Status of Offline

P.Gauzzi
on behalf of the Offline group

KLOE-2 General Meeting
22/9/2018

Datarec DBV-38

- Integrated Tracking 3.0 (2 track events, with $p_T > 10$ MeV)
- UIC scheme: new detectors and retracking after EVCL
- New FILFO (rejects 25 – 30 % of events; essentially mach. bckg/cosmics)
- New stream for Single Photon Trigger
- No DSTs produced
- No T0 corrections applied (T0-step1)
- DSTs will be replaced by the ROOT output (see Elena's presentation)

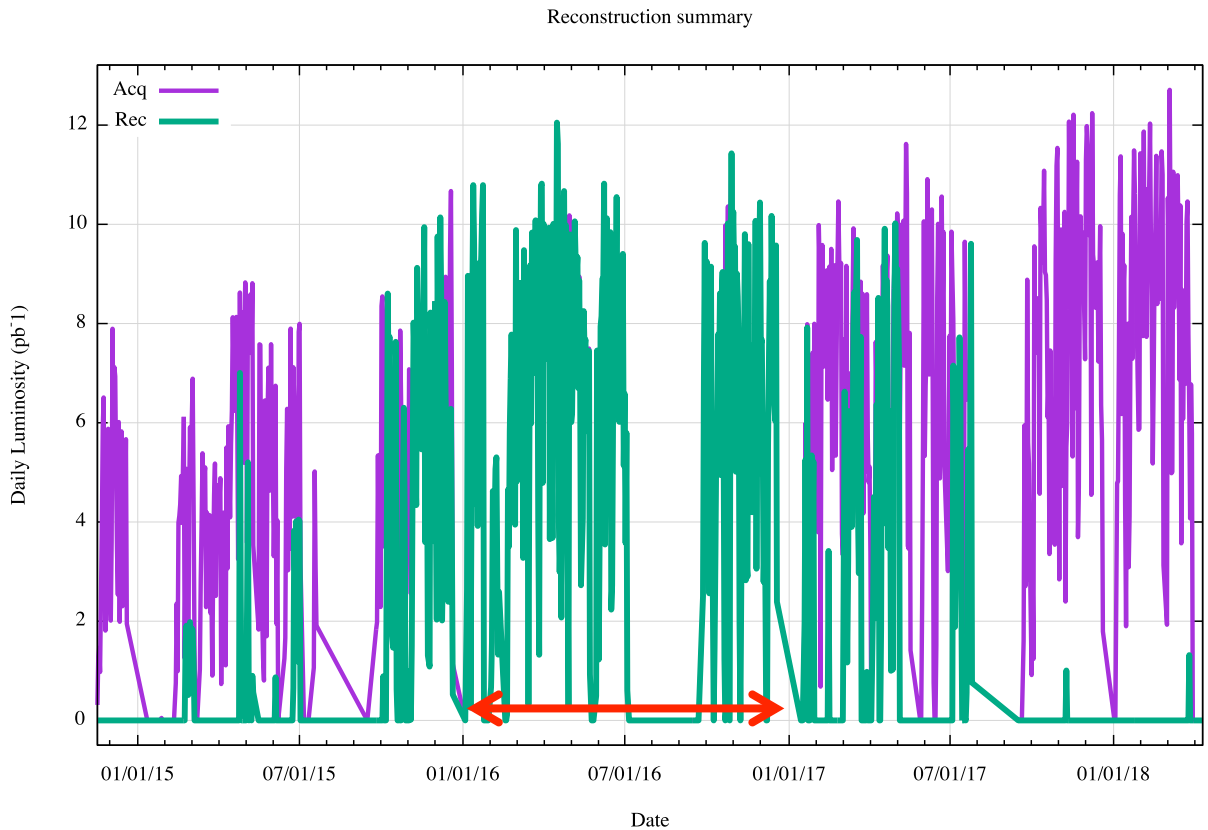
DBV-38

Reconstruction of 2016 data completed

- Started on March 21st, 2018
- End of April: bug found in CSPSMAKER (wrong energy scale factor for EMC)
- In May the reconstruction of the data between January and July 2016 has been restarted

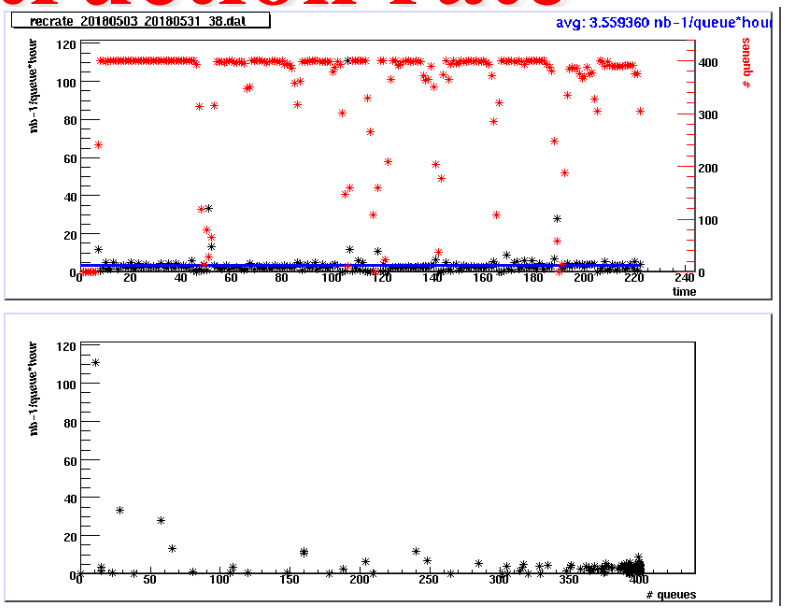
- At present: 1.5 fb^{-1} reconstructed

- We are now starting the reconstruction of 2017 data

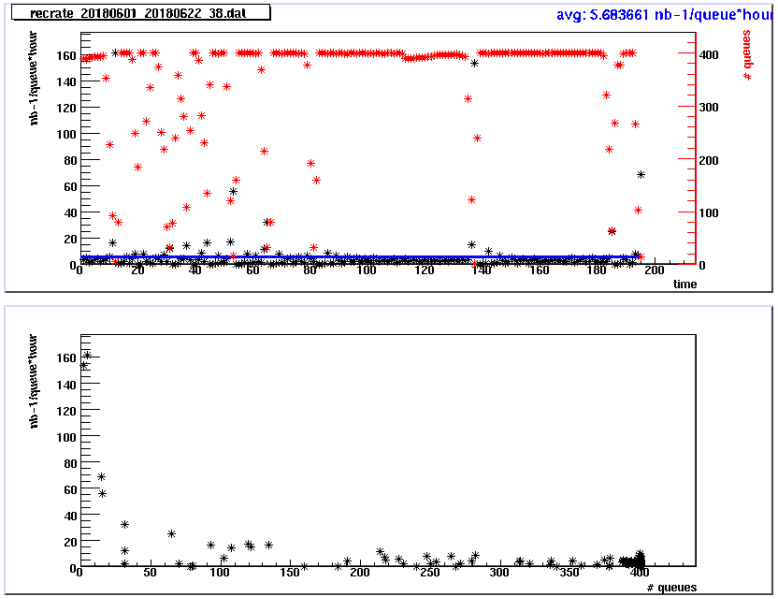


DBV-38 reconstruction rate

- **May 2018:**
  ~ 400 queues used
  ⇒ 21 pb^{-1}/day
  ⇒ 2.2 nb^{-1}/queue/h

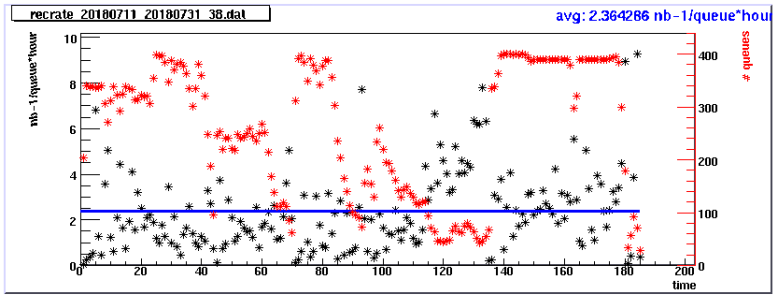


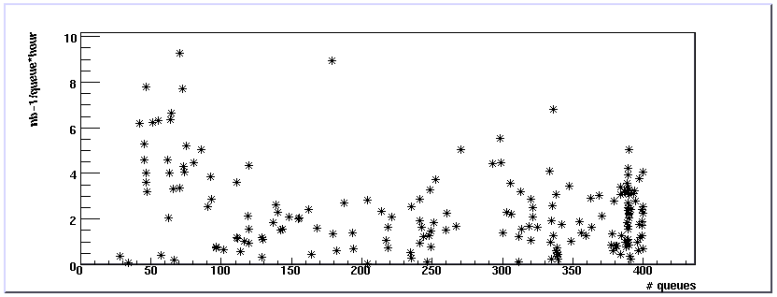
- **June 2018:**
  ~ 400 queues used
  ⇒ 20 pb^{-1}/day
  ⇒ 2.0 nb^{-1}/queue/h



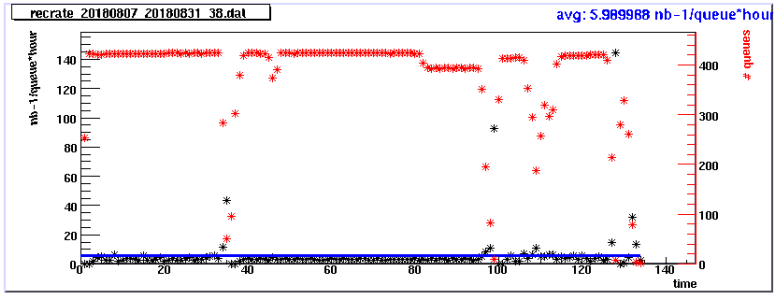
DBV-38 reconstruction rate

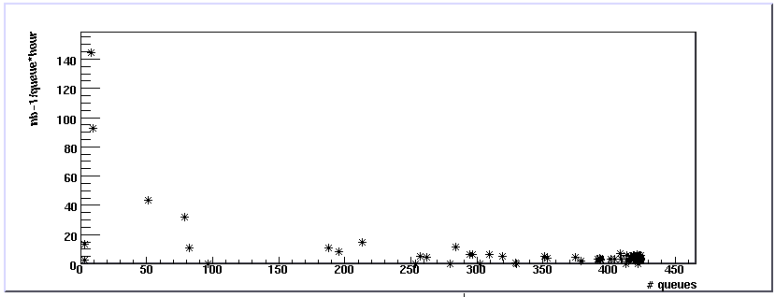
- **July 2018:**
  Problems with the Load_leveler
  (jobs not sumbitted also if the queues are free)
  Considering last part (27/7 → 31/7)
  ~ 400 queues used
  ⇒ 22 pb^{-1}/day
  ⇒ 2.3 nb^{-1}/queue/h





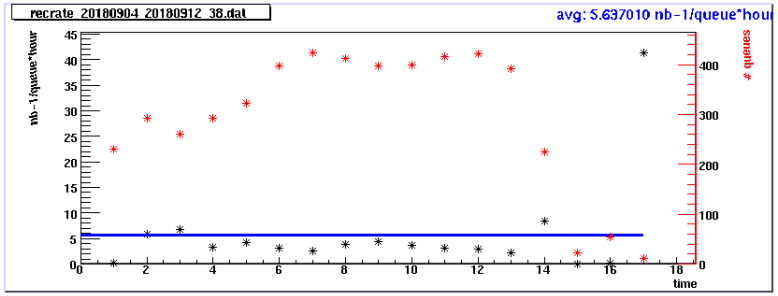
- **August 2018:**
  ~ 420 queues used
  First part (7/8 → 15/8)
  ⇒ 31 pb^{-1}/day
  ⇒3.1 nb^{-1}/queue/h
  Last part (17/8 → 31/8)
  ⇒ 40 pb^{-1}/day
  ⇒ 4.0 nb^{-1}/queue/h

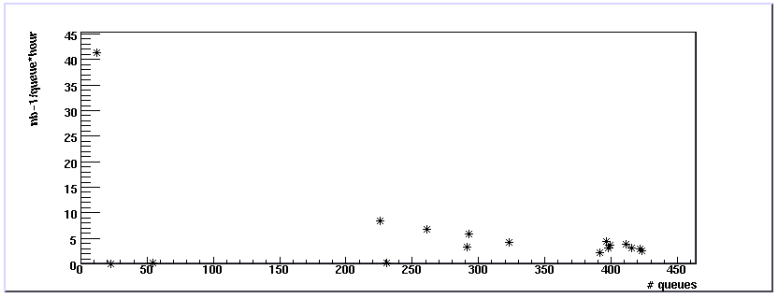




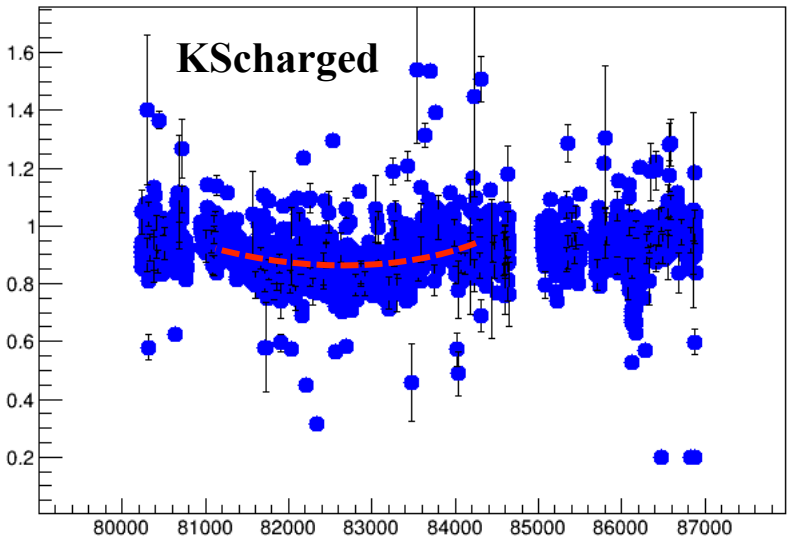
DBV-38 reconstruction rate

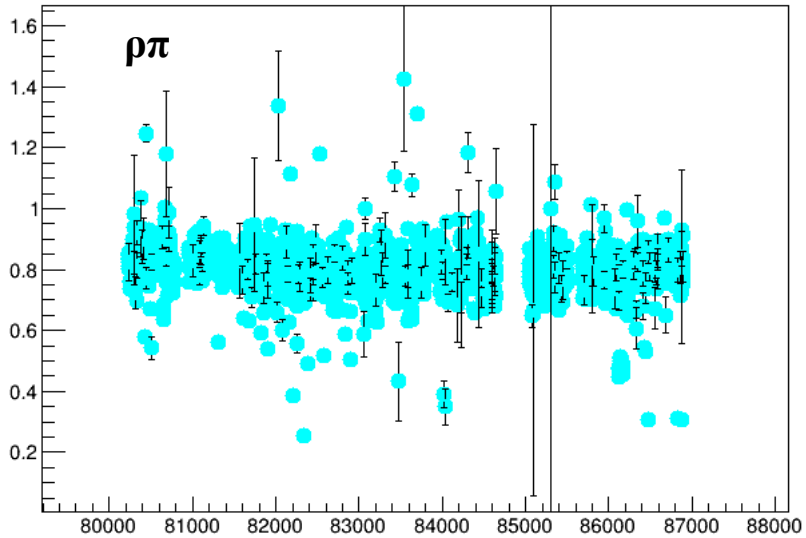
- **September 2018:**
  ~ 400 queues used
  Recovery of the error and missing runs of the 2016 sample
  ⇒ 29 pb^{-1}/day
  ⇒ 3.0 nb^{-1}/queue/h

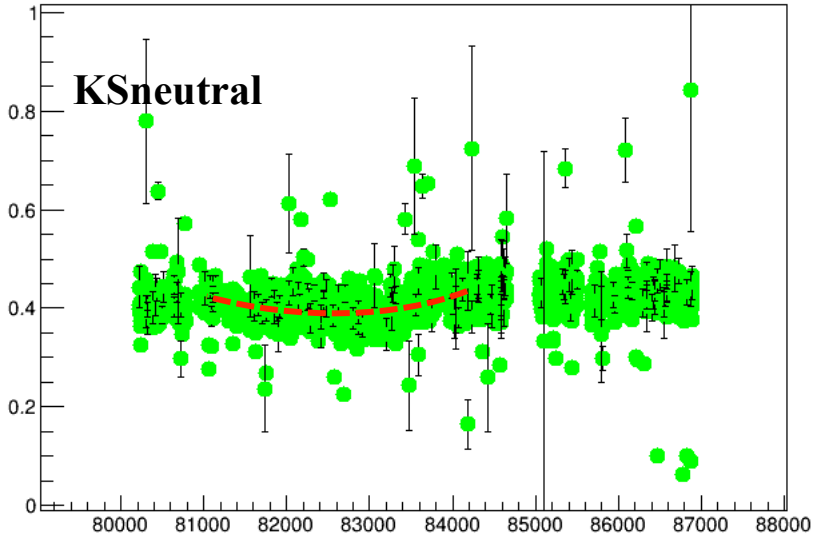


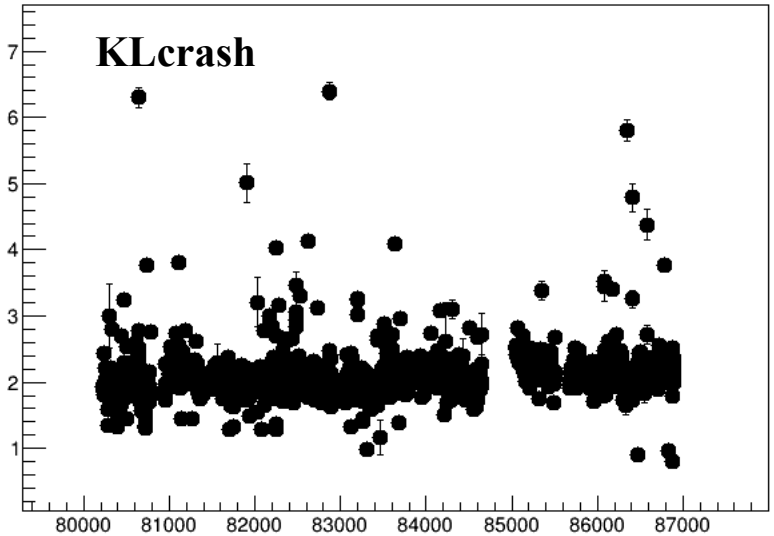


Counters DBV-38



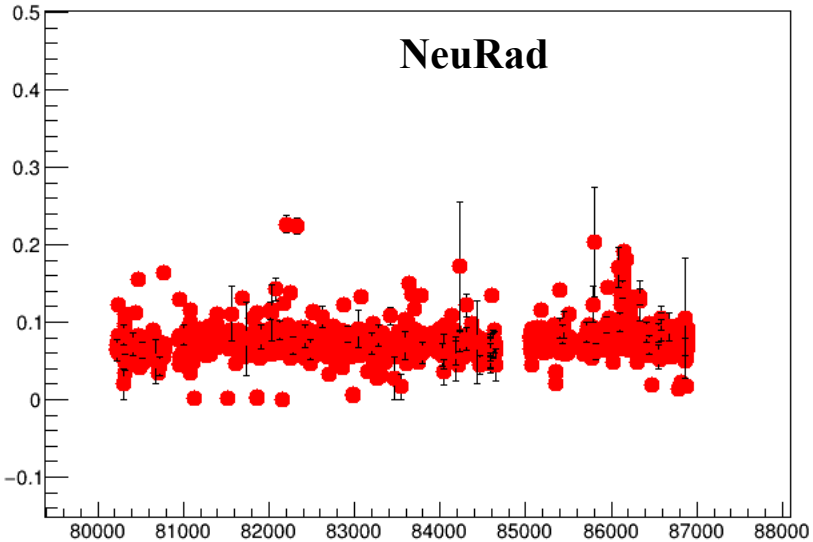


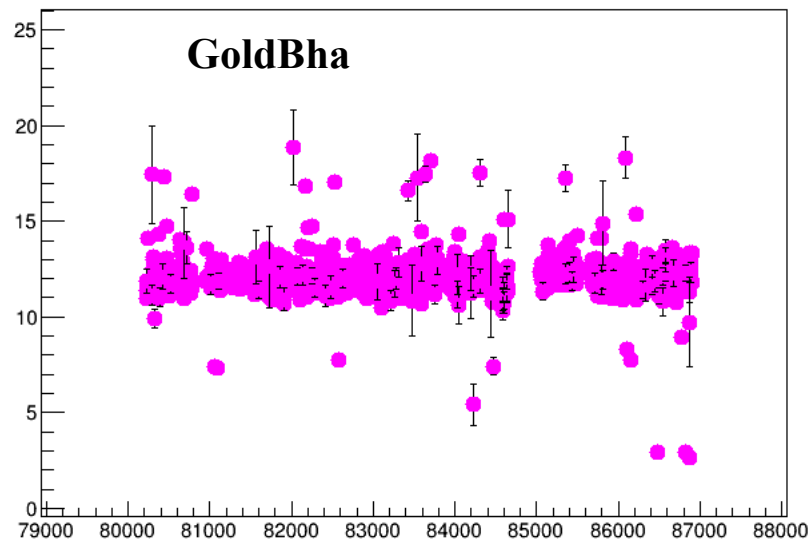


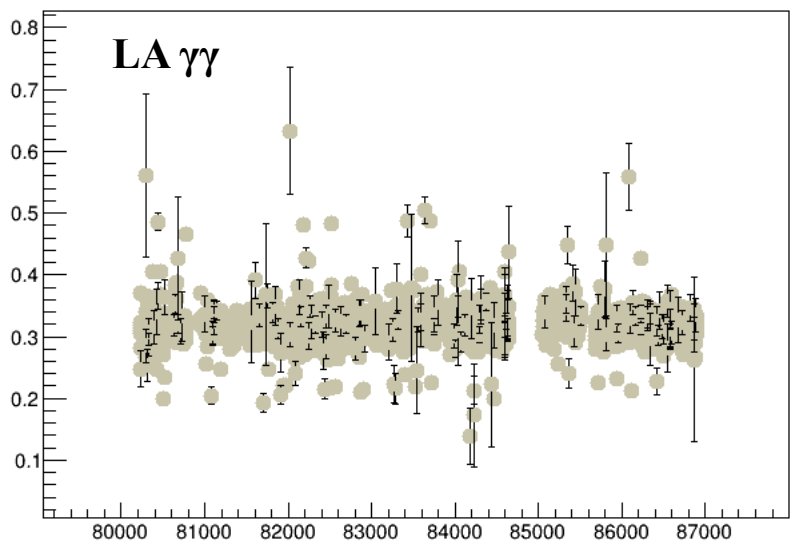


(normalized to VLAB)

Counters DBV-38







(normalized to VLAB)

Crashes and errors

- **Most frequent fatal errors**

Error type		Error message
crash	0.9 %	too many dead wires (in the DC) (> 1300 dead wires)
crash	0.6 %	file(s) LOST

- **Minor errors: some datarec file missing (generally kpm, rpi, lsb)**

Very Short Run (tipically less than 1 min)	2.0 %
Last files missing	2.0 %
DC tripping , standby, off	0.8 %
Beams lost during run	< 0.1 %

(percentage is in number of runs, not in luminosity)

Datarec perspectives and plans

- 2016 data: 1.5 fb^{-1} reconstructed from April 28 to September 12
- 168 days ⇒ 8.8 pb^{-1}/day (includes dead times, KID problems, calibration problems, etc.)
- "Duty cycle" ≈ 44 %
- If we continue in this way we'll need 14 months to complete the first round of reconstruction (without production of the ROOT files)
  ⇒ end of November 2019

- We must improve the efficiency, reducing the dead times, the goal is to complete the first round in the first half of 2019

- ROOT output implementation: start tests as soon as we finally choose the format
  Do we want to apply at this level the T0 corrections ?

- Second round of reconstruction: start soon after the end of the first one with an improved version of Datarec

Towards DBV-39 or 40

- New version of the Integrated Tracking fully tested and working is needed
- A HET output stream (if HET group needs it)
- In principle there is still room for background reduction (in FILFO)
- Other detectors (QCALT, CCALT) ?
- Tests on all these items should start at last on next January

The new Datarec must be ready before June 2019

MC status

MC DBV-38:

- **Geometry of the new detectors implemented**
- **Response of the new detectors implemented but not validated**
- **Inner Tracker simulation not implemented**

- **Production of 2016 data (All_phys) with LSF = 1 started on last April 20th**

- **117 queues used**
- **Production rate ≈ ≈ 15 pb^{-1}/day**

- **"MC efficiency" = = 94 %**

Run range	Rec (pb-1)	MC (pb-1)
88966-89096	53.21	49.46
85681-86024	95.89	91.55
86025-86511	104.96	96.97
86512-86900	110.44	104.67
80694-81537	98.24	92.77
81538-81973	102.05	93.64
81974-82398	119.13	113.2
82411-82850	111.55	103.09
82851-83145	89.55	84.46
83201-83498	102.82	100.86
TOTAL	**987.84**	**930.67**

Errors in MC production

- **Error 214 (run does not appear to have been entirely reconstructed) ≈ 3.5 %**

  This is due to runs with some Bhabha file missing in the reconstruction
  It is not a large fraction, but for example if we have a run with 60 raw files, also if only one is missing we skip the simulation and we completely loose that run (~ 1 pb^{-1})

- **Error 210 (Unable to obtain .mco filename from logged geanfi output) < 0.1 %**

MC – luminosity studies

- LAB events used for the evaluation of the tracking efficiency:
  - 2 clusters with cos α > − 0.9 (acollinearity)
  - One track (tag) with p > 400 MeV associated to one LAB cluster
  - Second LAB cluster with E > 400 MeV
  - Look for a second track from the IP, with p > 400 MeV and opposite charge w.r.t. the tag one, and with d > 50 cm
    (d is the distance from the first hits of tag and candidate tracks)
- MC samples generated with card = "goldbha" and LSF = 0.01
- No Integrated Tracking used for data (tracks from DTF1, if DBV-38)

Run #	Acq. date	DBV	ε (data)	ε (MC)
85681 – 86024	Nov. 2016	38	98.2 %	98.4 %
86583 – 86735	Dec. 2016	37	98.4 %	98.4 %
88966 – 89096	Apr. 2017	38	92.8 %	95.7 %
89080 – 89163	Apr. 2017	37	92.8 %	95.8 %
90885 – 91035	Jul. 2017	37	92.1 %	96.0 %

Offline shifts

- Offline shifts from April to end of July
- New shift schedule, until end of the year, in preparation

- Shifter's duties:

- Follow the reconstruction, check that the jobs are running
- When needed, start the reconstruction of a new bunch of data
- Check day by day the errors: run the scripts to find them and report about each single error by looking at the log files
  ⇒ put an entry in the Google sheet (see after)
- Check the reconstruction rate
- Check the data counters
- Be aware of the reconstructed luminosity
- Perform checks on Data Quality (see after)
- For any problem, contact Elena and me

- Be present and report at the weekly meeting (also from remote)

Error file

DatarecErrors

File Modifica Visualizza Inserisci Formato Dati Strumenti Componenti aggiuntivi Guida L'ultima modifica è stata apportata ieri alle 19:36

Run #

Run #	Files with errors		Error type	Error message	Acq. Date
80337	51		crashed	too many dead wires, > 1800	1/10/2016
80340	36		crashed	too many dead wires, > 1800	
80341	24		crashed	too many dead wires, > 1800	
80342	9		crashed	too many dead wires, > 1800	
80343	9		crashed	too many dead wires, > 1800	
80346	12		crashed	too many dead wires, > 1800	
80347	51		crashed	too many dead wires, > 1800	
80349	6		crashed	too many dead wires, > 1800	
80350	51		crashed	too many dead wires, > 1800	
80351	60		crashed	too many dead wires, > 1800	
80352	30		crashed	too many dead wires, > 1800	
80353	9		crashed	too many dead wires, > 1800	
80354	3		crashed	too many dead wires, > 1800	
80355	48		crashed	too many dead wires, > 1800	
80356	36		crashed	too many dead wires, > 1800	
80357	42		crashed	too many dead wires, > 1800	
80358	51		crashed	too many dead wires, > 1800	1/11/2016
80359	51		crashed	too many dead wires, > 1800	
80360	33		crashed	too many dead wires, > 1800	
80361	48		crashed	too many dead wires, > 1800	
80362	27		crashed	too many dead wires, > 1800	
80363	18		crashed	too many dead wires, > 1800	
80503	1	/48	crashed	raw080503N_ALL_f06_1_1_1.004 file unavailable --> LOST	1/16/2016
80544	1	/9	crashed	raw080544N_ALL_f06_1_1_1.000 file unavailable --> LOST	1/17/2016
80672	15		datarec full		1/20/2016
80673	9		datarec full		
80675	11		datarec full		
80676	42		datarec full		
80677	6		datarec full		1/21/2016
80678	3		datarec full		
80679	35		datarec full		
80681	23		datarec full		
80683	3		datarec full		
80684	13		datarec full		
80685	34		datarec full		
80686	6		datarec full		
80687	36		datarec full		
80688	39		datarec full		
80689	12		datarec full		
80691	18		datarec full		
80692	9		datarec full		
80693	8		datarec full		
81687	1	/18	crashed	raw081687N_ALL_f06_1_1_1.000 file unavailable --> LOST	3/5/2016
82058	3	/18	crashed	too many dead wires, > 2000	3/18/2016
82059	3	/6	crashed	too many dead wires, > 2000	
82297	1	/45	crashed	raw082297N_ALL_f07_1_1_1.008 file unavailable --> LOST	3/26/2016
82450	8	/48	crashed	raw files unavailables, but the files are in the recall areas	
82452	5	/27	crashed	raw files unavailables, but the files are in the recall areas	
82455	11	/48	crashed	raw files unavailables, but the files are in the recall areas	
82456	5	/30	crashed	raw files unavailables, but the files are in the recall areas	

Fatal errors (produce a crash of datarec)

Error file

DatarecErrors

File Modifica Visualizza Inserisci Formato Dati Strumenti Componenti aggiuntivi Guida L'ultima modifica è stata apportata

	A	B	C	D	E
1			Missing		VSR = very short run
2					
3	80227		kpm 000	DC standby	
4	80238		kpm 013, 014, 015	DC tripping	
5	80239			VSR	
6	80280		rpi last files		
7	80283		rpi last file		
8	80299			VSR	
9	80311		rpi last files		
10	80380			VSR	
11	80391			VSR	
12	80405		kpm 000	DC off	
13	80411		kpm 000	DC standby	
14	80412		kpm 000	DC standby	
15	80476			VSR	
16	80479			VSR	
17	80962		kpm 000	DC standby	
18	81008			VSR	
19	81015		rpi last files		
20	81023		kpm last files		
21	81040		rpi last files		
22	81098			VSR	
23	81195		kpm last files		
24	81297			no beams	
25	81333		kpm last files		
26	81344			VSR	
27	81499			VSR	
28	81563			VSR	
29	81571		kpm last files		
30	81573			VSR	
31	81662		kpm last files		
32	81734			VSR	
33	81741		kpm last files		
34	81748		kpm 000	DC off	
35	81820			VSR	
36	81828		kpm last files		
37	81893		kpm last file		
38	81933		kpm last file		
39	81938		kpm last file		
40	82057		kpm last file		
41	82121		kpm last file		
42	82130		kpm last file		
43	82176			VSR	
44	82179			VSR	
45	82197			no beams	
46	82213			VSR	
47	82217			VSR	
48	82226			VSR	
49	82309		kpm last file		
50	82332		rpi last files		
51	82341		kpm last file		
52	82418			VSR	

Minor errors (some datarec files missing)

Important to know, can have an impact on MC production if they are bha files

Data Quality

- We have now a large sample of reconstructed data, and a sizeable sample of MC production for the same period

- We strongly need Data Quality studies as well as Data-MC comparisons

- We need to validate the MC simulation looking at tracks, clusters, invariant masses and any other interesting information.
  We cannot produce any paper on new data without a reliable MC

- Every member of the Collaboration, in particular (but not only) people doing analysis, should contribute to this very important task

- Offline shifters can do also this during their week, but people doing analysis must make available scripts, macros, and code to be used for this task

- If we don't do that, we'll reconstruct 5 fb^{-1} of data, but they will be useless

Conclusions

- **2016 data reconstruction completed (1.5 fb^{-1})**
- **Starting reconstruction of 2017 data**
- **We plan to complete the whole KLOE-2 dataset reconstruction by mid of 2019**
- **Second round will start soon after, with an improved version of Datarec**
- **MC production of ~ 1 fb^{-1} of 2016 data**
- **We need to tune the MC to the run conditions of KLOE-2**
- **We really need an effort of the whole Collaboration on Data Quality and Data-MC comparison**

DBV-38 reconstruction rate

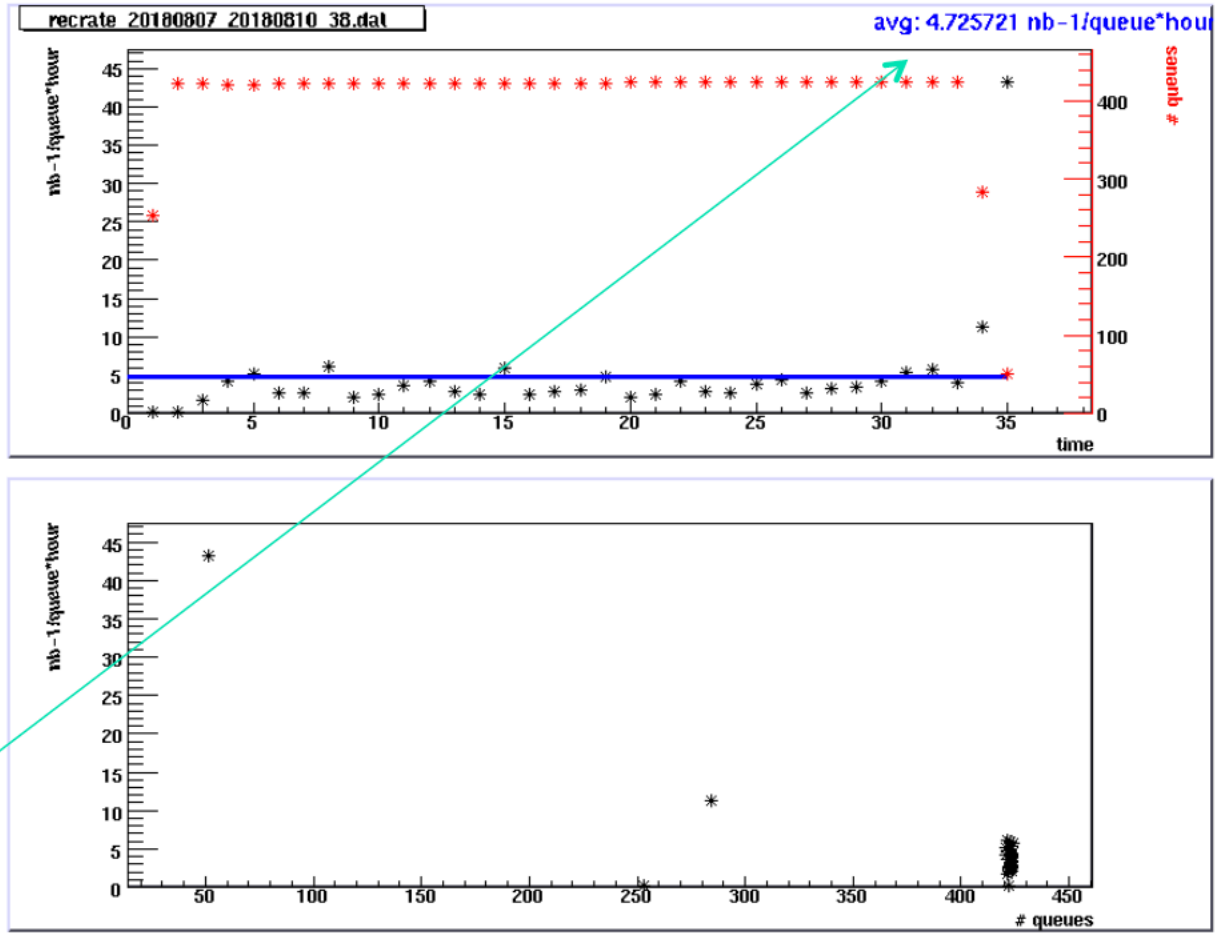
- **Example:**

  From log files:
  105 pb^{-1} in 2.8 days

  ⇒ 37 pb^{-1}/day

  ⇒ 3.7 nb^{-1}/queue/h

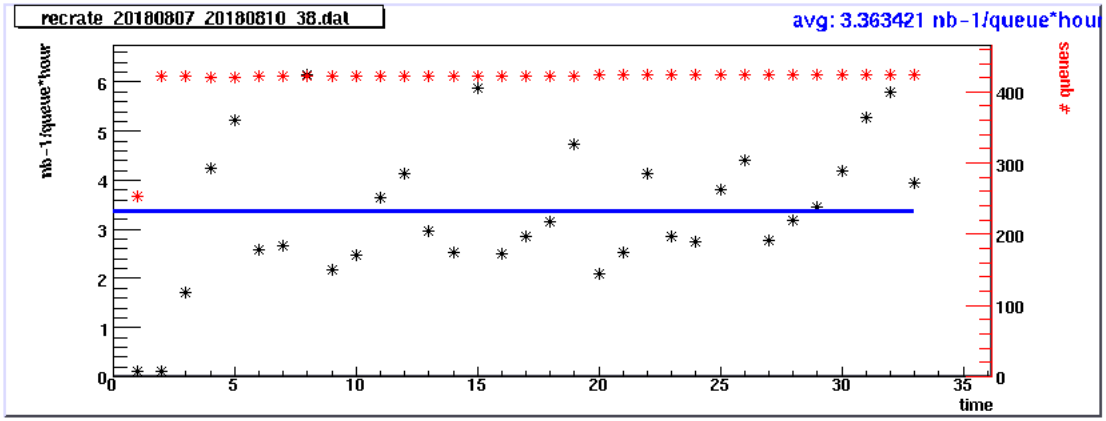
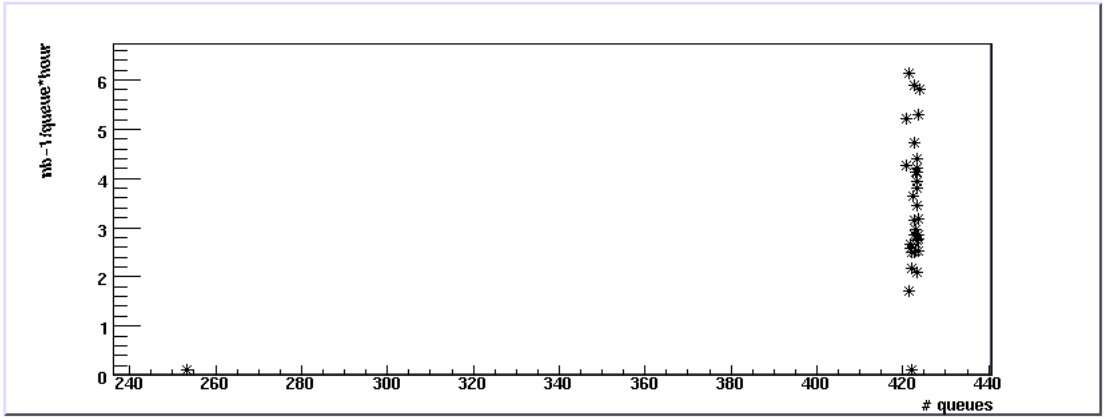
- **Average from the script: 4.7 nb^{-1}/queue/h, but this is due to last two black points**



DBV-38 reconstruction rate

- The same without last two points:

Average:

⇒ 3.4 nb^{-1}/queue/h

Consistent with 3.7