

Benutzerversammlung  
9. März 2011

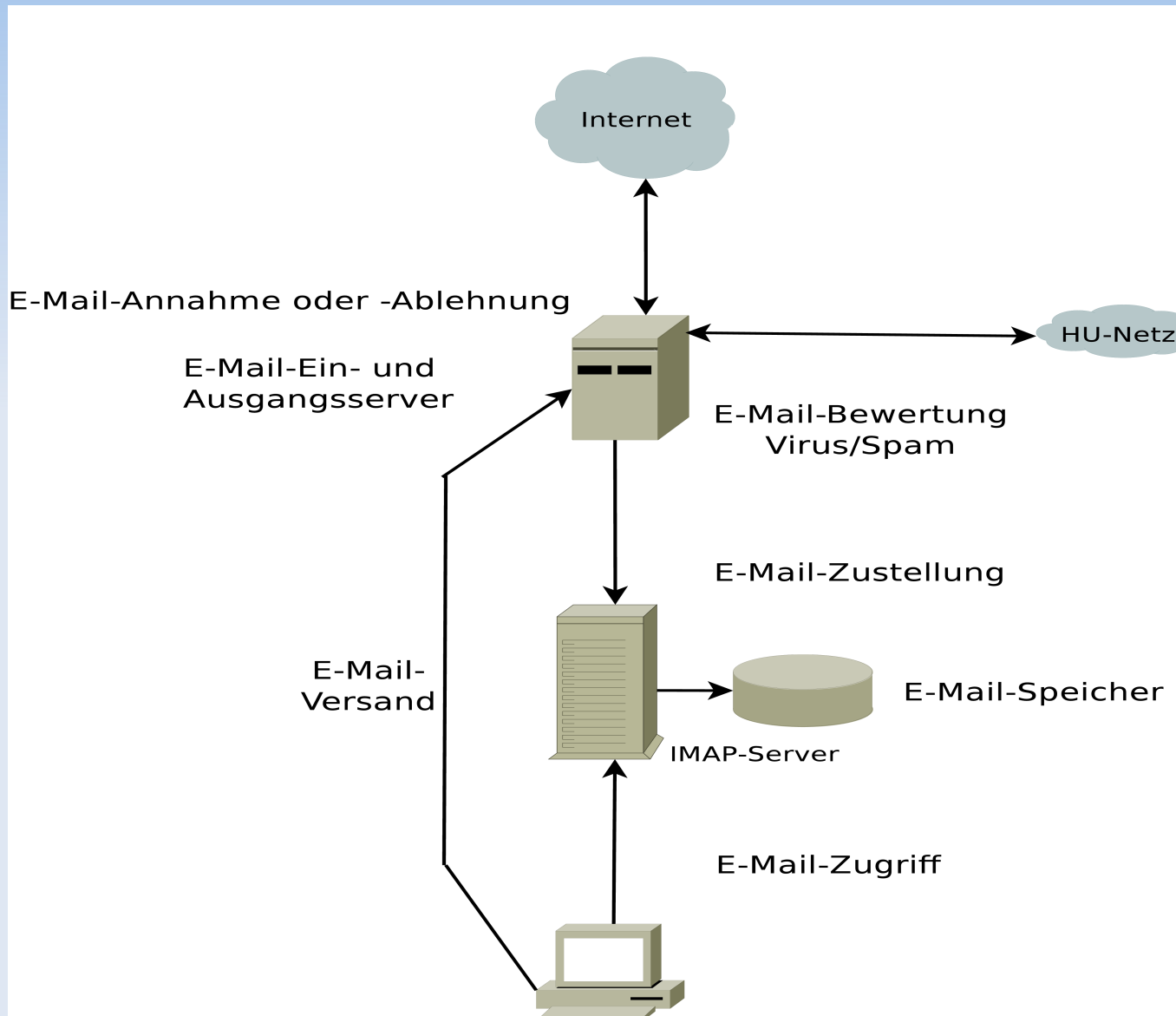
E-Mail:  
erneuerte Infrastruktur

Burckhard Schmidt

# Agenda

- E-Mail – das Prinzip bleibt
- Mailboxformat „maildir“
- Terabyte - sind wir Mail-Messis?
- IMAP-Servervielfalt – ein Ende in Sicht?
- fvf: forward, vacation, filter
- Wermutstropfen

# E-Mail – ein einfaches Prinzip



# E-Mail-Eingangsserver I

- zwei spezielle Server von Cisco/Ironport
- zuständig für externe und interne E-Mails
- Funktionen
  - E-Mail-Annahme oder Verweigerung der Annahme
    - Blacklists und Reputation von Mailservern
  - Spam-Bewertung (Kopfzeile X-Spam-Flag=YES)
  - Virencheck (positiv → E-Mail löschen)
  - Drosselung der E-Mail-Annahme (Anzahl/h)

# E-Mail-Eingangsserver II

- Weiterleitung der E-Mail an den zuständigen IMAP-Server mit der Mailbox
- E-Mails aus dem HU-Netz: „nur“ Virenprüfung
- Phishing Mails, Phishing-URLs
  - ein preisgebener Account reicht dem Spammer
- Versand von Spam aus dem HU-Netz passiert

# Die Mailbox

- Mailbox: account auf einem IMAP-Server
  - Account steht für Speicherplatz und Eindeutigkeit
  - Bsp.: *schmidwa@mailstore2.cms.hu-berlin.de* Alias: *walter.schmidt*
  - IMAP-Server ist somit hier „mailstore2.cms...“
- Diese Informationen stehen im LDAP-Server.

# IMAP-Server

- Server verwaltet E-Mails je Nutzer (account)
  - realisiert mit einer IMAP-Server-Software
  - Authentisierung gegen LDAP-Server bei Zugriff
- „viel“ Speicherplatz aus dem SAN
- mehrere IMAP-Server – also Nutzergruppen
- Vermeidung eines Totalausfalls für „alle“

# Mailboxformat maildir

- Eine E-Mail ist eine Datei!
  - mbox: ein Ordner mit vielen E-Mails ist eine Datei
- sehr viele Verzeichniseinträge (Filesystem)
- schneller Zugriff (u.a. Index-Strukturen)
- geringe Datenmengen sind zu lesen/schreiben
  - I/O-Durchsatz wird bei mbox zum Problem



# „maildir“ und Datenmengen

- Datenbestand 3.000.000.000 kB Mail (3 TB)
- maildir-Format
  - 10kB/Mail → 300 Mio. Dateien (Mails)
- mbox-Format
  - 40.000 Nutzer:
  - Mit je 100 Ordnern → 4 Mio. Dateien

# Datensicherung

- Dateisicherung:
  - Welche Dateien haben sich verändert?
  - Abgleich von 300 Mio. Dateien (mbox: nur 4 Mio.)
- Dateiwiederherstellung
  - 30 ms je Datei (Verzeichniseintrag, ohne Suche im Bandarchiv)
  - Dauer ca. 10 h für 1 Mio. Dateien!
- Fazit: klassische Sicherung ist ungeeignet
- Snapshots von ganzen Dateisystemen nutzen

# IMAP-Server-Software

## Dovecot

- Neuentwicklung
- Formate maildir, mbox
- shared mailbox
- umfangreiche Strukturen für die interne E-Mailverwaltung, z.T. „Selbstheilung“
- viele Optionen (autocreate, expire, sieve, ...)

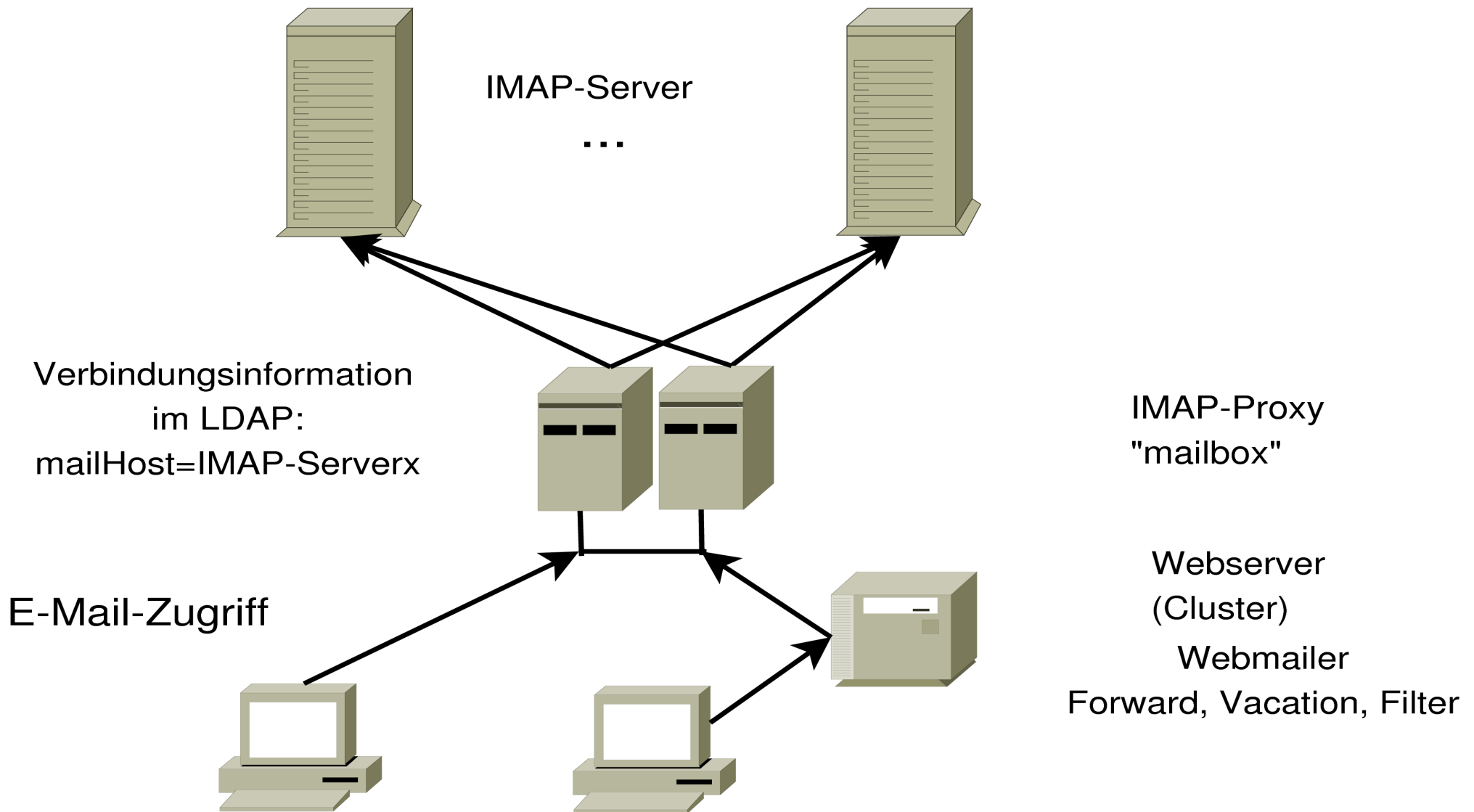
# IMAP-Server-Software im CMS

- Ziel: Dovecot für alle
  - Version 2.x, maildir, shared mailbox
  - Aufwand: Migration der Mailboxen
- aktuell:
  - UW-Server hauptsächlich für Studierende
  - Cyrus-Server für die Uni-Verwaltung
  - Cyrus-Server für Alumni
  - Dovecot-Server V 1.x (mbox) für alle anderen
  - Dovecot-Server V 2.x (maildir) für 80 „Tester“

# Proxy-Server

- mehrere IMAP-Server
  - Nutzer: Wie heißt mein IMAP-Server?
  - symbolisch: mailbox.cms.hu-berlin.de
- Der Proxy verbindet den E-Mailclient mit dem „echten“ IMAP-Server.
- LDAP hilft auch hier wieder
  - Die Information ist am Account „fest gemacht“.

# Proxy-Server steuert den IMAP-Zugriff



# Forward, Vacation, Filter

- <https://forward.cms.hu-berlin.de>
- Einwilligungserklärung, Widerruf
- Weiterleitung (forwarding)
- Abwesenheitsbenachrichtigung (vacation)
- Filter für E-Mails (filter)
  - Ordner AutoCleanSpam
  - X-Spam-Flag=YES
  - automatisches Löschen nach frühestens 12 Wochen

# GPR – Datenschutz - E-Mail

- Die Inbetriebnahme der erneuerten Infrastruktur darf noch nicht erfolgen.
- Dienstvereinbarung „Antispam“ (Versuche seit Ende 2008) kam nicht zu Stande.
- Seit Herbst 2010: Erfolg versprechende Bemühungen, die Betriebserlaubnis durch einen Mitbestimmungsantrag beim GPR zu erlangen.
- Der Datenschutzbeauftragte der HU ist natürlich einbezogen.



# Zusammenfassung

- zuverlässige E-Mail-Eingangsserver
- Vereinheitlichung der IMAP-Server (dovecot 2.x)
- IMAP-Proxy im Verborgenen
- Datensicherung – neue Methoden erforderlich
- Webserver mit den bekannten Webmailern
  - FVF-Interface in den Startlöchern
- alles nur mit Mitbestimmungsantrag beim GPR