

Analysis trains – Status & experience from operation

Mihaela Gheata

Summary of central trains – QA(1)

- PWG1 QA train ([PWG1/PilotAnalysis/PilotTrain*.C](#))
 - 2 utility tasks (CDB connect + physics selection)
 - 20 user tasks, evolving very fast (2 new introduced the last month)
 - 28 wagons, below 2GB resident memory (but close), output size below 30 MB
- QA runs automatically for each run after reconstruction but was started on demand for different simulation productions
 - The default QA train is replaced when a new AN tag is deployed (weekly). All new data is processed by it. Train are indexed to avoid clashes.
 - We will run simulation QA automatically from now on.

Summary of central trains – QA(2)

- The output of QA trains is stored in the AliEn paths:
[/alice/data/2010/LHC10d/<run#>/ESDs/passN/QA<NN>](#)
 - Output files: QAresults.root, event_stat.root, trending.root
 - The same run can be processed by several QA trains
 - On request, if new code was added to QA
 - Due to bugs or instabilities
 - Generally the last for a given run is the best
- The train is quite big (close to the 2GB limit) so we may need to limit in future the number of histograms
 - The output size per task is already limited to 50 MB, but there are many tasks and the number only increases
 - Now QA running stable but has to be closely followed for each new tag.
 - Some wagons (TPC, TRD, vertexing) are very important and the QA train cannot run without (otherwise need to be redone)
- Some runs with large number of events produce memory problems during merging
 - Sometimes well beyond the accepted limits

The train type can be selected

Terminate is run after task merging, producing event_stat.root and trending.root (only TPC currently)

Production description	Comment
) QA34_LHC10h_Merging: PWG1 QA train_Merging	
) QA34_LHC10h: PWG1 QA train	
) QA33_LHC10h_Merging: PWG1 QA train_Merging	
) QA33_LHC10h: PWG1 QA train	The same train as QA...
) QA32_LHC10h_Merging: PWG1 QA train_Merging	
) QA32_LHC10h: PWG1 QA train	Same as QA31 but add...
) QA31_LHC10h_Merging: PWG1 QA train_Merging	
) QA31_LHC10h: PWG1 QA train	
) QA30_LHC10h_TPC_Merging: PWG1 QA train_Merging	
) QA30_LHC10h_TPC	WG1 QA train
) QA30_LHC10h_Merging	g: PWG1 QA train_Merging
) QA30_LHC10h	QA train
) QA30_LHC10f_Merging	i: PWG1 QA train_Merging
) QA28_LHC10g2asim_Merging: PWG1 QA train_Merging	
) QA30_LHC10g_Merging: PWG1 QA train_Merging	
) QA30_LHC10f: PWG1 QA train	

One can check which runs were processed by a given QA train

Production details: QA31_LHC10h: PWG1 QA train

Job info		Software versions			Output directory		Job states					
PID	Run no.	ROOT	AliROOT			%	Total	Done	Active	Wait	Err.	Oth.
64713527	137137	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137137/ESDs/pass1/QA31		100%	43	43				
64703007	137136	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137136/ESDs/pass1/QA31		100%	18	18				
64723875	137135	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137135/ESDs/pass1/QA31		100%	51	51				
64677844	137133	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137133/ESDs/pass1/QA31		100%	16	16				
64677797	137132	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137132/ESDs/pass1/QA31		100%	6	6				
64677672	137125	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137125/ESDs/pass1/QA31		100%	5	5				
64677480	137124	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137124/ESDs/pass1/QA31		100%	8	8				
64703737	137045	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137045/ESDs/pass1_plus/QA31		100%	9	9				
64632689	137045	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137045/ESDs/pass1/QA31		100%	9	9				
64597198	137045	v5-27-06a-1	v4-21-03-AN	/alice/data/2010/LHC10h/000137045/ESDs_lowflux/pass1/QA31		100%	7	7				
10 jobs	8 runs						172	172	0	0	0	0

592	QAsim25_LHC10f7a_Merging	Completed	70%	143	112	143 (114931 - 126437)	QAsim25_LHC10f7a_Merging: PWG1 QA train_Merging	
591	QAsim25_LHC10f7a	Completed	99%	12481	12446	146 (114931 - 126437)	QAsim25_LHC10f7a: PWG1 QA train	
590	QA24_LHC10f_Merging	Completed	100%	109	109	106 (133004 - 135031)	QA24_LHC10f_Merging: PWG1 QA train_Merging	
589	QA24_LHC10f	Completed	99%	3971	3967	106 (133004 - 135031)	QA24_LHC10f: PWG1 QA train	

FILTERING trains

- Producing centrally main AODs and delta AODs for data and MC productions
 - Currently vertexing AOD, muon/dimuon filters, dielectron filter and centrality filter
- Three types so far:
 - **FILTERmuon** – produces AliAOD.root, AliAOD.Muons.root and AliAOD.Dimuons.root (no TENDER)
 - Started to run automatically after pass1
 - AliAOD.root also registered (!)
 - **FILTERpass1(2)** – produces in addition AliAOD.VertexingHF.root and AliAOD.Dielectrons.root (uses TENDER for pass1) and AliAODCentrality.root
 - **FILTERsim** – runs filtering for simulated data
- The process was not fully automatic so far due to the need of tenders with special configurations for pass1, special requests (conferences ,...) or bugs
- No merging done (memory problems in the past), so far the AOD size is chosen via *SplitMaxInputFileNumber* (which limits also execution time and memory use)
 - We will retry to merge AODs to see the limits

Production info					Jobs status							
	FILTER	- Any -										
ID	Tag	Status	Done%	Cfg	Out	Total	Done	Active	Waiting	Runs	Production description	Comment
682	FILTERmuon026_LHC10h	Running	53%			3111	1674	568	22	22 (137161 - 137546)	FILTERmuon026_LHC10h: Tender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
677	FILTERpass1_025_LHC10h	Technical stop	1%			1146	13			9 (137161 - 137370)	FILTERpass1_025_LHC10h: PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	Crash due to vertexi...
669	FILTERpass1_024_LHC10h_Merging	Completed	100%			7	7			7 (137135 - 137165)	FILTERpass1_024_LHC10h_Merging: PhysSel-> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
667	FILTERpass1_024_LHC10h	Completed	11%			659	469			7 (137135 - 137165)	FILTERpass1_024_LHC10h: PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	The same train as 02...
665	FILTERpass1023_LHC10e_Merging	Completed	100%			162	162			162 (127719 - 130848)	FILTERpass1023_LHC10e_Merging: PhysSel+Tender(v0, but not TRD) -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
661	FILTERpass1_023_LHC10e	Completed	95%			6138	5881			171 (127719 - 130848)	FILTERpass1_023_LHC10e: PhysSel+Tender(v0, but not TRD) -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
660	FILTERpass1_022_LHC10h_Merging	Completed	100%			11	11			10 (136833 - 137133)	FILTERpass1_022_LHC10h_Merging: PhysSel-> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
658	FILTERpass1_022_LHC10h	Completed	100%			79	79			10 (136833 - 137133)	FILTERpass1_022_LHC10h: PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
654	FILTERpass1021_LHC10h_Merging	Completed	100%			5	5			5 (136851 - 137045)	FILTERpass1021_LHC10h_Merging: PhysSel+Tender -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
653	FILTERpass1021_LHC10h	Completed	100%			30	30			5 (136851 - 137045)	FILTERpass1021_LHC10h: PhysSel+Tender -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
641	FILTERpass1019_LHC10e_Merging	Technical stop	100%			16	16			16 (129667 - 130848)	FILTERpass1019_LHC10e_Merging: PhysSel+Tender -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
635	FILTERpass1_019_LHC10e	Technical stop	84%			3389	2192			64 (127724 - 130848)	FILTERpass1_019_LHC10e: PhysSel+Tender -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	Crashing for many ru...
629	FILTERpass1_019_LHC10e	Software update	0%			35	0		35	1 (127719 - 127719)	FILTERpass1_019_LHC10e: PhysSel+Tender(incl.V0) -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	Train stopped becaus...
624	FILTERmuon014_LHC10g	Completed	99%			2182	2180			36 (135780 - 136377)	FILTERmuon014_LHC10g: Tender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
620	FILTERpass2017_LHC10b_Merging	Completed	100%			42	42			42 (114931 - 117222)	FILTERpass2017_LHC10b_Merging: VOTender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
619	FILTERpass2017_LHC10c_Merging	Completed	98%			56	55			56 (118506 - 121040)	FILTERpass2017_LHC10c_Merging: VOTender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
618	FILTERpass2018_LHC10d_Merging	Completed	100%			60	60			48 (122375 - 126437)	FILTERpass2018_LHC10d_Merging: PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons_Merging	
615	FILTERpass2017_LHC10c	Completed	96%			2053	1972			56 (118506 - 121040)	FILTERpass2017_LHC10c: VOTender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
614	FILTERpass2017_LHC10b	Completed	97%			764	746			42 (114931 - 117222)	FILTERpass2017_LHC10b: VOTender+PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
613	FILTERpass2018_LHC10d	Completed	97%			2995	2918			58 (122374 - 126437)	FILTERpass2018_LHC10d: PhysSel -> AODs: std(+jets)/(di)muon/vertexing/dielectrons	
608	FILTERsim+tasks016_LHC10f7a_Merging	Completed	100%			147	147			147 (114931 - 126437)	FILTERsim+tasks016_LHC10f7a_Merging: #PhysSel# AODs: std(+jets)/(di)muon/vertexing/dielectrons	

Production cycle

- An analysis branch is created at the beginning of the week
- Several tests are done via the alien handler for all types of trains
 - A tag is produced quite fast if OK
 - Reports posted in Savannah, fixes included in the branch
 - Not working wagons are excluded if the fixes come too late AND if the failing task is not among the critical ones. All QA trains missing these wagons (Vertexing, TPC, TRD) have to be redone.
- The AN tag is produced at the end of the week and the default trains are replaced in LPM
- New wagons are included only after extended tests (including memory profiling)

Central PWGn analysis trains

- The central analysis train was maintained in the development phase centrally, but the effort was taken over by PWG groups
 - Needed since the number of tasks is growing too fast to be handled centrally
- We discussed in the offline team the possibility to run PWG analysis trains centrally, for more efficiency, stability, availability of results and monitoring.
 - For all these reasons it is preferable to use the same procedures as for the current production trains (production cycle, reporting bugs, monitoring and registering results)
 - These trains should still be maintained by PWG groups, with full support from the central team
 - Using libraries, but par files could be still used at **small scale** to check the trunk if needed.
 - We will run this preferably on AODs
- We will describe the requirements (most already known) to the persons responsible and provide all support needed to put these trains in place
 - The long term maintenance will be in the hands of PWGn groups

Summary

- We have now more experience from several categories of trains which run centrally
 - Development and production have to work together, now we are using clear procedures to keep this under control
- Results are made available to the collaboration
 - PWG1 QA train results are now essential for checking run quality and finding problems in reconstruction
 - Most (almost all) data and recent MC productions have AODs
 - No excuse not to use them in analysis
- We will organize analysis trains that will be integrated with the central tools for production
 - Managed by PWG groups