Migrating DSpace to a new server

These instructions are meant as a general guideline to how you can migrate your DSpace site/data to a new server while also Upgrading DSpace to the latest release. Keep in mind that you MUST also review the Installing DSpace and Upgrading DSpace guides when performing a migration (e.g. you must ensure you have correct dependencies installed and you must ensure you perform all upgrade steps).

Overview of migration approaches

There are two main approaches to migrating your DSpace to a new server.

1. Install a fresh copy of DSpace & migrate the database/files into it - This is the approach documented on this page. It is the recommended approach as it ensures zero data loss. However, it does involve more steps to complete the migration.
2. Install a fresh copy of DSpace & use the AIP Backup and Restore - This is an alternative approach where you can use the AIP export tools to export AIPs from your old site, and then import them into the new site. While this also works, keep in mind that not all DSpace data can be exported to AIPs, so you will lose some data during this migration (namely any submissions not yet completed or still in workflow approval will be lost, see the AIP Backup and Restore documentation for more details on what data is currently missing from AIPs).

Step 1: Install a fresh copy of DSpace

On your new server, follow the Installing DSpace instructions and install a fresh (empty) copy of the latest version of DSpace. **BEFORE PROCEEDING**, ensure that this fresh copy of DSpace is correctly installed and shows no errors when you startup the site. (The site will obviously appear empty though, and that's OK)

You can also use this time to get your basic configurations setup properly for both the backend (local.cfg) and the frontend (config.prod.yml).

Step 2: Prepare your data to copy from the old DSpace to the new one

There are three main areas of data you need to migrate in order to ensure no data loss.

Perform these steps on the *old server*

1. First, you should STOP tomcat on the old server. These steps require the site to be down.
2. Update sequences *(optional)* - When migrating content, sometimes sites will find that database sequences will be outdated or incorrect. This can result in "duplicate key" errors during the database migration to the latest version. To avoid this, **before** you export your data, run this older copy of the "update-sequences" command on your database. This should ensure your database sequences are updated before you dump your data.

```
# If upgrading from DSpace 6 or below, run this on your old database
psql -U [database-user] -f [dspace]/etc/postgres/update-sequences.sql [database-name]
# e.g. psql -U dspace -f [dspace]/etc/postgres/update-sequences.sql dspace
```

   a. NOTE: It is important to run the "update-sequences" script which came with the OLDER version of DSpace (the version you are migrating from)! If you've misplaced this older version of the script, you can download it from our codebase & run it via the "psql" command above.
      i. DSpace 6.x version of "update-sequences.sql": https://github.com/DSpace/DSpace/blob/dspace-6_x/dspace/etc/postgres/update-sequences.sql
      ii. DSpace 5.x version of "update-sequences.sql": https://github.com/DSpace/DSpace/blob/dspace-5_x/dspace/etc/postgres/update-sequences.sql
3. The database data - Make sure to export the database data from your old DSpace site using a tool like "pg_dump" (for PostgresSQL). If you use "pg_dump", you'll end up with a large SQL file which contains all the data from your old database.

```
# Example of using pg_dump to export a database to an output file
pg_dump -U [db_username] [db_name] > [output_file.sql]
```

4. The "assetstore" folder - This folder is in your DSpace installation directory and it contains all the files stored in your DSpace. You will need **all the contents** of this folder (including all subdirectories), so you could choose to zip it up or you could copy it over directly.
5. The Solr data *(optional)* - Both DSpace authority and statistics are stored in Solr. If you want to keep these, you will want to export them from the old Solr and move them over. Use the "solr-export-statistics" tool provided with DSpace: see "Export SOLR Statistics" in the Solr Statistics Maintenance guide. *(Requires Solr to be running. Keep in mind, this may require you to start Tomcat back up if Solr is running in Tomcat.)*

Step 3: Copy over the prepared data and import it into the new DSpace

Copy the data you've prepared in Step 2 over to the new server.

Now, you'll import this data into your new installation of DSpace (created in Step 1).

Perform these steps on the *new server.*

1. First, you must STOP Tomcat on the new server.
2. The database data - Before you can import the data, you must delete the new, empty database.
   a. Delete/Clean the new, empty database (created in step 1) as you will have empty tables created during the installation. The easiest way to achieve this is to run the "./dspace database clean" command. Keep in mind it requires temporarily enabling it via "db. cleanDisabled=false" in your local.cfg. (After the "clean" command succeeds, make sure to remove this configuration.)

```
# Delete everything in your database
# Requires temporarily setting "db.cleanDisabled=false" in your local.cfg
./dspace database clean
```

      i. Alternatively, PostgreSQL users could delete the entire database (using dropdb command, e.g. "dropdb -U [db_username] [db_name]") and recreate it based on the "Database Setup" instructions in Installing DSpace.
   b. Import the database dump you created in Step 2 (above), which will recreate this with all your old data in it. For Postgres, you can use the "psql" command.

```
# Example of using psql to import data from a SQL file into a database
psql -U [db_username] [db_name] < [output_file.sql]
```

      (NOTICE the direction of the angle character... in this command you are telling Postgres to execute all the commands contained in your "output_file.sql", which will cause it to recreate all the database data in your new database.)
3. The "assetstore" folder - Delete the **empty** assetstore folder on the new server. Copy the entire assestore folder (and all subdirectories) from the old server to the new one. In the end, you should have a several subdirectory hierarchies (containing your files) under the [dspace] /assetstore/ folder on the new server.
4. The Solr data (optional) - If you exported the statistics or authority data in Step 2, then you can import this data from the exported files using the "solr-import-statistics" tool provided with DSpace, see "Import SOLR Statistics" in the Solr Statistics Maintenance guide. *(Requires Solr to be running)*

Step 4: Update the database, Start DSpace & Reindex

Now that all the data is copied over, you need to ensure it's updated and reindexed properly (for the new version of DSpace).

Perform these steps on the *new server.*

1. Migrate/Upgrade the database to the latest version - Now that your old data is migrated, you MUST ensure it's using the latest database updates based on the new DSpace you've installed. Review the database steps in Upgrading DSpace and follow the instructions there.

```
# Migrate your old data to the latest DSpace version
# WARNING: You must review the Upgrading DSpace docs to see if there are any additional database steps
listed there!
./dspace database migrate ignored
```

   NOTE: You should check the logs (dspace.log) for errors. Additional steps may be documented in the Upgrading DSpace guide.
2. Start Tomcat. This will bring your new DSpace back up, with the migrated data in place. Check the backend logs (dspace.log and Tomcat log) to ensure no errors occur on startup.
3. Reindex all content - This will ensure all search/browse functionality works in the DSpace site. Optionally, if you use OAI-PMH, you will want to reindex content into there as well.

```
# Reindex all your content in DSpace
./dspace index-discovery -b

# (Optionally) also reindex everything into OAI-PMH endpoint
./dspace oai import
```

   NOTE: Until this command completes (it may take a while for large sites), you will not be able to fully browse/search the content from the User Interface. To check the progress of the reindex, check your dspace.log file.

Step 5: Review the Upgrade Instructions and final cleanup

If you've changed the version of DSpace you are running, you should review the Upgrading DSpace guide for instructions related to necessary configuration changes or other necessary updates. You should perform any upgrade steps that you have NOT yet performed above (keep in mind you already should have upgraded your database, and reindexed your content).

At this time, you also may wish to review your configurations on your *old* DSpace site, and see if there are any configurations that you wish to copy over into your new DSpace site. This step is optional, as you can also choose to start "fresh" with a new local.cfg file.

FINALLY, test the new site and verify that all the content, user accounts, etc. have moved over successfully. If you encounter any issues, see our Troubles hoot an error guide for hints/tips on finding the underlying error message & reporting it to Support lists/channels. Also make sure to check our list of Comm on Installation Issues in the Installing DSpace guide.