Coordinated Volunteer Computing

David P. Anderson
Spaces Sciences Laboratory
University of California, Berkeley

7 October 2017

Abstract: Since its inception in 2004, BOINC-based volunteer computing has been decentralized: scientists create and promote their projects independently, and volunteers discover, evaluate, and choose from among these projects. This model has limited the adoption of volunteer computing. To move beyond these limits, we propose a new model in which a central "coordinator" provides a unified volunteer interface and manages the allocation of resources.

1   Models for volunteer computing

"Volunteer computing" (VC) does high-throughput computing using consumer resources volunteered by their owners. A platform for VC, such as BOINC, defines standard processes by which scientists get computing power, and by which volunteers learn of VC and make their resources available to scientists. These processes define a "model" for VC.

1.1   The project ecosystem model

BOINC's original model was based on these processes:

- Scientists get computing power by a) learning about BOINC and b) creating and operating a BOINC "project", comprising a web site and a job dispatcher. They recruit volunteers by publicizing their project, and by providing web content describing their research, scientific credentials, and past results.
- Volunteers discover VC via the publicity of a project P, which takes them to P's web site. This directs them to download the client software from the BOINC web site. When the BOINC client starts up, the volunteer is shown a list of projects, from which they select P, thus "attaching" the device to P. The volunteer, perhaps at a later time, surveys the other available projects, visits their web sites, and may choose to attach to some of them. (The BOINC client lets volunteers attach to multiple projects, and to control the division of resources among them.)

The intention of this model is 1) to create a dynamic "ecosystem" of projects that compete for computing power by promoting themselves and their science, and 2) to create a population of

volunteers that periodically evaluate the set of projects and make informed decisions, based on their personal views, about how to allocate their computing resources.

This "project ecosystem model" has had some success: there have been about 40 projects, many of them doing significant published research. However, this set is fairly static: only a few significant new projects have arisen in recent years. We think this is due to the model; in particular:

1. Creating and operating a VC project is harder than expected: it requires a combination of resource and skills (sysadmin, DB admin, web design, PR/outreach) that few academic research groups have.
2. For a research group, trying to use VC is a risk. There's a substantial investment, with no guarantee of any return, since no one may volunteer. Adding a VC component to a grant proposal adds uncertainty and weakens the proposal.
3. The computing needs of many research groups are sporadic; e.g. they need a big chunk of throughput every now and then. For such groups, buying computing time on a commercial cloud may be cheaper than using VC.

Similarly, although VC has attracted several hundred thousand volunteers, this population is slowly declining, it consists primarily of tech-savvy males, and most volunteers are "locked in" to a few projects and don't actively seek out new ones. Again, the underlying issues are inherent in the model:

1. The complexity and technical jargon of the BOINC interface confuses and drives away many computer owners.
2. Attracting volunteers is a marketing problem. It's difficult to do effective marketing when there are dozens of competing brands (i.e. projects names).
3. Most volunteers aren't interested willing to repeatedly survey and assess a large number of projects.

We recognized early on that it was unreasonable to expect volunteers to browse lots of project web sites, and so we added a feature called the "account manager architecture". An account manager (AM) is both a web site and a web RPC server. Instead of attaching directly to projects, a volunteer can attach their device to an AM. The BOINC client periodically issues an RPC to the AM. The RPC reply contains a list of projects to which the client should attach.

This architecture was used by independent developers to create two AMs – Gridrepublic and BAM! – that allow volunteers to browse and select projects on a single web site. This eliminates the need to browse lots of separate web sites in choosing projects, and it also makes it possible to efficiently manage the attachments of multiple devices. However, the AMs did not solve the basic problems of the model.

1.2   The coordinated model

To address the problems of the project ecosystem model, we propose a new model involving a central "coordinator" that performs several high-level functions, including the global allocation of

computing resources. The coordinator includes an XSEDE-type mechanism for determining the allocation policy.

In the coordinated model the scientist and volunteer processes are as follows:

- Scientists can create BOINC projects as they currently do. However, rather than having to recruit volunteers by creating a web site and publicizing their project, they can apply for registration with the coordinator. The coordinator allocates computing resources among registered projects.
- We encourage the addition of VC back ends to existing HPC providers such as supercomputing centers and science gateways. The large population of scientists using these providers benefit from VC (via shorter latencies and lower costs) with no effort on their part, and potentially with no knowledge that VC is being used.
- Volunteers interact through the coordinator web site. As part of the registration process, they indicate their "science preferences" – which areas of science they do or do not want to support – and their preferences for the location of the research. They attach their devices to the coordinator, which uses the AM architecture. They do not explicitly choose projects; rather, the coordinator dynamically decides what projects to attach each device to. It does this in a way that attempts to honors both volunteer preferences and allocation targets.

This model has a number of possible advantages:

- A scientist, group, or institution that is considering using VC can apply at the outset for an allocation. If this is approved they are guaranteed to get a certain level of resources. The risk of creating a project is eliminated.
- The coordinator's "brand" acts as a unified brand for VC. Publicity campaigns (mass media, social media, co-promotions, etc.) refer to this brand, rather than the brands of individual projects. This allows more effective promotion. It also eliminates the need for BOINC projects to invest a lot in creating attractive web sites; in principle they don't need a web site at all.
- It gives volunteers an interface defined in terms of science goals, which have been shown to be the most powerful incentive for participation.

2   The coordinator committee

We will establish a "coordinator committee" to determine coordinator policies, including its resource allocation policy. The committee will include representatives of the U.S. and European scientific funding agencies, the leaders of the coordinator project, and selected members of the volunteer community. It could potentially include representatives of projects, though this would be a potential conflict of interest.

2.1   Project selection

The committee must decide what projects to include. This should be based on published criteria such as:

- The project's computing is toward a scientific or technical goal (broadly interpreted to include things like mathematics and cryptography).
- The project is non-commercial.
- The project's leadership has a certain level of qualification (as demonstrated, e.g., by publications).
- The project can prove that it understands and follows certain security practices such as code-signing.

The committee will define a process by which potential new projects can apply for inclusion. For example, a scientist could apply for and receive inclusion, then submit a grant proposal to fund the development of the project.

The committee may choose to charge a fee for inclusion, with the possibility of waiving the fee in special cases. The proceeds would go toward the development and maintenance of the coordinator.