Hadoop 学习笔记

harli

2015/1/8

\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=6666,server=y,suspend=y"

# Hadoop 调试

<http://www.lifeba.org/arch/hadoop_debug.html>

**一、HDFS调试**

1、bin/hdfs 脚本中  
namenode下加入

 HADOOP\_OPTS="$HADOOP\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=7777,server=y,suspend=y"

secondarynamenode下加入

HADOOP\_OPTS="$HADOOP\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=7778,server=y,suspend=y"

**[](http://www.lifeba.org/wp-content/uploads/2012/05/lifeba.org_hadoop_debug_1.jpg)**

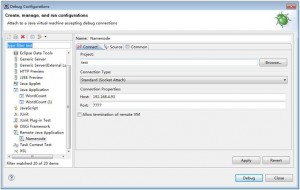
2、执行相应的脚本，然后通过Eclipse IDE连接来下断调试。

启动服务./start-dfs.sh，停止服务./stop-dfs.sh

启动后会停留在 Listening for transport dt\_socket at address: 7777 ，IDE连接后才会继续。

**二、eclipse hadoop 调试**

打开Eclipse的 Debug Configurations ，新建一个Remote Java Application。指定下project工程项目，以及需要调试的hadoop节点的ip和端口。

**[](http://www.lifeba.org/wp-content/uploads/2012/05/lifeba.org_hadoop_debug_3.jpg)**

在hadoop节点启动后，点击debug来连接远程的hadoop程序进行调试。hadoop节点在调试启动方式下，必须IDE连接上去后，才会继续，否则会停留在  
上面的Listening for transport dt\_socket at address: 7777，直到有IDE连接成功。

说明：  
1、如果需要调试datanode，在datanode节点下的bin/hdfs 中的datanode节点同样加入

HADOOP\_OPTS="$HADOOP\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=7777,server=y,suspend=y"

然后通过./hdfs datanode datanode启动，这样就可以连调。

2、不能使用外层hadoop启动或者start-all.sh启动，调试的话必须直接使用实际调用的sh启动。如  
如 ./hadoop namenode -format 要换为 ./hdfs namenode -format

**三、MapReduce调试**

1、调试MapReduce先启动HDFS服务，然后在bin/mapred 脚本中  
jobtracker下加入

HADOOP\_OPTS="$HADOOP\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=6666,server=y,suspend=y"

tasktracker下加入

HADOOP\_OPTS="$HADOOP\_OPTS -Xdebug -Xrunjdwp:transport=dt\_socket,address=6667,server=y,suspend=y"

[](http://www.lifeba.org/wp-content/uploads/2012/05/lifeba.org_hadoop_debug_2.jpg)

2、执行相应的脚本，同样通过Eclipse IDE连接来下断调试。

启动服务start-mapred.sh，停止服务stop-mapred.sh

如果需要驱动某个节点，可以执行单个启动命令。./mapred jobtracker 后面指定需要开启的服务的名称。

# 参考

<http://www.lifeba.org/arch/hadoop_debug.html>