

Volunteer Computing and Hubs

David P. Anderson

Space Sciences Lab
University of California, Berkeley

HUBbub
September 26, 2013



Computing resources

Consumer

**PCs
laptops
mobile devices
appliances**

volunteer computing

Organizational

desktop | data center

desktop
grid
computing

cluster,
cloud,
Grid,
super
computing

BOINC

How BOINC works

PCs
(Windows, Mac,
Linux, Android)



projects

nanoHUB

Einstein@home

WCG

attachment



It works best for...

- Large bags of tasks
- Throughput-oriented
- Moderate RAM, disk usage
- Moderate data/compute ratio
 - < 1 GB per hour

Volunteer computing status

- Volunteers: 400K people, 600K computers, 2M cores
- > 50 science projects, e.g.
 - Climateprediction.net (Oxford)
 - Einstein@home (Max Planck Inst.)
 - IBM World Community Grid
 - “umbrella” project; MGI is one
- > 15 PetaFLOPS, many million jobs/day
 - Potential: many ExaFLOPS

Adding BOINC to Hubs

Goal: increase the computing throughput available to Hub users

- A Hub has an associated BOINC project
- Large batches of jobs submitted to the Hub are run using the BOINC project
- Hub users attach their PCs to the project
- Project can be accessible to the public

BOINC Web RPC Interfaces

- Remote job submission
 - get batch completion estimate
 - submit/query/abort batches
- File management
 - content-based: only 1 copy of a given file
 - reference-counting, garbage collection

Identity and access control

- Each Hub user has a corresponding BOINC account
- BOINC enforces access control for job submission

User quotas

- Each user has a “quota” determining the fraction of computing they get
- May be set in proportion to
 - user’s computing contribution
 - other Hub contributions
- BOINC’s job scheduling policy
 - do small batches before large ones
 - don’t starve large batches
 - enforce quotas over the long term

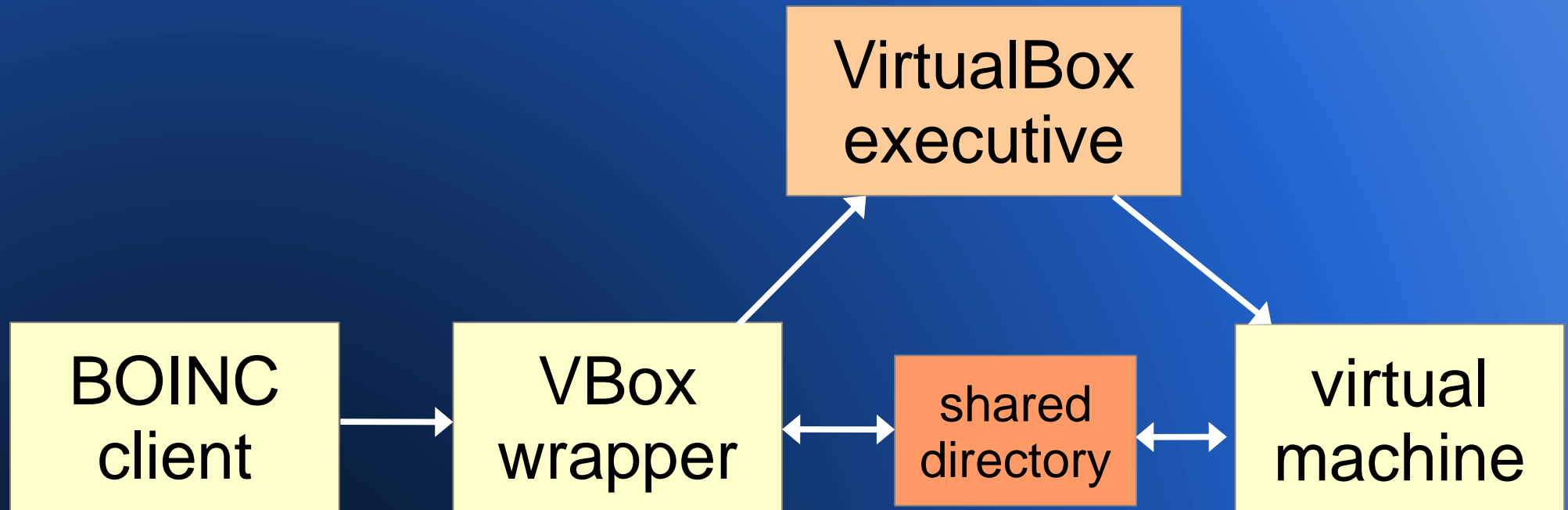
Using VM technology

- Heterogeneity of volunteer PCs
- Difficult to port applications to all platforms

BOINC VM-based apps

- Developers work in environment of choice
- App consists of
 - compressed minimal VM image (same across multiple apps)
 - executable
 - BOINC “Vbox wrapper”
- Volunteers need to install VirtualBox

How it works



VM bonuses

- Vbox wrapper creates a “snapshot” every few minutes
 - free checkpointing
- A VM is a strong security sandbox
 - can potentially deploy untrusted apps

Implementation

- Proof of concept in nanoHUB
 - VM with Tcl, Rappture library
 - Job submission via Web or workspace “submit”
 - Behind firewall for now
- To do
 - Add “batch” notion to HUBzero
 - Use BOINC’s job-submission interfaces
 - Automate identity management

Branding and marketing

- How to market to volunteers?
- Brand: nanoHUB@home or HUB@home?
 - Sometimes fewer brands are better



Account manager

Science@home

BOINC projects

nanoHUB@home

Conclusion

- HUBzero and BOINC are complementary
 - more computing power to scientists
 - more science options for the public
- Contact
 - <http://boinc.berkeley.edu>
 - davea@ssl.berkeley.edu
- BOINC Workshop
 - Grenoble, Sept 25-27 2013