Backup Plan

Backup Plan

1. Policy Statement

As part of Scholars Portal's implementation of the Bit-stream Copying Preservation Strategy (as detailed in the Preservation Implementation Plan), Scholars Portal is committed to regular backup procedures of both data storage areas and its operational areas (e.g. databases, application files). These backups are intended to serve as the basis for restoration of SP materials in the case of disaster recovery or corruption of data.

Data backup at Scholars Portal is coordinated through the University of Toronto Libraries' Information Technology Services. Since the data is stored on physical hardware located in the library server room, it uses the same backup hardware and software as the general library systems.

2. Implementation Examples

2.1. Data Backup

This backup strategy applies to the data storage areas containing primarily the Scholars Portal content objects. There are two backup strategies that apply to this group: one for content that is read/write, and one for content that is read-only.

2.1.1 Read/Write Space Backup Strategy

Storage space that is available to be written includes mounted storage locations which have not yet reached their maximum size. Since content is being written to these areas, they are mounted as a read/write filesystem on the machines performing ingest.

- Read/write filesystems are subject to incremental backups on a nightly basis.
- Data backed up on a given night is taken to off-site storage the following day, with a full copy of all data being maintained off site at all times.
- The backup system will maintain up to seven copies of any given file within a 30-day window.

2.1.2 Read-only Space Backup Strategy

Storage space that is considered "closed" is mounted read-only in order to minimize the chance of accidental tampering or deletion: for instance, ejournals volumes that have reached 2TB in size. Due to their unchanging nature, these backups are handled differently than read/write.

- At the time they are made read-only, an archive of the directory is made using the UNIX tar and gzip commands. This copy of the data is backed up to tape using backup software and is kept permanently, off site.
- The tapes containing these archives are subject to yearly testing and refreshment in order to avoid media degradation.

2.2 Database Backup

This backup strategy applies to content that is stored in a database. Primarily, this refers to objects located in Scholars Portal's MarkLogic database.

2.2.1 MarkLogic Database Backup Strategy

MarkLogic database backups are dependent on weekly database dumps. These dumps include all data and indexes in the database.

- Database dumps are backed up weekly and taken off site. Each backup is kept for three weeks.
- Transaction journals are also maintained for MarkLogic databases, allowing roll-back between weekly full database backups if these journals are available.

2.3 Application Backup

Application files are treated in the same manner as read/write data storage locations.

2.3.1 Application Files Backup Strategy

- Read/write filesystems are subject to incremental backups on a nightly basis.
- Data backed up on a given night is taken to off-site storage the following day, with a full copy of all data being maintained off site at all times.
- The backup system will maintain up to seven copies of any given file within a 30-day window.

3. Document History

Version	Date	Change	Author
0.1	09/06/11	First Draft	Steve Marks
0.2	09/19/11	Clarified wording per Amaz	Steve Marks
0.3	11/02/11	Minor textual corrections	Karl Nilsen

File	Modified
Microsoft Word 97 Document Scholars Portal Backup Plan.doc	Nov 02, 2011 by Karl Nilsen