Smart Data analysis of CERN Control Systems

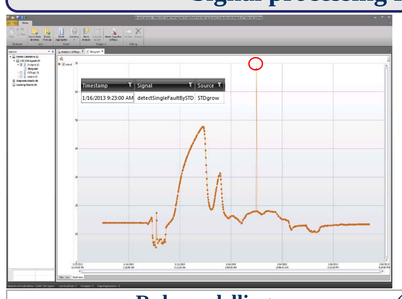
F. Tilaro, A. Voitier , M. Gonzalez Berges (CERN, Geneva, Switzerland)

ABSTRACT

The CERN Large Hadron Collider (LHC) is known to be one of the most complex scientific machines ever built by mankind. Its correct functioning relies on the integration of a multitude of interdependent industrial control systems, which provide different and essential services to run and protect the accelerators. Each of these control systems produces a huge amount of data, that holds an important potential value to optimize the entire system in terms of its operational efficiency and reliability. The extremely large control data sets, often termed as “Big Data”, cannot be processed through the use of traditional database tools. This is why it is necessary to design and develop new analytical frameworks, data models and data-mining algorithms scalable enough to cope with the huge volume of data; but at the same time they should be capable of fulfilling the timing constraints. Moreover due to the critical nature of industrial systems the analysis should be carried out without affecting the industrial processes and the system availability: no downtimes can be tolerated because of their inherent expensive consequences. CERN also represents a unique and critical environment to evaluate any analytical solution due to the size and heterogeneity of its industrial control systems. The main goal of the openlab collaboration between Siemens and CERN is to design and implement a Big Data platform which makes use of innovative analytical techniques and software packages flexible enough to match the challenging data analytics requirements of the various industrial control systems.

WatchCAT Analysis Tool for Offline Analysis

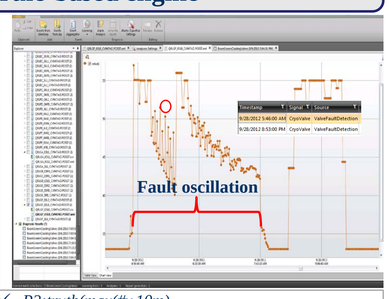
Signal processing rule-based engine



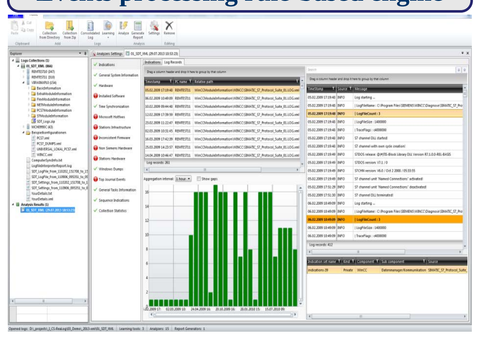
Rule modelling

- ✓ *R1:truth(sd(#value,10m!)>(sd(#value,4h!)*K))*

Events processing rule-based engine



- ✓ *R2:truth(max(#v,10m)-min(#v,10m)>sma(#v,10m)/100):duration(>=1h)*



WatchCAT installation in CERN OpenStack and its data integration

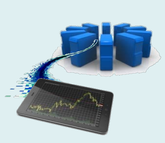
Analytics features

- Data Fusion of events & sensors
- Complex Event Processing (CEP)
- Automated Learning of fault patterns
- Logical Reasoning for Fault Detection & Isolation
- Fault prediction based on recognizable patterns
- Signal processing
- Graph visualization

Analyze

Detect abnormal pattern for Diagnosis and Prognostics based on domain knowledge

Learn



Provide experts with Root-cause and Gap Analysis using Rules and Patterns Mining

Diagnose

Forecasts, Trends and Early-Warning to increase Operating Hours

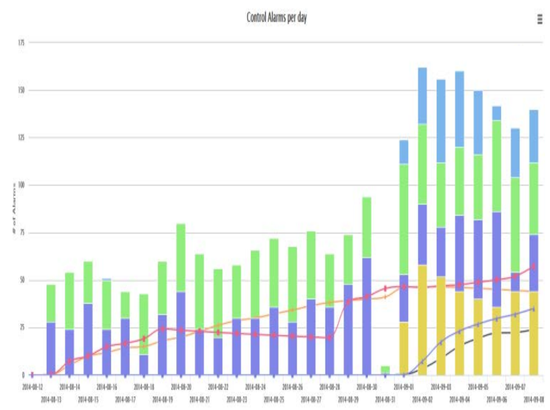
Offline Repository

Control Data

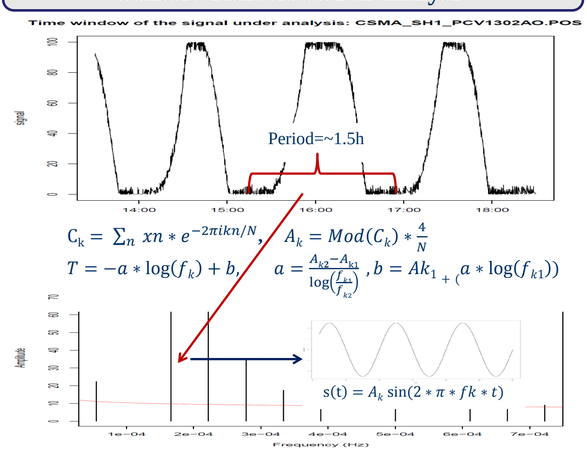
Analysis Feedback

ELVis Analysis Framework for Online Analysis

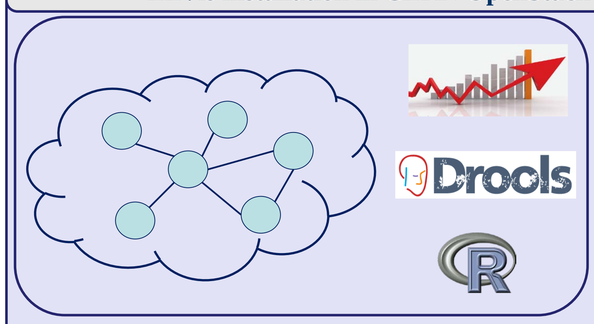
Alarms threshold learning analysis

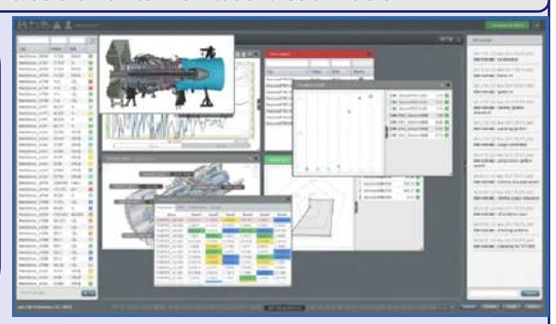


Sensors fault detection analysis



ELVis installation in CERN OpenStack cloud and its rich web-visualization





OPC Adapter | **WinCC OA Adapter** | **Stream Adapter**

Control Data Stream

Analysis Feedback

Analytics features

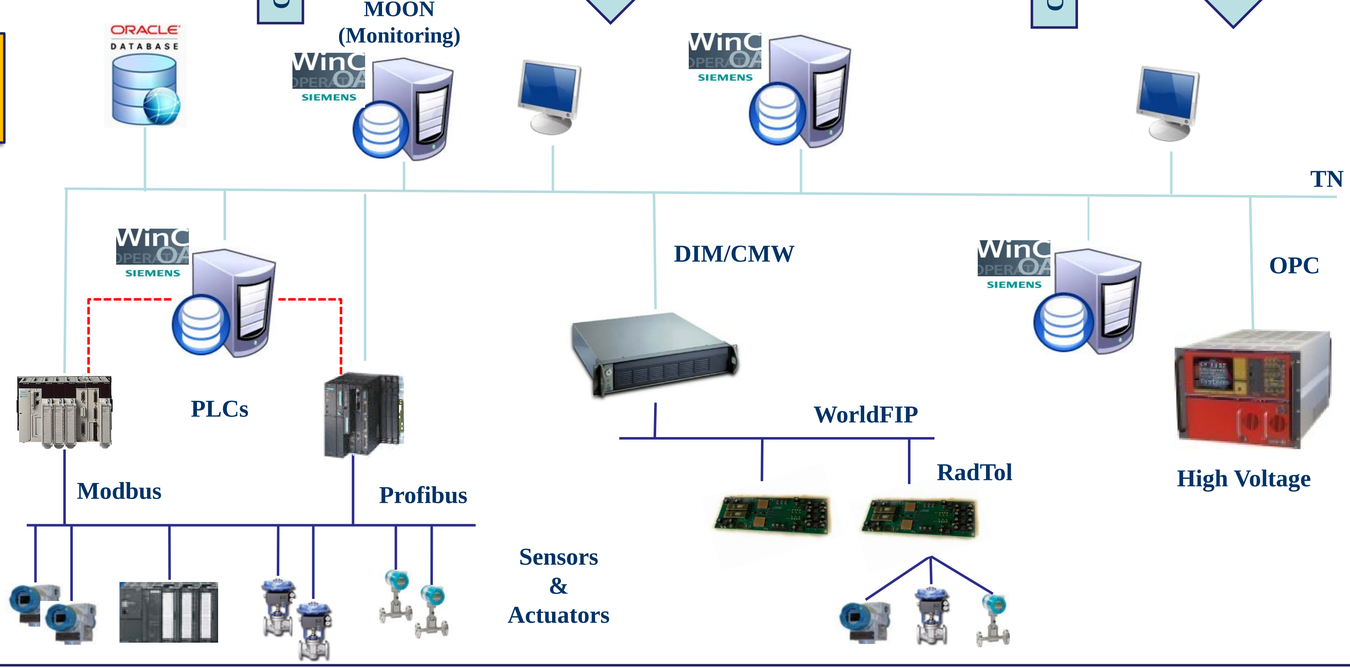
- Cloud-based Big Data Analytics
- Real-Time Stream Processing at customizable sampling rates
- High Performance Online Visualization in Rich Web-UI
- Sensor Data Validation Intelligence

CERN Industrial Control Architecture

Supervision layer

Process layer

Field layer



Summary Table

Industrial Control Application	WinCC OA Supervision Systems	Number of Parameters (I/Os)
ALICE	100	3×10^6
ATLAS	130	1.2×10^7
CMS	90	10^7
LHCb	160	10^7
Accelerators	200	10^7

CONCLUSIONS

Due to the complexity of CERN LHC industrial processes Smart Data technologies can play a fundamental role in their analysis to retrieve new insights and a better understanding of the entire control system. From an architectural standpoint the latter is, indeed, made of a multitude of different distributed control applications, which have been independently developed and updated. The activities carried out within the openlab collaboration have focused on analyzing the industrial data generated by the control systems with the main goal of increasing the reliability and improving the monitoring level. Another important benefit coming out from the used analytical frameworks is the advanced support to diagnose issues (root-cause analysis) and even detect anomalies that could eventually be the cause of future forced outages. Once these anomalous patterns have been discovered and defined, a prediction system can help the operator to take in time the proper actions. Nevertheless it is also worth mentioning the operation support enhancement thanks to the better performances achieved by running the analysis in a cloud-based framework. The initial encouraging results have motivated both CERN and Siemens automation industry teams to continue the data analytics activities in the future phase of their openlab collaboration.