GettingNutchRunningWithJboss

Running Nutch with JBoss AS 5.1

I only had to make minor changes beyond the basic Tomcat tutorials to get Nutch running on JBoss AS 5.1

Deployment

Make sure that your nutch-site.xml file is configured in your packaged .war file, or exploded .war directory.

Xerces

JBoss ships with a different version of Xerces installed and available to all deployed applications. I was getting an error about the conflict. Removing xerces-2_x_x-apis.jar and xerces-2_x_x.jar from the war file's lib directory fixed the problem.

Code changes

In the file:
/src/java/org/apache/nutch/plugin/PluginManifestParser.java

Nutch checks has a check to make sure it can find the plugin folder. The check looks like this:

```
} else if (!"file".equals(url.getProtocol())) {
  LOG.warn("Plugins: not a file: url. Can't load plugins from: " + url);
  return null;
}
```

This does not work in jboss because local files in the deployment directory have a protocol of vfsfile:// since vfsfile acts just like file, you just have to change this code to:

```
} else if (!"file".equals(url.getProtocol()) &&
        !"vfsfile".equals(url.getProtocol())) {
  LOG.warn("Plugins: not a file: url. Can't load plugins from: " + url);
  return null;
}
```