



链滴

Hadoop学习笔记

作者: [AnnPeter](#)

原文链接: <https://ld246.com/article/1463402946417>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

```

<div><div></div><h2 id="编译由于官网上下载的是32位版本的如果我们的服务器是64位的需要自己手动编译">编译, 由于官网上下载的是32位版本的, 如果我们的服务器是64位的, 需要自己手动编译</h2><ul>
<li><p>安装必要的包</p>
<blockquote>
<p><span>$</span> yum install autoconf automake libtool cmake ncurses-devel openssl-devel gcc* --nogpgcheck</p>
</blockquote></li>
<li><p>安装maven</p>
<blockquote>
<p><a href="https://ld246.com/forward?goto=http%3A%2F%2Fmirrors.cnnic.cn%2Fapache%2Fmaven%2Fmaven-3%2F3.3.9%2Fbinaries%2Fapache-maven-3.3.9-bin.tar.gz" target="_blank" rel="nofollow ugc">http://mirrors.cnnic.cn/apache/maven/maven-3/3.3.9/binaries/apache-maven-3.3.9-bin.tar.gz</a></p>
</blockquote>
<p>添加环境变量</p>
<blockquote>
<p>export MAVEN_HOME=/usr/local/apache-maven-3.3.9/ <br>
export PATH=<span>$</span>MAVEN_HOME/bin:$PATH</p>
</blockquote></li>
<li><p>安装protobuf</p>
<blockquote>
<p><a href="https://ld246.com/forward?goto=http%3A%2F%2Fpan.baidu.com%2Fs%2F1pJZubT" target="_blank" rel="nofollow ugc">http://pan.baidu.com/s/1pJZubT</a> <br>
<span>$</span> ./configure --prefix=/usr/local/protoc/ <br>
<span>$</span> make &amp;&amp; make install <br>
<span>$</span> protoc --version</p>
</blockquote></li>
<li><p>编译Hadoop(编译好的Hadoop位于当前目录下的hadoop-dist)</p>
<blockquote>
<p><span>$</span> mvn package -Pdist,native -DskipTests -Dtar</p>
</blockquote></li>
</ul><div></div><h2 id="安装">安装</h2><ol>
<li><p>安装前请先安装jdk</p></li>
<li><p>配置hadoop</p>
<ul><li>进入etc/hadoop</li>
<li>修改hadoop-env.sh, 添加java环境变量(根据Hadoop自己的方式找有可能找不到, 这里我们接指定)</li>
<li><p>修改core-site.xml</p>
<pre class="prettyprint linenums prettyprinted"><ol><li class="L0"><code class="language-xml"><span>&lt;property&gt;</span></code></li><li class="L1"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>fs.defaultFS</span><span>&lt;name&gt;</span></code></li><li class="L2"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>hdfs://centos100.annpeter.cn:9000</span><span>&lt;value&gt;</span></code></li><li class="L3"><code class="language-xml"><span>&lt;/property&gt;</span></code></li><li class="L4"><code class="language-xml"><span>&lt;property&gt;</span></code></li><li class="L5"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>hadoop.tmp.dir</span><span>&lt;/name&gt;</span><span></span><span>&lt;!--hadoop工作目录--&gt;</span></code></li><li class="L6"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>/usr/local/hadoop-2.7.2/tmp</span><span>&lt;/value&gt;</span></code></li><li class="L7"><code class="language-xml"><span>&lt;/property&gt;</span></code></li></ol></pre></li>
<li><p>修改hdfs-site.xml</p>
<pre class="prettyprint linenums prettyprinted"><ol><li class="L0"><code class="language

```

```
xml" > <span>&lt;property&gt;</span></code></li><li class="L1"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>dfs.replication</span><span></span></code></li><li class="L2"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>1</span><span>&lt;/value&gt;</span></code></li><li class="L3"><code class="language-xml"><span>&lt;/property&gt;</span></code></li></ol></pre></li>
```

<p>修改mapred-site.xml</p>

```
<pre class="prettyprint linenums prettyprinted"><ol><li class="L0"><code class="language-xml"><span>&lt;property&gt;</span></code></li><li class="L1"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>mapreduce.framework.name</span><span>&lt;/name&gt;</span></code></li><li class="L2"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>yarn</span><span>&lt;/value&gt;</span></code></li><li class="L3"><code class="language-xml"><span>&lt;/property&gt;</span></code></li></ol></pre></li>
```

<p>修改yarn-site.xml</p>

```
<pre class="prettyprint linenums prettyprinted"><ol><li class="L0"><code class="language-xml"><span>&lt;property&gt;</span></code></li><li class="L1"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>yarn.resourcemanager.hostname</span><span>&lt;/name&gt;</span></code></li><li class="L2"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>centos100.annpeter.cn</span><span>&lt;/value&gt;</span></code></li><li class="L3"><code class="language-xml"><span>&lt;/property&gt;</span></code></li><li class="L4"><code class="language-xml"><span>&lt;property&gt;</span></code></li><li class="L5"><code class="language-xml"><span>    </span><span>&lt;name&gt;</span><span>yarn.nodemanager.aux-services</span><span>&lt;/name&gt;</span></code></li><li class="L6"><code class="language-xml"><span>    </span><span>&lt;value&gt;</span><span>mapreduce_shuffle</span><span>&lt;/value&gt;</span></code></li><li class="L7"><code class="language-xml"><span>&lt;/property&gt;</span></code></li></ol></pre></li></ul></li>
```

<p>启动

在启动前，我们可以将Hadoop的命令加入环境变量</p>

<blockquote>

```
<p>export HADOOP_HOME=/usr/local/hadoop-2.7.2 <br>
export PATH=<span>$</span>HADOOP_HOME/bin:<span>$</span>HADOOP_HOME/sbin
<span>$</span>PATH</p>
```

</blockquote>

<p>关闭防火墙</p>

<blockquote>

```
<p><span>$</span> service iptables stop <br>
<span>$</span> chkconfig iptables off</p>
```

</blockquote>

<p>格式化文件系统</p>

<blockquote>

```
<p><span>$</span> hadoop namenode -format</p>
```

</blockquote>

<p>启动HFS</p>

<blockquote>

```
<p><span>$</span> start-dfs.sh</p>
```

</blockquote>

<p>启动后可以使用jps查看启动的进程</p>

<p>启动yarn(启动后，可以通过浏览器访问50070端口，查看你的DFS)</p>

<blockquote>

```
<p><span>$</span> start-yarn.sh</p>
```

</blockquote>

<p>测试文件存储功能，上传、下载一个文件(在写路径时，可以直接写uri，省略协议和域名)</p>

```
<blockquote>
  <p><span>$</span> hadoop fs -put XX.tar.gz hdfs://centos100.annpeter.cn <br>
  <span>$</span> hadoop fs -get /XX.tar.gz</p>
</blockquote>
<p>测试运行jar(Hadoop安装目录中有一个例子程序/hadoop-2.7.2/share/hadoop/mapreduce/h
doop-mapreduce-examples-2.7.2.jar)</p>
<blockquote>
  <p><span>$</span> hadoop jar hadoop-mapreduce-examples-2.7.2.jar pi 5 5 (计算圆周率p
)</p>
</blockquote> </li>
</ol> </div>
```