Volunteer computing

Nov. 14, 2016

Kate Cowles
374 SH, 335-0727
kate-cowles@stat.uiowa.edu

What is volunteer computing?

- opportunity for any computer owner to contribute to research in energy, climate change, medicine, astronomy, mathematics and more
- distributed system: multiple autonomous computers at different locations communicating via network
- distributed computing: use of a distributed system to solve a computational problem
  - must be possible to divide the computational problem into many small tasks
  - tasks must be able to run independently on separate computers
- volunteer computing: form of distributed computing in which individual computer owners donate unused computing power of their own computers to research projects

How does it work?

- requires a computer (Windows, Mac, or Linux) with an Internet connection available at least part of the time
- individual user downloads and installs middleware software program
- within the software, user chooses in which project(s) to participate
- software downloads work units from the project's server, runs them in the background while the computer is turned on, and uploads completed results to the project's server
- work units from (almost all) projects save intermediate results as they go along, so computer can be turned off and back on without losing everything

Berkeley Open Infrastructure for Network Computing (BOINC)

`http://boinc.berkeley.edu/`

- history: earliest volunteer computing projects wrote their own programs that combined the scientific computations and the communications management
  - Great Internet Mersenne Prime Search, 1996
  - SETI@home and Folding@home, 1999
- now: BOINC is middleware system that many projects use to drive their applications
- developed by an NSF-funded research project located at the UC Berkeley Space Sciences Laboratory
- open-source

- client software for Windows, Mac, Linux

World Community Grid

`worldcommunitygrid.org`

- an umbrella organization that hosts volunteer computing projects with clear potential benefit to humanity
- scientists propose project ideas to WCG
- WCG technicians develop the applications to run under BOINC
- WCG servers send out work units, receive results, and pass them on to scientists
- run by IBM
- WCG volunteers have contributed over 500,000 years of run time since 2004
- lots of information under "Help" on WCG web page

Current World Community Grid sciences

- OpenZika
- Help Stop TB
- Clean Energy Project Phase 2 (CEP2)
  - to find new materials for the next generation of solar cells and later, energy storage devices
  - to calculate the electronic properties of hundreds of thousands of organic materials and dermine which candidates are most promising for developing affordable solar energy technology
  - opt-in project because of demands on computer
    - 1 Gb of memory per running WU
    - large files to download and upload
    - infrequent checkpointing
    - lots of hard disk writing

- Fight Aids at Home
- Mapping Cancer Markers
- Outsmart Ebola Together

Controlling how WCG projects run on your PC

- under "Settings," "Device Manager" on WCG web page
- defaults work well for most people for unobtrusive use
- you probably will have to change your Windows settings for power management if you want BOINC WUs to keep running when computer is on and you're not using it
- separate opt-in for beta testing

Other BOINC projects in environmental science

- Climate Prediction
  `http://climateprediction.net/`
  - *very* long-running work units (100 to 500 hours on 2.5 GHz computer
- Quake Catcher Network – Seismic Monitoring
  `http://qcn.stanford.edu/sensor/`
  - requires either a laptop with an accelerometer or purchase of an external motion sensor
- Virtual Prairie
  `http://vcsc.cs.uh.edu/virtual-prairie/`

Finding other BOINC projects

- listing at Berkeley BOINC wiki
  `http://boinc.berkeley.edu/wiki/Project_list`
- some enable computing on graphics cards
- all projects that run under BOINC have similar user control features – just different look on website

The human side

- points
- teams
- badges
- forums

But does it actually accomplish anything?

- Clean Energy project (WCG)
  `http://www.nature.com/news/2011/110816/full/news.2011.481.html`
- Computing for Clean Water (WCG)
  `http://www.worldcommunitygrid.org/about_us/viewNewsArticle.do?articleId=161`
- Help Fight Childhood Cancer (WCG)
  `http://www.m.chiba-u.ac.jp/class/bioinfor/wcg/e/hfcc_e/news.html`
- Quake Catcher Network
  http://www.washingtonpost.com/national/health-science/earthquake-brought-seismometer-of-citizen-2011/08/25/gIQAgcqrnJ_story.html?hpid=z4

Validation: How do researchers know returned results are correct?

- redundant computing – 2 copies of each WU sent to different computers
- single validation (type 1) – after a computer has returned valid WUs under redundant computing, no validation except for occasional spot checks
- single validation (type 2)
  - for projects that do simulation, returned results from different computers for the same WU won't be identical
  - many copies (e.g. 19 for HPF2) of each WU sent out
  - results must agree to within pre-specified tolerance

How are work units for new projects tested?

- alpha testing by development team
- beta testing by users
  - users opt in
  - beta tester is expected to truly test and give feedback to project
    * suspend/resume
    * turn computer off and back on
    * check whether they are responding to BOINC control
    * test running concurrently with other software
    * check CPU temperature

What I recommend if you want to participate

- on Windows, download BOINC from World Community Grid website
- on Ubuntu Linux, install BOINC from the package repository
- use WCG default settings unless you want to get deeply involved in technical side
- check temperature (especially important on laptops)
  - install CoreTemp or TThrottle on Windows
  - install `lmsensors` on Linux
  - reduce % of processor usage or number of cores used if necessary
- participate in beta testing only if you have time to keep a close eye on running beta test units
- stick to well-established projects

- participate in GPU BOINC projects only if you have considerable technical expertise