

Status of the LEGO framework

Costin Grigoras,
Jan Fiete Grosse-Oetringhaus,
Markus Zimmermann

26/06/2014

- LEGO framework is the system for organized analysis in ALICE
key elements: MonALISA, LEGO backend, LPM and AliEn
- active = run since the last Offline week presentation on the 20th of March
- 58 active trains (out of 78 in the system) (for all PWGs)
- 107 active users (out of 168 in the system)
- Per month statistics (from 03.2014 to 06.2014)
(compared to status from 10.2013 to 03.2014)
 - 730 train runs (490)
 - 2.8 Million jobs (2 Million)
 - 365 years wall time (219 years)
 - 9.3 wagons per train run (9 wagons/train run)
 - average train duration: 14h (17h),
spread depending on the PWG: 8h - 17h (6h - 21h)

Very well used system

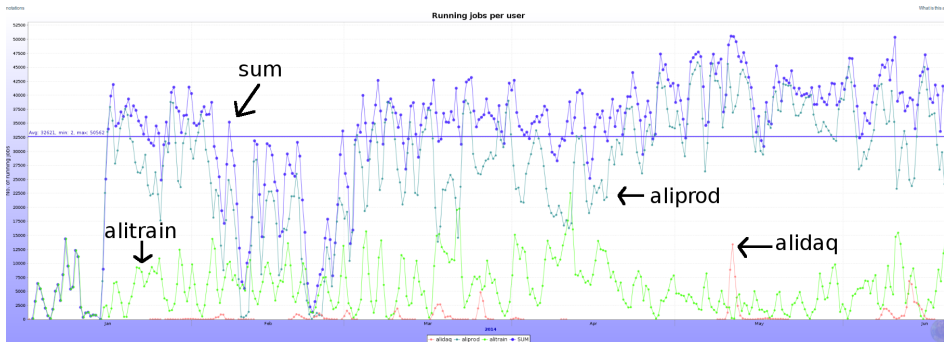


Figure 1: Performance of the GRID in the last 6 months.

user	avg number of jobs	relative to the sum
sum	37228 (up from 32200)	100%
alitrain	5778 (up from 5200)	15%
users	2290 (up from 2100)	6%

Table 1: Average number of jobs.

	10.2013 - 03.2014	03.2014 - 06.2014
Total wall time	220y	365y
Total number of train runs	489	730
Train duration	17:34	14:01
Submission	2:21	0:37
Analysis & per Run merging	14:20	12:40
Final merging	1:22	0:44
ESD	28y (13%)	57y (16%)
AOD	140y (64%)	197y (54%)
MC ESD	33y (15%)	51y (14%)
MC AOD	19y (8%)	42y (11%)

Table 2: Per month normalized status values of the train system.

Skip processing per run feature

- the 'Skip processing per run' feature which was introduced before the last offline week allows to process multiple run numbers in a single job
- flag to indicate that a wagon or dataset cannot use the skip processing per run feature was introduced
- need this because some analysis need runwise output, e.g. with $p_{T,hard}$ bins

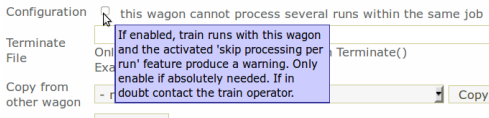


Figure 2: Wagon flag which defines if the 'Skip per Run processing' feature can be used in a train run with this wagon.

- finished internal note on the train system
- available at <https://aliceinfo.cern.ch/Notes/node/342>

Contents

1	Introduction	4
2	Definitions	5
2.1	Train	5
2.2	Used Infrastructure	6
2.3	People Working with the Train	7
2.4	Monitoring the Train	7
3	How to Use the Train System	8
3.1	Workflow Overview	8
3.2	Train Index	10
3.3	Train Overview	11
3.3.1	General Section	11
3.3.2	Wagon Section	12
3.3.3	Dataset Section	13
3.3.4	Configuration Section	15
3.3.5	Train Runs	16
3.4	Usage Statistics	20
3.5	Using Own Files with the Train	22
3.6	Local Testing	22
4	Components of the LEGO Train System	24
4.1	The Database	24
4.2	The LPM	30
4.3	The Backend Machine	30
4.4	ROOT and AliRoot Versions	31
4.4.1	Job States	31

Figure 3: Internal note on the train system.

- meeting with the train operators
- discussed which improvements the operators would like to have
- minutes at <https://indico.cern.ch/event/323428/>
- got 6 pages of suggested improvements, e.g.
 - change some interface features (e.g. allow users to switch groups, ...)
 - introduce META wagons (run on a dataset with multiple parameter settings of one wagon)
 - introduce META datasets (define multiple reference productions in one dataset)

- Implement discussed improvements
- Submitted abstract for the ACAT conference
- Talk on the Juniors Day

BACKUP

statistics for all trains

PWG	Number of Trains	Total Wall Time	Total Number of train runs	Total Number of jobs	average train duration (run_end)	average train duration (merging finished)	average train duration (submit to masterjobs submitted)	average train duration (masterjobs submitted to final merge submitted)	average train duration (final merge submitted to train finished)
CF	7	151y 122d	338	773725	164d 9:31	11:55	36m 20s	10:56	22m 32s
DQ	3	38y 351d	128	349633	118d 12:13	8:26	30m 53s	7:24	30m 25s
GA	11	133y 218d	307	1557391	160d 9:47	14:15	35m 34s	12:53	46m 44s
HF	10	244y 64d	493	2357752	184d 12:32	17:13	54m 29s	15:22	56m 11s
JE	8	247y 311d	438	1906528	213d 20:50	13:34	28m 19s	12:13	52m 45s
LF	10	199y 88d	219	1052932	333d 14:14	16:21	22m 11s	14:57	1:02
PP	3	24y 12d	111	238615	79d 18:05	15:29	40m 44s	13:59	49m 34s
UD	2	3y 351d	14	30985	103d 7:06	9:49	23m 8s	8:54	31m 5s
ZZ	4	58y 148d	137	317014	161d 12:00	9:04	23m 21s	8:30	10m 28s

sum	58	1101y 208d	2185	8584575	187d 21:14	14:01	36m 57s	12:40	44m 6s

Figure 4: Performance of the GRID in the last 3 months.

Total wall time	Total number of train runs	Train duration	Submission	Analysis & per Run merging	Final merging
1318y	2934	17:34	2:21	14:20	1:22

Table 3: Values from March 2014 over the previous 6 months.

statistics for all trains

PWG	Total Wall Time	Total ESD	% of total for ESD	Total AOD	% of total for AOD	Total MC ESD	% of total for MC ESD	Total MC AOD	% of total for MC AOD
CF	151y 122d	-	0	144y 34d	95	2d 22:35	0	7y 84d	5
DQ	38y 351d	19y 127d	50	17y 358d	46	-	0	1y 16d	3
GA	133y 218d	18y 20d	14	41y 354d	31	65y 327d	49	7y 246d	6
HF	244y 64d	31y 341d	13	152y 35d	62	11y 114d	5	48y 302d	20
JE	247y 311d	93y 107d	38	40y 234d	16	69y 57d	28	44y 276d	18
LF	199y 88d	5y 192d	3	170y 229d	86	7y 334d	4	15y 59d	8
PP	24y 12d	4y 331d	20	19y 27d	79	-	0	-	0
UD	3y 351d	-	0	3y 351d	100	-	0	-	0
ZZ	58y 148d	2d 9:56	0	196d 23:48	1	-	0	-	0

sum	1101y 208d	173y 28d	16	590y 363d	54	154y 125d	14	124y 257d	11

Figure 5: Performance of the GRID in the last 3 months.

ESD	ESD %	AOD	AOD %
168y	13%	840y	64%
MC ESD	MC ESD %	MC AOD	MC AOD %
198y	15%	111y	8%

Table 4: Values from March 2014 over the previous 6 months.