SCADA Statistics Monitoring Using the Elastic Stack

James Hamilton, Brad Schofield, Manuel Gonzalez Berges, Jean-Charles Tournier

TUPHA034

The Problem

The Industrial Controls and Safety systems group at CERN, in collaboration with other groups, have developed and currently maintain around 200 controls applications that include domains such as LHC magnet protection, cryogenics and electrical network supervision systems. Millions of value changes and alarms from many devices are archived to a centralised Oracle database but it is not easy to obtain high-level statistics from such an archive. A system based on Elasticsearch, Logstash and Kibana (the Elastic Stack) has been implemented in order to provide easy access to these statistics. This system provides aggregated statistics based on the number of value changes and alarms, classified according to several criteria (e.g. time, application domain, system, device). The system can be used, for example, to detect abnormal situations and alarm misconfiguration. In addition to these statistics each application generates text-based log files which are parsed, collected and displayed using the Elastic Stack, to provide centralised access to all the application logs.

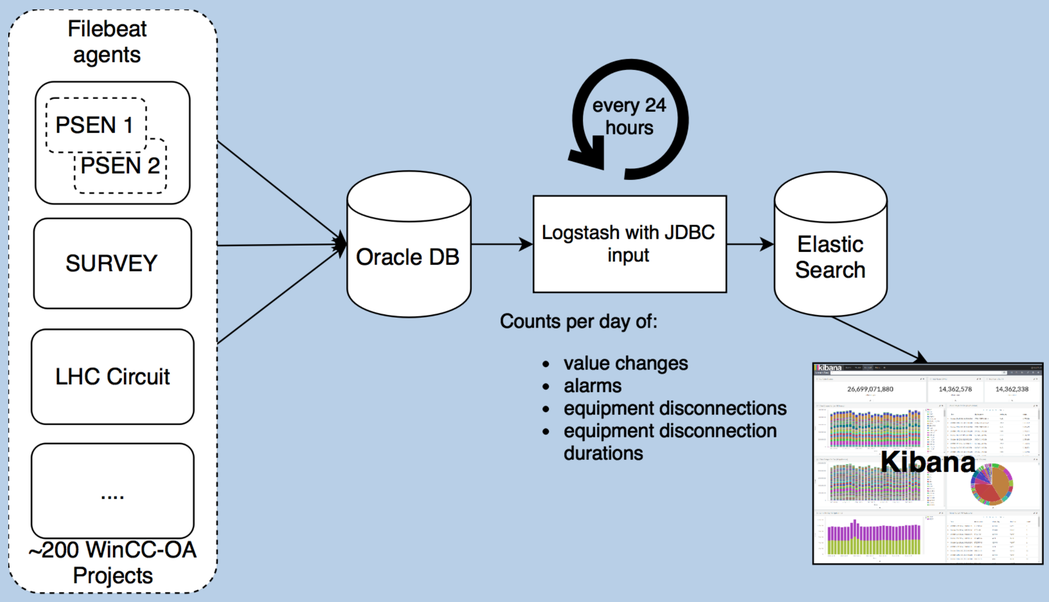
Technology

Elasticsearch Logstash Kibana Beats

Statistics

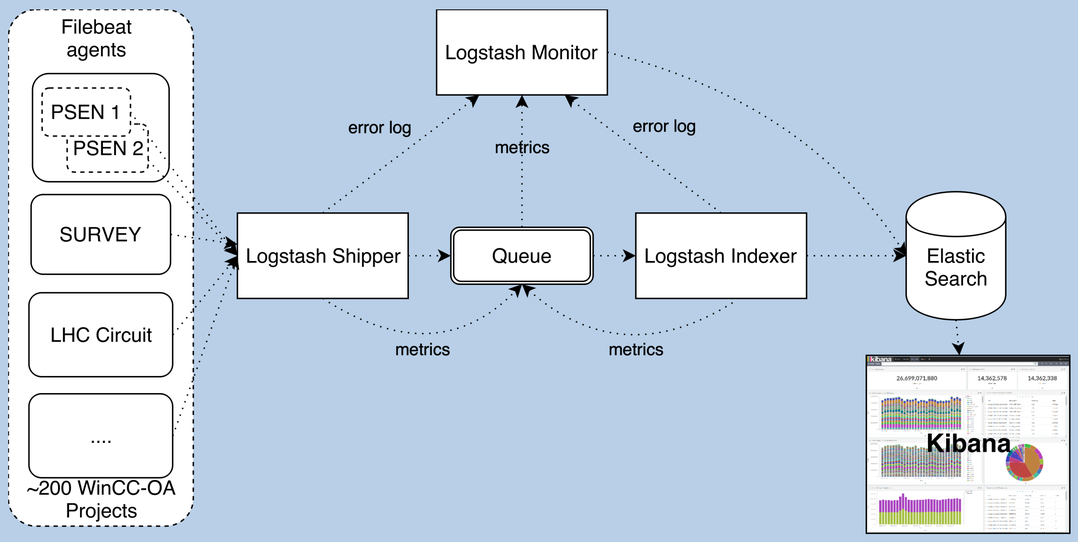
- 145 applications
- 800,000,000 value changes per day
- 600,000 alarms per day
- 1,000,000 log entries per day
- 3,000,000 Elasticsearch documents per day

Value Change & Alarm Statistics



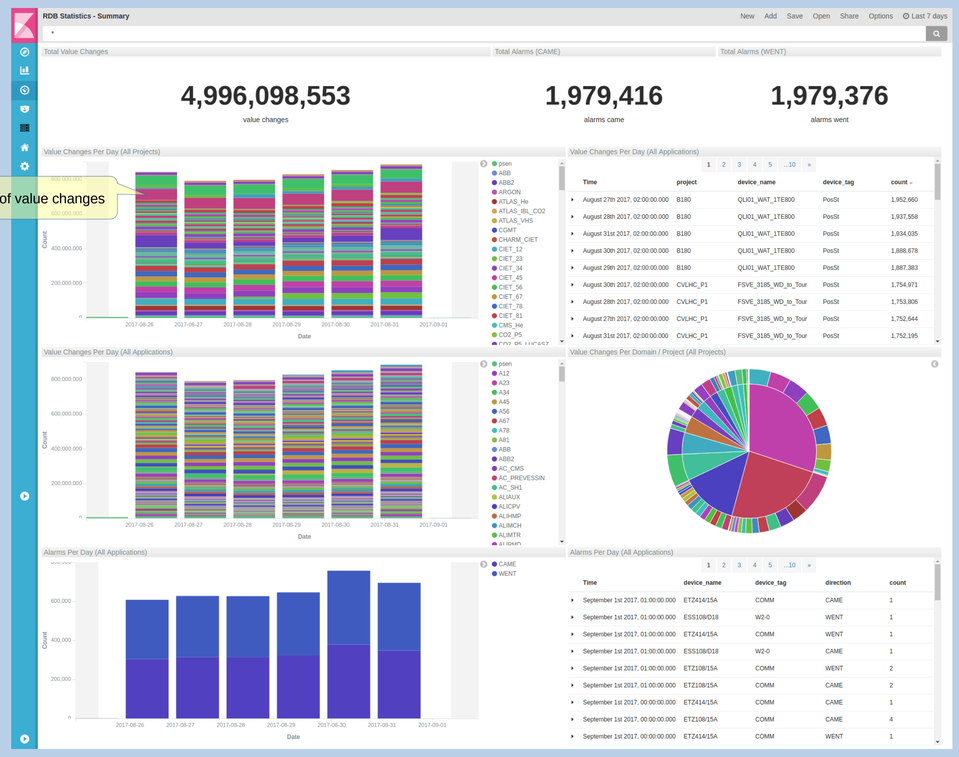
- Each Elasticsearch document contains the count of value changes/alarms for that day
- ~500 daily SQL to obtain aggregate statistics
- Users won't see live data in Kibana but aggregated data from previous days
- Queries are scheduled using Logstash with the logstash-input-jdbc plugin

Application Text Logs



- Filebeat – a lightweight application – runs on each production server
- Logstash shippers receive log entries and combine multiline log entries
- The queue holds log entries temporarily
- Logstash indexers read from the queue and parse the logs using regular expressions
- Easy to scale parsing by adding more indexers
- Logstash monitor reads error logs and statistics from shippers, indexers and queue

Value Change & Alarm Dashboard



Case Study: Archive Misconfiguration

The application LHC Circuit which monitors the roughly 1600 power converters for the LHC was found to have a faulty archiving setting for the reference voltages of many devices. As an analog value, the voltage reference would typically be expected to have the customary deadband filtering.

However, a number of devices had been configured with 'on change' archiving, meaning that voltage reference values were being sent to the archive at their sampling rate of approximately 2Hz. This translated to over one hundred thousand value changes per day, which stood out very clearly in the value change statistics. For example, in the figure on the left the bar charts are split into slices for each application and it can be seen which applications have higher numbers of value changes & alarms. Thanks to the service the archive settings could be corrected for all affected devices.

Conclusion

Industrial controls applications at CERN produce lots of data in different areas including an archive of process and system data and log files. There are many activities going on to analyse this data and get value from it. The developed service presents a high-level approach where the data is mainly viewed as aggregated statistics and correlations. Only in some cases is the actual data processed in detail. The service presented has been built on a modern state of the art set of tools - the Elastic Stack - that greatly facilitated the task.

There is a huge amount of 'noise' in the controls application logs which provides a good candidate for anomaly detection using machine learning. It is currently difficult to find the 'true' errors when looking through the log and an automated method of detecting anomalies would be very helpful for the application developers.

Furthermore, we believe that the combination of value & alarm statistics and application logs could lead to some interesting results to better understand the behaviour of CERN's controls applications. For example, if there are errors in the log there could be corresponding alarms.

CERN Beams Department
Industrial Controls and Safety Systems Group
Corresponding author: James Hamilton *james.hamilton@cern.ch*

ICALEPCS2017
Barcelona · Spain, October 8-13 Palau de Congressos de Catalunya