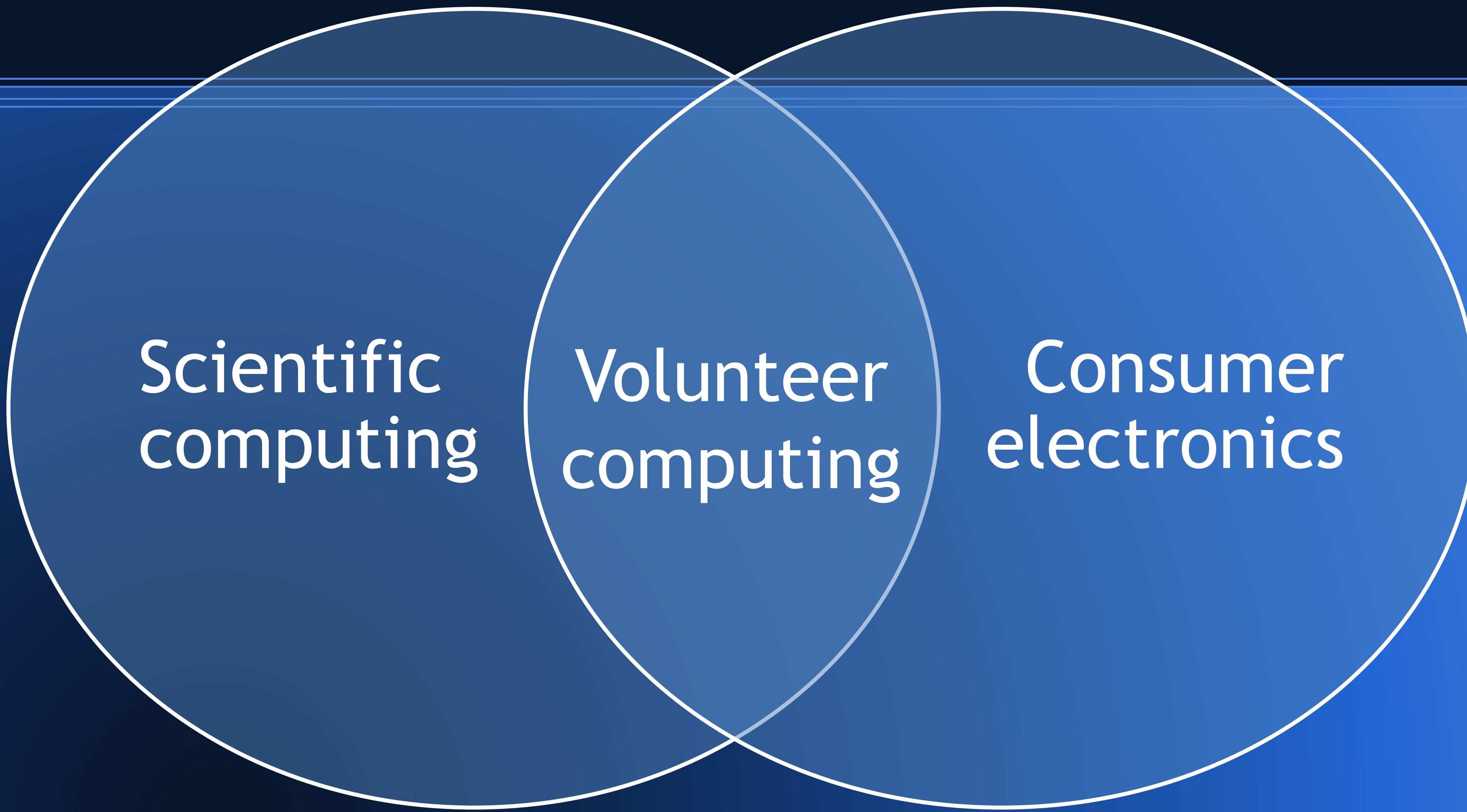
Volunteer Computing and BOINC

David P. Anderson

Space Sciences Lab
UC Berkeley

MAGIC
October 5, 2016

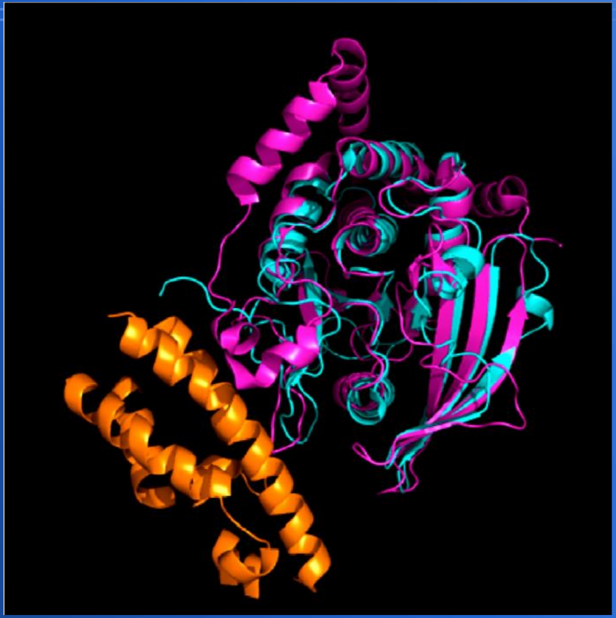

Scientific computing
Volunteer computing
Consumer electronics

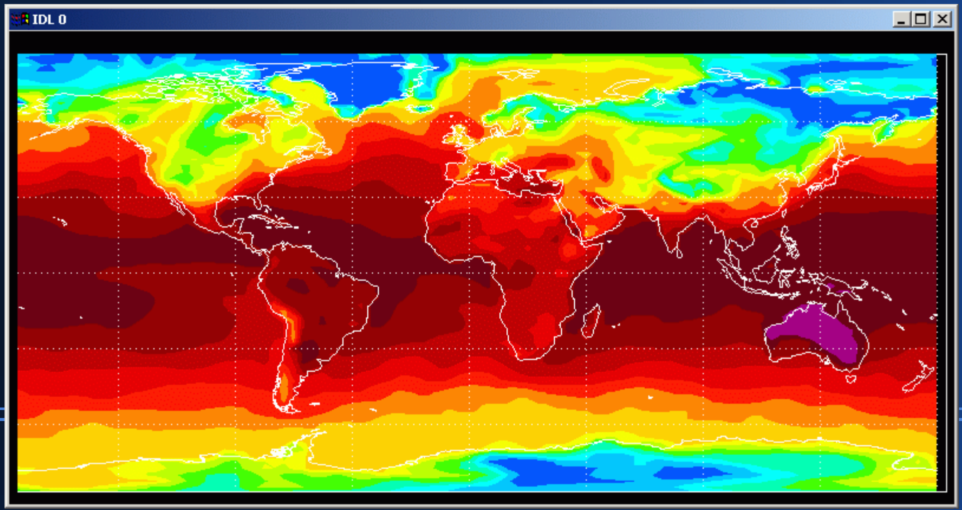
BOINC

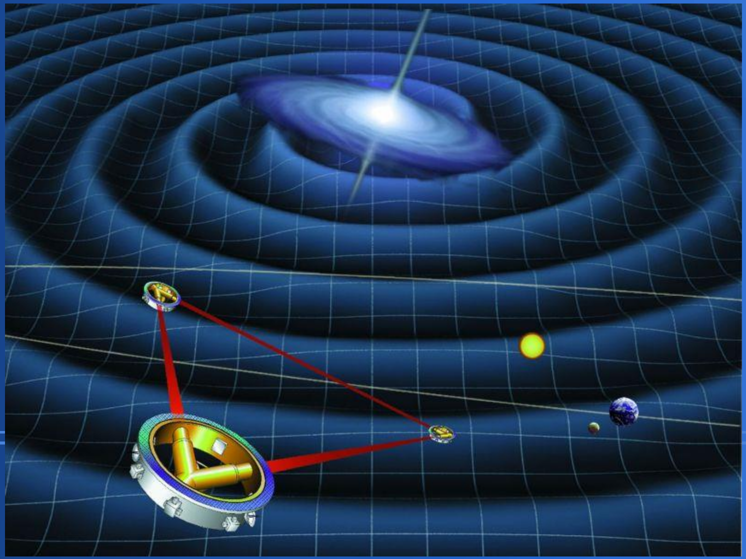
- Middleware for volunteer computing
  - Open-source, NSF-funded development
  - Community-maintained
- Server: used by scientists to make “projects”
- Client: runs on consumer devices
  - “attach” to projects
  - fetches/runs jobs in background

Example projects

- Climateprediction.net
- Rosetta@home
- Einstein@home
- IBM World Community Grid
- CERN







Current volunteered resources

- 500,000 active devices
  - BOINC + Folding@home
- 2.3M CPU cores, 290K GPUs
- 93 PetaFLOPS
- 85% Windows, 7% Mac, 7% Linux

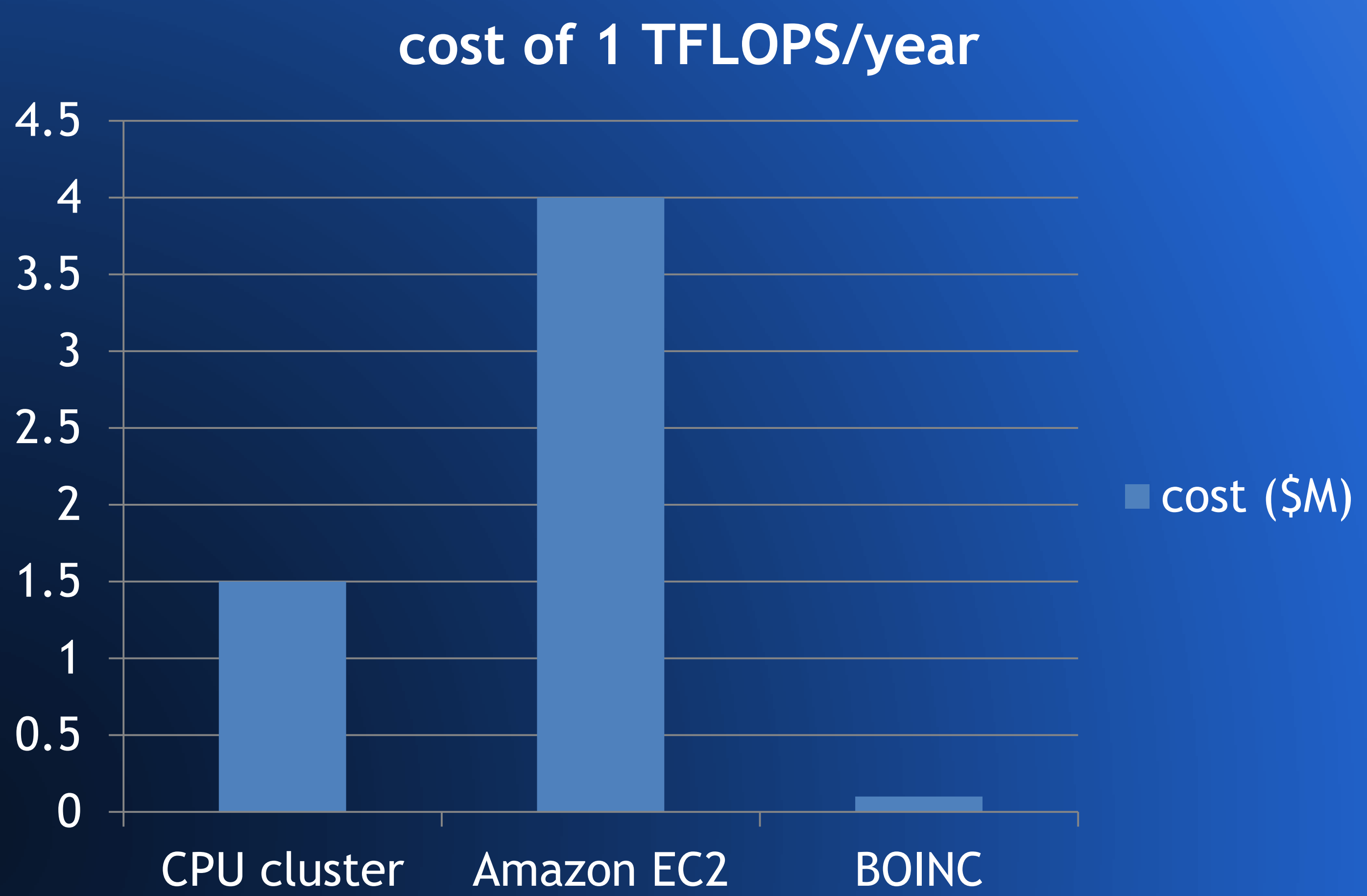
Performance potential

- 1 billion desktop/laptop PCs
  - CPUs: 10 ExaFLOPS
  - GPUs: 1,000 ExaFLOPS
- 5 billion smartphones
  - CPUs: 20 ExaFLOPS
  - GPUs: 1,500 ExaFLOPS

Realistic potential

- Study: 5-10% of people who learn about VC would participate
- Devices compute 60% of the time
- So if we can tell the world about VC, could get 100 ExaFLOPS

Cost


cost of 1 TFLOPS/year
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
CPU cluster
Amazon EC2
BOINC
cost ($M)

BOINC job model

- An app can have many versions
- Submit jobs to apps, not versions
- The BOINC scheduler decides what version(s) to use in response to a particular request

Per-platform apps

- Windows/Intel, 32 and 64 bit
- Mac OS X
- Linux/Intel
- Linux/ARM (works for Android too)

Other types of apps

- Multicore
- GPU apps
  - CUDA (Nvidia)
  - CAL (AMD)
  - OpenCL (Nvidia, AMD, Intel)

VM apps

- App is VM image + executable
- BOINC client interfaces via “Vbox Wrapper”
- Advantages:
  - No Win/Mac compilation
  - sandbox security
  - checkpoint/restart using Vbox “snapshots”
- Docker apps

Issues with VM apps

- Host must have VirtualBox installed
  - included with default BOINC install
- To run 64-bit VMs or Docker, host must have VM CPU features enabled
- Doesn’t work with GPUs (currently)
- Doesn’t work with ARM/Android

What workload can VC handle?

- Variable turnaround, so best for bags of tasks
  - but can handle DAGs/workflows too
- Moderate RAM, storage requirements
- Network: server capacity limits
- Data privacy

Areas where VC is useful

- Compute-intensive data analysis
  - particle colliders (LHC)
  - astrophysics
  - genomics
- Simulations of physical systems
  - nanosystems
  - drug discovery, protein folding
  - climate modeling
  - cosmology

BOINC system structure: the original model

- Dynamic “ecosystem” of projects
- Projects compete for computing power by publicizing their research
- Volunteers browse projects, support what they think is important
- Best science gets most computing power
- Public learns about science

Original model hasn’t worked

- Creating a project is too hard
- Creating a project is risky
- It’s hard to publicize VC: too many “brands”
- Volunteers are static

A new model

- Volunteers see “Science@home”, not separate projects
- Can express “science preferences”
- Science@home allocates computing power

New model, part 2

- Instead of single-scientist projects, “umbrella” projects that serve many scientists, and are operated by organizations
- Prototypes under development:
  - TACC
  - nanoHUB

Corporate partnerships

- Past/current:
  - IBM World Community Grid
  - Intel Progress Thru Processors
  - HTC Power to Give
  - Samsung Power Sleep
- In development:
  - Blizzard Entertainment (games)
  - EE (British cell phone provider)

Contact info

- BOINC
  - http://boinc.berkeley.edu
- Me
  - davea@ssl.berkeley.edu