A Social Cloud for Public eResearch

Koshy John
School of Engineering and
Computer Science
Victoria University of Wellington
Wellington, New Zealand
Email: Koshy.John@ecs.vuw.ac.nz

Kris Bubendorfer
School of Engineering and
Computer Science
Victoria University of Wellington
Wellington, New Zealand
Email: Kris.Bubendorfer@ecs.vuw.ac.nz

Kyle Chard
Computation Institute,
University of Chicago and
Argonne National Laboratory,
Chicago, IL, USA.
E-mail: kyle@ci.uchicago.edu

Abstract—Scientific researchers faced with extremely large computations or the requirement of storing vast quantities of data have come to rely on distributed computational models like cloud computing. However, distributed computation is typically complex and expensive. The Social Cloud for Public eResearch aims to provide researchers with a platform to exploit social networks to reach out to users who would otherwise be unlikely to donate computational time for scientific and other research oriented projects. In this paper we explore the motivations of users to contribute computational time and examine the various ways these motivations can be catered to through established social networks. We specifically look at integrating Facebook and BOINC, and discuss the architecture of the functional system and the novel social engineering algorithms that power it.

I. INTRODUCTION

Scientific research increasingly relies on complex computation and large scale storage of scientific data, the scale of which cannot be provided by individual personal computers or even small clusters. Distributed computing models based on clusters, grids and more recently clouds, provide large scale capacity to scientists. However, obtaining funding to support IT infrastructure is often difficult due to the current structure of the funding agencies and licensing and ownership requirements imposed by commercial organizations. Access to national Grid infrastructures (TeraGrid, OSG) only supports selected projects and imposes strict time/resource restrictions on them. Several studies have also shown that conducting scientific research on commercial clouds often costs more than purchasing local resources [1], [2] and funding agencies are only now exploring models by which researchers can access public cloud time. This combination of factors significantly limits the processing power available to researchers.

Volunteer computing [3] is an alternative means of obtaining large computing resources – by getting the public to support specific projects by donating their spare computational and storage resources. The amount of computational time available to researchers is a function of the number of volunteers contributing at any given point of time. While there are a sizable number of volunteers who participate in volunteer computing (e.g. 2.2 million BOINC participants [4]), this is insignificant when compared to the 500 million active Facebook users [5].

The primary goal of the research presented in this paper is to integrate Social Networking and Volunteer Computing and thereby bring eResearch to the masses. The way in which we do this is based upon our earlier work creating the Social Cloud [6]:

A Social Cloud is a resource and service sharing framework utilizing relationships and policies established between members of a social network.

We have named this fusion of social networking, social cloud and volunteer computing, the *Social Cloud for Public eResearch*. The potential for growth is significant, an uptake of only 0.5% of Facebook users would equal the entire existing BOINC user base. In addition, BOINC has no infrastructure by which new projects can be advertised, for example, of the 2.2 million BOINC users, over 1.1 million contribute to SETI@Home, while some newer projects have as few as 786 volunteers [4]. The public visibility of a project has a clear impact on the number of volunteers that it garners. We see this as another intrinsic advantage of adopting a social network like Facebook, where posts, news feeds, and social incentives can be used to bring a new project into the public eye.

II. WHAT IS A SOCIAL CLOUD?

A Social Cloud is a scalable computing model in which heterogeneous resources contributed by users are dynamically shared amongst a group of "friends" in a social network. A Social Cloud benefits from an implicit level of trust between users and the associated socially corrective mechanisms that exist due to the real-world basis of the relationships represented. The cloud-based usage model enables virtualized (elastic) resource sharing through service-based interfaces exposed by members of the network.

One way of thinking about the Social Cloud is to consider that social network groups are analogous to dynamic Virtual Organizations (VOs) [7]. Groups, like VOs, have policies that define the intent of the group, the membership of the group and sharing policies for the group.

This model is illustrated in Fig. 1, where user-specific groups, defined by relationship types, are shown in the context of a social network. In this example group A is composed of only co-worker members, whereas group B is formed by family members and group C includes only friends. Clearly the level of trust and mechanisms for *social correction* (identifying incentives and disincentives for users to participate) differ between groups. This figure also highlights that social clouds

are not mutually exclusive, that is, users may be simultaneously members of multiple social clouds. Whereas a VO is often associated with a particular application or activity, and is often disbanded once this activity completes, a group is longer lasting and may be used in the context of multiple applications or activities.
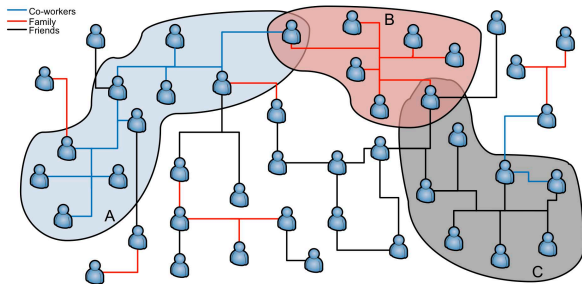

Fig. 1. Social Cloud overlay in a Social Network [6]. Three different Social Clouds are illustrated to highlight the use of relationships when establishing Social Clouds.

III. RELATED WORK

Well-known examples of volunteer computing include SETI@Home [8] and Folding@Home [9]. SETI@Home was started in 1999 to analyze radio signals coming from outer space in the hope of detecting signatures indicative of intelligent life. Folding@Home performs simulations of protein folding to provide a better understanding of the development of many diseases. Both projects far exceeded researchers' expectations, gathering huge resource pools and generating worldwide publicity. More recently *crowdsourcing* has emerged as a hugely successful distributed problem solving model in which large scale tasks are broadcast to, and solved by, an unknown group of amateur members of the public [10]. This is a more general model of volunteer computing in which the donated 'resources' are the contributors personal skills and their time.

The BOINC (Berkley Open Infrastructure for Network Computing) [11] platform was created as a generic volunteer computing middleware due to the widespread success of volunteer computing projects. As of June 2011, BOINC has over 50 supported projects, including many that contribute significantly to the global good. One example is Malariacontrol.net, that simulates the spread of malaria to determine minimal efficacy and duration of effects needed for a trial vaccine and also to optimize deployment of established treatments. Another is Rosetta@Home, that determines the shapes of new designs for three-dimensional proteins – in order to help find cures for intractable diseases such as cancer and HIV.

BOINC had more than double the processing power (5.6 petaflops [4], through its army of volunteers) of the fastest super-computer in the world (Tianhe-I of China with 2.6 petaflops [12]) in March 2011, although there are some inherent performance limitations in the volunteer model [13]. In July 2011 the top ranked super computer reached 8 petaflops, 3 times the power of the now second placed Tianhe-I.

Despite the notable success of BOINC, it has primarily relied on world-of-mouth publicity. There are also perceived (and real) barriers to entry for less-technical users. For example, it is difficult for a non-technical user to understand BOINC, discover projects, and install a BOINC environment for contribution. Nearly all current barriers stem from the fact that BOINC was originally created by, and for, a technical and knowledgeable audience.

BOINC took a major step towards addressing these barriers by collaborating with account management systems [14] like GridRepublic [15] to ease the management of multiple projects for volunteers. Before the introduction of account management systems, users who wanted to support multiple projects had to manually setup accounts with each of the projects and then manage contributions separately. This was a well-documented source of frustration for early users.

More recently BOINC has collaborated with Intel to create a Facebook application called Progress Thru Processors [16]. Progress Thru Processors has seen modest success, however we believe there is much greater success to be attained if the social aspects of social networking are fully embraced.

Several other middleware alternatives are also in public use, including Folding@Home and Distributed.net, which have fewer users than BOINC, and the commercial XGrid [17] from Apple and Grid MP [18] from Univa.

IV. MOTIVATION

The Social Cloud for Public eResearch represents a unique computing environment in which users of a social network are able to donate resources to scientific computing projects. In doing so, this work aims to take volunteer computing from more technically oriented users towards everyday users of a social network.

The adoption of Social Cloud computing provides a novel mechanism for leveraging social engineering to motivate and facilitate volunteer based sharing. Specific motivations for users and researchers are detailed in the following sections and referenced throughout the remainder of this paper.

A. Users

There are several reasons users participate in volunteer computing projects, in general users are motivated by altruistic or self-interested reasons. For example, altruistic users may have a desire to make a difference or have a strong interest in a particular field of research, whereas self-interested users are typically motivated by competition with regards to the contribution size (leaderboards) or a desire to be publicly recognized.

It is reasonable to assume that only a very small percentage of the general public is interested in volunteer computing. Generally, the projects themselves drive a user's desire to contribute, and in essence volunteer computing is just a means to an end for these users. Specific motivation of volunteer computing users have been studied in [19]. Briefly, the following key factors were identified as strong motivation for volunteers: potential impact of the science, probability of success, the

utility of the project, safety of the project, political signals, democratizing science, and personal benefits such as a sense of community, competition, personal interest and visual pleasure.

In current volunteer computing initiatives users are responsible for finding appropriate projects, weighing up which projects are most suited to their interests, and setting up and maintaining required volunteer software. Clearly this is a barrier for some, perhaps many, users. In a social cloud relationships between users can be used to share information and determine suitable projects for participation, while resources can be shared more easily through a simple social network application.

The social relationships defined in a Social Cloud present a unique way of providing many of these key motivating factors. Specifically targeted social algorithms and techniques can be used to both maximize the number of new volunteers, and, to keep them engaged and involved so that the computational time available to researchers grows quickly and sustainably. The Social Cloud provides a single integrated management view (within the social network) of all projects a user contributes to easily monitor current activity and also explore new projects.

B. Researchers

Project owners face similar challenges as they must advertise their projects, and, determine and implement appropriate motivation mechanisms to encourage user contribution. In addition to benefiting users, the Social Cloud for Public eResearch aims to provide increased publicity to different volunteer projects.

Through the Social Cloud researchers will be able to connect with the people (and groups) that support their research. It is therefore beneficial for all parties as research can truly become participatory and inclusive. Significant public interest in a project can result in new support and interest from external sources of funding. It can also be a tool to influence political interests in areas of research that polarize public opinion. It is conceivable that an increase in the number of volunteers may unintentionally favour high visibility projects and therefore discourage researchers from trying to exploit volunteer models for newer less established projects. It is therefore very important to ensure that such projects are not disadvantaged through appropriate algorithms and policies.

In addition to the benefits described, the Social Cloud model lowers the barriers of entry both for volunteers and researchers through its resource sharing framework, thereby providing a large amount of processing power that would otherwise be irrecoverably wasted.

V. Design Background

For volunteer computing to be revitalized there needs to be an amalgamation of a mature volunteer computing platform and a high-potential source of new volunteers. In designing a proof of concept for the Social Cloud for Public eResearch, we chose BOINC as our volunteer computing platform. This decision was based on the maturity, modularity, and proven scalability of the middleware.

A. Contributing with BOINC

BOINC offers multiple ways for users to contribute idle processing cycles on their systems to a project. Before contribution a user must download the BOINC client software on to their system, provide it with a project URL, email address and password. Users can choose to support more than one project by allocating *resource shares* [20] to each project. Resource shares are not percentages, they are a reflection of the portion of the total resource available to a specific project amongst all the projects attached to a given computer. Users use their discretion in determining resource shares and selecting projects. This process may overwhelm less technical users who don't know others who are already contributing to BOINC and are willing to offer advice and help.

When the BOINC client runs on a system, it downloads work units periodically from the various selected projects. It processes these work units, sends the results back to the project servers and claims credit for the work done. Each project has its own specific method of verifying the work done (e.g. quorum-based replication) - if the work unit returned is validated, the user receives credit. Credit is not granted if the result is returned after a set deadline or if the result was found to be inaccurate. The credit system is designed to discourage cheating and to encourage users to donate more by creating a sense of competition around credits earned. There are a number of credit statistics sites like BOINC Stats [21] that maintain user rankings.

The easiest way for users to manage multiple projects is to rely on an account management system [14]. Account management systems ease the process of joining and contributing resources by allowing the user to set up a 'meta-account' over multiple selected projects. The account management system creates accounts with the selected projects on behalf of the user. The user can direct the BOINC client on their system to connect to the account management system (rather than a BOINC project) along with their credentials, the account manager then acts as a proxy between the user and the project. They can use the account management system interface to add/remove projects and set resource shares. BOINC has published a set of WebRPCs [22] that specify how account management systems and project servers should communicate. There are also a set of account manager RPCs [14] specifying how the BOINC clients and account management systems communicate.

Account management systems do not have any social features that would bring in new users and keep existing users involved. The Progress Thru Processors Facebook application provides a streamlined way to create a GridRepublic account from within the Facebook platform. While there is some account management functionality within the Facebook application, users are redirected to GridRepublic to perform most non-trivial tasks. Overall, there is still a lot of scope to integrate volunteer computing models into Facebook, and leverage social connections between users and their associated incentives.

VI. SOCIAL ENGINEERING

Unlocking the true power of social networks requires delving into complex social engineering, for example motivating user behaviour based on social incentives. Current volunteer platforms like BOINC do not explicitly consider social engineering. However, it is our view that social engineering should be considered an important factor in volunteer computing. The underlying social network in the Social Cloud provides deep access to the social relationships between users and therefore an opportunity to exploit social mechanisms to encourage and maintain contribution.

We attempt to motivate individuals by tapping into their sense of competition, by playing to their desire for public recognition and by relying on social pressure from their friends. This is also known as *gamification* - the process of using mechanisms typically found in games to induce desired behaviours in a non-game application.

Appealing to a user's sense of competition and their desire for public recognition has traditionally involved individual leaderboards on BOINC statistics sites [21]. However, this is restricted to the credits earned and there is no real incentive to bring in new users. The exception is when teams are established [23] as the credits earned are pooled. This has proven to be a successful strategy for growing computational contributions, however, additional potential lies in tying it to social networks and extending it to more than just computational credits.

Given the social context created by the Social Cloud, users are able to compete directly with their friends - people they have far stronger social relationships with. This mechanism acts both to encourage increased individual contribution but also to encourage new friends to join in.

In the following sections we describe the structured incentives we have created to work towards the various goals of the Social Cloud for Public eResearch. At this point, it is important to distinguish between *the Social Cloud*, which is the overarching reference to the application itself and the various actors that actively participate, and *a user's social cloud*, which is the set of their friends that are members of the Social Cloud.

A. Easing the Process of Joining

When a potential user arrives at the Social Cloud Facebook application, there are a number of steps that they need to follow before they become valuable volunteers to BOINC:

1) Understanding how volunteer computing works and why it is important.
2) Selecting projects to support.
3) Choosing resource shares.
4) Installing the BOINC client and configuring it properly.

The first step is critical for capturing the user's interest and keeping them motivated through the subsequent stages. From a social engineering perspective, this can mean including compelling hooks from their friends and providing various *calls to action*.

To ensure that new users find it easy to select projects that match their own interests, we generate an *interest signature* for them based on their selections in a simple form. An interest signature defines a point in n dimensional space describing an individual user's specific areas of interest, each dimension represents a well-defined field of interest.

The interest signatures that are obtained from users are normalized to enable reliable comparison. We calculate the interest signature distance, D_{uf}, as defined in equation (1), between a user, u, and all their friends, U_f, based on their interest signatures, I_u and I_f, to identify friends with the most similar interests to that particular user. The shorter the Euclidean distance between two interest signature points, the more similar the interests of the two users.

$$\forall f \in U_f, D_{uf} = \sqrt{\sum_{i=0}^{n-1} (I_u[i] - I_f[i])^2} \qquad (1)$$

We also introduce a *project signature* which is the average interest signature of a set of contributors to a given project. By calculating *signature distances*, we are able to highlight projects that were chosen by users, either globally or within the user's set of friends, that have interest signatures similar to the new user.

We calculate the signature distance, D_{up}, between a user's interest signature, I_u, and a project signature, I_p, as follows:

$$D_{up} = \sqrt{\sum_{i=0}^{n-1} (I_u[i] - I_p[i])^2} \qquad (2)$$

This enables us to help the user easily pick projects that appeal to them.

Additionally, new users that are not sure of which projects to choose can align their choices with those of their friends. Users that want a reliable personal opinion on a project are directed to *project champions* within their group of friends.

Once users select projects, they are required to select individual resource shares for each of the projects. Recall, a resource share is a reflection of the percentage of the computational time that a project gets on a user's machine. Users can select their own shares but we also provide recommendations based on normalized averages of the shares of their friends. This helps both to reduce technical barriers to entry and also to catalyse engagement by encouraging competition.

Finally, users are given instructions on installing and configuring the BOINC client, and are directed to their friends in the event that they need assistance at any point.

Figure 2 shows a screen shot of the prototype, in this figure the user is shown a list of BOINC projects that are most suited to their interests. This list of projects are selected based on the distance between the project signature and the user's own interest signature. Users are also shown the number of their friends that support each project and may select projects and shares from this view.
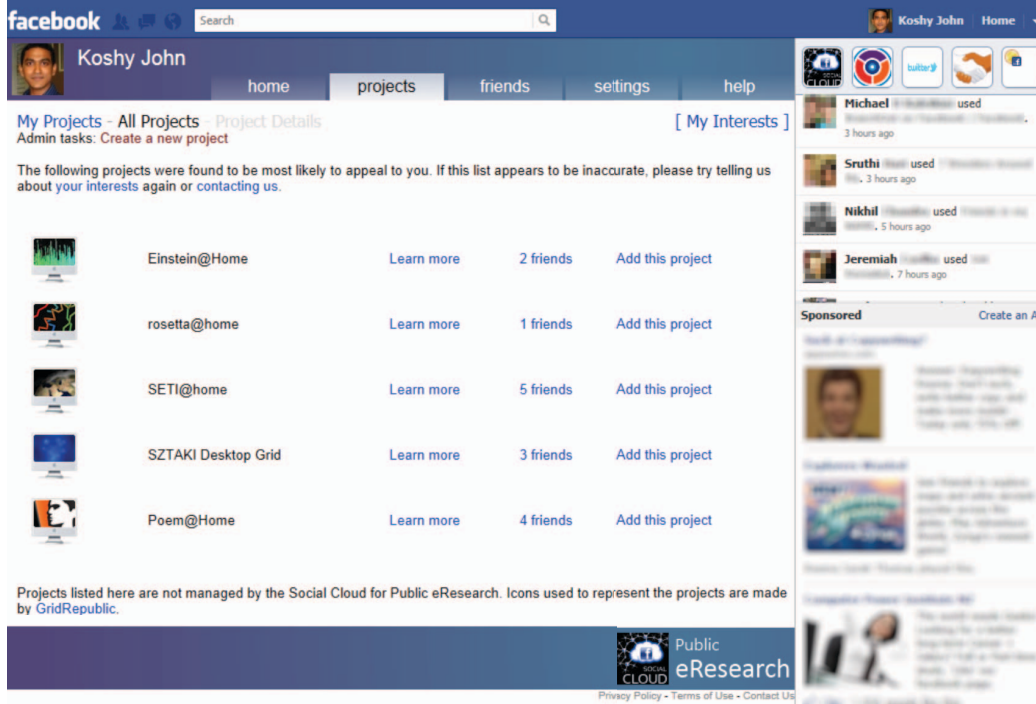
Fig. 2. A screen shot showing the project suggestions based on interest and project signature distances.

B. Incentivizing Involvement, Contribution and Growth

We have developed mechanisms akin to gamification to ensure that users contribute to growing the user base and the computational time available to BOINC projects. Scores related to various goals are calculated periodically to establish user rankings and to identify the best contributors within each user's set of friends.

We have introduced three high scoring sets of users - *project champions*, *social anchors* and *compute magnates* based off *project credits*, *social scores* and *compute scores* respectively. These labels are local to the view of each user and are reserved for friends that fall into the top bracket in a particular score. Project credits are routinely queried from project servers while social scores and compute scores are periodically recalculated using algorithms detailed in the following sections.

To motivate users and to enable them to measure their progress we utilize leaderboards specific to each user's personal social cloud. Significant changes in positions on the leaderboard are published to the user's Facebook wall – this in turn is expected to create interest in the application in addition to giving the user a sense of recognition.

There will be situations where a there might not be enough data to work with, like when a user with no friends joins the Social Cloud. In those situations, we fall back to values based off the Social Cloud as a whole.

Project Champions

Project champions are considered to be the "biggest" contributors to a given project and therefore represent the best go-to person for a user from within their friends list if they need to know more about a project. Due to their high contributions they are expected to champion the cause of the project. Project champions are identified based on the total credits that they have earned contributing to a given project.

It is easier to become a project champion of a less popular project, we therefore expect the desire to become a project champion to also increase the computational time that smaller projects receive. This is an important goal of the Social Cloud.

Social Anchors

Given the underlying social nature of the Social Cloud, it is expected that the majority of users will be introduced to it by their friends either explicitly (through an invitation or status message on their Facebook wall) or serendipitously where a potential volunteer may chance upon a report of the contributions of their friends in their Facebook news feed.

Existing Social Cloud users are incentivized to bring in their friends into the Social Cloud by tying actions in this direction to a *social score*. The top bracket of friends in terms of social scores are identified as *social anchors* in the user's personal social cloud. Breaking into that top bracket in their social cloud earns the user the title of social anchor.

Social Scores

$$\forall u \in U, Sv_u = \frac{1}{n_f + 1} \qquad (3)$$

$$\forall u \in U, Ss_u = \sum Sv_{uf} \qquad (4)$$

A user's social score represents a measure of their continuing contributions to the growth of the Social Cloud. Social scores are used to motivate existing users to encourage their friends to join the Social Cloud. New users with the least number of existing friends in the Social Cloud are of high value because they are less likely to have joined otherwise and have high potential to bring in new users, the users that are responsible for bringing them in are appropriately rewarded with a higher boost to their social score. New users with a lot of existing friends in the cloud are of less value because the effort required to get them to join is likely to be lower. The increase in the social score of the users associated with the new user would be consequently lower.

To this end, every user has an associated social value, Sv_u as defined in equation (3). The higher the number of friends, n_f, they have in the Social Cloud the lower their social value. Social values of a user's friends (in the Social Cloud), Sv_{uf}, add up to give the user their social score, Ss_u as defined in equation (4). For a user to maintain a high social score, they would need to keep recruiting users who are less likely to join.

The social value of users who have increased the number of friend connections in the cloud since they joined will decay by virtue of that fact. This serves the dual purpose of disincentivizing users adding existing Social Cloud users as friends on Facebook to boost their social score and to ensure that users don't rest on their social score achievements.

Social anchors are the key to growing the pool of users, and the title is recognition for their continuing contributions in this direction.

Compute Magnates

Compute magnates represent the top bracket of friends generating *valuable* computational time for the Social Cloud through their *own* social clouds. The computational time is based on the calculation of individual compute scores. This title was shaped to serve the dual purpose of incentivizing higher computational time contributions and application of social pressure on friends to maintain or improve on their contributions.

Compute Scores

$$\forall u \in U, Cr_u = \sum C_{u30} \qquad (5)$$

$$\forall u \in U, Cv_u = \frac{Cr_u}{n_f + 1} \qquad (6)$$

$$\forall u \in U, Cs_u = \sum Cv_{uf} \qquad (7)$$

Compute scores are a reflection of how much credit a user and their set of friends generate in a rolling 30 day window, C_{u30}. They are used to encourage users to ensure that their friends are contributing computational time with regularity. Like social scores, users are incentivized to focus on people with fewer friends in the Social Cloud, and disincentivized from adding friends simply to boost their compute scores.

The rolling credit value of a user, Cr_u, is used to generate their compute value, Cv_u, by dividing it by the number of friends, n_f, in their social cloud. Compute scores, Cs, for every user are then generated by summing up the compute values of all their friends as shown in equation (7).

Because of the method of calculation, users who have friends that contribute less than what they (the users) stand to gain in terms of their own compute value may feel incentivized to remove those friend connections. But this would work only through breaking the Facebook friend relationship itself and we feel that most relationships are strong enough for the user to work on getting their friend to contribute more instead.

VII. THE ARCHITECTURE

The Social Cloud for Public eResearch can be visualized as shown in Figure 3. It is essentially a privately hosted multi-faceted application designed to work with BOINC and Facebook.

From the perspective of project servers and volunteer PCs running the BOINC client, the Social Cloud for Public eResearch is an account management system. From the perspective of users (volunteers), the Social Cloud is a socially aware account management system which runs as a Facebook application. The existence of the Social Cloud is transparent to researchers and requires no additional effort on their part to support.

As discussed previously, the Social Cloud considers a number of factors that may prevent a new user from engaging in volunteer computing, and attempts to address those by tapping into their social circles.

Users can add and remove supported projects from within the Facebook application, they can also set relative resource shares (percentages of the total computational time donated) for the various projects that they choose to support. This information is used to communicate with the various project servers and to control the BOINC client running on the user's PC.

A. Facebook

The Social Cloud is built on Facebook for reasons detailed earlier. Facebook allows externally-hosted applications to run transparently within the Facebook UI. Access to social information is provided through the Facebook Graph API [24].

The Graph API exposes access to the underlying social graph that contains users and their connections with other nodes in the graph (people, photos, events, pages, etc.). To access the Graph API, both the user and the application must be authenticated by Facebook using the OAuth protocol [25]. The power and potential of the Graph API combined with the vast user base that Facebook has made it the obvious social network for this project.

The social data required to provide a meaningful experiences for our users is obtained through the Graph API. We also extend the Social Cloud experience back deep into Facebook by manipulating Facebook objects in the same manner.
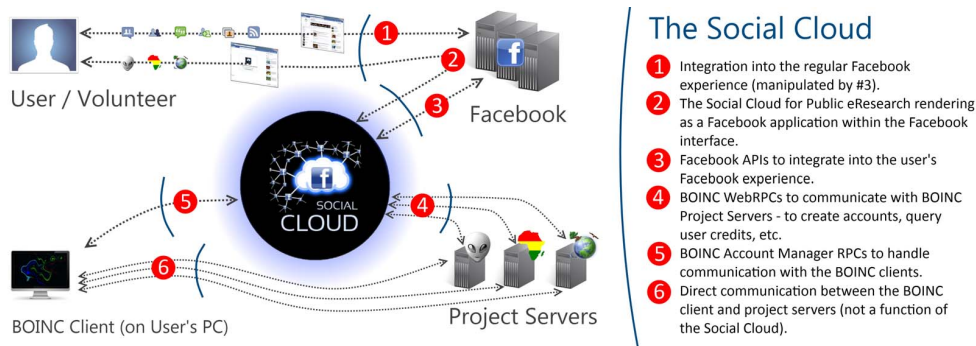
Fig. 3.   Architecture of the Social Cloud for Public eResearch

B. BOINC Project Servers

As far as BOINC project servers are concerned, the Social Cloud for Public eResearch is just an account management system. As long as a project supports BOINC's published Web Remote Procedure Calls (WebRPCs) [22], the Social Cloud account manager can support it.

The WebRPC model assumes every RPC to be an HTTP GET transaction, the input parameters are represented as a set of parameterized GET arguments. The resultant output is an XML document with well-defined fields that is parsed by the social cloud to let users monitor their contributions and also to feed our social engineering algorithms described in the previous section.

C. BOINC Clients (on Volunteer PCs)

Locally deployed BOINC clients can be attached to an account management system in various ways. Data about the account manager can be bundled with the installer, or the user can specify the account management system URL (in this case, it is the URL to the Social Cloud for Public eResearch). The user will provide authentication details for the BOINC client to obtain their project and resource share preferences from the Social Cloud.

The BOINC client communicates with the Social Cloud account management system using Account Manager RPCs published by BOINC. Once the client has processed data relating to the projects that the user supports, it attaches itself to each of the project servers directly and starts pulling information for processing.

D. Interactions

To help understand the architecture we describe a basic usage scenario for the Social Cloud through a sequence diagram in Figure 4.

The process starts when a Facebook user discovers the Social Cloud for eResearch application. If they choose to add the application, permissions for the required user data are requested through Facebook.

Once the application has been added, the user is presented a form to determine their interests in various research areas – this is used to generate their interest signature. The interest
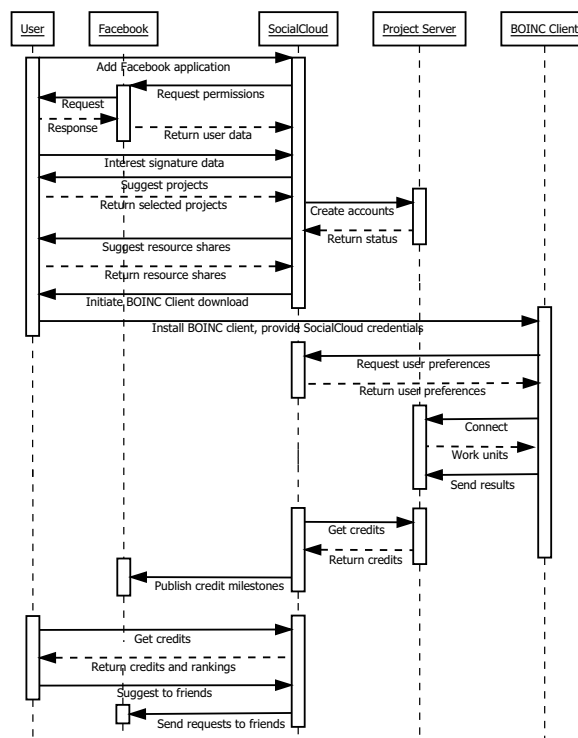

Fig. 4.   A simple example of interactions between all the actors associated with the Social Cloud for Public eResearch.

signature generated is then compared against the project signatures of all the available projects to determine appropriate projects ordered (by projected interest) for this user. The user can then select any projects that appeal to them and the Social Cloud proceeds to create accounts for them at each of the various BOINC project servers on their behalf.

Suggestions are made to the user on the resource shares that they should allocate to the various projects that they have selected. These suggestions are based on the normalized resource share values of their friends. This allows for meaningful competition in the future.

The user is then prompted to install the BOINC client and provide credentials to connect to the Social Cloud. The

BOINC client pulls information regarding the projects that the user has selected along with their resource shares from the Social Cloud. It connects to each of the project servers and downloads work units for processing. In due course, the results of the processing are sent back to the project servers. Each project server verifies the results obtained and grants credits as appropriate.

The Social Cloud routinely queries credits for every user from individual project servers. If a user is found to have achieved a credit milestone in a project, it is published to their Facebook wall. This is visible to friends and will hopefully generate interest. The user can also view the application at their convenience to check on their progress and that of their friends. They may also suggest the Social Cloud for Public eResearch to their friends to help drive its growth (and increase their social score).

In addition, the Social Cloud periodically processes data available to it to establish rankings and achievements for users based on the algorithms described previously.

VIII. Implementation Status

The core implementation is complete, and the Facebook application/user interface is currently being refined and tested. The screenshot shown in Figure 2 is taken from the running application, although populated with projects obtained from GridRepublic – as there is no equivalent set of projects exclusively managed by our system yet available.

IX. Conclusion

The major purpose of the Social Cloud for eResearch is to increase the uptake of public eResearch, or volunteer computing, through social influence applied via social networks. This involves identifying three different roles that incentivize users by rewarding contributors. The Social Anchor role is awarded to those who are most active at bringing new recruits into the social cloud. The Compute Magnate role is awarded to those who both through their own contribution, and that of their friends, bring the largest pool of resources into the entire social cloud. The third role is the Project Champion, which is awarded to those who contribute strongly to a specific project.

In addition to these roles, we have introduced interest signatures, making it easier for users to choose projects based on the interests of their friends. This will also contribute to the formation of communities embedded within the social cloud.

We have presented an architecture that acts as a Project Manager and integrates Facebook, BOINC project servers and clients. This exploits an existing mature middleware platform (BOINC), and Facebook – the world's largest Social Network. We have implemented this architecture and have a functional prototype that is entering a user testing phase.

We envision that with the unique strengths of this approach to volunteer computing we will see significant increases in the processing power available to BOINC based projects. If we are able to recruit 1% of the current user base of Facebook to become BOINC volunteers, we will effectively double or even triple the present number of BOINC volunteers. Given some

of the vital areas of research that many of the BOINC projects are involved in, we think this is a goal worth achieving.

References

[1] J. Wilkening, A. Wilke, N. Desai, and F. Meyer, "Using clouds for metagenomics: A case study," in *CLUSTER*, 2009, pp. 1–6.

[2] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good, "The cost of doing science on the cloud: the Montage example," in *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, ser. SC '08. Piscataway, NJ, USA: IEEE Press, 2008, pp. 50:1–50:12.

[3] Anderson, D.P. and Fedak, G., "The Computational and Storage Potential of Volunteer Computing," in *Cluster Computing and the Grid. CCGRID 06*, vol. 1, May 2006, pp. 73–80.

[4] BOINCstats, "BOINCstats | BOINC combined - Credits Overview," http://boincstats.com/stats/project_graph.php?pr=bo, June 2011.

[5] Facebook, "Statistics | Facebook," http://www.facebook.com/press/info.php?statistics, June 2011.

[6] Kyle Chard and Kris Bubendorfer and Simon Caton and Omer Rana, "Social Cloud Computing: A Vision for Socially Motivated Resource Sharing," *IEEE Transactions on Services Computing*, vol. 4, no. 4, 2011.

[7] Foster, Ian and Kesselman, Carl and Tuecke, Steven, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," *International Journal of High Performance Computing Applications*, vol. 15, no. 3, pp. 200–222, August 2001.

[8] E. Korpela, D. Werthimer, D. Anderson, J. Cobb, and M. Leboisky, "SETI@home-massively distributed computing for SETI," *Computing in Science Engineering*, vol. 3, no. 1, pp. 78–83, jan/feb 2001.

[9] A. Beberg, D. Ensign, G. Jayachandran, S. Khaliq, and V. Pande, "Folding@home: Lessons from eight years of volunteer distributed computing," in *Parallel Distributed Processing. IPDPS 2009.*, May 2009, pp. 1–8.

[10] J. Howe, "The Rise of Crowdsourcing," Wired Magazine http://www.wired.com/wired/archive/14.06/crowds.html, June 2006.

[11] D. Anderson, "BOINC: a system for public-resource computing and storage," in *Grid Computing. Fifth IEEE/ACM International Workshop on*, Nov 2004, pp. 4–10.

[12] Top 500 Supercomputer Sites, "November 2010 | Top 500 Supercomputing Sites," http://www.top500.org/lists/2010/11, June 2011.

[13] F. Costa, L. Silva, and M. Dahlin, "Volunteer cloud computing: Mapreduce over the internet," in *Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW)*.

[14] BOINC, "AccountManagement - BOINC," http://boinc.berkeley.edu/trac/wiki/AccountManagement, June 2011.

[15] GridRepublic, "GridRepublic | GridRepublic | BOINC Volunteer Distributed Grid Computing," http://www.gridrepublic.org/index.php?page=about, June 2011.

[16] Intel, "Progress Thru Processors | Facebook," http://www.facebook.com/progressthruprocessors, June 2011.

[17] Kramer, D. and MacInnis, M., "Utilization of a local grid of Mac OS X-based computers using Xgrid," in *High performance Distributed Computing.*, June 2004, pp. 264–265.

[18] Venkat, J., "Grid computing in the enterprise with the UD MetaProcessor," in *Peer-to-Peer Computing*, 2002, p. 4.

[19] Unofficial BOINC Wiki, "How to decide on Resource Share - Unofficial BOINC Wiki," http://www.boinc-wiki.info/How_to_decide_on_Resource_Share, June 2011.

[20] ——, "Resource Share - Unofficial BOINC Wiki," http://www.boinc-wiki.info/Resource_Share, June 2011.

[21] BOINCstats, "BOINCstats | BOINC Combined - User stats," http://boincstats.com/stats/boinc_user_stats.php?pr=bo&st=0, June 2011.

[22] BOINC, "WebRpc - BOINC," http://boinc.berkeley.edu/trac/wiki/WebRpc, June 2011.

[23] BOINCstats, "BOINCstats | BOINC Combined - Team stats," http://boincstats.com/stats/boinc_team_stats.php?pr=bo&st=0, June 2011.

[24] Facebook, "Graph API — Facebook Developers," http://developers.facebook.com/docs/reference/api/, June 2011.

[25] ——, "Authentication — Facebook Developers," http://developers.facebook.com/docs/authentication/, June 2011.