Performance Considerations

Key differences between Fedora 6 and earlier versions

While the user-facing API has not changed significantly from Fedora 5 to Fedora 6, the internal storage and indexing of data in Fedora 6 has been completely rewritten in Fedora 6.   As such there are important implications of which you should be aware.

1) Writes of RDF to the repository will be slower in some cases

In Fedora 5 and earlier, RDF resources as well as small (sub 10K) binary files were stored directly in the database rather than being written to disk. This was a performance optimization in Modeshape (Fedora 4 and Fedora 5's storage layer) that came at the cost of having an obscure representation of repository data.   Some resources were stored in a database, while others lived on disk. The user and developer community found this approach unsettling especially when trying to troubleshoot issues. It required intimate familiarity with both the bespoke database structure of Modeshape as well as its application logic for where and when to store repository objects.

In Fedora 6 all RDF resources are stored on disk in OCFL. If you're writing to a local SSD storage you may not notice a significant decrease in performance. However if you are using NAS (Network Attached Storage) you may observe slower performance than Fedora 5.   In order to support the data preservation sensibilities offered by OCFL - namely a transparent on disk representation of the repository - Fedora 6 must write all persistent resources to disk regardless their size and/or function.   So the difference in write performance is attributable to the fact that Fedora 6 is writing everything to NAS (which can have significant latency depending on the characteristics of your network) whereas Fedora 5 was writing directly to a database.