

CLOUD CONTAINERS, PRÊT-A-PORTER

Installed and tested between the end of 2017 and the beginning of 2018, the new GARR Cloud Container Platform will soon open to the research community. The new platform will complement the GARR standard cloud computing offering, based on virtual machines, which is in operation since last year and is now used by more than 500 users and growing. The container technology is one of the newest and most interesting trends in cloud and swiftly growing in the commercial world but it is not very popular yet within the research and education environment, and the GARR platform would be one of the first to be expressly dedicated to researchers.



Features and benefits

Compared to traditional cloud computing, the container technology offers a more agile approach to virtualisation, which is operated on a higher level of the virtualising stack: while a virtual machine abstracts from a physical machine, a container abstracts its operating system. Within the container, we can install fully autonomous software packages, each one isolated from the others and provided with all the components they need to run: code, runtime, tools and system libraries, settings. With containers, the applications share the operating system, and especially kernels, system files, and network and disk drivers. As they don't need to include all server functions to work, they are much lighter than virtual machines, require less CPU power and can be activated in few moments: typically, in the order of the tenths of a second, against several minutes for a virtual machine. A disadvantage is that all applications in a single container must be written for the same operating system they share. However, if applications are built with portability in mind that won't be a big constraint.

The implementation

To implement the container platform, we selected an architecture that allowed us to instantiate the containers directly at the bare metal level, and not above a virtual machines layer. This approach caters for better performance and offers a direct access to the interface of special devices as GPUs (Graphics Processing Units), which cannot be shared with traditional virtualisation techniques. The GPUs are processors specifically optimised for image processing and rendering and are therefore much faster than CPUs in parallel processing of large blocks of data, as for the tensor calculus used in Deep Learning.

More power to Deep Learning and Big Data

Having access to GPUs in the academic environment can be crucial to keep the pace with leading-edge international research: indeed, it is almost impossible to compete with someone who can perform 100 experiments in the same timespan you complete one. Among the applications that can benefit more of GPUs are those based on Deep Learning, that in the last decade had amazingly good results in the fields of image processing, strategic games, robotics, automated drive, voice recognition and synthesising, automatic translation and natural language processing, and AI in general, that can

Words
Giuseppe Attardi, GARR Computing and Storage Department

use computing capacity to process large amounts of learning data.

On our Cloud Container Platform, the most common Deep Learning libraries are available as containers, to facilitate users in developing similar applications. The platform is offered to researchers from any disciplines, and not only to technical and scientific ones, as this makes the creation of new applications easier and frees researchers from most of the complexities connected with installing machines and software: applications will be built through a simple GUI, by composing them from a wide catalogue of off-the-shelf containers. A brief demo is available here: <https://youtu.be/asxozHcandc>.

The value of community

The annotated learning datasets and the models that can be inferred from them are the cornerstones of Machine Learning applications. To stimulate developments in this field, GARR will grant free access to the platform for a given period of time to users who will agree to share their data and models with the community, thus enabling others to use them as a starting point of their own research. Accessing this kind of resources would be helpful both for those who are new to this application development techniques, who would not have to start from scratch, and for those who are already working on similar problems.

All these features make GARR Cloud Container Platform a one-stop shop, where researchers can find everything they need to start their experiments: 1) computing resources, 2) pre-set software components that can be easily reused and customised, 3) ready-to-use datasets and models that can be used or adapted to new use cases.

With this new tool, we hope to provide the Italian scientific community with a tool to become a leader in the new era of applications, and in particular of AI.