# Upgradating sample production data from Ginkgo.2 to Ironwood.master via Hawthorn.2 in Development Environment

This process is performed before migrating data to the live server.

**Following machines are used in the upgradation process:**

Ginkgo.2 10.129.103.62

Hawthorn.2 10.129.103.26

Ironwood.master 10.129.103.27

**Note:** Most of the time the “notifier-celery-workers” service raises a fatal error on Hawthorn.2 10.129.103.26 and Ironwood.master 10.129.103.27 machines, as

notifier-celery-workers FATAL Exited too quickly (process log may have details)

It is due to

*File "/edx/app/notifier/virtualenvs/notifier/local/lib/python2.7/site-packages/django/db/backends/sqlite3/base.py", line 318, in execute*

*return Database.Cursor.execute(self, query, params)*

*OperationalError: database is locked*

# Observations:

Following are the observations noted.

## Version of databases:

On the Ginkgo.2 machine:

Mongo: MongoDB shell version: 2.6.12

MySQL: mysql Ver 14.14 Distrib 5.6.32, for debian-linux-gnu (x86\_64) using EditLine wrapper

On the Hawthorn.2 machine:

Mongo: MongoDB shell version: 3.2.21

MySQL: mysql Ver 14.14 Distrib 5.6.32, for debian-linux-gnu (x86\_64) using EditLine wrapper

On the Ironwoord.master machine:

Mongo: MongoDB shell version: 3.2.16

MySQL: mysql Ver 14.14 Distrib 5.6.32, for debian-linux-gnu (x86\_64) using EditLine wrapper

## Gingko Sample Data

**The additional tables in Ginkgo sample production data dump:** Below are the extra tables found in the sample production data received

* *coursetalk\_coursetalkwidgetconfiguration (It is absent, course talk feature is a paid service and now IITBombayX is not using this feature)*
* *courseware\_course\_subject*
* *courseware\_organization*
* *courseware\_subject*
* *student\_teammember*

All 11 tables of celery is present in “edxapp” MySQL database.

**On the Fresh Hawthorn.2 machine**

The mongo collection *“system.indexes” in the “edxapp” database is* missing, it is due to the Mongo version change.

The *“cs\_comments\_service” database is* not present in the Mongo.

All 11 tables of celery are present in “edxapp” MySQL database.

**On the Fresh Ironwood.master machine**

The mongo collection *“system.indexes” in the “edxapp” database is* missing, it is due to the Mongo version change.

In the *“cs\_comments\_service” database, “users” collection is present with 1 record of "username" is "staff" and “external\_id” is 11.*

*All 11 tables of celery is present in “edxapp” MySQL database.*

**After Ginkgo.2 data restore on the Hawthorn.2 machine:**

The mongo collection “system.indexes” is missing in the “edxapp” and *“cs\_comments\_service” databases, it is due to the Mongo version change.*

*All 11 tables of celery is present in “edxapp” MySQL database.*

## Before migrating data to Hawthorn.2

The Ginkgo.2 data dump is restored on the Hawthorn machine. Following are the observations.

* To migrate data from Ginkgo to Hawthorn, it needs to drop the database tables used by djcelery. These tables should be empty in Ginkgo data, so it is safe to drop them. The edx-platform application has a management command to check that they are empty and drop them, the error occur at time of running the command is as follow

*DROP TABLE IF EXISTS `celery\_taskmeta` CASCADE*

*django.db.utils.IntegrityError: (1217, 'Cannot delete or update a parent row: a foreign key constraint fails')*

## After migrating data to Hawthorn.2

After porting the sample production data dump in Hawthon.2, Following are the observations.

The mongo collection is missing on Hawthorn.2 Machine.

*Collection “system.indexes” in the “cs\_comments\_service” database, it is due to the Mongo version change.*

As per the fresh machine instance, below listed tables are to be deleted but it is not deleted from the migrated data Hawthorn.2 Machine:

* assessment\_aiclassifier
* assessment\_aiclassifierset
* assessment\_aigradingworkflow
* assessment\_aitrainingworkflow
* assessment\_aitrainingworkflow\_training\_examples
* ccxcon\_ccxcon
* ccx\_ccxfieldoverride
* ccx\_customcourseforedx
* corsheaders\_corsmodel
* thumbnail\_kvstore
* coursetalk\_coursetalkwidgetconfiguration(It’s as per the previous result.)

**Note:** However If there are extra tables in the database, it does not affect the functionality of the machine.

## Before porting to Ironwood.master

The Hawthorn.2 data dump is restored on the Ironwood.master machine. Following are the observations.

* To migrate data from Hawthorn to Ironwood, it needs to drop the database tables used by djcelery. These tables should be empty in Hawthorn data, so it is safe to drop them. The edx-platform application has a management command to check that they are empty and drop them, the error occur at time of running the command is as follow

*LOCK TABLES `djcelery\_taskstate` WRITE, `djcelery\_periodictask` WRITE, `djcelery\_intervalschedule` WRITE, `djcelery\_crontabschedule` WRITE, `djcelery\_workerstate` WRITE, `celery\_tasksetmeta` WRITE, `celery\_taskmeta` WRITE, `djcelery\_periodictasks` WRITE*

*django.db.utils.ProgrammingError: (1146, "Table 'edxapp.djcelery\_taskstate' doesn't exist")*

**Git diff in "/var/tmp/configuration" directory on Ironwood.master Machine:**

*-EDXAPP\_CUSTOM\_COURSES\_EDX: false*

*+EDXAPP\_CUSTOM\_COURSES\_EDX: true*

## After porting to Ironwood.master

The data dump is ported from the Hawthorn machine to Ironwood.master. Following are the observations.

The error occur during running the migration process on Ironwood.master machine is as:

*ProgrammingError: (1146, \"Table 'edxapp.content\_type\_gating\_contenttypegatingconfig' doesn't exist\")",*

**Note**: Without any modification in the platform setting and in the database, the migration script ran next time successfully.

The same mongo collection is absent on Ironwood.master:

*Collection “system.indexes” in the “cs\_comments\_service” database.*

***Note:*** *It is missing on the Hawthorn.2 machine.*

As per the fresh machine instance, below listed tables are to be deleted but it is not deleted from the migrated data Ironwood.master Machine:

* course\_structures\_coursestructure
* coursetalk\_coursetalkwidgetconfiguration(It’s as per the previous result.)

**Note:** However If there are extra tables in the database, it does not affect the functionality of the machine.

Extra JWT authenticated user “login\_service\_user” present on the Ironwood.master machine.

**Note:** It is created due to setting in “./edx-platform/lms/envs/common.py” file.

The number of entries in the "django\_migrations " table is responsible for the whole migration process. The record count of the “django\_migrations” table in the “edxapp” and “edxapp\_csmh” databases on Ironwood.master Machine:

* Row count in django\_migrations table of edxapp database: 511
* Row count in django\_migrations table of edxapp\_csmh database: 511

**Note:** It was 510 records in the previous round result. (the previous round machines have been released, so unable to find exactly which migration file is added.)

**In Testing:**

The “custom fields on the registration page of the OpenedX” application shouldn't be installed on the Ironwood.master 10.129.103.27 machine. It would be install after the migration.

**After Testing:**

As per the celery tables on the Ginkgo.2 machine, the list of the celery tables are as follow:

* celery\_taskmeta
* celery\_tasksetmeta
* celery\_utils\_chorddata
* celery\_utils\_chorddata\_completed\_results
* celery\_utils\_failedtask
* djcelery\_crontabschedule
* djcelery\_intervalschedule
* djcelery\_periodictask
* djcelery\_periodictasks
* djcelery\_taskstate
* djcelery\_workerstate

The celery tables on the Howthorn.2 machine and the Ironwood.master machine are as follow:

* celery\_taskmeta
* celery\_tasksetmeta
* celery\_utils\_chorddata
* celery\_utils\_chorddata\_completed\_results
* celery\_utils\_failedtask
* djcelery\_crontabschedule
* djcelery\_intervalschedule
* djcelery\_workerstate