

GettingNutchRunningWithRedHatApplicationServer

This was written for RedHat Application Server 1.0. There may be new versions of the rpms listed, so you might need to translate the version number from what is listed here. I'm assuming that you're using the RedHat Network to get these packages. That should help you when it comes to getting all the dependencies.

These instructions have you installing software from multiple channels. If your server was not previously subscribed to one or both of these channels, make sure that you have rhn update the package list for the server after you've added the subscriptions to these channels.

Packages to Install from RedHat Application Server Channel

One big caveat is that the version of ant that comes from RedHat throws an error when you try to build the nutch war file. I suggest installing ant manually. I've added some instructions for doing that further down.

- ant-1.6.2-3jpp_3rh.noarch.rpm
- ant-apache-regexp-1.6.2-3jpp_3rh.noarch.rpm
- axis-1.2-0.rc2.1jpp_2rh.noarch.rpm
- jaf-1.0.2-3jpp_8rh.noarch.rpm
- jakarta-commons-beanutils-1.7.0-1jpp_2rh.noarch.rpm
- jakarta-commons-collections-3.1-1jpp_2rh.noarch.rpm
- jakarta-commons-daemon-1.0-2jpp_2rh.noarch.rpm
- jakarta-commons-dbcp-1.2.1-3jpp_3rh.noarch.rpm
- jakarta-commons-digester-1.6-2jpp_2rh.noarch.rpm
- jakarta-commons-discovery-0.2-1jpp_1rh.noarch.rpm
- jakarta-commons-el-1.0-2jpp_1rh.noarch.rpm
- jakarta-commons-fileupload-1.0-3jpp_2rh.noarch.rpm
- jakarta-commons-httpclient-2.0.2-1jpp_2rh.noarch.rpm
- jakarta-commons-launcher-0.9-3jpp_2rh.noarch.rpm
- jakarta-commons-logging-1.0.4-2jpp_2rh.noarch.rpm
- jakarta-commons-modeler-1.1-3jpp_2rh.noarch.rpm
- jakarta-commons-pool-1.2-2jpp_2rh.noarch.rpm
- jakarta-commons-validator-1.1.3-1jpp_2rh.noarch.rpm
- jakarta-regexp-1.2-15.i386.rpm
- jakarta-taglibs-standard-1.1.1-4jpp_2rh.noarch.rpm
- jakarta-taglibs-standard-javadoc-1.1.1-4jpp_2rh.noarch.rpm
- javamail-1.3.1-1jpp_4rh.noarch.rpm
- jta-1.0.1-0.b.3jpp_1rh.noarch.rpm
- jta-javadoc-1.0.1-0.b.3jpp_1rh.noarch.rpm
- junit-3.8.1-3jpp_3rh.noarch.rpm
- libgcj34-3.4.0-1.i386.rpm
- log4j-1.2.8-7jpp_1rh.noarch.rpm
- mx4j-2.1.0-1jpp_4rh.noarch.rpm
- oro-2.0.8-1jpp_1rh.noarch.rpm
- regexp-1.3-1jpp_3rh.noarch.rpm
- servletapi4-4.0.4-3jpp_3rh.noarch.rpm
- servletapi5-5.0.18-1jpp_3rh.noarch.rpm
- struts11-1.1-1jpp_4rh.noarch.rpm
- tomcat5-5.0.28-2jpp_5rh.noarch.rpm
- tomcat5-admin-webapps-5.0.28-2jpp_5rh.noarch.rpm
- tomcat5-webapps-5.0.28-2jpp_5rh.noarch.rpm
- tyrex-1.0.1-2jpp_2rh.noarch.rpm
- wsdl4j-1.4-2jpp_1rh.noarch.rpm
- xalan-j2-2.6.0-2jpp_3rh.noarch.rpm
- xerces-j2-2.6.2-4jpp_2rh.noarch.rpm
- xml-commons-1.0-0.b2.6jpp_3rh.noarch.rpm
- xml-commons-apis-1.0-0.b2.6jpp_3rh.noarch.rpm
- xml-commons-resolver-1.1-1jpp_2rh.noarch.rpm

Packages to Install from Extras Channel

This assumes you want to use IBM's Java.

- java-1.4.2-ibm-1.4.2.2-1jpp_9rh
- java-1.4.2-ibm-devel-1.4.2.2-1jpp_9rh

Replacing Ant

When I tried to build the nutch war file using the ant rpm listed above I got errors about "XslpLiaison". In the end I installed ant from scratch and got it to work for me. Here's what I did:

1. Download the ant source from <http://ant.apache.org/>
2. Verify the version you downloaded is legit (they offer pgp and other signatures).
3. Extract the package and copy it to /usr/share (creating /usr/share/apache-ant-[version]).

4. Replace the original version of ant with the fresh one. As root:
 - a. `cd /usr/bin`
 - b. `mv ant ant_orig`
 - c. `ln -s /usr/share/apache-ant-1.6.5/bin/ant ./`
 - d. `mv /etc/ant.conf /etc/ant.conf_rpm_orig`

Now you can download nutch (see [DownloadingNutch](#)) and run through the tutorial (<http://lucene.apache.org/nutch/tutorial.html>). One thing to note is that you'll want to put the segments directory under /usr/share/tomcat5. Once it's there, you can start tomcat with the command "service tomcat5 start".

<<< [FrontPage](#)