Data Transfer Observations

TREND use-case in the framework of FCPPL project

Fabio Hernandez

fabio@in2p3.fr

December 12th, 2011

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
MINISTÈRE DES AFFAIRES ÉTRANGÈRES ET EUROPÉENNES

中國科學院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

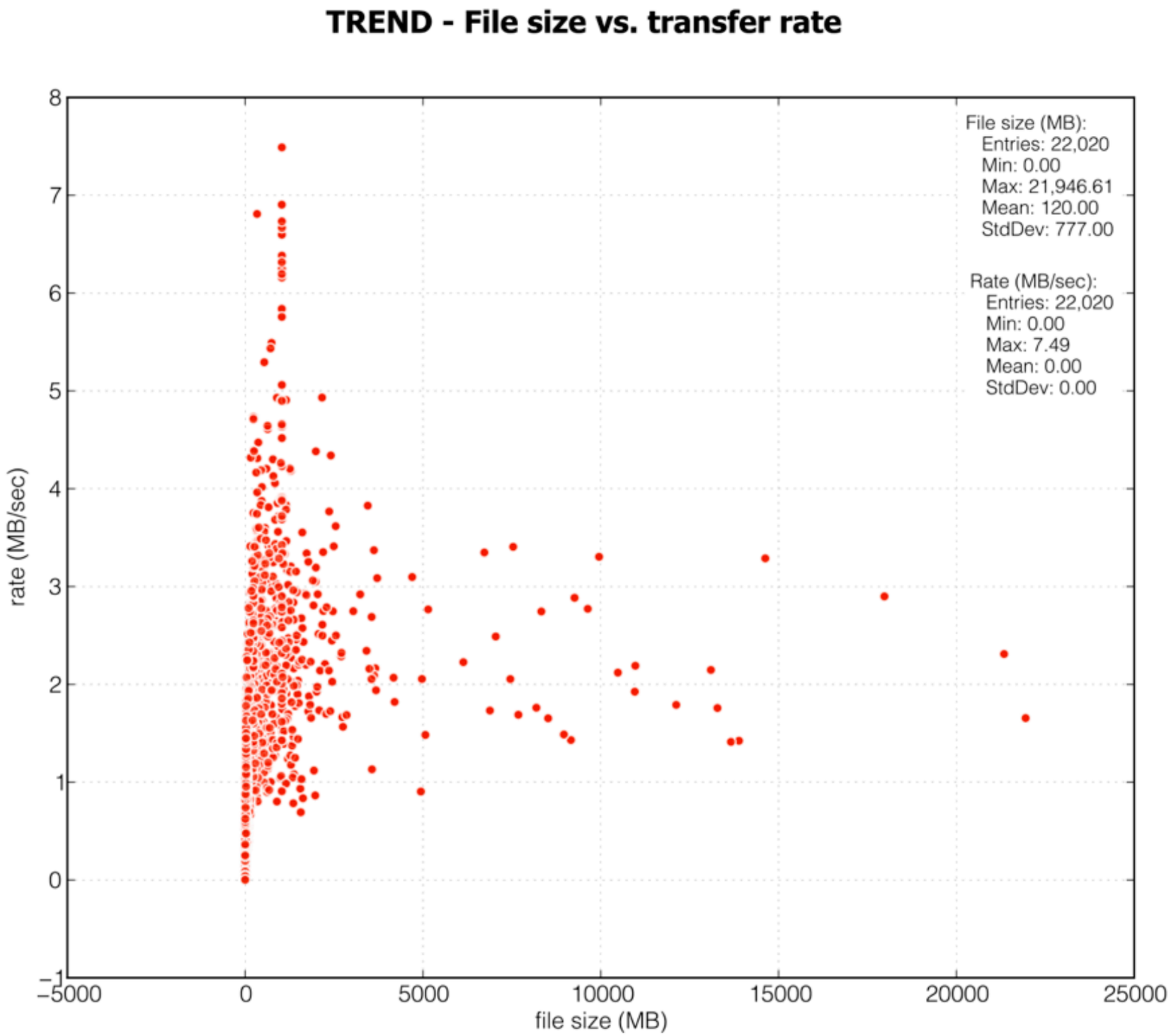
CCIN2P3

Context and Motivation

- CC-IN2P3 is the data repository for TREND
- Raw data is imported by removable disk to IHEP-CC, temporally stored locally and sent to iRods at CC-IN2P3 through the network
- Most recent transfer campaign: early December
- Observed low transfer rate

  this presentation shows the observations and the investigation so far

Observations



Observed transfer rate for 22.000+ TREND files early December 2011

Sender hosts:
lxslc502, lxslc503, lxslc507 (IHEP-CC login farm)

Receiver host: ccirods.in2p3.fr

Remarks:
- (very) small files
- surprisingly low transfer rate

Goals

- Investigate what is the realistically usable network bandwidth from IHEP-CC to CC-IN2P3
- Investigate if iRods client commands can fully exploit the available bandwidth

Progress so far

- First step: how much network bandwidth can we use?

  memory-to-memory, from a machine in the login farm at IHEP to a machine at CC-IN2P3 (not iRods server)

- Tools: iperf v2.0.5

  Server side (CC-IN2P3):

```
iperf --server --port 5001 --len 16M --window 16M
```

  Client side (IHEP):

```
iperf --client ccxfert02.in2p3.fr --port 5001 --len 32M --window 32M \
      --format M --time 120
```

iperf results

- Tests performed on Monday 12/12/2011 (i.e. today), morning Beijing time

```
iperf  --client ccxfert02.in2p3.fr --port 5001 --len 32M --window 32M
       --format M --time 180 --print_mss --nodelay
------------------------------------------------------------
Client connecting to ccxfert02.in2p3.fr, TCP port 5001
TCP window size: 32.0 MByte
------------------------------------------------------------
[  3] local 202.122.33.174 port 44454 connected with 193.48.99.202 port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0-199.4 sec   160 MBytes  0.80 MBytes/sec
[  3] MSS size 1460 bytes (MTU 1500 bytes, ethernet)
```

Transfer data during 180 seconds using a wide window size.

Observed rate: 0.8 MB/sec

Route

- Route from `lxslc507.ihep.ac.cn` to `ccxfert02.in2p3.fr`

```
traceroute to 193.48.99.202 (193.48.99.202), 30 hops max, 40 byte packets

 1  202.122.33.1 (202.122.33.1)  6.559 ms  6.550 ms  6.573 ms
 2  ntclient4.ihep.ac.cn (202.38.128.184)  0.241 ms  0.246 ms  0.198 ms
 3  cisco3640.ihep.ac.cn (202.122.32.130)  1.859 ms  2.019 ms  2.181 ms
 4  192.168.47.1 (192.168.47.1)  0.728 ms  0.721 ms  0.796 ms
 5  8.201 (159.226.254.221)  0.793 ms  0.842 ms  0.903 ms
 6  8.194 (159.226.254.194)  0.640 ms  0.631 ms  0.665 ms
 7  210.25.189.65 (210.25.189.65)  4.482 ms  4.698 ms  6.514 ms
 8  210.25.189.18 (210.25.189.18)  129.296 ms  139.688 ms  139.790 ms
 9  210.25.189.46 (210.25.189.46)  181.272 ms  181.280 ms  181.331 ms
10  so-2-0-0.rt1.ams.nl.geant2.net (62.40.112.78)  194.146 ms  194.081 ms  194.037 ms
11  as1.rt1.lon.uk.geant2.net (62.40.112.138)  202.308 ms  202.174 ms  202.284 ms
12  as0.rt1.par.fr.geant2.net (62.40.112.105)  209.650 ms  209.616 ms  209.743 ms
13  renater-gw.rt1.par.fr.geant.net (62.40.124.70)  214.677 ms  210.774 ms  212.693 ms
14  * * *

15  in2p3-lyon-vl3114-paris1-rtr-021.noc.renater.fr (193.51.186.177)  215.127 ms  215.135 ms  215.031 ms

16  Lyon-OPN.in2p3.fr (134.158.224.2)  215.273 ms  215.295 ms  215.355 ms

17  ccxfert02.in2p3.fr (193.48.99.202)  214.993 ms  214.984 ms  214.973 ms
```

Round-trip time

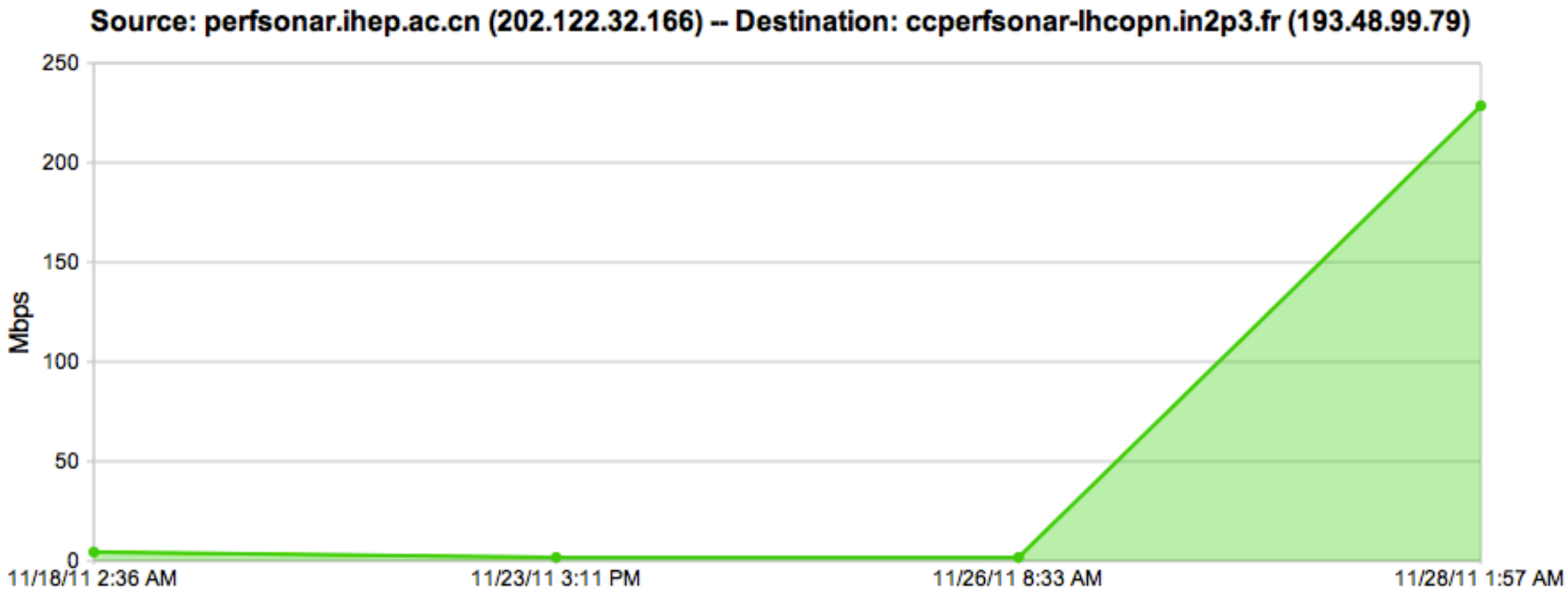
- PING from from lxslc507.ihep.ac.cn to ccxfert02.in2p3.fr

```
ping  ccxfert02.in2p3.fr

PING ccxfert02.in2p3.fr (193.48.99.202) 56(84) bytes of data.
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=1 ttl=48 time=215 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=2 ttl=48 time=214 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=3 ttl=48 time=214 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=4 ttl=48 time=214 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=5 ttl=48 time=215 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=6 ttl=48 time=215 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=7 ttl=48 time=215 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=8 ttl=48 time=214 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=9 ttl=48 time=214 ms
64 bytes from ccxfert02.in2p3.fr (193.48.99.202): icmp_seq=10 ttl=48 time=215 ms

--- ccxfert02.in2p3.fr ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9001ms
rtt min/avg/max/mdev = 214.886/214.984/215.037/0.550 ms
```

Perfsonar



Maximum **perfsonar.ihep.ac.cn -> ccperfsonar-lhcopn.in2p3.fr**	228.83 Mbps
Average **perfsonar.ihep.ac.cn -> ccperfsonar-lhcopn.in2p3.fr**	59.42 Mbps
Last **perfsonar.ihep.ac.cn -> ccperfsonar-lhcopn.in2p3.fr**	228.83 Mbps

Source: http://perfsonar.ihep.ac.cn

What is next?

- Are the machines in IHEP login farm the best choice for doing data transfers over wide-area network?
- Are the iRods client commands able to fully exploit the available bandwidth?

  This remains to be tested

Questions & Comments

fabio hernandez
fabio@in2p3.fr
CAS/IHEP Computing Centre
CNRS/IN2P3 Computing Centre