PICNIC
Portal-based Platform for MRI Processing of Neurodegenerative Diseases

J. Delgado-Mengual, Y. Vives-Gilabert, A. Sainz-Ruiz, M. Delfino-Reznicek
Port d'Informació Científica, UAB Universitat Autònoma de Barcelona, Edifici D Campus Bellaterra, Barcelona, Spain

B. Gómez-Ansón
Hospital Sant Pau i de la Santa Creu, Barcelona, Spain

Keywords: Web-applications, Web-portals, Neuroimaging, Telemedicine.

Abstract: The use of medical image processing techniques is increasing, especially those applied to the early diagnosis of diseases, like neurodegenerative diseases. The software tools involved in it, are sometimes hard to use for medical researchers and hospitals don't have neither the hardware resources nor the personnel expertise, to accomplish the requirements. With these necessities born PICNIC, a technological hardware-software platform with a web portal interface that integrates MRI processing tools with a user-friendly interface, a database to manage clinical data and other services like the de-identification, visualization images and job's monitoring.

1 INTRODUCTION

With the increase of the magnetic field generated by the scanners and the consequent increase in the resolution of MRIs (Magnetic Resonance Images), imaging techniques are becoming an important reference in the early diagnosis of a lot of diseases. For this reason, clinical researchers need technological platforms dedicated to the storage and post-processing of data for the analysis of medical images.

Thinking in a scientific community that travels around the world, web-portal technologies offer a lot of possibilities to work remotely, and it also reduces the required hardware resources and the informatics knowledge of the users: only an actual Internet browser and a Java Plug-in for the applets involved in the portal are required. The high portability is also an important issue taking into account that medical researchers have to be able to access with different browsers, regardless their operating system.

Moreover, some of these services have a strong graphical interaction, so a platform needs a mechanism to integrate the interfaces of the software packages on the web portal with a user-friendly interaction. Furthermore, the web portal has to personalize the menus according to the user's profile, to offer a highly safe and controlled environment.

The current state-of-the-art regarding web-portals and technological platforms in neuroimage processing is nowadays centred in complex data pipelines. However, some of these advanced platforms become complicated for medical researchers due to the huge amount of specialized software packages that they offer, which require a high degree of understanding before using them. GRID environments have been introduced in similar projects with a most global scope.

PICNIC (PIC NeuroImaging Center) platform is aimed to support research of neuroradiologist communities whose data is centralised, so they don't require the GRID technology to share data but a local cluster that allows to post-process MRIs in parallel.

2 REQUIREMENTS AND GOALS

PICNIC is designed with the main objective of unifying the access to the data and to the different post-processing tools, until now installed in different hardware systems and accessible via different operating systems. The idea is to make the

Delgado-Mengual J., Vives-Gilabert Y., Sainz-Ruiz A., Delfino-Reznicek M. and Gómez-Ansón B. (2010).
PICNIC - Portal-based Platform for MRI Processing of Neurodegenerative Diseases.
In *Proceedings of the First International Conference on Bioinformatics*, pages 239-244
DOI: 10.5220/0002745202390244

interaction with the data and the software as easy as possible for the medical researchers. In this section, the requirements in terms of management of the data, software, services and security are described.

2.1 Data Requirements

The users need to work with many types of images and formats (DICOM, ANALYZE, NIFTI or PFILES), depending on the project. Therefore, researchers need a database of images' metadata, where other information like the age or the gender can be also associated to the images (very useful to perform clinical studies). This database has to link with other applications of PICNIC like the post-processing tools or an image viewer, in order to allow the users to interact with this data.

2.2 Post-processing Tools and Services Requirements

The requirements associated to the software packages concern the integration of different hardware platforms and environments (operating systems) to a unique interface. It is also required that other tools can be implemented to PICNIC in the future. The interface between the user, the software and the hardware implies different issues: the design of a GUI (Graphical User Interface) for the tools that don't have any, the development of scripts to run the processes in parallel (when possible), to the export of the results to excel files, etc. This interfaces and the rest of web services have to be implemented following the last HTML and CSS standards, validated by the W3C to guarantee the portability between the actual browsers, like Internet Explorer, Mozilla Firefox, Safari, Opera, Google Chrome.

Regarding the services' requirements, it is necessary to implement an application to transfer the data to the centre remotely via internet and also to install a viewer to visualize MRIs. It is also necessary to ensure that the data stored in the system is correctly de-identified.

2.3 Security Requirements

It is crucial to keep personal data of the users and to perform the authentication process as safe as possible. Some restrictions have to be applied from the network point of view, in order to control the connections to the server.

3 NEUROIMAGING PLATFORM

The platform is designed with a modular structure to satisfy the needs of the medical community in the present as well as in the immediate future. As represented in Figure 1, there are three basic areas in the architecture scheme: User's area, Data area and Post-Processing area. Every area is interconnected to the others through a machine where the following packages are installed: a CentOS, an Apache server, mySQL, NX server and different post-processing tools.

The whole system is strictly controlled with a white list policy of security, where any action or access is not permitted by default and the user has to contact the Administrator to ask for permissions.

So, the "Administrator's" role is very important to mediate the interaction between the final user and the portal, in terms of permissions, job monitoring, and other aspects related to the database and the interaction with the hardware and the applications.

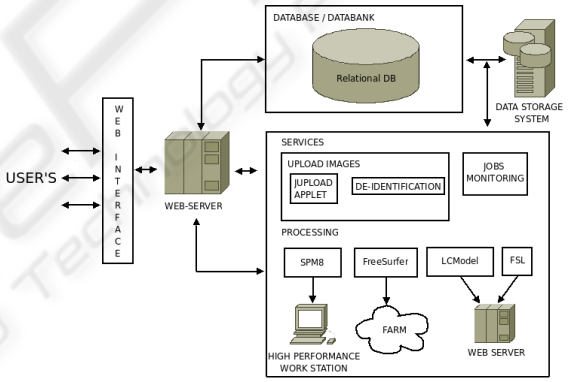

Figure 1: Architecture Scheme.

3.1 User's Area

The different user's profiles that will interact with the platform are: radiologists, neurologists, psychologists and the administrators of the portal.

Administrators have a full access to all functionalities, but the rest of the users have to ask for permissions to use the software, services and also to access to the data of research projects. The combination of all of these permissions makes the platform personal, in terms of access to the information stored in disk, sections, software tools, functionalities.

The advantages that PICNIC offers to the users are:

- The software packages and the system are correctly installed and updated in adequate machines with a properly operating system.

- The data (images, post-processing results...) is properly organized in high available hardware resources.
- The data is maintained in a secure and consistent manner, with an adequate backup policy.
- The software assistance, parametrization and configuration, and analysis of the jobs logs are covered by personnel formed with the software.

There is an opened communication line between the users and the administrators to send suggestions, issues and problems, to improve the platform and the services offered.

3.2 Data Area

The storage is one of the main problems of clinical researchers that need a large amount of data to obtain significant statistical results from their research. PICNIC offers different possibilities of storage systems according to the importance of the data (FAS2020 Cabin, Sun Thumpers...) and a backup policy, previously negotiated with the administrators.

In order to facilitate the access to the data, the platform has a relational database with a query interface, as a data bank of images for the different research projects, as well as information to administrate the platform. The connection between the database and the storage system is performed saving a link to the route of the files mounted with a NFS system in the database entries.

3.3 Post-processing Software Area

The software tools integrated in the portal at the moment (it will surely be extended in the future) are: FreeSurfer, FSL, LCModel and SPM8. Some of these tools allow running jobs in parallel in a cluster which accelerate the process.

So, according to the characteristics of the applications, the software tools are executed in the different hardware systems:

- Cluster of machines, with 115 worker nodes of two types: 1) Intel Xeon QuadCore 5355 @ 2.66MHz (8 cores) with 16GB of RAM and 2) Xeon Dual Core 5160 @ 3MHz (4 cores) with 8GB of RAM.
- High performance Work Stations Intel Core 2 Extreme Processor QX9650 (3.67 GHz OC, 1333MHZ FSB, 12 MB Cache) with 4GB of RAM.

Basically there are two types of post-processing tools included in PICNIC: tools with a graphical interface to which the researchers are already familiarized and “command line” tools, who require a high degree of knowledge in operating systems, usually Linux.

For the first case, a solution to virtualize applications is required. PICNIC uses a Java applet called NX Web Companion of the Nomachine Company with a Free License for a non commercial use. This applet allows virtualizing desktops to load applications remotely, becoming a proper tool we need to load remotely applications with a strong graphical interaction. In this case, the application runs in the machine at which it is installed, but virtualized with the NX applet. An advantage comparable to other techniques is that, with NX the user's sessions in the OS are independent. This means that the display interaction is completely independent. The applet is inserted into the required web pages and using a session file, the applet loads the preferences of the session, to start with the correct software and parameters.

Regarding the software tools that don’t have a GUI, it is the case of some modules of FreeSurfer. While the modules Tk-medit, Tk-surfer and Qdec run with the NX plug-in, to post-process images in the cluster the “command-line” mode is required. For this purpose, a web interface designed in close collaboration with clinical researchers (Figure 2), has been implemented, avoiding the command line functionality.

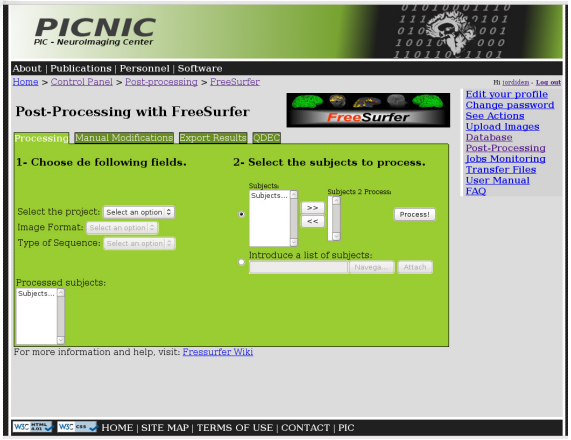

Figure 2: FreeSurfer interface in PICNIC.

4 SERVICES

Apart from the storage and the post-processing

requirements, researchers need other services related to the management of the data. These other services implemented in PICNIC are described in this section.

4.1 Upload Images

The portal includes "Jupload" and "Jdownload" applications to allow users to transfer files (images mainly) to the server. The images have to be in a well-defined directory tree, due the fact that the "Jupload" script verifies the tree directory to save the data correctly in a temporal destination. Taking into account that the images travel via Internet, even if HTTPS protocol is used, for security reasons the platform requires that they are correctly de-identified at the moment of the upload. However, to ensure that all the images in the platform are anonymous, a de-identification process is executed. Afterwards the images are added to the database.

4.2 Image De-Identification

After the upload process, the images are saved in a temporal directory and the de-identification process is automatically executed. The objective is to extract personal data of the subject from the header of the image. Some of the fields of the header that have to be erased are: subject's name, acquisition's date, gender, date of birth, hospital information...

When the de-identification process of a subject finishes, the de-identifier sends an e-mail to the administrator with the log of the process, indicating if upload has finished properly and if not, the problems that arose in the process. It also sends another e-mail to the user indicating if the process finished successfully. If everything was fine, PICNIC adds a new entry to the database with the image and the subject information.

The de-identification process is transparent to the user and it runs in background after the upload process has finished, independently if the user closes the browser or the session.

4.3 Image Visualization

For radiologists, it is very important to have the possibility to visualize the images. Therefore, PICNIC includes an image viewer, MANGO, with the possibility to visualize the images in three formats (DICOM, NIFTI and ANALYZE). Like other software commented before, the viewer uses NX to load the images, opened in a new window.

If multiple images are opened at the same time, the platform loads the viewer only once and the viewer will open a different window for every image to be loaded.

4.4 Jobs Monitoring

Some of the software tools used in PICNIC send the jobs to the cluster of machines of PIC. The job scheduler mechanism used at PIC is PBS (Portable Batch System). PICNIC offers the possibility to monitor the state of each job that has been sent to the cluster at every moment (queued, running, out). Moreover, it gives information like the number of the job and the number of action at which the job belongs to, useful information in case of problems. Finally, it is also possible to stop the job with only a one click and the system will automatically kill the job in the cluster and update the status of the job in the database.

5 SECURITY

The platform has to satisfy the requirements exposed previously to keep it as safe as possible. With this objective, some considerations have been taken into account.

- A module to encrypt the data transmission with SSL (Secure Socket Layer) protocol has been installed in Apache.
- Other options of the Apache have been configured to deny the access to paths or URLs of the server and also to put as default mode of access the HTTPS protocol.
- At the network level, only a minimum number of ports have been opened (22 for NX and 80 for the web connections).
- Finally, at the implementation level, two considerations have been taking into account:
  1. The database queries are protected against SQL Injection escaping characters (like ', "", \, /, |), using techniques that PHP offers.
  2. The access to the hardware and data is very restrictive to ensure that the user can only reach the contents at which he/she has privileges, noted in a white list policy.

6 VALIDATION

Some validation tests have been performed in the way to get indicative results of the operation of the

platform. These tests have been centred in five basic points: database, post-processing tools, transfer files, image visualization and more in general some attacks to the security of the system.

6.1 Database

The consistency of the data is very important, so it has been necessary to verify in detail that all the procedures that interact with the database work properly. To do it, the test has been focused with a double strategy: a preview test of the design of the database using a database manager as "phpmyadmin" and then in the implementation, a validation test to ensure that the web interfaces has been modified according to the actions sent to the database.

6.2 Post-processing Tools

To test all the post-processing tools included in the platform, the test was oriented in two ways:

- All the software tools that run with NX work correctly in terms of NX session files for every user, software and the parameters.
- For the post-processing tools where a web interface was developed, an exhaustive test has been performed to prove that all the features implemented, work correctly in the cluster and the results obtained are correctly added to the database and imported to a spreadsheet.

6.3 Transfer Files and Images

Regarding the transfer of data, firstly it has been tested that all the directory paths were correctly processed by the applet in the case of uploading images. Secondly, the de-identification process of every slice of DICOM images was tested in order to extract personal data of the subjects from the slices. Some of this information is used as input of some fields of the subject or image in the database, like for example the age of the subject calculated from the date of birth.

6.4 Image Visualization

In this case, Mango runs correctly and taking into account that this service is related to the database queries, it was also tested that the selected subjects/images of the query were correctly opened by the viewer. This was a triple test: database, NX and viewer.

6.5 Security

In the way to prove the robustness and security of the implementation, the following features were tested:

- For a different URL of the platform, the access to the derivative paths of the root path is not allowed by the Apache options.
- The PHP implementations have been attacked with SQL injection applications like "sqlier-0.8b.sh", obtaining negative results that confirm the strength of the SQL implementation.

6.6 Analysis Result

To make balance of the results obtained by the implementation and testing of the platform, some contributions and weaknesses are exposed.

6.6.1 Contributions

- It is a scalable solution that can be completed with more services.
- It is in accordance with the clinical research mechanisms, with the advantage of being accessible from everywhere.
- It is easy to use.
- The web-portal is dynamically loaded using PHP and AJAX techniques.

6.6.2 Weaknesses

- The service is at mercy of the Internet connection and traffic, so the transfer time can be limiting.
- The volume of data that can be transferred with the applets, is limited to 1GB.
- The authentication process in the platform is performed with a password and a user-name. In order to improve it, other solutions will be envisaged in the future.

7 COPYRIGHT FORM

The authors of this work let to reference the content of this paper, but not to make a derivative work of this or to copy it.

8 CONCLUSIONS

PICNIC represents a portable solution for

neuroimaging researchers to manage their data and post-process images easily without any previous knowledge of scripting. Based in a web technology, it offers a high availability of the services from everywhere.

PICNIC offers a robust infrastructure in terms of hardware, with the adapted resources according to the post-processing needs, running jobs in parallel in the cluster of machines or running others sequentially in powerful workstations. Periodical backups of the original and post-processed data are also performed.

Due to the modular structure of PICNIC, in the future new software tools will be implemented in the platform, according to the demands of the medical community. Moreover, the authentication of the platform will be improved by implementing a more robust authentication system through certificates. In a long term period and if the requirements of the researchers are oriented to the data sharing, PICNIC could be slightly modified to be used as interface for a GRID infrastructure.

REFERENCES

UCLA – University of California Los Angeles (2009). *LONI - Laboratory of Neuro Imaging.* http://ww.loni.ucla.edu

UTAH Scientific Computing and Imaging Institute (2009). *UCNIA - Utah Center for Neuroimage Analysis (UCNIA)* http://www.ucnia.org/

Martinos Center for Biomedical Imaging (2009) – *FreeSurfer Wiki* http://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferWiki

Shawn N. Murphy, MD Ph.D. et al. (2006) - *A Web Portal that Enables Collaborative Use of Advanced Medical Image Processing and Informatics Tools through the Biomedical Informatics Research Network (BIRN)* http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1839506

SPM: Statistical Parametric Mapping (2009) - http://www.fil.ion.ucl.ac.uk/spm/

LCModel (2009) – http://s-provencher.com/pages/lcmodel.shtml

W3Schools (2009)- Online Tutorials http://www.w3schools.com/

PHP: Hypertext Preprocessor (2009) http://www.php.net

Apache (2008) – HTTP Server Project http://httpd.apache.org/

University of Florida (2008) – *Installing Apache, PHP and Mysql on Fedora core*. http://www.flmnh.ufl.edu/linux/install_apache.htm

MySQL (2009) – *The world's most popular open source database* http://www.mysql.com/

MySQL Hispano (2009) – *La Comunidad de usuarios de MySQL* http://www.mysql-hispano.org/

Andrew Crabb (2009) – *Imaging Software* http://www.idoimaging.com/index.shtml

NIH Blueprint for Neuroscience Research - *NITRC The Source for Neuroimaging Tools and Resources* - http://www.nitrc.org/

ImageJ (2009) – *Image Processing and Analysis in Java* http://rsbweb.nih.gov/ij/

Jack L. Lancaster, Ph.D., Michael J. Martinez - Research Imaging Center, University of Texas Health Science Center San Antonio (2009). *Multi-Image Analysis GUI Mango* http://ric.uthscsa.edu/mango/index.html

Jupload (2009) – *Jupload the multi-files upload*er http://jupload.biz/

Jdownload (2009) – *Download Manager multiple files and folders* http://jdownload.jupload.biz/

Dr Allen D. Malony – *Computer & Information Science Department, Computational Science Institute CIBER University of Oregon (2008). Distributed Computational Architectures for Integrated Time-Dynamic Neuroimaging.*

Nomachine (2008) – *Desktop Virtualization and Remote Access Management Software.* http://www.nomachine.com/