Enterprise Desktop Grids*

Evgeny Ivashko

Institute of Applied Mathematical Research,
Karelian Research Centre of Russian Academy of Sciences,
Petrozavodsk, Russia,
ivashko@krc.karelia.ru
WWW home page: http://www.krc.karelia.ru/HP/ivashko

Abstract. The paper describes a tool to perform high-performance computing using idle resources of desktop computers of an organization. The tool – Enterprise Desktop Grid – strengthens ideas of the Desktop Grids. The main characteristics and application areas of the Enterprise Desktop Grids are presented.

Keywords: Enterprise Desktop Grid, BOINC, distributed computing

1 Introduction

Due to its huge potential computing power and ease of use, Desktop Grids are the valuable part of the domain of scientific high-performance (high–throughput) computing.

Desktop Grid is the concept developing from the 1990-ths [1]. It is widely known as the tool to perform long-term computing intensive scientific projects like development the new climate models and computational experiments [2], validation of the mathematical hypothesis [3], astronomy projects [4] and many others.

By now, there are more than 50 active worldwide-known projects, the most active of which exceeds 700 TFlops of peak performance [5].

Also, Desktop Grids can be used as an additional pool of computing resources which allow to save the expensive supercomputing resources (for example, Desktop Grid project of CERN' Large Hadron Collider [6]).

There are different software middleware developed to establish Desktop Grid projects. But the most known of them is Berkeley Open Infrastructure for Network Computing (or BOINC [7]).

The main problem which limits the use of Desktop Grid is a narrow class of suitable problems. Because of its nature, Desktop Grids have very slow and unreliable communications between computing nodes and central server. So, the only class of suitable problems is so-called "bag-of-tasks" — computing–intensive problems which consist of huge number of independent computing tasks do not requiring transfers of big amounts of data.

* The work is supported by grants of Russian Fund for Basic Research 15-29-07974, 13-07-00008 and 15-07-02354

One of the ways to widen the class of suitable problems is development the technologies of high-performance (and high–throughput) computing based on desktop computers (and laptops) connected by local-area communication network of an organization.

2 Enterprise Desktop Grid

Enterprise Desktop Grid is a set of geographically distributed desktop computers, connected with a server by a communication network of an organization with the aim of solving a common computational problem.

There are some examples of Enterprise Desktop Grids: for example, Entropia [12], Aneka [14], Alchemi [13] and Fraunhofer Enterprise Grids [11]. The comparison of several Enterprise Desktop Grid systems is given in [15].

Enterprise Desktop Grids allow organizations to harness the unused computing power of their existing desktop computers to build virtual supercomputers with no additional investment in hardware. It can reduce the cost of hardware by improving utilization rates of desktop computers (like cloud computing concept reduces the cost of hardware by improving utilization rates of servers).

According to [12], the Enterprise Desktop Grid should meet the following requirements:

1. **Efficiency** — The system must harvest unused cycles efficiently, collecting virtually all of the resources available.
2. **Robustness** — The system must complete computational jobs with minimal failures, masking underlying resource and network failures. In addition, the system must provide predictable performance to end-users despite the unpredictable nature of the underlying resources.
3. **Security** — The system must protect the integrity of the distributed computation. Tampering with or disclosure of the application data and program must be prevented. The system must also protect the integrity of the computing resources that it is aggregating. Distributed computing applications must be prevented from accessing or modifying data on the computing resources.
4. **Scalability** — The system must scale to the use of large numbers of computing resources. Systems must scale both upward and downward performing well with reasonable effort at a variety of system scales.
5. **Manageability** — Any system involving thousands to hundreds of thousands of entities must provide management and administration tools.
6. **Unobtrusiveness** — The system typically shares resources (both computing, storage, and network resources) with other systems in the corporate IT environment. As a result, the use of these resources should be unobtrusive, and where there is competition, non-aggressive. The computing system must manage its use of resources so as not to interfere with the primary use of the desktop computers' resources and networks for other activities. This includes both the use due to system activities as well as use driven by the application.

7. **Openness/ease of application integration** — Fundamentally, the Enterprise Desktop Grid is a platform on which to run applications. The number, variety, and utility of the applications supported by the system directly affects its utility. The system must support applications developed with all kinds of models, with many distinct needs and with minimal effort.

Together, these seven criteria represent the key requirements for Enterprise Desktop Grid systems.

Below is given a short comparison of Personal computers (PC), Desktop Grid systems (DG), computing clusters (Cluster) and Enterprise Desktop Grid systems (EDG).

	PC	**DG**	**Cluster**	**EDG**
Source	PC	anonymous PCs over Internet	high-performance computers	PCs over local-area network
Dedication	yes/no	no	yes	no
Resource level	low	high	medium	medium
Class of problems	non-HPC	narrow	wide	narrow/wide
Heterogeneity	no	high	medium/low	medium/low
Trust	high	low	high	medium/high
Complexity	low	high	medium	medium
Costs	low	low	big	low
Users	single user	science	business	SME

The table shows that Enterprise Desktop Grid concept could be useful.

BOINC software platform seems to be a good choice to develop an Enterprise Desktop Grid middleware. It is well-known, easy to use, widely used and has good support and documentation as well as open source code. But the problem is that BOINC is based on POP model of client-server interaction (see Fig. 1). It means that client is responsible for initiation all the interactions with the server. In case of Enterprise Desktop Grids it, in particular, leads to significant waste of resources. PUSH model (see Fig. 2) is the model of client-server interaction, which should be used for Enterprise Desktop Grid.

3 Applications

From the point of view of users, it is very important to clearly understand the possible applications of the Enterprise Desktop Grids. The some of them are the following:

– Rendering [10]: rendering is the process of generating an image from a 2D or 3D models. It is a very CPU/GPU-consuming process that can be easily divided between multiple computers. Rendering can be performed as an Enterprise Desktop Grid project to produce some advertising materials, video-models and so on.

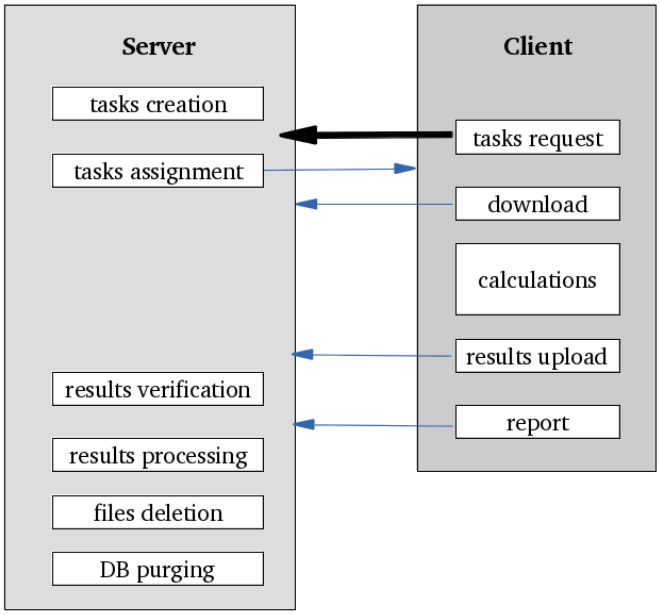

Fig. 1. POP model of client-server interaction.

- Data analysis, including Data Mining methods [9]: analysis of data is a process of inspecting, cleaning, transforming, and modelling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. Data subsets can be analyzed independently using Data Mining methods. The results of the subsets' analysis are combined to get the whole picture of data. Data analysis usually is not suitable application for a Desktop Grid project because of low-speed communications between server and the clients. But Enterprise Desktop Grid can use a local area network which is usually much more reliable and much faster.
- Local (private, organization-wide) scientific research: in case of short-term or organization-wide research it is very hard to gather big enough community to set up a Desktop Grid project. The same time, the use of Enterprise Desktop Grid doesn't need to pay attention to a community of the project.

In addition, Enterprise Desktop Grid can be used as a cheap and scalable accelerator for a high-performance computing cluster.

4 Conclusion and discussion

Desktop Grid concept is an important part of the scientific high-performance computing. But the main problem limiting the use of Desktop Grids is narrow class of problems and necessity to keep a contact with community.

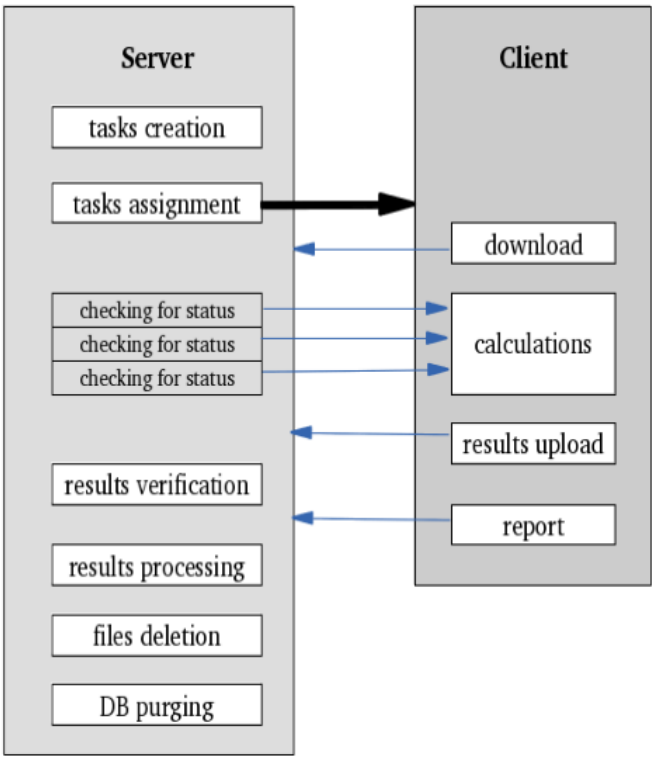

Fig. 2. PUSH model of client-server interaction.

Because of much faster and more reliable communication network, Enterprise Desktop Grid widens the concept of Desktop Grids.

Comparing to other concepts of high-performance computing, Enterprise Desktop Grid has some fruitful characteristics, such as resource level, class of suitable problems, trust and complexity and costs.

Enterprise Desktop Grid systems have some useful applications for small and medium enterprises as well as small and medium research groups and organizations.

References

1. Ian Foster, Carl Kesselman. The History of the Grid. [Online] http://www.ianfoster.org/wordpress/wp-content/uploads/2014/01/History-of-the-Grid-numbered.pdf
2. Climate Prediction [Online] http://www.climateprediction.net/
3. PrimeGrid [Online] http://www.primegrid.com/
4. SETI@Home [Online] http://setiathome.berkeley.edu/
5. BOINCstats. [Online] http://boincstats.com/
6. Volunteer computing for the LHC. [Online] http://lhcathome.web.cern.ch/
7. BOINC: Open-source software for volunteer computing. [Online] URL: http://boinc.berkeley.edu/index.php
8. S. J. Choi, H. S. Kim, E. J. Byun, M. S. Baik, S. S. Kim, C. Y. Park, and C. S. Hwang. Characterizing and Classifying Desktop Grid. IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID 2007), Sixth International Workshop on Global and Peer to Peer Computing (GP2P), pp. 743-748, May 2007.
9. E. Ivashko, A. Golovin. Partition Algorithm for Association Rules Mining in BOINC–based Enterprise Desktop Grid. Parallel Computing Technologies. LNCS Vol. 9251. 2015.
10. Renderfarm.fi The Publicly Distributed Rendering Service. [Online] http://www.renderfarm.fi/
11. Anette Weisbecker, Franz-Josef Pfreundt, Johannes Linden, Steffen Unger (Hrsg.). Fraunhofer Enterprise Grids. Business Cases. Fraunhofer IRB Verlag, Stuttgart, 2008. ISBN:978-3-8167-7803-5
12. Chien A., Calder B., Elbert S., Bhatia K. Entropia: Architecture and performance of an enterprise desktop grid system. J. Parallel and Distributed Computing 63, 597-610 (2003).
13. A. Luther, R. Buyya, R. Ranjan, S. Venugopal, Alchemi: A .NET-Based Enterprise Grid Computing System, Proceedings of the 6th International Conference on Internet Computing (ICOMP'05), June 27-30, 2005, Las Vegas, USA.
14. Chu X., Nadiminti K., Jin C., Venugopal S., Buyya R. Aneka: Next-Generation Enterprise Grid Platform for e-Science and e-Business Applications. Proceedings of the 3th IEEE International Conference on e-Science and Grid Computing (e-Science 2007), Bangalore, India (2007).
15. Sirajul Haque. Enterprise Grid Computing: A Potential IT Infrastructure for Small Business Enterprises. Journal of Management and Social Sciences. Vol. 6, No. 1, (Spring 2010) 44-51