

From sensor to cloud

Big Data is now present in all sectors of economy and society. Its management raises major scientific and technical challenges. Connected objects bring an important contribution to the current explosion of collected data. Clermont Auvergne University is one of the first campus in France offering to its researchers and engineers a completely open IoT platform. Since 2017, the team is designing with the support of LPC engineering services a set of innovative solutions that allow communicating nodes to transmit through an open and secured communication protocol (LoRa) data from all types of sensors to a cloud hosted by Clermont-Auvergne Mesocentre.

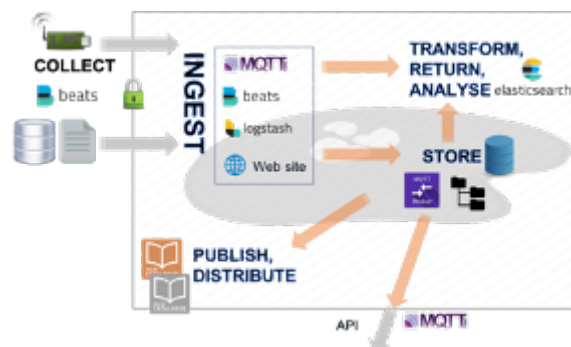
Power Autonomous Communicating Objects

The search for energy efficiency and the transition to eco-responsible industries is one of the major societal challenges of our time. In order to minimize our carbon footprint, it is necessary to think about using or reusing the surrounding energy in the devices of our daily life, but also to minimize the ecological impact of the manufacture of these energy recovery elements. To achieve these ecological objectives, companies and states rely on the development of communicating solutions to make our objects “smart” and thus allow a better measurement and control of energy expenditure. But the Internet and its infrastructure now represent the equivalent of a country in the 5 th World rank in terms of electrical power consumption, and we have gone from 12.5 billion connected objects in 2010 to a projection of 25 billion in 2025¹(<https://www.gsma.com/newsroom/press-release/new-gsma-study-operators-must-look-beyond-connectivity-to-increase-share/>) with the development of these smart solutions. It is therefore imperative that:

- these objects are as invisible as possible in terms of energy cost on the electrical power grid network
- the energy cost of production and recyclability are considered in the beginning of the design of this smart object

The increase in data volume, variety and complexity requires to completely revise our philosophy about data management and analysis. One promising approach is to migrate data treatment close to the connected objects. The objective is to analyze data as early as possible after it is produced and to develop automatic learning approaches for embarked and autonomous systems. The aim is to develop new generation of IOT devices which integrate intelligent system based on neural networks for classifying and sorting the data and so allowing the selection of data to be transmitted. This method leads to a significant reduction in the energy requirement of the IOT sensors (wireless transmission being the most energy-consuming part of electronic devices).

Environmental Cloud



The Environmental Cloud (CEBA) is a data lake designed as an answer to the needs of the academic research community to collect, store, and display environmental data coming from connected objects. The data packets are sent by the different LoRaWAN networks through the internet and ingested as Json files into the data lake using the Elastic Stack (Beats, Logstash and Elasticsearch products from Elastic, Mountain View, CA, USA). As shown on the Figure above, Beats is a set of tools comprising the data shippers that transfer the data from collection points (the LoRa server, in our case) into the data lake. Logstash as a data processing pipeline performs data transformation and shipping (logging of data received from Beats, labeling and sending to storage). Elasticsearch is the primary search engine of the data lake. It is coupled to the visualization tool Grafana (Grafana Labs, New York, NY, USA) for real-time data visualization on dashboards, automatic monitoring, and alarm triggering. Data can be also exported outside of the data lake using a Message Queuing Telemetry Transport protocol (MQTT) to share them with data users.

The data packets are stored for later usage either as flat files or in a relational PostgreSQL database, where their consistency is checked. A website allows data lake users to enrich their datasets with metadata, while publication tools are made available, including a GeoNetwork data catalog.

The CEBA data lake is currently used for the monitoring of radon emission by the Etna volcano in Sicily (for more information see <https://www.mdpi.com/1424-8220/20/10/2755/htm> (<https://www.mdpi.com/1424-8220/20/10/2755/htm>)).

A PhD, started on February 2020 and funded by Financement : CPER / Projet CONNECSSENS-2, supervised by LIMOS and LPC laboratories, is working on data query and integration applied on CEBA. It aims at building a query tool with the same functionalities as those proposed by standard spatial data warehouses.

Contact person

Emmanuel BERGERET, professor

[Emmanuel\(mailto:Emmanuel%2EBergeret%40clermont%2Ein2p3%2Efr?Subject=&body=\)_Bergeret@clermont.in2p3.fr](mailto:Emmanuel%2EBergeret%40clermont%2Ein2p3%2Efr?Subject=&body=)_Bergeret@clermont.in2p3.fr)(mailto:Vincent%2EBreton%40clermont%2Ein2p3%2Efr)

David Sarramia, associate prof.

[David.sarramia@uca.fr](mailto:David%2Esarramia%40uca%2Efr)(mailto:David%2Esarramia%40uca%2Efr)

<https://see.lpc.uca.fr/energy-environment/from-sensor-to-cloud>(<https://see.lpc.uca.fr/energy-environment/from-sensor-to-cloud>)