限定机器访问Hadoop集群

随着使用集群用户规模的增大，Hadoop集群安全问题就摆在我们面前；如何来防止恶意用户访问Hadoop集群？这是很多人都在思考的问题。本文主要是通过用防火墙的功能来实现简单的安全控制，只能限定到IP范围，不能实现控制目录级别的控制，如果你想了解更多关于Hadoop集群安全问题，可以阅读Kerberos安全。

以CentOS为例，在命令行里面查看现有防火墙的配置可以通过下面命令实现：

```
[wyp@iteblog /etc/sysconfig]$ sudo iptables -L -n --line-number
Chain INPUT (policy ACCEPT)
num  target     prot opt source               destination

Chain FORWARD (policy ACCEPT)
num  target     prot opt source               destination

Chain OUTPUT (policy ACCEPT)
num  target     prot opt source               destination
```

如果没有做任何防火墙配置输出应该如上述所示。

Hadoop集群中HDFS默认的是开启8020端口提供给外面访问；同样，ResourceManager是开启8042端口来给外面访问的。我们可以用下面的方法限制外面机器访问集群。

第一步：先关掉先前开启的8020端口

```
[wyp@iteblog /etc/sysconfig]$ sudo iptables ₩
                              -I INPUT -p tcp --dport 8020 -j DROP
```

第二步：把子群的IP加入到8020的允许访问列表里面：

```
[wyp@iteblog /etc/sysconfig]$ sudo iptables ₩
                              -I INPUT -s 192.168.24.69 -p tcp --dport 8020 -j ACCEPT
[wyp@iteblog /etc/sysconfig]$ sudo iptables ₩
                              -I INPUT -s 192.168.24.70 -p tcp --dport 8020 -j ACCEPT
[wyp@iteblog /etc/sysconfig]$ sudo iptables ₩
                              -I INPUT -s 192.168.24.71 -p tcp --dport 8020 -j ACCEPT
```

第三步：保存上面所有的设置，并重启防火墙

```
[wyp@iteblog /etc/sysconfig]$ sudo service iptables save
[wyp@iteblog /etc/sysconfig]$ sudo service iptables restart
Flushing firewall rules:                                   [  OK  ]
Setting chains to policy ACCEPT: filter                    [  OK  ]
Unloading iptables modules:                                [  OK  ]
Applying iptables firewall rules:                          [  OK  ]
```

经过上面的设置，只有IP地址192.168.24.69-71可以访问集群的8020端口，其他的机器全都不行。上面的配置全部都写在/etc/sysconfig/iptables文件中，如果你想看看上面的配置是否保存了，可以用下面的命令实现：

```
[wyp@iteblog /etc/sysconfig]$ sudo iptables -L -n --line-number
Chain INPUT (policy ACCEPT)
num  target     prot opt source               destination
1    ACCEPT     tcp  --  192.168.24.71        0.0.0.0/0           tcp dpt:8020
2    ACCEPT     tcp  --  192.168.24.70        0.0.0.0/0           tcp dpt:8020
3    ACCEPT     tcp  --  192.168.24.69        0.0.0.0/0           tcp dpt:8020
4    ACCEPT     tcp  --  192.168.24.101       0.0.0.0/0           tcp dpt:8020
5    DROP       tcp  --  0.0.0.0/0            0.0.0.0/0           tcp dpt:8020

Chain FORWARD (policy ACCEPT)
num  target     prot opt source               destination

Chain OUTPUT (policy ACCEPT)
num  target     prot opt source               destination
```

当然，你也可以直接查看/etc/sysconfig/iptables文件

```
[wyp@iteblog /etc/sysconfig]$ sudo vim /etc/sysconfig/iptables
[wyp@iteblog /etc/sysconfig]$ sudo vim /etc/sysconfig/iptables
# Generated by iptables-save v1.3.5 on Fri Jan  3 17:44:59 2014
*filter
:INPUT ACCEPT [4083424:1150372085]
:FORWARD ACCEPT [0:0]
:OUTPUT ACCEPT [10065:1809008]
-A INPUT -s 192.168.24.71 -p tcp -m tcp --dport 8020 -j ACCEPT
-A INPUT -s 192.168.24.70 -p tcp -m tcp --dport 8020 -j ACCEPT
-A INPUT -s 192.168.24.69 -p tcp -m tcp --dport 8020 -j ACCEPT
```

```
-A INPUT -s 192.168.24.101 -p tcp -m tcp --dport 8020 -j ACCEPT
-A INPUT -p tcp -m tcp --dport 8020 -j DROP
COMMIT
# Completed on Fri Jan  3 17:44:59 2014
```

如果你想添加允许其他机器访问集群的8020端口，请运行上面的
步骤二和步骤三。如果你想删掉已经设定好的配置，可以运行下面命令来实现

```
[wyp@iteblog /etc/sysconfig]$ sudo iptables -D INPUT 1
```

其中1是sudo iptables -L -n --line-number命令输出列的num 值。可以参考上面的命令。