The 8th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID 2018)

Contribution ID: **335** Type: **Sectional reports**

THE GAME CHARACTER OF COLLABORATION IN VOLUNTEER COMPUTING COMMUNITY

Monday, 10 September 2018 15:00 (15 minutes)

The paper shows the emergence of a new form of online scientific collaboration, the collaborative networks of volunteer computing (VC) participants. And it examines what makes a collaborative VC-project successful and determines the formation of VC-community. We report on data from a statistic online study of volunteers'activities and an online survey of VC-participants on several online forums, and discuss and analyze the emerging type of collaboration network of VC-volunteers using Brandeburger and Nallebuff's (1996) notion –the "co-competition".
The results can be significant for optimizing VC-management for solving problems that require large computational resources.

Summary

INTRODUCTION
The notion of "collaboration"is nowadays a central issue in many fields including business, science and even everyday speech. So, for example the result of the request in the Google browser gives greater than eight hundred millions links in 0,44 sec. We can say that this notion has turned into a buzzword. The usual definition of collaboration as it was done in Oxford dictionary is the action of working with some one to produce something. This notion is used in a variety of phrases to describe different phenomena. So, in this paper it will be used to glean knowledge that will deepen our understanding of an emerging type of scientific distributed computing and online communication.
One of the forms of the distributed computing is the volunteer computing (VC), a specific form of online activity of volunteers (> 4,5 mln.), usually ordinary computer users which have no relation to scientific or professional computer activities, and participants do not receive compensation for their work. So, we use the notion of volunteer computing as a type of a distributed computing in which unskilled computer owners can donate their spare computer resources to perform a computation of one or more large-scale research projects. VC relies on the client-server architecture; and volunteersdesktops, notebooks, mobile phones can be connected to form equivalent of a huge and super-power supercomputer (greater than 24,032 Teraflops).
The VC paradigm is divided into three main points:
The idea David Gedye of using multiple integrated computers for distributed computing in scientific tasks (David Gedye, 1994);
The idea of the network organization of communications (M. Castells, 1996; J. Arquila & D. Ronfeld, 2001);
The idea of distributed network of computers (>1,200,000) implemented in a BOINC software/platform (David P. Anderson, 2004)
As a whole VC-projects is based on two pillars:
 Technical task (computational): slicing a problem into thousands of tiny pieces which are then allocated (with a help of downloading project software) to a large number of individuals who volunteer their PC (notebooks, etc.) to scientific computing;
 Social task (participation): recruiting and retaining a large number of individuals to volunteer their computer resources, time and energy to the project, and facilitating their continuous contribution.
And while the computational aspect of VC received much research attention; the participative aspect remains largely unexplored. In order to understand the phenomenon of VC we ought to look in side it. And we discover the two spaces.
At first, it is computational VC-space. As it was mentioned earlier VC relies on the standard client-server architecture where clients (software programs) "talk"only to the project server and do not talk to one another

directly. So, in order to participate in a VC project, a user must download the project software from the VC project website and install it on their computer. Once installed, the project software contacts the project website and is assigned a task. Upon solving the task, the result is sent back to the project server and a new task is downloaded if the user so desires.

And the second, it is VC social structure. VC-participants are not a "sack"of volunteers. VC projects are by their nature decentralized and so, volunteer's communication through online form a communal virtual place. Participants in VC projects can either work alone or join a team and work together on the same project. Also all VC-projects have their own websites that are used to host the project software, disseminate information about performing the project and host a ranking table of volunteers activity. Also there are discussion forums and online groups in social media where they can discuss different aspects and technical problems of distributed computing. So, we can say that VC is a dispersed network of individuals and teams, project and team sites, forums. And the VC organizational framework consists of projects –typically academic-based research assignments (there are 57 projects, http://www.boincstats.com); individuals –Internet-connected low-powered computers = average person or organization (volunteers), not paid for their work, anonymous, working alone (just about 16,9%, survey May, 2018); VC teams –a long-term strategic alliance formed by volunteers, individual users of PC, in order to participate in solving large scientific computational task (just about 83,1%, survey in May, 2018); Internet resources –project websites with statistics in rank tables, discussion forums, etc

The implementation of VC project directly depends on the number of participants, PC (other machines) and time of their work. So, the main question of the project managers –Why do volunteers decided to participates in VC-projects? Most sociological studies of the motives of volunteers' behavior were based on the representation of the community as a set of participants and as a result, the research focuses only on the motivation of individual participants [Holohan A. and Garg A., 2006; Nov O., et al., 2014; Andreev. A., 2014; Kurochkin et al., 2015]. The background of such view is based on an idea that VC participants provide something for others (processing data for scientific projects) at their own cost (time, energy, opportunity costs, use of PC resources). E.g., factors which need for self-oriented motivations.

But we think that a motive of self-actualization is not enough to explain why millions of common people regular (day by day) participate in scientific computation. Our research indicates that the answer lies at the intersection of self-oriented motivation and the interactional and organizational possibilities emerging through the Internet. We suggest: online collaboration can capture peoples motivation better than only intrinsic motives. And the questions to be answered are:

- Is there any collaboration in volunteer computing?
- Does this collaboration affect a performance computing?

For answering these questions some statistical approaches (mainly clustering, graph visualization, etc.) are applied on several networks.

AIMS, OBJECTIVES AND METHODOLOGY

It is obviously that the implementation of VC-project directly depends on the number of participants, PC (other machines) and time of their work , and so we will show that along the project managers, VC-community figures out collaboration ways to speed up and expand the capability of their computer resources. In the study as a metric (measuring the amount of computational time and resources/efforts donated by each volunteer/team), we use conditional points (crédits) that are charged by project managers in the team rating tables posted on the statistical website (www.boincstats.com). The object of the study was Russian VC-community (www.boinc.ru). The object of the study is a lot of participants in the BOINC.RU community. Out of 107,227 teams, uniting more than 4,5 million participants in the VC projects, there are 821 Russian teams. Of 57 active BOINC-projects using the BOINC platform, there are 11 Russian projects.

The Data were collected in online survey, statistic analysis assembled in database; network of Boinc.ru community was visualized with a help of "Gephi"and "Force Atlas 2". To conduct a statistical analysis of the behavior of Russian participants in the VC, we used the data obtained with the websites www.boincstatistic.com and www.boinc.ru. Content of the database were unique participant identifiers, participant names, unique project identifiers, project names, number of units (crédits) of participants/teams for the last week, month, year and all the time, participants' ownership of projects, unique team IDs, team names, team memberships.

THE COLLABORATIVE NETWORKS EFFECT

As it was mentioned before participation in VC is voluntary and participants do not receive compensation for their work. So project managers/organizers in search of a mechanism to encourage participation in VC-projects use conditional points accrual mechanism; the number of these points («credits») depends on the provided capacities, the time of participation in projects, and other characteristics of the activity of volunteers and their teams. The availability of constant statistics for all projects, in addition to tracking various ratings, provokes various virtual competitions ("challenges") between participants and teams. And many of the projects create an environment for the competitors by volume computations are done, both individually and in the team event. Thus modus operandi of VC –is spirit of competition. If volunteers are members of a team, they are simultaneously competing with the other teams on the project on the more immediate goal of racking up the most contributions and coming out on top of the table of statistics documenting contributions. This form of cooperation and competition demonstrates a new type of online scientific collaborative network. As an example, we can cite the Russian project SAT@home [Andreev A., 2013], which is actively supported by national volunteers. For the period of each of these competitions the project's productivity increased by about 7,4 teraflops with the average value in 4,3.

Picture 1 The performance of the SAT@home project. The peaks correspond to the moments of the competition among teams of volunteers.

In order to highlight traces of collaborative networks we will allocate links between the active volunteers in a whole Boinc.ru community. The background of looking for the active participants is based on the idea that only a few participants of virtual communities are the most effective in obtaining the most important results. For example, in the study of the OSS projects was showed that only a small subset is significant, 85-90%% of the code was written by about 15 developers [Weber S., 2004].

For determining an active group of participants in VC-projects in Boinc.ru community we have used the accrued credits, assumed that a person took part in the project if he collected at least 50,000 points in it. So it appeared –200 volunteers, who participated in 2176 projects. To visualize and analyze links (simultaneous participation of two volunteers in a project) between volunteers we use the Gephi software and the Force Atlas 2 algorithm.

Having constructed a connectivity table, where aij –the number of general projects for two volunteers we construct a graph.

Picture 2 Visualization of links in group active volunteers (>50 000 credits, 2176 projects)

Increasing the participation threshold in the project to 100 000 crédits, followed by up to one million crédits hardly changes the picture evaluation of the participation of the 200 most active volunteers. In spite increasing the threshold the form of graphs does not change. It means that almost all active users are connected with some projects.

Picture 3. Distribution of active participants in the projects

In spite increasing the threshold the distribution of the participants in the projects are very similar. And as can be seen from the distribution and height of the peaks, which corresponds to the number of participants in the projects, these active volunteers participate in almost the same projects. The most active users are not inclined to participate in projects "on trial", they thoroughly approach the choice of the project and actively participate in the accruals.

If we look for the communication between participants of the largest team in Russian projects, ("Russia Team") and construct a graph, we see the graph of the links (participation) of volunteers in 7 projects (orange cluster is dedicated to the collaborative community).

Picture 4. Collaborative network

The unification of volunteers into this group, without a provocation, shows that these are not only active, but also interested community members who follow the BOINC space, analyze the information and agree on the preference for new projects. Such behavior of the community members testifies, very likely, about their interaction with each other on third-party network resources.

CONCLUSION

The most active and productive participants in VC exemplify "co-competition"(Brandenburger & Nalebuff, 1996) –collaboration within the network organization of a VC project. Our research indicates that the motivation of participation of millions of unskilled volunteers in VC-projects lies at the intersection of intrinsic motivation and the organizational possibilities emerging through the collaboration. In providing the means for channeling participants' motivations to compete and cooperate, VC-projects provide powerful insights into a new type of collaborative network

VC community model –collaborative network

REFERENCES

David P. Anderson. Boinc: A system for public-resource computing and storage. In Proceedings of the Fifth IEEE/ACM International Workshop on Grid Computing. IEEE (2004), 4-10. Доступно: https://boinc.berkeley.edu/grid_paper_04.pdf (дата обращения 13.06.2018)

Holohan, A. et al. Collaboration Online: The Example of Distributed Computing. // Journal of Computer-Mediated Communication. 2006, 10(4). Доступно: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2005.tb00279.x (дата обращения 13.04.2018).

Андреев А. Социология ГРИД. Троицкий вариант, № 166, 04 ноября 2014 года. с. 8. Доступно: http://trv-science.ru/2014/11/04/sotsiologiya-grid/ (дата обращения 13.04. 2018)

Oded Nov et al. Scientists@Home: What Drives the Quantity and Quality of Online Citizen Science Participation? PLOS One, April 1 2014. Доступно: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0090375/ (дата обращения 13.04.2018)

Якимец В. Н., Курочкин И. И. «Рейтинг проектов добровольных распределенных вычислений» // VII-я Международная конференция «Распределенные вычисления и грид-технологии в науке и образовании» (GRID 2016), 4-9 июля, 2016, г. Дубна, Россия. (Book of abstracts)

Victor I. Tishchenko. Behavioral patterns of volunteer computing communities//BOINC-FAST 2017: International Conference BOINC: FAST 2017. Proceedings of the Third International Conference BOINC-based High Performance Computing: Fundamental Research and Development (BOINC:FAST 2017) Petrozavodsk, Russia, August 28 - September 01, 2017. Edited by Evgeny Ivahsko, Alexander Rumyantsev//CEUR-WS. –2017. –P. 55-60/ Доступно: https://ceur-ws.org)/ (дата обращения 13.04.2018)

Kurochkin I. I., Posypkin M. A., Andreev A. A., Vatutin E. I., Zaikin O. S., Putilina E. V., Manzuk M. O. The activity of Russian chapter of international desktop grid federation // Distributed Computing and Grid-Technologies in Science and Education 2016. C. 36.
Тищенко В. И., Прочко А. Л. Виртуальное сообщество российских участников добровольных распределенных вычислений на платформе BOINC // Сборник презентаций и статей докладов НСКФ'2017// Доступно: http://2017.nscf.ru/prezentacii (дата обращения: 13.04.2018)

Primary author: Dr TISHCHENKO, Victor (RCF "Computer Science and Control" RAS)

Presenter: Dr TISHCHENKO, Victor (RCF "Computer Science and Control" RAS)

Session Classification: 7. Desktop grid technologies and volunteer computing

Track Classification: 7. Desktop grid technologies and volunteer computing