A Data-Warehouse Architecture supporting Energy Management of Buildings

H.U. Gökçe, Y. Wang, K.U. Gökçe, K. Menzel
IRUSE, University College Cork, Ireland

ABSTRACT: Environmental legislative and economical drivers require more efficient and accurate energy management of buildings. Current building management systems do not have the capabilities of energy specific monitoring and management. Therefore, this paper addresses this deficit by introducing a holistic N-dimensional information management system. The proposed system consists of: a data warehouse core to store, integrate and analyse the complex data sets, an extraction, transformation, loading tool and context sensitive information representation components for decision support and monitoring. The proposed system is implemented in a test building at University College Cork, Ireland.

1 INTRODUCTION

Current residential and office buildings provide a significant contribution to total energy consumption and CO_2 emissions. Reports by the Intergovernmental Panel on Climate Change (IPCC) and the U.S. Department of Energy note that buildings account for 25-30% of total energy related CO_2 emissions (Price et al. 2006). The overall building stock in the U.S. accounts for 40% of the total energy consumption of the country (Flippn, 2000). It has been estimated that the operation of buildings (space lighting, heating and cooling) is responsible for about 50% of primary energy use and a slightly lower share of CO_2 and green house gas emissions in the EU (Cohen et al. 2004). In the case of Ireland energy consumption of buildings accounts for approximately 30% of green house gas emissions and approximately 40% of total energy consumed (Howley et al. 2006).

The dramatically increasing effect of building stock on global warming and energy consumption creates an international stimulus for legislative drivers. The European Energy Performance of Building Directive came into effect from January 2003 and requires owners to certify the energy performance of all buildings in excess of 50 m^2 (Energy Performance Building Directive 2002/91/EC-2006/32/EC). Several case studies suggest (Herzog et al. 1992) that energy savings of between 15% and 40% could be made in commercial buildings with closer monitoring and supervision of energy related data (Salsbury & Diamond 2000). According to the European commission, research has shown that more than 20% of present energy consumption could be saved by 2010 based on 2000 levels by applying stricter standards to buildings undergoing refurbishment and to new buildings (EC 2005). A further directive on energy end-use efficiency and energy services (2006/32/EC), also referred to as the Energy Services Directive (ESD) (EU 2006) came into effect in May 2006. The directive requires the establishment of an indicative target of 9% improvement in energy efficiency by 2016 for all member states.
As a result of these research outputs and legislative drivers launched by the national and international organizations, the role of monitoring and analyzing building performance is gaining significant importance for reducing the energy consumption and operational costs.

This paper discusses a holistic information management approach based on a data warehouse technology in order to support multi-dimensional analysis of building performance data.

Performance based assessments of building operation provides huge potentials for energy saving more structured and detailed monitoring and analysis capabilities to the end-users. Current Building Management Systems (BMS) primarily focus on building control but lack analysis and monitoring capabilities. Also, modern building management systems incorporate data archival but energy managers need assistance in extracting useful information from large volume of data compiled (Piette et al. 2001). In order to analyze and monitor building performance data end users have to define and calculate relevant metrics based on limited information provided by the building management system. Considering the required competencies it is most likely that efficient energy analysis is beyond the capabilities of the end-users. Furthermore, current building management systems can not address relational conclusions associated with detailed location, time, organization and building systems information which are crucial for more accurate and faster analysis of building performance data. In order to overcome this obstacle a data warehouse is integrated to the core of the system. Moreover, during the life cycle, buildings evolve with regards to different end-user requirements. These kinds of adjustments such as redesigning office spaces for different tenants, directly affect buildings' energy performance. Therefore, an integrated ETL tool has also been developed which updates information automatically from multiple sources such as CAD tools, energy simulation tools etc. Finally, current building management systems do not provide the required Graphical User Interface (GUI) for performance monitoring. The proposed system incorporates context sensitive Graphical User Interfaces specifically developed for building performance monitoring and analysis.

The Environmental Research Institute (ERI) building is used as a demonstrator in order to validate the system. The ERI building is a 4500 m^2 office and laboratory facility located on the campus of University College Cork, Ireland. The Building is equipped with geothermal heat pumps, under floor heating systems and different types of solar panels. It has a Building Management System with 180 wired sensors and meters. Additionally, a wireless sensor network along with wireless sensors and meters has been installed since April 2008. Sensors include temperature, humidity, CO_2 and lighting sensors. Meters include devices to measure electricity, mains water, cold water, gas, lighting, energy consumption, boiler heat, solar heat and under floor heat.

2 HOLISTIC N-DIMENSIONAL INFORMATION MANAGEMENT SYSTEM

Recent advancements in building technologies and building control strategies coupled with the introduction of new building codes have contributed to the improvement of poor energy performance in commercial and residential buildings.

The objective of the proposed system is to provide a holistic information management system to store, integrate and analyse complex data sets from multiple information sources such as CAD tools, energy simulation tools and performance framework specification tools. These data sets will be integrated with data streams collected from wired and wireless sensors and meters in order to perform N-dimensional analysis of building performance data and to support decision making process of the end users.

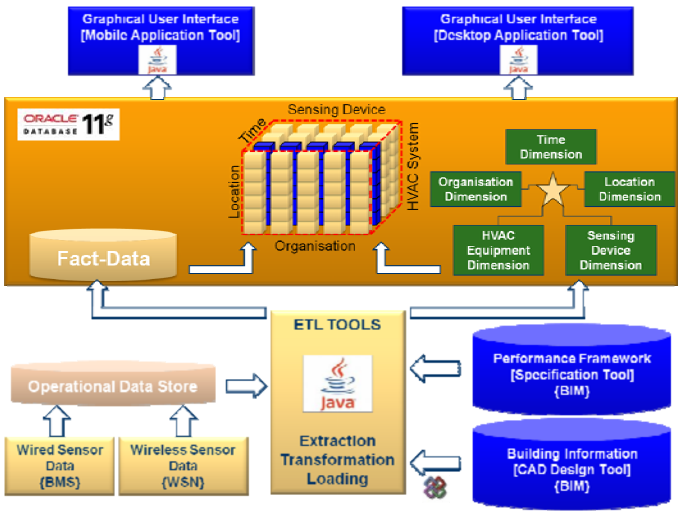

Figure 1: Holistic N-Dimensional Information Management System Architecture

It is obvious that a holistic approach to building performance monitoring requires simultaneous access to the data extracted from different information sources. Energy and efficiency reports and trend analysis should be accessible to energy managers but this is often not the case (Piette et al. 2005). Systematic procedures to address inefficient building operation are beginning to emerge (Mills et al. 2004).The proposed system enables continuous monitoring to tune building systems for optimal comfort and peak efficiency based on current operational requirements. These methods have saved an average of over 20% of the total energy cost and over 30% of the heating and cooling cost in more than eighty American buildings (Claridge et al. 1994). It is also possible to achieve saving of $18 billion or more annually if a systematic commissioning was applied to the entire U.S. commercial building stock (Mills et al. 2004). As a result the holistic N-dimensional information management system supported with data warehouse technology which provides required tools and methods to perform building performance monitoring, enables more efficient performance analysis and dramatic energy savings.

The proposed system extracts sensor data from building management system and from wireless devices. Collected sensor data is stored in the operational data store for data cleansing and redundancy check processes. This pre-processed data is loaded to the fact data section of the data warehouse system via an ETL tool. Simultaneously, data gathered from the building information model is loaded to the dimensional data section of the data warehouse. Loaded fact data and dimensional data is aggregated with regards to different stakeholder requirements in the data warehouse system and presented through specific Graphical User Interfaces.

A simplified architecture is as shown in Figure 1. The system consists of three integrated main components which will be explained in the following sections. These components are:

- Data warehouse core
- Extraction, Transformation, Loading Tool (ETL)
- Information Representation Tools

2.1 *Data Warehouse Technology*

The topic of data warehousing encompasses architectures, algorithms, and tools for bringing together selected data from multiple databases or other information sources into a single repository called a data warehouse, suitable for direct querying and analysis (Widom, 1995).

A data warehouse is a subject oriented, integrated, time varying, non-volatile collection of data that is used primarily in organisational decision making (Inmon, 1992).

Data warehousing systems allow a number of alternative ways to integrate and query information stored in it. Thus, a data warehouse coupled with On-Line Analysis Processing (OLAP) enables end-users to creatively approach, analyze and understand the building performance under different circumstances. The data warehouse system is used to provide solutions for building performance monitoring, since it transforms operational data into strategic decision-making information. The data warehouse stores summarized information instead of operational data. This summarized information is time-variant and provides effective answers to queries such as ‘‘Energy consumption of a particular room in a particular building when the outside temperature is 21°C.”

The aim of the data warehouse component of the system is to:

1. Collect dynamic data from different sources such as wired/wireless sensors and meters.
2. Map the dynamic data with data extracted from CAD tools, energy simulation tools and performance specification tools.
3. Perform N-dimensional data aggregation to support decision making process.

Data Warehouse component consists of three sub-components. These are:

- Operational Data Store
- Fact Data and Dimensional Data
- Aggregated Data

2.1.1 *Operational Data Store*

The operational data store (ODS) is a database designed to integrate current valued subject oriented, volatile and real time data from multiple sources such as building management system and wireless sensor network. An ODS is usually designed to contain low level or atomic (indivisible) data (such as transactions and prices) with limited history that is captured "real time" or "near real time" as opposed to the much greater volumes of data stored in the data warehouse generally on a less frequent basis. According to Bill Inmon (Inmon, 1999), the originator of the concept, an ODS is "a subject-oriented, integrated, volatile, current-valued, detailed-only collection of data in support of an organization's need for up-to-the-second, operational, integrated, collective information." In our case this temporarily stores the data collected by the sensor network. Also, ODS provides input to the data warehouse by storing regularly changing electricity and gas prices that are necessary for energy cost calculation. ODS deals with multiple sources therefore this process includes data cleansing, redundancy corrections and integrity checks. Figure 2 below depicts operational data store and its data sources.

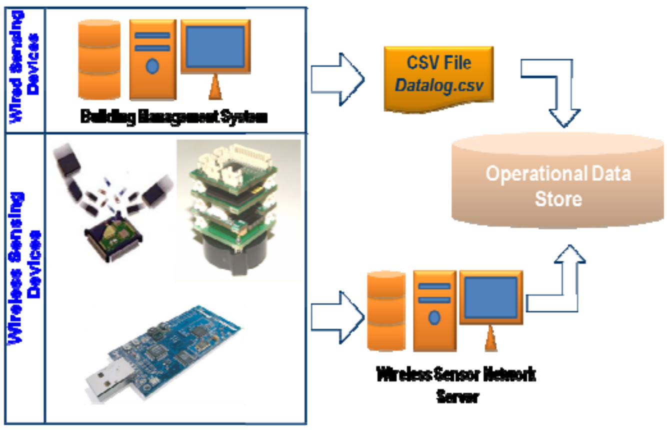

Figure 2: Operational Data Store

2.1.2 *Fact Data and Dimensional Data*

Fact data is the main repository for long term storage of dynamic data. A fact table is the primary table in a dimensional model where the numerical performance measurements of the business are stored (Kimball, 2002). For the proposed system, fact data subcomponent stores and updates the data collected by wired/wireless sensors and meters. Also, it stores and updates the energy unit prices which are used for energy cost calculations. Data collected and temporally stored in the operational data store, are received at certain intervals. Fact data can be further evaluated through different views and needs by integrating dimensional data from multiple sources such as time, sensing device, HVAC systems, location and organisation.

Dimension is a data element that categorizes each item in a data set into non overlapping regions. According to Kimball (2002), “dimension tables are the integral companions to a fact table and contain the textual description of the business.” It provides the means to "slice and dice" data. Dimensions provide structured labelling information to otherwise unordered numeric measures. For example, "Location", "Time", and "Organization" are all dimensions that could be applied meaningfully to the fact data collected from the measurement devices. Each dimension in a data warehouse may have one or more hierarchies applied to it. For example for the “HVAC Equipment” dimension the possible hierarchy is structured as HVAC System>HVAC Equipment Category>HVAC Equipment.

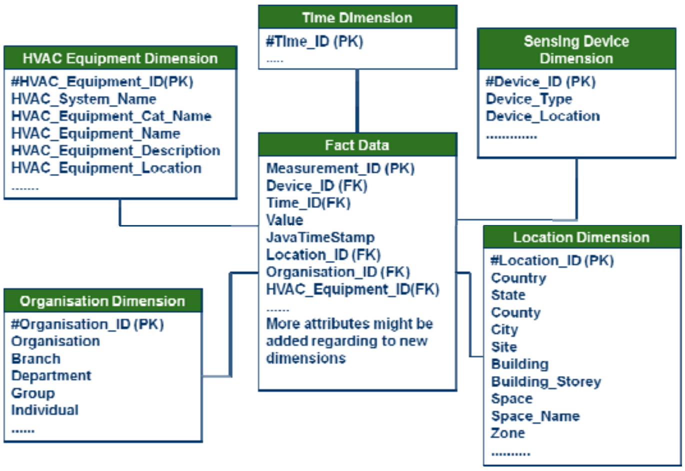

Figure 3: Data Warehouse Star Schema

Figure 3 above depicts a data warehouse star schema. There are four dimensions illustrated. These dimensions are:

- *Time_Dimension:* This consists of time and date information. It enables end-users to analyze the historical data with regards to time and date basis, e.g., “energy consumption of a particular building in August 2007.”
- *Location_Dimension*: This consists of all location information of a particular building including all the individual zones. It enables end users to analyse information according to different locations, e.g., “humidity level of the offices located at the second floor in a particular building.”
- *HVAC_Equipment_Dimension*: This consists of the HVAC system information of a particular building. It enables end users to analyse information according to specific HVAC equipment installed in the building, e.g., “return water temperature of the vacuum tube solar collectors.”
- *Organisation_Dimension:* This consists of organizations and individuals that are occupied in the building. It enables end users to analyse information according to different organizations and individuals, e.g., “Monthly energy consumption of the Civil and Environmental Engineering Group.”
- *Sensing_Device_Dimension:* This consists of sensing device information such as wired/wireless sensors and meters. It classifies the devices with regards to their types. It enables end-users to analyse information according to different sensing device type, e.g., “Humidity level of the second floor for August 2009”

2.1.3 *Aggregated Data*

Aggregated data is the decision support level of the multi-dimensional data warehouse. Every data warehouse contains pre-calculated and pre-stored aggregated data. Sensed raw data collected from wired and wireless sensor infrastructure populates the fact data section of the data warehouse. Fact data becomes meaningful when it is associated with the dimensional data and provides the end user the means to "slice and dice" data. The results are stored in the data cube. A cube can be interpreted as a data view in a multi-dimensional model. Regarding the business requirements aggregated data emphasises on one or more dimensions: e.g., computing and ranking the monthly energy consumption of each floor. Time is a dimension that is of particular significance to decision support (e.g., trend analysis). Other possible operations include comparing two measures: e.g., current building performance compared with the performance objectives defined during the design phase. The scale of the data cube can be adapted to the different requirements hence Performance Metrics provided by the building information model (BIM) can be calculated in multiple dimensions and granularities. Figure 4 depicts a sample of data cube and aggregated data with regards to sensing device and organisation dimensions.

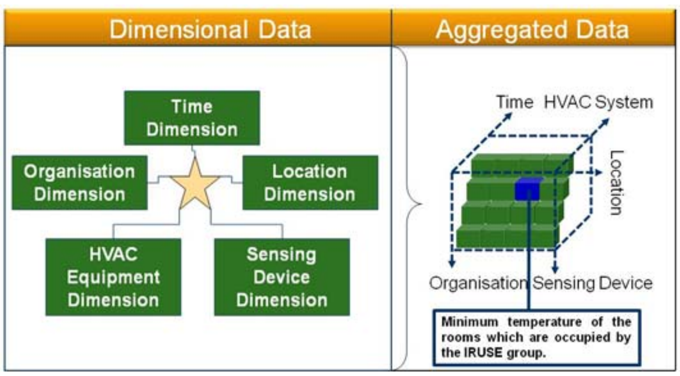

Figure 4: Aggregated Data

2.2 *Extraction, Transformation and Loading Tool*

Data need to be loaded to the data warehouse regularly. To do this, data from one or more operational systems needs to be extracted and loaded into the warehouse. The process of extracting data from source systems and bringing it into the data warehouse is commonly called ETL, which stands for Extraction, Transformation, and Loading (Loney, 2004).

In the data warehouse, raw operational data is transformed into a warehouse deliverable fit for user query and consumption (Kimball, 2002). This is executed by a set of processes called ETL processes which involves:

- Extracting data from multiple sources such as sensor/meter readings and building information model.
- Transforming it to fit data warehouse requirements which might be inconsistent with the outside data sources, e.g., data type inconsistencies.
- Loading it to the data warehouse.

The advantages of efficient and consistent data warehouses make ETL very important as the way data actually gets loaded. For the proposed system the ETL tool is used to populate the fact data table which stores long term dynamic data such as measurement stream Also, ETL Tool can be used to populate Dimensional Tables which stores relatively static data such as architectural data and building HVAC systems data.

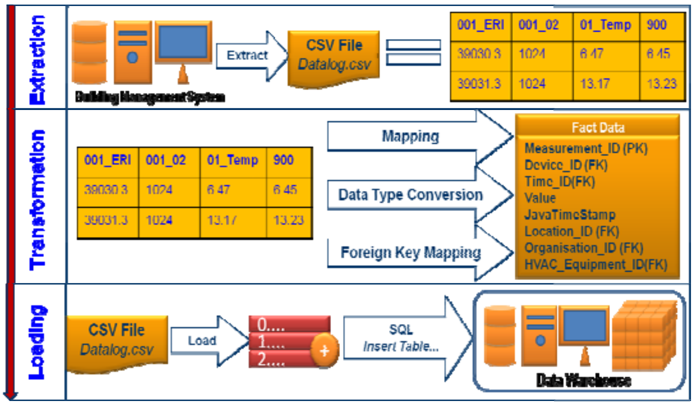

Figure 5: Extraction, Transformation and Loading

Figure 5 depicts an ETL process developed for the "Buildwise" project in order to populate fact data table which:

- Extracts data from the current building management system (BMS) comma separated values (CSV) file archive,
- Eliminates inconsistencies such as duplicate rows,
- Transforms the CSV file structure to the BuildWise data warehouse fact data table structure,
- Loads the CSV files to the BuildWise data warehouse fact data table.

Also, the ETL tool is used to populate the dimensional tables. HVAC equipment dimension, location dimension and sensing device dimension tables can be populated by extracting the data from the CAD tool.

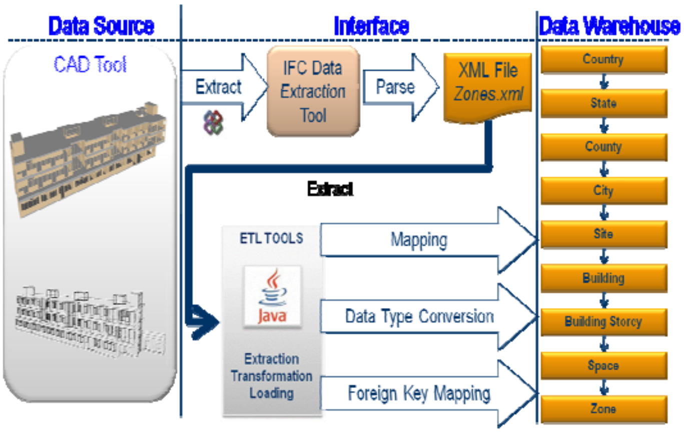

Figure 6: The ETL Process for Dimensional Data

Figure 6 depicts an ETL process to populate the location dimension. This process involves:

- Extracting architectural building information by using IFC (Industry Foundation Classes) data extraction tool.
- Parsing the extracted data to a file in XML format, e.g., Zones.xml.
- Extracting the data from the XML file.
- Transforming the extracted data by proving necessary mappings, data type conversions and foreign key associations.
- Loading the pre-processed data to the data warehouse's location dimension table.

2.3 *Information Representation*

The common goal of the graphical user interfaces is to represent the building performance information to the end users with regards to their roles and functions. The aim of the proposed system's information representation section is designing and implementing a user friendly Graphical User Interface (GUI) with cubes. In order to achieve this, a Java based interface is developed which enables end users easy querying without dealing with complex SQL statements. Also, this GUI is capable of representing query results both in graphical format and/or tabular format regarding to end user preference. There are two main utilization scenarios considered for the system such as performance monitoring and on-site diagnostics. Therefore, two different types of user interface need to be applied:

- Desktop Application for performance evaluation
- Mobile Application for onsite facility management and building diagnostics

Figure 7 depicts a Graphical User Interface developed for the BuildWise project (Wang, 2009). This is designed to track electricity, gas and water consumption of specific zones within a specific facility with regards to selected time intervals. The proposed GUI also provides consumption cost and material unit prices. At the data warehouse level, system queries meter readings for the specific zone selected by the end user. Simultaneously, the system queries spot unit prices which are updated in a regular interval. The graphical user interface represents the data in graphical format and/or in tabular format with regards to end user preferences.

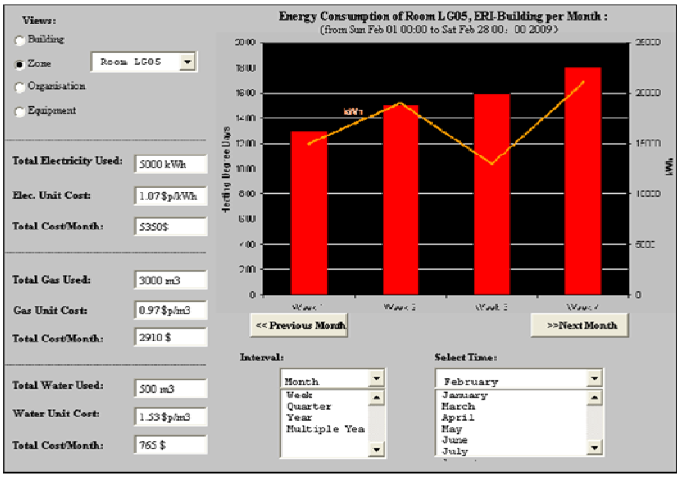

Figure 7: Graphical User Interface for Energy Consumption

3 CONCLUSIONS

Holistic N-dimensional information management system described in this paper provides necessary tools for efficient energy management and monitoring of buildings. Implementing N-dimensional modelling utilizing data warehouse technologies provides effective decision support capabilities to end users. Also, implementing the extraction transformation and loading tool make it possible to integrate different data sources in one repository. Finally, development of context sensitive Graphical User Interfaces represents aggregated data in desired granularity in order to meet different stakeholder requirements.

4 REFERENCES

Claridge, D. et al. 1994. Can You Achieve 150% Predicted Retrofit Savings? Is it Time for Re-Commissioning. American Council for and Energy Efficient Economy.

Cohen, R. & Bordass, W. & Field, J. 2004. Energy performance of non domestic buildings: closing the credibility gap. In Building Performance Congress.

Herzog, P. & LaVine, L. 1992. aridge, D. et al. 1994. Identification and quantification of the impact of improper operation of midsize Minnesota office buildings on energy use: a seven building case study. American Council for and Energy Efficient Economy.

Howley, M. & F.O. & O'Gallachoir, B. 2006. Energy in Ireland 1990 – 2004. Sustainable Energy Ireland (SEI).

Inmon, W. 1999. Building the Operational Data Store. second Edition, John Wiley & Sons.

Kimball, R. 2002. The Data Warehouse Toolkit Second Edition. New York: John Wiley and Sons, Inc.

Loney, K. 2004. Oracle Database 10g The Complete Reference. California: Oracle Press McGraw-Hil/Osborne.

Mills, E. et al. 2004. The Cost-Effectiveness of Commercial-Buildings Commissioning: A Meta-Analysis of Energy and Non-Energy Impacts in Existing Buildings and New Construction in the United States, Lawrence Berkeley National Laboratory Portland Energy Conservation Inc. Energy Systems Laboratory, Texas A&M University.

Piette, M.A. & Kinney, S.K. & Haves, P. 2001. Analysis of an information monitoring and diagnostic system to improve building operations. *Energy and Buildings* 33 (8): 783-791

Piette, M.A. et al., 2005. Development and Evaluation of Fully Automated Demand Response in Large Facilities. California Energy Commission.

Price, L. et al. 2006. Sectoral Trends in Global Energy Use and Greenhouse Gas Emissions. Lawrence Berkeley National Laboratory.

Salsbury, T. & Diamond, R., 2000. Performance validation and energy analysis of HVAC systems using simulation. Energy and Buildings 32(1): 5-17

Wang, Y., 2009. Graphical User Interface for Facility management. MEngSc. Thesis, University College Cork, Ireland. (to be submitted September 2009)