Data Analysis as a Service: an infrastructure for storing and analysing the Internet of Things

Martin Lehmann

9th April 2015

Abstract

The Internet of Things (IoT) is perhaps the fastest emerging new technology trend, and must be deeply addressed through research. While we have many small-scale, single-component solutions for connecting parts of the Internet of Things, we have seen very few research oriented full-stack implementations for gathering, storing, and analysing data. We present Data Analysis as a Service (DAaaS), a full-stack sample implementation of a complete IoT application that accepts, stores, and provides both real-time and static access to the data. The main considerations of DAaaS include storage of time-series data in a regular document database, and machine-to-machine communication through standardised APIs. One common research challenge in the Internet of Things, *security*, is considered only briefly, and is of utmost importance in future research.

1 Introduction

- The Internet of Things (IoT) is perhaps the fastest emerging technology trend at the present time. The IoT technologies and applications are still in their infancy (Xu, He & Li, 2014), and so the academic community must thoroughly address the area. This article...
- introduces and describes the most important concepts and areas to address in the field of IoT,
- takes a look at the state of the art in the form of a brief literature review (considering mostly other literature reviews and survey papers) and a brief look at papers relevant to building a full/stack application designed for the IoT,
- presents the major architectural choices in the development of the research artifact (Data Analysis as a Service - DAaaS) with an integration-oriented focus,
- reviews the major untouched areas of interest (mainly security, but also integrations for two-way 2nd and 3rd party integrations),
- and finally concludes with a summary of the results and experiences made throughout the rapid development of the prototype and how it fits into the current research on the field of IoT.

2 The artifact: Data Analysis as a Service (DAaaS)

This section considers the artifact developed in the writing of this article: Data Analysis as a Service (DAaaS).

2.1 Brief description of the implementation

- DAaaS is a full-stack application with several example integrations.
- Core application:
  - Meteor (JS) application built around MongoDB for time series (could/should have used different)
  - REST endpoints everywhere for simplicity (refer to book Architecting the Internet of Things) Both provider and consumer endpoints (for pushing and analysing data)
  - Realtime output endpoints (DDP because there is no de facto standard yet and Meteor provides DDP support out of the box)
- Data providers:
  - Native Android application sending light sensor data (with simple extendability)
  - Cross-platform mobile app written in Ionic to track toilet visits (extremely simple proof of concept)
  - Spark Core (http://spark.io) for any sensor that can be connected to it (e.g., temperature, light, humidity, whatever) with polling rate (e.g., 500 ms between data push)
- Data consumers:
  - Realtime analytics dashboard (just connect via DDP)
  - REST data providing endpoints
- Security is not considered in this application: a simple random token is the only used "authentication", and this is sent in the URL when launching dashboard (refer to Medium.com tragedy)

2.2 Experiences

-

References

Xu, L. D., He, W. & Li, S. (2014). Internet of things in industries: a survey. *IEEE Transactions on Industrial Informatics*, 2233–2243.