





IHEP connectivity: network throughput tests report

Document Revision 1

Fabio Hernandez

Beijing, China January 6th, 2012

Introduction

This report is a summary of the ongoing test campaign which aims to understand the cause of the low performance observed in the connectivity of IHEP computing centre with CC-IN2P3 in France and potentially with other European sites.

This work is performed under the frame of the France-China Particle Physics Laboratory, FCPPL [1].

Context

As reported in [2], while transferring of data for the TREND experiment from IHEP computing centre to CC-IN2P3 we observed transfer rates of a few MB/sec, which is considered extremely low compared to the theoretical capacity of the shared international link from China to Europe used by these transfers (provided by TEIN3/Orient [3]), which has a bandwidth of 2.5 Gbps.

We therefore decided to conduct some tests to understand the cause of this limitation which severely reduce the data export capability from IHEP to France. This report presents the results of the first campaign of memory-to-memory tests between IHEP and several IN2P3 hosts.

Test conditions

This campaign of tests consisted of performing memory-to-memory transfers from hosts at IHEP computing centre to hosts at IN2P3.

At the IN2P3 side, 3 machines were configured to receive the data using the iperf [4] throughput test tool.

At the IHEP side, 2 machines were used to send the data: one machine in the login farm and another machine near the border router of the site. The data sent by the machine in the login farm has to traverse the LAN while the machine near the border router has not. Our goal was to understand the potential effect of the LAN in these transfers. All the performed tests showed that the same quantitative results are obtained with both machines, so for brevity in this report we present only the results of the tests performed by sending the data from the border machine, which happens to be the same host running the Perfsonar toolkit at IHEP.

In addition, two different set of tests were performed on different days: the first was performed on December 29th 2011 around 12h CST and the second one on January 5th 2012 around 15h CST. Again, in those two set of tests the observed results were similar, so in this report we only include the most recent results.

Three different tests were performed against each of the 3 receiving machines in the IN2P3 network:

- iperf test using TCP
- iperf test using UDP

traceroute

The IHEP machines used for these tests were all using iperf version 2.0.5.

In the following section we present a summarized version of the results and in the annexes we include all the details for each one of them.

Test results

Sender host: perfsonar.ihep.ac.cn (202.122.32.166)

Operating system:

uname -a

Linux perfsonar.ihep.ac.cn 2.6.18-274.7.1.el5.web100 #1 SMP Wed Nov 2 08:35:05 PDT 2011 i686 i686 i386 GNU/Linux

Destination Host	iperf TCP	iperf UDP
lallhcone02.in2p3.fr	connexion refused	N/A
cclhcone01.in2p3.fr	2.81 Mbits/sec	302 Mbits/sec
		WARNING: did not receive ack of
		last datagram after 10 tries.
ccxfert02.in2p3.fr	2.50 Mbits/sec	302 Mbits/sec
		WARNING: did not receive ack of
		last datagram after 10 tries.

As shown in the table above, the receiving host lallhcone02.in2p3.fr does not allow the sending machine to connect to port number 5001, so the transfers to this target machine could really be tested. In addition, we are not sure we can draw conclusions from the results of the UDP tests as iperf sender reported it did not receive acknowledge of the packets sent.

The throughput by these memory-to-memory tests which is less that 3 Mbits/sec, is consistent with the low performance observed while transferring TREND data from IHEP to the iRods disk servers at CC-IN2P3. There is therefore some intrinsic limitation in the connectivity that we need to understand and that is the motivation of this work.

An additional test was performed from IHEP to a host in University of Minnesota (USA) on January 6th around 14h CST. The observed throughput using iperf over TCP is 557 Mbits/sec. Although this result is not directly comparable with the ones presented above because of the different network paths used to transport the data from IHEP to the destination, it is interesting to see that there seems not to be an internal limitation at the sending site that would explain the low throughput with the IN2P3 sites. The details of this test are also included in the annexes.

Perspectives

In order to understand whether the throughput limitation is observed while transferring data from IHEP to other European sites, we would like to perform equivalent throughput tests using the existing Perfsonar infrastructure deployed at all the WLCG tier-1s. Unfortunately, in the current situation, the only European tier-1 allowing IHEP machines to perform tests against it is CC-IN2P3. We therefore intend to ask other tier-1s to reconfigure their Perfsonar setup so that we can perform the desired tests. We will report the results in a timely manner.

Conclusion

The only conclusion we can draw so far is that there is indeed some limitation in the IHEP connectivity to IN2P3 hosts as shown by the presented results. We need to perform additional tests to understand what is the root cause of this issue. The next step is to perform the same tests with other European sites, in particular the WLCG tier-1s.

Acknowledgements

Many thanks to Jérôme Bernier on the IN2P3 side and Fazhi Qi on the IHEP site for setting up the environment required to perform these tests. Their intervention has greatly contributed to make this work easier.

Annex 1 - detailed iperf results

IN2P3 Hosts

```
lallhcone01.in2p3.fr (134.158.195.1) -- TCP
iperf -c lallhcone01.in2p3.fr -p 5001 -w16M -i5 -t60
connect failed: Connection refused
lallhcone01.in2p3.fr (134.158.195.1) -- UDP
iperf -c lallhcone01.in2p3.fr -p 5001 -w16M -i5 -t60 -u -b300M
_____
Client connecting to lallhcone01.in2p3.fr, UDP port 5001
Sending 1470 byte datagrams
UDP buffer size: 32.0 MByte (WARNING: requested 16.0 MByte)
_____
[ 3] local 202.122.32.166 port 58728 connected with 134.158.195.1 port 5001
[ ID] Interval
                  Transfer
                             Bandwidth
[ 3] 0.0-5.0 sec 24.2 GBytes 41.5 Gbits/sec
[ 3] 5.0-10.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 10.0-15.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 15.0-20.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 20.0-25.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 25.0-30.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 30.0-35.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 35.0-40.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 40.0-45.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 45.0-50.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 50.0-55.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 55.0-60.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 0.0-60.0 sec 246 GBytes 35.2 Gbits/sec
[ 3] Sent 1538421 datagrams
read failed: Connection refused
[ 3] WARNING: did not receive ack of last datagram after 2 tries.
cclhcone01.in2p3.fr (193.48.100.201) -- TCP
iperf -c cclhcone01.in2p3.fr -p 5001 -w16M -i5 -t60
_____
Client connecting to cclhcone01.in2p3.fr, TCP port 5001
TCP window size: 32.0 MByte (WARNING: requested 16.0 MByte)
_____
[ 3] local 202.122.32.166 port 45520 connected with 193.48.100.201 port 5001
[ ID] Interval
                  Transfer
                             Bandwidth
[ 3] 0.0-5.0 sec 23.6 MBytes 39.6 Mbits/sec
[ 3] 5.0-10.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 10.0-15.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 15.0-20.0 sec 0.00 Bytes 0.00 bits/sec
```

```
[ 3] 20.0-25.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 25.0-30.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 30.0-35.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 35.0-40.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 40.0-45.0 sec 7.88 MBytes 13.2 Mbits/sec
[ 3] 45.0-50.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 50.0-55.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 55.0-60.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 60.0-65.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 65.0-70.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 70.0-75.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 75.0-80.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 80.0-85.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 85.0-90.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 0.0-94.5 sec 31.6 MBytes 2.81 Mbits/sec
cclhcone01.in2p3.fr (193.48.100.201) -- UDP
iperf -c cclhcone01.in2p3.fr -p 5001 -w16M -i5 -t60 -u -b300M
_____
Client connecting to cclhcone01.in2p3.fr, UDP port 5001
Sending 1470 byte datagrams
UDP buffer size: 32.0 MByte (WARNING: requested 16.0 MByte)
_____
```

[3] local 202.122.32.166 port 36450 connected with 193.48.100.201 port 5001

```
[ ID] Interval
              Transfer
                              Bandwidth
[ 3] 0.0-5.0 sec 180 MBytes 302 Mbits/sec
[ 3] 5.0-10.0 sec 180 MBytes 302 Mbits/sec
[ 3] 10.0-15.0 sec 180 MBytes 302 Mbits/sec
[ 3] 15.0-20.0 sec 180 MBytes 302 Mbits/sec
[ 3] 20.0-25.0 sec 180 MBytes 302 Mbits/sec
[ 3] 25.0-30.0 sec 180 MBytes 302 Mbits/sec
[ 3] 30.0-35.0 sec 180 MBytes 302 Mbits/sec
```

- [3] 35.0-40.0 sec 180 MBytes 302 Mbits/sec
- [3] 40.0-45.0 sec 180 MBytes 302 Mbits/sec
- [3] 45.0-50.0 sec 180 MBytes 302 Mbits/sec
- [3] 50.0-55.0 sec 180 MBytes 302 Mbits/sec
- [3] 0.0-60.0 sec 2.11 GBytes 302 Mbits/sec
- [3] Sent 1538455 datagrams
- [3] WARNING: did not receive ack of last datagram after 10 tries.

ccxfert02.in2p3.fr (193.48.99.202) -- TCP

```
iperf -c ccxfert02.in2p3.fr -p 5001 -w16M -i5 -t60
_____
Client connecting to ccxfert02.in2p3.fr, TCP port 5001
TCP window size: 32.0 MByte (WARNING: requested 16.0 MByte)
```

```
[ 3] local 202.122.32.166 port 56297 connected with 193.48.99.202 port 5001
[ ID] Interval
                  Transfer
                              Bandwidth
[ 3] 0.0-5.0 sec 23.6 MBytes 39.6 Mbits/sec
[ 3] 5.0-10.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 10.0-15.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 15.0-20.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 20.0-25.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 25.0-30.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 30.0-35.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 35.0-40.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 40.0-45.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 45.0-50.0 sec 7.88 MBytes 13.2 Mbits/sec
[ 3] 50.0-55.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 55.0-60.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 60.0-65.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 65.0-70.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 70.0-75.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 75.0-80.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 80.0-85.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 85.0-90.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 90.0-95.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 95.0-100.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 100.0-105.0 sec 0.00 Bytes 0.00 bits/sec
[ 3] 0.0-106.0 sec 31.6 MBytes 2.50 Mbits/sec
ccxfert02.in2p3.fr (193.48.99.202) -- UDP
iperf -c ccxfert02.in2p3.fr -p 5001 -w16M -i5 -t60 -u -b300M
_____
Client connecting to ccxfert02.in2p3.fr, UDP port 5001
Sending 1470 byte datagrams
UDP buffer size: 32.0 MByte (WARNING: requested 16.0 MByte)
-----
[ 3] local 202.122.32.166 port 47416 connected with 193.48.99.202 port 5001
[ ID] Interval
                  Transfer
                              Bandwidth
[ 3] 0.0-5.0 sec 180 MBytes 301 Mbits/sec
[ 3] 5.0-10.0 sec
                   180 MBvtes
                               301 Mbits/sec
[ 3] 10.0-15.0 sec
                   180 MBytes 302 Mbits/sec
[ 3] 15.0-20.0 sec
                   180 MBytes 302 Mbits/sec
[ 3] 20.0-25.0 sec
                   180 MBytes 302 Mbits/sec
[ 3] 25.0-30.0 sec
                               302 Mbits/sec
                   180 MBytes
```

180 MBytes 302 Mbits/sec

180 MBytes 302 Mbits/sec

180 MBytes 302 Mbits/sec

302 Mbits/sec

302 Mbits/sec

302 Mbits/sec

180 MBytes

180 MBytes

180 MBytes

[3] 30.0-35.0 sec

[3] 35.0-40.0 sec

[3] 40.0-45.0 sec

[3] 45.0-50.0 sec

[3] 50.0-55.0 sec

[3] 55.0-60.0 sec

```
[ 3] 0.0-60.0 sec 2.11 GBytes 302 Mbits/sec
```

- [3] Sent 1538449 datagrams
- [3] WARNING: did not receive ack of last datagram after 10 tries.

```
University of Minnesota
ssb-dtv1.nts.umn.edu (134.84.88.5) -- TCP
iperf -c ssb-dtv1.nts.umn.edu -p 5001 -w16M -i5 -t60
_____
Client connecting to ssb-dtv1.nts.umn.edu, TCP port 5001
TCP window size: 32.0 MByte (WARNING: requested 16.0 MByte)
-----
[ 3] local 202.122.32.166 port 59411 connected with 134.84.88.5 port 5001
[ ID] Interval
                 Transfer
                            Bandwidth
[ 3] 0.0-5.0 sec 23.6 MBytes 39.6 Mbits/sec
[ 3] 5.0-10.0 sec 256 MBytes 429 Mbits/sec
[ 3] 10.0-15.0 sec 332 MBytes 557 Mbits/sec
[ 3] 15.0-20.0 sec 389 MBytes 653 Mbits/sec
[ 3] 20.0-25.0 sec 380 MBytes 638 Mbits/sec
[ 3] 25.0-30.0 sec 377 MBytes 633 Mbits/sec
[ 3] 30.0-35.0 sec 387 MBytes 649 Mbits/sec
[ 3] 35.0-40.0 sec 335 MBytes 562 Mbits/sec
[ 3] 40.0-45.0 sec 377 MBytes 632 Mbits/sec
[ 3] 45.0-50.0 sec 377 MBytes 633 Mbits/sec
[ 3] 50.0-55.0 sec 373 MBytes 626 Mbits/sec
[ 3] 55.0-60.0 sec 379 MBytes 635 Mbits/sec
[ 3] 0.0-60.0 sec 3.89 GBytes 557 Mbits/sec
ssb-dtv1.nts.umn.edu (134.84.88.5) -- UDP
iperf -c ssb-dtv1.nts.umn.edu -p 5001 -w16M -i5 -t60 -u -b300M
______
Client connecting to ssb-dtv1.nts.umn.edu, UDP port 5001
Sending 1470 byte datagrams
UDP buffer size: 32.0 MByte (WARNING: requested 16.0 MByte)
_____
[ 3] local 202.122.32.166 port 54341 connected with 134.84.88.5 port 5001

    IDl Interval

                 Transfer
                            Bandwidth
[ 3] 0.0-5.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 5.0-10.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 10.0-15.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 15.0-20.0 sec 20.2 GBytes 34.7 Gbits/sec
[ 3] 20.0-25.0 sec 20.2 GBytes 34.7 Gbits/sec
```

[3] 25.0-30.0 sec 20.2 GBytes 34.7 Gbits/sec
[3] 30.0-35.0 sec 20.2 GBytes 34.7 Gbits/sec
[3] 35.0-40.0 sec 20.2 GBytes 34.7 Gbits/sec

- [3] 40.0-45.0 sec 20.2 GBytes 34.7 Gbits/sec
- [3] 45.0-50.0 sec 20.2 GBytes 34.7 Gbits/sec
- [3] 50.0-55.0 sec 20.2 GBytes 34.7 Gbits/sec
- [3] 55.0-60.0 sec 20.2 GBytes 34.7 Gbits/sec
- [3] 0.0-60.0 sec 242 GBytes 34.7 Gbits/sec
- [3] Sent 1538449 datagrams

read failed: Connection refused

[3] WARNING: did not receive ack of last datagram after 1 tries.

Annex 2 - detailed traceroute results

IN2P3 Hosts

1. From perfsonar.ihep.ac.cn (202.122.32.166) to **lallhcone01.in2p3.fr** (134.158.195.1):

```
traceroute lallhcone01.in2p3.fr
```

```
traceroute to lallhcone01.in2p3.fr (134.158.195.1), 30 hops max, 40 byte packets

1 * * * *

2 192.168.1.25 (192.168.1.25)  0.337 ms  0.356 ms  0.377 ms

3 (159.226.253.73)  21.625 ms  21.642 ms  21.627 ms

4 (159.226.253.42)  0.618 ms  0.594 ms  0.586 ms

5 210.25.189.65 (210.25.189.65)  2.288 ms  3.778 ms  3.064 ms

6 * * *

7 210.25.189.54 (210.25.189.54)  177.897 ms  178.048 ms  177.761 ms

8 as0.rt1.par.fr.geant2.net (62.40.112.105)  185.348 ms  185.148 ms  185.136 ms

9 renater-gw.rt1.par.fr.geant.net (62.40.124.70)  190.011 ms  187.594 ms  187.566 ms

10 * * *

11 in2p3-lal-vl133-orsay-rtr-021.noc.renater.fr (193.51.188.161)  186.005 ms  185.988 ms  185.972 ms

12 lallhcone01.in2p3.fr (134.158.195.1)  185.936 ms  185.945 ms  185.914 ms
```

2. From perfsonar.ihep.ac.cn (202.122.32.166) to **cclhcone01.in2p3.fr** (193.48.100.201):

traceroute cclhcone01.in2p3.fr

```
traceroute to cclhcone01.in2p3.fr (193.48.100.201), 30 hops max, 40 byte packets

1 ppp1.ihep.ac.cn (202.122.32.161) 0.590 ms 0.751 ms 0.935 ms

2 192.168.1.25 (192.168.1.25) 76.391 ms 76.377 ms 76.387 ms

3 (159.226.253.73) 8.447 ms 8.441 ms 8.418 ms

4 (159.226.253.42) 0.659 ms 0.656 ms 0.639 ms

5 210.25.189.65 (210.25.189.65) 4.446 ms 3.626 ms 2.826 ms

6 * * *

7 210.25.189.54 (210.25.189.54) 177.861 ms 177.960 ms 177.943 ms

8 as0.rt1.par.fr.geant2.net (62.40.112.105) 185.447 ms 185.256 ms 185.148 ms

9 renater-gw.rt1.par.fr.geant.net (62.40.124.70) 189.591 ms 190.656 ms 190.646 ms

10 * * *

11 in2p3-lyon-vl3114-paris1-rtr-021.noc.renater.fr (193.51.186.177) 190.858 ms

12 cclhcone01.in2p3.fr (193.48.100.201) 190.753 ms 190.583 ms !X 190.754 ms
```

3. From perfsonar.ihep.ac.cn (202.122.32.166) to ccxfert02.in2p3.fr (193.48.99.202):

traceroute ccxfert02.in2p3.fr

```
traceroute to ccxfert02.in2p3.fr (193.48.99.202), 30 hops max, 40 byte packets 1 ppp1.ihep.ac.cn (202.122.32.161) 0.566 ms 0.759 ms 0.945 ms
```

- 2 192.168.1.25 (192.168.1.25) 0.339 ms 0.334 ms 0.314 ms
 3 (159.226.253.77) 0.361 ms 0.383 ms 0.374 ms
 4 (159.226.253.46) 0.722 ms 0.674 ms 0.647 ms
 5 210.25.189.65 (210.25.189.65) 3.901 ms 2.291 ms 3.118 ms
 6 * * *
 7 210.25.189.54 (210.25.189.54) 177.898 ms 177.803 ms 177.871 ms
 8 as0.rt1.par.fr.geant2.net (62.40.112.105) 185.274 ms 185.430 ms 185.138 ms
 9 renater-gw.rt1.par.fr.geant.net (62.40.124.70) 189.742 ms 189.721 ms 189.712 ms
 10 * * *
- 11 in2p3-lyon-vl3114-paris1-rtr-021.noc.renater.fr (193.51.186.177) 190.686 ms 190.721 ms 190.806 ms
- 12 lyon-opn.in2p3.fr (134.158.224.2) 190.795 ms 190.788 ms 190.831 ms
- 13 ccxfert02.in2p3.fr (193.48.99.202) 191.024 ms 190.888 ms 191.017 ms

University of Minnesota

From perfsonar.ihep.ac.cn (202.122.32.166) to **ssb-dtv1.nts.umn.edu** (134.84.88.5):

traceroute ssb-dtv1.nts.umn.edu

traceroute to ssb-dtv1.nts.umn.edu (134.84.88.5), 30 hops max, 40 byte packets

- 1 ppp1.ihep.ac.cn (202.122.32.161) 0.551 ms 0.712 ms 0.920 ms
- 2 192.168.1.25 (192.168.1.25) 0.365 ms 0.357 ms 0.369 ms
- 3 (159.226.253.77) 0.354 ms 0.369 ms 0.355 ms
- 4 (159.226.253.54) 0.735 ms 0.721 ms 0.710 ms
- 5 8.192 (159.226.254.254) 38.216 ms 38.196 ms 38.184 ms
- 6 8.199 (159.226.254.214) 38.285 ms 38.065 ms 38.072 ms
- 7 (192.31.99.221) 179.679 ms 179.671 ms 178.954 ms
- 8 abilene-1-lo-jmb-706.sttlwa.pacificwave.net (207.231.240.8) 177.914 ms 177.807 ms 177.823 ms
- 9 xe-0-1-3.0.rtr.salt.net.internet2.edu (64.57.28.104) 194.337 ms 194.339 ms 194.312 ms
- 10 xe-0-1-1.0.rtr.kans.net.internet2.edu (64.57.28.25) 218.972 ms 218.967 ms
- 218.952 ms
- 11 telecomb-gr-01-vl-430.northernlights.gigapop.net (146.57.253.57) 227.563 ms
- 227.553 ms 227.616 ms
- $12 \quad telecomb-br-01-te-4-3. northernlights. gigapop.net \ (146.57.252.177) \quad 227.871 \ ms$
- 227.699 ms 227.731 ms
- 13 telecomb-bn-01-v3710.ggnet.umn.edu (128.101.58.145) 227.700 ms 227.680 ms
- 227.612 ms
- 14 ssb-dtv1.nts.umn.edu (134.84.88.5) 227.603 ms 227.632 ms 227.682 ms

References

- 4. France China Particle Physics Laboratory: http://fcppl.in2p3.fr
- 5. http://indico.in2p3.fr/conferenceDisplay.py?confld=6204
- 6. TEIN3: http://www.tein3.net
- 7. iperf: http://en.wikipedia.org/wiki/lperf