Governing Bot-as-a-Service in Sustainability Platforms – Issues and Approaches

Hong-Linh Truong\textsuperscript{a}, Phu H. Phung\textsuperscript{b}, Schahram Dustdar\textsuperscript{a}

\textsuperscript{a}Distributed Systems Group, Vienna University of Technology, Austria
\textsuperscript{b}ProSec Security Group, Chalmers University of Technology, Sweden

Abstract

The emerging cloud computing models for Internet-of-Things have fostered the development of lightweight applications using cloud services for monitoring and optimizing devices and equipment hosted in distributed facilities. Such applications – called bots in our work – can be composed and deployed with multiple types of governance policies from cloud platforms to distributed hosting environments and they can access not only local data and devices but also cloud data and features. Therefore, it is a great challenge to govern them. In this paper, we discuss governance issues and state-of-the-art on supporting the emerging Bot-as-a-Service in sustainability governance platforms. Based on that we outline our approaches to policy development and enforcement for the Bot-as-a-Service model.

Keywords: bot-as-a-service, governance, policy enforcement, runtime monitoring, cloud computing

1. Introduction

Advances in networking, sensor, data management and cloud computing techniques have fostered the integration of Internet-of-Things into cloud platforms for facility management. Cloud-based sustainability governance platforms \cite{1} have been developed and provided under the cloud computing models for facility monitoring and management. Various sensors and applications for analyzing diverse types of data and for managing devices and equipment for different stakeholders have been introduced in such platforms.

1.1. Motivation

While many techniques have concentrated on sensor integration, data integration and data analytics on top of these platforms \cite{2}, we are interested in the development, deployment and execution of (intelligent and context-aware) lightweight applications that can be developed, composed and deployed on-the-fly based on context-specific situations captured from the monitoring and analysis of near-realtime sensor data. Together with techniques for connecting and monitoring devices and equipment, which are concentrated on getting information from these devices and equipment, such lightweight applications can be used to maintain and optimize these devices and equipment. They are crucial as nowadays several types of equipment and devices in large-scale facilities, such as freezers, chillers, and backup power systems, are expensive to maintain.

\textsuperscript{☆}The work mentioned in this paper is partially funded by the Pacific Control Cloud Computing Lab. We thank our colleagues in the Pacific Control Cloud Computing Lab for fruitful discussion about bots and their application domains.

\textit{Email addresses:} truong@infosys.tuwien.ac.at (Hong-Linh Truong), phung@chalmers.se (Phu H. Phung), dustdar@infosys.tuwien.ac.at (Schahram Dustdar)
and optimize. With a huge amount of monitoring data managed in sustainability governance platforms and multiple stakeholders (e.g., manufactures, maintainers, and operators) using such platforms to monitor and maintain their devices and equipment, it is possible for these applications to automatically learn the operational history of, detect errors of, and to automate repair and support of devices and equipment.

However, these lightweight applications can be developed and deployed to cloud platforms by third-parties, can be created by composing existing applications and functions offered by cloud services and then configured and deployed to facility sites on-the-fly, or may be pre-installed in hosting environments deployed at facility sites while being reconfigured for specific context. Thus the development, deployment and execution of these applications must be monitored and controlled at runtime to ensure they comply with different governance policies. In our work, these lightweight applications are called bots and they can be developed and executed in the cloud via a Bot Platform-as-a-Service (BoP), offering the Bot-as-a-Service (BaaS) business model for sustainability governance.

1.2. Related work

While many research effort have concentrated on security for monitoring sensors and data protection in the facility side, such as smart Grids [3], and in the cloud side [4, 5], very little effort has been spent on understanding issues in governing such bots in the cloud. In fact, the concept of combining cloud computing models and bots development has just been emerging: such concepts are not similar to, e.g., mobile agents and their applications for building management [6], due to the complexity and diversity of cloud data, devices, equipment, and stakeholders, although the concept of BaaS will naturally be built upon existing work. Related work on understanding governance issues either focus only on part of the BaaS model (e.g., hosting environments for mobile agents [7] and application stores [8]) or one aspect of governance issues (e.g., security [3, 9, 10], data access [11], or performance [12]). We do a comprehensive review of several governance issues in different phases of bot development, deployment and execution in clouds. Most importantly, we consider relevant governance issues in an emerging model - BaaS in which governance issues span across layers and places as well as associated with business models in the cloud.

1.3. Contributions and paper structure

In this paper, we systematically motivate the need for policy enforcement for bots in BoP. Our contributions are (i) a deep analysis of the scenarios of the bot-as-a-service architecture and its governance issues, (ii) a review of the state-of-the-art enforcement techniques in literature and identified the emerging issues that have not been solved, and (iii) approaches to the policy definition, management, and enforcement framework for the identified requirements and potential future developments.

The rest of this paper is organized as follows: Section 2 overviews BoP. Section 3 characterizes components and interactions in BoP. Section 4 discusses governance issues and state-of-the-art. Section 5 presents our approaches to the governance of bots. We conclude the paper and outline our future work in Section 6.

2. Overview of Bot Platform-as-a-Service

We are focusing on facility monitoring in smart cities, such as using the Galaxy platform (http://pacificcontrols.net/products/galaxy.html). In our work, several sensors are deployed in different buildings to monitor building Mechanical, Electrical and Plumbing (MEP) systems and surrounding environments, such as temperature and air quality. In a cloud facility management model, at each facility site, sensor data are aggregated and significant information is propagated to cloud services where online monitoring is performed. At the moment, most of monitoring tasks for devices and equipment are conducted by operators who are going to fix problems detected in facilities. In a recent emerging concept, calling intelligent and context-aware bots, bots can be deployed at the facility sites to detect problems and fix them automatically.

A bot is a lightweight application that is executed by a hosting environment. In real systems, bot can be built based on mobile agent platforms, OSGI (http://http://www.osgi.org), or plain Java objects. Bot functionality can also be described using composition languages, e.g., lightweight workflows. To support
the development and execution of bots in the cloud, a Bot Platform-as-a-Service (BoP) can be provided. Figure 1 describes a high level view of BoP, including the following conceptual components:

- **Bot-as-a-Service (BaaS):** supports the development, composition of bots, management, and deployment of bots, and the definition and management of governance policies for bots.
- **Bot Store:** stores bots and templates for building bots. Bot Stores can be hosted by the sustainability governance platform or third parties (e.g., a specific company which manages its chillers).
- **Bot Hosting Environment:** deploys, verifies and executes bots, and monitors bot execution.

Bots and BaaS interact with several other services, sensor/actuator integration gateways, devices and equipment at the cloud and facility sides. For example, when the Monitoring Service detects a possible problem, it can control BaaS to find suitable rules and algorithms to build a bot, which is deployed to the Bot Hosting Environment to analyze device’s data and control the device. An important point is that, at runtime, bots are instantiated based on existing problems and bots can be launched from bot stores in cloud service to bot hosting environments at facility sites.

The BoP we present here conceptually describes how bots, BaaS, and other components in the cloud platform and sensor integration gateways interact. From the implementation perspective, bots and their hosting environments can be implemented using different frameworks, such as mobile agents, OSGI or plain Java objects. For the sake of understanding general governance issues of BoP, we will not examine languages and frameworks for bots, but focus on governance issues associated with BoP. Furthermore, bots can be context-aware and intelligent in the way they handle data and control devices. However, we consider these features are bot’s internal and out of scope of this paper.

### 3. Characterizing Components and Interactions in Bot Platform-as-a-Service

In order to analyze possible issues in governing bots during their lifecycle – from the development to the execution – it is important to understand requirements and scenarios for which bots are required. In BoP described in Figure 1, the lifecycle of a bot has three main phases (i) **Development** – bots are compiled from source code or bots are composed from existing objects/bots, (ii) **Deployment** – bots are transferred from clouds to hosting environments for execution, and (iii) **Execution** – bots are running in hosting environments. Due to the diversity of devices and equipment, on the one hand, multiple stakeholders will develop different types of bots, specific for particular devices and equipments. On the other hand, multiple stakeholders will have different service contracts for using bots. A single BoP must be able to support such diverse stakeholders, bots, and hosting environments.
Table 1. Scenarios in the development, deployment, execution and dissolution of bots

<table>
<thead>
<tr>
<th>Phase</th>
<th>Scenarios</th>
<th>Governance support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development</td>
<td>Bots are written by third parties for specific types of devices/equipment</td>
<td>Bot integrity must be ensured. Many bots service contract terms can only be defined for specific cases, e.g., which types of devices and data analytics algorithms supported</td>
</tr>
<tr>
<td></td>
<td>and are stored either in cloud-based Bot Store or third-party Bot Store</td>
<td>Governance policies for composite bots must be compatible with bots to be composed.</td>
</tr>
<tr>
<td></td>
<td>Bots are composed, potentially from other bots, by cloud consumers using BaaS</td>
<td>Policies for bots are defined on the fly, e.g., for specific hosting environments, devices, local sensor data access and remote cloud features.</td>
</tr>
<tr>
<td>Deployment</td>
<td>Bots are deployed from Bot Stores to Bot Hosting Environments.</td>
<td>We need to verify if bots are trusted and if bots are really sent by trusted BaaS.</td>
</tr>
<tr>
<td></td>
<td>Bots are deployed to Bot Hosting Environments on-the-fly by BaaS to work with specific types of devices</td>
<td>Access to specific types of data is monitored and controlled.</td>
</tr>
<tr>
<td></td>
<td>Bots can only deal with specific devices and can utilize data and services from the cloud platform.</td>
<td>Bad performance of bots should not bring down the hosting environment. Bot’s CPU and memory consumption must be controlled.</td>
</tr>
<tr>
<td>Execution</td>
<td>Multiple bots are concurrent executed</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 describes some representative scenarios in the development, deployment, and execution of bots. In these scenarios, several governance support are required. We categorize types of governance issues that should be consider in BoP in the following:

1. **System/network security and access control**: protect systems and networks in order to prevent unauthorized access that can compromise BoP.
2. **Application integrity and service verification**: ensure that the bot content is sent by the trusted party and is unchanged during bot transfer processes, e.g., from Bot Store to Bot Hosting Environment.
3. **Service contract management**: due the pay-per-use model of cloud computing, bot capabilities are depending on a service contract. This service contract can cover different terms related to, e.g., application performance, and data acquisition and devices to be controlled.
4. **System and application performance**: ensure that the execution of bots will not prevent the correct operation and the availability of hosting environments.
5. **Data acquisition and control**: Bots will access data from local hosting environments and sensor integration gateways as well as data from the cloud platform. Furthermore, bots will be able to control building MEP (via actuators) and initiate certain features in the cloud platform (e.g., launch new bots or escalate an emergency response processes).

Several governance support require integrated solutions due to the intersection between cloud computing model, mobile code, and Internet-of-thing. For example, except service contract management, most governance issues look similar to that for mobile code. However, we see that the issues of application performance and cloud access are different, not to mention that the business service contract has a strong influence on all other governance issues.

4. Governance Issues and State-of-the Art

4.1. Current solutions for governance policy enforcement

In order to examine how existing techniques could support governance of bots in BoP, we need to examine identified governance issues in different places in BoP. Table 2 summarizes main techniques that
could be used for governing bots in BoP. Basically, a bot is built from a number of existing pieces of binary code which might come from untrusted and trusted sources and the function and runtime of bots are dependent on specific context and service contract between the consumer and the cloud provider. The main issue is that the execution of bots must be monitored and controlled at runtime in order to prevent unintended behaviors. Several techniques can be ready to support BoP, such as secured network connection, message signature, and role-based access controls. Therefore, we do not discuss them here. Instead, we will focus on main critical points that are specific to BoP.

### 4.1.1. Governing application integrity, service identity, and application security

There are several techniques that can be applied to bots, however, these techniques are suitable only certain phases of bots.

- **Development phase**: Static analysis can check bots before executing so that only those not violating are allowed to be executed. However, it cannot check runtime violations.

- **Deployment phase**: Code signing can certify the integrity of the code but cannot prevent bad behavior of the code execution. Using certification authority, signatures of bots and BaaS can be checked again with hosting environment policies, e.g., based on concepts in [18]. However, it can only allow to bots to be installed/deployed/executed or not but no other steps.

- **Execution phase**: Execution monitoring techniques, e.g., based on inlined reference monitoring [17], can enforce bot-specific security policies to prevent bad behaviour at runtime. However, they are designed specific for governing security only and are not targeted to our BaaS model. Model-Carrying Code (MCC) [14] can be adaptable to Bot Hosting Environments, however, it requires the base system to be modified. Sandboxing techniques can confine a bot instance within a Bot Hosting Environment that can only access to a limited functionality. The confinement mechanism is based on “all-or-nothing” approach which is not suitable for fine-grained monitoring. Safe interpreter can also be used to control bots, but it requires bots and Bot Hosting Environments to support interpretation languages.

### 4.1.2. Governing application performance

Governing application performance occurs mainly in execution phase. However, policies for application performance can be defined in the development or deployment phases. There exist several techniques to measure the performance of hosting environments and their processes. Many techniques, such as based on sandboxing model [12], to control CPU/memory usage by applications require external tools to interact with OS and applications but such tools are heavy. It is also possible to use virtual machine for hosting environments and control the performance of virtual machines [15]. However, we will not be able to prioritize

<table>
<thead>
<tr>
<th>Governance issues</th>
<th>Places</th>
<th>Existing techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application integrity and service authenticity</td>
<td>BaaS; Bot Store; Bot Hosting Environment</td>
<td>code signing; certification authority</td>
</tr>
<tr>
<td>Application composition based on service contract</td>
<td>BaaS</td>
<td>dynamic software product lines [13]</td>
</tr>
<tr>
<td>Service contract enforcement</td>
<td>BaaS; Bot Hosting Environment</td>
<td>static analysis; sandboxing; mediation; security by contract [9, 10]; Model-Carrying Code (MCC) [14]</td>
</tr>
<tr>
<td>System/network security and access control</td>
<td>BaaS; Bot Store; Bot Hosting Environment</td>
<td>secured connection; message signature; role-based access control</td>
</tr>
<tr>
<td>Application performance</td>
<td>Bot Hosting Environment</td>
<td>sandboxing [12]; virtual machine [15]; specific runtime systems [16]; reference monitoring [17]</td>
</tr>
<tr>
<td>Data acquisition and control</td>
<td>Bot Hosting Environment</td>
<td>inlined reference monitoring [17]; safe interpreter; Model-Carrying Code (MCC) [14]; security by contract [9, 10]</td>
</tr>
</tbody>
</table>

Table 2. Summary of main existing techniques for governing bots in BoP
bots. Specific runtime systems [16] also enable bot-specific performance monitoring but it means that Bot Hosting Environments must be based on these systems.

4.1.3. Governing service contract

Some contractual terms can be used to limit functions when a bot is developed and the other terms are used to check runtime properties, such as number of data points, types of devices and specific facilities. In particular, for on-the-fly composition and deployment, contract terms can strongly influence bot's functions. Potentially, dynamic variability techniques for supporting building limited functions, e.g., based on context [13], can be utilized. However, most dynamic composition and variability models are designed for service-oriented software systems and large-scale software product lines, not lightweight applications. Furthermore, there is a lack of techniques to generate bot software features from specific service contracts in specific situations. At runtime, while existing techniques discussed in other sections can be used to enforce certain contract terms, such as security and performance, we need to apply different techniques for enforcing a service contract for bot instances of a consumer, as the contract covers multiple types of governance policies.

4.1.4. Governing data acquisition and control

Governing data acquisition and control is mainly related to the execution phase, although it might need some support, such as policy generation and code rewriting, from the development phase. The key point is that, dependent on service contract, data acquisition and control can be governed at the system APIs for accessing data and sending control commands or at specific types of data/commands. For example, a bot might be allowed to use \texttt{read()} API for reading any sensor data type or to use \texttt{read} with only chiller data of chiller manufactured by a specific company (e.g., \texttt{read(datatype=chiller, chillerType=company\_A)}). While APIs can be protected and checked by using existing techniques, e.g., static analysis, sandboxing and inlined reference monitoring, they do not support well, e.g., how much data or which control commands a specific bot instance should be allowed. Thus, advanced techniques, e.g., application-level data access monitoring [11], could be used. However, this may require extra components to intercept bot level data acquisition and control messages.

4.2. Current solutions for governance policy definition and management

In order to govern bot executions, the unintended behaviors must be specified in policies for each bot so that its execution can be controlled at runtime. As we have multiple governance issues, we also need to consider multiple types of policies, including data acquisition and control, safety, and service contract policies. These policies are bot-specific and the policies are established in case-to-case basis depending on each particular service contract between the consumer and the cloud platform provider. However, BoP must be able to handle multiple types of policies for multiple bots from different consumers while these bots are possible executed in the same hosting environment for the same facility (e.g., a bot for chiller and a bot for electricity backup system can come from different stakeholders but are executed in the same hosting environment at the same time). Another important point is that policies are used at different places, such as BaaS and Bot Hosting Environments, at different phases, as inputs for different enforcement techniques. Although there have been a number policy languages including research prototypes such as [19, 20, 21], and industry standards such as WS-Policy (http://www.w3.org/Submission/WS-Policy/), XACML (http://labs.oracle.com/projects/xacml/), none of these considers all of the above multiple types of governance policies. In mobile application development, such as Android (http://developer.android.com), applications are associated with use permission which specifies resources and functions can be accessed from the hosting environment. However, such policies are static while BoP need dynamic policies.

4.3. Discussion

From the analysis in the previous section, on the one hand, we observed that while certain techniques can in principle be used, it is not clear if they can be engineered in BoP, in particular, BoP needs to support multiple types of governance and diverse types of hosting environments whose capabilities are limited. For example, Proof-carrying code techniques require hosting environments to inspect bots; if using safe
interpreters, hosting environments must support interpretation languages; or using encryption to transfer bots would require a heavy cryptosystem in hosting environments; or limiting CPU and memory of bots might not be implemented in all hosting environments. On the other hand, we see that we need multiple types of governance support carried out by cloud services and Bot Hosting Environments at different phases of bot’s lifecycle. As a result, BoP needs to have governance policy specifications that allows different types of governance. Furthermore, to enforce such policies, various techniques must be integrated, as typically a type of techniques is suitable only for a specific governance support at a specific phase of bot’s lifecycle.

5. Towards Multi-phased Policy Management and Enforcement for Bots

![Conceptual framework for multi-phased policy management and enforcement for bots](image)

As discussed in the previous section, different types of policies need to be enforced for bots. Our approaches are to address (i) policy definition and management and (ii) policy enforcement for multiple types of governance through different phases of bot’s lifecycle in an integrated framework. Figure 2 depicts our framework which is divided into two main blocks for policy definition and management and for policy enforcement. In these two blocks, policy definitions are evolved through different phases and in each phase, at different places, different components will use different techniques to enforce policies.

In summary, the framework works as follows. After a bot is built and stored, we already have static, common bot-specific policies – bp_s. Such policies are obtained from the development and defined for a bot. When a bot named b is selected for deployment for a particular consumer under a particular situation, BaaS can take bot-specific policies and combine with consumer’s business service contract (e.g., cost and service level agreements) to generate bot context-specific policies – bp_cxt. In cases, bots are composed and deployed on the fly, then the above-mentioned way is still valid, except that BaaS can just consider bp_s as a part of bp_cxt. Many policies p_i ∈ bp_cxt can be checked by BaaS before BaaS starts to deploy b. Before deploying b, BaaS produces two subset of policies, bp_cxt(h) and bp_cxt(r). The first set of policies bp_cxt(h) can be checked by Bot Hosting Environment before running b, while bp_cxt(r) can be used to check b at its runtime. In our framework, bp_cxt(r) is attached to b and bots will self-regulate their operations based on bp_cxt(r). For checking policies, suitable techniques will be employed at different places. Thus, our framework will provide extensible mechanisms to enable plug-ins of different techniques.

With this approach, we will provide templates to define such policies. The policy templates for bot-specific policies and context-specific policies are based on API calls provided by the hosting environment and by cloud services. Templates for bot instance policies need further investigation to combine between event sequences and their parameters, together with business rules. In our framework, an enforcement service is integrated into BaaS to enforce desired policies. A bot is first checked by a static analysis technique to ensure that it does not violate the defined static policies, and then context-specific policies. A bot passing
the check will be rewritten to embed inlined reference monitor into the code so that the execution will be controlled and monitored by the monitor at runtime to ensure that its execution will not violate the runtime policies. When deploying into the hosting environment, the authenticity and integrity of the bot code must be ensured. Our framework will support this feature by employing a code signing technique. Our proposed framework provides an end-to-end enforcement solution for bots construction and execution. While several feasible techniques are chosen, the challenging issue is how to integrate these techniques into the framework so that they can work together to enforce desired policies for the bot architecture.

6. Conclusions and Future Work

The emerging Bot-as-a-Service model for monitoring and managing devices and equipment by utilizing cloud computing offerings calls for a careful investigation on governance issues. In this paper, we analyze possible governance issues and existing techniques that can be reused and should be improved in order to support governance of bots in sustainability platforms. Our future work is to focus on the development our policy definition, management and enforcement framework that support cross governance issues for bots.

References


