# Improvement of Job Scheduling for Automatic Chain Processing in Radio Occultation Context

Lorenzo Mossucca, Olivier Terzo
*Istituto Superiore Mario Boella*
*Via P. C. Boggio 61,*
*Torino, Italy*
*Email: {mossucca, terzo}@ismb.it*

Manuela Cucca, Riccardo Notarpietro
*Politecnico di Torino*
*Corso Duca degli Abruzzi 24,*
*Torino, Italy*
*Email: {manuela.cucca, riccardo.notarpietro}@polito.it*

***Abstract*—The new Italian GPS receiver for Radio Occultation has been launched from Satish Dhawan Space Center (Sriharikota, India) on board of the Indian Remote Sensing OCEANSAT-2 satellite. The Italian Space Agency has established a set of Italian universities and research centers to develop the Web Science Grid, an infrastructure based on grid computing, that is implemented for the overall processing Radio Occultation chain. In consideration of the complexity of our scenario due to the modules involved and difficulties of geographically dispersed nodes, after a brief description of the algorithms adopted, that can be used to characterize the temperature, pressure and humidity, the paper presents an improvement of job scheduling in order to further decrease the elaboration time. Two applications to manage automatically the Radio Occultation data are described: Local and Global scheduler, one for worker nodes and one for the master node. Also the estimated processing time and actual processing are shown.**

***Keywords*-radio occultation; grid computing; local scheduler; global scheduler; job scheduling.**

## I. INTRODUCTION

The GPS Radio Occultation (RO) is an emerging remote sensing technique for the profiling of atmospheric parameters (first of all refractivity, but also pressure, temperature, humidity and electron density, see [1] and [2]). It is based on the inversion of $L_1$ and $L_2$ GPS signals collected by an ad hoc receiver placed on-board a Low Earth Orbit (LEO) platform, when the transmitter rises or sets beyond the Earth's limb. The relative movement of both satellites allows a ”quasi” vertical atmospheric scan of the signal trajectory and the profiles extracted are characterized by high vertical resolution and high accuracy. The RO technique is applied for meteorological purposes (data collected by one LEO receiver placed at 700 km altitude produce 300÷400 profiles per day, worldwide distributed) since such observations can easily be assimilated into Numerical Weather Prediction models. Anyway, it is also very useful for climatological purposes, for gravity wave observations and for Space Weather applications. Starting from the first operational RO mission on board the German CHAMP satellite [6], there are presently several other satellite missions carrying on-board a RO payload. The most important are RO experiments on-board the European METOP-1 mission [3] and on-board the USA/Taiwan COSMIC constellation mission. Several other missions are planned for the next future. In particular, during the 2009 autumn season, the Indian OCEANSAT-2 mission carrying on-board the Italian ROSA (Radio Occultation Sounder of the Atmosphere) GPS receiver was launched. In the framework of this opportunity, the Italian Space Agency [4] funded a pool of Italian Universities and Research Centers for the implementation of the overall RO processing chain, which is called ROSA-ROSSA (ROSA-Research and Operational Satellite and Software Activities). The ROSA-ROSSA was integrated in the operational ROSA Ground Segment by an Italian Software enterprise (INNOVA, located in Matera, Italy), and the ROSA ground segment is operating in Italy (at the ASI Space Geodesy Center, near Matera) and in India (at the Indian National Remote Sensing Agency [5], near Hyderabad) starting from the 2009 autumn season. This version implements RO state-of-the-art algorithms and, for the first time, it was developed and it runs on a distributed hardware and software infrastructure exploiting a grid computing strategy, which is called Web Science Grid (WSG). The paper is structured as follows: Section 2 is devoted to a more detailed description of the ROSA-ROSSA software. This section is given in order to better set up the scientific application which exploits grid processing strategies. Section 3 describes motivations. Section 4 presents the structure of our system and scheduling description. Section 5 contains considerations about the time execution obtained by the system based on grid computing. Section 6 draws the conclusions.

## II. THE PROCESSING CHAIN OF RO OBSERVATIONS

The ROSA-ROSSA software implements state-of-the-art RO algorithms which were already available from the scientific group and are during the validation phase before their final transfer inside the official Ground Segment of the ROSA Radio Occultation receiver. The processing chain, which is subdivided into seven different software modules (namely Data Generators-DG), is executed in a sequential

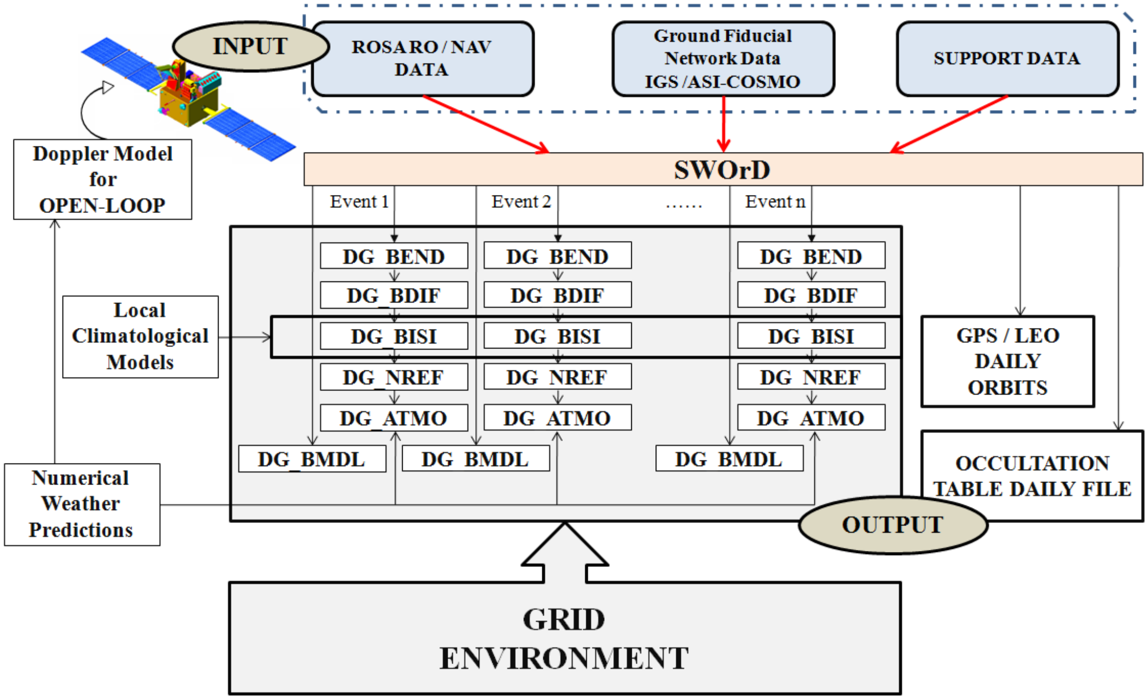

INPUT
ROSA RO / NAV DATA
Ground Fiducial Network Data IGS / ASI-COSMO
SUPPORT DATA
SWOrD
Doppler Model for OPEN-LOOP
Event 1
Event 2
……
Event n
DG_BEND
DG_BDIF
DG_BISI
DG_NREF
DG_ATMO
DG_BMDL
Local Climatological Models
GPS / LEO DAILY ORBITS
OCCULTATION TABLE DAILY FILE
Numerical Weather Predictions
OUTPUT
GRID ENVIRONMENT

Figure 1. ROSA-ROSSA Overall Chain.

mode. Figure 1 shows a simple diagram of the processing chain and of the corresponding data-flow. Before delving into how the various parts work, a DGs explanation is given, in order to focus the data types to process. Starting from ROSA Level 1.a engineered data coming from the ROSA on-board OCEANSAT-2 platform observations, from the ground GPS network and from other support data, the ROSA-ROSSA is able to produce data at higher levels, using a data processing chain defined by the following Data Generators. SWOrD is a software module that fully supports orbit determination, orbit prediction, and implements Level 2 data generation connected with the ROSA sensor on-board OCEANSAT-2. Input data for SWOrD are ROSA GPS navigation and Radio Occultation observations, ground GPS network data and other support data. It generates the following output data:

- Estimated rapid orbits and predicted orbits for the GPS constellation;
- Estimated rapid orbits and predicted orbits for the OCEANSAT-2 platform;
- 50 Hz closed-loop and 100 Hz open-loop excess phases and signal amplitude data for each single occultation event;
- Tables showing estimated and predicted (up to 6 hours in advance) occultation (Data Level 2.c).

The BMDL Data Generator predicts a bending angle and impact parameter profile (Level 2.d data) usable as input in the ROSA on-board software excess doppler prediction module for open-loop tracking. For each ”predicted” occultation event, latitude and longitude of the geometrical tangent points (the nearest point of each trajectory to the Earth’s surface, evaluated through predicted orbits) is used to compute bending angle and impact parameter profile from interpolated numerical weather prediction models (bending angle and impact parameter are geometrical parameter univocally identifying each trajectory followed by the RO signal. See Figure 2 for details). Predicted bending angle and impact parameter profiles $\alpha$ (a) (2.d Data Level) are stored in ASCII data files containing bending angles and impact parameters together with the UTC time stamp, one file for each event. Input data for DG_BMDL are 1b.a, 1b.b (predicted GPS and LEO orbits, respectively) and 2.c (Predicted Occultation Tables), together with ECMWF world forecasts for the synoptic times valid for the future observed occultation event. The BEND Data Generator provides ”raw” bending angle and impact parameter profiles $\alpha$(a) computed on GPS occulted signals on both GPS frequencies $L_1$ and $L_2$, by using a Wave Optics approach below a certain altitude (generally in troposphere). Above that altitude threshold, standard Geometrical Optics algorithms are applied. Raw bending angle and impact parameter profiles $\alpha$(a) (Data Level 3.a) are stored for each event in ASCII data files. Inputs for DG_BEND are 2.a data ($L_1$ and $L_2$ excess-phases and related orbit data) and 2.b data ($L_1$ and $L_2$ signal amplitudes).

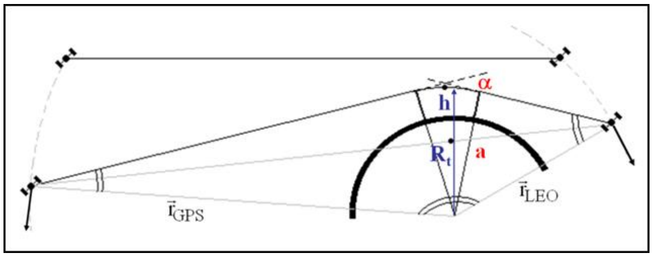


Figure 2. Radio Occultation geometry. The quasi instantaneous trajectory can be defined by the following geometrical parameters: the bending angle $\alpha$, the impact parameter a. $R_t$ is the local Earth's radius and h the tangent point height.

The BDIF Data Generator provides (for each event) a bending angle and impact parameter profile, on which the ionospheric effects have been compensated for. This DG processes both $L_1$ and $L_2$ bending angle and impact parameters profiles (Data Level 3.a) given as input, in order to minimize the first order ionospheric dispersive effects. Outputs for DG_BDIF are bending angle and impact parameter iono-free profiles (Data Level 3.b). The BISI Data Generator provides profiles of bending angle versus impact parameter optimized in the stratosphere above 40 km. In the ROSA-ROSSA, data coming from a Numerical Weather Prediction Model (ECMWF analysis) are used in place of climatological data for implementing the statistical optimization procedure necessary to reduce the high noise level left to the signal after ionospheric first order compensation applied by the previous DG_BDIF. DG_BISI processes bending angle and impact parameter profiles obtained from Data Level 3.b. Output for DG_BISI are bending angle and impact parameter profiles optimized in the stratosphere (Data Level 3.c). The NREF Data Generator provides (for each event) the refractivity profile and dry air temperature and pressure profiles. This DG is able to process iono-free and properly initialized bending angle and impact parameter profiles (Data Level 3.c) in order to compute the corresponding dry air ”quasi” vertical atmospheric profiles (Data Level 3.d). The ATMO Data Generator allows to evaluate the temperature and the water vapour profiles using forecasts or analysis obtained by numerical weather prediction. This DG receives on input Level 3.d data files and produces on output Level 3.e data files, which contain the total temperature and total pressure profiles in terms of wet and dry components.

## III. Architecture Motivations

The main purpose is to create a flexible architecture in order to manage the radio occultation data and to reduce their processing time. The system guarantees the entire processing chain automatically that consists of seven DGs executed sequentially as explained before. In a learning phase, we evaluated that for each day, the events number to process are about 250, on a single machine the elaboration time for the entire chain processing, is approximatively 40 hours. The idea of using a distributed environment arose from the need to reduce this processing time because these makes it difficult to use the results. The WSG has been developed with the goal of simplifying this task, by providing implementations of various core services provided by Globus Toolkit and deemed essential for high-performance distributed computing. Furthermore, it allows engineers and physicists of the project to have a tool for processing and sharing data, independently from the university in which they are.

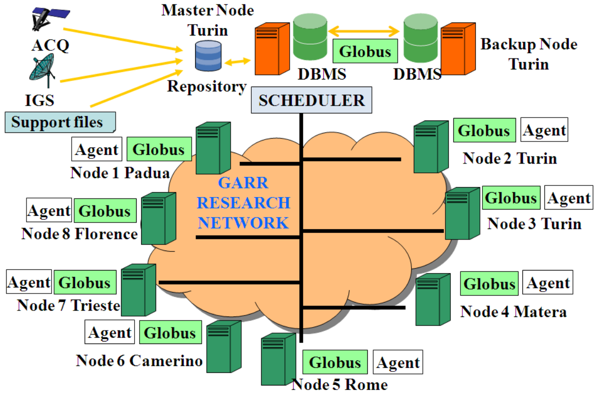


Figure 3. Web Science Grid.

## IV. Architecture Design

The Web Science Grid (WSG) is an integrated system devoted to handle and process RO data of the OCEANSAT-2 ROSA on board sensor.

### A. Web Science Grid architecture

The WSG is composed by the subsystems(see Figure 3): middleware, central repository, relational database, scheduler, agents and applications. The general purpose of our project is: sharing the computational resources, transferring a great amount of files and submitting jobs from several different organizations of the scientific community located in different places in Italy. All these operations are processed in an automatic way without any user interaction. The pool of nodes consists of 10 nodes with 2 processors each, 2 GB RAM, 64 bit machines , and on all these machines run Linux (Ubuntu). The nodes are located geographically in Italy, for accuracy to:

- Istituto Superiore Mario Boella (Turin);
- Polytechnic University of Turin (Turin);
- University of Padua (Padua);
- Sapienza University (Rome);
- University of Camerino (Macerata);
- International Center of Theoretical Physics (Trieste);
- Italian Space Agency (Matera);
- Institute for Complex System (Florence).

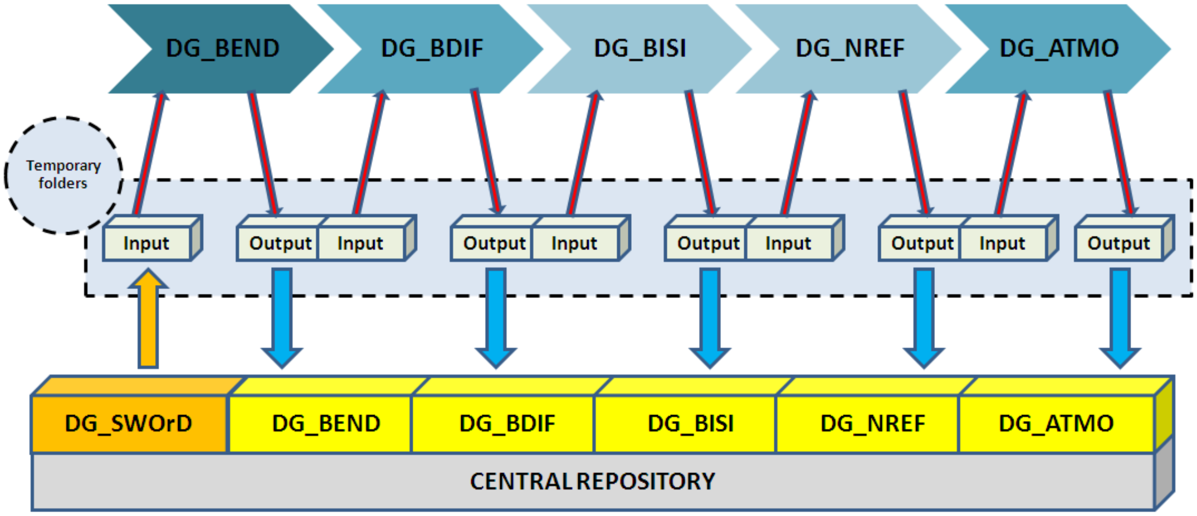


Figure 4. Automatic processing chain

The Globus Toolkit has been used as middleware [7] and [8], since it allows obtaining a reliable information technology infrastructure that enables the integrated, collaborative use of computers, networks and databases. The Globus toolkit is a collection of software components designed to support the development of applications for high-performance distributed computing environments, or computational grids [11].

### B. Automatic chain

Our software allows to run the chain automatically; it is composed from two schedulers: one who listens to the master node, called global scheduler, and checks for files ready for execution and sends them to worker nodes, according to the scheduling rules, the other, called local scheduler, listens on the worker nodes, and when it receives a file executes and returns the result file obtained on the master node [9] and [10]. In Figure 4, data flow is depicted, the first transaction takes place on master node; it receives the files directly from the satellite and performs the first step in the chain, i.e., SWOrD, generating about 256 files that are placed in the folder the next step, DG_BEND. When there are files in the folder DG_BEND, the global scheduler checks nodes available by querying the database, and sends files to them. Global scheduler provides for automated scheduling of any input files. It uses all machines belonging to the grid to distribute work load and to provide a backup system for all critical tasks within the system. The choice of how to share the file to run is based on 2 sets of scheduling rules, one concerning the available nodes and one derived from an analysis of the file to run. An agent is installed on each node, is used to monitor the availability of each service on the node and periodically, it sends its general status to the database on master node, if all services are active the node is in condition to receive a job. For the selection of nodes available and ready to run, the global scheduler checks on the database directly instead of querying each machine. When the worker node sees a file in its folder, starts the processing procedure that will generate an output file that will be sent to master node in the folder next step, i.e., DG_BDIF. This procedure is performed for every steps of the chain, the operation is as follows: from SWOrD, the DG n-1 generates the output file that will be the input files of DG n, and so on. On worker nodes, each execution is performed in a temporary folder, so that, in case of error, identify the type of error made and then to reprocess the file. Two types of errors can occur: the first for lack of data in the file due to the satellite reception, the second for network failures or node crash. Only in the last case it is worth recover the process, and it is enough reprocess il file. Anyway, each process has a timeout, if within a fixed time processing has not been completed, the process is killed. An important component of this architecture is the database, which allows us to monitor any action of the grid. Regarding the automatic chain, each transaction is stored on the database when it starts running, when it ends, input files, output files, the node that has run and type of error, if it has generated them. The database also contains information on the status of each node and are available to receive the file to run, this allows us to understand whether there are network problems, so if the node is reachable.

## V. Improving performances

All DGs of the processing chain have been tested during a learning phase; for a single event it obtained the percentage values in Figure 5, and for a daily events in Figure 6. In the two graphs, the difference is due from that, for each hour SWOrD generates only one event for DG_BMDL and instead from DG_BEND to DG_ATMO it generates about nine events. Number of input, output and time for elaboration have been considered. It has been assumed that SWOrD has already been executed, then it is outside the calculation processing. The Eq. 1 and Eq. 2 represent an estimation time of elaboration and cover both a non-distributed (N = 1) and distributed architecture ($N > 1$).

$$T_p = T_e * \eta \tag{1}$$

$$T_p = \frac{1}{N}\sum_{i=1}^{\eta}(T_{ei} + \beta) \tag{2}$$

Where:

$$T_p = TotalProcessTime$$
$$T_e = EventProcessTime$$
$$T_{ei} = EventiProcessTime$$
$$\eta = NumberofROEvents$$
$$N = NumberofGridNodes$$
$$\beta = FileTransferTime$$

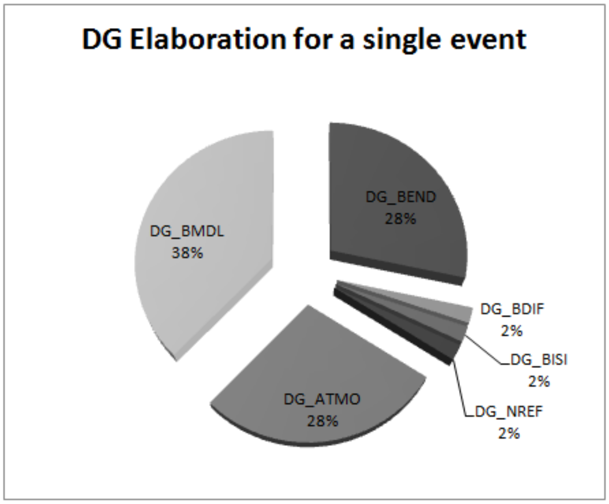


Figure 5. DG Elaboration for a single event.

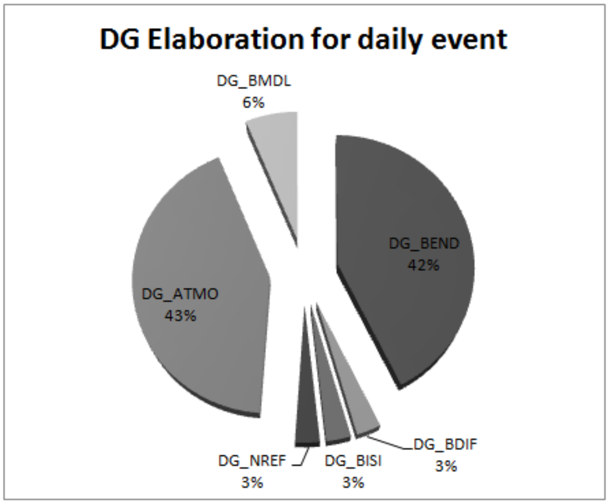
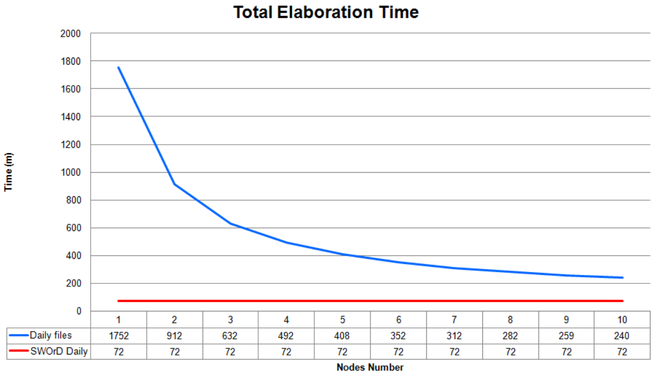


Figure 6. DG elaboration for daily events.



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Daily files | 1752 | 912 | 632 | 492 | 408 | 352 | 312 | 282 | 259 | 240 |
| SWOrD Daily | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 |

Figure 7. Estimated processing time for daily data (about 250 events)

In Figure 7, the execution time trend is estimated, when the number of nodes and events is increased. When only one node is available, the total execution time for a daily files is 1752 minutes (about 29 hours), instead increasing the number of nodes, the execution decrease further, just note that with 2 nodes is 912 (about 15 hours). An important point when a single event is processed is that there is no gain time in grid environment; rather time is higher because we must consider the transfer time; it has a sizeable gain time only when a set of files are processed.

In Table I, processing time detected for daily data elaboration is considered; it depicts how to change the processing time when worker nodes increase. Certainly, the benefits of the grid is ensure the elaboration the overall chain in less time, instead, in distributed system where worker nodes are geographically located, it can have disadvantage in the network layer, in case of network failures or slow connections, to overcome this problem only internal nodes are available for elaboration.

| NODES NUMBER | PROCESSING TIME |
|---|---|
| 10 | 2h20m |
| 8 | 3h06m |
| 6 | 4h00m |
| 4 | 6h30m |

Table I
DETECTED PROCESSING TIME FOR DAILY DATA (ABOUT 250 EVENTS)

## VI. CONCLUSIONS

The ROSA-ROSSA software implements Radio Occultation technique, which run for the first time on a grid computing infrastructure, called Web Science Grid and elaborations time are described. This paper want to be an example of application where you can use grid computing. In frameworks such as Radio Occultation, where the amount of data to be processed is significant, the use of a distributed architecture as the grid can be the best choice. We have focused on a way to manage the assignment nodes for execution in automatic way without any human interaction through a local and a global scheduler. As future works we plan the extension of the proposed architecture to clusters available across the European Grid Infrastructure (EGI) and we are studying a solution for EC2 environment by Amazon to allow to further increase available computing power.

## VII. Acknowledgments

The authors are grateful the Italian Space Agency (ASI) for supporting this project within contract I/006/07/0 and to all the ROSA-ROSSA partners for their contributions.

## References

[1] Melbourne, W., *The application of spaceborn gps to atmospheric limb sounding and global change monitoring*, JPL Publ, pp. 18-94, 1994

[2] Kursinski, E.R., *Observing Earth's atmosphere with radio occultation measurements*, Journal Geophys. Res., pp. 429-465, 1997

[3] Luntanama, J.P., *Prospects of the EPS GRAS Mission for operational atmospheric applications*, Bull. Am. Met. Soc, pp. 1863-1875, 2008

[4] Italian Space Agency (ASI), *http://www.asi.it/*, 2010

[5] Indian National Remote Sensing Agency (ISRO), *http://www.isro.org/*, 2010

[6] Wickert, J., *The radio occultation experiment aboard CHAMP: Operational data processing and validation of atmospheric parameters*, Journal Meteorol. Soc. Jpn, pp. 381-395, 2004

[7] Foster, I. and Kesselman, C., *The Grid2: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann, pp. 38-63, 2003

[8] Berman, F., Fox, G. and Hey, G., *Grid Computing Making the Global Infrastructure a Reality*, Wiley, pp. 117-170, 2005

[9] Dimitriadou, S. and Karatza, H., *Job Scheduling in a Distributed system Using Backfilling with Inaccurate Runtime Computation*, The international conference on complex, intelligent and software intensive system, pp. 329-336, 2010

[10] Xhafa, F., Pllan, S. and Barolli, L., *Grid and P2P Middleware for Scientific Computing Systems*, The international conference on complex, intelligent and software intensive system, pp. 409-414, 2010

[11] Brunett, S., Czajkowski, K., Fitzgerald, S., Foster, I., Johnson, A., Kesselman, C., Leigh, J., Tuecke, S., *Application Experiences with the Globus Toolkit*, The Seventh International Symposium on High Performance Distributed Computing, pp.81-88, 1998

ISBN: 978-1-61208-133-5