## **Open Source: BOINC**

Laurence Field



# **Volunteer Computing**

- Computer owners donate computing capacity
  - To a cause or project
- Spare cycles on computers
  - Away from office/desk, unused cores
- Other opportunistic resources
  - Idle machines in data centers
  - IoT?
  - Opportunistic energy?
  - Cars e.g. V2G?







## How It Works

- Download and run the client
- Select a project
- Enter an email address and password
- Run the application and earn credit













#### climate prediction.net

the world's largest climate modelling experiment for the 21st century







Microbiome Immunity Project OpenZika Help Stop TB FightAIDS@Home - Phase 2 Mapping Cancer Markers









## Motivation

- Free\* resources
  - 100K hosts achievable for large projects
  - Actual job slot count (number of cores) is higher
- Community engagement
  - Outreach channel
    - Explaining the purpose and value of the science
  - Participation
    - Offering people a chance to contribute
    - Engagement forms a strong bond
- Community support

\*There are costs associated with their use



# Challenges

- The cost of using the free resources
  - Initial integration requires investment
- Operations and Maintenance
  - Public facing support on all levels
    - Lowered by community supports
- Attracting and retention of volunteers
  - Advertisement and engagement
  - Communications cost for capacity building
- Low level of trust
  - Anyone can register as a volunteer
  - Not the same as internally owned systems
- Running scientific software on Windows
  - ~80% of the resources



CERN Accelerating science	Sign in	Directory
Volunteer computing		
@home for the LHC Search		
HOME ABOUT - PROJECTS - JOIN US! HELP & FAQ CONTACT		



#### 



http://cern.ch/lhcathome

## **Computing Workflow**

Processed Data (Active tapes)





# Applications





# Our Usage of BOINC

- One project with multiple applications
  - Reduce operational costs
    - Single forum
    - One service
  - Simplify for the volunteers
    - One project
    - Single user name and password
- Classic and virtualized applications
  - Sixtrack
  - Test4thoery, ATLAS, CMS, LHCb and ALICE
    - Because HEP software only runs on Linux



## **BOINC** with Virtualization





## CernVM and CVMFS



Small image size but need to "bake" the images to reduce unnecessary downloads



### CDN: Cloudflare/OpenHTC.io





### CMS@home Job Speedup

CloudFare





https://indico.cern.ch/event/587955/contributions/2936824/

# VM Capacity Gap



July 2018



# Native App

- Remove the VM layer for Linux
  - Requires CVMFS and Singularity
- Persistent CVMFS Cache
  - Jobs directly from the BOINC sever
    - No additional infrastructure required
- Run native on internal resources
- Windows and Mac
  - Run in a Linux VM
    - Provide one pre-configured
  - No local resource management



## Platform

- Berkeley Open Infrastructure for Network Computing
  - Started in 2002
    - By a team based at the Space Sciences Laboratory
    - University of California, Berkeley
    - Led by David Anderson
  - Now an open source project
- Provides the toolkit for volunteer computing
  - Client (Mac, Windows, Linux, Android)
  - Graphical User Interface
  - Application runtime system
  - Server software
  - Project Web site





# **BOINC Community**

- Transforming into an open community
  - https://github.com/BOINC/boinc-policy
- Meetings
  - Quarterly projects calls
  - Biweekly dev calls
- GitHub
  - Focal point
  - Facilitating communications
  - Merge Requests from more individuals
    - Actively included
- New server release
  - boinc-docker-server



## Open-Source ≠ Open Source

- Open-Source is more than provision
  - It's about empowering
- BOINC Transistion
  - Master  $\rightarrow$  Branch $\rightarrow$  Release
  - Branch  $\rightarrow$  Master  $\rightarrow$  Release
  - Managed Team  $\rightarrow$  Collaborating Community
- Community forced change
  - Desired open-source best-practice approach
    - Community defined policies
- More information
  - https://opensource.guide
  - https://www.youtube.com/watch?v=uzxclLudFWM



## The Merge Request

- Every change can be peer-reviewed
  - Branch the code from Master
  - Make an atomic change
  - Create a Merge Request
  - Request can be discussed
  - Request accepted/rejected
- Communication
  - Suggestion contribution clearly shown
    - Can add explanations
  - Discussion linked to suggested change
    - Code can be annotated
    - Block changes and make requests
    - Archived for future reference



## People Before Code

- Build a community first
  - They will build the code
- Code of Conduct
  - Be respectful
- Create a welcoming environment
  - Clear policies and adherence
    - Documentation!
- Welcome contributions
  - Provide prompt feedback on merge requests
- Quickly identify contentious issues
  - Process for achieving consensus
    - Or have a zero-consensus process



## **Open-Source As A Model**

- Global citizens collaborating to achieve a common goal
  - Demonstrates what is possible
  - Not just volunteers but people from organizations
- How can this be applied to other areas?
  - People
  - Commodity
  - Process
  - Tooling



# Summary

- Volunteer Computing can and is providing
  - Significant additional computing resources
    - Potentially O(100K) machines
- Virtualization, Containers and CVMFS enable HEP applications
  - VM capacity gap
- Using CernVM
  - Baked VM to reduce downloading each time the VM is restarted
    - Use openhtc.io for better CVMFS performance
- BOINC is a healthy Open-Source community
  - The Merge Request
  - People Before Code
- Come and join the fun!
  - http://lhcathome.web.cern.ch/join-us

