BROADSEA - Docker Containers For Cross-Platform Installation of OHDSI Tools

Lee Evans[1,4] , Marc Suchard[2,4], Jon Duke[3,4]

[1] LTS Computing LLC, West Chester, PA; [2] Department of Biomathematics, David Geffen School of Medicine, University of California, Los Angeles, CA; [3] Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN; [4] Observational Health Data Sciences and Informatics (OHDSI)

LTS Computing LLC

Abstract

We packaged the OHDSI open source standard software stack into Docker containers with the aim of simplifying cross-platform OHDSI software installation on a range of Operating Systems, DBMSs and hardware infrastructure. Docker is the open source technology we used to deploy the OHDSI software stack as portable, self-sufficient software containers.

We believe this simpler software deployment option will encourage OHDSI Community members to install the full OHDSI software stack for research on their own Version 5 Common Data Model (CDM) databases and more easily participate in OHDSI Network studies.

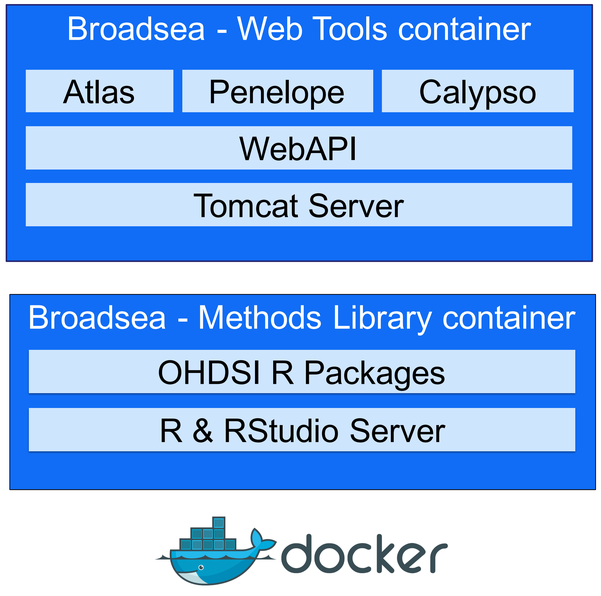
The OHDSI software stack Docker container image build process and the OHDSI container images are collectively known as Broadsea.

Background

The OHDSI software stack deployment process requires some technical familiarity with the underlying Java/HTML5/R development technologies. Further, a Tomcat application server and an R environment must be made available along with the CDM database. Deployment is a typical Java/HTML5/R application process but there are a number of dependencies and a number of steps that need to be performed.

We use continuous integration and Docker container technology to build and package the open source OHDSI web applications and OHDSI R packages into easily deployed cross-platform software containers.

Methods



The Tomcat application server and OHDSI web applications are included in the Broadsea Web Tools Docker container.

R, RStudio Server & the OHDSI R packages are included in the Broadsea Methods Library Docker container

The containers are configured by editing the docker-compose.yml file to connect to the local CDM V5 database, and the database source tables are populated by editing and executing a SQL script file.

Docker Engine & the Docker Compose utility must be installed on the host machine.

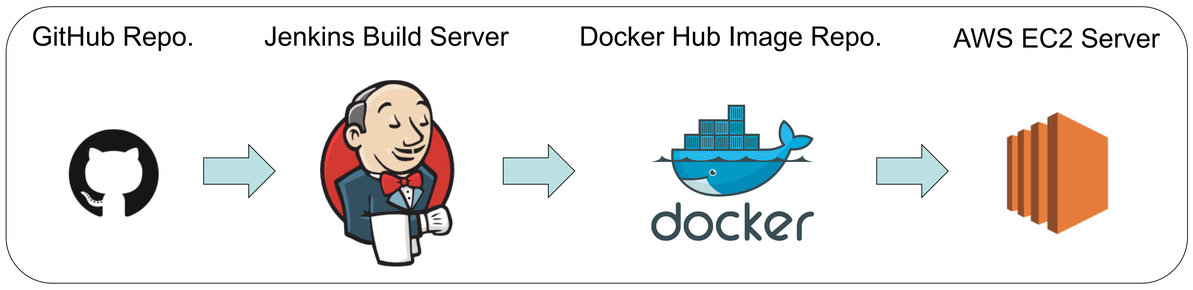
Results

The public Atlas, Calypso & Penelope OHDSI cloud apps are continuously deployed in Broadsea Docker containers on Amazon. They are available on the OHDSI.org website.

The Broadsea container images have been downloaded from Docker Hub a total of **495** times.

Broadsea containers (running within a Virtual Machine) are being used for training in the ETL & the OHDSI technology stack tutorials for the OHDSI 2016 Symposium.

Instructions for deploying and configuring the OHDSI Broadsea Docker containers are available here: https://github.com/OHDSI/Broadsea



OHDSI Cloud apps continuous deployment process: from Github source code to deployed Docker containers

Conclusion

Broadsea simplifies the installation of the full OHDSI standard software stack for the OHDSI Community.

Broadsea containers provide a standard release version of the OHDSI software stack that may be distributed/archived to facilitate OHDSI study reproducibility.

Deploying Broadsea containers is a simple way to provide hands-on training with the OHDSI software stack.

Broadsea will evolve over time to include additional OHDSI tools and backend services (e.g. R Service Bus / RPooli) as they are made available to the OHDSI community.

Broadsea Docker containers enable more advanced deployments of the OHDSI software stack using Docker engines running on cluster computing platforms such as Mesos, Kubernetes & Docker Swarm on cloud services such as Amazon Web Services or Google Cloud Platform. Some examples include:

- Multi-tenant deployment using multiple instances of the OHDSI software stack
- Horizontal performance scaling by adding additional servers to a Docker engine cluster

Contact: levans@ltscomputingllc.com