

Using and Debugging Rya Examples



Rya Working Group
April 6, 2016

Environment

Apache Maven 3.3.1

Java version: 1.8.0_45

Windows 7 and Mac OS X

Eclipse IDE for Java Developers

Luna Service Release 2 (4.4.2)

Rya Version 3.2.10-SNAPSHOT

Branch: develop

Latest commit efc815e

Checking out from Git and Building Rya

Check out Rya:

```
git clone https://github.com/apache/incubator-rya.git
cd incubator-rya/
git checkout develop
```

Building Rya:

```
mvn clean install -DskipTests
```

May take ~30 minutes to download dependencies.

Last build took 26 minutes

My build took 3.5 minutes

Rya Modules

Apache Rya SAIL

[incubator-rya/sail/](#)

Apache Rya Common API

[incubator-rya/common/rya.api/](#)

Apache Rya Accumulo DAO

[incubator-rya/dao/accumulo.rya/](#)

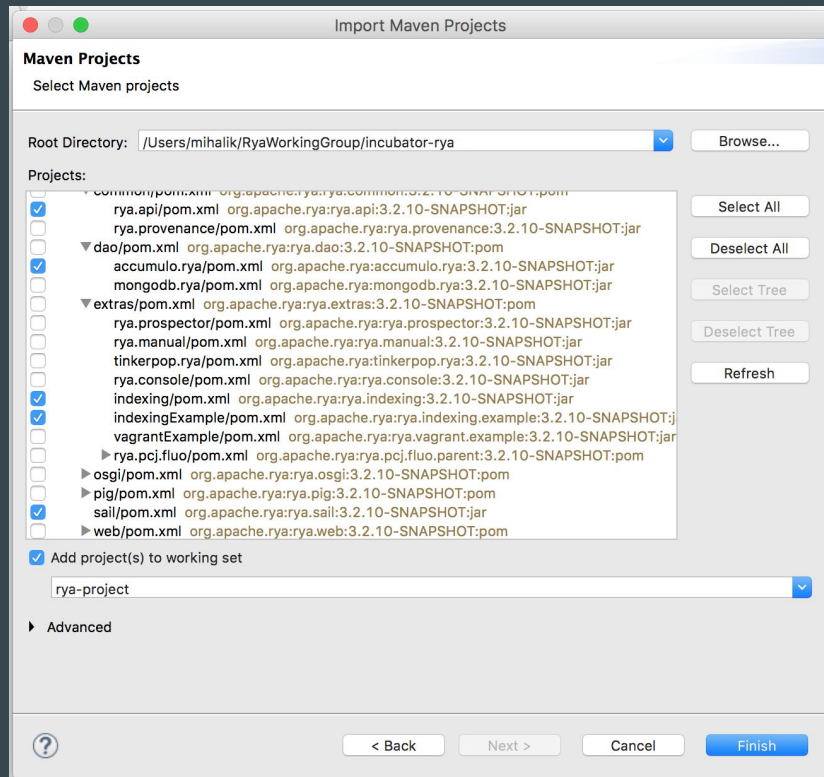
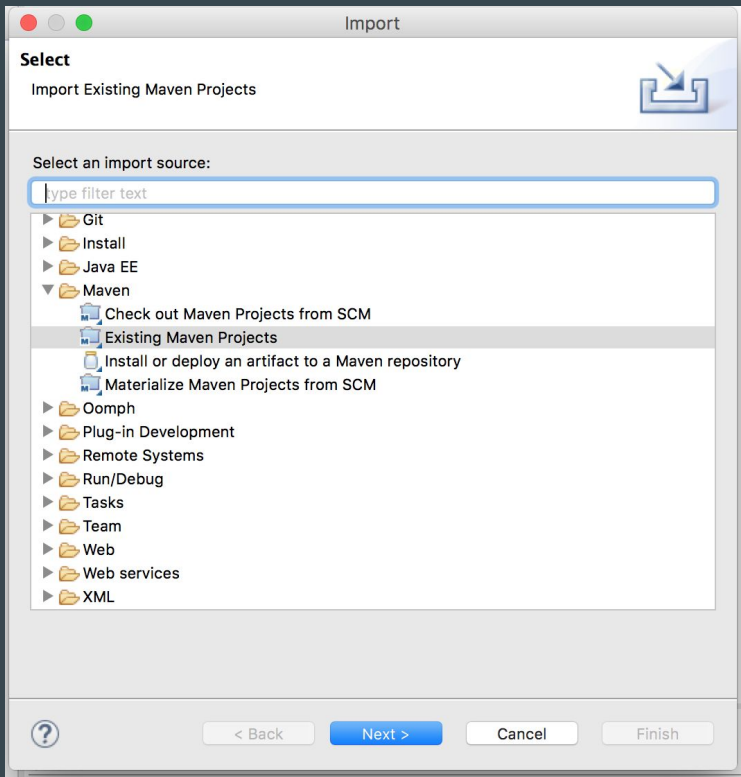
Apache Rya Secondary Indexing

[incubator-rya/extras/indexing/](#)

Apache Rya Secondary Indexing Example

[incubator-rya/extras/indexingExample/](#)

Importing Rya into Eclipse



Rya Direct Example (Background)

Secondary Indexing Example project contains basic examples for getting up to speed with the Rya SAIL/Java Client and Secondary Indexing

“Java Mains” in `extras/indexingExample/src/main/java/`

Uses “Mock” (In-memory) Accumulo Instance

Can be configured to use other Accumulo Instance

Note fields at the top of the class and the `getConfig()` method

Currently, `RyaDirectExample` is broken for “Mock”

Comment out ln 97 “`createPCJ()`”

Configuring the Rya Direct Example

Accumulo Connection Configuration

Set the table Prefix:

CONF_TBL_PREFIX

Use in memory Accumulo:

USE MOCK_INSTANCE

Accumulo Connection Params:

CLOUDBASE_USER

CLOUDBASE_PASSWORD

CLOUDBASE_ZOOKEEPERS

CLOUDBASE_INSTANCE

CLOUDBASE_AUTHS

Secondary Indexing Configuration

Enable Secondary Indices:

USE_PCJ

USE_GEO

USE_FREETEXT

USE_TEMPORAL

Partitions for Indices (Geo, Freetext):

NUM_PARTITIONS

Predicates to index:

GEO_PREDICATES_LIST

Connecting to Rya

From RyaDirectExample line 92:

```
final Sail extSail = RyaSailFactory.getInstance(conf);  
  
SailRepository repository = new SailRepository(extSail);  
  
repository.initialize();  
  
SailRepositoryConnection conn = repository.getConnection();
```


Rya Direct Example

Understanding Rya Architecture by debugging:
SPARQL Example: Add and Delete

Try to answer the following questions:

How does Rya interact with OpenRdf libs?

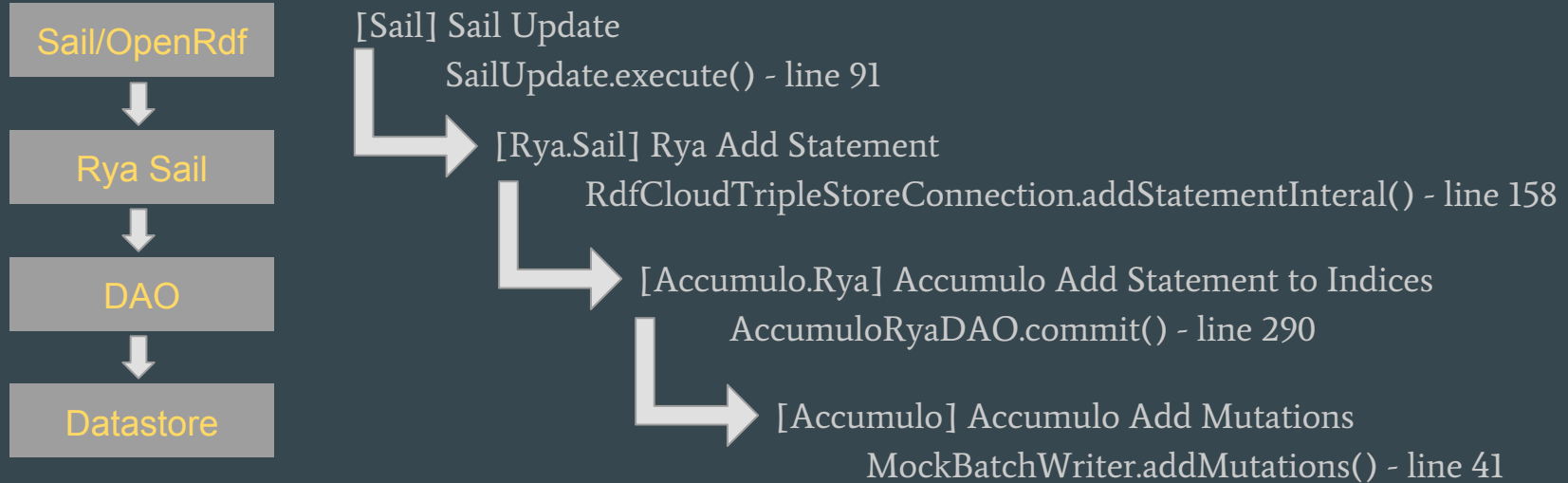
Where does it become Rya Specific?

Debugging Approaches:

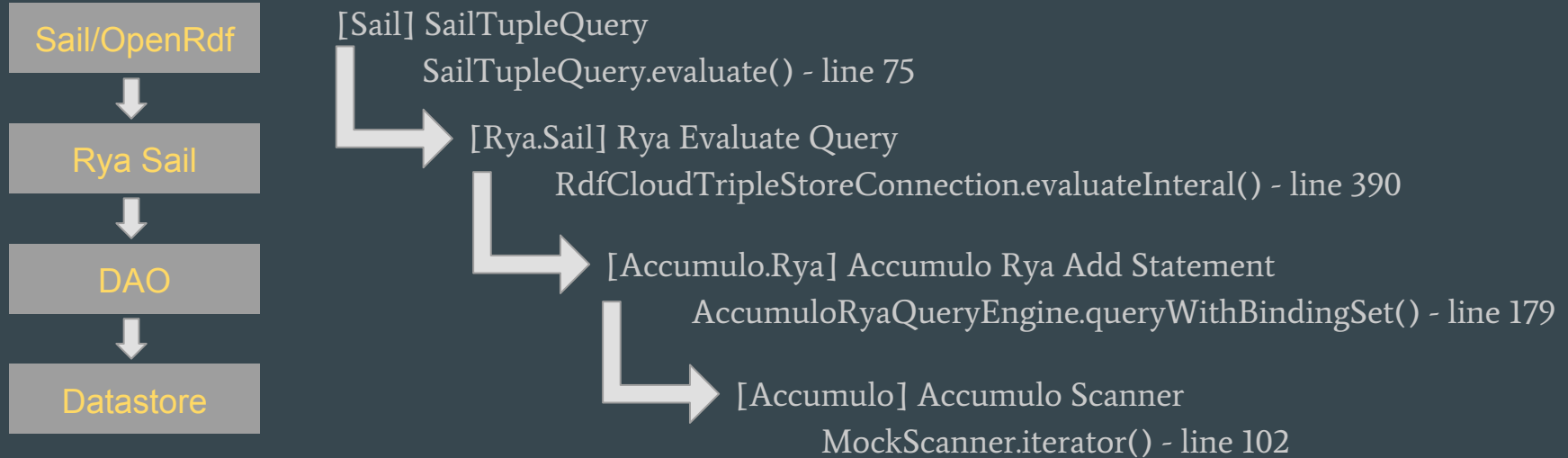
Top Down: Step through Code

Bottom Up: Set breakpoints in Accumulo Scanner and BatchWriter

Rdf Updates



Rdf Query



RdfCloudTripleStoreConnection.evaluateInternal()

Collect Provenance/Auditing (line 212)

Configure the query parameters (line 222)

Query processing, Accumulo parameters, etc.

Optimize the query

Match for Pre-Computed Join (line 280-293)

Out of the Box Open Rdf Optimizers (line 299-337)

Rya Secondary Indexer optimizers (line 314)

Apply Backwards Chaining inferencing (line 346-351)

Optimize using Statistics (line 365)

Perform Query (line 389)

Resources

Rya Papers and Presentation:

[Rya a scalable rdf triple store for the clouds](#)

[SPARQL in the Cloud using Rya](#)

[Rya: Optimizations to Support Real Time Graph Queries on Accumulo](#)

Mailing list:

dev@rya.incubator.apache.org