



THURSDAY, OCTOBER 26, 2017

E-INFRASTRUCTURES FOR LARGE-SCALE COLLABORATIONS

VI-SEEM Virtual Research Environment

Dušan Vudragović¹, Petar Jovanović¹, and Antun Balaž¹

¹*Scientific Computing Laboratory, Center for the Study of Complex Systems,
Institute of Physics Belgrade, University of Belgrade*

In the last decade, a number of initiatives were crucial for enabling high-quality research in both South-East Europe and Eastern Mediterranean region. This was achieved by providing e-Infrastructure resources, application support and training in these two areas. VI-SEEM project brings together these e-Infrastructures to build capacity and better utilize synergies, for an improved service provision within a unified virtual research environment for the inter-disciplinary scientific user communities in those regions. The overall aim is to provide user-friendly integrated e-Infrastructure platform for regional cross-border scientific communities in climatology, life sciences, and cultural heritage. This includes linking computing, data, and visualization resources, as well as services, models, software and tools. The VI-SEEM virtual research environment provides the scientists and researchers with the support in a full lifecycle of collaborative research: accessing and sharing relevant research data, using it with provided codes and tools to carry out new experiments and simulations on large-scale e-Infrastructures, and producing new knowledge and data. The VI-SEEM consortium brings together e-Infrastructure operators and scientific communities in a common endeavor that will be presented in this talk. We will also point out how the audience may benefit from this newly created virtual research environment.

Underlying e-Infrastructure of the VI-SEEM project consists of heterogeneous resources - HPC resources - clusters and supercomputers with different hardware architectures, Grid sites, Clouds with possibility to launch virtual machines (VMs) for services and distributed computing, and storage resources with possibility for short and long-term storage. The heterogeneous nature of the infrastructure presents management challenges to the project's operational team, but is also an advantage for the users because of its ability to support different types of applications, or different segments of the same application. These are modern, state-of-the-art technologies for computing, virtualization, data storage and transfer.

Efficient management of the available computing and storage resources, as well as interoperability of the infrastructure is achieved by a set of operational tools. Static technical information, such as name, geographical location, contact and downtime information, list of service-endpoints provided by a particular resource center within the infrastructure etc., is manually entered and made available through the VI-SEEM GOCDB database. Based on this information, project monitoring system is able to automatically trigger execution of monitoring service probes, and to enable efficient access to results of the probes via a customized monitoring web portal. Using standardized metrics, the VI-SEEM accounting system accumulates and reports utilization of the different types of

THURSDAY, OCTOBER 26, 2017

E-INFRASTRUCTURES FOR LARGE-SCALE COLLABORATIONS

resources. User support and service-related problems are resolved mainly through the helpdesk system, but via a technical mailing list as well. The VI-SEEM source code repository contains all codes developed within the project, while the technical wiki collects technical documentation, know-hows, best practices, guidelines, etc.

A solid but flexible IT service management is one of the keystones of the foundation for the service-oriented design. The specifics of the federated environment, such as the one found in the VI-SEEM consortium, impose requirements for service management tools that cannot be met using common off-the-shelf solutions. Hence, special care is taken in the design and the implementation of easy to use, custom solutions that are tailor-made for the scientific communities. Our application-level and data services are managed through the VI-SEEM service portfolio management system. It has been developed to support the service portfolio management process within the project as well as to being usable for other infrastructures, if required. The main requirements for the creation of this tool have been collected from the service management process design, and it is designed to be compatible with the FitSM service portfolio management.

The VI-SEEM authentication and authorization infrastructure relies on the Login service. It enables research communities to access VI-SEEM e-Infrastructure resources in a user-friendly and secure way. More specifically, the VI-SEEM Login allows researchers whose home organizations participate in one of the eduGAIN federations to access the VI-SEEM infrastructure and services using the same credentials they are using at their home institutions. Furthermore, the VI-SEEM Login supports user authentication through social identities, enabling even those users who do not have a federated account at home institutions to be able to seamlessly access the VI-SEEM services without compromising the security of the VI-SEEM infrastructure.

The provided infrastructure resources and services are mainly used through the development access, as well as through the calls for production use of resources and services. The VI-SEEM development access facilitates the development and integration of services by the selected collaborating user communities: climatology, life sciences, and cultural heritage. In this process, applications are given access to the infrastructure and necessary computational resources for a six-month period, during which application developers are expected to develop and integrate relevant services. The calls for production use of resources and services target specific communities and research groups that have already began development of their projects. These calls are intended for mature projects, which require significant resources and services to realize their workplans. Therefore, a significant utilization of the VI-SEEM resources comes from the calls for production use of resources and services, and an order-of-magnitude smaller utilization comes from the development access.