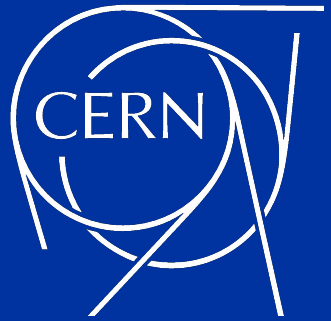


CMS MiniDAQ-3

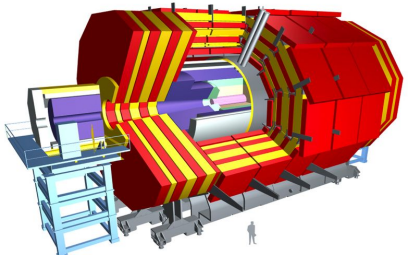
providing concurrent independent subdetector data-taking on CMS production DAQ resources

Philipp Brummer on behalf of the CMS DAQ group – philipp.brummer@cern.ch

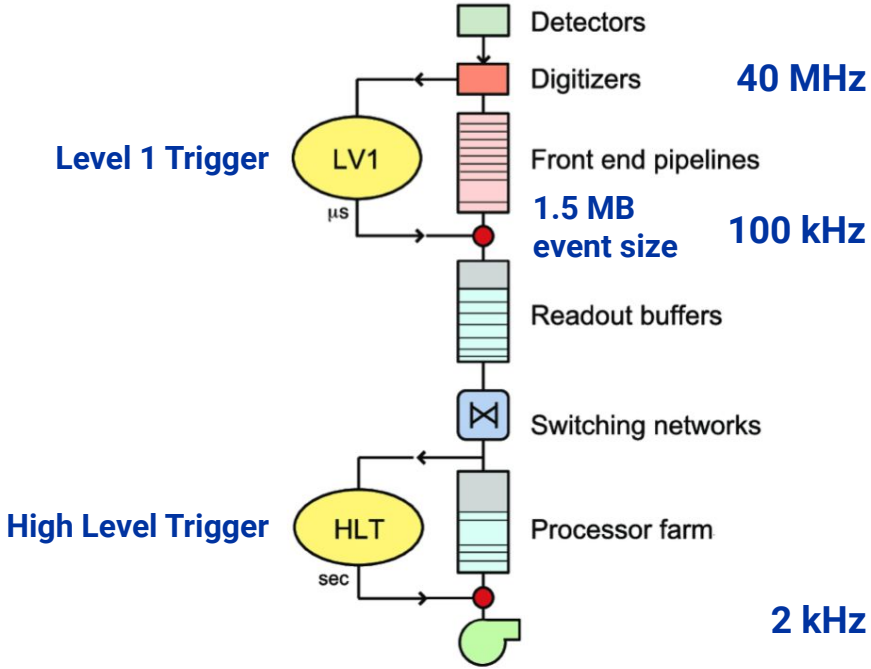
CHEP2023 – May 11, 2023 – Track 2 - Online Computing

CMS DAQ-3

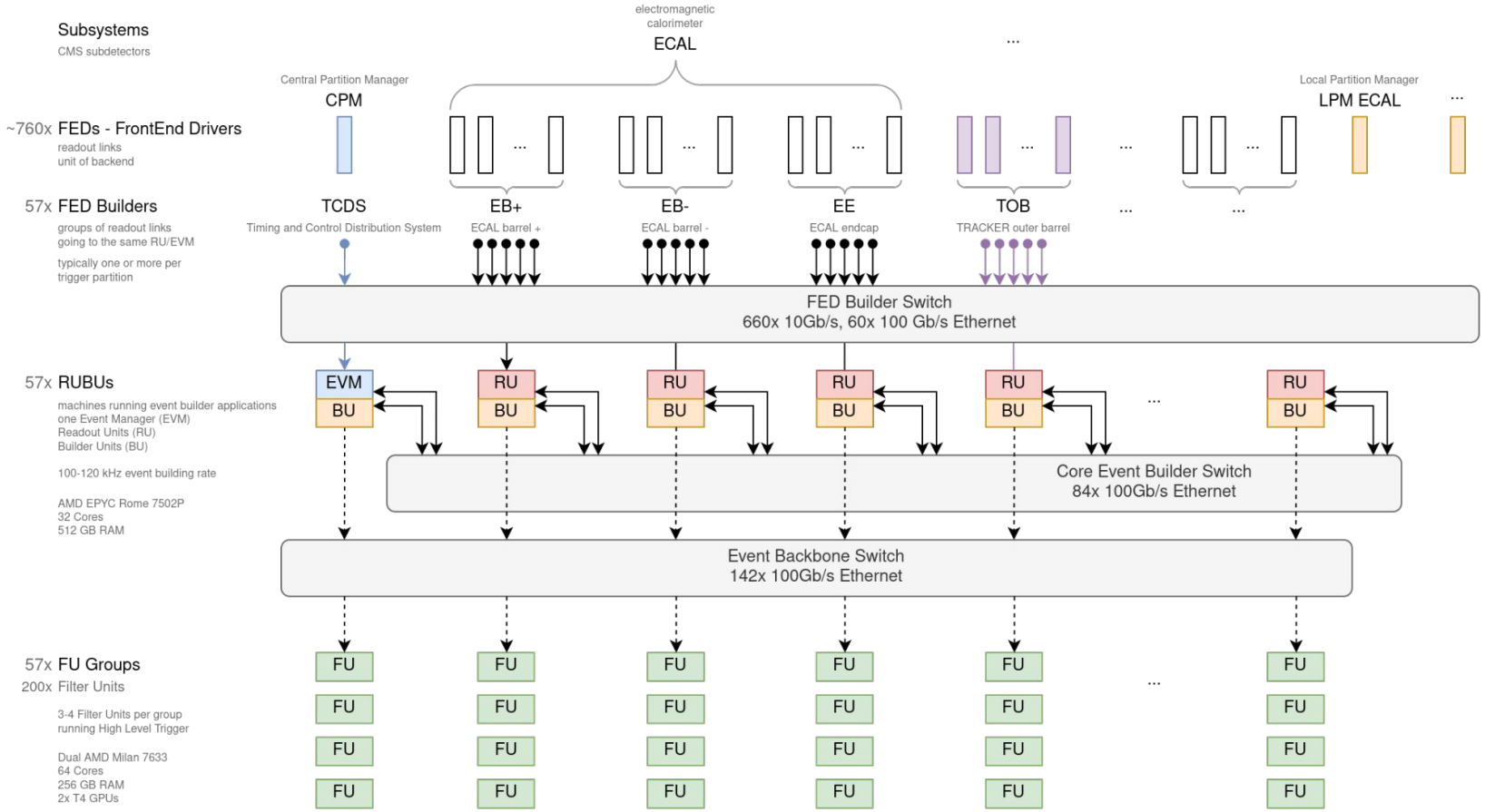
CMS Data Acquisition during Run 3 of the LHC



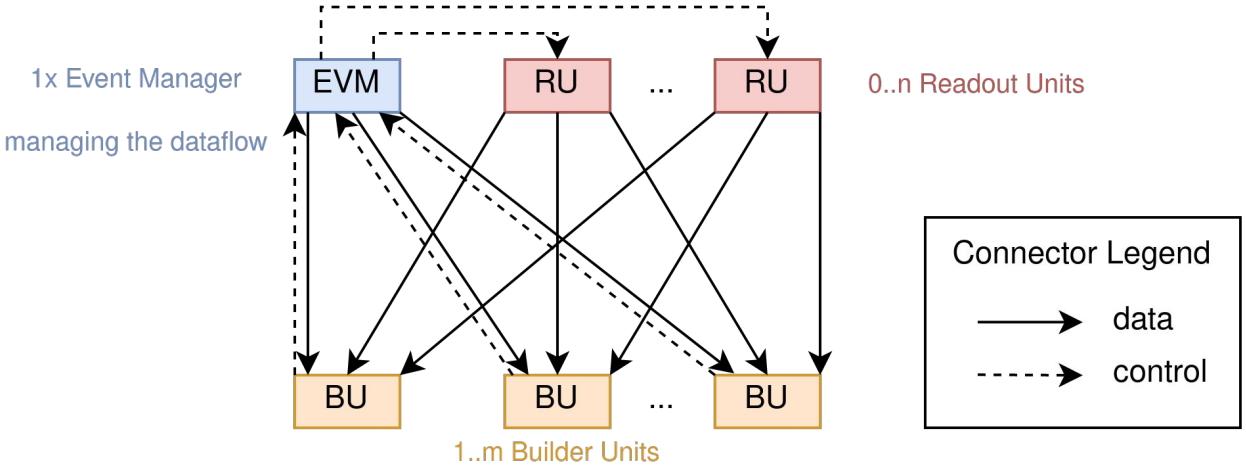
Compact Muon Solenoid



CMS Data Acquisition for Run 3 (DAQ-3) Overview



CMS DAQ-3 Event Builder (EVB) Overview

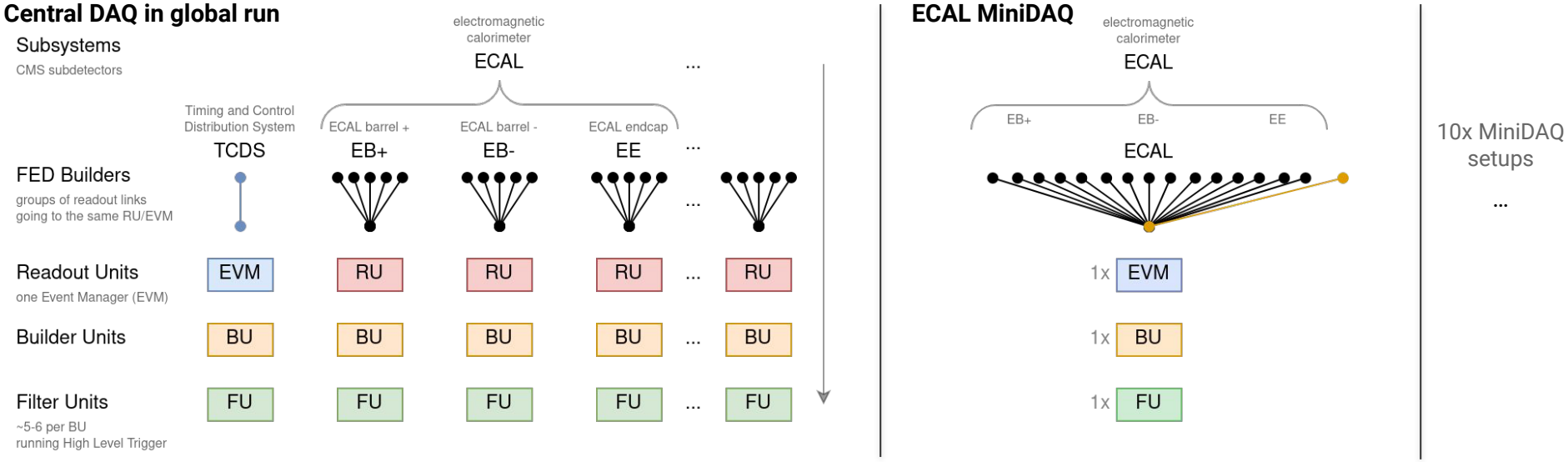


- **Builder Units request data from the Event Manager when they have free processing capacity**
- **The EVM then informs all Readout Units to send data to a BU with free capacity**
- **Once an event is built, the Filter Units connected to the Builder Unit can process it**

What are MiniDAQs and why do we need them?

- **During CMS data taking, subsystems are in the global run with the central DAQ**
- **Between LHC fills, MiniDAQs allow subsystems to run with the full DAQ chain outside of the global run**
  - This includes Event Building, the High Level Trigger and transfer to storage
  - Used for calibration runs and integration tests with the DAQ and HLT
  - As long as there is no overlap in enabled FED Builders or trigger partitions, MiniDAQs can run in parallel to the global DAQ and each other
  - The global trigger is not available to MiniDAQs, they instead use a local trigger input via the Local Partition Manager (LPM)

MiniDAQ during Run 2 (2015-2018)

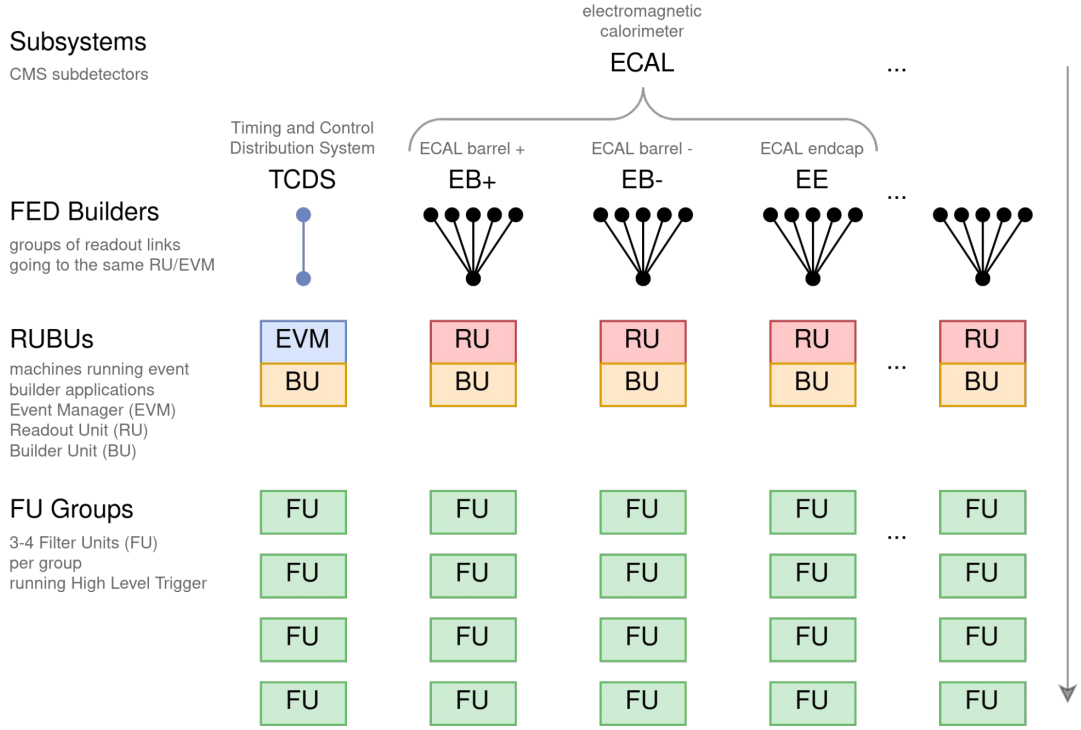


- **One MiniDAQ setup per subdetector (10 total)**
- **MiniDAQs are using machines separate from global DAQ resources**
  - Event Building: 1 RU (Readout Unit)/EVM (Event Manager) and 1 BU (Builder Unit) machine
  - HLT: 1 FU (Filter Unit) machine
- **Very limited performance compared to the central DAQ**
  - Most subsystems are unable to take data at the full event rate in their MiniDAQ
- **MiniDAQ resources are independent from the central DAQ**

MiniDAQ for Run 3 (since June 2022)

Differences compared to the Run 2 MiniDAQ:

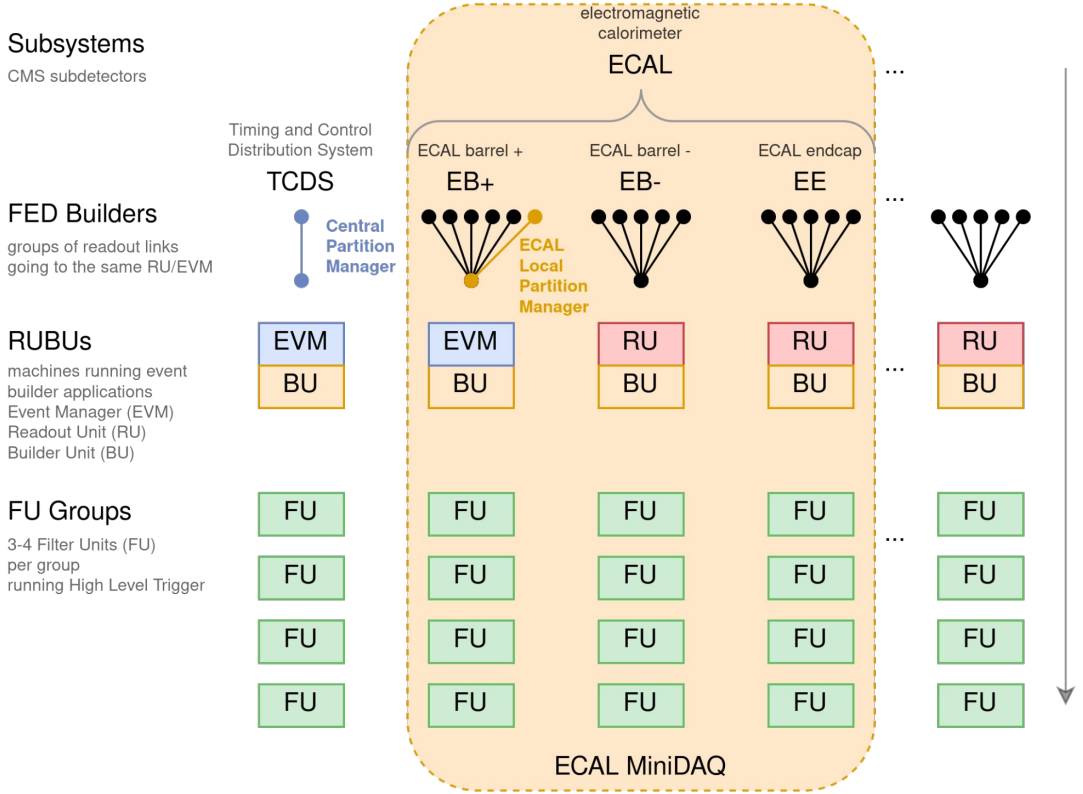
- **MiniDAQs share resources with the central DAQ**
- **Subsystem-specific FED Builders and the connected RUs, BUs and FUs are assigned to RUBU machines**



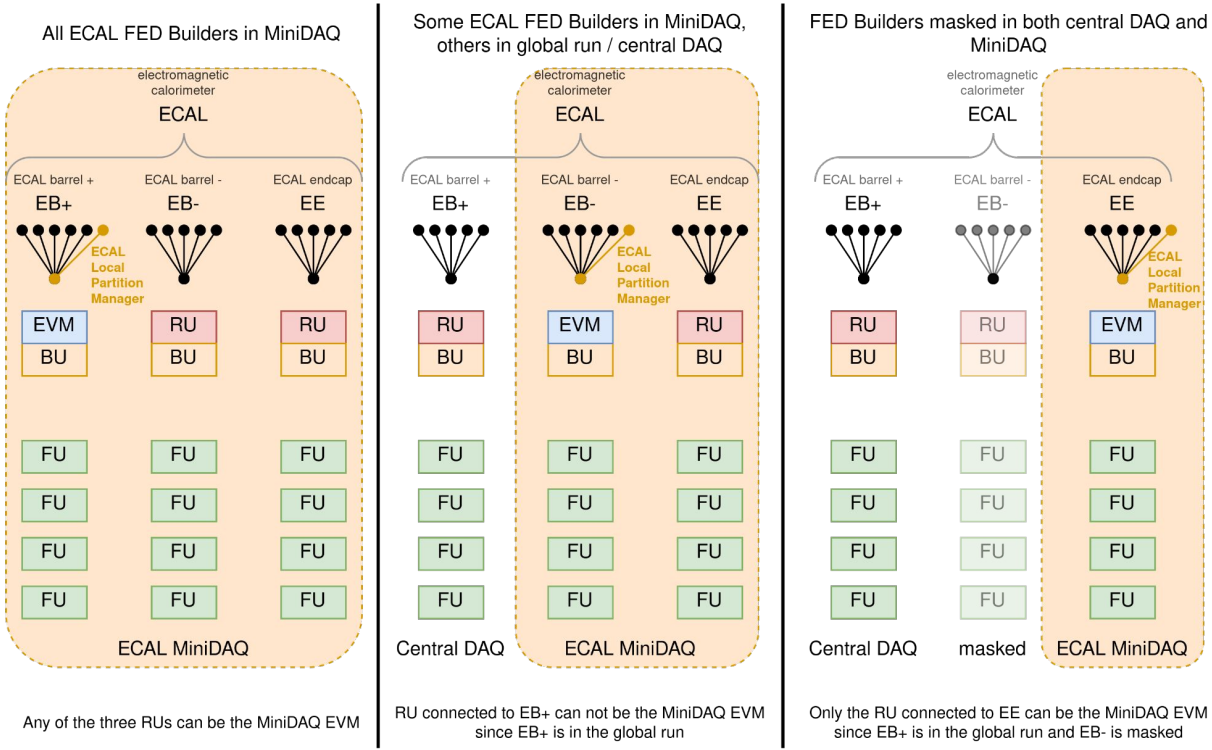
MiniDAQ for Run 3 (since June 2022)

Differences compared to the Run 2 MiniDAQ:

- **MiniDAQs share resources with the central DAQ**
- **Subsystem-specific FED Builders and the connected RUs, BUs and FUs are assigned to RUBU machines**
- **When a FED Builder is removed from the global data taking, the assigned DAQ resources (RUs, BUs, FUs) are removed from the central DAQ as well**
- **The assigned resources are available to be used in the subsystem's MiniDAQ**



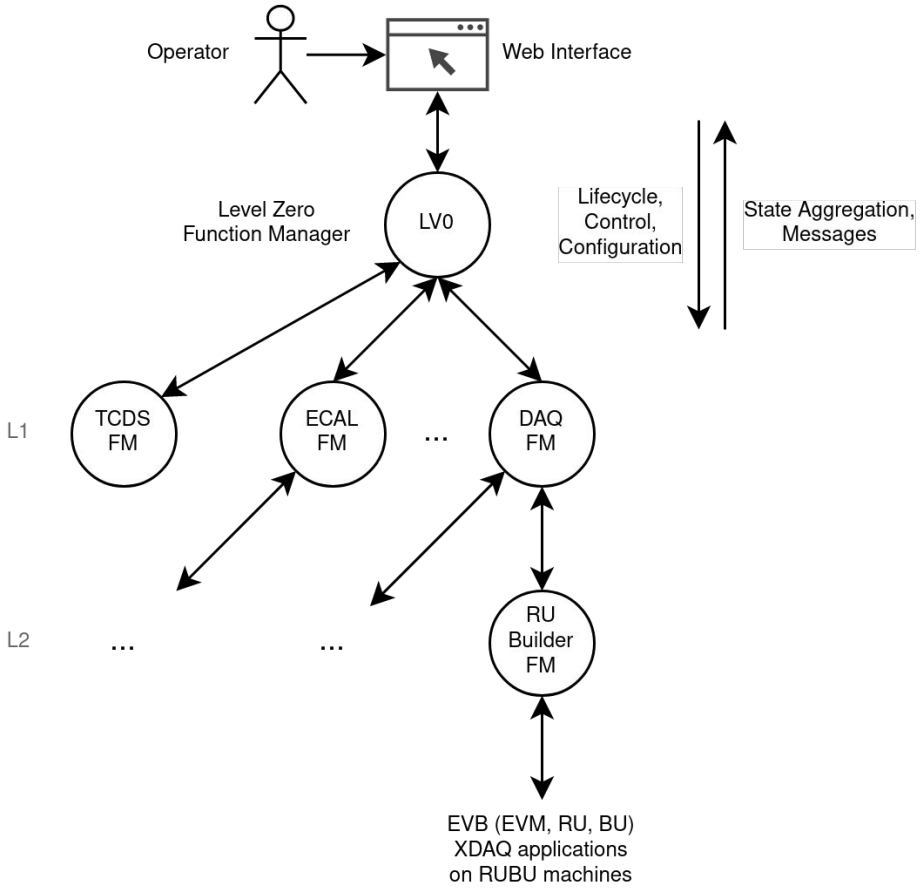
MiniDAQ for Run 3 – Event Manager



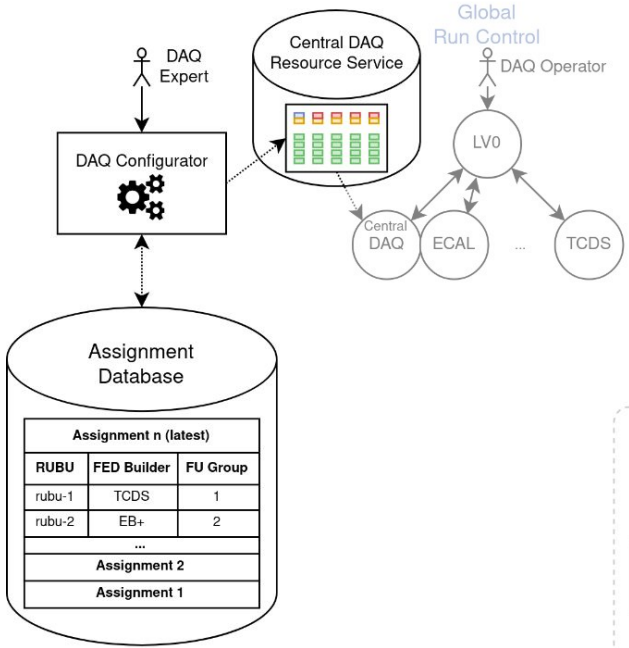
- **In every run, there needs to be one Event Manager (EVM)**
- **The Local Partition Manager of the subsystem is read out by the MiniDAQ's EVM**
- **A different Readout Unit might need to become the EVM, depending on which FED Builders are enabled in MiniDAQ**
- **Moving the EVM requires a different DAQ configuration**

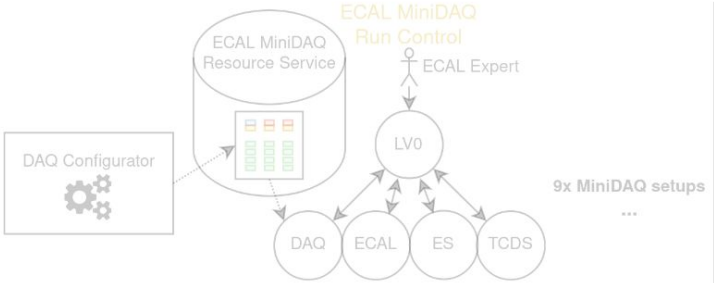
Run Control Hierarchy

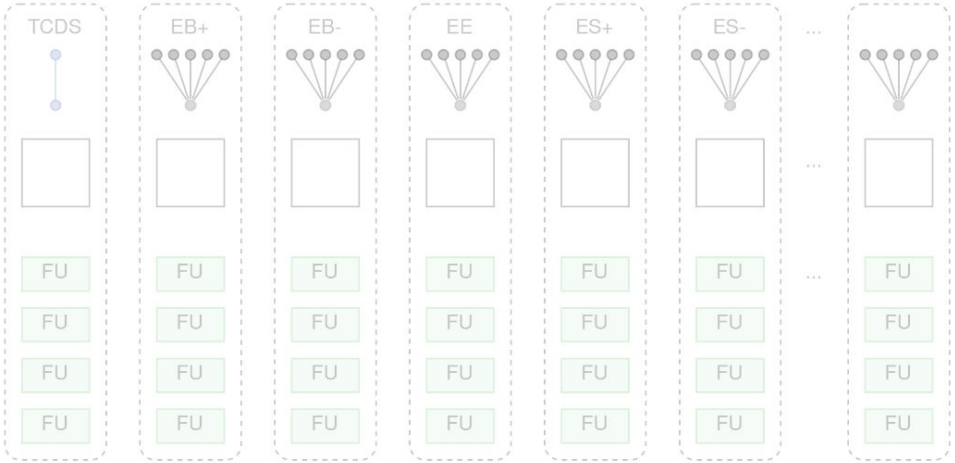
- Run Control nodes are organized in a hierarchical structure
- The operator controls the hierarchy using the web interface of the Level Zero node
- Nodes on the first level represent various subsystems, including subdetectors and the DAQ
- The DAQ node configures and controls the DAQ resources, including event building on the RUBU machines



MiniDAQ-3 – Resource Assignment



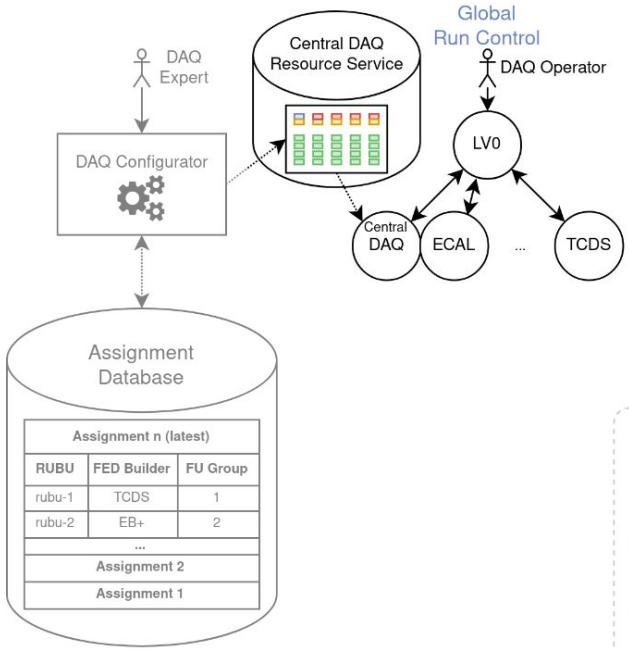


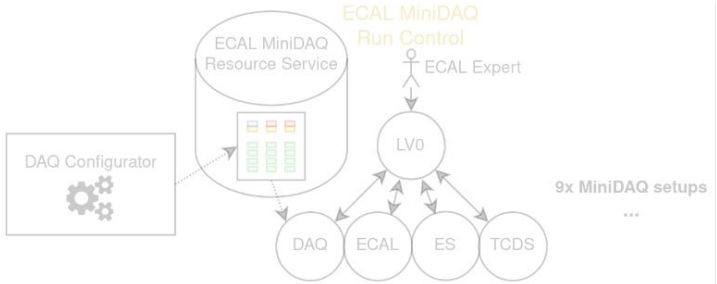


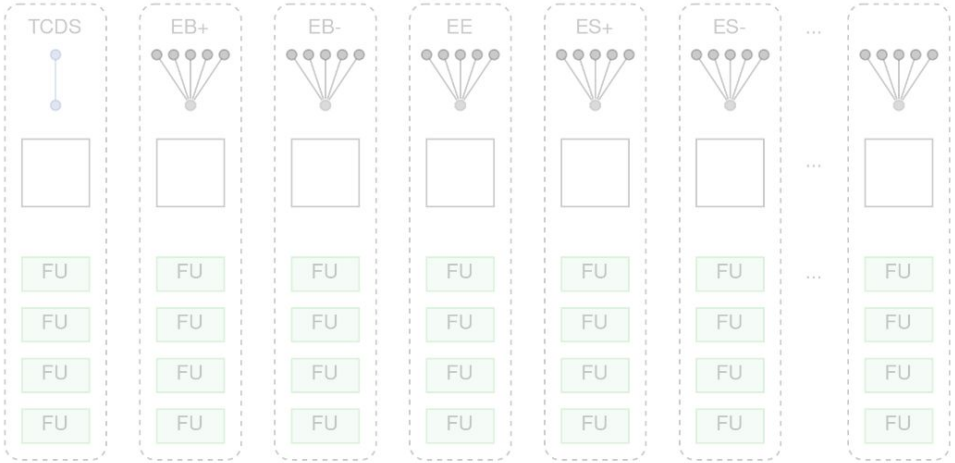
Using the Configurator tool, DAQ experts can:

- Modify the resource assignment
- Make new DAQ configurations

MiniDAQ-3 – Resource Assignment

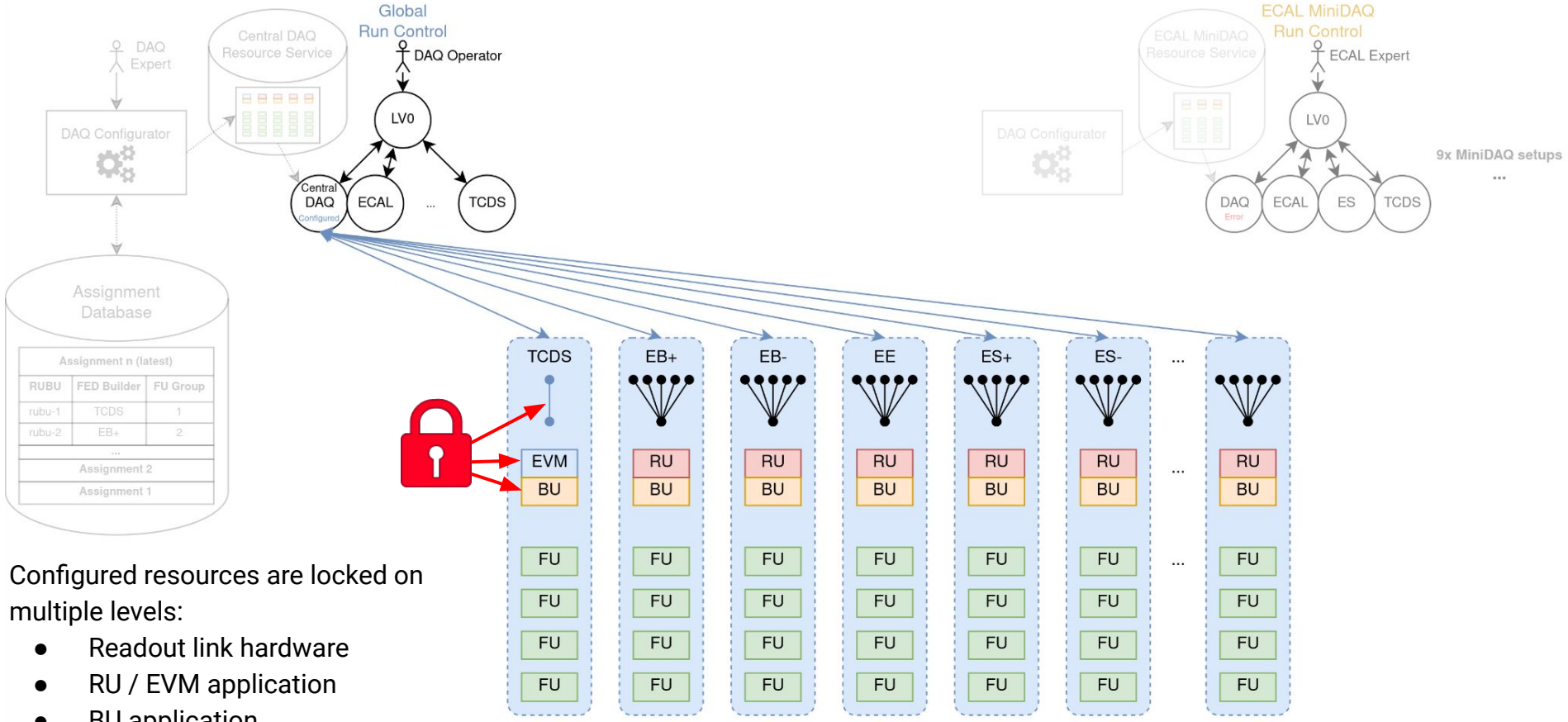






- The DAQ run control node loads the latest DAQ configuration from the resource service

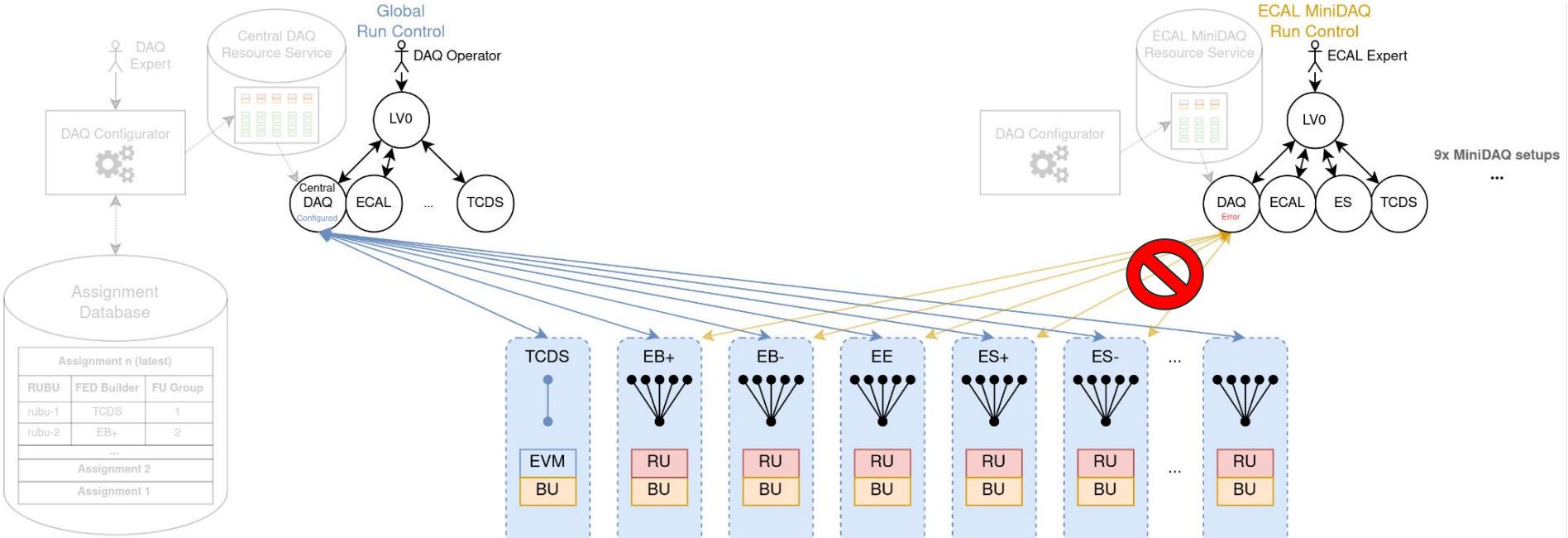
MiniDAQ-3 – Resource Locking



Configured resources are locked on multiple levels:

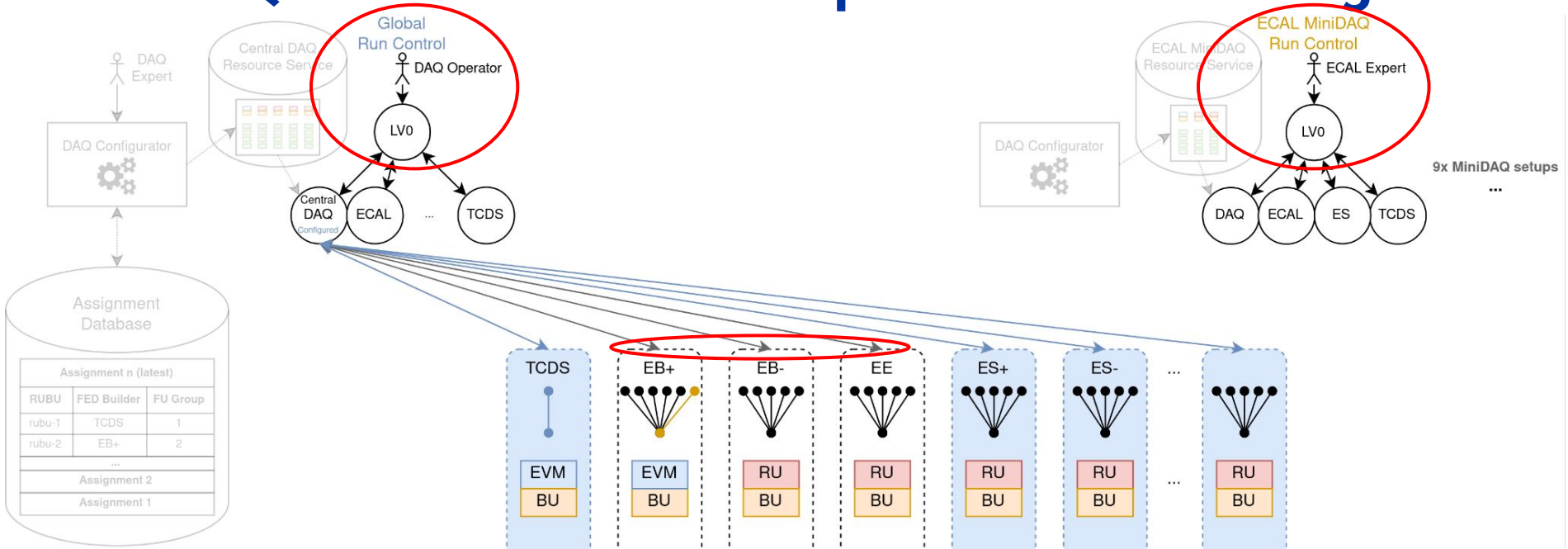
- Readout link hardware
- RU / EVM application
- BU application

MiniDAQ-3 – Resource Locking



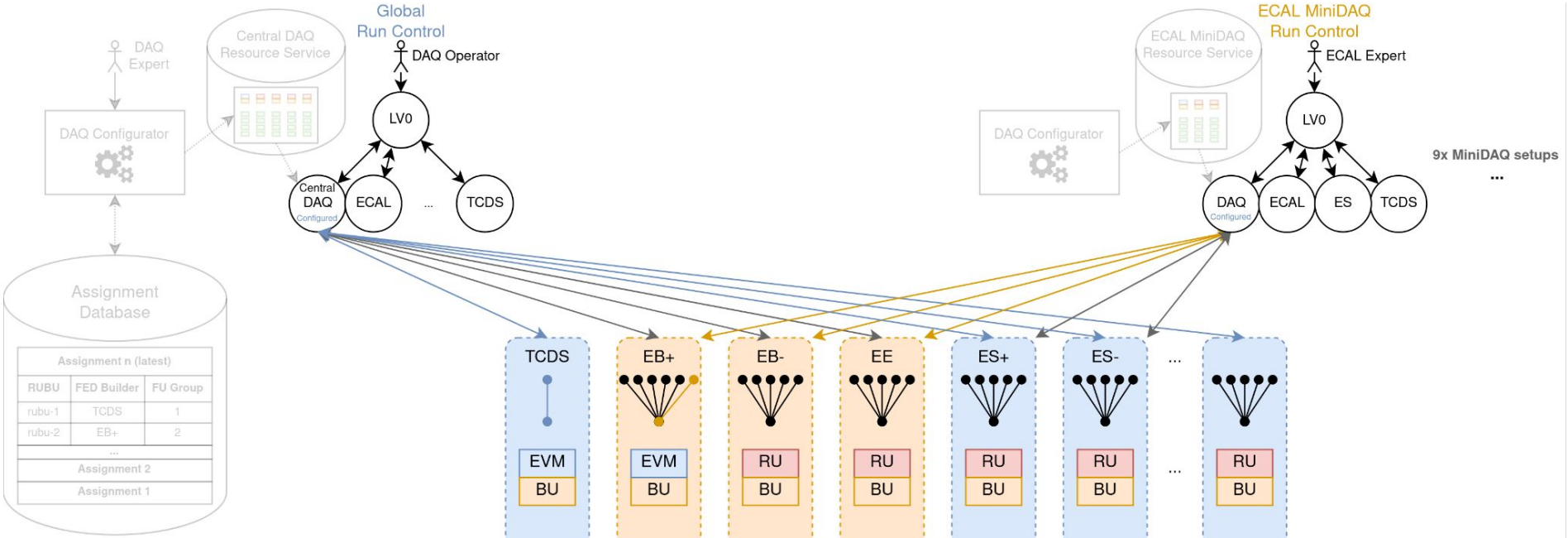
- MiniDAQs have their own run control hierarchy
- During configuration, the MiniDAQ can not lock already locked resources

MiniDAQ-3 – Concurrent independent data-taking



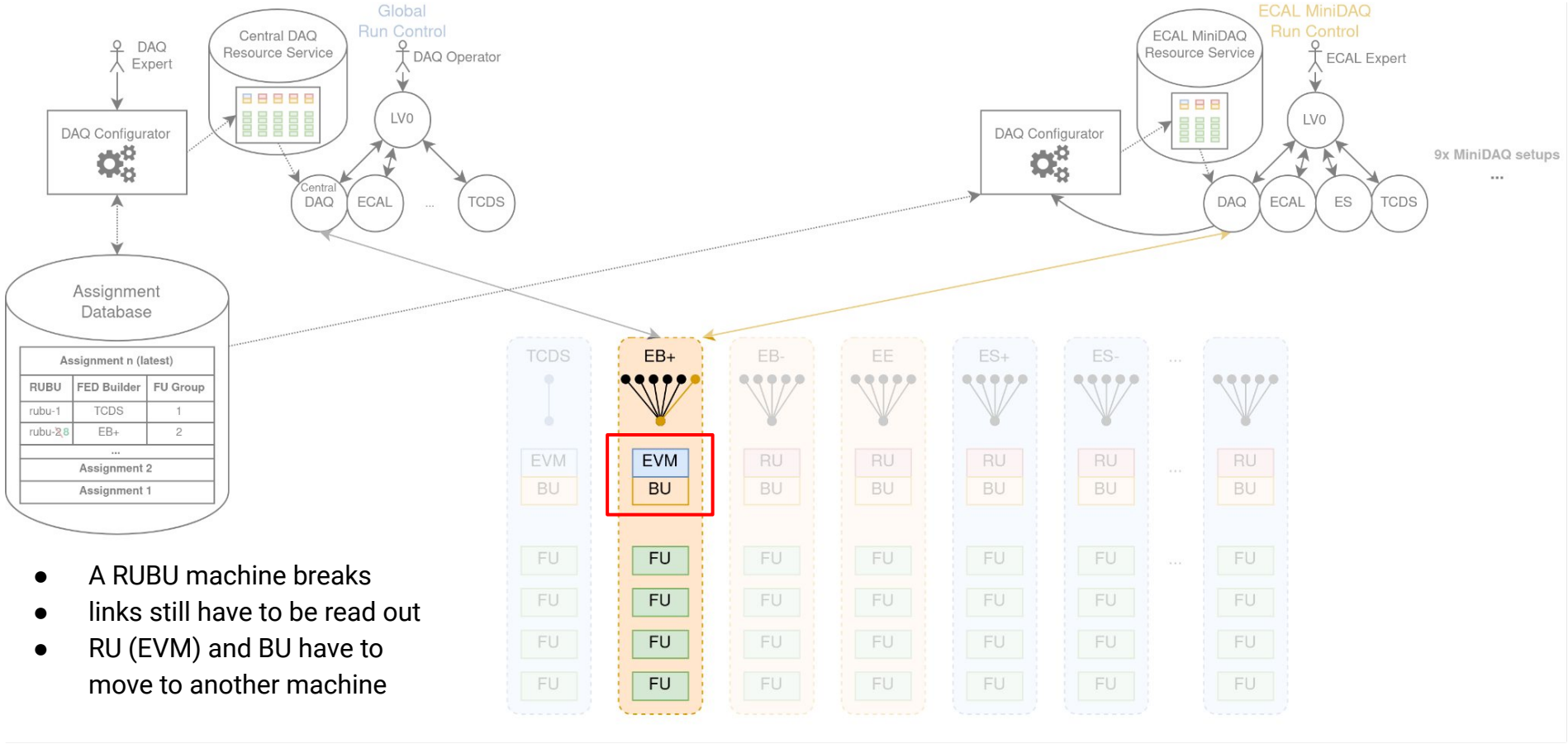
- Operators can mask FED Builders in the LV0 web interface
- Resources of masked FED Builders are not locked

MiniDAQ-3 – Concurrent independent data-taking



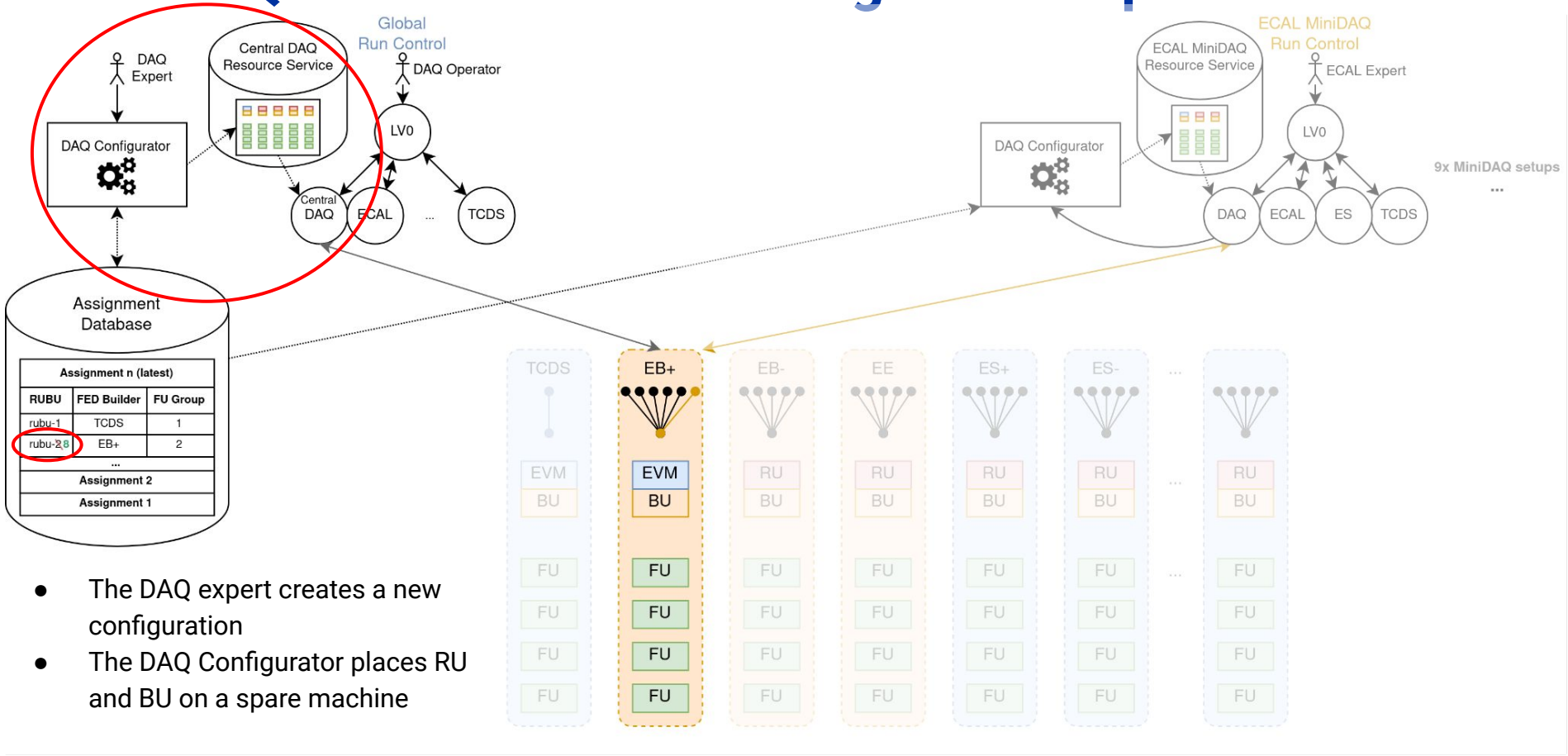
- If there is no overlap in FED Builders and trigger partitions, central DAQ and MiniDAQs can be configured concurrently

MiniDAQ-3 – Automated configuration updates



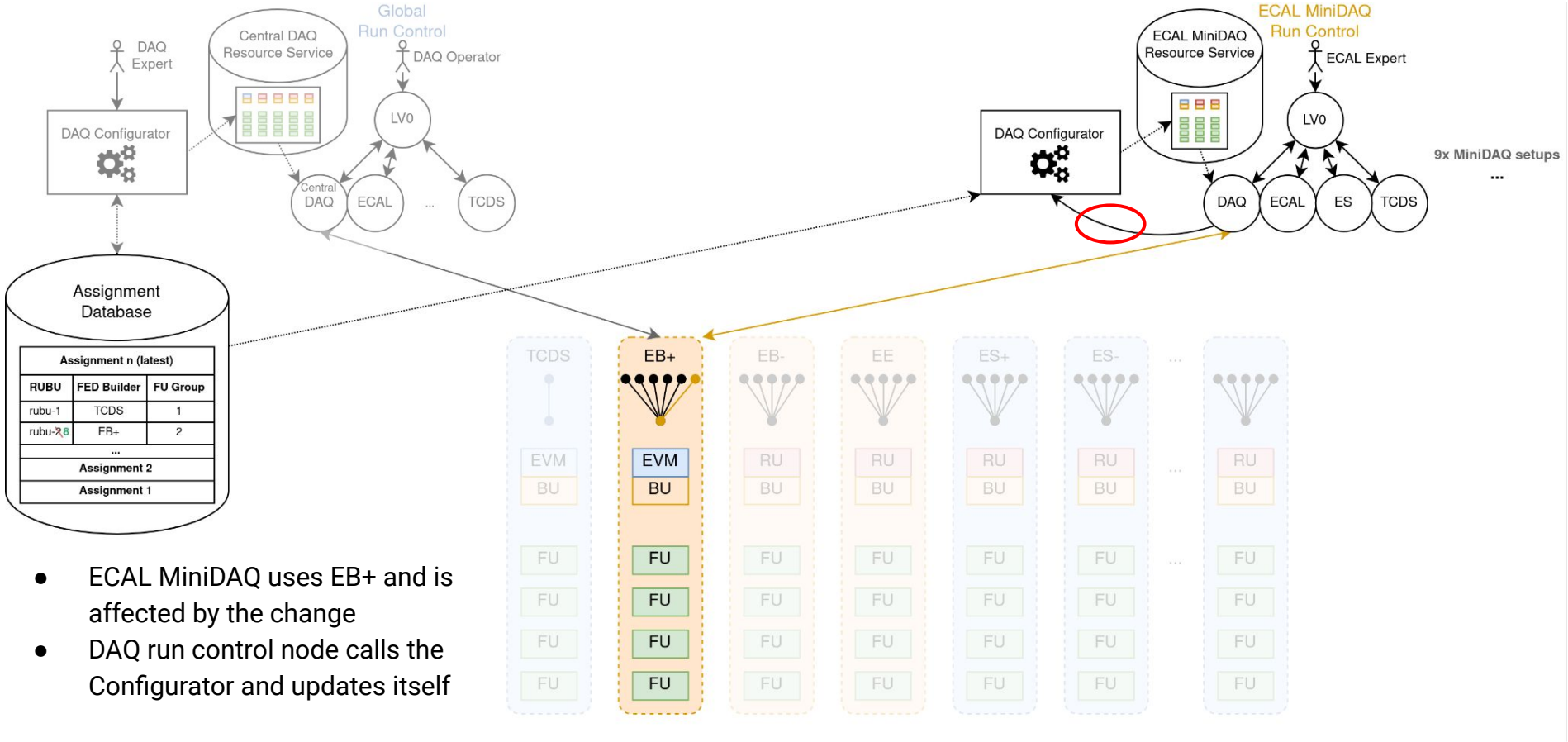
- A RUBU machine breaks
- links still have to be read out
- RU (EVM) and BU have to move to another machine

MiniDAQ-3 – Automated configuration updates



- The DAQ expert creates a new configuration
- The DAQ Configurator places RU and BU on a spare machine

MiniDAQ-3 – Automated configuration updates



- ECAL MiniDAQ uses EB+ and is affected by the change
- DAQ run control node calls the Configurator and updates itself

Operational experience during Run 3

- **Sharing production resources works as planned**
- **The automated configuration updates result in a greatly reduced workload for DAQ experts**

MiniDAQ for Run 3 - Summary

- **MiniDAQ allows subsystems to take independent runs with the full DAQ chain**
  - There are nine MiniDAQ setups available
- **With MiniDAQ-3, these setups share resources with the global DAQ**
  - Data taking is possible at a much higher rate, in many cases comparable to the rate possible with the central DAQ
  - Saves cost compared to the approach for MiniDAQ-2
  - The added complexity due to the need to synchronize the shared resources is covered by the automated configuration update procedures

Questions?

Authors

Sakulin, Hannes (CERN); **Brummer, Philipp (CERN)**

Co-Authors

Amoiridis, Vassileios (CERN); Behrens, Ulf (RICE); Bocci, Andrea (CERN); Branson, James (UCSD); Cano, Eric (CERN); Cittolin, Sergio (UCSD); Da Silva Almeida Da Quintanilha, Joao (CERN); Darlea, Georgiana-Lavinia (MIT); Deldicque, Christian (CERN); Dobson, Marc (CERN); Gigi, Dominique (CERN); Glege, Frank (CERN); Gomez-Ceballos, Guillelmo (MIT); Gutic, Neven (CERN); Hegeman, Jeroen (CERN); Izquierdo Moreno, Guillermo (CERN); Kartalas, Miltiadis (CERN); Krawczyk, Rafal (RICE); Li, Wei (RICE); Long, Kenneth (MIT); Meijers, Frans (CERN); Meschi, Emilio (CERN); Morovic, Srecko (UCSD); Orsini, Luciano (CERN); Paus, Christoph (MIT); Petrucci, Andrea (UCSD); Pieri, Marco (UCSD); Rabady, Dinyar Sebastian (CERN); Racz, Attila (CERN); Schwick, Christoph (CERN); Simelevicius, Dainius (Vilnius University); Vazquez Velez, Cristina (CERN); Zejdl, Petr (CERN); Zogatova, Dominika (CERN); Zhang, Yousen (RICE)

Presentations at CHEP2023

CMS Data Acquisition

May 9, 11:30 AM – Track 2 – Andrea Petrucci
First year of experience with the new operational monitoring tool for data taking in CMS during Run 3

May 9, 4:30 PM – Track 2 – Rafal Krawczyk
Event Building studies for CMS Phase-2 at CERN

May 11, 12:00 PM – Track 2 – Dainius Simelevicius
Towards a container-based architecture for CMS data acquisition

May 11, 12:15 PM – Track 2 – Philipp Brummer
MiniDAQ-3: providing concurrent independent subdetector data-taking on CMS production DAQ resources

CMS High Level Trigger

Poster Session – Abdulla Mohamed
Optimising the configuration of the CMS GPU reconstruction

May 9, 2:30 PM – Track X – Ganesh Parida
Run-3 Commissioning of CMS Online HLT reconstruction using GPUs

May 9, 5:00 PM – Track 2 – Andrea Bocci
Adoption of the alpaka performance portability library in the CMS software

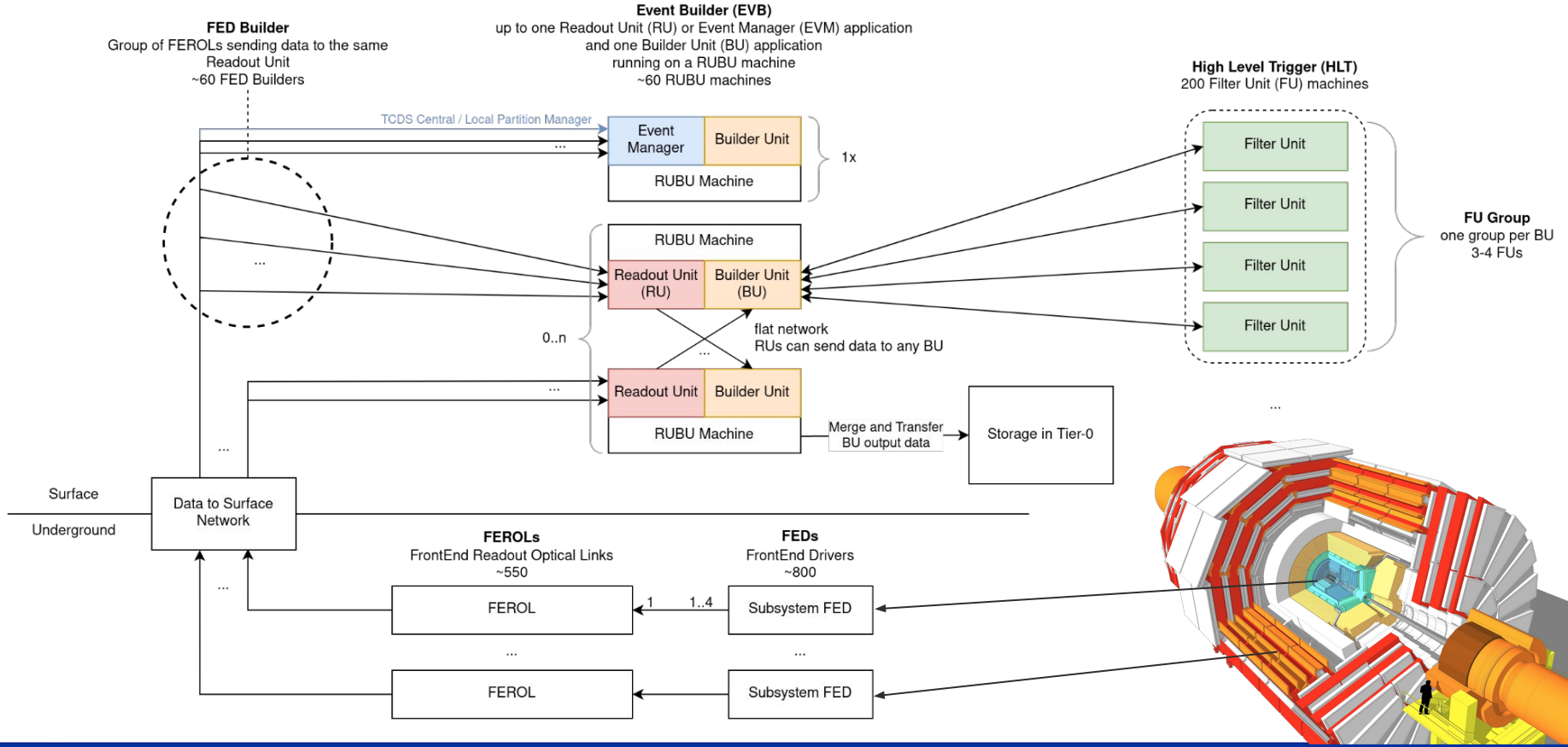
Backup

Structure of this presentation

- Overview of the CMS Data Acquisition for Run 3 (DAQ-3)
- What are MiniDAQs and why do we need them?
- MiniDAQs in Run 2 and changes for Run 3
- Operational experience with MiniDAQs during Run 3 (since mid-2022)

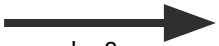
CMS DAQ-3 overview (simplified)

Cutaway diagram of the CMS detector
Tai Sakuma and Thomas McCauley, "Detector and Event Visualization with SketchUp at the CMS Experiment" *J. Phys.: Conf. Ser.* 513 022032, doi:10.1088/1742-6596/513/2/022032 - http://cds.cern.ch/record/2665537/



MiniDAQ for Run 3 - Static Assignment

RUBU Machine	FED Builder RU configuration	FU Group BU configuration
rubu-1.cms	TCDS	FUGroup1
rubu-2.cms	EB+	FUGroup2
rubu-3.cms	EB-	FUGroup3
rubu-4.cms	EE	FUGroup4
rubu-5.cms	*<spare>*	*<spare>*

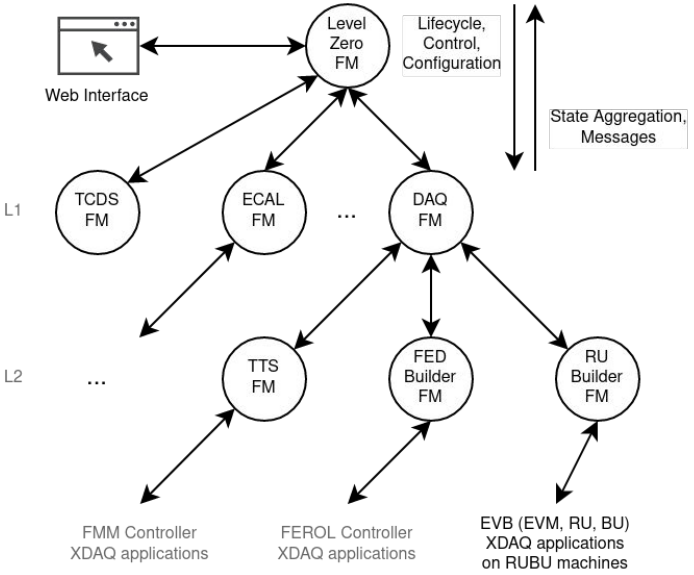


rubu-3.cms disabled due to hardware defect

RUBU Machine	FED Builder RU application	FU Group BU configuration
rubu-1.cms	TCDS	FUGroup1
rubu-2.cms	EB+	FUGroup2
rubu-3.cms	*<blocklisted>*	*<blocklisted>*
rubu-4.cms	EE	FUGroup4
rubu-5.cms	EB-	FUGroup3

- **RUBU machines might need to be excluded from the run, for example due to hardware defects, making a new DAQ configuration necessary**
- **When a RUBU machine is removed from the configuration, assigned Readout Units and Builder Units and the connected FED Builders and FU Groups, need to be assigned to a different RUBU machine**
- **Static Assignment minimizes the changes when generating a new configuration, keeping existing assignments where possible**
- **Minimizing changes in assignment keeps synchronization efforts between central DAQ and MiniDAQs to a minimum and also simplifies debugging of previous configurations**

Run Control – Hierarchy

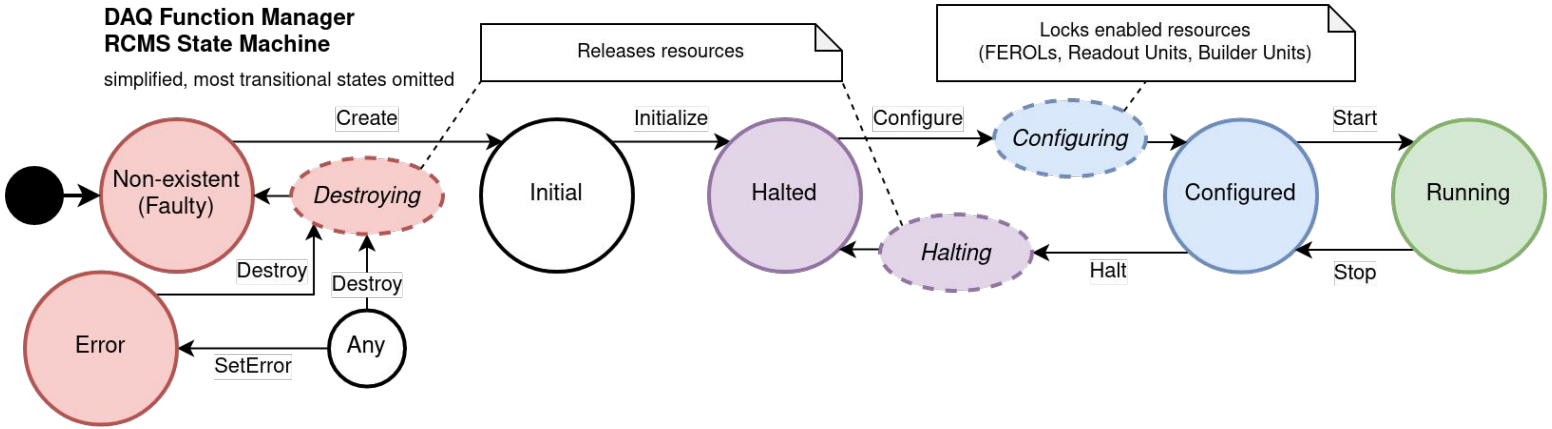


- Subsystems (subdetectors, DAQ, TCDS, …) are organized in a hierarchy, represented by Function Managers (FM)
- Function Managers (FMs) control child resources, which can be other FMs, XDAQ applications etc.
- The top-level FM is called the Level Zero FM (LV0) and is used by the operators to control CMS data taking
- Run Control FMs manage the lifecycle of their child resources and send state machine inputs and configurations to them
- A parent FM usually aggregates the state of its child resources
- The DAQ FM's child FMs control the Event Builder applications on the RUBU machines, as well as FEROL and FMM controllers

MiniDAQ setups have a separate LV0 and their own child FM configurations. Currently, there are nine MiniDAQ setups.

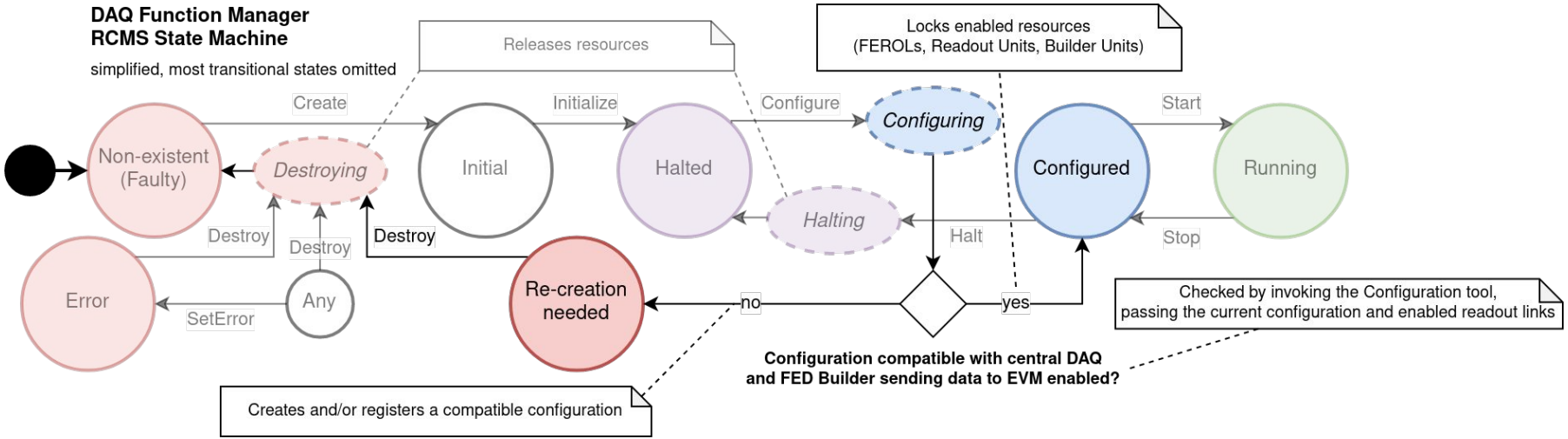
Setup	Subsystems
Global Run (central DAQ)	PIXEL, TRACKER, ES, ECAL, HCAL, DT, GEM, CSC, RPC, CTPPS_TOT, CTPPS, TCDS, TRG, L1SCOUT, DAQ, DQM, DCS
ECAL MiniDAQ	ECAL, ES, TCDS, DAQ
GEM MiniDAQ	GEM, TCDS, DAQ
…	…

Synchronization – resource locking



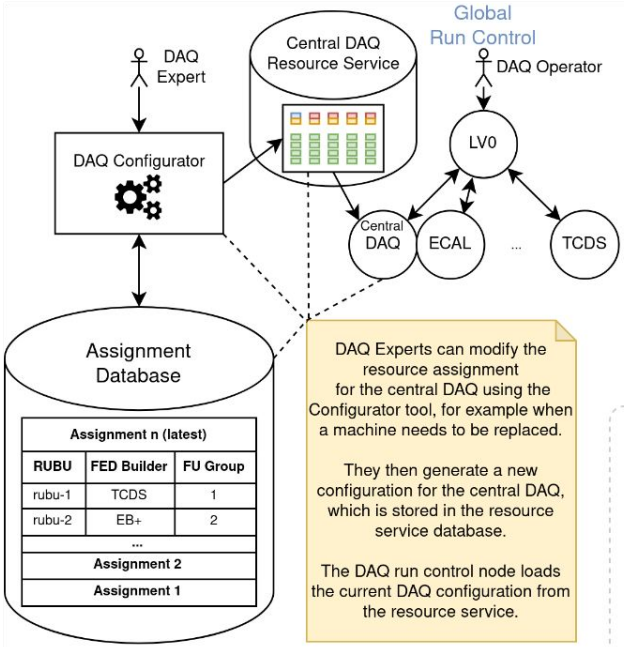
- Since resources are shared between the central DAQ and MiniDAQs, synchronization between the setups is required
- FED Builders, Readout Units, Builder Units and FU Groups are either configured in the central DAQ or in a MiniDAQ; locks ensure that once they are configured, they have to be released before they can be configured in another setup
- Hardware locks on the level of the FEROLs ensure that a FED Builder can only be configured to send data to one Readout Unit
- Locks on the level of the Event Builder ensure that at most one Readout Unit and Builder Unit can be configured on each RUBU machine
- When a change in assignment is required, the assignment of the central DAQ for the global run is changed; MiniDAQs follow this global assignment

Synchronization – MiniDAQ-3 configuration checks



- **When the DAQ FM of a MiniDAQ setup is being configured, it invokes the DAQ configuration tool to check if the current configuration is compatible with the central DAQ and if the FED Builder sending data to the Event Manager (EVM) is enabled**
  - if yes, the system configures (assuming the resources are not locked by the central DAQ), and a run can be started
  - if not, a compatible configuration is generated and/or registered by the configuration tool and the MiniDAQ has to be re-created to pick it up
- **For changes outside of the assignment and EVM placement, a script is provided that allows DAQ experts to update all MiniDAQ configurations, using information managed in Git**

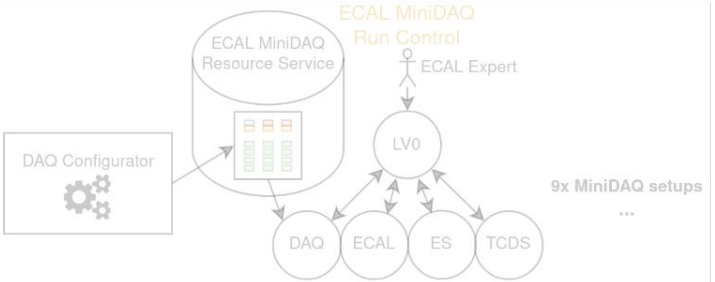
MiniDAQ-3 – Resource Assignment

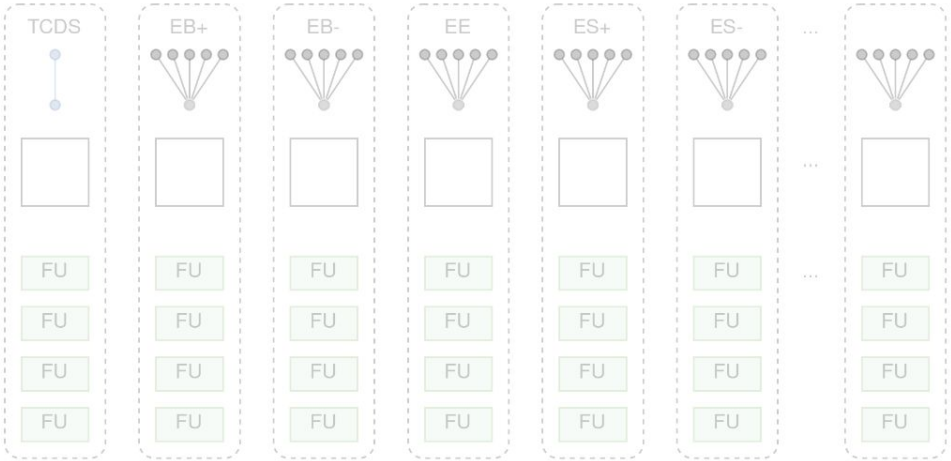


DAQ Experts can modify the resource assignment for the central DAQ using the Configurator tool, for example when a machine needs to be replaced.

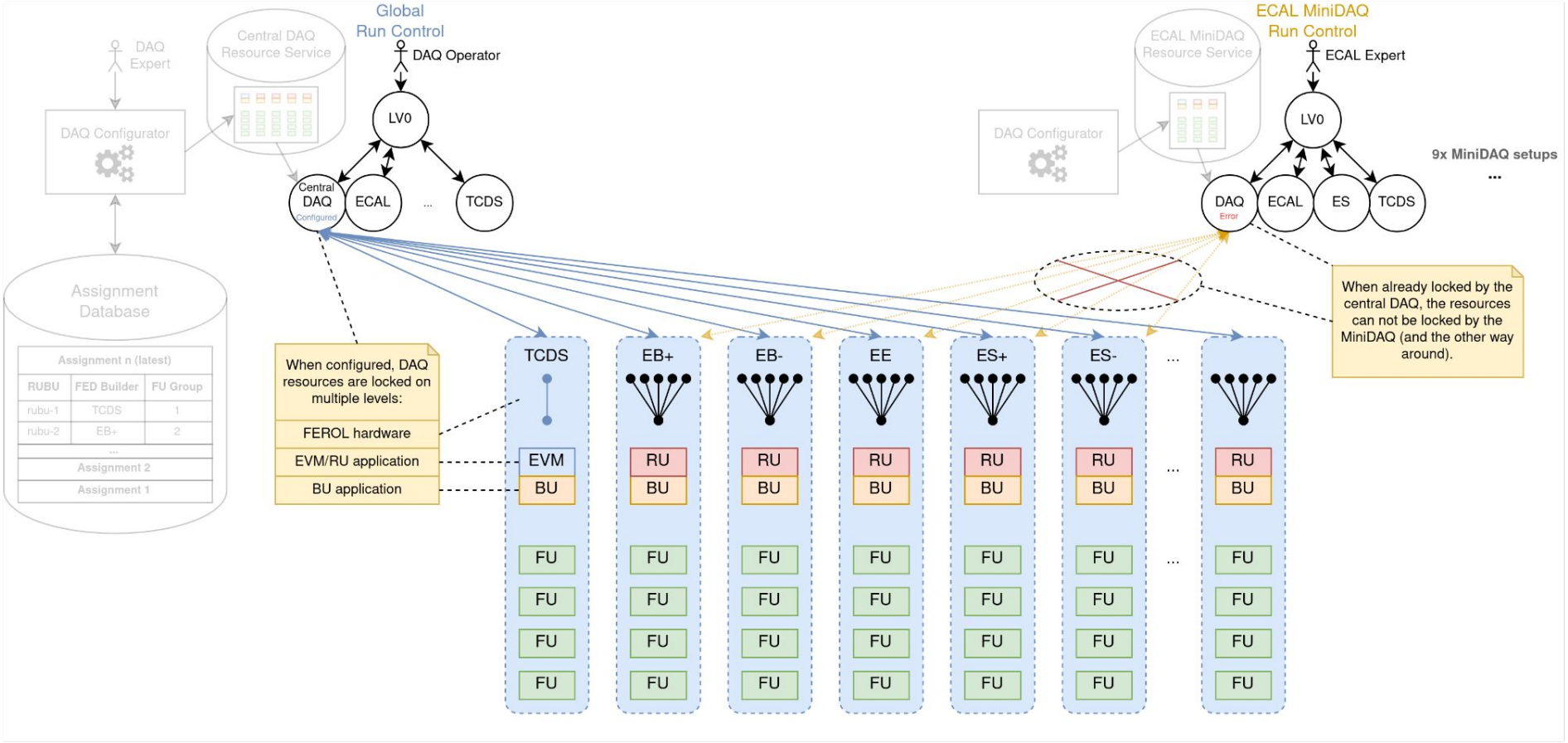
They then generate a new configuration for the central DAQ, which is stored in the resource service database.

The DAQ run control node loads the current DAQ configuration from the resource service.

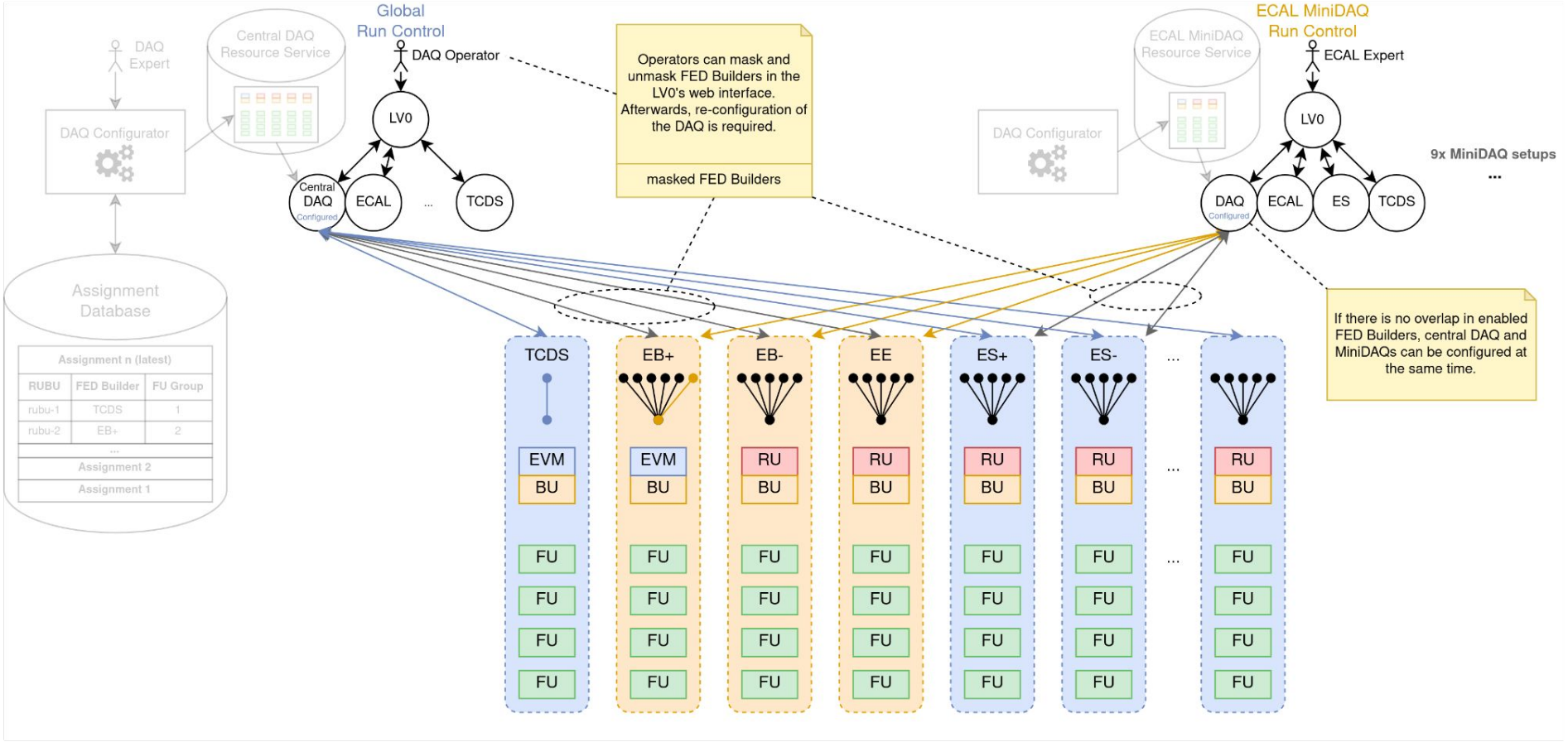




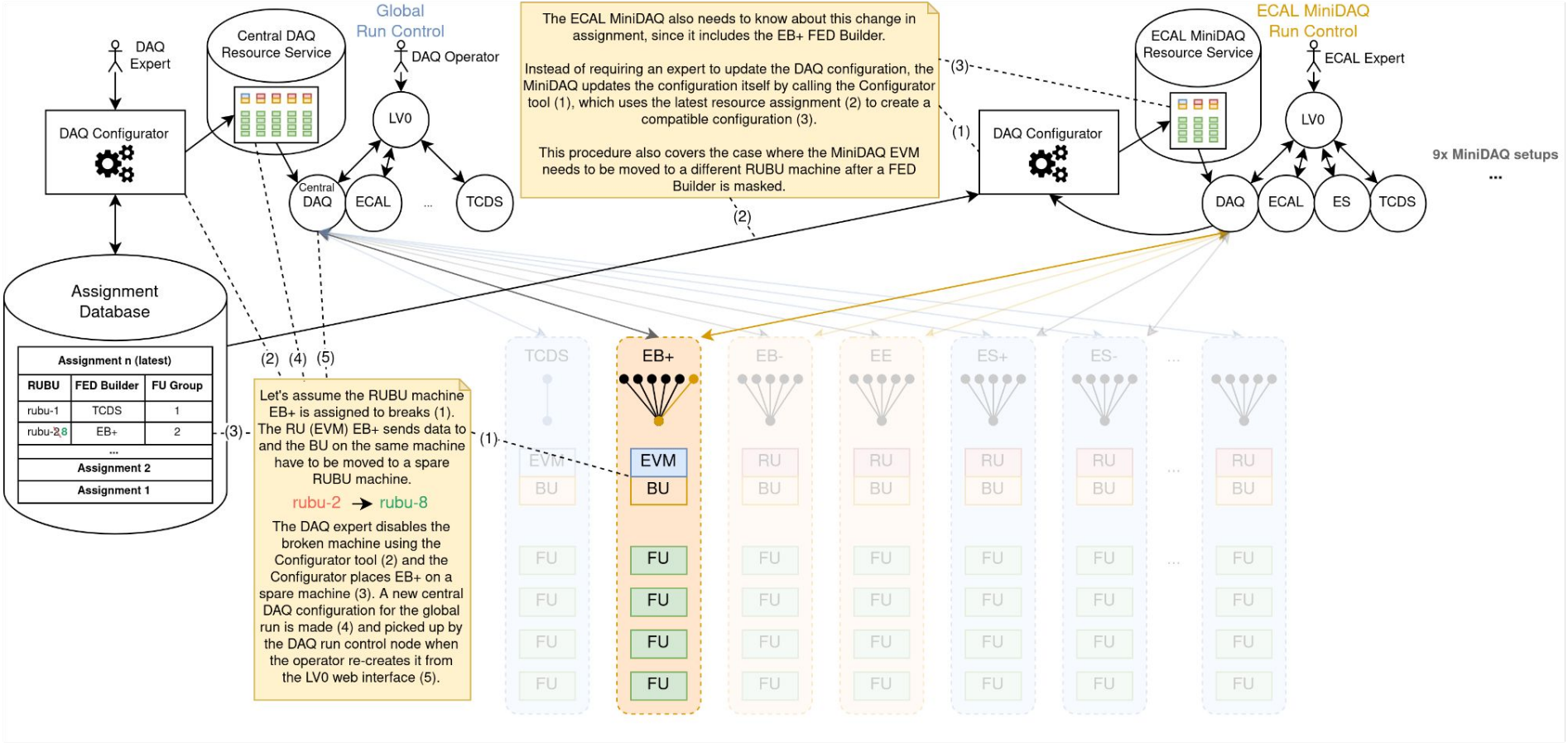
MiniDAQ-3 – Resource Locking



MiniDAQ-3 – Concurrent independent data-taking



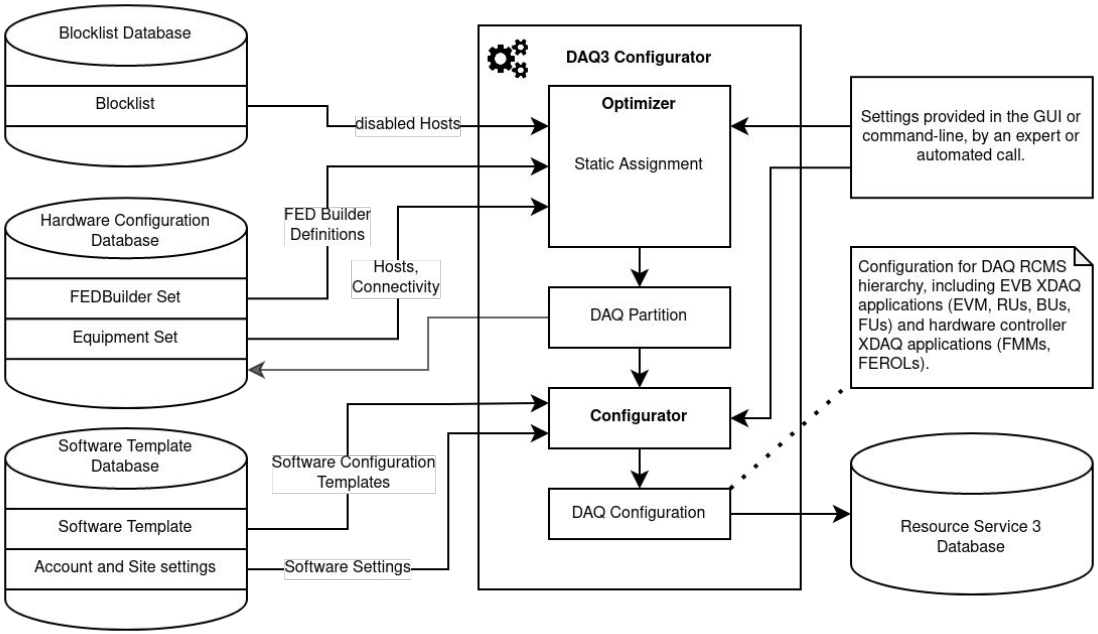
MiniDAQ-3 – Automated configuration updates



Operational experience during Run 3

- **Sharing production resources works as planned**
- **The automated configuration updates result in a greatly reduced workload for DAQ experts**
- **During global high-rate test runs with just a small number of subsystems enabled, there can be backpressure from the DAQ due to a lack of resources (Builder Units and Filter Units)**
  - This is because removing a subsystem from the global run also removes its assigned resources from the central DAQ, as they need to be available to be used in MiniDAQ
  - The issue can be avoided by creating a special DAQ configuration in preparation for these tests

MiniDAQ for Run 3 - Configuration Workflow



- **Generated DAQ Configurations are registered in the Global Configuration Map (GCM)**
- **The DAQ RCMS node picks up the registered configuration on creation**

MiniDAQ setups

- **Nine setups for different subdetectors**
- **The number of FED Builders equals the number of possible EVM placements and therefore configurations that can be switched between by the automated procedures**

MiniDAQ	**number of FED Builders = number of configurations for the same assignment**
TRIGGER	2
PIXEL	10
ECAL	5
HCAL	9
CSC	3 (including GEM)
RPC	1
DT	1
GEM	1
TOTEMCTPPS	2

About me

- **M.Sc. in Computer Science**
- **Junior Fellow at CERN since June of 2021**
  - At CERN as Technical Student in 2015 and occasionally as User
  - Computer Science Bachelor Thesis at CERN in 2018
- **In the CMS Data Acquisition (DAQ) group**
- **Working on the Run Control and Monitoring System (RCMS)**
- **E-Mail: philipp.brummer@cern.ch**