CernVM: ~~15~~ 17 Years After

Predrag Buncic, Gerardo Ganis
(CERN/EP-SFT)

Introduction

- This is actually 17th (!) anniversary of CernVM project
  - It officially started in 2007 when funding for a software R&D project in the Physics department was secured (prototyping started one year earlier)

- What follows is a short history of the project

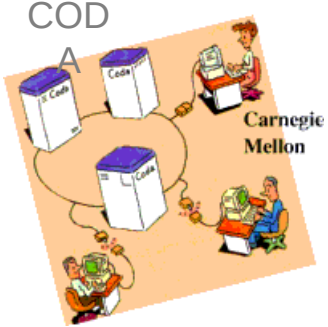
The Problem (2008)

- LHC was still under construction
  - Amazon WS was just starting, about two years old
  - Grid developments in a full swing, everything is based on traditional batch
- Software @ LHC
  - Millions of lines of code that frequently change
  - Different packaging and software distribution models
  - Complicated software installation/update/configuration procedure
  - Long and slow validation and certification process
  - Very difficult to roll out major OS upgrade (SLC4 -> SLC5)
- Novel approaches to virtualization might have provided a solution
  - Build minimal OS & bundle OS and applications into Virtual Machine image
- Not really
  - Updating, building and distributing large (>10GB) images was equally complicated problem
- Still, an opportunity to do some innovative R&D and look for practical and pragmatic solutions

Distributed File Systems

CernVM Software Appliance

CODA

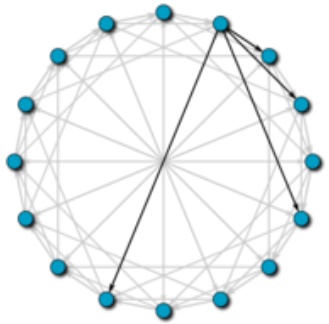


http://www.coda.cs.cmu.edu/

Chirp/Parrot/Grow-FS

https://ccl.cse.nd.edu/

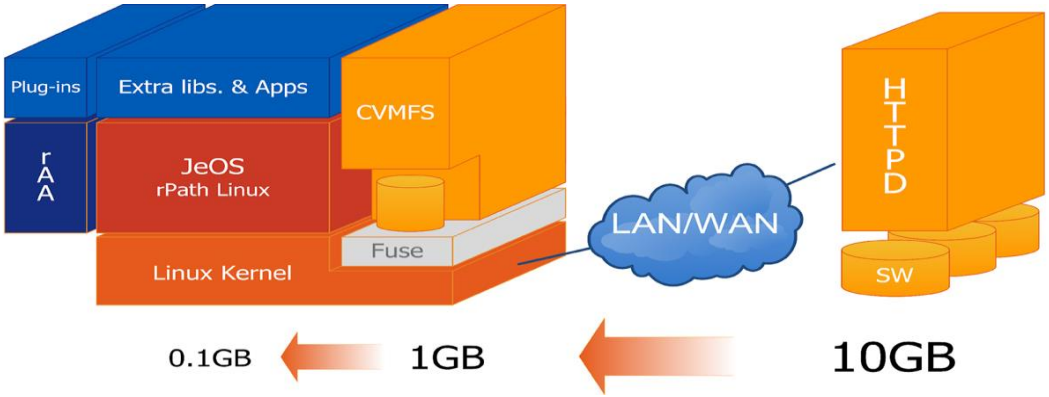
https://www.openafs.org/



IgorFS

http://doi.ieeecomputersociety.org/10.1109/P2P.2008.19

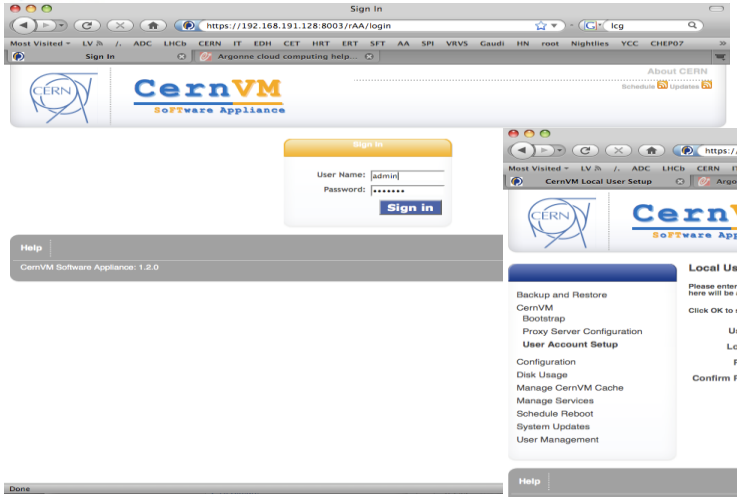
CernVM R&D



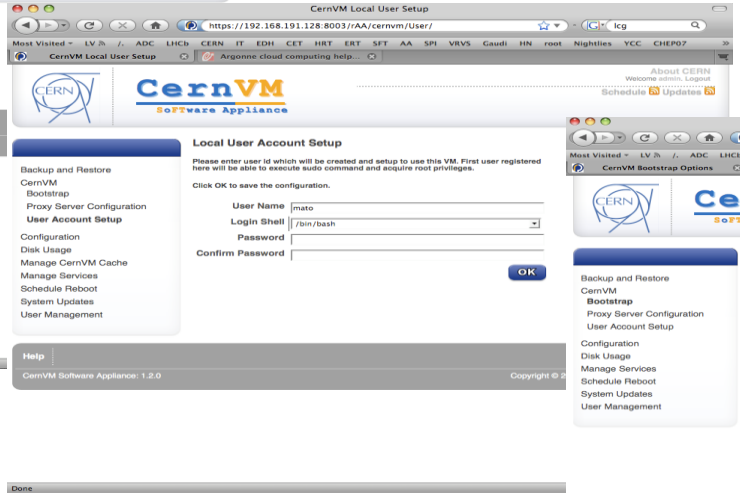
- Prototype CernVM File System (CernVM-FS) was derived from Parrot (http://www.cctools.org) and its GROW-FS code base and adapted to run as a FUSE kernel module adding extra features like:
  - possibility to use multiple file catalogs on the server side
  - transparent file compression under given size threshold
  - dynamical expansion of environment variables embedded in symbolic links
- Using rPath tools, packaged with minimal OS into CernVM appliance

Leandro Franco

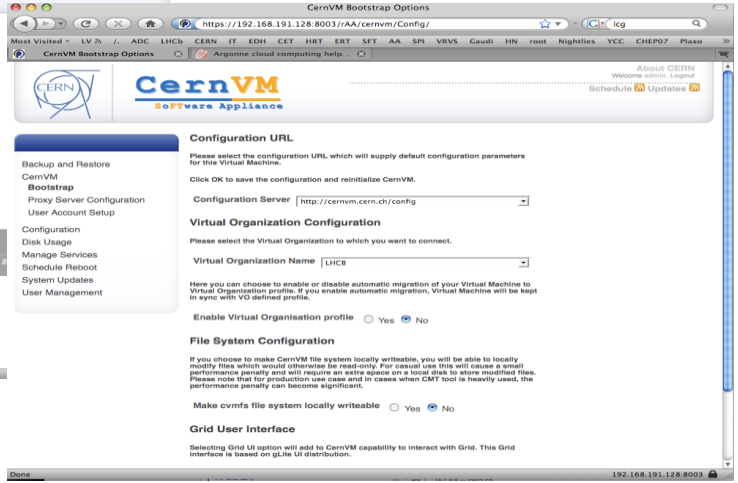
As easy as 1,2,3



1. Login to Web interface



2. Create user account



3. Select experiment, appliance flavor and preferences

CernVM Team (2008)

Carlos Aguado

Service Infrastructure Development and Support

System Development Engineer at Amazon Web Services

Artem Harutyunyan

Co-Pilot Architect and Developer

Co-Founder and CTO at bardeen.ai

Jakob Blomer

CernVM-FS developer

Now CERN staff in EP-SFT

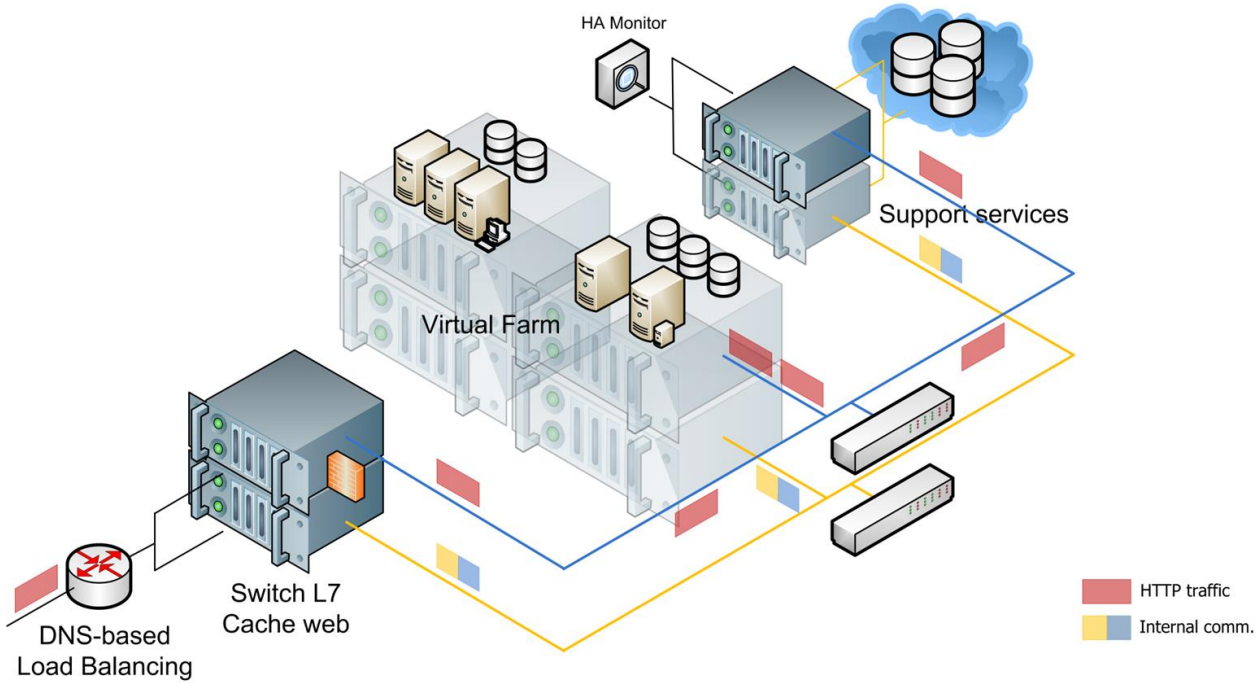
Ioannis Charalampidis

CernVM Online, micro CernVM, Test4Theory..

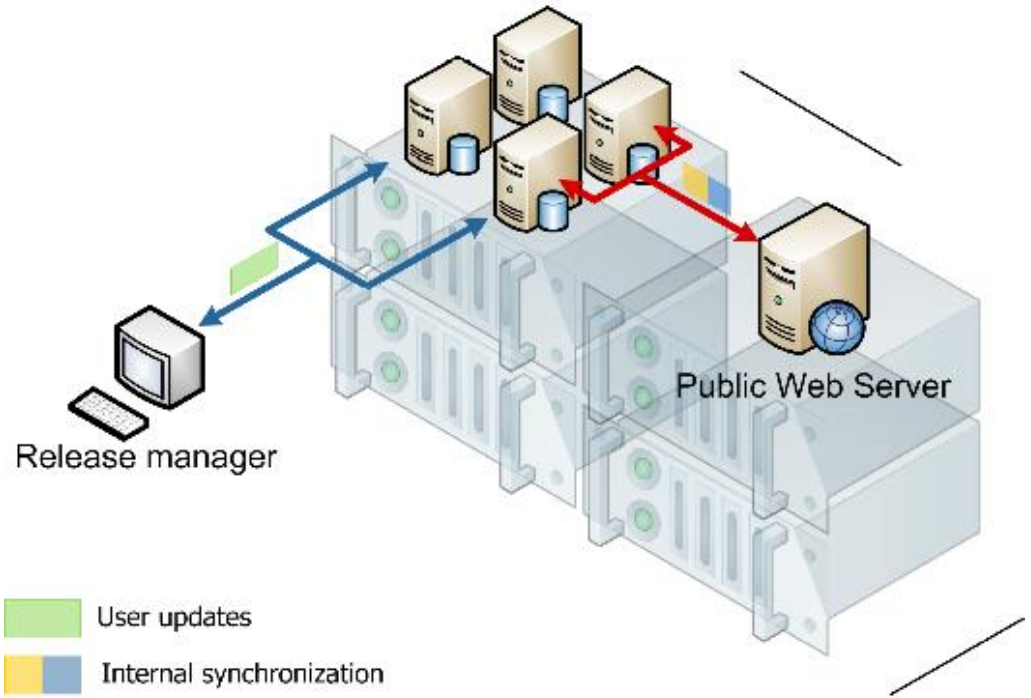
Now Software Engineer at bardeen.ai

Under the wise supervision of **Predrag Buncic**

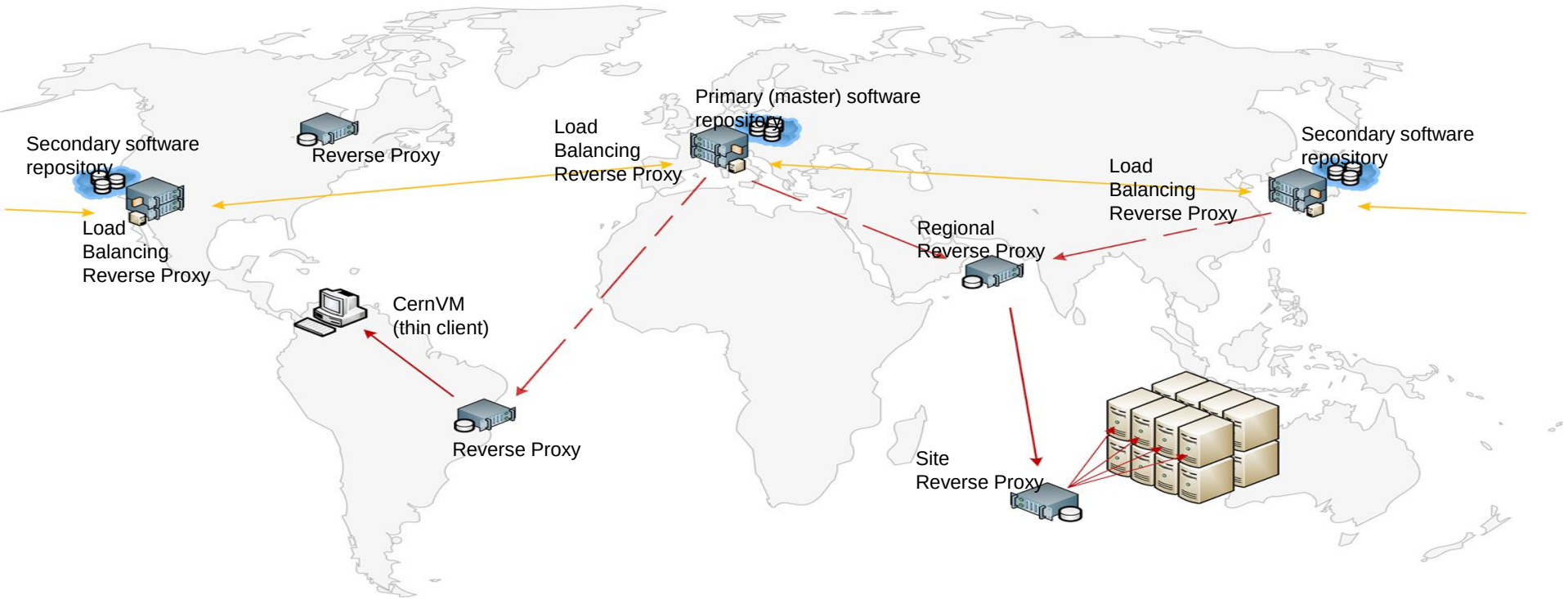
Scalable infrastructure



Publishing releases



Deployment picture



Sales pitch (2009)

- *If you believe that it is time to jump on bandwagon and explore direction in which major software industry players are going and if you have one of these problems:*
  - *You work for LHC experiment and its software is not compatible with your favorite hardware or s/w platform running on your laptop/desktop*
  - *You do not want to spend time to manually keep software up to date*
  - *Your want to profit from the latest developments in CPU technology and use your multi/many core CPU to its maximal potential without modifying your application*
  - *You want to share spare CPU cores/cycles with others*
  - *You want to run your software on voluntary resources beyond the current Grid*

…then CernVM might be what you are looking for.

CernVM Users

~6000 unique IP addresses
(~2009?)

Expanding use cases

- How can we profit from the user workstations' modern (multi and many-core) CPU power and use them at least as a development platform?
- How do we decouple application and infrastructure lifecycles and ensure a homogeneous job execution environment compatible with the one in which the application was developed?
- How can we preserve experiment software and keep it usable and accountable over many years?
- How can unmodified HEP software be used to harvest CPU cycles in a volunteer computing environment?

CernVM-FS

- By the end of 4 four-year R&D project, we were still struggling to get clear support from at least two LHC experiments
  - We had many users from across all four collaborations using CernVM, but only LHCb would speak loudly about that
- Ian Collier from RAL approached us suggesting to **factor out** CernVM-FS from CernVM project
  - It was a hard pill to swallow, but we accepted it under the condition that it remains officially called the CernVM File System.
  - He did excellent work on convincing the HEPIX community that CernVM-FS meets the security requirements. He put convincing arguments backed up by operational evidence that it solves many problems for large sites.
  - That is what finally cut it. CernVM-FS was accepted across the board, and CernVM just won't die ☺

CernVM/FS Today

- LHC Run3 is in a full swing
  - Experiments have collected exabytes of data, routinely processed across WLCG sites that are seamlessly virtualized or containerized.
- The problem of distributing software is almost universally solved using CernVM-FS
  - All LHC experiments, many others in HEP and the broader scientific community rely critically on it.
- Today, fully virtualized infrastructure is often replaced with lightweight containers
  - CernVM-FS can be used from within the container image to reduce image size or to distribute the container root file system itself.
- CernVM still has an important role in the context of data preservation
  - That is how we can access ALEPH data today, and the same might happen to LHC data in 30 years.
- There is a growing interest outside HEP to apply the same technology to possible different use cases
  - Using CernVM-FS as a scalable namespace while accessing multi-PB datasets via other channels.

Summary

- CernVM (including CernVM-FS) is one of the baseline projects within the EP-SFT group at CERN
  - That assures necessary support for the LHC experiments and continuous development to address performance improvement and possible new requirements.
  - The aim is to keep the support and development team at about 4 FTEs.
- We are engaging with external partners (e.g. JumpTrading) to help them address some of the use cases and performance improvements beyond what we need in traditional deployment scenarios for LHC experiments
  - That is beneficial for our community as it helps to uncover and fix the bugs that would otherwise rarely be seen.
- While CernVM–FS was designed with scalability and resilience as primary goals, improving performance is for everyone's benefit
  - We are willing to extend engagements in joint projects aimed at performance improvements, providing that this is not disruptive to our primary goals and obligations towards the LHC experiments and HEP community.

CernVM Core Team Today

Valentin Völkl

Developer and Project lead

EP-SFT staff

Jakob Blomer

Developer and advisor

EP-SFT staff

Laura Promberger

CernVM-FS developer

EP-SFT Fellow funded by JT

Amal Thundiyil

CernVM-FS containerization developer

EP-SFT Technical student, co-funded by JT

+ crucial external contributors and invaluable help from Stratum {0,1} operation teams