CVS services for ATLAS (and others)

Helge Meinhard

CERN

Outline

- Introduction
- User perspective
  - Services offered
  - Gotchas/niggles
- System perspective
  - Design goals
  - Implementation
- Steps taken, experience so far; what next?

Introduction (1)

- Large software projects require
  - a version control system
  - a configuration management system
  - a build system
  - developer communication
  - change control, workflows, process model
  - automated testing
  - management
  - …

Introduction (2)

- cvs only addresses the first point (version control system)
  - Repository of (specially formatted) files
  - Checkout/commit
  - Concurrent editing, conflict resolution
  - Tags, branches, …
- All developers need access to repository
  - by using machines that see the repository in the file system (‘local’ mode)
  - by using cvs client/server

Introduction (3)

- ‘Local’ access requires write permission to all developers
  - No guarantee that everybody uses cvs commands
  - Even rm –rf $CVSROOT is possible…
- Large organisations have used AFS
  - Additional problems with stale locks, corrupted files, …

Introduction (4)

- (Obvious) alternative: cvs in client/server mode
- Repository can be on local disk
- Enforces usage of cvs (no rm –rf possible)
- Login to server required for small number of people only

User perspective

User perspective: Services

- cvs server
- cvs notification
- Web server with cvsweb
- Mirroring tool

Services: cvs server (1)

- Access methods:
  - kserver
    - export CVSROOT=:kserver:atlas-sw.cern.ch:/atlascvs
    - Kerberos 4 (same as for AFS)
    - Works automatically when logging into CERN machines
    - No special registration

Services: cvs server (2)

- Access methods (2):
  - pserver
    - export CVSROOT=:pserver:user@atlas-sw.cern.ch:/atlascvs
    - Requires special registration on the server, with password, requires cvs logon on client side
    - Weak security
  - Others: Possible, not currently implemented
    - ssh
    - Kerberos 5
    - ...

Services: cvs server (3)

- Commit and tag access controlled
  - by username
  - for any subdirectory of the repository (implies all subdirectories thereof)

Services: cvs notification

- Mail digest sent to mailing list
- Contains details about all commit and tag operations

Services: Web server

- Web server to provide services requiring local access to repository
- Cvsweb: interactive browsing of repository
  - Check file tree
  - Look at revision history of a file
  - Look at any specific revision of a file
  - Run diff on any pair of revisions of a file
- Editing capabilities of cvsweb not implemented

Services: Mirroring tool

- cvsupd server running
  - Allows for efficiently updating a 1:1 copy of the repository
  - Useful for outside labs
  - … and for our own mirror on AFS

User perspective: Gotchas (1)

- cvs checkout –d <dir> module
  - -d . , -d /an/absolute/path, -d two/levels all don’t work
  - “Long standing, hard to fix design defect in cvs client/server”
- Spurious changes of CVSROOT
  - Workaround in cvs-acl.pl scripts
- Problems with accounts in multiple groups
  - cvs failed if ATLAS was not the primary group
  - Another consequence: Some repository files, after copying over from AFS, had wrong group assignment
  - Being resolved now

Gotchas (2)

- ‘Connection refused’ by cvs server
  - Too many connections within 60 second interval, limit now 240
  - Cures itself after 10 minutes

System perspective

System perspective: Design goals

- Use standards wherever possible
  - Standard machine, OS, management tools, …
  - Well-known and established tools for specific services
- Serving more than one repository from same server should be possible
- Assurance of data integrity
- Service view decoupled from physical implementation

System perspective: Implementation

- System
  - Hardware, OS, disk layout, backups, …
- User administration
- cvs servers
- cvs commit/tag controls, notification
- Web server, cvsweb
- Mirror server
- Data safeguards

System (1)

- 4 year-old PC
  - Tyan Tahoe II, 66 MHz FSB
  - 2 Intel Pentium II 300 MHz
  - 256 MB
  - System disk: 10 GB IDE non-mirrored
  - Intel Pro/100
  - Adaptec AHA 2940 UW
- In air-conditioned server room, under UPS
- Connected to CERN backbone via 100 Mbit/s switch

System (2)

- OS: CERN Linux 6.1.1
  - with AFS, SUE (default/CERN), sshd
- Data disks: 2 Seagate Barracuda 9 GB 7200 rpm
  - Software RAID (level 1 – mirrored)
  - 8.5 GB mounted as /local

System (3)

- (Simplified) layout of /local

/local/atlas/cvs		Repository
	cvslock	Reserved for locks
	cvsup	Conf. files for cvsupd
	httpd	cvsweb conf. and hooks
/local/home/atlascvs		cron jobs running under ATLAS account
/local/home/httpd		cvsweb scripts

/atlascvs $\rightarrow$ /local/atlas/cvs

System (4)

- Backup
  - Every night
  - / (including /etc), /local, /var
  - using TSM (ex-ADSM)
- Unnecessary services disabled
  - ftp, telnet, shell (rsh), login (rlogin)
  - Only connection: ssh
- DNS aliases defined: atlas-sw, chorus-sw

User administration

- All cvs users need to be known on server
- Daily synchronisation with lxplus cluster
  - /etc/passwd, /etc/group, /etc/account
  - Can add local users, otherwise no change
- Interactive access blocked for most users at HEPiX shell script level
  - SUE feature project_pdp_acl with local configuration files

cvs servers

- Started by inetd on request
- /etc/services: Added cvskserver on port 1999 (cvspserver already defined on port 2401)
- /etc/inetd.conf: Added cvspserver and cvskserver
  - running /usr/local/bin/cvs pserver/kserver
  - Parameters: stream tcp nowait.240 root

Cvs: Controlling commits/tags

- Perl scripts hooked into commitinfo and taginfo
  - All calling cvs_acls.pl
  - Separate configuration for commits and tags
- Non-zero exit aborts commit or tag operation
- Another perl script to detect and block attempts to move or delete tags

Cvs notification

- taginfo, loginfo files make sure every tag and commit gets recorded
  - one file per operation in a subdirectory of CVSROOT
- cron job (running as atlascvs) every 10 minutes
  - Sorts and reformats all files
  - Sends mail [to atlas-cvs-notify@cern.ch]
  - Deletes all files

Web server: Apache

- Apache 1.3.14-2.6.2 rpm (from RedHat 6.1 update area)
- Configuration changes:
  - cgi scripts enabled
  - cvsweb registered as cgi script
  - No automatic server signatures, no version numbers from server
  - Virtual host atlas-sw.cern.ch running from same IP address
    - running cvsweb as user atlascvs, group zp

Web server: cvsweb

- Revision 1.112
  - From stud.fh-heilbronn.de/~zeller/cgi/cvsweb.cgi
  - Newer incarnations exist with additional functionality, eg. for committing and tagging
- Linked with virtual Apache host atlas-sw, tied with ATLAS repository

Mirror server: cvsupd (1)

- Supports mirroring a file tree
  - Particular support for cvs repositories
- Version 16.1 downloaded from www.polstra.com/projects/freeware/CVSup
  - installed in /usr/bin, /usr/sbin

Mirror server: cvsupd (2)

- Configuration:
  - Run daemon as nobody
  - Allow connections from everywhere, but require a password
  - Mask out sensitive information from CVSROOT
- Client is run all 30 mins in order to create a mirror of ATLAS repository to AFS

Data safeguards

- All relevant user data on mirrored disk
  - cvs repository, configuration of auxiliary tools
- Frequent (30 min) mirroring of repository to AFS
- Daily TSM backup of all machine data

Steps taken, experience so far; what next?

Steps taken, experience so far

- ATLAS migration: done in several steps in June and July 2001
  - Write permissions removed from AFS except for cvs server
  - Moved repository from AFS to local disk
- Some problems fixed, workarounds provided, or left open
- Heavy developer activity since, some (proto-) releases of full ATLAS software built
- Last week: Chorus repository migrated

Steps taken, experience so far (2)

- No major problem so far
  - Machine stuck due to /var/log/lastlog filling up /var
  - /var almost filling up because /var/log/httpd/*log not rotated (sue purge feature disabling RedHat logrotate…)
  - Bad performance when running cvsup client for mirroring onto AFS on server
    - Now moved to separate machine

What next?

- Move to IT provided machine
  - Faster hardware, hardware RAID disks
  - Fully monitored by operators
- Improvements to existing stuff
  - Per-directory notification of commits/tags
  - Access control with cvsupd
- Other accompanying services
  - LXR, Bonsai, …
- Discuss/agree with more potential users
- Define a service (?)