GPUs parallel computing exploitation for neuroimaging

Poster No.: C-1125

Congress: ECR 2017

Type: Scientific Exhibit

Authors: A. Giuliano, S. Arezzini, P. Bosco, A. Ciampa, S. Coscetti, D. Fabiani, M. E. Fantacci, E. Mazzoni, A. Retico; Pisa/IT

Keywords: Computer applications, Neuroradiology brain, MR, Computer Applications-General, Image registration, Image verification, Speech disorders

DOI: 10.1594/ecr2017/C-1125

Aims and objectives

Dedicated computing environments are becoming increasingly important in neuroimaging applications. Indeed we are experiencing a fast development of non-invasive technologies that progress the research on human brain. These advanced techniques, such as magnetic resonance imaging (MRI), positron emission tomography (PET) and electroencephalography (EEG), give us the possibility to visualize and analyse brain function and structure in exceptional detail, but they have led to the necessity of storing and processing very large amounts of data. Moreover, the growth in the complexity of algorithms developed to analyse brain images, has involved an augmenting demand of high-performance resources for data storage and management, and computing systems for image processing and quantitative analysis, e.g. Graphics Processing Units (GPU). The main motivations for using GPU in neuroimaging are the time saving and the possibility to apply advanced algorithms instead of simple ones.

The aim of this study is to describe the issues related to the development of a computing environment for neuroimaging applications. The dedicated farm we built, consisting at the moment of a computing node and a storage unit, has been implemented in the Pisa INFN computing centre. It has been designed to guarantee the secure data handling, storage and the access to fast cloud-based computational resources.

Methods and materials

GPUs can reduce computational times of neuroimaging algorithms with respect to CPUs, thanks to their high-performance, dataparallel architecture. A GPU is the computational component of a graphics card used in ordinary computers, with larger number of cores than a traditional CPU. In general, a CPU core is more powerful due to a higher clock frequency and a larger cache memory. At the beginning, a GPU could only be controlled using computer graphics programming languages (e.g. OpenGL and DirectX), which made it hard to use for wide sets of operations. Nevertheless, employing GPUs for general purpose computing (GPGPU) has been popular for several years [1]. Furthermore, the release of the CUDA programming language in 2007, using nVidia GPUs to accelerate arbitrary calculations, has simplified the process since CUDA is very similar to the widely used C programming language. A number of reports on large speedups, compared to optimized CPU implementations, have been presented so far [2,3]. The performance of a GPU implementation considerably depends on how much a certain angorithm is suitable to be runned in parallel. Fortunately, neuroimaging data are often analysed in the same way for each image element (pixel or voxel). For this reason, a lot of the algorithms commonly used for neuroimaging puposes are appropriate for parallel implementations. In neuroimaging, GPUs have been used for a wide range of applications, for example to speedup reconstruction of data from magnetic resonance (MR) scanners, and to accelerate algorithms such as image registration, image segmentation and image denoising. Some of these applications are integrated in very well-known neuroimaging software toolkits such as SPM (Statistical Parametric Mapping) (http://www.fil.ion.ucl.ac.uk/spm) and FreeSurfer (http://freesurfer.net/) [4].

In semptember 2016, a dedicated computing node for neuroimaging applications has been set up in the Pisa INFN computing centre (Fig.1). It is organised as follows:

- CPU: 2x 10 cores Intel Xeon E5-2640v4 @2.40 GHz;
- GPU: 4x nVidia K80, with 2x GPUs Tesla GK120 GB RAM and 2496 CUDA cores each;
- memory: 64 GB RAM.

Alongside this computation facility, users have 72 TB of total disk space for the storage of the processed images. This environment has been designed to guarantee the secure data handling, storage and the access to fast grid/cloud-based computational resources.

Preliminary tests have been performed on a dataset of 30 T1-weighted MRI brain scans using FreeSurfer recon-all tool, one of the most widespread neuroimaging softwares for brain tissues segmentation and parcellation (Fig.2).

Images for this section:

Fig. 1: The Computing Centre of INFN Pisa Section.

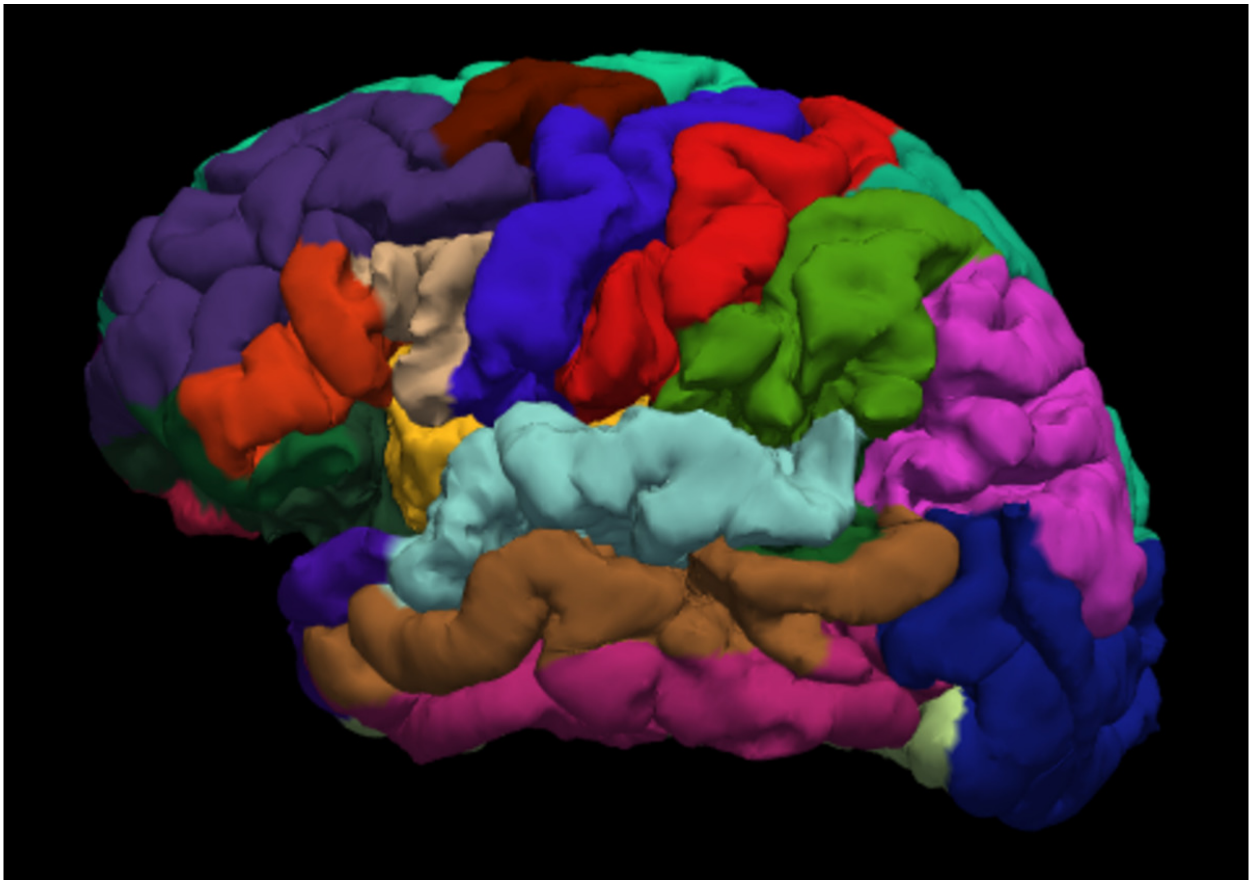

Fig. 2: Cortical parcellation with Freesurfer software v5.3 according to Desikan-Killiany-Tourville (DKT) atlas.

Results

The whole infrastructure is up and running and the main neuroimaging tools have been tested on the grid environment. A web interface is now available that allows specialists to upload MRI scans, classifying them according to prearranged criteria. The results of this classification are submitted in a database, available to the local neuroimaging researchers, that are so able to extract images related to what they are interested in. The extracted images are then ready to be processed on the computing facility described above.

The computing node has been equipped with some graphical coprocessors in order to accelerate the parallel run of the most common neuroimaging softwares used nowadays, such as FreeSurfer, AFNI, ANTs, MatLab and SPM.

The recon-all execution time has been reduced by a actor of almost two exploiting the GPUs parallel computation on the computing node (Fig.3). Moreover, the analysis time of a single scan is reduced by a factor of six with respect to a standard workstation.

Images for this section:

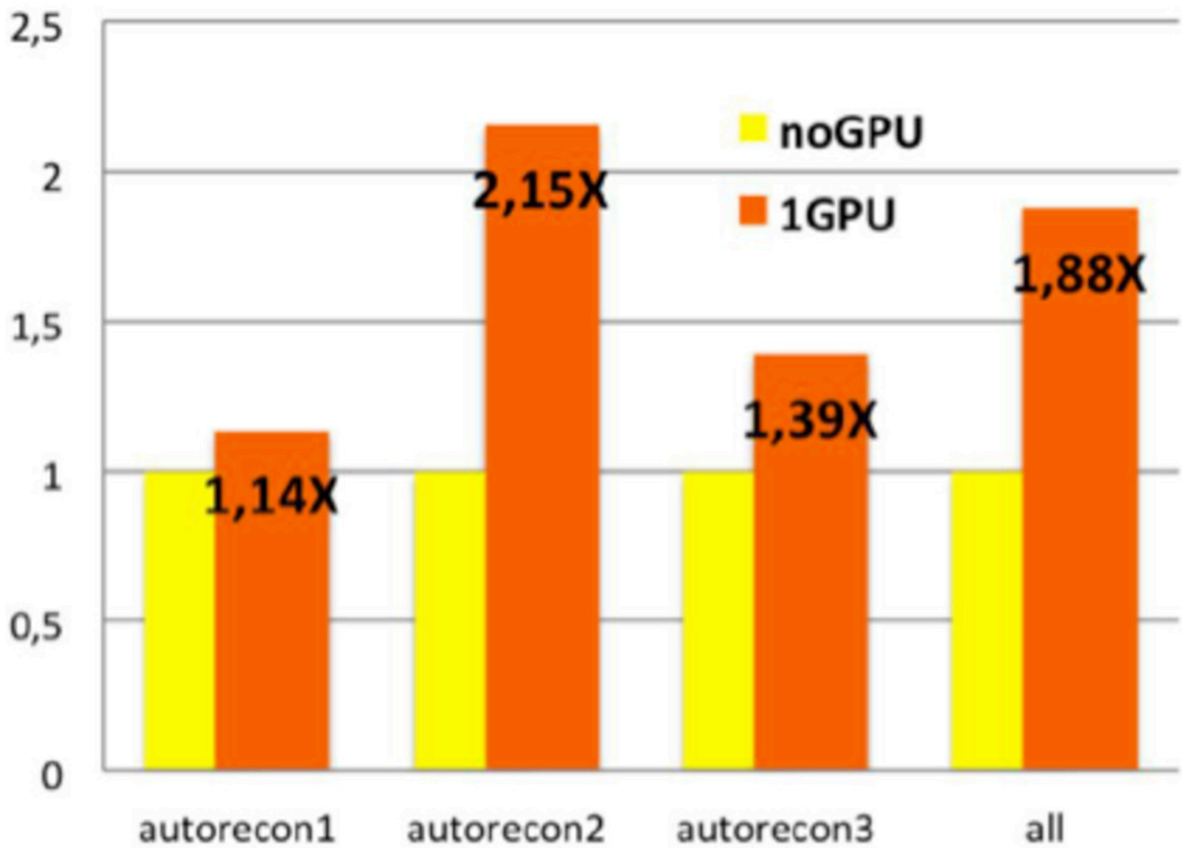

Fig. 3: FreeSurfer recon-all tool performance comparison using a GPU.

Conclusion

Due to fast growing use of MRI as a reliable diagnostic and research tool and to the increasing need of highly performing computing environment for neuroimaging data processing, we set up a dedicated computing facility, and we demonstrated its functionality. The whole infrastructure is up and running and the main neuroimaging tools have been tested on the grid environment. GPUs parallel computing can be fruitfully applied to the preprocessing of neuroimaging data.

Personal information

This work has been partially founded by the Tuscany Government (PAR-FAS 2007-2013 FAS Salute, ARIANNA Project), the University of Pisa (PRA_2016_39 VIP project), and INFN-CNS5 (nextMR project).

We would like to aknowledge the team of researchers of the project *ARIANNA* and the researchers of the computing centre of the INFN Pisa Section.

References

[1] Owens, J. D., Luebke, D., Govindaraju, N., Harris, M., Kruger, J., Lefohn, A. E., and Purcell, T. J. (2007). A survey of General-Purpose Computation on Graphics Hardware. In Computer graphics forum, volume 26, pages 80-113. Wiley Online Library.

[2] Che, S., Boyer, M., Meng, J., Tarjan, D., Sheaffer, J. W., and Skadron, K. (2008). A Performance Study of General-Purpose Applications on Graphics Processors using CUDA. Journal of parallel and distributed computing, 68(10):1370-1380.

[3] Garland, M., Le Grand, S., Nickolls, J., Anderson, J., Hardwick, J., Morton, S., Phillips, E., Zhang, Y., and Volkov, V. (2008). Parallel Computing Experiences with CUDA. Micro, IEEE, 28(4):13-27.

[4] Fischl, B. (2012). Neuroimage 62(2):774-781.