BOINC: Middleware for Volunteer Computing

The world's largest computing resource is the pool of consumer products – desktops, laptops, tablets, mobile devices – owned by the general public. There are currently about 1 billion privately-owned PCs and 100 million GPUs capable of general-purpose computing, with a total capability of about 20 ExaFLOPS, growing to 100 ExaFLOPS by 2012. The consumer resource pool has several advantages relative to traditional HPC resources:

- Its power and hardware are paid for by consumers, and the hardware is continuously upgraded to state-of-the-art components.
- It's self-maintaining: consumers fix their own software and hardware problems.
- Its interconnection network – the commodity Internet – has a rapidly expanding capacity.
- Consumer products are the main focus of computing research and development. Thus, for example, the fastest and cheapest GPUs are those in consumer-market boards.

To reach Exa-scale capabilities quickly, to maintain rapid growth, and to minimize cost, the infrastructure for scientific computing must focus on consumer resources rather than traditional HPC resources such as clusters and supercomputers.

Volunteer computing is a mechanism by which consumer resources can be used for scientific computing. It allows computer owners to volunteer for particular computational research projects or groups of projects. Compute jobs are then executed in the background on the volunteer's computers, and data files are stored on their disks. Volunteers are motivated by various factors: support for the goals of the research, participation in online communities, and competition based on computational contribution. There are currently about 50 volunteer computing projects, 500,000 volunteers, and 1 million volunteered computers, providing an average throughput of 10 PetaFLOPS.

Volunteer computing is a paradigm shift: scientists gain computing power not by getting grants and buying hardware, but by persuading the public that they're doing good science.

BOINC, an NSF-funded project at UC Berkeley, develops the primary middleware platform for volunteer computing. BOINC provides server software that lets scientists create volunteer computing projects, and client software (available for all major platforms) that lets volunteers participate in any combination of these projects. The BOINC software addresses issues such as security, heterogeneity, scale, sporadic availability, and unreliability. The project's staff consists of 3 programmers; many of its functions (testing, documentation, customer support, translation) are done by volunteers.

BOINC is used by projects from many institutions, doing research in many areas, including astrophysics, cosmology, climate study, biochemistry, epidemiology, environmental science, cognitive science, particle physics, nanotechnology, quantum computing, genetics, mathematics, and seismology. Some of these projects involved compute-intensive analysis of large data sets; for example, the Einstein@home project analyzes Petabyte-scale data from the LIGO gravitational-wave observatory and the Arecibo radio observatory.

Almost all computational scientists could benefit from volunteer computing, but only a tiny fraction currently do. The main barrier, we believe, involves organizational level. Projects should be operated not by individual research groups, but by organizations such as universities, research labs, and funding agencies, in which case they would serve many research groups and could leverage the organization's existing IT and publicity resources.