Technical Document on Sensor Networks

GDD-06-19

GENI: Global Environment for Network Innovations

September 15, 2006

Status: Draft (Version 1.0)

Note to the reader: this document is a work in progress and continues to evolve rapidly. Certain aspects of the GENI architecture are not yet addressed at all, and, for those aspects that are addressed here, a number of unresolved issues are identified in the text. Further, due to the active development and editing process, some portions of the document may be logically inconsistent with others.

This document is prepared by the Wireless Working Group.

Editors:

Ramesh Govindan, *USC*

John Heidemann, *ISI/USC*

Matt Welsh, *Harvard*

Deborah Estrin, *UCLA*

1. Introduction

Sensor networks will be pervasive in the future Internet. If we want to be able to design and experiment with the next generation Internet architecture, it is essential to understand how sensornets will affect that architecture, so we must involve sensornets in the GENI facility. This document proposes the design of the sensor network component of the GENI facility.

An essential aspect of sensornets is that they are physically coupled to the environment and therefore effective experimental facilities must encompass *realistic applications*. When separated from the requirements of applications and other real-world constraints, embedded sensing systems risk becoming an academic exercise. With authentic requirements and constraints they have a much greater opportunity to influence computer science, others science, and education.

Because of this need, an important requirement underlying this design is *realism*. By this, we mean that the sensor networks in GENI should not merely be thought of as "generic" sources of traffic for the GENI backbone. Rather, these networks must be carefully designed in consultation with application domain experts (e.g., have realistic sensors, node placements, and uses). This level of realism will appropriately inform the development of Internet architectural elements, will advance sensor networks research, and will lead to increased use and impact of this technology on the relevant domain experts.

We envision three kinds of sensor network testbeds connected to the GENI facility:

- **Deployed regional testbeds** (2-3): These will be deployed on a campus or metropolitan scale, and may consist of several sensor "sub"-networks, each of which is designed with a particular application in mind.
- **Sensor subnet Kits** (50): These will consist of up to 100 nodes each, and wil be each designed with a particular application in mind. Each kit would be acquired, possibly customized, and managed by a PI or research group. Each kit would be connected to the GENI backbone, and accessible to users with some locally enforced policy.
- **Indoor reconfigurable labs** (2): This testbed would contain a large number of sensor nodes (about 1000), and would allow sensor network researchers to experiment at scale with a greater variety of traffic patterns and network configurations than possible with the deployed regional backbone, and in a repeatable and controlled manner.

From the GENI perspective, the division between kits and the regional testbeds mirrors the expected topological distribution of sensor networks. The future Internet will contain a large number of small to moderately sized networks connected directly to the Internet infrastructure (to, say, a regional or campus network), and a smaller number of sensor networks each with large spatial extent and connected to the Internet infrastructure at perhaps more than one or two locations. From the sensor network perspective, the large number of kits will allow sensor network researchers to experiment with a diversity of applications, the deployed regional testbeds will provide realistic data and enable experimentation at representative geographic scales, and the indoor reconfigurable labs will push scale in numbers of nodes and provide some measure of repeatability and control.

A major component of the GENI sensor network infrastructure will be shared testbed facilities that will permit experimentation at significant physical scale. Each testbed will be interfaced to GENI and will be open to the research community, with the goal of encouraging multiple research groups to leverage the infrastructure. In any testbed setting, there is a tension between the desire to tie the testbed to a specific application domain and the need for retasking for varying experimental needs. For this reason, we have proposd multiple testbeds with partially overlapping goals that encompass a range of physical scales and node capabilities.

The following sections describe the proposed facilities in more detail.

2. Deployed regional testbeds

The first class of GENI sensor testbed will permit studies at very different geographic scales and with more capable sensor nodes than the proposed indoor testbeds. This approach will push sensor networks to span new application domains and consider new design tradeoffs. In contrast to the indoor testbeds, urban-scale testbeds will focus on direct tie-in with application requirements and offer a maintained experimental facility with a high degree of realism.

An urban sensor testbed will consist of 200 or more nodes mounted in an outdoor, urban setting, such as on the sides of buildings or on utility poles. Each node will be permanently powered and will consist of an embedded PC-class device with a moderately powerful radio interface (e.g., 802.11a or g). Multiple nodes will have a wired Internet connection, acting as gateways between the testbed and the rest of the GENI infrastructure. The remaining nodes will rely on wireless communication.

The sensor types used by the urban testbeds will vary depending on the specific domain applications considered, but will typically consist of climate sensors (temperature, humidity, barometric pressure, sunlight, wind speed, rainfall, and snowfall), air pollution sensors, and so forth. These sensors can be useful in a range of environmental, pollution, and public health studies.

We also anticipate that these urban testbeds can be integrated into colocated fixed or mobile sensor deployments by other groups. For example, a vehicular sensor network experiment could collect environmental data both from lightweight, inexpensive sensors mounted on vehicles, but combine this data with that collected from more elaborate, fixed sensors as part of the urban testbed. The presence of the urban testbed will facilitate development of experiments that use the city environment as an "active laboratory" with new sensors and capabilities being added over time.

3 Testbed federation and management

An important aspect of building open testbeds is the framework and policies for managing shared resources. We anticipate being able to draw on previous work in developing solutions for GENI sensor network facilities that permit both space- and time-sharing of the testbed systems as well as varying user needs. For example, batch scheduling may be appropriate for lengthy experiments whereas non-exclusive, "interactive" use may be preferable for debugging and initial experimentation.

A number of projects, such as PlanetLab, the various Grid efforts, and the sensor network indoor testbeds mentioned below, have proposed varying approaches to resource management in a shared testbed environment. MoteLab makes use of a simple time-slot scheduler with a per-user quota on outstanding jobs. The Intel Mirage testbed requires that users "bid" on needed resources in a virtual market environment to support differentiated service during periods of contention. Each of these approaches is tailored for somewhat different usage scenarios and degrees of user sophistication. Striking the right balance between ease-of-use and flexibility for advanced experimentation.

The proposed sensor infrastructure also offers the exciting potential to federate multiple testbeds into a single, shared infrastructure for conducting sensor-based research at unprecedented scales. Rather than viewing each testbed as an independent entity, these systems can be linked into a "Sensor Grid" with software services for deploying applications, collecting data, monitoring status, and debugging. Other sensor-based infrastructures, such as EarthScope, NEON, and CASA, can be linked into the facility, providing new opportunities for real-time data collection supporting numerous eScience efforts. The GENI indoor and urban sensor testbeds will serve as a prototyping environment for such a federated sensor platform.

4. Sensor subnet Kits

Realistic applications and application traffic are one of the most important things sensornets bring to GENI. Augmenting traditional testbeds, we see Sensor subnet "kits" as essential to encourage a diverse set of realistic applications. Kits accomplish this by lowering the barrier to entry in using GENI for sensornet research by applications researchers. As a secondary goal, they also can serve as a distributed resource for GENI researchers.

Sensornet kits have two components, one intellectual and the other physical. The intellectual component of sensornet kits is a common set of designs for a sensor platform, a standard set of sensor packages, and the systems software necessary to make this hardware reasonably easy to use. The physical component will be approximately 50 "sensornet clouds" at diverse sites that have a moderate number (20-100) of these devices, connected to the GENI backbone either through a regional sensornet testbed (described below) or connected directly via a gateway host.

We expect GENI to select a few default sensor packages for applications of known interest (examples might include seismic monitoring, habitat monitoring, and simple localization and tracking). In addition, we expect a standard sensor interface board with analog and digital I/O capabilities will enable easy addition of other sensors as required for unanticipated or specialized applications. The sensor hardware platform would be based on a standard family of processing and communications modules, supporting two or three flavors with different capabilities. An example might be today's MicaZ and Intel Stargates. The software platform should be open source with minimal licensing restrictions (BSD-style license or similar), with a group that maintains the point and coordinates bug reports, fixes, and releases. The gateway device that connects a sensornet cloud to GENI will be responsible for allowing sensornet traffic into the broader GENI network, and for mediating use of the cloud by external researchers, as discussed below.

The primary purpose of kits is to reflect sensornet applications into GENI, while a secondary purpose is to serve as a resource for general GENI researchers. These goals influence how we see kits being used. We therefore see kits as owned by small groups of researchers and dedicated for that group's application. We expect individual sensor nodes to be deployed as required by a given application, and so different clouds will have very different spacing, connectivity, duty cycles, sets of sensors, etc. The projects would remain tightly integrated with GENI by requiring that the data from and the API to the specific application be available to GENI researchers. The kit applications would both use the GENI backbone, and making available public models or traces that represent that application. The secondary goal of kits is to provide a resource for general GENI researchers. Sharing access to kits would be optional, but strongly encouraged to the extent it doesn't conflict with the application. We expect the cloud's gateway host would coordinate access to nodes in the cloud. If made available for external, individual sensors in a cloud can be "virtualized" using similar mechanisms as other wireless testbed components, using time-slicing and frequency-based partitioning.

The advantage of kits comes from centralizing development of the hardware and software components common to a number of users. We therefore see kits as requiring a central development lead. However, kits also must be very reactive to the needs of applications. Kit design must therefore promote extensibility and make it easy to add custom sensors and software by sites fielding kits. We envision two models of kit acquisition. We expect GENI to seed a moderate number (50 or so) of moderate size (20-100) clouds. To leverage other GENI investments, these clouds should ideally be coordinated and co-located with regional sensornet testbeds. However, the primary selection criteria should be to involve a diverse set of applications. The second model of kit acquisition will be through existing NSF research funding opportunities. We encourage general researchers to purchase kits from a third-party manufacturer and incorporate them in their work. GENI can encourage these researchers to attach to the GENI backbone. Even though these kits are not directly funded by GENI, promulgation of the GENI designs maximizes the impact of GENI, and we expect that researchers will be interested to attach to the GENI backbone to collaborate and gain from networking effects. (A similar "ecosystem" has arisen around today's Planetlab.)

Kits will require a number of hardware and software components. Some of these are common with general sensor networking research: stable hardware platform designs and systems software. Others are common to GENI: software for allocation, sharing, virtualization, remote reprogramming, data anonymization and privacy management, data archiving and modeling. The only component unique to kits is the need for policy control and deployment management. To allow both application-driven use and remote access, policy controls are needed that insure that remote users do not interfere with the primary application. Deployment management tools are needed to make the deployment transparent to remote users. Primary, local users are easily aware node positions and status because they presumably have physical access to the nodes or participated in their deployment. Similarly, they can easily find out what nodes are unavailable by checking the nodes or asking the researcher down the hall. Remote users lack these benefits. We expect that new deployment tools will be needed that make it very easy to document sensor deployment and current sensor status. We expect that localization services, needed for other sensornet testbeds, will be beneficial here.

5. Indoor Reconfigurable Labs

Indoor sensor network testbeds, such as Kansei, Mirage, MoteLab, the Emstar emulator, and Emulab, are currently providing the sensor network community with a valuable resource for controlled experimentation in an indoor setting. These testbeds each operate on a scale a few hundred nodes deployed in either a very dense laboratory setting or more widely distributed throughout a building. Each sensor is typically attached to a serial or network interface for programming and debugging, and is permanently powered. A front-end interface provides scheduling, programming, debugging, and logging capabilities.

For GENI, the primary focus of indoor testbeds will serve as a development, testing, and benchmarking sandbox to experiment with sensor networking primitives at a significant scale. As such, these testbeds are not intended to represent a specific application domain, but rather constitute an "artificial" environment for large-scale measurements in a controlled setting.

We anticipate that existing sensor testbeds will interface to GENI to seed this effort. In addition, we intend to develop an additional 2-3 indoor reconfigurable labs at significantly higher scales than currently exist, with as many as 1000 sensor nodes deployed over a large building or multi-building campus. Such an environment will enable experimentation at unprecedented scales as well as space-sharing of the testbed to support multiple concurrent experiments. Federation of multiple testbeds will offer researchers the unique opportunity to measure proposed protocols, algorithms, and systems in a range of controlled settings. By providing a standardized environment for measurements, direct comparisons across competing techniques can be made.

A typical indoor reconfigurable lab will consist of 500-1000 sensor nodes such as the current MicaZ or TMote Sky, equipped with an embedded microcontroller and 802.15.4 radio. However, some labs may include or be dominated by other radios (e.g., 802.11g or a software defined radio). Each unit will be connected to an Ethernet or serial/USB backchannel for programming, debugging, and data logging. Each sensor will be interfaced to a suite of sensors, depending on the application tie-in focus of the testbed; one lab can also support nodes with varying sensor types. Each lab will include the order of 20-40 more powerful gateway devices, such as embedded PCs, permitting studies of multitier architectures, and which act as local aggregation points. Finally, each lab will have a front-end server providing a user interface, scheduler, and multiterabyte database for logging traces collected from the network.

6. Synergy with other efforts in GENI

We see two points of synergy with other efforts in GENI.

First, our large testbeds will contain wireless mesh backbones that are functionally similar to those being proposed by the wireless working group. As such, it may be possible to design a testbed that is used both for experimentation with wireless mesh infrastructures, and for sensor network research. We note, however, that the sensor network testbed requirements are slightly different from those of the wireless mesh testbed:

- the sensor network testbed needs a stable mesh backbone since that is essential for providing connectivity between different sensor subnetworks.
- the sensor network mesh backbone will need to be able to run higher-level services on the wireless nodes (storage, computation), not merely use the backbone nodes for packet forwarding.

Second, the virtualization mechanisms implemented on the mesh backbone will be similar to (and perhaps be more resource constrained versions of) the virtualization software developed on the GENI backbone. Thus, if the GENI virtualization software were designed with these requirements in mind, it would avoid some duplication.