

OMERO in the Cloud: A BD2K Project

Douglas P.W. Russell

Harvard Medical School

OMERO Annual Users Meeting 2017

31st May 2017

Disclosure

Peter Sorger is a founder and shareholder of Glencoe Software

LINCS BD2K Supplement

- Library of Integrated Network-Based Cellular Signatures
- At many of the consortia, a lot of imaging involved
- Some sites have OMERO and are using it to publish data, but not all have the infrastructure and we only share it for browsing, not for reuse.
- BD2K (Big Data to Knowledge): Remit to facilitate broad use of biomedical big data and introduce FAIR (Findable, Accessible, Interoperable and Reuseable) standards.
- Granted a BD2K supplement to LINCS to develop a prototype method of addressing some of those challenges for LINCS imaging data.

Sharing data for reuse: Have your cake and eat it too

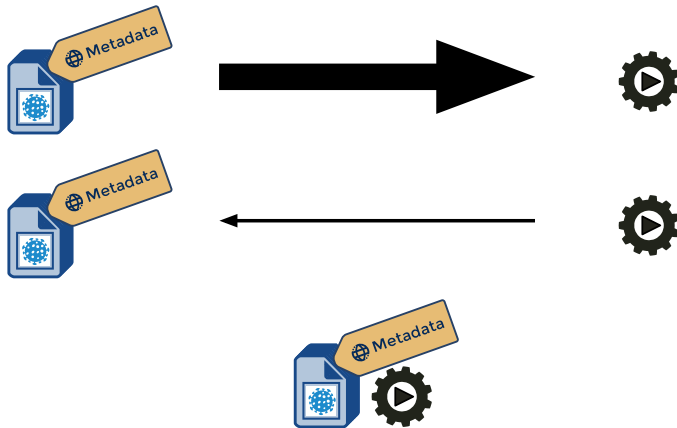
Without OMERO

- Collection of files, not well organized and annotated
- Terrible for browsing, inspection and querying
- Third party could download data for reuse

With OMERO

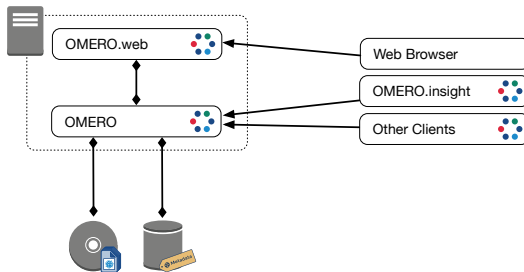
- Data available, organized and annotated
- Great for browsing, inspection and querying
- Not very useable for third party reuse

Public Data Sharing

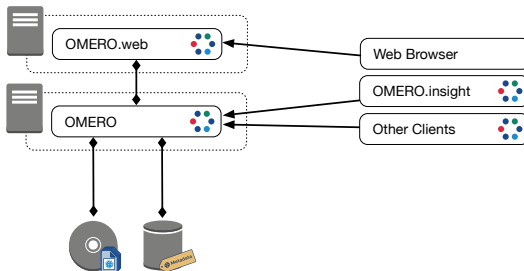


- Dataset(ish) as an instance
- Static
- Containerization
- Scalability

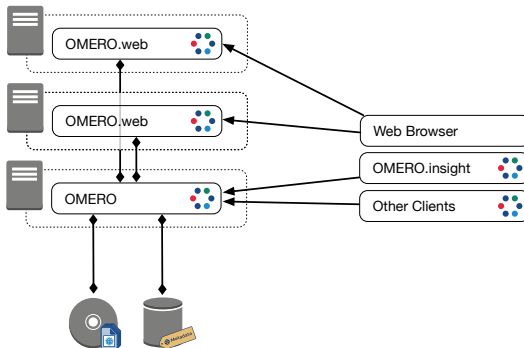
All-In-One



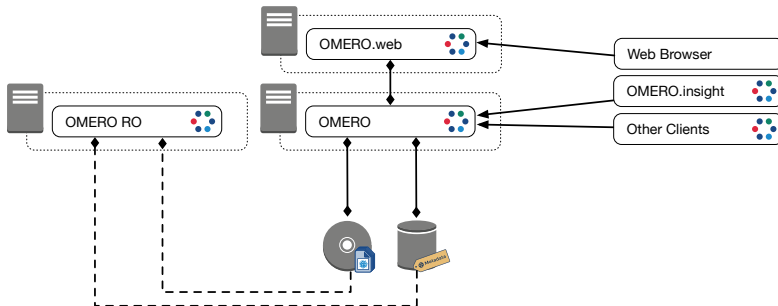
Split



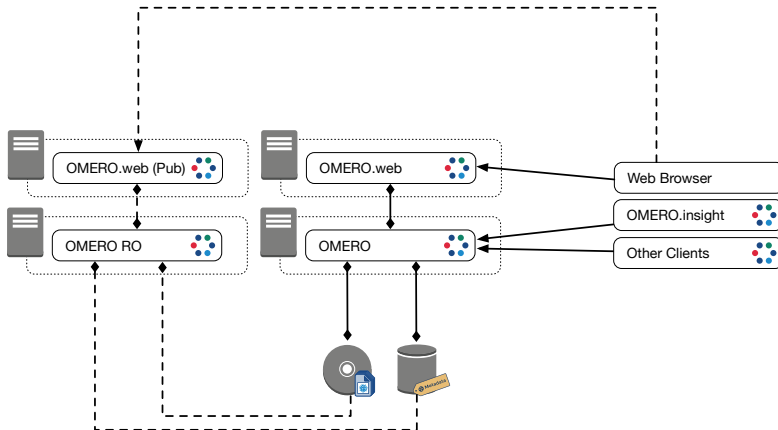
Extra OMERO.web



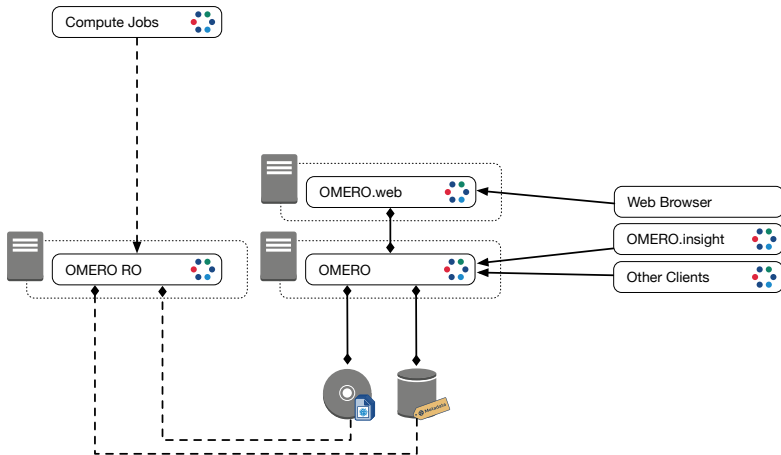
Extra OMERO (Utilising Glencoe Horizontal Scaling)



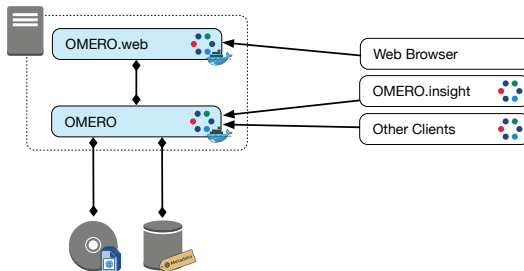
Public Example (Utilising Glencoe Horizontal Scaling)



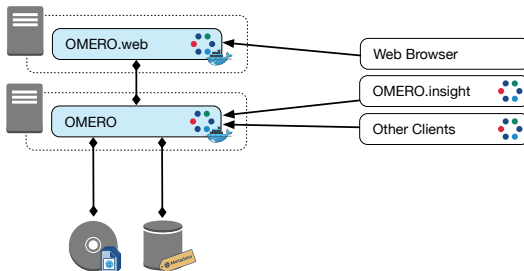
Compute Example (Utilising Glencoe Horizontal Scaling)



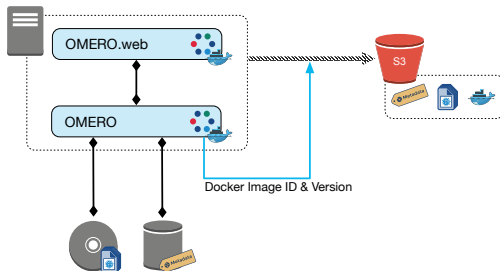
Docker All-In-One



Docker Split



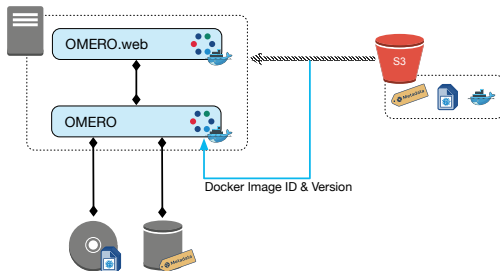
Local Dehydrate



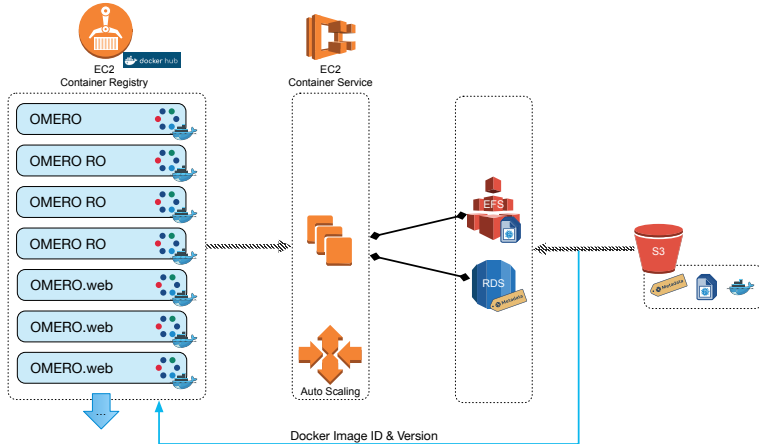
Static



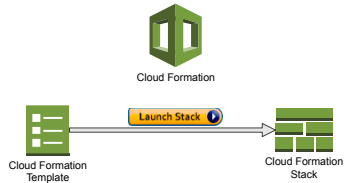
Local Hydrate



Hydrate



Cloudformation



Cloudarchive Cloudformation Template

CloudFormation ▾ Stacks > Create Stack

Create stack

Select Template

Specify Details

Options

Review

Specify Details

Specify a stack name and parameter values. You can use or change the default parameter values, which are defined in the AWS CloudFormation template. [Learn more.](#)

Stack name

Parameters

AsgMaxSize Specifies the number of instances to launch and register to the cluster. Defaults to 2.

DeviceNameForData Specifies the device mapping for the Volume for data storage. Ignored if not using an EBS data volume.

EBSVolumeSizeForData Optional - Specifies the Size in GBs, of the newly created Amazon Elastic Block Store (Amazon EBS) volume for data storage

EBSVolumeSizeForECS Optional - Specifies the Size in GBs, of the newly created Amazon Elastic Block Store (Amazon EBS) volume for ECS container storage

EBSVolumeTypeForData Specifies the Type of (Amazon EBS) volume for data storage. Ignored if not using an EBS data volume.

EBSVolumeTypeForECS Specifies the Type of (Amazon EBS) volume for ECS container storage. Ignored if not using an EBS ECS volume.

EcsEndpoint Optional - Specifies the ECS Endpoint for the ECS Agent to connect to

EcsInstanceType Specifies the EC2 instance type for your container instances. Defaults to t2.medium

EFSOn Switches on EFS use in place of EBS for OMEROCLOUD storage. Will override an EBS Volume for Data, if also set.

Cloudarchive Cloudformation Stack

Services

Resource Groups

SUDO @ syspharm

N. Virginia

Support

CloudFormation

Stacks

Create Stack

Actions

Design template

C

Filter: Active

By Stack Name

Showing 1 stack

Overview

Outputs

Resources

Events

Template

Parameters

Tags

Stack Policy

Change Sets

2017-05-31

Status

10:56:00 UTC+0100

CREATE_COMPLETE

Type

AWS::Logs::LogGroup

Logical ID

CloudwatchLogsGroup

Status reason

10:56:00 UTC+0100

CREATE_IN_PROGRESS

AWS::Logs::LogGroup

CloudwatchLogsGroup

Resource creation Initiated

10:55:59 UTC+0100

CREATE_IN_PROGRESS

AWS::IAM::Role

ECSServiceRole

Resource creation Initiated

10:55:59 UTC+0100

CREATE_IN_PROGRESS

AWS::Logs::LogGroup

CloudwatchLogsGroup

10:55:59 UTC+0100

CREATE_IN_PROGRESS

AWS::IAM::Role

IamRole

Resource creation Initiated

10:55:59 UTC+0100

CREATE_IN_PROGRESS

AWS::EC2::SecurityGroup

EcsSecurityGroup

10:55:58 UTC+0100

CREATE_IN_PROGRESS

AWS::IAM::Role

ECSServiceRole

10:55:58 UTC+0100

CREATE_IN_PROGRESS

AWS::IAM::Role

IamRole

10:55:58 UTC+0100

CREATE_COMPLETE

AWS::ECS::Cluster

ECSCluster

10:55:57 UTC+0100

CREATE_IN_PROGRESS

AWS::ECS::Cluster

ECSCluster

Resource creation Initiated

10:55:57 UTC+0100

CREATE_IN_PROGRESS

AWS::ECS::Cluster

ECSCluster

10:55:52 UTC+0100

CREATE_IN_PROGRESS

AWS::CloudFormation::Stack

OMEROCLOUDARCHIVEPOTN

User Initiated

Feedback

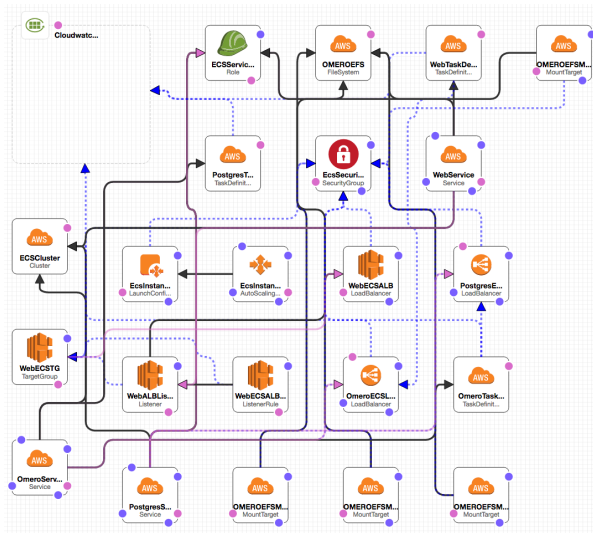
English

© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Cloudarchive Cloudformation Stack



Result

The screenshot displays the OMERO webclient interface in a web browser. The address bar shows the URL: `omero-webec-1ek7gdgaysv1r-1663956677.us-east-1.elb.amazonaws.com/webclient/`. The browser's bookmark bar includes links to 'Code', 'Sys', 'House', 'Technology', 'Finance', 'Sports', 'Harvard', 'Leisure', 'AWS', 'router5 | flexible an...', and 'Baseimage-docker:...'. The OMERO interface has a top navigation bar with 'Data', 'History', 'Help', and 'Admin' links, along with a search bar and a user profile dropdown for 'root root'. The main interface is divided into three panels. The left panel, titled 'Explore', shows a file tree under 'root root' with a folder named 'Orphaned Images' containing five files: '00kingspades.jpeg', '38CTR_01_5_R3D_D3D.jpeg', 'kingspades.jpeg', '...des_underscore-hyphen space.dot.jpg', and 'train.jpg'. The middle panel, titled 'Filter: Add filter', displays five image thumbnails: four identical King of Spades playing cards and one image of a train. The right panel, titled 'General', contains a message: 'This is a virtual container with orphaned images. These images are not linked anywhere. Just drag them to the selected container.'

Future Work

- On-demand image sync
- On-demand image S3 urls
- Auto-generated manifests
- Integration with future Bioformats S3 Reader

Project

- <https://github.com/sorgerlab/omero.cloudarchive-docker>
- <https://github.com/sorgerlab/omero.cloudarchive-cloudformation>

Acknowledgements



HARVARD
MEDICAL SCHOOL



Glencoe
SOFTWARE

Hi+ S



Laboratory of
Systems Pharmacology



NIH LINCS
PROGRAM



OME

Sorger Lab
HMS Research Computing
HMS Image Management Core