Wireless Data Logger Networks

[Extended Abstract]

Javier González
IT University of Copenhagen
Rued Langgaards Vej 7 2300 - Copenhagen, Denmark
jgon@itu.dk

ABSTRACT

Research in the area of wireless sensor networks is predicated on the advent of small, cheap and energy efficient sensors attached to programmable motes. Still Today, many data acquisition systems rely on configurable data loggers that store measurements from a few monolithic sensors and make them available via a serial interface. Examples include the rugged data loggers deployed for in-situ sensing in ecological monitoring programs together with expensive and possibly custom-made sensors, or the electronic energy meters deployed for building energy monitoring. Components are emerging that transform standalone data loggers into a networked system integrated with a data management infrastructure. Such wireless data logger networks are quite different from the mote-based sensornets envisaged a decade ago. Can the results obtained in the wireless sensor network community impact the design of these components? In my thesis, I focus on two components of a wireless data logger network: (i) a serial data transmission component that connects a data logger to a wireless network, and (ii) a data management integration component that supports efficient loading and usage control. In the process, I revisit the need for a generic programming framework in the context of wireless data logger networks.

Categories and Subject Descriptors

J.2 [**Physical Sciences and Engineering**]: Earth and atmospheric sciences; J.7 [**Computers in other Systems**]: Real Time; C.2 [**Computer-Communication Networks**]: Network Protocols, Network Operations; E.0 [**Data**]: General

Keywords

WSN, High-Arctic, Data Acquisition, Wireless Serial Data Transmission (WSDT), Renewable Energies, Smart Grids, Data Management, Access Control, Usage Control

IPSN'12, April 16–20, 2012, Beijing, China.

1. BIOGRAPHY

My name is Javier González, I am 24 and I started my Ph.D (3 years) in November 2011 at the IT University of Copenhagen. My PhD is partly funded by the FP7 infrastructure project INTERACT[1], under the supervision of associate professor Philippe Bonnet. I obtained both my Bachelor and Master's at the University of Valladolid(Spain), being awarded with the Best Student Award in 2011. In summer 2010, I moved to the IT University of Copenhagen (Denmark) to carry out my Master's Thesis in the context of the MANA[1]. In that context, I deployed a wireless data logger network at the Zackenberg Research Station[2] in North-East Greenland. My teaching duties at ITU includes supervising three groups of M.Sc. students at ITU, collaborating with students at Strathmore University Kenya, in the context of their Global Software Development projects[3].

2. PROBLEM

In this thesis, we revisit some of the early assumptions that still shape the WSN community: (i) despite the MEMS revolution those sensors whose development is not driven by consumer electronics, aviation or medical industries are still bulky, expensive and energy inefficient, (ii) despite its success in the research community, TinyOS has not taken over embedded programming practitioners, and (iii) despite some recent breakthroughs with sMap[2] and readingdb[4], there is still a long way to go to reach a flexible yet efficient integration of sensor networks and data management systems.

Example 1 In the context of INTERACT - that gathers all Arctic monitoring stations - we are working together with ecologists that collect data from custom-made sensors attached to stand-alone data loggers. Part of their research consists in defining new sensor modalities (e.g., to measure C02 or CH4 flux), they are more interested in precision than spatial coverage. The data loggers are high bandwidth serial devices. Today, data is collected manually via a laptop connected to the data logger's serial port. However, ecologists are interested in wireless data logger networks for telemetry and tele-command purposes. Note that they are interested in the data, they are certainly not interested in programming the data loggers. Finally, in terms of data management, ecologists rely on MS Access/Excel or other proprietary tools for storing, visualizing, and manipulating the time series they

[1] http://www.eu-interact.org/
[2] http://www.zackenberg.dk/
[3] http://gsd.wikit.itu.dk/home
[4] https://github.com/stevedh/readingdb

collect. Funding agencies are now requiring ecologists to share their data. However, there is no infrastructure that allows them to share their data sets while retaining some form of auditing or control of how these are used.

Example 2 In the context of Building Energy Management construction companies still pick reliable powered sensors and actuators for monitoring and control while energy usage is measured via a hierarchy of meters and submeters. The sensors, actuators and meters embedded in a building are not programmable, they are configurable at best. In terms of data management, building energy management systems are proprietary and tightly integrated with the existing sensors.

In my thesis, I focus on three specific challenges that arise from those examples:

1. How to integrate stand-alone digital devices with a serial interface (e.g., a data logger or a meter) into a networked system - while this is trivial at low baud-rates, efficient data transmission between the UART and the radio interfaces at high baud-rates (115200 bps) has proven a challenge for commercially available EPIC-based systems. We refer to the problem of transmitting serial data over a wireless medium as Wireless Serial Data Transmission (WSDT).

2. How should a wireless data logger network be programmed? Actually the first question is: who is programming a data logger network? What abstractions do they need? Put another way, how generic should a data logger network (or for that matter a sensor network) programming framework actually be? We argue that there is a mismatch between the current incarnation of TinyOS –whose goal is genericity and hardware independence–, and the actual needs of sensor net programming (e.g., forwarding data between a high-speed serial interface and a radio). We also revisit the importance of hardware design on embedded network programming (e.g., while TelosB or Epic motes equipped with TinyOS lack an agile data forwarding capability among their serial interfaces, radio and FLASH -due to shared SPIs- Arduino motes can easily achieve this).

3. While it might seem obvious, there are still many obstacles to using sensor data with a DBMS (be it SQL or NoSQL). We focus on two issues here: (i) The first issue is flexible and reliable data loading: how do we deal with data sets stemming from diverse sources and structure? How do we support provenance when deriving new data sets from existing ones? How to support annotations on existing data sets? (ii) The second issue is how to enforce usage control? While neither scientific data or energy data are sensitive, they are valuable to their owners who want to maintain a form of control over who accesses their data and how it is used. Defining usage control policies for the data generated in the arctic ecologists community or for building energy management is an open issue. Designing an infrastructure for enforcing these policies is also an open issue.

3. APPROACH

In the first part of my Thesis, I focus on integrating stand-alone data loggers into a networked system. In the context of INTERACT, our initial goal is to equip data loggers with wireless capabilities so that they can forward serial data at high baud-rates (115200 bps) in single-hop and multi-hop topologies. Interestingly, we have found that high speed serial data forwarding is a problem for EPIC or TELOS motes running TinyOS (be it with a commercial or a custom-programmed application). For a mote to sustain a 115200 baud data rate, the UART must be available every 0.07ms. Since (i) the UART is interrupt driven and since (ii) there are components in TinyOS that either are non-interruptible or disable interrupts, this cause the UART to miss interrupts and thereby data. This phenomenon occurs in point-to-point links, and increases when extending to a multi-hop network with routing capabilities. Also, (iii) the SPIs are a limited resource which is shared, and has to be acquired and released as such. We associate this overhead with an observed high frequency of motes hanging. In contrast, the problem of forwarding serial data at high baud-rates (115200 bps) is easily solved with an out-of-the box Arduino system. More specifically, this solution relies on the following capabilities from an Arduino Mega 2560 equipped with XBee: (i) multiple serial interfaces (UART) with separated SPIs, (ii) out-of-the box configurable point-to-point WSDT, (iii) easily programable multi-hop WSDT. Generalizing from this example, we are now investigating the interplay between hardware design and programming framework for specific classes of network embedded systems.

The second part of my research is focused on how to handle the vast amount of diverse data sets generated by a given instrumentation infrastructure. The need to organize the management, publication and archival of these data sets, originally pointed out by Jim Gray[3]. While there is a consensus that a collaborative data infrastructure is needed to allow researchers in different domains to collaborate on the same data sets in order to find new insights, there are significant barriers to the realization of this vision. We have a prime example in the arctic ecologists community, where the International Polar Year[5] has not allowed to make a significant breakthrough in terms of data sharing. One of the key challenge is to allow scientists to share their data widely while retaining some form of control over who accesses this data (access control) and more importantly how it is used (usage control). Access and usage control is necessary to enforce existing open data policies. I aim at defining, implementing and evaluating such an infrastructure in the context of INTERACT.

4. REFERENCES

[1] M. Chang and P. Bonnet. Monitoring in a high-arctic environment: Some lessons from mana. *Pervasive Computing, IEEE*, 9(4):16 –23, october-december 2010.

[2] S. Dawson-Haggerty, X. Jiang, G. Tolle, J. Ortiz, and D. Culler. smap: a simple measurement and actuation profile for physical information. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 197–210, New York, NY, USA, 2010. ACM.

[3] J. Gray, D. T. Liu, M. Nieto-Santisteban, A. Szalay, D. J. DeWitt, and G. Heber. Scientific data management in the coming decade. *SIGMOD Rec.*, 34:34–41, December 2005.

[5] http://www.ipy.org/