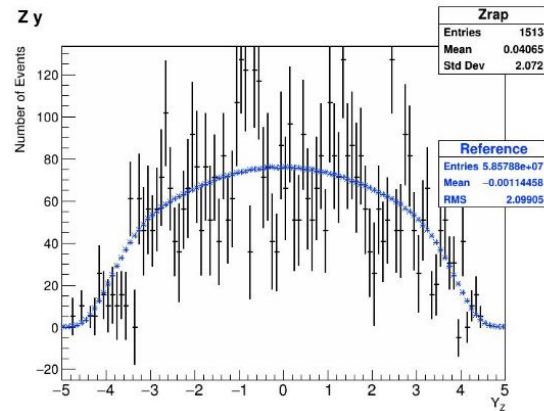
- Produce GEN validation plots for each McM request
- Use DQM GUI: automatic upload of root files from McM

GOALS:

- Give a set of reference plots at a small scale, before sample production at large scale
- Avoid production of buggy samples (from a physics point of view)

STATUS:

- Feature fully implemented in McM
- Working on reference production and definition of histograms for “special” requests (signals/BPH..) together with GEN
- **Extensive tutorial on Wednesday ([McM/pmp](#))**



[PPD dev. meeting presentations: Link](#)

ReReco web-based tool

Injection and monitoring web-based database for data production (“a-la-McM”)

STATUS:

- Tool already in place and fully working
- First (real) workflows already injected

..to replace (error-prone) twikis ([Link](#)) with a database

ReReco

SUBCAMPAIGNS

SUBCAMPAIGN TICKETS

REQUESTS

Requests

Columns

- | | | | | | |
|--|---|--|--|---|---|
| <input checked="" type="checkbox"/> PreplD | <input checked="" type="checkbox"/> Actions | <input checked="" type="checkbox"/> Status | <input checked="" type="checkbox"/> Memory | <input checked="" type="checkbox"/> CMSSW Version | <input checked="" type="checkbox"/> Notes |
| <input type="checkbox"/> Energy | <input type="checkbox"/> Step | <input type="checkbox"/> Sequences | <input type="checkbox"/> History | <input type="checkbox"/> Input dataset | <input type="checkbox"/> Subcampaign |
| <input type="checkbox"/> Output datasets | <input type="checkbox"/> Priority | <input type="checkbox"/> Processing String | <input type="checkbox"/> Runs | <input type="checkbox"/> Size per Event | <input type="checkbox"/> Time per Event |
| <input type="checkbox"/> Computing Workflows | <input type="checkbox"/> Dataset Name | | | | |

PreplD	Actions	Status	Memory	CMSSW Version	Notes
ReReco-Run2018A-Charmonium-12Nov2019_UL2018-00001	Edit Delete cmsDriver Job dict Next	approved	14000 MB	CMSSW 10.6.4.patch1	Single sequence!
ReReco-Run2018A-Charmonium-processing_string-00001	Edit Delete cmsDriver Job dict Next	approved	8000 MB	CMSSW 10.6.4.patch1	Multiple sequences

PLAN: move all injections to the new tool and communicate campaigns/samples through the new page (Tutorial soon!)

Per-group-sample page

- **IDEA:**

- provide McM non-experts with a list of samples created for a particular campaign + POGs/PAGs

- **GOAL:**

- Find samples relevant to particular analyzers/analyses and their status
 - Help PC to plan budget/priority increase
 - Avoid difficult navigation in McM
 - Use lists of samples for a given production cycle (e.g. "Run-2") to define baseline production for next cycle (e.g. "Run3").

- **STATUS**

- First working prototype in place: <https://cms-pdmv-dev.cern.ch/samples>
 - Ready to be distributed to POGs for tagging and organizing samples relevant to them
 - Groups will be asked to tag only UL17 campaign, we will do the rest ;)
 - > Tags will be propagated to other UL campaigns
 - > List for Run3 will be automatically created

Per-group-sample page: how it works

WORK IN PROGRESS

PdmV Samples

Existing samples

Campaign group
RunII Summer19UL16
RunII Summer19UL17
RunII Summer19UL18

[GJets_HT-600ToInf_TuneCP5_13TeV-madgraphMLM-pythia8](#)

[JME-RunII Summer19UL18wmLHEGEN-00005](#)

Status: submitted
Priority: 110004
Events: 4905665/5000000 (98.1%)
[Output in DAS](#)

[JME-RunII Summer19UL18MiniAOD-00029](#)

Status: submitted
Priority: 110004
Events: 4741551/5000000 (94.8%)
[Output in DAS](#)

Premix

- JME
- [Add PPD](#)

List of relevant campaigns (e.g. UL16, 17 and 18)

- Possibility to tag per group and per campaign

PdmV Samples - RunII Summer19UL18wmLHEGEN

I am going to tag samples for **physics working group**

- New page per group automatically created

Groups will find their relevant list in a page like this: [JME example](#)

Status and plans

Legacy data and MC production is well in track:

- Data processing almost completed
- MC POG samples ready for 2017/2018, PAG samples growing
- MC for 2016 waiting for final validation

Run II samples to be replaced by Legacy at the end of the year

Active development group for improvement of current tools and creation of new ones to optimize the injection workload and the monitoring for the users:

- Page containing the list of samples per group
- Data injection web-based tool
- Monte Carlo validation from McM

TARGET: continuous development and interaction with users, with delivery of “products” during 2020

Per-group-sample page



Tagging a sample from the list is very easy, BUT.. requires action from conveners/users

GJets_HT-600ToInf_TuneCP5_13TeV-madgraphMLM-pythia8

JME-RunIISummer19UL18wmLHEGEN-00005	JME-RunIISummer19UL18MiniAOD-00029
Status: submitted	Status: submitted
Priority: 110004	Priority: 110004
Events: 4905665/5000000 (98.1%)	Events: 4741551/5000000 (94.8%)
Output in DAS	Output in DAS

Premix

- JME
- [Add PPD](#)

Each group has its own dedicated page which requires minimal maintenance

- only tagging new samples when requested in McM

Existing samples per PWG

Campaign group	PWGs																														
RunIISummer19UL16	B2G	BPH	BRI	BTV	CPF	CSC	DTS	ECA	EGM	EXO	FSQ	FWD	GEN	HCA	HGC	HIG	HIN	JME	L1T	LUM	MUO	PPD	PPS	RPC	SMP	SUS	TAU	TOP	TRA	TRK	TSG
RunIISummer19UL17	B2G	BPH	BRI	BTV	CPF	CSC	DTS	ECA	EGM	EXO	FSQ	FWD	GEN	HCA	HGC	HIG	HIN	JME	L1T	LUM	MUO	PPD	PPS	RPC	SMP	SUS	TAU	TOP	TRA	TRK	TSG
RunIISummer19UL18	B2G	BPH	BRI	BTV	CPF	CSC	DTS	ECA	EGM	EXO	FSQ	FWD	GEN	HCA	HGC	HIG	HIN	JME	L1T	LUM	MUO	PPD	PPS	RPC	SMP	SUS	TAU	TOP	TRA	TRK	TSG

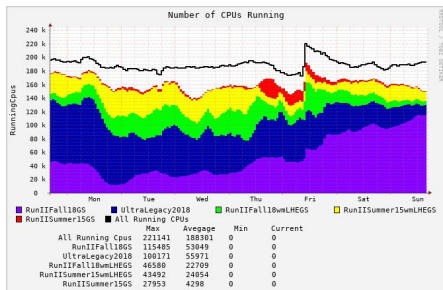
PdmV home page

What is CMS currently processing?

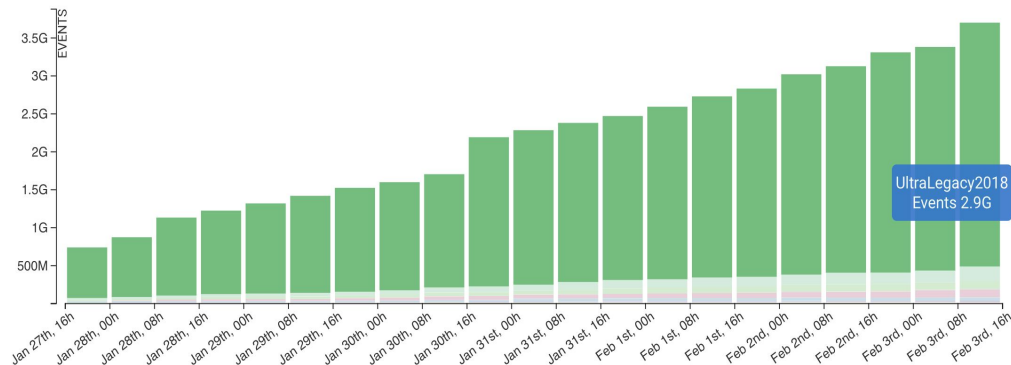
<http://cms-pdmv.cern.ch/home/>

+ Summary plots of number of produced events vs time/campaign/physics group

(different from CPU running per campaign, used in computing)



Event Production



+ Portal to the standard Monte Carlo and data monitoring tools:

- McM (Monte Carlo management)
- pMp (production monitoring platform)
- ValDB (validation database)
- RelMon (release monitoring)

Campaign	Events	pMp Links	McM Links
UltraLegacy2018	3.21G	Present Historical	

Priority Block Filter

Campaign Filter (selected 10 out of 24)

+ Collection of useful links to documentation, activities in PdmV, mandate of the group..