

UQ eSpace Storage and Backup

UQ eSpace currently uses a multi-tier architecture. The public-facing web application (a Fez instance) runs on a web server, dynamically generating content from a MySQL database server and provides access to digital objects stored in a Fedora repository system.

Digital objects stored in eSpace exist as files stored in a filesystem attached to the Fedora repository server. This is called the Fedora datastore. Associated metadata is stored in MODS XML files - also in the Fedora datastore. The metadata is uploaded to the MySQL relational database (and indexed into SOLR) to facilitate discovery and access through the web tier.

The web server, Fedora repository and MySQL database server currently exist as separate virtual machines on the Library's primary vSphere cluster housed in Prentice Data Centre 2. A corresponding set of three virtual machines exist in our second site vSphere cluster (currently in the Biological Sciences Library - soon to be in the Duhig Building) to provide Business Continuity/Disaster Recovery capabilities for UQ eSpace.

The Fedora datastore is synced from the primary site to the second site. It is also backed up to tape via our BackupExec enterprise backup system. It is also synced to a ZFS filesystem in BSL and to a SAM-FS archiving system in the Duhig Building . SAM-FS is a hierarchical storage management system whose current policy makes two copies of all files to separate tapes in our tape archive. So there are multiple copies of the datastore - in DC2, Duhig and in the Biological Sciences Library. Currently a copy is also being synced to a virtual machine in the QERN infrastructure and the use of Amazon Web Services cloud infrastructure (especially Glacier) is being investigated to provide an off-site dark archive - and potentially a complete off-site eSpace capability.

The MySQL database is replicated to two 'slaves' - one in the second site. One slave is used to backup the database daily using a SQL dump - which is compressed and stored on a ZFS filesystem in BSL and also backed up to tape via BackupExec. The SQL dumps are also protected using daily ZFS snapshots to provide multiple restore points for the database without recourse to restoration from tape. In the unlikely event that we lost all copies of the MySQL database (and all slaves, snapshots and backups), the critical metadata describing the objects in the Fedora datastore can be regenerated from the MODS XML files in the datastore itself.