WindS@UP: The e-Science Platform for WindScanner.eu

Filipe Gomes[1] **& João Correia Lopes**[2] **& José Palma**[1] **& Luís Frölén Ribeiro**[3]

[1] FEUP, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

[2] INESC TEC, DEI, FEUP, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

[3] IPB, Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

jlopes@fe.up.pt

Afiliação (1,3)

Abstract

The WindScanner e-Science platform architecture and the underlying premises are discussed. It is a collaborative platform that will provide a repository for experimental data and metadata. Additional data processing capabilities will be incorporated thus enabling *in-situ* data processing. Every resource in the platform is identified by a Uniform Resource Identifier (URI), enabling an unequivocally identification of the field(s) campaign(s) data sets and metadata associated with the data set or experience. This feature will allow the validation of field experiment results and conclusions as all managed resources will be linked. A centralised node (Hub) will aggregate the contributions of 6 to 8 local nodes from EC countries and will manage the access of 3 types of users: data-curator, data provider and researcher. This architecture was designed to ensure consistent and efficient research data access and preservation, and exploitation of new research opportunities provided by having this "Collaborative Data Infrastructure". The prototype platform—WindS@UP—enables the usage of the platform by humans *via* a Web interface or by machines using an internal API (Application Programming Interface). Future work will improve the vocabulary ("application profile") used to describe the resources managed by the platform.

Introduction

The whole WindScanner system recurs to three WindScanner units synced and coordinated to measure the flow field along a certain path in the measuring volume. This is the core technology of the WindScanner infrastructure[1].

The *WindScanner.eu*—The European WindScanner Facility, is an ESFRI project, under the FP7-Infrastructures-2012-1; it is a mobile, distributed, facility with 6 to 8 nodes in European countries. The project aims to be an Open Access distributed research infrastructure promoting the dissemination of results including innovation products and their exploitation [2].

WindScanners are deployed at existing or planned test facilities, covering different climate conditions and terrains. Each *Campaign Site* handles low-level communication with several WindScanner units at the measurement site.

The WindScanner.eu e-Science Platform

The platform has a centralised architecture where all data migrates to the Hub node that provides data storage and computing capabilities.

The data is collected at measuring campaign by the windscanners and other wind energy related instrumentation such as, for instance, tower mounted cup or sonic anemometers. The data is transferred to a local node, local/regional manager of the windscanner facility, who will upload the whole experiment information to the central node.

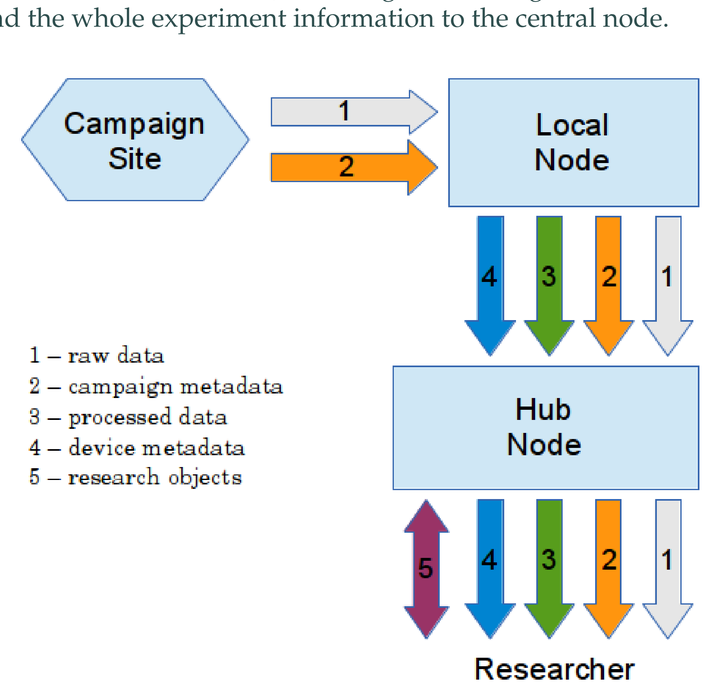

Figure 1: Data flow between nodes

- Multiple *Local Nodes*, distributed among countries or institutions, collect data from its affiliated *Campaign Sites*, ~~that is, campaigns carried out by a given institution researchers.~~
- A Quality Assurance (QA) process is performed ~~in order to validate the~~ received data.
- ~~Depending on the campaign, the processing of raw data to produce "clean data" may be reuired. These data are similar to the raw data—a time series—while some portions may be purged due to detected errors, or some mathematical operations may be performed.~~
- Usually, researchers will ~~use~~ these data rather than the raw data.
- Further documentation ~~should~~ be added, for instance, detailing the clean data.
- The *Hub Node*—the proposed e-Science platform itself—receives these data and metadata, making them available for other researchers.
- Resorting to these data, the researches carry on their research, producing *Research Objects*.

- validation of
may
i.e., a pdf or png files

Platform Architecture and Implementation

The collaborative platform provides:

- A central repository for experimental data ~~(data set with three simultaneous measurements of the wind vector's three components)~~ and metadata (e.g. the experimental setup)
- Data processing capabilities ~~are to be incorporated th~~us enabling *in-situ* data processing
- A Uniform Resource Identifier (URI) to identify every resource in the platform~~, enabling researchers to cite them from outside the platform.~~

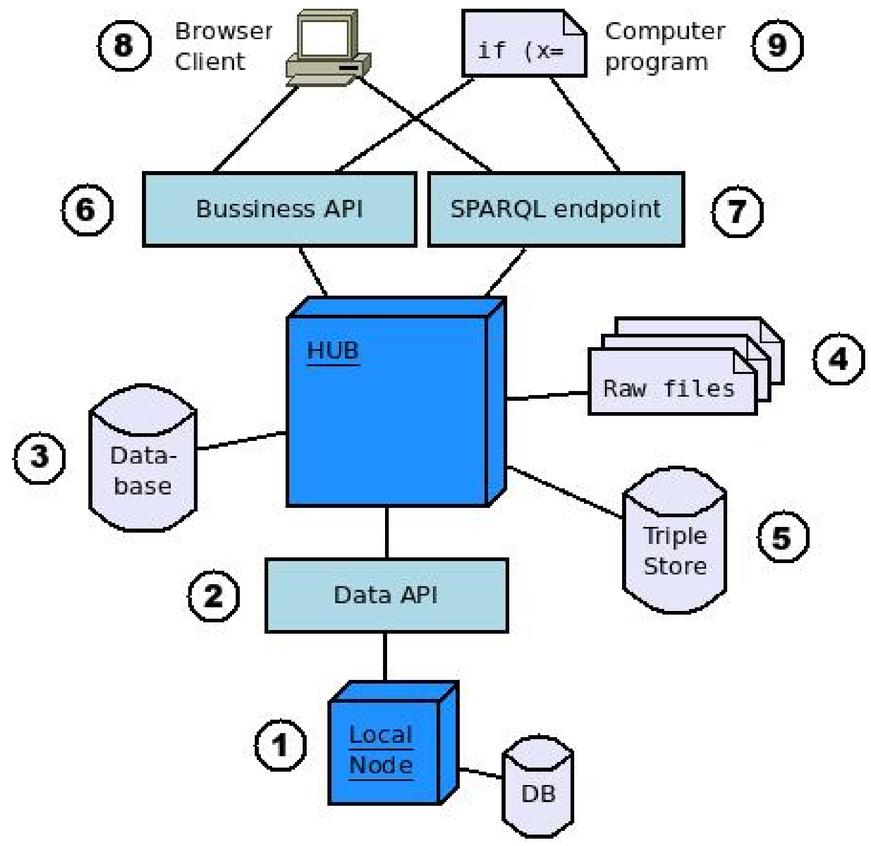

Figure 2: Platform Architecture

The platform provides an Application Programming Interface (API) to interact with most of its components (see Figure 2). This enables the development of 3rd party applications that garner the capabilities of the e-Science platform. The API can be divided in three areas targeted at: data upload (**n. 2**), Business logic common operations (**n. 6**) and access to the Knowledge base comprising the semantic Web resource descriptions (**n. 7**). Machine programs have broad access to the platform features through the API (**n. 9**) and the Web interface we developed uses the API likewise (**n. 8**).

Conclusions

- Researchers' needs are provided by an e-Science platform where they may collect data sets with campaign measurements, run their computations against existing data to produce new data sets, upload other research objects (such as research papers in PDF), describe any research object using a controlled vocabulary (metadata), make comments on any research object and upload procedures to be run against data sets.
- The procedures for data quality will be prior to the upload of data (and metadata) to the platform to certify the data to an agreed format before storing it for long term persistence.
- The e-Science platform, which stores and provides access to researchers to data from multiple campaigns from several sites, provides elastic storage to accommodate the growing amount of data originated from the equipment and also the data generated by the researches themselves.
- The access to the platform to upload of the various resources (data sets, research objects, etc.) is facilitated through a set of Web Services, with a documented API, that enforces the verification of access privileges and interoperable data formats.

Forthcoming Research

- In order to accomplish its purpose, the e-science platform must scale to accommodate the ever-growing experimental and derived data and the users' computing needs.
- An internal API to isolate the platform from storage and computational needs provides by a cloud layer has been designed and it is being prototyped at the moment.

References

[1] Torben Mikkelsen, Søren Siggaard Knudsen, Mikael Sjöholm, Nikolas Angelou, and Anders Tegtmeier. WindScanner.eu—a new Remote Sensing Research Infrastructure for On- and Offshore Wind Energy. In *International Conference on Wind Energy: Materials, Engineering, and Policies (Hyderabad, India, Nov 22-23)*, 2012.

[2] WindScanner.eu. The European WindScanner Facility: Annex I—Description of Work. Technical report, Seventh Framework Programme, 2012.

Acknowledgements

The *WindScanner.eu*—The European WindScanner Facility—is an ESFRI project (N: 312372) under the FP7-Infrastructures-2012-1. The authors are grateful to all colleagues in WP5 for the fruitful discussions, namely Dimitri Foussekis (CRES), Doron Callies (IWES Fraunhofer), Hans Verhoef (ECN), Harald Svendsen (Sintef), Jan Willem Wagenaar (ECN), Javier Sanz Rodrigo (CENER), Martin Bitter (Forwind), Mikael Sjöholm (DTU), Steen Arne Sørensen (DTU) and Teresa Simões (LNEG).