

<http://www.blogjava.net/yongboy/archive/2013/04/11/397677.html>

在服务端，连接达到一定数量，诸如50w时，就有些隐藏很深的问题，就会不断出来。通过查看dmesg命令查看，发现大量TCP: too many of orphaned sockets错误，下面到了需要调整tcp socket参数的时候了。

1. 第一个需要调整的是tcp_rmem，即tcp读取缓冲区，单位为字节，查看默认值。

```
[root@izwz92dpr9gpgatl1u1qm8z logs]# cat /proc/sys/net/ipv4/tcp_rmem
4096      87380    6291456
[root@izwz92dpr9gpgatl1u1qm8z logs]#
```

默认值为87380 bytes 约等于 86k，最小为4096 bytes 约等于 4k，最大值为4064k

2. 第二个需要调整的是tcp_wmem，发送缓冲区，单位为字节，默认值

```
[root@izwz92dpr9gpgatl1u1qm8z logs]# cat /proc/sys/net/ipv4/tcp_wmem
4096      16384    4194304
[root@izwz92dpr9gpgatl1u1qm8z logs]#
```

解释同上

第三个需要调整的tcp_mem，调整tcp的内存大小，其单位是页，1页等于4096字节。系统默认值为：

```
[root@izwz92dpr9gpgatl1u1qm8z logs]# cat /proc/sys/net/ipv4/tcp_mem
42525     56702     85050
[root@izwz92dpr9gpgatl1u1qm8z logs]#
```

tcp_mem(3个整型变量)：low, pressure, high

low：当TCP使用了低于该值的内存页面数时，TCP不会考虑释放内存。

pressure：当TCP使用了超过该值的内存页面数量时，TCP试图稳定其内存使用，进入pressure模式，当内存消耗低于low值时推出pressure状态。

high：允许所有的tcp sockets用于排队缓冲数据报的页面量，当内存占用超过此值时，系统拒绝分配socket，后台日志输出“TCP: too many of orphaned sockets”。

一般情况下这些值是在系统启动时根据系统内存数量计算得到的，根据当前tcp_mem最大内存页面数为85050，当内存为 $(85050 * 4) / 1024 = 332\text{M}$ 时，系统将无法为新的socket连接分配内存，即tcp连接将被拒绝。

实际测试环境中，据观察大概在99w个连接左右的时候，进程被杀死，出发out of memory错误。每一个连接大致需要占用7.5k内存(这个数字需要再查)，大致可算得此时内存占用情况为 $990000 * 7.5 / 1024\text{k} = 7251\text{M}$ 。

所以tcp_mem也需要进行调整。

三个调整语句为：

```
echo "net.ipv4.tcp_rmem = 4096 4096 1677216" >> /etc/sysctl.conf
```

```
echo "net.ipv4.tcp_wmem = 4096 4096 1677216" >> /etc/sysctl.conf
```

```
echo "net.ipv4.tcp_mem = 786432 2097152 3145728" >> /etc/sysctl.conf
```

备注：为了节省内存，设置tcp读、写缓冲区都为4k大小，tcp_mem三个值分别为3G、8G、16G，tcp_rmem和tcp_wmem最大值也是16G

实验见上链接