Vaasan yliopisto
UNIVERSITY OF VAASA

OSUVA Open Science

This is a self-archived – parallel published version of this article in the publication archive of the University of Vaasa. It might differ from the original.

Advanced utilization of big data for real-time monitoring and data analytics in Sundom smart grid

Author(s):   Hovila, Petri; Monot, Aurelien; Laaksonen, Hannu; Rita-Kasari, Matti

Title: Advanced utilization of big data for real-time monitoring and data analytics in Sundom smart grid

Year: 2019

Version: Publisher's PDF

Please cite the original version:

Hovila, P., Monot, A., Laaksonen, H., & Rita-Kasari, M., (2019). Advanced utilization of big data for real-time monitoring and data analytics in Sundom smart grid. In: *Proceedings of 25th International Conference on Electricity Distribution: CIRED 2019 : Madrid, 3-6 June 2019*. Conference proceedings CIRED 0744. https://www.cired-repository.org/handle/20.500.12455/134

ADVANCED UTILIZATION OF BIG DATA FOR REAL-TIME MONITORING AND DATA ANALYTICS IN SUNDOM SMART GRID

Petri HOVILA
ABB – Finland
petri.hovila@fi.abb.com

Aurelien MONOT
ABB Corporate Research – Switzerland
aurelien.monot@ch.abb.com

Hannu LAAKSONEN
University of Vaasa – Finland
hannu.laaksonen@uwasa.fi

Matti RITA-KASARI
Jubic – Finland
matti.rita-kasari@jubic.fi

ABSTRACT

Modern technologies are increasingly used in many protection, control and monitoring solutions for power distribution grids. Voltage and current measurements are nowadays usually converted to digital signals, which are further processed and stored in different locations depending on the application. At the same time, when data processing and storing capacity is increasing, the number of new data sources also increases. This means that the data transfer capacity may also need to be increased at different levels in future distribution network management and protection architectures. This paper presents how IEC 61850-9-2 standard based raw data streamed from the Sundom Smart Grid can be processed and utilized for power quality monitoring and analytics applications. Digitalization also enables several other parallel applications since the same data stream can be used for several targets when the connectivity, interoperability and data quality are sufficient. In this paper some of these potential parallel applications are also introduced.

INTRODUCTION

In the future, distributed energy resources (DER) and real-time measurements in active network management (ANM) will have a key role in improving electricity supply reliability and local distribution grid resiliency. Accurate real-time measurements and utilization of fast communication technologies will enable the real-time response of the DER as well as improved monitoring and state-estimation. Accurate state estimation is an important input for different ANM functionalities in distribution grids, as well as for potential future big data analytics (BDA) based enhanced monitoring and predictive protection solutions.

Handling the data generated in smart grids with traditional analyses is challenging, in particular to produce actionable information within required timeframes. In general the objective of BDA is to investigate very large volumes of data from multiple smart grid components and transform the data into meaningful inputs like patterns, trends, fault detection and control commands for machine learning applications. The outcome of these applications may identify operating trends leading to failure patterns of devices or components and in this way enable new proactive protection and predictive maintenance solutions [1].

Power Quality (PQ) monitoring plays an important role in ensuring reliable operation and electricity supply of distribution grids [2]. In the future, BDA based proactive protection functionalities will be increasingly based on advanced, real-time utilization of centralized monitoring of many different PQ indices simultaneously from multiple measurement points. Centralized monitoring of PQ parameter trends, variations and temperature measurements from different feeders and measurement points can potentially be enabled to detect faults before they happen, improve fault location and detect measurement errors. In addition, it can enable supply level differentiation for customers and verification of the realized power quality for these purposes. In the future, analytics comprising BDA, machine learning and artificial intelligence has a significant role in making the smart grids more intelligent, cost- and energy efficient.

One possibility is to utilize calculations which are integrated into IEDs and send pre-processed values by GOOSE messages. This approach reduces the amount of data, but all calculations must be integrated into the IEDs. As a proof of concept, we developed DAGR (Data Analytics for Smart Grids), a prototype solution for PQ monitoring based on IEC 61850-9-2 Sampled Value (SV) stream data.

SUNDOM SMART GRID

The developed DAGR platform is running in the Sundom Smart Grid (SSG) pilot of ABB Oy, Vaasan Sähkö (local DSO), Elisa (telecommunication/IT company) and University of Vaasa in Vaasa, Finland [3]-[6]. The goal of the SSG pilot has been to enhance the electricity supply reliability by new grid automation solutions for more accurate earth-fault detection and localization in mixed compensated distribution grids (overhead, OH-line & cable). In addition, the SSG focus has been to integrate and management of renewables and other DER units in an intelligent way.

The Sundom Smart Grid consists of one primary HV/MV substation and four secondary MV/LV substations (Fig. 1). Today there are two distributed generation (DG) units

connected to SSG (Fig. 1). One full-power-converter based wind turbine (3.6 MW) connected to MV network with own MV feeder J08 (Fig. 1) and another LV network connected inverter based PV unit (33 kW) at MV/LV substation TR4318 (Fig. 1).

On top of the IEDs required to protect and operate the SSG, additional functionality in IEDs have been activated in the substations in order to capture the IEEE 1588 time-synchronized measurements of current and voltage values from multiple points and transfer them to a data center using the IEC 61850-9-2 standard protocol. Additional IEDs have also been installed to secondary substations. This system sends its data in real time through the optical fiber network (Fig. 1). All the data are sent to a centralized server (Fig. 1). In total, 20 IEDs are sending IEC 61850-9-2 SV streams with a sampling rate of 4 kHz, as well as GOOSE data.
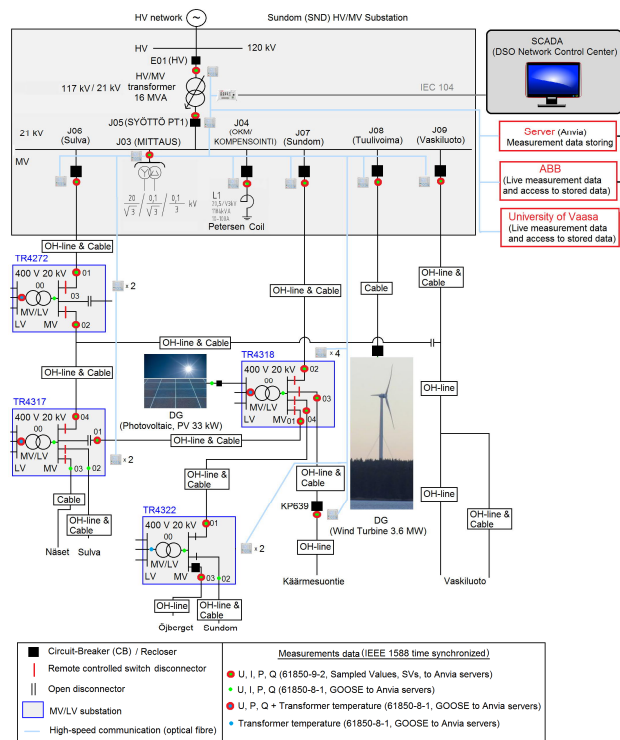

Figure 1. Sundom Smart Grid single line diagram (SLD) and data measurement points.

REAL-TIME POWER QUALITY MONITORING

To demonstrate that we could handle big data processing in real-time, we developed DAGR a solution for PQ monitoring. As defined by the IEC 61000 standards, we are computing the following parameters every ten periods: power, reactive power, frequency, total and fundamental signal RMS as well as THD and harmonics (up to the 40th) for current and voltage.

DAGR software architecture

A combination of open-source libraries, existing ABB libraries, as well as newly developed software components
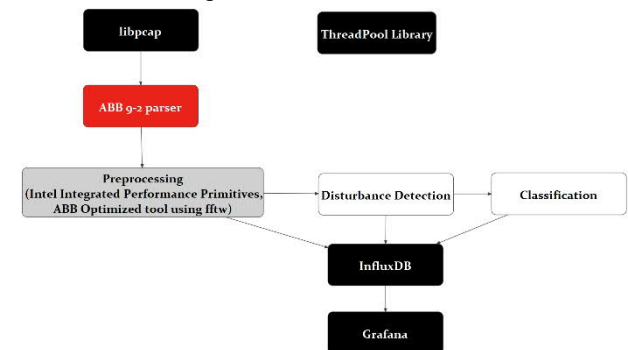
are used in our implementation.

Figure 2. DAGR software components. Open-Source software is in black, existing ABB code in red, code developed reusing some existing code in grey, and code developed from scratch in white.

Fig. 2 shows how the data flows from one software component to another. The incoming sample value streams are captured on the Ethernet port using libpcap (open source). The Ethernet frames are then parsed using an existing ABB internal library. We adapted and extended a library from ABB, which contains optimized implementations of preprocessing functions such as FFTW (Faster Fourier Transform in the West). The resulting computed parameters are stored using the open source time-series database InfluxDB. A parallel data flow evaluates the resulting parameters to detect power quality disturbances using a software component developed during the project. If a disturbance is detected, the raw data corresponding to the detected disturbance is stored in another running instance of InfluxDB dedicated to power quality disturbances. At the same time, we also run another software component to classify the type of disturbance, as described next.

PQ disturbance classification

We implemented a mechanism to detect and classify disturbances based on [7]. Specifically, a rule-based decision tree is used for classification, where the rules are established based on electrical engineering knowledge. This classifier was chosen because it is efficient and can therefore be applied in real-time, and because it uses parameters, namely FFT and RMS, that are already computed. A test with simulated signals resulted in classification accuracy of 94% for a sound to noise ratio of 30dB. Furthermore, we compute the duration of the detected disturbance and store the corresponding raw data so that further analyses can be run.

Handling big data in real-time

The efficient processing of the power quality parameters was obtained by using multi-threading while avoiding concurrency issues by design, and thus avoiding the utilization of mutex (mutual exclusion) primitives which would have slow down the process significantly.

The raw data is extracted using an existing IEC61850-9-2

parser on the raw network data. Data coming from each sample value streams (i.e., measurement point) is copied in its corresponding ring buffer. There are as many ring buffers as sample value streams which is static configuration parameter described in a file.

Fig. 3 describes the different threads running in parallel as well as the data of the batch of sample each thread is handling. The insertion "fill" threads (in green) are running on batches of size X. The "processing" threads (in orange) are working with batches of size Y. And the "store" threads (in red) are working on batches of size Z. A fourth type of "query" threads (in blue) are responsible of sending data to the databases. As doing single insertions in the database is costly, doing insertions by batches allows to use the time-series database more efficiently.
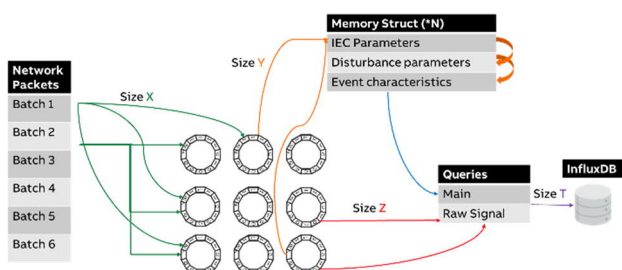

Figure 3. Batch processing of the raw data.

Doing such batch processing allows to improve performance for two reasons. First, processing data in batches allows to decrease the fixed overhead of certain operations such as doing requests to the database. Secondly, this allows to take advantage of the CPU cache and get a significant speed-up, in particular when processing the raw data to compute power quality metrics using math algorithms such as a Fourier Transform.

Fig. 4 shows the batch sizes used in DAGR starting from the raw data. For a 50Hz line, each input stream amounts to 4000 samples per second corresponding to 800 samples for 10 cycles, the window defined in the standard to compute the relevant metrics for PQ monitoring. Each processing thread operates on batches of between 4 000 and 12 000 samples. Finally, storage is triggered when metrics are computed on the equivalent of between 320 000 and 640 000 samples.
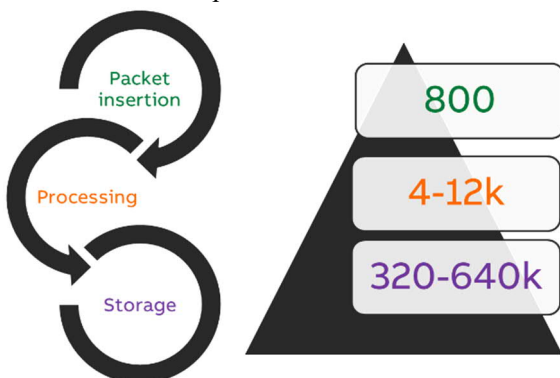

Figure 4. Batch processing of the raw data.

Performance and sizing

These implementation optimizations allow to scale our solution in multiple ways. Test experiments have shown that our implementation can handle up to 80 SV streams on a COTS PC (Quad-Core i5), which is 4 times the size of the Sundom Smart Grid, or up to 10 streams on a raspberry pi 3.

Table 1 describes the required amount of raw data to transfer and process per stream and for our smart grid pilot. We focus here on the SV streams as the GOOSE bandwidth is negligible in comparison. For this estimation, we use SV packets of 140 bytes (B), the maximum size, and a sampling rate of 4 kHz.

Table 1. Amount of raw data to be transferred and processed for SV streams with current and intensity with three phases and neutral.

Time span	Per SV stream	For the Sundom Smart Grid
1 second	~ 500 KB	~ 10 MB
1 day	~ 40-45 GB	~ 800 – 900 GB
1 year	~ 15 TB	~ 300 TB

When running DAGR, we are storing around 30GB of data per day in our database. The computed metrics and raw signal (in case of disturbance) are stored using InfluxDB, an open-source real-time database and are displayed using Grafana as shown in Fig. 5.

Figure 5. Screenshot of the interface of the DAGR tool displaying power quality parameters.

Although we observe a reduction of a factor around 30 in data size between the raw signal and the resulting metrics, this results in a significant amount of data. This amounts to around 10 TB per year. Since the metrics are computed every 10 cycles, we have a granularity of 200ms for a grid operating a 50 Hz. Depending on the use case, such granularity might not always be needed. Down-sampling the output of our solution would allow to reduce the required storage space greatly, while still being detailed enough for the considered use cases.

FUTURE APPLICATIONS

In this paper we address the possibility to utilize SV streams from multiple locations as a source for advanced analytics and real-time monitoring of distribution network. Since the amount of raw real-time data output as specified in IEC 61850-9-2 is large, using these sources for big data applications requires smart filtering and pre-processing of data for different applications (Table 2).

Table 2. Different type of data (data filtering and pre-processing) and data storage needed in future big data applications in smart grids.

Data storage time	Application examples	Type of data
ms - sec	Real time protection	samples or pre-processed
sec - min	Local monitoring	samples or pre-processed
hours - days	Local statistics	e.g. averages
years	Historically based statistics and analytics	Snapshot of samples, pre-processed or average

In addition to PQ measurements, other kind of digitized data (on-load-tap-changer, OLTC, number of operations, temperature and thermal and noise, voice sensors measurements of network components like feeders, transformers, circuit-breakers etc.) could be stored and utilized with advanced IoT and big data-analytic tools and cloud-based platforms to enable future proactive protection and predictive maintenance functionalities of smart digital primary and secondary substations. This will enhance the electricity supply reliability as well as cost-material- and energy-efficiency of future smart grids.

In general, IEEE 1588 time-synchronized IEC 61850-9-2 SV current and voltage measurements could be also used in a similar manner as synchrophasor or phasor measurement unit (PMU) measurements. Synchrophasor applications are generally categorized as 1) Offline applications (disturbance analysis and power system dynamic model tuning and validation), 2) Online monitoring applications (enhanced power system state estimation, frequency and phase angle monitoring, line thermal loading monitoring, voltage instability monitoring, and oscillation monitoring), and 3) Real time control/protection applications (power flow control, oscillation and damping control, and emergency control against voltage, rotor angle, or frequency instability). Recent developments in ICT technologies and infrastructures increased the feasibility of synchrophasor applications. Through improved real-time state-estimation and topology verification by accurate, time-synchronized synchrophasor-/PMU measurements, distribution network the DG hosting capacity could be increased without network reinforcements, as well as enable the utilization of active network management (ANM) functionalities in large-scale.

CONCLUSIONS

In general, centralized monitoring of PQ parameter trends, variations and temperature measurements from different feeders and measurement points could be increasingly used in the future for proactive protection and predictive maintenance functionalities by big data analytics tools. In terms of data transfer and processing, it is important to take into account the needs of different applications. In the future, local and distributed control architectures can provide solutions that can reduce the data transmission load and computational resources required by fully centralized solutions.

This paper presented how IEC 61850-9-2 standard based raw data streams could be processed hierarchically and utilized in PQ monitoring and analytics application. As a proof of concept, DAGR (Data Analytics for Smart Grids), a prototype solution for PQ monitoring based on IEC 61850-9-2 Sampled Value (SV) stream data was developed for the Sundom Smart Grid pilot. At the end of the paper some other potential parallel applications utilizing the same SV data stream were also introduced.

REFERENCES

[1] IEEE Smart Grid Big Data Analytics, Machine Learning and Artificial Intelligence in the Smart Grid Working Group, 2017, "Big Data Analytics in the Smart Grid, IEEE White paper", (https://smartgrid.ieee.org/resources/white-papers/big-data-analytics-in-the-smart-grid)

[2] E. Gasch, J. Meyer, P. Schegner, K. Schmidt, 2015, "Efficient Power Quality Analysis of Big Data (Case Study for a Distribution Network Operator", 23rd International Conference on Electricity Distribution (CIRED 2015), Lyon, France.

[3] A. Monot, M. Wahler, J. Valtari, M. Rita-Kasari and J. Nikko, 2016, "Real-time research lab in the Sundom smart grid pilot", CIRED 2016 Workshop, Helsinki, Finland.

[4] H. Laaksonen, P. Hovila, 2016, "FlexZone Concept to Enable Resilient Distribution Grids – Possibilities in Sundom Smart Grid", CIRED 2016 Workshop, Helsinki, Finland.

[5] H. Laaksonen, P. Hovila, 2017, "Future-proof Islanding Detection Schemes in Sundom Smart Grid", 24th International Conference on Electricity Distribution (CIRED 2017), Glasgow, Scotland.

[6] H. Laaksonen, P. Hovila, K. Kauhaniemi, K. Sirviö, 2018, "Advanced Islanding Detection in Grid Interactive Microgrids", CIRED Workshop 2018, Ljubljana, Slovenia.

[7] M. Zhang, K. Li and Y. Hu, 2011, "A Real-Time Classification Method of Power Quality Disturbances", Electric Power Systems Research vol. 81:2, 660-666.