Diversity of Grid Traffic: A Survey-based Study

Yehia El khatib, Christopher Edwards

Computing Department
InfoLab21
South Drive
Lancaster University
Lancaster LA1 4WA UK

{yehia, ce} @ comp.lancs.ac.uk

***Abstract*- Grid computing offers the prospect of harnessing huge amounts of computational resources. However, it is being argued that such potential cannot be fully exploited due to the nature of the Internet architecture which is not suitable for high-speed communication of large volumes of data. This has motivated EC-GIN, a European project which aims to exploit the network in a way that better suits the needs of Grid applications. In order to reach such a goal, the network requirements of Grid applications first need to be understood. We have conducted a survey to investigate the requirements and characteristics of a number of Grid applications used in scientific research. Among other things, the survey results have revealed the diversity of Grid traffic, suggesting that there is more to Grid traffic than just transfers of huge bulks and tiny control signals. In this paper, we present these results and identify different classes of traffic behaviour that have been observed within the results. We then validate our findings by looking in detail at two of the applications that we have surveyed.**

I. INTRODUCTION

Grid computing offers the prospect of gaining huge amounts of computational resources for a fraction of the cost that would be paid to actually own such resources. It is a safe investment for any organisation whose work involves computations that require abundant resources, such as processing power, storage space, etc. Grid computing, therefore, enables organisations to harness the most out of their resources regardless of how geographically scattered or locally administrated these resources are.

However, it is being argued that present-day networking technologies are not suitable for the kind of traffic that is transmitted in Grids [21]. In particular, it is suggested that the limitations of the TCP/IP stack prevent Grids from working to their full potential [2, 7, 19]. However, this argument is largely based on the assumption that Grid traffic is mostly large bulks of data [13, 15, 19]. While there is indeed sufficient evidence that TCP is not suitable for high-speed bulky data transfers [2, 3, 9, 10, 11, 12, 18, 20], there is not much to support the assumption that large bulks of data dominate the traffic of Grid applications. We believe that this assumption has been based on speculations and forecasts of how Grid applications work.

The argument that current Internet technologies are inapt for Grid communications provided the inspiration behind Europe-China Grid InterNetworking (EC-GIN) [23], a European-funded research project for improving the ability of the network to support Grid applications. One of the objectives of EC-GIN is to introduce GIN-TONIC, a comprehensive networking API that provides new programming abstractions designed to improve the performance of network communication across the Grid. For the architectural design of GIN-TONIC, understanding the requirements of Grid applications is crucial. It might be easy enough to predict these requirements according to our perception of Grid applications. However, a close look at some applications that are currently in operation would yield a more realistic set of requirements.

This has motivated us to conduct a survey of current Grid applications. In this survey, we look at different characteristics of the Grid applications, their middleware environments, their traffic footprints, and most importantly their network requirements. The survey also presents us with evidence that Grid traffic is not necessarily “mice and elephants” [17], i.e. very small control signals and very large bulks of data transfer.

The rest of this paper is arranged as follows. Section II summarises the results of the survey. In section III, we define five different classes of Grid applications according to their traffic footprint. To illustrate how this classification scheme is implemented, we apply it to two of the surveyed applications in section IV. We discuss future work in section V, and conclude in section VI.

II. SURVEY OF APPLICATION REQUIREMENTS

This section gives an overview of the conducted survey. We briefly discuss the aim and process of conducting the survey and then comment on its significant results.

A. Aim

The aim of the survey is to draw a clearer picture of what the network requirements of Grid applications are, based on the specifications of deployed applications. The results give a recommendation of the services that need to be included in the API design. The results also describe some aspects of the applications such as scale, composition, dataset granularity, delay-sensitivity, middleware, accounting metrics, etc. The output of the survey, however, is not intended to be a comprehensive statistical analysis of the different aspects of Grid applications.

B. Process

The survey was conducted by circulating a 2-page questionnaire amongst projects employing or in the process of developing Grid applications for scientific research. Due to the technical nature of some of the questions, we targeted people who have adequate experience with Grid applications. This included the developers, administrators, and advanced users who have used the system enough to know about its behaviour and requirements. In a small number of cases, participants were asked for a short interview to get more details or to clarify their responses. A set of 16 individual results was collected and analysed.

C. Results

1. Research Field

18% of the applications we surveyed are used for particle physics research; 13% are used for astronomy; 13% for engineering; 13% for some form of mathematical analysis computations; 13% for social sciences; while the remaining 30% were equally used for the following purposes: environmental sciences, medicine, meteorology, software development, visualization. Fig. 1 illustrates this distribution.

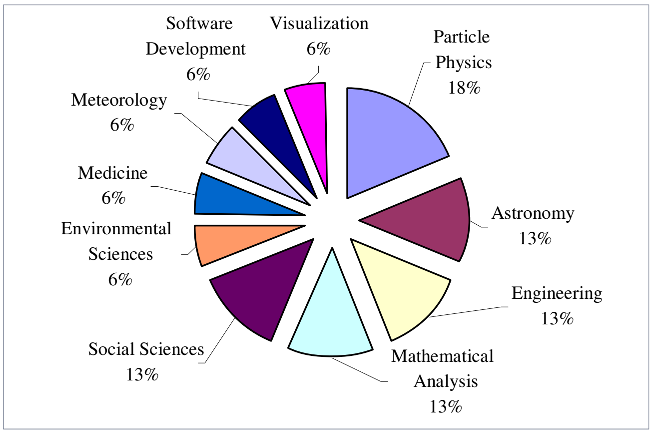

Fig. 1. Pie-chart illustrating the research fields of the surveyed applications

2. Scale

Of the surveyed applications, 13% are deployed over Grids made up of 10 nodes or less, 54% are deployed over 100-400 nodes, 20% are deployed over 400-1000 nodes, and 13% are deployed over Grids of more than 1000 nodes.

The majority of these Grids (71%) span across 3-10 administration domains, while 21% have nodes in 10-100 different domains. Only 8% of the surveyed applications are deployed over a Grid that has nodes in more than 1000 different domains.

3. Composition

47% of the surveyed applications are deployed solely on dedicated clusters. Only 7% of the surveyed applications are deployed on a Grid free of dedicated clusters, consisting only of desktop computers. The remaining Grids (46%) are almost equally composed of dedicated clusters and desktop machines.

It is worth noting that only one application uses small devices (such as embedded processors) and they only constitute 1% of the total number of devices in that Grid. In addition, there is no application that has mobile phones as nodes in its Grid.

4. Dataset Granularity

Based on the (approximate) values given by the participants, the survey revealed that the three most common dataset sizes are 10 kB, 10 MB, and 100 GB. These are visible as the peaks in Fig. 2 which depicts the logistic distribution of dataset sizes.

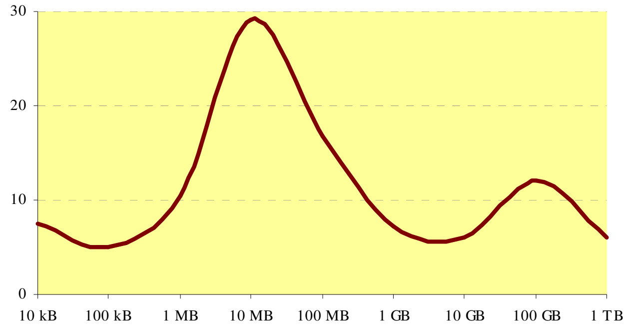

Fig. 2. The probability density function of dataset sizes

A closer look at the numbers shows that almost 12% of the datasets of all surveyed applications are in bulks smaller than 100 kB in size, 55% are in bulks of 1-100 MB, and 18% are in bulks of 10 GB or more. Fig. 3 illustrates the sigmoid curve of the distribution of dataset sizes.

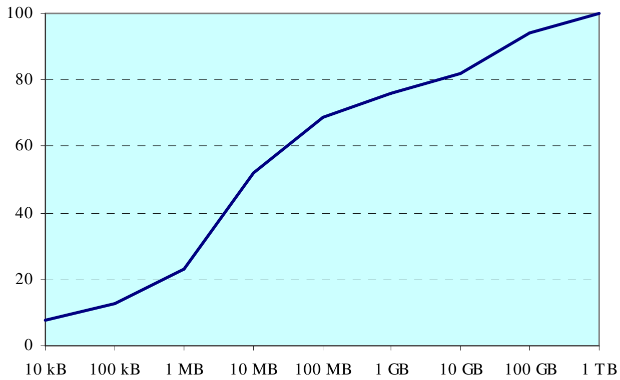

Fig. 3. The cumulative distribution function of dataset sizes

Although only to a limited extent, these numbers show how different Grid traffic is when compared to generic IP traffic (such as Web traffic). Moreover, they illustrate how mixed the dataset sizes are. Such diversity imposes a huge challenge on the design of an API that is to enhance the communication of Grid traffic.

This observation, however, is equally important because it refutes the common concept that Grid traffic consists predominantly of large bulks. The results clearly demonstrate that the majority of dataset sizes are below the gigabyte limit.

5. Data Timeliness

Time-critical applications need to enforce deadlines on the delivery of their packets. Packets that arrive later than the deadlines are considered of no use and are discarded. Embarrassingly parallel applications, on the other hand, do not typically impose such deadlines.

One of the applications we surveyed is being used for forecasting Alpine watersheds and thunderstorms based on parameter measurements from data collection points deployed in the field. Data that arrives late has to be discarded in order to process the data that is due. Besides this application, only one more of the surveyed applications imposes a deadline on the delivery of non-multimedia-stream data packets. Although this latter application involves Web service invocations which are asynchronous by nature, the application imposes strict deadlines on the delivery of Web service results. This reflects the essence of promptness in this application.

Interestingly, the time-sensitive part of the data in the two applications discussed above is mainly the part that is transferred in bulks of gigabytes or more.

6. Encryption

Although security is a major concern of Grid applications [4], only 44% of the surveyed applications encrypt their data prior to sending it over the network. Of these applications, 57% rely on the middleware to provide the encryption as opposed to encryption being carried out entirely by the network transport layer.

7. Data Path

The created traffic that has more than one recipient amounts to 22% of the total traffic of all surveyed applications.

Only 44% of the surveyed applications employ one-to-many communication schemes. These applications all use multicast one way or another. Two thirds of these applications integrate a multicasting mechanism into their code, while the other third employs middleware multicasting services.

Besides multicast, only one application uses an anycast scheme [16], which is provided by the middleware. The same application also implements its own means of scavenging, a more advanced anycasting scheme where the recipients of the data are chosen according to specific criteria set forth by the application and verified by the resource brokering element of the middleware.

III. Analysis of Traffic Behaviour

The survey results indicate the non-conformance of Grid traffic to common belief. In order to emphasise this contrast and highlight the different ways Grid applications utilise the network, we classify the surveyed applications according to their traffic behaviour as identified by the sizes of their datasets.

The aim of this effort is not to typecast Grid applications. Our aim is to distinguish the magnitude of difference in traffic behaviour and to investigate the causes of such differences. Moreover, the recognition of such differences will help in the design of a more realistic network-enhancing API.

From the survey results, we noted five main classes of traffic behaviour. We list these different classes below and comment on the common aspects observed within the applications of each class. Fig. 4 displays the range of average dataset sizes for each of the different classes. The x-axis in the figure is to logarithmic scale.

A. *Class A*

The first class of traffic behaviour is the most noticeable due to the fact that although it might be the least anticipated behaviour of Grid applications, 34% of all the surveyed applications fall under this class. This includes applications that deal mostly with lightweight datasets; ones which are never larger than 5-10 MB. These applications are used for either mathematical calculations or distributed data management in projects related to mathematical analysis, engineering, and social sciences.

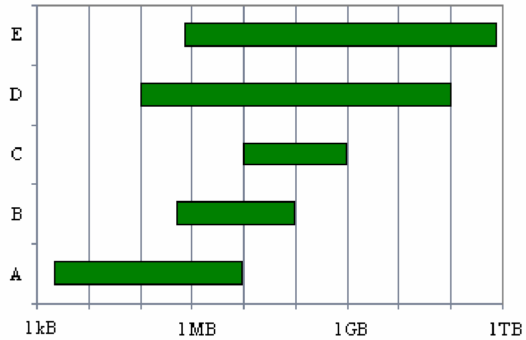

Fig. 4. Average dataset sizes of the different traffic behaviour classes

B. *Class B*

Applications which can be grouped under this class have datasets between 0.5 and 100 MB in size. Moreover, the variance in dataset sizes in this class is relatively large, i.e. dataset sizes of each application under this class tend to be at both ends of the mentioned range. Of all surveyed applications, 20% fell under this class. These included image analysis as well as simulation applications. Remarkably, all applications in this class are used for either astronomy or meteorology. Furthermore, all these applications were deployed over Grids of 100-300 nodes across 6-8 administrative domains.

C. *Class C*

13% of the surveyed applications had all their dataset sizes in the relatively narrow range of 10 MB – 1 GB. These applications are used for advanced software development techniques and distributed data management. All applications in this class are deployed on Grids that are made up mostly of desktop machines, making this class the only one with a vivid relationship between the composition of the Grid and the traffic behaviour.

D. *Class D*

The fourth class contains applications whose dataset sizes vary within a wide range from 100 kB to 100 GB. However, the majority of the datasets are between 10 MB and 10 GB in size. 13% of the surveyed applications fall under this class, and they are used for simulations, mathematical modeling, calibrations and complex computations.

E. *Class E*

The fifth and final class contains the heavyweight applications that have received the most attention in Grid computing literature. The main focus of these applications is the analysis of very large datasets, in the order of tens to hundreds of gigabytes, as well as other datasets as small as a few megabytes in size. The 20% of our survey population that fall under this class are being used for particle physics, engineering and social sciences in order to perform complex numerical analysis and/or large-scale simulations. These applications run over huge Grids made up of thousands of nodes, including clusters, desktop machines, and small devices (such as embedded processors), spanning across a large number of administrative domains.

F. Comments

With the exception of class C, there seems to be no clear relationship between the average size of the datasets and the composition of the Grid. There are applications that handle datasets of sizes in the order of gigabytes (such as class D or E applications) and there are others which have the majority of the datasets in the order of a few kilobytes (such as class A). Nonetheless, there is enough evidence in the survey results to suggest that the all applications are capable of running on Grid systems that are entirely made up of clusters or desktop machines. Furthermore, there is no solid association, other than the one observed in class B, between the research field for which the application is used and the amount or pattern of traffic it creates.

It is interesting that the ranges of dataset sizes of classes D and E are much wider than those of the other classes. This can be easily discerned from Fig. 4. Perhaps this is due to the fact that applications in classes D and E are used for intensive computations, such as simulations and complex numerical processing, which, according to the evidence presented by the survey, require the transfer of very large bulks of data as well as a significant amount of small datasets.

Our survey has included applications in various fields of scientific research including particle physics, meteorology, astronomy, engineering, mathematical analysis, social sciences, and medicine. However, they do not all involve the transfer of large-scale data volumes. In fact, the most common class of traffic behaviour is class A (see Fig. 5) which involves datasets no larger than 10 MB.

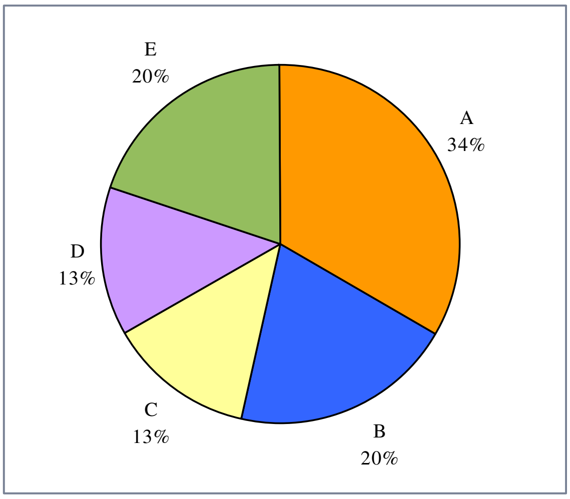

Fig. 5. The distribution of the traffic behaviour classes amongst the surveyed applications

IV. EXAMPLES OF THE GRID APPLICATIONS SURVEYED

In order to validate our findings, we report in some detail on two of the applications that have been included in the survey and apply the classification scheme to them. First, we give an outline of the purpose of each application, how it works and its typical traffic pattern. Then, in view of that, we place each application under one of the aforementioned classes.

A. ATLAS for LHC

LHC, or the Large Hadron Collider [26], is the world's largest particle physics experiments to date, costing a total of £2.6 billion. Located at CERN near the Switzerland-France border, LHC will see its first particle collisions by November 2007. The experiment is planned to run for nine consecutive months and then cease for three months before commencing again. During the first active period of nine months, the LHC experiment is expected to trigger huge amounts of raw data in the neighbourhood of 10 petabytes. This harvested data will then be processed by the Grid and the results obtained will be compared to those of simulated experiments. ATLAS (A Toroidal LHC ApparatuS) [22] is one of the five particle detector experiments that will run at LHC, and it is "the largest volume detector ever constructed for particle physics" [29]. ATLAS brings together almost 2000 scientists from around the world.

The Grid infrastructure for this project, the LHC Computing Grid (LCG) Project [27], is made up of 165 scientific organisations, universities and government bodies connected together using a dedicated 10Gbps lightpath. These sites are organised using a three-tier distribution architecture. *Tier 0* is the particle physics laboratory at CERN where part of the data analysis will take place. However, all particle physics aside, the main function of the laboratory is to farm out the raw data over the Grid to the *Tier 1* sites.

There are ten Tier 1 sites scattered across France, Germany, Italy, Japan, the Netherlands, the UK and the US. Each of these sites has a cloud of *Tier 2* sites associated with it. Tier 1 sites are responsible for splitting up the raw data they receive, from Tier 0, between their respective Tier 2 sites. Each Tier 2 site processes the data upon receiving it, stores the results locally on magnetic disk, and sends a copy of the results to its respective Tier 1 site where it is stored on tape. Hence, there are always at least two copies of every result file in the Grid.

Most data is sent as huge chunks, in the order of a few gigabytes, at scheduled times during which sufficient network resources would have been reserved in advance. However, some datasets of special importance are transferred on demand regardless of the pre-scheduled times. Also, data processing results returned from Tier 2 to Tier 1 are sent whenever they are ready (ad hoc submission).

Data is stored in physical files of 1-4 gigabytes each. Such large file sizes are necessary in order to reduce the overhead on magnetic and tape storage devices, as a small number of large files are easier to manage and index than a large number of smaller files.

A dataset can be a collection of any sort of raw data or processed information. It is written only once and is then never modified nor deleted (except on rare occasions). A single dataset can be composed of one or more physical data files. To the physicists (i.e. the users), however, the physical file partitioning is invisible; they deal only with datasets. In contrast, *catalogues* are transactional collections of metadata that may be modified by more than one user at a time.

When a job is submitted to the system it is typically executed at the location where the required datasets reside, regardless of where the job was initiated. The aim, of course, is to minimise the overhead induced by moving large datasets across the network. This strategy, however, relies on the over-provisioning of processing power. After the job is executed, a dataset of results is returned to the user. This dataset is reasonably small in comparison to those containing raw data.

Hence, the traffic created by ATLAS is mostly made up of datasets in the order of a few gigabytes (i.e. the raw data), but it also consists of smaller datasets in the order of a few megabytes (i.e. the job results). Accordingly, ATLAS is classified under class D.

B. GROWL for GeSRM

SABRE (Software for Analysis of Binary Recurrent Events) [28] is an application developed to process very large amounts of longitudinal data. Such data is typically made up of millions of observations per dataset, with a large number of parameters associated with every dataset. SABRE employs fast numerical algorithms, running them in a parallel fashion across the Grid.

GROWL (Grid Resources On a Workstation Library) [25] is a toolkit that facilitates the use of client-server legacy applications on the Grid, by employing SABRE in order to submit jobs to the legacy server. GROWL, thus, enables any pre-built service to be run over the Grid without the need for any modification to the service. All that is needed is to build a thin client that translates the users' jobs into SABRE Web services to be sent to a GROWL server. In turn, this server will then translate these Web service invocations into calls that are recognisable by the legacy server which resides at the same site as the GROWL server. Such separation of server logic from the client application makes it easy and flexible to distribute more than one copy of a client-server application across a Grid.

GeSRM [24] is a research project intended to develop a method of spatial analysis known as GWR, or Geographically Weighted Regression, to run over the Grid. GeSRM employs GROWL to submit a large number of computational tasks over the Grid using Web services. This approach minimises the client footprint on users' machines. Nevertheless, GeSRM transfers very large datasets of spatially dispersed data between clusters. It is not uncommon for these datasets to stretch to 100 GB or more in size. At the other end of the scale, Web service invocations are quite lightweight and, although they are usually smaller than 1 MB, they do constitute almost 30% of the transferred traffic. This is because a large number of user jobs might induce only little data processing on the clusters. Therefore, we classify the traffic created by GROWL in this instance under class E.

V. Future Work

The results obtained from the survey have encouraged us to conduct the survey once again in pursuit of a larger result set. There is no doubt that a greater set of results will help in scrutinising any relationships discovered between the classes and the aspects of the applications (such as Grid scale, Grid composition, middleware, etc. as well as dataset granularity).

The survey offers an overview of the size and nature of traffic exchanged in Grids. The results illustrate how traffic differs from one application to another. However, the results do not show how the traffic of one application fluctuates from one time to another. In a further effort to study Grid traffic, we intend to monitor the traffic of a number of Grid applications. Through carrying out detailed analysis and mathematical modeling of the monitored traffic, we are hopeful of providing a realistic representation of Grid traffic that can then be used in Grid simulators. However, monitoring any distributed system, including Grids, is no simple task [1, 8, 14], as several factors must be taken into consideration.

VI. Conclusion

We conducted a survey in which we focused on network-related features of Grid applications. The survey presented information about the applications such as the network functionality, network demands, middleware interaction, etc.

From the results, we have suggested, for illustrative purposes only, a classification scheme that distinguishes different traffic footprints. This classification points out the diversity in Grid traffic; 34% of all the surveyed applications have datasets under 10 MB in size, 54% of all surveyed applications have datasets under 100 MB in size, and 74% of all surveyed applications have datasets under 1 GB in size. With these numbers in mind, the survey fables the belief that the majority of Grid traffic is made up of enormous volumes of data. Quite the opposite, the results demonstrate that Grid traffic comes in all shapes and sizes.

Acknowledgments

The authors would like to thank the European Commission for providing the funding for the work under which this paper was completed. The authors are grateful to Mr. Brian Davies of The Particle Physics Research Group at Lancaster University, Dr. Roger Jones of LHC at CERN, Dr. Daniel Grose of The Centre for e-Science at Lancaster University, and Prof. Robert Crouchley of The Centre for e-Science at Lancaster University.

References

[1] "Grid Network Monitoring: Demonstration of Enhanced Monitoring Tools". Deliverable D7.2, EU Datagrid Document: WP7-D7.2-0110-4-1. 30 April, 2007.

[2] W. Feng and P. Tinnakornsrisuphap. "The Failure of TCP in High-Performance Computational Grids". Proceedings of the 2000 ACM/IEEE conference on Supercomputing, article no. 37, Dallas, Texas, United States, 2000. ISBN: 0-7803-9802-5.

[3] S. Floyd. "HighSpeed TCP for Large Congestion Windows". Request for Comments: 3649. December 2003.

[4] I. Foster, C. Kesselman, G. Tsudik, and S. Tuecke. "A Security Architecture for Computational Grids". Proceedings of the 5th ACM conference on Computer and Communications Security, pp: 83-92, San Francisco, California, United States, 1998. ISBN: 1-58113-007-4.

[5] I. Foster, C. Kesselman, and S. Tuecke. "The Anatomy of the Grid: Enabling Scalable Virtual Organizations".

International Journal of Supercomputer Applications, vol. 15, no. 3, 2001.

[6] I. Foster, and C. Kesselman (ed). "The Grid: Blueprint for a New Computing Infrastructure". Second edition, Morgan Kaufmann Publishers, 2004. ISBN: 1-55860-933-4.

[7] Y. Gu and R. Grossman. "UDT: An Application Level Transport Protocol for Grid Computing". Second International Workshop on Protocols for Fast Long-Distance Networks, Argonne, Illinois, United States, 2004.

[8] D. Gunter, B. Tierney, K. Jackson, J. Lee, and M. Stoufer. "Dynamic Monitoring of High-Performance Distributed Applications". Proceedings of the 11th IEEE Symposium on High Performance Distributed Computing (HPDC-11), Edinburgh, UK, July 2002.

[9] E. He, R. Kettimuthu, Y. Gu, S. Hegde, M. Welzl, P. Vicat-Blanc Primet, J. Leigh, and C. Xiong. "Survey of Protocols and Mechanisms for Enhanced Transport over Long Fat Pipes". Global Grid Forum White paper Draft (work in progress), Data Transport Research Group, 2003. Available at: http://www-unix.mcs.anl.gov/~kettimut/Survey.pdf

[10] C. Jin, D. X. Wei, S. H. Low, J. Bunn, D. H. Choe, J. C. Doyle, H. Newman, S. Ravot, S. Singh, F. Paganini, G. Buhrmaster, R. L. A. Cottrell, O. Martin, and W. Feng. "FAST TCP: From Theory to Experiments". IEEE Network, vol. 19, no. 1, pp. 4–11, January-February 2005.

[11] D. Katabi, M. Handley, C. Rohrs. "Congestion Control for High Bandwidth-Delay Product Networks". Proceedings of the 2002 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, session: Congestion Control, pp: 89-102, Pittsburgh, Pennsylvania, USA, 2002. ISBN: 1-58113-570-X.

[12] T. Kelly. "Scalable TCP: Improving Performance in Highspeed Wide Area Networks". ACM SIGCOMM Computer Communication Review, vol. 33, issue 2 (April 2003), column: Technical papers, pp: 83-91, 2003. ISSN: 0146-4833.

[13] T. Lavian, D. Hoang, J. Mambretti, S. Figueira, S. Naiksatam, N. Kaushik, M. Inder, R. Durairaj, D. Cutrell, S. Merrill, H. Cohen, P. Daspit, and F. Travostino. "A Platform for Large-Scale Grid Data Service on Dynamic High-Performance Networks". First International Workshop on Networks for Grid Applications (GridNets), San Jose, California, United States, October 2004.

[14] C. Lee, J. Stepanek, R. Wolski, C. Kesselman, and I. Foster. "A Network Performance Tool For Grid Environments". Proceedings of the 1999 ACM/IEEE Conference on Supercomputing, Portland, Oregon, United States, 1999. ISBN: 1-58113-091-0.

[15] L. Paisley and J. Sventek. "Real-time Detection of Grid Bulk Transfer Traffic". 10th IEEE/IFIP Network Operat-ions and Management Symposium (NOMS), pp 66-72, 2006. ISSN: 1542-1201, ISBN: 1-4244-0142-9.

[16] C. Partridge, T. Mendez, and W. Milliken. "Host Anycasting Service". Request for Comments: 1546. November 1993.

[17] F. Travostino, J. Mambretti, and G. Karmous-Edwards (ed). "Grid Networks: Enabling Grids with Advanced Communication Technology". John Wiley & Sons Ltd, 2006. ISBN: 0-470-01748-1.

[18] D. Wei, C. Jin, S. Low, and S. Hegde. "FAST TCP: Motivation, Architecture, Algorithms, Performance". IEEE/ACM Transactions on Networking (TON), vol. 14, issue 6 (December 2006), pp: 1246-1259, 2006. ISSN: 1063-6692.

[19] E. Weigle and W. Feng. "A Case for TCP Vegas in High-Performance Computational Grids". Proceedings of the 9th IEEE International Symposium on High Performance Distributed Computing (HPDC'01), San Francisco, California, United States, August 2001.

[20] M. Welzl. "Traceable Congestion Control". International Workshop on Internet Charging and QoS Technologies (ICQT 2002), Zürich, Switzerland, 16-18 October 2002.

[21] M. Welzl and M. M. Yousaf. "Grid-Specific Network Enhancements: A Research Gap?". IEEE/IFIP International Workshop on Autonomic Grid Networking and Management (AGNM'05), Barcelona, Spain, October 2005.

[22] The ATLAS (A Toroidal LHC ApparatuS) Experiment. http://atlas.ch

[23] Europe-China Grid InterNetworking (EC-GIN), a Framework 6 STREP project, Action Line: IST-2005-2.6.5.1.c Grids China, Project Reference: 045256. http://www.ec-gin.eu/

[24] Grid-enabled Spatial Regression Models (GeSRM), a small grant project funded by the National Centre for e-Social Science (NCeSS), UK. http://www.ncess.ac.uk/research/sgp/gesrm/

[25] Grid Resources On a Workstation Library (GROWL), a collaborative project (JISC VRE programme) between CCLRC Daresbury Laboratory and the Universities of Cambridge and Lancaster. http://www.growl.org.uk/

[26] The Large Hadron Collider. http://lhc.web.cern.ch

[27] The LHC Computing Grid Project. http://lcg.web.cern.ch/LCG

[28] Software for Analysis of Binary Recurrent Events (SABRE). http://sabre.lancs.ac.uk/

[29] CERN press release PR17.06, "World's largest superconducting magnet switches on", 20.11.2006 http://press.web.cern.ch/Press/PressReleases/Releases2006/PR17.06E.html