

Update



Anil Panta
March 2, 2022

Ole Miss

Lifetime in gb2_ds_rep

- gb2_ds_rep is used for making any extra replicas by user/operation.
- **--lifetime xh(our) /xd(ay) /xw(eek) /xm(onth)**
- Not specified : default 1 Month.
- Defined MaxRuleLifetime : x s
- Tape SE: Unavailable for replication.
- Local SE: lifetime is None

```
[apanta@ccw05 ~]$ gb2_ds_rep /belle/user/anil123/sigma2ppi0_1010_01/sub02 -d DESY-TMP-SE -u anil123 --lifetime 2h
Do you want to proceed with the replication?:
Please type [Y] or [N]: Y
Replication rule will be associated with account anil123
Rule lifetime : 0.0833333333333333 days
RuleID: ['u'f8a837f095434558bec902f0a4c7bfa1']
Replication submitted
```

```
[apanta@ccw05 ~]$ gb2_ds_rep_status /belle/user/anil123/sigma2ppi0_1010_01/sub02 -l
```

RuleID	account	LFN/LPN	OK	Replicating	Stuck	Dest_SE	Src_SE	Created (UTC)	Expires at (UTC)
2682e3a781614da29f96469283f713f4	anil123	/belle/user/anil123/sigma2ppi0_1010_01/sub02	116	0	0	ANY=true	None	2021-11-11 03:45:21	None
f8a837f095434558bec902f0a4c7bfa1	anil123	/belle/user/anil123/sigma2ppi0_1010_01/sub02	0	116	0	DESY-TMP-SE	None	2022-03-01 21:56:35	2022-03-01 23:56:35

Collection searchable:

- Short Term: very few collection
 - But the slide suggest:
 - We can have more collection.
 - Leads to : confusion on which is what.
-
- Idea to make collection searchable.
 - Like we do for datasets with dataset-searcher.

Collection summary

- For each experiment/bucket
 - Data
 - For each beam energy
 - MCri
 - Same set for all data
 - Granularity for integrated lumi ($100+100+\dots+100=1000/\text{fb}$)
 - qqbar, mixed, charged, low mult, ... (*14)
 - MCrd
 - Same granularity as Data
 - qqbar, mixed, charged, low mult, ... (*14)
 - Skim
 - For each Data, MCri, MCrd collection
 - For each Skim (*70)
- Grand total of collection can be huge!
 - Data ($10 \text{ exp} + 10 \text{ bucket}$)* $(2 \text{ Beam energy}) = \sim 20$ collections
 - MCrd $\sim \text{Data} * 14 \sim 250$ collections
 - Skim ($\text{Data} + \text{MCri} + \text{MCrd}$)* $(\# \text{skims}) \sim O(1000)$ collections ???



But wait: that's more!
Versioning

100+ collection
w/o a DB and
queries are very
hard to use

Stefano Lacaprarà, INFN Padova - Umberto Tamponi, INFN Torino

11

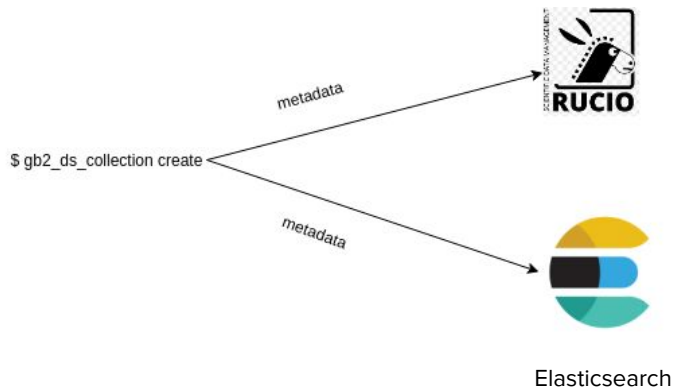
How?

- We need to **store more metadata of datasets in collection's metadata.**
- Currently : [campaign, generalSkimName , skimDecayMode , description, int_lum]
- Proposal:
 - Store all the metadata attributes that we store for datasets in DSS.
 - E.g.

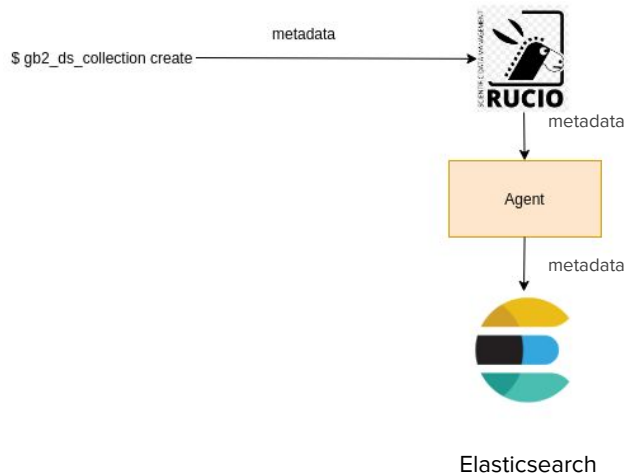
```
dataType : ['data']
dataLevel: ['mdst']
generalSkimName: ['hadron']
campaign: ["proc12", "bucket 18"]
experimentLow/High : 12
runLow/High : 11
globalTag : ['db000911', 'db00911']
mcEventType: [""]
skimDecayMode: []
beamEnergy: ["4S" , "4S_offres"]
release: ["release-05-04-01", "release-05-06-01"]
description : "some description"
Int_lum : 100.22
```

Dataset-Searcher extension:

- **New Elasticsearch index (DB)** to store collection metadata.
- Set of Command line tool and WebApp to search for collection.



OR



Example:

-> Two collection:

-> 1: campaign ["bucket16"] , beam_energy ["4S_offres"]

-> 2. campaign ["proc12", "bucket16"], beam_energy :["4S"]

```
Singularity> python insert.py campaign bucket16 -c y  
/belle/collection/Data/my_data_coll_1  
/belle/collection/Data/my_data_coll_2
```

```
Singularity> python insert.py beam_energy 4S -c y  
/belle/collection/Data/my_data_coll_2
```

```
Singularity> python insert.py campaign proc12 -c y  
/belle/collection/Data/my_data_coll_2
```

Insert , delete , update and search all working in development server.

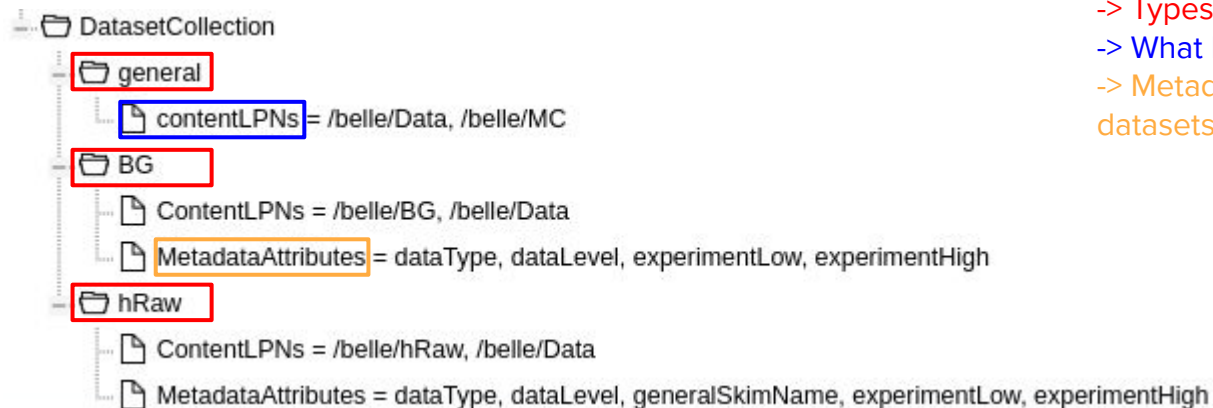
Conclusion:

- Collection searching working properly in Elasticsearch.
- But Elasticsearch is not yet available in production.
- We need to wait.

Back up

Configuration:

- The type of collection and its contents restriction are configurable.



-> Types of collection

-> What kind of LPN in collection

-> Metadata attr consistent among datasets.

/belle/collection/general/<collection_name>

/belle/collection/BG/<collection_name>

/belle/collection/hRaw/<collection_name>

gb2_ds_collection update:

- Only to **update integrated luminosity value and description.**

```
$ gb2_ds_collection update <collection_path>  
  --int_lum 200
```

```
$ gb2_ds_collection update <collection_path>  
  --description 'updated description of collection'
```

Note: -n option for multi-campaign collection will result in globaltag failure

Gbasf2 with collection

- Collection path will work as it will for any LPN.
- `gbasf2 my_script.py -i <collection_path> -p <projectname>`

```
[apanta@ccw07 ~]$ cd project/project sigma2ppi0/sigmapipi/
[apanta@ccw07 sigmapipi]$ gbasf2 analyze.anil.sigma.py -i /belle/collection/general/collection_cert_v5r2_pre4_v2 -s release-05-01-24 -p anil_cert_v5r2-pre4_v1
gbasf2 release v5r2-pre4 is newer than the recommended: v5r1p3
Requesting metadata of input files (3968 files). Please wait.
*****
***** Project summary *****
** Project name: anil_cert_v5r2-pre4_v1
** Dataset path: /belle/user/anil123/anil_cert_v5r2-pre4_v1
** Steering file: analyze.anil.sigma.py
** Job owner: anil123 @ belle (23:36:33)
** Preferred site / SE: None / None
** Input files for first job: /belle/Data/release-05-01-22/DB00001627/proc12/prod00017730/e0007/4S/r03221/hadron/mdst/sub00/mdst_000001_prod00017730_task6322000001.root
** Number of input files: 3968
** Number of jobs: 3968
** Processed data (MB): 2071582
** Processed events: 488392932 events
** Estimated CPU time per job: 32842 min
*****
Are you sure to submit the project?
Please enter Y or N: Y
Scout jobs will be submitted
Initialize metadata for the project:
No attribute. Initialize Dataset...
Dataset initialization: OK
Dataset metadata attributes already exist (30): OK
Successfully finished.
<=====v5r2-pre4=====>
Scout JobID = 282948 ... 282957 (10 jobs)
Initialize metadata for the project:
No attribute. Initialize Dataset...
Dataset initialization: OK
Dataset metadata attributes already exist (30): OK
Successfully finished.
<=====v5r2-pre4=====>
JobID = 282958 ... 283957 (1000 jobs)
<=====v5r2-pre4=====>
JobID = 283958 ... 284957 (1000 jobs)
<=====v5r2-pre4=====>
JobID = 284958 ... 285957 (1000 jobs)
<=====v5r2-pre4=====>
JobID = 285958 ... 286925 (968 jobs)
```

Auto Refresh: Disabled ▾		
Selected Statistics ▾		
Status ▾		
	Key	Value
	Checking	3239 (-2)
■	Done	15
■	Running	1
■	Waiting	548

gb2_ds_list

- gb2_ds_list is now compatible with collection.
- Use collection as you use it for LPNs.
- All the options of gb2_ds_list works with collections.

```
[apanta@ccw02 ~]$ gb2_ds_list /belle/collection/BG/test BG collection bucket16
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00694/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00695/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00696/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00698/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00699/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00722/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00723/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00724/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00726/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00727/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00728/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00729/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00730/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00731/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00732/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00733/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00734/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00735/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00736/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00741/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00742/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00743/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00744/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00745/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00746/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00747/beamgb
/belle/BG/release-05-01-10/DB00001410/BG0Exp12/prod00016763/e0014/4S/r00748/beamgb
```