

CVMFS status within CERN PROD

WLCG GDB

6th April 2011

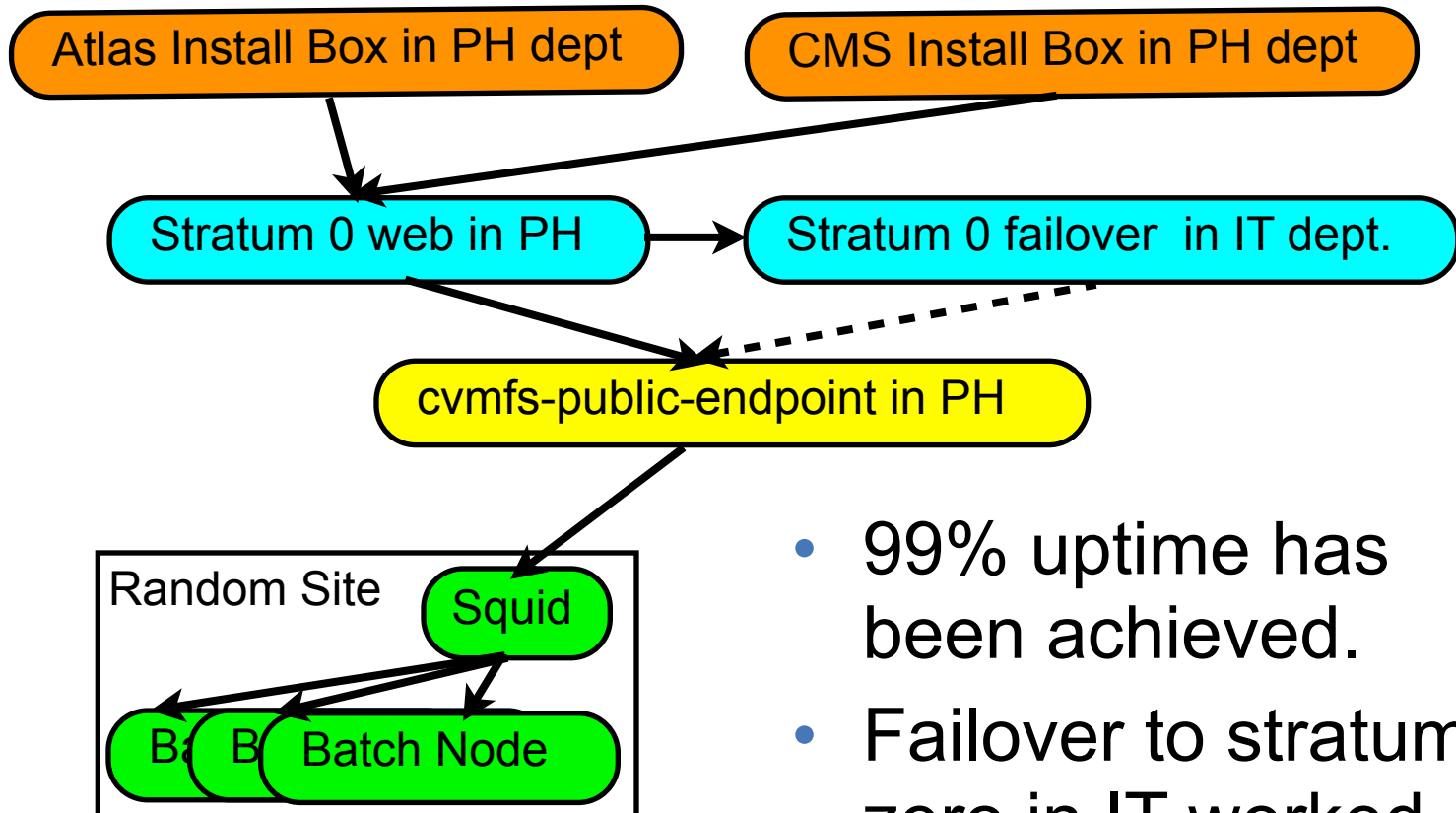
steve.traylen@cern.ch

- One slide intro to CVMFS.
- CVMFS within WLCG.
- Current deployment setup
- Future deployment setup
 - Including motivation for migration.
- Three components for CVMFS at CERN
 - Stratum one deployment at BNL, CERN & RAL.
 - Stratum zero service within CERN-PROD.
 - Client deployment in lxbatch.
- Summary
 - Next steps, timescales.

- CVMFS clients:
 - Provides a /cvmfs/ filesystem area.
 - Files served from a web-server.
- File accesses are intercepted by fuse.
 - Metadata operations (ls, cd, ...) work on an sqlite database - downloaded from web-server.
 - File operations (cat, open, ...) trigger download of file from web server.
- Lot's of caching:
 - On batch workers between jobs.
 - On intermediate squid servers.
- Everything transparent to user.

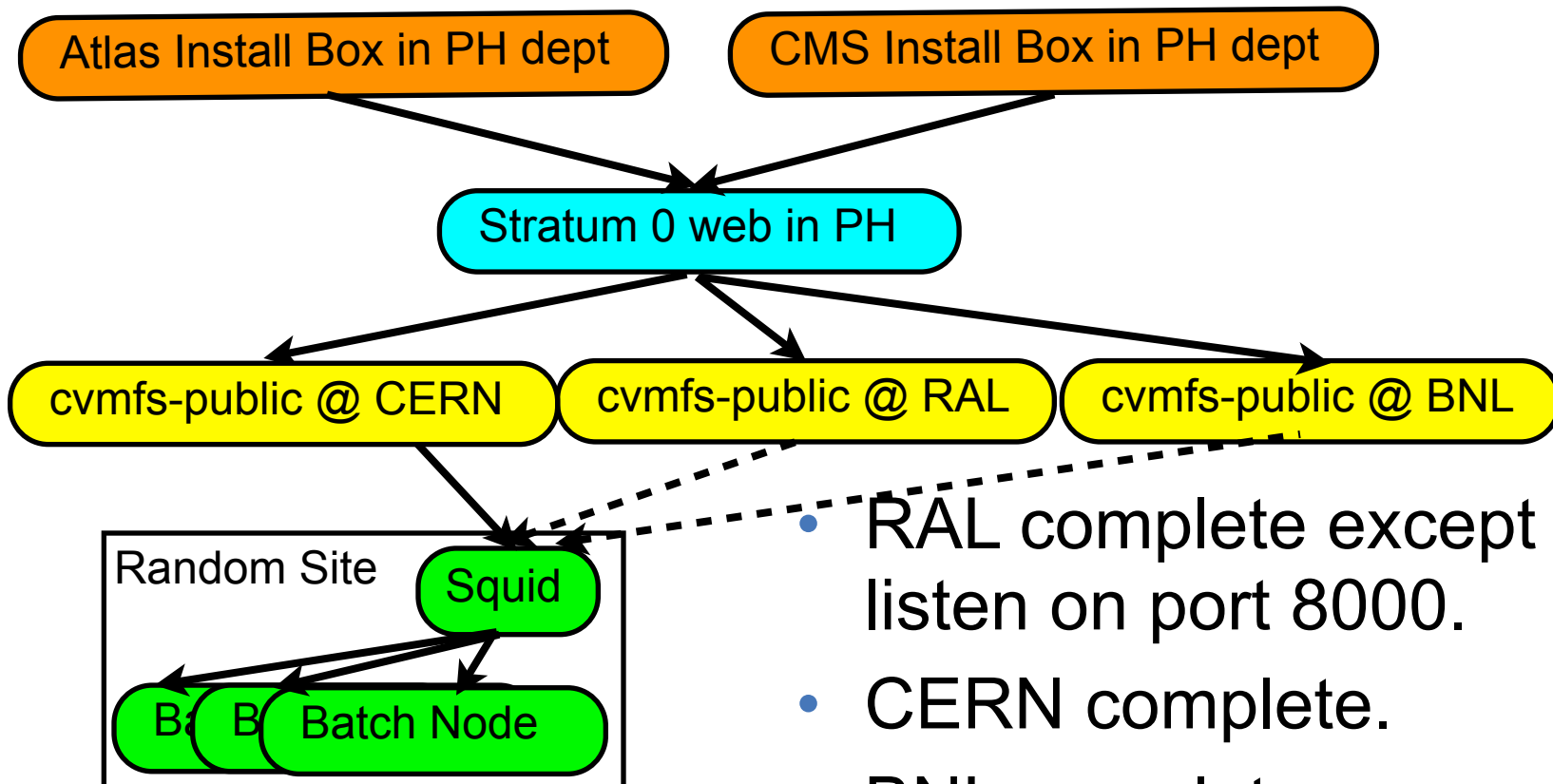
- CVMFS is set to replace current software installation methods.
- Current method:
 - Run <VO>sgm job via grid
 - Write files within job to some shared storage.
 - Validate software.
 - Publish tag in BDII.
 - Process has to be repeated at every site.
 - Process has to be debugged at every site.

- Install once (on stratum zero)
 - Files appear everywhere across WLCG.
- This can be many days faster.
- Hopefully less variation across sites.
 - Common path /cvmfs/...
 - Very few variables:
 - Cache size, Squid QOS.
- Same install bugs everywhere - fix once.
 - e.g LHCb are having problems with CMT usage on /cvmfs but at least it's everywhere.
- Some sites are struggling to provide scaled NFS/AFS.
 - One shared area read by every batch worker.



- 99% uptime has been achieved.
- Failover to stratum zero in IT worked last week and xmas.
- Current setup won't survive CERN outage.

- Introduce a stratum 1 layer.
 - Full replicas at CERN, BNL and RAL.
 - lessens the CERN dependence.
 - Will protect the stratum 0 from direct access.
 - Needs to work smoothly for installations.
 - Devolves responsibility for lower tiers.
 - Stratum 1 is for resilience not scaling.
 - Three is enough, other countries/sites use simple squids.
- Removes service operation from PH group.
 - Allows them do more R&D.



- RAL complete except listen on port 8000.
- CERN complete.
- BNL complete on testbed.
- All sites will have to re-config to use stratum 1.
- Eventually it will be forced.

- Each of CERN, RAL and BNL
- One disk server pulling full copies from the zero.
 - Currently an hourly CRON.
 - Future solutions may use a cgi trigger or messaging.
 - Two load balanced failover reverse squid proxies.
 - Reverse proxies look like a web-server.
 - Stratum 1s will listen on 80 and 8000.
 - Allows sites to choose what is best for them.
- Challenges
 - An individual large site can connect directly.
 - Access may need to be managed if needs be.

- PH department currently runs:
 - A large OpenSolaris ZFS filesystem.
 - One virtual machine per 14 CVMFS repositories.
- IT department currently runs:
 - A large NetApp NFS filesystem.
 - One virtual machine for the it.cern.ch repository.
 - This is very recent - not yet served.
- Two solutions proposed for migration:
 1. Provide new empty repositories in IT.
 - Installers install in both locations for a time.
 2. Serve IT disk space to PH install machines.
 - Migrate behind the scenes
 - Bring up install machines in IT all pre-populated.
- Which ever can be done repo by repo.

- Ramp up over last three weeks complete.
- Monitoring in place.
- Problems encountered:
 - 2 GB was too small for LHCb. Now at 5GB.
- Current status
 - CVMFS is available across all of lxbatch.
- Since Monday LHCb has been configured to use it.
 - Configured via VO_LHCB_SW_DIR
- ATLAS & CMS are invited to use:
 - CMS waiting for some configuration.

- The stratum one service seems close.
 - We enter a testing phase at a few sites.
 - Finally recommend all sites use it.
- Work with ATLAS & CMS to allow them to use lxbatch /cvmfs.
- Complete the it.cern.ch repository to verify the full chain from IT.
- Finalize plan for stratum 0 migration.
- Create a WLCG CVMFS service homepage:
 - Details of recommended cache, squid sizes.
 - Details of Stratum 1 services.
 - Support lines, SLA, SLD.