IoT PaaS: Intelligent IT infrastructure for smart cities

FEI LI, MICHAEL VÖGLER, SANJIN SEHIC, SOHEIL QANBARI, HONG-LINH TRUONG, STEFAN NASTIC, SCHAHRAM DUSTDAR

Current systems in the Internet of Things (IoT) generally offer services for individual vertical domains such as building management. In systems like this, domain-specific requirements form the basis for the design of all system components and play an important role in everything from the technical selection of hardware devices to application logic.

The process that is used for this is highly dominated by providers. For this reason, subsystems provided by different interest groups are integrated via proprietary interfaces. This model typically leads to vertical systems in which hardware, networks, middleware and applications are closely linked. As a result, the scalability and the costs of IT infrastructure are among the biggest problems in systems like this.

This article describes IoT PaaS, a cloud-based IoT platform that provides environments for IoT services in which many different applications, clients and vertical solutions are supported. The platform offers a horizontal and generic middleware layer with all the necessary features that supports the service provision process of the various parties.

Two models for the provision of services are supported. The first is a virtual, vertical model that not only follows the traditional vertical process but also effectively uses cloud and middleware features. The second is an open model that completely decouples application logic from infrastructure and in this way encourages the development of new services based on existing IoT solutions and other publicly available services.

Use cases

In this section, we will first analyse two representative use cases in smart cities.

Building management

Building management is one of the most important services in intelligent buildings. A project usually begins with a survey of the target building by the solution provider. It can be either a building that already exists and has been in use or a building that is under construction. In some cases, existing buildings must of course be retrofitted with certain components such as heating, ventilation, air conditioning, thermostats, motion sensors and lighting control. These projects can range in size from one small building to a large university campus with a number of different building types. For this reason, the number of devices, the volume of data to be processed and the complexity of the analysis of the data collected can vary considerably.

After the survey and planning are finished, the provider starts to procure the necessary hardware components from manufacturers. These are then integrated into the infrastructure, and applications are developed based on this.

This process generates many vertically isolated building management systems (also known as "silos"), which leads to two types of scalability problems. The first problem is **maintenance scalability**. The more silos that are made available, the more system instances must be administered by the provider. The administration of systems in this context can become a very difficult task because hardware systems must be observed in separate systems and software systems must be adapted separately and tested on site using special hardware. Another problem is **expansion scalability**. Many building complexes are constantly growing in order to accommodate more and more users. The building management systems that are installed must therefore manage more and more buildings and facilities. For this reason, building management systems should be scalable and be able to adapt based on the prevailing circumstances. In current silo-based service models, an expansion like this can require the complete revision of the entire system or the addition of isolated building management systems due to the close coupling of devices, IT infrastructure, middleware and applications. In addition, the energy consumption data that are collected from individual building management systems cannot be used in full due to the heterogeneity of the individual solutions. These days, the painstaking process of data cleaning and integration is an important prerequisite for a fine-grained energy consumption analysis on a large scale.

Building management systems are an excellent example of a vertical model for the provision of services: A solution provider installs and integrates devices, offers IT infrastructure and develops applications. The provider is generally responsible for the procurement and maintenance of the entire system during its life cycle. This model is frequently used in areas in which highly specialised device solutions and a high level of expertise are required such as intelligent transport systems.

Event planning

Public events represent an important part of urban life. Some of these events take place regularly, such as annual city marathons or national parades, and others may occur spontaneously, such as demonstrations. An event, whether big or small, has certain effects on city residents. Participants or visitors would like to know, for example, how they can get to and away from a certain location. Others who are not interested would like to know how they can avoid events. All of these concerns and many more must be taken into consideration by the event organiser in the planning stage.

A typical application for event planning comprises data collection and the creation of organisational and emergency plans. The process generally occurs in four phases: planning, preparation, operation and follow-up work. A variety of information is needed for an efficient application, including persistent information such as maps, public transport routes, accessibility of facilities and real-time information such as weather, capacity utilisation of public transport and parking spaces. Although a lot of this information is already available through public services, these services are usually isolated and highly domain-specific. Every service has its own infrastructure, and the data are processed independently and provided via various channels.

The main challenge in the development of such applications is dealing with the unique aspects of every event in terms of the scale and the necessary information. Developing a service for an event takes a lot of time and money. And it may be that the service will only be used once or will have to be updated periodically for regular events. In addition, the resources that are used to provide such services are only required during the events. So the costs and efforts that go into developing a service like this are only justified if development goes fast enough and the necessary IT resources are only provided as necessary.

Public event planning and organisation is a prime example of the provision of services by third parties: A provider receives access to existing IoT services and other information sources in order to develop applications for certain purposes. The information sources vary immensely and range from public to commercial services. The provider's focus is on the development of the application logic because the provider does not own the information sources and in most cases does not have direct control over them. However, the provider must deliver and manage IT resources in order to ensure the quality of the service.

IoT PaaS

In this section, we would like to present the IoT PaaS cloud, which was developed with the aim of delivering efficient and effective IoT services. See Figure 1 (next page). The infrastructure in the Internet of Things comprises networked sensors, actuators and other intelligent devices. Gateways are a reliable method for connecting heterogeneous and resource-dependent devices with each other. Generally, gateways provide device drivers and special protocols. If several protocols need to be used, there is usually a standardised interface for applications that require access to devices. Intelligent devices such as mobile phones are becoming more and more powerful and are now able to communicate with these applications. For this reason, many gateways and devices now offer web-based interfaces (for example, oBix and OGC) to facilitate integration with applications. We refer to this part as *IoT infrastructure virtualisation* because it is where the first conversion between network/hardware and software interfaces actually occurs.

IoT resource management offers a point for managing virtualised devices, gateways and field solutions. The provider decides on the granularity of the virtualised resources that are offered. The component monitors the status of the resources and the enforcement of the access policies. Although most existing gateway solutions try to reduce the heterogeneity of the hardware levels to a certain degree, the variety of gateway products and their areas of application have actually led to another layer of heterogeneity. Generally, the interfaces of the various gateways in the industrial sector have to be connected by domain mediators. Since these gateways are driven by specific requirements, the mediators are an important component of the software.

Data services are an essential part of the PaaS cloud. The service effectively hides the underlying database system. However, IoT applications rely heavily on the production and consumption of real-time data, which are not sufficiently integrated into existing data services. The IoT PaaS cloud therefore offers *real-time data service* in order to minimise problems that arise in the use of heterogeneous real-time data processing systems. The service supports the processing and analysis of events that are sent by devices from lower levels. Applications are capable of specifying how events should be handled and which events or patterns are of interest. The real-time data service also ensures that the use of events complies with the access policies of the event provider.

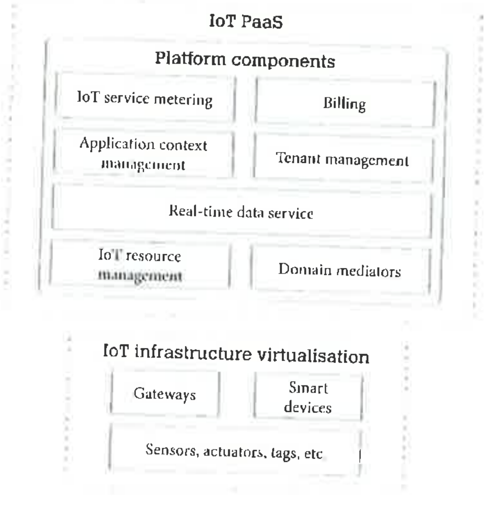

Figure 1: IoT PaaS platform

Tenant management offers a consolidated view in order to divide virtualised IoT resources between application providers. In contrast to IoT resource management, which governs access to resources for IoT solutions, tenant management is designed for platform and application providers to see all of the resources that can be accessed by the tenant. Because device functions and IoT solutions are made available for several tenants through virtualisation, two different methods for the shared use of resources must be examined. Firstly, the shared use of information in real time. In principle, this method can be compared with read access to a database. However, this kind of use becomes difficult when providers want to share information of different granularity and tenants require different data quality levels . This should be handled by the data service component. The second method is shared/multiple access. Because access in IoT systems generally results in a change in status attributes in the real world, there may be problems with the policies for shared access with the aims of the services or the limits of the IoT solutions. It is therefore important for the platform to provide a generic framework for related interest groups in order to work on strategies and solutions to conflicts during run time.

In the Internet of Things and the cloud, every application runs in a complex and dynamic environment and involves the available IoT and cloud resources, the software configurations and the fluctuating requirements of users. *Application context management* focuses on the optimal use of run time resources and software configurations for applications. Based on the resources that were acquired through tenant management, this administration helps applications to select the necessary resources during run time in order to comply with the required SLAs and costs. In this way, components can react to constantly changing requirements by quickly acquiring and releasing IoT and cloud resources. Together, tenant and application context management give all applications and their providers a virtual dedicated working environment.

Quantitative measurements of IoT services help providers to understand how services are used and, based on this, how services can be optimised. In addition, they offer important information for those involved in the process of providing services when it comes to deciding how revenues based on the actual use of services should be managed. *IoT service metering* measures the use of various services by an application. This is mainly performed by observing sent and received messages and how many times a service is called up. There are three types of measurements for this:

- time-based, with the use of a service measured based on the start and end of a service instance,
- call-based, which monitors the number of specific service call ups, and
- volume-based, which measures the volume of real-time data or certain events that are collected by a service.

At the end of the life cycle of the service provision, the *billing* component then links the various business models to the run time. The billing component creates invoices by analysing the measured data in line with the policies on charging configured by the parties. A flexible billing mechanism encourages service providers with different capacities to take part in the process of providing services.

Service provision in IoT PaaS

The emergence of cloud-based IoT service platforms generates various types of interest groups that are involved in the process of providing services.

- Platform providers are a new kind of PaaS provider. They offer cloud services that integrate and manage IoT solutions, plan both cloud and IoT resources, and support various models for the provision of IoT services. The main platform and application components were described in the previous section. The details and components of the IoT infrastructures and cloud resources are hidden from application providers. In addition, platform providers are responsible for providing all of the IT resources for everyone involved.
- Solution providers developed out of the traditional IoT service providers, which generally specialise in one application domain and offer vertical IoT solutions. They are familiar with the selection of field devices and understand domain-specific use cases. IoT PaaS allows providers to offer solutions as services in the cloud so that these can be used by various application developers.
- Application providers use IoT solutions with the help of the cloud platform and create applications for end users. Domain-specific knowledge is not required because IoT solutions are virtualised and the most important components are provided by the platform. As a result, application providers can focus on the application logic and use cloud and IoT resources as necessary.

The collaboration between these parties is the key to more flexible models. We would like to briefly introduce two of these models.

The model of virtual verticals

This model corresponds to traditional vertical solution development: Solution providers develop solutions for field devices and also application logic. The main difference is that virtual verticals use the IT resources and other necessary software on a cloud platform. Applications are linked to devices via special mediators developed by the solution provider. In the PaaS paradigm, isolated system environments for virtual verticals are ensured by multitenant mechanisms. As a result, virtual verticals are easy to scale because, as the number of device solutions increases, new computing resources can simply be added. In addition, the elasticity of the cloud can be used to deal with the fluctuation in user requirements. This model can be used for many existing vertical IoT solutions. The scenario of building management systems is a typical area where this model can be applied. The cloud platform offers building management for virtual vertical environments.

Applications by third-party providers

In traditional vertical IoT solutions, the development of applications by third-party providers is not a common practice. One reason for this is the closely linked system architecture on the one hand and the lack of flexibility in the provision of resources on the other. The development of cross-domain applications is even more complicated. The PaaS paradigm offers several important components that make it easier to open up IoT solutions to applications by third-party providers. Applications by third-party providers can simply use services offered by the platform such as data services. The usual problems, including the maintenance of devices and data collection, are decoupled from application development. For this reason, the application provider can focus on the development of the business logic and deliver a lot of new services in a relatively short amount of time. Like with the model of virtual verticals, applications can have their own application context with the necessary solutions, software services and other resources. In addition, flexible use models and billing systems can be configured for IoT solutions and applications. As a result, costs and revenues can be assigned to applications and monitored in real time. The platform therefore provides an environment that third-party providers can use to develop new applications. The organisation of public events is a typical application for this model in which devices and domain-specific resources can be accessed by the public. Application developers can access this public data via the platform and use the specific information they need. With this model, application providers no longer have to worry about the provision of IT resources. Another important aspect of this use case is that the PaaS paradigm offers specialised IoT solution providers an environment where they can serve multiple users through virtualisation.

Conclusion

The platform described here is the result of our ongoing research at the Pacific Controls Cloud Computing Lab (PC3L). The platform represents the core of our long-term goal of encouraging an ecosystem that allows for and simplifies the open creation of services in smart cities. The platform is meant to encourage specialised small and medium-sized enterprises to participate in developments in smart cities. We will examine further specific challenges in the corresponding fields of cloud computing, the Internet of Things and software architectures in order to continuously incorporate the research results into the platform.

Sources

I Cook, J., Smith, D. and Meier, A.: "Coordinating Fault Detection, Alarm Management, and Energy Efficiency in a Large Corporate Campus", in *2012 ACEEE Summer Study on Energy Efficiency in Buildings*, 2012, pp. 83–93.

II Drobi, S.: "Play2: A New Era of Web Application Development", *IEEE Internet Computing*, Vol. 16, No. 4, pp. 89–94, July 2012. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6257378

III Sehic, S., Li, F., Nastic, S. and Dustdar, S.: "A Programming Model for Context-Aware Applications in Large-Scale Pervasive Systems", in *The 8th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2012)*, 8–10 October 2012, Barcelona, Spain.

IV Li, F., Nastic, S. and Dustdar, S.: "Data Quality Observation in Pervasive Environments", in *The 10th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC 2012)*, Paphos, Cyprus, 2012.

V Batini, C. and Scannapieco, M.: *Data Quality Concepts, Methodologies and Techniques*, Springer, Berlin, Heidelberg, New York, 2006.

Endnotes

1 https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=obix

2 http://www.opengeospatial.org/projects/groups/sweiotswg

3 http://pc3l.infosys.tuwien.ac.at

AUTHORS

Fei Li, Michael Vögler, Sanjin Sehic, Soheil Qanbari, Hong-Linh Truong and Stefan Nastic are part of the Distributed Systems Group led by Schahram Dustdar at the Institute of Information Systems at the Vienna University of Technology. The group conducts research and teaches about distributed systems with the following emphases: service-oriented computing; autonomic, complex and context-aware computing; cloud computing; and mobile and ubiquitous computing.