

Some utilities for ZMQ throughput tests

by Andy Götz

This document describes how to use some utilities for testing the performance of network and disk throughput with ZMQ for the Local Buffer Storage (LBS).

The source code for these tools is derived from the performance examples in the *perf* subdirectory of the zmq source code distribution:

<http://download.zeromq.org/zeromq-3.2.3.tar.gz>

The modified code is available on the local lbs disks (lbs291) in `~goetz/zmq/zeromq-3.2.3/perf` or on my github account:

<https://github.com/andygötz/myzmq/tree/master/perf>

The binaries are in `/tmp_14_days/goetz` for Debian 64 bit and Windows 32 bit.

The two utilities used are `local_thr` and `remote_thr` for measuring network throughput using a PUSH-PULL socket for messages of different sizes. They have been modified to write data to disk in one or more separate thread(s) to measure network + disk performance. Performance is measured and printed out every 1000 transfers in a gnuplot friendly format. The resulting data can also be imported and plotted in a spreadsheet.

Syntax

(1) `remote_thr` :

Linux :

`remote_thr <bind-address> <message size> <number of messages>`

`bind-address` = `tcp://ip-address/socket` where ip-address must be 10 Gbps endpoint

`message-size` = size in bytes e.g. 8000000 for 8 MB

`number-of-messages` = large number for many calls from client(s) e.g. 10000000

Example running on kvm22 talking to lbs291 :

./remote_thr tcp://160.103.197.16:8001 8000000 100000000&

Windows :

remote_thr_hwr <bind-address> <message size> <messages> <high-water-mark>

bind-address = tcp://ip-address/socket where ip-address must be 10 Gbps endpoint

message-size = size in bytes e.g. 8000000 for 8 MB

messages = large number for many calls from client(s) e.g. 10000000

high-water-mark = number of messages to buffer (< 50 on 32 bit machines)

Example running on computertest1 talking to lbs291 :

./remote_thr_hwr tcp://160.103.197.16:8001 8000000 100000000 50&

(2) local_thr :

Linux :

local_thr <bind-address> <message size> <messages> <directory> [<no-of-threads>]

bind-address = tcp://ip-address/socket where ip-address must be 10 Gbps endpoint

message-size = size in bytes e.g. 8000000 for 8 MB

messages = number for messages to transfer e.g. 100000

directory = disk where to store data e.g. /nobackup, /lbsram or /network

if /network the data is not written to file i.e. only measures network performance.

NOTE: the directory **data** must be created beforehand on this disk

no-of-threads = optional number of threads to create for writing (default is 1, maximum 10)

Example running on lbs291 :

./local_thr tcp://160.103.197.16:8001 8000000 1000000 /nobackup

Output :

```
#local_thr local host lbs291 disk /nobackup bind to tcp://160.103.197.16:8004 message size  
8000000 message count 1000000 writer threads 1
```

```
#mean throughput for 1000 msg:
```

```
9 [s] 123 [msg/s] 984.000 [MB/s] 7872.000 [Mb/s]  
16 [s] 140 [msg/s] 1120.000 [MB/s] 8960.000 [Mb/s]  
23 [s] 128 [msg/s] 1024.000 [MB/s] 8192.000 [Mb/s]  
32 [s] 121 [msg/s] 968.000 [MB/s] 7744.000 [Mb/s]  
40 [s] 119 [msg/s] 952.000 [MB/s] 7616.000 [Mb/s]  
47 [s] 140 [msg/s] 1120.000 [MB/s] 8960.000 [Mb/s]  
56 [s] 115 [msg/s] 920.000 [MB/s] 7360.000 [Mb/s]  
64 [s] 117 [msg/s] 936.000 [MB/s] 7488.000 [Mb/s]  
72 [s] 133 [msg/s] 1064.000 [MB/s] 8512.000 [Mb/s]  
79 [s] 133 [msg/s] 1064.000 [MB/s] 8512.000 [Mb/s]  
87 [s] 123 [msg/s] 984.000 [MB/s] 7872.000 [Mb/s]  
96 [s] 120 [msg/s] 960.000 [MB/s] 7680.000 [Mb/s]
```

Examples

Here are some example scripts to run the tests for various buffer sizes:

remote_thr on kvm22 :

```
./remote_thr tcp://160.103.197.16:8001 1000000 1000000000&  
./remote_thr tcp://160.103.197.16:8002 2000000 1000000000&  
./remote_thr tcp://160.103.197.16:8003 4000000 1000000000&  
./remote_thr tcp://160.103.197.16:8004 8000000 1000000000&  
./remote_thr tcp://160.103.197.16:8005 16000000 1000000000&  
./remote_thr tcp://160.103.197.16:8006 32000000 1000000000&  
./remote_thr tcp://160.103.197.16:8007 64000000 1000000000&
```

local_thr on lbs291 (note all files in /nobackup/data **MUST** be removed beforehand to measure the disk performance on a clean directory) :

```
rm /nobackup/data/*
```

```
./local_thr tcp://160.103.197.16:8001 1000000 10000 /nobackup | tee linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8002 2000000 10000 /nobackup | tee -a linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8003 4000000 10000 /nobackup | tee -a linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8004 8000000 10000 /nobackup | tee -a linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8005 16000000 10000 /nobackup | tee -a linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8006 32000000 10000 /nobackup | tee -a linux_nobackup_1t.dat  
rm /nobackup/data/*  
./local_thr tcp://160.103.197.16:8007 64000000 10000 /nobackup | tee -a linux_nobackup_1t.dat
```

Known problems

(1) if the performance is not what you expect (10 Gb/s) then check the network performance using the /network parameter and make sure iperf is giving the expected network performance.