Performance

The Fedora community has conducted a number of performance tests. The details of each test, including testing parameters and results, are included in the following sections.

A number of factors influence Fedora 4 performance, such as:

- Server hardware (CPU, disk, memory, network bandwidth)
- Repository configuration (storage, indexing and sequencers configured)
- Data (number of child resources)

For best performance, we recommend:

- Configure Java to have a large heap size, e.g., `-Xmx2048m`
- Limit the number of children under a single resource – ideally to 1,000 or less. Consider using a hierarchy to organize resources, e.g., storing the resource with ID `12345678` at the path `/12/34/56/78`. Automatically-generated identifiers will use this pattern.
- Particularly when working with files larger than available memory, use Infinispan file storage (e.g. `-Dfcrepo.modeshape.configuration=config/single-file/repository.json`) or filesystem federation. Federation avoids having to transfer files using HTTP – and with larger file sizes (or with larger numbers of files being processed), this can improve performance significantly.

Single Node Performance Testing

Ingest Testing

Ingesting content into Fedora 4 is generally faster than ingesting into Fedora 3 on the same hardware (tested up to 10,000 containers with 50MB binaries).

- See Ingest Test Matrix.

Create/Read/Update/Delete Testing

Updating content in Fedora 4 is generally faster than in Fedora 3.

- See Create/Read/Update/Delete Test Results.

Large File Ingest and Retrieval

Arbitrarily-large files can be ingested (or projected into the repository using filesystem federation) and downloaded via the REST API (tested up to 1TB). The only apparent limitations are disk space available to store the files, and a sufficiently large Java heap size (tested with -Xmx2048m).

- See Large File Ingest and Retrieval.

Clustered Performance Testing

Ingest and read/update/delete testing on multiple servers configured as a cluster.

- See Response Time Comparison of Single Fedora VS Cluster on a Single machine.

CRUD testing on multiple servers configured as a cluster with a load balancer.

- See Response Time Comparison of Single Fedora VS Cluster in AWS

Case Studies

Reports of real-world ingest and performance testing.

- See Stanford SALT Collection