CernVM[-FS] Status and Development Plans

GDB March 2014

Jakob Blomer, Ian Collier

Agenda

1. CernVM-FS Versions and Components
2. Transition from CernVM-FS 2.0 to CernVM-FS 2.1
3. CernVM-FS Development Roadmap
4. CernVM 3 Status
5. CernVM 3 Roadmap

CernVM-FS Versions and Components

- Components
  1. **Client**
  2. **Stratum 1** software (replication tools)
     Runs at CERN, RAL, FNAL, BNL, ASGC
  3. **Installation Box / Release Manager Machine** software (cvmfs server)
     Runs at CERN for WLCG repositories
     Independent installations at OSG, EGI, DESY, and others

- Compatibility

2.0 2.1	Stratum 0	Stratum 1	Client
Stratum 0	---	OK	OK
Stratum 1	no	---	OK
Client	no	OK	---

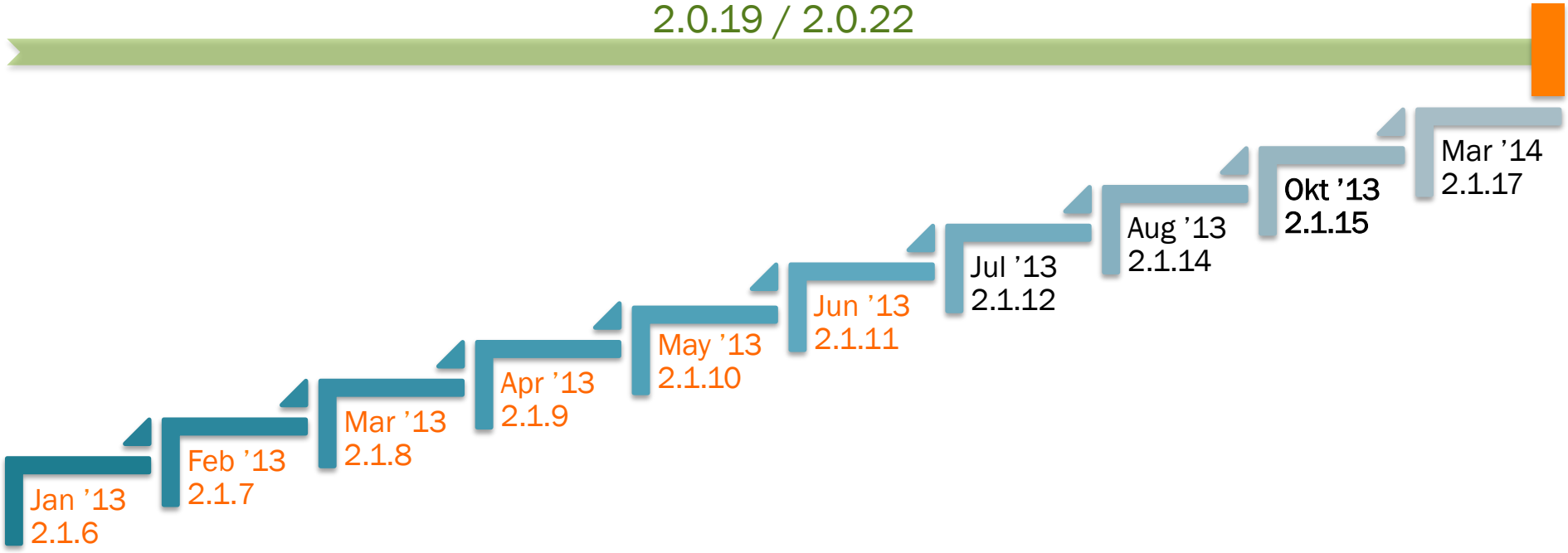
- Need to upgrade all clients and stratum 1s before stratum 0s can be updated

CernVM-FS 2.0 / 2.1 Branches

1. CernVM-FS 2.0
   - The 2.0 branch started in 2009 as a prototype
   - Served us well on the client side,
     however, there are some known “won’t fix” bugs and some feature requests were just not doable with the code base
   - Latency problems at Stratum 1s
   - Serious risks on the Stratum 0: number of inodes, backups, large files
   - Large files in the repositories are an issue on grid sites
   - Platforms: SL5/6 (Client, Stratum 1), SL5 only (Stratum 0)

2. CernVM-FS 2.1
   - Rewrite (2012/’13) based on experience and feature wishes from the 2.0 branch
   - Platforms:
     SL5/6, SuSE, Fedora, Ubuntu, Mac OS X (client only)
     Intel, ARM
   - New features: shared local cache, NFS support, hotpatching,
     file system snapshots, file chunking, …
   - Some of the features require **the full chain** to be on 2.1 branch

CernVM-FS Client Releases 2013/'14



Thanks a lot to all grid sites for patience and help in debugging issues!

Now 100% grid coverage

As a result we are now in much better shape to manage change:

- Fully automated, cloud based unit and integration test suite
  6 platforms, several configurations, multi-hour stress test
- Well-defined release procedure
  http://cernvm.cern.ch/portal/filesystem/release-procedure
- RPM hot-patch updates

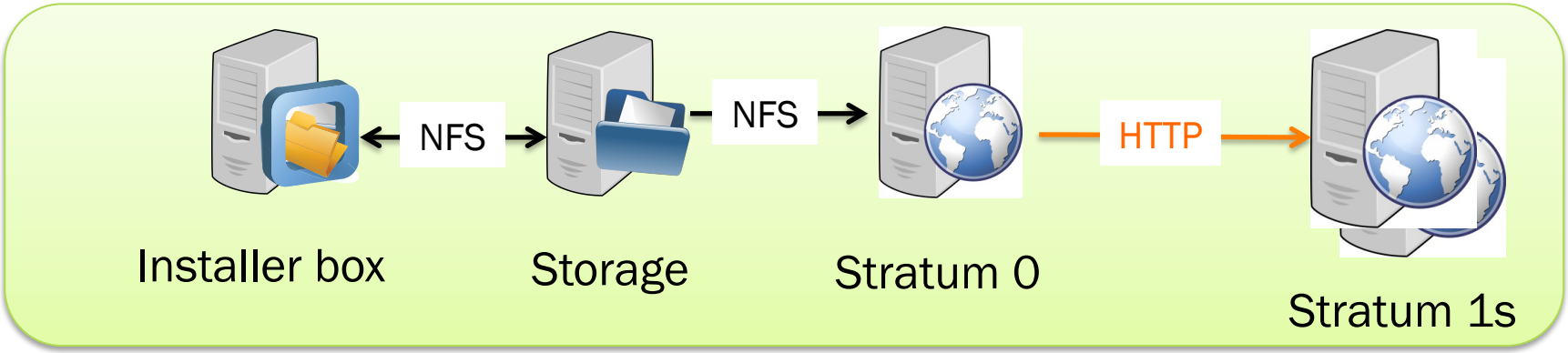
CernVM-FS 2.1 Deployment Roadmap

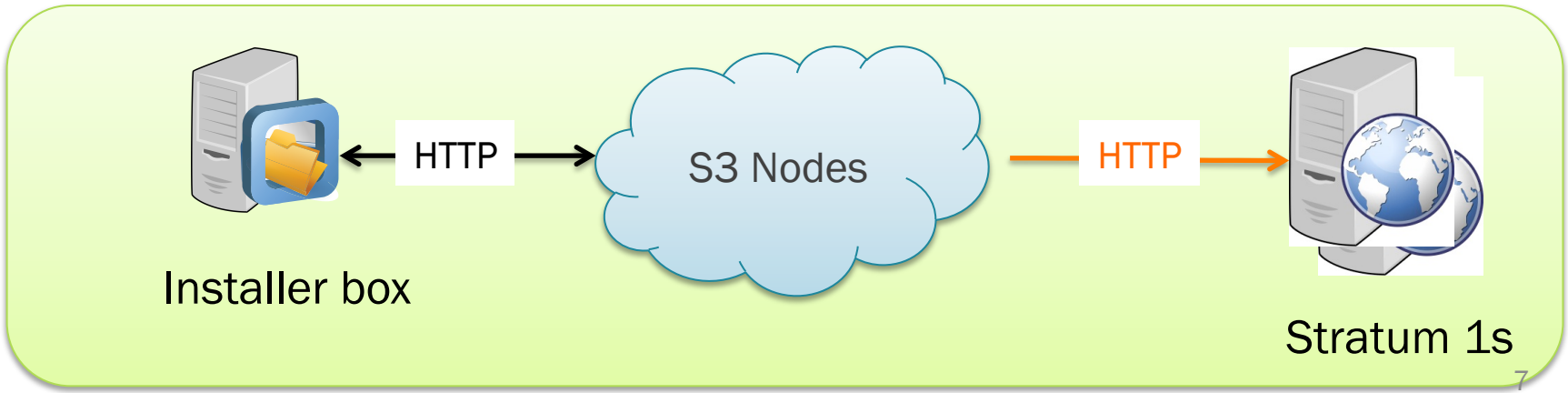
- CernVM-FS 2.1 Deployment Status
  - WLCG recommended client version is 2.1
  - Stratum 1, version 2.1:
    - FNAL, BNL, CERN: production
    - RAL, ASGC: pre-production
  - Installation boxes, version 2.1: outside WLCG, new repositories at CERN, in deployment for ATLAS/LHCb nightlies
  - Reminder: CernVM-FS 2.0 end-of-life end of March 2014

- **April—June 2014: step-wise migration of CERN installation boxes**
  - Good results from load test with LHCb nightly builds (thanks to Ben Couturier!) Version 2.1.18 will contain all bug fixes found during load test
  - Automated migration from 2.0 repository → 2.1 repository in place
  - Test migrations start March/April 2014
  - Step-wise migration starting with smaller WLCG repositories (grid.cern.ch, sft.cern.ch, geant4.cern.ch, …)
  - Last step is the migration of LHC VOs

CernVM-FS Development Plans 2014 – I

1. Support for S3 Backend (contribution by **Seppo Heikkila / CERN Openlab**)
   - Standard with many open source implementations
     Ceph, Riak CS, OpenStack SWIFT, ...
   - Well-suited for many small objects
   - Redundant storage on commodity hardware
   - Possibility to use commercial storage provider (e.g. Amazon)
   - HTTP server built-in





CernVM-FS Development Plans 2014 – II

2. Better support for super computing environments
   Request from CHEP'13
   - Super computing environments are often sealed
   - Idea: create repository replica into the SC cluster file system
   - CernVM-FS replication tools now support replication into a client cache format
   - CernVM-FS client supports "alien cache", an **unmanaged** cache directory on a cluster fs
   - Released with 2.1.17

3. Alternative content hashes
   - Liberates us from SHA-1 dependency
     (note: SHA-1 is still perfectly fine for CernVM-FS use case)
   - As of next release:
     support for RIPEMD-160, support for other hash algorithms can be very easily added
   - Rehashing of existing repositories and 2.1.X client roll-out will probably still require a campaign

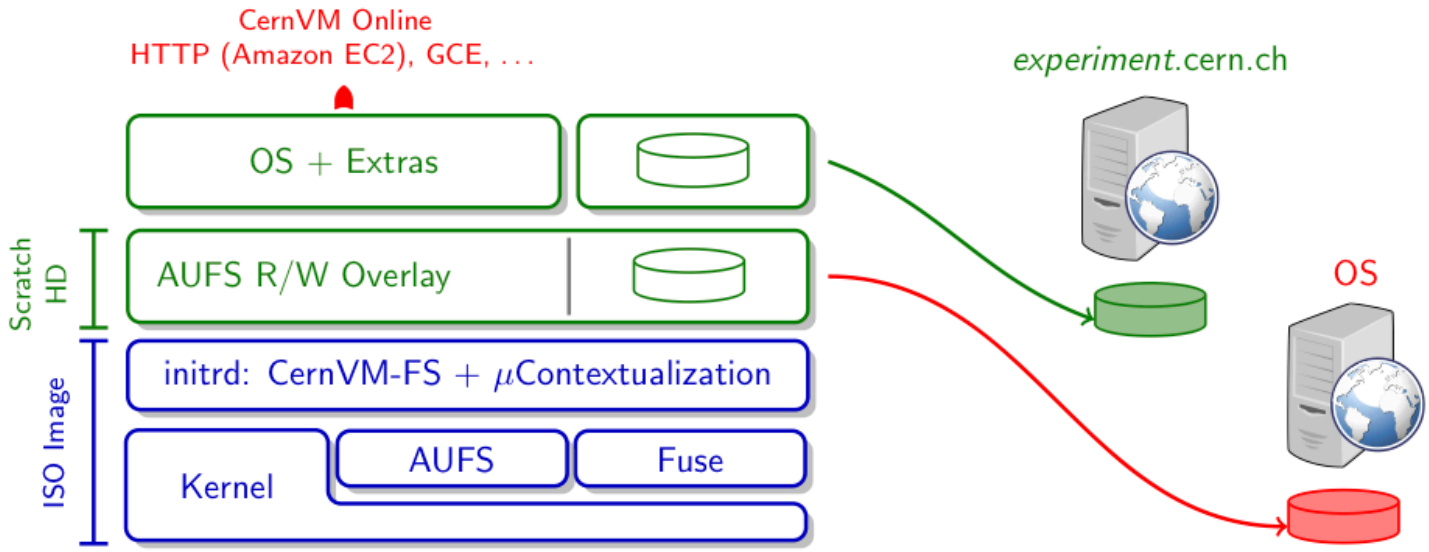
CernVM-FS Development Plans 2014 – III

4. Better support for auto-configuration and small VOs
   Coordination among CERN, EGI, OSG, DESY: https://indico.cern.ch/event/303632
   - Automatic discovery of site proxy servers
     We follow the WLCG proxy discovery task force
     https://twiki.cern.ch/twiki/bin/view/LCG/HttpProxyDiscoveryTaskForce
   - Automatic ordering of Stratum 1 servers
     Initial design for an extension of Stratum 1 web interface exists
   - Distribution and registry of CernVM-FS keys and configuration
     - We expect to see a growth in independent Stratum 0
     - Some can probably survive under the umbrella of a software installation service, e.g. using the EGI service provided by RAL
     - Others, e.g. lcd.desy.de, will be maintained fully independently and develop links to a subset of available Stratum 1s
     - Central key and configuration repository approach: we can provide a central yum and a cvmfs repository (with identical content) to track Stratum 0s and Stratum 1s for software supposed to run on the grid
   - CernVM-FS Monitor: http://cernvm-monitor.cern.ch/cvmfs-monitor/matrix

5. Support for active repository replication (requested by ALICE)
   Replacing the current 15 minutes replication cron job

CernVM Components and Versions

1. CernVM 2.X
   - CernVM 2.X based on SL5
   - Based on Conary and other rPath tools
   - rPath does not exist anymore, we continue to operate in “auto-pilot” mode
   - CernVM 2 EOL: 31. September 2014
2. CernVM 3: denominates a bundle of independent components
   μCernVM bootloader + SL6 subset + CernVM extras
   See https://indico.cern.ch/event/214784/contribution/213, http://arxiv.org/pdf/1311.2426



A single ~20MB virtual machine image using the **same contextualization** across clouds

CernVM Roadmap

1. CernVM 3.1
   - Released January 2014
   - First stable CernVM 3 release
   - Supports both cloud-init and amiconfig (our own) contextualization
   - Using µCernVM 1.16 and SL 6.4 + many extra/non-standard packages: cvmfs, condor, ganglia, puppet, squid, xrootd, cloud clients (nova, ec2-..., gcutil, azure, ...)
2. CernVM 3.2
   - Planned for end of March
   - Based on Scientific Linux 6.5, µCernVM 1.17
   - Google Compute Engine support
   - Integration with cloud-scheduler (VM/job scheduling) and Shoal (dynamic web proxy provisioning)
     https://github.com/hep-gc/cloud-scheduler
     https://github.com/hep-gc/shoal
   - We are happy for feedback!
3. Long-term data preservation (DPHEP) support, work in progress
   - Scientific Linux 4 and Scientific Linux 5 environments based on µCernVM