
Guía de Instalación DSpace

García Rojas Alan
Liévana Poy Erick
Lima Estrada Efraín
Reyes Reyes Julián
BUAP Puebla - México



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Facultad de Ciencias de la Computación

Administración de Redes

Guía de Instalación DSpace

García Rojas Alan

Liévana Poy Erick

Lima Estrada Efraín

Reyes Reyes Julián

Profesor Larios Gomez Mariano

Índice general

Presentación	I
1. Resumen	1
2. Dependencias del Backend	2
2.1. Sistema Operativo	2
2.2. Java Development Kit	2
2.3. Apache Maven	2
2.4. Apache Ant	2
2.5. Base de Datos	3
2.5.1. PostgreSQL	3
2.5.2. Oracle DataBase	3
2.6. Apache Solr	3
2.7. Apache Tomcat	3
3. Instalación de Backend	4
3.1. Preinstalación	4
3.2. Descarga de DSpace	4
3.3. Base de Datos	5
3.4. Configuración Inicial	5
3.5. Directorio de Instalación	6
3.6. Construcción del paquete	7
3.7. Instalación del Backend de DSpace	7
3.8. Inicialización de la Base de Datos	7
3.9. Implementar la aplicación web del servidor	7
3.10. Nucleos de Solr	8
3.11. Crear una cuenta de Administrador	8
3.12. Primer inicio	8
3.13. Agregando soporte https	9
3.14. Dependencias del Frontend	9
3.15. Node.js	9
3.16. Yarn	10
3.17. PM2	10
4. Instalación del Frontend	11
4.1. Preinstalación	11
4.2. Creando archivo de Configuración	11
4.3. Compilar la Interfaz de Usuario	12
4.4. Configuración de PM2	12
4.5. Configuraciones Adicionales	13

Capítulo 1

Resumen

- La presente tiene como finalidad servir como un manual de instalación para el software de repositorios abiertos DSpace.
- Esta guía es funcional para la versión 7.0 de DSpace.
- A partir de la versión 7 (y superior), la aplicación DSpace se divide en un "frontend"(interfaz de usuario) y un "backend"(API de servidor). La mayoría de las instituciones querrán instalar AMBOS. Sin embargo, puede decidir si ejecutarlos en la misma máquina o en máquinas separadas.
- Recomendamos instalar primero el Backend, ya que el Frontend requiere un Backend válido para ejecutarse correctamente.

Capítulo 2

Dependencias del Backend

2.1. Sistema Operativo

- DSpace puede ser instalado y ejecutado en sistemas operativos basados en UNIX(Linux/Mac OSX) o basados en Microsoft Windows.
- Muchas distribuciones de Linux / Unix vienen con algunas de las dependencias a continuación pre-instaladas o instaladas fácilmente mediante actualizaciones. Debe consultar la documentación de su distribución particular o los administradores del sistema local para determinar lo que ya está disponible.

2.2. Java Development Kit

- Java es un lenguaje de programación desarrollado originalmente por Sun Microsystems lanzado en 1995 como un componente central de Sun Microsystems' Java Platform.
- Las instrucciones de descarga e instalación de OpenJDK se pueden encontrar aquí <http://openjdk.java.net/install/>
- Es importante asegurarse de instalar el paquete de desarrollo completo(JDK).

2.3. Apache Maven

- Apache Maven es una herramienta de comprensión y gestión de proyectos de software. Basado en el concepto de un modelo de objetos de proyecto (POM), Maven puede administrar la construcción, los informes y la documentación de un proyecto a partir de una pieza central de información.
- Maven es necesario en la primera etapa del proceso de compilación para ensamblar el paquete de instalación para su instancia de DSpace.
- Apache Maven puede ser descargado desde la siguiente liga: <http://maven.apache.org/download.html>

2.4. Apache Ant

- Apache Ant es una biblioteca de Java y una herramienta de línea de comandos cuya misión es impulsar los procesos descritos en los archivos de compilación como objetivos y puntos de extensión que dependen unos de otros.

- Se requiere Apache Ant para la segunda etapa del proceso de compilación. Primero, Maven se usa para construir el instalador, después de lo cual Ant se usa para instalar / implementar DSpace en el directorio de instalación.
- Apache Ant puede ser descargado desde la siguiente liga: <http://ant.apache.org>

2.5. Base de Datos

- DSpace al ser un repositorio, requiere de una base de datos para poder manejar la información. DSpace puede funcionar con:
 - PostgreSQL
 - Oracle DataBase

2.5.1. PostgreSQL

- PostgreSQL es un potente sistema de base de datos relacional de objetos de código abierto con más de 30 años de desarrollo activo que le ha ganado una sólida reputación por su fiabilidad, solidez de funciones y rendimiento.
- Este manual instalara DSpace con PostgreSQL.
- PostgreSQL puede ser descargado de la siguiente liga: <http://www.postgresql.org/>

2.5.2. Oracle DataBase

- Oracle Database es un sistema de gestión de base de datos de tipo objeto-relacional desarrollado por Oracle Corporation.
- Tenga en cuenta que todo el desarrollo activo de DSpace se produce en PostgreSQL en este momento.
- Oracle DataBase puede ser descargado de la siguiente liga: <http://www.oracle.com/database/>

2.6. Apache Solr

- Solr es un motor de búsqueda de código abierto basado en la biblioteca Java del proyecto Lucene, con API en XML/HTTP y JSON, resultado de resultados, búsqueda por facetas, caché, y una interfaz para su administración.
- Es posible que desee leer partes del manual de inicio rápido para familiarizarse con el diseño y el funcionamiento de Solr.
- Apache Solr puede ser descargado en la siguiente liga: <https://solr.apache.org/>

2.7. Apache Tomcat

- Apache Tomcat es una tecnología de plataforma Java para extender servlets. Todos los motores de servlet tienen una API común. El componente Conector de servlet de ArcIMS requiere un motor de servlet si su servidor web no admite servlets de forma nativa.
- Los servlets de Java son programas que se ejecutan en un servidor web o de aplicaciones y actúan como una capa intermedia entre las solicitudes que provienen de un navegador web u otro cliente HTTP y las bases de datos o aplicaciones en el servidor HTTP.
- Apache Tomcat puede ser descargado desde la siguiente liga: <http://tomcat.apache.org>

Capítulo 3

Instalación de Backend

3.1. Preinstalación

- Se deben instalar todas las dependencias del BackEnd antes mencionadas.
- Este manual para la instalación con la base de datos PostgreSQL.
- Este manual sera enfocado a una instalación realizada a un Sistema Operativo basado en UNIX, por lo cual tendrá comandos propios a este tipo de sistemas.
- Algunas de las dependencias como Apache Tomcat, deben ejecutarse con un usuario que tengas permisos de lectura y escritura en el directorio root de DSpace.
- Para lograr eso debe asegurarse de que el propietario de Tomcat también sea propietario del directorio root de DSpace, ó puede crear una nueva cuenta de usuario ”*dspaceUsuario*” asegurarse de que Tomcat también se ejecute desde esa cuenta.

Crear un usuario

```
useradd -m dspaceUsuario
```

3.2. Descarga de DSpace

- Descargar la ultima versión de DSpace. Se puede descargar como .zip o como .tar.gz.
- Se puede descargar desde el repositorio en Github de Dspace que se encuentra en la siguiente liga:
<https://github.com/DSpace/DSpace/releases>
- Para poder descomprimir DSpace se pueden usar los comandos:

Archivo zip

```
unzip dspace-7.0.zip
```

Archivo tar.gz

```
gunzip -c dspace-7.0.tar.gz | tar -xf -
```

- Para facilitar la referencia, nos referiremos a la ubicación de esta versión descomprimida de la versión de DSpace como ”*dspace-source*.” en el resto de estas instrucciones. Después de descomprimir el archivo, es posible que el usuario desee cambiar la propiedad de la carpeta dspace-7.x al usuario ”*dspace*”.

Cambio de propietario

```
chown -c -R dspace-source
```

3.3. Base de Datos

- Como mencionamos este manual hará una instalación de DSpace con PostgreSQL.
- Como requerimiento para la base de datos es necesario instalar y activar la extensión *pgcrypto*.
- La extensión pgcrypto permite a DSpace crear UUID (identificadores universalmente únicos) para todos los objetos en DSpace, lo que significa que los identificadores de objetos (internos) ahora son únicos a nivel mundial y ya no están vinculados a las secuencias de la base de datos.
- En la mayoría de los sistemas operativos Linux, esta extensión se proporciona en el paquete ”postgresql-contrib.” en su administrador de paquetes. Por lo tanto, asegúrese de haber instalado ”postgresql-contrib”
- Cree un usuario en la base de datos de PostgreSQL para DSpace (este usuario puede tener cualquier nombre, pero asumiremos que lo llama ”*Usuario*”). Esto es completamente independiente del usuario del sistema operativo *dspace* creado anteriormente.

Crear usuario

```
createuser --username=postgres --no-superuser --pwprompt Usuario
```

- Se le pedirá (dos veces) una contraseña para el nuevo usuario de *dspace*. Luego se le pedirá la contraseña del superusuario de PostgreSQL (*postgres*). Este ultimo fue creado durante la instalación de PostgreSQL.
- Cree una base de datos de *dspace*, propiedad del usuario de *dspace* PostgreSQL. Al igual que en el paso anterior, esto solo puede hacerlo una cuenta de ”superusuario.” en PostgreSQL.
- El dueño de la base de datos es el usuario que creamos en el paso anterior. Y la base de datos se llamará *dspace*

Crear usuario

```
createdb --username=postgres --owner=Usuario --encoding=UNICODE dspace
```

- Finalmente se debe habilitar en la nueva base de datos de DSpace la extensión *pgcrypto*. Esto solo puede ser realizado por un superusuario.

Activar extensión pgcrypto

```
psql --username=postgres dspace -c "CREATE EXTENSION pgcrypto;"
```

3.4. Configuración Inicial

- DSpace usará la configuración localizada en ”*dspace-source/dspace/config/local.cfg*”. El archivo *local.cfg* puede ser usado para guardar cualquier configuración que se desee hacer a la instalación local.
- DSpace provee un archivo de configuración de ejemplo ubicado en: ”*dspace-source/dspace/config/local.cfg.EXAMPLE*”

- CUALQUIER configuración se puede copiar en este archivo local.cfg desde dspace.cfg o cualquier otro archivo * .cfg para anular la configuración predeterminada. Para la instalación inicial de DSpace, hay algunas configuraciones clave que probablemente querrá anular.
- dspace.dir - Es el directorio donde DSpace es instalado.
- dspace.server.url - URL completa de este backend de DSpace incluido el puerto y cualquier subruta.
- dspace.name - nombre .^adecuado” legible por humanos de su servidor, p. ej. ”Mi Biblioteca Digital”.
- solr.server - URL completa del servidor Solr. DSpace utiliza Solr con fines de indexación.
- default.language - idioma predeterminado para todos los valores de metadatos (el valor predeterminado es .en_US”)
- db.url - la URL completa de JDBC a su base de datos (se proporcionan ejemplos en local.cfg.- EXAMPLE)
- db.driver - qué controlador de base de datos usar, en función de si está utilizando PostgreSQL u Oracle
- db.dialect - qué dialecto de la base de datos usar, en función de si está usando PostgreSQL u Oracle
- db.username - el nombre de usuario de la base de datos utilizado en el paso anterior.
- db.password - la contraseña de la base de datos utilizada en el paso anterior.
- db.schema - el esquema de la base de datos que se utilizará (se proporcionan ejemplos en local.cfg.- EXAMPLE)
- mail.server - nombre de dominio completo de su servidor de correo saliente.
- mail.from.address - la dirección ”De:” para poner en el correo electrónico enviado por DSpace.
- mail.feedback.recipient - buzón de correo para comentarios.
- mail.admin - buzón para el administrador del sitio de DSpace.
- alert.recipient - buzón de correo para errores / alertas del servidor
- registration.notify- buzón de correo electrónico cuando se registran nuevos usuarios

3.5. Directorio de Instalación

- Cree el directorio para la instalación del backend de DSpace (es decir, ”*dspace*”). Como root (o un usuario con los permisos adecuados), ejecute:

Creación del directorio

```
mkdir dspace
```

- Asigne al usuario del sistema *dspace* o el que haya creado al inicio de la instalación la propiedad del directorio:

Cambiar propietario

```
chown dspaceUsuario dspace
```

3.6. Construcción del paquete

- Como el usuario del sistema operativo creado para dspace (*dspac.Usuario*) genere el paquete de instalación de DSpace con el siguiente comando:

Generación del paquete de instalación

```
cd dspace-source  
mvn package
```

3.7. Instalación del Backend de DSpace

- Como el usuario del sistema operativo creado para dspace (*dspac.Usuario*) instale DSpace en el directorio elegido para la instalación de DSpace, en este caso *dspace*.

Instalación

```
cd dspace-source/dspace/target/dspace-installer  
ant fresh\_install
```

- Para ver distintas opciones para la instalación se puede ejecutar:

Opciones de instalación

```
ant help
```

3.8. Inicialización de la Base de Datos

- Si bien este paso es opcional (ya que la base de datos de DSpace debe autoinicializarse en el primer inicio), siempre es bueno verificar por última vez que la conexión de la base de datos funciona correctamente. Para inicializar la ejecución de la base de datos:

Iniciar base de datos

```
dspace/bin/dspace database migrate
```

3.9. Implementar la aplicación web del servidor

- El backend de DSpace consta de una única aplicación web de "servidor" (en *dspace/webapps/server*). Debe implementar esta aplicación web en su contenedor de servlets para esta instalación se usara Tomcat.
- Existen 2 formas de realizar esta acción:

- Dígale a su instalación de Tomcat dónde encontrar su aplicación web DSpace. Como ejemplo, en el directorio *tomcat/conf/Catalina/localhost* puede agregar archivos similares a los siguientes (pero reemplace *dspace* con su ubicación de instalación):

Definir una ruta de contexto para webapp de DSpace: server.xml

```
$<$?xml version='1.0'?$$>$\\\n$<$Context\\\n docBase="dspace/webapps/server"/$$>$\\
```

- 2. Sencillo y completo. Copia solo (o todas) las aplicaciones web de DSpace que desea usar desde el directorio *dspace/webapps* al directorio apropiado en su instalación de Tomcat.

Copiar todas las aplicaciones web a Tomcat

```
cp -R dspace/webapps/* tomcat/webapps
```

Copiar solo la aplicación web del servidor a Tomcat

```
cp -R dspace/webapps/server tomcat/webapps*
```

3.10. Nucleos de Solr

- La instalación de DSpace crea un conjunto de cuatro núcleos Solr vacíos ya configurados. Para su uso es necesario copiarlos del directorio *"dspace/solr"* a un directorio donde la instancia de Solr pueda detectarlos esto depende del directorio de Solr por el momento nos referiremos a este directorio como *"solr"* ademas reiniciar el servicio de Solr.

Copiar archivos de Solr

```
cp -R dspace/solr/* solr/server/solr/configsets
```

```
chown -R solr:solr solr/server/solr/configsets
```

Iniciar o reiniciar Solr

```
solr/bin/solr restart
```

- Puede verificar el estado de Solr y sus nuevos núcleos DSpace utilizando su interfaz web administrativa. Vaya a <http://localhost:8983/solr/> para ver si Solr está funcionando bien, luego mire los núcleos seleccionando (a la izquierda) Core Admin o usando la lista despegable Core Selector.

3.11. Crear una cuenta de Administrador

- Se puede crear una cuenta de administrador para DSpace con el siguiente comando:

Crear Cuenta Administrador

```
dspace/bin/dspace create-administrator
```

3.12. Primer inicio

- Reinicie Tomcat para cargar de nuevo las configuraciones:

Reinicio de Tomcat

```
tomcat/bin/tomcat restart
```

- Ingrese a las siguientes urls:

- REST API: <http://dspace.myu.edu:8080/server/>
- OAI-PMH: <http://dspace.myu.edu:8080/server/oai/request?verb=Identify>

- Para saber si su instalación funciona correctamente puede entrar a la siguiente liga para compararla: <https://api7.dspace.org/server/>

3.13. Agregando soporte https

- La ejecución del DSpace Backend en HTTP y el puerto 8080 solo se puede usar para entornos de desarrollo local, donde se ejecuta la interfaz de usuario y la API REST desde la misma máquina, y solo se accede a ellos a través de las URL del host local.
- Para la compatibilidad con HTTPS, recomendamos instalar Apache HTTPD o Nginx, configurar SSL en ese nivel y enviar todas las solicitudes a su instalación de Tomcat. Tenga en cuenta que si desea alojar tanto el DSpace Backend como el Frontend en el mismo servidor, puede usar una instalación de Apache HTTPD o Nginx para administrar HTTPS / SSL y el proxy para ambos.
- En esta guía se usara Apache HTTPD.
- Descargue e instale Apache HTTPD desde la siguiente liga: <https://httpd.apache.org/>
- Instale los módulos *mod_proxy* y *mod_proxy_ajp* con el siguiente comando:

Instalación de los módulos

```
sudo en2mod proxy  
sudo a2enmod proxy\_ajp
```

- Reinicie Apache
- Para que mod_proxy_ajp se comunique con Tomcat, deberá habilitar el conector AJP de Tomcat en el server.xml de su Tomcat.

Habilitar Tomcat AJP

```
$<$Connector protocol="AJP/1.3" port="8009" redirectPort="8443" URIEncoding="UTF  
→ -8"/$>$
```

- Ahora, configure un nuevo VirtualHost para su sitio (usando HTTPS / puerto 443) que envía todas las solicitudes al conector AJP de Tomcat (que se ejecuta en el puerto 8009)

Configurando el virtual host

```
$<$VirtualHost \_default\_:_443>$\\  
SSLEngine on\\  
SSLCertificateFile [full-path-to-PEM-cert]\\  
SSLCertificateKeyFile [full-path-to-cert-KEY]\\  
ProxyPass /server ajp://localhost:8009/server\\  
ProxyPassReverse /server ajp://localhost:8009/server\\  
$<$/VirtualHost$>$
```

- Después de cambiar a HTTPS, asegúrese de volver atrás y actualizar las URL (por ejemplo, dspace.server.url) en su local.cfg para que coincida con la nueva URL de su backend. Esto requerirá reiniciar Tomcat brevemente.

3.14. Dependencias del Frontend

3.15. Node.js

- Node.js es un entorno en tiempo de ejecución multiplataforma, de código abierto, para la capa del servidor basado en el lenguaje de programación JavaScript, asíncrono, con E/S de datos en una arquitectura orientada a eventos y basado en el motor V8 de Google.

- Puede descargar Node.js así como sus instrucciones de instalación en la siguiente liga: <https://nodejs.org/en/>

3.16. Yarn

- Yarn es un administrador de paquetes para su código. Permite usar y compartir (por ejemplo, JavaScript) código con otros desarrolladores de todo el mundo. Yarn hace esto de forma rápida, segura y confiable.
- DSpace funciona solo con la versión 1.x de Yarn.
- Puede instalar Yarn, después de instalar Node.js con el comando:

Instalar Yarn

```
npm install --global yarn
```

3.17. PM2

- En escenarios de producción, recomendamos encarecidamente iniciar/detener la interfaz de usuario con un administrador de procesos Node.js. En esta guía recomendamos usar *PM2*.
- Al igual que Yarn puede instalar PM2 después de instalar Node.js con el siguiente comando:

Instalación PM2

```
npm install --global pm2
```

Capítulo 4

Instalación del Frontend

4.1. Preinstalación

- Primero, instale todos los requisitos de frontend enumerados anteriormente y verifique que la API de backend/REST sea de acceso público.
- Descargue la última versión de dspace-angular del repositorio de GitHub de DSpace. Puede elegir descargar el archivo zip o tar.gz proporcionado por GitHub, o puede usar "git" para verificar la etiqueta apropiada (por ejemplo, dspace-7.0) o rama.
- Instale todas las dependencias locales necesarias ejecutando lo siguiente desde el directorio "dspace-angular" descomprimido

Instalando Dependencias

```
yarn install
```

4.2. Creando archivo de Configuración

- Cree un archivo de configuración de producción en *dspace-angular/src/environments/environment.prod.ts*.
- Es posible que desee utilizar *environment.template.ts* como punto de partida. Este archivo *environment.prod.ts* se puede utilizar para anular cualquiera de las configuraciones predeterminadas especificadas en *environment.common.ts*
- Como mínimo, este archivo DEBE incluir las secciones *ui* y *rest* similares a las siguientes:
- La sección UI define dónde desea que Node.js se ejecute / responda. A menudo es una URL de host local, especialmente si está utilizando un proxy.

Configuración de UI

```
ui: \{\\\
  ssl: false,\\
  host: 'localhost',\\
  port: 4000,\\
  nameSpace: '$/$'\\
\},
```

- La configuración de REST DEBE corresponder a la URL principal del backend. Por lo general, esto significa que deben mantenerse sincronizados con el valor de "dspace.server.url." en el archivo local.cfg del backend.

Configuración de REST

```
rest: \{\\
  ssl: true,\\
  host: 'api.mydspace.edu',\
  port: 443,\\
  nameSpace: '$/$server'\
\}
```

4.3. Compilar la Interfaz de Usuario

- Cree la interfaz de usuario para la producción. Esto usa su environment.prod.ts y el código fuente para crear una versión compilada de la interfaz de usuario en la carpeta dspace-angular/dist
- Puede hacerlo con el siguiente comando:

Compilación de la Interfaz

```
yarn run build:prod
```

- Si cambia ó actualiza su *environment.prod.ts*, deberá reconstruir la aplicación de la interfaz de usuario, vuelva a ejecutar este comando.

4.4. Configuración de PM2

- Cree un archivo de configuración JSON que describa cómo ejecutar nuestra aplicación de interfaz de usuario.
- No es necesario que esté en el mismo directorio que la base de código dspace-angular.

Archivo de configuración PM2: dspace-angular.json

```
\{ "apps": [ \{\\
  "name": "dspace-angular",\
  "cwd": "/home/dspace/dspace-angular",\
  "script": "yarn",\
  "args": "run serve:ssr",\
  "interpreter": "none"\
\} ] \}
```

- Ahora, inicie la aplicación usando PM2 usando el archivo de configuración que creó en el paso anterior.
- Suponemos que el archivo de configuración es llamado dspace-angular.json

Iniciando la interfaz

```
pm2 start dspace-angular.json
```

- En este punto, la interfaz de usuario debería estar disponible en la URL que configuró en su archivo *environment.prod.ts*.
- Puede compararla con la interfaz de ejemplo que puede encontrar en la siguiente liga: <https://demo7.dspace.org/>

4.5. Configuraciones Adicionales

- Las configuraciones adicionales de la interfaz de usuario se describen en *environment.common.ts*.
- De igual forma puede encontrar mas opciones de configuración y mayor información sobre el funcionamiento de DSpace en la siguiente liga: <https://github.com/DSpace/dspace-angular/blob/main/docs/Configuration.md>

Capítulo 5

Referencias

- [1] Página DSpace: link
- [2] Especificaciones DSpace: link
- [3] Diagrama DSpace: link
- [4] Smith M., Barton M., Bass M., Branschofsky M., McClellan G., Stuve D., Tansley R. & Harford J.. (2003). DSpace An Open Source Dynamic Digital Repository. Septiembre 2021, de D-Lib Magazine Sitio web: link
- [5] Guía de Instalación de DSpace: link
- [6] Opciones extra de configuración: link
- [7] Repositorio de DSpace-Angular: link
- [8] Repositorio de DSpace: link
- [9] OpenJDK: link
- [10] Apache Maven: link
- [11] PostgreSQL: link
- [12] Apache Solr: link
- [13] Apache Tomcat: link
- [14] Node.js: link
- [15] Yarn: link
- [16] PM2: link