

Minerva: A Cloud Native Microscopy Platform

Douglas P.W. Russell

Harvard Medical School

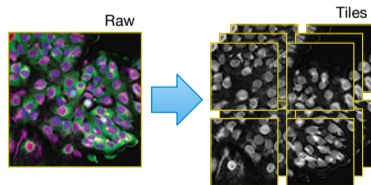
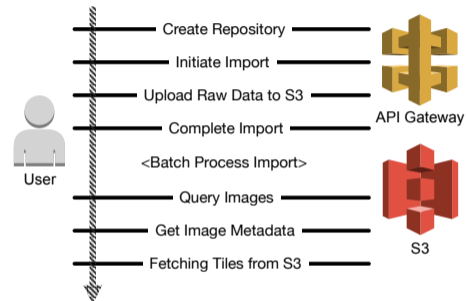
OMERO Annual User Meeting
30th May 2018

Goal: Enable us to leverage the public cloud

- Bring computation to the data
- Native scalability
- Create building-blocks to allow powerful bespoke tools to be built
- Share raw data (as original files or tiled)
- Foster ad-hoc collaboration

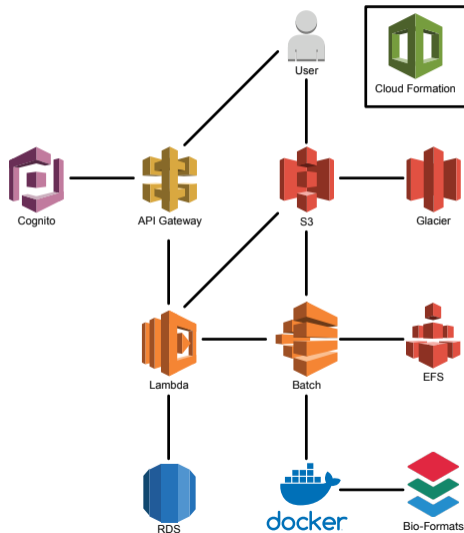
User Workflow

- User interacts with REST web services to create repositories and initiate imports
- Raw data is uploaded wholesale directly to S3 using temporary credentials
- When complete, user initiates the import process for the uploaded data
- Batch processes extract a pyramid of tiles and the metadata XML from the raw data
- User can then fetch metadata of imported images via REST and be granted temporary credentials to pull tiles directly from S3



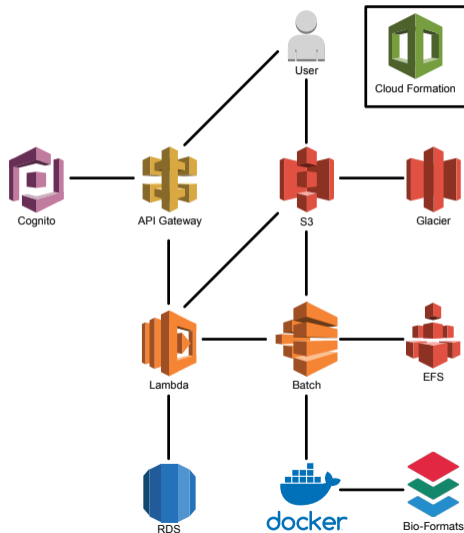
Architecture Overview

- API Gateway: RESTful APIs
- Cognito: User authentication and management
- Lambda (Serverless API backend and internal use):
 - Creating repositories
 - Initiating imports
 - Listing images
 - Requesting temporary credentials for tile access
 - Registering detected Bio Formats units (BFUs)
 - Registering detected images
 - Lifecycling raw data to Glacier



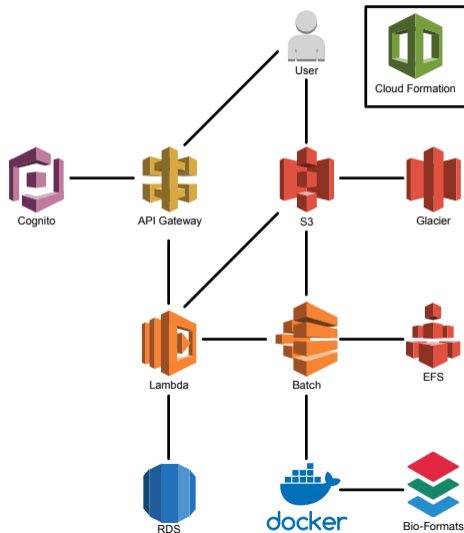
Architecture Overview

- Database: Persistent state management
- S3: Object storage for raw files, tiles and metadata documents
- Glacier: Cheap tiered storage for raw data once processed

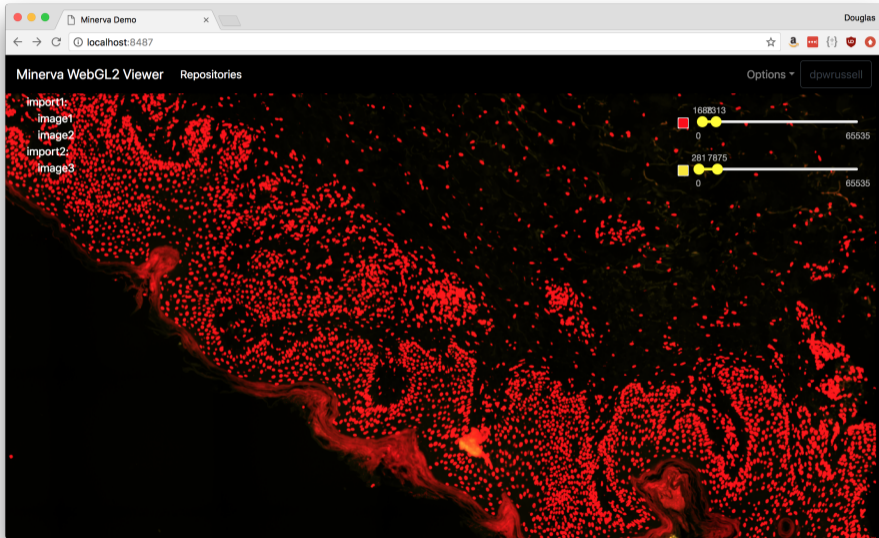


Architecture Overview

- Bio-Formats: Microscopy image detection and extraction
- Docker: Containerised tasks, e.g.
 - Metadata extraction
 - Tile generation
- Batch: Managed execution of containerised jobs
- EFS: Used Internally by Batch for file centric operations
- Cloudformation: Reusable infrastructure provisioning



Browser Based WebGL2 Rendering



Future

- Project is in early stage development
- Integration with analysis tools
- Infrastructure to reprocess data if new Bio-Formats release adds value to that originally processed
- Continue to improve client side rendering in particular. We would like to ultimately build and contribute a WebGL2 drawer for Open Seadragon along with some other enhancements.

Thanks

- Thanks for the support over the last 6+ years!
- Moving Cambridge to Cambridge later in 2018!
- Will work for salary!

Acknowledgements

- Sorger Lab
- Laboratory of Systems Pharmacology
- Harvard Program in Therapeutic Science (HiTS)
- HMS Image Management Core (IMC)
- HMS Research Computing
- HMS Research IT Solutions (RITS)
- NIH LINCS Program



Laboratory of
Systems Pharmacology

HiTS



NIH LINCS
PROGRAM



HARVARD
MEDICAL SCHOOL