Building Automation System Data Integration with BIM: Data Structure and Supporting Case Study

Caroline Quinn[1], Ali Zargar Shabestari[2], Marin Litoiu[2] and J.J. McArthur[1,*]
[1] Department of Architectural Science, Ryerson University
[2] Department of Electrical & Computer Engineering, York University
* email: jjmcarthur@ryerson.ca

Abstract

Buildings Automation Systems (BAS) are ubiquitous in contemporary buildings, both collecting room condition, equipment operational data and sending control points developed by integrated sequences, but are limited by prescribed logic, possess only rudimentary visualizations, and lack broader system integration capabilities. Advances in machine learning, edge analytics, data management systems, and Facilitates Management Building Information Model software (FM-BIM) permit a novel approach to cloud-hosted building management. This paper presents an integration technique for mapping the data from a building Internet of Things (IoT) sensor network to an FM-BIM. The sensor data ontology and time series analysis strategies integrated into the data structure are discussed and presented, including the use of a 3D nested list to permit time-series data to be mapped to the FM-BIM and readily visualized. The developed approach is presented through a case study of an office living lab consisting of a local sensor network mimicking a BAS, which streams to a cloud server via a virtual private network connection. The resultant data structure and key visualizations are presented to demonstrate the value of this approach, which permits the end-user to select the desired timeframe for visualization and readily step through the spatio-temporal building performance data.

Keywords: FM-BIM, System Integration, Sensor data streaming, Internet of Things (IoT), Smart and Continuous Commissioning

1. Introduction

Parametric design tools, Building Information Models (BIMs), digital fabrication, and virtual construction scheduling generate a wealth of digital data on the built environment through the design and construction phases. As a building is put into operation, the magnitude of this information increases exponentially and FM-enabled BIMs (FM-BIMs) are of specific value in this digital context, having demonstrated time- and cost-savings benefits for soft (Arayici, Onyenobi, & Egbu, 2012) and hard facilities management activities (Volk, Stengel, & Scultmann, 2014; Love, Matthews, Simpson, Hill, & Olatunji, 2014; Arayici, Onyenobi, & Egbu, 2012; Bryde, Broquetas, & Volm, 2013). When Computer Aided Facility Management (CAFM) data is integrated into a BIM, this data provides significant benefits such as utility cost reductions, comfort management, space optimization, improved inventory management, and energy management (Love, Matthews, Simpson, Hill, & Olatunji, 2014). Over time, however, the volume of this information becomes extremely large and cumbersome, resulting in barriers to adoption associated with the level of effort necessary to maintain accurate BIMs and manage the information (Volk, Stengel, & Scultmann, 2014). BIM's are one of many "standalone information systems" used by building stakeholders. By creating integrations between systems (ex: IoT sensor network and FM-BIM), and working toward aggregating data in a single model, negative effects of standalone information systems can be minimized. A central FM-BIM is the system that should be built out to integrate to all building data sources, to be used through the building project and operational lifecycle.

To address the challenge of managing this information exchange, several researchers have developed strategies to integrate building information into FM-BIMs, including several who have used Dynamo or other visual programming languages to this end (Preidel, Daum, & Borrmann, 2017; Khaja, Seo, & McArthur, 2016; Gerrish, 2017). Others have developed ontologies to support data streaming to BIM using the Semantic Sensor Network (SSN) ontology (Kučera & Pitner, 2016). Once sensor data is available in the BIM it can be further integrated to other systems. COBie has been recommended for data integration from a BIM to a Computerized Maintenance Management System (Pishdad-Bozorgia, Gao, Eastman, & Selfa, 2018). There remains a paucity of literature regarding data streaming from sensor networks, as well as the type of summary information relevant to the FM-BIM (Kassem, Kelly, Dawood, Serginson, & Lockley, 2015). This research paper aims to fill this gap by presenting a complete IoT data acquisition, management, and BIM mapping ontology.

This paper presents a database architecture to facilitate the integration of IoT sensor data to an FM-BIM, effectively taking the first steps toward FM-BIM as a central data source for all building data. This paper further implements the proposed integration on a sensor data stream from a test-bed office, suggesting appropriate batch analytics and queries for the sensor data stream. Each of the sensor data streams are mapped to fields within the FM-BIM and using Dynamo sensor data values for desired time frames can be navigated. The research offers a database architecture which could act as the data source for cloud based SCCx through machine learning, as well as an accessible portal to BAS sensor related data typically hidden within a proprietary system. This would facilitate the creation of a full digital twin of a building in the form of an FM-BIM. Facility managers could use this model to test and plan maintenance projects, observe building conditions over time given outdoor conditions as well as system settings, retrieve specific sensor data using a visual interface, among other things. The next phase of this research will be scaling the architecture to a Smart Building level, as which point SCCx algorithms can begin to be explored and full building visualizations within the FM-BIM will be possible. Alternatively, future phases could include adding additional data integrations to FM-BIM.

2. Methodology

An IoT sensor data integration to FM-BIM requires three data processing components: (1) data acquisition and ingestion, where by the raw sensor data is collected and transferred to a cloud database; (2) batch analytics, a set of business rules to be executed on sensor data to summarize data for faster querying; and (3) integration, the querying process that summarizes data and maps to the appropriate FM-BIM fields. Figure 1 shows how these processes act on the various system elements (sensor points, database of historical records, and the FM-BIM).

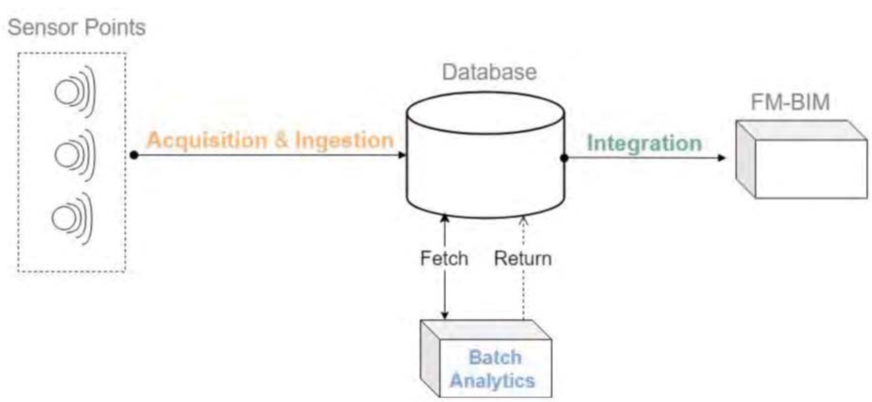


Figure 1: System architecture for IoT data streaming to FM-BIM*)*

2.1 Data Acquisition and Ingestion

In order to provide an efficient and flexible architecture minimal processing will occur during acquisition and ingestion. Data point values are organized in a hierarchical schema where the building is broken into successively more granular components . The hierarchy for a building would be: “building”, which contains “systems”, which have “elements”, which have data collecting “points”. For controls system mapping, relevant elements are assigned a Building Automation System Identifier (BASID), which may be either the equipment name or room number, depending on the element type. Each sensor point type is then assigned a PointID. The specific point assignments are based on a systematic review of the BAS schematics and controls sequences for the building. This hierarchical schema will result in each sensor being tagged with a human interpretable ID, allowing FM users to interpret the relevance of each sensor point to the building. Controllers periodically push data to a buffer which synchronizes the time series data of the points, this is a more efficient approach than synchronizing each data point stream directly in the buffer as the latter requires multiple network connections. The controller will allow data to be streamed more quickly from sensors to the FM-BIM, giving FM users a near live view of building operations (Ramprasad, et al., 2018). This process can be seen in Figure 2.

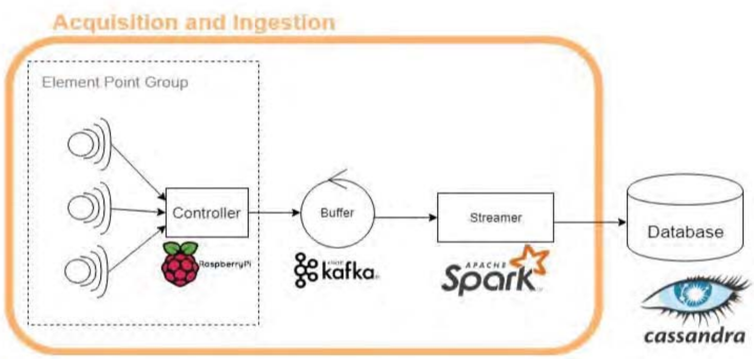


Figure 2: Data acquisition and ingestion process

Because a hierarchical data schema is used for integration of summarized point data to the FM-BIM, a variety of timestep granularities can be supported for visualization within the FM-BIM using aggregating run time queries. By decoupling individual points from the element group, batch analytics can run in parallel and therefore more efficiently.

2.2 Batch Analytics

Once raw data has been acquired and ingested to the cloud-based database centralized time series analytics can run in batch jobs to summarize data. Summarizing the data to single human readable metrics will mean less time spent by FM users interpreting building data. The cloud base environment is the appropriate place to run these analytics as it is a stable environment with high capacity and enough computational power to deal with high complexity summary calculations. Appropriate time series analysis methods depend greatly on the desired FM-BIM visualizations. Counts can be used as an effective means to visualize building deficiencies and therefore deficiency percent changes over time. Averages can be used to show conditions and system performance in a building, they can also be presented with their point value range for the selected timeframe. Figure 3 shows the general process using Spark for batch analytics.

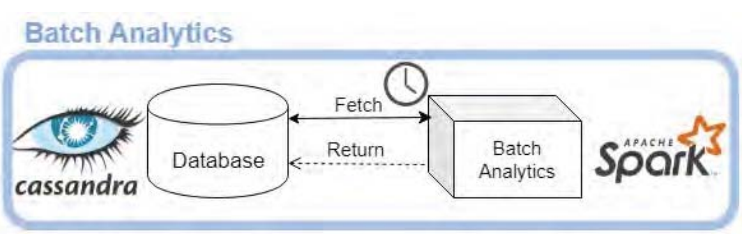


Figure 3: Batch analytics process

2.3 Integration

The data transfer from the database to the FM-BIM consists of three key steps:

1. For each visualization, a .csv file (format in Table 1) must be maintained from the summary tables that can be passed to the FM-BIM using a visual programming Language (VPL)(Dynamo) script. This file would be updated regularly as batch analytics are run with new incoming sensor data and saved with a fixed filename. This .csv file is a 2D list, columns headed with BASID and rows with timestamp, cells contain ordered comma delimited point values for the intersecting ID and timestamp.
2. The Dynamo script imports the .csv file, sorts it to match the FM-BIM element ordering and creates a 3D list by converting the ordered point values from a string to a list. Each index of the list contains a Revit element ID ordered list of point values.
3. The user selects the desired time for display in the FM-BIM using a slider, which inputs to a function that filters the transposed 3D list using a python script and outputs a 2D further transposed list of sensor data for each BASID at the specified timestep. This list is then mapped to the FM-BIM using the Dynamo Element.SetParameterByName node, updating the parameter in Revit of the element.

Table 1: Sample CSV format for BAS mapping to BIM (truncated)

	BASID1	*BASID2*
2019-04-24 17:00:00	*0.23, 22.3, 5*	*0.20, 24.3, 1*
2019-04-24 16:00:00	*0.22, 22.3, 0*	*0.23, 24.2, 1*

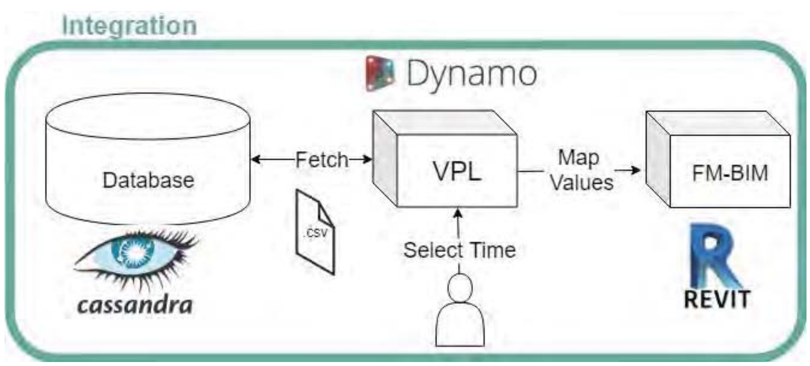


Figure 4: Database to FM-BIM integration process

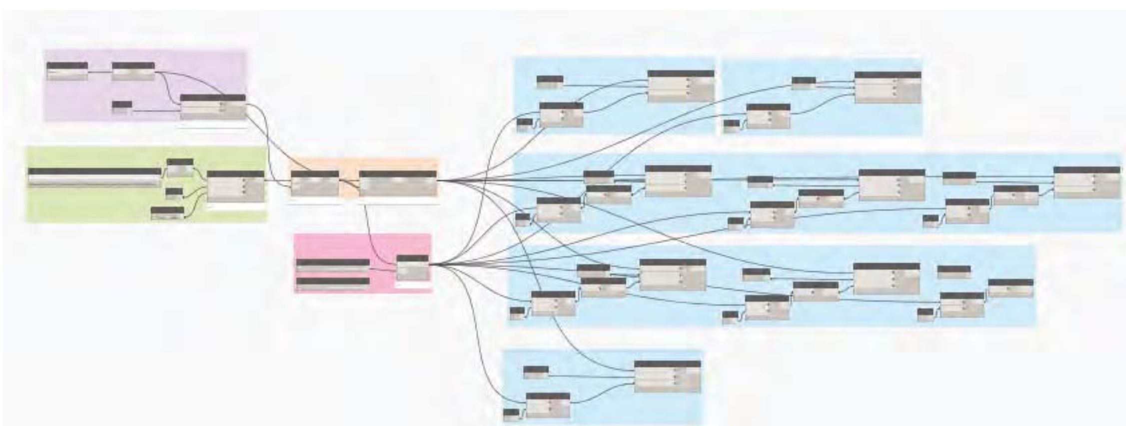
These .csv files can be used to map times series data such as averages or counts, a sample .csv can be seen in Table 1. A time slider to navigate over the time frame can be used to select data for mapping, for example a slider with values from 0-72 hours can be used to access average point (sensor) values in the past 72 hours. These are shown in the BIM, giving FM users a visual and interactive platform to interpret sensor data. Furthermore, daily averages or counts can be mapped for monthly trend visualization using a less granular slider that ranges from 0-30 days, or monthly averages or counts for annual or seasonal visualization with. Figure 5 shows the Dynamo implementation for hourly navigation. The code for creating the 3D list of inputs for mapping is as-follows:

```
    import clr
clr.AddReference('ProtoGeometry')
from Autodesk.DesignScript.Geometry import *
importList = IN[0] # the main list imported from excel
elementList = IN[1] # the list of elements in revit
newList = []
for i in range(1, (importList[0])):
    for j in range(0, len(elementList)):
```

```
            if elementList[j] == importList[0][i]:
            element_to_add = []
            element_to_add.append(importList[0][i])
            for k in range(1, len(importList)):
                tuple_to_add = []
                tuple_to_add.append(importList[k][0])
                temp = importList[k][i].split (",")
                for z in range(0, len(temp)):
                    tuple_to_add.append(temp[z])
                element_to_add.append(tuple_to_add)
            newList.insert(i-1, element_to_add)
            break;
        else:
            continue
OUT = newList
```

```
    import clr
clr.AddReference('ProtoGeometry')
from Autodesk.DesignScript.Geometry import *
elementList = IN[0]
ImportList = IN[1]
newElementList = []
for i in range(0, len(ImportList)):
    for j in range(0, len(elementList)):
        if (ImportList[i][0] == elementList[j].GetParameterValueByName("Number")):
            newElementList.insert(i, elementList[j])
    else:
        continue
#Assign your output to the OUT variable.
OUT = newElementList
```

Top:



Bottom:

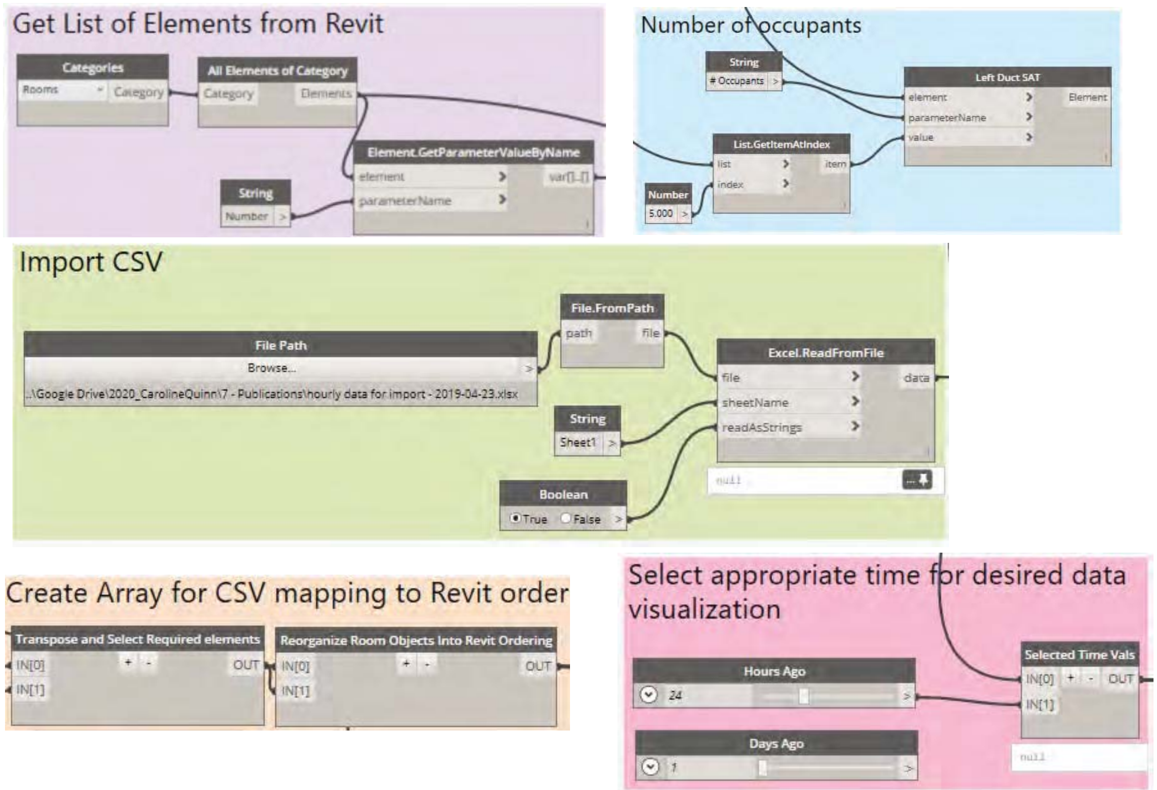


Figure 5: Dynamo script for time-specific data mapping: (top) high-level summary showing node relationships; (bottom) individual code blocks)

The code for selecting the time is as-follows:

```
    import clr
clr.AddReference('ProtoGeometry')
from Autodesk.DesignScript.Geometry import *
allData = IN[0]
selectedTime = IN[1]
valsToMap = [[] for _ in range (len(allData[0][1])-1)]
for j in range (0, len(allData)):
    for k in range (1, len(allData[0][int(selectedTime)])):
            valsToMap[k-1].insert(j , allData[j][int(selectedTime)][k])
OUT = valsToMap
```

Note that the scripts presently do not support null point values as this causes data type errors within the Dynamo and therefore a large dummy value must be inserted for any missing data points. Future work will implement the necessary logic to permit null value mapping so that sparse matrices of point value data can be handled, as these are computationally the most efficient for data storage.

3. Case Study from an Office Living Lab

A living lab test cell has been established within a single faculty office, which incorporates a local sensor network. Sensors measuring occupancy, lighting state, door position, and HVAC are integrated with an Arduino Mega 2560 and streamed via an Ethernet connection to a private cloud. Other systems in this office but not discussed in this paper measure ambient temperatures in the office and adjacent spaces (direct cloud streaming) and thermocouples for surface temperature measurement. Figure 6 shows sample instantaneous data accessible through a web interface.

Ryerson (ryerson)

Timestamp	values
2019-04-25 13:28:12	occupancy: 1 lights: 1 doorPosition: 1 airVelocity_left: 10.87 airVelocity_right: 1061.07 humidity_left: 50.20 temperature_left: 21.50 humidity_right: 47.10 temperature_right: 24.20

	A	B	C
1	timestamp	ARC-0309	ARC-0314
2	'14-Mar-2019 19:44:00'	0.2,0.2,0.1,12.1,36.2,20.8,34.6,24	0,0,0,10.4,33.1,20.4,33.8,22.6
3	'15-Mar-2019 19:44:00'	0,0,0,18.7,41.9,18.9,35.3,23	0,0,0,10.6,56.5,18.9,43.7,23.5
4	'16-Mar-2019 19:44:00'	0,0,0,12.4,26.3,20.9,27.7,23.3	0,0,0,10.9,36.5,20.2,34.6,23.5
5	'17-Mar-2019 19:44:00'	0,0,0,11,21,21.2,23.8,23.6	0,0,0,10.9,45.4,19.7,38.4,23.5
6	'18-Mar-2019 19:44:00'	0.1,0.1,0.1,10.9,26.6,20.6,27.4,23.6	0,0,0,10.9,48.9,19.1,40,23.7
7	'19-Mar-2019 19:44:00'	0.3,0.3,0.2,13.4,30.4,20,29.7,23.4	0,0,0,11,21,21.2,23.8,23.6

Figure 6: Sample screenshot from live streaming of office data to the cloud (top) and pre-processed daily average CSV rows (bottom); the highlighted data is mapped in Figure 7

This sensor data has been mapped to an FM-BIM for the host building, created in Revit 2019. This model uses the room number as the BASID as in this simplified case, all sensors are room-mounted. ARC-309 is the sensor test room showing actual data; dummy data has been populated into another room on the same floor (ARC-314) for visualization purposes. The Dynamo script presented in Section 2 has been integrated with this FM-BIM and visualizations of the daily average sensor data for two sample days are shown in Figure 7.

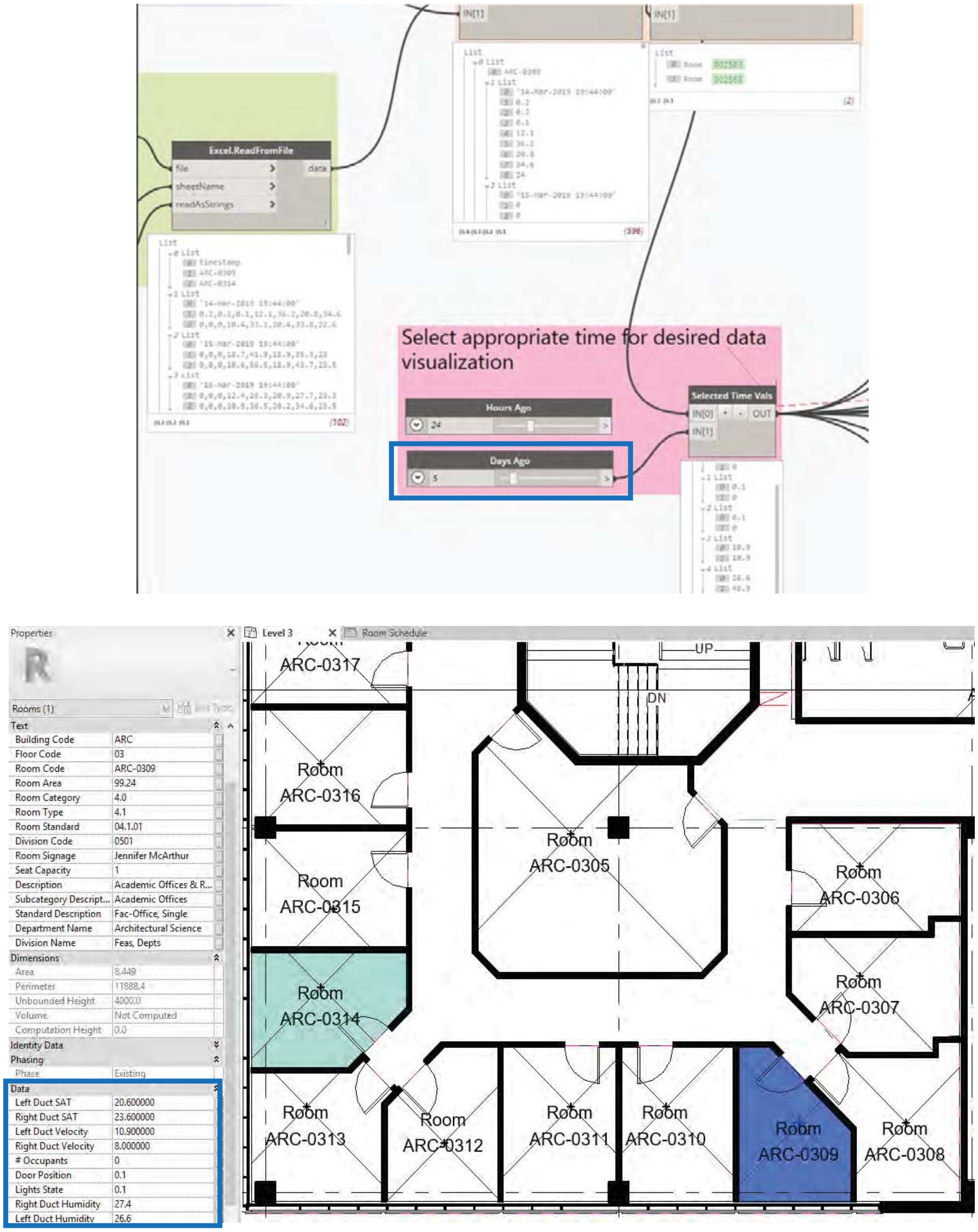


Figure 7: Top: Dynamo slider set to "5 days ago". Bottom FM-BIM showing daily average BAS sensor data mapped from March 17th, 2019 19:44 – March 18th 2019 19:44

4. Discussion and Conclusion

BAS data is highly valuable within an FM-BIM, however due to the high measurement frequency, the supporting database architecture and data visualization must be carefully considered in order to facilitate BAS to FM-BIM integration. This paper presents a database architecture that allows for high-frequency low-latency data transmission by using sensor point controllers, dedicated database buffers and streamers, a cloud-based database, and batch analytics. This architecture permits efficient summarization of heterogeneous BAS data on the cloud where scalable computational resources are available. The data summarization techniques proposed include averages and counts on an hourly, daily, and monthly basis. Consistent with building energy modeling norms, data regarding occupancy or lighting state are indicated in real time as integers, however these must be stored as floats for hourly, daily, and monthly averages to better reflect the percentage time that a space is occupied or illuminated. The summarization of sensor data using this method allows the FM-BIM to be the single building data model for FM users to consult during the operational life of a building, while maintaining a data structure that can also be used by building applications for energy management and other building controls.

The use of Dynamo, a VPL, was found to be highly effective in mapping time series data integration from a .csv file to an FM-BIM, provided proper structure. The presented format contains sensor data for each point as a string to form a tuple, which is then inserted into a 2D list with time in rows and BASIDs in columns to create a 3D list with indices of time, BASID, and point values. From this list, sub-arrays mapping BASID vs POINTID can be accessed for a given time using the slider and code presented. This study has been limited to room-hosted data points and a more complex mapping algorithm will be required to update both room-hosted and element-hosted time-series data, as the existing script can only map to a single element type at once. An additional limitation of this work requiring further development is that null values cannot be processed by the script; as a temporary solution, a dummy value (555555) has been substituted for missing data, but code refinement is necessary in future work to overcome this limitation. This will reduce the computational cost by permitting the use of spare matrices for data storage.

This work has yet to be scaled to the full-building or tested by facility-users and this scaling and testing will form the long-term future work in this field. In addition, incremental functionality will be developed to enhance the analytics to integrate deficiency alarms using the trended data as defined in accordance with ASHRAE Guideline 36-2018 High Performance Sequences of Operation for HVAC Systems (ASHRAE, 2018). A more refined graphical user interface should be developed, were users can interact with sliders outside of the VPL to initiate IoT sensor data integration with an FM-BIM. Further work on defining relevant BAS data summarization techniques should be done in collaboration with end-users such as Facility Engineers. Topics for consideration include additional measures to be mapped, for example data maxima and minima, different timeframes, and dashboard integration.

Cyber physical systems are concerned with the feedback loop of sensing, evaluating, and acting on building conditions (Schmidta & Åhlund, 2018). Schmidta describes building automation as a three-layer architecture facilitating this feedback loop. This research focuses on making the bottom layer of this architecture- the field level comprised of sensors and actuators- available to the top layer – the management level comprised of building management system and visualizations tools-while maintaining the ability to engage the middle layer where applications such as predictive control are applied. Most current research in cyber physical systems focuses on optimizing operational energy cost or consumption (Schmidta & Åhlund, 2018) through the development of the middle layer. However, sensor data visualization should be considered and further researched in cyber physical systems. The middle layer is not easily interpreted by FM users, and without a thoroughly developed top management layer there will be a loss of agency for FM users as cyber physical system development progresses. This is increasingly important for cyber physical systems with tightly coupled sensing and actuating embedded systems as described by Kleissl & Agarwal (2010) where there is no discussion of visualization at the management level.

This research lays the foundation for a long-term project to develop a cloud-hosted BIM-integrated FM platform, permitting data analytics with complex predictive modeling and classification algorithms to support applications such as Smart and Continuous Commissioning and Model Predictive Control. The visualization of the summarized sensor data provides an integrated view for facility managers and building operators to support integrated asset management and optimization.

Acknowledgements

This research was financially supported by the BRAIN Alliance - Big Data Research, Analytics, and Information Network, funded by the Ontario Research Fund Centres of Excellence program and FuseForward.

References

Arayici, Y., Onyenobi, T., & Egbu, C. (2012). Building Information Modelling (BIM) for Facilities Management (FM): The Mediacity Case Study Approach. *International Journal of 3-D Information Modeling, 1*, 55-73.

ASHRAE. (2018). ASHRAE Guideline 36-2018: High Performance Sequences of Operation for HVAC Systems. Atlanta: American Society for Heating, Refrigeration and Air-Conditioning Engineers.

Bryde, D., Broquetas, M., & Volm, J. M. (2013). The project benefits of building information modelling (BIM). *International Journal of Project Managmeent, 31*(7), 971-980.

Gerrish, T. (2017). *Exploring the effectiveness of BIM for energy performance management of non-domestic buildings (Doctoral dissertation).* Loughborough: Loughborough University.

Kassem, M., Kelly, G., Dawood, N., Serginson, M., & Lockley, S. (2015). BIM in facilities management applications: a case study of a large university complex. . *Built Environment Project and Asset Managemen, 5*(3), 261-277.

Khaja, M., Seo, J. D., & McArthur, J. J. (2016). Optimizing BIM Metadata Manipulation Using Parametric Tools. *Procedia Engineering*, *145*, pp. 259-266. Tempe.

Kleissl, J., & Agarwal, Y. (2010). Cyber-Physical Energy Systems:. *Design Automation Conference* (pp. 749 - 754). Anaheim: IEEE.

Kučera, A., & Pitner, T. (2016). Semantic BMS: Ontology for Analysis of Building Automation Systems Data. In L. M. Camarinha-Matos, A. J. Falcão, N. Vafaei, & S. Najdi (Ed.), *7th Doctoral Conference on Computing, Electrical and Industrial Systems (DOCEIS)* (pp. 46-53). Costa de Camperica, Portugal: IFIP Advances in Information and Communication Technology. doi:10.1007/978-3-319-31165-4_5

Love, P., Matthews, J., Simpson, I., Hill, A., & Olatunji, O. (2014). A benefits realization management building information modeling framework for asset owners. *Automation in Construction, 37*, 1-10.

Pishdad-Bozorgia, P., Gao, X., Eastman, C., & Selfa, A. P. (2018). Planning and developing facility management-enabled building information model (FM-enabled BIM). *Automation and Construction, 87*, 22-38.

Preidel, C., Daum, S., & Borrmann, A. (2017). Data retrieval from building information models based on visual programming. *Visualization in Engineering, 5*(1), 1-14.

Ramprasad, B., McArthur, J., Fokaefs, M., Barna, C., Damm, M., & Litoiu, M. (2018). Leveraging existing sensor networks as IoT devices for smart buildings. (pp. 452-457). Singapore: IEEE 4th World Forum on Internet of Things (WF-IoT).

Schmidta, M., & Åhlund, C. (2018). Smart buildings as Cyber-Physical Systems: Data-driven predictive control. *Renewable and Sustainable Energy Reviews*, 742-756.

Volk, R., Stengel, J., & Scultmann, F. (2014). Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction, 38*, 109-127. doi:http://dx.doi.org/10.1016/j.autcon.2013.10.023