SMART BUILDINGS: USE CASE FOR MIDDLEWARE FOR DATA VISUALIZATION AND DATA ANALYTICS

Claudia CHITU[1], Valentin SGÂRCIU[2]

Buildings are becoming more intelligent and in the same time more complex, in due time with the advance in IoT developments and innovations. This means more interconnected devices and more challenges to enhance tenant comfort, to reduce energy consumption and to assert the financial well-being of the space. Therefore, research on Smart Buildings is relevant to address opportunities for IoT innovation in intelligent spaces, increase operational performance and improve efficiency. This work presents data platforms for dashboards and predictions trends, emphasizing multiple provocations in the field. Our objective is to build a middleware for scalable implementation to explore and visualize spatial and temporal data across buildings. We propose the understanding of the meaning behind data - using an open source backend tool as TICK stack with the front end of Chronograf or even Grafana which runs as a web application to create smart buildings sensing service.

Keywords: Building Management System, data platform, dashboard, data visualization, data analytics

1. Introduction

Smart buildings as a research topic faces complex challenges, from the lack of common approaches in industry standards to the need of dramatically reducing the effort of adding new functions and control capabilities in the building management system. Why is this important? Ubiquitous wireless networks, sensing grids, scalable data architectures, cheaper data storage and processing platforms would enable groups of buildings to behave cooperatively more rapidly.

Firstly, instead of defining a smart building concept as it is given in [1], we illustrate it with an example: the campus of University of California, Merced, committed to triple zero: zero net energy, zero waste and zero net greenhouse gas emissions [2], including 1 megawatt solar array, 1.4MW rooftop solar, 1M gallon thermal storage tank, 1MW hydro plant. A smart space within the building contains sensors to automatically adjust the room's light in concordance with the daylight, using adequately HVAC (Heating, Ventilation, Air Conditioning)

[1] PhD student, Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: claudia@imtt.pub.ro; Fulbright Visiting researcher at UC Merced, California, e-mail: cchitu@ucmerced.edu

[2] Prof., Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: vsgarciu@aii.pub.ro

systems [3] when a room is not fully occupied, running management strategies to reduce peak demand charges and also involving the human factor in the loop [4], [5]. Summarizing the attributes of a smart building, the definition of such intelligent building, which in [6] is considered to be an 'emerging' technology, should envelop more than remote control and automation, but highlighting the prediction and learning of user's behavior.

The reason given why these efforts are in place is that 90% of time is spent inside the buildings. In addition, another factor is that HVAC systems accounts for 42% of energy consumed by US buildings [7]. Looking at Europe, on the agenda for 2020 key objectives stand as: reducing greenhouse gas emissions by 20% comparing with 1990 levels and increasing with 20% the energetic efficiency [8].

This work comes to help researchers to better understand the concept of a smart building as a big picture and better understand the opportunities that come with the challenges on improving current building management systems, services and strategies. Our objective is to approach data driven decisions, by building a middleware to explore and visualize data at different granularity across the building.

In the next section, we introduce technologies for Smart Buildings; then, in section 3, we present data analytics approaches showing the use case using a time series data base suitable for IoT events, illustrating a dashboard. Section 4 will be dedicated to discussing this topic in a monitored environment including our contribution to one example of dashboard. The last section of this paper concludes with some remarks about the entire work and further plans.

2. Technologies for Smart Buildings

First of all, an abstracted layer of what smart building usually is based on, s shown in Fig. 1:

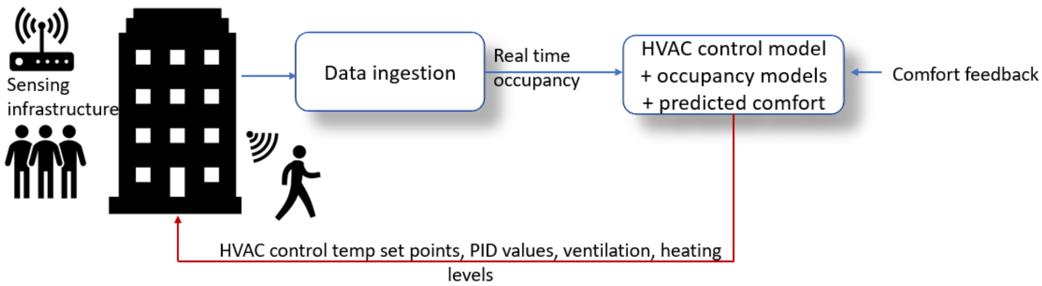

Fig. 1. Abstracted general architecture of smart building

Inside the building, sensing grids are deployed to collect data from occupants and to measure ambient values (temperature, humidity, air gases concentration etc.), which then feed the system. This data is then processed, sometimes with human feedback, and inputted into machine learning algorithms and then, the HVAC systems are controlled based on this information grabbed from the prediction and control algorithms. Reducing energy consumption, increasing energy efficiency and tenant comfort are objectives attained with the aid of granular data collected from the building data platform. The drivers that support business intelligence goals are sometimes on premises, and more common, on cloud data storages.

Environmental monitoring and building automation are two of the applications of wireless sensor networks which are defined as small devices with processing units [9]. The authors propose a LabVIEW interface to integrate this WSN (Wireless Sensor Network) with a smart house, using the XMesh networking protocol. Often, lighting systems and thermostats use Zigbee protocols (2.4GHz mesh LAN) [10]. Rather than being an IoT application protocol like ZigBee, 6LoWPAN is a network protocol [11], primarily used for home automation. In other applications, depending on the type of sensors, Wi-fi and Bluetooth are used. In [12], a complex comparative analysis and evaluation for Wi-fi (over IEEE 802.11), Bluetooth (over IEEE 802.15.1) and ZigBee (over IEEE 802.15.4) are performed. The last 2, are suitable for low data rate applications. Also, discussing security and monitoring in this context of deployment, integrating WSN is relevant. Some of the most relevant and practiced attacks have proved to be: sink hole attack, selective forwarding, wormholes, the Sybil attack. They could be very impactful, leading up to data loss. Thus, the sensor platforms should be featured with alerting systems to stop in the incipient phase the attack.

The next layer, in a building architecture, is represented by data and more precise, by the exchange of data. The Web RESTful service is used in practical approaches to enable the data exchange between physical devices, based on a time series database. In this type of architecture, measurement occurs in a 'measurement point' (ex.: a flow meter) and in a 'channel' (stream of scalar values coming from a measurement). In fact, this is the layer between Wi-Fi powered devices and web applications, mobile applications, being a common match with TinyOS. It is being utilized in many research groups to build on top of HVAC control. The HVAC control is then built using algorithms as Model Predictive Control (MPC) [13], [14].

3. Data analytics and visualization

Data analytics refer to applied statistics, probability theory (decision trees or Markov chain algorithms), machine learning (classification, clustering, predictive algorithms). A proper model to control a cyber physical system such as a smart building would engage optimization and prediction, not only anomaly detection, including data analytics approaches also.

Gigabytes of data from traditional and timeseries databases as InfluxDB [15] are streamed into numerous cloud repositories. To enable data driven decisions, complex data platforms are built by combining time series databases with RDMS (Relational Database Management System) and Big Data. When building a model for occupancy prediction, usage consumption etc., tenant comfort is analyzed from historical data storages often visualized on real time dashboards [16].

Data visualizations based on big data sources will be vital as the reality will present collaborative building groups. These visualizations will be consumed not only by the engineers, building facility managers, business owners, but also by the alert systems of the buildings, to control cyber physical systems which implies massive amounts of IoT data and time series format databases. With this perspective in plan, one has to take into consideration not only the interactive visualizations and dashboards but also the latency which could block automation or near-real time response to prevent critical situations. IoT data is characterized by three main attributes: real-time, time series and streaming. We have been previously discussing an architecture about streaming IoT data in cloud with automation in [16]. The arguments for which Timeseries databases are preferable for IoT applications over the traditional databases are numerous and they come from practice: massive scalability and performance, reduced downtime and improved business decisions because they provide real time data monitoring suitable for prediction and data driven decisions.

Depending on the need of keeping the data on disk or in memory (for faster queries), there are different solutions such as InfluxDB. This is part of an open source project called TICK (Telegraf, InfluxDB, Cronograpf and Kapacitor) stack – a platform to collect, analyze and action on time series data (see Fig. 2).

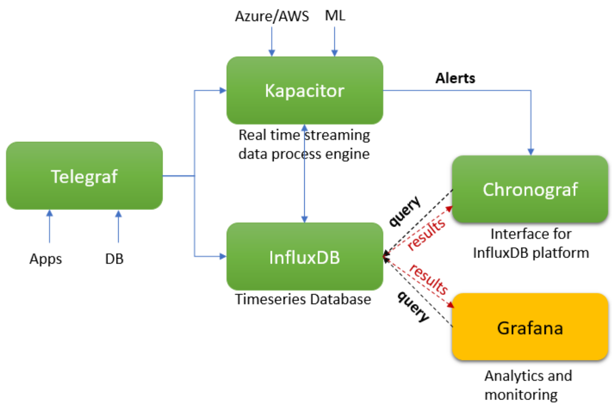

Fig. 2. Components of the TICK stack + Grafana (adapted from [20])

TICK stack illustrated in Fig. 2 presents the advantages of having multiple features (some of them vital for real time control): statistical analysis, prediction tool, data visualization, alerting system. InfluxDB presents the advantage of horizontally scale, without taking complexity of a situation similar to manage Hadoop clusters, supporting multiple ingestion protocols and data retention policies. In InfluxDB, data is stored in a format of key-value and has a SQL resembling query language. Although it is considered to be one of the most popular timeseries databases by a recent systematic survey [17], it deprecated the join operations and to better handle these types of situations it is indicated to use functions in Spark as some alternatives. The same ranking is demonstrated for 2018 in Fig. 3 by the DB-engines, comparing with other competitors for timeseries databases. The popularity score is calculating based on the following parameters: number of apparitions on the websites, Google trends, frequency of technical discussions about the time series, relevance of professional networks.

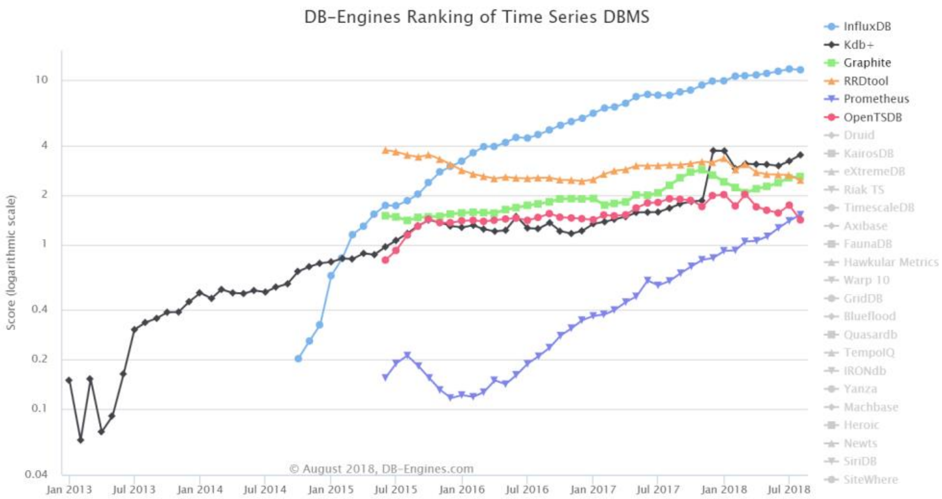

Fig. 3. Timeseries popularity [21]

Chronograf offers several types of visualizations: line graph, stacked graph, step-plot graph, single stat, gauge, bar graph, line graph + single stat, with options to select the color and add other features on the dashboard. We have created 2 of these charts in our dashboard presented in Fig. 5. The TICK stack is often integrated with Grafana which is more suitable for Big Data situations [15].

We have created one dashboard built in Chronograf, that illustrates the visual representation possibilities. This open source monitoring solution offers also features to set alerts and catch outliers, helping to short the time to get the value out of data and find useful insights, using a SQL resemble query style.

We used the version 1.6.1 (released in 2018) for Windows, also to test that the solution works well for this OS, because it is designed primarily for Linux and MAC OS.

4. Discussion

In this section, we present the results of our use case. Having an infrared sensing grid from Panasonic, named GridEye, deployed in a room, we collect temperature values from the grid containing 64 values in Celsius degrees. Based on these values we have modeled prediction data for occupancy detection and counting [18], [19]. The experiment deployment looks like in the Fig. 4, with the sensing grid placed on the top of the doorcase. For data collected from our sensors, we used an event platform for common metrics and alerts to detect threshold crossed values.

Fig. 4. Deployment topology

The platform is based on the TICK stack previously presented and the infrastructure monitoring is very useful, because often, when cleaning data for algorithms processing, visualization is very powerful way to fast capture anomalies and find patterns.

The interface from Chronograf presents data time with RFC3339-formatted timestamps, whilst the CLI (Command Line Interface) shows it in UTC format epoch nanosecond timestamps. This is suitable for our time series events, since we log data with a frequency of one event/second, recording 64 temperature values per frame, each second. Anomalies from the hardware or communication transmission could be simply detected with a plot on the Chronograf dashboard, having as a backend the Influx database. In Fig. 5, we have a screenshot of our dashboard. The plot from the right side of the dashboard, with the average of the 64 temperature values grabbed from the 64 sensors from the GridEye sensing device, has an alert system connected to send email when an outlier is detected.

Data was collected in 2018 as it is shown in Fig. 5 with a sample and the alerting system sends notification when the average values for the 64 grid cells from the sensor are not within common bounds (e.g. more than 40 °C). This phase is important in the cleaning stage as the accuracy of potential machine learning implementation for occupancy detection based on these data collection, is dependent on the data quality. So, this step is the foundation of storing rich data and limit the error coming from physical factors by taking action at the right moment once the notification is received from the system.

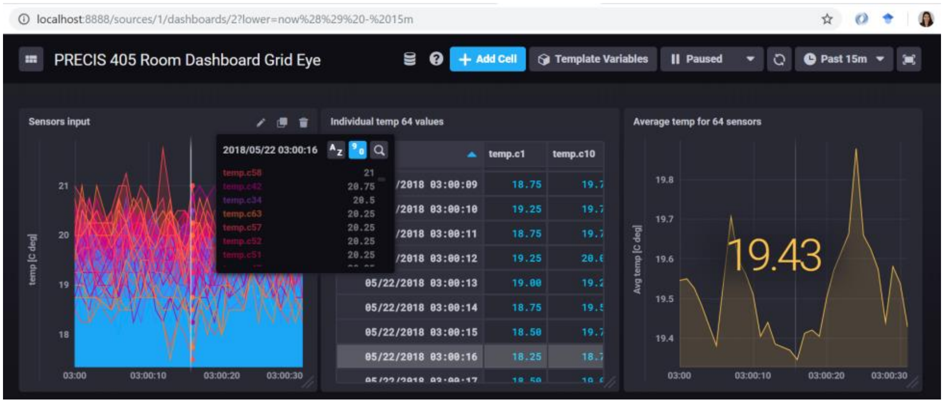

Fig.5 Dashboard built in Chronograf with time series data

This middleware based on open source tools is dedicated to scalable implementation for data exploration and visualization of spatial and temporal data streams across buildings using an online dashboard at intelligent cities scale.

In the past we used MysSQL database, a traditional RDBMS (Relational Database Management System), but with our recording frequency at one second sampling, we understood it will pile up very soon and then it will become difficult to scale with a long-term experiment. InfluxDB offers us the same SQL resemble query language, but it is more appropriate for time series event collection.

5. Conclusions

For data storage, it is profitable to project a scalable infrastructure to enable easy addition of more sensors/devices, preferably in the cloud, with a replication factor because, soon, collaborative buildings will be a reality.

We learned that researchers developing products or services for smart buildings, will face many challenges. Among these, we enumerate: the lack of enough data for understanding the desired comfort, consensus in communication systems or architectures, and especially concerns related to privacy and security in what regards the monitored spaces and users.

This paper presented our use case of data visualization approach for the infrastructure deployed in a room in our university building, for a sensing platform used with applications in occupancy prediction. In addition, we included a comprehensive overview of most used technologies, highlighting key ideas in this area, on all layers of a smart building (sensing, communication, data storage, machine learning algorithms, security, systems). For future work, we will study the main characteristics of data platforms supported by a larger wireless sensor network storage. Thus, non-intrusive middleware to leverage data across buildings with possibility to benchmark data from multiple rooms and more important, multiple buildings, was presented. This aspect is very important because nowadays, even in the same campus, we face the situation of several buildings being managed by using different Building Management Systems (BMS), employing lack of integration.

Acknowledgement

This work was conducted during the Fulbright Student Program in the LEED gold certified smart buildings of University of California Merced using historical data from deployment in Bucharest. We would like to express our gratitude to prof. Alberto Cerpa and PhD. Eng. Daniel Winkler for lessons learned, knowledge and passion sharing.

R E F E R E N C E S

[1] *Tanasiev, Vladimir, and Adrian Badea*, "Energy Efficiency through Intelligent Management of Building Equipment," University Politehnica of Bucharest Scientific Bulletin, Series C: Electrical Engineering, pp. 49-56, 2012.

[2] *University of California Merced*, "University of California Merced - Triple zero commitment," 2018. [Online]. Available: https://www.ucmerced.edu/triple-zero-commitment . [Accessed on 30.08.2018].

[3] *Erickson, Varick L., and Alberto E. Cerpa*, "Occupancy based demand response HVAC control strategy" in Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, New York, 2010.

[4] *Winkler, Daniel A., et al.*, "FORCES: feedback and control for occupants to refine comfort and energy savings" in Proceedings of the 2016 ACM International Joint Conference on

Pervasive and Ubiquitous Computing. *ACM,* 2016.

[5] *Erickson, Varick L., and Alberto E. Cerpa*, "Thermovote: participatory sensing for efficient building hvac conditioning" in Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings. *ACM*, 2012.

[6] *Eugeny Batov*, "The distinctive features of “smart” buildings". Procedia Engineering, 2015.

[7] *Beltran, Alex, Varick L. Erickson, and Alberto E. Cerpa*, "Thermosense: Occupancy thermal based sensing for hvac control" in Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings, 2013.

[8] *Eurostat*, "Statistics Explained - Europe 2020 headline indicators," 2018. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_headline_indicators. [Accessed on 30.08.2018].

[9] *Stamatescu, Grigore, and Valentin Sgârciu*, "Integration of wireless sensor networks with virtual instrumentation in a residential environment" in arXiv preprint arXiv:*1305*, 2013.

[10] *Stamatescu, Grigore, et al.*, "Analytical and experimental sensor node energy modeling in ambient monitoring" in Industrial Electronics and Applications *(ICIEA)*, 2014.

[11] *Shelby, Zach, and Carsten Bormann*, "6LoWPAN: The wireless embedded Internet" 2011.

[12] *Lee, Jin-Shyan, Yu-Wei Su, and Chung-Chou Shen*, "A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi" in Industrial Electronics Society, 2007.

[13] *Neculoiu, Giorgian, et al.* "Model Predictive Control applied for building thermal control." Electrical Machines & Power Electronics (ACEMP), 2015 Intl Conference on Optimization of Electrical & Electronic Equipment (OPTIM) & 2015 Intl Symposium on Advanced Electromechanical Motion Systems (ELECTROMOTION), 2015 Intl Aegean Conference on. IEEE, 2015.

[14] *Neculoiu, Giorgian, et al.* "Increasing energy efficiency of a building with minimal computational demands using model predictive control." RoEduNet International Conference-Networking in Education and Research (RoEduNet NER), 2015 14th. IEEE, 2015.

[15] *Kalluri, Balaji, et al.* "A Cyber-Physical Middleware Platform for Buildings in Smart Cities." Advances in Informatics and Computing in Civil and Construction Engineering. Springer, Cham, 2019. 645-652.

[16] *Chitu, Claudia F., Grigore Stamatescu, and Valentin Sgârciu*, "Scalable architectures for stream analytics and data predictions dedicated to smart spaces" in Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments, 2017.

[17] *Bader, Andreas, Oliver Kopp, and Michael Falkenthal*, "Survey and comparison of open source time series databases" in Datenbanksysteme für Business, Technologie und Web, 2017.

[18] *Chitu, Claudia, Mokhloss I. Khadem, and Valentin Sgarciu*. "Predictive modeling of occupancy patterns in smart buildings." Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2017 9th IEEE International Conference on. Vol. 2. IEEE, 2017.

[19] *Chiţu, Claudia, et al.* "Wireless system for occupancy modelling and prediction in smart buildings." Control and Automation (MED), 2017 25th Mediterranean Conference on. IEEE, 2017.

[20] “Open source time series platform”. [Accessed on 30.08.2018]. Online: https://www.influxdata.com/time-series-platform/

[21] “DB Engines ranking – Trend of time series DB popularity.” [Accessed on 30.08.2018]. Online: https://db-engines.com/en/ranking_trend/time+series+dbms