Milestone M15

Work Package 4, Milestone 4.5: Final deployment of the consolidated Platform

Besides operating and maintaining the existing application portals and infrastructure, the consolidation work of WP4, after what already described in the previous documents delivered for Milestones MS12, MS14 and Deliverables D4.3, D4.5, was progressing in the following directions:

- integration of the Onedata service storage solution
- virtualization of portals using emerging cloud orchestration technology
- final implementation of the West-Life SSO solution and integration with other EOSC-Hub AAI supported solutions

The work progress is tracked at the WP4 wiki pages [1]. This document is a snapshot of the information available there at the time of this milestone, i.e. August 2018. A more detailed description will be provided in the next project deliverable document D4.6: *Final report on deployment of consolidated platform* and the overall architecture due at month 36, i.e. by October 2018.

The following interfaces/services have been integrated in the consolidated infrastructure and available for use by other WPs.

Onedata integration with West-Life SSO and Virtual Folder

Onedata is a global data access solution for science developed by INDIGO-DataCloud project. With Onedata, users can perform heavy computations on huge datasets; access their data in a dropbox-like fashion regardless of its location; publish and share their results with public or closed communities.

As already described in the previous documents, INFN partner in 2017 had made available a Onedata experimental storage space to all West-Life users. The storage available at INFN-PADOVA data centre consists of 14.6 TB. West-Life users can obtain a storage space by login at the Onezone server [2] hosted at INFN-CNAF centre.

While initially only Username/Password and INDIGO-IAM SSO were implemented as authentication methods, in 2018 the support to West-Life SSO was added. To achieve this, INFN developed a new plugin specific for the West-Life SSO service, which was contributed upstream to Onedata software.

After that, INFN worked in collaboration with STFC on the integration of the West-Life Virtual Folder with the Onedata storage back-end.

Virtual Folder acts as a gateway for many storage systems, such as WebDAV, Dropbox and B2Drop. Each protocol requires a specific plugin loaded by the engine of the framework. The plugin is responsible for retrieving metadata for files and directories stored in the storage provider. All the files tracked by Virtual Folder are made available through a WebDAV

interface; the framework maintain a WebDAV area synchronized with the remote storage system; the mechanism for implementing the synchronization depends on the remote storage provider, ranging from mounted network filesystems to specific tools for mirroring.

Onedata is a distributed storage system based on a network of storage providers (Oneprovider service) federated together through a central manager (Onezone service). The Onezone service is responsible for the user authentication and authorization; it delivers a token to the authenticated user which grants him the access to the spaces on any federated Oneprovider. The user must ask for spaces on a provider and provider administrators can grant the access to that user.

The integration between Onedata and Virtual Folder consists on a plugin for the Virtual Folder framework which is able to get any information about files and directories from a given Oneprovider. The plugin makes remote invocations to the web API interface (RESTful) of the Oneprovider service; it passes the token that the user retrieves from the Onezone service for the authentication step. Moreover, the plugin is able to mount on demand any user space available on the provider. The operation is carried out with the execution of the specific command line tool, Oneclient. The plugin for Onedata contains also an extension of the Virtual Folder dashboard; the extension allows the user to specify the configuration attributes, such as the endpoint of the service and the token to be used for authetication, when a new provider is created in the framework. The plugin for Onedata has been developed in C#; the extension of the Virtual Folder dashboard is based on Aurelia, a javascript framework.

The documentation with the instructions for West-Life users on how to login and use the Onedata service has been updated on the West-Life main web site.

Cloud orchestration service and portal virtualization

Since MS14 the cloud orchestration recipes were updated to the new major version 4 of Cloudify. The new version is more robust in general, installation of the client is straightforward (packaged for all major Linux distributions) compared to the previous versions.

In addition, the individual portal recipes were migrated from standalone Cloudify to Cloudify Manager service. This approach adds several features suitable for production use:

- continuous monitoring of the cloud deployment,
- automatic healing of failed deployment nodes
- evaluation of service load, connected with automatic scale up and down (e.g. adding/removing worker node replicas)

We provide Cloudify Manager service for the use by other portal providers in the project. It is available at http://cfm.westlife.dyn.cerit-sc.cz/. Access to both EGI Federated Cloud services via OCCI as well as the use of pre-allocated nodes (via hostpool plugin) are supported.

The most elaborated pilot service using this approach is the Gromacs portal. Work on similar deplouyment of the Scipion portal is in progress.

Milestone M15

West-Life SSO service and other EOSC-Hub AAI supported solutions

The activities over the reporting period focused on provisioning of a stable environment and accessing the key features of West-Life AAI to the end users and service providers. Effort was dedicated to extend the utilization of AAI, the number of user registered with the identity management system has been growing since the service was put in production in summer 2017, with 62 users being registered with the system at the moment. The AAI system was used to provision access to the West-Life services during training events (e.g. Scipion course at MU), during which the registration procedure was verified to work for a number of attendees.

Development and preparation of West-Life AAI is pursued with close links to similar developments in the field of life-science (esp. the Life-Science pilot in AARC, development in ELIXIR and BBMRI) with the vision of providing a compatible solution that will be potentially be possible to merge with a more general solution applicable for the Life-Science domain. First communication about potential utilization of the EGI Checkin service was performed also.

In addition, all Utrecht based portals were migrated to a central user management system. As part of the EOSC-Hub project this system was connected to the EGI-CheckIn service with the OpenId Connect (OIDC) protocol and the support will be enabled in the production version of the User portal later this year. A development version is already available [3].

The key concept and current state of the West-Life AAI service was presented at the Digital Infrastructures for Research 2017 (DI4R) conference in Brussels (November 2017).

With the adoption of EU regulations on data protection and privacy, the West-Life AAI components were reviewed to ensure compliance with the existing requirements. The “West-Life AAI Acceptable Usage Policy” (AUP) and “West-Life AAI Privacy Policy” documents were updated and published from the West-Life portal [4]. Both the documents follow closely rules stipulated by other life-science infrastructures, trying to retain compatibility in the domain. The registration process was extended to clearly communicate the procedures to new users. Existing users were notified about information that is collected and processed for the need of West-Life AAI operations.

The West-Life AAI technology was extended with support for the OpenId Connect (OIDC) protocol, which extended current SAML-based interfaces. OIDC is becoming a popular technology to integrate new services and is widely supported. Introducing the new interface extended the flexibility of the service and eased its adoption by end services. At the moment, the system registers 12 end services using SAML and 8 using OIDC, which are exposed to users. In addition to that, a number of testing and development services are registered, too.

WP5 portals upgraded for using WP4 services

A number of WP5 portals have been upgraded to make use of the above services:

Gromacs

The original Gromacs portal by UU was previously modified to support Gromacs 5, which also provides efficient GPU computation.

Recently, the portal was transformed to the Westlife project visual style, and it was fully integrated with the Westlife AAI.

A pilot deployment is available to the users at http://gromacs.westlife.dyn.cerit-sc.cz/. It uses the cloud orchestration service described above. It runs on MU cluster of Intel 6138 CPUs with NVidia GeForce 1080Ti GPUs, cluster nodes are added and released dynamically according to the current portal load and node availability.

Scipion

Scipion Web Tools

In December 2017 access to Virtual Folder (file picker) was integrated on Scipion Web Tools, which made possible to upload files from the VRE VF and used them as input to the web tools services.

Virtual appliances

Since June 2018 two new versions of ScipionCloud virtual appliances are available on EGI Applications Database: ScipionCloud and ScipionCloud-GPU. These images were built using packer with the configuration available [5].

Cloudify deployment

Deployment of full Scipion application through installation with Cloudify recipes is now fully supported [6]. Besides installing the Scipion software and all its dependencies, the deployment can utilize Nvidia GPU, when available in the cloud machine, for both computation and accelerated 3D rendering through VirtualGL.

Despite the Cloudify deployment automates the setup of the Scipion node completely, running it is still non-trivial task, uncomfortable for the non-expert user. Therefore we provide a simplified web front-end interface to manage the cloud deployments, and to provide the user with access to the remote application desktop through web browser, without the need to install any client software. Details are given in the internal technical report [9], and the solution will be presented at the DI4R 2018 conference.

Work on the cloud deployment continues, the setup of the web front-end will be integrated to the cloudify deployment, and internal communication of the components will be streamlined. These changes are expected to be released by the end of the project, together with a "catch all" service instance available to the user community, using MU cloud.

PowerFit/DisVis

Due to discontinuation of the ui-enmr.lsg.bcbr.uu.nl endpoint, which was used to submit all grid jobs of DisVis and PowerFit, the grid submission machinery of these services has been migrated to the new enmr-ui.grid.surfsara.nl machine. In connection with this migration, the submission machinery has been adapted to handle submission from multiple hosts, thereby increasing the portability of the grid-enabled portals.

Some efforts have also been made to automate the deployment of the portals on cloud resources. A setup relying on docker-compose to handle 3 docker images (nginx, a PostgreSQL database and the web portals flask-based frameworks) has been developed. This fills in the gap for on-demand deployment of web portals during workshops, courses, etc. This is already in test for DisVis on a local cluster and available [8] (submission to GRID resources only at this time).

Virtual Folder

WP6 - Virtual Folder local deployment configuration was integrated into West-life SSO. There is need to amend existing user's configuration in public VF deployment therefore migration is in the process of design and technical decision. An implementation of the Virtual Folder access has been demonstrated in the devel version of the DisVis web portal, in Scipion Web Tools portal (production portal in FedCloud) and FANTEN web portal.

Other new features:

- Import Settings from another Virtual Folder instance - Using asymetric encryption RSA + AES, #67
- generate obfuscated public webdav URL in order to prevent full path disclosure using symetric AES encryption #64
- added Jupyter notebook option to start instance on local deployment

Updated features:

- UI - basic UI as SPA with router navigation (#/setting #/filemanager), updated angular.js 1.6.10 used by EBI PDB components.
- Backend installation into standard linux paths - shared data at /srv/virtualfolder/, database at /var/lib/westlife/, configuration /etc/westlife/, logs /var/log/westlife/

Known issues:

- West-Life SSO not working in CernVM4 h2020-westlife-eu/wp6-repository#42 as a workaround use latest Scientific Linux 7.5

Further info at [9].

FANTEN

The FANTEN service has been extended by implementing a direct connection to the HADDOCK server that allows FANTEN-generated models of protein-protein adduct to be directly submitted to HADDOCK for flexible refinement.

AMPS-NMR

The AMBER portal for NMR structure refinement was modified to include user authentication using West-Life SSO. Connection to the Virtual Folder is under way..

References

[1] http://internal-wiki.west-life.eu/w/index.php?title=WP4.
[2] https://onezone.cloud.cnaf.infn.it
[4] https://csbdevel.science.uu.nl/
[5] https://auth.west-life.eu/
[7] https://github.com/I2PC/scipion-cloud/tree/master/scipion-ubuntu-16.04
[8] https://github.com/ICS-MU/westlife-cloudify-scipion
[9] http://internal-wiki.west-life.eu/images/8/87/Scipion-in-the-cloud.pdf
[10] https://nestor.science.uu.nl/disvis/
[11] https://github.com/h2020-westlife-eu/virtualfolder/releases/tag/v18.06.01