Nordic Testbed for Wide Area Computing and Data Handling

10/1/01

PERFORMANCE EVALUATION OF THE GRIDFTP WITHIN THE NORDUGRID PROJECT

Balázs Kónya, Oxana Smirnova

Abstract

This report presents results of the tests measuring the performance of multithreaded file transfers, using the GridFTP implementation of the Globus project over the NorduGrid network resources. Point to point WAN tests, carried out between the sites of Copenhagen, Lund, Oslo and Uppsala, are described. It was found that multiple threaded download via the high performance GridFTP protocol can significantly improve file transfer performance.

1. Introduction

A secure, reliable, efficient and high performance data transfer mechanism over high-bandwidth wide area networks is a key component of any kind of Grid infrastructure. The Globus metacomputing project [1] has proposed the GridFTP protocol [2] which contains extensions to the standard highly popular FTP protocol, in order to meet the requirements of high performance wide area data movement. Their extended file transfer protocol supports the following features:

- Grid Security Infrastructure (GSI) [3]
- partial file transfer
- third-party (from server to server) transfer
- automatic negotiation of TCP buffer sizes
- parallel (multi-threaded) data transfer.

The Globus team has provided software implementation of their protocol in terms of production libraries and tools. The GridFTP code used was the alpha-4 release, checked out from the Globus CVS [4]. All the new features except for the TCP buffer negotiation, has been implemented, and the code will be a part of the Globus 2 release [5].

The purpose of this investigation was to examine the alpha release of the GridFTP code, test and compare its performance to standard FTP file transfer within the framework of the Nordic Testbed for Wide Area Computing and Data Handling (the NorduGrid project) [6]. Because the GridFTP will be the underlying data transfer engine of many Grid projects, it is very important to have a first-hand experience of its performance and capabilities over the NorduGrid hardware and network resources. Particularly interesting issue is the expected performance improvements due to the new parallel transfer mechanism of the GridFTP protocol. It is well-known that the utilization of parallel streams in data transfer over high speed WAN connections is a feasible solution to overcome the TCP buffer size limitations without the necessity of the modifications of any sensitive TCP system parameters [7], therefore parallel streams can lead to a significant improvement in data throughput. Although several multi-threaded file transfer clients exist based on the single-threaded FTP protocol, the Globus Project's GridFTP implementation is one of a very few, which support parallel streams on the protocol level. In the tests presented here, the parallel GridFTP performance over the NorduGrid network have been measured.

2. Test environment

Four NorduGrid sites, Copenhagen, Lund, Oslo and Uppsala, have been participating in the tests. The sites are connected via the high speed Nordunet network [8], the actual network configuration is given in Figure 1.

At each site a dedicated Linux server (the Globus gatekeeper of the local Grid cluster) as a GridFTP server was used in the investigation. The servers of Lund and Uppsala are connected through a 100 Mbit/s link to the LAN, the Copenhagen server has a Gigabit connection, while the Oslo server at the time of the investigation had only a 10 Mbit/s LAN connection.

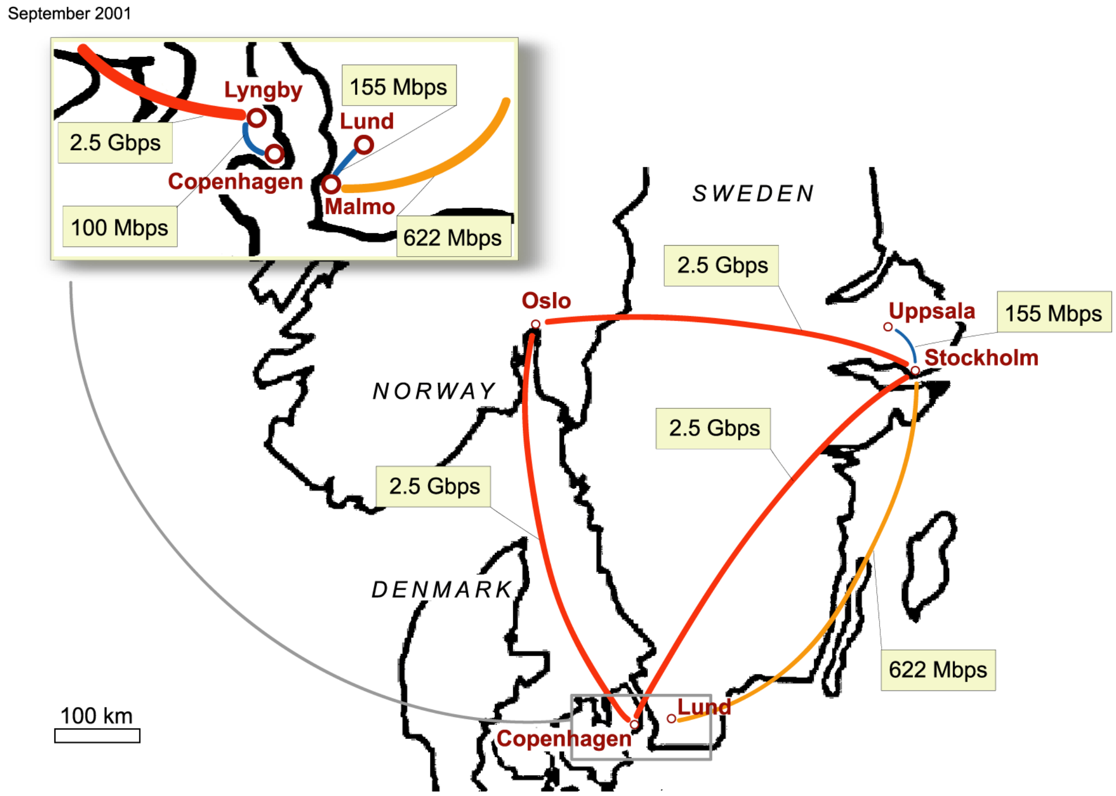

Figure 1. Connectivity map of the NorduGrid participating sites

The hardware configurations of the GridFTP servers are listed in Table 1; more detailed description can be found on the NorduGrid Web-site [9].

The Globus version 1.1.3b14, shipped with the *globus-url-copy* GridFTP client, was installed and configured on the machines. Each site ran the GSI-enabled Globus-modified version of the WU-FTPD server [10]. The particular software configuration of the servers is listed in Table 2.

GridFTP server	***CPU***	***Memory***	***NIC***	***LAN connectivity***
Lund	P-III 1 GHz	512 MB	Intel Pro 100 VM	100 Mbit/s
Uppsala	P-III 866 MHz	512 MB	Intel Pro100 VE	100 Mbit/s
Copenhagen	Dual PIII-933 MHz	512 MB	AlteonAceNIC Gigabit Ethernet	1 Gbit/s
Oslo	P-III 866 MHz	128 MB	EtherExpress Pro100	10 Mbit/s

Table 1. Hardware specifications of the GridFTP servers

GridFTP server	***Platform***	***Globus toolkit***	***GridFTP software***
Lund	Mandrake 8.0	1.1.3b14	gsi-wuftpd-0.5
Uppsala	Redhat 7.1	1.1.3b14	gsi-wuftpd-0.5
Copenhagen	Redhat 6.2	1.1.3b14	gsi-wuftpd-0.5
Oslo	Redhat 7.1	1.1.3b14	gsi-wuftpd-0.5

Table 2. Installed software versions

3. Method of measurement

The *globus-url-copy* tool and the *gsi-wuftpd* server, as the GridFTP client and server from the Globus alpha release 4, were used, respectively. A test file of a size of 100 Mbytes was transferred among the sites, since it was found that this was large enough to average out most of the network fluctuations. During the tests, the default TCP settings were not modified. Throughput measurements with respect to different number of parallel threads were performed over three different network connections:

- Lund-Copenhagen (15 router hops)
- Lund-Uppsala (9 router hops)
- Lund-Oslo (11 router hops)

During the investigation, it was found that the background load of the network link had a significant influence on the download performance: for example, a single-threaded download time of the test file from Uppsala to Lund could change from 64s in the morning to 284s in the afternoon. However, for a reasonable period of time (approximately half an hour) the network conditions could be considered being stable. During the measurements, it was found that most of the cases over any given link the downloading time did not vary more than 5%, provided the downloads were made within 30 minutes. Therefore, in order to eliminate the effect of changing network load, download times were compared only if the downloads were performed over the same network link and taken within the same short time interval; furthermore, all the unreconsctructable and outstanding values were discarded. For a specific collection of data it was required that the difference in the download time of the test file measured at the begining and at the end of the test session did not exceed 5%: in this way one could assume that the test downloads were carried out over a relatively constant network load of the particular network link.

4. Results

4.1. FTP clients

Before investigating the parallel performance of the GridFTP protocol and the *globus-url-copy* client, tests comparing different Linux and MS Windows based FTP clients were carried out. Downloads were performed from the Uppsala server using different FTP clients installed on a dual booting (Mandrake 8.0 Linux and W2K) P-III 1 GHz machine in Lund. Under Linux, the FTP, NcFTP, lftp, SSH copy and the Prozilla [11] clients were tested. The Prozilla application is a multi-threaded Linux FTP client. It is capable of opening multiple connections to a server, where each of the connections downloads a part of a file, and upon completion of downloads, the partial files are merged.

After rebooting the machine to W2K, the tests were repeated with FTP, Windows Commander built-in client, and CuteFTP Pro. The latter client allows up to 4 parallel connections per transferred file.

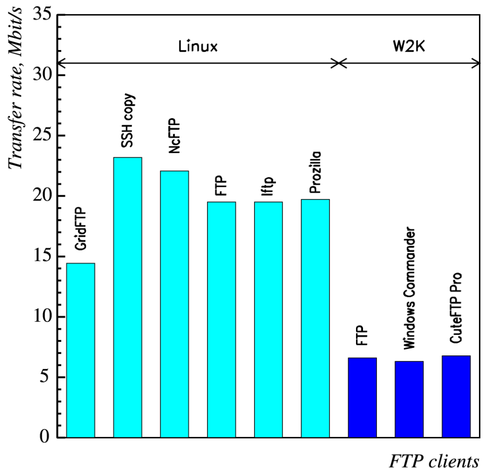

Figure 2. Download performance of conventional FTP clients

The download performance of the clients were compared to single and multiple threaded GridFTP downloads. The results are shown in Figure 2 and Figure 3. It can be seen that multithreaded clients generally improve the transfer time, while Linux clients routinely outperform those of W2K. The Linux multithreaded downloader Prozilla, although providing higher transfer rates with increasing thread number, generally consumes too much time during the reconstruction phase, slowing down the performance. GridFTP provides the fastest downloads already with 4 threads, and can improve further, providing LAN performance and the network load allow for it.

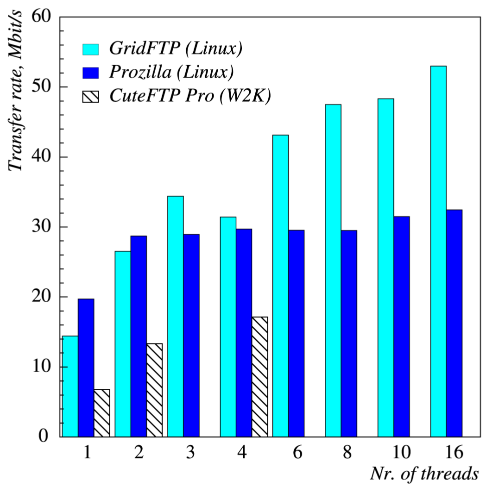

Figure 3. Download performance of multi-threaded FTP clients

4.2. Lund – Uppsala transfer

The main GridFTP tests were carried out between Lund and Uppsala. At both sites, the GridFTP servers are connected to their LAN with a fast 100 Mbit/s Ethernet link. Lund has a 155 Mbit/s connection to the Malmo backbone, which is connected via a 622 Mbit/s line to Stockholm, and Stockholm in turn has a 155 Mbit/s line to Uppsala (see Figure 1).

The network load of the datapaths on the day of the measurements is shown in Figures 4 to 6 [12]. Most of the day (9:00 thru 1:00 CET), the Lund-Malmo link is usually overloaded (running at 90-100% of its total capacity), while the other two links have a moderate load of ~40%.

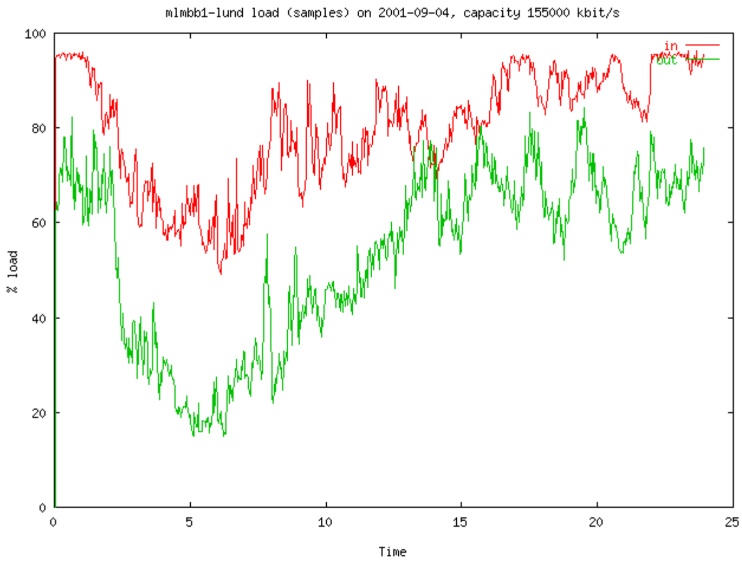

Figure 4. Daily network load of the Malmo-Lund link (percentage of total capacity, in- and outgoing traffic)

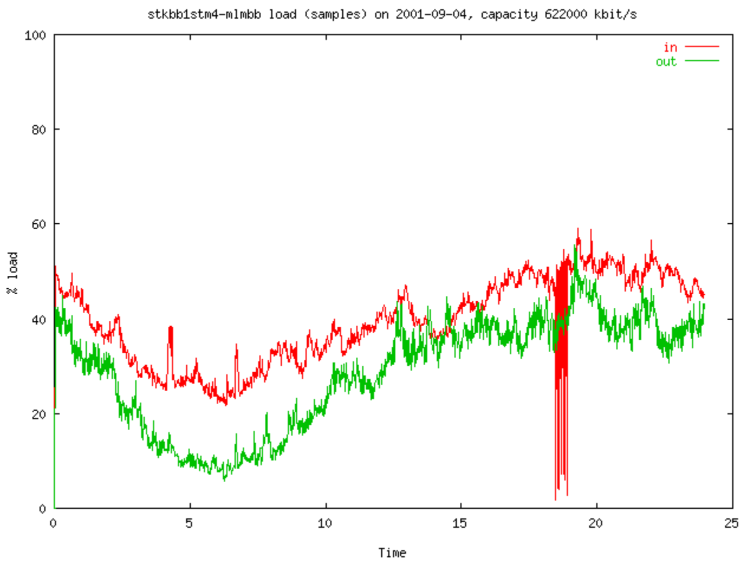

Figure 5. Daily network load of the Stockholm-Malmo link (percentage of total capacity, in- and outgoing traffic)

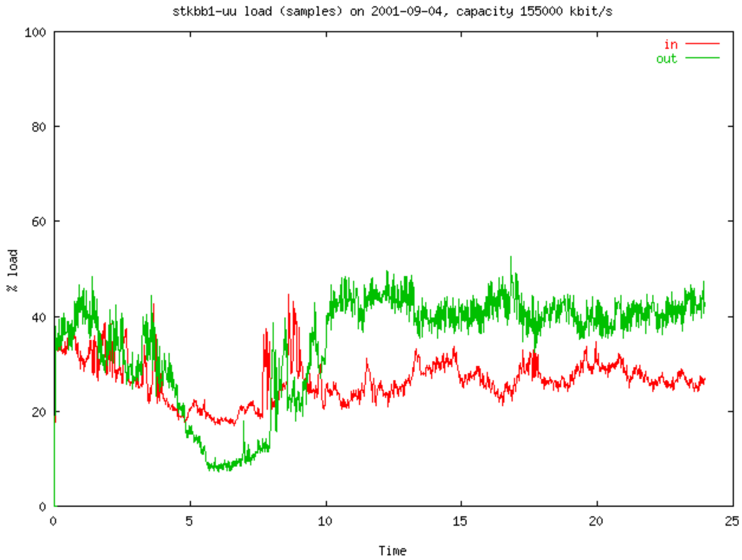

Figure 6. Daily network load of the Stockholm-Uppsala link (percentage of total capacity, in- and outgoing traffic)

Two sets of tests were performed: the first series of transfer measurements were taken in the morning at a relatively low network load, while the second was performed over a congested network at peak time. The throughput performance of the GridFTP with respect to the number of parallel threads via transferring the 100 Mbyte test file between Lund and Uppsala was measured (see Figure 7).

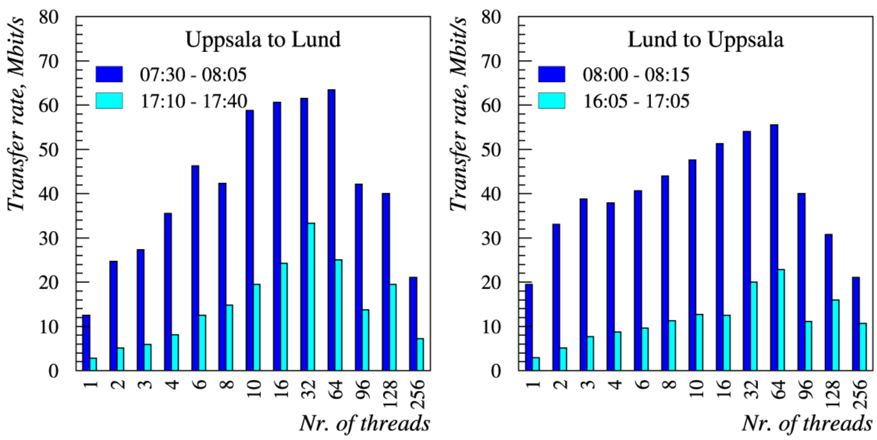

Figure 7. Transfer rates for GridFTP downloads between Uppsala and Lund, measured at different network load

Usage of parallel threads radically increased the transfer rate both over the congested and uncongested network, regardless of the load. The performance was steadily increasing for up to ~64 threads; over 96 threads instabilities and lower transfer rates were experienced. Over a congested network, the throughput with 64 threads increased 780% and 980% respectively in the two directions, compared to the normal single- threaded transfer. In case of the unloaded network, the transfer rate increased from 20 Mbit/s to 60 Mbit/s (Lund to Uppsala), which is actually around the maximum throughput of a 100 Mbit/s LAN. This means that over the Lund-Uppsala datapath already the LAN performance represents the bottleneck in the multi-threaded GridFTP downloads.

4.3. Lund – Oslo transfer

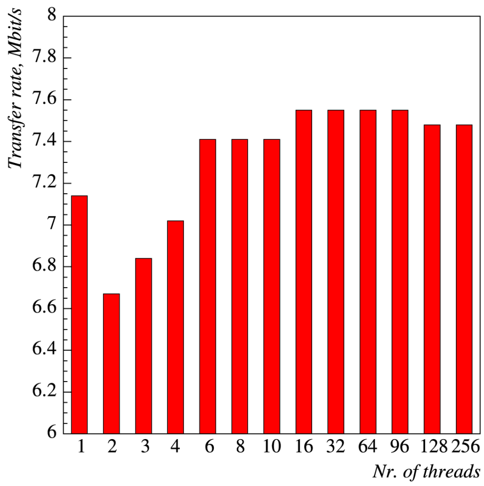

Figure 8. GridFTP transfer rates for download from Oslo to Lund

Lund is connected to Oslo through the 155 Mbit/s Lund-Malmo, 2.5 Gbit/s Malmo-Stockholm and 2.5 Gbit/s Stockholm-Oslo links. However, the 10 Mbit/s LAN connection of the Oslo server was the real bottleneck on the Lund-Oslo datapath. The result of the multithreaded test downloads from Oslo to Lund is shown in Figure 8. The 7 Mbit/s single-threaded transfer rate is already close to the 10 Mbit/s theoretical maximum, thus the 10 Mbit/s networking bottleneck seriously limits the functionality of the parallel downloads.

4.4. Lund – Copenhagen transfer

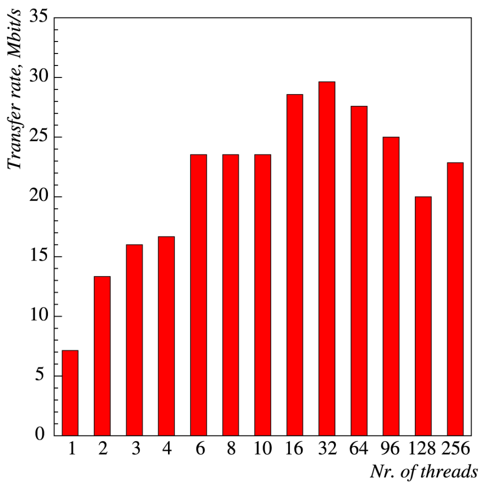

Figure 9. GridFTP transfer rates for download from Lund to Copenhagen

The Lund – Copenhagen datalink, despite the geographical neighborhood of the two sites, represents the longest network connection in these GridFTP tests (15 router hops). The data travel from Lund to Stockholm, then from Stockholm through a 2.5 Gbit/s link to Lyngby, the Danish Nordunet gateway, and finally arrives to Copenhagen (see Figure 9). Performing parallel downloads from Lund to Copenhagen, a throughput gain from 7 Mbit/s to 30 Mbit/s (over a congested network) was experienced, obtained with 32 threads compared to the single download rate.

4.5. Third Party Transfers

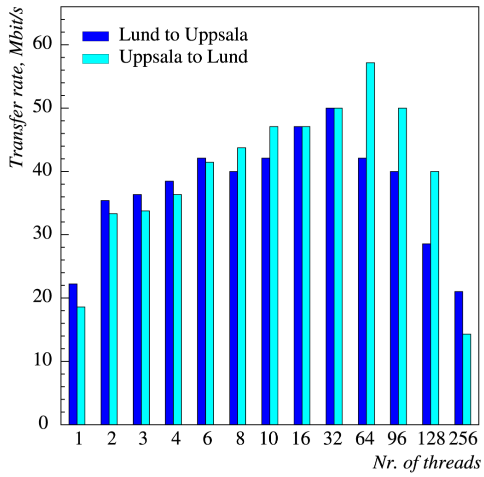

Figure 10. GridFTP transfer rates for the third-party (Copenhagen) transfer between Lund and Uppsala

Finally, third party transfer tests from Copenhagen between the servers of Lund and Uppsala were made (see Figure 10). The parallel threads resulted in a performance gain similar to one of the case of the direct transfer (compare Figure 7 and Figure 10). The transfer rate increased from the single-threaded 15-20 M bit/s up to 50-55 Mbit/s, achieved with 32 threads.

4.6. Stability issues

In spite of the alpha status of the GridFTP software implementation, it was found that the *globus-url-copy* client and the *gsi-wuftpd* server are rather stable. However, using more than hundred parallel streams can sometimes cause instabilities. On several occasions, with large number of used threads, the start-up phase of the download froze for several seconds. Restarting the transfer usually solved the problem. More common unstable behavior of the above-100-threads downloads was the extremely slow completion of the last few hundred kilobytes. On few occasions, the download process in its finishing state (retrieving the last few hundred bytes) suddenly restarted from the beginning. A general conclusion is that the software implementation for above ca. 100 threads became unreliable.

5. Summary

In this report the high performance GridFTP data transfer mechanism was evaluated over the NorduGrid resources. The tests were mainly focused on the performance gain due to the usage of parallel streams. Another purpose was to compare the performance of the Globus GridFTP implementation to several ordinary FTP clients. It was found that the performance of the conventional FTP clients differ not more than 20 percent and that the single threaded GridFTP could deliver in the same range.

The multithreaded GridFTP transfers resulted in a remarkable performance increase of about 600-800 percent, compared to a single threaded GridFTP (or conventional FTP) downloads. Parallel threads lead to increased throughput over both congested and unloaded networks. It's worth to point out that in some of the cases the LAN or the actual hardware configuration became the bottleneck in the GridFTP tests. Parallel third party (server to server) transfers were tested and similar performance enhancement to direct transfers was achieved.

The tests were by no means exhaustive, many other cases could have been considered. The NorduGrid network resources are going to be considerably upgraded at the end of 2001 (the Lund Malmo and the Copenhagen-Lyngby links) and the tests are planned to be extended using the upgraded network; moreover a dedicated 1GBit/s CERN-Copenhagen link is about to be set up for tests purposes in the near future.

The alpha status software is proved to be relatively stable, since irregularities occurred only for downloads involving exaggerated amount of threads (above 100). The conclusion is that the multithreaded GridFTP significantly boosted the data throughput over the NorduGrid network and it is certainly a very promising solution for high performance data transfer.

References

1. The Globus project, www.globus.org
2. Definition of the GridFTP protocol, features and implementation, www.globus.org/datagrid/gridftp.html
3. Globus Security Infrastructure, www.globus.org/security
4. The GridFTP alpha 4 release, www.globus.org/gsiftp-alpha/release-alpha4.html
5. Globus toolkit version 2, www.globus.org/gt2
6. NorduGrid project, www.quark.lu.se/grid
7. TCP Tuning Guide for Distributed Application on Wide Area Networks, www-didc.lbl.gov/tcp-wan.html
8. Nordunet, www.nordu.net
9. NorduGrid hardware resources, www.quark.lu.se/grid/hardware.html
10. WU-FTPD project, www.wu-ftpd.org
11. Multi-threaded Linux FTP client, prozilla.delrom.ro
12. Nordunet network statistics, www.nordu.net/stats, stats.noc.kth.se/stats