Summary of slow-control till LS1

Belle II

+

DAQ

Takuto KUNIGO

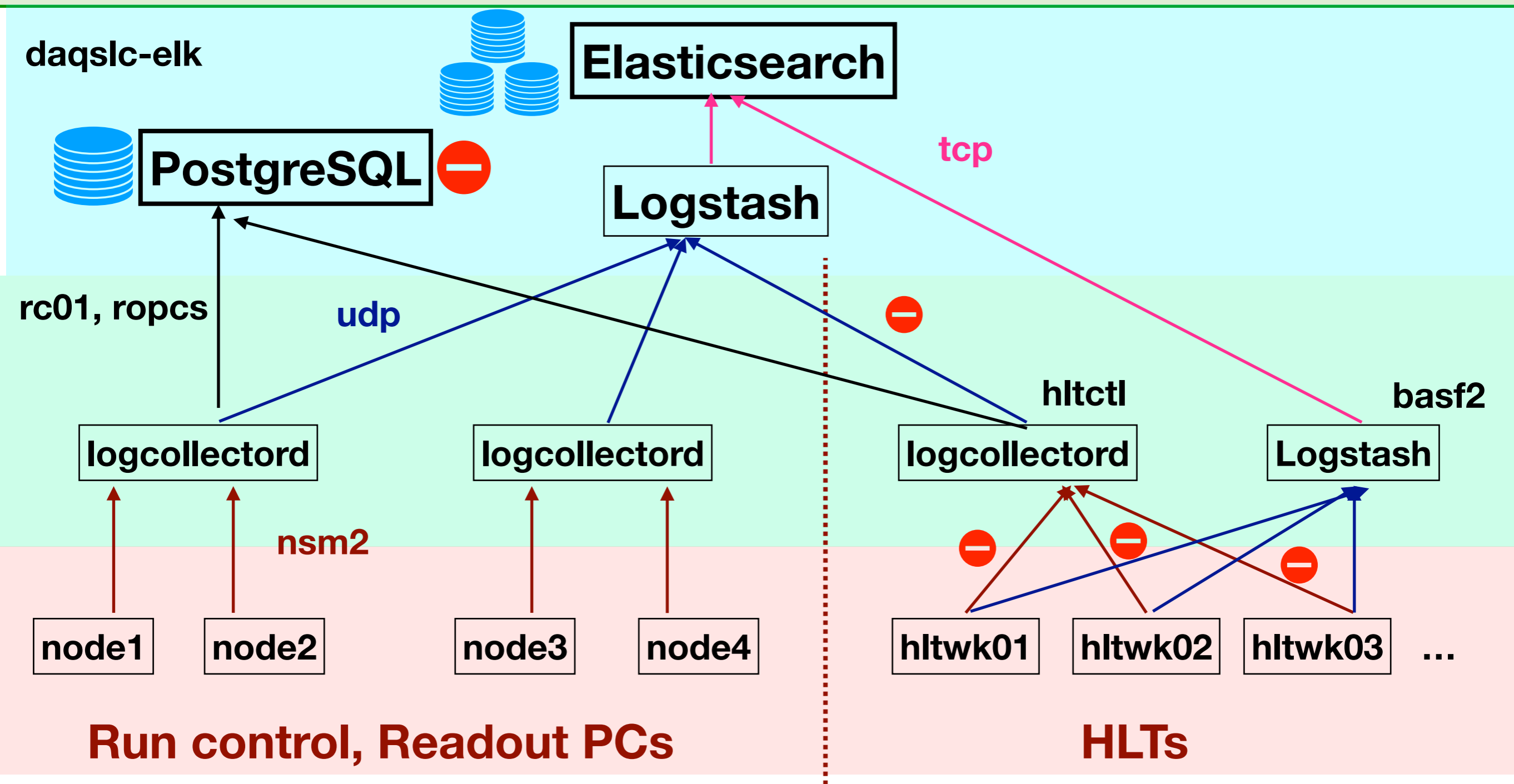December 1, 2022

Belle II TRG/DAQ Workshop 2022

KEK

Brief summary

After the slow-control group dissolved, no intense development

Remaining/on-going topics

- Development related to the PCIe40 integration
  - ‣ Discussed yesterday
- Renovation of the detector control system
  - ‣ To be discussed in the latter part of this session
- NSM2-related errors
- Software instabilities
  - ‣ Need debugging not only by a single person but in a more collaborative way
- More automatic/convenient recoveries
  - ‣ To stack up tolerable errors into a list during a run, then perform a recovery at a next run stop (e.g. due to beam dump)
- Clean-up of log-messages
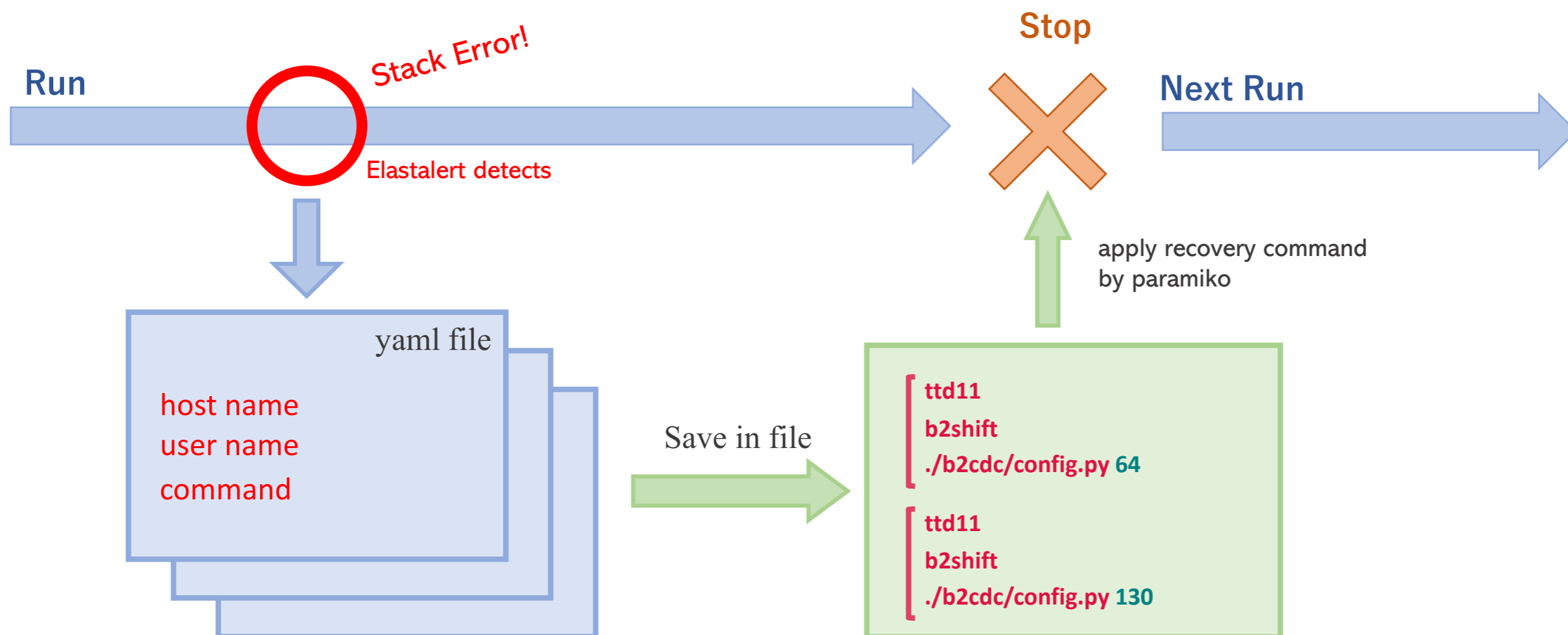 - ‣ Modifications ready & under review

Log-messages

daqslc-elk

Elasticsearch

PostgreSQL ⊖

Logstash

tcp

rc01, ropcs

udp

logcollectord

logcollectord

hltctl

logcollectord

⊖

basf2

Logstash

nsm2

node1   node2

node3   node4

⊖   ⊖   ⊖

hltwk01   hltwk02   hltwk03   …

Run control, Readout PCs   **HLTs**

+ proper category (sub-system) assignment
++ clean-up (proper severity, aggregating of a burst)

Stack up tolerable errors

By Junewoo Park

Reminder

Run

Stack Error!

Stop

Next Run

Elastalert detects

apply recovery command
by paramiko

yaml file

host name
user name
command

Save in file

ttd11
b2shift
./b2cdc/config.py 64

ttd11
b2shift
./b2cdc/config.py 130

- I try to improve the auto-mode
- Simple test was done with Kunigo-san

Our efficiencies

Period	Time based efficiency	Luminosity weighted efficiency	Good-data-quality efficiency
February 21 - June 22, 2022	87%	87%	86%
October 19 - December 23, 2021	87%	87%	87%
February 16 - July 5, 2021	90%	89%	88%
October 19 - December 18, 2020	89%	88%	87%
February 25 - July 1, 2020	84%	87%	86%
October 18 - December 12 2019	(Not monitored)	(70%)	(66%)
March 25 - June 15 2019	(Not monitored)	(67%)	(67%)

- Data-taking efficiencies reached a *plateau* (**85-90**%)
- Official target is at **90**%
- Even 1.0 % is very expensive in terms of time, money, manpower, energies…

We should achieve **higher efficiency after LS1** and
further higher after LS2

- We'll overview some visualisations to seek for further improvements
- Only for 2020ab, 2020c, 2021ab, 2021c, and 2022ab runs
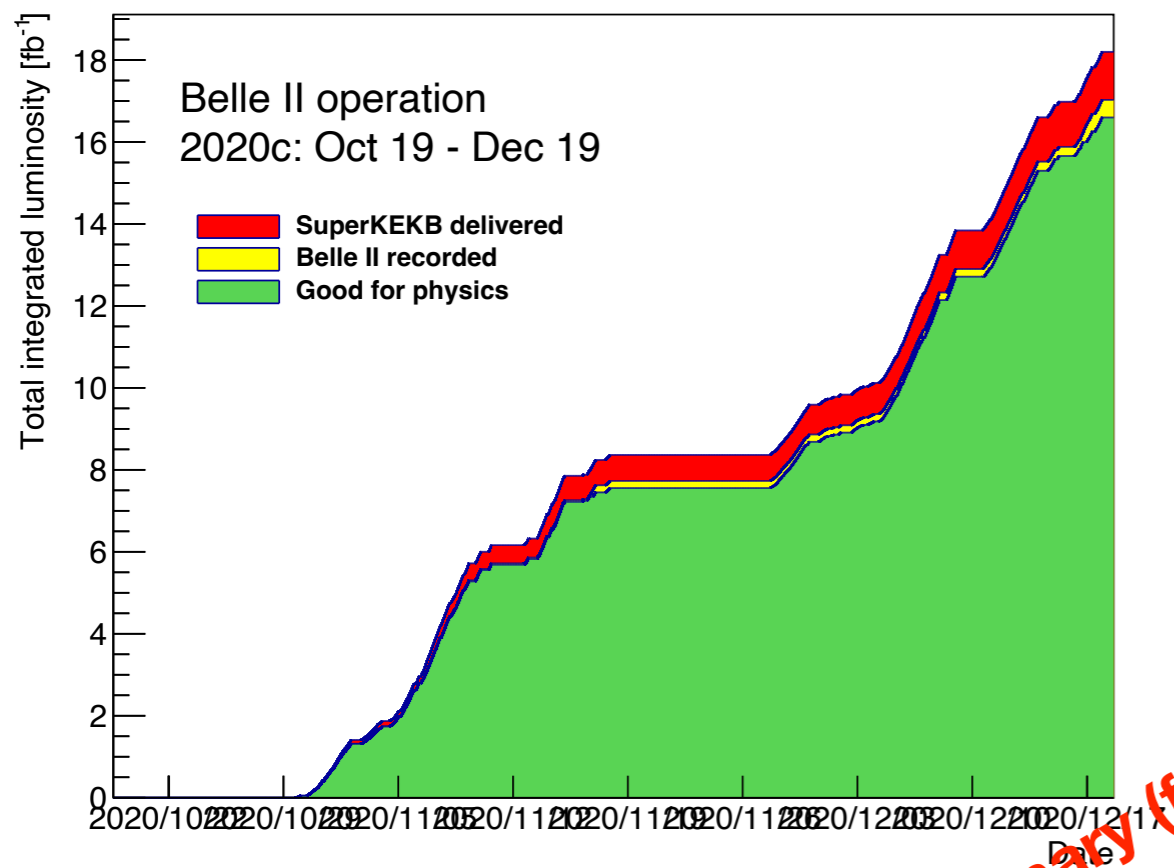- Only online information used without precise correction, thus the results differ from the offline calculation by a few percent

2020ab run

Belle II operation
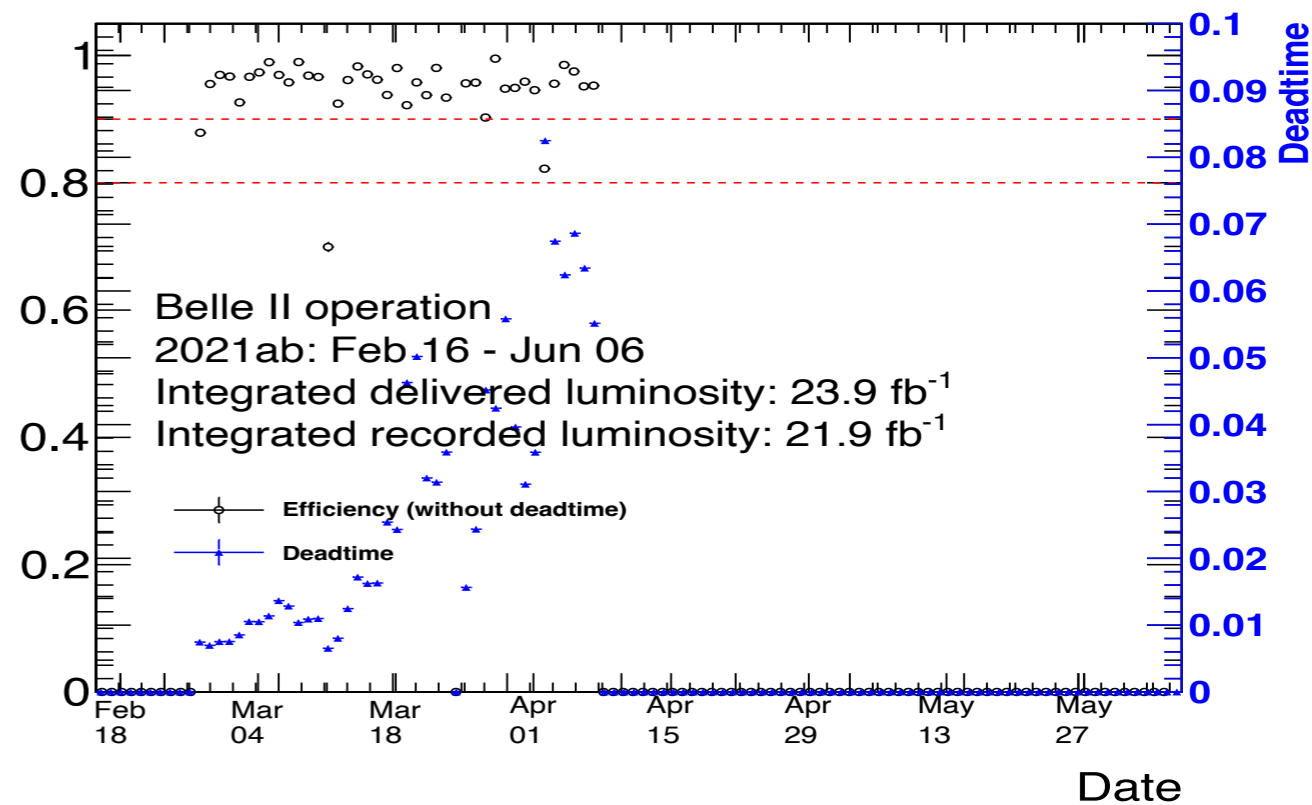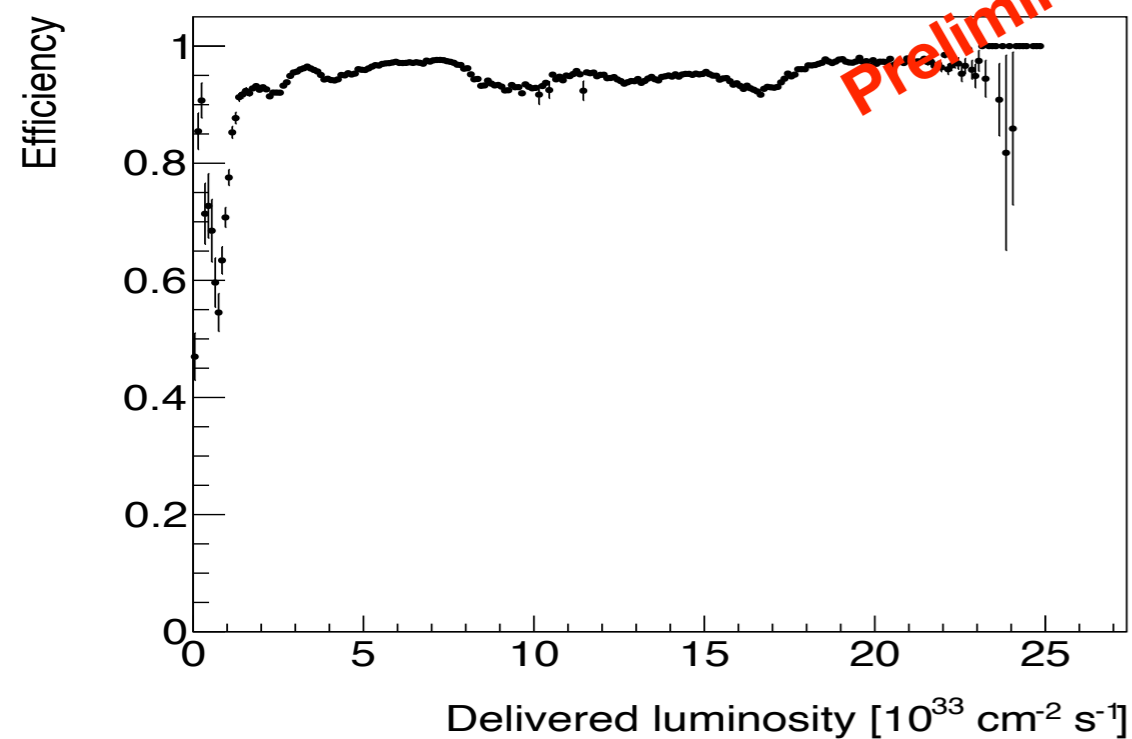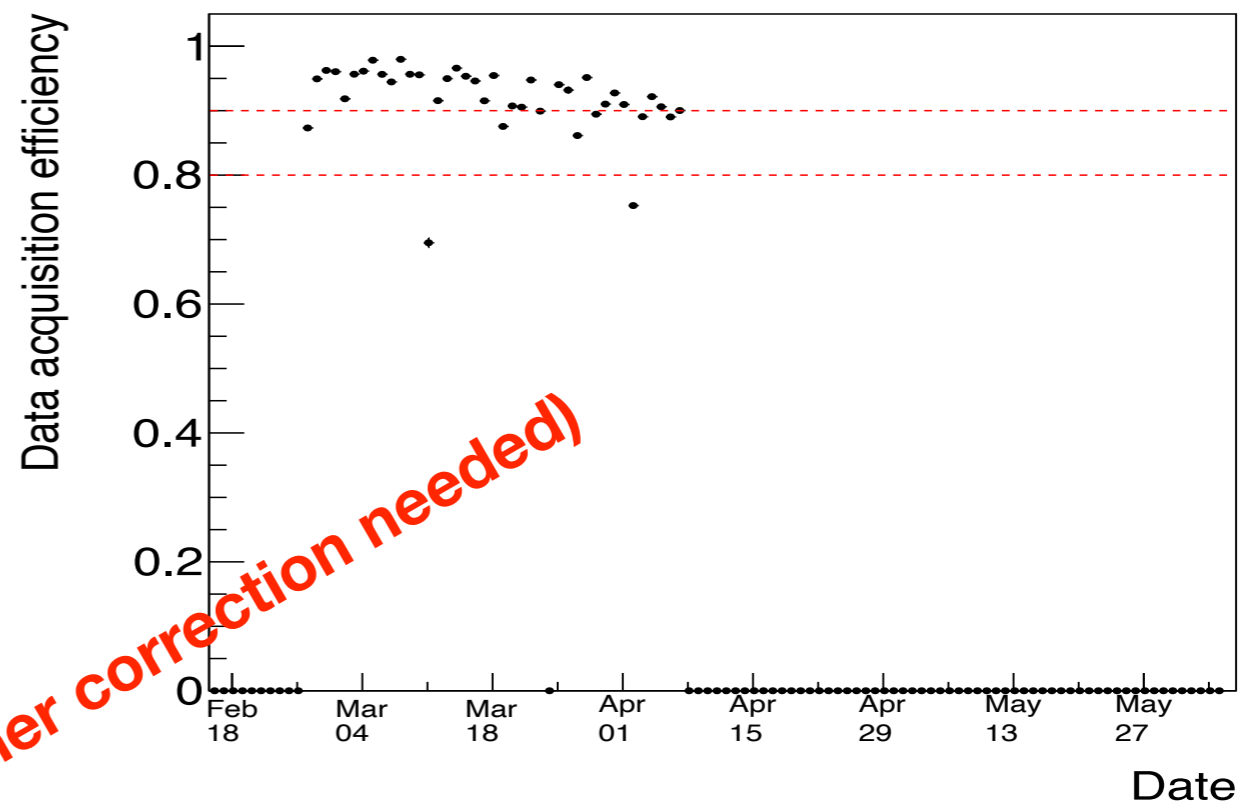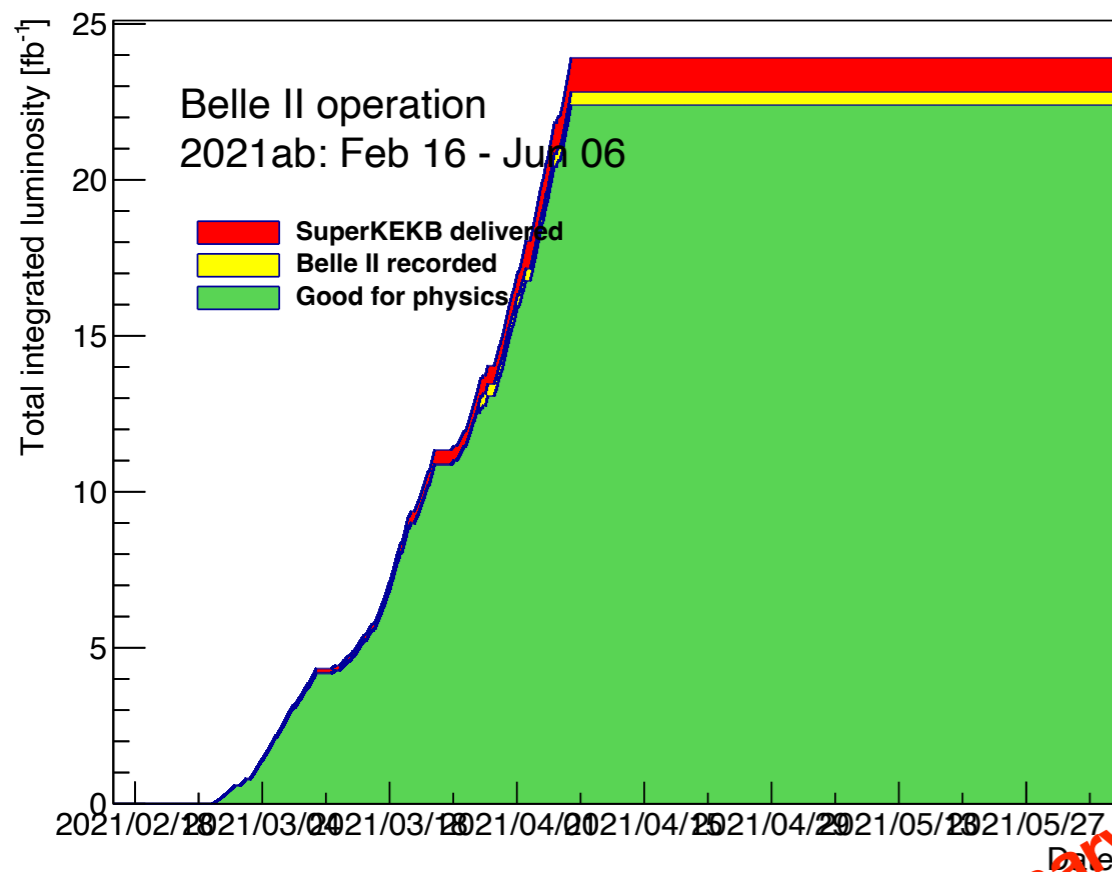2020ab: Mar 01 - Jul 03

SuperKEKB delivered
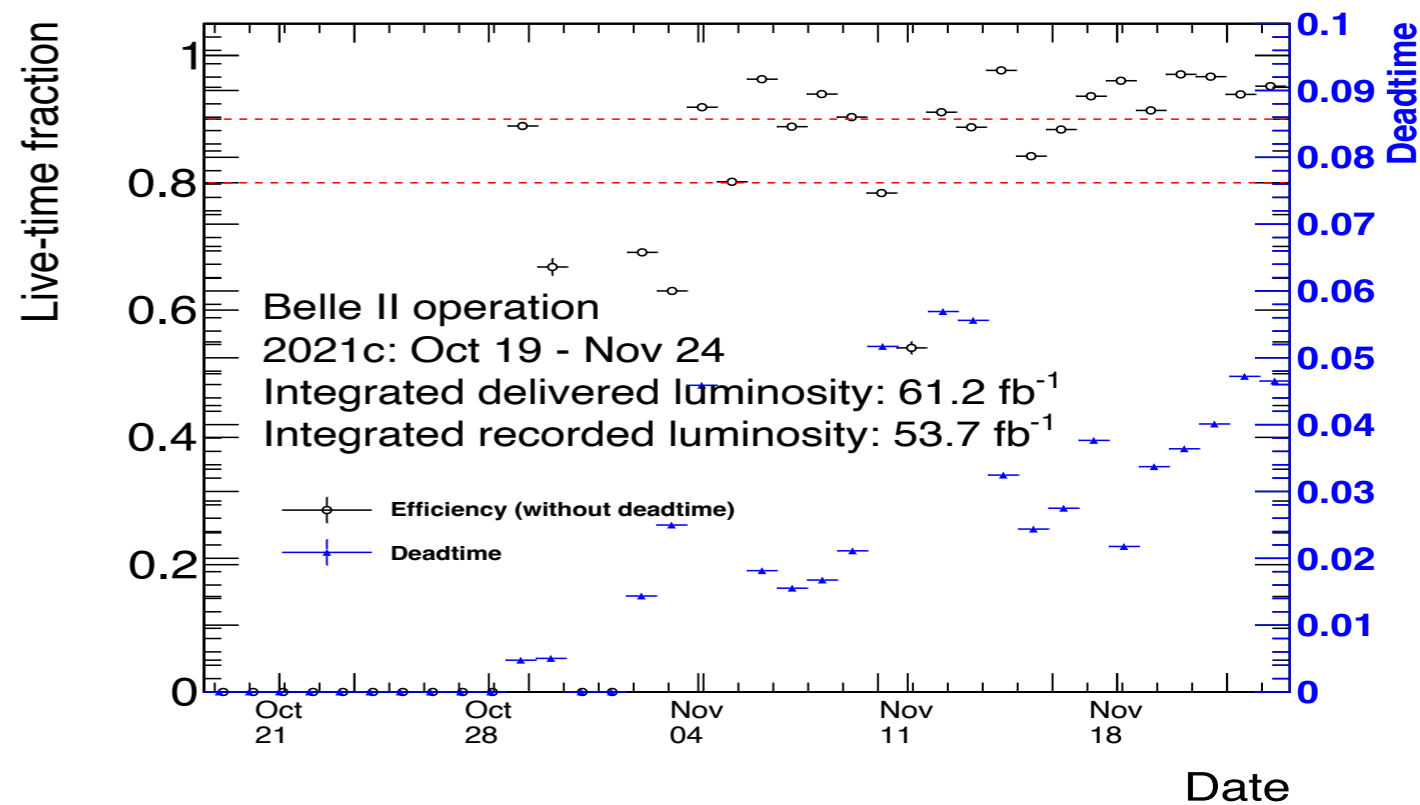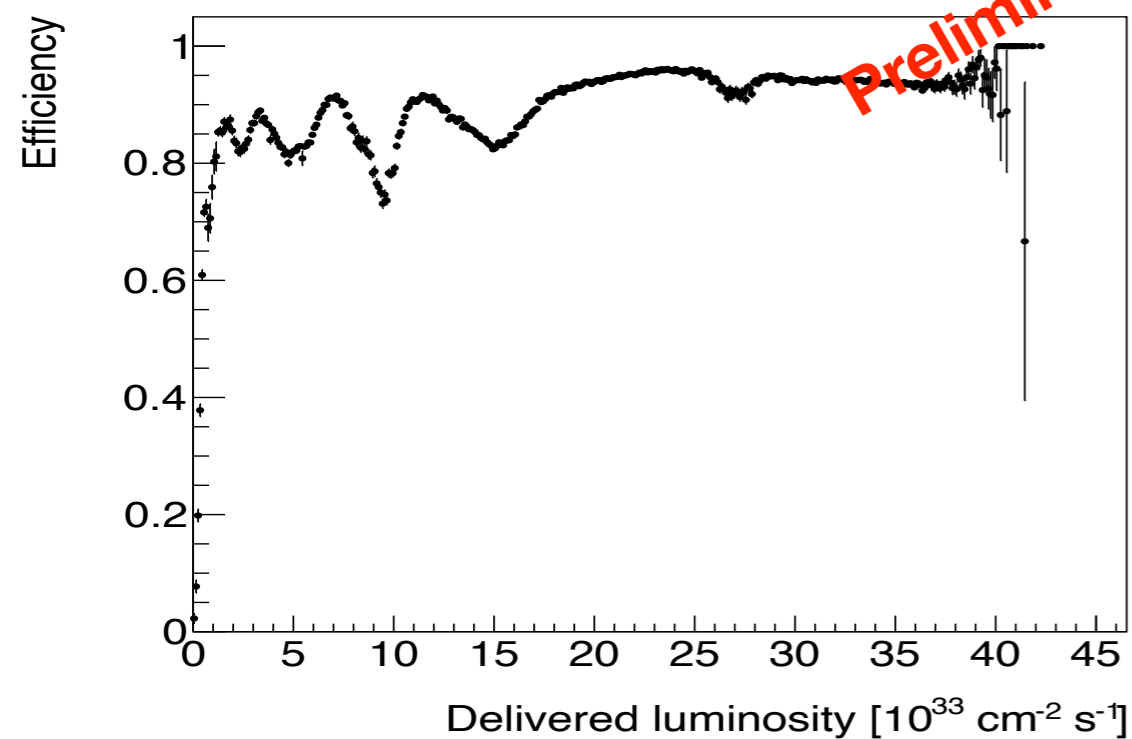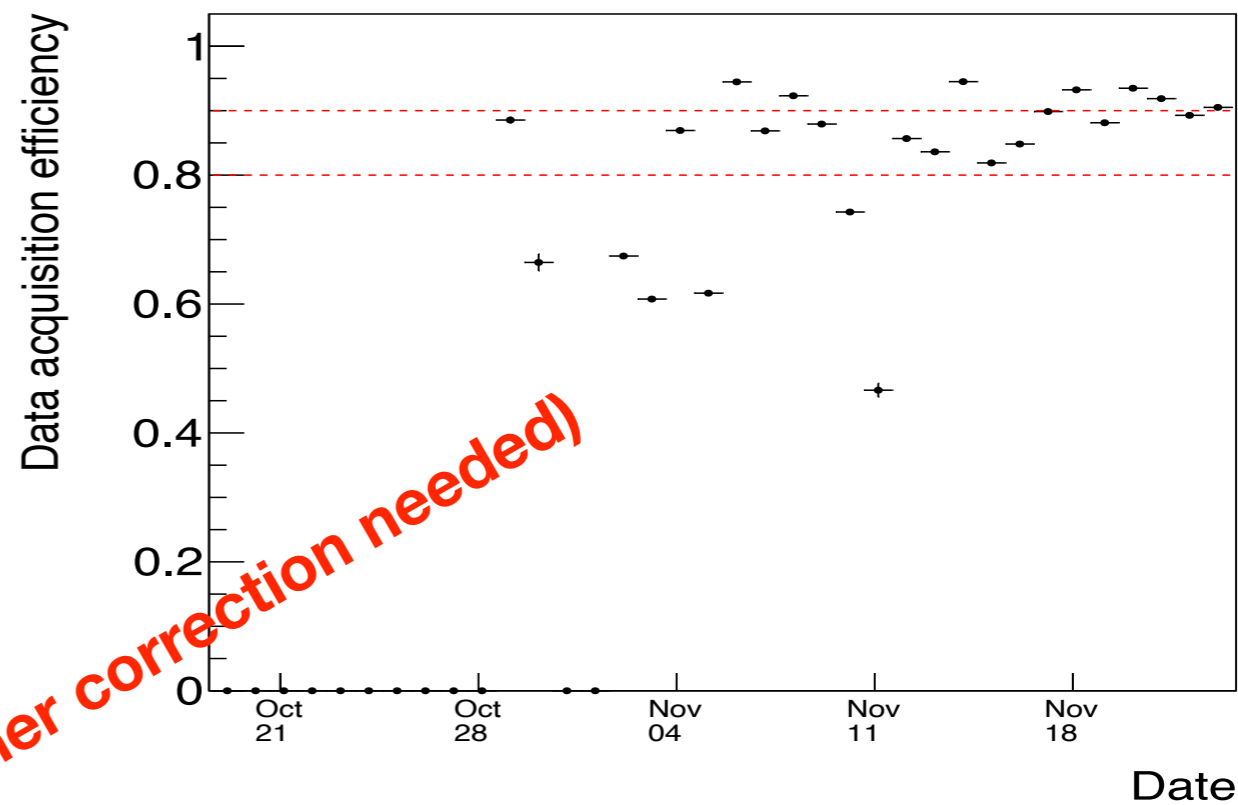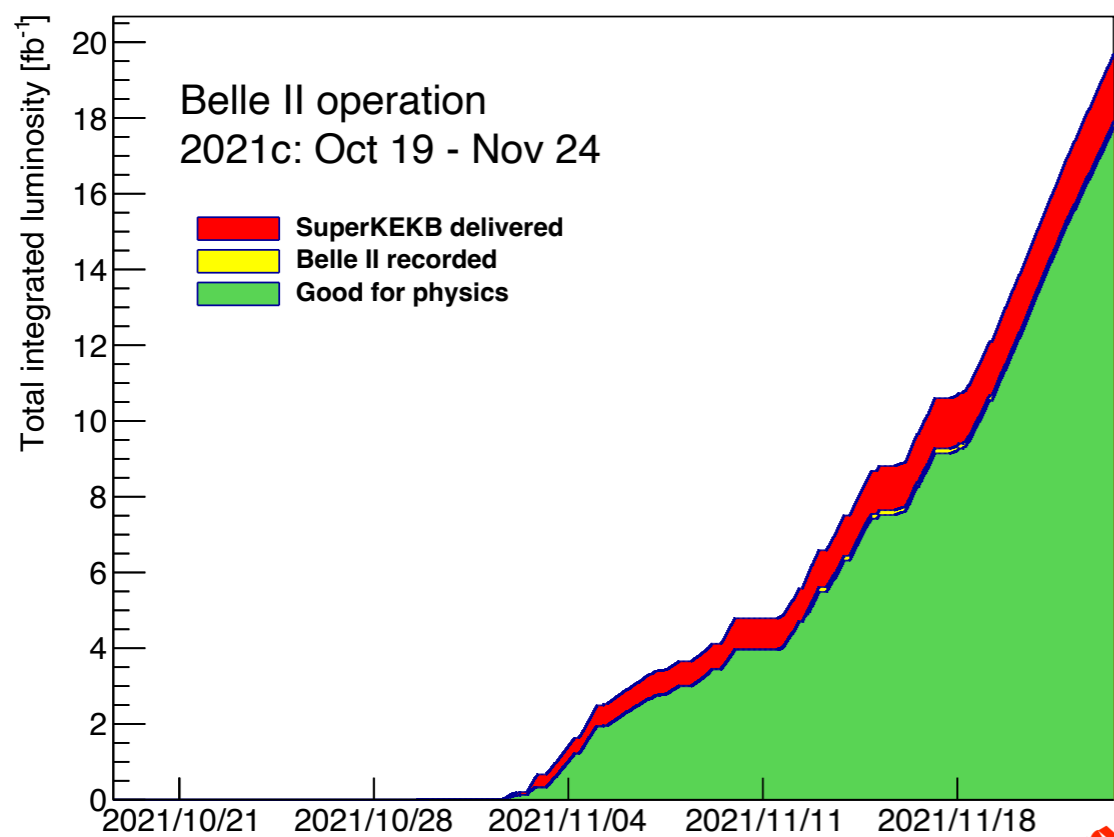Belle II recorded
Good for physics

Preliminary (further correction needed)

Belle II operation
2020ab: Feb 25 - Jul 02
Integrated delivered luminosity: 72.6 fb^{-1}
Integrated recorded luminosity: 63.1 fb^{-1}

Efficiency (without deadtime)
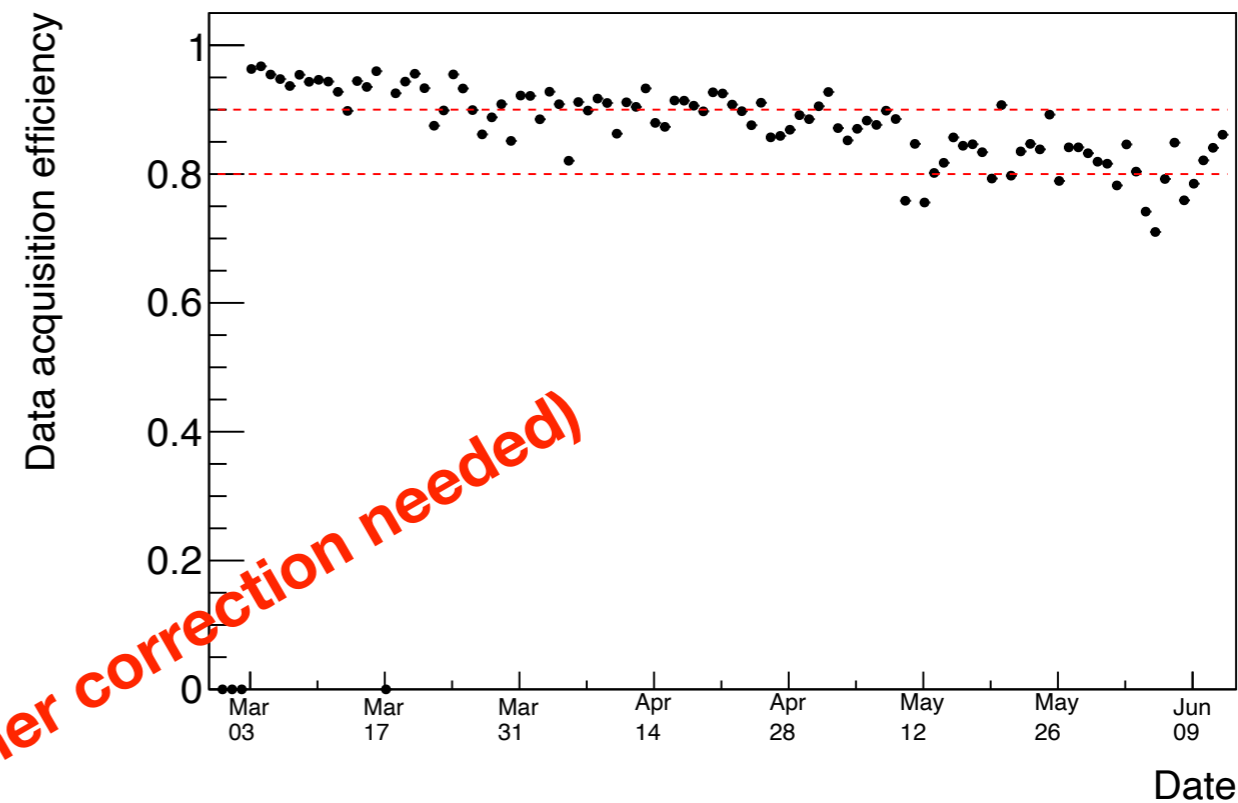Deadtime

2020c run
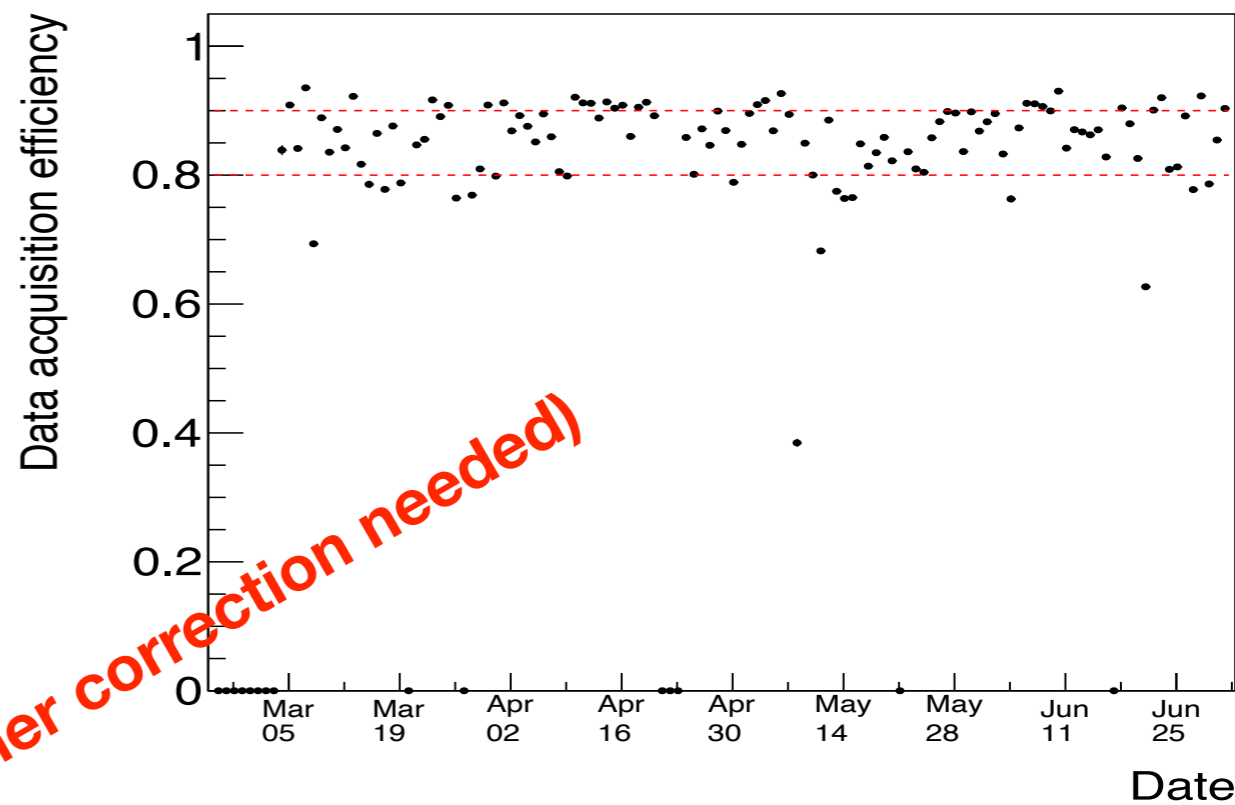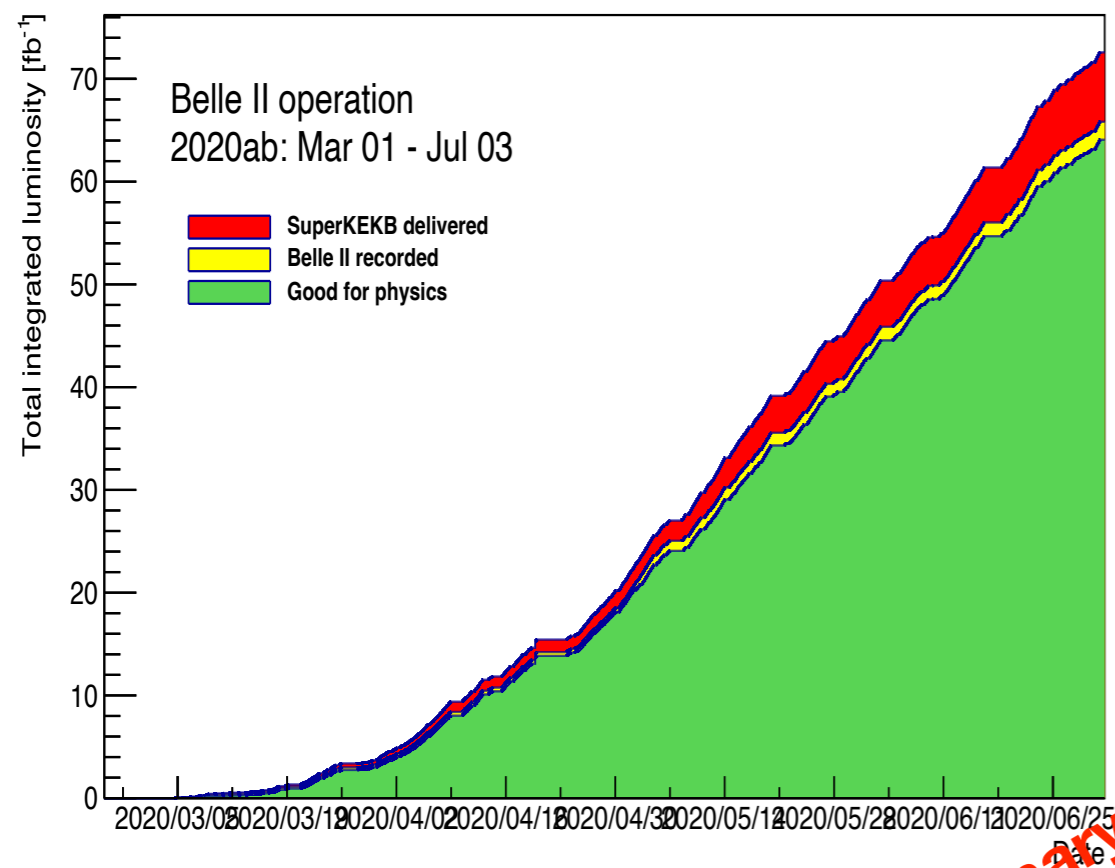
2021ab run

Preliminary (further correction needed)
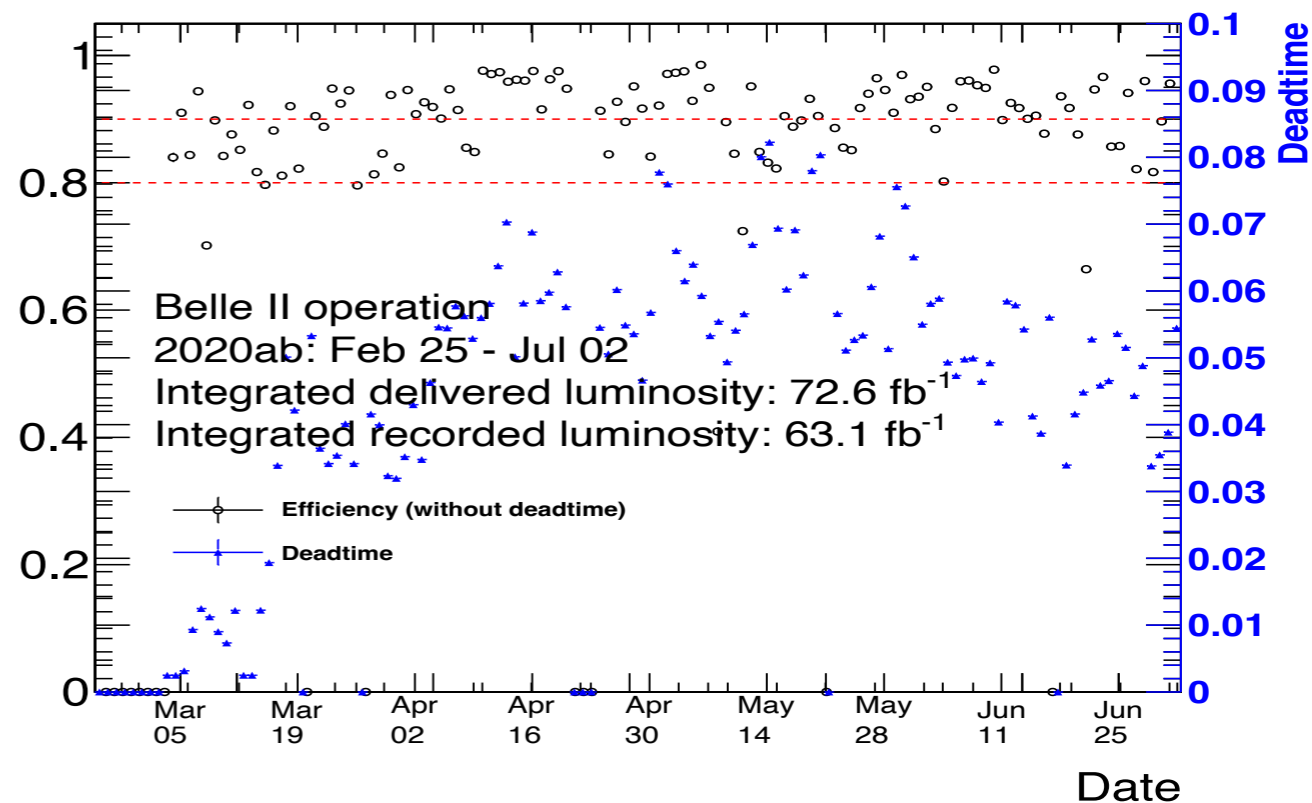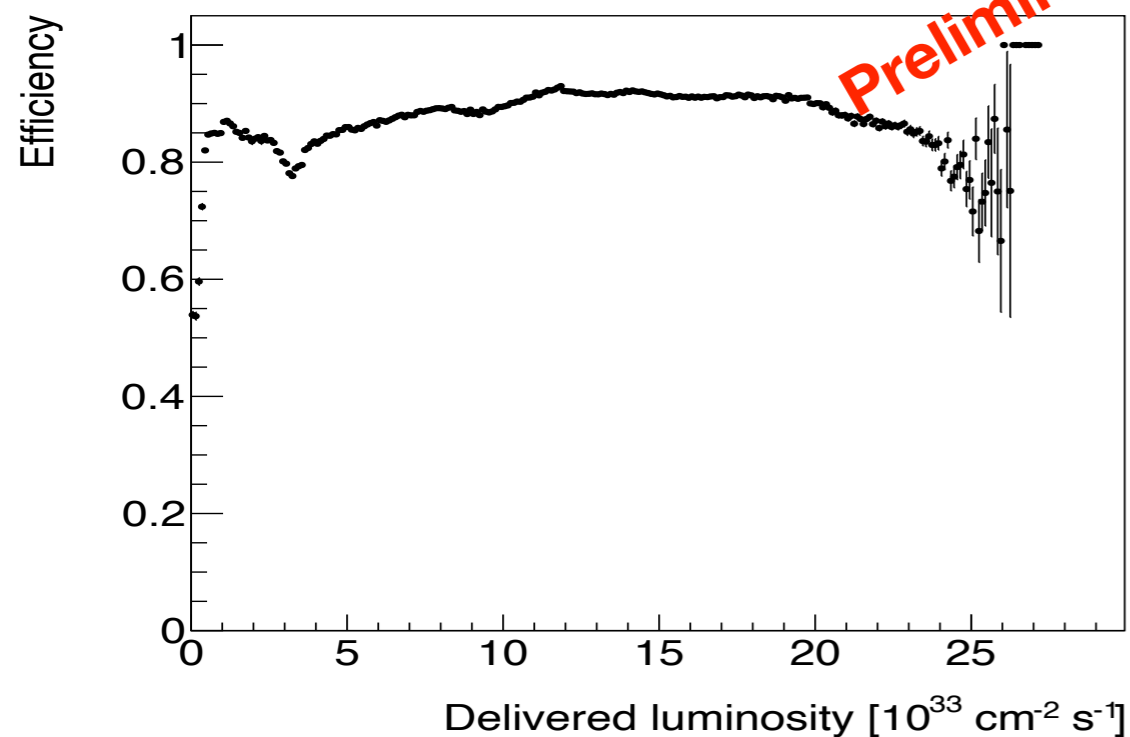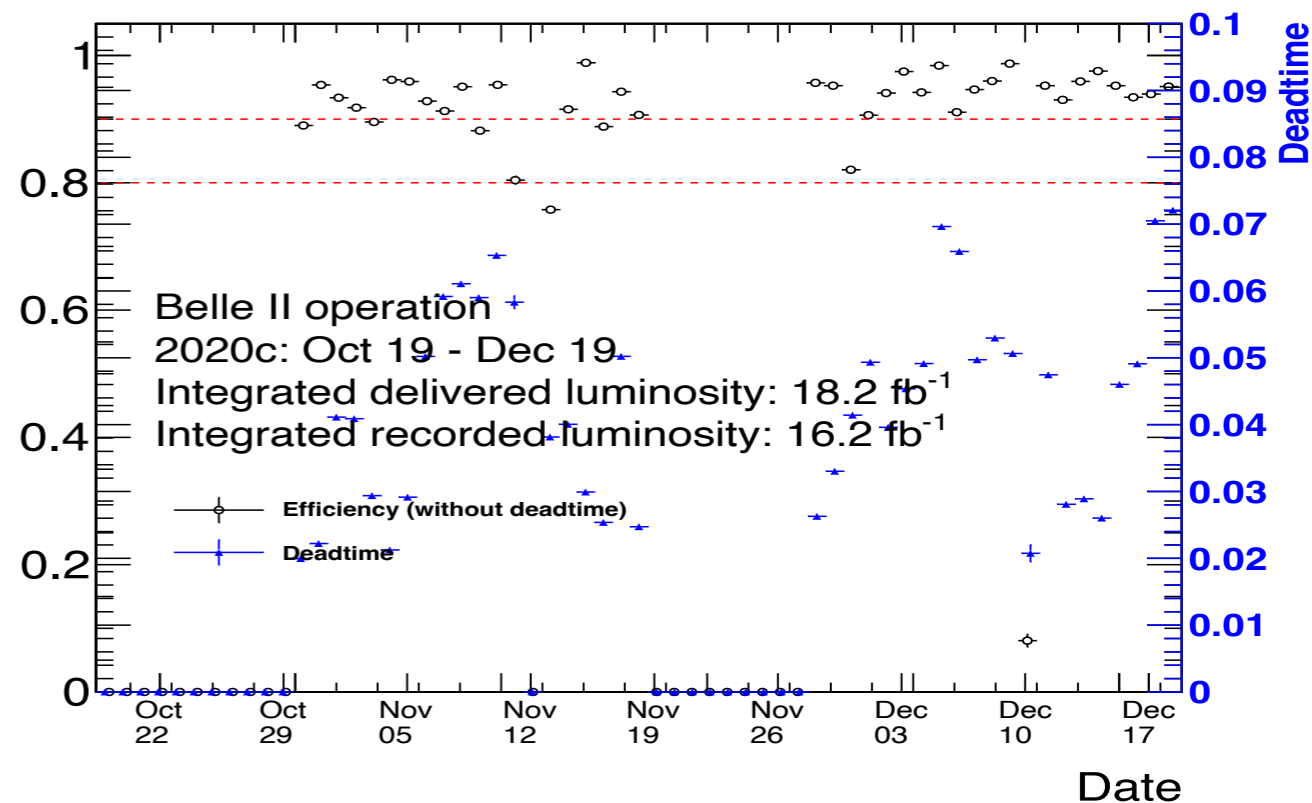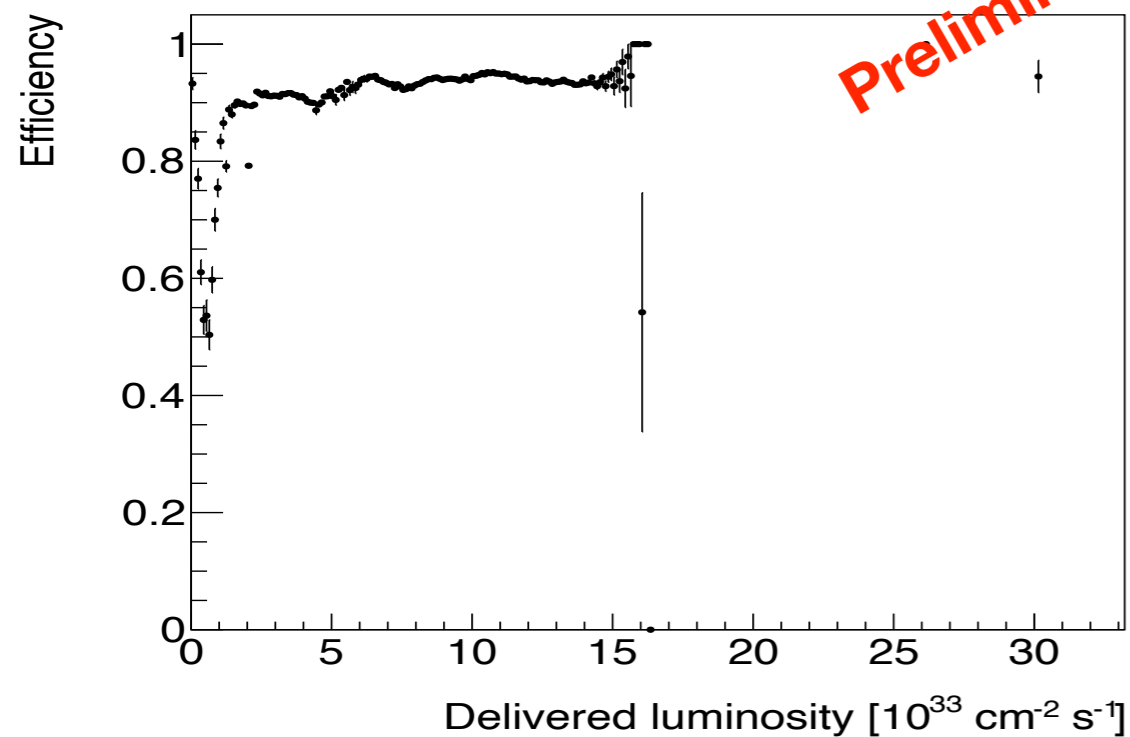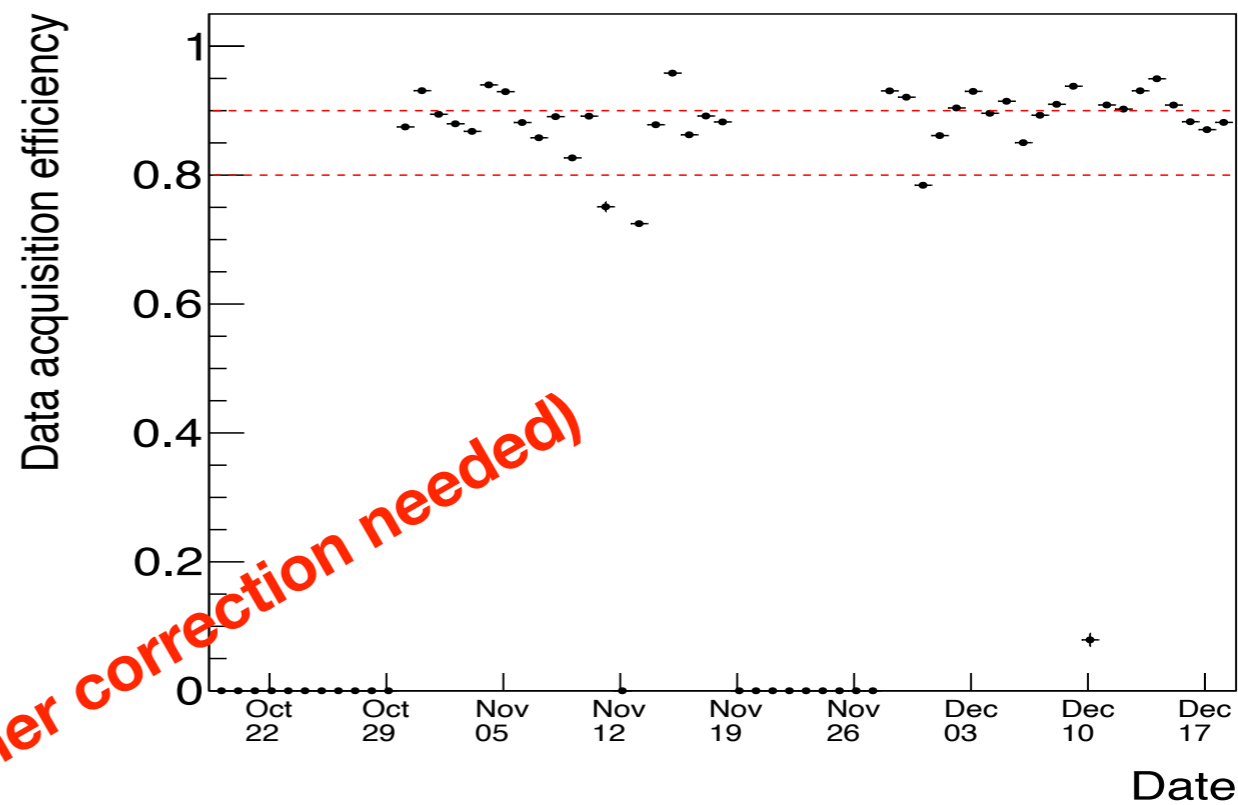
2021c run

2022ab run
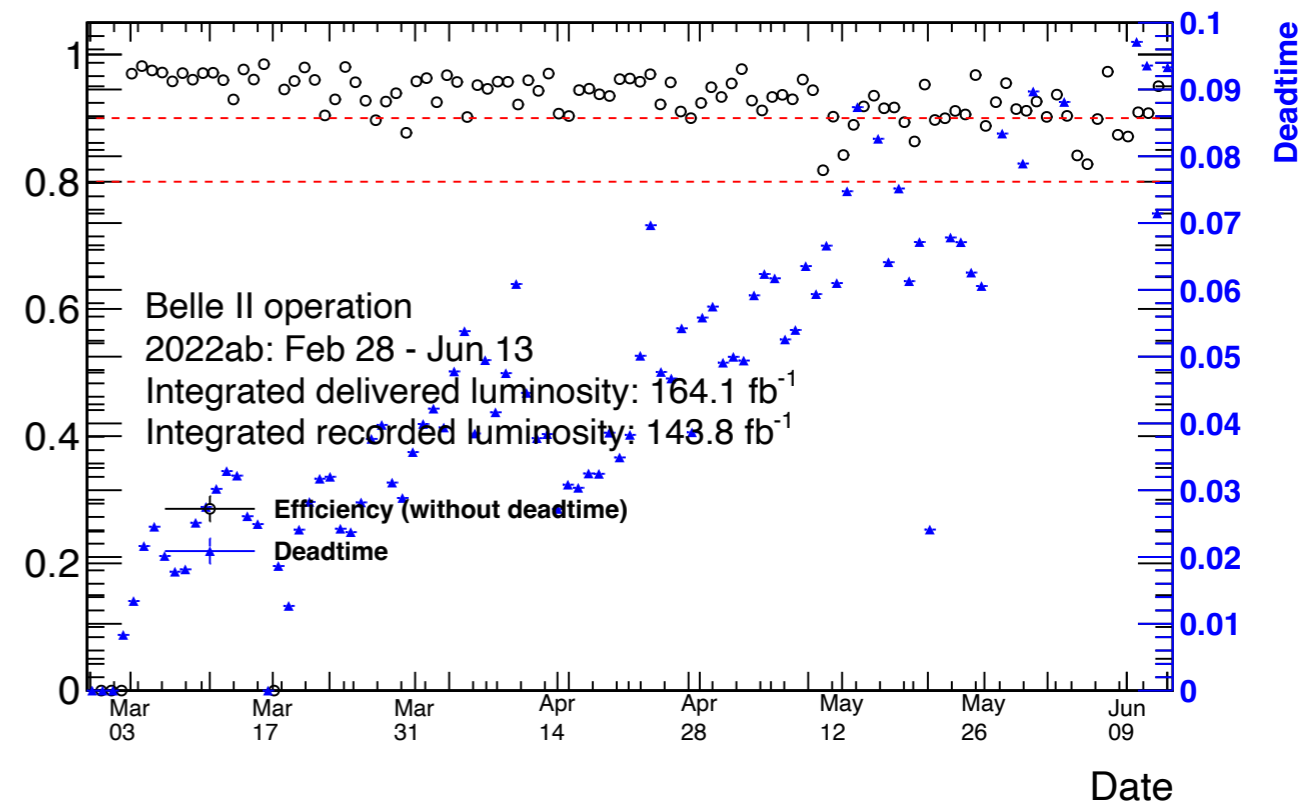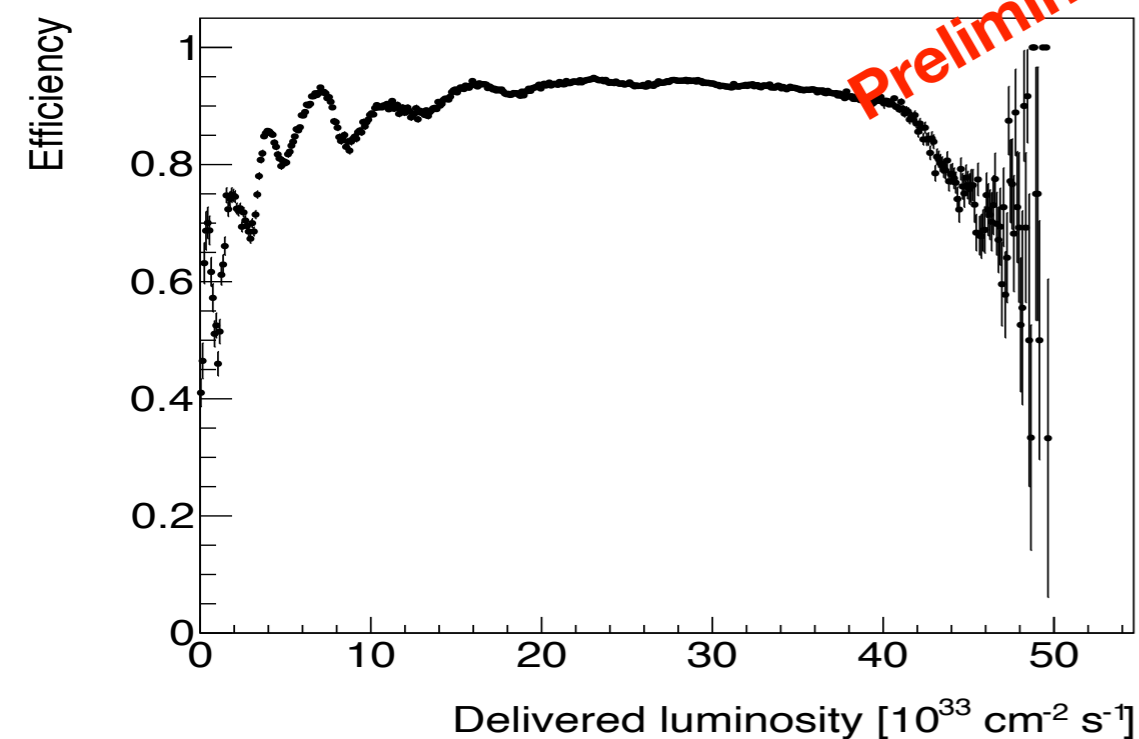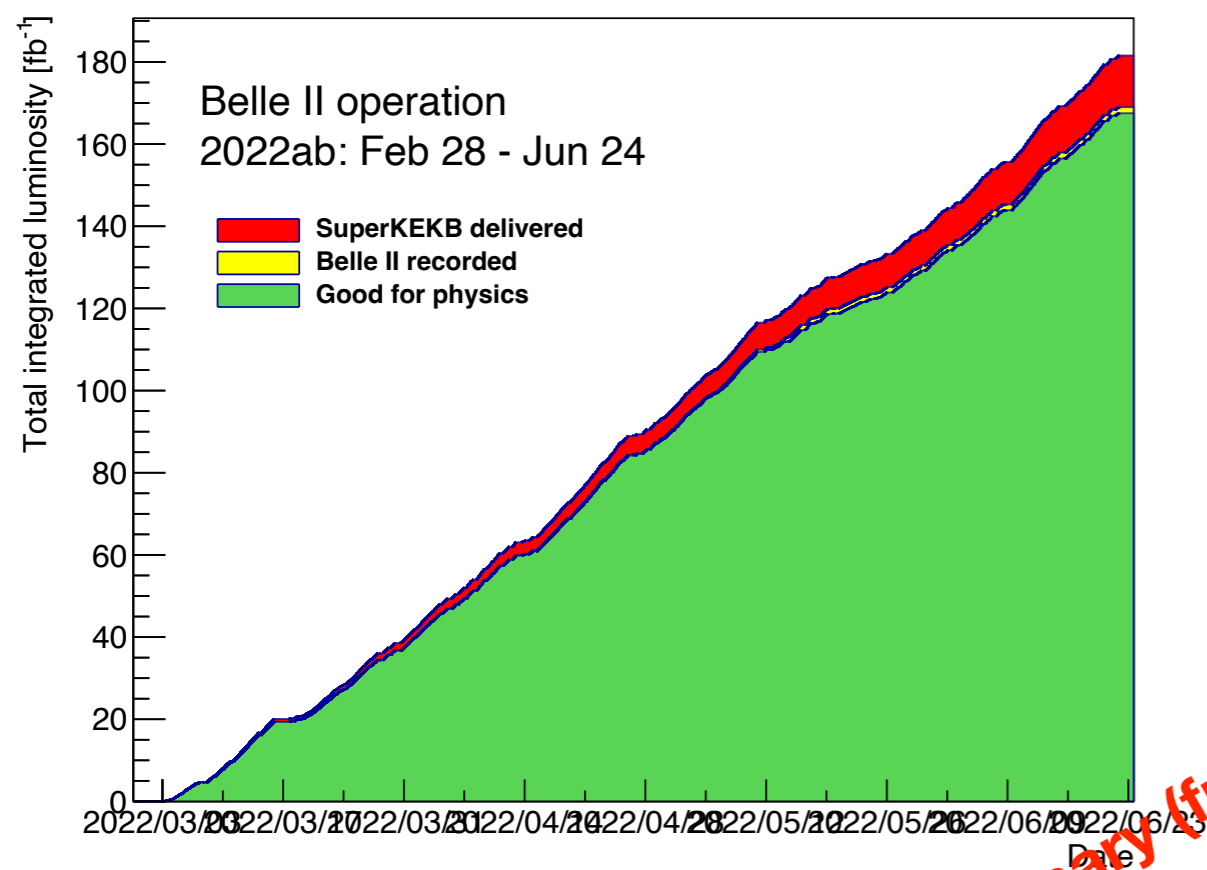
Evolution of cumulative efficiencies

Preliminary (further correction needed)

Efficiency

Delivered cumulative luminosity [fb⁻¹]

- - - 2020ab recorded   —— 2020ab good for physics
- - - 2020c recorded   —— 2020c good for physics
- - - 2021ab recorded   —— 2021ab good for physics
- - - 2021c recorded   —— 2021c good for physics
- - - 2022ab recorded   —— 2022ab good for physics
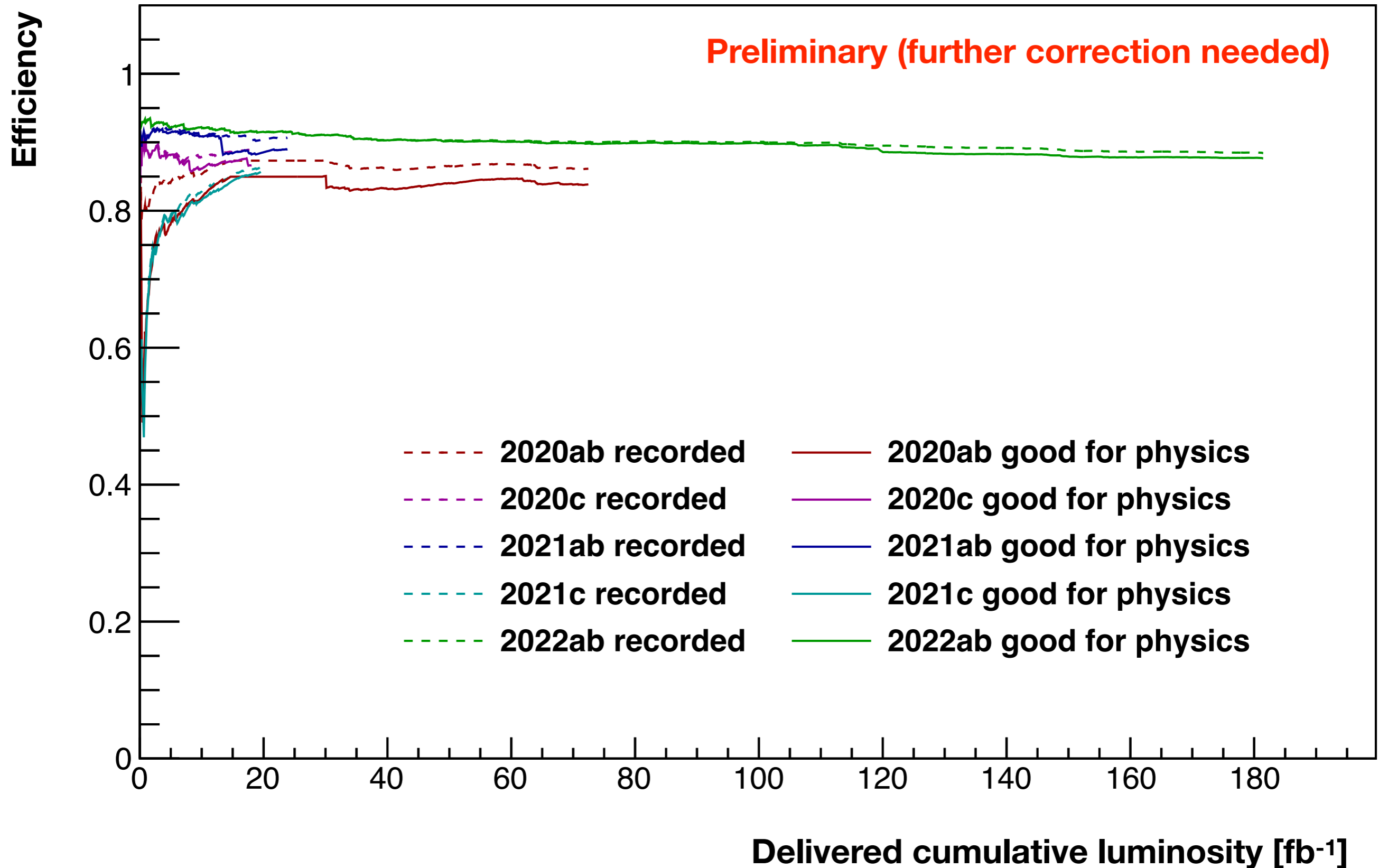
Run duration
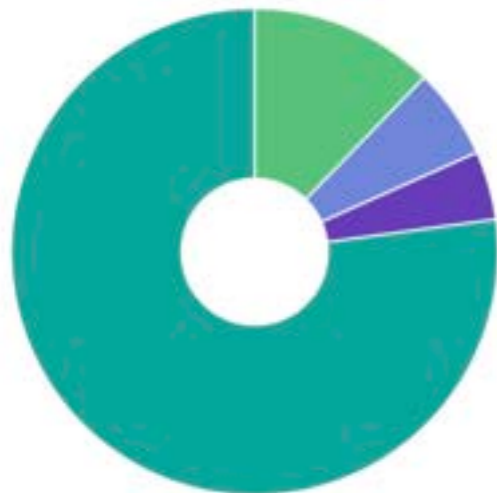
2020ab

2020c

2022ab
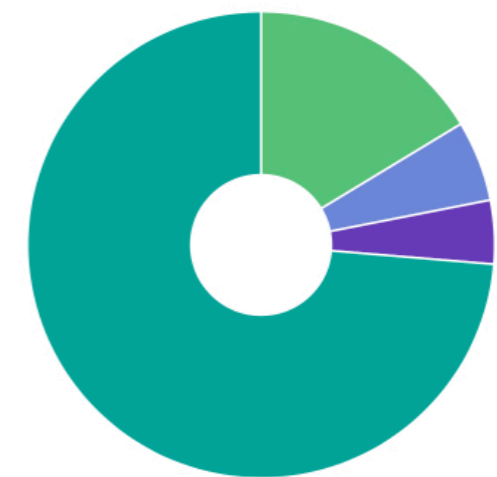
2021ab

2021c

- 1 hour < duration < 2 hour
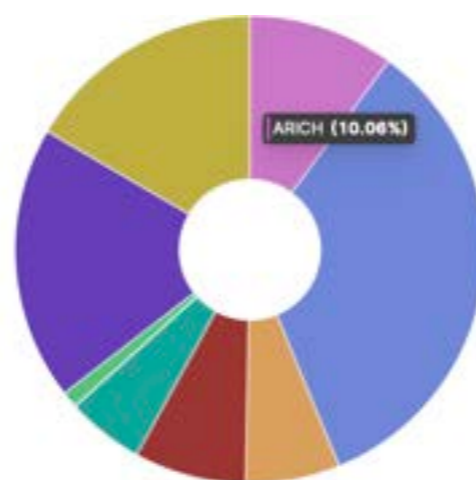- 2 hour < duration < 3 hour
- 3 hour < duration
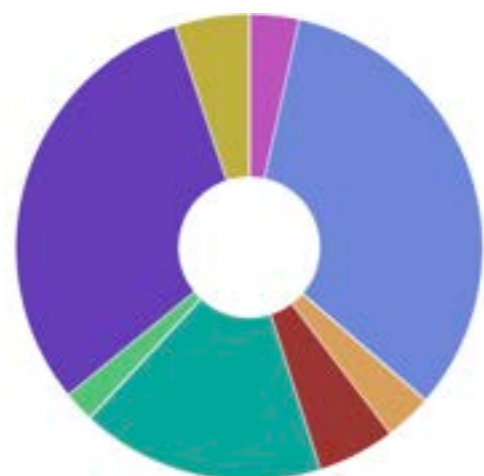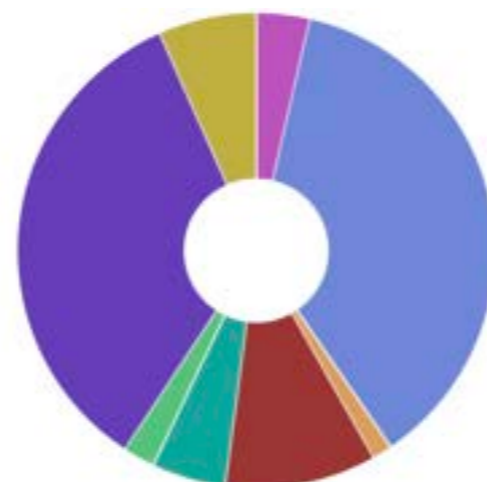- Less than 1 hour

Run stoppers categorised by sub−system

2020ab

2020c

ARICH (10.06%)
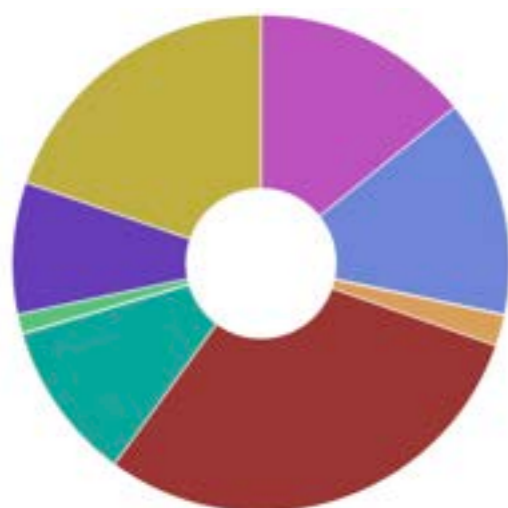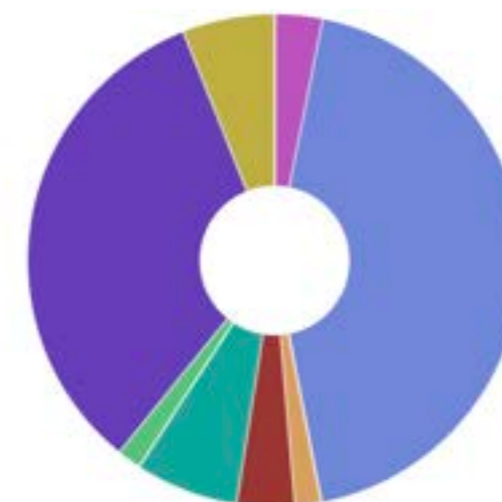
2022ab

Categorised using elog entries

2021ab

2021c

- ● ARI...
- ● CDC
- ● ECL
- ● KLM
- ● PXD
- ● SVD
- ● TOP
- ● TRG

TTD errors 2020ab

Legend (left chart):
- ttlost
- b2llost
- ttdown
- fifoerr
- semmbe,s...
- busy
- tagerr
- rerr
- b2ldown
- clklost

Right chart labels:
- KLM 43.55%
- TOP 23.88%
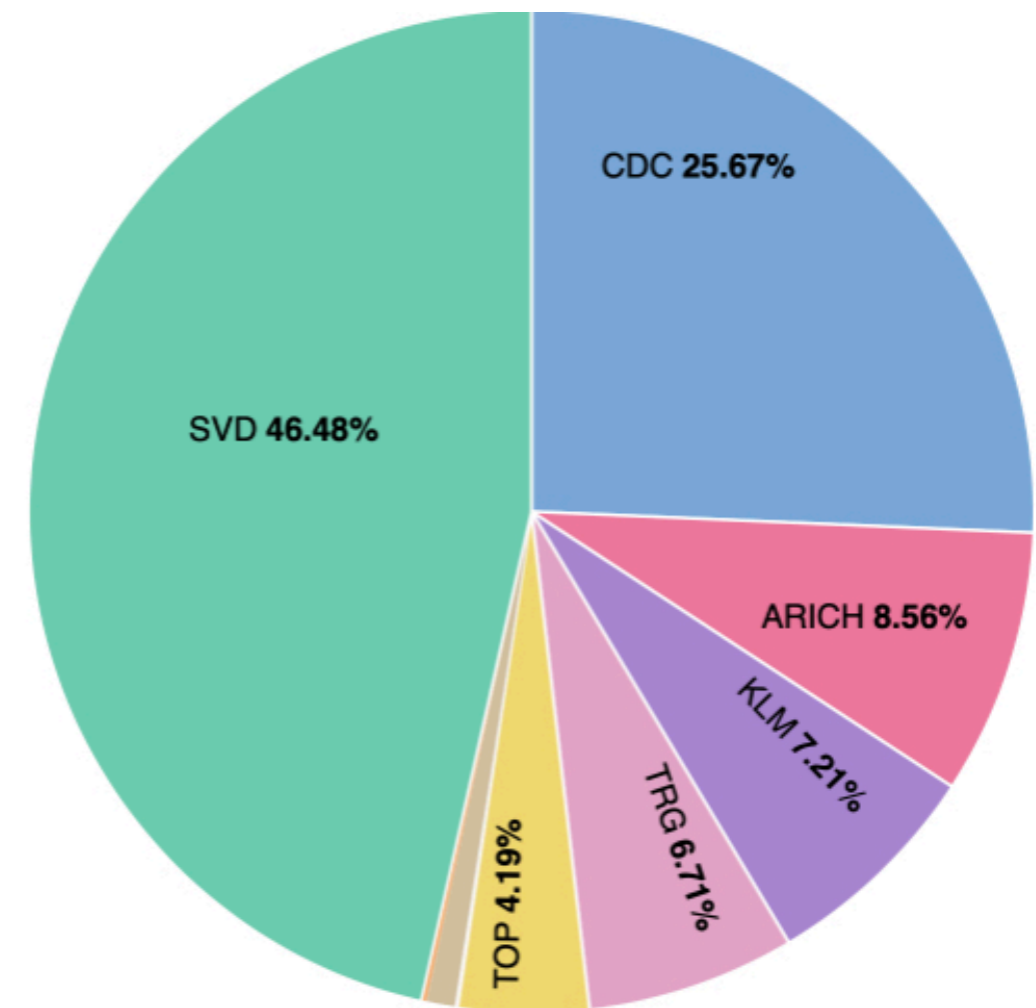- CDC 13.06%
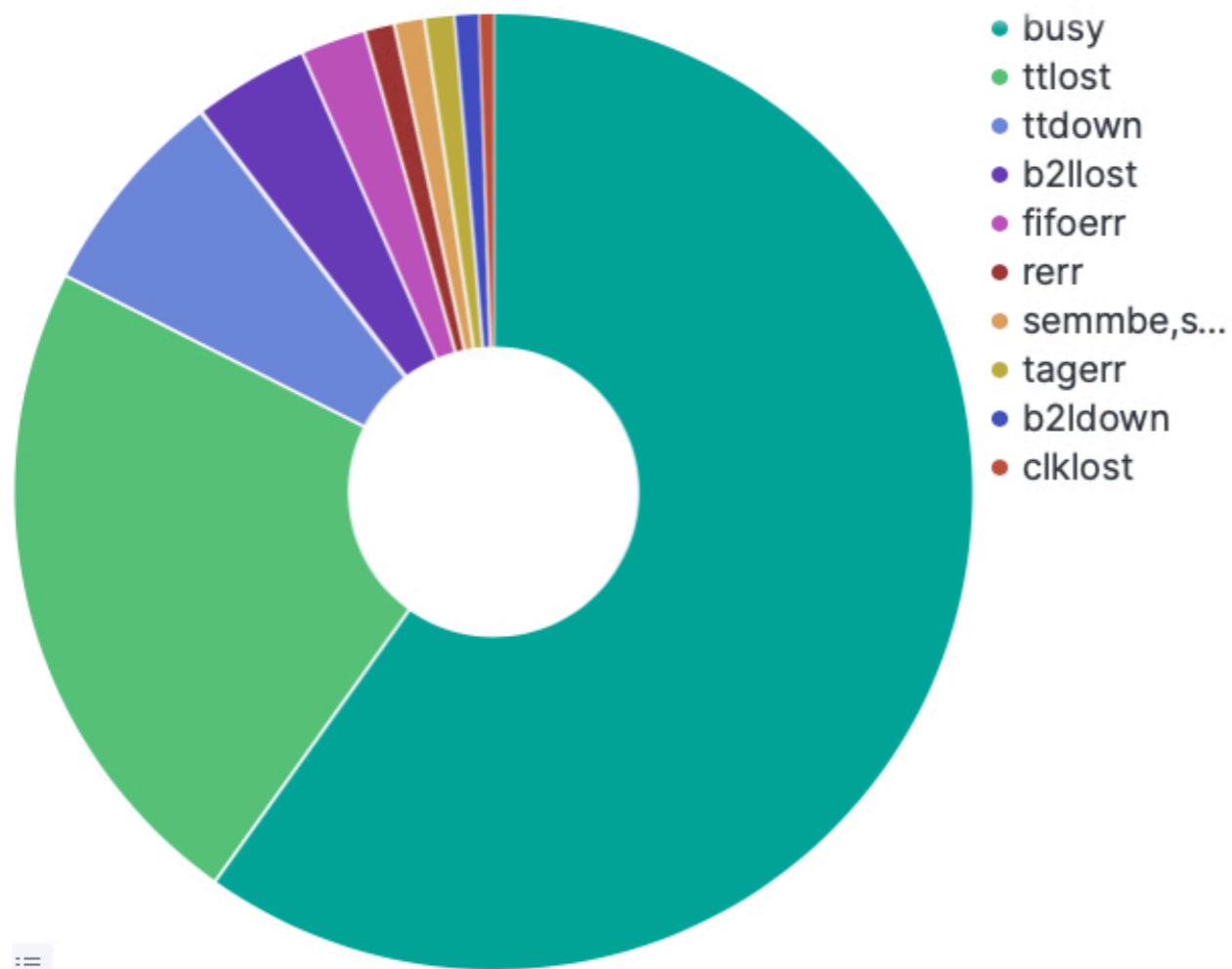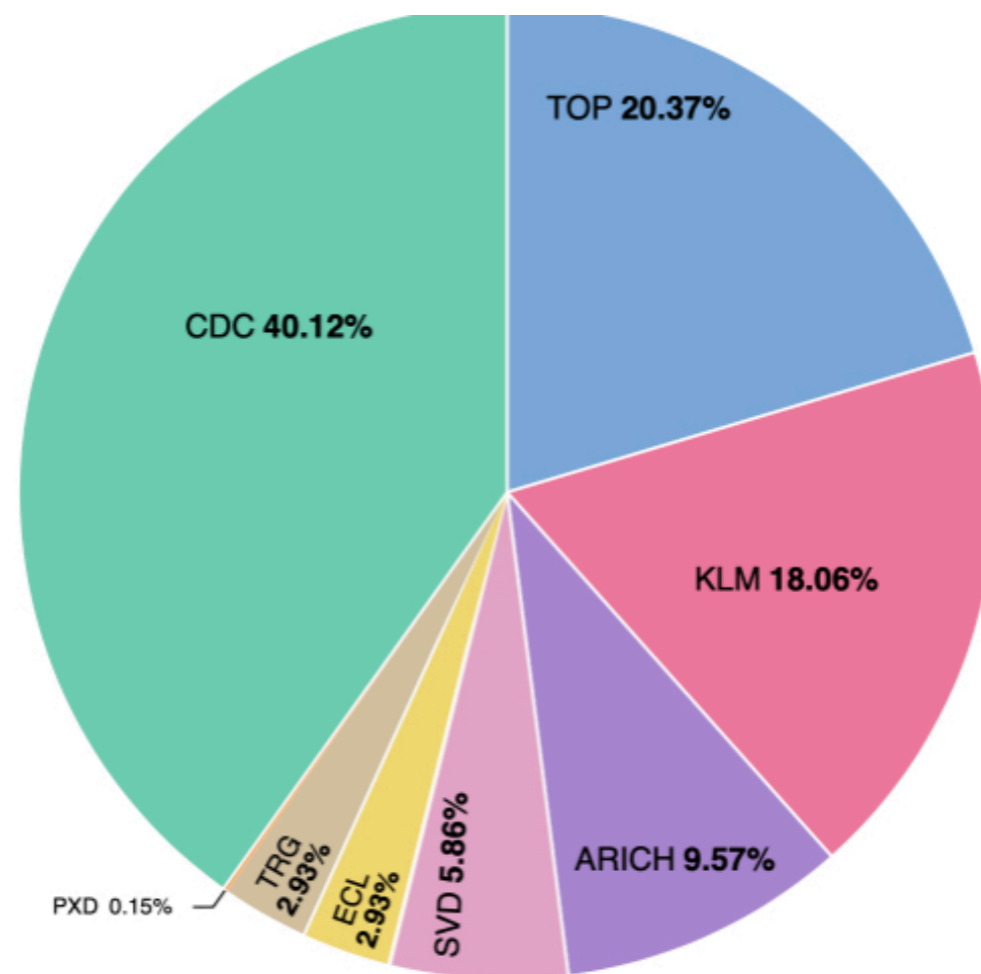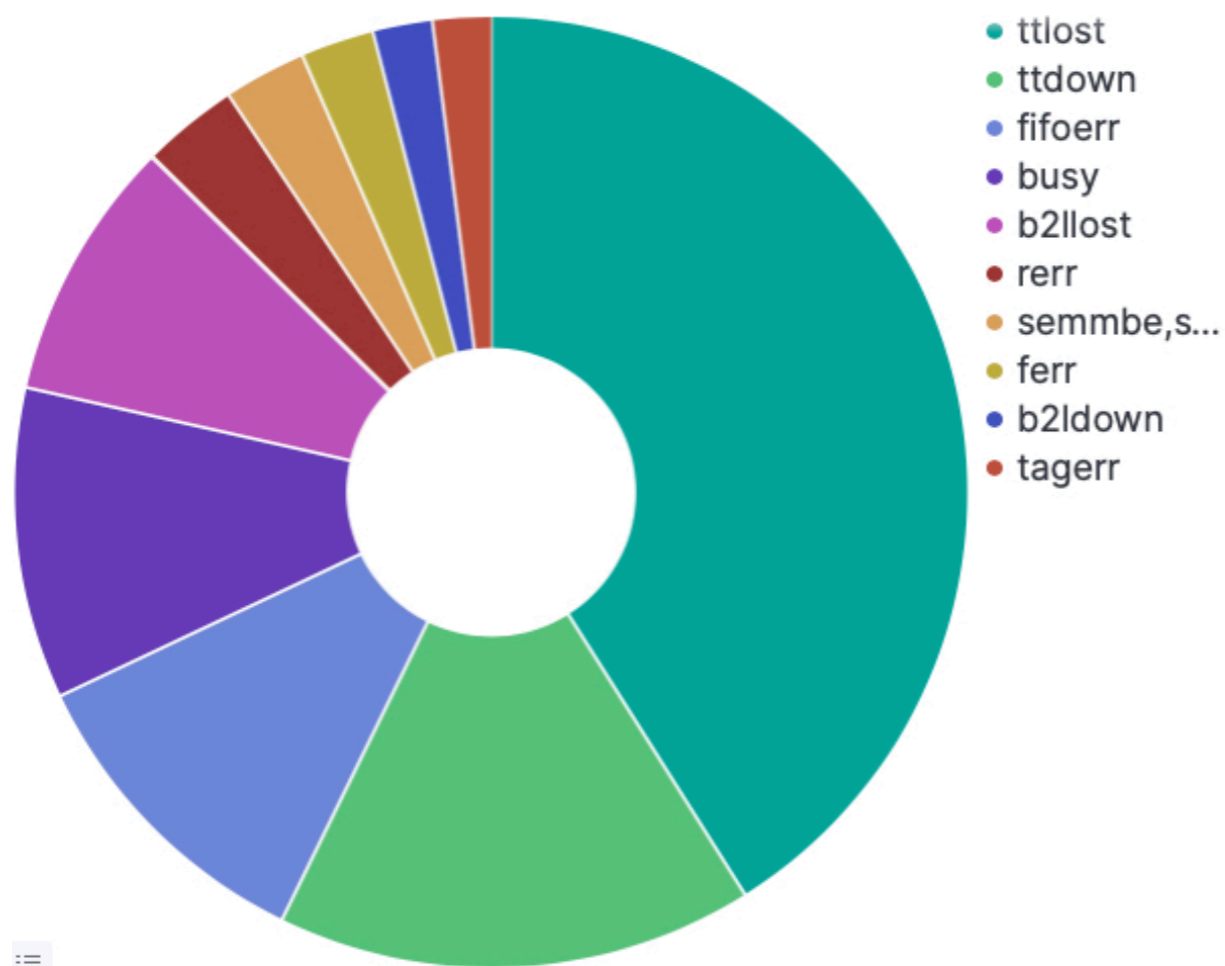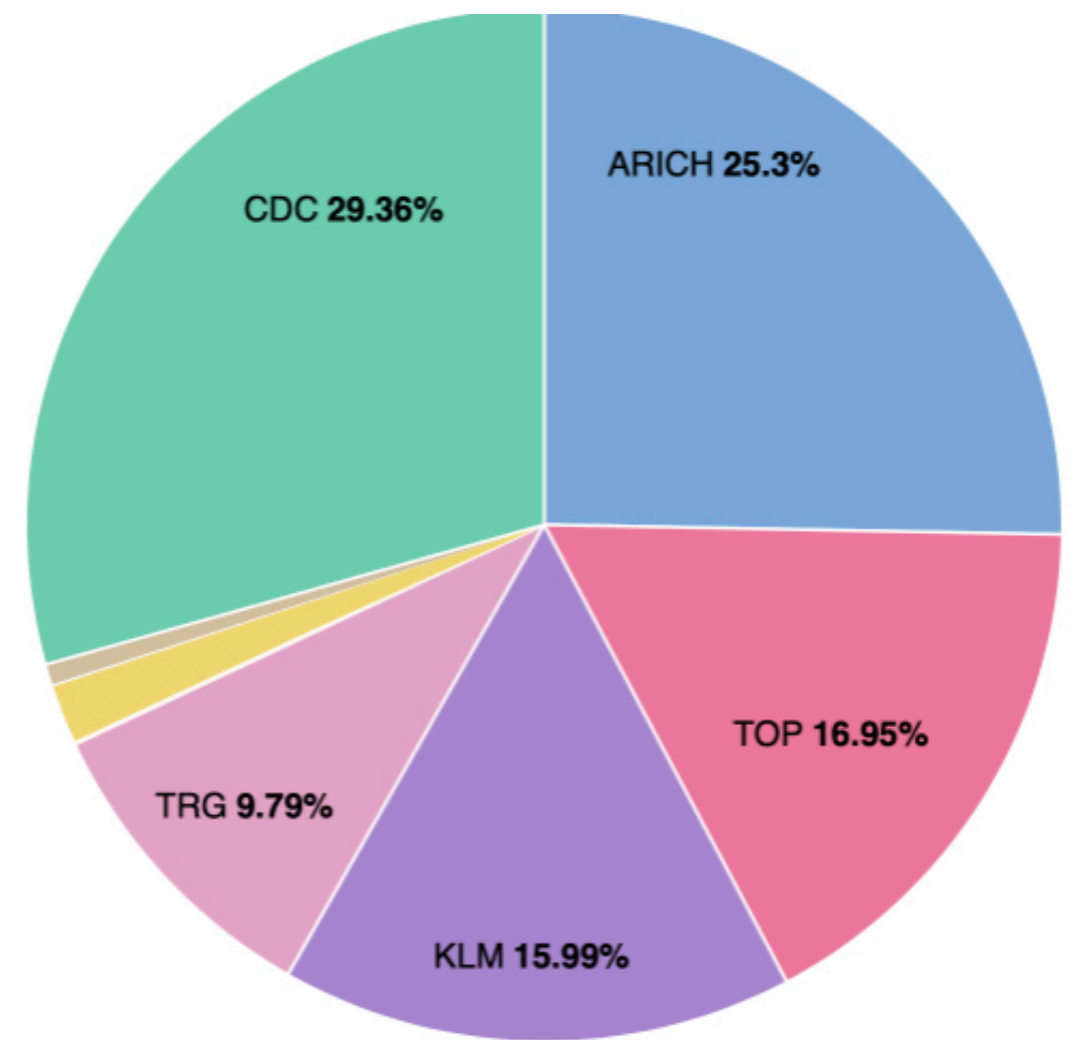- ARICH 8.8%
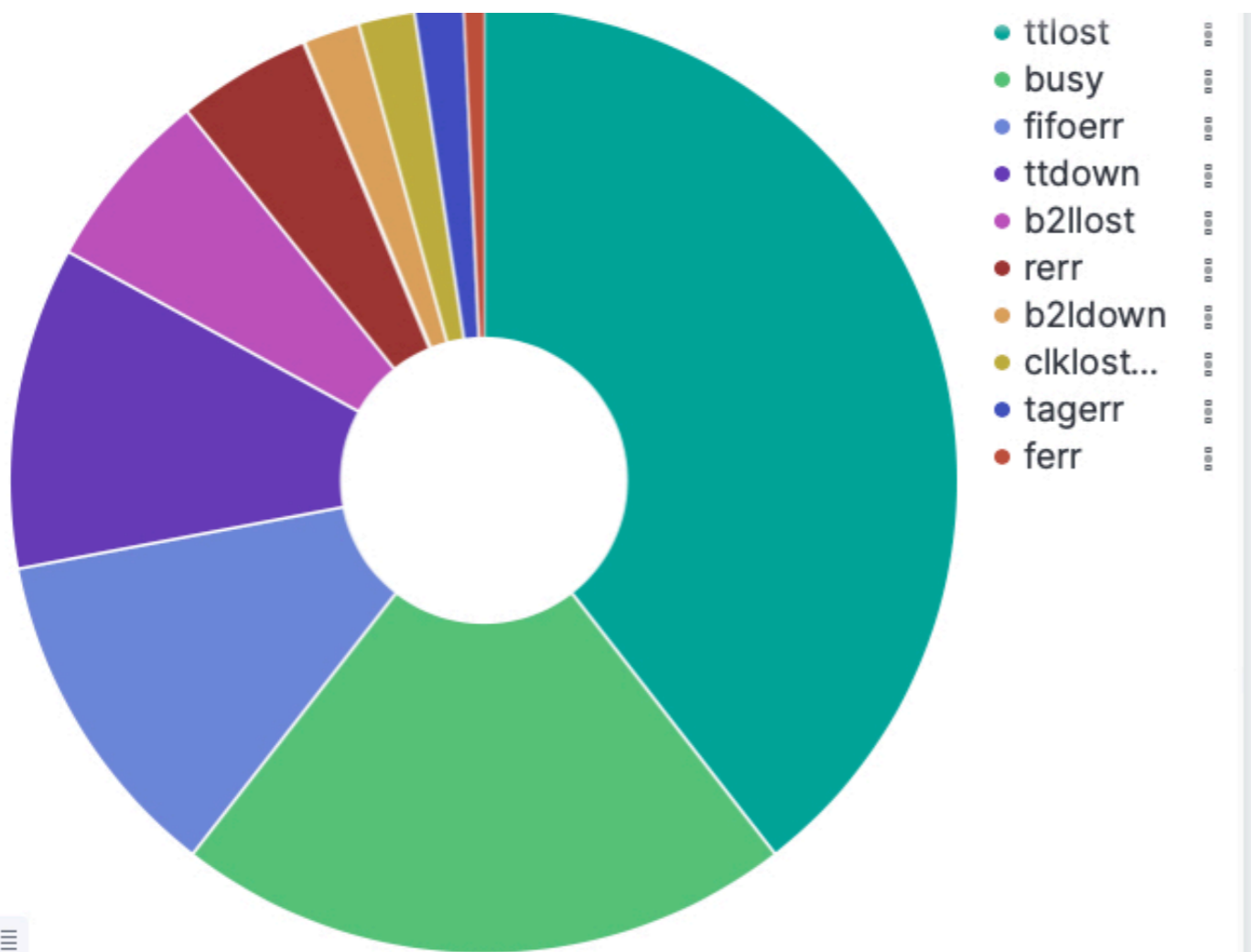- TRG 4.66%
- ECL 4.18%
- PXD 0.16%
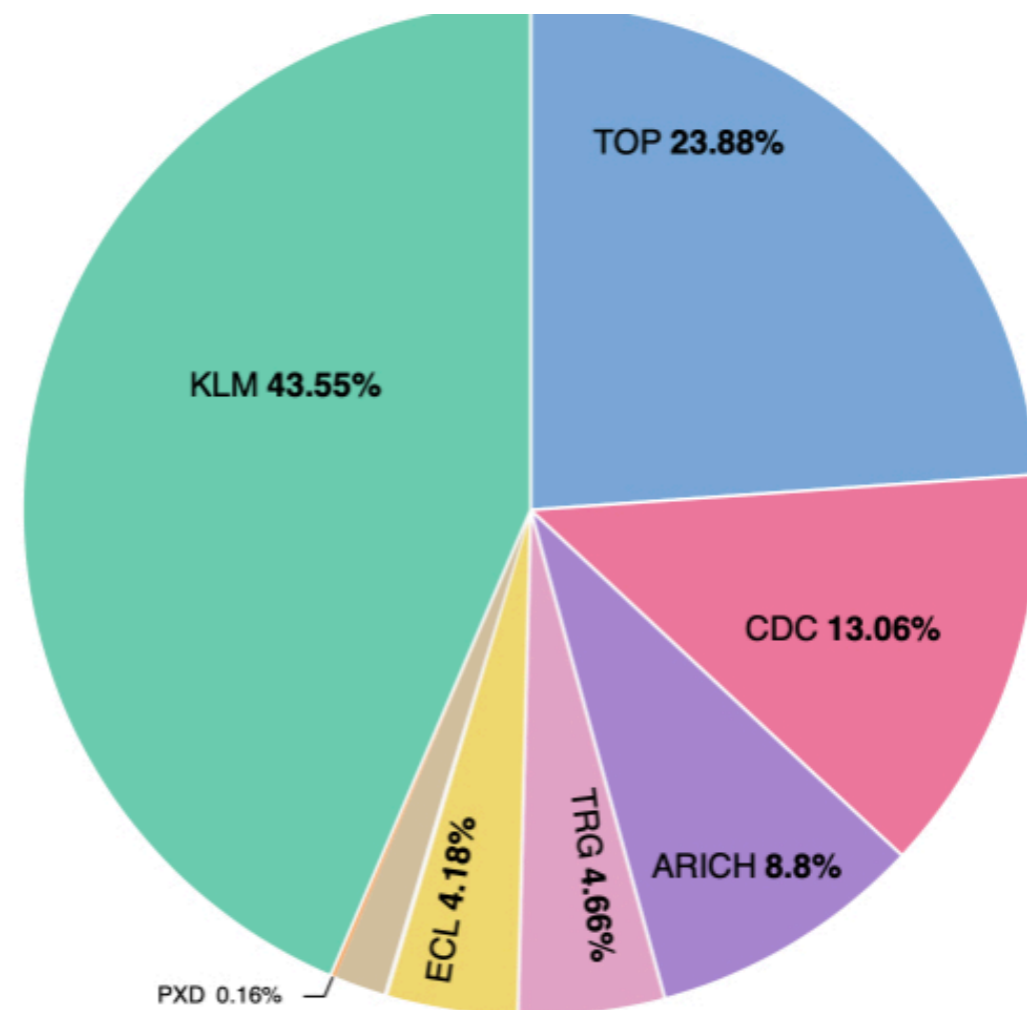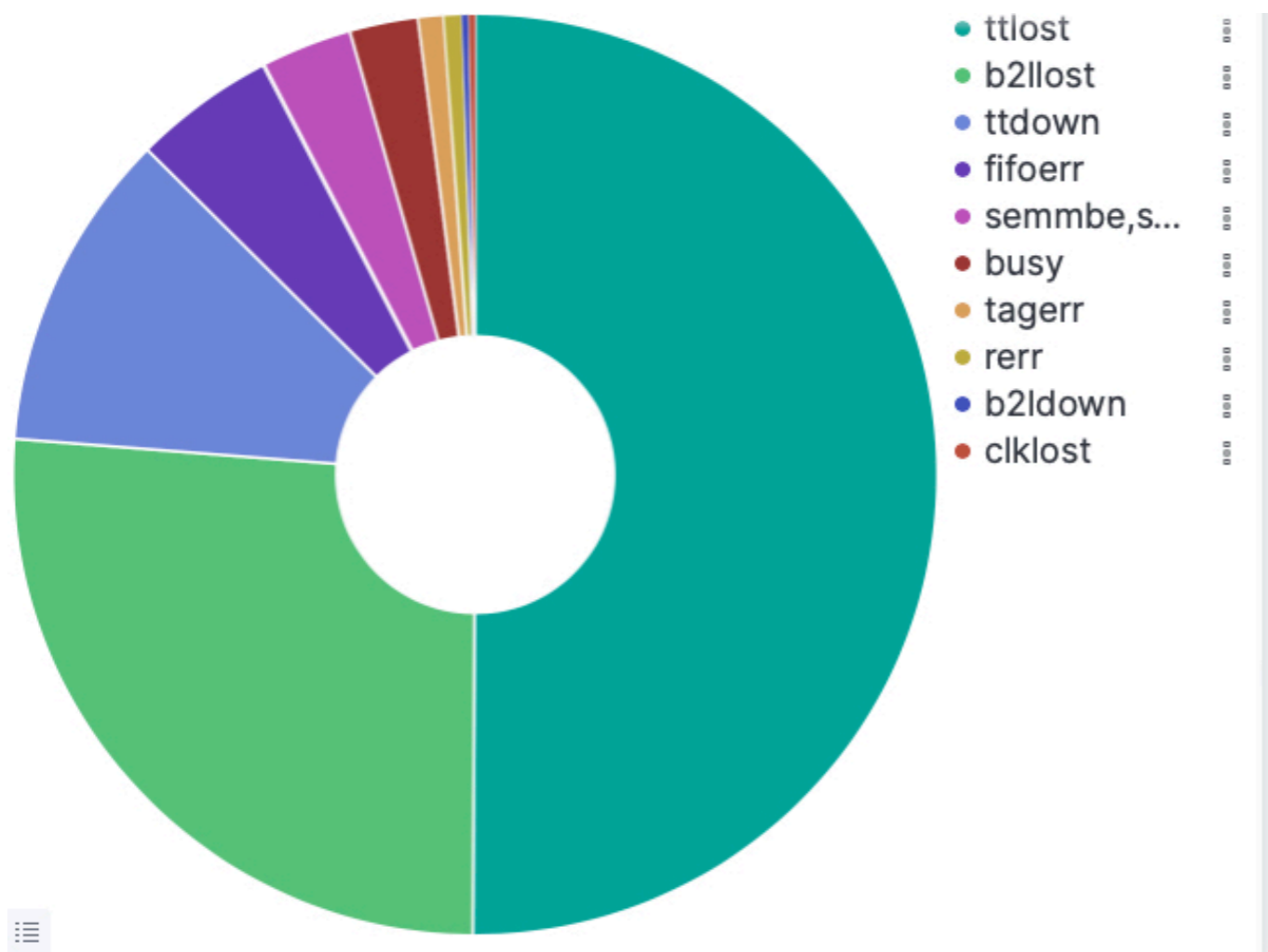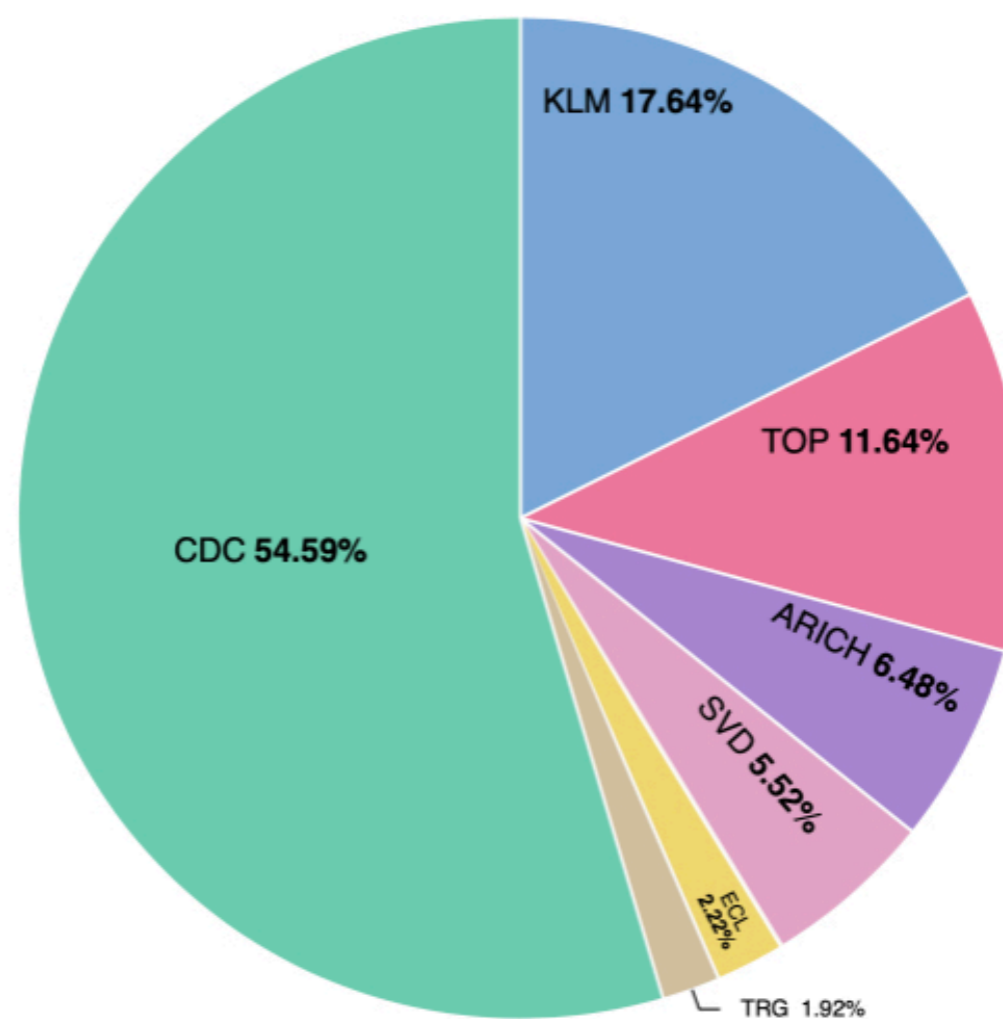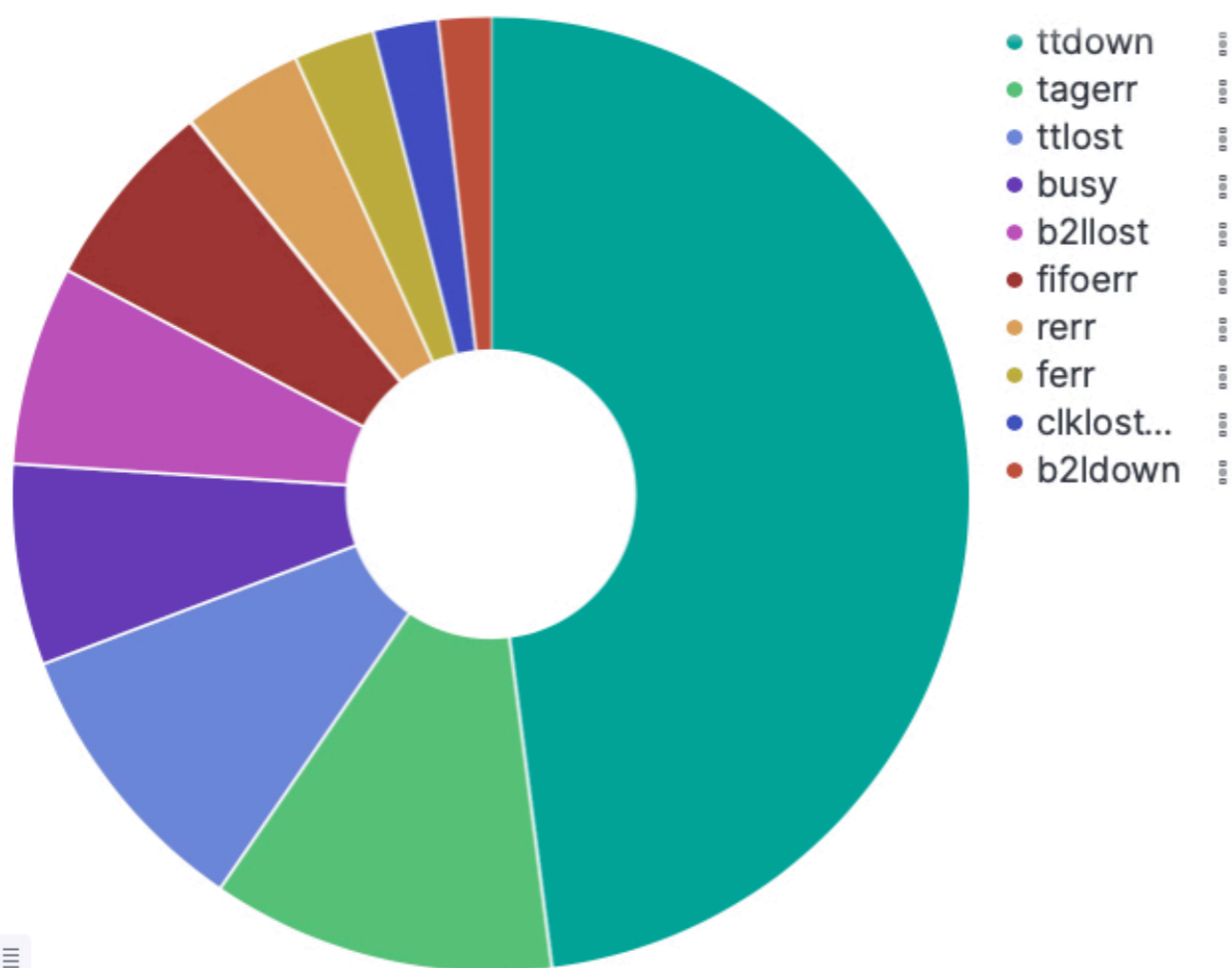
TTD errors 2020c

TTD errors 2021ab

TTD errors 2021c

TTD errors 2022ab

To where we should assign resources?

Machine learning

Why machine learning?

- **Forecasting**
  - We can catch "sign"s of errors and predict future errors
- **Root Cause Analysis in DevOps**
  - We still have multiple "UNKNOWN" errors
  - Massive multi variate approach by clustering, periodicity, etc
- **(Machine Learning in SIEM)**

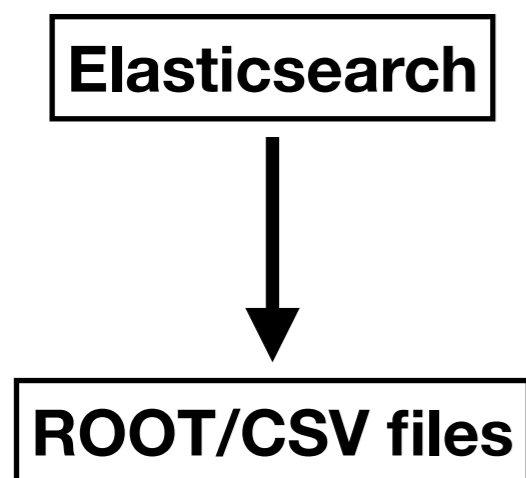**Tracks in the inter-experiment machine learning workshop**

1. ML for object identification and reconstruction
2. ML for analysis : event classification, statistical analysis and inference, including anomaly detection
3. ML for simulation and surrogate model : Application of Machine Learning to simulation or other cases where it is deemed to replace an existing complex model
4. Fast ML : Application of Machine Learning to DAQ/Trigger/Real Time Analysis
5. ML infrastructure : Hardware and software for Machine Learning
6. ML training, courses, tutorial, open datasets and challenges
7. ML for astroparticle
8. ML for phenomenology and theory
9. ML for particle accelerators
10. Other

Feasibility study?

- **Machine learning in Elasticsearch**
  - Convolution of multiple techniques (e.g. Clustering, various types of time series decomposition, Bayesian distribution modelling, Correlation analysis etc)
  - **Expensive (O(M JPY/year))**
  - Feasibility study started in a simple setup (e.g. using pycaret)

Elasticsearch → **ROOT/CSV files**

Can we use this group space?