

Yu Feng

原創文章，轉載請註明：轉載自 [系統技術非業餘研究](#)

本文鏈接地址: [qperf測量網絡帶寬和延遲](#)

我們在做網絡服務器的時候，通常會很關心網絡的帶寬和延遲。因為我們的很多協議都是request-reponse協議，延遲決定了最大的QPS，而帶寬決定了最大的負荷。通常我們知道自己的網卡是什麼型號，交換機什麼型號，主機之間的物理距離是多少，理論上是知道帶寬和延遲是不少的。但是現實的情況是，真正的帶寬和延遲情況會有很多變數的，比如說網卡驅動，交換機跳數，丟包率，協議棧配置，光實際速度都很大的影響了數值的估算。所以我們需要找到工具來實際測量下。

網絡測量的工具有很多，netperf什麼的都很不錯。我這裡推薦了qperf，這是RHEL 6發行版裡面自帶的，所以使用起來很方便，只要簡單的：

```
| yum install qperf
```

就好。

我們看下man qperf的介紹：

```
| qperf measures bandwidth and latency between two nodes. It can work over TCP/IP as well as the RDMA transports. On one of the nodes, qperf is typically run with no arguments designating it the server node. One may then run qperf on a client node to obtain measurements such as bandwidth, latency and cpu utilization.
```

```
| In its most basic form, qperf is run on one node in server mode by invoking it with no arguments. On the other node, it is run with two arguments: the name of the server node followed by the name of the test. A list of tests can be found in the section, TESTS. A variety of options may also be specified.
```

使用起來也相當簡單：

在其中一台機器上運行qperf,不帶任何參數就好，這台機器就充當服務器角色：

```
$ uname -r
```

```
2.6.32-131.21.1.tb477.el6.x86_64
```

```
$  
qperf
```

在另外一台機器上運行qperf，測量tcp的帶寬和延時，順便看下雙方機器的配置情況：

```
$ qperf 10.232.64.yyy tcp_bw tcp_lat  
conf
```

```
tcp_bw:
```

```
    bw    = 118  
MB/sec
```

```
tcp_lat:
```

```
    latency = 31.9  
us
```

```
conf:
```

```
    loc_node    =  
xxx.sqa.cm4
```

```
    loc_cpu     = 16 Cores: Intel Xeon L5630 @  
2.13GHz
```

```
    loc_os      = Linux 2.6.32-  
131.21.1.tb477.el6.x86_64
```

```
    loc_qperf   =  
0.4.6
```

```
    rem_node    =  
yyy.sqa.cm4
```

```
    rem_cpu     = 16 Cores: Intel Xeon L5630 @  
2.13GHz
```

```
    rem_os      = Linux 2.6.32-  
131.21.1.tb477.el6.x86_64
```

```
    rem_qperf   =  
0.4.6
```

是不是很方便？典型情況下我們的帶寬是118M，延遲是32us，在標準的千M環境下是符合預期的。

當然qperf有很多高級參數，可以設置socket buffer的大小，綁定CPU親緣性等，很讚的一個特性是可以通過持續改變某個重要參數的值，來觀察臨界點：

Run a test multiple times sequencing through a series of values. Var is the loop variable; Init is the initial value; Last is the value it must not exceed and Incr is the increment. It is useful to set the `-verbose_used (-vu)` option in conjunction with this option.

比如我們可以透過改變消息的大小 (`msg_size`)，比如從1個字節到64K，每次倍增的方式，來觀察帶寬和延遲的變化情況，演示下：

```
$ qperf -oo msg_size:1:64K:*2 10.232.64.yyy tcp_bw  
tcp_lat
```

```
tcp_bw:
```

```
    bw  =  2.43  
MB/sec
```

```
tcp_bw:
```

```
    bw  =  4.69  
MB/sec
```

```
tcp_bw:
```

```
    bw  =  9.12  
MB/sec
```

```
tcp_bw:
```

```
    bw  = 18.5  
MB/sec
```

```
tcp_bw:
```

```
    bw  = 33.1  
MB/sec
```

```
tcp_bw:
```

```
    bw  = 61.4  
MB/sec
```

```
tcp_bw:
```

```
    bw  = 114  
MB/sec
```

tcp_bw:

bw = 118
MB/sec

tcp_bw:

bw = 113
MB/sec

tcp_bw:

bw = 114
MB/sec

tcp_bw:

bw = 114
MB/sec

tcp_bw:

bw = 118
MB/sec

tcp_bw:

bw = 117
MB/sec

tcp_bw:

bw = 118
MB/sec

tcp_bw:

bw = 118
MB/sec

tcp_bw:

```
    bw = 117
MB/sec
```

```
tcp_bw:
```

```
    bw = 117
MB/sec
```

```
tcp_lat:
```

```
    latency = 31
us
```

```
tcp_lat:
```

```
    latency = 31.1
us
```

```
tcp_lat:
```

```
    latency = 31.1
us
```

```
tcp_lat:
```

```
    latency = 31.4
us
```

```
tcp_lat:
```

```
    latency = 30.8
us
```

```
tcp_lat:
```

```
    latency = 32.1
us
```

```
tcp_lat:
```

```
    latency = 32.6
us
```

tcp_lat:

latency = 33.3
us

tcp_lat:

latency = 35.5
us

tcp_lat:

latency = 38.6
us

tcp_lat:

latency = 50.1
us

tcp_lat:

latency = 69.6
us

tcp_lat:

latency = 88
us

tcp_lat:

latency = 128
us

tcp_lat:

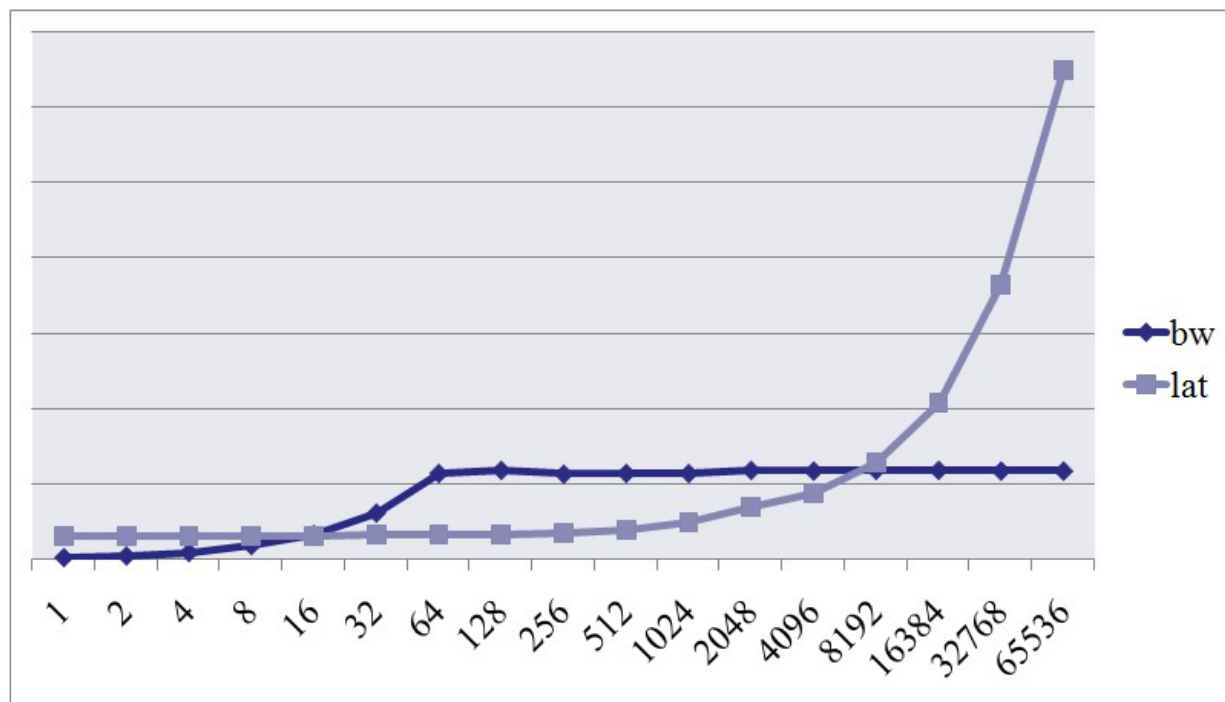
latency = 209
us

tcp_lat:

```
latency = 365  
us
```

```
tcp_lat:
```

```
latency = 650  
us
```



我們可以看到當包的大小達到64字節的時候，帶寬就上不去了；包到達1K的時候，延遲有了很大的變化。這些臨界點對我們的服務器編程時候對性能的估計和預期非常有幫助。

qperf除了測量tcp的，還可以測試rdma, udp, sctp等主流網絡協議的帶寬和延遲，算是個很新的工具，推薦大家使用。

祝玩得開心！

Post Footer automatically generated by [wp-posturl plugin](#) for wordpress.