SDG Volunteer Computing

("LHC@home enhancement")

openlab summer project 2018

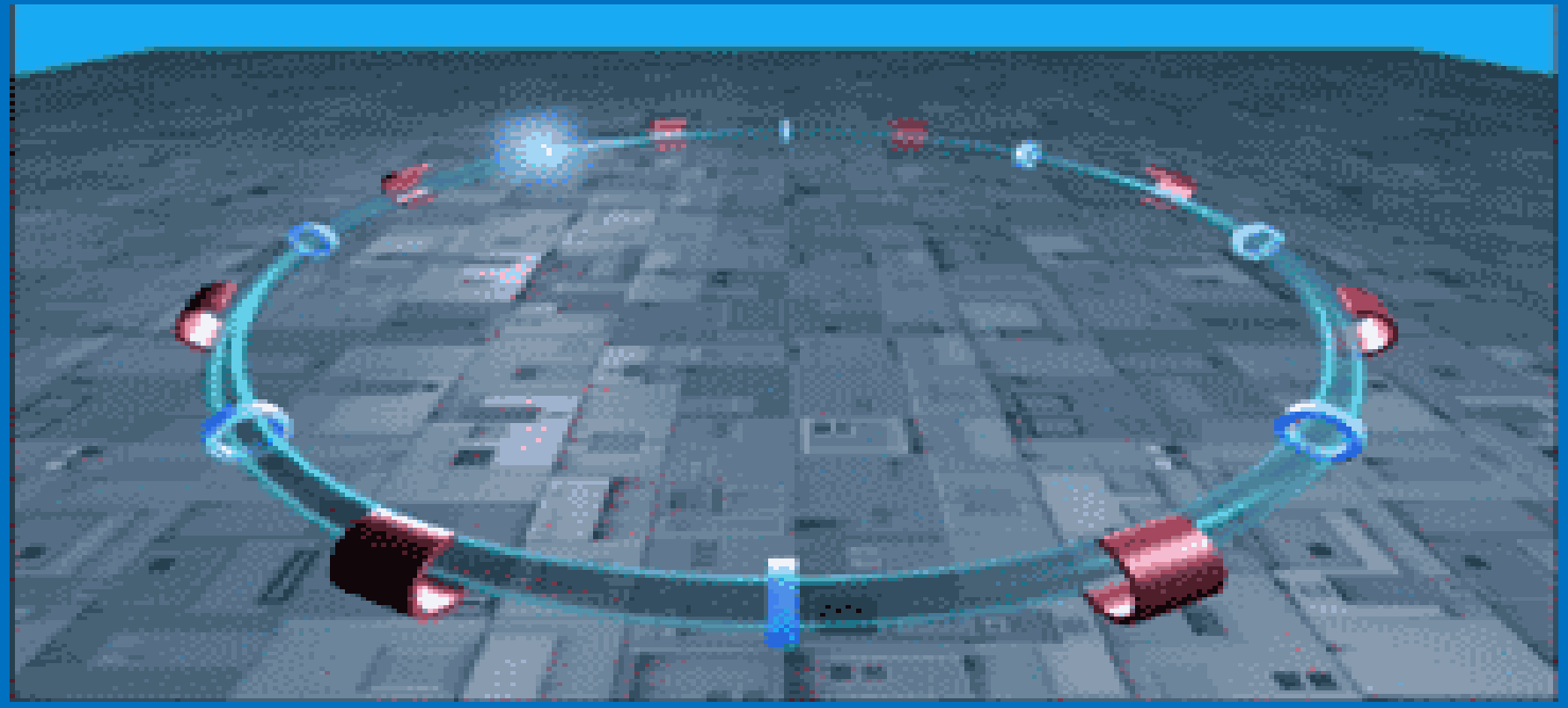
Laurence Field & Ben Segal
CERN

Volunteer Computing

- A type of distributed computing using public volunteers
- SETI@home (1999), Folding@home (2000)
  - LHC@home launched in 2004
- Computer owners donate computing capacity
  - To a cause or project
- Not necessarily only spare cycles on Desktop PC's
    - Idle machines in data centers
    - Home clusters
    - Tablets or phones …

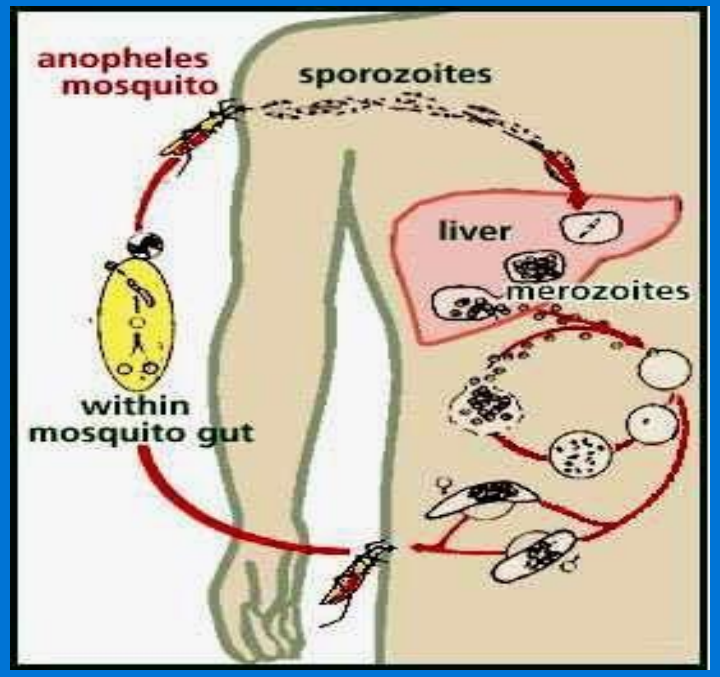
LHC@home (2004)

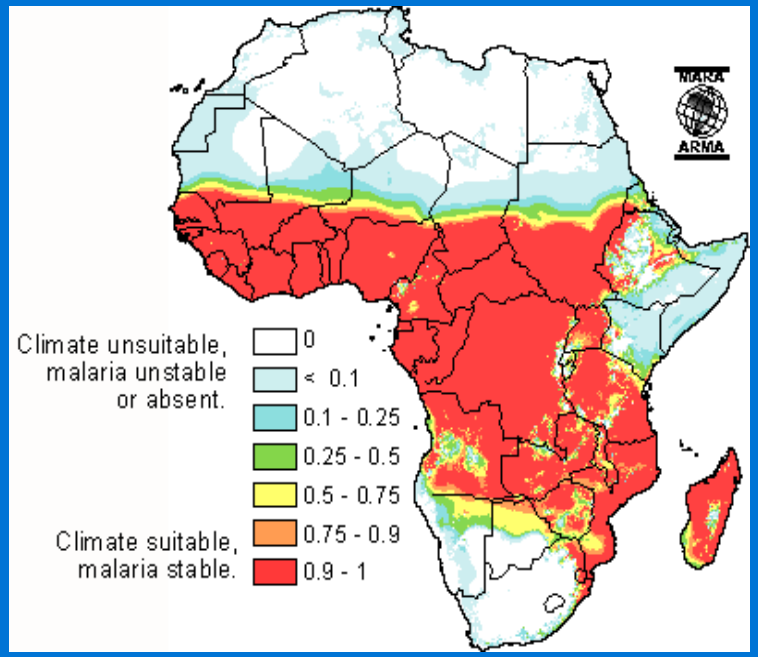
Accelerator design by beam simulation

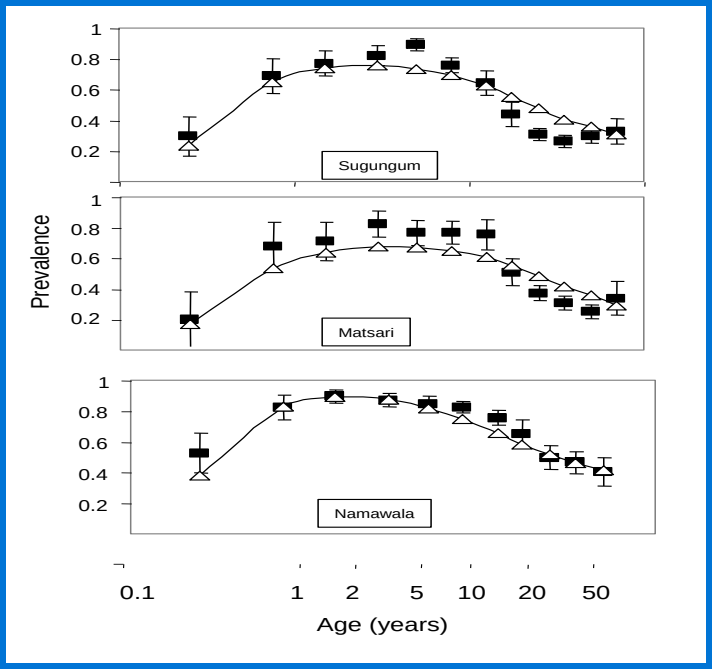


Africa@home (2006)

Modelling the epidemiology of malaria in Africa with Swiss Tropical Institute

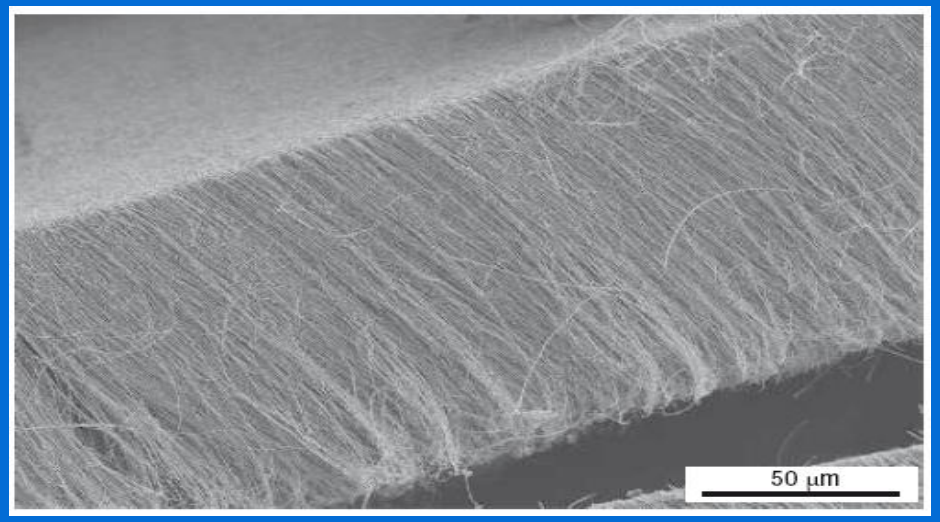


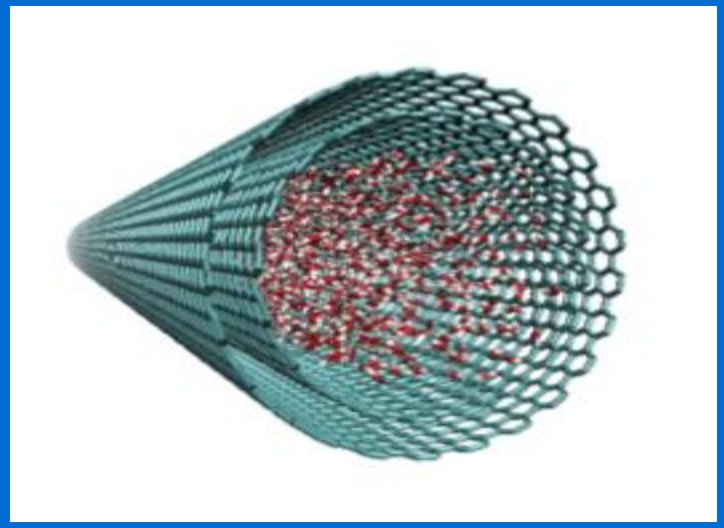




CAS@home (2009)

Simulation of nanotech water filters



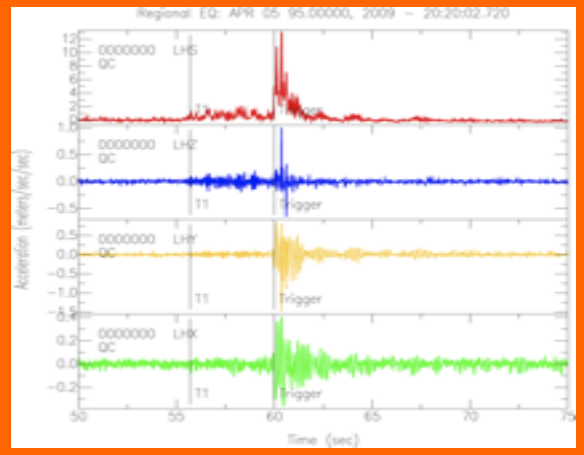


Project with Tsinghua Univ, IBM World Community Grid

1) **Simulating enhanced water flow through nanotubes**
2) **Accuracy at low v needs large samples (~10^5 CPU-years)**
3) **IBM WCG projects preloaded on Sony Vaios in USA**

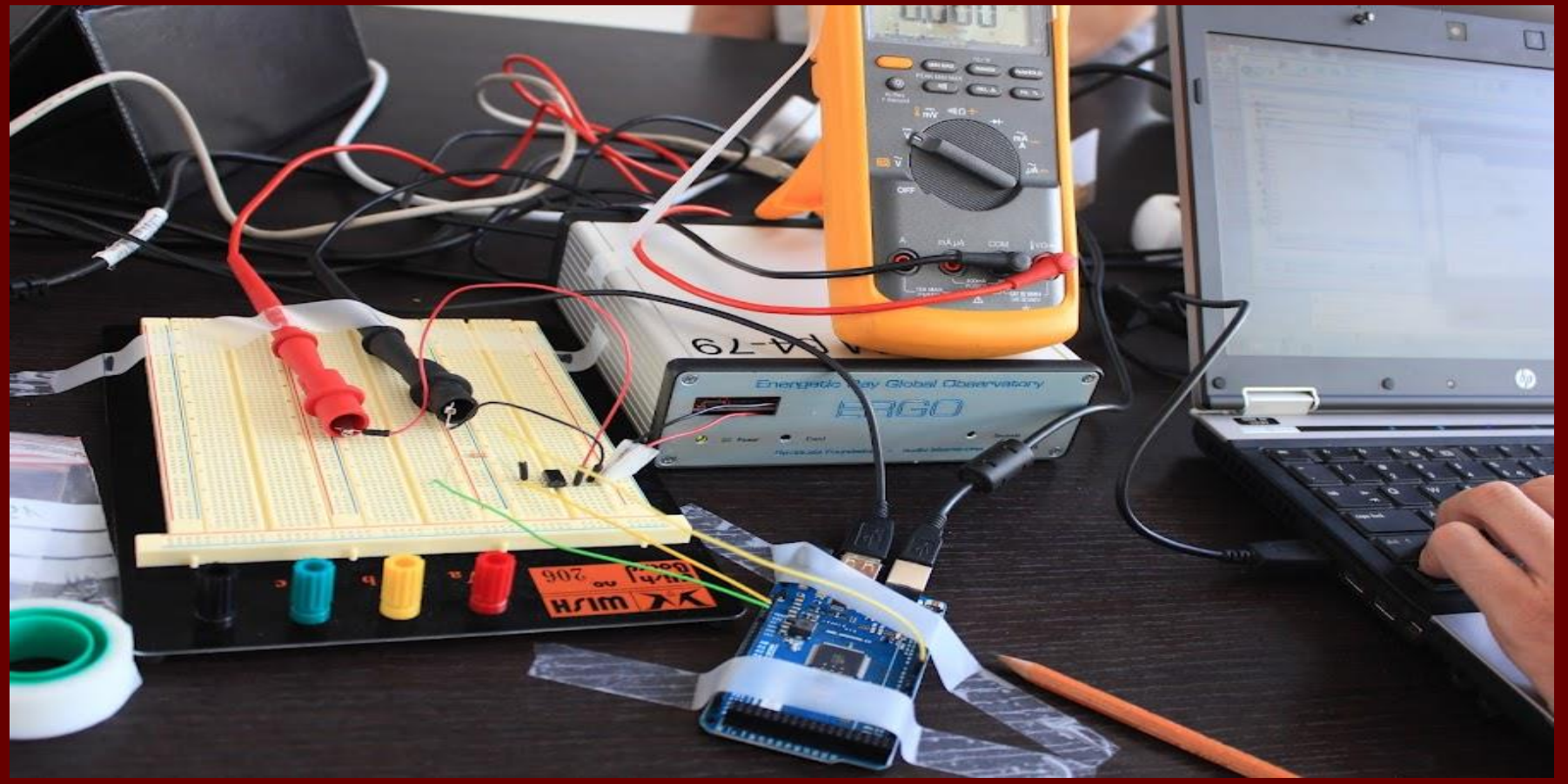
Asia@home (2010)

volunteer seismic detection and science



Citizen Cyberlab

Learning and creativity in citizen cyberscience (EU FP7)

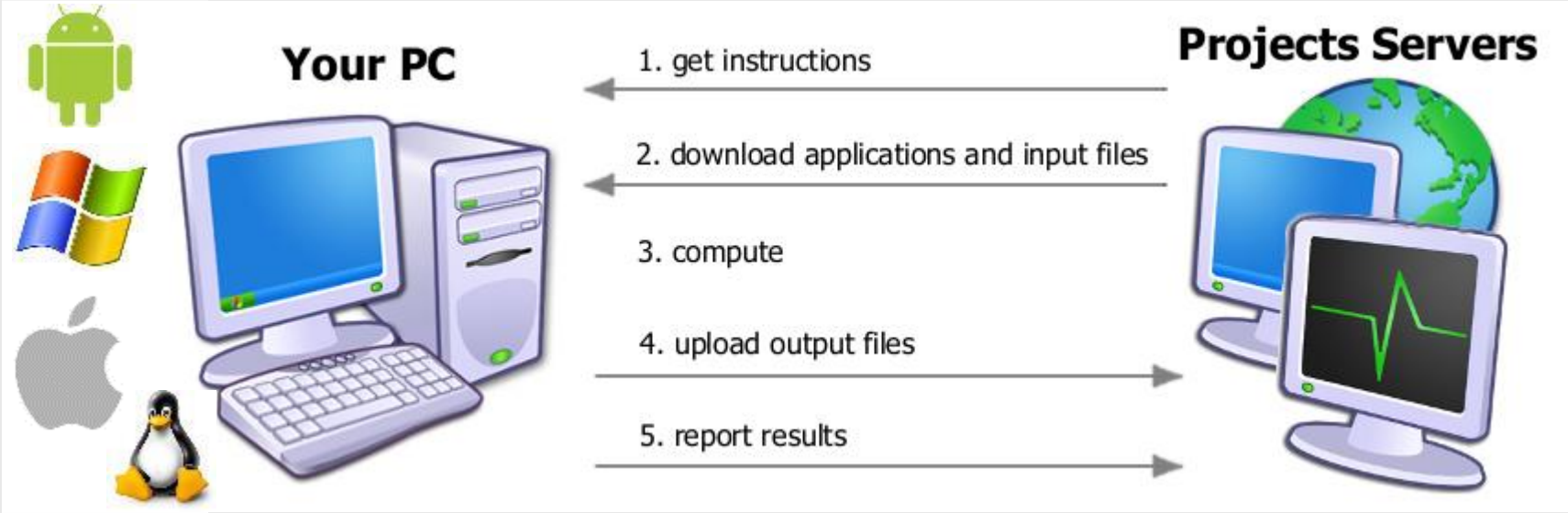


BOINC

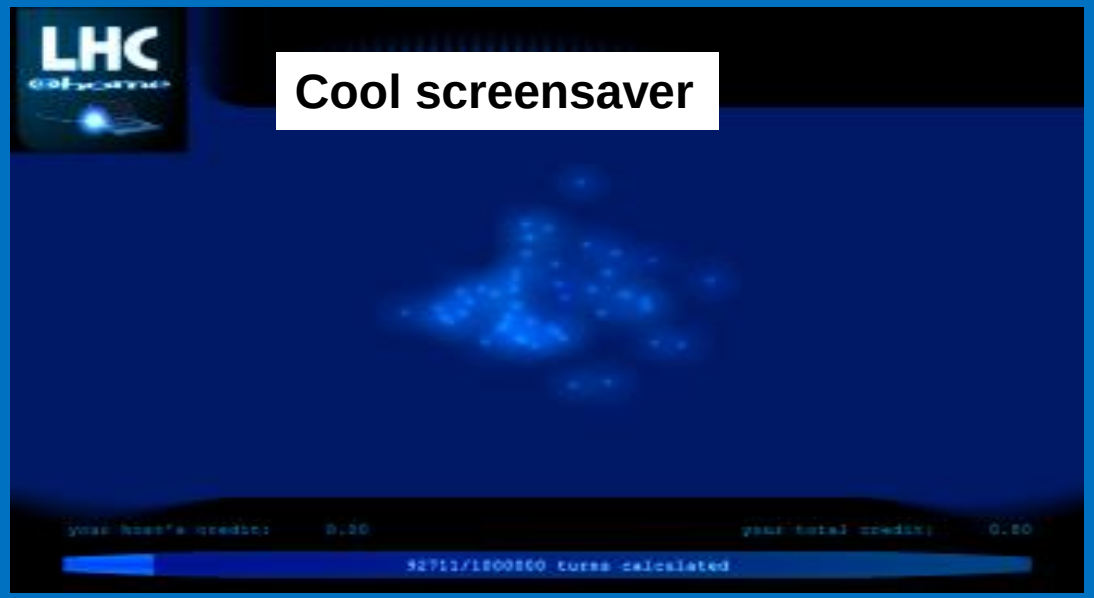
- Berkeley Open Infrastructure for Network Computing
  - Started in 2002 by SETI@home team
  - Funded by the National Science Foundation (NSF)
  - Developed by a team based at the Space Sciences Laboratory
    - University of California, Berkeley
    - Led by David Anderson
- Provides open middleware for volunteer computing
  - Client (Mac, Windows, Linux, Android) with CLI
  - GUI
  - Application runtime system
  - Server software
  - Project Web site

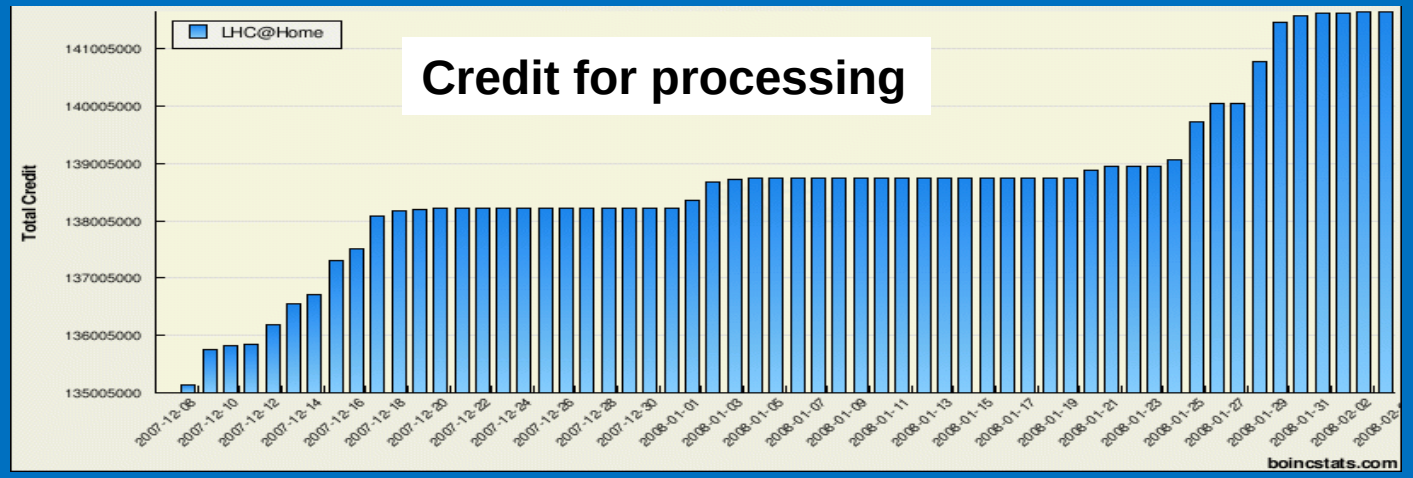
Volunteer Perspective

- Download and run the BOINC client
- Choose a project
- Enter an email address and password
  - Or silent connection with a key
- Run the application and earn credit



Why do volunteers participate?





Motivation for a BOINC project

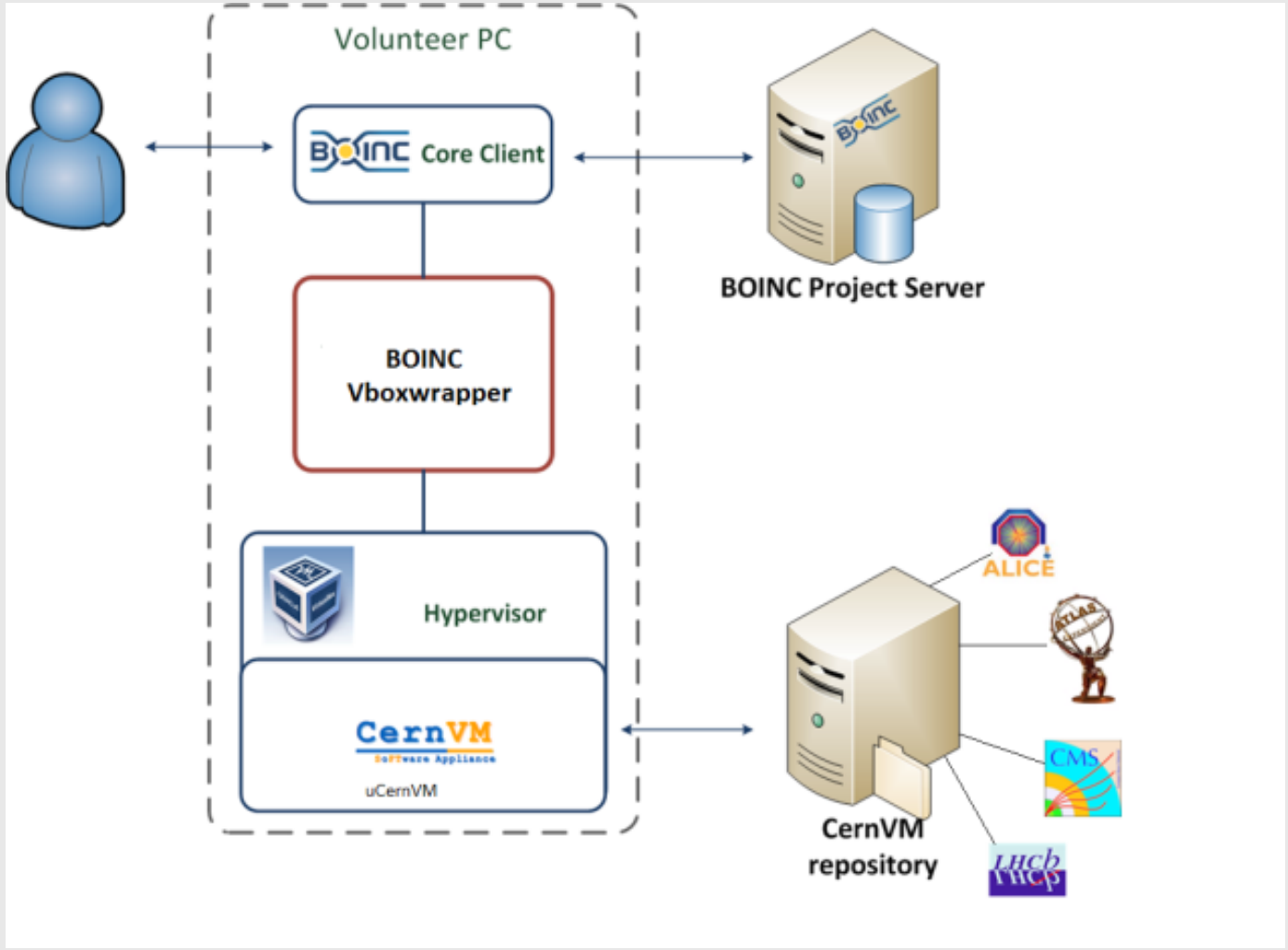
- Free* resources
  - 100K hosts achievable for large projects
    - Actual job slot count (number of cores) maybe even higher
- Community engagement
  - Outreach channel
    - Explaining the purpose and value of the science
  - Participation
    - Offering people a chance to contribute
      - Engagement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

* ***(But there are some costs required to use them!)***

Challenges to solve

- The cost of using the “free” resources
  - Initial software integration requires investment
- Operations and Maintenance
  - Public facing support on all levels
    (but lowered by support from the community itself!)
- Attracting and retaining volunteers
  - Advertisement and engagement
  - Communications cost for capacity building
- Low Level of Assurance
  - Anyone can register as a volunteer
    (not the same level of trust as with the Grid)
- Running HEP software on Windows
  - Windows systems are till 85% of the resources !!
  - CERN solved this problem by using VIRTUALISATION

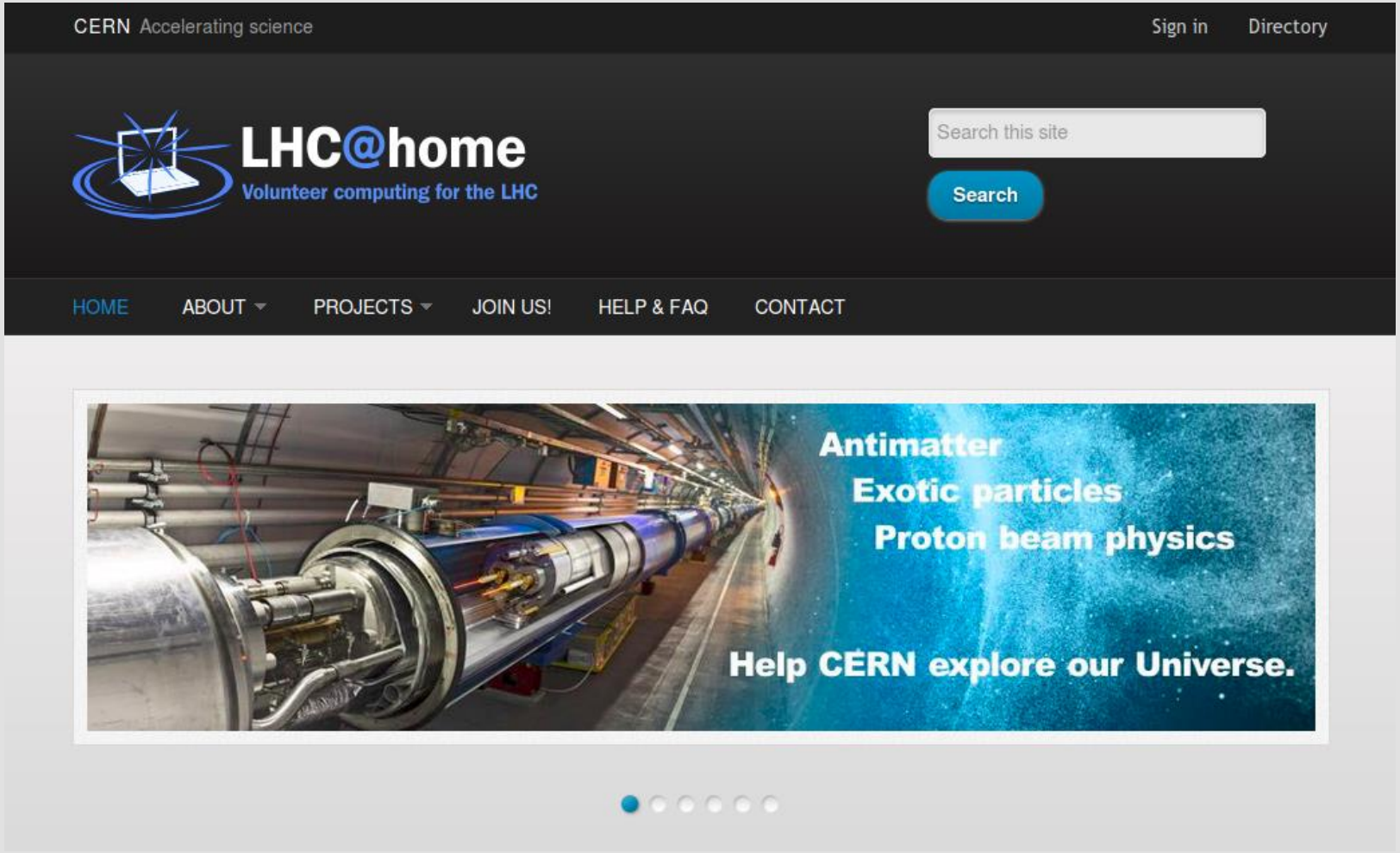
BOINC with Virtualization



Our Usage of BOINC at CERN

- A single project “LHC@home” with multiple applications
  - Reduced operational costs
    - Single forum
    - One service
  - Simplified for the volunteers
    - One project to attach to instead of several
    - Single user name and password
- Both classic and virtualised applications run together
  - Sixtrack (“classic”)
  - Test4theory, ATLAS, CMS, LHCb and ALICE
    - **(“virtualised”)** - because HEP software only runs on Linux

http://cern.ch/lhcathome



Current LHC@home Applications

	**ATLAS@home** "Known" physics and "new" phenomena - want to create alternative models of the universe?
	**Beauty** "b" is for "Beauty" - the gorgeous little particle in antimatter physics.
	**CMS@Home** CMS is on the lookout for completely new, unpredicted phenomena.
	**SixTrack** Help CERN accelerator engineers to run intensive simulations to check the stability of the twin proton beams circulating in the LHC machine.
	**Test4Theory** Simulate high-energy particle collisions and help tune the theory to the experimental results.

2018 Summer Student project

- The summer student project:

➔Create new applications for SDG researchers:

- Working with University of Geneva teams
- Demonstrate R language and Machine Learning capabilities
- Prototype the applications first in a private cloud cluster
- Port prototypes to BOINC / LHC@home for Volunteer Computing

2018 Summer Student project

- The summer student project:
  - **First time Google's system "Tensor Flow" used in BOINC**
    - Build, train and exploit Neural Networks for SDG applications
    - Allows the volunteers' GPU's to be used as well as CPU's

LHC@home
Volunteer computing for the LHC

Summary

- Volunteer Computing can and is providing:
  - Significant additional computing resources
  - Potentially O(100K) machines
- Virtualisation enables HEP applications
  - To run on multiple platforms: Windows, Mac and Linux
    - Can therefore reach more volunteers
- LHC@home is a common platform
  - Supporting multiple applications, now including SDG applications
- Come and join the fun!
  - http://lhcathome.web.cern.ch/join-us