

Hadoop web页面的授权设定

一、相关概念

在默认情况下，Hadoop相关的WEB页面（JobTracker, NameNode, TaskTrackers and DataNodes）是不需要什么权限验证就可以直接进入的，谁都可以查看到当前集群上有哪些作业在运行，这对安全来说是很不合理的。我们应该限定用户来访问Hadoop相关的WEB页面，只有授权的用户才能看到自己授权的作业等信息，而不应该看到他不该看到的。其实Hadoop提供了简单的web页面的授权设定。

和Hadoop的RPC授权一样，Hadoop相关的WEB页面也可以通过配置Kerberos利用HTTP SPNEGO协议来进行相应的授权控制，HTTP SPNEGO协议是支持Firefox和Internet Explorer浏览器。不过Hadoop WEB页面也是支持Hadoop的Pseudo/Simple授权控制，如果相应的授权开启了，用户需要在查看某个页面的时候，需要在URL的后面带上一个查询参数，如下：

```
http://localhost:50030/jobtracker.jsp?user.name=babu
```

这样，jobtracker.jsp中将会获取到user.name参数的值，从而达到简单的授权控制。如果需要达到比较好的授权机制，用户可以自己写个Hadoop WEB页面授权的插件（比如我们可以实现自己的AuthenticatorHandler，可以参考Hadoop自带的Hadoop-auth.jar工具包）。

二、如何配置

上面说了Hadoop支持对相关的WEB页面进行简单的授权，那么如何来设定呢？下面就来进行说明。

下面的所有配置都是在\$HADOOP_HOME/etc/hadoop/core-site.xml里面进行设定，设定好后，需要将core-site.xml同步到集群中的所有机器。

1、hadoop.http.filter.initializers的值设定为org.apache.hadoop.security.AuthenticationFilterInitializer，它的默认值如下：

```
<property>
  <name>hadoop.http.filter.initializers</name>
  <value>org.apache.hadoop.http.lib.StaticUserWebFilter</value>
  <description>A comma separated list of class names. Each class in the list
  must extend org.apache.hadoop.http.FilterInitializer. The corresponding
  Filter will be initialized. Then, the Filter will be applied to all user
  facing jsp and servlet web pages. The ordering of the list defines the
  ordering of the filters.</description>
```

```
</property>
```

2、设定hadoop.http.authentication.type的值，这个属性的值支持simple、kerberos、#AUTHENTICATION_HANDLER_CLASSNAME#；默认配置是simple。

```
<property>
  <name>hadoop.http.authentication.type</name>
  <value>simple</value>
  <description>
    Defines authentication used for Oozie HTTP endpoint.
    Supported values are: simple | kerberos | #AUTHENTICATION_HANDLER_CLASSNAME#
  </description>
</property>
```

3、hadoop.http.authentication.token.validity这个值设定本次授权的token多长时间有效，单位是秒。默认配置是3600秒。

4、hadoop.http.authentication.signature.secret.file设定signature secret file文件存放的绝对路径，这个文件需要同步到集群中所有的JobTracker, NameNode, DataNode和TastTracker，它的默认值是\${user.home}/hadoop-http-auth-signature-secret。需要保证signature secret file文件可以被启动JobTracker, NameNode, DataNode和TastTracker进程的用户可读。

5、hadoop.http.authentication.cookie.domain：以Cookie存储认证Token的域名，为了能够使得授权得以生效，需要配置一个正确的值，该属性是默认值为空。

6、hadoop.http.authentication.simple.anonymous.allowed：在simple授权模式下，是否允许匿名用户请求。默认值是ture，也就是允许。

7、hadoop.http.authentication.kerberos.principal：在kerberos授权模式下，是否允许HTTP终端用 Kerberos principal，这个属性的值必须是以'HTTP/'开头。默认的值是HTTP/_HOST@\$LOCALHOST。

8、hadoop.http.authentication.kerberos.keytab：keytab file文件的存放地址，默认值是\$user.home/hadoop.keytab.i。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】（）](#)