



ICALEPCS 2015

International Conference on Accelerator
& Large Experimental Physics
Control Systems

How Cassandra Improves Performances and Availability of HDB++ TANGO Archiving System

Reynald Bourtembourg, Jean-Luc Pons, Pascal Verdier
ESRF, Grenoble, France

WEM310



ICALEPCS 2015

International Conference on Accelerator
& Large Experimental Physics
Control Systems

Project goal

HDB++ = **TANGO** collaborative project (Ref: WED3O04)



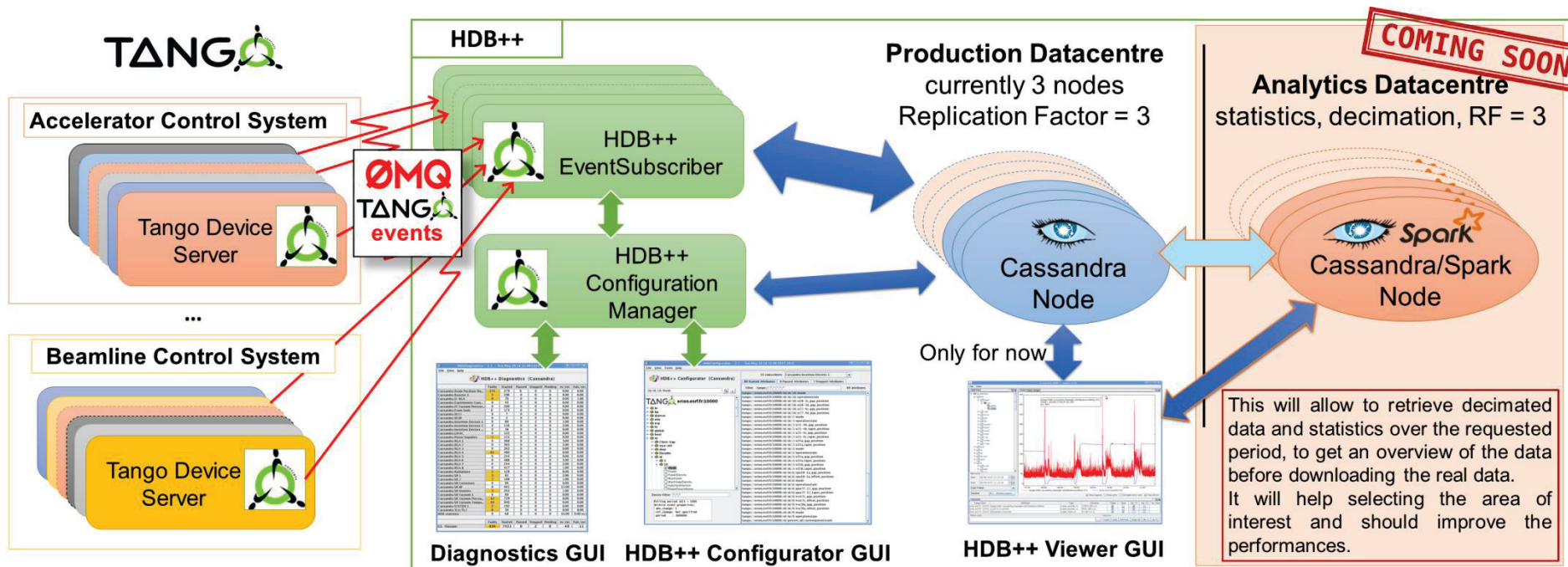
- Continuous availability with no SPOF
- Linear Scalability
- Multi Datacentre replication
- Analytics with **Spark**



ICALEPCS 2015

International Conference on Accelerator & Large Experimental Physics Control Systems

Solution outline and future plan







ICALEPCS 2015

International Conference on Accelerator & Large Experimental Physics Control Systems

Conclusion

- ✓ **Great HDB++ design!**
 - ✓ Easy to add support to new DB back-end
- ✓ **Cassandra works!**
 - ✓ Continuous availability (when well used)
 - ✓ Write Performances
 - ✓ Scalability
- **Next Step:**
 - Reduce partitions size
 - Analytics datacentre with 

Come and see the poster for more details! ➔ WEM310



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 part of a collaboration between the ESRF and Elettra.

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

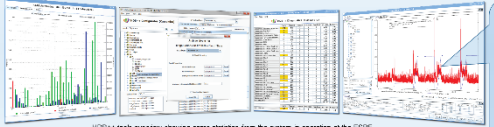
Using Cassandra adds high-availability with no single point of failure and scalability to the new TANGO archiving system.

HDB++ Design

HDB++ (see WED3004) 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++ 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 running showing some statistics from the system in operation at the ESRF.

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 with 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 deconvolution tables.

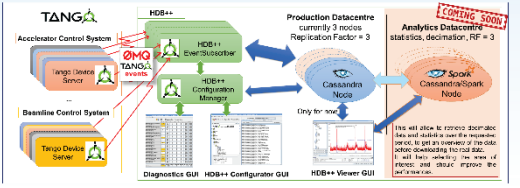



Number of archived events versus storage time during machine dedicated time.

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

Maximum peak (March -> October 2014): 742 events/s (at full beam current)

Average -> 10 events/s



HDB++ Cassandra architecture and future evolution.

TANGO

<http://www.tango-controls.org>

ESRF

<http://www.esrf.fr>

ESRF

<http://www.esrf.fr>

ESRF

<http://www.esrf.fr>