1. Aims of the Project
Implementation of the Enhanced Resource Aware clustering algorithm into Sun SPOT Distributed Query Processor.
- Generic resource monitor
- Framework for stream processing
- Adapting to resource availability

2. Introduction
This project is part of the research conducted by the School of IT on data processing in wireless sensor network. The objective for this project is to develop an extension to the Sun SPOT Distributed Query Processor (SSDQP) in order to evaluate the performance of the Enhanced Resource Aware (ERA) clustering algorithm in life operation of a typical wireless sensor network.

In implementing the ERA clustering algorithm, a number of supporting infrastructures have been developed. These supporting infrastructures provide the capability to sense available resources (CPU utilisation, free memory and remaining battery level) as well as to retrieve the instantaneous state of an ERA cluster.

3. Background
A wireless sensor network (WSN) is a network of wirelessly connected, battery operated embedded devices equipped with one or more sensors. A WSN is designed to operate without operator intervention.

Radio communication is the most energy-expensive operation on a sensor node. In order to reduce the amount of data transmitting across the network, the ERA clustering algorithm is used to summarise sensor data.

SSDQP provides a SQL-like language for data abstraction, running on WSN nodes and the host computer. Other features include in-network processing, GUI and multi-tasking support.

4. Platform
The research prototype runs on the novel Sun Microsystems\textsuperscript{TM} Small Programmable Object Technologies (Sun SPOT) platform.

Sun SPOT nodes run the Squawk Java Virtual Machine, a small virtual machine specially designed for resource-limited devices which allows rapid prototyping.

5. System Overview
The extension introduces the design concept of implementing a stream processing operator as a data container, allowing its instantaneous state be retrievable by an external query. Each element of the data stream is processed as they are added to the container.

To support this concept, a generic storage point (similar to variables) was developed to store multiple containers. In this project 2 types of data containers have been implemented – buffering and clustering. This storage point is generic such that new containers can easily be used. This allows new stream processing operators be implemented easily.

A generic resource monitor using the standard publisher-subscriber interface was also developed. It is comprised of two components – a recurring query that retrieves the current resource level, and a resource table that is modified by the query and notifies any subscriber of changes to the values.

SSDQP’s sensing operations have also been extended to enable resources to be determined through “internal sensors”. The sensing operation was also overlaid to allow queries to access data stored in the storage points.

6. ERA Clustering Algorithm
The ERA clustering algorithm produces a summary of the measurements by grouping the values into different ranges. Only the mean of the attributes in each cluster and their count are returned.

Measurements are allocated a cluster by locating the cluster whose mean is closest to the measured value and is within a specified radius. N-dimensional distance is obtained using Manhattan distance. If no such cluster is found, a new cluster is created.

The algorithm adapts to low memory level by increasing the cluster radius and removing outlier and unused clusters. High CPU load is adapted by searching a random subset of clusters to fi nd the closest node instead of searching the entire set.

The results from the simulation of this algorithm presented by Phung et al in their paper are included below:

7. Evaluation
Code correctness was evaluated through the use of unit testing, argument for correctness and integration testing.

Tests have shown that the modules added/modified in this project are operating as expected.

8. Conclusions
- Implementation of the ERA clustering algorithm into SSDQP.
- Development of a framework for future stream processing algorithms.
- Implementation of generic mechanism for resource monitoring.
- Future work includes the upgrade of GUI, clustering algorithm and porting to newer hardware (eSPOTs).

The University of Sydney