e-ScienceTalk

ANNUAL UPGRADED VERSION OF THE RTM

EU DELIVERABLE: D2.3

Document identifier:	e-ScienceTalk_D2_3_GridGuide
Date:	**31/7/2012**
Work package:	**WP2**
Lead Partner:	**Imperial**
Document Status:	**FINAL**
Dissemination Level:	**PUBLIC**
Document Link:	https://documents.egi.eu/document/1299

Abstract

The Real Time Monitor has been upgraded during the course of the second year of e-ScienceTalk and a new version launched to meet D2.3 Annual Upgraded Version of the RTM. This document describes the work carried out on the upgraded version during PY2 as part of this Deliverable.

I. DELIVERY SLIP

	Name	Partner/Activity	Date
From	J. Martyniak	Imperial/WP2	25/07/2012
Reviewed by	**Moderator:** **Reviewers:**	Various	25/07/2012
Approved by	**PMB:**		31/07/2012

II. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	30/07/2012	First draft	J Martyniak & N O'Neill / Imperial & QMUL
2			
3			

III. APPLICATION AREA

This document is a formal deliverable for the European Commission, applicable to all members of the e-ScienceTalk project and its beneficiaries and collaborating projects.

IV. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors.

V. PROJECT SUMMARY

Over the last 10 years, the European Commission and governments have invested substantial funds in distributed computing infrastructures. Scientists have access to state-of-the-art computational and data resources located around the world, putting European research into a leading position to address the greatest challenges facing us today, such as climate change, pandemics and sustainable energy. The advent of the European Grid Infrastructure, combined with the blurring of boundaries between grids, clouds, supercomputing networks and volunteer grids, means that a clear consistent source of information aimed at non-experts is now more important than ever, through dissemination projects such as e-ScienceTalk, that cross national boundaries.

Objectives:

- e-ScienceTalk will build on the achievements of the GridTalk project in bringing the success stories of Europe's e-Infrastructure to policy makers in government and business, to the scientific community and to the general public.

- e-ScienceTalk will work with EGI-InSPIRE and other collaborating projects to expand the scope of the existing GridTalk outputs, and to report on the interactions of grids with e-Infrastructures such as cloud computing and supercomputing.

- The project will explore options for the sustainability of e-ScienceTalk's products.

- e-ScienceTalk will produce a series of reports aimed at policy makers to disseminate key policy issues underpinning grid and e-Infrastructure development in Europe. The project will also coordinate e-concertation activities.

- The GridCafé, GridCast and GridGuide suite of websites will cover new topics and explore novel web technologies; they will integrate closely with GridPP's Real Time Monitor, combining live views of grid activity with the human aspects of computing.

- The growing weekly publication, International Science Grid This Week (iSGTW) will bring news and events to the existing and potential e-Science community under a new name of The Digital Scientist.

VI. EXECUTIVE SUMMARY

In the first year of the e-ScienceTalk project the GridGuide and Real Time Monitor were upgraded and integrated as described in D2.1 GridGuide Upgraded Integration with the RTM. The Real Time Monitor has been upgraded during the course of the second year of e-ScienceTalk and a new version has been launched as D2.3 *Annual Upgraded Version of the RTM*. This document describes the work achieved during PY2 as part of this Deliverable.

The RTM is a real time visualisation of activity on the grid computing infrastructure. The RTM overlays the movement of site activity and job transfers onto the 3D globe, giving users the ability to see the current state of the grid infrastructure. RTM development has focused on four areas; the website, maintenance, user support and extending the application's functionality. In PY2, 66 countries are included in the RTM and the team has visited 10 events where the RTM has been demonstrated. GridGuide gives a human face to the grid, showing the sites and sights of grid computing. Users can listen to podcasts from grid sites worldwide, read about the ongoing work and watch interviews with researchers. There are currently 59 sites on the GridGuide, including 36 EU sites and 23 non-EU in the Americas, Africa and the Asia-Pacific region. Of these, 34 are also currently included in the Real Time Monitor.

During PY2, the e-ScienceTalk team added a new RTM logo designed by the WP2 team to the website. Grid technology is still evolving and sourcing the information about the status of the infrastructure varies from experiment to experiment. During PY2 the middleware used by the majority of sites monitored by the RTM changed subtly and so required some work. Since the start of e-ScienceTalk the RTM has kept abreast of developments of WorldWind and moving to the next full version of the software is being investigated and should be completed during PY3.

During PY2, e-ScienceTalk has contacted the "silent" users of the RTM who run the application during talks or at their institutes to demonstrate the grid. This has highlighted previously unknown usage in the UK, Germany, Greece, Taiwan and New Zealand. There have also been requests to be added to the RTM map from the EUMED grid as well as the Far East and Kuwait.

The functionality of the RTM has also been upgraded. The RTM's developer was approached by the GÉANT to show data traffic on their network, the pan-European data network dedicated to the research and education community, which is now available as a new layer in the RTM. During PY1 the ATLAS job submission system PANDA was added to the RTM. Throughout PY2 the RTM team has been working on implementing the file transfers for the CMS experiment. CMS is also based at the Large Hadron Collider at CERN and uses a system called PhEDEx to monitor and manage data movements. This work is being done concurrently with the upgrade to WorldWind 1.4 so that the visualisations being used are slightly different to the grid traffic layer. Data is presented in a form of pulsating 3D cylinders, two per site, one for incoming data and a second for outgoing traffic.

The aim for PY3 will to complete the transfer to the latest version of WorldWind, fully integrate the CMS data transfers, continue to ensure that the application displays up-to-date and accurate information and investigate displaying data from new sources and infrastructures.

TABLE OF CONTENTS

1   INTRODUCTION

In the first year of the e-ScienceTalk project the GridGuide and Real Time Monitor were upgraded and integrated as described in D2.1 GridGuide Upgraded Integration with the RTM [R1]. The second year focused on improving the content of the GridGuide by expanding the institutes and countries covered while the RTM has incorporated new data sources and infrastructures. There has also been work done on upgrading the code the RTM runs.

The current version of the GridGuide is available at http://www.gridguide.org and the Real Time Monitor can be downloaded as a standalone application from http://rtm.hep.ph.ic.ac.uk/ or launched as a Java webstart version. This document describes the work achieved during PY2.

1.1   The Real Time Monitor

sThe RTM is a real time visualisation of activity on the grid computing infrastructure. The High Energy Physics e-Science group at Imperial College London has been developing it since 2002. Initially funded by the UK-based GridPP collaboration, work on the RTM became a part of the e-ScienceTalk project in September 2010. The latest version utilises the NASA World Wind virtual globe, which is based on OpenGL and Java. The RTM overlays the movement of site activity and job transfers onto the 3D globe, giving users the ability to see the current state of the grid infrastructure. It is modularly designed, making it easy to add and change various aspects of the application depending on the demands placed on it.

The application also includes more detailed information about the status of the individual sites on the grid including:

- Current number of jobs running/queued at a site
- The load on a site's computing elements and workload management systems
- Graphs detailing the work done at the site over various time periods
- Information from the GridGuide about that site, if available

Running the RTM is also relatively straightforward, with no special installation required, on computers with a modern up-to-date operating system. This has widened the use beyond a small community of experts to anyone wanting to look at or discuss the grid infrastructure.

The tools on which the RTM is based are used for more than generating the live visualisations. They also generate important statistic about sites, jobs and virtual organisations, which can be used to help in diagnosing problems or monitoring performance of the infrastructure.

Since becoming a part of the e-ScienceTalk project the RTM development has focused on four areas: the website, maintenance, user support and extending the application's functionality. This has included input from other areas of the project to help give a fresh look to the application.

In PY2, 66 countries are included in the RTM and the team has visited 10 events where the RTM has been demonstrated. These included the EGI Technical and User Forums in Lyon and Munich, and SuperComputing'11 in Seattle, which attracted 10,000 delegates. Partners have also used it at numerous talks and events including Manchester Science Week, Healthgrid 2012 and the TuringFest.

1.2 The GridGuide

GridGuide (www.gridguide.org) is the youngest of the e-ScienceTalk products and gives a human face to the grid, showing the sites and sights of grid computing. Users can listen to podcasts from grid sites worldwide, read about the ongoing work and watch interviews with researchers. As well as giving a visual overview of current grid work, GridGuide enables users to drill down to more detail about an individual scientist's work and how the grid has produced results. For these reasons, the GridGuide is useful for engaging with policy makers who are able to find out more detail about work going on in their local regions or areas of responsibility, as well as the general public and other scientists.

The GridGuide complements the GridCafé by providing a more in-depth guide to institutions across the globe that are involved in grids and distributed computing. GridGuide has become increasingly interactive and accessible through co-development with the Real Time Monitor (RTM), which shows traffic on the worldwide grid in real time. The current integration with GridGuide allows a visitor to click on a site and view both the technical statistics from the RTM as well as the pages from GridGuide. The RTM is widely used for demonstrating the grid at conferences and events and is an accessible and engaging way to understand more about the grid.

Feedback on GridGuide was gathered through email and one-to-one interviews with science communicators and scientists. The GridGuide map has been cited as a good way of demonstrating what the grid does, and provides a draw for people researching grids and e-Infrastructure in their own locality. E-ScienceTalk has focused its efforts to encourage more interest in the GridGuide site. One suggestion would be to provide the information in different formats and to cross-promote by including a link from the GridCafé website. Reviewers have also suggested creating a short video clip showing a few minutes of grid activity on the RTM for marketing and demo purposes. This could include a human aspect with a young scientist uploading data to the grid.

There are currently 59 sites on the GridGuide, including 36 EU sites and 23 non-EU in the Americas, Africa and the Asia-Pacific region. Of these, 34 are also currently included in the Real Time Monitor.
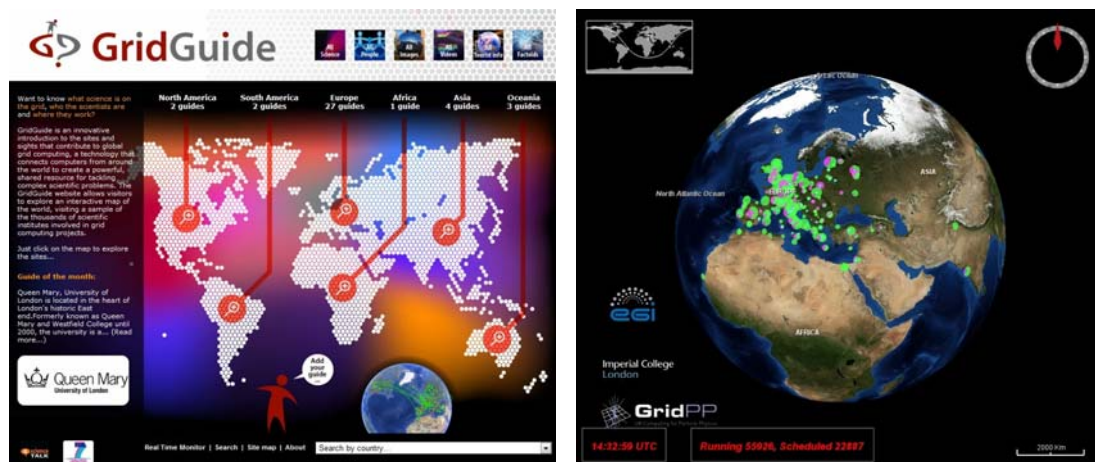
Fig. 1: Screenshot of the GridGuide (left) and Real Time Monitor (right)

2   PROGRESS IN YEAR TWO

2.1   Website and logo

In the last 12 months the RTM website has not undergone any major changes. However as well as keeping the information up-to-date and providing the most recent version of the application, the new RTM logo (http://rtm.hep.ph.ic.ac.uk/headtext.jpg) was added to the site. This was designed by the APO team in WP2 to ensure consistency across the project's different products.

2.2   Maintenance

For the RTM to remain a useful and reliable application it has to monitor two very different areas of technology development. The first is the grid side, where providers and users are still settling on the tools and programmes they are using. The second is the WorldWind code, which is being constantly updated by the developers.

Grid technology is still evolving and sourcing the information about the status of the infrastructure varies from experiment to experiment. During PY1 the RTM integrated the information from the PANDA system used by the ATLAS experiment at CERN (see D2.1 GridGuide Upgraded Integration with the RTM [R1]). During PY2 the middleware used by the majority of sites monitored by the RTM changed subtly and so required some work. This is only temporary as the latest version of the middleware fixes the issue.

Prior to the start of the e-ScienceTalk project the RTM was upgraded to the latest version of the WorldWind software, this was a major undertaking. Since the start of e-ScienceTalk the RTM has kept abreast of developments of WorldWind and the possibility of moving to the next full version of the software is being investigated and should be completed during PY3.

2.3   User Support

Currently the RTM has a wide range of users, with varying levels of computer literacy. This is in stark contrast to early versions, which were used by a small number of technical users. The RTM website is designed to help the new users, with a dedicated mailing list for "power" users to inform them of new developments. There is also an email address for feature requests, bug reporting and support questions.

During PY2 more emphasis was put on contacting and eliciting response from the "silent" users, people running the RTM during talks or at their institutes with whom the team has had little to no contact previously. This has highlighted and alerted the e-ScienceTalk team to the use of RTM in many institutes across the world including the UK, Germany, Greece, Taiwan and New Zealand. There have also been requests from administrators to add them to the RTM map, most from the EUMED grid as well as the Far East and Kuwait.

2.4 Extending the RTM Functionality

2.4.1 GÉANT display

GÉANT is the pan-European data network dedicated to the research and education community. During the EGI Technical Forum in September 2011, the RTM's developer was approached by the GÉANT team about using the RTM to show data traffic on their network.

The grid RTM uses XML formatted files provided by various sources to gather the information needed. Once GÉANT explained what they would like to display, they were provided with a template for an XML file that would work with the existing RTM framework. The file includes a site's geographic position and router activity ie the amount of data flowing in and out of the site.

The RTM server downloads the XML file provided by GÉANT and serves it to any running instances of the network RTM. The file is updated every 24 hours. The new display looks similar to the grid version with pie-charts marking the sites on the globe. The colours and pulsing of the pie charts represents incoming and outgoing router traffic. There is also a histogram with historical data displayed when a site is clicked on.

2.4.2 Upgrading software

The latest stable release of NASA WorldWind is 1.4. During PY2 significant work has been carried out to move the RTM over to this latest version of WorldWind. This is still on-going work in PM23, but the upgrade will improve the user experience, the graphical abilities of the RTM and add more options for new features. The aim is to launch the upgraded version early in PY3.

2.4.3 CMS data transfers

During PY1 the ATLAS job submission system PANDA was added to the RTM. Throughout PY2 the RTM team has been working on implementing the file transfers for the CMS experiment. CMS is also based at the Large Hadron Collider at CERN and uses a system called PhEDEx to monitor and manage data movements.

This work is being done concurrently with the upgrade to WorldWind 1.4 so that the visualisations being used are slightly different. Data is presented in a form of pulsating 3D cylinders, two per site, one for incoming data and a second for outgoing traffic. The cylinder size is related to the amount of data transferred. When a mouse is hovered over an individual site, a summary of the current transfers is displayed. As for the RTM grid job display, the actual file transfers are presented as lines with bullets, with a histogram of transfer activity displayed when a site is clicked.

PhEDEx already published its data online in a XML format so the converting this to something that the RTM uses was relatively straightforward. The only issues is that it does not provide geographical

location information, so a conversion table was needed to map CMS sites to the geographical data provided to the original RTM, allowing the sites to be mapped accurately.
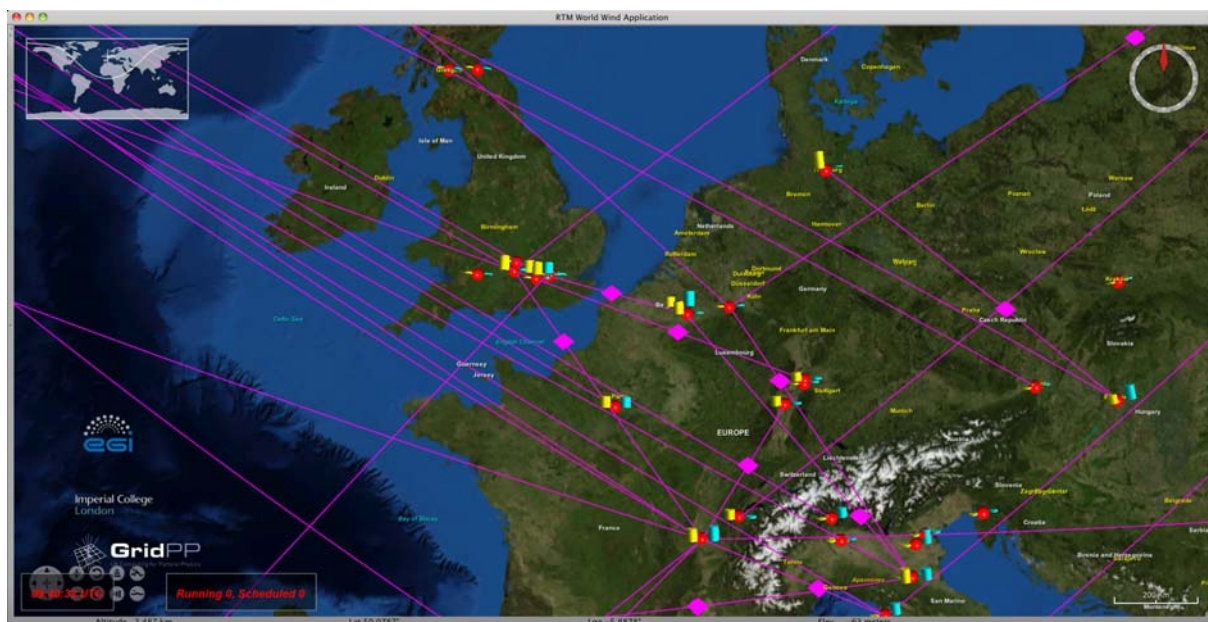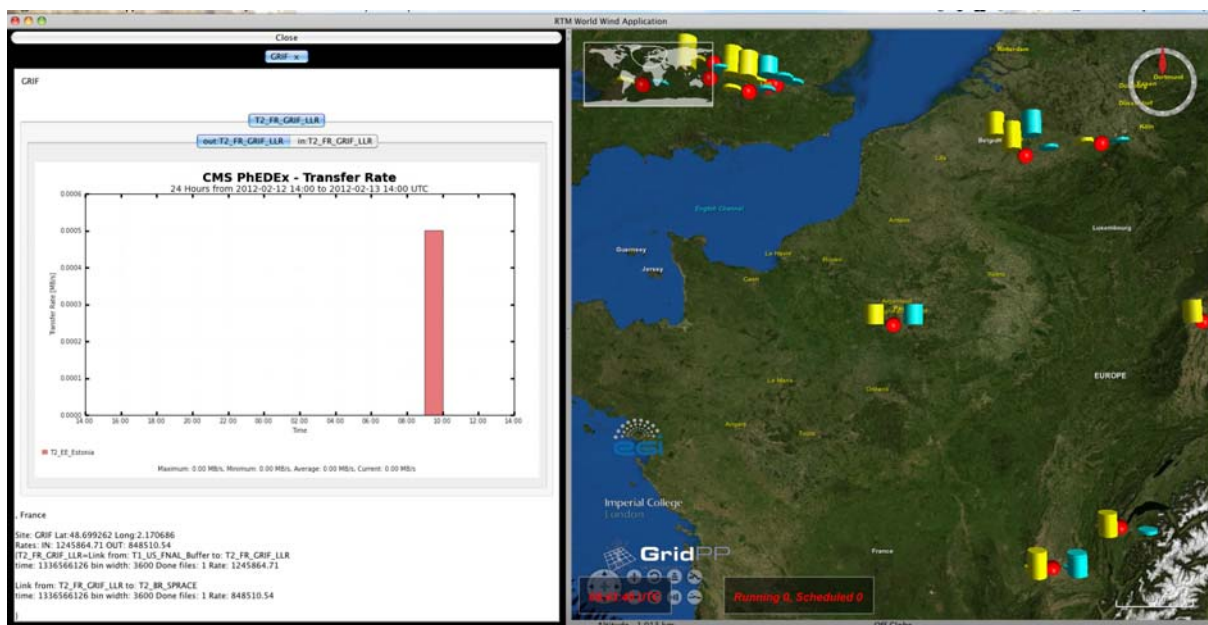

Fig. 2: Screenshot of the CMS data transfers

Fig. 3: Screenshot of the CMs data transfers with transfer rate histogram

3 CONCLUSION

During PY2, the team worked on improving the amount of information provided by the GridGuide, increasing the overall number of guides but also improving coverage outside Europe. The RTM has been upgraded with software updates and the team has integrated new data sources and infrastructures into the previously grid only display. The aim for PY3 will to complete the transfer to the latest version of WorldWind, fully integrate the CMS data transfers, continue to ensure that the application displays up-to-date and accurate information and investigate displaying data from new sources and infrastructures.

4 REFERENCES

R 1	D2.1 GridGuide Upgraded Integration with the RTM https://documents.egi.eu/document/766
R 2	
R 3	
R 4	