Versionsmanagement jenseits von CVS

Dipl.-Chem. Rainer Orth
Technische Fakultät
Universität Bielefeld
ro@TechFak.Uni-Bielefeld.DE

Übersicht

- Klassische Versionsmanagement-Systeme (RCS, SCCS, CVS) und ihre Probleme
- Neue Ansätze: Konfigurationsmanagement mit Subversion, BitKeeper oder OpenCM
- Exkurs: WebDAV und DeltaV

Wie alles begann: SCCS und RCS

- SCCS: Marc Rochkind, Bell Labs, 1972
- RCS: Walter F. Tichy, Purdue, 1985
- Versionierung einzelner Files
- Symbolische Namen
- Branches
- Locking zur Konfliktvermeidung

Der nächste Schritt: CVS

- Dick Grune, VU Amsterdam, 1986
- Brian Berliner, Jeff Polk, Prisma Inc., 1989
- `http://www.cvshome.org/`
- optimistische Konfliktvermeidung: copy-modify-merge
- hierarchische Operationen
- Vendor-Branches zum Tracken externer Software
- Trennung von Repository und Workspaces
- Client-Server-Betrieb mit einem Server: Standalone-Daemon oder rsh/ssh
- Authentisierung mit Passwort oder GSS-API/Kerberos V5

Probleme mit CVS

- keine Unterstützung von Dateiattributen (z.B. Permissions, Symlinks)
- kein Tracking von File- und Directory-Renames, Copy
- damit kein Konfigurationsmanagement, keine benamsten Konfigurationen
- keine Gruppierung zusammenhängender Änderungen: Change Sets, atomare Commits
- fehlende Skalierbarkeit: nur ein Repository pro Projekt, keine zuverlässige Replikation
- ineffiziente Implementierung: Branches, Tagging
- kein disconnected Betrieb, keine lokalen Branches
- große, schwer wartbare Code-Basis, kaum noch Weiterentwicklung

Subversion: ein besseres CVS

- Reimplementierung der CVS-Funktionalität ohne die bekannten Schwächen
- von vornherein modulare Architektur: lokaler oder Netzwerk-Betrieb
- Transaktionssupport: Commits sind atomar
- allgemeine versionierte Objekt-Attribute
- versionierte Copies, Renames
- Branches und Tags in konstanter Zeit
- Command-Line-Client möglichst wie CVS
- basiert auf WebDAV (siehe Exkurs)

Subversion-Implementierung

- `http://subversion.tigris.org/`
- Status: self-hosting, Alpha-Release: 15. Juli 2002, Beta: ca. August
- Client: bisher: Command-Line
- Server: Apache 2.0 mit `mod_dav`, `mod_dav_svn`
- Repository: Berkeley DB 4.0
- Übertragung von XML-Dokumenten
- hält Kopien der unmodifizierten Basis-Revisionen

Subversion-Probleme

- momentan kein disconnected Betrieb, Replikation, lokale Branches
- Webserver-Integration erzwungen: Resource-Management über Projekt-Grenzen problematisch
- keine granulare Access-Kontrolle
- Authentisierung?

Exkurs: WebDAV: Versionsmanagement für „das Web“

- WWW Distributed Authoring and Versioning
- webdav WG: Requirements: RFC 2291, Spec: RFC 2518
  - Locking zur Konfliktvermeidung
  - Properties: Metadaten auf Web-Resourcen
  - Namespace-Operationen
- DeltaV: das „V“ in WebDAV
- deltav WG: Versioning Extensions to WebDAV: RFC 3253

Funktionsweise von WebDAV

- HTTP/1.1-Extensions: neue Methoden für WebDAV
  - Properties: PROPFIND, PROPPATCH
  - Locking: LOCK, UNLOCK
  - Namespaces: MKCOL, COPY, MOVE
- weitere Methoden für DeltaV
  - VERSION-CONTROL, REPORT
  - CHECKOUT, CHECKIN, UNCHECKOUT
  - MKWORKSPACE, UPDATE, LABEL, MERGE
  - BASELINE-CONTROL, MKACTIVITY

BitKeeper: SCCS on steroids

- kommerzielles Produkt: BitMover Inc., Larry McVoy
- `http://www.bitkeeper.com/`
- verteilter/disconnected Betrieb durch Repository-Cloning
- Repositories bilden einen Baum
- aber: beliebige push/pull-Operationen unabhängig von der Hierarchie
- ChangeSets: Konfigurationsmanagement jenseits von Einzelfiles
- Command-Line- und GUI-Client (Tcl/Tk)
- Server: standalone Daemon mit eigenem Protokoll, über HTTP oder per rsh/ssh

BitKeeper (cont.)

- frei für single-user-Betrieb und mit open logging
- Einsatz in größeren Projekten: Linux-Kernel, NTP, MySQL

OpenCM: ein echtes verteiltes CM-System

- Konfigurationsmanagement für EROS-Betriebssystem
- `http://www.opencm.org/`
- Jonathan S. Shapiro, John Hopkins University
- Verteiltes System mit starker Authentisierung, granularer Access-Kontrolle
- Unterstützung für disconnected Betrieb, lokale Branches
- Storage: bisher RCS, zukünftig XDFS (XDelta file system)
- kryptographische Namen: SHA-1-Hashes der Inhalte
- nicht auf Files beschränkt