Visualization Element: towards the definition of a new Grid service

Giuseppe Andronico[1]

Italian National Institute of Nuclear Physics , Division of Catania
Via S. Sofia 64, 95123 Catania, Italy
E-mail: giuseppe.andronico@ct.infn.it

Roberto Barbera

Department of Physics and Astronomy of the University of Catania and INFN
Viale A. Doria 6, 95125 Catania, Italy
E-mail: roberto.barbera@ct.infn.it

Andrea Fornaia

Italian National Institute of Nuclear Physics, Division of Catania
Via S. Sofia 64, 95123 Catania, Italy
E-mail: andrea.fornaia@ct.infn.it

Marcello Iacono Manno

Consorzio COMETA
Via S. Sofia, 64, 95123 Catania, Italy
E-mail: marcello.iacono@ct.infn.it

Giuseppe La Rocca

Italian National Institute of Nuclear Physics, Division of Catania
Via S. Sofia 64, 95123 Catania, Italy
E-mail: giuseppe.larocca@ct.infn.it

Nowadays, global Grid infrastructures offer production quality services to a large number of Virtual Research Communities (VRCs) and several scientific and industrial applications produce, after long runs, complex output patterns that can be collected and encoded as a unique video which can be easily shown to a wide audience and monitored with the strong impact of multimedia communication. In this contribution we describe the design and prototypal implementation of a new Grid service: the Visualization Element which has been first tested on the Sicilia Grid of COMETA and allows non-expert users to control and show their jobs' evolution using a simple and intuitive web-based interface.

---

[1]   Speaker

1. Introduction

Grid technology, based on open standards and protocols, allows users to share a wide and heterogeneous amount of distributed hardware resources such as supercomputers, storage systems, data repositories and complex instruments regardless their geographical location. This new technology has introduced a new paradigm which has become extremely important for many scientific and industrial applications belonging to several domains such as Astronomy, Computational Chemistry, Earth Science, Financial Simulation, High Energy Physics, Biomedicine, etc. In this scenario, we now have an increasing set of grid applications, especially those producing complex output patterns after long runs (e.g. Fluent [1]), which can be easily monitored and shown to a wide audience with the strong impact of multimedia communication. In this paper we show the technical details of the a new Grid service: the Visualization Element for on-line streaming.

2. The COMETA Grid Infrastructure

The COMETA Grid infrastructure is managed by the COMETA Consortium, a not-for-profit organization established in Catania (Italy) in 2005 and formed by the Universities of Catania [2], Messina [3] and Palermo[4], the Italian National Institute of Nuclear Physics (INFN) [5], the Italian National Institute of Astrophysics (INAF) [6], the Italian National Institute of Geophysics and Volcanology (INGV) [7] and the SCIRE Consortium [8] which brings together both public research institutions and private industrial companies. Fig. 1 shows the locations of the COMETA sites in the three main Sicilian cities of Catania, Messina and Palermo.

Fig. 1 – The COMETA Grid Infrastructure.

The aims of this infrastructure are to: 1) create a virtual laboratory in Sicily, both for scientific and industrial applications; 2) connect the Sicilian e-Infrastructure to those already

existing in Italy, Europe and elsewhere in the world; 3) foster the adoption of Grid computing for massive computations, improve the competitiveness of e-Science and e-Industry "made in Sicily"; 3) trigger the start-up of spin-off in the ICT area; 4) address the endemic problems of "digital divide" and "brain drain" of brilliant people to other parts of Italy and beyond.

Overall, about 2500 CPU cores and more than 250 TB of disk storage space are currently available on the Sicilian e-Infrastructure. The computational and storage resources of the e-Infrastructure are exposed to the end users through the virtual services of the gLite middleware [9], developed in the context of the EC funded project EGEE [10]. The adoption of this middleware makes the Sicilian e-Infrastructure fully interoperable with others already existing in Italy and abroad. In three years of the project many scientific and commercial software (e.g. ABAQUS [11], Fluent, IDL [12], Matlab [13]) have successfully been integrated on the Sicilian e-Infrastructure. Among them, many applications (e.g.: Fluent, Blender [14] and VisIVO [15]) produce, sometimes after several CPU months, complex output images. Up to now, the related long runs are usually monitored by specific tools such as the Watchdog utility [16] that allows the on-line inspection of some selected output files. On the footprints of this technique, a more sophisticated tool was developed in order to encode the image collections generated by an application into a single video that is regularly and automatically updated during the execution of the program. The new tool, which will be better explained in the next sections, integrates the multimedia contents with the Grid technologies in order to provide a more intuitive way to monitor Grid applications. Moreover it supplies a valuable tool for demos, as it can show the birth and evolution of complex structures, even in 3D vision.

3. The Visualization Element service

The design and implementation of the Visualization Element has been driven by the possibility to develop a new Grid service that can be used by non-experts in order to control and show the evolution of long lasting jobs. The implementation of this new service in the COMETA Grid infrastructure has been pursued in two steps.

First, a list of useful additional software packages has been identified. During this phase we focused on open-source solutions released under the GNU General Public License [17]. To encode videos starting from media content, we choose MPlayer MENCODER [18], a free and powerful command line video encoding/decoding and filtering tool able to convert all the formats that MPlayer reads into a variety of compressed and uncompressed formats using different codecs. The ffMPEG [19] software package has been used only when a conversion is needed prior to start the streaming to the client. For this task we relied on the VLC VideoLAN [20] media player which is available on different platforms and supports most of the existing media codecs and video formats. In order to configure the computing resources as on-demand video streamers, this software package has been successfully deployed on the COMETA infrastructure and, last but not least, the `VO-cometa-VLC-0.9.9a` software tag has been created to identify the COMETA sites with this support.

Once the infrastructure was set up, we focused towards the implementation of web-based services exploiting the functionalities of the EnginFrame [21] framework, developed by the Italian company Nice srl [22]. The web-based EnginFrame technology, on which the GENIUS

(Grid Enable web eNvironment for site Independent User job Submission) Grid portal [23][24][25][26] is built on, has been used to provide end-user with a transparent layer to call the new service and make it accessible to all the members of the COMETA Virtual Organisation.

4. Service Implementation

In the first release of the service, two different implementations of the Virtualization Element have been considered. They are shown in figures 2 and 3.
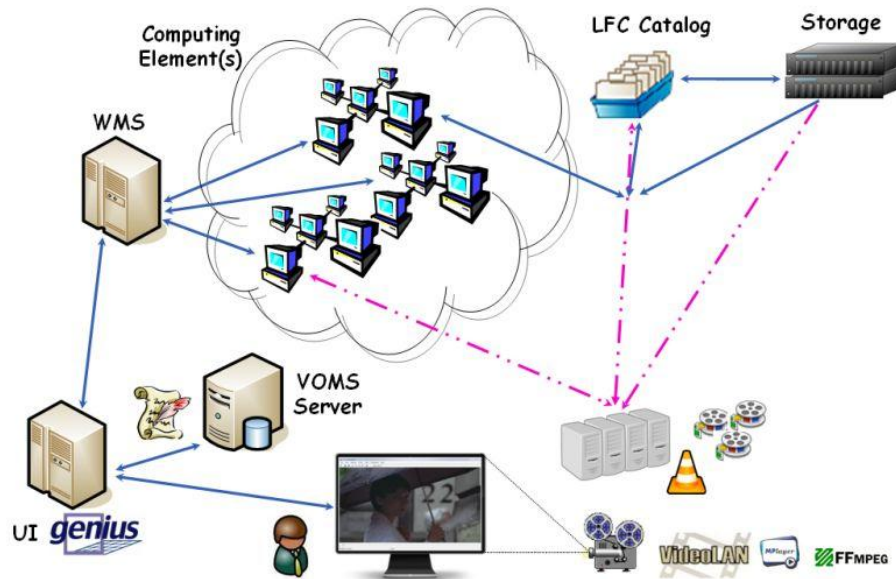

Fig. 2 – Generation of the video output submitting a simple job.

Figure 2 shows the application workflow designed to encode the video output made of the images generated by a producer application running as a Grid job. The client side is represented by the user workstation running a web browser (at the bottom in the figure). The server side is represented by a gLite User Interface (UI), equipped with the latest gLite middleware services to submit jobs and do data management, the Apache Web Server, EnginFrame and the GENIUS portal (on the left in the figure). Once the user has logged in, he/she can provide the input settings before running the service (e.g.: frame rate, video output format, grid location of the TAR file containing the images produced by the application and registered in the file catalogue, see figure 3). When the service is instantiated, the job running on a Worker Node (WN) downloads from the Storage Element (SE) the TAR file produced by application during its execution. The C-based GFAL API has been adopted to speed up the download process of files from the SE. The TAR file downloaded from the SE is processed by MENCODER and an output video is thus created. Before the end of the run, the WN opens an UDP network connection through the default port 1234 to the public IP of the user desktop. This UDP connection is used by VLC to stream the video. An important requirement for the users is to open the firewall in order to receive the UDP traffic on this port.

Fig. 3 – Visualization Element input settings.

Figure 4 shows the application workflow designed to generate the video output with a Direct Acyclic Graph (DAG). In this case, depending on the number of images processed by each job, a DAG (Nx1) is created on-the-fly and submitted by the portal to the Grid.

All the video chunks reconstructed by each node of the DAG are collected by a Collector job. This job is in charge of starting the video streaming to the user. In order to improve the performance of the service, the following additional attributes have been considered: Max_Running_Nodes and NodesCollocation. The first one indicates the maximum number of nodes of the workflow can be simultaneously submitted by DAGMan. This attribute has been set to N, where N is the number of DAG nodes. The second one has been set to 'true' in order to avoid the WMS submitting the nodes of the DAG on a different Computing Element (CE). We need this attribute as the chunks produced by each DAG node will be transferred as inputs by the Collector job.
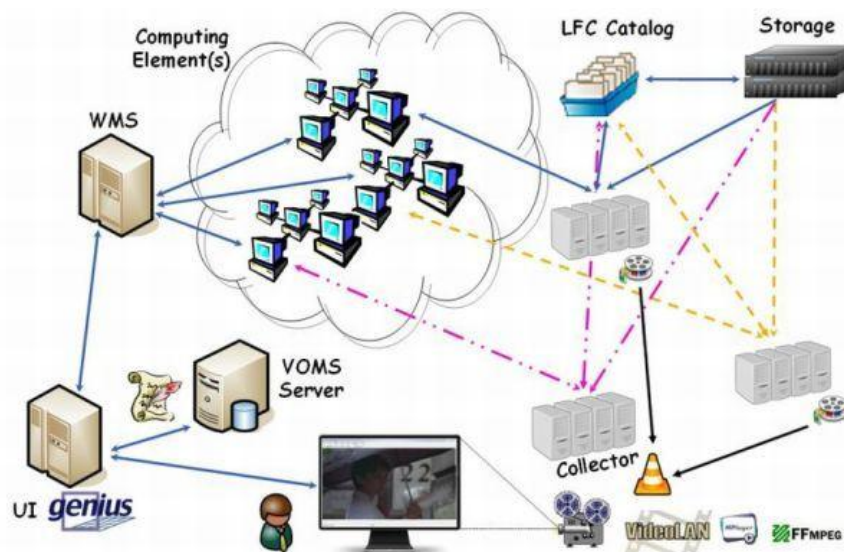

Fig. 4 – Generation of the video output submitting a DAG.

4.1 Stereoscopy

To enhance the visualisation impression of multimedia contents produced by the Grid applications, we have also extended the Visualization Element system with the possibility to create on-demand stereoscopic videos in 3D. The use of this technology, that has much progressed in the last few years, may be of greatl help when one has to show dissemination information to wide audiences, taking advantage on the attractive stereoscopic effect. Furthermore, this technology can also provide a new approach to the simulation's data analysis, watching the produced 3D model in a way never done before, i.e. while the producer application is creating or refining the image, the user can change on the fly the parameters of the simulation as computational errors can be immediately detected instead of waiting until the job ends.

Grid applications, in order to use the stereoscopy feature of the Visualization Element, have to produce not only one stream, as in the solution shown in the previous section, but two, corresponding to different points of view. Each application has to get two different streams from the simulation, as there were two displaced cameras that generate separated images, one for each eye. The way in which this goal is reached can really make the difference between a good stereoscopic video and one that is not [27]. The two methods shown in figure can be used in alternative to produce the stereoscopic vision. The toe-in method simply adopts two cameras with non-parallel optical axes. This is easy to implement, as many rendering software support it, but introduces some vertical parallax that leads to a false reproduction of depth. The off-axis method is preferable as it causes only minor distortions.
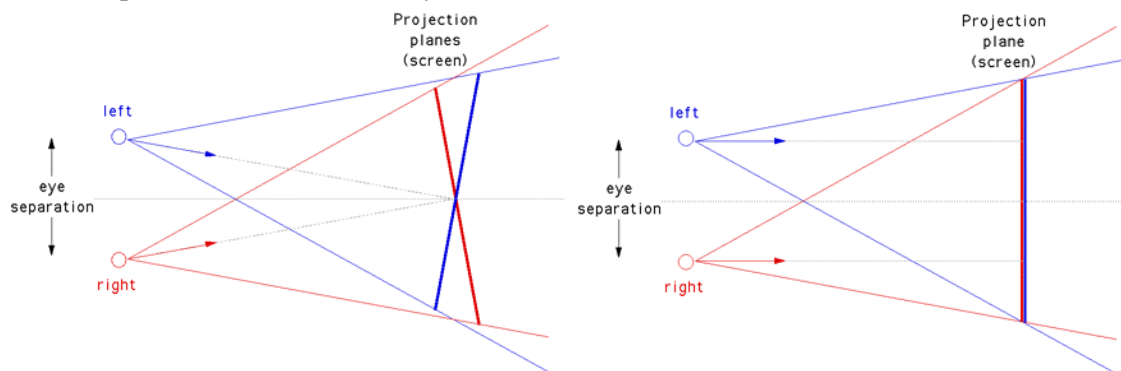
Fig. 5 – Camera parallax and orientation methods: toe-in (on the left) and off-axis (on the right).

Once a producer application has created the left and the right streams, it must store them coupled on the same Storage Element (SE): The Visualization Element loads them in the same way as for the single stream solution, but with an important difference: before releasing the final VLC stream to the end user, the system has to combine the two offset streams in a single one, using appropriated video drivers in according to the chosen stereoscopic technique.

There are several methods to obtain a stereoscopic effect from a pair of left and right images [28], that differ in the optical method used to drive each image in the correct eye. Among them, we mention Head-Mounted Display (HDM), Polarized light, Shutter glasses, Auto-stereoscopy displays, and Anaglyph. These methods not only differ for the optical technique used to drive the separated streams, but also in the equipment required and,

consequently, in the overall cost and portability. In any case, anyway, the practical difference from a developer point of view consists in choosing the right *codec* to use, in order to allow a correct understanding of the image streams (generated from the Grid applications) by the used stereoscopic reproduction system.

In our work, we decided to focus on the Anaglyph method because of his portability, cheap distribution costs and easy-to-watch characteristics

In figure 3 we can see how we have introduced the "Stereoscopic Type" field that in the GENIUS-based interface of the Visualization Element that allows users to choose one of the supported complementary colour couples, and the "Method" field to choose among True Colour, Half Colour and Optimized. In this version we have enhanced the use of the Anaglyph support, due to the benefits described above, but features can be easily extended to other stereoscopic systems by simply implementing the appropriate codec (like the side-by-side one), with no need to change the rest of the application, neither the stream acquisition method nor the way to convey the stereoscopic content to the final user.

5. Results

Fluent is a very well know commercial application (a de facto standard) for computing fluid dynamics and it is deployed on the COMETA Grid Infrastructure. This software has been used as one of the producer applications to test the current version of the Visualization Element. With Fluent has been simulated the Le Marmore water falls in Italy (see Fig. 6) for environment protection purposes. For this simulation, 64 CPUs cores of the COMETA Grid Infrastructure were used in parallel (through MPI) for 2 entire weeks.
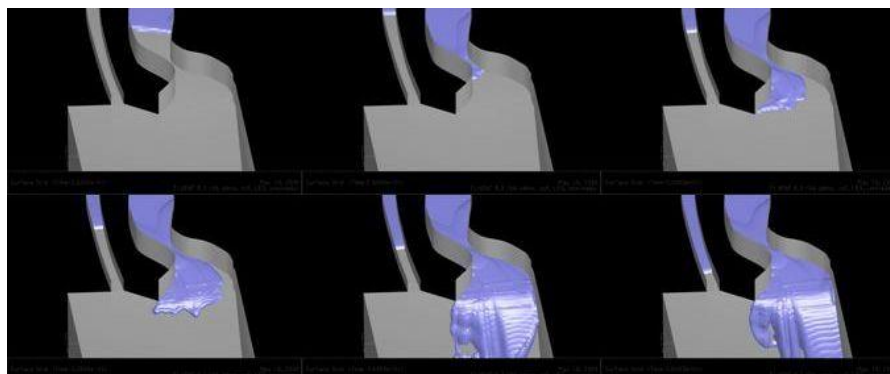
Fig. 6 – Some frames of the sequence produced by FLUENT and
streamed by the Visualization Element.

The Visualization Elements, with the support for stereoscopy, has been successfully tested with two other producer applications: Blender and VisIVO. In the first case, the features of this free open source 3D content suite have been used to create the left and right images of the fixed scenario. The media content produced by Blender has been encoded by the Visualization Element in order to produce a stereoscopic video which was sent back to the calling client. VisIVO is the Visualization Interface used by many astrophysicists involved in the COMETA Consortium to render the output of astrophysical simulations. VisIVO can handle both

observational and theoretical data. Among others, the FITS and HDF5 data formats are supported. VisIVO can be used both as a stand-alone application, that manipulates local files, and as an interface to the Virtual Observatory framework, from which it can retrieve data. In this second use case, the Visualization Element was used to encode astrophysical data of stars and shows the results to the members of the community.

Conclusions

The on-line streaming tool introduced in this paper, configured as a new Grid service and embedded in the GENIUS Grid portal, allows users to perform a better monitoring of their applications that produce complex output patterns. The service has been built on top of the gLite-3.2 stack. Thanks to this new service, the result of the Grid calculation can be easily shown to wide audiences of taking advantage from the strong impact of multimedia communication also with 3D effects.

Acknowledgments

References

[1]   Fluent – www.fluent.com

[2]   University of Catania – www.unict.it

[3]   University of Messina – www.unime.it

[4]   University of Palermo – www.unipa.it

[5]   The National Institute of Nuclear Physics (INFN) – www.infn.it

[6]   The National Institute of Astrophysics (INAF) – www.inaf.it

[7]   The Italian National Institute of Geophysics and Volcanology (INGV) – www.ingv.it

[8]   The SCIRE Consortium – www.consorzioscire.it

[9]   gLite – Lightweight middleware for Grid Computing – glite.cern.ch

[10]   The EGEE Project – http://www.eu-egee.org/

[11]   ABAQUS software package – www.simulia.com/products/abaqus_fea.html

[12]   IDL software package – http://idlastro.gsfc.nasa.gov/idl_html_help/home.html

[13]   Matlab– The Language of Technical Computing – www.mathworks.com/products/matlab/

[14]   Blender software package – www.blender.org

[15] The Visualization Interface to the Virtual Observatory (VisIVO) – http://visivo.oact.inaf.it/index.php

[16] The Watchdog Utility – https://grid.ct.infn.it/twiki/bin/view/ConsorzioCOMETA/WatchdogUtility

[17] The GNU General Public License – www.gnu.org/licenses/gpl.html

[18] Mencoder – www.mplayerhq.hu/design7/news.html

[19] ffMPEG – www.ffmpeg.org

[20] The VLC VideoLAN Client – www.videolan.org

[21] The XML/Java-based EnginFrame framework – www.enginframe.com

[22] NICE srl – www.nice-italy.com/web/nice/home

[23] R. Barbera, G. Donvito, A. Falzone, G. La Rocca, G. P. Maggi, L. Milanesi, *GENIUS Grid Portal and robot certificates to perform phylogenetic analysis on large scale: an experience within the Italian LIBI Project*. Managed Grids and Cloud Systems in the Asia-Pacific Research Community – Lin, Simon C., Yen, Eric (Eds.), 2010, XII, ISBN 978-1-4419-6468-7

[24] R. Barbera, G. Donvito, A. Falzone, G. La Rocca, G. P. Maggi, L. Milanesi, *The GENIUS Grid Portal and Robot Certificates* – Final Workshop of Grid Projects "PON RICERCA 2000-2006, AVVISO 1575", March 2009, Catania – Italy– ISBN: 978-88-95892-02-3

[25] R. Barbera, G. Donvito, A. Falzone, G. La Rocca, G. P. Maggi, S. Vicario, L. Milanesi, *The GENIUS Grid Portal and robot certificates: a new tool for e-Science* — BMC Bioinformatics 2009, 10 (Suppl 6) doi:10.1186/1471-2105-10-S6-S2

[26] R. Barbera, G. Andronico, G. Donvito, A. Falzone, J. J. Keijser, G. La Rocca, L. Milanesi, G. P. Maggi and S. Vicario, *A grid portal with robot certificates for bioinformatics phylogenetic analyses*, Concurrency and Computation Practice & Experience, Special Issues IWPLS 2009 – John Wiley & Sons. Ltd, Nov. 2010, Vol. 23, Issues 3, pages 246-255, doi:10.1002/cpe.1682

[27] P. Bourke, *Calculating Stereo Pairs* – http://paulbourke.net/ – July 1999

[28] L. Lipton, *The Stereoscopic Cinema: From Film to Digital Projection* – SMPTE Journal, September 2001