BIGDATA CONTRIBUTION TO THE MANAGEMENT OF THE FRENCH DISTRIBUTION NETWORK

Danièle BUI
ERDF - France
daniele.bui@erdfdistribution.fr

Marc DE POSSESSE
ERDF - France
marc.de-possesse@erdfdistribution.fr

ABSTRACT

In 2012 ERDF launched an IT reorganisation project aimed at implementing a massively parallelised solution for the data management of energy measures, contractual descriptions of consumption and production sites as well as the technical description of its network. This project addresses the growing need for data analysis in all DSO activities, and especially network planning and operations.

INTRODUCTION

ERDF is the major Distribution System Operator (DSO) in France. It manages the distribution network for 95% of continental France. With a grid of nearly 1.3 million km, it represents the largest European DSO.

The 35 million consumption points are generating growing data volumes: energy measures, power measures, power outage events, power quality … This data is now essential to reaching the distributor's operational excellence on its historic missions. Moreover, the proper functioning of new and evolving electricity markets relies on a swift and efficient treatment of the detailed data available through the deployment of the "LINKY" smart meters.

The renewed need for data processing by the distributor, be it for industrial, administrative or commercial activities, has shown the necessity for the decompartmentalisation of the IT structure. Implementing a data infrastructure common to all of these activities relies on solutions based on centralised storage and massive data processing also known as Big Data.

By explaining the causes for the rising data volumes and the need for consolidating the data repository throughout the company, this article will introduce the Big Data solution being implemented within ERDF and the first results obtained thanks to this project.

DEVELOPMENT

Data processing / analysis a requirement guided by the need to master the network

ERDF is dedicated to operational excellence in its sovereign missions as an electric distribution operator. Network development accompanying new use patterns and production methods, network operations in regards to "Smart Grid" as well as controlling the quality of the electric signal all rely on the generation of more and more information through the instrumentation of the network, new metering systems and exogenous information sources (weather for instance).

An ever more precise modelling of local load curves is a prerequisite to the efficient development of the distribution network

Historically, the sizing of distribution network works (primary substations, MV feeders, MV/LV substations) was based on historical peak load information. The Electric Power System, characterised by centralised production capacities (Nuclear and thermal plants, hydraulic installations), routed energy globally from the transmission network towards the distribution network and end-users.

Today, this structure is questioned by the increasing integration of Distributed Renewable Energy Sources (DRES). At the start of 2014, wind and photovoltaic power capacities connected to the distribution network operated by ERDF reached respectively 7GW and 3 800 MW. These evolutions in energy flows on the network need to be anticipated so as to minimise any rejection from the distribution network towards the transmission network and encourage local consumption of energy produced on the distribution network. The grid connection studies of wind / solar farms and network development must therefore take these new constraints into account and rely on a more precise modelling of consumption and production curves. This detailed modelling requires:

- load curves from a representative (large) sample of consumption sites downstream from the elements being resized ;
- logs of local meteorological data to quantify the thermo-sensitivity of production and consumption
- dimensioning climatic scenarios needing to be applied to determine the extreme cases requiring anticipation.

Network operations: a growing need for local consumption and production forecasts

ERDF is the holder of a contract giving it access to the transmission network of RTE (Réseau de Transport d'Electricité) for an annual invoice of 3 billion Euros. The control of the budget path therefore implies the optimisation of its subscriptions and the monitoring of energy flows to respect these subscriptions.

The ever increasing integration of distributed renewable energy sources (wind / solar power …) and the development / evolution of uses (electric mobility, stationary batteries …) are rendering the management of the network more complex. The planning of network outages, maintenance acts and the minimisation of the invoice to access the transmission network use more and more forecasting models to produce:

- local consumption curves (per primary substation, MV feeder or MV/LV substation)
- production curves per wind / solar farm or even by aggregating dispersed production capabilities (residential wind and solar power) per primary substation, MV feeder or MV/LV substation.

These half hourly active power forecasts are based on: econometric models relying on local consumption logs (several years), production per wind or solar farm and meteorological characteristics (cloud cover, wind speed and direction, temperature …). They must be produced on approximately 20 000 grid cells several times a day so as to take into account both network events and the update of weather forecasts.

Controlling the electric signal's quality requires the collection of hundreds of millions events each year

ERDF is subject to regulatory obligations regarding signal quality. Consequently, ERDF must deliver the number of clients poorly supplied and account for all events, shortages or power violations on its network. The qualification of these events requires the cross-checking of this data with the network tree to identify the origin of these events at each level:

- primary substations
- MV feeder
- MV/LV substation
- Transmission network

These hundreds of millions of annual events (with daily peaks reaching over 10 million) are collected by customer meters.
The increasing instrumentation of the network therefore participates in the growing generation of data.

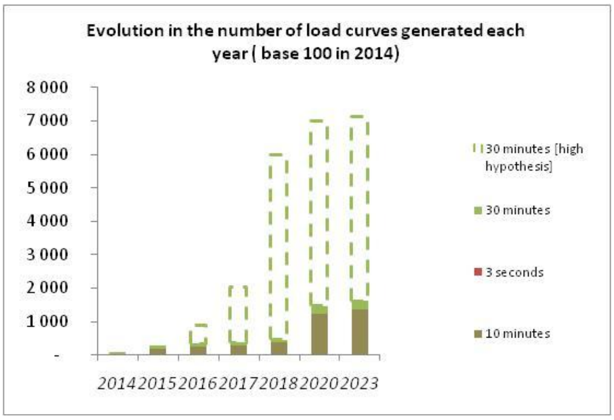


Figure 1: Load curve generation

The needs are reinforced by the new market mechanics and the will to control energy losses

The mechanics of demand response (on the distribution network) and the diminution of energy losses on the network (for which ERDF supports the cost) are national activities demanding the consolidation of various data sets coming from all geographic regions.

Demand response: a new contingency when managing the network

The French electric consumption is characterised by the existence of a particularly important daily and yearly peak. This specificity in the French system is reinforced by the strong thermo-sensitivity of French demand (roughly 2300MW per Celsius degree in winter). In addition to the development of peak production capabilities, demand response allows the modulation of demand and offers added flexibility when addressing extreme consumption peaks.

Demand response mechanisms give actors the possibility of regrouping and modulating the consumption of large numbers of small consumers and value such groups on the market. The challenge behind these mechanics consists, for the distributor, in reconstituting the consumption of all the nodes that were capped, in correcting these power limitations to properly deliver energy to those sites' suppliers and apply limitations to the relevant actors. These energy allocations use consumer panels and statistical models. This new mechanism implies a need for vast data sets representing the diversity of energy uses.

The need to control energy losses on the network

The closing of the yearly electric balance at the ERDF level shows losses of approximately 1TWh. This gap can be explained by losses due to warming (joule effect), iron losses (transformers), copper losses (coils) and by energy not accounted for (no metering systems, abnormalities, fraud …). To counter these losses (or balance differences), ERDF faces two issues:

1. Optimise forecasting regarding the energy balance so as to buy the corresponding energy, to respect its regulatory obligations, while avoiding any penalties imposed by the French Balancing Market rules (“Mécanisme d’Ajustement”);
2. Reduce the volume of non technical losses (i.e. excluding joule, iron and copper losses) in order to avoid unnecessary costs.

The management of energy purchases to compensate for losses relies on forecasting tools for each item in the electric balance. These econometric models use meteorological and calendar based data from the past several years (aggregated at the national level) as exogenous variables.

The reduction of non technical losses is a major objective for ERDF. The identification of these non losses relies on:

- cross-checking technical data with energy measures to identify sites where abnormal demand can be observed in comparison with consumption panels;
- cross-checking geographical and contractual data to identify consumption sites without a contract or without a metering system;
- The elaboration of local balances (at a regional, primary substation or smaller level) to identify zones with singular loss levels.

The aggregation of this data globally points towards the necessity to implement transverse data analysis infrastructures for all ERDF business lines.

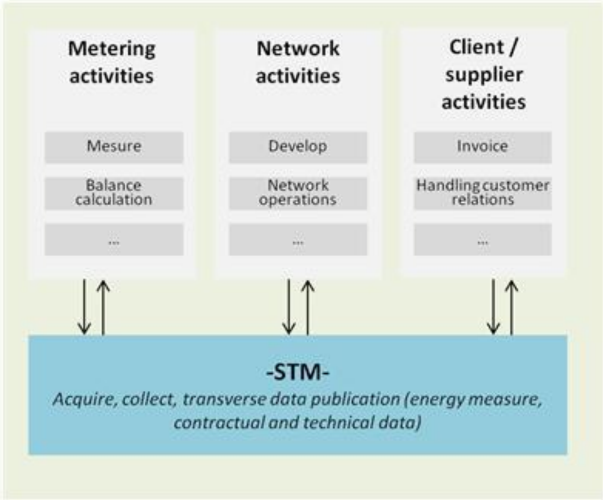
Reshaping the IT infrastructure around Big Data

In order to offer an efficient response to the problems exposed here above, ERDF started the STM (Système de traitement de masse) project in April 2012.

STM: effective management of all transverse data

This project aims at accompanying the transformations of all business lines of the network operator by implementing a market technology based Big Data solution. The deployment of such an infrastructure handling: energy, contractual, weather and network data will be done over several years and numerous versions. This infrastructure will support: the generation of energy balances, consumption / local production forecasting, signal quality analysis and the publication of energy data aimed towards the users / customers of the distribution network.

Figure 2: Simplified view of the ERDF IT architecture



Step 1: rationalisation of the ERDF IT system

The first version of STM was launched in March 2014 with the goal of offering the basis for transverse data handling. This first step has allowed the capitalisation of all the work done on data modelling by the network and metering teams. It has also permitted the aggregation of 10 years of data regarding:

- Metering and contractual information for all consumers and producers linked to the distribution network
- Description of the connection and metering systems of all sites (consumers and producers).

The other versions planned for 2014 – 2015 (2 versions per year) are aimed at simplifying the applicative structure of the IT system: 9 applications should be replaced by STM.

Next steps: extension of the functional perimeter

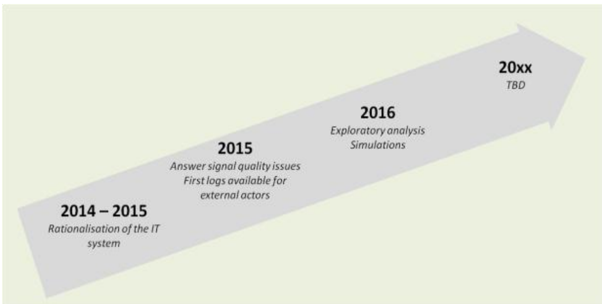


Figure 3: The major steps of the STM project

After the rationalisation of the IT system and the drastic reduction in the number of available applications with redundant functions, the STM trajectory plans:

- In 2015, answering the need for a complete collection of outage events / voltage violations on all network nodes equipped with smart metering systems. By this date STM should be able to qualify these events and identify their origins.
- In 2016, have analytical tools available for operational teams to help them reduce non technical energy losses. There will also be means of simulating electric balances and local forecasts (for producers and consumers)

Some first tangible results

The first financial gains related to the infrastructure and the reduction in the number of applications are expected as soon as 2014. Furthermore, the extended performance of the STM solution will also allow the first publication of data for external actors (10 half hourly electric balances can be produced each week for the balance responsible parties and the Transmission System Operator). Moreover, the first queries/analyses were produced 10 times faster than what the historic system could provide. These studies where in particular aimed at detecting abnormalities and therefore reducing non technical losses.

This calculating capacity, in time, should facilitate network operations thanks to the daily production of local forecasts at D+1 (per primary substation, MV feeder, or MV/LV substation)

CONCLUSION

The transformations of the ERDF IT system based on the incorporation of Big Data technologies is guided by the will to secure all of the traditional distributing activities while controlling the network and implementing new market mechanisms. Finally, this reshaping of the IT system revolving around better data management puts ERDF in a position where it can encourage the energy transition by offering new energy services to all actors.