Workflows and Science Gateways for Astronomical Experiments

Giuliano Taffoni,[1] Ugo Becciani,[2] Mel Krokos,[3] Eva Sciacca,[2] Alessandro Costa,[2] Fabio Vitello,[2] Claudio Vuerli,[1] and Fabio Pasian[1]

[1]*INAF - Astronomical Observatory of Trieste, Via G.B. Tiepolo 11, 34143, Trieste, Italy*

[2]*INAF - Astrophysical Observatory of Catania, Via Santa Sofia 78, 95123, Catania, Italy*

[3]*School of Creative Technologies, University of Portsmouth, Portsmouth, United Kingdom*

Abstract. Workflow and science gateway technologies have been adopted by scientific communities as a valuable tool to carry out complex experiments. They offer the possibility to perform computations for data analysis and simulations, whereas hiding details of the complex infrastructures underneath. In this paper we describe our experiences in creating workflows oriented science gateways based on gUSE/WS-PGRADE technology that allow to build user–friendly science gateways for Astronomers.

1. Introduction

Workflows have emerged as a new paradigm for researchers to formalise and structure complex scientific experiments in order to enable and accelerate scientific discoveries. Workflows system combined with Science Gateway (SGW) technologies are used to provide a technological framework that integrates an enriched web user interface with a solid computational engine to orchestrate scientific applications and tools. The computational processes executed bye SGWs are organised as scientific workflows that explicitly specify dependencies among underlying tasks for orchestrating distributed computational resources[1] appropriately. SGW as defined here is a community-developed set of tools, applications, and data that is integrated in a web-based graphical user interface.

A large number of workflow management systems can be used to orchestrate workflows, these systems differ in terms of workflow description languages and workflow engines (Deelman et al. 2009). Moreover, research communities have developed a large numbers of workflows to run their experiments. This often has a profound impact on the resulting workflow performance, development effort, management and portability. It takes significant effort and time to learn how to use workflow systems, and requires specific expertise and skills to develop and maintain workflows. Moreover, this results in duplicating efforts when the same or similar workflows are developed using different workflow engines.

[1]e.g.: clusters, HPC resources, grids or cloud computing resource

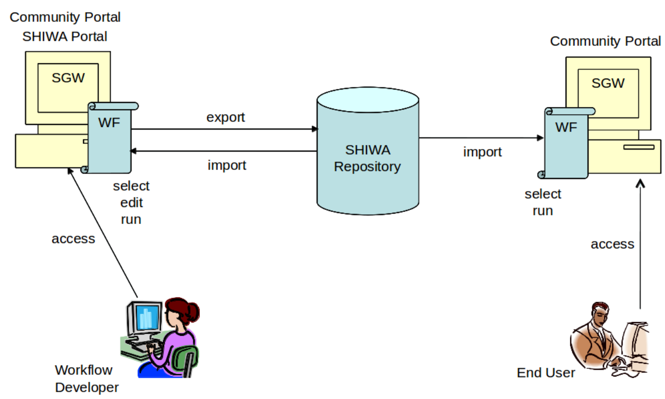

Figure 1. The SHIWA Simulation Platform. In this picture the main SHIWA simulation platform components are presented: the repository and the portal.

In this paper we describe our experience in creating workflows oriented science gateways based on gUSE/WS-PGRADE technology (Kacsuk et al. 2012). This technology allows astronomers to develop workflows and science gateways using any workflows engine (e.g. Kepler, Taverna, Mouter) and workflow description languages. gUSE/WS-PGRADE allows researches to recycle workflows already developed speeding up the development and implementation of new workflows and SGWs.

2. Workflow Technology

A workflow is a formal specification of a scientific process, which represents, streamlines and automates the analytical and computational steps that scientists need to go through from data selection and integration, computation and analysis to final data presentation and visualization.

Workflows technology has been widely adopted by Astronomers. For example The European Southern Observatory (ESO[2]) implement an environment to automate data reduction workflows, the Recipe flexible execution workbench (Reflex). In the Workflow4Ever project[3], the Astronomers developed more than 50 workflows using Taverna[4] and the AstroTaverna[5] plugin.

In our work, we exploit the results of the SHIWA project[6] to implement sharing and exchanging of workflows between workflow systems and distributed computing infrastructures resources through the SHIWA simulation platform.

[2] ESO: http://www.eso.org

[3] Workflow4Ever project: http://www.wf4ever-project.org

[4] Taverna: http://www.taverna.org.uk

[5] AstroTaverna: http://amiga.iaa.es/p/290-astrotaverna.htm

[6] http://www.shiwa-workflow.eu

From the user's perspective the most important services offered by the SSP are shown in Figure 1:

- The SHIWA Repository[7]: A database where workflows and meta-data about workflows can be stored. The database is a central repository for users to discover and share workflows within and across their communities.
- The SHIWA Portal[8]: A web portal integrated with the SHIWA Repository that enables the execution of SHIWA repository workflows.

The main reason we adopt this technology, it that SSP allows workflow developers to design a workflows that combines together modules written for different workflow management systems. The SHIWA Portal is based on the gUSE/WS-PGRADE technology and the portal technology is based on Liferay[9] portal framework.

The generic WS-PGRADE portal instance is easily customisable into a research domain specific science gateway. Using JAVA portlets it is possible to develop a web user interface to provide input parameters, to execute applications and to display the results in a user-friendly way. Each application specific portlet contains the details of the related underlying workflows.

3. Workflows and Gateways for Astronomy

Using the SSP and the gUSE/WS-PGRADE technology we developed several SGWs and associated workflows focused on different applications. COMCAPT (Capture of comets from the interstellar space by the Galactic tide) is a SGW that focuses on applications related to studies of small bodies in the Solar system. FRANEC (Frascati Raphson Newton Evolutionary Code) allows to execute stellar evolutionary code on a grid distirbuted computing infrastructure (Taffoni et al. 2010). LaSMoG is a portal to visualise Large Simulation for Modified Gravity (Zhao et al. 2011). MESTREAM (Modelling the dynamical Evolution of meteoroid stream) is a SGW that allows Astronomers to calculate the dynamical Evolution of meteoroid streams. Planck (Simulations of the ESA Planck satellite mission) SGW is designed to execute simulation of Planck satellite mission developing a web application of the Planck simulation software (Reinecke et al. 2006). VisIVO (Visualization Interface for the Virtual Observatory) (Sciacca et al. 2013) provides visualization and data management services to the scientific community exploiting the functionalities of VisIVO.

Recently a federation of Astrophysics-oriented science gateways, named STARnet, has been designed and implemented by Becciani et al. (2014). STARnet is based on SHIWA technology and it envisages sharing a set of services for authentication, distributed computing infrastructure (clusters or DCIs), data archives and workflow repositories.

---

[7] `http://shiwa-repo.cpc.wmin.ac.uk`

[8] `http://shiwa-portal2.cpc.wmin.ac.uk/liferay-portal-6.1.0`

[9] `http://www.liferay.com`

4. Conclusion

In the last four years, a large effort has been devoted to develop SGWs and workflows for Astronomical and Astrophysical applications. Our efforts allow to set up a number of SGWs that provide the Astronomers with a set of tools towards facilitating the use of complex computing resources (e.g. cloud or HPC resources) by demonstrating benefits of using this approach in doing science.

In this paper we briefly describe a technology to develop SGWs and workflows that allow to speed up the process of designing, implementing, deploying and maintaining workflows and SGWs. This technology allows astronomers to develop workflows and science gateways using any workflows management system (e.g. Kepler, Taverna, Mouter) and even combining them together. Astronomers can use their preferred workflows system and recycle workflows. Although the major obstacle of workflow recycling is that workflow systems are not normally compatible, our adoption of this framework allows to overcome this limitation. We notice that this approach improves efficiency and reliability by reusing tested methodologies, it increases the lifetime of workflows and it reduces development time for new workflows and consequently science gateways.

This approach allows to easily develop science gateways and it could represent an effective solution for projects that needs to provide web based data analysis applications, visual analytics or hiding to the Astronomers all the complexity of the software and focus only on the scientific analysis.

NOTA: CITA VISIVO E WS-PGRADE

Acknowledgments. The research leading to these results has received funding from the FP7 project under contract no 312579 ER-flow (Building an **E**uropean **R**esearch Community through Interoperable Work**flow**s and Data)

References

Becciani, U., et al. 2014, in Science Gateways for distributed computing infrastructures (Springer International Publishing), XIX, 301

Deelman, E., Gannon, D., Shields, M., & Taylor, I. 2009, Future Generation Computer Systems, 25, 528 . URL `http://www.sciencedirect.com/science/article/pii/S0167739X08000861`

Kacsuk, P., Farkas, Z., Kozlovszky, M., Hermann, G., Balasko, A., Karoczkai, K., & Marton, I. 2012, Journal of Grid Computing, 10, 601

Reinecke, M., Dolag, K., Hell, R., Bartelmann, M., & Enßlin, T. A. 2006, Astronomy & Astrophysics, 445, 373. `astro-ph/0508522`

Sciacca, E., Bandieramonte, M., Becciani, U., Costa, A., Krokos, M., Massimino, P., Petta, C., Pistagna, C., Riggi, S., & Vitello, F. 2013, in 5th International Workshop on Science Gateways, IWSG 2013 (CEUR Workshop Proceedings)

Taffoni, G., Cassisi, S., Manzato, P., Molinaro, M., Pasian, F., Pietrinferni, A., Salaris, M., & Vuerli, C. 2010, Journal of Grid Computing, 8, 223

Zhao, G.-B., Li, B., & Koyama, K. 2011, Physical Review D, 83, 044007