Independent University **Bangladesh (IUB)**

IUB Academic Repository

IUB Library Workshop

2014-05-01

DSpace Manual for BALID Trainee

Ahammad, Nur

IUB Library

https://ar.iub.edu.bd/handle/11348/419

Downloaded from IUB Academic Repository

Training on DSpace Institutional Repository
Organized by
BALID Institute of Information Management (BIIM)

1-2 May 2014
Venue: CIRDAP

- **Installation of DSpace on Debian**
- **Configuration of DSpace**
- **Customization of Dspace**
- **Cron Jobs setup for production system**
- **MTA Setup for DSpace**
- **Some Important Commands of PostgreSQL**
- **DSpace Discovery Setup**

Prepared By
Nur Ahammad
Junior Assistant Librarian
Independent University, Bangladesh

DSpace Installation on Debian

Add root password and create dspace as a user and enter password for dspace at the time of Debian installation. While installing Debian, select Desktop Environment, SQL database, SSH server & Standard System options.

First create a dspace user in debian for Dspace, if you did not add at the time of Debian installation.

root@dir:~#adduser dspace

Add source list

root@localhost:~#nano /etc/apt/sources.list

Comment or erase all and add the followings:

deb http://security.debian.org/ squeeze/updates main contrib
deb-src http://security.debian.org/ squeeze/updates main contrib non-free
deb http://http.us.debian.org/debian/ squeeze contrib non-free main
deb-src http://http.us.debian.org/debian/ squeeze contrib non-free main

Upgrade the Software.

```
root@localhost:~#sudo apt-get update
root@localhost:~#sudo apt-get dist-upgrade
```

Then install other softwares as prerequisites for Dspace installation. It is a good approach to install sun-java6-jdk.

root@dir:~# apt-get install sun-java6-jdk
root@dir:~# apt-get install tomcat6
root@dir:~#apt-get install maven2
root@dir:~#apt-get install postgresql-8.4

(In training we are skipping the above four packages installation. We are going to install the above five packages with one command from CD. In Training kits, we provide a CD which has all pre-requisites of Dspace installation in .deb file extension. You can install all .deb packages with one command.)

First in enter in the directory which contain .deb files

root@localhost:~#cd /media/cdrom/ dspacepackages/
root@localhost:/media/cdrom/dspacepackages# dpkg -i *.deb

After installing prerequisite softwares, create dspace database in postgresql for Dspace.

root@dir:~# cd /home/dspace/
root@dir:/home/dspace# su postgres
postgres@dir:/home/dspace$ createuser -U postgres -d -A -P dspace
Enter password for new role:
Enter it again:
Shall the new role be allowed to create more new roles? (y/n) n
postgres@dir:/home/dspace$ exit
exit

root@dir:/home/dspace# su dspace
dspace@dir:~$ createdb -U dspace -E UNICODE dspace
dspace@dir:~$exit

Set Up Environment Variables:
root@dir:/home/dspace#sh -c 'echo export PATH=$PATH:\"/usr/lib/jvm/java-6-sun/bin\" >> /etc/profile'
root@dir:/home/dspace# sh -c 'echo export JAVA_HOME=\"/usr/lib/jvm/java-6-sun\" >> /etc/profile'
root@dir:/home/dspace# sh -c 'echo export JRE_HOME=\"/usr/lib/jvm/java-6-sun/jre\" >> /etc/profile'

After that download the latest version of Dspace
root@dir:/home/dspace# wget
http://skylink.dl.sourceforge.net/project/dspace/DSpace%20Stable/4.1/dspace-4.1-src-release.tar.gz

Untar the downloaded dspace
root@dir:/home/dspace#tar –zxvf dspace-4.1-src-release.tar.gz

We have dspace in training CD. We are not going to download it
root@dir:/home/dspace#tar –zxvf /media/cdrom/dspace-4.1-src-release.tar.gz

Set right configuration
root@dspace:/home/dspace# cd dspace-4.1-src-release/
root@dir:dapace-src-release#nano dspace/config/dspace.cfg

Thereafter build Dspace by running the following command
root@dir:/home/dspace/dspace-src-release#cd dspace
root@dir:/home/dspace/dspace-src-release/dspace# mvn package

[INFO] Scanning for projects...
Downloading: http://repo1.maven.org/maven2/org/dspace/dspace-pom/10/dspace-pom-10.pom
[WARNING] Unable to get resource 'org.dspace:dspace-pom:pom:10' from repository central (http://repo1.maven.org/maven2): Error transferring file: repo1.maven.org
[INFO] --
[ERROR] FATAL ERROR
[INFO] --
[INFO] Error building POM (may not be this project's POM).
Project ID: org.dspace:dspace-parent:pom:1.8.2
Reason: Cannot find parent: org.dspace:dspace-pom for project: org.dspace:dspace for project org.dspace:dspace-parent:pom:1.8.2

[INFO] --
[INFO] Trace
org.apache.maven.reactor.MavenExecutionException: Cannot find parent: org.dspace

If you have problem in your Internet connectivity, you will get the above messages. These messages
also can get, if you have proxy in your Internet connectivity. To solve this problem, set proxy in you
maven configuration file.

root@dir:/home/dspace/dspace-src-release/dspace#nano /etc/maven2/settings.xml

Uncomment and set you proxy in the following section-

```
<proxies>
<proxy>
<id>optional</id>
<active>true</active>
<protocol>http</protocol>
<username></username>
<password></password>
<host>proxy.iub.edu.bd</host>
<port>3000</port>
<nonProxyHosts>localhost</nonProxyHosts>
</proxy>
</proxies>
```

root@dir:/home/dspace/dspace-src-release/dspace# mvn package
root@dir:/home/dspace/dspace-src-release/dspace# cd target/dspace-4.1-build/
root@dir:/home/dspace/dspace-src-release/dspace/target/dspace-4.1-buildr# ant fresh_install

To complete installation, you should do the following:
Setup your Web servlet container (e.g. Tomcat) to look for your DSpace web applications in:
/home/dspace/webapps/
OR
copy any web applications from /home/dspace/webapps/ to the appropriate place for your servlet container. (e.g. '$CATALINA_HOME/webapps' for Tomcat)

root@dir:/home/dspace/dspace-src-release/dspace#cp -r /dspace/webapps/*
/var/lib/tomcat6/webapps/

Create Dsapce administrator by using the command
root@dir:/home/dspace/dspace-src-release/dspace# /home/dspace/bin/dspace create-administrator

Assign permission to dspace installation directory
root@dir:/home/dspace/dspace-src-release/dspace#chmod -R 777 /dspace/*

Finally you can get your Dsapce by pointing the address in your web browser:
http://localhsot:8080/xmlui

If ant fresh_install fail to download GeoLiteCity do the following-

In order to use DSpace Solr Usage Statistics, you will need to
manually re-run:
ant update_geolite

OR
You may manually install this file by following these steps:
(1) Download the file from
http://geolite.maxmind.com/download/geoip/database/GeoLiteCity.dat.gz
(2) Unzip it to create a file named 'GeoLiteCity.dat'
(3) Copy that file to '/dspace/config/GeoLiteCity.dat'

If you cannot find your item after upload Items, do the following
root@dir:/home/dspace/dspace-src-release/dspace#/dspace/bin/dspace index-init

Configuration and Customization of Dspace

We have to change the following sections at the time of DSpace installation

The location of main Dspace configuration file is –
dapace-src-release/dspace/config/dspace.cfg

Basic information

# DSpace installation directory
dspace.dir = ${dspace.install.dir} ***Put your Dspace installation Directory***

# DSpace host name - should match base URL. Do not include port number.
dspace.hostname = ${dspace.hostname} ***Assign your hostname here***

# DSpace base host URL. Include port number etc.
dspace.baseUrl = ${dspace.baseUrl}

# DSpace base URL. Include port number etc., but NOT trailing slash
# Change to xmlui if you wish to use the xmlui as the default, or remove
# "/jspui" and set webapp of your choice as the "ROOT" webapp in
# the servlet engine.
dspace.url = ${dspace.baseUrl}/xmlui

# Name of the site
dspace.name = ${dspace.name} ***Write your Dspace name here***

# Default language for metadata values
default.language = ${default.language}

Database settings

# Database name ("oracle", or "postgres")
db.name = ${db.name}

# Database username and password
db.username = ${db.username}
db.password = ${db.password} ***Insert your database password here***

# From address for mail
mail.from.address = ${mail.from.address} ***Put your E-mail Address here***

# Currently limited to one recipient!
feedback.recipient = ${mail.feedback.recipient} ***Put your E-mail Address here***

General site administration (Webmaster) e-mail

mail.admin = ${mail.admin} ***Put your E-mail Address here***

# Recipient for server errors and alerts
alert.recipient = ${mail.alert.recipient} ***Put your E-mail Address here***

# Recipient for new user registration emails
registration.notify = ${mail.registration.notify} ***Put your E-mail Address here***

# Boolean search operator to use, current supported values are OR and AND
# If this config item is missing or commented out, OR is used
# AND requires all search terms to be present
# OR requires one or more search terms to be present
search.operator = OR ***Define your default search operator here***

Handle settings
#handle.canonical.prefix = ${dspace.url}/handle/ ***Uncomment it to use your own URL***
# If omitted, the canonical URL prefix will be http://hdl.handle.net/
handle.canonical.prefix = ${handle.canonical.prefix} ***Comment it to use your own URL***

# CNRI Handle prefix
handle.prefix = ${handle.prefix} ***Comment it to use your own URL***

Settings for item count (strength) information
webui.strengths.show = false ***Change this to true*** for enabling Item-counter
#
The default is to use a cache
webui.strengths.cache = true ***Uncomment this line***

# Set to true to use local server URLs (i.e. http://myserver.myorg/handle/123456789/1)
webui.feed.localresolve = false ***Set it to true to use your own URL***

Controlled Vocabulary Settings
Enable or disable the controlled vocabulary add-on
Warning: this feature is not compatible with WAI (it requires javascript to function)
#
webui.controlledvocabulary.enable = true ***Uncomment this line to enable controlled vocabulary***

Cron Jobs setup for production system

Add crob jobs with the following lines in crob jobs
root@dspace#crontab -e

```
@hourly        /dspace/bin/dspace itemcounter

@hourly        /dspace/bin/dspace stats-util -f
@hourly        /dspace/bin/dspace stats-util -i
@daily         /dspace/bin/dspace stats-util -u
@daily         /dspace/bin/dspace stats-util -m

@daily         /dspace/bin/dspace update-discovery-index -o
@weekly        /dspace/bin/dspace update-discovery-index -b -f
@monthly       /dspace/bin/dspace update-discovery-index -c

@weekly        /dspace/bin/dspace checker -u -d2h -p
@weekly        /dspace/bin/dspace checker-emailer -a
@monthly       /dspace/bin/dspace cleanup

10 0 * * * /dspace/bin/dspace embargo-lifter

0 1 * * * /dspace/bin/dspace sub-daily

10 1 * * * /dspace/bin/dspace stat-general
20 1 * * * /dspace/bin/dspace stat-monthly

10 2 * * * /dspace/bin/dspace stat-report-general
20 2 * * * /dspace/bin/dspace stat-report-monthly

@weekly    /dspace/bin/dspace filter-media
0 3 * * * /home/dspace/bin/dspace index-update
0 6 * * * /home/dspace/bin/dspace generate-sitemaps
```

Running DSpace on Standard Port 80 instead of 8080

Redirecting with IP tables

You can do this simply by port forwarding in your IP table

Use Apache HTTPD (mod_proxy_ajp) + Tomcat (port 8009)

This assumes you have the following working:

1. httpd is running and listening on port 80
2. tomcat is running and listening on port 8009,8080

By convention, web-servers listen on port 80 to deliver content such as static html files. So that web browsers can

```
## use the more familiar url
 http://www.dspace-instance.org

 ## instead of....
 http://www.dspace-instance.org:8080/jspui
 http://www.dspace-instance.org:8080/xmlui
```

- Set Tomcat to serve up DSpace by default

```
cd /var/lib/tomcat6/webapps
mv ROOT ROOT_hold
ln -s /dspace/webapps/jspui ROOT
## for the Manakin interface replace jspui with xmlui
```

* Configure /etc/httpd/conf.d/ssl.conf or proxy_ajp.conf

```
# Put in VirtualHost element
#
ProxyPass  /do_not_touch  !
ProxyPass  /  ajp://localhost:8009/
ProxyPassReverse  /  ajp://localhost:8009/
```

Note:

1. You'll need to reload or restart the httpd service
2. The "!" <bang> sets Apache web-server to NOT REDIRECT everything under /do_not_touch

Exim4 MTA Setup

Install exim4

```
# sudo aptitude install exim4
# dpkg-reconfigure exim4-config
```

- Choose the option for "internet site" and select "Ok" to continue.
- Enter your system's FQDN (fully qualified domain name) on the "mail name" configuration screen.
- Enter "127.0.0.1" when asked which IP address to listen on for SMTP connections
- Make sure you list your FQDN, hostname, and localhost entries when you're asked which destinations mail should be accepted for.
- Leave the relay domains and relay machines fields blank.
- Select "No" when asked whether to keep DNS queries to a minimum.
- Select your mail directory
- "non-split" option for your mail configuration file: No

Test your mail configuration

```
echo "This is a test." | mail -s Testing nurahammad@hotmail.com
```

Some Important Commands of PostgreSQL

To show list of databases
postgresql: \l
postgresql: SELECT datname FROM pg_database;

To connect/use a particular database:
mysql: use database_name;
Postgresql: \c database_name

To show list of tables in your database:
mysql: SHOW TABLES
postgresql: \d
postgresql: SELECT table_name FROM information_schema.tables WHERE table_schema = 'public';

To show list of columns in a particular table / schema use :
mysql: SHOW COLUMNS
postgresql: \d table_name
postgresql: SELECT column_name FROM information_schema.columns WHERE table_name ='table';

To create a database
createdb -U dspace -E UNICODE dspace

To drop a database:
dropdb dspace

Backup database:
Pg_dump –E UNICODE dspace>/home/dspace/dspace.sql

Create user:
createuser -U postgres -d -A -P dspace

DSpace Discovery

The Discovery Module enables faceted searching & browsing for your repository. Although these techniques are new in DSpace, they might feel familiar from other platforms like Aquabrowser or Amazon, where facets help you to select the right product according to facets like price and brand. DSpace Discovery offers very powerful browse and search configurations that were only possible with code customization in the past.

Enabling Discovery

Because Discovery was adopted as the default infrastructure for search and browse in DSpace 4, no manual steps are required to enable Discovery. If you want to enable Discovery on older versions of DSpace, please refer to the DSpace documentation for that particular version.

Configuration files

The configuration for discovery is located in 2 separate files.

- General settings: The `discovery.cfg` file located in the `[dspace-install-dir]/config/modules directory`.
- User Interface Configuration: The `discovery.xml` file is located in `[dspace-install-dir]/config/spring/api/` directory.

"Did you mean" spellcheck aid for search configuration

DSpace 4 introduces the use of SOLR's SpellCheckComponent as an aid for search. When a user's search does not return any hits, the user is presented with a suggestion for an alternative search query.



The feature currently only one line of configuration to discovery.xml. Changing the value from true to false will disable the feature.

```
<property name="spellCheckEnabled" value="true" />
```

Routine Discovery Solr Index Maintenance

It is strongly recommended to run maintenance on the Discovery Solr index daily (from crontab or your system's scheduler), to prevent your servlet container from running out of memory:

```
[dspace]/bin/dspace index-discovery -o
```