

2016-05 Performance and Scalability Report

Goals

Initial goals of the Performance and Scalability group:

- Document the current performance of targeted operations in Fedora 4 (creating binaries of various sizes, creating and updating containers)
- Identify performance bottlenecks to inform future development
- Identify performance changes over time with infrastructure updates (e.g., the upgrade to Modeshape 5)

Current Work

- Creating testing scenarios for common operations: [Performance and Scalability Test Plans](#)
 - Creating small files
 - Creating large files
 - Creating empty containers
 - Creating containers with RDF properties
- Creating JMeter test plans to test scenarios in a reproducible way: <https://github.com/fcrepo4-labs/fcrepo4-jmeter>
 - Periodic checks of small-scale CRUD operations to check the health of the repository as the test progresses
 - Clearly defined failure criteria
 - Command-line options to run different scenarios and fine-tune test parameters
- Running several of the tests by several different institutions, including repeated tests with different storage backends (LevelDB vs. PostgreSQL) or other relevant options:
 - [Princeton Results - Test 2](#)
 - [Princeton Results - Test 3](#)
 - [Princeton Results - Test 4](#)
 - [UW Madison Results - Test 4](#)
 - [Virginia Tech Results - Test 4](#)
 - [York University Results - Test 4](#)

Future Work

- Have multiple institutions run all of the identified tests with Fedora 4.5.1 to provide a solid baseline measurement of current performance.
- Repeat the same tests with a Modeshape 5 branch to compare performance of the new version of Modeshape.
- Run tests with profiling enabled to gather information most relevant to identifying and addressing performance bottlenecks in Fedora code.