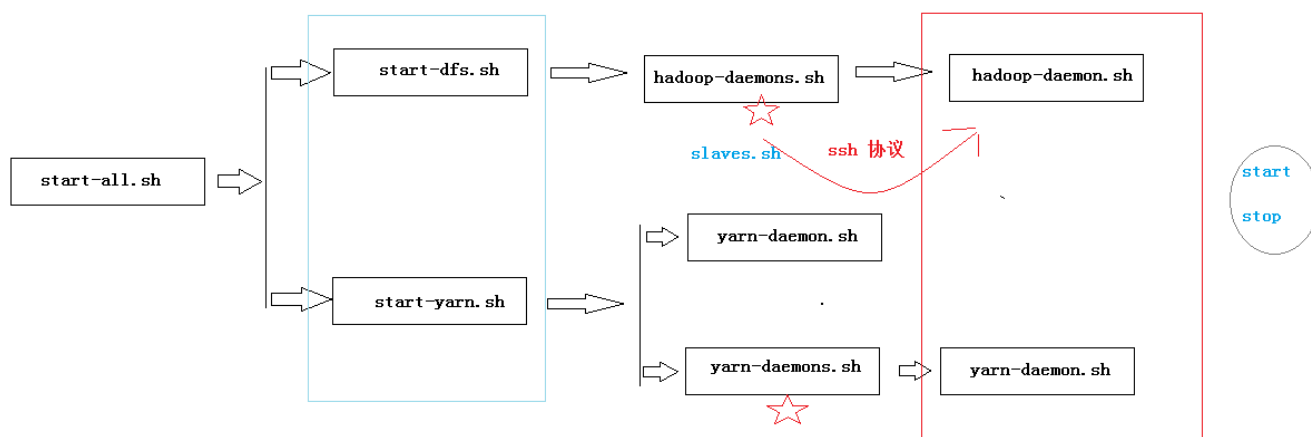


Hadoop 2.x 深入浅出基础课程笔记之 Lesson 04

内容提纲:

- 1) 回顾 Hadoop 2.x 伪分布式环境安装部署（配置文件的内容）
- 2) 启动 Hadoop Deamons 三种启动方式，并分析启动脚本
- 3) 三种启动方式的关系
- 4) SSH 协议、公钥登录、无密钥登录和配置详解
- 5) Hadoop 2.x 默认配置文件讲解、对 Hadoop 2.x 配置进行调优配置
- 6) MapReduce 程序相关日志讲解

1、三种启动方式



2、SSH 协议与配置

SSH 协议

- ◆ SSH是一种网络协议，用于计算机之间的加密登录。
- ◆ 如果一个用户从本地计算机，使用SSH协议登录另一台远程计算机，可以认为，这种登录是安全的，即使被中途截获，密码也不会泄露。
- ◆ 公钥加密
 - (1) 远程主机收到用户的登录请求，把自己的公钥发给用户。
 - (2) 用户使用这个公钥，将登录密码加密后，发送回来。
 - (3) 远程主机用自己的私钥，解密登录密码，如果密码正确，就同意用户登录。
- ◆ 命令

```
$ ssh localhost
```

配置 SSH 无密钥登陆

- ◆ 使用密码登录，每次都必须输入密码，非常麻烦。SSH还提供了公钥登录，可以省去输入密码的步骤。
- ◆ 所谓“公钥登录”，就是用户将自己的公钥储存在远程主机上。登录的时候，远程主机会向用户发送一段随机字符串，用户用自己的私钥加密后，再发回来。远程主机用事先储存的公钥进行解密，如果成功，就证明用户是可信的，直接允许登录shell，不再要求密码。
- ◆ 生成公钥/私钥对

```
$ ssh-keygen -t rsa
```
- ◆ 在\$HOME/.ssh/目录下，会新生成两个文件：id_rsa.pub和id_rsa。前者是你的公钥，后者是你的私钥。
- ◆ 拷贝公钥至authorized_keys文件

```
$ cp ~/.ssh/id_rsa.pub ~/.ssh/authorized_keys
```

3、Hadoop 2.x 配置文件

Hadoop 2.x 配置文件

◆ 默认的

- 位置: \$HADOOP_HOME/share/hadoop 中对应project文件夹中的Jar文件中
- 文件名: core-default.xml、hdfs-default.xml、yarn-default.xml、mapred-default.xml

◆ 自定义的

- 位置: \$HADOOP_HOME/etc/hadoop
- 文件名: core-site.xml、hdfs-site.xml、yarn-site.xml、mapred-site.xml

◆ 查阅 默认配置文件与说明

4、MapReduce 程序相关日志

Hadoop 2.x 配置文件

◆ 默认的

- 位置: \$HADOOP_HOME/share/hadoop 中对应project文件夹中的Jar文件中
- 文件名: core-default.xml、hdfs-default.xml、yarn-default.xml、mapred-default.xml

◆ 自定义的

- 位置: \$HADOOP_HOME/etc/hadoop
- 文件名: core-site.xml、hdfs-site.xml、yarn-site.xml、mapred-site.xml

◆ 查阅 默认配置文件与说明