

Islandora Batch

Overview

The Islandora Batch module provides the ability to ingest multiple objects at a time, either from a .zip file or a directory on the server. The ingest is a three-step process:

- **Preprocessing:** The data is scanned, and a number of entries created in the Drupal database. There is minimal processing done at this point, so it can complete outside of a batch process.
- **Ingest:** The data is actually processed and ingested. This happens inside of a Drupal batch.
- **Cleanup:** The batch entries in the Drupal database need to be deleted, so the associated temp files can be purged. This can be configured to happen automatically, or can be done manually.

Dependencies

This module requires the following modules/libraries:

- [Islandora](#)
- [Tuque](#)

Additionally, installing and enabling [Views](#) will allow additional reporting and management displays to be rendered.

Downloads

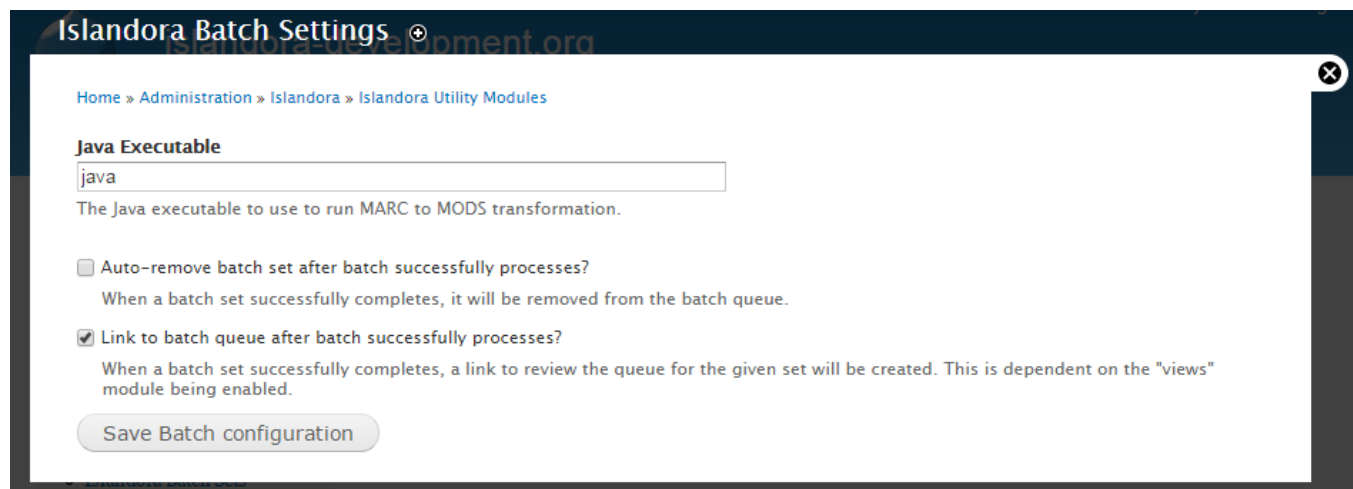
[Release Notes and Downloads](#)

Installation

Unzip this module into your site's modules directory as you would any other contrib module. See [this](#) for further information.

Configuration

After you have installed and enabled the Islandora Batch module, go to Administration » Islandora » Islandora Utility Modules » Islandora Batch Settings (admin/islandora/tools/batch) to configure the module.



The screenshot shows the 'Islandora Batch Settings' configuration page. At the top, there is a breadcrumb trail: 'Home » Administration » Islandora » Islandora Utility Modules'. Below this, the 'Java Executable' field is set to 'java'. A description states: 'The Java executable to use to run MARC to MODS transformation.' There are two checkboxes: 'Auto-remove batch set after batch successfully processes?' (unchecked) and 'Link to batch queue after batch successfully processes?' (checked). A description for the second checkbox states: 'When a batch set successfully completes, a link to review the queue for the given set will be created. This is dependent on the "views" module being enabled.' At the bottom, there is a 'Save Batch configuration' button.

You should make sure that the path to your java executable is correct. Optionally, if you have the Drupal Views module enabled, you can also have the module link back to the Batch Queue in its results messages.

Usage

Detailed usage instructions for Islandora Batch are available [here](#).

Drush Integration

The Islandora Batch module can also be run using Drush. This is a two stage process. First you populate the Batch Queue by running the `islandora_batch_scan_preprocess` command.

Drush made the `target` parameter reserved as of Drush 7. To allow for backwards compatability we use `target` in Drush 6 and below, but `scan_target` in Drush 7 and up.

Drush 7 and above:

```
drush -v -u 1 --uri=http://localhost islandora_batch_scan_preprocess --type=zip --scan_target=/path/to/archive.zip
```

Drush 6 and below:

```
drush -v -u 1 --uri=http://localhost islandora_batch_scan_preprocess --type=zip --target=/path/to/archive.zip
```

Available options include (Drush 7 and above):

<code>--content_models</code> objects.	A comma-separated list of content models to assign to the
<code>--namespace</code> Fedora configuration.	Namespace of objects to create. Defaults to namespace specified in
<code>--parent</code> to the root Islandora	The collection to which the generated items should be added. Defaults
PID.	repository
<code>--parent_relationship_pred</code> "isMemberOfCollection".	The predicate of the relationship to the parent. Defaults to
<code>--parent_relationship_uri</code> to	The namespace URI of the relationship to the parent. Defaults
external#".	"info:fedora/fedora-system:def/relations-
<code>--scan_target</code> (or <code>--target</code>)	The target to directory or zip file to scan. Required.
zip	Requires the full path to your archive from root directory, e.g. <code>/var/www/drupal/sites/archive.</code>
<code>--type</code>	Either "directory" or "zip". Required.

After the Batch Queue has been populated, you start the actual ingest process with the `islandora_batch_ingest` command. For example:

```
drush -v -u 1 --uri=http://localhost islandora_batch_ingest
```

This will process the batch, ingesting the objects and creating their derivatives.

Batch Ingest Cleanup

It is important to clean up completed batches periodically, as the record of the batch prevents the associated Drupal temp files - i.e. any uploaded payload - from being automatically deleted. There is a configuration option that will do this immediately after a successful ingest (see config screenshot, above). It is possible (but not in scope here) to create a cron task to do this after a period of time. It is also possible to delete batches manually through the interface. An illustrated explanation of this process is found at [How to Batch Ingest Files#BatchIngestCleanup](#).

Customization

Custom ingests can be written by [extending](#) any of the existing preprocessors and batch object implementations. Checkout the [example implementation](#) for more details.

Clearing the semaphore table

If a user kills a Drush batch ingest, or a batch ingest initiated via the web GUI dies for some reason, it is impossible to start another batch ingest until the Islandora Batch entry in Drupal's `semaphore` table expires. You may clear this entry manually within your database, but doing so may impact other batch ingest jobs that are running. If you are sure no other batch ingest jobs are running, delete the row from Drupal's `semaphore` table where the `name` is 'islandora_batch_ingest'.