Astronomical Data Analysis Software and Systems XVII **P1-8**
ASP Conference Series, Vol. XXX, 2008
J. Lewis, R. Argyle, P. Bunclarck, D. Evans, and E. Gonzales-Solares, eds.

Web services at TERAPIX

Olivier Ricou[1], Anthony Baillard[2], Emmanuel Bertin[2], Frédéric Magnard[2], Chiara Marmo[2] and Yannick Mellier[2]

[1] *EPITA Research and Development Laboratory (LRDE), Le Kremlin Bicêtre, France*

[2] *Institut d'Astrophysique, Paris, France*

Abstract. We present an implementation of V.O.-compliant web services built around software tools developed at the TERAPIX centre. These services allow to operate from a remote site several pipeline tasks dedicated to astronomical data processing on the TERAPIX cluster, including the forthcoming EFIGI morphological analysis tool.

1. Introduction

The EFIGI[1] (Extraction de Formes Idealisées de Galaxies en Imagerie) project hosted at TERAPIX[2] proposes to address both the computational and algorithmic aspects of the extraction of useful morphological information from a large number of galaxies. Measuring galaxy morphology is computationally intensive, but requires only modest amounts of data to be analysed for each detection (the necessary data bandwidth is typically 20kB/s); it is therefore particularly suited to implementation as a web service.

Instead of developing a web service restricted to galaxy morphology measurements, we investigated the possibility to offer a simple and generic mean to access several services related to astronomical image analysis/processing. These services have in common that they are all directly related to executables that work on or generate FITS data files and XML-VOTable metadata files in batch mode. The executables have been developed and are maintained in-house, which means that they can easily be modified if needed. Since different users have different needs, we provide three ways to run our services:

- a web form for testing or occasional runs.
- a web service to pipe to other services or to include in programs.
- a Globus interface — the gate to Grid computing — for more intensive uses.

[1] `http://www.efigi.org`

[2] `http://terapix.iap.fr`

2. Technology and conformance to standards

The astronomical community is currently involved in an international effort to normalise the format of metadata and web service protocols: the Virtual Observatory (VO). It is therefore logical to design a system which conforms as much as reasonably possible to the VO recommendations:

- The tools designed at TERAPIX provide support for VOTables in output (Bertin & Tissier 2007), although it is not yet clear whether this standard will remain popular in the future.
- Our web service prototype transfers files with MTOM according the new VO recommendations. Unfortunately this new protocol is rarely included in Web service libraries. Java and .NET can run MTOM but not Python nor Perl.
- The Grid service, based on Globus, offers Reliable File Transfer (RFT), security, accounting and the ability to connect to other VO Grid like the Japanese VO (Ohishi et al. 2004) or the German AstroGrid (Henke et al. 2007).

Our software architecture relies on Condor[3] to dispatch the jobs on the TERAPIX cluster according to user authorisations, providing the same interface for internal and external use. TERAPIX insiders can directly operate Condor while external users must connect to one the three services which hand the job to Condor (see Fig. 1). Of course, priority is given to internal users, however regular external users can be registered to increase their privileges. A future extension might be to connect to the registry of the VO for authentication.

The services, as displayed in Fig. 1, rely on Apache for the CGI, on Axis 2[4] for the SOAP web services using MTOM for the file transfer, and on Globus for the grid access, including authentication with GSI and file transfer with RFT.

3. How to use it

The Web form `http://efigix.iap.fr/ws/efigi_wi.html`. Fill the form, press "Run EFIGI". A mail is sent at the end of your run, containing a temporary link to the results.

The Web service `http://efigix.iap.fr/ws/ws.html`. Get the WSDL description of the EFIGI service to generate the stubs and data types in your favorite language as long as it knows MTOM (essentially restricted to Java and .net at this time), and write your own client to call the service. An example of a client program, written in Java, is provided on the EFIGI web site. It uses the Axis 2.1.3 version of WSDL2Java.

The Grid service `http://efigix.iap.fr/ws/globus.html`. Install the Globus Toolkit[5] version 4. Get a one week certificate as a guest user or ask us for a

[3] `http://www.cs.wisc.edu/condor`

[4] `http://ws.apache.org/axis2`

[5] `http://www.globus.org/toolkit`

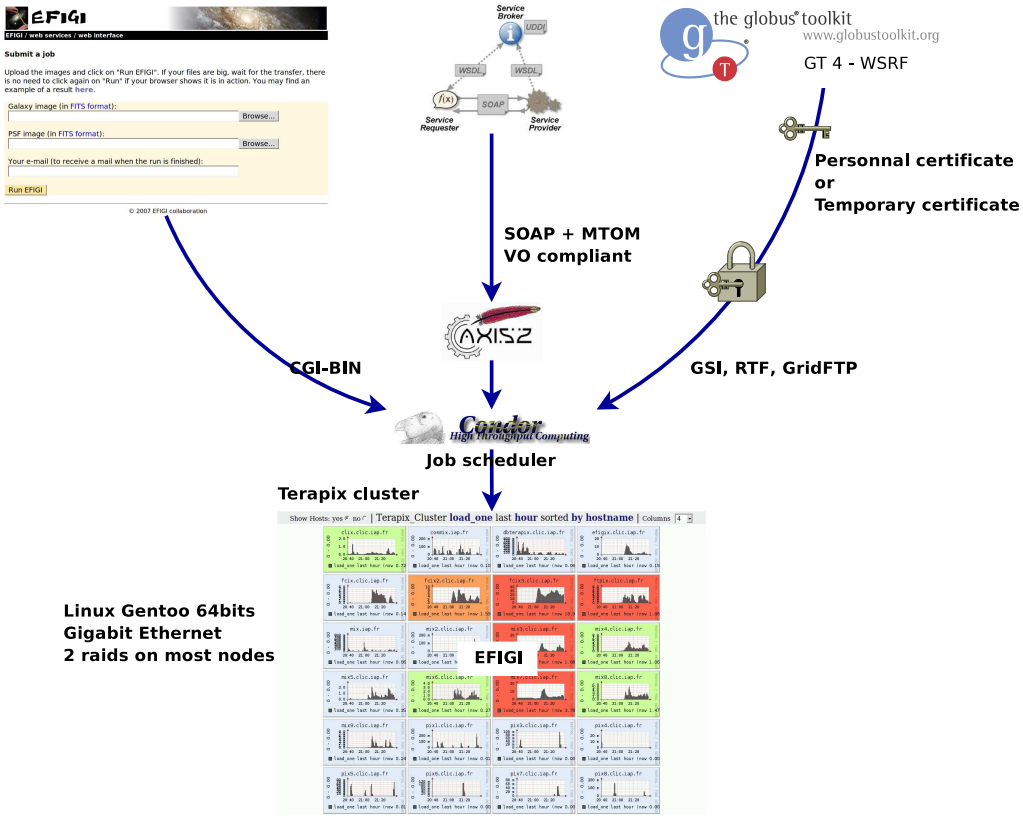

Figure 1. The software architecture of the different services

long-term certificate. Prepare your submission file and submit it. In principle only the client part of Globus is required, but in practice it is advised to install the full distribution to take advantage of all Globus features.

4. The current EFIGI prototype

The current EFIGI pipeline prototype is based on modified versions of several TERAPIX tools[6] including nFIGI, PSFEx, SExtractor, and STIFF. It requires two FITS images in input: the first image contains one or several galaxies to be measured and the second image contains the Point Spread Function (it may actually contain a mix of galaxies and point sources from which a proper model will automatically be derived). The final product consists of an XML VO-Table containing the morphological measurement vectors and diagnostic images. An XSLT filter is provided to present the results in a user-friendly way (Fig.

[6] http://terapix.iap.fr/soft

2). The current version of the EFIGI web service has limited functionalities and is only meant for testing. We expect the final version and additional services to be online in December 2007. Announcement will be made on the TERAPIX forums (`http://terapix.iap.fr/forum`).

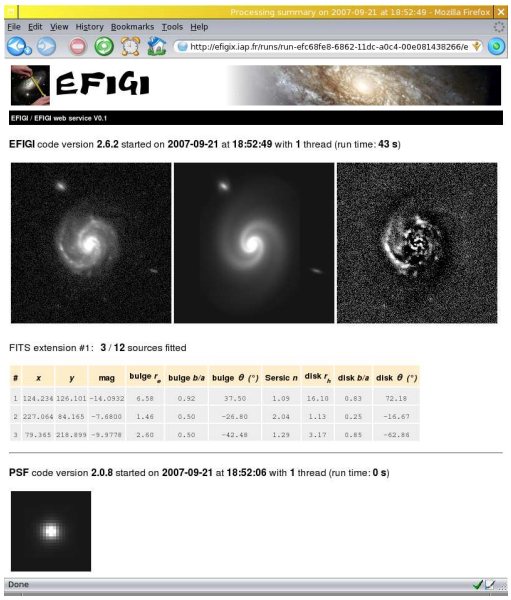

Figure 2. Example of an XSL-transformed result from the current EFIGI prototype.

5. The future

We intend to provide web services based on other CPU-bound TERAPIX tools in the near future. Managing efficiently data-intensive tasks on a grid-computing environment is more problematic, and will be investigated at later times.

Acknowledgments. This work is supported by grant 04-5500 ("ACI masse de données") from the French Ministry of Research.

References

Baillard A., Bertin E., Mellier Y., McCracken H.J., Géraud T., Pelló R., Leborgne J.-F., Fouqué P., 2006, in ASP Conf. Ser. 351, ADASS XV, ed. C. Gabriel, C. Arviset, D. Ponz, & E. Solano (San Francisco: ASP), 236

Bertin et al., 2002, in ASP Conf. Ser. 281, ADASS XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco: ASP), 228

Bertin & Tissier, 2007, in ASP Conf. Ser. 376, ADASS XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco: ASP), 507

Enke H., Steinmetz M., Radke T., Reiser A., Röblitz T., Högqvist M., 2007, in Proceedings of the German e-Science Conference, May 2-5, 2007, Baden-Baden, Germany

Ohishi et al., 2004, in ASP Conf. Ser. 314, ADASS XIII, ed. F. Ochsenbein, M. Allen, & D. Egret (San Francisco: ASP), 296