Installation and testing of selected Digital Asset Management Systems (work area E)

Mark Philips
Serhiy Polyakov

April 3, 2007
Draft 3

Table of Contents

1. Introduction

This document contains installation procedures for the DAMS identified as the most suitable candidates for the deployment by the THDI partners (see document *Comparison of Digital Asset Management Systems).* Procedures include prerequisites, installation steps, problems and solutions. Also this document contains final recommendations about implementation of the DAMS.

2. DSpace

2.1. Overview

DSpace is an open source software package which provides the tools for management of digital assets, and is commonly used as the basis for an institutional repository. It is also intended as a platform for Digital preservation activities.

2.2. Prerequisites

The list below describes the third-party components and tools needed to run a DSpace server.
1. UNIX-like OS (Linux, HP/UX etc)
2. Java 1.4 or later (standard SDK is fine, you don't need J2EE)
3. Apache Ant 1.5 or later (Java make-like tool)
4. PostgreSQL 7.3 or later, an open source relational database, or Oracle 9 or higher.

2.3. PostgreSQL installation

PostgreSQL might be already installed, check about that.
Note: Fedora Core 4 distributive includes PostgreSQL, so just start postgresql service.

Related directories and files in Fedora Core 4 installation:
```
/var/lib/pgsql/
/usr/lib/pgsql/
/usr/share/pgsql/
/usr/include/pgsql/
/etc/sysconfig/pgsql/
/usr/bin/postgres
```

If PostgreSQL is not installed, follow the following steps:
```
gunzip postgresql-version.tar.gz
tar xvf postgresql-version.tar
./configure --prefix =      // default path is /usr/local/pgsql
gmake
su
gmake install
```

Creating the PostgreSQL User
It is recommended to create a separate user to own the PostgreSQL files and processes that will be installed. The user name is typically postgres. By default, POSTGRESQL allows database access only to users logged into the computer running the database server.
```
adduser postgres
```

Check if directory `/var/lib/pgsql/data` exists.

If owner of the directory is not postgres, change it:
```
chown postgres /var/lib/pgsql/data
```

Starting the server and Initialization
Initialization creates a database called template1 in the PostgreSQL home directory. This database is used to create all other databases.

This step typically involves running the program called postmaster.

```
su - postgres

/usr/bin/initdb -D /var/lib/pgsql/data
/usr/bin/postmaster -D /var/lib/pgsql/data >logfile 2>&1 &
```

Creating the database:
```
su - postgres
/usr/bin/createdb test
```

Test if client runs:
```
/usr/bin/psql test
```

Tuning for DSpace
Once installed, TCP/IP connections need to be enabled (DSpace uses JDBC).
For PostgreSQL 8.0+, in postgresql.conf uncomment the line starting:
```
listen_addresses = 'localhost'
```

Then tighten up security a bit by editing pg_hba.conf and adding this line:
```
host  dspace  dspace  127.0.0.1  255.255.255.255  md5
```

Then restart PostgreSQL.

2.4. Tomcat
Installing Tomcat

Tomcat service has been installed on test server on port 8080:

Stop tomcat5 (came with Fedora)

Before installing tomcat55 needed modules (dependencies) have been installed with yum:
jakarta-commons-daemon
xerces-j2
struts

Download 11 tomcat55 rpms from
http://www.jpackage.org/browser/rpm.php?jppversion=1.6&id=3553

rpm -ihh tomcat55*

Start
service tomcat55 start
OK

Test that Tomcat is working by typing http://localcost:8080 into the browser. Tomcat welcome page should come.

The Tomcat webbapps directory is:

/var/lib/tomcat55/webapps

There is no need to install the connector between Apache web Server and Tomcat because Tomcat can be functioning in standalone mode. Port number (8080) should be specified in the URL in this case.

Tuning Tomcat for DSpace

Note that DSpace will need to run as the same user as Tomcat (not really!)

Ensure that Tomcat has
   a) enough memory to run DSpace
   b) uses UTF-8 as its default file encoding for international character support. So ensure in startup scripts (etc/profile/) that the following environment variable is set:

```
JAVA_OPTS="-Xmx512M -Xms64M -Dfile.encoding=UTF-8"
run
export JAVA_OPTS
```

Alter Tomcat's default configuration to support searching and browsing of multi-byte UTF-8 correctly. Add a configuration option to the <Connector> element in etc/tomcat55/server.xml:

```
URIEncoding="UTF-8"
```

Default Tomcat config should read:

```
<!-- Define a non-SSL HTTP/1.1 Connector on port 8080 -->
<Connector port="8080"
           maxThreads="150" minSpareThreads="25" maxSpareThreads="75"
           enableLookups="false" redirectPort="8443" acceptCount="100"
           connectionTimeout="20000" disableUploadTimeout="true"
           URIEncoding="UTF-8" />
```


2.5. Installing DSpace
Directories:
   - the source directory, referred to as [dspace-source]
        in this test installation = /usr/local/src/dspace-1.4.1-souce
   - the install directory, referred to as [dspace]
        in this test installation = /dspace
   - the web deployment directory. If you're using Tomcat, this will be [tomcat]/webapps/dspace (with [tomcat] being wherever you installed Tomcat--also known as $CATALINA_HOME). This directory is generated by the web server when it unpacks dspace.war, and should never be edited.

1. Create the DSpace user. This needs to be the same user that Tomcat will run as. e.g. as root run:

```
useradd -m dspace
```

2. Download the latest DSpace source code release and unpack it:
```
gunzip -c dspace-source-1.x.tar.gz | tar -xf -
```
(to /usr/local/src/)

3. Copy the PostgreSQL JDBC driver (.jar file) into [dspace-source]/lib. It can be downloaded directly from the PostgreSQL JDBC site. Driver should be for the installed version of PostgreSQL (8.0) and for JDBC2.

Alternatively copy
```
/usr/share/java/postgresql-8.0-311.jdbc2.jar
```
to
```
[dspace-source]/lib
```

4. Create a dspace database, owned by the dspace PostgreSQL user:

```
createuser -U postgres -d -A -P dspace ; createdb -U dspace -E UNICODE dspace
```

Enter a password for the DSpace database [dsp]. (This isn't the same as the dspace user's UNIX password.)

Note: to drop a database from linux command line (under dspace user):
dropdb [dbname]

5. Edit [dspace-source]/config/dspace.cfg, setting the following properties:
```
dspace.url
dspace.hostname
dspace.name
db.password    (the password entered in the previous step)
mail.server
mail.from.address
feedback.recipient
mail.admin
```

6. Create the directory for the DSpace installation (/dspace).
As root, chance to directory usr and run:
```
mkdir [dspace] ; chown dspace [dspace]
```

(Assuming the dspace UNIX username.)

7. As the dspace UNIX user, compile and install DSpace:

cd [dspace-source] ; ant fresh_install

Note: this operation writes files into `/dspace` and `/usr/local/src/dspace-1.4.1-souce/build` directories. Delete them before trying again. Also dspace postgresql database needed to be dropped/recreated.

8. Copy the DSpace Web application archives (.war files) to the appropriate directory in Tomcat installation.

```
cp /usr/local/src/dspace-1.4.1-souce/build/*.war var/lib/tomcat55/webapps
```

9. Create an initial administrator account:
/dspace/bin/create-administrator

email: sp0065@gmail.com
Serhiy
Polyakov
password: dspadm

---

10. Start up (or restart) Tomcat.

Visit the base URL of the server to see the DSpace home page:
http://spmachine.lis.unt.edu:8080/dspace

In order to set up some communities and collections administrator UI needed to be accessed:
http://spmachine.lis.unt.edu:8080/dspace/dspace-admin

2.6. References

DSpace System Documentation: Installation (DSpace) <http://www.dspace.org/technology/system-docs/install.html>

PostgreSQL Installation <http://www-css.fnal.gov/dsg/external/freeware/pgsql_Install.html>

3. Greenstone Digital Library Software

3.1. Overview

Greenstone is a suite of software for building and distributing digital library collections. It provides a new way of organizing information and publishing it on the Internet or on CD-ROM. The aim of the Greenstone software is to empower users to build their own digital libraries, to encourage the effective deployment of digital libraries to share information and place it in the public domain.

3.2. Installation

1. Download Unix distribution of Greenstone Digital Library (GSDL) from the web site www.greenstone.org/english/download.html

2.
```
gunzip -c gsl-2.72-unix.tar.gz | tar -xf -
```
(to /usr/local/src/)

3.
Change to the directory
```
/usr/local/src/gsl-2.72-unix/Unix
```

and install:
```
sh Install.sh
```

4.
You will be asked a number of questions which can be answered intuitively. However, some of the important questions and the answers are given below

Q: Enter the directory to install Greenstone into. A gsdl directory will be created in this directory

/var/www/html

Q: You may either install pre-compiled, statically linked linux [b]inaries (i386 only) or install and [c]ompile the Greenstone source code. Enter "[b]" or "c".

Ans : Enter b or just press the enter key (default).

Q: The default Greenstone cgi-bin directory
/var/www/html/gsdl/cgi-bin … Enter [1] or 2

Ans: Enter 1 or just press the enter key (default)

Q: Enter the web address of the gsdl/cgi-bin directory
…
Ans: http://spmachime.lis.unt.edu/gsdl/cgi-bin

Q: Enter password for the administration pages. A user with the username 'admin" will be created with the password you provide.

Ans: green

Add the following options to /etc/httpd/httpd.conf file

```
#Configure your webserver to treat
#/var/www/html/gsdl/cgi-bin as a cgi executable directory.
#For the Apache webserver this means adding the following
#ScriptAlias directive to your httpd.conf configuration file.

 ScriptAlias /gsdl/cgi-bin/ "/var/www/html/gsdl/cgi-bin/"
  <Directory "/var/www/html/gsdl/cgi-bin">
    AllowOverride None
    Options None
    Order allow,deny
    Allow from all
  </Directory>

#Configure your webserver to treat
#/var/www/html/gsdl as a web accessible directory.
#For the Apache webserver this means adding the following
#Alias directive to your httpd.conf configuration file.

Alias /gsdl/ "/var/www/html/gsdl/"
  <Directory "/var/www/html/gsdl">
    Options Indexes MultiViews FollowSymLinks
    AllowOverride None
    Order allow,deny
    Allow from all
 </Directory>
```

5.
Restart web server.

6. Access Grenstone:
http://spmachime.lis.unt.edu/gsdl/cgi-bin/library

3.3. References

Greenstone Digital Library Software (New Zealand Digital Library Project): <http://www.greenstone.org/>

4. CONTENTdm

4.1. Overview

CONTENTdm is powerful and flexible digital collection management package. It provides tools for everything from organizing and managing to publishing and searching digital collections over the Internet.

CONTENTdm consists of three main elements:
- The Acquisition Station where data and images are imported or added
- The CONTENTdm Server where data and images are stored
- The customizable Web-based search client

DiMeMa offers two convenient, free 60-day evaluation options. With hosted evaluation user are able to build, organize, search, and manage test collections of media items on the DiMeMa host. A full evaluation allows complete testing in user's environment (one copy per organization).

4.2. Installing evaluation version

Evaluation account for hosted evaluation has been obtained from UNT School of Library and Information Science. This account allows downloading Acquisition Station and managing collections on the DiMeMa host.

In order to proceed with the hosted evaluation the following parameters needed to be known:
Institution ID
License code
Hosted CONTENTdm URL
User Account and password

1. Setup an account at User Support Center:
http://www.dimema.com/USC/index.asp

2. Download Acquisition Station from User Support Center. Acquisition Station is used to catalog and publish digital objects to the CONTENTdm Web site. The published objects may be viewed with any standard Web browsers.

3. Install Acquisition Station.

Demonstration site comes with 27 pre-built collections. In hosted environment at DiMeMa "admin" accounts are not allowed to add collections or users.

5. Fedora

5.1. Overview

Fedora is a general purpose repository system. The Fedora Project is devoted to the goal of providing open-source repository software and related services to serve as the foundation for many types of information management systems. The Fedora Project is based on previous research at Cornell University Computer Science that was funded by DARPA and the National Science Foundation. The Fedora Project is currently supported by generous grants from the Andrew W. Mellon Foundation.

5.2. Install Prerequisites

Required:
    - Java SE Development Kit (JDK) 5.0.
Optional:
    - Database
    The Fedora server is backed in part by a relational database. To simplify installation, the Fedora installer includes and can configure an embedded instance of the McKoi SQL Database 1.0.3. Fedora also supports MySQL, Oracle 9 and PostgreSQL.
    - Servlet Container
    The Fedora installer includes Tomcat 5.0.28.
Additional requirements for building Fedora from source:

    Ant 1.6.5 or later

5.3. Download Fedora

The latest version of the software can always be found at http://www.fedora.info/download/.
There are two download options: the Fedora Installer and the source code distribution.

5.4. Prepare Environment Variables

To set an environment variable run:
```
[VARIABLE_NAME] ="variable value"
export [VARIABLE_NAME]
```

Also, add the lines above to
```
/etc/profile
```

The following environment variables must be correctly defined:

JAVA_HOME
This should point to the base directory of Java installation.

FEDORA_HOME
This is the directory where Fedora will be installed, for example, /usr/local/fedora.

PATH
This must include the Java and Fedora bin directories. For UNIX derivatives, this will be $FEDORA_HOME/server/bin, $FEDORA_HOME/client/bin and usually $JAVA_HOME/bin.

If Fedora is configured to use Tomcat, CATALINA_HOME must be set before starting Fedora. If using the quick install option, CATALINA_HOME should be set to $FEDORA_HOME/tomcat.

5.5. Types of Installation

The Fedora Installer provides three installation options: quick, custom, and client.

To start the installer, change to the directory where you downloaded the installer and at a command prompt, enter
```
java -jar fedora-2.2-installer.jar
```

Please ensure that the user account that is running the installer has sufficient permissions to write to the directories where Fedora will be installed (if deploying to an existing Tomcat installation, this includes permissions to the Tomcat directory). Installer created files will usually be owned by the user running the

installer. Consequently, for example, after installation users of the Fedora Admin client will need write permissions to the installer created FEDORA_HOME/client log directories.

5.5.1. Quick

The quick option is designed to get Fedora up and running as quickly as possible, with a minimum of advanced options. The quick install will automatically install a servlet container and database. Neither SSL support nor XACML policy enforcement is enabled by the quick install.

5.5.2. Custom

The custom option provides the most flexibility in configuring an installation. Options include the choice of servlet container, database, the host and ports Fedora will be running on, as well as security options including SSL and XACML policy enforcement.

Servlet Container
The installer will automatically configure and deploy to Tomcat 5.0.x and 5.5.x servlet containers. However, if an existing Tomcat installation (as opposed to the Tomcat bundled with the installer) was selected, the installer will not overwrite existing server.xml, but rather, place a modified copy at FEDORA_HOME/install so that you may review it before before installing it yourself.

5.5.3. Client

Both the quick and custom options will install the Fedora client software in addition to the Fedora server. The client option, however, will install only the Fedora client software.

5.6. Installation

Custom installation option with Tomcat (v 5.0.28) bundled with Fedora has been selected. This decision has been made after unsuccessful attempts to run Fedora Administration utility with Tomcat 5.5 currently installed on the system.

Run:
```
java -jar fedora-2.2-installer.jar
```

Options:

Administrator account
fedoraAdmin
password: [fedoraAdmin]

Authentication for users:
no

SSL availability:
no

Tomcat:
included

Path to Tomcat:
/usr/local/fedora/tomcat

Port: 8081
(Existing tomcat55 uses port 8080, so two tomcat servers will be running on different ports)

Policy enforcement enabled:
false

Database:
use included McKoi

Deploy local services and demos
true

5.7. Running Fedora

Tomcat which came with Fedora has been installed in FEDORA_HOME/tomcat. To run Fedora, start Tomcat by entering
$FEDORA_HOME/tomcat/bin/startup.sh

5.7.1. <u>Demo Objects</u>

Demonstration objects which came with Fedora provide an idea of how Fedora works.

NOTE: If, during a custom install, you entered values other than the defaults for fedoraServerHost (localhost) or fedoraServerPort (8080), you must run the demo object converter utility script to change the host and/or port in the demo object ingest files. The demo object conversion is only required if you are ingesting demo objects. If the demo objects are already ingested into the repository (e.g. from a previous installation), there is no need for conversion.

To ingest the demos, at a command prompt, enter:
fedora-ingest-demos [hostname] [port] [username] [password] [protocol]

Run:
```
fedora-ingest-demos spmachine.lis.unt.edu 8081 fedoraAdmin fedoraAdmin http
```

Once demo objects are ingested into the repository, they can be viewed via a web browser:
http://{hostname}:{port}/fedora/get/{objectPID}.

For example, Simple Image Demos (demo:6 and demo:7) demonstrate the UVA Simple Image behaviors by associating more complex disseminators with the objects:
http://spmachine.lis.unt.edu:8081/fedora/get/demo:6
http://spmachine.lis.unt.edu:8081/fedora/get/demo:7

See The Fedora™ Demo Documentation for more information.

5.8. References

Fedora (Cornell University Information Science and the University of Virginia Library)
<http://www.fedora.info>

Fedora Administrative GUI Client (Fedora) <http://www.fedora.info/download/2.0/userdocs/client/admin-gui/index.html>

The Fedora™ Demo Documentation (Fedora)
<http://www.fedora.info/download/2.2/userdocs/distribution/demos.html#d0e96>