VML and CernVM Virtualization for ATLAS

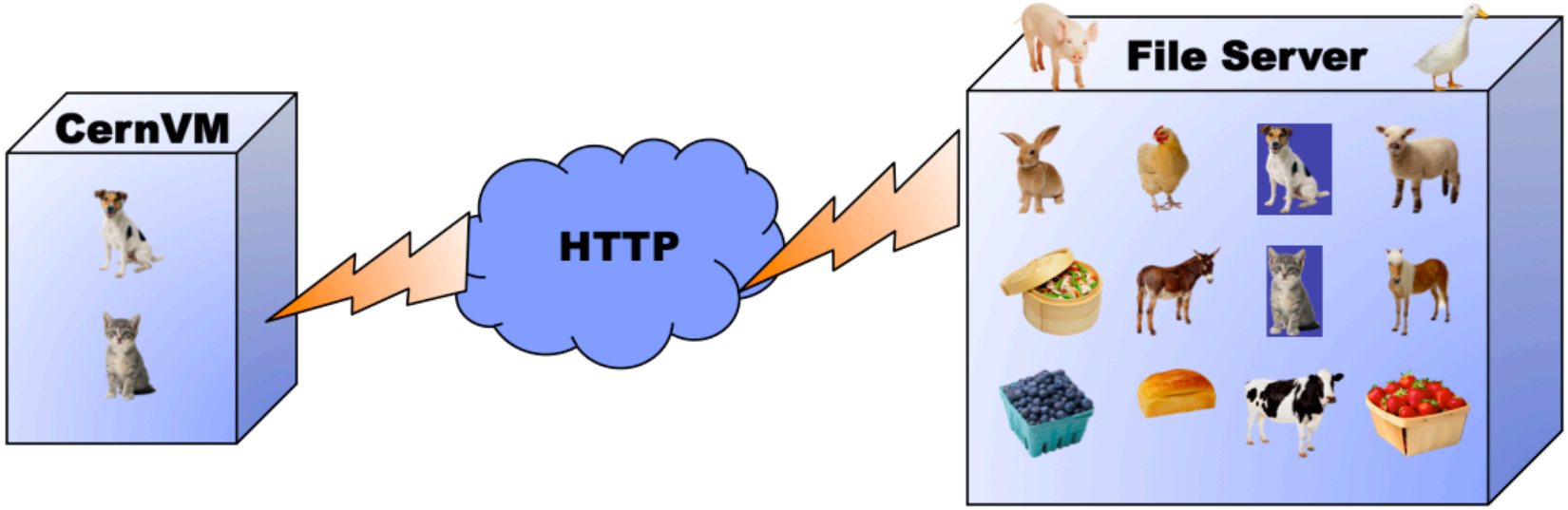
Yushu Yao, Paolo Calafiura, Charles Leggett (LBNL)
Andrea Cavalli, Julien Poffet, Frederic Bapst (EIA, Fribourg)

In collaboration with the CernVM Team

What is CernVM

- Our Goal: do all development/analysis on your own laptop.
- why? Why not lxplus?
  - Always need to connect with ssh, lost connection = lost where I was
  - Always need to compete resources with other users
  - Always need a Network Connection, we are business travellers, we need to work when flying
- Why not download an ATLAS kit into your own laptop?
  - It's TOO BIG (8 GB each release and one release every month, for nightly builds, there is 8GB every day)
  - I don't have SLC on my laptop, so a kit simply doesn't run
- CernVM is a Virtual Machine, which runs on every laptop
- CernVM is Small, but big enough to run ATLAS
- CernVM give you access to ATLAS software, without downloading them.

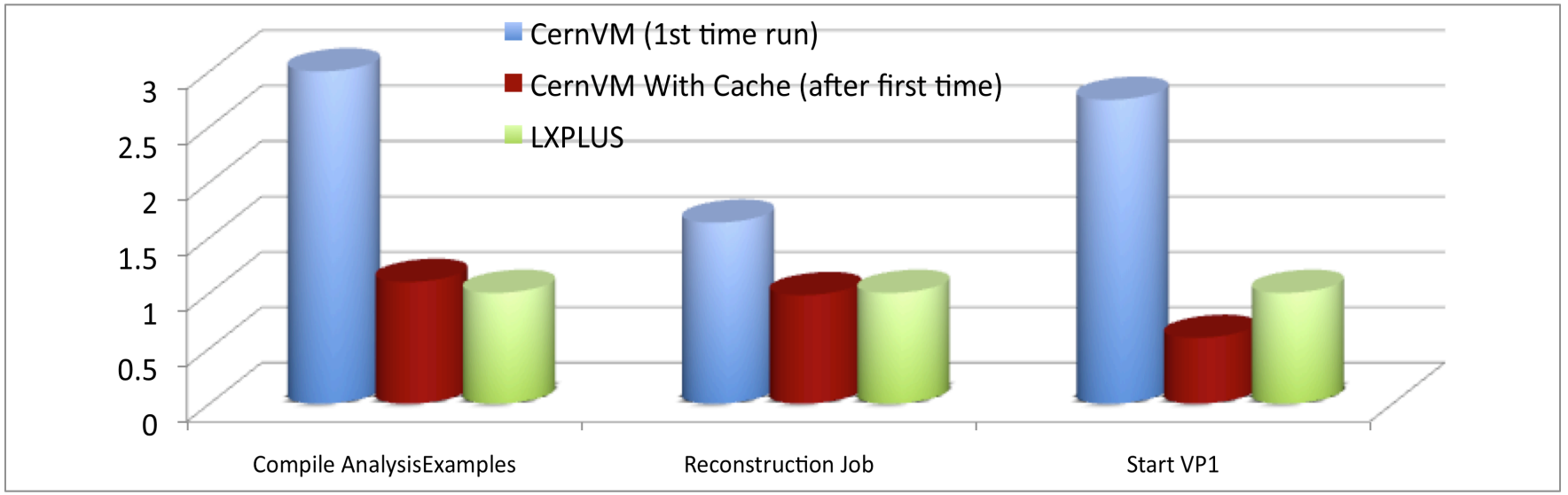
CernVM File System (CVMFS)



- There might be a lot of files inside an ATLAS software kit (8GB)
- Only a fraction of them are needed for certain development/analysis
- We put the ATLAS kits on a HTTP server
- CVMFS will go online to fetch the files when they are first accessed
- After the files are fetched, they are saved locally for future access

Only 700MB download is needed to run full reconstruction, and it is done automatically

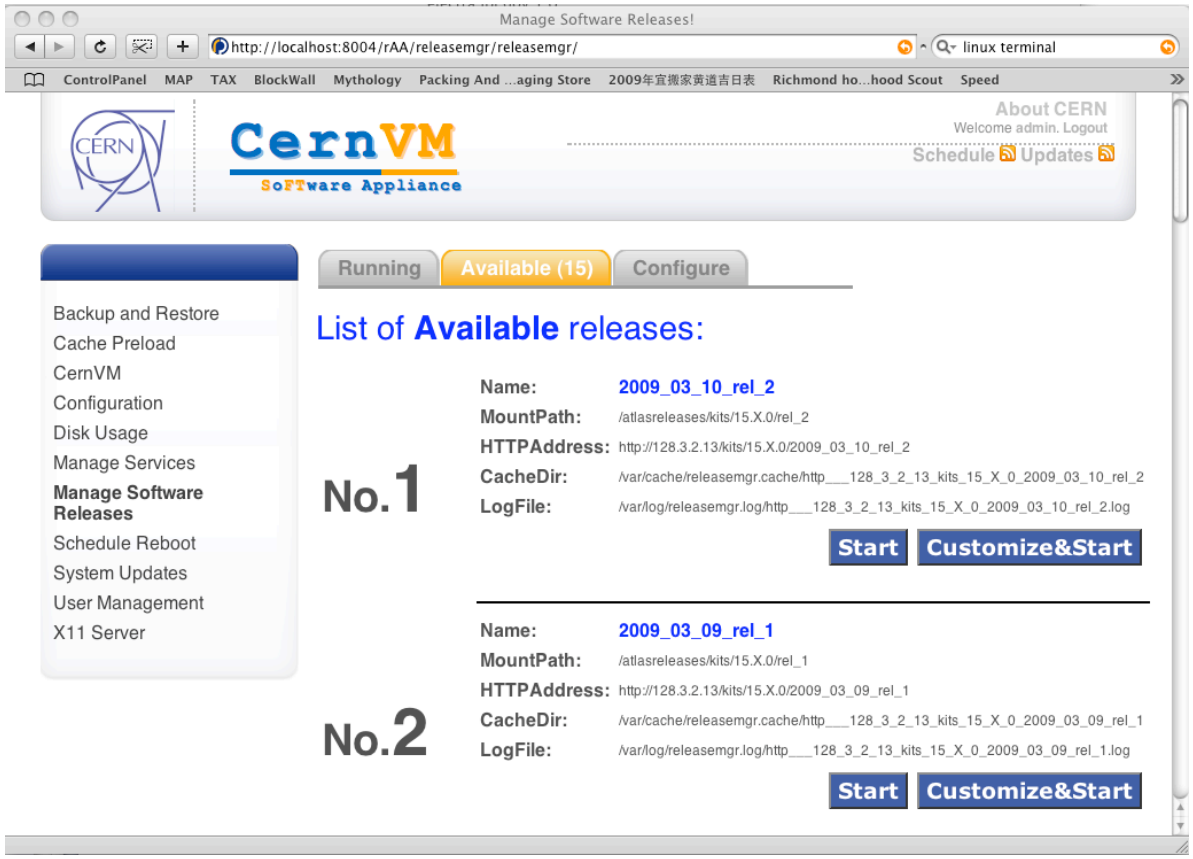
Performance CernVM vs.. LXPLUS



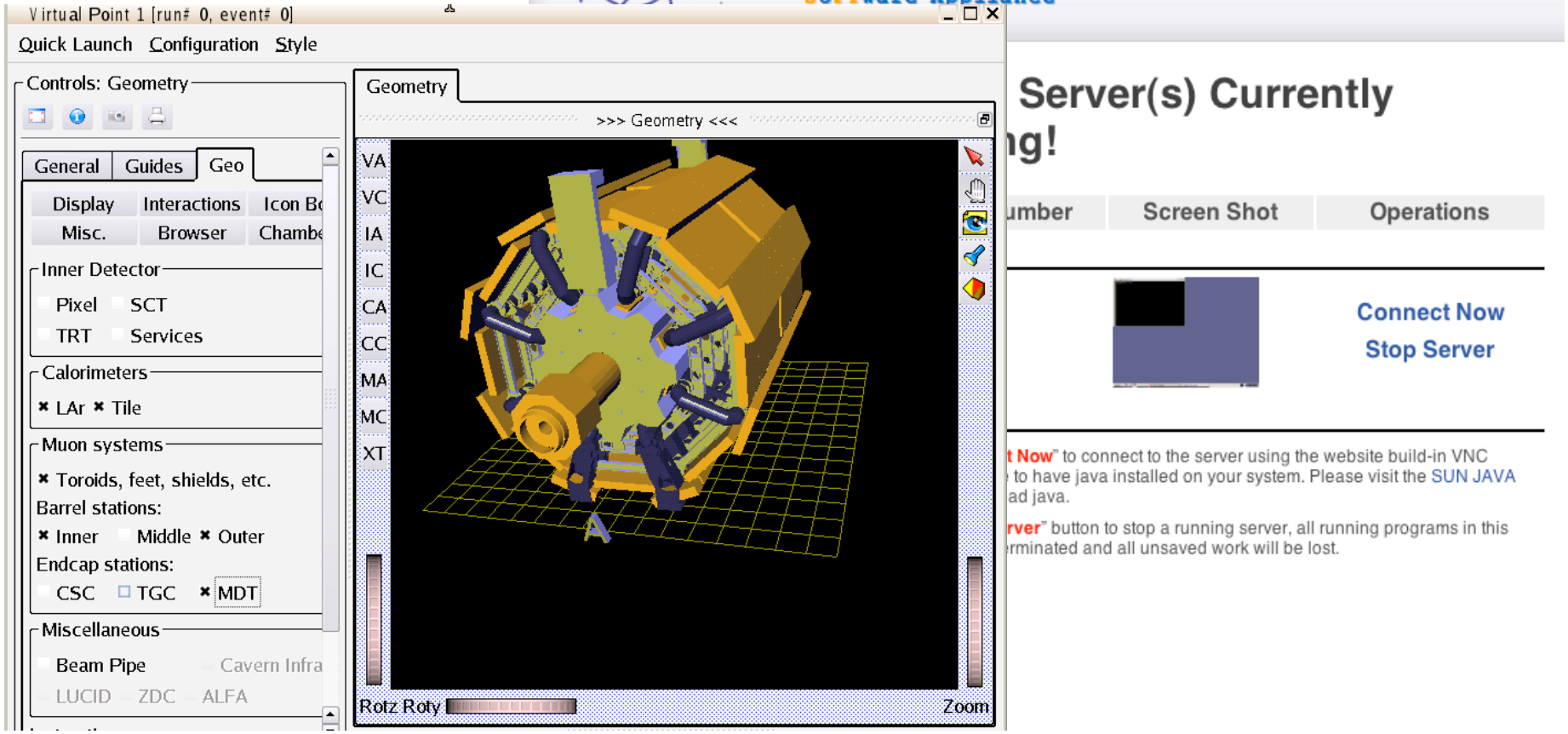
- For the first time running, CernVM is 2-3 times slower, depending on the network speed, since it needs to cache files
- Once the files are cached, the speed on CernVM is roughly the same as on LXPLUS. CernVM is faster when reading large files (since they are local)
- Note: LXPLUS has a higher CPU rate than my test computer (3GHz vs. 2.66GHz)

ATLAS Extension to CernVM

- Release Manage
  - Now user have access to NIGHTLY BUILDS
  - Individual releases could be mounted via a simple web interface



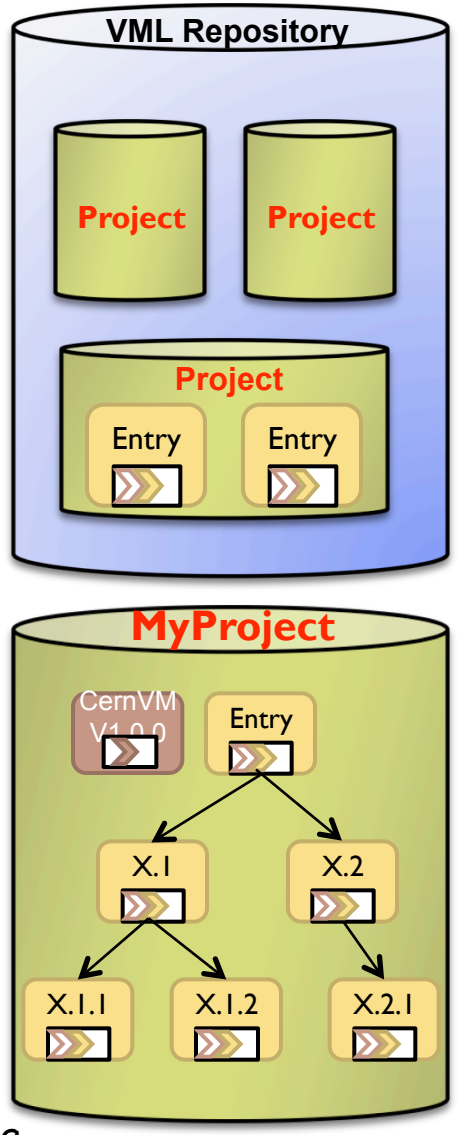
1-click X-Windows and Event Display



Next time LHC is turned on, remember to Launch CernVM with VP1 and see the events live on your laptop.

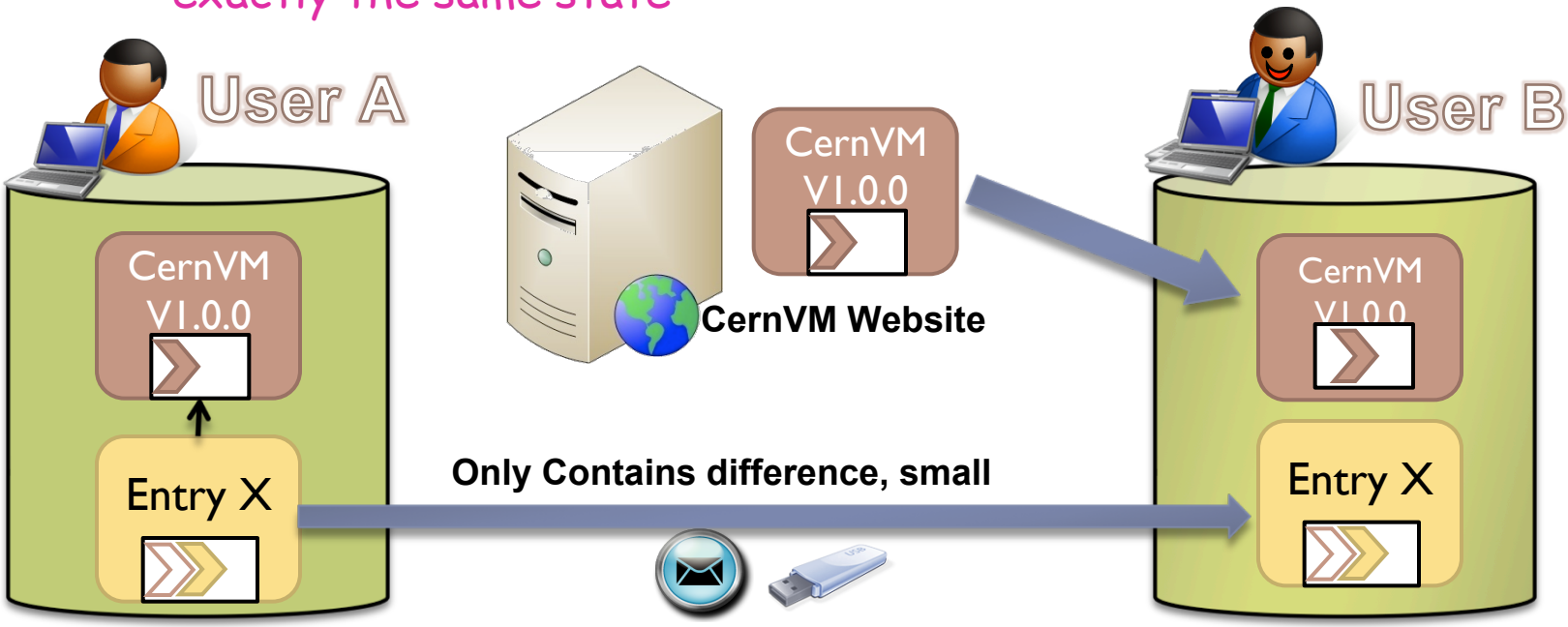
Virtual Machine Logbook (VML)

- VML is a tool that helps you organize and share your virtual machines
  - Work on multiple projects at the same time:
    - Analyze data in Project1
    - Develop new code in Project2
    - Bug fix in Project3
  - Save the states of a project, even build a tree of states.
  - Quickly switch between different projects
  - Automation of CernVM initialization
    - Though CernVM setup is already simple, one can make it simpler with VML. One simple command, VML will download, update and setup a ready to use CernVM.
- Available for Linux/Windows with VMware
  - Support for Mac and other hypervisors are in development



Sharing Work with VML

- All your CernVM projects are based on certain version of CernVM (e.g. 1.01)
- VML can save a state of your work
  - the saved state contains only the difference from its base
  - The difference can be very small, that you can send it over email.
- To share your work, just send the difference to another user
  - VML will obtain the base CernVM automatically and reconstruct exactly the same state



CVMFS to Distribute Software

- We have tested CVMFS not only inside CernVM, it also works in SLC4/5
- Comparing CVMFS to AFS/NFS
  - Although AFS/NFS caches data as well, CVMFS gives more flexibility in managing cache, and can work totally offline.
  - Read-only CVMFS might also give better performance over the read/write AFS/NFS.
- CVMFS can be used to distribute ATLAS software
  - It reduces the work load of institution's administrators
    - No need to download a new kit each time one comes out. When a new Release arrives, it's ready-to-use with CVMFS.
  - Currently we are trying to deploy CVMFS to tier 3's, since they normally don't have high speed network drive to distribute the software.
- We are also investigating the possibility of using CVMFS to deploy software for Batch Systems.

Call for Proxies

- A proxy can let your local user have 10X faster CernVM
- Anybody can setup a proxy server:
  - We will provide help on SQUID install and configuration
  - The server can use port 80, which is normally not blocked by firewall
  - You can setup a proxy which only allow user from your own institution
  - The proxy server only have limited traffic.

Summary

- CernVM
  - Do All Development / Analysis on your Laptop
- Virtual Machine Logbook
  - Organize your CernVM projects
  - Share your work with others
    - Not only sharing the code
    - But also sharing the whole environment (state-of-work)
- Standalone CVMFS to distribute software
  - Save a lot of disk space
  - Much faster than AFS (for outside CERN access)
- Please contribute by setting up your own proxy server.