HOW CASSANDRA IMPROVES PERFORMANCES AND AVAILABILITY OF HDB++ TANGO ARCHIVING SYSTEM

Reynald BOURTEMBOURG, Jean-Luc PONS, Pascal VERDIER, ESRF, Grenoble, France

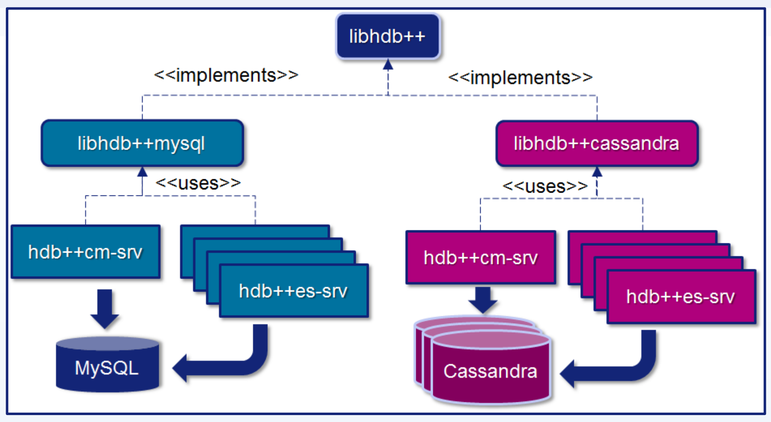
HDB++, the new TANGO event-driven archiving system is being developed as a collaboration between the ESRF and Elettra.

Specific libraries have been developed, giving the possibility to store HDB++ data into Apache Cassandra, the widely used and popular NoSQL database.

Using Cassandra adds high-availability with no single point of failure and scalability to the new TANGO archiving system.

HDB++ Design

HDB++ (see WED3O04) was designed in a modular way, using TANGO device servers to handle the configuration and the events subscribing/archiving process.
These device servers can be compiled with C++ libraries inheriting from the *libhdb++* library and implementing access to the database of your choice.
Up to now libraries for MySQL (developed by Elettra) and Apache Cassandra have been implemented.

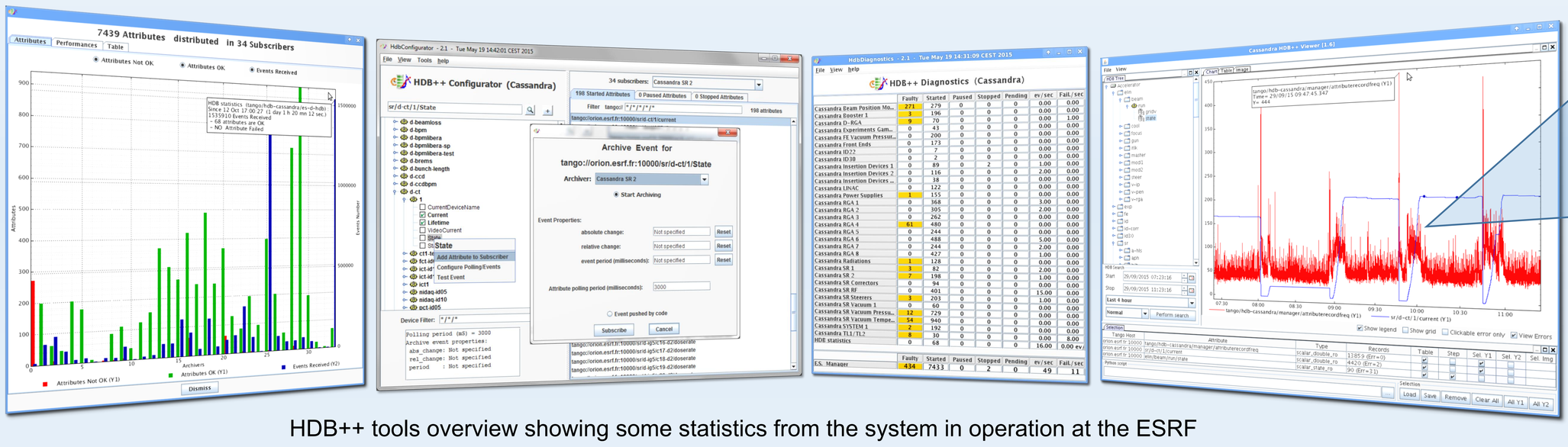


HDB++ design

HDB++ Tools for free

Implementing specific Cassandra HDB++ libraries inheriting from the HDB++ C++ abstract libraries or implementing the Java HDB++ interfaces was enough to be able to create and manage a Cassandra-based HDB++ system.

All the already developed HDB++ tools (device servers, configurator GUI, diagnostics GUI, extractor GUIs) could be reused directly.



HDB++ tools overview showing some statistics from the system in operation at the ESRF

Number of archived events versus storage ring current during machine dedicated time

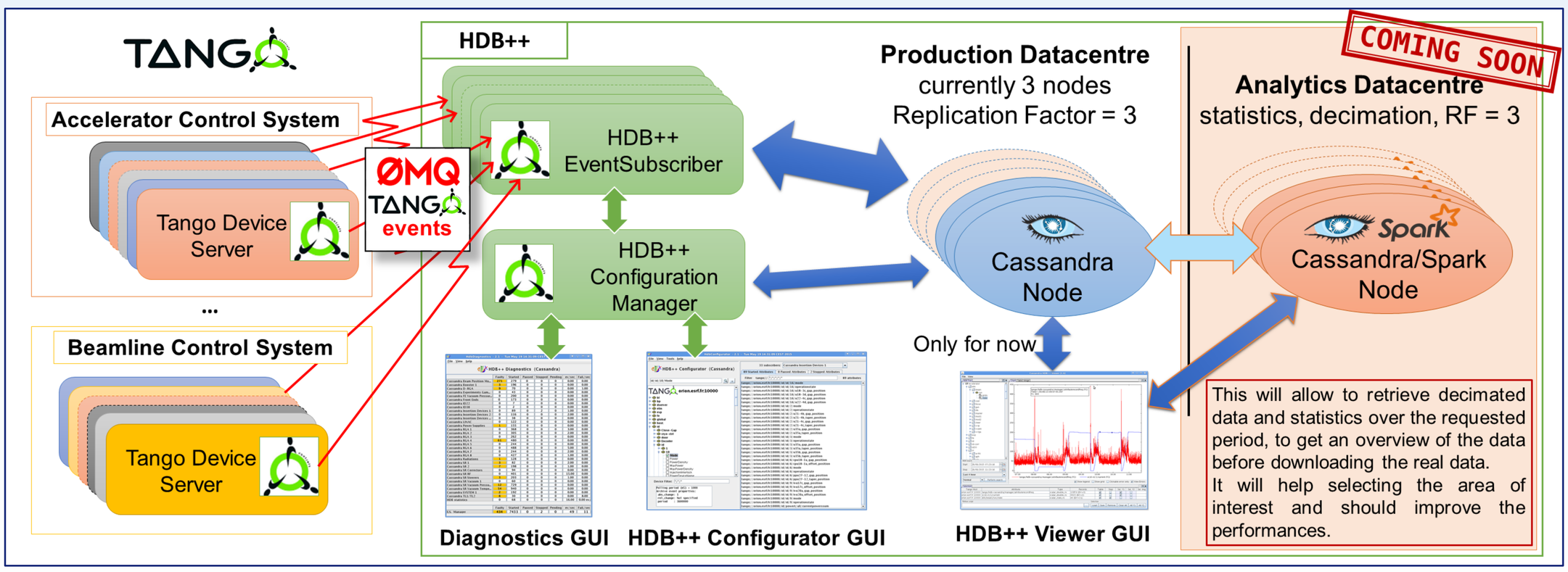
There are peaks when the beam is killed and an increase of the archive events frequency during injections.

Maximum peak
(March → October 2015)
=
762 events/s
Average ~= 55 events/s

7439 archived attributes

HDB++ Cassandra @ ESRF

At the ESRF, 3 Cassandra nodes with a replication factor of 3 are in operation since October 2014, in parallel with the old HDB system and the HDB++ MySQL version. It is planned to add a new datacentre composed of 3 Cassandra nodes soon. This datacentre will be dedicated to analytics and will be using Apache Spark to compute statistics and fill in decimation tables.



HDB++ Cassandra architecture and future evolution