Alarm and recovery system in the Belle II operation

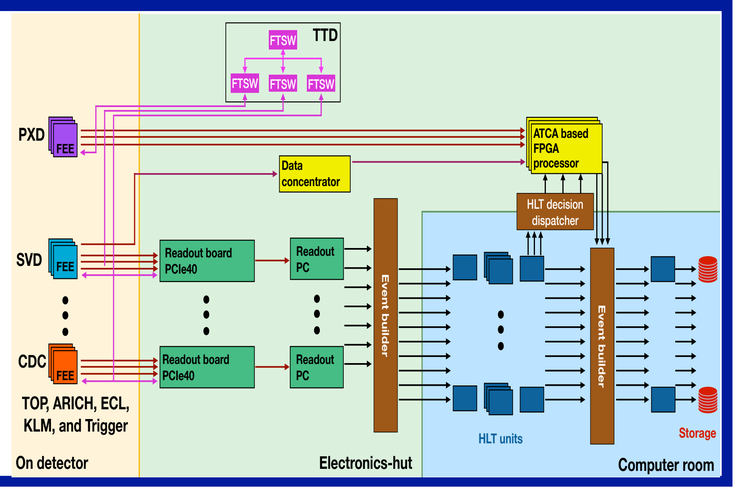
Takuto Kunigo[1], Ryosuke Itoh[1], Mikihiko Nakao[1], Soh Suzuki[1], Satoru Yamada[1], Seokhee Park[1], Dmytro Levit[1], Jaeyoung Kim[2], Chanho Kim[2], Tommy Lam[3]

[1]KEK, Japan, [2]Yonsei University, Korea, [3]Virginia Tech, USA

Belle II DAQ system

Our system consists of various components
- Detector front-end electronics
- Readout system + Event builder
- 2-staged trigger system (Level-1: hardware, HLT: software)
- Networks: EPICS, Network Shared Memory 2, general purpose network
- Storage

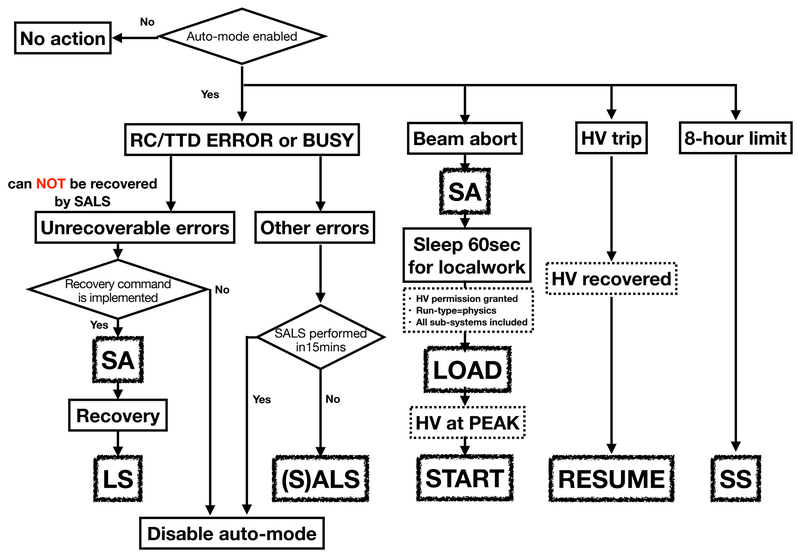
We need to monitor each component effectively; **automatic error diagnoses** is a key



Automatic recovery

Auto-pilot Process

We have an auto-pilot process which takes the required actions in case of:
- Beam dump
- High Voltage trip
- ERROR and BUSY



Automatic Recovery

When the auto-pilot process detects a run stop, it checks whether there is an activated EPICS alarm PV. If there is an activated PV, the auto-pilot process reads the corresponding Elastalert rule file to load the implemented recovery command. The auto-pilot process executes the loaded recovery command automatically.

```
55  command: ["/home/vagrant/bin/caput.py", "B2_alarm:CDC:FEEREPROGRAM:ID", "{FEEs}"]
56  recovery_command:
57    host: "ttd11"
58    user: "b2shift"
59    rcrequests_before_recovery: "STOP-ABORT"
60    command: "cd scripts/cdc_config && ./config_parallel.py --ID caget(B2_alarm:CDC:FEEREPROGRAM:ID)"
61    reset:
62      PV: "B2_alarm:CDC:FEEREPROGRAM:ID"
63      value: "-1"
```

Monitoring system based on Elastic Stack

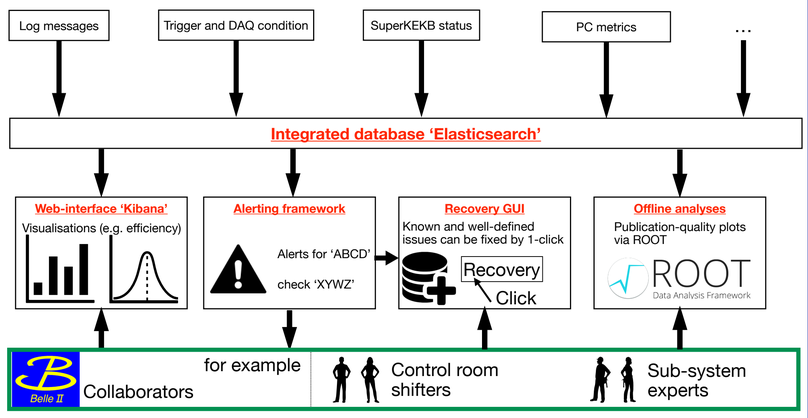
We develop a monitoring system based on Elastic Stack. We feed various operational data into Elasticsearch, provide various applications using it.

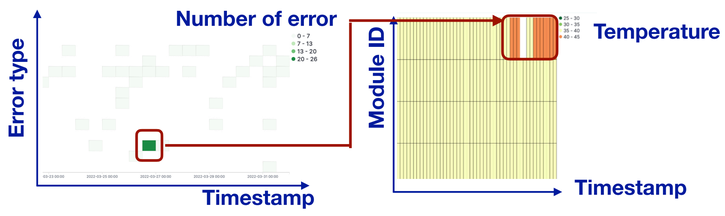
Web-interface

We analyse the data on a web-interface, named Kibana. We check correlation between different data sources. (e.g. Log message vs PC metrics)

Alerting Framework

We use a third-party tool, Elastalert, to diagnose errors automatically. Elastalert periodically performs queries on Elasticesarch, and checks if the queried data satisfy the alert rules. When there is a match, Elastalert issues a set of alerts.





Alarm system

Phoebus Alarm system

In real operation, one problem can cause multiple errors. We use Phoebus Alarm System to help the operators to take correct actions.

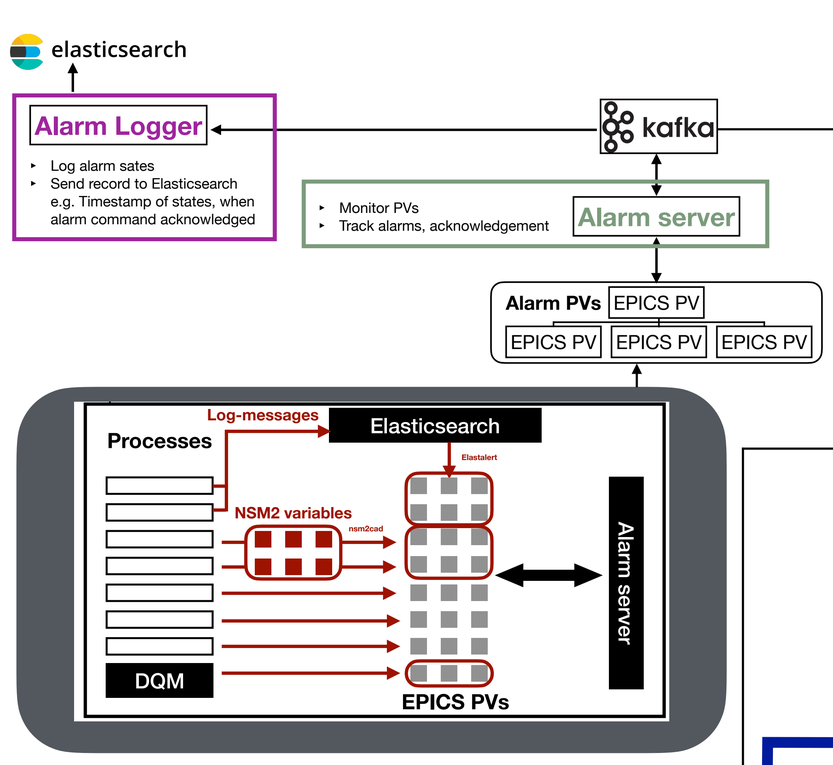
We set alarm limit to each EPICS PV, which can trigger alarm. The alarm server process monitors all the alarm trigger PVs, and track the alarm status. The information is shared via Apache Kafka; the alarm logger process save the information to Elasticsearch, alarm client GUI accesses the information to visualise it.

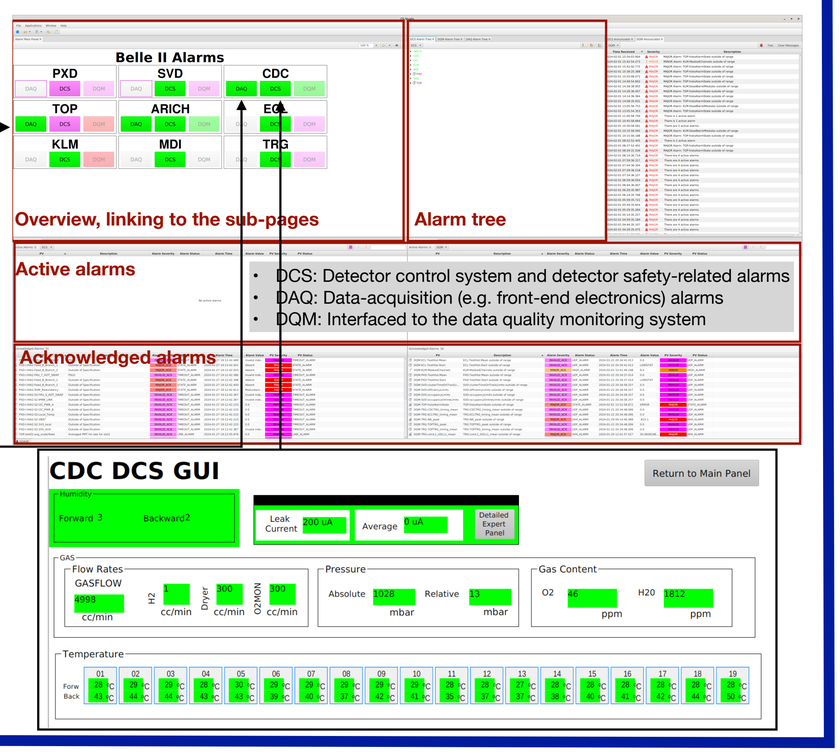
Alarm Categories

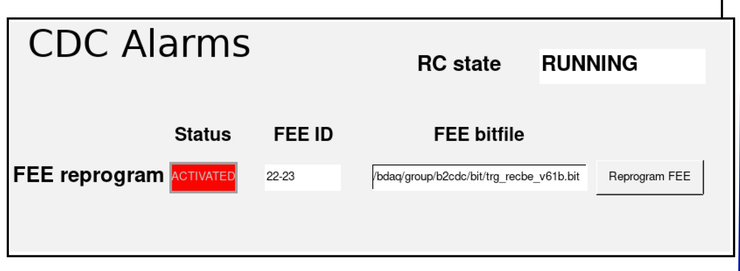
We have three types of alarms:
- DAQ: Data acquisition alarms (e.g. error of front-end electronics)
- DCS: Detector control system and detector safety-related alarms (e.g. High Voltage error)
- DQM: Interfaced to the data quality monitoring system. (e.g. tracking efficiency)

Some alarms cannot be activated directly by an EPICS PV. For example, in order to activate an alarm using log-messages, we exploit Elastalert.

We make a set of GUI windows; from the top window, more detailed sub-windows are linked. The operators access more detailed information naturally. For the DAQ alarms, we have a one-click recovery button for each alarm, this is associated to an appropriate recovery action, and is enabled only when the corresponding alarm is activated and data-taking is not on-going.







Summary and Plan

Summary

- We develop an efficient monitoring and recovery system for Belle II operation
- We have an auto-pilot process, which handles the automatic recovery actions
- To help the operators take correct actions, we exploit Phoebus Alarm System
- The alarm system already helped us to identify several detector issues

Plan

- We check any missing items in the alarm system
- More sophisticated alarm condition (dynamically change the alarm threshold)
- We will study machine learning based alarms

Status

We started physics data-taking since February, 2024. The alarm system already helped to identify several detector issues. We are adding more alarms and optimising the alarm limit;
- Dynamic limit; change the alarm limit, for example, depending on the accelerator status
- Alarm delay; only alarm if the PV remains in alarm for X seconds
- Alarm delay&count; alarm when the PV becomes alarm more often than Y times in X seconds

24th IEEE Real Time Conference - ICISE, Quy Nhon, Vietnam