DATA STREAMING WITH APACHE KAFKA FOR CERN SUPERVISION, CONTROL AND DATA ACQUISITION SYSTEM FOR RADIATION AND ENVIRONMENTAL PROTECTION

Adrien Ledeul, Gustavo Segura Millan, Alexandru Savulescu, Bartlomiej Styczen
CERN, Geneva, Switzerland

MOMPL010

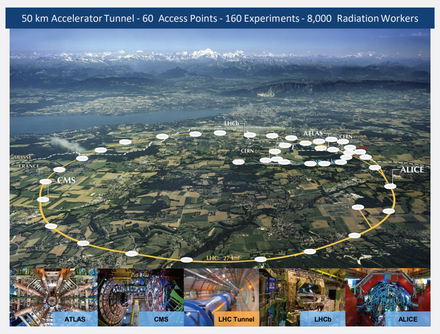
RADIATION AND ENVIRONMENTAL MONITORING AT CERN

OPERATIONAL RADIATION PROTECTION

CONTAMINATION

VENTILATION

GATE



540 Monitoring Stations:
- **27** Station Types
- **COTS** & **CERN instruments**
- **Surface** and **Underground**
- Geographically **Distributed**

3,200 Measurement Channels:
- **3,500,000** measurement to archive every hour

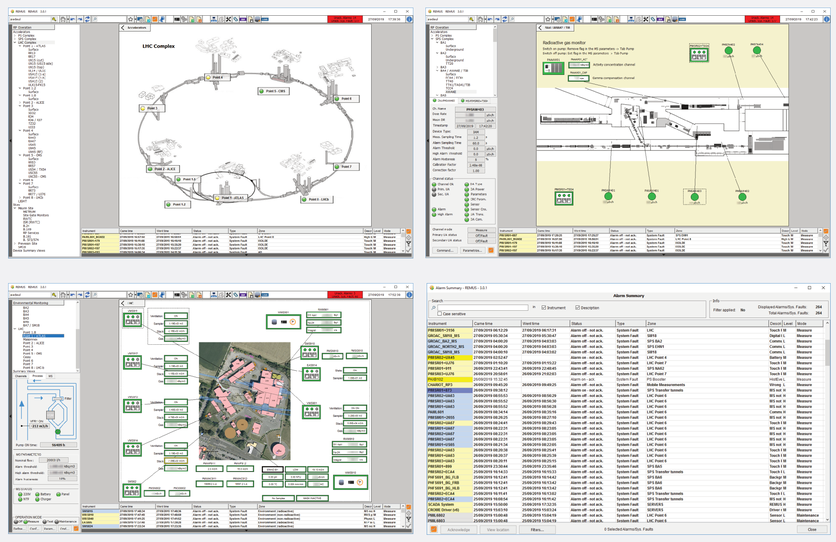
AIR

WEATHER

WATER

REMUS – RADIATION AND ENVIRONMENT MONITORING UNIFIED SUPERVISION

The **Occupational Health & Safety and Environmental Protection Unit** at CERN provides a **SCADA** system for the **Radiation Protection** and **Environment Monitoring** of particle accelerators, experiments and the environment



- **250+** Users
- **8** Control Rooms
- **700+** Synoptic
- **Tailorable** Layouts
- **Multiplatform**

- **650,000** Data Tags
- **85,000** Alarms
- **55,000** Parameters
- **3,700** Changes /sec

- **Continuous Operation**
- **>99.99%** Availability
- Full **Redundancy**

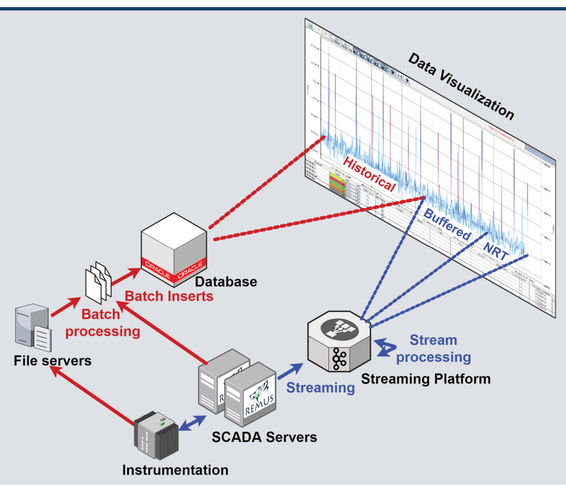
REMUS DATA STREAMING POWERED BY APACHE KAFKA

Data Visualization Tools for Control Rooms (1)

REMUS Data Streaming infrastructure allowed for a **Secured**, **Reliable** and **Scalable** near-real time data source

Latency of less than **1 second** between the data acquisition from the distributed instrumentation, the stream processing and the near real-time display of measurements in **CERN Control Rooms**

In operation since March 2019 in CERN Control Rooms with **100% Availability**



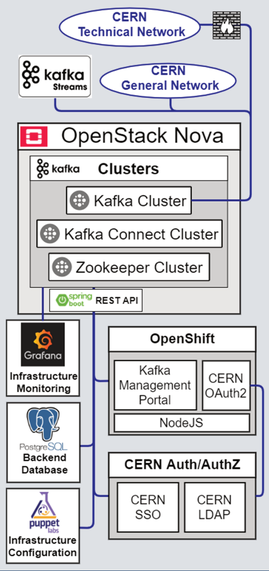
Kafka Connect (8)

Kafka Connect is an **open-source framework** providing an **API** for building Kafka connectors. **Connectors** can either take data from a **Data Source** and push it into a Kafka cluster or pull the data from a Kafka cluster into a **Data Sink**

Available Connectors:

PostgreSQL · mongoDB · elastic · hadoop HDFS · MQTT.ORG · influxdb · Apache ACTIVEMQ · ...

Kafka @ CERN (7)

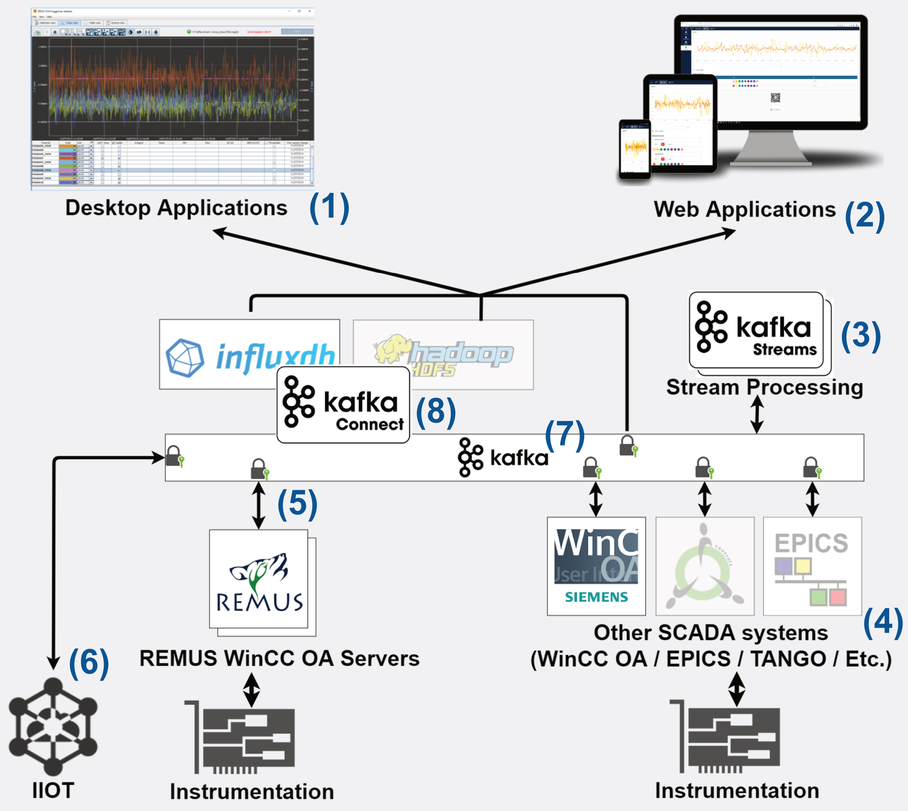


CERN IT and Beams Departments operate **Kafka clusters**, used by **NXCALS** (Next CERN Accelerator Logging Service, Beams Department), **IT Security Event Operations**, **CERN Data Centre Infrastructure Monitoring** (IT Department) and **REMUS**

The Kafka Cluster used by REMUS is available from **CERN General Purpose Network** and **CERN Technical Network**

Kafka & Industrial IoT (6)

Utilizing a bi-directional connector between WinCC OA and Kafka permits **communication with any IIoT layer** having Kafka connecting capability

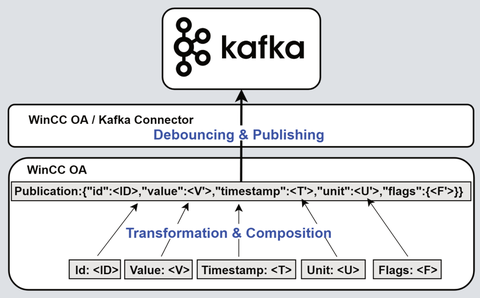


WinCC OA To Kafka (5)

All **Measurements** retrieved by the REMUS supervisory system are published in near-real time through Kafka, in a dedicated **Kafka Topic**

Write access to REMUS Kafka Topics is secured through **TLS Encryption** and **Kerberos Authentication**

REMUS contains **3,300** publication tags, sending about **600** messages per second to CERN Kafka brokers



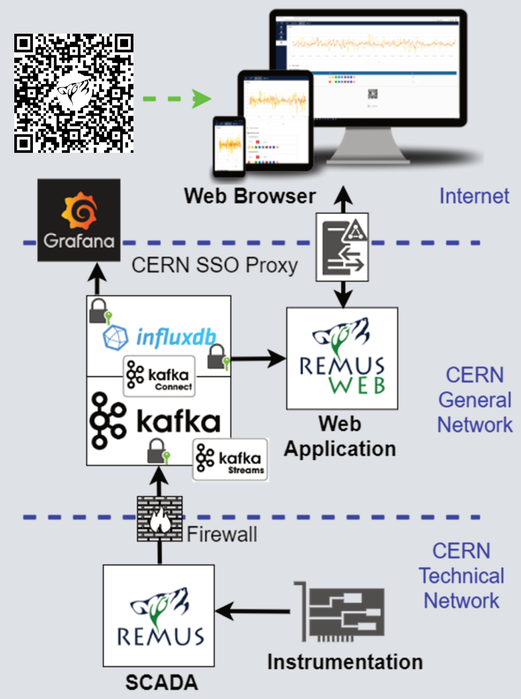
Web & Mobile (2)

The streams of data generated from REMUS are exploited by a responsive web application, **REMUS Web**, able to display measurements in near real-time, among other information about the SCADA processes

Web-based applications such as **Grafana** or **Chronograf** provide data visualization and analytics tools out-of-the-box

Automatic generation of **QR Codes** from the web interface provides a fast, user-friendly yet secure way for accessing near real-time data from **any terminal** device having internet connectivity

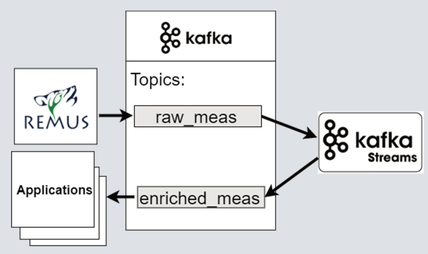
Access to the web application is protected by the **CERN SSO** - Single Sign On - service



Kafka Streams (3)

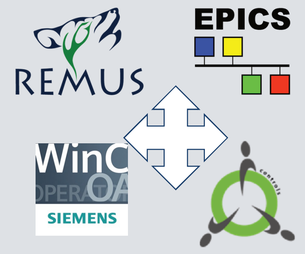
REMUS relies on Kafka Streams (released in 2016) for its **Stream Processing** functionalities

Kafka Streams provides **Exactly Once** semantics and benefits from Kafka scalable and distributed architecture advantages



Link with External Control Systems (4)

The **wide adoption** of Apache Kafka as a solution for **Asynchronous Data Streaming** opened up possibilities for REMUS to communicate with **external SCADA systems** based on different technologies such as Tango, EPICS, WinCC OA etc.



CERN Occupational Health & Safety and Environmental Protection Unit
Author: Adrien Ledeul
e-mail: *adrien.ledeul@cern.ch*

ICALEPCS 2019