

CROSS-Fire: a risk management decision support system on the Grid

António Pina¹, Bruno Oliveira¹, Joel Puga¹, António Esteves¹, Alberto Proença¹

¹*Departamento de Informática da Universidade do Minho
{pina ,boliveira ,joel,esteves, aproenca}@di.uminho.pt*

Abstract

The CROSS-Fire project aims to develop a grid-based risk management decision support system for the Civil Protection (CP) authorities, using forest fires as the main case study and FireStation as a standalone CAD application to simulate the fire spread over complex topography.

CROSS-Fire approach is based in an architecture that includes: information models, encodings, and metadata that represent the scientific knowledge associated to FireStation execution models and standards to enable the discovery and access of Web services, data repository, sensor networks and data processing facilities.

To achieve the desired integration of information and services we use: i) EGEE to provide raw technological capability provision, including data management and storage, access to meta-data data bases and high-performance computing and ii) a Geospatial Information Infrastructure based on OGC-WS and SWE Web services to provide the access and management of remote geospatial data and virtualized sensor networks.

This article, stresses the relevance of standards adoption of OGC-WS by describing the work that is been done to provide G-FireStation with: i) a standard-based SDI layer, based on Geoserver to exploit/enable geospatial services for data access/processing and ii) a 52N's implementation of a OGC-SWE compatible layer, to address sensors CP data sources, such as meteorological stations data and satellite images and iii) the development of G-FireStation graphical user interface to access the platform facilities.

The core of the CROSS-Fire Platform is a WPS 52North OGC standard layer divided into three interoperable components, respectively, the CROSS-Fire Business Logic, the Grid Services and Geospatial Services. WPS serves as an interface to a wide range of distributed computing resources provides the mechanism to access the grid facilities for processing and data management and including all the algorithms, calculation, or model that operates on spatially referenced data, also mediating all the communication with the portal and other GUI clients.

The G-FireStation user interface that is currently under development is an open-source desktop with GIS and CAD capabilities that exploits an SDI client complying with OGC-WS and EU INSPIRE directives. It provides facilities to locate and access the spatial data infrastructure and to visualise the fire propagation, based on the native facilities of gvSig, it was also extended to support a OGC WPS client that mediate all the interactions with the core WPS service layer.

1. Introduction

The Control of forest fires is a relevant Civil Protection (CP) activity that involves many different and autonomous actors, from public bodies to research centres. It requires a fast and reliable risk management support system, with real-time or near real-time availability of critical geo-referenced data and settings-based forecasts for fire spreading.

CP applications require a strict integration of human and physical resources that must be shared in a coordinated and effective way, and available for the whole emergency procedure.

The GRID and Virtual Organizations (VO) enables such integration by providing the coordination and the sharing of the available interconnected resources (computing, storage, communication, sensors and actuators) geographically scattered across national borders.

In another hand, OGC based geo-web services are being adopted worldwide, as the technology to support the development of complex distributed applications over grid platforms, to deal with data from many different sources, including meteorological station and satellites. Recent work clearly showed the advantage of the OGC proposals for open standards for geospatial interchange formats, over past legacy formats and applications.

1.1 CROSS-Fire Approach

The development of a CP application requires not only the availability of high-performance computing resources and data management at remote Grid sites, along with other requirements such as data interoperability, and user authentication and authorization, but also the ability to access, to integrate, to analyze and present geospatial, available data repositories and sensor networks data across a distributed computing environment.

To achieve all those goals, the software development approach follows and adheres to most of the latest normalization initiatives such as EU/INSPIRES and OGC-WS: OGC-WS WCS/WFS/WMS/WPS, to exploit/enable geospatial services for data access processing, and OGC-SWE SOS to address other CP data sources, such as meteorological station networks (MSN) or satellites.

The overall software development is made of several components: client application, which request geo-referenced data and fire spread simulation, Spatial Data Infrastructures (SDI), which provide geo-referenced data, and the GRID, which gives support to the computational and data storage requirements.

To give support to each of the above component, and using afore mentioned norms, a web services layer was created, based on WPS. The use of web services allows different components to interact with each other in a standardised mode, acting each one as clients of the web services layer. It also allows the connection of heterogeneous clients, as long as they comply with the standards.

1.3 FireStation

The CROSS-Fire project aims to develop a grid-based risk management decision support system, using the EGEE infrastructure, for the Civil Protection (CP) authorities, using forest fires as the main case study and FireStation (FS) as a standalone CAD application to simulate the fire spread over complex topography [1].

FS integrates a module for wind field generation, as well as a module for the computation of the Fire Weather Index (FWI), from the Canadian System. The proprietary software includes built-in subroutines for the design of window-based interfaces and generation of visualization elements, both in 2D and 3D space.

FS needs three different kinds of input data to simulate fire propagation: i) the terrain divided into cells, each one characterized by its altitude and fuel type, complying to the BEHAVE model [2], ii) the wind conditions, at mid-flame high, affecting that terrain and ii) some control parameters, such as the ignition points and the stopping simulation criteria (see figure 1).

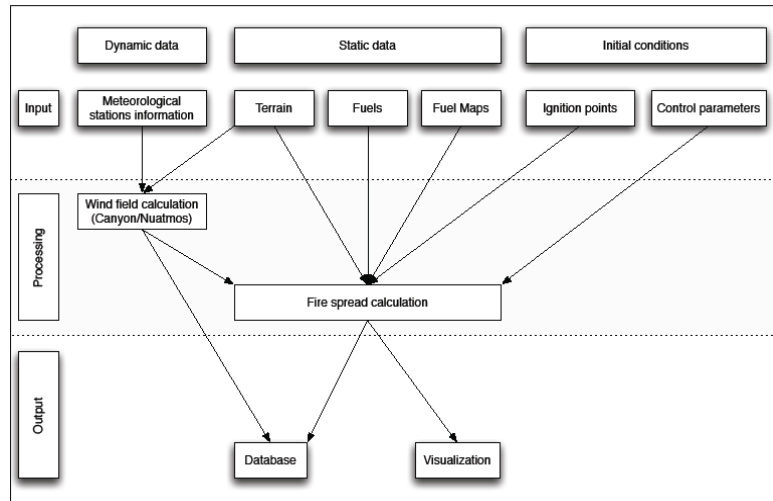


Figure 1. Components of FS Simulation model

The information about the wind conditions affecting the terrain is previously generated by a Wind Field Module. The simulation stopping criteria can be the maximum simulation time, the maximum area burned or the maximum number of cells burned.

1.3. Previous Work

The CROSS-Fire initial tasks focussed on the development of: i) P-FireStation, a parallel version of the fire simulator engine and ii) G-FireStation the FS porting into the EGEE grid environment, to support higher processing/storage capabilities, improved I/O data resolution/models complexity, faster multi-simulation execution and wider simulation areas.

The P-FireStation version explores the inherent parallel environment offered by clusters at each site of the EGEE grid to support larger data sets and to improve the accuracy of the predictions. This parallel version relies on the MPI protocol and supports larger data sets taking advantage of the MPI parallel I/O facilities [3].

The G-FireStation version integrates EGEE grid facilities, namely the gLite data management services and tools to access data, the AMGA[4] grid metadata catalogue to manage the simulations meta-data, and the WatchDog [5] tool to monitor and provide data for the interactive execution of simulations [6].

2. Cross-Fire Platform

The CROSS-Fire platform is composed of a central core, a WPS layer, and two external infrastructures: a SDI platform and the GRID.

The core WPS is divided in three parts: Business Logic, Grid Services and Geospatial services. The Business Logic is an abstract layer configured to handle the specific algorithms that provide all the functionality of FireStation, namely forest fire propagation, wind field calculation and fire weather index.

The Grid Services is the component that interfaces with the GRID infra-structure. Amongst its responsibilities one can find proxy delegation, job creation and management, and data movement to and from the GRID.

The Geospatial Services act as an interface between the clients who request geo-referenced data and the available collection of SDIs. It represents single entry point to query data, managing on behalf of the clients the geo-referenced data sources and performing queries on all known SDIs, returning a single, unified, result set, from which users can select the desired data.

3. Web services layer

Web services are one of the fundamental layers of the CROSS-Fire platform, implemented as a set of WPS algorithms that deal with most of the functionalities of the three components of the platform. Figure 2 show all the components of the Cross-Fire platform, the SDI and Grid infrastructures on the top, the WPS layer one the centre and a client layer at the bottom; the white rectangles details the algorithms/services relevant to the presentation of the platform.

The WPS is been developed as a general WPS server protocol implementation, on top of WPS 52 North. In the following sections we will describe most of its algorithms.

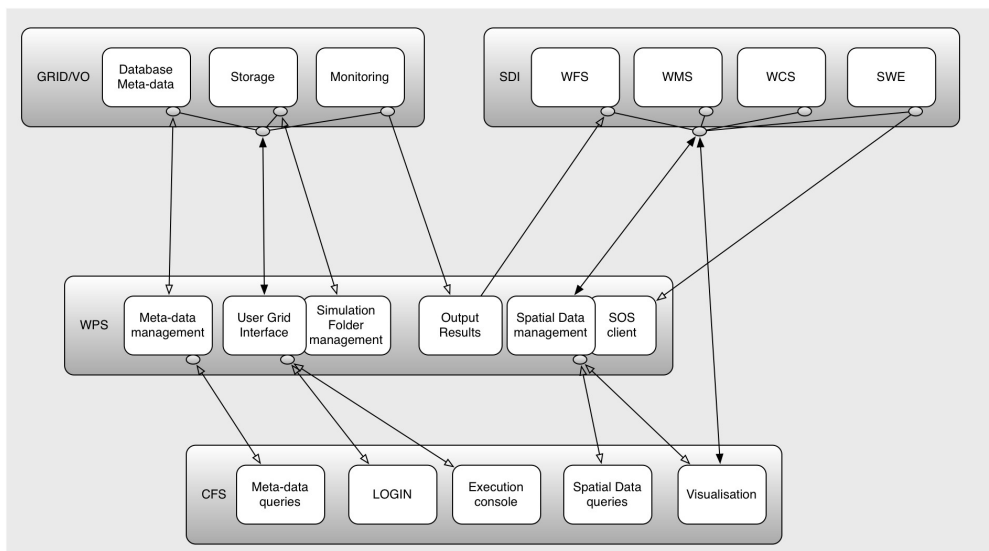


Figure 2 The CROSS-Fire platform

3.1. Security

As expected, user authentication and authorization are a very important matter when dealing with a CP application, when referring to both the access to the resources available through the GRID infra-structure and the geo-referenced data stored in the SDIs repositories. As so, before looking in further in the platform itself, we present the security approach we follow when designing the platform.

In EGEE, access to the GRID is always associated with a X.509 certificate used to identify and authorize every user when accessing resources in the infra-structure. The same metaphor may also be used to secure the access to sensible geo-referenced data, within the SDI infra-structure, solving the security problem with a single solution.

The WPS layer must then act on behalf of any authenticated users, using a method called delegation, in order to be able to access and use the GRID/SDI services. By delegation we mean the user's certificated, or a valid proxy created with this same certificate, that can be used by an application on his/her behalf.

Actually, the platform uses the library GridSite to implement authentication and delegation. The library provides both an Apache module that grants X.509 based security to https calls, to use when dealing with OGC web protocols, and a delegation library that can be used when submitting job or managing data in the GRID infra-structure, on behalf of the clients and users.

3.2 Meta-data management

In CROSS-Fire one expects not only to be able to run simulations as a component of the a risk management decision system but also to access and manipulate the results produced in the past. To accomplish this requirements the platform must be able to register in a convenient way all of the input data used in the simulation, as well as its outputs.

To achieve those goals, all the simulation input/output data along with the meta-data that characterize the executions, such as the date or the execution time or the resources used is recorded and accessed though specially designed data-base. This data-base is supported by a Postgresql data-base server with GIS extension, implemented on top of PostGis.

In order to have these data available on the GRID and, at the same time, to have access to the replication, redundancy and security facilities offered by the GRID, the data-based is interfaced via AMGA, which allows a uniform view over a set of data-bases under its management, regardless of the underlying data-base server architecture.

3.3 User Grid Interface

This algorithm is used whenever a user needs to run a simulation or some data is made available on the GRID. User interacts with this algorithm on the assumption that he/her already posses as valid proxy to the GRI/SDI infra-structures. Functionality includes the management of GRID jobs, along with creation of the JDL files, the job submission and the execution control and status.

To launch a simulation, the user identifies a set of geo-referenced files within the SDI, along with a set of control parameters, such as ignition points, time and date, and meteorological conditions, obtained both statically or dynamically from the SDI.

To ease the implementation of this algorithm we are currently profiting from the possibilities of the GANGA library to manage job submissions. The library may also be used to manage different types of security schemes and other GRID back-ends (not only gLite).

4. CFS – Client

The desktop version of FS offers a nice and user friendly CAD environment that we decide to keep as close as possible to the original after gridification and augmenting it with facilities to handle geo-referenced data and to interface with both the GRID and the SDI infrastructures.

Currently CFS (console FireStation) is based on gvSIG an Open Source GIS full feature desktop solution, funded by EC, under active development and with a solid user base, that conforms to requirements of the INSPIRE framework for managing geospatial information.

It has support for multiple file formats, DBMS and web protocols, and includes a SDI client that permits the connection, through the use of standards, to OGC Services like OGC WMS, OGC WFS, and OGC WCS, accessing data and being able to overlap it and combine it in gvSIG map views. Discovery service client is also provided within gvSIG which can be used to localize data resources within a SDI.

4.1 Execution console

This module establishes the connection between users and the GRID, being responsible to negotiate with the WPS layer the user proxy generation and delegation. It also organizes and collect all the data needed to submit to the grid the jobs that represent all the execution models that integrate the overall CROSS-Fire Business Logic.

Any other interactions with GRID as it is the case of the access to the Data Services, for instance, to retrieve a file from a Storage Element is also managed by this module.

4.2 Visualization

To be able to interact with the WPS server layer of the CROSS-Fire Platform, more recently, we add to CFS a WPS client taking advantage of the plug-in facilities of gvSig that is been used for built several other modules, including the one used to visualise fire spread simulations.

This module provides a GUI to discover and obtain the information produced by past executed simulations and visualize it as an animation of the fire spread. Users are presented with a slider corresponding to the total time of the simulation, that may walk back and forth in time, to show fire spread evolution.

4.3 Meta-data queries

This module allows users to interact with the WPS algorithm Meta-data management, with the purpose of seeking and obtain data, by means of temporal or geographic queries, or determine the running status of the simulation. This last example is of high importance to the Execution Console module – the module responsible for the execution of the simulations – because it needs to know when new data is made available through all the execution time.

Using a GUI based on dialog boxes, the user may submit queries to the meta-data data-base to explore and use the meta-data that represents the simulations returned by the query. Upon receiving the answer, the user can select which simulation he/she wants to review/resubmit, and the system will automatically perform all the necessary tasks, downloading all data corresponding to that simulation from the geo-referenced data-base.

4.5 Spatial Data Queries

In order to obtain all the necessary data to run the simulations from the SDI, this algorithm interacts with the WPS algorithm Spatial Data Management, allowing users to query the SDI about the availability of data in a certain geographical region.

This module interacts with user giving then the possibility to choose the kind of data to be obtained – terrain, fuel, meteorological data, among others – and to specify parameters, such as geospatial co-ordinates or time intervals. The chosen data is then made available in other layers inside the application.

The CAD facilities may also be used to allow for example, drawing a rectangle in a region of the map, obtaining all the available information regarding that region.

5. Data flow

To better understand the overall CROSS-Fire project approach we depicted in figure 3 the complete data flow and interactions between all the different components of the platform.

We start by a data request issued by the client (CFS), who contacts the WPS layer querying for geo-referenced data related to a certain area under simulation (1). The WPS then contacts a know/discovered SDI to obtain a list of possible data (2). After selected and downloaded the required data (3), the user issues a simulation request to the WPS layer (4). This request can either trigger the G-FS or G-Wind execution models or both simulations.

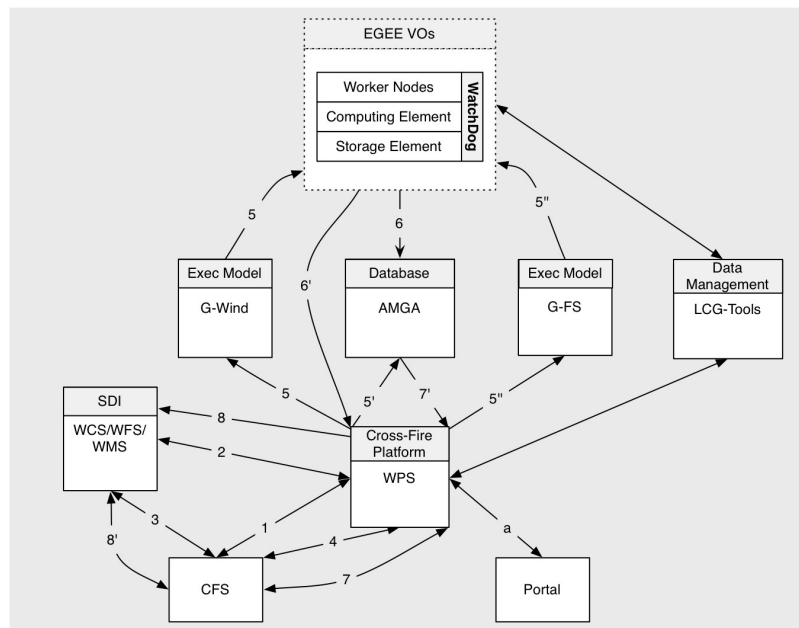


Figure 3 Data Flow

In response to the request the WPS spawns the following procedures: First, it creates the necessary JDL files to run the simulations, and then it contacts the AMGA server to record a new simulation entry and updates its state (5'). Simultaneously, it initiates the task of finding a suitable GRID site to run the simulation and then copies the necessary information from the SDI to the appropriate place, performing data conversions when needed. When all the input data is made available, one or more related jobs are automatically submitted to the GRID (5, 5'').

The job(s) includes a monitoring process, based in WatchDog(WD) that will report the progress of the simulation, to the WPS layer. The WD will update AMGA (6), changing the status of the simulation. To allow near-real time visualisation of the fire spread simulation we use a WD embedded WPS client that runs locally to return the partial computational results to the WPS layer (6') that stores it in the SDI, making it automatically available to clients (8).

The totality of the result data produced by the simulation is stored, at end of the simulation, in the folders where the input data was previously store in order to make all the data available for analysis and further processing.

Once the user initiates a simulation, a polling mechanism is launched in the client to check the simulation's executing status, using the WPS layer (7, 7''). When the state switch to running, the application also receives a link to the place where new data will become

available, entering then in another polling cycle used to progressively retrieve the simulation's results (8').

Two types of clients are depicted in the figure: the CFS with GIS and CAD capabilities complying with the OGC standards, and a Portal we plan to develop to allow users access, through a WEB browser to most of the CFS facilities. The (a) labelled arrow that connects the Portal with WPS is a mean to represent all the interactions described above, between the CFS and the WPS.

6. Current and Future Work

Current work is centred on: 1) the development of a parallel implementation of CANYON, a 3D Navier-Stokes Wind simulation execution model to complement G-FireStation capabilities, 2) the enhancement of CFS with the implementation of a login module to allow user authentication and authorization on the Grid and Geospatial infrastructures and a user grid interface module to allow a consistent and easy way to launch fire and wind simulations on the grid and 3) the design and implementation of a decision-support system based on a web portal where many players can connect, to request services through the core WPS core layer.

The portal will allow for example CP authorities and other entities (e.g., the City Council officers) to request new simulations, to update information concerning the input data required to estimate the risk of the natural hazard, such as fuel distribution and terrain maps, or to access the past simulations to validate the predictions with actual field data.

In another direction, a standard-based SDI layer Geo-server based, is been exploited to provide FS with static data, and to publish data for further processing, while to provide FS with the dynamic data, we are presently using the 52° North implementation of the OGC-SWE compatible layer. The dynamic data includes meteorological information and satellite images, coming from sensors in weather stations (such as DAVIS Vantage Pro2) and sensors MODIS from satellites (such as Terra/Aqua).

7. References

- [1] A. Lopes, "FireStation User's Manual", Universidade de Coimbra, 2000
- [2] A. Lopes, M. Cruz, and D. Viegas. "FireStation - an integrated software system for the numerical simulation of fire spread on complex topography". *Environmental Modelling and Software*, 17(3):269285, 2002
- [3] W Gropp, R Thakur, E Lusk, "Using MPI-2: Advanced Features of the Message Passing Interface", MIT Press Cambridge, MA, USA, 1999
- [4] N. Santos, B. Koblitz, "Metadata services on the Grid", *Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, 2005
- [5] R. Bruno, "A watchdog utility to keep track of job execution", INFN, <https://grid.ct.infn.it/twiki/bin/view/PI2S2/WatchdogUtility>
- [6] A. Pina, B. Oliveira, J. Puga, R. Marques, A. Proença, "An OGC-WS framework to run FireStation on the Grid ", *3d Iberian Grid Infrastructure Conference*, Valencia, Espanha, 2009