Industry, Production and Logistics

BBData, a Big Data platform for Smart Buildings

Lucy Linder, Jean Hennebert, Julien Esseiva

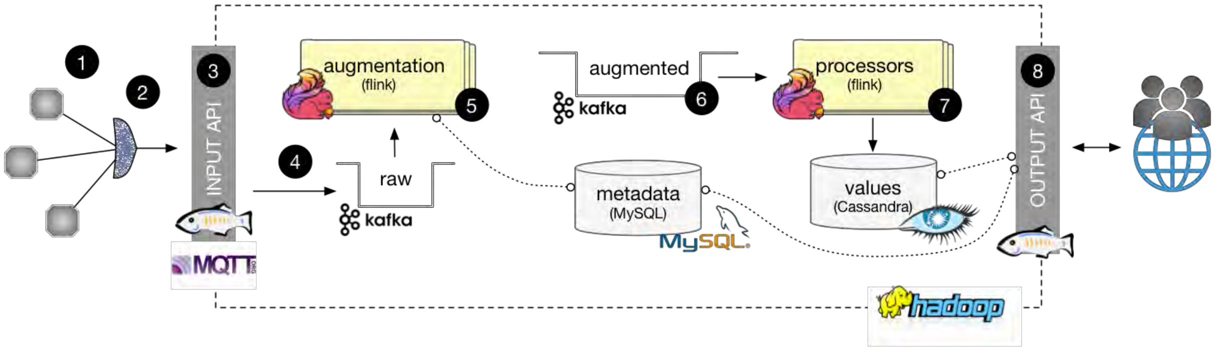
HES-SO, HEIA-FR, iCoSys - Institute of Complex Systems

Abstract

The handling of Smart Building data in a Big Data perspective is challenging in many ways. First, the building automation domain has a long history with diverse systems and technologies creating interoperability issues [1] [2]. Second, the models behind current advanced controls systems become more and more complex, triggering the need for a constant source of data, both historical and in real time, as well as increasing computing power and storage capabilities [3]. Finally, efficiently applying Big Data technologies is a challenge in itself [4], added to the fact that there are few comprehensive approaches to support the collection of data from building sensors and to allow their exploitation.

The BBData – Big Building Data – project [5], is an attempt to answer those challenges by providing a full-featured data processing platform for Big Building Data. Still in its early stages of development, it has been running continuously for one year to gather data from more than two thousand sensors installed in the Blue Hall of the blueFACTORY innovation site in Fribourg. The platform is designed following four major principles: genericity, open- source technologies, scalability and performance. It serves two major purposes: a centralised and safe storage for building data and a highly efficient platform for real time, think-ahead data processing.

The journey of a measure inside BBData is shown in the Figure below. In (1), different kind of sensors produce measures. Those sensors might come from different providers and use various data encodings. To ensure the compatibility of the platform with any kind of equipment, BBData uses the concept of *virtual objects*. A virtual object has metadata, such as name, unit and type, and can be mapped to a real sensor through an ID. This mapping is handled by *collectors* (2). A collector creates a bbdata record for each measure and sends it to the input api in a structured JSON format. In case some sensors lack an internal clock, the collector can also produce a timestamp. The *input api* (3) is the entry point to the system. Available in two flavors, a standard REST API or an MQTT broker, it validates the incoming measure and ensure its authenticity using the object's ID and a secure token. If the check succeeds, the security information is dropped and the resulting JSON is added to an Apache Kafka *message queue* (4). Before processing, the measure is first *augmented* with the virtual object's metadata pulled from a MySQL database (5), such as its unit and type. The result is stored in a second message queue using a compressed format (6).



In BBData, *processing* covers a wide area of tasks, from the saving of raw values into a persistent store to the detection of anomalies or the computation of time aggregations. Each of these tasks is handled by a specific processor. Processors are independent streaming applications running in a hadoop cluster. They subscribe to the augmented Kafka topic, carry their task and save their output, if any, in a Cassandra database. This design makes it possible to add or remove processors without any impact on the system. We have currently two kinds of processors: the first saves the raw records as is, the second computes live time aggregates (mean, max, last measure, standard deviation) with a granularity of fifteen minutes and one hour.

Users and building automation applications can access the data and manage virtual objects through a standard REST interface called the *output api* (7) or via HTML5 web applications [6]. Those applications also offer graphing and visualisations technologies, allowing users to monitor their building easily and build flexible dashboards in minutes.

Our future works include the development of more complex processors such as machine learning predictive systems and the inclusion of a MQTT broker at the output API.

References

[1] M. V. Moreno, «Applicability of Big Data Techniques to Smart Cities Deployments,» 2016.

[2] J. Ploennigs, «A modular, adaptive building automation system ontology,» 2012.

[3] A. Khan, «Big data from the built environment,» 2011.

[4] U. a. K. M. M. a. I. Z. a. W. V. Sivarajah, «Critical analysis of Big Data challenges and analytical methods,» *Journal of Business Research,* vol. 70, pp. 263-286, 2017.

[5] L. Linder a. D. V. a. J.-P. B. a. J. Hennebert, «Big Building Data - a Big Data Platform for Smart Buildings,» *Energy Procedia,* vol. 122, pp. 589 - 594, 2017.

[6] G. Bovet et J. Hennebert, «Offering web-of-things connectivity to building networks,» chez *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, Zurich, 2013.