Benchmarking tools for NextGen Archiver for WinCC OA

Speaker:
Jayaditya Gupta (BE-ICS)
jayaditya.gupta@cern.ch

SUPERVISOR(S):
Rafal Lukasz Kulaga
Anthony Hennessey

Openlab Collaborator:
ETM (Siemens)
Problem Statement

Objective

What and Why are we doing this?

- Record state changes
- Experiment control system

CERN electric Grid
CERN Cryogenics
Interlocking system
Problem Statement

Objective

What and Why are we doing this?

Woops !!!!!
Problem Statement

Currently used:
Oracle-only archiver

Upcoming:
NextGen Archiver with support for Oracle, InfluxDB and others through pluggable backends.
Challenges

• Large amount of data
  Around 3 years of CERN electric control system data
  Signals count: ~1.4 million
  Total rows: ~11.5 billion
  Per row: ~120 bytes

• Varied frequency

• Complex data retrieval

• Data analytics
  ONLY FROM 1 subset of schema.
  Total: 700
Workflow & Tools

1. Fetch Data From Oracle
2. Transform & Store data in InfluxDB
3. Run queries on Oracle & InfluxDB
4. Write Results

Original DB which contains all the data

Temporary file

Write to file

Read data from file

Measure write performance

Query Benchmark

Input Query Types & Config

Archive Pump Tool

Measure read performance
First Results and Future work

Some numbers

Results
- Easily achieved write rate of ~50,000 events/s in test
- Certain queries are not finished because influx is running out of memory.

Future Work
- Benchmark on Apache kudu
- Read tests
QUESTIONS?

Jayaditya.gupta@cern.ch
jayadityagupta11@gmail.com

Twitter: @jayssj1, Github: hackertron