[**每天一个linux命令（47）：iostat命令**](http://www.cnblogs.com/peida/archive/2012/12/28/2837345.html)

Linux系统中的 iostat是I/O statistics（输入/输出统计）的缩写，iostat工具将对系统的磁盘操作活动进行监视。它的特点是汇报磁盘活动统计情况，同时也会汇报出CPU使用情况。同vmstat一样，iostat也有一个弱点，就是它不能对某个进程进行深入分析，仅对系统的整体情况进行分析。iostat属于sysstat软件包。可以用yum install sysstat 直接安装。

**1．命令格式：**

iostat[参数][时间][次数]

**2．命令功能：**

  通过iostat方便查看CPU、网卡、tty设备、磁盘、CD-ROM 等等设备的活动情况, 负载信息。

**3．命令参数：**

-C 显示CPU使用情况

-d 显示磁盘使用情况

-k 以 KB 为单位显示

-m 以 M 为单位显示

-N 显示磁盘阵列(LVM) 信息

-n 显示NFS 使用情况

-p[磁盘] 显示磁盘和分区的情况

-t 显示终端和CPU的信息

-x 显示详细信息

-V 显示版本信息

**4．使用实例：**

**实例1：显示所有设备负载情况**

**命令：**

iostat

**输出：**

[root@CT1186 ~]# iostat  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.30    0.02    5.07    0.17    0.00   86.44  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda              22.73        43.70       487.42  674035705 7517941952  
sda1              0.00         0.00         0.00       2658        536  
sda2              0.11         3.74         3.51   57721595   54202216  
sda3              0.98         0.61        17.51    9454172  270023368  
sda4              0.00         0.00         0.00          6          0  
sda5              6.95         0.12       108.73    1924834 1677123536  
sda6              2.20         0.18        31.22    2837260  481488056  
sda7             12.48        39.04       326.45  602094508 5035104240

**说明：**

**cpu属性值说明：**

%user：CPU处在用户模式下的时间百分比。

%nice：CPU处在带NICE值的用户模式下的时间百分比。

%system：CPU处在系统模式下的时间百分比。

%iowait：CPU等待输入输出完成时间的百分比。

%steal：管理程序维护另一个虚拟处理器时，虚拟CPU的无意识等待时间百分比。

%idle：CPU空闲时间百分比。

**备注：**如果%iowait的值过高，表示硬盘存在I/O瓶颈，%idle值高，表示CPU较空闲，如果%idle值高但系统响应慢时，有可能是CPU等待分配内存，此时应加大内存容量。%idle值如果持续低于10，那么系统的CPU处理能力相对较低，表明系统中最需要解决的资源是CPU。

**disk属性值说明：**

rrqm/s:  每秒进行 merge 的读操作数目。即 rmerge/s

wrqm/s:  每秒进行 merge 的写操作数目。即 wmerge/s

r/s:  每秒完成的读 I/O 设备次数。即 rio/s

w/s:  每秒完成的写 I/O 设备次数。即 wio/s

rsec/s:  每秒读扇区数。即 rsect/s

wsec/s:  每秒写扇区数。即 wsect/s

rkB/s:  每秒读K字节数。是 rsect/s 的一半，因为每扇区大小为512字节。

wkB/s:  每秒写K字节数。是 wsect/s 的一半。

avgrq-sz:  平均每次设备I/O操作的数据大小 (扇区)。

avgqu-sz:  平均I/O队列长度。

await:  平均每次设备I/O操作的等待时间 (毫秒)。

svctm: 平均每次设备I/O操作的服务时间 (毫秒)。

%util:  一秒中有百分之多少的时间用于 I/O 操作，即被io消耗的cpu百分比

**备注：**如果 %util 接近 100%，说明产生的I/O请求太多，I/O系统已经满负荷，该磁盘可能存在瓶颈。如果 svctm 比较接近 await，说明 I/O 几乎没有等待时间；如果 await 远大于 svctm，说明I/O 队列太长，io响应太慢，则需要进行必要优化。如果avgqu-sz比较大，也表示有当量io在等待。

**实例2：定时显示所有信息**

**命令：**

iostat 2 3

**输出：**

[root@CT1186 ~]# iostat 2 3  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.30    0.02    5.07    0.17    0.00   86.44  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda              22.73        43.70       487.42  674035705 7517947296  
sda1              0.00         0.00         0.00       2658        536  
sda2              0.11         3.74         3.51   57721595   54202216  
sda3              0.98         0.61        17.51    9454172  270023608  
sda4              0.00         0.00         0.00          6          0  
sda5              6.95         0.12       108.73    1924834 1677125640  
sda6              2.20         0.18        31.22    2837260  481488152  
sda7             12.48        39.04       326.44  602094508 5035107144  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.88    0.00    7.94    0.19    0.00   83.00  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda               6.00         0.00       124.00          0        248  
sda1              0.00         0.00         0.00          0          0  
sda2              0.00         0.00         0.00          0          0  
sda3              0.00         0.00         0.00          0          0  
sda4              0.00         0.00         0.00          0          0  
sda5              0.00         0.00         0.00          0          0  
sda6              0.00         0.00         0.00          0          0  
sda7              6.00         0.00       124.00          0        248  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           9.12    0.00    7.81    0.00    0.00   83.07  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda               4.00         0.00        84.00          0        168  
sda1              0.00         0.00         0.00          0          0  
sda2              0.00         0.00         0.00          0          0  
sda3              0.00         0.00         0.00          0          0  
sda4              0.00         0.00         0.00          0          0  
sda5              0.00         0.00         0.00          0          0  
sda6              4.00         0.00        84.00          0        168  
sda7              0.00         0.00         0.00          0          0

**说明：**

每隔 2秒刷新显示，且显示3次

**实例3：**显示指定磁盘信息

**命令：**

iostat -d sda1

**输出：**

[root@CT1186 ~]# iostat -d sda1  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda1              0.00         0.00         0.00       2658        536

**说明：**

**实例4：显示tty和Cpu信息**

**命令：**

iostat -t

**输出：**

[root@CT1186 ~]# iostat -t  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
Time: 14时58分35秒  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.30    0.02    5.07    0.17    0.00   86.44  
  
Device:            tps   Blk\_read/s   Blk\_wrtn/s   Blk\_read   Blk\_wrtn  
sda              22.73        43.70       487.41  674035705 7517957864  
sda1              0.00         0.00         0.00       2658        536  
sda2              0.11         3.74         3.51   57721595   54202216  
sda3              0.98         0.61        17.51    9454172  270024344  
sda4              0.00         0.00         0.00          6          0  
sda5              6.95         0.12       108.73    1924834 1677128808  
sda6              2.20         0.18        31.22    2837260  481488712  
sda7             12.48        39.04       326.44  602094508 5035113248

**说明：**

**实例5：以M为单位显示所有信息**

**命令：**

iostat -m

**输出：**

[root@CT1186 ~]# iostat -m  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.30    0.02    5.07    0.17    0.00   86.44  
  
Device:            tps    MB\_read/s    MB\_wrtn/s    MB\_read    MB\_wrtn  
sda              22.72         0.02         0.24     329119    3670881  
sda1              0.00         0.00         0.00          1          0  
sda2              0.11         0.00         0.00      28184      26465  
sda3              0.98         0.00         0.01       4616     131848  
sda4              0.00         0.00         0.00          0          0  
sda5              6.95         0.00         0.05        939     818911  
sda6              2.20         0.00         0.02       1385     235102  
sda7             12.48         0.02         0.16     293991    2458553

**说明：**

**实例6：查看TPS和吞吐量信息**

**命令：**

iostat -d -k 1 1

**输出：**

[root@CT1186 ~]# iostat -d -k 1 1  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
Device:            tps    kB\_read/s    kB\_wrtn/s    kB\_read    kB\_wrtn  
sda              22.72        21.85       243.71  337017916 3758984340  
sda1              0.00         0.00         0.00       1329        268  
sda2              0.11         1.87         1.76   28860797   27101108  
sda3              0.98         0.31         8.75    4727086  135012508  
sda4              0.00         0.00         0.00          3          0  
sda5              6.95         0.06        54.37     962481  838566148  
sda6              2.20         0.09        15.61    1418630  240744712  
sda7             12.48        19.52       163.22  301047254 2517559596

**说明：**

tps：该设备每秒的传输次数（Indicate the number of transfers per second that were issued to the device.）。“一次传输”意思是“一次I/O请求”。多个逻辑请求可能会被合并为“一次I/O请求”。“一次传输”请求的大小是未知的。

kB\_read/s：每秒从设备（drive expressed）读取的数据量；

kB\_wrtn/s：每秒向设备（drive expressed）写入的数据量；

kB\_read：读取的总数据量；kB\_wrtn：写入的总数量数据量；

这些单位都为Kilobytes。

上面的例子中，我们可以看到磁盘sda以及它的各个分区的统计数据，当时统计的磁盘总TPS是22.73，下面是各个分区的TPS。（因为是瞬间值，所以总TPS并不严格等于各个分区TPS的总和）

**实例7：查看设备使用率（%util）、响应时间（await）**

**命令：**

iostat -d -x -k 1 1

**输出：**

[root@CT1186 ~]# iostat -d -x -k 1 1  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
Device:         rrqm/s   wrqm/s   r/s   w/s    rkB/s    wkB/s avgrq-sz avgqu-sz   await  svctm  %util  
sda               0.44    38.59  0.40 22.32    21.85   243.71    23.37     0.04    1.78   4.20   9.54  
sda1              0.00     0.00  0.00  0.00     0.00     0.00    18.90     0.00    8.26   6.46   0.00  
sda2              0.36     0.43  0.11  0.01     1.87     1.76    63.57     0.01   63.75   1.94   0.02  
sda3              0.00     1.24  0.04  0.95     0.31     8.75    18.42     0.04   39.77   8.73   0.86  
sda4              0.00     0.00  0.00  0.00     0.00     0.00     2.00     0.00   19.67  19.67   0.00  
sda5              0.00     6.65  0.00  6.94     0.06    54.37    15.67     0.26   36.81   4.48   3.11  
sda6              0.00     1.71  0.01  2.19     0.09    15.61    14.29     0.03   12.40   5.84   1.28  
sda7              0.08    28.56  0.25 12.24    19.52   163.22    29.28     0.27   21.46   5.00   6.25

**说明：**

rrqm/s：  每秒进行 merge 的读操作数目.即 delta(rmerge)/s

wrqm/s： 每秒进行 merge 的写操作数目.即 delta(wmerge)/s

r/s：  每秒完成的读 I/O 设备次数.即 delta(rio)/s

w/s：  每秒完成的写 I/O 设备次数.即 delta(wio)/s

rsec/s：  每秒读扇区数.即 delta(rsect)/s

wsec/s： 每秒写扇区数.即 delta(wsect)/s

rkB/s：  每秒读K字节数.是 rsect/s 的一半,因为每扇区大小为512字节.(需要计算)

wkB/s：  每秒写K字节数.是 wsect/s 的一半.(需要计算)

avgrq-sz：平均每次设备I/O操作的数据大小 (扇区).delta(rsect+wsect)/delta(rio+wio)

avgqu-sz：平均I/O队列长度.即 delta(aveq)/s/1000 (因为aveq的单位为毫秒).

await：  平均每次设备I/O操作的等待时间 (毫秒).即 delta(ruse+wuse)/delta(rio+wio)

svctm： 平均每次设备I/O操作的服务时间 (毫秒).即 delta(use)/delta(rio+wio)

%util： 一秒中有百分之多少的时间用于 I/O 操作,或者说一秒中有多少时间 I/O 队列是非空的，即 delta(use)/s/1000 (因为use的单位为毫秒)

如果 %util 接近 100%，说明产生的I/O请求太多，I/O系统已经满负荷，该磁盘可能存在瓶颈。

idle小于70% IO压力就较大了，一般读取速度有较多的wait。

同时可以结合vmstat 查看查看b参数(等待资源的进程数)和wa参数(IO等待所占用的CPU时间的百分比，高过30%时IO压力高)。

另外 await 的参数也要多和 svctm 来参考。差的过高就一定有 IO 的问题。

avgqu-sz 也是个做 IO 调优时需要注意的地方，这个就是直接每次操作的数据的大小，如果次数多，但数据拿的小的话，其实 IO 也会很小。如果数据拿的大，才IO 的数据会高。也可以通过 avgqu-sz × ( r/s or w/s ) = rsec/s or wsec/s。也就是讲，读定速度是这个来决定的。

svctm 一般要小于 await (因为同时等待的请求的等待时间被重复计算了)，svctm 的大小一般和磁盘性能有关，CPU/内存的负荷也会对其有影响，请求过多也会间接导致 svctm 的增加。await 的大小一般取决于服务时间(svctm) 以及 I/O 队列的长度和 I/O 请求的发出模式。如果 svctm 比较接近 await，说明 I/O 几乎没有等待时间；如果 await 远大于 svctm，说明 I/O 队列太长，应用得到的响应时间变慢，如果响应时间超过了用户可以容许的范围，这时可以考虑更换更快的磁盘，调整内核 elevator 算法，优化应用，或者升级 CPU。

队列长度(avgqu-sz)也可作为衡量系统 I/O 负荷的指标，但由于 avgqu-sz 是按照单位时间的平均值，所以不能反映瞬间的 I/O 洪水。

**形象的比喻：**  
       r/s+w/s 类似于交款人的总数  
      平均队列长度(avgqu-sz)类似于单位时间里平均排队人的个数  
      平均服务时间(svctm)类似于收银员的收款速度  
      平均等待时间(await)类似于平均每人的等待时间  
      平均I/O数据(avgrq-sz)类似于平均每人所买的东西多少  
       I/O 操作率 (%util)类似于收款台前有人排队的时间比例  
  
       设备IO操作:总IO(io)/s = r/s(读) +w/s(写) =1.46 + 25.28=26.74  
      平均每次设备I/O操作只需要0.36毫秒完成,现在却需要10.57毫秒完成，因为发出的 请求太多(每秒26.74个)，假如请求时同时发出的，可以这样计算平均等待时间:  
      平均等待时间=单个I/O服务器时间\*(1+2+...+请求总数-1)/请求总数   
       每秒发出的I/0请求很多,但是平均队列就4,表示这些请求比较均匀,大部分处理还是比较及时。

**实例8：查看cpu状态**

**命令：**

    iostat -c 1 3

**输出：**

[root@CT1186 ~]#  iostat -c 1 3  
Linux 2.6.18-128.el5 (CT1186)   2012年12月28日  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.30    0.02    5.07    0.17    0.00   86.44  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           8.64    0.00    5.38    0.00    0.00   85.98  
  
avg-cpu:  %user   %nice %system %iowait  %steal   %idle  
           7.62    0.00    5.12    0.50    0.00   86.75

**说明：**