

## **How a virtual appliance facilitates the data management, the SeaDataCloud project experience.**

**Leda Pecci**, ENEA (Italy), leda.pecci@enea.it  
**Marcello Galli**, ENEA (Italy), marcello.galli@enea.it

The SeaDataCloud project has in its main objectives the innovation of the SeaDataNet infrastructure, reached by means of the adoption of cloud computing as well as the High Performance Computing (HPC). In the framework of the project's activity there is also a challenge to simplify the management of the network nodes, especially for new data centres, of the semi-distributed system architecture, that is the base of the infrastructure. The SeaDataNet infrastructure allows to access to a very large number of ocean and marine data sets stored by more than 110 data centres deployed in a territory that extends beyond Europe. The central portal plays the role of providing data visualisation, discovery and access services, the latter service is based on the CDI (Common Data Index) system which includes a key component that allows data centres to manage data sets requests. The component, called Download Manager (DM), is a Java application that communicates with the central portal and provides data to SeaDataNet users.

A solution that helps the system managers of data centres from installation, configuration and maintenance of the SeaDataNet node has been carried out. Data providers that intend to become a node of the infrastructure have to install, configure and maintain a server that contains the D. M. application, that manages the delivery of the data files ordered, and the software required to run the application. Virtualization technology can optimise the connection of data centres to the infrastructure and delivers an alternative over traditional hardware servers. We make available a pre-built virtual appliance, containing the D. M. component, that is a special purpose virtual machine, based on a linux operating system and contains only the necessary software applications (e.g. web and application servers) with an operating system that is enough for it to run optimally and, thanks to which the system is less vulnerable to security breaches. This virtual appliance allows delivering and deployment of the D. M. application, which requires not easy configuration after being installed, in an easy way.

The SeaDataNet virtual appliance includes all necessary data management tools, it is easily deployable into a compatible virtualisation environment, ready to use with minimal setup, which eases the process of maintenance and management of the distributed nodes. We have also prepared a bash procedure for a semi-automatic installation of the appliance software and some testing.

The virtual appliance is realised using a TurnKey Linux 14 (TKL), based on Debian 8 (Jessie) a stable version of the operating system, suitable for the production environment, well known for its robustness and reliability. The operating system updates and software patches contained in the SeaDataNet virtual appliance are automatically downloaded and installed. The virtual appliance contains only open source software.

Virtual appliances eliminate the need for physical hardware because they run on virtual platform solutions. The advantages of adopting this solution are different, data centres save money on

hardware, software, the availability and reliability of the global infrastructure increases, because the disaster recovery of the nodes is simplified, the operating system and the data can be easily recovered by duplicating images of the virtual appliance and migrating them in another host system. Furthermore virtualization provides the ability to adapt the hardware to real needs, increasing the amount of RAM, the size of hard disks, etc., via software.

On a single physical platform multiple virtual machines can run, this allows to have lower energy consumption, that means lower costs and benefits for the environment.

The appliance is usually managed via a ssh shell, but can also be managed by a web interface, which uses the tools shellinabox and webmin. Currently, the time required to install and configure the download manager is reduced and the maintenance of the system is easier.