Before you start...

There are several facilities that support a running Fedora repository. Each of them may be supplied by prepackaged default tools, or by externalized replacements. The choice of which to use for your installation depends on your local resources as well as what you expect to do with your repository.

Storage Configuration

TBD.

The Relational Database

TBD.

The Resource Index

The first question to ask about the Resource Index is whether or not you want to enable it at all. The Resource Index provides for fast SPARQL/iTQL/SPO queries against the RDF graph of objects in your repository. Some advanced features of the Fedora platform depend on its availability, such as using RDF queries to supply resource attributes for making policy-based access decisions. If you are not using these advanced features, and you do not expect to offer access to the RDF graph of your repository as a service, then you may wish to disable the Resource Index, which you can do as an option in the Fedora installer.

If you do decide to enable the Resource Index, you may either use the prepackaged Mulgara triplestore, or you may supply an external triplestore to your repository. Trippi, the Fedora API for accessing triplestores, currently supports Mulgara and MPTStore.The advantages of the prepackaged Resource Index include:

- It is simple. You need do very little or nothing more to have a working Resource Index.
- It may be all you need. With appropriate computing resources supplied to your repository, the built-in Mulgara instance can scale to several millions of triples.

The disadvantages include:

- The built-in triplestore is difficult or impossible to upgrade independently of Fedora itself. This means that if you would like to use functionality provided in a version of Mulgara later than that packaged with your repository (which some users have discovered to be the case) you will find it very difficult. Using a version of Mulgara as a built-in triplestore that was not packaged with your version of Fedora is not a supported configuration.
- It is not possible to supply computing resources to the triplestore or to Fedora alone. Since they run in the same JVM and as part of the same application, they share memory and processor access and other computing resources. If you find that you are having a problem scaling one or the other, you will only be able to offer resources to both.

The alternative to using the built-in triplestore is to install your own triplestore and connect it to Fedora. The advantages include:

- You are able to use the triplestore of your preference. For example, if you would like to use a version of Mulgara that is more advanced that the one packaged with your version of Fedora to get access to new features, you can.
- You can configure your triplestore the better to meet your needs. This might include adding extension functionality, adjusting indexing, etc.
- You can supply computing resources to the triplestore or to Fedora independently, as appropriate to your circumstances.

The disadvantages include:

- You will have to install, configure, and maintain a triplestore. This may or may not be onerous, depending on your local support environment. A highly-customized triplestore can be a complex piece of software to configure and deploy.
- You will probably have to supply more computing resources to meet a given level of service, as compared with using the built-in triplestore. For example, you may need to support two servers instead of one, or you may need to supply extra memory to your server to allow it to support two application containers. This may or may not be a significant burden, again depending on your local environment.
- You will have to configure Fedora to connect to your triplestore. This isn't usually very difficult, unless you have some exceptional circumstance (e. g. a very unusual network configuration), but it is more work to do.

There isn't a single right choice for everyone in this decision. Some questions you may wish to consider include:

- Is supplying query access to the RDF graph of objects in your repository important to you? (If not, you may not need to enable the Resource Index.)
- Do you expect to use any of the advanced features of the Fedora service platform that require the presence of the Resource Index? (These include the use of Enhanced Content Models and RDF queries to supply resource attributes for making policy-based access decisions.)
- Do you expect to have a very large and sophisticated graph of objects? (If your graph is going to scale beyond a few million triples and you would like to be able to query it, you may want to consider an external triplestore.)
- How much time and effort can you commit to supporting your repository? (An external triplestore offers more functionality than does the built-in triplestore, but also requires more work.)
- Are you very interested in exploring Fedora's semantic graph capabilities? (An external triplestore is the most flexible and customizable configuration for the Resource Index.)

It is to be noted that creation of a Resource Index and migration from a built-in triplestore to an external triplestore are supported by Fedora's command-line utilities, so the decision you make at installation need not be final. If you do not have a clear reason in your mind why you would want to use the Resource Index, you can safely leave it off, and if you do know that you will want to use the Resource Index but do not have a clear reason to use an external triplestore, you can safely use the built-in triplestore.