



BalticGrid-II TCP Performance tests

www.balticgrid.org

Project info:

BalticGrid project Second Phase started on 1 May 2008 to increase the impact, adoption and reach of the Grid technologies, and to further improve the support of services and users of the recently created e-Infrastructure in the Baltic States and Belarus.

BalticGrid-II:

- 7 countries
- 13 partners
- 23 sites, 5266 processors
- 127 TB storage



Fig. 1 - BalticGrid-II e-Infrastructure

Problem:

Many grid applications require transfer of large data volumes. Although most partner sites are connected to GÉANT2 with at least 1Gbps links, in most cases a single TCP session is not able to transfer data at a rate higher than about 50-120 Mbps. Although high throughput can be achieved using UDP streams or multiple TCP connections (e.g. GridFTP, FDT), not all applications are able to natively use these features.

Measurements and Results:

During the first year of the project there have been three major achievements:

1. Reliable gigabit test-lab for TCP performance measurement at variable RTT created

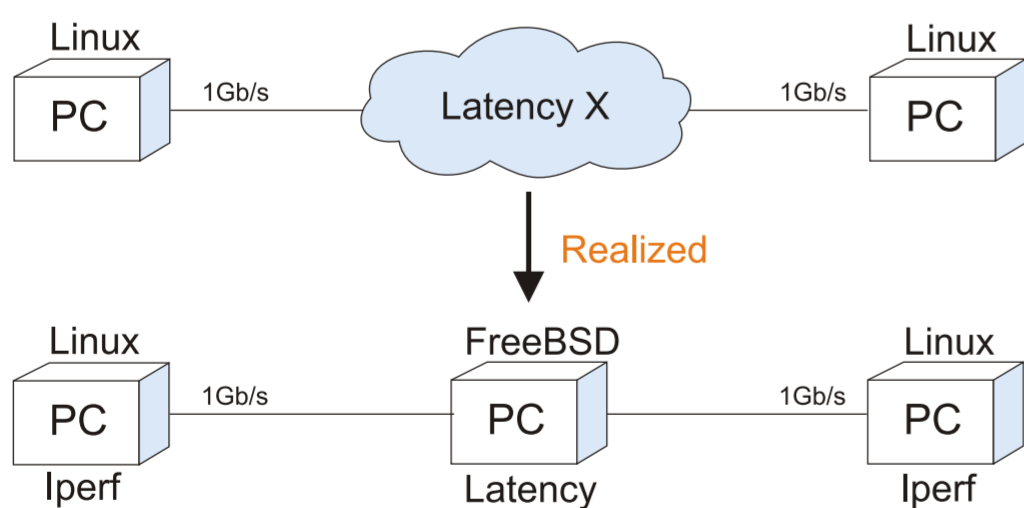


Fig. 2 - TCP performance test lab

We have found out that FreeBSD firewall ('ipfw') configured as Layer 2 bridge can reliably simulate latency at 1Gbps rates.

To investigate TCP/IP performance it was necessary to create a test laboratory. The laboratory has to provide possibility to simulate network latencies and optionally – jitter and packet loss. The created laboratory consists of three servers, one of them is used as a middle ethernet traffic interceptor, and its task is to introduce latency and packet loss. The remaining two servers are connected to the middle server and TCP performance is measured between them.

Contact us:

Project Director:
Ake Edlund, KTH
e-mail: edlund@nada.kth.se

Network Manager:
Guntis Barzdins, IMCS UL
e-mail: guntis@sigmanet.lv

2. TCP performance variations for different versions of Linux kernels measured

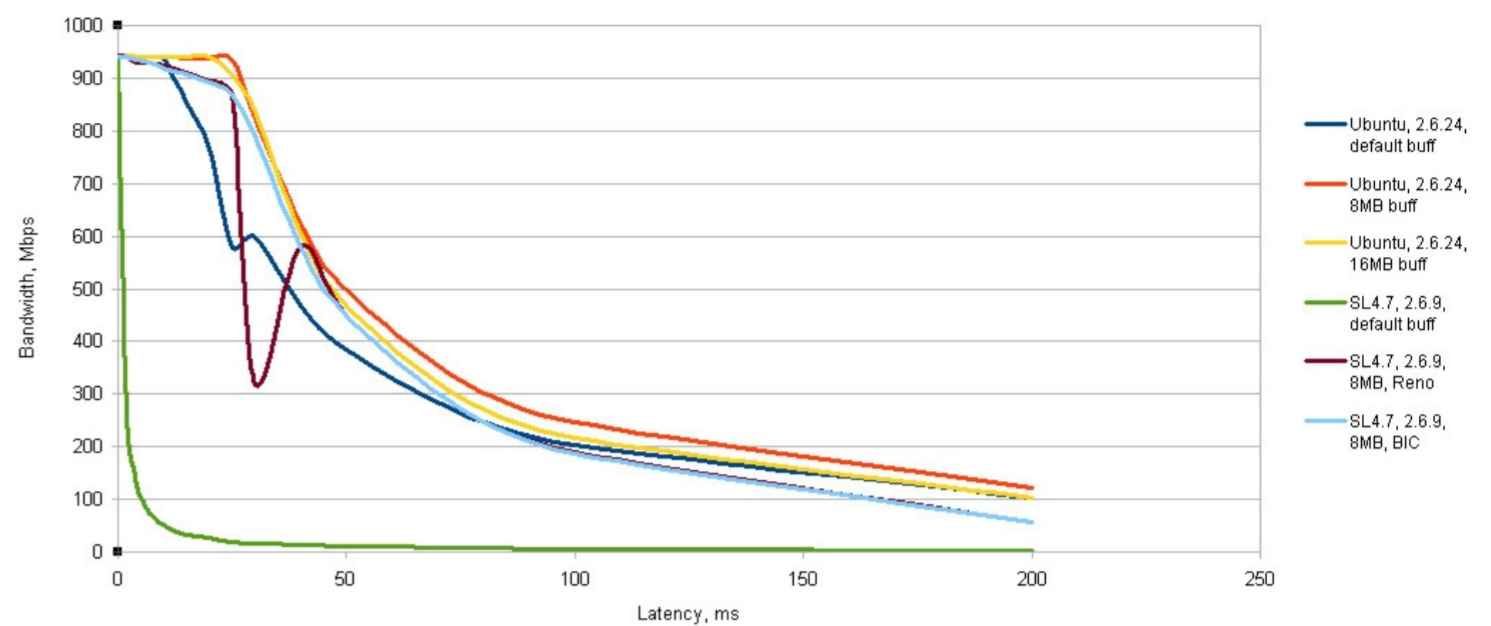


Fig. 3 - TCP bandwidth depending on latency (RTT)

After performing tests it was found that Scientific Linux 4, with default kernel and TCP/IP stack configuration performs very poorly at latencies above 1ms. Results were compared with other Linux kernels stack parameters and TCP/IP. The graph shows TCP bandwidth for various Linux kernels, latencies (RTT) and TCP send/receive buffer sizes. As it can be seen, default Scientific Linux 4 installation under-performs at latencies greater than 1-2ms. Typical network latency between sites in BalticGrid-II is 8-35ms.

3. Solution for TCP performance improvement proposed

It was traced down that as few as 4 configuration lines for TCP stack inserted in system configuration file resolves the TCP performance bottleneck up to gigabit speeds.

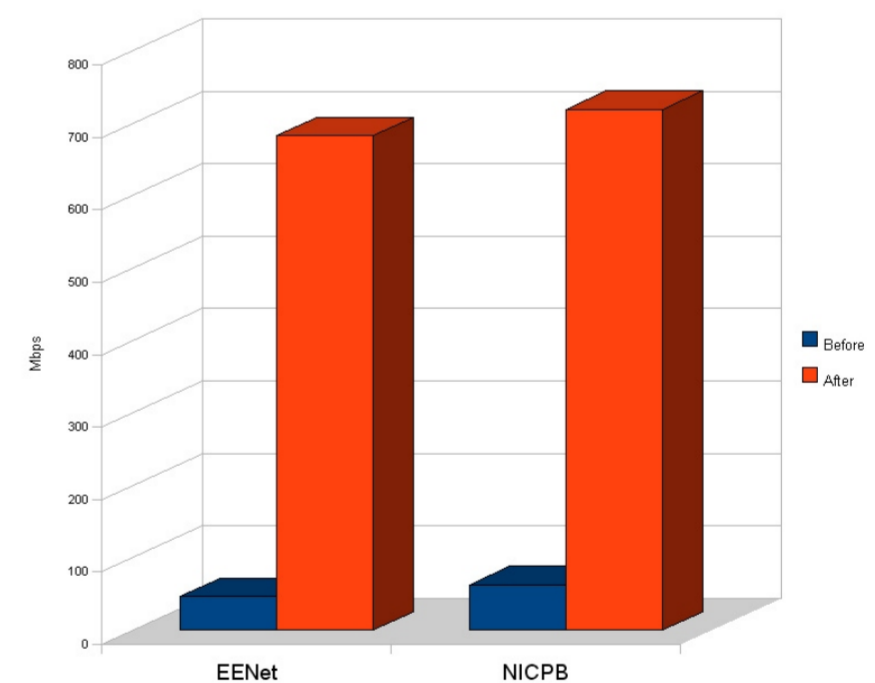


Fig. 4 - TCP performance improvement

```

In /etc/sysctl.conf:
#TCP stack tuning
net/core/rmem_max = 8738000
net/core/wmem_max = 8738000

net/ipv4/tcp_rmem = 8192 4369000 8738000
net/ipv4/tcp_wmem = 8192 4369000 8738000
  
```

Fig. 5 - The 4 configuration lines

After tuning basic TCP/IP stack parameters, Scientific Linux 4 with default Linux kernel 2.6.9 and BIC or Reno congestion control algorithm performs up to **10-20 times faster** in latency range of 6-40ms, the important latency range for the BalticGrid-II applications.

