ESCIENCE CENTER
UNIVERSITY OF COPENHAGEN

MANA Project

http:://mana.escience.dk/

Ph. Bonnet (bonnet@diku.dk), Marcus Chang, Kirsten Christoffersen

Abstract
MANA is a research project whose goal is to is to improve scientific data acquisition in polar regions. The aim of the MANA project is to develop a new generation of data logger adapted to extreme weather conditions, and limited bandwidth that are characteristics of harsh, remote environments. The innovation we propose consists in (a) leveraging wireless sensor networking technology to connect densely deployed sensors to the data logger, and (b) designing an autonomous data logger that controls the quality of the collected data, and that adjusts its sampling strategy to improve the relevance of the collected data while optimizing resource utilization.

Project Structure

Mana is a collaboration between the computers science department and the fresh water biology lab at University of Copenhagen, the School of Computing at Reykjavik University, Arch Rock corp and Dan-systems.

The project is funded by the Strategic growth technologies program from the Danish Strategic Research Council.

The MANA project started on Feb 1st 2008. It will last three years. Our first deployment is planned for August 2008.

The Goals

The overall goal of the MANA project is to improve scientific data acquisition in remote, harsh environments such as the arctic. Because physical access and communication bandwidth are limited, manual measurements are costly, manually tapped data loggers are unreliable, and remote supervised control is impractical.

In MANA, we focus on the monitoring of limnic parameters in the Zackenberg region, North- East Greenland. The goal is to document the effects of climate change on lake environment, in particular in the winter season that has been neglected so far because of logistics constraints.

In high arctic Greenland, the limnic environment is highly sensitive to climate changes. Many aspects are however poorly known, and especially the winter season has previously been neglected due to logistic constraints. At Zackenberg, BioBasis monitors the water chemistry and occurrence of freshwater biota in two small lakes during summer. These measurements, however, do not adequately describe the processes in the lakes, especially not in the periods when the lakes are covered with ice. We thus aim at increasing the temporal resolution of the measurements.

In the context of the Biobasis program, measurements are performed either manually, or using simple data loggers that store the raw measurements they obtain from one or several sensors. In particular TinyTags are deployed to monitor abiotic parameters. The measurements are tapped from the data loggers once a year. The collected data is then transferred to a computer to be postprocessed and analyzed. Post processing consists in time stamping the data, and identifying potential anomalies due to sensor errors, sensor drift, cable problems, or hardware/software problems on the data logger. The biological variables of interest are then derived from these time series. We aim at ensuring that the raw measurements are most useful for deriving the biological variables of interest. This is why our data loggers will (1) check the collected data, and (2) autonomously adapt their sampling strategy to optimize the quality of the raw measurements. In this sense, our data loggers constitute a significant advance with respect to the most advanced data loggers on the market, and a step towards the autonomous experimentation platforms mentioned in the Science 2020 report.

We aim at enhancing sensors and data loggers with computation and communication capabilities so that we can program them to be reliable and autonomous. We plan to develop sensor network-based data loggers that (a) check the data they collect and correlate measurements in time and space, and (b) autonomously adapt their sampling strategy in order to optimize data quality as well as ressource utilization.

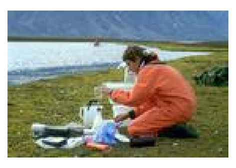
The Challenges

The key challenge when designing a data acquisition system for year-round lake monitoring in a high-Arctic environment is to take measure of the extreme weather conditions, specially in winter (September-July) where the wind is blowing, temperature reaches -40C and the lakes are covered with an ice lid of approximately 2 m as well as a layer of snow.

A wire between sensors inside the lake and a data logger located on the shore would be exposed to the forces that are applied on the ice forming at the surface of the lake (during freeze-thaw periods or when the wind is blowing).

A popular option in the context of ocean and coastal waters monitoring is to attach a sensors to a buoy that integrates a data logger, as well as communication capabilities. We investigate a slight variation of this approach, where one data logger located on the shore is connected via mid-range wireless links to a collection of buoys, each equipped with various sensors.

We believe that our wireless sensor network-based approach is more flexible, as each sensor is transformed into an Internet device. Our approach is also more cost effective for dense sensor deployments, as several sensor nodes can be arranged in a multi-hop network rooted at the data logger e.g., with several multi-sensor probes per lake, or dozens of soil moisture sensors deployed around a weather station.



Vexcel Microserver

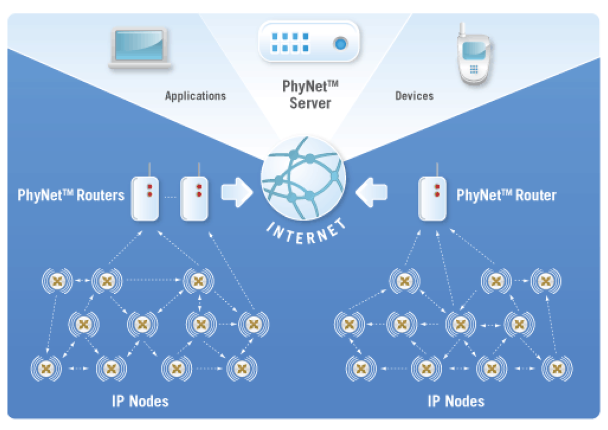
The Technology

We base our data acquisition system on commercial off-the-shelf components:

• Multi-sensors probe: Multi-sensors probe provide a uniform interface for several co-located sensors. Each sensor may have a different modality (e.g., chlorophyll, oxygen, temperature sensors can be attached to the same multi-sensors probe). The key challenge when transitioning from manual to unattended sensors is to deal with the biological and chemical debris that accumulate in time. This is particularly important for optical sensors such as chlorophyll sensors. We are relying on the Water Quality Monitor (WQM) from Wetlabs. It includes anti fouling solutions and wipers that clean the sensitive surface of optical sensors and thus allow long-term monitoring.

WQM from Wetlabs

• Sensor Networking: We rely on the sensor networking technology from Arch Rock: (1) each multi-sensor probe is connected to an Arch Rock Sensor Node located inside a buoy, (2) the border router is co-located with the data logger on the shore, and (3) an external antenna is used so that the mote can communicate via 6lowPan with the flash-based PC located on the shore.



Arch Rock Sensor Networking Technology

• Low-power, flash-based PC: This is the low-power server connected to the sensors via the Arch Rock border router. It is the hardware component on which the data logger executes.
The key characteristic of this component are (i) that it can be aggressively duty cycled, and (ii) that its secondary storage is flash-based. Indeed, hard drives cannot operate in the extremely low temperatures of a high-Arctic environment.We are using the Vexcel Microservers, developed in the context of the SEAMONSTER project.

Autonomous Data Logger

The problem with traditional data loggers is twofold. First, they implement a best effort approach: raw measurements are stored as they are received. Scientists need to clean the collected data to find out about outliers, missing values, or mis-calibration. At that point, the best scientists can hope for is to interpolate the valid data points in order to compensate for the unusable data points.

Our idea is that the data logger should check the data it collects, and take compensating actions when necessary. This way, the data logger can control the quality of the data it collects.

Second, traditional data loggers implement a static sampling strategy, i.e., the timing (sampling rate) and modality of the sensor measurements are fixed and remain unchanged throughout the data collection campaign. A fixed sampling rate is either set high enough to capture all interesting events -- in which case too much energy is spent in the phases where the monitored system behaves as expected, or it is set too low -- in which case some interesting events will not be captured. Our idea is to let the data logger decide on whether to reduce the sampling rate to save energy (e.g., the data logger might choose to reduce the sampling rate of the temperature sensors if it can predict the measured temperature), or to improve the sampling rate when it has detected an anomaly in the measurements (e.g., if chlorophyll concentration typically remains constant during the winter season then a slight increase should trigger an increase of the sampling rate to enhance the quantity of data points related to this potentially interesting phenomenon).

The data logger should check the data, and take actions *autonomously*. We adopt the classical three-tier architecture for our autonomous systems.
(1) The functional layer interfaces with the system functions, i.e., the functionality provided by the Arch Rock border router, as well as the flash-based secondary storage.
(2) The execution layer implements the quality check and anomaly detection functions. The traditional technique to check the data collected by a sensor is to monitor either changes in the mean value of the measurements, or spikes that exceed some threshold value. Such systems control each sensor separately. An improvement consists in correlating the measurements from several (non-redundant) sensors. These techniques, commonplace in remote sensing,have so far never been used in the context of in-situ data loggers.
(3) The deliberative layer implements the planner/ scheduler. Our data acquisition system is relatively simple (compared to fleet of earth observing satellites or a Mars rover). The key challenge is to leverate the limited ressources (energy, computation, storage).