Distributed edge computing paradigm with dedicated devices for energy efficiency and predictive maintenance applications

Month 05, 2018

Sarantis Kotsilitis[1,2,3], Effie Marcoulaki[1], Emmanouil Kalligeros[2], Yorgos Mousmoulas[3]

[1] System Reliability and Industrial Safety Laboratory, National Centre for Scientific Research “Demokritos”, Athens, Greece; [2] Department of Information and Communication Systems Engineering, University of the Aegean, Samos, Greece; [3] Plegma Labs, Athens, Greece

ABSTRACT

This paper presents a new project to deliver an advanced and innovative solution for energy disaggregation intended for industrial applications, commercial buildings and households. The PREDIVIS project aims to address the problem of predictive maintenance by providing a cost-efficient solution compared to traditional solutions with multiple sensor systems. This solution involves the measurement and analysis of aggregated signals, and the application of suitable disaggregation technologies to retrieve individual device signals. The paper discusses the proposed architecture plan involving intelligent edge devices, hardware accelerated algorithms and cloud computing services using advanced data processing analytics.

Keywords: Edge Computing, Internet of Things, Predictive maintenance, Energy efficiency, Energy disaggregation, Device health monitoring.

INTRODUCTION

During the last decade most ICT enterprises and organizations were focused on migrating to cloud solutions and architectures. Cloud solutions are still dominating the field of ICT to provide adequate storage, computational power and centralization features. Cloud computing technologies have been used on a plethora of cases, which has led huge corporations (e.g. IBM, Microsoft etc.) to migrate or implement their own solutions. In most cases, the cost for the infrastructure rises exponentially when new features or additional data are integrated into systems, leading to hybrid solutions that use cloud infrastructures solely for data storage and centralization, and private ones for data processing and intense computational applications.

Monitoring of energy consumption is crucial for companies, organizations, industries and individuals towards energy efficiency. Project PREDIVIS, herein presented, aims to deliver a novel solution for energy disaggregation, energy efficiency and predictive maintenance based on major ICT technologies such as IoT, Edge computing, Microelectronics, Digital Circuits, and Deep Learning. Energy disaggregation, also referred as Non-Intrusive Load Monitoring (NILM), is the processing of an aggregated electric load signal metered at a single point. By analyzing the transitions of load at that point (line), we can identify the devices used at a certain time.

The PREDIVIS project is a collaboration between three partners: an IoT company (Plegma Labs S.A.), a research institution (NCSR “Demokritos”) and a university (University of the Aegean, Dept. of Information and Communication Syst. Engineering). The goal of this project is to deliver an advanced and innovative solution for energy disaggregation intended for industrial applications, commercial buildings and households. The project will create a smart energy analyzer sampling device working at a high frequency sampling rate (64KHz) to preprocess the majority of the data on site. The following sections present the proposed architecture plan, its benefits and how the project aims to address the problem of predictive maintenance by providing a cost-efficient solution, compared to multiple sensor systems.

STATE OF THE ART

IoT devices are producing, on a daily basis, vast amounts of data that are transferred to central nodes/ units for storage and processing. According to CISCO’s forecast [1], the traffic will reach

3.3 ZB/year by 2021. Due to the huge amount of data produced, IoT technology is often accompanied by Big Data [2] techniques to handle the data storage and processing. The most common practice for sensor data acquisition systems is to harvest the data and transmit them in raw format to some central infrastructure, demanding, most of the times, huge upload speeds to deliver near real-time feedback to users.

Since Hart [3] first introduced the concept of NILM in the 90's (1992), numerous techniques have been developed to address the problem [4, 5, 6]. The main concept of data analysis has been to utilize smart devices such as energy meters, at various sampling rates, from 1 measurement of total power every 15 minutes up to a couple of MHz of electric waveforms' samples. Most of the times, the collected data are transferred to a cloud infrastructure in order to be processed. To minimize the amount of data transferred, various compression techniques [7, 8, 9] and system architectures have been used. The type of applied data analysis varies with the sampling rate, due to the different amount of information that can be extracted from data of higher resolution.

Most systems designed for energy disaggregation [10, 11] and industrial monitoring [12, 13] consist of a sensor network sampling various kind of data, and a gateway system that collects the data and either stores them locally or sends them to a cloud infrastructure. The collected data are processed to extract features and knowledge. The majority of previous applications transferred the entire volume of data to the cloud, where the data were stored and analyzed in real time. This procedure required a lot of resources for data transmission, remote storage, analysis and processing, and led to delays of up to a couple of days.

Since 2013, major companies in ICT and Industry 4.0 have realized the limitations of cloud infrastructures and moved towards distributed architectures referred as edge computing or fog computing, powered by the advances in telecommunications (4G, LTE, 5G) and broadband connection speed. Edge and fog systems differ by the way their architecture assigns the system's computing power: edge systems assign power to data gathering devices, whereas fog computing systems assign the power to a local area network, where the data are processed within a hub, node or a gateway [14,22]. Distributed architectures maintain the main features of cloud computing, but they offload the overall system in terms of storage, bandwidth and computational needs. They are also more tolerant to internet service failures [17], more secure, and can shift control and intelligence away from central nodes to the devices or even the sensors.

By offloading the system, a company/organization can lower the infrastructure costs because less data will be transmitted remotely. By processing data on site, the system can share information between sensors or deployed devices within seconds or even milliseconds. The recent increase in power of microcontrollers, gateways and even low-power processors makes these types of architectures capable of delivering high quality of services by processing large amounts of data.

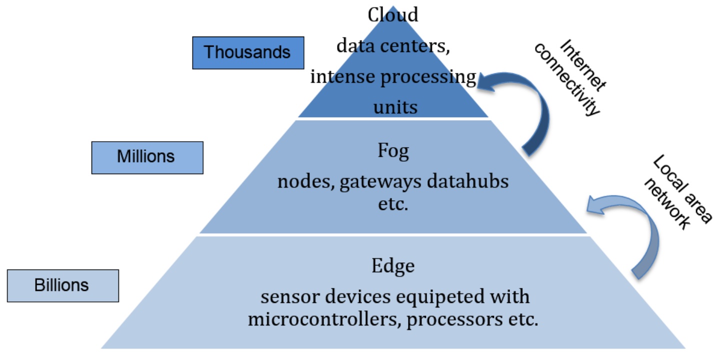

Figure 1 Cloud, fog and edge architectures pyramid, in terms of number of devices and type of connectivities

Approaches on industrial monitoring for preventive and predictive maintenance [10, 11] with NILM methods have been developed with promising results on the field. Such approaches feature high frequency sampling, computing power level characteristics like active and reactive power, power factor, harmonics voltage imbalances etc. Furthermore, NILM systems can deliver adequate power quality parameters to analyze electric powered machinery. By monitoring simultaneously multiple machinery of the same group or with adequate related components, it is easier to observe anomalies and faults. Aside from that, power analysis on the production chain can give insights and support actions of load balancing, energy efficiency and job scheduling.

SYSTEM ARCHITECTURE

This work will combine existing advanced ICT technologies to deliver a new ecosystem based on edge computing appliance detection, using a custom Field Programmable Gate Array (FPGA)-based design to accelerate neural networks and a centralized unit for intense analysis towards predictive maintenance and energy efficiency.

This section analyses the architecture of PREDIVIS as an edge-cloud hybrid computing system, as well as specific technologies used for data collection, storage, analysis and processing.

The proposed system’s architecture consists of several components working together to deliver a set of services. It comprises high performance data sampling and processing devices referred to as agents, a central node to deliver the service to users and a software suite. The agents of the system are compact, low-energy devices that act as data collectors and data processors, while they also communicate with the central node. An Analog-to-Digital (ADC) conversion component is used as a multichannel sensor to sample the voltage and current waveforms. A System-on-Chip (SoC) FPGA collects the data and implements advanced Neural Network (NN) structure and algorithms for data processing and energy disaggregation. This device integrates both processing cores and programmable logic in the same piece of silicon, and is responsible for backing up data for a certain amount of time and transmitting results to the central node. The flexibility to reprogram the FPGA with different NNs and increase the systems accuracy allows Continuous Delivery and Integration of more accurate models for each deployment site, compared to the case of using dedicated Application Specific Integrated Circuits (ASICs). FPGAs offer some important advantages compared to ASICs. A high-level comparison is provided in Table 1.

Property	**FPGA**	**ASIC**
Time to market	Fast	Slow
Reconfigurable	Yes	No
Application type	General	Specific
Unit cost	High	Low
Performance	Medium	High
Connection with multiple of-the-self electronics	Yes	No

Table 1 High level comparison between FPGA and ASIC devices

The central node is responsible for storing the data transmitted by the agents and for delivering a visualization platform to the users. Additionally, it is used to fine-tune the NN models on the agents by receiving feedback from the users and data from other agents. Moreover, the central node analyzes further the received data to deliver advanced insight and is responsible for helping an agent distinguish different appliances with similar behavior and electrical characteristics. On the central node, the Business Intelligence of energy efficiency analyzes data to provide personalized suggestions to users, in order to lower their carbon footprint. Finally, in cases where the NNs on the network edge fail to provide results, raw data are transferred to the central node to be analyzed and compared with other agents’ data to retrain the model.

Property	EDGE	Cloud
Internet connection	Optional	Mandatory
Bandwidth	Low	High
Storage	Small	Large
Security	High	Medium
Architecture	Decentralized	Centralized

Table 2 Comparison of EDGE and Cloud systems

The above architecture has many advantages compared to widely used cloud systems (Table 2), even though there are certain limitations on what we can achieve, mainly in terms of the local system availability. As the number of deployed devices increases, so do the points of failure, calling for advanced techniques to monitor and mitigate failures. The agent also requires power outage measures in order to provide the service without interruptions that might affect the devices on the deployment site, their health and the quality of service. Beyond the availability drawbacks, edge devices are capable of processing finite volumes of data, so different hardware/software combinations for different locations and deployment sites may be required.

Another component of our system is the library of disaggregation and predictive maintenance algorithms, where multiple cases can run in parallel to search for superior alternatives, when the accuracy of a node is below a certain threshold or if the user specifies it. This module is responsible for keeping metrics for the devices and monitor their health. Such metrics can be the working hours of a device, power shortages / anomalies, etc. A simpler, less sophisticated approach of the predictive maintenance algorithms will reside on the edge devices. This architecture will reduce the amount of data transferred to the central node from some GBs per day to a few keypair timestamp-labeled values, minimizing the volume of data, as illustrated in the next section.

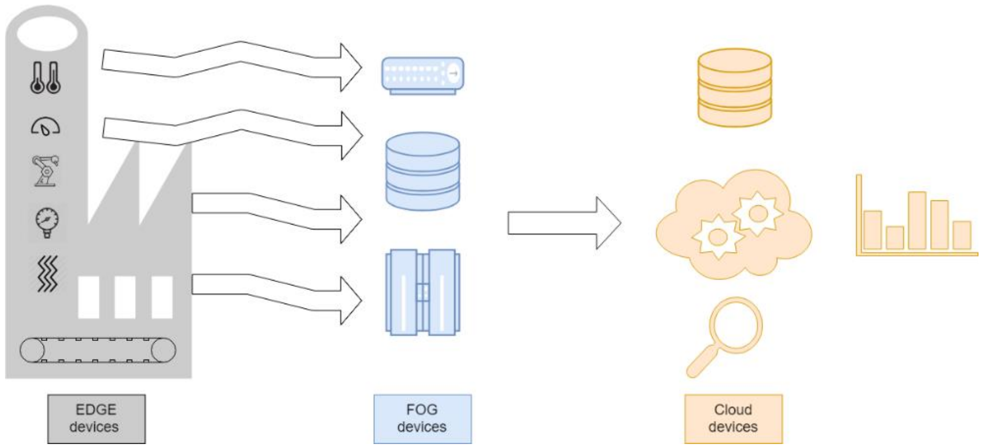

Figure 2 Predictive maintenance application architecture

Most of the times, industrial production systems are using advanced and complicated sensor systems to harvest data from machinery like conveyor belts and motor drive equipment [18,19] described in Figure 2. Data collected from sensors will further on be analyzed using tailored models to produce knowledge on the equipment health. Most common practices for motor drive equipment are vibration sensors on multiple points, temperature and soundwave monitoring devices. Approaches on industrial health monitoring through current analysis have been made using Motor Current Signature Analysis [20] (MCSA), Instantaneous Power Signature Analysis [21] (IPSA), power quality etc., to predict machine fault. Theses sensors need to be installed on every device to harvest data. NILM technology can provide a Non-Intrusive method to harvest this kind of data on newer and older machinery devices, giving the advantage of retrofitting to older machinery and outdated systems decreasing the overall cost of installation and monitoring.

DATA LOAD ESTIMATION

Preliminary analysis of the system's tradeoff has shown that the expected reduction of transmitted data, using the proposed edge computing architecture, is reduced by at least six orders of magnitude. The amount of raw data to be transferred is in the magnitude of several GBs per day.

The current load, L, of transmitted data per time unit can be calculated as follows:

$$L = F * (Np * Nw * R)$$

where *Np* is the number of alternative current phases sampled, *Nw* is the number of waveforms sampled for each phase, *F* is the sampling frequency used, and *R* is the resolution in bits for each channel. Assuming a typical 3-phase household installation, and setting the sampling frequency at 64 kHz and the resolution at 16 bits per channel to sample both current and voltage waveforms of each phase, the data load becomes:

$$L = \frac{64000}{s} * (3 * 2 * 16\, bits\,) = 768\frac{Kbytes}{s}$$

Most household internet connections have a fixed limit of 1 Mbit/s upload and larger upload speed connections often increase exponentially the internet connection cost. This means that every second the connection can transmit 125 Kbytes, which is insufficient to stream the raw data load to the cloud. By shifting the processing of data to the edge, the data that will finally be transferred are significantly limited. These data comprise the device on/off events, the transitions in device working states (e.g., a washing machine transition from washing to rinsing). Consider a typical household with a number of active devices ranging from 7 to 20, and a time window, *T*, equal to one hour. During *T*, each active device will produce zero (if activated before the beginning of *T* and stayed at the same state the whole time) to five (for a washing machine) events. The additional data include the working time of each device, the active and reactive power, voltage imbalances etc. In total, for each device, the metadata volume to be transmitted to the cloud during *T* is normally up to 120 Kbytes per hour. It is therefore shown that, data preprocessing on site can reduce the amount of data to be transferred by at least 6 orders of magnitude, leading to a few Kbytes per day.

CONCLUSION

This paper discussed the architecture of PREDIVIS project. The paper presented both software and hardware component architectures that can address the problem of data transfer and storage by processing data on the Edge. NILM technology can be a cost efficient Non-Intrusive method for predictive and preventive maintenance and machinery health monitoring for industrial, commercial and household environments.

Future work will concentrate on finalizing the design of the architecture on both software and hardware components, deciding which data should be collected, and installing agents in different sites to collect data. The major challenge will be then to develop appropriate data processing algorithms in order to complete the ecosystem for energy disaggregation and predictive maintenance.

This work will collect a vast amount of high-frequency data (64KHz) over the next years to create a data warehouse of multiple and heterogeneous instances of households, commercial and industrial sites. The data will include manufacturer specifications, historical data (time to failure and repair/maintenance), and data collected during system operation. The latter will consist of time series of carefully selected performance indicators depending on the device. The collected data will be published online to be used by interested parties.

ACKNOWLEDGEMENTS

This work is financially supported by the Programme of Industrial Research Fellowships of the Stavros Niarchos Foundation.

REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper.

[2] Ahmed, E., Yaqoob, I., Hashem, I. A. T., Khan, I., Ahmed, A. I. A., Imran, M., & Vasilakos, A. V. (2017). The role of big data analytics in Internet of Things. Computer Networks, 129, 459-471.

[3] Hart, G. W. (1992). Nonintrusive appliance load monitoring. Proceedings of the IEEE, 80(12), 1870-1891.

[4] Kolter, J. Z., Batra, S., & Ng, A. Y. (2010). Energy disaggregation via discriminative sparse coding. In Advances in Neural Information Processing Systems (pp. 1153-1161)

[5] Barsim, K. S., Streubel, R., & Yang, B. (2014, June). An approach for unsupervised non-intrusive load monitoring of residential appliances. In Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring.

[6] Bochao Zhao, Lina Stankovic, and Senior Member. On a Training-Less Solution for Non-Intrusive Appliance Load Monitoring Using Graph Signal Processing. IEEE Transactions on Smart Grid, 4,2016

[7] Clark, M., & Lampe, L. (2015, December). Single-channel compressive sampling of electrical data for non-intrusive load monitoring. In Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on (pp. 790-794). IEEE.

[8] Reinhardt, A. (2017, December). Adaptive load signature coding for electrical appliance monitoring over low-bandwidth communication channels. In 2017 Sustainable Internet and ICT for Sustainability (SustainIT) (pp. 1-8). IEEE..

[9] He, D., Lin, W., Liu, N., Harley, R. G., & Habetler, T. G. (2013). Incorporating non-intrusive load monitoring into building level demand response. IEEE Transactions on Smart Grid, 4(4), 1870-1877.

[10] Nguyen, T. K., Azarkh, I., Nicolle, B., Jacquemod, G., & Dekneuvel, E. (2018) Applying NIALM Technology to Predictive Maintenance for Industrial Machines.

[11] Panten, N., Abele, E., & Schweig, S. (2016). A Power Disaggregation Approach for Fine-grained Machine Energy Monitoring by System Identification. Procedia CIRP, 48, 325-330.

[12] Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23.

[13] Bagheri, B., Yang, S., Kao, H. A., & Lee, J. (2015). Cyber-physical systems architecture for self-aware machines in industry 4.0 environment. IFAC-PapersOnLine, 48(3), 1622-1627.

[14] Klonoff, D. C. (2017). Fog Computing and Edge Computing Architectures for Processing Data From Diabetes Devices Connected to the Medical Internet of Things.

[15] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637-646.

[16] Shi, W., & Dustdar, S. (2016). The promise of edge computing. Computer, 49(5), 78-81.

[17] Satyanarayanan, M. (2017). The emergence of edge computing. Computer, 50(1), 30-39.

[18] Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. Mechanical systems and signal processing, 42(1-2), 314-334.

[19] Goyal, D., & Pabla, B. S. (2016). The vibration monitoring methods and signal processing techniques for structural health monitoring: A review. Archives of Computational Methods in Engineering, 23(4), 585-594.

[20] Bravo-Imaz, I., Ardakani, H. D., Liu, Z., García-Arribas, A., Arnaiz, A., & Lee, J. (2017). Motor current signature analysis for gearbox condition monitoring under transient speeds using wavelet analysis and dual-level time synchronous averaging. Mechanical Systems and Signal Processing, 94, 73-84.

[21] Irfan, M., Saad, N., Ibrahim, R., & Asirvadam, V. S. (2017). Condition monitoring of induction motors via instantaneous power analysis. Journal of Intelligent Manufacturing, 28(6), 1259-1267.

[22] Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. Future Generation Computer Systems, 78, 680-698.