



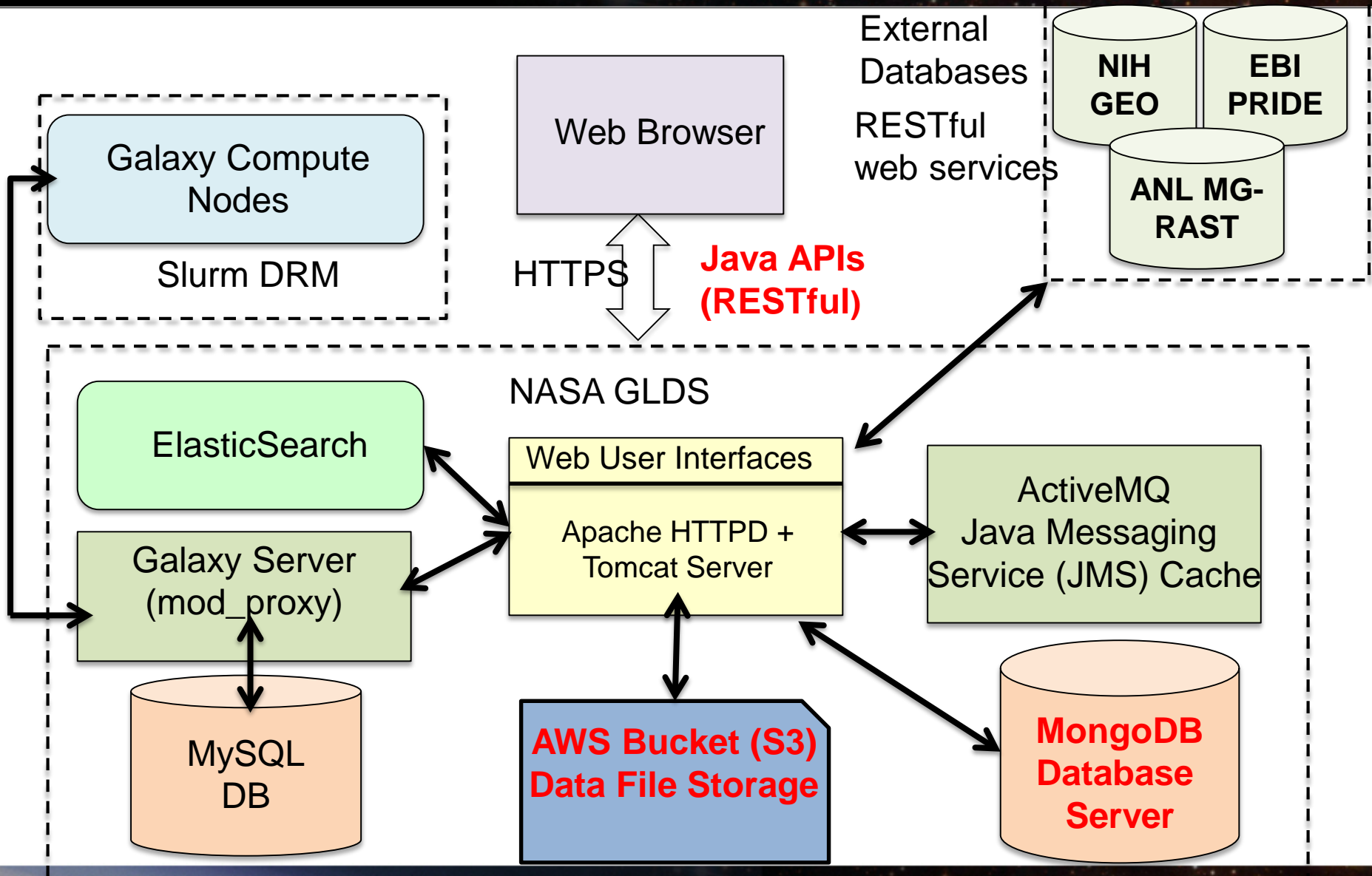
GeneLab Data System (GLDS)

Dan Berrios
Visualization Working Group Meeting
April 2019





GLDS 3.0 System Components





Search APIs

- /genelab/data/search/
 - This API calls the GldsSearch.java servlet
 - Reads from the MongoDB Collections:
 - GLDS_Study, GLDS_StudyMetadata
- /genelab/data/lookahead
 - This API calls the GldsLookahead.java servlet
 - Reads from the MongoDB Collections:
 - GLDS_Study, GLDS_StudyMetadata



Searching Data



<https://genelab-data.ndc.nasa.gov/genelab/data/search/>

{DATA SOURCE}/{TERM}/{FILTER}/{STARTING RANGE}/{COUNT}

(only searches limited metadata)

paramter	Data Type	Notes	Examples
{TERM}	String	search keyword	term=liver
{STARTING RANGE}	integer	starting range of the result	from=0
{HOSTNAME}	String	Server Hostname	
{FILTER}	URL encoded JSON String	filter field name and field values pairing	ffield=
{DATA SOURCE}	String	multiple values should be comma separated	type=cgene
{COUNT}	integer	Number of matches being returned	size=10



Search API Examples



- <https://genelab-data.ndc.nasa.gov/genelab/data/search?term=liver&from=0&type=cgene>



Metadata API

/genelab/data/glds/meta/<ACCESSION #>

- Query for ALL study metadata
- Returns ISA metadata in json format using isa2json library
- <https://genelab-data.ndc.nasa.gov/genelab/data/glds/meta/47>
 - ...
 - a100000samplename: FLT 4-liver
 - a100017rawdatafile: *epigenomics_C-FLT-4.tar.gz



GLDS RESTful APIs (cont'd)



Data File API

https://genelab-data.ndc.nasa.gov/genelab/data/glds/files/<glds_id>

- Query for ALL study data files
- Returns json listing of files, with URLs to access them

<https://genelab-data.ndc.nasa.gov/genelab/data/glds/files/47>

file_name	"GLDS-47_epigenomics_C-FLT-4.tar.gz"
remote_url	"/datamanager/file/Home/Public/genelab/GLDS-47/epigenomics/GLDS-47_epigenomics_C-FLT-4.tar.gz"
file_size	24429284700
date_created	1506008635156



Additional Info





Background Information:

GLDS has the following Java-based customized web applications and shell scripts developed and deployed since Release 2.0 (aka Phase 2) in Sept. 2017 timeframe and have been in continuous improvements since then and into Release 3.0.

- GLDS Data Visualization: Data viz for correlations between organisms, GLDS accession numbers, and assay types <https://genelab-data.ndc.nasa.gov/dataviz>
- GLDS Administrator: Web admin tool for curation process <https://genelab-data.ndc.nasa.gov/gldsadmin> (internally accessible to nasa.gov networks only; Requires LaunchPad PIV login and access controlled via NASA email address)
- GLDS GeneLab: Main web app for search, navigation, and browse data assets <https://genelab-data.ndc.nasa.gov/genelab>



- **GLDS S3 Monitor:** Real-time S3 bucket monitoring service for security checking whitelisted file content types, file move/rename operations, etc. leveraging from AWS Simple Queue Service (SQS) and Simple Notification Service (SNS)
- **GLDS Data Synchronization:** Series of shell scripts used for routine S3 data sync and database export/import process (e.g., MongoDB, MySQL) via manual or scheduled cronjobs process
- **GLDS QA/Test Process:** Utilizes JUnit testing framework from GenomeSpace platform and integrate with TI babelfish's Bamboo Continuous Integration (CI) framework



Customized GS Web Apps



File Structure and Setup Layout:

Server location: `$TOMCAT_HOME/webapps/genelab`

Java code locations: `x-gene-genomespace/glds-genelab/src/main/java/gov/nasa/arc/glds`

`/resource` → Contains the main java servlet endpoints

`/GldsSearch.java`

`/GldsLookahead.java`

`/GldsRestApi.java`

`/ManageStudyResource.java`

`/StudyResourceAccessionFilter.java`

`/StudyResourceFileRedirect.java`

`/persistence` → Contains all the java servlet helper functions

`/elasticsearch/ElasticsearchPersistenceManager.java`

`/mongodb/MongoStudyPersistenceManager.java`



Customized GS Web Apps



File Structure and Setup Layout – Continued:

x-gene-genomespace/glds-genelab/src/main/webapp/

/js → Contains JavaScript

/img → Contains the GLDS Study icons and images

/css → Contains Cascading Style Sheets

/help → GLDS Tutorials dropdown menu pdf documents

/ISA → Addt'l GLDS Tutorials dropdown menu pdf documents

/html → Static html web pages (e.g. Environmental Radiation)

/jsp → Dynamic html web pages (e.g. Data Repository, GLDS Study pages)

/META-INF → Contains Tomcat specific web app's metadata information, such as a security context, etc.

/WEB-INF → Contains Java classes, libraries, web.xml configuration file, and JSP tags (global header and html container)