

3 REPOSITORY ARCHITECTURE

The architecture is based on a virtualized platform implementing KVM open-source hypervisor² running on top of Ubuntu server 16.04 in conjunction with OpenVSwitch to manage networking and iSCSI with multipath to manage SAN storage.

Current architecture requires three virtual machines: a) the repository to manage digital objects, b) the image server to render image files and c) the front-end, hosting the interface for public access to the repository objects and data. All virtual machines are running Ubuntu server 16.04 LTS.

Finally, a highly reliable storage (SAN) completes the architecture; it supports the repository for objects preservation and the virtual machine for system backups.

3.1 Back end server

The back-end virtual machine hosts Fedora Repository, BlazeGraph and Solr, each one of them running in a separated Java servlet container; the first two, in a couple of Tomcat instances, while Solr is running on self-included Jetty servlet container.

Fedora Repository manages the Byterfly conservation task, which includes API's for programmable ingesting, semantic description of the relationships among the objects and their management based on models.

Unfortunately, the Fedora triple-store manager Mulgara has low performance and poor versatility; that is why we chose to replace it with BlazeGraph, a powerful triple-store manager. You can follow our steps³ to modify Fedora configuration for the new triple-store based on tripi-sail module (thanks to Discovery Garden!).

Finally, Solr is the search and indexing engine, which is in charge of indexing all ingested data stream contents and provides very fast search results. Configuration is customized to index date, full-text data stream and all Dublin Core elements in order to provide right fields to be used with Views and search facet block into Islandora front-end.

3.2 Image Server

The Islandora Large Image and Book modules use Openseadragon and Internet Archive Bookreader to show images, pages and books. These two components are very powerful and provide the front-end with a better user experience for accessing to digital objects. Both of them are based on JPEG2000 image format and they need an image tile source to show properly.

In the past, we used Adore-djatoka servlet to feed the image viewer components with the required derivatives; it run into a container which in most cases was the same container used by Fedora repository. For a few months, Islandora Openseadragon head version has started to support IIIF (International Image Interoperability Framework) standard, while simultaneously Cantaloupe open-source dynamic image server has reached a good maturity and reliability. Moreover, starting on a first implementation of Islandora Internet Archive Bookreader version 2 developed by Diego Pino Navarro⁴, we added it some features and made code ready to be used properly with Cantaloupe IIIF server.

The result has been an Islandora front-end full based on IIIF server, without any need of adore-djatoka. Cantaloupe is a high performance image derivative generator with powerful cache features, so we decided to build a third virtual server in order to host exclusively the IIIF server, which is able to server more than one Islandora at time, i.e. our Cantaloupe server is providing images to three distinct repository front-ends.

Byterfly development code of the repository is available to the open community in dev site⁵.

² <http://dev.digibess.it/doku.php?id=hypervisor:start>

³ http://dev.digibess.it/doku.php?id=reloaded:be_repmulg

⁴ https://github.com/DiegoPino/islandora_internet_archive_bookreader/tree/7.x-2dev

⁵ <http://dev.digibess.it/doku.php?id=reloaded>