A simulation tool for ALICE storage and computing resource usage

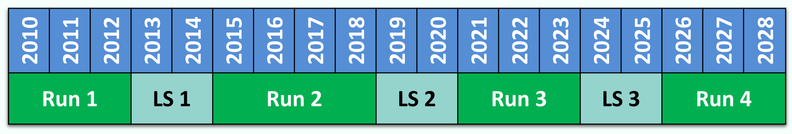
#392

Authors: **Armenuhi Abramyan**[1], **Narine Manukyan**[1], **Tim Hallyburton**[2]
Co-authors: Predrag Buncic[2], Latchezar Betev[2], Ara Grigoryan[1]

[1] A.I. Alikhanyan National Science Laboratory (Yerevan Physics Institute) Foundation
[2] The European Organization for Nuclear Research (CERN)
{Armenuhi.Abramyan, Narine.Manukyan, Tim.Hallyburton}@cern.ch

Introduction

Since its beginning, the lifecycle of LHC [1] has been composed of the data taking (Run) and upgrade (Long Shutdown, LS) periods.

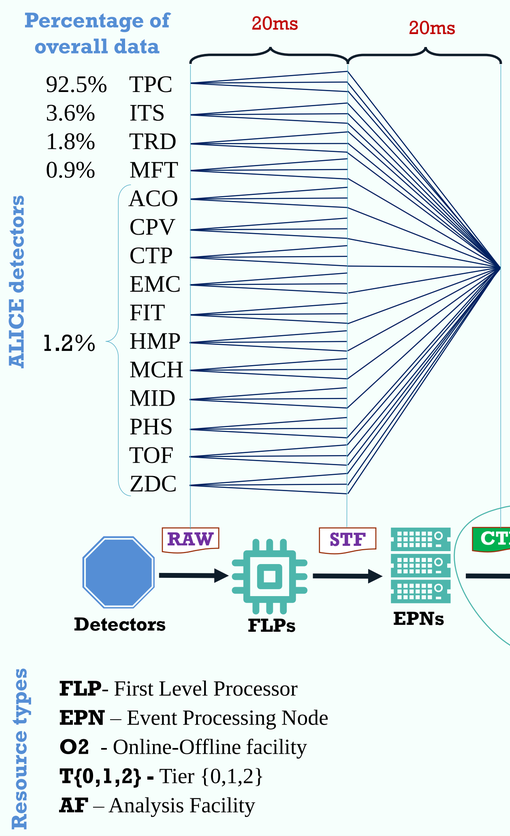


During the Run periods, the LHC experiments, such as ALICE [2], require a significant amount of storage and computing resources to store, process and analyze the experimental and simulated data. In addition to that, the requirements for these resources are increasing over the years, at each LHC running period.
In order to predict the resource requirements of the ALICE Experiment for a particular LHC Run period, we developed a flexible and highly configurable simulation tool, which does this prediction by Discrete-Event Simulation (DES) of ALICE data flow processes. The tool provides a Web GUI, which allows entering of all the necessary parameters and to graphically visualize the results of simulations.

The steps of the current developments can be summarized as follows:

- Description of the ALICE data flow processes for certain LHC Runs
- Definition of input parameters necessary for simulations.
- Simulation of ALICE data flow for Run 3 and estimation of required storage resources
- Same simulation for 2017 year of Run 2 and comparison of the results with real numbers for the validation of tool results.

ALICE data flow during LHC Run3/Run4 [3]



RAW - Data coming from detector

STF - Sub Time Frame, containing RAW data from a single **FLP** for a period of time ~20ms.

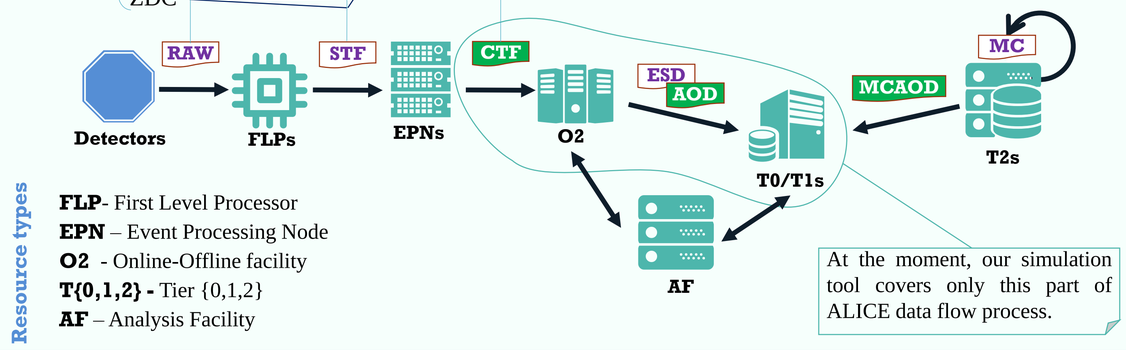
CTF - Compressed Time Frame, containing processed RAW data of all active detectors for a ~20 ms period of time.

ESD - Event Summary Data, auxiliary data to CTF, containing the output of the reconstruction process linking the tracks and clusters.

AOD - Analysis Object Data, containing the final track parameters in a given vertex and for a given physics event.

MC - Monte Carlo, Simulated energy deposits in sensitive detectors. Removed once the reconstruction of MC data is completed.

MCAOD - Analysis Object Data for a given simulated physics event.



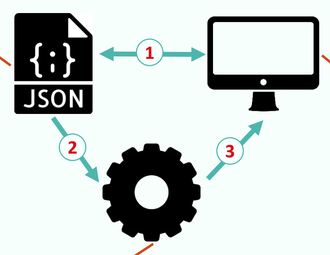
FLP- First Level Processor
EPN – Event Processing Node
O2 - Online-Offline facility
T{0,1,2} - Tier {0,1,2}
AF – Analysis Facility

At the moment, our simulation tool covers only this part of ALICE data flow process.

The simulation tool architecture

config.json
Default values of all input parameters.

Most important part of the software is its configuration with appropriate values of Run 2 and Run 3.



Simulation software
Sets the computing environment and defines data and job types and performs simulation of resource usage at each resource type.

Web GUI
Allows to manipulate with the default input parameters and run the simulation with different scenarios.
The interface is also used for the graphical representation of simulation results and allows to export these results in **PDF** format.

Input parameters

General LHC conditions

Time granularity is a **day**.

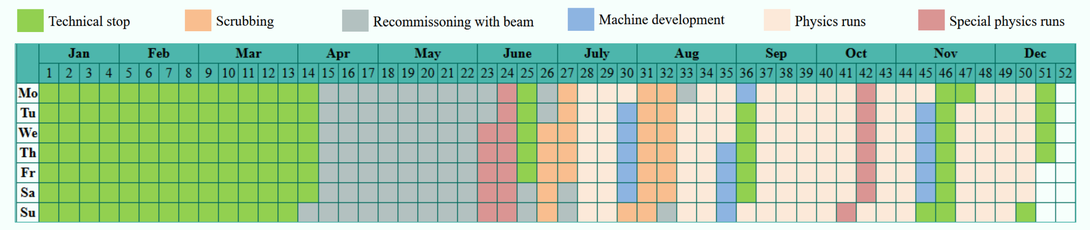
$$CTF\_size\_per\_day = (C_{rate} * E_{size} * Efficiency/100) * Seconds\_in\_a\_day$$

Planned for Run 3

Collision type	Collision Rate (C_{rate}) (N of collistions/sec)	Data taking efficiency (%)	CTF size per event (E_{size}) (kB)
pp	22396.5	50	50
pp-ref	30864	50	50
pPb	108421	50	100
PbPb	11057.5	50	1600

LHC running schedule

In order to have more realistic picture, the real running schedule of LHC is taken into account. We have used LHC running schedule of Run 2 (2015) [4], since presently the LHC schedule for Run3/Run4 is not available.



Resource types and their capacities

Total CPU and storage capacities of resource types involved in the data flow process

As of 2017

Site Type	Site Name	CPU resources (N of CPU cores)	Storage Resources		Tape resources	
O2	CERN-1	25000	600000	TB	80000	TB
T0	CERN-2	126000	23000	TB	37000	TB
T1	T1-6	44000	22000	TB	30000	TB
T2	T2-80	113485	23000	TB	0	TB
AF	AF-1	25000	10000	TB	0	TB

These numbers serve for the graphical representation of resource thresholds to see the differences between consumed and pledged resources.

Job types and their CPU consumption

These parameters are required for the estimation of computing resource usage.

The amount of CPU required for the transition of certain data types for a given collision type.

The percentage of CPU of each resource type pledged for the transition of certain data types.

Planned for Run 3

Data type transition	CpuTransformations (HS06 - consumed CPU seconds per event)				CpuShare (%)			
	pp	pPb	PbPb	pp-ref	O2	T1	T2	AF
RAW -> CTF	0	0	0	0	100	0	0	0
CTF -> ESD -> AOD	300	710	3800	300	67	33	0	0
MC -> MCAOD	1000	3000	45000	1000	0	0	100	0
AOD -> HISTO	200	700	3700	200	0	0	0	100

Data management policies

The data replication and removal policies are significantly influence on the resource usage estimation results.

Amount of disk and tape replicas for each data type

The percentage of each data type kept on each storage resource

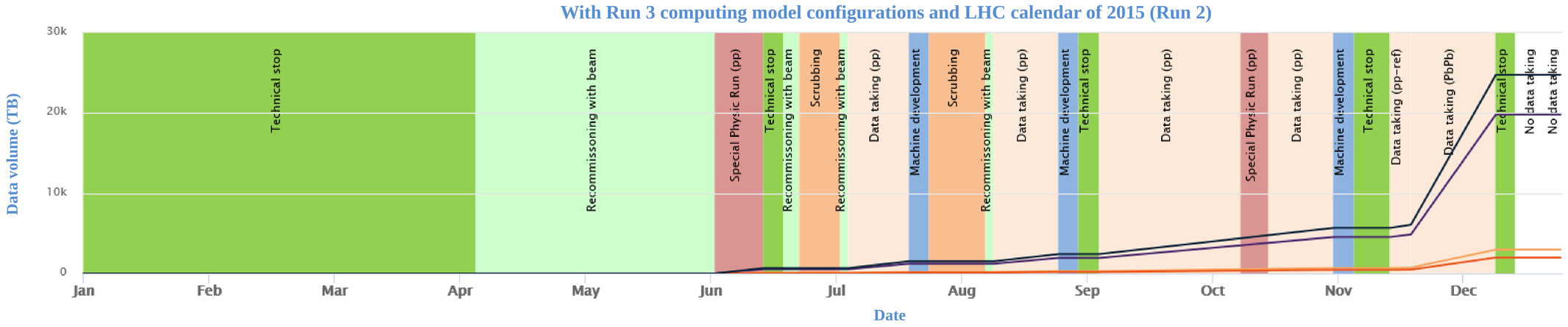
The number of days that each data type is kept on each storage resource, and then moved to tape or removed.

Planned for Run 3

Data Types	Replication factor		Storage Sharing (%)					LifeTime on Disk (days)				
	Disk	Tape	O2	T0	T1	T2	AF	O2	T0	T1	T2	AF
CTF	1	1	66.6	0.0	33.3	0.0	0.0	270	0.0	270	0.0	0.0
ESD	1	0	75.0	0.0	25.0	0.0	0.0	10	15	25	0.0	0.0
AOD	1	1	15	0.0	0.0	1.0	0.0	150	250.0	100	0.0	100
MCAOD	1	1	0	25	75	0.0	1.0	100	100.0	100.1	5.0	100
HISTO	1	0	10.0	5.0	75.0	0.0	10.0	10.0	100.0	150.0	0.0	50.0

Initial results (Estimated amount of created data during 1 year of Run 3)

With Run 3 computing model configurations and LHC calendar of 2015 (Run 2)



Monte Carlo simulations are not included in these results.

Data type	Volume (PB)
CTF	~ 19
ESD	~3
AOD	~2
SUM	**~24**

The results of same simulations for the 2017 year of Run 2 showed that the total amount of created **RAW** data is **~9 PB**, meanwhile the real number is **~8 PB** [5]. The works on the development of the tool are in progress to make the simulations as realistic as possible.

Conclusion and future work

This simulation software is a flexible and highly configurable tool, which presently allows to estimate and graphically visualize the volume of storage resources, necessary to store each type of data of ALICE experiment during LHC Run 3 period.
In order to validate the initial results of simulations, another simulation has been done for the storage resource usage during the 2017 year of LHC Run 2 and compared with real numbers.
The works on the development of the tool are continued to make the simulations more precise, as well as with the vision to make the software applicable to the other experiments of LHC.

References

[1] The LHC (Large Hadron Collider) - http://home.web.cern.ch/about/accelerators/large-hadron-collider
[2] The ALICE experiment - http://aliweb.cern.ch
[3] Alice O2 Upgrade Technical Design Report - https://cds.cern.ch/record/2011297/files/ALICE-TDR-019.pdf
[4] LHC Schedule 2015 - https://indico.cern.ch/event/367052/contributions/1782854/attachments/729631/1001227/LHC_Schedule_2015.pdf
[5] ALICE Computing Resources Usage in 2017-2018 and Resource Requirements for 2018, 2019 and 2020, Page 2 - https://indico.cern.ch/event/718770/contributions/2954296/attachments/1627208/2591768/Resources_review_and_requirements_2017-2020.pdf