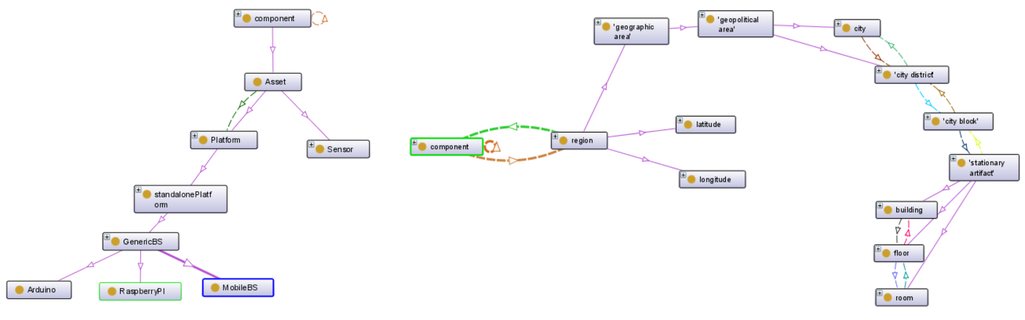
Ph.D. Research Activity in Energetic, Systems, Computer and Telecommunications Engineering

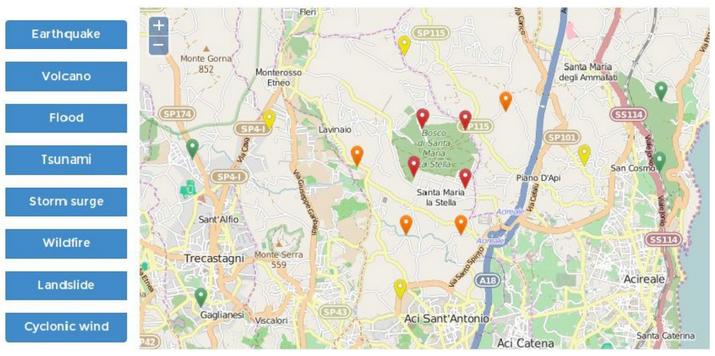
Ph.D. Student: Marco Cavallo Tutor: prof. Vincenzo Catania
DIEEI – University of Catania
Joint Open Lab WAVE

A semantic model for the geographical characterization of sensor networks

Context

The widespread availability of **IoT** technologies allows detailed monitoring of environmental parameters that can be used for effective prevention and forecasts of natural disasters. When such events occur the primary concern of authorities responsible for the **public safety** is to organise and enforce the relief in an as quick as possible and effective way.
However, the huge amount of information that an IoT infrastructure is able to provide is useless if not correlated and contextualised on the territory where the hazard has occurred.





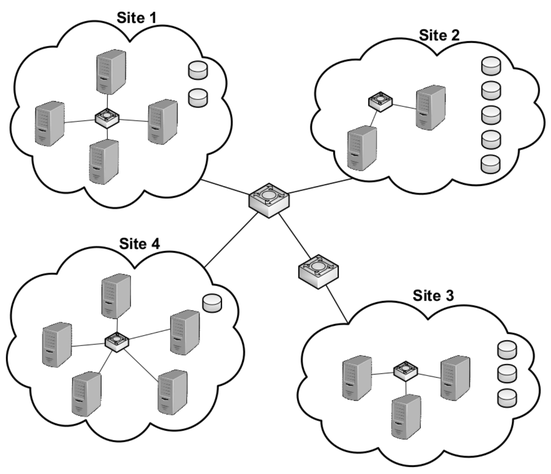
Proposal

A semantic approach allows to create a common language for the representation, classification and description of any device, able to fetch data about the surrounding environment.
The designed **Semantic Information Model** integrates different information domains into a unified framework. The resulting knowledge base can be used to build monitoring services capable of identifying on the territory those sensing units whose sensed data are really useful for the set-up of a relief plan.

Hierarchical Map Reduce for geo-distributed big data

Context

The diffusion of information technology services (social computing applications or smart city services) produces a significant increase of the amount of digital data, distributed among data centers located in different geographical locations. **MapReduce** is probably the most known distributed programming model used in cloud computing for large-scale data analysis applications. **Hadoop** (open-source implementation of the MapReduce model) assumes that every node in a cluster has the same computing capacity and that data are local to tasks.
Standard MapReduce implementation performance is low in **geo-distributed** environments.

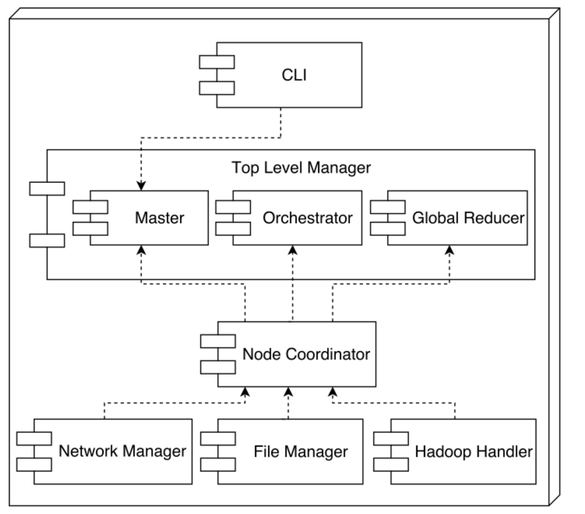


Proposal

This research project proposes a **hierarchical** MapReduce programming model where a top-level scheduling system is aware of the underlying computing **contexts heterogeneity**. System can monitor the distributed context's **resources**. Each site periodically advertises its overall computing capacity to the Orchestrator.
SDN-enabled switches are capable of measuring the instant **bandwidth** occupied and retrieve information about allocated bandwidth.

H2F: Hierarchical Hadoop Framework

The implemented solution is based on a **modular** architecture, made up of several modules that are in charge of the steps of the job's execution flow.
The framework enforces a **scheduling strategy** by crossing information concerning the submitted job and the execution context in the form of Sites' overall available computing capacity and inter-Site bandwidth capacity.
The objective is to elaborate an execution plan that ensures the highest **job performance** in terms of completion time.



Results