Environmental Monitoring 2.0*

Sebastian Michel#, Ali Salehi#, Liqian Luo+, Nicholas Dawes◇, Karl Aberer#, Guillermo Barrenetxea#, Mathias Bavay◇, Aman Kansal+, K. Ashwin Kumar#, Suman Nath+, Marc Parlange#, Stewart Tansley+, Catharine van Ingen+, Feng Zhao+, Yongluan Zhou#

#*Ecole Polytechnique Fédérale de Lausanne, Switzerland*
+*Microsoft Research, Redmond, USA*
◇*WSL, Swiss Federal Institute for Snow and Avalanche Research (SLF), Davos, Switzerland*
contact: sebastian.michel@epfl.ch

***Abstract*— A sensor network data gathering and visualization infrastructure is demonstrated, comprising of Global Sensor Networks (GSN) middleware and Microsoft SensorMap. Users are invited to actively participate in the process of monitoring real-world deployments and can inspect measured data in the form of contour plots overlayed onto a high resolution map and a digital topographic model. Users can go back in time virtually to search for interesting events or simply to visualize the temporal dependencies of the data. The system presented is not only interesting and visually enticing for non-expert users but brings substantial benefits to environmental scientists. The easily installed data acquisition component as well as the powerful data sharing and visualization platform opens up new ground in collaborative data gathering and interpretation in the spirit of Web 2.0 applications.**

I. Introduction

In recent years, sensor networks have gained wide popularity in a variety of application scenarios, ranging from monitoring applications in production chains to more sophisticated sensor deployments in the environmental sciences.

In science, the requirement to develop a bespoke acquisition, databasing and querying infrastructure for each application adds a layer of expense and a requirement for skills which may not be present within the team. A generic infrastructure which addresses all of these issues whilst remaining open and flexible enough to allow the scientist to carry out any data processing required, allows costs to be reduced and allows more "science" to take place. Such a generic infrastructure to support environmental science projects is presented here, focusing on environmental monitoring inside the Swiss Experiment (SwissEx)[1] project.

SwissEx is a collaboration of environmental science and technology research projects. These projects cover a range of environmental hazards from sustainable land use, to earthquakes and avalanches. In the details of these projects, there is however a large range of overlap where they may benefit from sharing data, particularly if experiments can be arranged to take place on common sites. Measurements such as meteorological parameters, soil temperature/conductivity/humidity and hydrological parameters are common across many projects and some projects even have synergies on much larger scales.

Scientific projects have in the past been very isolated, data has seldom been reused within departments, opportunities for data sharing within institutions are missed and collaboration across institutions has generally only taken place when the expertise did not exist in-house. E-science is changing this and Swiss-Experiment is one such e-science project. The Swiss Experiment collaboration will encourage data sharing and preservation of knowledge across projects and institutions through the use of a common, state-of-the-art databasing and data processing infrastructure. The addition of a spatially aware interface, combined with advanced querying tools is aimed at making scientists aware of what data exists and encouraging them to re-use data and/or collaborate on data acquisition. Through the re-use of data across projects, SwissEx aims to bridge the traditional scientific domains, broadening scientific knowledge on the interdisciplinary process interactions with the aim of eventually exploiting these links in large scale sensor deployments to improve environmental hazard forecasting and warning.

The same visual interface, utilising common tools such as spatial interpolation, is aimed at allowing scientists to easily try out various techniques on their data. Visualisation of the results on a map/digital topographic model allows scientists to better understand the relationship between the 2D results and the real processes that are occurring. This interface can also be used in publishing scientific results in an interactive electronic form, providing greater public interest and hence awareness of environmental research and the processes occurring in the environment around them.

II. Application Scenarios

The infrastructure is aimed at assisting throughout the life cycle of environmental monitoring. This is demonstrated by the following partially fictive application scenarios:

1. Planning: *Marc, a renowned hydrologist, is in his office and wishes to review existing datasets that have been captured in the past year at the Le Genepi field deployment*

* This research is supported in part by Microsoft External Research, the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322, and the Competence Center Environment and Sustainability of the ETH Domain (CCES).

[1] http://www.swiss-experiment.ch

of a wireless sensor network in order to plan the deployment for this year's campaign. He wants to better understand the interaction between the rock glacier and the atmosphere, in particular how the wind patterns drive the ventilation of the rock glacier. To do this he retrieves the data on rock and air temperature from the Swiss Experiment data repository and generates a visualization of the temperature differences on the SensorMap[1] interface. He is surprised by the large deviations at some locations and decides to concentrate more stations there. When visualizing the measurements of the rain gauges, he observes that they gave mostly uniform measurements and decides to reduce the number of rain gauge sensors. This year he also received new satellite data on temperature, accessible through Web Services, and uses a visualization of this data to decide on the placement of some stations at a larger scale surrounding the core area of the measurement campaign.

2. Monitoring the deployment: *The sensor stations have been deployed as planned. Over the Web he and his group can at any time observe the current measurements through SensorMap. One evening Marc receives a warning email generated by the underlying data stream processing middleware that the measurements of some sensors are out of the expected ranges. After inspecting some graphs of recent measurements Marc realizes that some sensors are malfunctioning and decides to go to the field next day by helicopter. In the field his team discovers that some of the wind sensors have frozen and they fix the problem. In order to maximize the benefit of the field trip they also visualize model data generated from the real-time measurements through a handheld device and use this information to optimize the placement of some stations. The metadata on the new positions and time of deplacement is immediately updated and fed back to the Swiss Experiment data repository so that later models are correctly computed.*

3. Analyzing the data: *After the campaign, as more stations have been placed in critical regions Marc can refine the resolution of his energy balance model. Simulating the models that are implemented in Matlab requires several hours of computation on his large workstation. As a result he obtains visualizations of the energy flows that can be overlayed in SensorMap. After looking at the map while sliding back and forth in time, and comparing it to the model results from last year's data, he realizes that he has to revise some assumptions of his models. He annotates the regions exhibiting strange behavior on the map. He will hand over his data and observations to a Postdoc who will be in charge of next years campaign. In the meantime the measurement data, the model data and annotations are archived in the Swiss Experiment repository. Browsing in the repository a PhD student in another research group discovers that she could apply her new risk model for landslides on Marc's energy flow model. Though not perfect data the surrounding communities are highly interested in these risk assessments and so she decides to make them available to selected decision makers through SensorMap.*

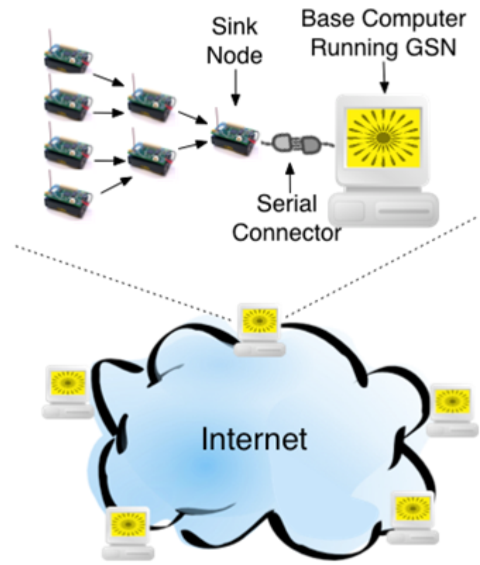

Fig. 1. A standard application of GSN (top) showing GSN gathering data from a sensor network via a serial forwarder that is itself connected to the network's sink node. GSN also offers the functionality of connecting several GSN instances to allow advanced query processing.

III. System Description

Our system comprises of two components: a sensor middleware component that handles the data acquisition and a data visualization and data sharing component. We will briefly review the fundamental concepts behind these technologies and then focus on the interactions between them and the challenges that arise in the integration process.

A. Data Acquisition: Global Sensor Networks (GSN)

GSN[2], [3] is a Java environment that runs on one or more computers composing the backbone of the acquisition network. A set of wrappers allow live data to be imported into the system. The data streams are processed according to XML specification files. The middleware system is built upon a concept of sensors (real sensors or virtual sensors - new data sources created by processing or repeating live data in software) that are connected together in order to build the required processing path (cf. Figure 1). For example, one can imagine an anemometer that would send its data into GSN through a wrapper (various wrappers are already available and writing new ones is quick), this data stream could then be sent to an averaging virtual sensor, the output of this virtual sensor could then be split and sent to a database for recording or to a visualization layer for displaying the average measured wind in real time.

GSN obtains the data directly from sensor network deployments and provides the capability of replaying previously measured data, for demonstration or exploration purposes.

B. Data Sharing and Exploration: SenseWeb/SensorMap

Once measurements about the physical world have been collected through GSN, it is advantageous to share the data, allowing multiple projects to share the instrumentation costs and deployment and maintenance effort. Sharing of large volumes of scientific data imposes challenges in data exploration techniques to efficiently discover a subset of data containing phenomena of interest to scientists. To tackle the challenges, an extensible infrastructure for data sharing (called *SenseWeb* [1]) has been designed as well as a map-based front-end (called *SensorMap*) to visually explore the shared datasets on geocentric interfaces such as maps and 3D terrain topographies.

The overall infrastructure allows scientists to share their sensor data acquisition systems over the common, programmable interface supported by SenseWeb, thus making the collected data available for researchers globally. Scientists share the sensors by adding their descriptions to SenseWeb. Such shared sensors then can be discovered based on location, type, or other characteristics. To efficiently support spatial queries of sensor metadata, SenseWeb indexes sensors by using a hierarchical triangular mesh (HTM) indexing scheme [4], which is particularly suitable for geographic queries.

SensorMap further enables scientists to explore the spatio-temporal distributions and correlations of the shared sensor data. SensorMap allows a user to directly specify the area of interest based on a browsable map, by drawing polygons or typing in geonames. Sensors within the specified geographical region are automatically aggregated at an appropriate granularity based on the zoom level of the map. SensorMap directly depicts the sensors on maps as image icons with different color schemes indicating the real-time readings.

Besides the real-time view, a user can explore sensor data streams in historic or spatial views. Via SensorMap, they can select a list of sensors of interest and visualize their temporal distributions in a single comparison chart or in multiple side-by-side time series charts. A third feature of SensorMap is to generate map/image-overlaid contours of selected sensors in view, which can be zoomed or panned together with the underlying map/image.

C. Integration

Figure 2 illustrates the architecture of the integrated system. At the bottom of the architecture are multiple GSN servers that acquire data from deployed sensors such as weather stations. Data streams collected by GSN are registered with the SenseWeb infrastructure to share among environmental scientists across multiple deployments. SensorMap accesses the shared data and visualizes their temporal and spatial correlations on top of maps and topological terrains.

One of the challenges in the integration process is to design a suitable communication protocol between GSN and SensorMap, on which we can pose the following requirements:

- It has to be pull based: many sensors have the capability of producing large data volumes and we want to minimize the communication between the GSN servers and the SensorMap servers. The protocol should be able to retrieve the data on demand when it is requested by a user (e.g., when somebody zooms-in on a specific point on the map).

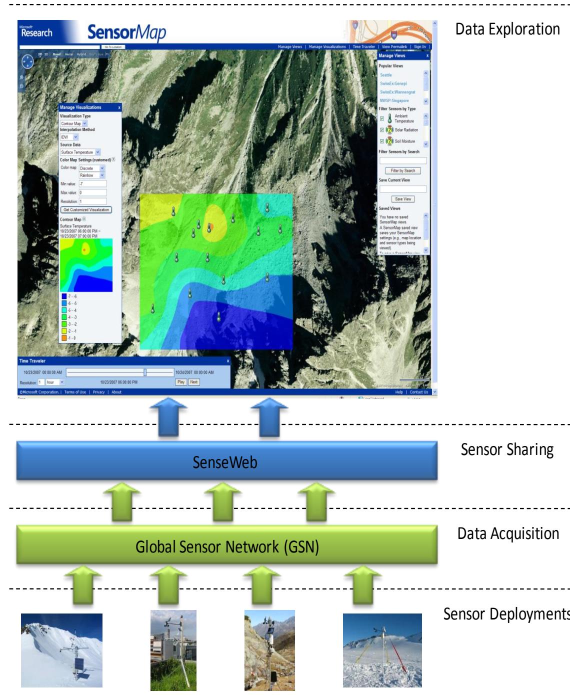

Fig. 2. Architecture of the integrated system

- It has to offer support for aggregation queries: being able to get aggregated values greatly reduces the communication traffic. SensorMap shows the high level picture of the data (e.g., aggregated every 6 or 12 hours) and once a user decides that they require a greater temporal resolution of data, SensorMap contacts the responsible GSN instance and asks for the high resolution data set.
- It should be location aware (GPS latitude and longitude)
- Its output should be machine parseable and preferably also human readable.
- It has to be simple enough to get adopted by a sufficiently large user community.

In order to address these requirements, the GeoRSS[2] standard was initially selected. GeoRSS is a geographically coded RSS output generated by GSN. One can specify the aggregation parameters in a simple REST request and retrieve the desired data stream. The output of GeoRSS is in XML format, making it convenient for other softwares to parse the output and produce their own visual interfaces over the GSN instances. This approach is what was used in version 2 of SensorMap.

The major shortfalls associated with a GeoRSS based solution are listed below:

- Extendibility: requests had to be modified to provide new parameters such as measurement types, sensor output rate, etc which are not part of GeoRSS.
- Interface: the solution has a push like interface for people who are interested in having a real-time view of a low traffic sensors (a value every few seconds).

In order to address these two issues, we decided to go

[2]http://www.georss.org/

to the Web Services interface which is flexible enough to handle both set of requirements. Version 3 of SensorMap is designed to use the new interface and GSN was also adopted accordingly. Using the new interface, the parameters are passed as method arguments (e.g., aggregating period, time range, etc) and since one can perform multiple calls over one Web services connection, the servers for SensorMap may maintain an open connection to their desired GSN servers in order to reduce the latency time. The push behavior can also be implemented using a simple call back interface.

IV. Demonstration Description

The SensorMap server is hosted at Microsoft Research in Redmond, both GSN instances are located in Switzerland, at EPFL in Lausanne and at SLF in Davos.

During the demonstration, participants will be able to view real-world streaming data obtained from sensor network deployments in the Swiss alps, focusing on the following two deployments.

The Le Genepi Deployment: The "Le Genepi" field deployment of a wireless sensor network was a campaign held between August and September 2007 conducted by the SensorScope team at EPFL. SensorScope [5] provides low cost, wireless and reliable sensor network systems for environmental monitoring to a wide community. It improves present data collection techniques with the latest technology, while meeting the requirements of the environmental scientists. The "Le Genepi" experiment was deployed on a glacier in the canton Valais (Switzerland) close to Martigny. In the three week experiment, 16 weather stations were deployed, measuring *air temperature, surface temperature, air humidity, wind direction, wind speed, precipitation and solar radiation.*

The Wannengrat Deployment: Above the town of Davos, Switzerland, at the Wannengrat alpine observatory, seven sensor stations have been installed for studying environmental processes involving snow. The project is maintained by environmental engineers from SLF in Davos. We use this installation as a valuable permanent test scenario for Global Sensor Networks (GSN). This scenario is significantly different from the SensorScope scenario, for instance the Genepi glacier experiment(as described above), since the stream data is inserted into the system as a periodic bulk import. The installation of GSN is at SLF in Davos and is connected to a GSN installation in Lausanne.

The demonstration will be based on these two real-world sensor streams, and will demonstrate the real-time and historic views, as Figure 3 depicts. Scientists can inspect real-time data as well as virtually go back in time to search for interesting events or analyze the temporal dependencies of the data. We further demonstrate contour visualizations of snapshot data of any selected time point, which helps users understand spatial correlations among dispersed measurements. Figure 4 illustrates one such contour plot overlaid over 3D terrain maps. Moreover, users are able to request an animation of contour plots for customized time durations and resolutions, through which a preliminary understanding of the spatio-temporal characteristics of selected data streams is obtained.

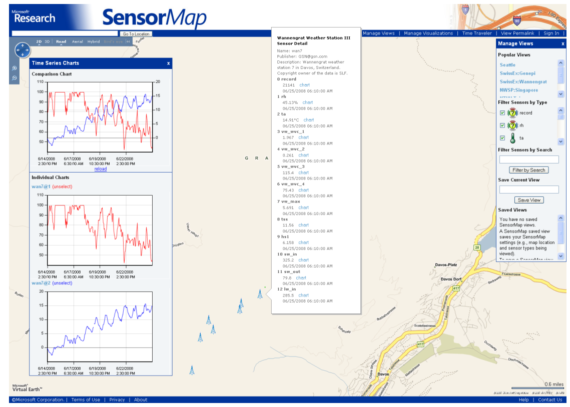

Fig. 3. Real-time and historic views of sensor data streams from the Wannengrat Deployment. Time series charts clearly illustrate the temporal correlations between humidity (red curve) and air temperature (blue curve).

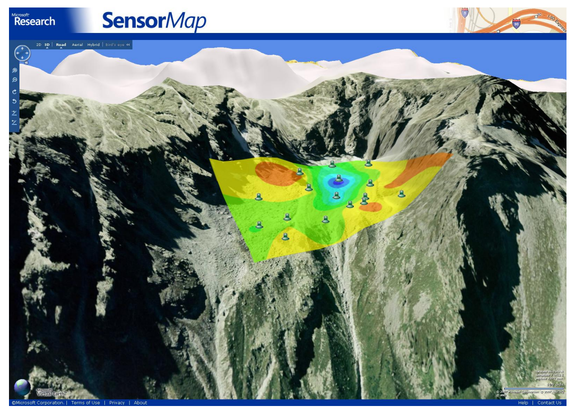

Fig. 4. Spatial visualization of ambient temperature measurements from the Le Genepi deployment. SensorMap generates this type of contour map on request and overlays it on top of 3D high resolution maps. Users clearly see the interaction between ambient temperature and the terrain.

V. Requirements for the Demonstration

Our described demonstration setup requires a standard broad-band Internet connection. We will bring our own laptops.

References

[1] S. Nath, J. Liu, and F. Zhao, "Sensormap for wide-area sensor webs," *IEEE Computer*, vol. 40, no. 7, pp. 90–93, 2008.

[2] K. Aberer, M. Hauswirth, and A. Salehi, "Infrastructure for data processing in large-scale interconnected sensor networks," in *MDM*, 2007, pp. 198–205.

[3] ——, "A middleware for fast and flexible sensor network deployment," in *VLDB*, 2006, pp. 1199–1202.

[4] A. Szalay, J. Gray, G. Fekete, P. Kunszt, P. Kukol, and A. Thakar, "Indexing the sphere with the hierarchical triangular mesh," in *MSR-TR-2005-123*, September 2005.

[5] G. Barrenetxea, F. Ingelrest, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, "Sensorscope: Out-of-the-box environmental monitoring," in *IPSN*, 2008, pp. 332–343.