
Smart Cloud of Things: An Evolved IoT Platform for Telco Providers

Mário Antunes, João Paulo Barraca*, Diogo Gomes,
Paulo Oliveira and Rui L. Aguiar

Instituto de Telecomunicações, Aveiro, Portugal
DETI, Universidade de Aveiro, Aveiro, Portugal
**Corresponding Author: jpbarraca@ua.pt*

Received 9 September 2015; Accepted 30 September 2015;
Publication 20 November 2015

Abstract

Although many environments are powered by smart solutions, users do not have a simple way to gather their collective knowledge and program devices' behaviour. Telecommunication (Telco) providers wish to act as the facilitator, but still lack proper components for enabling integrated services over their networks, even though this work is already in active standardisation. In this work we present the architecture of a IoT solution developed in close collaboration with Telco providers, its evolution, which we call SCoT (Smart Cloud of Things), and the best practices from the several pilots we have been running. The architecture builds upon ETSI standards, combined with a rich service execution environment, providing facilitated orchestration of services and devices, as well as integrated portals for end users access.

Keywords: IoT, IoS, Service Orchestration, Telco, M2M.

1 Introduction

When we think about the Internet we mostly consider servers, routers and fixed broadband providers. Internet grows by the addition of these devices, as well as mobile phones, tablets and laptops. These devices, and in particular, the

ones that are added for automation and sensing processes are named “things” and their number is bound to exponentially increase in the years to come.

This presents huge opportunities to the development of both smart environments and increasingly efficient business processes, as control and knowledge of our processes increases. According to the ICT Knowledge Transfer Network, the number of mobile devices is expected to increase worldwide from 4.5 billion in 2011 to 50 billions by 2020 [45].

With so many sensors collecting data about equipment status, environmental conditions and human activities the industry is growing rich with data. One can argue that instead of getting *rich*, it is being *buried* by the amount of data.

Several authors have previously demonstrated the potential of these untapped sources of data, as well as the wide field of possibilities for smart devices. Fusing information from several sensors makes it possible to predict a driver’s ideal parking spot [30, 38]. Projects such as Pothole Patrol [17] and Nericell [25] use vehicular accelerations to monitor road conditions and detect potholes. The TIME (Transport Information Monitoring Environment) project [7] combines data from mobile and fixed sensors in order to evaluate road congestion in real time. LOFAR_Argo [22, 41] was the first large scale experiment in precise agriculture at Netherlands. They were able to predict the presence of *phytophthora* in the crops by fusing information from 150 sensor nodes, each node measures the soil temperature and humidity. A curious aspect of this project is that LOFAR stands for Low Frequency ARay, and in reality is a radio telescope, showing how once sensors are deployed, the data they produce can be used for many scenarios. Also, some precision agriculture projects combine wireless sensor networks with Unmanned Aerial Vehicles (UAV) to monitor the crops [9, 44].

Gathering massive amounts of data is not enough, the next step is taking advantage of the untapped source of information. One area that greatly benefits from this collective knowledge is the area of smart environments. A smart environment means many things to numerous people. Yet, one thing remains constant: the part of being “smart”. In this context being “smart” implies self-adaptation based on the current context. The collective information, gathered from “things”, provides relevant knowledge about the environment. Take into account several sensors (air and leaf humidity, air and soil temperature among other sensors) scattered throughout a greenhouse. Currently several platforms are able to automatically water the plants, based only in soil humidity. However by combining the data from all the sensors, and crop information, it’s possible to detect the plants’ growth, if the plants are being attacked by any infestation or suffering any disease. Also, it becomes possible to correlate weather or other

forecasts in order to increase process efficiency. Other areas can be found in every production, distribution, or industrial process where metrics need to be observed (e.g., temperature of refrigerated vehicles), or actions need to be taken dynamically.

The ETSI M2M standard [18, 19] is now mature enough to provide solutions for such massive sensing and acting scenarios, and is now being supplemented by the worldwide OneM2M initiative [28], believed to define the best guidelines for integration of IoT devices. Both provide components and interfaces for low-level communication and management of devices, as well as integration with Telecommunication providers' infrastructures. However, these initiatives do not address the processing and visualisation needs of IoT. As a complement to these approaches, we propose a novel Machine-to-Machine (M2M) platform that merges the Internet of Things (IoT) [39] with the Internet of Services (IoS) [34], providing the necessary components to create a useful platform for service creation in IoT and M2M scenarios.

The platform [4] combines the ETSI M2M low level communication and management components, with a higher-level data manipulation layer that follows a Service Oriented Architecture (SOA). It provides several services to analyse, process and manipulate sensor data, routing data to multiple tenants, and advanced machine learning processes. Furthermore, it gives users a Graphical User Interface (GUI) to design process workflows that can be instantiated by the platform. As validation, after being tested in two different scenarios, our base solution is now being deployed as the reference M2M platform for a major European Telecommunication provider. Our enhanced solution, the Smart Cloud of Things (SCoT) platform continues its innovation by providing enhanced data mining and storage capabilities, which are vital for many IoT/IoS scenarios.

The remainder of the paper is organised as follows. The most relevant M2M platforms are discussed in Section 2. In Section 3 we present the conceptual architecture of the SCoT architecture, and its different domains. The evaluation use cases and performance evaluation are described in Section 4 and Section 5 respectively. Finally, the future scenarios and conclusions are presented in Section 6 and Section 7 respectively.

2 Related Work

M2M applications and services have reached the market through various vertical solutions that have hindered interoperability and the realization of a true smart-environment. We therefore focus on state of the art M2M platforms that acknowledge the need for such interoperability.

Metrosense [14] is a generic platform for urban sensing, focused on a specific scenario: people-centric applications. It provides the ability to sense data related to interactions between people and their surroundings. This project was specially designed for urban sensing and was not intended to be a general proposed communication infrastructure.

Sensor Andrew [32] aims to create a ubiquitous, multi purpose, sensing network. The primary goal of the project is to have a living laboratory where applications can be rapidly prototyped. Furthermore, the platform enforces the re-utilisation of technology and components. However, this project falls short of modern approaches as it lacks methods to do preliminary processing of sensory information coming from the gateways, and standardised, scalable methods of storing, representing, and analysing the information collected.

SensorAct [5] is an open-source middleware with objectives similar to Sensor Andrew. It supports applications that perform operations on sensors and actuators, e.g. query current and historical sensor data or trigger customised notifications. Nevertheless, it fails to provide a stronger Applications Layer, therefore not suitable for doing processing tasks and not allowing external entities to interact with the sensor and actuator data.

Several projects have tried to integrate web services with M2M platforms and provide a richer integration. SenseCampus [43] aims to connect places and entities in the physical world with objects in the cyber-space. Through that coupling, it creates a ubiquitous service that not only promotes the distribution of information among the various users, but also gives support to activities that take place in the environment.

BuildingDepot [1, 46] provides an extensible and distributed system, enabling storage, access control and management of sensory information. It's architecture focus on three components: data, directory and user service. Data Service (*DataS*) stores information generated by sensors together with meta-data that identifies the context and the appropriate sensor. Directory Service (*DirS*) links institutions with *DataS*, and stores meta-data from child *DataS* to allow searching. User Service (*UserS*) stores information on users and groups from a particular institution.

Both solutions provide access to sensory data through web services, but these platforms do not provide any flexible mechanism to process, analysed or transform sensory data. Both projects communicate with the sensors through a non-standard protocol, which makes it difficult to add third party sensors. Also, they lack a proper management and operation platform with the capability to integrate and audit the operation of devices.

The industry is also very active, with several startups in this area. Xively [23] is an IoT solution based on Web technologies, it provides a management solution for device provisioning and for the creation of communication pipes. ThingSpeak [42] is an open source data platform for the IoT that allow users to record sensor data and plot them in real time. Open.Sen.se [35] is a platform that enables sensory information collection, processing and actuation. The platform enables users to integrate existing applications based on flows of information. Nimbits [26] and BeeBotte [10] are Platform as a Service (PaaS) that allows users to develop their own services and deploy them into the cloud. These solutions provide specialised back-ends to disseminate sensory information, but only provide basic data manipulation services. These solutions lack integration with pre-existing management infrastructures.

OpenMTC [29] and Eclipse IoT [16] are middlewares optimized for new types of M2M communication. These projects aim to provide a middleware optimized for IoT/M2M scenarios. However, they only address communication and device management, letting behind data storing and manipulation, service creation and orchestration.

Telecommunication operators are interested in providing M2M platforms to their clients. However, most solutions are focused in the provisioning of SIM cards, and of basic managed connectivity services. While useful, they do not present an environment that fully leverage the existing communication and management infrastructure, in particular the OSS and IMS subsystems.

3 The SCoT Platform

The SCoT platform is an evolution over our previous work, APOLLO [4], aiming at the development of a generic platform for integration of IoT/IoS scenarios. Like its predecessor, it covers aspects related to network, device management, services and applications overcoming the shortcomings of the solutions previously identified, and presenting novel data mining concepts. An important aspect is that we considered the entire M2M ecosystem, and its stakeholders.

We assume that an existing Telecommunication provider infrastructure, and their Operation Support Systems (OSS) is present and capable of being fully integrated, through the standardised interfaces. Such integration is desirable as OSS provide many of the desirable functionality and enables full integration of our platform into an existing environment. SCoT aims to allow multiple tenants to deploy their services with agility and reduced time to market, over a wide range of scenarios and using different sensors.

The platform abides to ETSI M2M and can be divided in four major domains: Sensor, Network, Service, and Data (see Figure 1). These domains are closely related with IoT/IoS, enabling the Telecommunication operator to act as the vital glue holding both concepts together, and presenting an offer with added value to its clients.

In the following subsections the several domains are discussed with greater detail.

3.1 Sensor Domain

The Sensor Domain (SD) is composed by sensors, actuators, and gateways that enable integration of physical environments into the management platform. These devices can range from micro-controllers used in low power sensing scenarios, to appliances, cell phones, and other devices with M2M capabilities.

This domain is responsible for enabling smart devices to communicate with the remaining M2M network, abstracting the communication with sensors, and managing the communication facilities at each M2M enabled site through its gateway, or directly to each smart device. Particular scenarios may use different (non ETSI aligned) devices, and we observe that support for these sensors is vital for current deployments as it greatly reduces cost, complexity, and the required power budget. Effectively increasing the adoption of M2M solutions. Still, in both approaches (smart vs dumb), sensed values are reported to the upper layers following strict rules, and using lightweight protocols. In this sense, we considered that sensors shall primarily communicate through

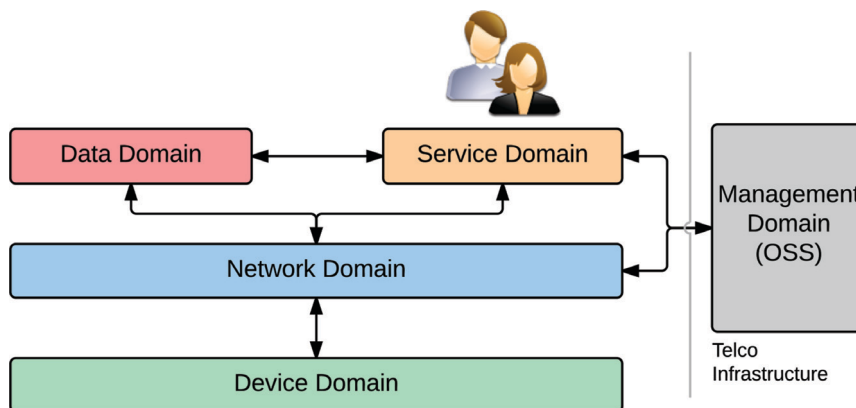


Figure 1 Architecture of the SCoT platform.

standard M2M methods, such as CoAP [36], MQTT [8], or even SCADA [12]. The platform also considers the integration of legacy communication solutions, ranging from proprietary RF protocols (e.g. Shockburst, MIWI), to higher level application protocols such as XMPP [33]. Still, the actual payload of each sensor in SCoT was standardised as being JSON [13], with gateways having the role of adapting content from sensors that do not support this content type.

SCoT takes advantage of the ETSI M2M/One2M specifications to support seamless integration between heterogeneous sensors and the services present in the upper domains, while supporting integration with a Telco OSS platform. In this sense, strong authentication, authorization and privacy mechanisms are mandated so that all communications are secured. The SD (see Figure 2), composed by devices, is mostly organised around Service Capability Layers (SCLs). Each SCL enables a smart sensor or gateway device, fully supporting the management capabilities of the SCoT framework, and allowing further deployment of specific applications. SCLs do not actually perform any measurement, only enable applications and provide management, authorization and routing functions. Applications are deployed into SCLs and are the ones that sense, process and receive M2M data.

3.2 Network Domain

The Network Domain (ND) consists of the device and network management components, hosted by a Telecommunication operator platform. Under the ETSI approach, this reflects the Network Service Capability Layer (NSCL), which in our case is integrated with the existing Operation Support Systems (OSS). The main function of these components is to serve as aggregation points for devices to connect and disseminate information. A relevant aspect is that tenant information must be mapped from the higher layers into the ND. This effectively enables the NSCL to enforce unified access control and

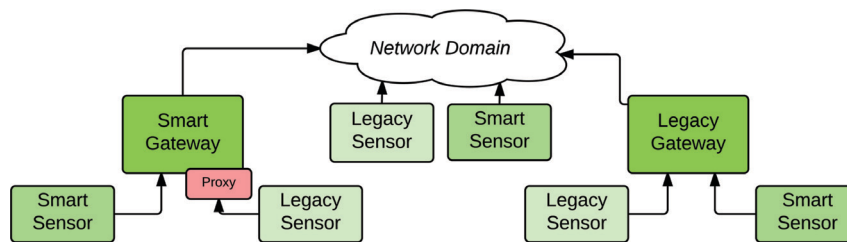


Figure 2 Components of the sensor domain.

accounting, as well as auditing and extended debugging, due to the integration with the Telecommunication operator OSS (see Figure 3).

Components can be shared amongst Telecommunication operators/providers as well as by clients with legacy M2M infrastructures. In this case, the NSCL considers the existence of external brokers, which can be directly integrated, or communicate through Network Interworking Proxies (NIP) as defined in ETSI M2M.

As part of an enhanced M2M platform, components of the network domain are also responsible for the management and auditing of devices, providing programmatic interfaces that facilitate device provisioning and debugging. M2M device management is vital as it provides the means for integration of devices, which are heterogeneous by nature and can belong to a multitude of tenants. OMA-DM [27] is our base line for Operations Administration and Management (OAM) support, and can map into each device accordingly to the inherent individual characteristics.

3.3 Service Domain

At the Service Domain there is little notion of the device characteristics, and only data objects are exchanged between service endpoints. The components in this domain connect to the Telecommunication operator OSS and to the NSCL component of the Network Domain, and exchange service information in order to compose a rich SOA environment. SOA allows components to be modelled as independent services, creating a loosely-coupled environment. The platform allows that multiple tenants could make use of the M2M platform, while keeping low latency and tenant isolation. Figure 4 depicts the general architecture of the Service Domain and its most relevant components.

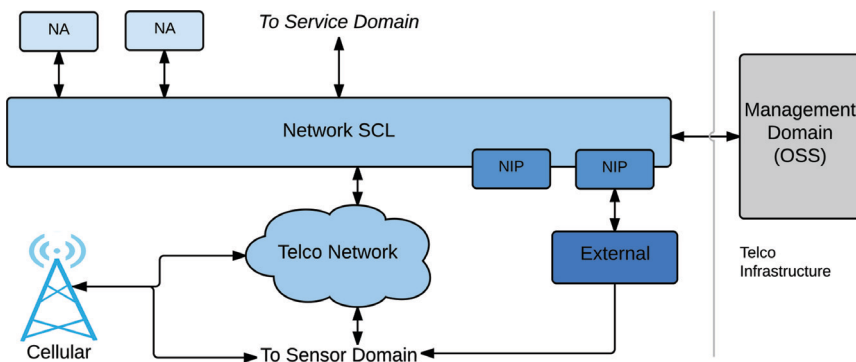


Figure 3 Components of the network domain.

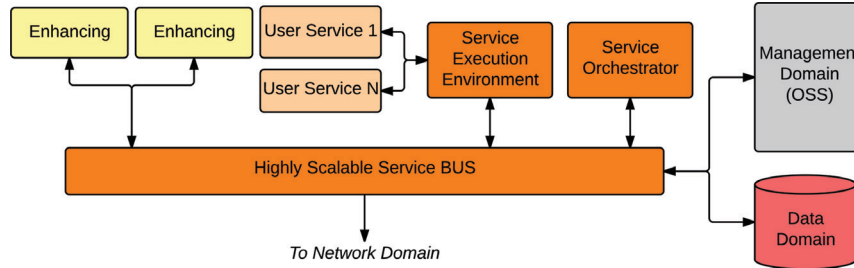


Figure 4 Components of the service domain.

This domain is based on the concept of a Highly Scalable Service Bus (HSSB). An internal component of the bus acts as a Network Application (NA) and registers the currently active topics with the NSCL. Therefore, all information relevant for services and users that reaches the NSCL is injected into the service bus as documents.

Due to scalability reasons, we consider the service bus actually to be composed by several instances, subscribed to different groups of devices, and with some level of routing between them. From our perspective, as we also consider the existence of multiple NSCLs, the platform can easily be scaled horizontally by adding more instances that deal with a subset of the topics published by sensors. Each HSSB contains multiple Enhancing Services (EN), providing additional functionality over the documents that are published to the HSSB. As an example, an EN can take the temperature, humidity and wind from a Weather Station and enrich the document with the indication that there is a risk of frostbite to plants. Some other ENs can provide richer documents to authorised services on demand, and effectively play the role of development accelerators and product enhancers, created by Telecommunication providers, to facilitate service development and deployment to their clients.

3.3.1 Service orchestration

From Section 2 we observe that most M2M platforms are based on vertical solutions. These solutions were designed to meet a very specific need, which for efficiency purposes have been deployed with a hard-coded implementation. Others, implemented as web platforms are too generic to be useful in specialized scenarios.

As a counter part to these approaches, the SCoT platform allows tenants to develop, deploy and orchestrate services (User Services) on the platform, with clear benefits of being closer to the data (lower latency). A simple workflow is depicted in Figure 5. Tenants may deploy two kinds of services: developed

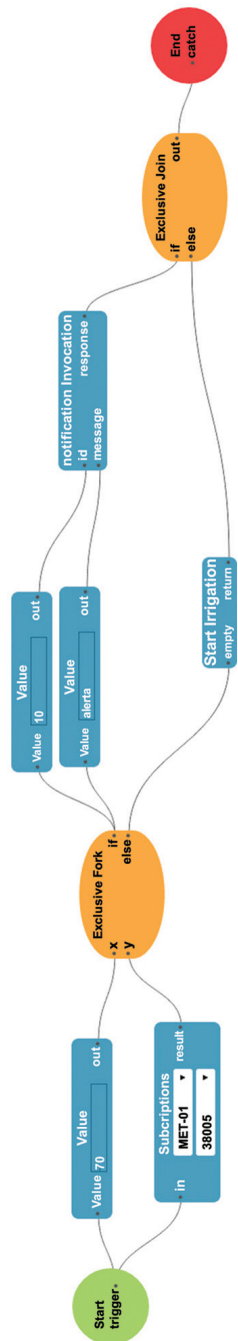


Figure 5 Flow based workflow design tool.

on their own following basic web services guidelines and API, or orchestrated through the supplied graphical user interface. The graphic orchestration tool allows users to design services/ processes (in BPMN format [15]), without any knowledge of the software development process. This line of work is similar to others, followed by entities such as IBM [40], although we focused in a generalized support for process description. Moreover, the SCoT orchestration facility is not tied to specific messaging communications. The clear advantage is that higher level processes and even business workflows can be directly mapped into a service composition instantiation.

When the process is submitted, it is converted to BPEL documents, that can be deployed immediately or used in future orchestrations as sub-components. Both service kinds are deployed in the Service Execution Environment (SEE) and made available to all other services through Web Services. BPMN and BPEL are standards, but are nonetheless complex notations and languages, which greatly benefit from flow based design interfaces.

3.4 Data Domain

As the number of connected devices grows, it becomes increasingly difficult to store and share all these new sources of information. Several representation schemes have been proposed, however none of them have been widely adopted [2]. Usually each platform defines a representation scheme that suits their specific needs. This hinders compatibility between platforms and limits the quantity of context information that can be shared/used in M2M scenarios.

According to a 2011 IDC study [20], unstructured data will account for 90% of all data created in the next decade. As a new, relatively untapped source of insight, unstructured data can reveal important relations/patterns that were previously difficult or impossible to determine. The vast majority of M2M platforms rely on structured data, however, the tools and techniques that have proved so successful transforming structured data into business intelligence and actionable information simply do not work when it comes to unstructured data [11]. Extracting knowledge from unstructured data has been active an research area [24, 31], but the majority of these techniques were developed to find structure in large corpus of documents, emails or web pages. Also, the characterization of M2M data sources can vary along the time and location (most of the data sources are mobile).

In M2M scenarios, a data source is an entity that produces a possibly infinite stream of multi-dimensional, potentially correlated, data. Generalizing the storing process is relatively easy: several databases can store binary blobs

or textual fields. The main challenge is classifying data in a way that provides discriminative retrieval, and cross stream context enrichment, and does not force a specific representation at storage time.

Some authors [6, 21, 37] point out that top-down classifications induce a bias into their model of the world. According to the authors, the signal loss brought about by the unification process of top-down classification is enormous. Top-down classifications limit the dimension along which one can make distinctions, and local choices at the leaves are constrained by global categorizations in the branches. It is therefore inherently difficult to put things in their hierarchical places, and the categories are often forced. The same authors explain that probabilistic models on top of bottom-up (user-centred) characterization produce better results than binary schemes built on top of top-down classification. Moreover, bottom-up characterization is massively dimensional, and there is no global consistency imposed by current practice.

According to these authors, the best solution to classify context information is through bottom-up characterization. Although sensor information is not manually tagged by users, we can model bottom-up characterization as an information retrieval problem. Organizing documents based on its content is one of the major objectives of information retrieval research: information retrieval informs on the existence (or non-existence) and whereabouts of documents related with user's query (similar to a web search engine). There are several methods that provide discriminative retrievals such as relational models, semantic web, ontologies/taxonomies among others. However, these methods either require knowledge about the context structure (relational model) or manually defined relations amongst entities (semantic web, ontologies/taxonomies).

To overcome these issues we developed a context storage solution [2, 3] optimized for M2M scenarios, that is agnostic to the representation scheme and provides advanced search capabilities. Our solution combines a NoSQL database with an information retrieval system, specially optimized for sensor data, exploiting the flexibility of dynamic context definition through bottom-up-characterization.

4 Evaluation Use Cases

The platform was instantiated into multiple scenarios for testing, of which we highlight three: Precision Agriculture focusing in low latency sensing and actuation; Road condition assessment focusing in massive number of events

in a Smart City; and Public Lighting focusing in LED based lighting with integrated sensors.

4.1 Smart Agriculture

In the Smart Agriculture scenario we equipped a local agriculture school (ESAC in Coimbra) greenhouse with smart sensors and actuators. Sensors were based on low power μ C (ATMEGA1281), battery/solar powered, capable of monitoring parameters from soil, water, air and solar radiation. Sensor operation relied on a variable duty cycle, adapted to the energy left in their Li-Ion batteries. This was required in order to maintain the network operational in days with reduced solar intensity. Communication between sensors and the gateway used ZigBee radios with mesh capabilities, and the CoAP protocol. The GSCL component reported information through a 3G/GPRS network. Several Gateway Applications closely interacted with the sensors creating richer information, or enabling low latency direct actuation. In our case, farmers were interested in detecting leaks and avoiding frostbite. Moreover, the flow based service creation interface allowed the definition and analysis of workflows controlling several aspects of the greenhouses. Each sensor produced reports every few seconds, generating a huge amount of documents, all handled in real time as actuation could be required.

More recently, this prototype is being further deployed in a distant agriculture area, covering 10.000 acres. The focus is to integrate concepts from Delay Tolerant Networks (DTN), through the use of data mules consisting of unmanned aerial vehicles. Due to its flexibility, the platform is being plugged with components able to programmatically instruct UAVs to cover the fields, collecting data from the sensors scattered over the entire area.

4.2 Road Surface Monitoring and Pavement Analysis

In a second scenario we targeted road condition assessment through pothole detection, recurring to crowd sourcing, massive data collection, using off-the-shelf mobile devices and machine learning techniques. An Android App was created and made available to citizens who would place their monitoring phones in the dashboard of their cars.

Each monitoring phone would monitor the location, speed, and 3 axis acceleration with a frequency of 15 Hz. The system assesses the road surface condition of several vehicles (use case similar to [17]). Sensors report information every 5 hours using their 3G connection, or immediately if a Wifi

connection was available. Data flows to an intermediate gateway, and then is dispatched to the components at the network layer. Finally, information is stored in several databases for the purpose of benchmarking, analysis and context enrichment.

The documents generated by the vehicles are filtered in order to detect high peaks in the Z (vertical) axis. After we leveraged our cluster based storage for detecting anomalies based on high Z peaks events, and a machine learning approach for determining anomalies based on a reference road segment. As a result we obtained 82% in determining potholes under realistic conditions. We had no control over the vehicle, driving style, vehicle condition, or cell phone location. We processed tens of million reports per month, which enabled us to build a detailed map covering the entire Aveiro municipal region, and even part of the centre region of Portugal.

4.3 Smart Public Lighting

In the third scenario we aimed at instrumenting the public lighting infrastructure of a ~ 300 m bridge, frequently used by pedestrians to reach the south part of our campus. Each group of luminaries was equipped with a smart object incorporating a 3 axis accelerometer, temperature probe, current probe, Passive Infra Red sensor, and a microwave proximity sensor. The sensors were used to detect pedestrians and dynamically adjust the light intensity for some seconds, reverting to 10% intensity when no one was present. Luminaries communicate through an Enhanced ShockBurst network, using multihop capabilities to the platform. The Gateway Applications are responsible for adapting content, creating aggregated metrics (e.g., total consumption), and even applying calibration algorithms. This approach allows for the deployment of low cost luminaries. The gateways, having higher flexibility and processing capabilities, create the appropriate information documents. Due to high sampling frequency and number of luminaries, this single system has generated hundreds of millions of samples over the course of 11 months.

5 Performance Evaluation

The platform described in this paper is still under heavy development, with new management and connectivity components in the near future, nevertheless we devised an evaluation scenario with key performance indicators. In our scenario we deployed a service, running in the service domain which listened to data from its sensors. We then created five data producers, simulating sensors,

each using a different communication method: CoAP, ETSI M2M, HTTP REST, WebSocket and MQTT. This environment will allow us to determine the best latency in our platform and characterize the inherent latency of each communication protocol.

Sensors sent 1100 messages, with payload sizes ranging from 2 to 2048, duplicating the payload size after each 100 messages. It is important to mention that this size corresponds to the size of useful (e.g., sensed) data, and each packet will have more information. In this case it was added nearly 60 bytes of overhead information used to timestamp the events. Information would flow through to the correct end-point, and after authentication it would be validated, enriched, re-routed to the correct tenant context, stored in the databases for logging and auditing purposes, and then delivered to the service for processing. During its course, and especially after the enrichment process, messages will grow with meta-data in a variable manner according to the characteristics of the sensor, its tenant and the data being measured. It should be noticed that this scenario exercises all components of our data plane, and also includes authentication and auditing components, therefore reflecting the performance of a typical application deployed over our system.

We establish that the underlying network imposes an average minimum latency of 0.952 ms between sensors and the components of the network domain, and 0.524 ms between the service exposure components and the services consuming data. Therefore, the total network latency is of 1.476 ms. If we consider sensors reporting information through a different communication technology (e.g., GPRS) this value will grow accordingly.

For the purpose of analyzing system performance, and because the platform is under a production environment, we discarded 25% of the population in order to exclude outliers. From the remaining values we calculated the average and the associated deviation, which is depicted in Figure 6. These values include the network latency previously determined.

It is clear that, solely from the perspective of latency, the selection of a particular M2M protocol is of great importance. Both in terms of minimum latency and of variability. Considering the end-to-end latency, when using the multiple protocols CoAP provides the lowest latency of 6.4 ms, followed by MQTT with 6.8 ms. If we discard the network latency, each protocol stack contributes with 4.9 ms and 5.4 ms, a value that is rather low, allowing the development of low latency applications and services. Other protocols have much higher latency values, near 18 ms when using HTTP, 35 ms when using the ETSI M2M interfaces, and approximately 50 ms when using WebSockets. The latency variation of all protocols is also rather low, except when using

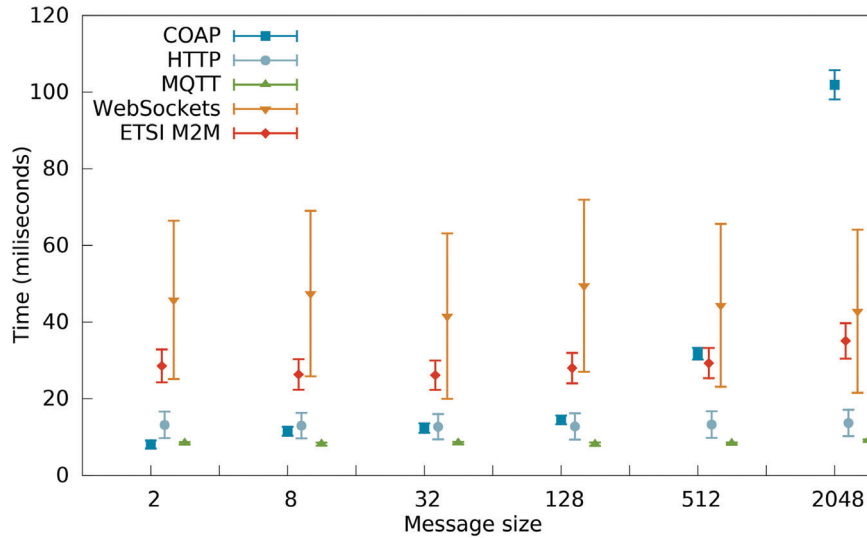


Figure 6 Protocol latency in the SCoT infrastructure.

WebSockets, which present a much higher variation. We postulate that this has to do with the higher complexity of the communication stack, and in particular the need to enter and leave browser context and the JavaScript environment. The next solution with higher latency variation is the ETSI M2M endpoint, which also is more complex than the remaining solutions.

As depicted, most protocols show only a slight increase in latency when we increase the payload size. From our perspective this shows that maximum system capacity was not reached, in particular, the effects of bandwidth constraints are not present. This is however not valid for CoAP. While CoAP provides very low latency for small messages, its behavior is not stable as the message size increases. As observed, the latency observed with CoAP increases linearly with the message size, surpassing WebSocket latency between 512 and 1024 bytes, and presenting a latency of 101 ms, twice the next protocol, when messages reach 2048 bytes.

From this analysis we can determine that our platform presents stable latency over multiple protocols. However, there are relevant differences between the latency values observed, leading to the need for a careful choice by developers and integrators. In particular, while CoAP is designed for sensors and uses the much simpler UDP transport, its adoption should take in consideration the message size to be transmitted.

6 Future Scenarios

Many consider that IoT is in the verge of becoming a mature technology. In our view, we are in its infancy. All mentioned scenarios and architectures follow designs and techniques that have been around for years, yet almost none of them have been fully adopted in real use-cases, resorting to the integration of legacy devices and solutions. One of the big obstacles to the coming of age of IoT is security, in terms of authentication and privacy of the communications, as well as security on the values acquired. The pervasiveness of IoT creates a severe security risk to any organization or individual if proper authentication and privacy is not handled properly. Unfortunately, for economic reasons, security is mostly an optional add-on to IoT solutions as it greatly increases device cost or power budget, development time, and debugging flexibility (among many others). In our platform we followed “security first” approach where all devices were provided cryptographic keys, and communications occurred inside secure tunnels. However, this was not an easy war to fight as it brought many implications.

A very debatable scenario is the penetration of IP to every connected device. Although solutions such as 6LowPan can provide such connectivity, it is in our view something not advisable to do. By being exposed to the Internet, IoT devices will become targets of security probes that will increase costs per device (communication and processing). Solutions such as the one described in this paper can provide similar levels of connectivity while protecting the devices from unintended communications. In fact, during the execution of our trials, several gateway systems were attacked by botnets trying to gain system access. Albeit they didn’t succeed, the amount of bandwidth wasted and system load, in some circumstances was not residual. In the near future we plan to scale our platform to building automation in highly heterogeneous environments, incorporating new technologies such as Bluetooth 4.0 (BLE), which will certainly provide a better playground for determining the risks of exposing sensors.

Another aspect we find critical, is that future systems must consider the quality of the values sampled and reported. Sensor technology is prone to various failure scenarios, in which data is reported but the values reported do not reflect reality. Current approaches based in economies of scale, using many sensors to report the same phenomena add more complexity to the system and do not necessarily solve the issues at hand.

7 Conclusions

In this paper we presented project SCoT and its novel architecture based on ETSI standards. SCoT combines ETSI M2M low level communication and management components with higher-level data manipulation that follows a SOA approach. The SCoT platform allows users to develop innovative M2M services that can be deployed in it by providing not only interfaces to the platform but also runtime engines in which services described in high level languages such as BPEL can easily be created with a graphical tool developed by the project. SCoT has been implemented and tested in the field together with a major Telecom operator and with two different use case scenarios: Smart Agriculture and Smart Road Monitoring. The results of the project such as datasets collect during the project are to be made available to the research community and parts of the developed code have been released using Open Source Licenses. In the future, we plan to extend the project with an identity layer that can support multi-party IoT scenarios.

Acknowledgements

This work has been partially funded by the Portuguese AdI QREN under grant agreements No. 2011/021580 (APOLLO) and CENTRO-07-ST24-FEDER-002031 (CLOUD THINKING), under research grant SFRH/BD/94270/2013.

References

- [1] Y. Agarwal, R. Gupta, D. Komaki, and T. Weng. Buildingdepot: an extensible and distributed architecture for building data storage, access and sharing. In *Proceedings of the Fourth ACM Workshop on BuildSys*, 64–71 (2012).
- [2] M. Antunes, D. Gomes, and R. Aguiar. Context storage for m2m scenarios. In *IEEE International Conference on Communications (ICC)*, 3664–3669, June (2014).
- [3] M. Antunes, D. Gomes and R. Aguiar. Semantic-based publish/subscribe for m2m. In *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)* (2014).
- [4] M. Antunes, J. P. Barraca, D. Gomes, P. Oliveira, and R. Aguiar. Unified platform for m2m telco providers. In *Workshop On IoT Everywhere: Towards smart objects everywhere, 8th International Conference on Ubiquitous Computing Ambient Intelligence (UCAmI)* (2014).

- [5] P. Arjunan, N. Batra, H. Choi, A. Singh, P. Singh, and M. B. Srivastava. Sensoract: a privacy and security aware federated middleware for building management. In *Proceedings of the Fourth ACM Workshop on BuildSys*, 80–87 (2012).
- [6] G. Avram. At the crossroads of knowledge management and social software. *Electr J Knowl Manage* **4**(1): 1–10 (2006).
- [7] J. Bacon, A. Bejan, A. R. Beresford, D. Evans, R. J. Gibbens and K. Moody. Using real-time road traffic data to evaluate congestion. In *Dependable and Historic Computing*, volume 6875 of *Lecture Notes in Computer Science*, C. B. Jones and J. L. Lloyd, eds. Springer, Berlin, Heidelberg, 93–117 (2011).
- [8] A. Banks and R. Gupta. Mqtt version 3.1.1. <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/csprd01/mqtt-v3.1.1-csprd01.html>, December 2013. OASIS Committee Specification Draft 01/Public Review Draft 01.
- [9] A. Barrientos, J. Colorado, J. del Cerro, A. Martinez, C. Rossi, D. Sanz, and J. Valente. Aerial remote sensing in agriculture: a practical approach to area coverage and path planning for fleets of mini aerial robots. *J Field Robotic* **28**(5): 667–689 (2011).
- [10] BeeBotte.com. Available at <https://beebotte.com/>, October 2014.
- [11] R. Blumberg and S. Atre. The problem with unstructured data. *DM Rev.*, 13:42–49, 2003.
- [12] S. A. Boyer. *Supervisory Control and Data Acquisition*. ISA, 2nd ed. (1999).
- [13] T. Bray. The JavaScript Object Notation (JSON) Data Interchange Format. RFC 7159 (Proposed Standard), March (2014).
- [14] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric urban sensing. In *The Second Annual International Wireless Internet Conference (WICON)*, 2–5 (2006).
- [15] M. Chinosi and A. Trombetta. Bpmn: an introduction to the standard. *Comput Stand Interfaces*, **34**(1): 124–134, (2012).
- [16] Eclipse. [iot.eclipse.org](http://iot.eclipse.org/frameworks.html). <http://iot.eclipse.org/frameworks.html>, December (2014).
- [17] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan. The pothole patrol: using a mobile sensor network for road surface monitoring. In *Proceedings of the 6th International Conference on MobiSys*, 29–39 (2008).
- [18] ETSI. Machine-to-machine communications (m2m); functional architecture, October (2011).

- [19] ETSI. Machine-to-machine communications (m2m); mia, dia and mid interfaces, February (2012).
- [20] J. Gantz and D. Reinsel. Extracting value from chaos. *IDC Rev.* 1–12 (2011).
- [21] T. Gruber. Ontology of folksonomy: a mash-up of apples and oranges. *Int J Semantic Web Inform Syst*, **3**(2): 1–11 (2007).
- [22] K. Langendoen, A. Baggio, and O. Visser. Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture. In *20th International Parallel and Distributed Processing Symposium, 2006, IPDPS 2006*, 8, April (2006).
- [23] LogMeIn. Xively. <https://xively.com/> (2014).
- [24] M. Michelson and C. A. Knoblock. Creating relational data from unstructured and ungrammatical data sources. *J Artif Int Res* **31**(1):543–590 (2008).
- [25] P. Mohan, V. N. Padmanabhan, and R. Ramjee. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems, SenSys '08*, 323–336, New York, NY, USA (2008).
- [26] Nimbits.com. <http://www.nimbits.com/>, October (2014).
- [27] OMA. OMA device management v2.0. <http://technical.openmobilealliance.org/Technical/technical-information/release-program/current-releases/oma-device-management-v2-0> (2014) (Accessed: 11 December 2014).
- [28] oneM2M. onem2m. <http://www.onem2m.org/>, July (2014).
- [29] OpenMTC. Openmtc. <http://www.open-mtc.org/>, December (2014).
- [30] T. Rajabioun, B. Foster, and P. Ioannou. Intelligent parking assist. In *21st Mediterranean Conference on Control Automation (MED)*, 1156–1161, June (2013).
- [31] R. Rao. From unstructured data to actionable intelligence. *IT Prof.* **5**(6): 29–35 (2003).
- [32] A. Rowe, M. E. Berges, G. Bhatia, E. Goldman, R. Rajkumar, J. H. Garrett, J. M. F. Moura, and L. Soibelman. Sensor andrew: large-scale campus-wide sensing and actuation. *IBM J Res Dev* **55**(1.2): 6:1–6:14, (2011).
- [33] P. Saint-Andre. Extensible messaging and presence protocol (XMPP): Core. RFC 6120 (Proposed Standard), March (2011).
- [34] C. Schroth. The internet of services: global industrialization of information intensive services. In *2nd International Conference on Digital Information Management, 2007. ICDIM '07*, Vol. 2, 635–642 (2007).

- [35] Sen.se. Open.sen.se. <http://open.sen.se/> (2014).
- [36] Z. Shelby, K. Hartke, and C. Bormann. The constrained application protocol (CoAP). RFC 7252 (Proposed Standard), June (2014).
- [37] C. Shirky. Ontology is overrated: categories, links, and tags. http://shirky.com/writings/ontology_overrated.html, May (2005) (Accessed on 22 July 2013).
- [38] J. K. Suhr and H. G. Jung. Sensor fusion-based vacant parking slot detection and tracking. *IEEE Trans Intell Transport Syst*, **15**(1):21–36, February (2014).
- [39] L. Tan and N. Wang. Future internet: The internet of things. In *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, Vol. 5, 376–380, 2010.
- [40] IBM Emerging Technology. Node-red. <http://nodered.org/>, July (2014).
- [41] J. Thelen, D. Goense, and K. Langendoen. Radio wave propagation in potato fields. In *1st Workshop on Wireless Network Measurements*, Vol. 2, 331–338, (2005).
- [42] ThingSpeak. Internet of things – thingspeak. <https://thingspeak.com/>, October (2014).
- [43] H. Tokuda and J. Nakazawa. Sensecampus: Sensor enabled cyber-physical coupling for ubiquitous services. *J Inform Process* **20**(1): 45–53 (2012).
- [44] R. J. Tomczak, J. Rudowicz-Nawrocka, S. Kujawa, and W. Mueller. Autonomous flying system for grasslands and fields monitoring. In *Proceedings of Fourth International Conference on Digital Image Processing (ICDIP 2012)*, Vol. 8334, May (2012).
- [45] E. Townsend. Future internet report. Technical report, Information Communications Technology KTN, May (2011).
- [46] T. Weng, A. Nwokafor, and Y. Agarwal. Buildingdepot 2.0: an integrated management system for building analysis and control. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, 1–8. ACM (2013).

Biographies



M. Antunes received his computer and telematics engineering MSc degree in 2011, with first class honours, from the Electronics, Telecommunication and Informatics Department, University of Aveiro, Portugal. He is currently working as a researcher for the Advanced Telecommunications and Networks Group at IT – Aveiro His main research areas focus on Knowledge Extraction and Context Storage in Internet of Things (IoT) Scenarios using Machine Learning techniques and Big Data repositories. His works include developing efficient ways to deal with unstructured information. These techniques are being implemented and evaluated in advance machine-to-machine projects, such as APOLLO and SCOT.



J. P. Barraca graduated in Computers and Telematics Engineering from the University of Aveiro, concluded a Masters in Electronics and Telecommunications, and finally a PhD in Informatics Engineering by the same university. He's currently an Invited Assistant at the University of Aveiro, teaching courses related to networking, programming and security. He also is member of Instituto de Telecomunicações, where he coordinates the ATNoG group. His research activities show a strong participation in research and innovation projects, as well as participation in the program committee of several journals and conferences. Currently he is focused in technologies and solutions for Internet of Things, as well as Network Virtualization through Software Defined Networking concepts.



D. Gomes graduated in Computers and Telematics Engineering from the University of Aveiro in 2003 with first class honours, and concluded his PhD by the same University on Resource Optimization for Broadcast Networks in 2009. He's currently an Auxiliar Professor at the University of Aveiro. In the last 10 years has participated in several EU funded projects such as IST-Mobydick, IST-Daidalos, IST-Akogrimo, IST-C-MOBILE, ICT-C-Cast, ICT-Onelab2 and ICT-Medieval where besides conducting research on QoS, IP Mobility, Multicast/Broadcast and Service & Application Development has always been deeply involved in the deployment of prototypes and demonstrations. Recently his research interest are related to Knowledge Extraction and Context Storage in Internet of Things (IoT) Scenarios using Machine Learning techniques and Big Data repositories.



P. Oliveira is finishing his Information Systems Master's degree from the Electronics, Telecommunication and Informatics Department, University of Aveiro. He is currently working on his Master's Thesis, proposing a improved elearning platform for a better guidance on Programming Courses. He is mostly interested on Web, Cloud and IoT.



R. L. Aguiar is full Professor at the University of Aveiro where he received his PhD degree in 2001 in electrical engineering. He has been an adjunct professor at the INI, Carnegie Mellon University and is invited researcher at Universidade Federal de Uberlandia. His current research interests are centered on advanced communication systems and he has more than 400 published papers. He is a member of the steering Board of the Networld 2020 ETP. He has served as Technical and General Chair of multiple conferences and is Associate Editor of several journals. He is a member of ACM and a senior member of IEEE.