The DSpace Course - Technical Basics

Lewis, Stuart

http://hdl.handle.net/10673/70

Downloaded from DSpace Repository, DSpace Institution's institutional repository
By the end of this module you will:

- Understand the DSpace application architecture
- Understand the DSpace server architecture
- Know what and when to back up within DSpace
- Understand the role of the repository administrator and the technical staff in configuring, managing and maintaining the repository (this will be discussed later in the course)
The DSpace system is organised into **three tiers** which consist of a number of **components**.

- Each layer only invokes the layer below it i.e. the application layer may not used the storage layer directly.
The storage layer is responsible for physical storage of metadata and content

DSpace uses a relational database to store all information about the organization of content, metadata about the content, information about e-people and authorization, and the state of currently-running workflows.
The business logic layer deals with managing the content of the archive, users of the archive (e-people), authorization, and workflow.
The Application Layer

<table>
<thead>
<tr>
<th>Statistics Tools</th>
<th>Web UI</th>
<th>OAI-PMH Data Provider</th>
<th>Simple Importer/Exporter</th>
<th>Media Filter</th>
<th>METS Exporter</th>
<th>Application Layer</th>
</tr>
</thead>
</table>

DSpace Public API

- The **application layer** contains **components** that communicate with the world outside of the individual DSpace installation, for example the Web user interface and the Open Archives Initiative protocol for metadata harvesting service.

- The **DSpace Web UI** is the **largest** and **most-used** component in the **application layer**. Two versions:
  1. JSPUI: Built on Java Servlet and JavaServer Page technology
  2. XMLUI (Manakin): Built on XML and Cocoon technology
These systems may reside on a single server or be hosted separately on dedicated servers.
- DSpace is split into three directory trees:
  - **Source Directory** [dspace-src]
    - Surprisingly, this is where the source code resides
  - **Install Directory** [dspace]
    - Populated during install & during normal operation
    - Contains:
      - Configuration files
      - Command line tools
      - Libraries
      - DSpace archive (depending on configuration)
  - **Web Deployment Directory** [tomcat]/webapps/dspace
    - Contains the JSPs and Java classes and libraries necessary to run DSpace
Source Directory Layout

- [dspace-source]
  - dspace/
    - build.xml
    - bin/
  - config/
    - controlled-vocabularies/
    - crosswalks/
    - emails/
    - language packs/
    - registries/
    - templates/
  - docs/
  - etc/
    - oracle/
  - modules/
    - jspui/
    - lni
    - oai
    - sword
    - xmlui
  - src/
  - target/
Installed Directory Layout

- [dspace]
  - assetstore/
  - bin/
  - config/
  - handle-server/
  - history/
  - lib/
  - log/
  - reports/
  - search/
  - upload/
  - webapps/
Log Files

- **[dspace]/log/dspace.log**
  - Main DSpace log file
  - Provides logging of events and errors that occur within the DSpace Code
  - Verbosity controlled by editing the [dspace]/config/templates/log4j.properties

- **[tomcat]/logs/catalina.out**
  - Tomcat standard output is written here
  - E.g. If Tomcat can’t find the DSpace.jar the error would be written to catalina.out
What to backup?

- Asset Store
  - This is where the bitstream files are located

- Database
  - This is where information about organization of content, metadata about the content, information about e-people and authorization, and the state of currently-running workflows is stored

- Source Directory
  - This is where the DSpace source code is located

- Installation Directory
  - This is where the files are located which are used by DSpace as it runs
Repository managers generally will manage the repository via the DSpace user interface.

Technical staff will be required to configure, customise and manage many features of the repository via the back end.

Examples of features that require configuration through the back end will be discussed throughout the course.
Practical exercise: Browse DSpace

- Open a terminal window and browse the DSpace structure and log files to familiarise yourself. The location of these can be found in the local instructions sheet.
These slides have been produced by:

- Stuart Lewis & Chris Yates
- Repository Support Project
  - http://www.rsp.ac.uk/
- Part of the RepositoryNet
- Funded by JISC
  - http://www.jisc.ac.uk/