LHC@home and CernVM – a new approach to porting large-scale applications to BOINC

Daniel Lombraña González, Ben Segal
and Artem Harutyunyan

Citizen Cyberscience Centre
CERN

March 20, 2011

Outline

The beginning

The challenge

Why don't you run real LHC physics on BOINC?

The challenge

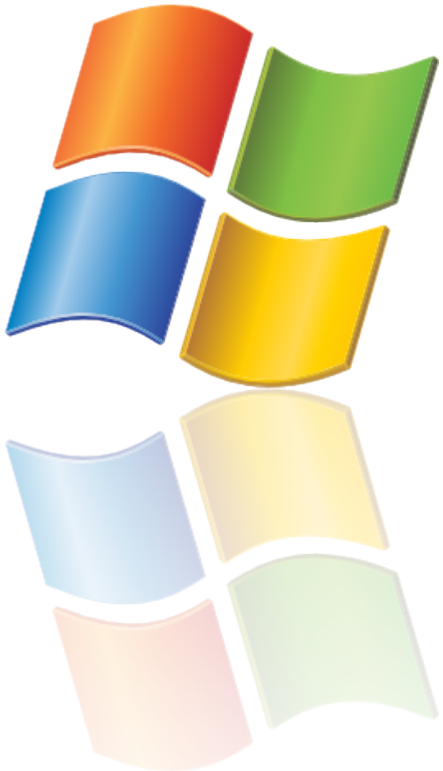
- Allow “any” PC to run a full LHC physics application.
- Make those commodity computers look like a “standard” CERN Data Center.

We did it!

Outline

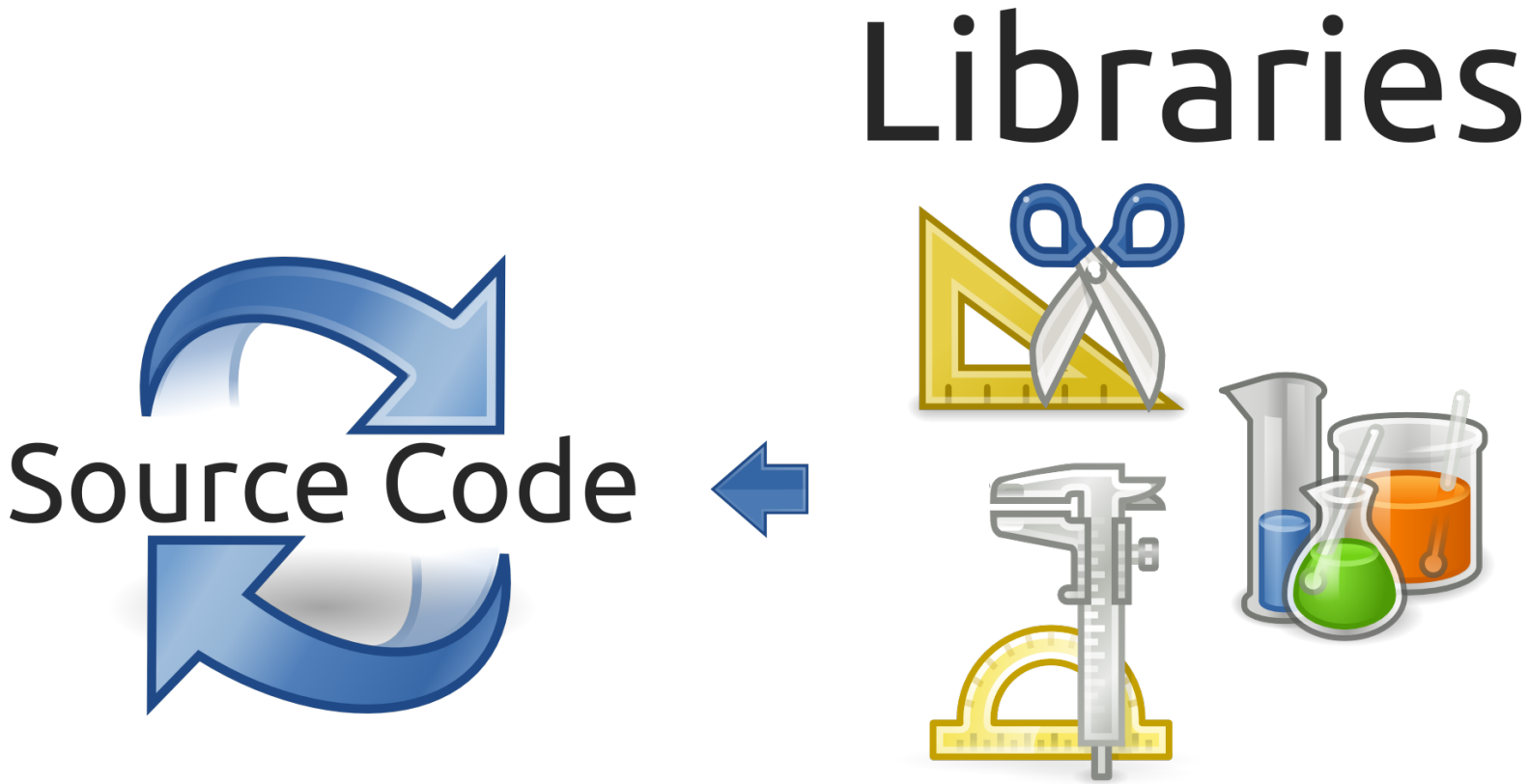
Challenge 1: Porting source code





Challenge 1: Porting source code

Challenge 1: Porting source code



Challenge 2: Job management systems

- Jobs must be fed into BOINC PCs, but CERN physics experiments have their own and don't want to use the BOINC distributions system.
- Volunteer computing resources are not "managed" and thus cannot be "trusted".

Outline

CernVM + Co-Pilot

- Using Virtualization and CernVM it is possible to solve the challenges (1) for application deployment on volunteers' machines.
- With CernVM's Co-Pilot it is possible to connect existing Grid infrastructures of LHC experiments with the BOINC volunteer resources, solving challenges (2).

CernVM
SoFTware Appliance

CernVM solution

CernVM is

a baseline Virtual Software Appliance for the participants of CERN LHC experiments.

CernVM solution

CernVM is

a baseline Virtual Software Appliance for the participants of CERN LHC experiments.

The goal is

to remove a need for the installation of the experiment software and to minimize the number of platforms.

CernVM Co-Pilot solution

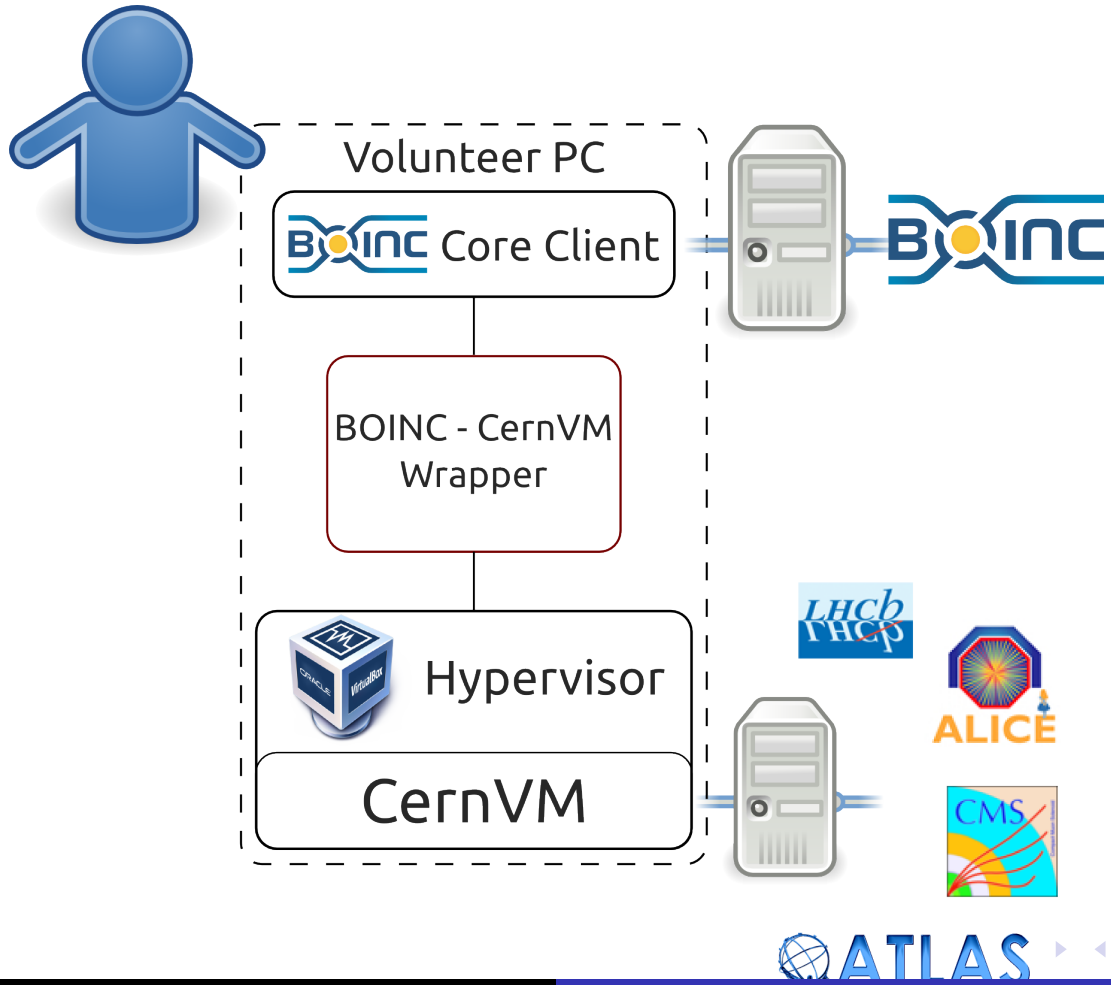
- A framework for the delivery and execution of the workload on the remote virtual machines.
- Consists of components developed to ease the integration of cloud resources into existing Grid infrastructures.
- Components communicate using Jabber/XMPP instant messaging protocol.

Outline

CernVM

- Using virtualization and CernVM solve the problem of porting the source code and deploying over PCs.
- CernVM Co-Pilot can connect existing Grid infrastructure of LHC experiments with BOINC resources.

BOINC + CernVM infrastructure

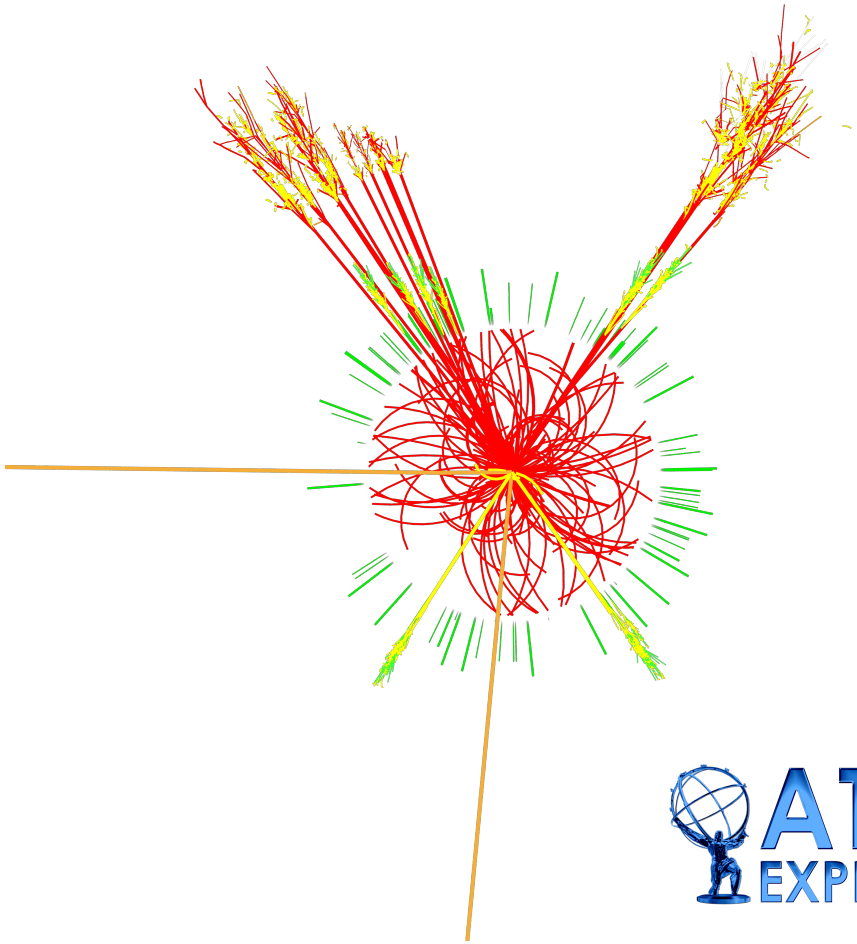


Co-Pilot Adapters

- Each LHC experiment that wants to connect to a set of CernVM machines via Co-Pilot, needs an adapter.
- Each adapter fetches jobs from the experiment's preferred job scheduler, submits them to the CernVM machines, and returns the results to the scheduler.
- Co-Pilot has a built-in security barrier, preventing untrusted CernVM machines to access Grid resources.

Outline

Monte Carlo events generation



Alpha testing

Active forums

- More than 500 posts in the forums so far.
- Very active volunteers, reporting all the problems they find.

High schools

- Three young students were invited to test the project.
- One of them became very helpful in the first stages of the testing phase.
- The obtained feedback allowed us to improve the project, knowing that young students could collaborate with complex projects.

Numbers

- Since the beginning of the project 25000 jobs have been successfully executed.
- 80% of the jobs were set to generate 300000 events (the rest were smaller ones that generated 50000 events).
- Totaling 6250 millions of events simulated since December 2010.
- There are on average 20 on-line users at any given point in time (peaks up to 30 users).
- We are running 1200-1300 jobs daily.

Outline

Conclusions

- A custom virtualized execution environment has been added to BOINC.
- Very complex and unmodified applications can be run within the system.
- Adapting Grid services to Co-Pilot is now understood.

Conclusions

We have built a Volunteer Cloud for BOINC!

Questions

teleyinex@gmail.com