SIEMENS

CERN openlab

SCADA Evolution and Data Analytics

CERN openlab Technical Workshop

BE-ICS

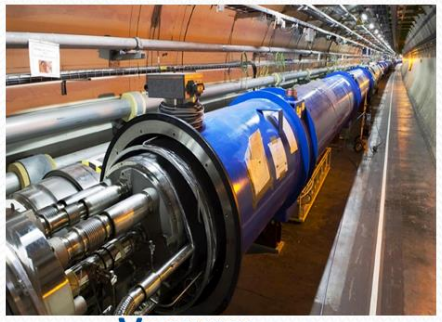
Piotr Golonka, Jakub Guzik, Anthony Hennessey, Rafal Kulaga, Filippo Tilaro, Fernando Varela (BE-ICS)

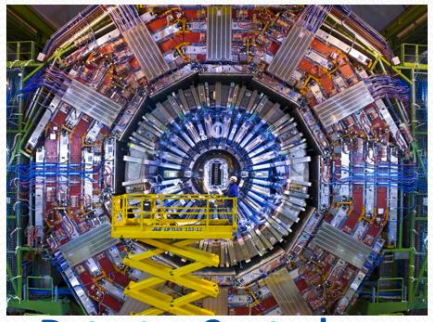
in collaboration with Siemens/ETM

23/01/2019

Hundreds of WinCC OA SCADA systems at CERN: a major challenge for archiving


Cooling & Ventilation


Vacuum


Detector Controls


Electric Grid


Cryogenics


Gas Distribution


Environment & Radiation

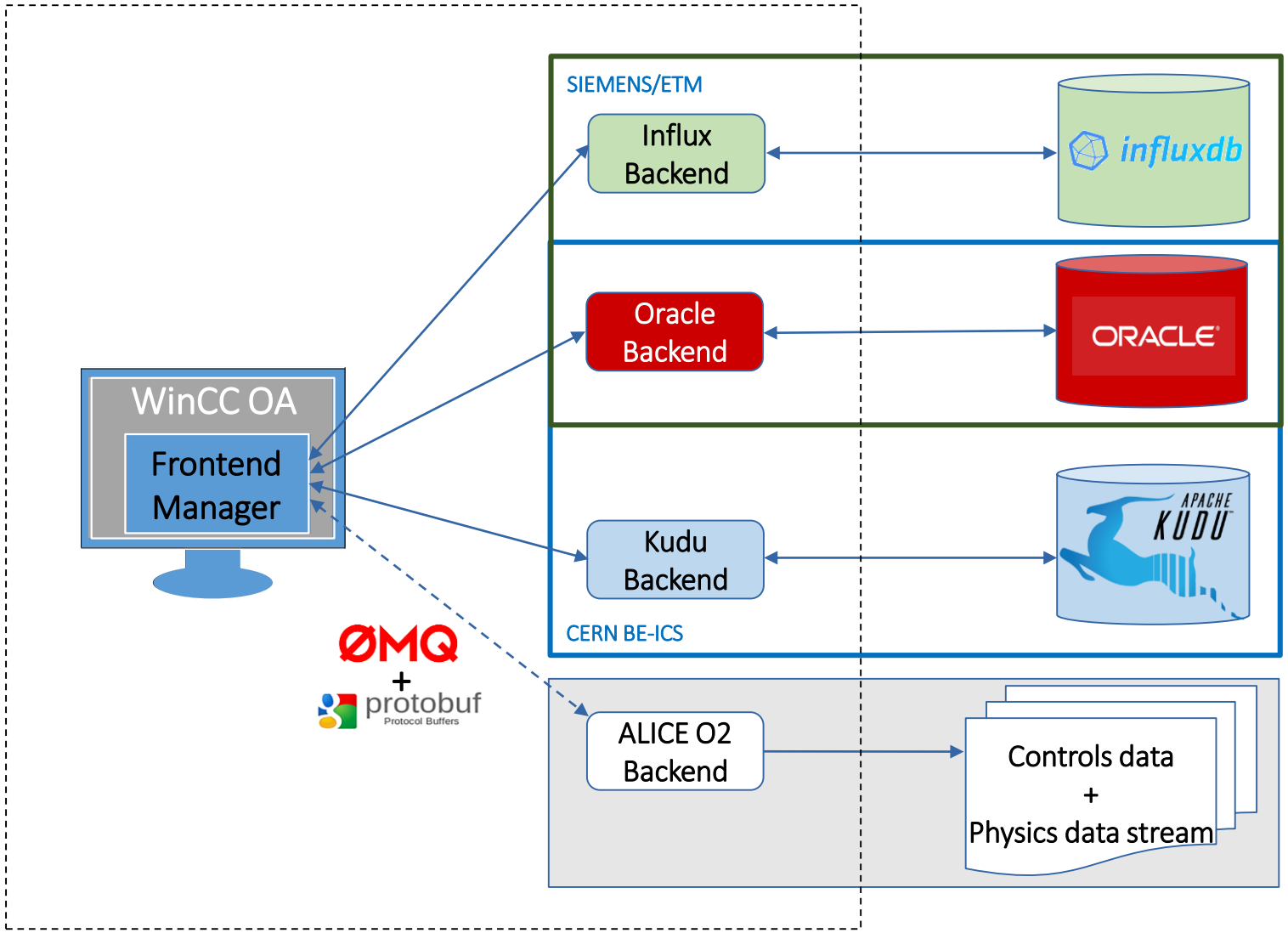

Interlocks and Safety

Requirements: reliability, performance and scalability, openness, new use cases

Data rates will increase in the HiLumi LHC

Need for a new, future-proof archiver with support for multiple database technologies

NextGeneration Archiver for WinCC OA



- Frontend Manager (ETM)
  - Connects to a WinCC OA system
  - Exposes ZMQ + protobuf API for backends
- InfluxDB Backend (ETM)
  - Both local and centralized InfluxDB installations supported
- Oracle Backend (CERN/ETM)
  - Compatible with the current archiver
- Apache Kudu Backend (CERN)
  - Potential next database technology for use at CERN (evaluation ongoing)
  - Enables data analytics
- Open architecture – custom backends (e.g. ALICE O2)

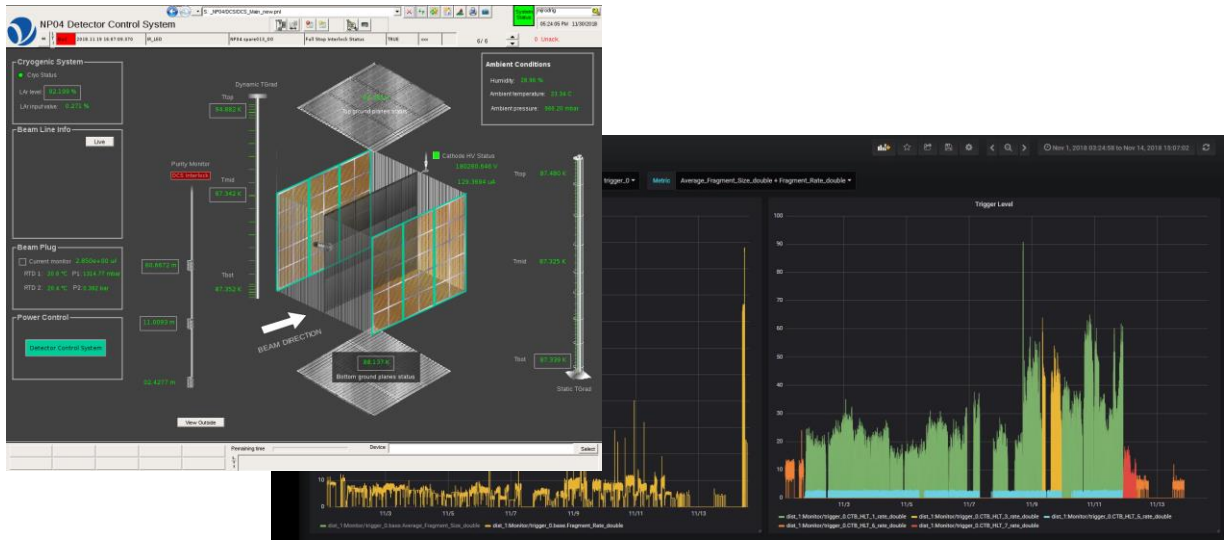
CERN openlab SIEMENS

Project status

- Frontend
  - Most important features already implemented
  - Performance improvements in progress (based on results of large-scale tests)
  - Release for early adopters in Spring
- InfluxDB Backend
  - The most feature-complete backend at the moment
  - Release for early adopters in Spring
- Oracle Backend
  - A bit behind InfluxDB Backend in terms of functionality
  - Takeover of work by the new openlab Fellow
  - Migration and rollback scenarios well understood and in test
  - Production-grade version required in May by ALICE (production: September)
- Apache Kudu Backend
  - Development currently on hold (focus on Oracle Backend)
  - Pending performance evaluation of Kudu

Early adopters at CERN

- ALICE
  - Oracle Backend and custom O2 Backend streaming changes to include them in the physics data stream
  - First tests of the O2 Backend successful
  - Preparing for large-scale tests in ALICE infrastructure



- protoDUNE
  - Making archived data available online in Grafana
  - InfluxDB Backend used in selected systems

- Streaming of data from WinCC OA to external applications (e.g. through Apache Kafka) – interest from the radiation monitoring group

Large-scale system tests at CERN

- Goal: evaluate the NextGeneration Archiver at the scale of CERN systems (hundreds of interconnected systems)
  - Write and query performance of different backends
  - Long-term reliability: no resource leaks or performance degradation
  - Handling of event avalanches
  - Recovering from network issues and database disconnections
  - Upgrade and rollback procedures for existing Oracle schemas

- First results have already influenced the readout architecture of the NGA

Industrial Data Analytics

Machine Learning and Big **Data Analytics** to build **smarter control systems**

- **Multiple Benefits** proven by R&D activities in 6 years of openlab collaboration with Siemens:
  - **Extend the monitoring capabilities of the control systems**
  - **Reduce operational and maintenance costs**
- **+40 use-cases** identified and many in progress:
  - LHC circuit monitoring, Linac3 beam source optimization, electron-cloud heat in Cryogenics, vacuum leak detection, Linac4 accelerator…

Control System

Data Analytics

Smart Data for Industrial Control Systems

2 Different groups of data analytics activities

Use-Cases and algorithms

Design and development of data analytics algorithms to match use-case requirements

- Expert system / condition monitoring
  - LHC Circuit Monitoring
  - Condition Monitoring for Cryogenics
  - Analysis of control systems alarms based on KPIs
- Machine Learning
  - Leak detection in cooling & ventilation
  - Linac3 optimization

Analytical Platforms

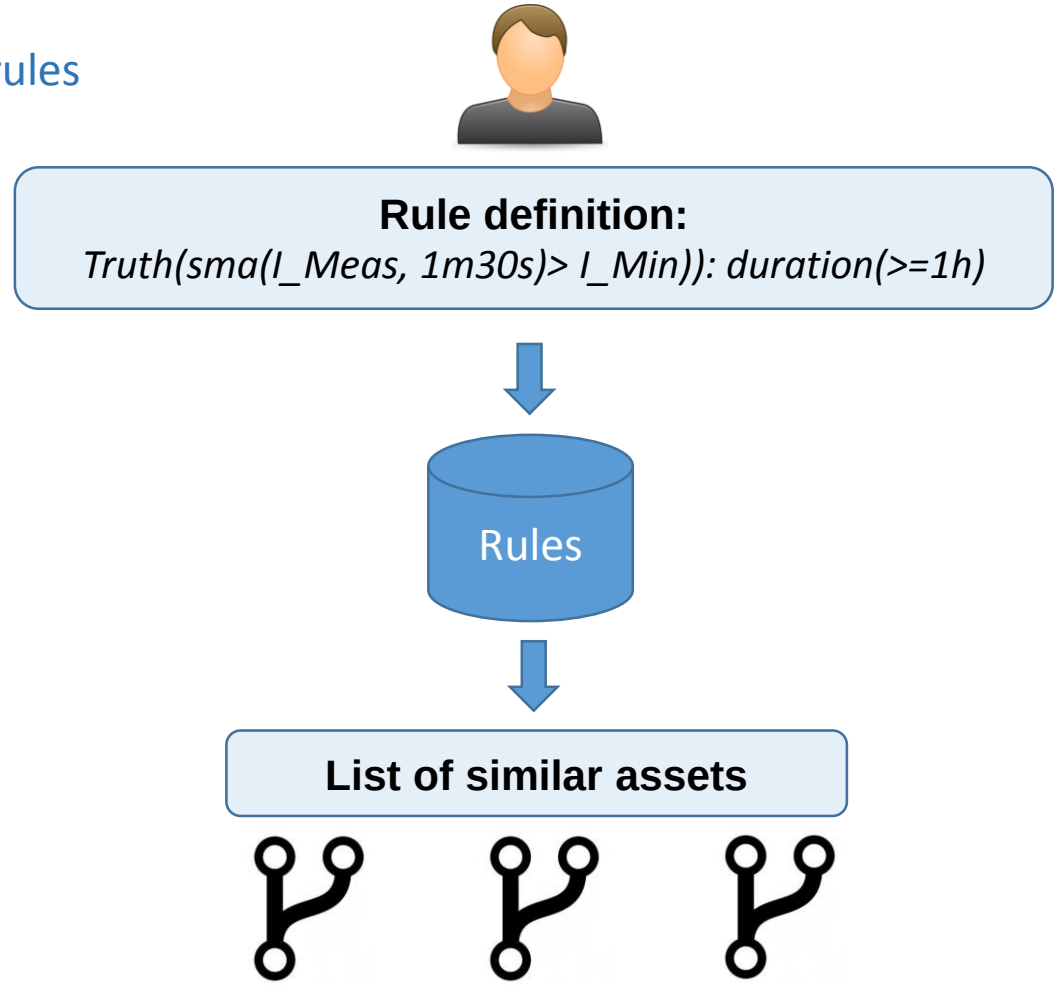
Design, development and evaluation of the data analytics platform for control systems

Siemens platforms:

- Smart Industrial IoT (Smart IIoT)
- Signal Event Processing Language (SEPL) and SEPLab
- PeregrineDB

Condition monitoring analysis

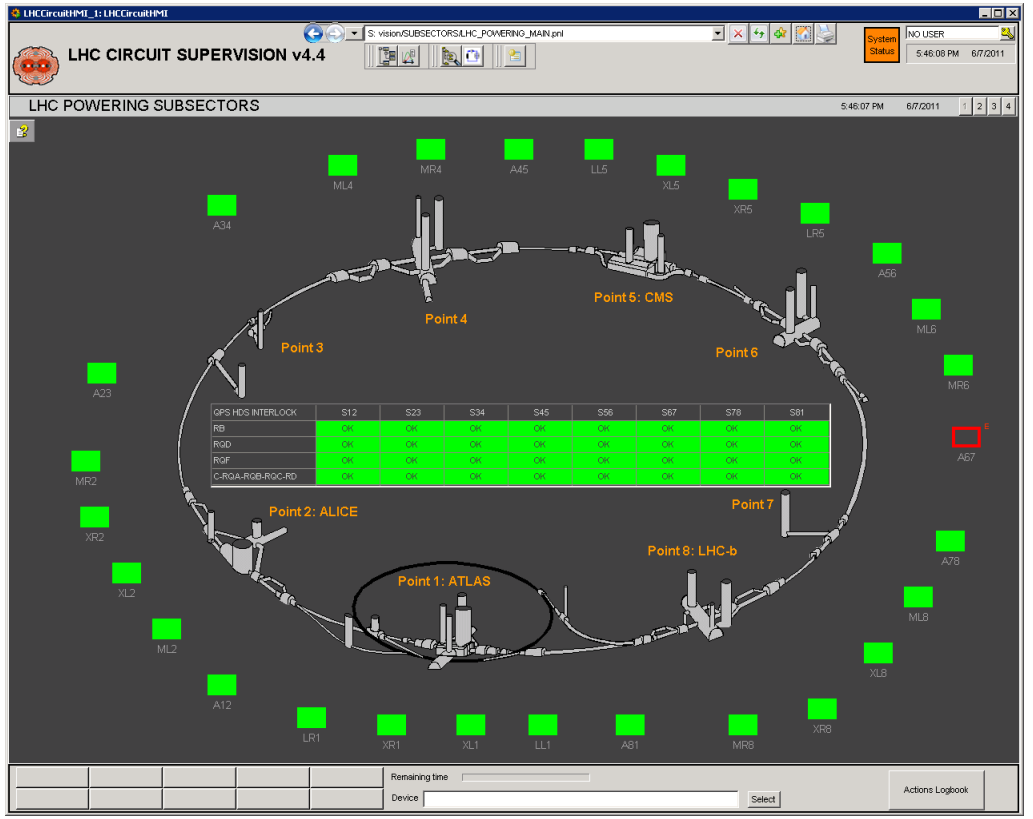
- Expert system
  - Translate experts' knowledge into formulation sets / rules
  - Rules central storage
  - Rule template to be reused, parametrized, validated
- Signal Processing Language (SPL):
  - Domain specific language (DSL)
  - Simple formulation
  - Time reasoning and temporal expression
  - Mathematical and logical functions



LHC circuit monitoring

Condition monitoring analysis *(in collaboration with TE-MPE)*

- **Evaluation of the LHC circuits health**
  - Degradation after many years of operations
  - Monitoring conditions:
    - anomalous change of current flows, impedance, circuit functioning …
- Challenges
  - Assessment of the system status involves
    ~ 500K Signals (electrical circuits, magnets, power converters, switches)
  - Readout (from 10KHz to 1Hz)
  - Time reasoning over desynchronized streams



16 WinCC OA servers, 44 industrial FECs, 2800 radiation-hard devices

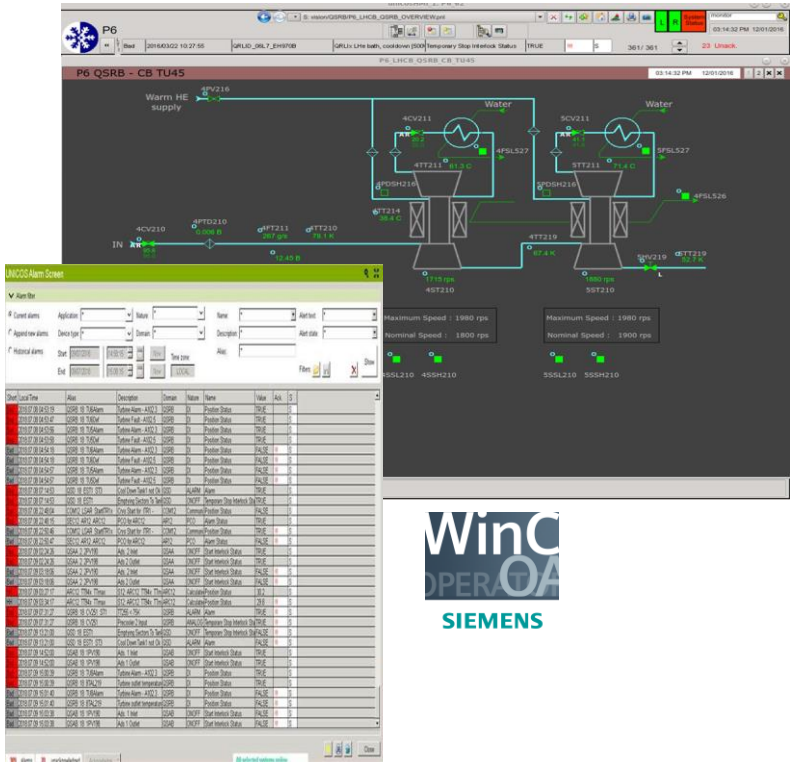
Events/Alarms generated by LHC control system

Condition monitoring of control alarms and events

- Industrial Control Systems for accelerator and technical infrastructure:
  - 220 WinCC OA apps over ~150 physical host with 25M Data Point Elements
  - ~5M I/O channels with 0.5M defined alarms
- Huge amount of events & alarms generated by control processes

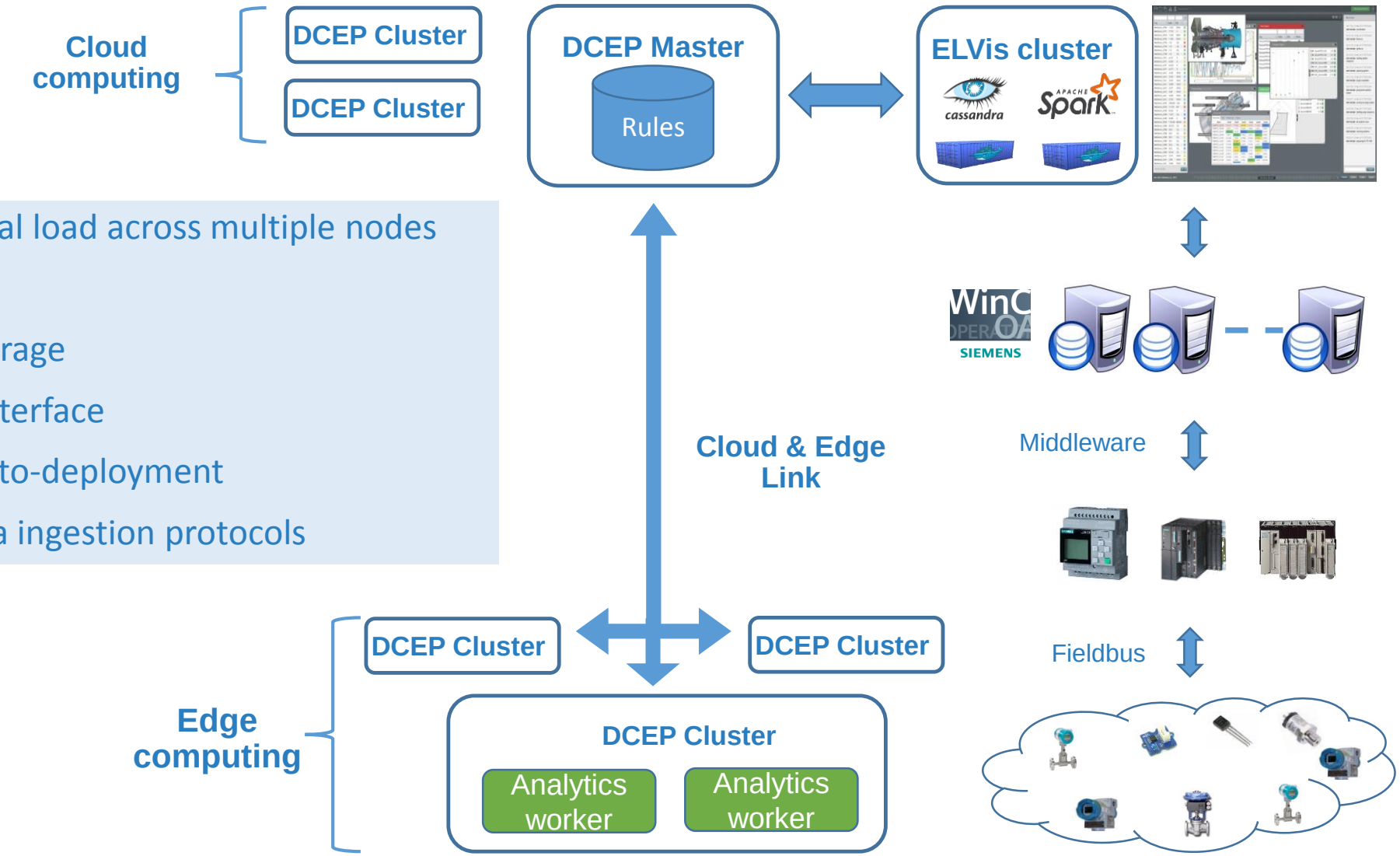
- **<u>GOAL:</u> support system experts/operators to identify critical conditions**

- Anomaly detection based on KPI and specific conditions:
  - Alarms number, integration in time, distribution, frequency, …
  - Detect in/decrease of KPI for time intervals
  - Multi-index/KPI analysis
  - Pattern mining analysis described by variables specifications (rules)
  - Identification of outliers for the predefined KPIs



Siemens Smart Industrial IoT analytical platform

- Distributed computational load across multiple nodes
- Faster rule deployment
- Improved knowledge storage
- Web-based multi-user interface
- Service discovery and auto-deployment
- Support for multiple data ingestion protocols

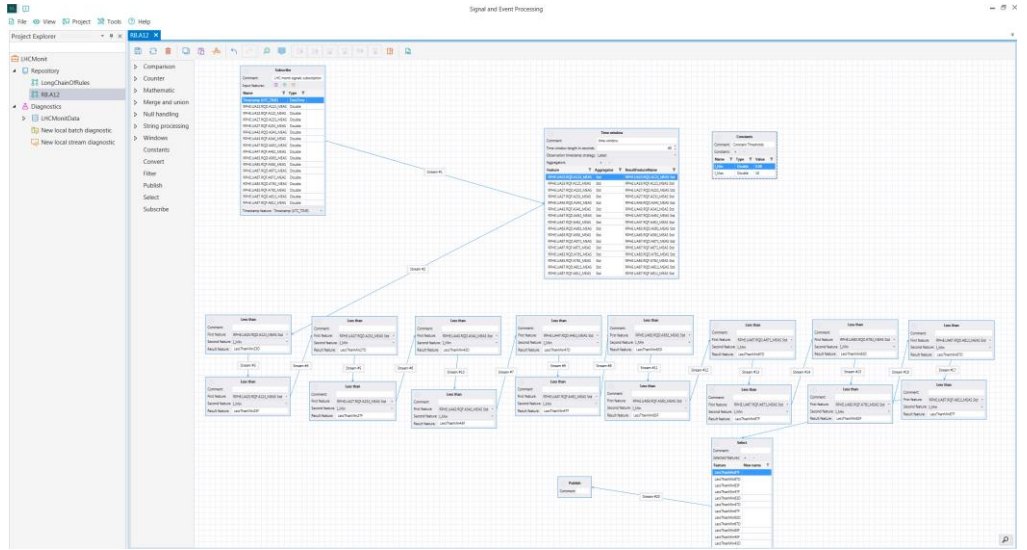


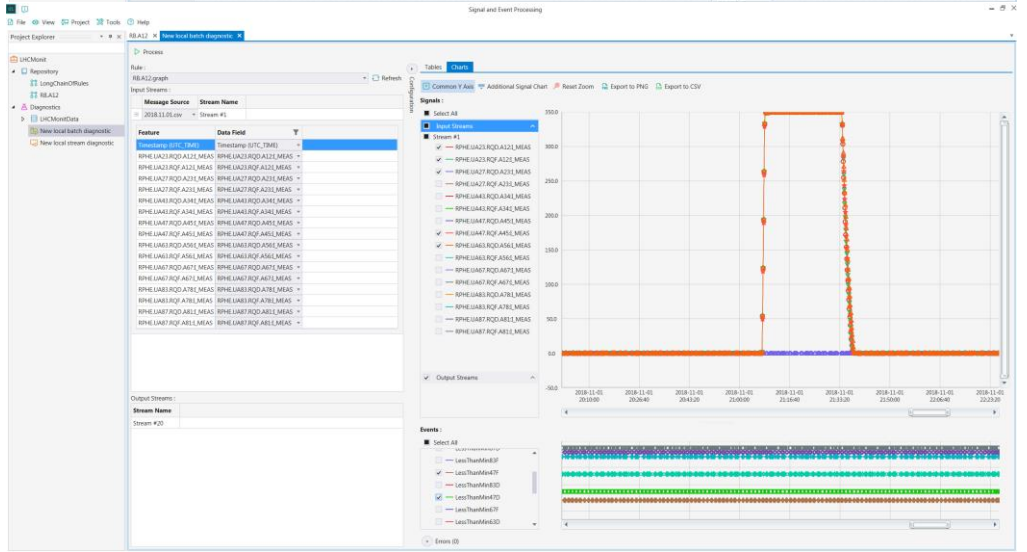
Signal Event Processing Language (SEPL)

Engineering tool to define analytical workflows

- Status:
  - Multiple versions of SEPLab released and evaluated
  - Handle numerical data and events at the same time
  - Automatic input data parsing

- Next
  - Integration with new Smart IIoT interface/API
  - Macro for simple script and template
  - Support for both online and historical data (rule validation)

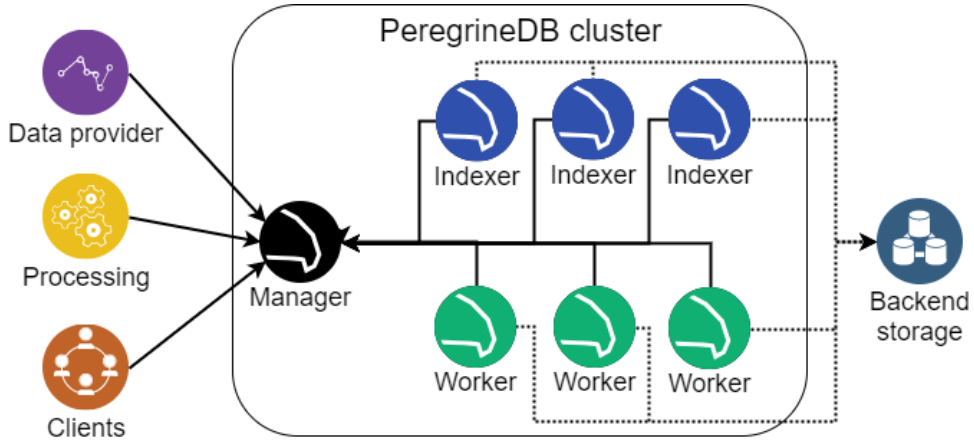




PeregrineDB for batch condition monitoring

In collaboration with ITMO university

- **Database optimized for fast , high availability storage and retrieval time for time series data**
- Extracts relevant inputs from big historical data sets to feed CEP
- Main features:
  - ✓ Aggregation and sampling
  - ✓ Data compression
  - ✓ Support for different backends, data formats
  - ✓ Lightweight index
  - ✓ High speed at ingestion and extraction
  - ✓ Distributed architecture



Summary & outlook

- Next Generation Archiver:
  - Good progress on all components
  - Promising results of pilot deployments at CERN
  - Considerable testing efforts required before deployment in ALICE in 2019
- Data Analytics:
  - **New use-cases** identified for both **Conditioning Monitoring** and **ML**:
    - LHC Circuit, Cryo, alarms KPI analysis, optimization of linac3 beam source
  - Various versions of **Smart IIoT** platform and **SEPLab** tested
  - Cloud computing: initial **ELVis integration** with Smart IIoT platform

- **Successful collaboration** with **Siemens** advancing at good pace
- Widening the collaboration scope: **new resources**, **new activities**
  - New fellow: Anthony Hennessey
  - Evaluation of **Siemens PLC AI module** against CERN use-cases
- A **big thanks** to **Siemens** for the fruitful collaboration and continuous support!

Thank you!

CERN BE-ICS

https://be-dep-ics.web.cern.ch/