

DSpace Installation Documentation for Edinburgh University Library

Introduction

This document will explain the main procedures that must be completed in order to build and install a working version of DSpace. The documentation is specific to our Red Hat 8.0 Linux install, but should be a good guide to getting it working on most Unix or Unix-based platforms.

Note that there are a number of things conspicuously missing from this document. These include: Configuring SSL for DSpace, Installing the Handle Server, Setting the Java Environment, and detailed guides on installing the pre-requisite software. At the end of this document there are short sections about the first three issues, and we hope to add more to all of this in the future. Your feedback would be most helpful.

Notation

Throughout the course of this document we will use the following notation:

[`dspace`] – The target installation directory of your DSpace system. For example, we use `/u01/dspace`

[`dspace-src`] – The directory into which you have loaded the DSpace code ready for building and installation. For example, we use `/u01/dspace-1.1`

[`postgres-src`] – This is the source from which you will build and install the PostgreSQL database system. For example, we use `/u01/postgres/postgresql-7.3.2`.

[`postgres`] – The directory in which you will/have install(ed) PostgreSQL. For example, we use `/u01/postgres`

[`postgres-data`] – This is the location of the PostgreSQL database, which will then contain the data from DSpace. For example, we use `/u01/postgres/data`. Note that this directory must have `rwX` permissions for the PostgreSQL user (to be created in the section *Installing PostgreSQL 7.3.2*).

[`tomcat`] – The home directory of your Tomcat installation. For example, we use `/usr/local/jakarta-tomcat-4.0.6`

[`apache`] – The home directory of your Apache installation. For example, we use `/usr/local/apache`

Prerequisites

Java 1.3

JavaBeans Activation Framework – place `activation.jar` in the `[dspace-src]/lib` directory.

JavaServlet 2.3 & JSP 1.2 – place `servlet.jar` in the `[dspace-src]/lib` directory.

JavaMail API – place `mail.jar` in the `[dspace-src]/lib` directory.

Jakarta Tomcat 4.0.6

Compile `mod_webapp` (hard).

Apache 1.3

The developers have not yet tested it with Apache 2.0, but believe that it should work fine.

Ant

If using PostgreSQL 7.2.3 or above Ant 1.5 is required, otherwise Ant 1.4 will suffice (see *Installing PostgreSQL 7.3.2* for more information)

PostgreSQL

We use 7.2.3 – see the installation guide for more information

Installing PostgreSQL 7.3.2

Note that DSpace comes packaged with a `postgresql.jar` file for an older version of PostgreSQL. To make DSpace work with PostgreSQL 7.3.2 it is necessary to replace their version of `postgresql.jar` with a new one. That you will be performing this procedure is assumed, and is explained in this section.

Note also that you require Ant 1.5 in order to build PostgreSQL 7.3.2 using the `--with-java` option required to create the drivers for DSpace.

- Create a user called `postgres` and a group of the same name.
- Login as `postgres` in order to do the installation.

- Configure the source:

Go to the directory [postgres-src], and run the following command (but see the note at the bottom first):

```
./configure --prefix=[postgres]
            --enable-multibyte
            --enable-unicode
            --with-java
```

Note that we have not yet built our version of PostgreSQL with the --enable-multibyte and --enable-unicode options, although we will do so before we produce a live service. As such we cannot vouch for any problems that might occur.

- gmake PostgreSQL:

In [postgres-src] run the command:

```
gmake
```

- Install PostgreSQL:

In [postgres-src] run the command:

```
gmake install
```

- Create the database:

Use the following commands (you may be in any system directory provided that you use the full path names of each directory):

```
mkdir [postgres-data]
chown postgres [postgres-data]
su - postgres
[postgres]/bin/initdb -D [postgres-data]
```

- Start the database service:

Use the following command, whilst logged in as the PostgreSQL user:

```
[postgres]/bin/postmaster -i -D [postgres-data]
```

Note that we use the -i option to enable TCP/IP connections – without this enabled, DSpace will be unable to use the database.

Installing the Source Code

The DSpace home page is available at:

`http://www.dspace.org/`

The source code for DSpace is available at:

`http://sourceforge.net/projects/dspace/`

Download the source file, which comes as a `.tar.gz` compressed file.

Decompress this file into `[dspace-src]`, and it will build all the relevant sub-directories beneath this.

You may now perform the DSpace installation, See the *DSpace Installation* section for more information.

DSpace Installation

- Create a `dspace` user and a group of the same name, containing that user.
- Log in as `dspace`.
- Create the `dspace` database:

Use the following commands in the directory `[postgres]/bin`:

```
createuser -U postgres -d -A -P dspace
createdb -U dspace dspace
```

- Customise the `dspace.cfg` file:

See the section *DSpace Configuration File*.

- Build the DSpace source:

In `[dspace-src]`, run the command:

```
ant
```

- Install the DSpace code:

In `[dspace-src]`, run the command:

```
ant fresh_install
```

- Connect DSpace webapps to Tomcat:

We use symbolic links within tomcat to point to the relevant DSpace directories:

Go to the tomcat directory, which will be something like:

```
/usr/local/jakarta-tomcat-4.0.6/
```

and go into the `webapps` directory beneath this. Here, make the relevant symbolic links using the commands:

```
ln -s [dspace]/jsp dspace
ln -s [dspace]/oai dspace-oai
```

- Install the config files:

Go to `[dspace]/bin` and run:

```
./install_configs
```

- Create the administrator account:

Ensure that PostgreSQL is running (see PostgreSQL installation guide for more information), then in `[dspace]/bin`, run:

```
./create_administrator
```

This will create the user details for the DSpace system administrator. Note that the password you enter will appear on the screen as you type it.

- Initialise Lucene search indices:

In `[dspace]/bin`, run:

```
./index_all
```

- Set up email subscription feature:

In the `dspace` user's `crontab` insert the following:

```
#Send out subscription emails at 01:00 every day
0 1 * * * [dspace]/bin/sub-daily
```

To create the `crontab`, edit a temporary file of your choice, containing the above information, then use the command:

```
crontab <your temporary file>
```

- Start Tomcat and Apache:

Start Tomcat first, followed by Apache - for best results wait around 5 seconds for Tomcat to initialise properly before starting Apache. The following commands are required:

```
[tomcat]/bin/startup.sh  
sudo [apache]/bin/apachectl start
```

Note that we use `sudo` as both Tomcat and Apache must run as `dspace`, but it is advisable that the main Apache process is owned by `root`.

DSpace Configuration File

Basic Settings

`dspace.dir` – This is the full path of the dspace installation [`dspace`].
For example, we use `/u01/dspace`. Note that we do not use a trailing slash (/).

`dspace.url` – The URL that will be used to access the DSpace system online. For example we use `banshee.lib.ed.ac.uk`

`dspace.hostname` – This should match `dspace.url`

`dspace.name` – String used to identify your DSpace installation on the website and some of the site emails (e.g., `Theses Alive!` at `Edinburgh University Library` – note that we do not use quotation marks).

Config Files

Following the above configuration options, there is the option to set all the template file paths – it is easiest to leave these all in, and ensure that the path is correct for the final destination of each one. It will be of the form:

```
[dspace]/config/<config file>
```

Database Settings

`db.url` – We use `jdbc:postgresql://localhost:5432/dspace`

`db.driver` – Class-path to the database driver java class. For PostgreSQL, this should normally be `org.postgres.Driver`

`db.username` – Username to access the DSpace database with (should be `dspace`). Note: this is different from the DSpace administrator that was created during the DSpace install.

`db.password` – password to access the DSpace database with. *Note that this is stored in plain text in the config file.*

Email Settings

`mail.server` – Your institution's mail server (eg `mailrelay.ed.ac.uk`).

`mail.from.address` – Email will come from this address specified here. Should probably be set up as the dspace mailing address. For example we use `dspace@srv4.lib.ed.ac.uk`

`feedback.recipient` – This email address will receive any mail sent using the feedback feature of DSpace.

`mail.admin` – Webmaster email account.

`alert.recipient` – This email account will receive notification of any Internal System Errors generated by DSpace, plus any other system alerts.

File Storage

`assetstore.dir` – Location that you would like your assets to be stored. Assets include all of the bitstreams for each item all bundled together in one file, as well as some other peripherals.

`history.dir` – Locations for history serialisations (*what exactly is this?*)

`search.dir` – Location for the search index files.

`log.dir` – location of DSpace log files.

`upload.temp.dir` – location uploaded files should be stored during submission procedure. This directory, should, for neatness, be the `/tmp` file for the server.

`upload.max` – Maximum size of uploaded files in bytes

Handle Settings

`handle.prefix` – This is the prefix given by the handle server to your institution. See “Installing the Handle Server” for more information. Note that this value must be set correctly before live usage of the DSpace system begins, as it is written into the database. It will be of the form `Prefix.Suffix` (eg `1721.1`).

`handle.dir` – Installation directory of the handle server. For example, we use `/u01/dspace/handle/handle-server`

Web UI Settings

`webui.site.authenticator` – must be `org.dspace.app.webui.SimpleAuthenticator`

`webui.cert.ca` – we use `/u01/dspace/etc/certificate-ca.pem`

`webui.cert.autoregister` – true if you want the system to set up eperson accounts automatically when necessary.

`webui.submit.blocktheses` – Prevent the UI from accepting submissions marked at “theses”. Takes `true/false`, and should be set to `false` for our purposes.

SFX Server

`sfx.server.url` – we have left this commented out.

Ingest Settings

`default.language` – the default language for the content of submissions. Set to `en` for UK English.

Making Changes to the Configuration

If changing any of the configuration options in `dspace.cfg` it is necessary to restart the web-service in order that they take effect, since Tomcat will cache these “application” level variables. To do this, run the following commands:

```
sudo [apache]/bin/apachectl stop
[ tomcat ]/bin/shutdown.sh
[ tomcat ]/bin/startup.sh
sudo [apache]/bin/apachectl start
```

It is recommended that you wait approximately 5 seconds between starting Tomcat and starting Apache, to give Tomcat sufficient time to initialise

properly. Note that we use `sudo` as we want to run these commands as the `dspace` user, but require `root` privileges to start Apache as a `root` owned process.

Basic Apache Configuration

To your standard Apache `httpd.conf` file you will need to ensure that the following code is included. The code is supplied here in the order that it will probably go into your current `httpd.conf` file, and it should be relatively easy to see similar lines grouped together to which the following can be appended:

```
LoadModule webapp_module libexec/mod_webapp.so

AddModule mod_webapp.c
```

After setting up these in the `httpd.conf`, you then need to ensure that Apache runs as the correct user, so beneath the lines which read:

```
# If you wish httpd to run as a different user or group,
# you must run httpd as root initially and it will
# switch.
```

include the lines:

```
User dspace
Group dspace
```

(Assuming that you created the user and group the same way that we have)

Set up your document root to serve DSpace:

```
DocumentRoot "[dspace]/jsp"
```

You also need to set up Apache to understand the mime-type:

```
AddType text/jsp .jsp
```

To deal with OAI requests, you must include the following lines, which can be found in the `httpd.conf` file that DSpace creates on install, and places in the directory you specified for `config.template.apache13.conf` in `dspace.cfg`, built for your own configuration. Ours look like this:

```
RedirectMatch ^/$ https://banshee.lib.ed.ac.uk/
RedirectMatch ^(/[^o].*) https://banshee.lib.ed.ac.uk$1
RedirectMatch ^(/.[^a].*) https://banshee.lib.ed.ac.uk$1
RedirectMatch ^(/..[^i].*) https://banshee.lib.ed.ac.uk$1
```

We include these lines directly rather than including the DSpace file in which they can be found.

Next we need to configure the behaviour of `mod_webapp`, which requires the following code:

```
<IfModule mod_webapp.c>
    WebAppConnection warpConnection warp localhost:8008
    WebAppInfo /webapp-info
    WebAppDeploy examples warpConnection /examples
    WebAppDeploy webdav warpConnection /webdav
    WebAppDeploy tomcat-docs warpConnection /tomcat-docs
    WebAppDeploy ROOT warpConnection /ROOT
    WebAppDeploy dspace warpConnection /dspace
</IfModule>
```

Note, at this point, that the DSpace documentation suggests including their copy of `httpd.conf` into your own in order to configure the application properly. If you follow the instructions in this section, there is no need to do this.

Setting the Java Environment

Many problems that you will encounter in the DSpace installation will be related in some way to the underlying Java support required. Mostly, you want to ensure that your `$PATH` environment variable is set with the correct Java and Tomcat paths in all the contexts that the `dspace` user will need them. This section will be addressed once we have more experience with the problems.

Configuring SSL with Apache

We have not yet configured SSL to run with DSpace. This causes a number of problems with workaround solutions, pending setting up the SSL. The primary problem is that when registering your email address for an account on DSpace, the email you receive will ask you to visit a URL which starts with `https://`. Simply remove the “s” to obtain the true URL, and you will be able to register.

Installing the Handle Server

We have not yet installed the handle server, but DSpace will function properly without it (you just can't resolve handles of the form `http://hdl.handle.net/xxxx.x`). Note that the handle that each item has is written into the database, so don't set up a genuine repository until you have the handle server installed and running, with your own number provided by the CNRI – otherwise you will need to manually edit the relevant database fields when you do have your own handle.