This paper is under submission for publication and is made available online to ensure timely dissemination of scholarly and technical work. For comments and suggestions, contact truong@infosys.tuwien.ac.at.

A Survey on Cloud-based Sustainability Governance Systems

Hong-Linh Truong and Schahram Dustdar
Distributed Systems Group, Viennna University of Technology, Austria
{truong,dustdar}@infosys.tuwien.ac.at
Last revised: 17 Feb 2012

Abstract

Sustainability governance for resource consumption by humans, such as greenhouse gas (GHG) and electricity consumption in facilities, is a complex issue that requires extensible platforms to manage and analyze multiple types of data over the time. In this paper, we will examine how cloud-based information systems and services can support emerging and future requirements for sustainability governance, in particular, of facilities. Our analysis is based on real-world cloud platforms and research initiatives that can be used to support and manage stakeholders, data, processes, and applications for sustainability governance. Based on that, distinguishable and common characteristics of cloud computing platforms for sustainability governance will be identified as well as the differences among cloud-based sustainability governance platforms with general-purposed cloud platforms will be provided. Then we analyze in detail aspects related to data governance, monitoring and analysis techniques, and service integration. Finally, we will present some research challenges and approaches for the development of cloud-based sustainability governance.

1 Introduction

Over the last few years, devices, equipments, cars, houses and buildings have been increasingly instrumented with smart meters and monitoring sensors to provide different types of data for not only monitoring and detecting abnormal status but also supporting sustainability development. However, to support sustainability development in the ecosystem of facilities, we must have adequate governance processes for sustainability. For example, to monitor resource consumption at near real-time in a large commercial/resident building, several types of monitoring data of equipments and spaces in the building, such as electricity consumption, temperature, water consumption, chillers, etc., have to be gathered and combined. Then, we need to store and to share these types of data over the time for long-term data analysis, reporting and auditing of sustainability measurements, such as trend analysis of Greenhouse Gas (GHG) emissions and electricity consumption. In particular, to meet sustainability compliance rules (e.g., for GHG emission and air quality) and to maintain the sustainability of these systems, various complex analysis methods need to be conducted to understand the behaviors of monitored systems. The complexity of data storage, sharing, analysis and application integration poses several challenges for any sustainability governance platform. Especially, sustainability monitoring and analysis of large facilities involve different stakeholders and multi-objective optimization (e.g., to meet law compliance and economical factors). While several information systems have been built for management of energy consumption of facilities in home and enterprise contexts, and their features may be accessed via the Internet, such systems are typically hosted and managed by or dedicated for only the facility owner. They do not support well multi-stakeholder and multi-objective optimization in compliance with diverse regulations.

In our focus on sustainability governance of GHG and energy consumption, we believe that a cloud computing model naturally would be a candidate for overcoming the above-mentioned challenges due to several reasons, such as reducing cost, easing data access and sharing, and enabling complex analysis and compliance assurance. However, to date, most cloud systems are targeted to generic computational resources and storage, and other domains, such as for small and medium enterprises (SMEs), rather than to facility sustainability governance. Only a few industrial systems have been focused on facility management, such as Galaxy (Pacific Control Systems, 2011), generic sensor data sharing and electricity data management, such as Pachube (Pachube, 2011), and carbon footprints analysis, such as AMEE (AMEE, 2011). Although several enabling techniques have been developed in research communities, they are not well integrated into cloud-based solutions for sustainability governance of facilities.

While existing industrial and research cloud systems enable certain sustainability governance features, they still cover only certain aspects in the ecosystem of sustainability governance. Therefore, we examined how cloud computing offerings can support sustainability governance from the perspective of data integration, sharing, and management, data analytics capability, and open cloud platforms. In this paper we analyze three aspects for sustainability governance with a focus on carbon footprints and energy consumption: (i) a detailed analysis model of sustainability governance based on a cloud computing model, (ii) comparison of existing cloud systems enabling sustainability governance, and (iii) open research issues.

The rest of this paper is organized as follows: Section 2 discusses background and related work. Section 3 discusses the model of sustainability monitoring and analysis in the cloud. We present a detailed analysis of cloud production systems for sustainability monitoring and analysis in Section 4. Section 5 discusses research prototypes. Open issues are discussed in Section 6. Section 7 summarized the paper and gives an outlook to our future work.

2 Background and Related work

2.1 Sustainability Governance in the Context of Facility Management

We consider sustainability for human, which is defined as "*the potential for long-term maintenance of well-being*" and which has "*environmental, economic and social dimensions*" (see http://en.wikipedia.org/wiki/Sustainable development). In this paper, we will focus on sustainability in the context of facility management. A facility can be a building, a home, a car, or an equipment, and its sub components/elements. In order to support sustainability development for facilities, we will focus on techniques for capturing, monitoring and analyzing sustainability measurements that characterize human consumption and for examining whether such measurements can meet compliance rules and can support the utilization of resources in a sustainable way.

Current sustainability measurements are diverse (see http://en.wikipedia.org/wiki/Sustainability measurement). However, in our paper, we consider sustainability measurements related to facility resource consumption by human, in particular, GHG and energy consumption. In this paper, sustainability governance (see http://en.wikipedia.org/wiki/Governance for what does it mean governance) applied in facility management is related to models, techniques and processes *to maintain monitoring, analysis, management and compliance assurance of sustainability measurements* to meet both consumer's expectation and regulation requirements. Concretely, to support sustainability governance for facilities, platforms for monitoring, analysis, management and operation of sustainable facilities should consider:

- Service governance: considered as a part of IT governance (see http://en.wikipedia.org/wiki/IT_governance) which has multiple facets (Willson & Pollard, 2009), can cover several aspects, such as service lifecycle management, quality of service, service change management and service contract (Leusse et al., 2009). In our work, we focus *on possible services and their quality that support the monitoring, analysis of sustainability measurements for sustainability standards or laws*.
- Data governance: Data governance (see http://en.wikipedia.org/wiki/Data_governance) is complex but in our work, we will examine *processes and policies that ensure the quality of data, data security and privacy of the sensory data and sustainability measurements in these platforms, and the data lifecycle to comply with sustainability regulations*.
- Stakeholder governance: reflects the role of stakeholders, e.g., how stakeholder access data. This is based on interests and roles of stakeholders in corporate governance (see http://en.wikipedia.org/wiki/Corporate_Governance). In our work, we will examine *how well existing platforms support stakeholders in the ecosystem of sustainable facilities*.

2.2 Facility Monitoring Using Service-oriented Architecture

The service-oriented architecture (SOA) model has been applied to monitor facilities over the past few years. However, to date the main use of the service model in this respect is focused on the use of Web services to remotely monitor and control these *monitored objects*[1] with a basic facility management model. In this model, typically the owner of monitored objects monitors and controls her objects. Supports for sustainability governance are negligible, as all monitored data is owned and managed by the owner. For example, in (Acker & Massoth, 2010) home is monitored via Web services. A SCADA system accessed from Web service is given in (Lipnickas, Rutkauskas & Cerkauskas, 2009). Web-based systems for buildings and energy management have been demonstrated to be very useful (Granderson, Piette & Ghatikar, 2011; Capehart & Capehart, 2005).

Several frameworks have been developed to support the integration of different monitoring sensors to provide data for buildings, houses and transportation vehicles, such as (Tompros, Mouratidis, Draaijer, Foglar & Hrasnica, 2009; Krishnamurthy, Anson, Sapir, Glezer, Rois, Shub & Schloeder, 2008; Choi, Shin & Shin, 2005). Such monitoring data can also be exposed through Web services and integrated into business processes (Guinard, Trifa, Karnouskos, Spiess & Savio; 2010). While middleware can be used to relay monitoring data to consumers, such as (Broering, Foerster, Jirka & Priess, 2010; Beywatch Consortium, 2010), these systems are typically limited to the boundary of a single organization. It means that a system can be used to monitor objects in distributed facilities but there is only a single owner and consumer[2] of the system. In our view, techniques for integration sensors are enabling technology for providing data which can be stored and processed by the cloud model but they do not support the cloud computing model in which they act as a platform for multiple organizations.

Recently, several cloud-based platforms to support the monitoring of energy consumption have been introduced, such as Tendril (Tendril, 2011) and AlertMe (AlertMe, 2011). These systems, handling only data from their own devices, act as a platform to store electricity consumption information which is updated and accessed from different homes. However, they are mostly for near real-time monitoring rather than for sustainability governance. Going beyond these monitoring systems, generic cloud-based services have been provided to store different types of monitoring data to facilitate sustainability monitoring and analysis, such as Pachube (Pachube, 2011). Furthermore, there are systems supporting sustainability governance for buildings such as ECView (Thyagarajan et al., 2010) and Galaxy (Pacific Control Systems, 2011). Some systems have supported generic ways to determine carbon footprints based on standard profiles for sustainability governance, such as the AMEE platform (AMEE, 2011). While there are several research reports on generic computational and data storage cloud systems (Armbrust et al., 2010; Wang et al., 2008), and for e-science (Sullivan, 2009), we are not aware of any work discussing how sustainability governance in general utilizes cloud computing offerings and how cloud computing could be useful for sustainability governance.

[1] Monitored objects in this paper should be understood as any facility and its sub-components/elements that can be monitored in order to characterize their well-being status and determine resources they consumers. To monitor an object directly, one or multiple sensors will be used. To monitor an object indirectly, one or multiple sensory data sources will be analyzed.

[2] We distinguish between a consumer and an objective of a consumer. In these systems, a consumer can have different objectives, e.g., to monitor facility, to analyze energy consumption, and to improve air quality. Thus, the consumer will need different types of data and processes. However, the entire system is designed for only a consumer.

3 Towards Cloud-Based Sustainability Governance

3.1 The Ecosystem of Facility Sustainability Governance

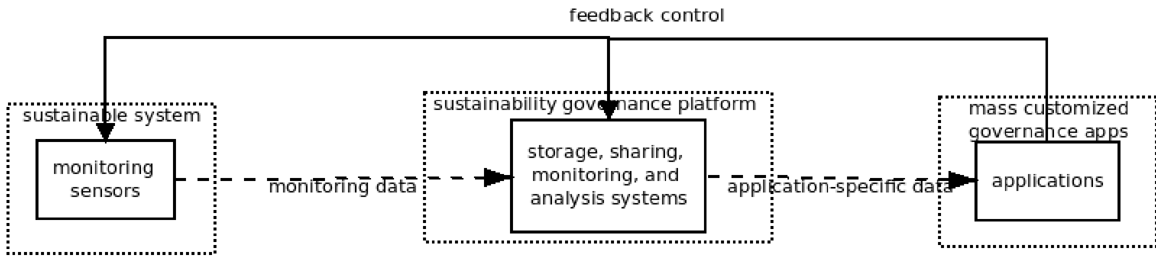

Figure 1: Main building blocks of an end-to-end system for sustainability governance

In sustainability governance, we have to consider the complexity of data. There exist many different types of monitoring data, each type for a kind of (e.g., a chiller) or a part of a monitored object (e.g., a room). Methods and algorithms for sustainability analysis are complex due to the huge amount of different monitored objects and sustainability measurements (e.g., even GHG has 10 primary different types (Center for Sustainable Systems, 2010)). These methods and algorithms will rely on a large set of reference/standard profiles.

Moreover, for large-scale facilities, several stakeholders conduct different activities that are inherent to the evolution of the ecosystem. Conceptually, in an end- to-end view of sustainability governance, shown in Figure 1, first, monitoring sensors are used to monitor (sustainable) systems to provide monitoring data for data analysis to determine sustainability measurements. Second, monitoring data will be stored for the analysis of sustainability measurements which involves complex calculation, estimation and prediction methods and utilizes various reference models. Third, application-specific sustainability measurements will be provided to governance applications. Along these paths, different activities are performed by different stakeholders and data will be stored, analyzed and shared due to sustainability governance rules. From the system architecture and integration perspective, these main building blocks can be distributed and provided by different providers. Furthermore, interactions among these building blocks can be carried out via the Internet.

In order to understand why cloud computing could offer benefits for sustainability governance, we must analyze the stakeholders and their roles in the ecosystem, and the evolution of the ecosystem, by considering the above-mentioned end-to-end data flows. We have observed several activities required for sustainability governance that are performed by different stakeholders, shown in the Figure 1. Main activities are:

- *gather and store monitoring data for sustainability measurement*: this can be done automatically or manually using different methods, such as monitoring sensors push data to the platform or the platform pulls data from monitoring sensors.
- *retrieve data for sustainability analysis*: data inside the platform and from other platforms can be retrieved, such as via querying or subscription, and combined, such as via data composition/mashup, for sustainability monitoring and analysis using different methods.
- *gather and manage reference models*: several reference models used in analysis methods (e.g., calculation, estimation and prediction) are developed by standard organizations, companies, and domain experts. Such reference models can be gathered and provisioned to different consumers.
- *monitor and analyze sustainability measurements*: near real-time monitoring features, such as alarm services, can be conducted based on monitoring data. Measurements and behaviors of monitored objects, such as estimating GHG emission or predicting electricity consumption, can be performed based on monitoring data gathered over time.
- *share monitoring data and sustainability measurements*: both monitoring data and measurements can be made available for other purposes, such as chiller data can be shared among chiller manufacturer, maintainer, and owner.
- *develop applications and analysis algorithms*: applications to utilize monitoring data and sustainability measurements as well as analysis methods/ algorithms for sustainability measurements can be developed.

- *examine and certify data, techniques and applications*: certain data collections and models, techniques and applications can be certified to comply with existing regulations.

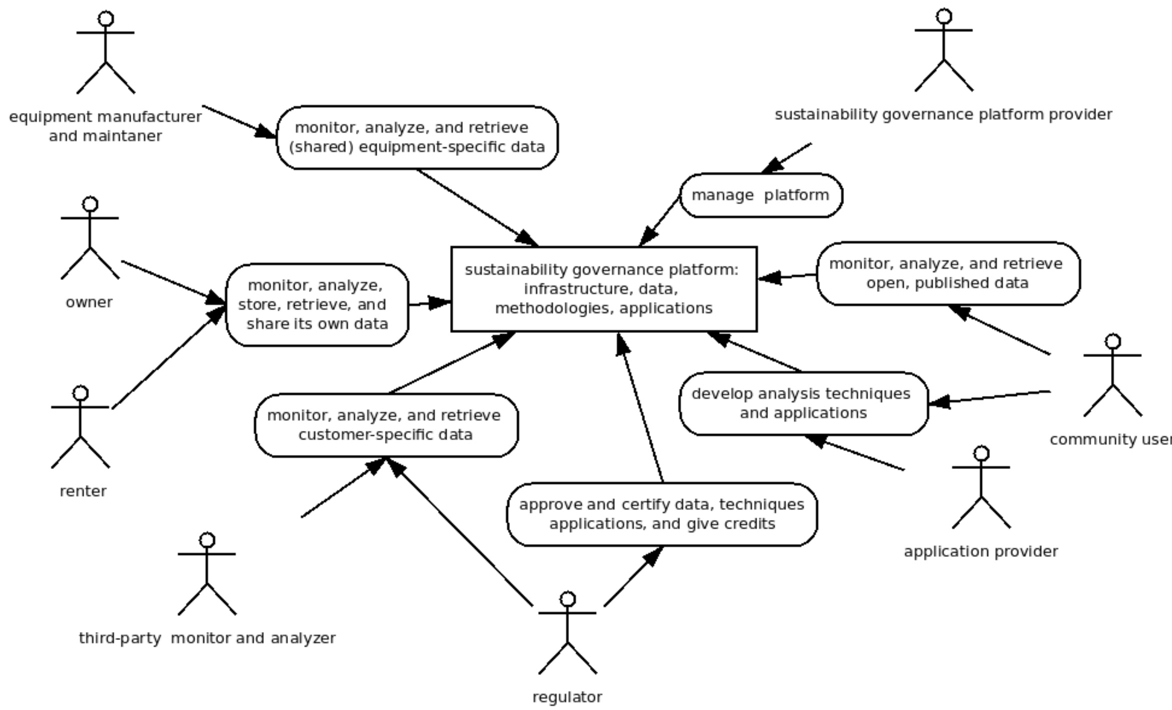

Figure 2: Stakeholders and their activities in sustainability governance systems

The above-mentioned activities are performed by different stakeholders, explained in Table I. Figure 1 also describes which activities can be performed by which stakeholders in the ecosystem of the sustainability governance. Note that all stakeholders might conduct different activities for the same monitored object. Therefore, multiple views should be supported in sustainability governance for monitored objects. Obviously, existing single-organizational facility monitoring systems discussed in Section 2.2, even based on SOA, are not capable of support of multiple stakeholders as identified in the ecosystem of sustainable facilities. Most of them just support the owner of the facilities. Furthermore, most of them do not consider data governance based on compliance regulations.

Types of Stakeholders	**Description**
owner	the owner of the monitored (sustainable) system. The owner can provide data about her own equipments, buildings, etc. The owner can perform various monitoring, analysis and sharing activities
renter	the renter of certain parts of the monitored (sustainable) system. The renter can provide data about her own spaces and equipments as well as can perform monitoring and analysis of her own spaces and equipments
equipment manufacturer and maintainer	those who have a right to access data about equipments in a facility in order to perform the maintenance

third-party governance monitor, analyzer, and auditor	the third-party has a contract with the owner, renter or the manufacturer to perform activities on behalf of them. Therefore, she has the right to monitor, analyze, and retrieve certain customer-specific data.
regulator	the regulator can examine claimed data, techniques and applications and certify them for their compliance with regulations. Furthermore, the regulator can also give credits to other stakeholders to obtain rewards for their achievements (e.g. saving energy)
community user	the user in a community who is interested in utilizing offerings for the community. Therefore, she can access (unlimited) open, published data as well as functionality of the governance platform. Not that we do not consider limited usage for developers as a community aspect.
application provider	any provider offering analysis techniques and applications for sustainability governance
governance platform provider	the provider of the governance platform

Table 1: Main stake holders in the ecosystem of sustainability governance

3.2 Cloud Computing as a Solution for Facility Sustainability Governance

With the ecosystem of sustainability governance and its evolution, consumers (end-users or enterprises) will face complexity in the management and sharing of monitoring data, analysis methods, and reference profiles required in governing sustainability. Therefore, systems supporting the enterprise model, such as (Swords et al., 2008; Beywatch Consortium, 2010), will not be adequate.

We believe that the cloud computing model will be a solution for sustainability governance due to several reasons:

- cloud computing offers a better long-term data management and sharing: large types and amount of monitoring data and reference profiles are supported well with the cloud model. Furthermore, the cloud computing model would hide the complex infrastructure from the end users who are typically not in the IT sector.
- cloud computing offers a better way to collect and develop reference profiles and complex analysis algorithms: such profiles and algorithms require a strong participation of different vendors, standard organizations, companies and domain experts. The cloud computing model would offer good means supporting strong participation of different stakeholders (e.g., via community model and sharing platform).
- cloud computing is able to handle large, elastic demands: the sheer number of monitored objects and complex data analysis methods makes the execution of data handling and analysis processes very complex and requires varying workload.

In the following section, we will outline an architectural view for cloud-based sustainability governance systems.

3.3 Architectural Views for Cloud-based Sustainability Governance Systems

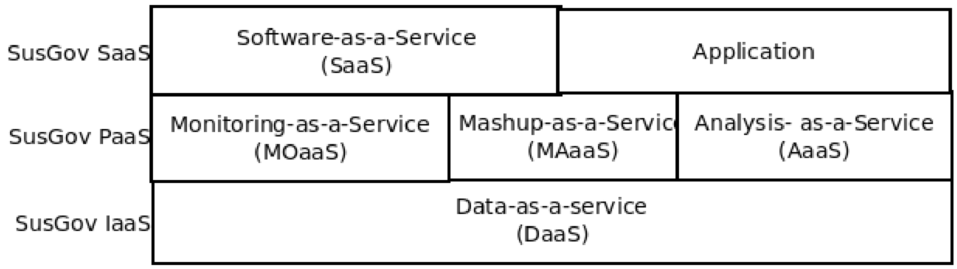

Figure 3: Layered architecture of cloud-based sustainability governance

Figure 3 presents a high-level architectural view for cloud-based sustainability governance systems. In general, the types of cloud systems can be classified into cloud Infrastructure as a service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) (Mell & Grance, 2009). In the context of facility sustainability governance (SusGov), We envisage five main types of cloud services in the ecosystem of sustainability governance. These types of services can be organized into the typical IaaS/PaaS/SaaS model in cloud computing as follows:

- SusGov DaaS (Data-as-a-Service) will provide storage for monitoring data and reference profiles for sustainability gov- ernance. We consider SusGov DaaS belonging to IaaS as data in SusGov DaaS will include monitoring data, reference profiles, and algorithms.
- SusGov MOaaS (Monitoring-as-a-Service), MAaaS (Mashup-as-a-Service) and AaaS (Analysis-as-a-Service) are in the PaaS category providing programming capabilities for sustainability governance, such as data mashup and enrichment, algorithm development, and data analysis workflows. SusGov MOaaS, MAaaS and AaaS will provide several common features for acquiring and analysis data to support high-level applications. They also will support among SusGov cloud systems.
- SusGov SaaS will include several customized applications that rely on SusGov IaaS and PaaS services suppo rting different potential applications, such as online visualization, dashboard, energy optimization, and GHG compliance.

In the following subsections, we will elaborate these types of services.

3.3.1 SusGov IaaS

SusGov DaaS will be used to store sensor data, reference profiles and algorithms so essentially it has certain requirements can be fulfilled by existing cloud data storages, such as large-scale storage capability. However, SusGov IaaS in our view should be designed for sustainability governance of facilities. It has to include (i) storage mechanisms suitable for facility monitoring data and (ii) data governance policies and processes suitable for sustainability governance. Therefore, in our view, basic storage-based IaaS, like Amazon S3 or Microsoft Azure, could provide basic storage and query mechanism but it is open to see if these basic IaaS can be used for large-scale and dynamic sensor data. Furthermore, whichever basic storage-based IaaS to be used, we need to develop specific data governance processes and policies for sustainability governance.

Another important issue is that SusGov IaaS has to store and mange reference profiles for sustainability analysis of facilities. Basically, for each facility model its reference profile provides basic information that can be used for determining sustainability measurements for the facility or provides reference sustainability measurements for that model. Currently, several reference models are provided, such as for cars in the AMEE platform. With respect to reference profile management and sharing, the cloud model is very suitable as reference models are typically defined by (standard) organizations and agreed by many parties.

3.3.2 SusGov PaaS

SusGovPaaS includes a collection of services that can be composed and utilized for monitoring and analysis of sustainability measurements, which are fed to different applications. In our view, on the one hand, SusGov MOaaS and SusGov AaaS are specially designed to deal with sustainability monitoring and measurements. Hence, they are domain-specific services that include techniques for analyzing sustainability measurements, such as GHG and energy consumption, based on facility sensor monitoring data. As there are several analytics and estimation models for different monitored objects, there could be several types of SusGov MOaaS and AaaS. SusGov MAaaS, on the other hand, could utilize common techniques in data and service mashups. In particular, SusGov MaaS also will support data mashup among SusGov cloud systems for SusGov MOaaS and AaaS.

3.3.3 SusGov SaaS

Different stakeholders require different governance features, thus different applications for sustainability governance that must be developed and provided. Such applications can be offered via the well-known SaaS model in the cloud. SusGov SaaS could provide features commonly found in enterprise facility monitoring systems, such online visualization of sustainability measurements, trend analysis, what-if analysis, etc. However, other advanced SaaS could be developed based on the analysis of large-scale, cross-facility monitoring data, such as recommendations of facility configuration and equipments for architecture, construction and engineering of new buildings.

4 Production Cloud Systems for Facility Sustainability Governance

To examine how production cloud systems can be used for sustainability governance, we selected Galaxy (Pacific Control Systems, 2011), Pachube (Pachube, 2011), and AMEE platform (AMEE, 2011) for further detailed analysis. They were chosen because, first, they are the only few cloud systems facilitating sustainability governance, and, second, they are complementary to providing different cloud-based services for facility sustainability governance.

Table 3 presents overall characteristics of the selected systems. These systems are diverse, offering different functionalities and are being targeted to different sectors. We observe several points. First, except Galaxy, existing systems do not provide an end-to-end solution and focus on the platform only. Those systems with non end-to-end solution, such as Pachube, can enable a wide range of data consumers and data providers, however, they also have some drawbacks, such as it is difficult to verify if monitoring data is authentic to be complied with regulations or it is difficult to enable data sharing among many stakeholders and sustainability analysis. Second, while existing systems can be used by multiple stakeholders, sharing the same data object to different stakeholders for different purposes is not well supported. In fact, existing systems tend to allow sharing the whole object or nothing. Third, existing systems currently do not support regulator roles. There is no such a concept that the regulator can certify data models, policies, algorithms and techniques for sustainability compliance in these systems.

Feature	**Galaxy**	**Pachube**	**AMEE**
end-to-end system	yes	no	no
type of cloud systems	SusGov DaaS, MOaaS, MAaaS, AaaS, SaaS	SusGov DaaS	SusGov DaaS, AaaS
integration with other cloud systems	no	no	no
monitored objects supported	buildings, building elements, equipments	neutral	neutral
sustainability measurements	energy consumption, GHG (carbon footprints)	no (only for storage and exchange)	GHG (carbon footprints)
types of stakeholder	all, except community user, application provider, and regulator	all, except regulator	all, except regulator
pricing model	subscription	free, subscription	subscription

Table 2: Overall features of cloud systems for sustainability governance

Table 4 examines different capabilities of the systems studied. In terms of DaaS, first of all, while using different terms, overall, most services use a similar level of hierarchical abstraction for modeling data concepts. Most systems arrange monitoring data into three levels associated with monitored objects, different types of time series data, and data values. While such abstractions may be enough for describing individual monitored objects, e.g., a chiller or a car, they are not capable of representing complex dependencies among monitored objects. All services support REST APIs and Galaxy specially supports JMS (Java Message Service). Currently, low-level monitoring data is provided using several data models, such as Open Building Information Exchange (oBIX) (oBIX, 2011), Extended Environments Markup Language (EEML) (EEML, 2011), and vendor-specific XML specifications. There is little support on reference models and algorithms catalog. For example, AMEE is a rare example that supports reference profiles for carbon footprints. Another important point is that we did not find how existing systems support data provenance that can be useful for regulators to verify the authenticity of monitoring data, in cases, giving credits for sustainability achievement.

Concerning capabilities of SusGov PaaS, most systems do not support data mashup capabilities - to combine different data sources for different purposes, and studied systems do not support customer-provided analysis methods. An exception is Galaxy, which provides data mashup and user-provided analysis methods by defining equations based on monitoring data, however, these methods do not support complex mashup data and analysis. Second, one important feature is the search capability for analysis methods and reference profiles, but among studied systems, AMEE is the only one supporting reference profiles based on several standards. Third, systems without PaaS and SaaS features, such as Pachube, can only support the analysis via third parties or only provide simple facility monitoring to show aggregated monitoring data. Overall, even with systems supporting analysis, the analysis is relied only data in the system. Four, so far, we have not observed that regulation compliance processes are integrated into these systems.

Concerning application support capabilities, since Pachube and AMEE are mainly DaaS, they do not support several forms of applications but third party applications can be developed by using their APIs, thus some can also support the app store development model: applications are developed by third parties but rely on platform APIs. With most systems, similar types of applications are generally supported, such as historical usage visualization, near real-time monitoring, and dashboard.

In overall, first, we see that it is possible that a SusGov system will have an end-to-end control, from sensor monitoring, sensor integration to IaaS, PaaS and to SaaS, like in the case of Galaxy. Contemporary cloud systems for enterprise computing tend to be either IaaS, PaaS or SaaS. The selection of an end-to-end control model, like Galaxy, or just be a DaaS, PaaS, or SaaS, like Pachube, has a strong impact on the development of service, data and stakeholder governance policies and processes as well as on SusGov interoperability. However, we believe that the selection is due to the business model of different vendors. From the technical perspective, to enable the integration of different SusGov systems and stakeholders, we need to rely on open protocols. Second, SusGov systems in our studied mainly have certain services to support sustainability governance but they do not focus on specific governance requirements for sustainability governance. For example, while several data security and privacy rules and techniques have been developed, it is not clear how privacy and security would be applied to which types of sensor data for which types of monitored objects.

Feature	**Galaxy**	**Pachube**	**AMEE**
core data concepts	Facility → Asset → Point → Value	Environment→ Datastream→ Datapoint	Data Category → Data Item → Data Value
monitoring data integration			
data model	oBIX (*Open Building Information Exchange* n.d.)	EEML(*Extended Environments Markup Language (EEML, 2011)*)	specific
data formats	XML, CSV	CSV, XML and JSON	JSON/XML
API	REST, SOAP	REST, Socket	REST
push-to-platform	yes	yes	yes
pull-by-platform	yes	yes	no
monitoring data publishing			
data model	specific (dependent on roles)	RSS/ATOM with EEML	specific JSON/XML
data publishing model	specific	private, community public	private
historical data retrieval	yes	yes	yes
API	REST, SOAP, JMS	REST	REST
pull-by-consumer	yes	yes	yes
push-by-platform	yes	yes	no

programming features for data and analysis			
composition of existing monitoring data	yes (but limited by using rules for fusing events in GUI of SaaS)	no	no
composition of analysis algorithms/data analytics workflows	yes (but limited by using equation editors in GUI of SaaS)	no	no
application programming framework	no	no	yes
built-in analysis algorithms	yes	no	yes
reference profiles	no	no	yes
application store	no	yes	yes
monitoring features			
notification/trigger	yes	yes	no
near real-time monitoring/observing	yes	yes	yes
sustainability measurements			
energy consumption	yes	no	yes
GHG	yes	no	yes
application support			
Web-based SaaS	yes	no	no
Mobile application	no	yes (via third parties)	no
Web widget	yes	yes	no

Table 3: Sustainability governance support in industrial products

5 Research Cloud Systems for Sustainability Governance

Consider building blocks of an end-to-end system for sustainability governance in Figure 1, several techniques have been developed for different purposes but can be used, as parts of, for sustainability governance. For example, for sensors and Sensor Web several techniques have been developed to capture different types of monitoring data and related monitoring data to the central places or allows monitoring data to be accessed. They are however do not follow the cloud computing model, either they follow the Web, everyone can access the data, such as Sensor Web (Gibbons et al., 2003) or designed for specific purposes (Wang et al., 2010; Fairgrieve et al., 2009). Furthermore they focus are on monitoring, rather than sustainability governance (although data are usually archived). Although several techniques are common, such data query, data integration, realtime monitoring, etc., we will not discuss them in this survey since we focus on cloud platforms, rather enabling techniques for sensor data integration. In SOA, different techniques have been developed for integrating sensors via the service model, such as in (Guinard, Trifa, Karnouskos, Spiess & Savio; 2010). However, although we consider the integration of sensors into SOA-based platforms as a fundamental part of the whole end-to-end facility governance, it is not the focus of our study in this paper.

Considering our architectural view of SusGov systems, different techniques have been developed atop cloud infrastructure that can be utilized. Table 5 describes relevant techniques for cloud-based sustainability governance:

- *SusGov DaaS*: investigation of cloud computing for storing and processing sensor data has been conducted recently. Rolewicz et al. present techniques to access sensor data stored in their cloud using HBase, etc. (Rolewicz et al., 2011). Although it has not been tested with facility sustainability governance, it could be useful for the development of SusGov DaaS wrt monitoring data. However, several issues related to stakeholders and how they access data have not been addressed.
- *SusGov MOaaS*: With respect to sensor data and its processing, (Simmhan, Giakkoupis, Cao & Prasanna, 2010) presents how sensor data from smart meters can be processed using cloud virtual machines. Their work introduces neither a SusGov system nor establish SusGov specific governance techniques. However, they present ideas about a platform (PaaS) for processing data on the fly. This is similar to recent stream data analysis frameworks that handle events using cloud infrastructures ( ). Since handling events on the fly is one goal of SusGov MOaaS, these works could be useful for the design of SusGov MOaaS atop cloud infrastructures.
- *SusGov AaaS*: Several research works have been investigated for the so-called computational sustainability (Gommes 2009), in which processes for analyzing sustainability measurements are developed, e.g., utilizing archived data with workflow, data analytics, etc. There are several workflow systems, which are able to analyze data from different sources, such as Kepler (Ludäscher et al., 2006), Taverna (Hull et al., 2006), Trident (Trident, 2011). However, they have not been tested and integrated for sustainability analysis. Still many analysis algorithms are implemented in sequential programs, R and MathLab scripts. Patnaik et al. show data analytics techniques for analyzing chillers in data centers but this work is just focused on data analysis aspect isolated from cloud sustainability governance systems. ~~In sustainability governance of facilities,~~ Overall, a research system based on cloud computing model for sustainability governance has not been observed. Several research efforts has spent to develop enabling techniques that allow us to easy connect sensor data to cloud but there is a lack of integrated system, the lack of techniques to support multiple stakeholders and governance and lack of integration of analysis workflows.

Feature	End-to-end system	Type of SusGov	Integration with other clouds	Monitored object support	Sustainability measurements	Stakeholders	Pricing
Rolewicz et al., 2011	No (store data and provide access to the data)	SusGov DaaS	no	No particular objects (generic storage)	No (only for storage)	No discussion	No discussion
Simmhan, Giakkoupis, Cao & Prasanna, 2010	No (it is a software architecture)	SusGov DaaS and PaaS	no	Smart grids	Energy consumption	No discussion	No discussion
Patnaik et al., 2011	No (only data analytics)	Only data analytics	no	Chillers	CO_2, water consumption, electricity consumption	No discussion	No discussion

Table 4: Examples of relevant researches for cloud-based sustainability governance systems

6 Open Research Issues on Sustainability Governance Using Cloud Computing

What we have observed in the previous section is that existing systems have basic support for data storage and data retrieval but they are still limited to basic monitoring and analysis, with very little, if at all, support for multiple stakeholders, complex analysis and compliance processes. To improve the support of sustainability governance using the cloud computing model, we believe we need to address the following points:

Linked Data Concepts and Monitored Object Dependencies: Currently, for example, our studied systems mainly support a hierarchical structure of monitored objects, data streams and individual data points based on different specifications. We observed two issues: (i) data exchange among cloud systems for sustainability governance and (ii) complex dependency among monitored objects in the analysis of sustainability measurements. The first issue is the difficulty to utilize different cloud systems due to the different data models in different cloud systems. The second issue is that it is difficult to support complex analyses requiring multiple types of data. Due to the diversity of types of monitored objects, we do not expect a single data model to represent reference profiles and monitoring data. However, as complex sustainability analysis requires different types of monitoring data, we expect data models to be linked to reflecting the dependency of monitored objects they represent. Currently, generic data models are used for measurements produced by monitoring sensors but they represent individually monitored objects. We lack a mechanism to specify the above-mentioned dependencies. In particular, most systems provide monitoring data but metadata about monitored objects are inadequate.

Analysis Algorithms and Reference Profile Management: The management of reference profiles for monitored objects and analysis algorithms is very challenging. These algorithms are diverse and complex in terms of structures, input parameters, reference models as well as computational resource requirements. First, analysis algorithms also have different properties: they can contain or require simple or complex analytics, e.g., based on calculation, estimation, prediction methods implemented with different algorithms. Moreover, algorithms can require different computational resources backed by different libraries and data sources. Second, reference profiles for diverse monitored objects are also complex and they need to be collected and maintained over time. Algorithms for calculating, estimating and predicting sustainability measurements have been on the focus of researchers only recently. Currently, research on how to manage analysis algorithms and how to provide an open platform for third parties to develop, search and share algorithms is quite open. Furthermore, algorithms have a strong dependency on reference profiles and compliance rules. So far, how to manage and link them is still an open issue.

Cloud Interoperability: In the ecosystem of sustainability governance, we have observed complementary cloud services for sustainability governance. For example, SusGov DaaS, like Pachube, can be utilized by any consumer to store monitoring data, while SusGov AaaS, like AMEE and Galaxy, can offer data analysis capabilities, such as complex carbon footprints estimation. Although a few applications have been developed to utilize different SusGov cloud systems, the issue of SusGov cloud interoperability and dependency has not been studied yet. To support the interoperability, the issues of data modeling and algorithms and reference profiles discussed previous will also play a crucial role.

Data governance: as the amount of monitoring data for sustainability governance of facilities is huge, to ensure the quality of data and to determine suitable data lifecycle for which types of data are challenging. This challenging issue is due to the fact that it is not clear where and when we can apply there techniques. For example, quality of sensor (Klein and Lerner, 2009) can have a strong influence in sensor level, gateway and in the system. We need to classify which types of sensors that should be applied for privacy and security, which types of sensor must be kept for a long time while other can be short. Furthermore, how to apply these rules and processes are still open.

Programmable PaaS: currently, PaaS, like PCS, is close, not open, we seek for programmable PaaS in which scientists, stakeholder, can program analyses based on monitoring data. This requires existing SusGov PaaS to open not only data APIs but also mechanisms to develop and execute different analytics implemented in different languages, such as Java, Matlab and R.

7 Conclusions

Sustainability governance for resource consumption by humans, such as in buildings and personal transportation, needs scalable and interoperable platforms to support the integration, sharing, monitoring, analysis of different types of monitoring data and sustainability measurements. Due to the complexity of data management, monitoring and analysis, cloud computing emerges as a potential candidate for supporting sustainability governance. While there are many open issues, as we have analyzed some production cloud systems enabling the sustainability governance, with their potentials being able to deal with complex data storage, sharing and analysis for multiple stakeholders in Internet-scale, we believe that cloud computing services specially designed for sustainability governance will play an important role in the future.

However, we need to provide several techniques to address the management of large-scale monitoring data and to link different types of monitoring data with monitored object description in order to support advanced analysis. Furthermore, we need to provide high-level data management for multiple types of stakeholders and APIs for analytics development. Currently, our future work is focused on linked data models for stakeholders, monitored objects, monitoring data and analytics applications in sustainability governance platforms. Moreover, we are working on a SusGov PaaS for facility governance.

Acknowledgements

The work mentioned in this paper is partially funded by the Pacific Control Cloud Computing Lab. We thank Vivek Sundaram for his discussion and support on an early draft of this paper. We thank our colleagues in the Pacific Control Cloud Computing Lab for providing detailed information about the Galaxy platform.

References

Acker, R. & Massoth, M. (2010), Secure ubiquitous house and facility control solution, in ‘Proceedings of the 2010 Fifth International Conference on Internet and Web Applications and Services’, ICIW ’10, IEEE Computer Society, Washington, DC, USA, pp. 262–267. URL: http://dx.doi.org/10.1109/ICIW.2010.45

AlertMe (2011). http://www.alertme.com.

AMEE (2011). http://www.amee.com/, Last access: 7 Feb 2011.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I. & Zaharia, M. (2010), ‘A view of cloud computing’, *Commun. ACM* **53**(4), 50–58.

Broering, A., Foerster, T., Jirka, S. & Priess, C. (2010), Sensor bus: an intermediary layer for linking geosensors and the sensor web, *in* ‘Proceedings of the1st International Conference and Exhibition on Computing for Geospatial Research & Application’, COM.Geo ’10, ACM, New York, NY, USA, pp. 12:1–12:8. **URL:** *http://doi.acm.org/10.1145/1823854.1823870*

Capehart, B. L. & Capehart, L. C., eds (2005), *Web Based Energy Information and Control Systems: Case Studies and Applications*, Fairmont Press, Inc. Choi, J., Shin, D. & Shin, D. (2005), Research and implementation of the context-aware middleware for controlling home appliances, *in* ‘Consumer Electronics, 2005. ICCE. 2005 Digest of Technical Papers. International Conference on’, pp. 161 – 162.

Beywatch Consortium(2010), ‘Deliverable d2.2: End-to-end platform specification beywatch data model (annex)’. http://www.beywatch.eu/docs/D22 Annex.pdf.

Extended Environments Markup Language (EEML) (n.d.). http://www.eeml.org.

Fairgrieve, S.M.; Makuch, J.A.; Falke, S.R.; , "PULSENet™: An implementation of Sensor Web standards," Collaborative Technologies and Systems, 2009. CTS '09. International Symposium on , vol., no., pp.64-75, 18-22 May 2009
doi: 10.1109/CTS.2009.5067463
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5067463&isnumber=5067423

Klein, A., Lehner, W., (2009). Representing Data Quality in Sensor Data Streaming Environments. J. Data and Information Quality 1, 2, Article 10 (September 2009), 28 pages. DOI=10.1145/1577840.1577845
http://doi.acm.org/10.1145/1577840.1577845

Phillip B. Gibbons, Brad Karp, Yan Ke, Suman Nath, and Srinivasan Seshan. 2003. IrisNet: An Architecture for a Worldwide Sensor Web. IEEE Pervasive Computing 2, 4 (October 2003), 22-33. DOI=10.1109/MPRV.2003.1251166
http://dx.doi.org/10.1109/MPRV.2003.1251166

Carla P. Gomes (2009) , Computational Sustainability: Computational Methods for a Sustainable Environment, Economy, and Society, Frontiers of Engineering 39:4, Winter, 2009, http://www.nae.edu/Publications/Bridge/17281/17286.aspx

Center for Sustainable Systems, University of Michigan (2010). "U.S. Greenhouse Gases Factsheet." Pub. No. CSS05-21. http://css.snre.umich.edu/css_doc/CSS05-21.pdf

Granderson, J., Piette, M. & Ghatikar, G (2011), 'Building energy information systems: user case studies', *Energy Efficiency* **4**, 17–30. 10.1007/s12053-010-9084-4. **URL:** *http://dx.doi.org/10.1007/s12053-010-9084-4*

Guinard, D., Trifa, V., Karnouskos, S., Spiess, P. & Savio, D. (2010), 'Interacting with the soa-based internet of things: Discovery, query, selection, and on-demand provisioning of web services', IEEE Transactions on Services Computing 3, 223–235.

Duncan Hull, Katy Wolstencroft, Robert Stevens, Carole Goble, Mathew R. Pocock, Peter Li, Tom Oinn, Taverna: a tool for building and running workflows of services,
Nucleic Acids Res. 2006 July 1; 34(Web Server issue): W729–W732. Published online 2006 July 14. doi: 10.1093/nar/gkl320

Krishnamurthy, S., Anson, O., Sapir, L., Glezer, C., Rois, M., Shub, I. & Schloeder, K. (2008), Automation of facility management processes using machine-to-machine technologies, in 'Proceedings of the 1st international conference on The internet of things', IOT'08, Springer-Verlag, Berlin, Heidelberg, pp. 68–86. URL: http://portal.acm.org/citation.cfm?id=1793060.1793067

Pierre de Leusse, Theo Dimitrakos, and David Brossard. 2009. A Governance Model for SOA. In Proceedings of the 2009 IEEE International Conference on Web Services (ICWS '09). IEEE Computer Society, Washington, DC, USA, 1020-1027. DOI=10.1109/ICWS.2009.132 http://dx.doi.org/10.1109/ICWS.2009.132

Lipnickas, A., Rutkauskas, R. & Cerkauskas, R. (2009), Interoperability of scada system applications with web services, *in* 'Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2009. IDAACS 2009. IEEE International Workshop on', pp. 196 –200.

Ludäscher B., Altintas I., Berkley C., Higgins D., Jaeger-Frank E., Jones M., Lee E., Tao J., Zhao Y. 2006. Scientific Workflow Management and the Kepler System. Special Issue: Workflow in Grid Systems. Concurrency and Computation: Practice & Experience 18(10): 1039-1065

Mell, P. & Grance, T. (2009), 'The nist definition of cloud computing', http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc. V15.

Open Building Information Exchange (n.d.). http://www.obix.org/.

Pachube (n.d.). http://www.pachube.com/, Last access: 7 Feb 2011.

Pacific Control Systems (2011), 'Galaxy'. http://pacificcontrols.net/cloudservices/galaxy.html, Last access: 24 Oct 2011.

Debprakash Patnaik, Manish Marwah, Ratnesh K. Sharma, and Naren Ramakrishnan. 2011. Temporal data mining approaches for sustainable chiller management in data centers. ACM Trans. Intell. Syst. Technol. 2, 4, Article 34 (July 2011), 29 pages. DOI=10.1145/1989734.1989738 http://doi.acm.org/10.1145/1989734.1989738

Simmhan, Y., Giakkoupis, M., Cao, B. & Prasanna, V. K. (2010), On using cloud platforms in a software architecture for smart energy grids, *in* 'International Conference on Cloud Computing Technology and Science (CloudCom)', IEEE. Poster. **URL:** *http://salsahpc.indiana.edu/CloudCom2010/EPoster/cloudcom2010 submission 269.pdf*

B. Swords, E. Coyle, B. Norton (2008), An enterprise energy-information system, Applied Energy, Volume 85, Issue 1, January 2008, Pages 61-69, ISSN 0306-2619, 10.1016/j.apenergy.2007.06.009.

Sullivan, F. (2009), 'Guest editor's introduction: Cloud computing for the sciences', *Computing in Science and Engineering* **11**(4), 10–11.

Ian Rolewicz, Michele Catasta, Hoyoung Jeung, Zoltán Miklós, Karl Aberer: Building a Front End for a Sensor Data Cloud. ICCSA (3) 2011: 566-581

Tendril (2011), http://www.tendrilinc.com. Last access 29 October 2011.

Thyagarajan, G, Sarangan, V., Sivasubramaniam, A., Suriyanarayanan, R., Chitra, P. & Y, A. (2010), 'Managing carbon footprint of buildings with ecview', *Computer* **99**(PrePrints).

Tompros, S., Mouratidis, N., Draaijer, M., Foglar, A. & Hrasnica, H. (2009), 'Enabling applicability of energy saving applications on the appliances of the home environment', Netwrk. Mag. of Global Internetwkg. 23, 8–16. URL: http://dx.doi.org/10.1109/MNET.2009.5350347

Project Trident: A Scientific Workflow Workbench, 2011. http://research.microsoft.com/en-us/collaboration/tools/trident.aspx. Last access: 28 October 2011.

Wang, L., Tao, J., Kunze, M., Castellanos, A. C., Kramer, D. & Karl, W. (2008), Scientific cloud computing: Early definition and experience, in 'HPCC', IEEE, pp. 825–830.

Demin Wang, Dharma P. Agrawal, Wassana Toruksa, Chaichana Chaiwatpongsakorn, Mingming Lu, and Tim C. Keener. 2010. Monitoring ambient air quality with carbon monoxide sensor-based wireless network. Commun. ACM 53, 5 (May 2010), 138-141. DOI=10.1145/1735223.1735257 http://doi.acm.org/10.1145/1735223.1735257

Phyl Willson and Carol Pollard. 2009. Exploring IT Governance in Theory and Practice in a Large Multi-National Organisation in Australia. Inf. Sys. Manag. 26, 2 (March 2009), 98-109. DOI=10.1080/10580530902794760 http://dx.doi.org/10.1080/10580530902794760