

Lego Train Optimization

Costin Grigoras,
Jan Fiete Grosse-Oetringhaus,
Markus Zimmermann

07/11/2013

- improvement suggestions
- stop to slow jobs
- Dataset consolidation
- Optimizing splitting parameters

Improvement suggestions

- analyze performance of a train run and notify the operator automatically if something could be improved for the next train runs
- notify operator if
 - more than 5% of the jobs are in the same error state
 - more than 10% of all the jobs are in an error state
 - the train jobs finish very quickly: the splitting parameter could be increased and the ttl decreased
 - the train jobs finish very slowly: the splitting parameter could be decreased and the ttl increased

Improvement suggestions

- 95% of the jobs are finished after 56m 56s, the number of files per job (SplitMaxInputFileNumber) can be increased.
- Most of the jobs (95% after 56m 56s) are finished faster than 1/3 of the time to life (TTL) of the jobs. To get the results faster the TTL could be reduced in the dataset configuration. Now the TTL is 12:00. We suggest twice the time after which 95% of the jobs had finished: 1:53 (= 6833s).

Figure 1: Improvement suggestion for a finished train run.

Stop to slow jobs

- If a train has $>90\%$ of jobs in a final state and $>85\%$ of jobs in DONE, the remaining waiting jobs can run at any site (not only at the sites where the data is)
- if a train has $>98\%$ jobs in a final state, the remaining running jobs (at max 2%) are stopped.

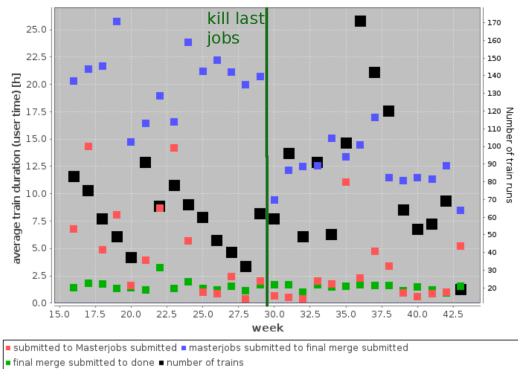


Figure 2: Train performance of the last 6 months.

Dataset consolidation

- each file is saved at **replica** different sites
- combination of sites at which a file is saved is called a **basket**
- only files from the same basket can run together in one job (according to the splitting algorithm)
- reduce number of baskets per dataset → reduce number of jobs to analyze this dataset
- increase input files per job: e. g. mixed event analysis gets more statistics

file	SE 1	SE 2	SE 3
A	x	x	-
B	x	-	x
C	-	x	x
D	x	x	-

- reduce number of baskets by moving some files to other sites
- while doing this could also unify the number of replicas per basket
- example reason: lost a disk in a SE

file	SE 1	SE 2	SE 3
A	x	x	-
B	x	x	-
C	x	x	-
D	x	x	-

reduction of jobs

dataset	files	moved files	jobs previous	jobs after	splitting	reduction
LHC10c_AOD137	4695	-	885	200	10	77%
LHC10d_AOD135	4205	-	1528	418	10	73%
LHC10h_AOD086	27716	9693	7413	4377	7	41%
LHC11h_AOD115	130830	1111	17893	13032	5	27%
Hijing_LHC10h	56543	2434	3963	2054	20	48%

Table 1: Consolidation of different datasets

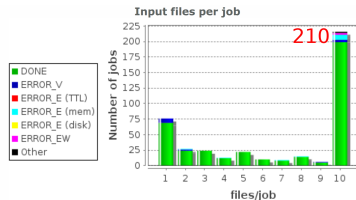
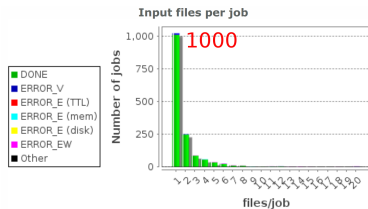


Figure 3: Splitting before the consolidation up of

Figure 4: Splitting after the consolidation of

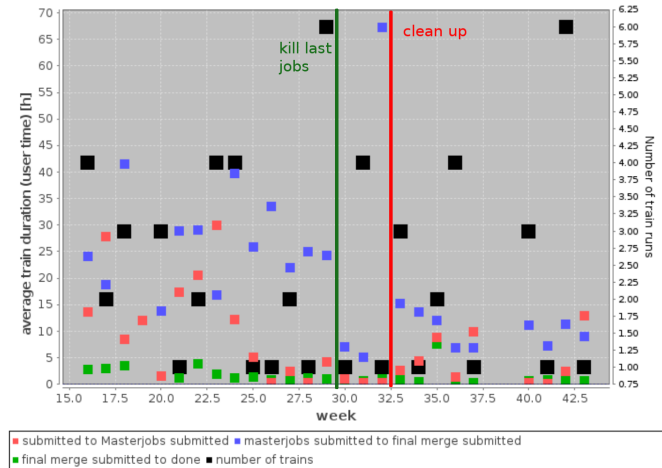


Figure 5: Running time of the datasets LHC10h_AOD086. Shown are the calendar weeks in 2013.

- some sites are not allowed to be removed (GSI)
- some sites shouldn't be the aim of a copy operation
- take into account the different types of storage elements
- until mid October only very few consolidation operations were successful executed
- executed quite some operations in October → small statistics

- dataset consolidation reduces the number of jobs by the factor 2 - 5
- improve the statistics per job for the mixed event analysis
- don't have enough statistics to see significant changes in the running time

Optimizing splitting parameters

- give the train operators an educated guess about a good ttl and splitting parameter
- try to find a correlation between the test time and the average running time
- automatize the suggestions for the ttl and splitting parameter
- find parameters which describe a good standard test
- adjust test environment so that in most cases a standard test is executed

- plot the test and the running time of the train runs for each train separately
- fit the running time dependent on the test time
- for future train runs:
extrapolate the running time on the basis of the fit

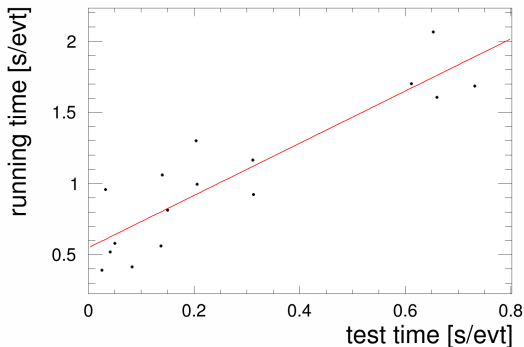


Figure 6: Example of the fitting procedure for train 18 (CF_PbPb_MC_AOD).

For the determination of the fits only certain train runs are used for each train

- with more than 500 test events
- from the last 4 months
- with a test time less than 1 s/evt
- with a running time less than 7 s/evt
- fit only with 5 or more points

- use the train runs which fulfill these requirements to get the fit for each train
- for a new train test use the correspondent fit to extrapolate the running time

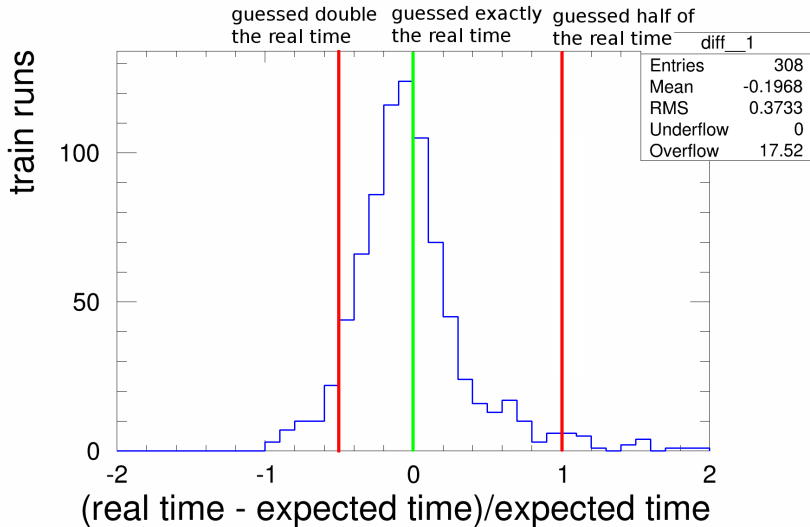


Figure 7: Validate the extrapolation.

- identical: the training and testing datasets are identical and from a 4 month period
- independent: the training dataset is from a 4 month period while the testing dataset is from a different 1 month period

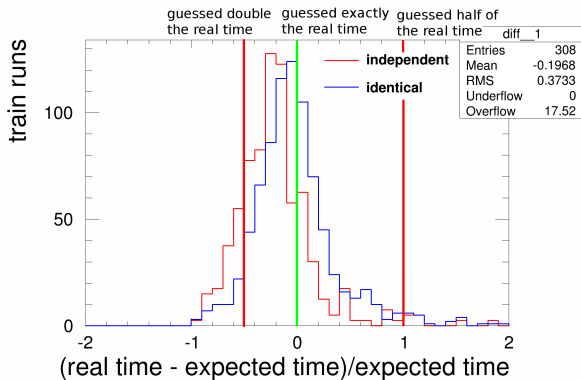


Figure 8: The independent curve is scaled to the train runs of the 4 month dataset.

- found correlation between testing time and average running time
- write automatic script which runs every month and determines the new variables for each train
- put the suggested splitting parameters for the train operators on the train pages

- Improvement suggestions should help the train operators to improve the train setup
- allowing all sites to pick up the last jobs and killing the last jobs reduces the real time of a train
- Dataset consolidation leads to a significant reduction of the number of jobs
- Dataset consolidation improves statistics for analysis like the mixed event
- found correlation between testing time and average running time

- consolidate more datasets and look into the improvements
- put optimizing splitting parameters into production
- develop an automatic update process for the fit values

- nano AODs

BACKUP

One Fit for all trains

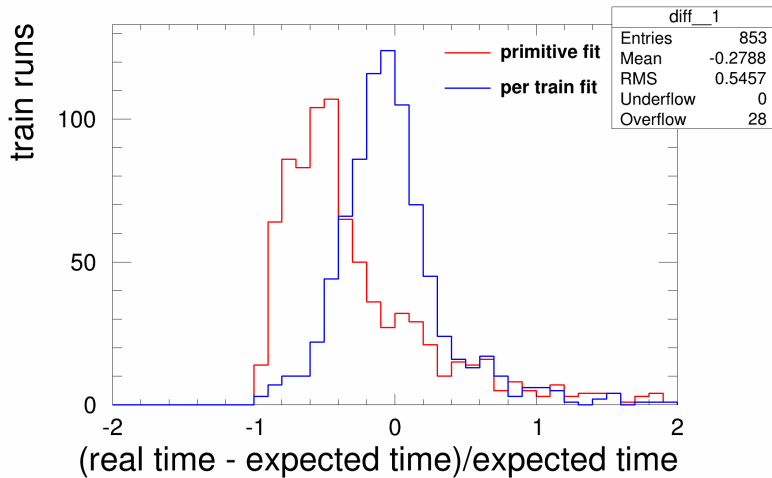


Figure 9: Validation of the primitive assumption that all train runs could be fitted with one fit.