Data intensive scientific analysis with grid computing

**Original**

**Availability:**
This version is available at: 11583/2426391 since:

**Publisher:**
University of Zielona Gora press

**Published**
DOI:10.2478/v10006-011-0016-z

**Terms of use:**
openAccess
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

(Article begins on next page)
At the end of September 2009, a new Italian GPS receiver for radio occultation was launched from the Satish Dhawan Space Center (Sriharikota, India) on the Indian Remote Sensing OCEANSAT-2 satellite. The Italian Space Agency has established a set of Italian universities and research centers to implement the overall processing radio occultation chain. After a brief description of the adopted algorithms, which can be used to characterize the temperature, pressure and humidity, the contribution will focus on a method for automatic processing these data, based on the use of a distributed architecture. This paper aims at being a possible application of grid computing for scientific research.

Keywords: grid computing, GPS radio occultation, scheduler, agent, e-science.

1. Introduction

GPS Radio Occultation (RO) is an emerging remote sensing technique for profiling atmospheric parameters (mainly refractivity, but also pressure, temperature, humidity and electron density (Melbourne et al., 1994; Kursinski et al., 1997)). It is based on the inversion of $L_1$ and $L_2$ GPS signals collected by an ad hoc receiver placed on 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 a high vertical resolution and a high accuracy. The RO technique is applied for meteorological purposes (data collected by one LEO receiver placed at an altitude of 700 km produce 300-400 profiles per day, worldwide distributed) since such observations can be easily assimilated into numerical weather prediction models. Anyway, it is also very useful for climatological purposes (the accuracy of inferred tropopause parameters is one of the most attractive aspects of the RO technique), gravity wave observations and space weather applications. Figure 1 (taken from the work of Luntama et al. (2008)) sketches a concept view of the radio occultation technique.

Starting from the first operational RO mission on board the German CHAMP satellite (Wickert et al., 2004), there are presently several other satellite missions carrying on board an RO payload. The most important are RO experiments on board the European METOP-1 mission (Luntama et al., 2008) and on board the USA/Taiwan COSMIC constellation mission. Several other missions are planned for the 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 (ASI, 2010) 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). ROSA-ROSSA was integrated in the operational ROSA ground segment by an Italian software enterprise (INNOVA, located in Matera, Italy), while the ROSA ground segment has been operating in Italy (at the ASI Space Geodesy Center, near Matera) and in India (at the Indian National Remote Sensing Agency (ISRO, 2010), near Hyderabad) starting from the 2009 autumn season.

Two ROSA-ROSSA releases are actually available: the operational one, which implements well consolidated RO algorithms and runs on the Matera and Hyderabad ground segments, and the extended ROSA-ROSSA version. The latter implements RO state-of-the-art algorithms and, for the first time, was developed and runs on a distributed hardware and software infrastructure exploiting a grid computing strategy, which is called the Web Science Grid (WSG). The ROSA-ROSSA extended version was validated through the WSG, giving on input to the overall chain CHAMP raw data (related to one day of observations, those performed on 22 January 2004) formatted to ROSA-ROSSA level 1.a data. Output data, which were extracted at different processing stages, were compared with reference data. In particular, both CHAMP products (at the same ROSA-ROSSA output levels) and co-located data taken on independent sources (we used numerical weather analysis from the European Center for Medium range Weather Forecast, ECMWF) were adopted as reference output data. Statistical comparisons were carried out in order to validate the ROSA-ROSSA VE performances.

The paper is structured as follows. Section 2 is devoted to a more detailed description of the ROSA-ROSSA extended version software (in terms of algorithms and the data level) and to the discussion of some results obtained after WSG processing of one-day CHAMP observations from a scientific point of view. This section aims to better set up the “scientific” application which exploits grid processing strategies. Section 3 describes motivations and related work. Section 4 presents the structure of our system, data model, agents and scheduling description. Section 5 explains the applications for elaboration. Section 6 contains a discussion of the time execution obtained by the system based on grid computing. Sections 7 draws conclusions and includes proposals for future work.
constellation in the conventional terrestrial reference frame (data level 1b.a);

- estimated rapid orbits and predicted orbits for the OCEANSAT-2 platform in the conventional terrestrial and celestial reference frame (data level 1b.b);

- 50 Hz closed-loop and 100 Hz open-loop excess phases (data level 2.a) and signal amplitudes data (data level 2.b) for each single occultation event;

- tables showing estimated and predicted (up to 6 hours in advance) occultation (data level 2.c).

**DG_BMDL** 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, the latitude and longitude of geometrical tangent points (the nearest point of each trajectory to the Earth’s surface, evaluated through predicted orbits) is used to compute the bending angle and impact parameter profile from interpolated numerical weather prediction models (the bending angle and the impact parameter are geometrical parameters univocally identifying each trajectory followed by the RO signal, cf. Fig. 3 for details). The predicted bending angle and impact parameter profiles $\alpha(\alpha)$ (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 synoptic times valid for a future observed occultation event.

**DG_BEND** 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. Input data 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).

**DG_BDIF** 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 parameter profiles (data level 3.a) given as input, in order to minimize first order ionospheric dispersive effects. Outputs for DG_BDIF are bending angle and impact parameter iono-free profiles (data level 3.b).

**DG_BISI** provides profiles of the bending angle versus the impact parameter optimized in the stratosphere above 40 km. In the ROSA-ROSSA extended version, 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).

**DG_NREF** provides (for each event) the refractivity profile as well as 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).

**DG_ATMO** allows evaluating the temperature and 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.

Leaving the description of the grid architecture and procedures implemented in the framework of the Web Science Grid for the following sections, we want here to describe some statistical results obtained by processing one day of CHAMP raw data (22 January 2004) using the implemented ROSA-ROSSA extended version. This dataset contains about 120 setting events whose data are available for POD evaluations, at the $L_1$ and $L_2$ excess-phase levels, at the bending angle and impact parameter iono free and optimized in the stratosphere profile level, up to the refractivity level. Such data were converted from the native format (CHAMP) to the ROSA-ROSSA one in order
Fig. 4. ROSA-ROSSA scenario for validation purposes using CHAMP reference data as input to the ROSA-ROSSA software. (POD: precise orbit determination; PHD: excess phases; IF: iono-free; SO stratospheric optimization.) Yellow areas define products that were validated against corresponding CHAMP data at the same level. Green areas define products that were validated against independent and co-located atmospheric data (we used data meteorological analysis obtained from ECMWF).

to make them directly available on input to the ROSA-ROSSA chain for immediate processing and to ease the comparisons with ROSA-ROSSA products at the corresponding level. Statistical comparison between data generated at different levels from the ROSA-ROSSA software using WSG and the corresponding CHAMP original products at the same level was carried out (the scenario is shown in Fig. 4). Also independent atmospheric data were used to validate our ROSA-ROSSA products. ECMWF global analyses were co-located in time and space to the generic atmospheric ROSA-ROSSA product, differences were computed, and statistical comparisons on several data files (up to 120, the total number of available occultation events) were performed. Figure 5 shows the statistics on fractional differences between bending angle iono-free and stratosphere optimized profiles obtained by the ROSA-ROSSA extended version and the corresponding CHAMP original profiles. The observed bias is probably due to the incomplete knowledge of the center of the local atmospheric symmetry computation strategy performed by the CHAMP software. Statistics on fractional differences between refractivity profiles obtained by ROSA-ROSSA ($N_{RSA}$) and the corresponding CHAMP original profiles ($N_{CHA}$) are shown in Fig. 6.

Statistical validation of the refractivity profiles given on output to DG$_{NREF}$ and ECMWF co-located data was also carried out. Results are given in Fig. 7.

Finally, co-located ECMWF temperature profiles ($T_{ECM}$), specific humidity profiles ($q_{ECM}$) and refractivity profiles ($N_{ECM}$) derived from global analysis for January 22, 2004 are the reference data for the corresponding products generated by the ROSA-ROSSA extended version DG$_{ATMO}$ data generator. Statistics on the differences are shown in Figs. 8–10.

3. Related work

The increasing number of research projects on Earth observation, such as GPS radio occultation, produces a large amount of geographical datasets. The analysis, storage and management of these data are a very hard and bor-
Data intensive scientific analysis with grid computing

Fig. 8. Statistical analysis of fractional differences between the ROSA-ROSSA profile and the corresponding ECMWF co-located profile. Profiles here are the refractivity profiles $N(h)$.

Fig. 9. Statistical analysis of fractional differences between the ROSA-ROSSA profile and the corresponding ECMWF co-located profile. Profiles here are the temperature profiles $T(h)$.

Fig. 10. Statistical analysis of fractional differences between the ROSA-ROSSA profile and the corresponding ECMWF co-located profile. Profiles here are the specific humidity profiles $q(h)$.

Fig. 11. SWOrD description.

As discussed by Kurowski et al. (2006), scheduling jobs on grid and resource management are the major challenge in a grid environment. In general, a lot of effort and features used for grid scheduling have been partly developed by the enabling grid for e-science project (EGEE), now the European Grid Infrastructure (EGI) and the Globus Alliance. Due to their complexity, scientific applications need efficient algorithms to manage resources and jobs. In the e-science environment, in many cases some effort is required in terms of the scheduling approach. In this paper, due to the complexity of the analysis of radio occultation as mentioned by Dimitriadou and Karatza (2010), a global and local scheduling approach was studied and implemented. As mentioned in previous sections, the processing chain consists of seven DGs and these involve complex algorithms that use Matlab. Executable files were created with the Matlab Compiler, which allows reducing further processing time and discarding source files. SWOrD use a set of languages like Fortran and C++. It is elaborated on the master node and generates about 256 files per day; only these files are processed sequentially within the grid (see Fig. 11).

4. System design

4.1. Overview. Grid Processing Management (GPM) is an integrated system devoted to handle and process RO data of the OCEANSAT-2 ROSA on board sensor. GPM is composed of the following subsystems (Fig. 12): middleware, central repository, relational database, adapter,
scheduler, agents, and application. The general purpose of
our project is sharing the computational resources, trans-
ferring a great amount of files and submitting jobs from
several different organizations of the scientific commu-
nity located in different areas of Italy. The nodes are lo-
cated geographically in Italy, namely, at Istituto Superiore
Mario Boella (Turin), the Polytechnic University of Turin
(Turin), the University of Padua (Padua), Sapienza Uni-
versity (Rome), the University of Camerino (Macerata),
the International Center of Theoretical Physics (Trieste),
the Italian Space Agency (Matera) and the Institute for
Complex Systems (Florence). Hence, we used the Globus
Toolkit as middleware (Berman, 2003; Globus, 2010a;
2010b), since it allows obtaining a reliable information
technology infrastructure that enables integrated, collab-
orative use of computers, networks and databases. The
repository is the component responsible for storing data
in directories and files; it stores executable files, source,
input/output and support files.

4.2. Data model. The data model, based on a MySQL
Database Management System, is a fundamental compo-
nent of the system. It has three main functions: one for
grid configuration and user accounting, one for data gen-
erator algorithm description, and one for log data. The
data model is composed of six domains in order to define
the main function (Fig. 13).

GRID CONFIGURATION: It is dedicated to the nodes’
description, the technical characteristics of each node, the
path where are stored files and the network parameters
such as IP, MAC address and location. This information
is useful for the scheduler and middleware.

DATA GENERATORS DESCRIPTION: This set of tables
describe each data generator. Each table is deputed to
store information about files, softwares and libraries list
needed for its execution, the path where all sources and
support files can be found. All DGs are characterized by
an XML file, and we also describe the XML file structure
for each DG; in this way, we are able to dynamically generate the XML file for each event.

GRID STATUS: This set of tables are dedicated to
the scheduler; each table contains all information concerning
the general status of each node and how many services are
up/down.

GRID STATISTIC: This set of tables store information
about the behavior of some parameters (CPU and RAM)
on each node and the total processing time for each DG
on each node. A policy rule for the scheduler is based on
the estimation of the total elaboration time.

USER ACCOUNT AND GRID CREDENTIAL: On the
grid, each user has different permission, such as executing
DGs, reading data and showing output.

4.3. Scheduling models. The scheduler provides for
automated scheduling of any input files, processes as well
as import and export of data (Buyya et al., 2000). It uses
all machines belonging to the grid to distribute the work
load and provide a backup system for all critical tasks
within the system (Kurowski et al., 2006). It has the abil-
ity to run more than one job at the same time and run
multiple steps within a single job in parallel to complete
jobs faster and more efficiently (Leonid, 2004). Its tasks
can also be run with a configuration parameter if desired,
creating only a part of the chain. Two schedulers were developed:
a global and a local scheduler. The global scheduler is in-
stalled only on the master node and makes it possible to
check the number of files to be processed and the num-
ber of available nodes, and send files to these machines
(see Fig. 14). This scheduler decides how to share the
computer among the processes that are ready to run. The
choice of how to share the file to run is based on two sets
of scheduling rules, one concerning the available nodes
and one derived from an analysis of the file to run. The
global scheduler is split in modules, one module for each
data generator. Each module is a daemon and is in charge
of the scheduling of jobs. Each module is independent.

Six modules have been implemented, each of them
based on Java implementation from the Globus Toolkit.
Data intensive scientific analysis with grid computing

During the requirements of RO analysis, a job is defined by the presence of data files to be analyzed. Each module checks the availability of files on a directory and has a specific directory, one for each Data Generator. The DG module is responsible for getting the list of files for analysis. The next step is to check the availability of the node by querying the grid status table on the data model. A list of available nodes is returned; for each node the schedule module checks if it is in condition to run a job by sending a globus command.

The last condition is to check how many nodes are running; the module sends files by choosing free nodes. If all nodes are running, the files are in a queue condition. Finally, each node has a threshold for its computational capacity and, according to this criterion, one or more files are sent for the analysis. The local scheduler is a daemon and has in charge the run execution of each DG. Each node has a set of directories that correspond to each repository of files for each DG. The local scheduler is implemented with a round robin policy in order to give a rotation for the elaboration of each DG algorithm. The local scheduler takes the list of files and for each of them runs the correct DG algorithm. DG algorithms are developed in a Matlab compiled file. When the execution job is accomplished, the output file is sent to the master node and is located in the next DG directory. In this case, it will be managed by the global scheduler for the next elaboration.

4.4. Agent. In spite of the success of grid computing in providing solutions for a number of large-scale science fields, one of the problems is the scalability of the system. The agents are emerging as a solution to provide flexibility and scalability (Gradwell, 2002). The agent is installed on each node, is used to monitor the availability of each service on the node and periodically sends its general status to the database on the master node. If all services are available, the node is in condition to receive a job. The scheduler, for the selection of nodes available and ready to run, instead of querying each machine, queries the database directly. A specific function checks if the feedback information from all nodes has been sent and, in the case of a missing status, the node is considered unavailable.

5. Java application

5.1. Single job request. At a preliminary step of development, where the algorithms were continuously improved, it was necessary to have a user friendly Java application. It was born for university researchers in order to allow them to test their own algorithms, but it also provides some functionalities:

- to grant secure access to the grid;
- to allow choosing where a DG will be executed, in a local machine or in a grid through the scheduler;
- to display the evolution of the submitted DG;
- to allow monitoring the available machine belonging to the grid;
- to plot the final result of the DG.

5.2. Automatic chain. Our software allows running the chain automatically. It is composed of two daemons: one that listens to the master node, checks for files ready for execution and sends them on worker nodes, according to the scheduling rules, and one that listens to the worker nodes, and on a receipt of a file, executes and returns the result file obtained on the master node. In Fig. [15], the data flow is depicted. The first transaction takes place on the master node, which 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 DG_BEND folder, the scheduler daemons check nodes available by querying the database, and sends them necessary information. When the worker node sees a file in its folder, it starts the processing procedure that will generate an output file that will be sent on the master node in the folder next step, i.e., DG_BDIF. This procedure is performed for each step of the chain, and 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 the case of an error, it is possible to identify the type of error made. 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 in the database: when it starts running, when it ends, input files, output files, the node that has run and the type of error, if it has generated them. The database also contains information on the status of each machine and is available to receive the file to run.
This allows us to understand whether there are network problems and hence whether the machine is reachable.

6. Experimental results

The DGs for the overall processing chain were tested during a learning phase. For a single event we obtained the percentage values in Fig. 16 and for daily events those in Fig. 17. The amounts of input, output and time for elaboration were considered. The following equations represent the estimation time of elaboration and cover both a non-distributed ($N = 1$) and a distributed architecture ($N > 1$):

$$T_p = T_e \cdot \eta,$$

(1)

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

(2)

respectively, where $T_p$ is the total process time, $T_e$ is the event process time, $T_{ei}$ is the process time for event $i$, $\eta$ denotes the number of RO events, $N$ stands for the number of grid nodes, $\beta$ is the file transfer time.

We have assume that SWOrD has already been executed, and then it is outside the calculation.

In Fig. 18 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 daily files is 1752 minutes (about 29 hours). Instead of increasing the number of nodes, the execution decreases further, just note that with 2 nodes it is 912 (about 15 hours). This means that, when a single event is processed, there is no gain time in the grid environment. The time is higher rather because we must consider the transfer time. It has sizeable gain time only when a set of files are processed. In Table 1 detected processing time is considered.

Certainly, the benefit of the grid is ensuring the elaboration of the overall chain in less time. In a distributed system where worker nodes are geographically located, it can have disadvantages in the network layer in the case of network failures or slow connections. To overcome this problem, only internal nodes are available for elaboration.

7. Conclusions and future work

The radio occultation technique, its implementation on the ROSA-ROSSA extended version software which runs for the first time on a distributed hardware and software in-
Data intensive scientific analysis with grid computing

Fig. 17. DG elaboration for daily events.

Fig. 18. Estimated processing time.

Infrastructure based on grid computing (the Web Science Grid), as well as some results obtained during the validation activity were described. It has to be noted that, besides the impressive decrease in computation time, which allows the availability of radio occultation products in quasi real time, one of the benefits of using a grid approach is the possibility to open the door to each researcher which is interested in developing his or her own processor (or even a small subroutine) of radio occultation data at each level of the processing chain. Such a researcher could take data from a certain level, locally process such data and put the corresponding output into the grid, exploiting the processing given by the previous and, (if interested) the following data generators. This is a very big added value of such an approach, which can permit in the future to improve radio occultation processing strategies without the necessity to implement the complicated processing chain.

This paper attempts to be also an example of an application of 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. Increasing the number of the nodes belonging to the grid, the processing time is reduced exponentially. For the future, we are evaluating the possibility of creating a cluster around each node, thereby increasing computational power, so that the scientific community would have a tool for sharing the results and a collaborative engineering approach.

Acknowledgment

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

References


Olivier Terzo was born in Dijon, France, in 1970. He obtained his degrees in electronic engineering from the Polytechnic University of Turin, Italy, in 2004. His research interests include grid computing, cloud computing technologies and in e-science/engineering data intensive applications.

Lorenzo Mossucca is a student of information engineering at the Polytechnic University of Turin, Italy. He works on the project Software ROSA for Oceansat-2, related to the development of the ground segment for the processing of GPS data, collected by the Italian payload ROSA on board the Indian OCEANSAT-2 mission. He implemented the grid computing environment in the system. His scientific interests include grid computing, database development and Java applications.

Manuela Cucca works on the project entitled Software ROSA for Oceansat-2, related to the development of the ground segment for the processing of GPS data, collected by the Italian payload ROSA on board the Indian OCEANSAT-2 mission. Furthermore, she carries on a research activity on GNSS reflectometry. In particular, she is designing a GNSS receiver for scatterometry purposes, with the goal of developing a low cost, portable and easily reconfigurable receiving system; a measurement campaign is soon foreseen both on ground and on a small UAV (Unmanned Aerial Vehicle) aircraft. She is involved in teaching some courses holding practical exercises.

Riccardo Notarpietro is an assistant professor at the Polytechnic University of Torino. He belongs to the Remote Sensing Group, operating at the Electronics Department. His fields of activities include also electromagnetic wave propagation and radar meteorology. He has been involved in several national and international research projects. Presently he is the scientific coordinator of a sub-contract related to the Italian Space Agency project Software ROSA for Oceansat-2. This project is concerned with the development of the ground segment for the processing of GPS data collected by the Italian payload ROSA on board the Indian OCEANSAT-2 mission. Within the field of GPS meteorology, he is also involved in the project METAWAVE (Mitigation of Electromagnetic Transmission Errors Induced by Atmospheric Water Vapour Effects); in this framework he is developing a tomographic algorithm for three dimensional reconstruction of high resolution wet refractivity fields using GPS signals collected by a dense receiver network. He is the author or coauthor of 11 peer-reviewed papers and of more then 40 contributions to international and national conferences and workshops.

Received: 28 June 2010
Revised: 4 December 2010
Re-revised: 21 January 2011