Integrating Future Large-scale Wireless Sensor Networks with the Internet

Marco Zúñiga Z. and Bhaskar Krishnamachari
Department of Electrical Engineering, University of Southern California,
Los Angeles, CA 90089,USA.
E-mail: {marcozun,bkrishna}@usc.edu

***Abstract*— Advances in hardware technology and wireless communications will enable the development of large-scale wireless sensor networks (WSN). Due to the variety of applications and their importance, WSN will need to be connected to the Internet. We discuss the issues involved in this integration. In particular, we point out why all IP-sensor networks are infeasible, and suggest the feasible alternatives in the case of both homogeneous and heterogeneous networks.**

I. Introduction

Wireless Sensor Networks (WSN) are promising to change the way we obtain information from the physical environment. It is envisioned that WSN will consist of thousands to millions of tiny sensor nodes, with limited computational and communication capabilities. When networked together, these unattended devices can provide high-resolution knowledge about sensed phenomena. According to a recent National Research Council report, the use of such networks "could well dwarf previous milestones in the information revolution [1]. Possible applications of these networks range from habitat and ecological sensing, structural monitoring and smart spaces, emergency response and remote surveillance. In recent years there has been a great surge of interest in WSN, focused on developing the hardware, software, and networking architectures needed to enable such applications [2].

In general, WSN can operate as stand-alone networks or be connected to other networks. Real-world experiments have been done with both types of network architectures, though at much smaller scales than envisioned for the future. As an example of the stand-alone network architecture, consider the example of the target tracking application described in [3]. In that project, several wireless nodes were deployed in the desert to track the route of a tank. An unmanned areal vehicle (UAV) was then flown over the nodes in order to collect the information. For many important applications, however, it makes a lot of sense to integrate these sensor networks somehow to the existing IP networks. An in-depth study of applying wireless sensor networks in a real-world habitat monitoring application is presented in [4]. In this project, some Berkeley sensor motes [5] were deployed in the Great Duck Island, off the coast of Maine, to monitor the microclimates in and around nesting burrows used by the Leach's Storm Petrel bird. The nodes would periodically sample and relay their sensor readings to a gateway connected to the Internet through a satellite link, allowing researchers around the world to access real-time environmental data.

The task of connecting WSN to the existing Internet brings with it several challenges. Any network wishing to be connected to the Internet needs to address the question of how it will interface with the standard protocols like the Internet Protocol (IP). In this article, we describe the characteristics of WSN that differentiate them from traditional IP-based networks: chief among these are that WSN are large-scale unattended systems consisting of resource-constrained nodes that are best-suited to application-specific, data-centric routing. As we shall see, these fundamental differences rule out the possibility of all-IP sensor networks and recommend the use of application-level gateways or overlay IP networks as the best approach for integration between WSN and the Internet.

II. Characteristics of Sensor Networks

Data Flow Patterns: The most basic use of sensor networks is to treat each node as an independent data collection device. Periodically, each node in the network sends its readings to a central warehouse/data sink. Alternatively, it is possible to treat sensor networks as essentially distributed databases - users interested in specific information insert a query into the network through a node (or nodes) usually called the sink, as shown in figure 2. This query is propagated into the network. Then nodes with the data – called sources in WSN jargon – respond with the relevant information. Thus one-to-many and many-to-one data flows dominate the communications in sensor networks. This can be contrasted with the arbitrary one-to-one addressable flows that are typical of most IP-based networks.

Energy Constraints: The nodes in unattended large-scale sensor networks are likely to be battery powered, with limited recharging capabilities. Under these conditions, the primary network performance metric of interest is the energy efficiency of operation (a related metric is the lifetime of the network - measurable in terms of the time when a significant portion of nodes in the network fail due to energy depletion). Typically, communication is significantly more

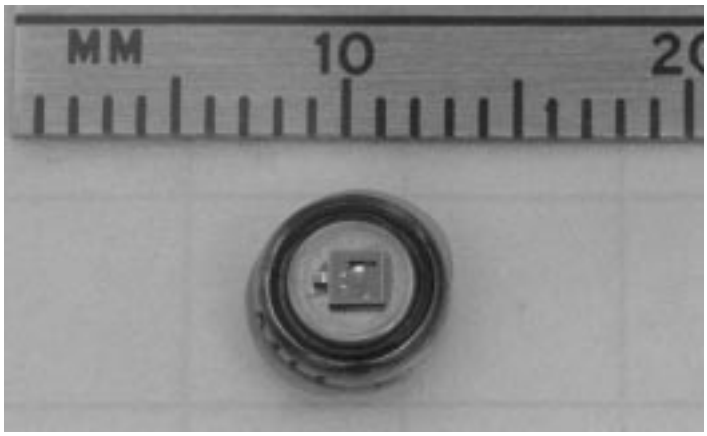

Fig. 1. The Smart Dust project at UC Berkeley [7] aims to create sensor nodes of a size of a grain of sand. These devices will contain sensors, computational ability, bi-directional wireless communications, and a power supply. The picture shows the current size of the sensor nodes. Image reproduced with permission from http://www-bsac.eecs.berkeley.edu/~warneke/SmartDust/.

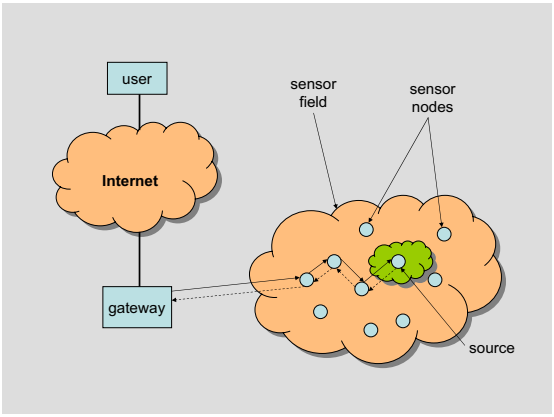

Fig. 2. Communication Architecture using a gateway: the full arrows represent the dissemination of the query, and the dashed arrows, the data routed back. In this case the gateway is the single point of access to the WSN and it performs the conversion of the necessary protocols including IP.

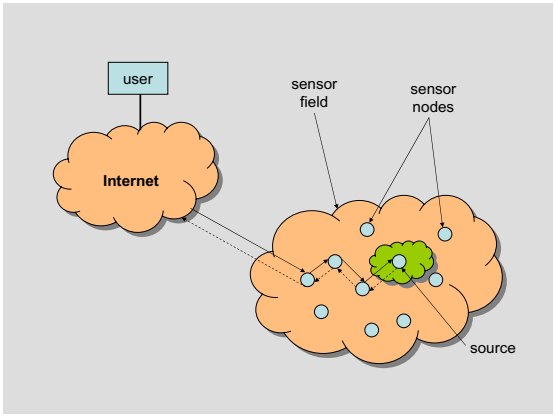

Fig. 3. Communication Architecture with direct connection: the difference with figure 2 is that in this case every node has an IP address and can be directly access from any point in the Internet that has wireless capabilities.

energy-expensive than computation. The Berkeley motes, for example, can process 100 instructions with less energy than the amount needed to transmit a single bit. This has led researchers to espouse communication-minimizing design principles for sensor networks that are directly at odds with the application-independent networking methodology that underlies traditional networks.

Application-specific networking and data-centric routing: Traditional IP-based networks follow the layering principle which separates the application level concerns from network-layer routing. This is necessary because a multitude of applications are expected to run over a common networking substrate. By contrast, sensor networks are likely to be quite limited in the applications they perform. This calls for cross-layer optimizations and application-specific designs. One design principle that exploits application-specificity to significantly reduce communication energy is the use of in-network processing to filter out irrelevant and redundant information. For example, intermediate nodes may be allowed to look at the application-level content of packets in order to aggregate them with information originating from other sources [8].

Related to this is the distinction between address-centric and data-centric routing. The Internet was designed around an address-centric ideology, which works when data is usually

	Traditional IP-Based Networks	Large Scale Wireless Sensor Networks
Networking Mode:	Application-independent	Application-specific
Routing Paradigms:	Address-centric	Data-centric, Location-centric
Typical Data Flow:	Arbitrary, One to one	To/from querying sink, One-to-many and many-to-one
Data Rates:	High (Mbps)	Low (kbps)
Resource constraints:	Bandwidth	Energy (battery-operated nodes), Limited processing and memory
Network Lifetime:	Long (years-decades)	Short (days-months)
Operation:	Attended, administered	Unattended, Self-configuring

Fig. 4. Key differences between traditional IP-based networks and large scale wireless sensor networks.

attached to a specific host. It requires a prior knowledge of which host to contact. Almost all transactions (ftp, http, email etc) in the Internet have this characteristic – it is known *a priori* where the data is located. For this reason, communication on the Internet is usually point-to-point, and this requires the ability to uniquely identify each host through IP addresses.

In sensor networks, however, the query is most likely to be for named data. For example, in a WSN application, the question is unlikely to be: "What is the temperature at sensor number 271?" Rather, the question would be: "Where are the nodes whose temperature exceeds 45 degrees?" [6]. The Directed Diffusion protocol [9] has shown that it is possible to do data-centric querying and routing without the use of globally unique IP-like addresses for all nodes in the network. One advantage of doing without globally unique ID's is that each packet need not carry address information in the header. Many applications for low-rate sensing will result in small amounts of data per packet (on the order of a few bytes). IPv6, for example, has 40 bytes of header per packet. Doing without this header addressing information can result in a significant reduction in communication overhead – and thus energy.

Another argument for doing without globally unique ID's for all nodes in sensor networks is the complexity of address-management in such large-scale, unattended , self-configuring networks. Keeping in mind the limited lifetime of disposable sensor nodes and their large numbers, it would otherwise be necessary to implement complex, dynamic address allocation schemes (similar to DHCP). These schemes may present an additional energy-burden on the network.

Finally, the implementation of the full IP stack on sensor networks may not be feasible due to the limited computational and memory resources on component nodes. Sensor networks are thus in many ways fundamentally different from traditional IP-based networks. For these reasons, all-IP large-scale sensor networks are neither desirable nor feasible.

III. Gateway-based Integration

We have discussed why giving an IP address to every sensor node is not the right approach to integrating sensor networks with the Internet. While it is desirable to not have to develop new protocols or perform protocol conversion at gateways, the application specific property of wireless sensor networks demands this type of solution. Single or multiple independent gateways are called for in *homogeneous networks*, where all the nodes have the same capability in terms of processing, energy and communication resources. In addition to gateways,

an overlay IP network may be utilized in *heterogeneous networks*, where some nodes may be more capable than the majority of nodes (for example, when some laptop computers can be part of the network).

A. Homogeneous Wireless Sensor Networks

The basic solution for integration in the case of a homogeneous wireless sensor network is to use an application-level gateway to interface the sensor network to the Internet. The gateway may be implemented in the form of a web-server for example. In the case of simple sensor networks where nodes are providing information continuously, they can be stored and displayed on a dynamic web-page from the gateway node. This is more or less the approach taken, for example, in the Great Duck Island experiment [4]. In the case of more sophisticated sensor networks, the gateway can viewed as a front-end to a distributed database. The users accessing the gateway server may issue SQL-type queries. The query optimization is performed through data-centric in-network processing and the response is obtained from the network and displayed to the user. One drawback of this approach is that a lot of data has to be routed from and back to the gateway, implying that all the nodes nearby the gateway will exhaust their energy resources sooner, if they are not rechargeable.

Another possibility is to deploy wireless sensor networks with more than one independent gateway used as points of interface between the network and the Internet. Having several points of access to the network would have two important advantages: eliminating a single point of failure and distributing evenly the energy consumed by the nodes (assuming the queries on the different gateways can be load-balanced).

In homogeneous WSN, where all the nodes have the same capabilities, the flexibility for other communication architectures is limited. We now turn to the case of heterogeneous WSN.

B. Heterogeneous Wireless Sensor Networks

Heterogeneous networks allow for the possibility of giving an IP address to the more capable nodes in the network. In general, capable devices could perform more tasks, and hence carry more of the burden in the network. There may also be application-specific reasons why these more capable devices should be addressable from within and without the network. For example, if the more capable devices are capable of actuation, they may need to be addressed in order to be tasked. In other scenarios, the higher capability nodes may act as addressable cluster-heads. In such networks, it may be possible to construct an overlay IP network that sits on top of underlying wireless sensor network. The technical challenge in this approach is to construct some kind of tunnelling mechanism to allow the devices with IP addresses to communicate among themselves in an address-centric manner (figure 5). In general, the IP-addressable nodes in

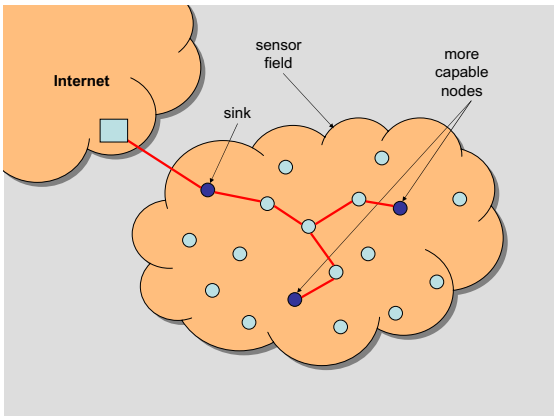

Fig. 5. Heterogeneous Network: the red lines show the tunneling communicating the nodes with IP addresses (blue circles).

the network may not be adjacent to each other. To create an overlay IP network, then, it will be necessary to create some form of a link-abstraction from the multiple hops between nearby IP-addressable nodes. If the intermediate nodes do not have any global identifiers, the link-abstraction will need to be formed in a data-centric manner.

The problem of creating tunnels depends on the characteristics of the wireless sensor network. If the application is more likely to have a high IP-traffic inside the wireless sensor network, then, multiple paths among the IP-addressable nodes would be preferred, in order to load balance the consumption of energy in the less-capable sensor nodes. On the other hand, if the IP-traffic is going to be low, then, a single route can be enough. We describe briefly how to build an overlay network based on a flooded-query approach (Directed Diffusion [9]), which is going to be suitable for high traffic, and also how to build it using a directed-query approach (ACQUIRE [10]), which is suitable for low traffic conditions.

We believe that Directed Diffusion is a good candidate to build up the overlay structure in high-IP-traffic applications. Directed Diffusion is a data-centric communication paradigm that is quite different from the address-centric ideology in traditional networks. The goal of Directed Diffusion is to establish efficient n-way communication between one or more sources and sinks. In basic Directed Diffusion, an *interest* for named data is first distributed through the network via flooding. The interest description is done by attribute-value pairs. In our case it could be described as:

```
type: IP-addressable // detect nodes that have an IP address
interval: 20ms       // send message every 20ms
duration: 200ms      // ... a total of 10 messages
```

This initial interest can be seen as exploratory and the data rate should be low. As the interest is propagated, the nodes set up gradients from the source back to the sink. Upon

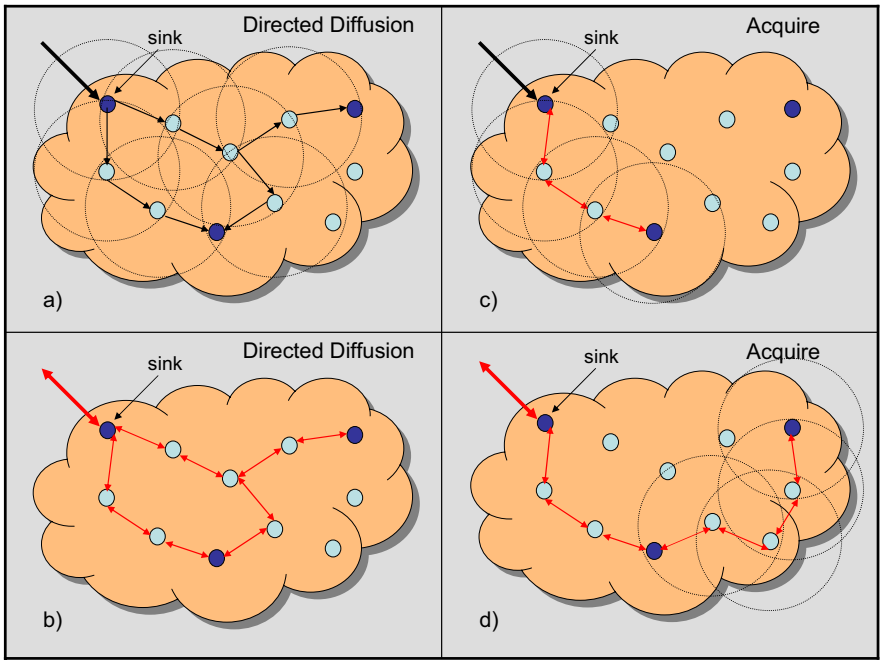

Fig. 6. Heterogeneous WSN with and Overlay IP network: a) shows the first stage in directed diffusion where the query is flooded to find all the IP-addressable nodes. b) shows that multiple routes are obtained with this mechanism. c) ACQUIRE is used to build up the overlay IP network, observe that the query is sent through a path - that can be randomly chosen. d) shows the overlay obtained by ACQUIRE where only one path is obtained

reception, the sources with relevant data (IP-addressable nodes) respond with the appropriate information stream. For our example the response could be of the form:

type: IP-addressable // the sensor node has an IP address
address: IP-address // IP address of replying node

This data is sent back through the interest's gradient path. After reception, the sink must refresh and reinforce the most efficient paths. Finally, the sink can select n-paths depending on the expected IP-traffic, higher traffic would imply more paths in the overlay structure. Figure 6 a) and 6 b) show the directed diffusion mechanism. Note that once the overlay paths are created, they may be used for arbitrary communication between the IP-addressable nodes, not only between the nodes and the sink.

The main advantage of implementing the overlay structure with Directed Diffusion is that it can provide multiple paths; however, the amount of energy consumed by the network is high. If the IP-traffic is expected to be low and the number of IP-addressable nodes is known a priori, then a more energy efficient routing method can be used to construct the overlay. The basic idea is to send an agent to traverse the network and find all the IP-addressable nodes, instead of flooding. One proposed routing mechanism that uses agents in WSN is ACQUIRE [10]. We now briefly explain how an overlay structure can be set using ACQUIRE.

ACQUIRE is a novel resource discovery mechanism that presents significant savings in terms of energy compared with flooding – at a cost of longer delays. ACQUIRE is suitable for one-shot, complex queries. For creating an overlay IP network, a one-shot query could be sent to find routing information about nodes X, Y, Z which are known to have IP-addresses. In ACQUIRE, the active query is forwarded step by step through a sequence of nodes. At each intermediate step the node which is currently carrying the active query does a lookahead of d hops in order to resolve the query partially. Once the resource is found the required data is sent back. For our purposes, routing information to these nodes must be included in the data that is sent back (e.g. by including the intermediate routing nodes in the data, as in source-routing). Figure 6 c) and 6 d) show the ACQUIRE mechanism.

We have observed that depending on the characteristics of the network the algorithms required to solve the tunnelling problem may be different. This highlights, once again, the application-specific characteristic of WSN.

IV. Conclusions and Comments

The numerous and important applications of wireless sensor networks demand for an integration with existing IP networks, especially the Internet. An all-IP-network will not be viable with this new technology, due to the fundamental differences in the architecture of IP-based networks and sensor networks that we have outlined in this paper. We

foresee that the integration of sensor networks with the Internet will need gateways in most cases.

Finally we should point out that even though IP addresses are not suitable to identify every sensor node in WSN, some applications may require at least a subset of the nodes to possess a unique ID within a wireless sensor network. For example, sensors that are also responsible for controlling an actuator or higher-capability sensors acting as cluster-heads may need a unique global address identifier, so that they may be individual tasked or to facilitate direct access to the Internet. In this case, either a specific internal identification scheme, or the overlay IP network solution (that we described for heterogeneous networks) may be used.

REFERENCES

[1] D. Estrin *et al. Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers*, National Research Council Report, 2001.

[2] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam and Erdal Cayirci, "A Survey on Sensor Networks," IEEE Communications Magazine, vol. 40, no. 8, August 2002.

[3] "http://tinyos.millennium.berkeley.edu/29Palms.htm"

[4] Alan Mainwaring, Joseph Polastre, Robert Szewczyk, David Culler, John Anderson, "Wireless Sensor Networks for Habitat Monitoring," In the 2002 ACM International Workshop on Wireless Sensor Networks and Applications. WSNA '02, Atlanta GA, September 28, 2002.

[5] TinyOS Homepage. http://webs.cs.berkeley.edu/tos/

[6] Deborah Estrin, Ramesh Govindan, John Heidemann and Satish Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," Mobicom 1999.

[7] Kris Pister, The Smart Dust Project, "http://robotics.eecs.berkeley.edu/ pister/SmartDust/"

[8] B. Krishnamachari, D. Estrin, S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks," *International Workshop on Distributed Event-Based Systems, (DEBS'02)*, Vienna, Austria, July 2002.

[9] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *ACM/IEEE International Conference on Mobile Computing and Networks (MobiCom 2000)*,August 2000, Boston, Massachusetts

[10] N. Sadagopan, B. Krishnamachari, and A. Helmy, "The ACQUIRE Mechanism for Efficient Querying in Sensor Networks," *First IEEE International Workshop on Sensor Network Protocols and Applications (SNPA'03)*, May 2003.