

Email Protokolle

Thomas Starzacher
Christian Prähauser

29. Jänner 2003

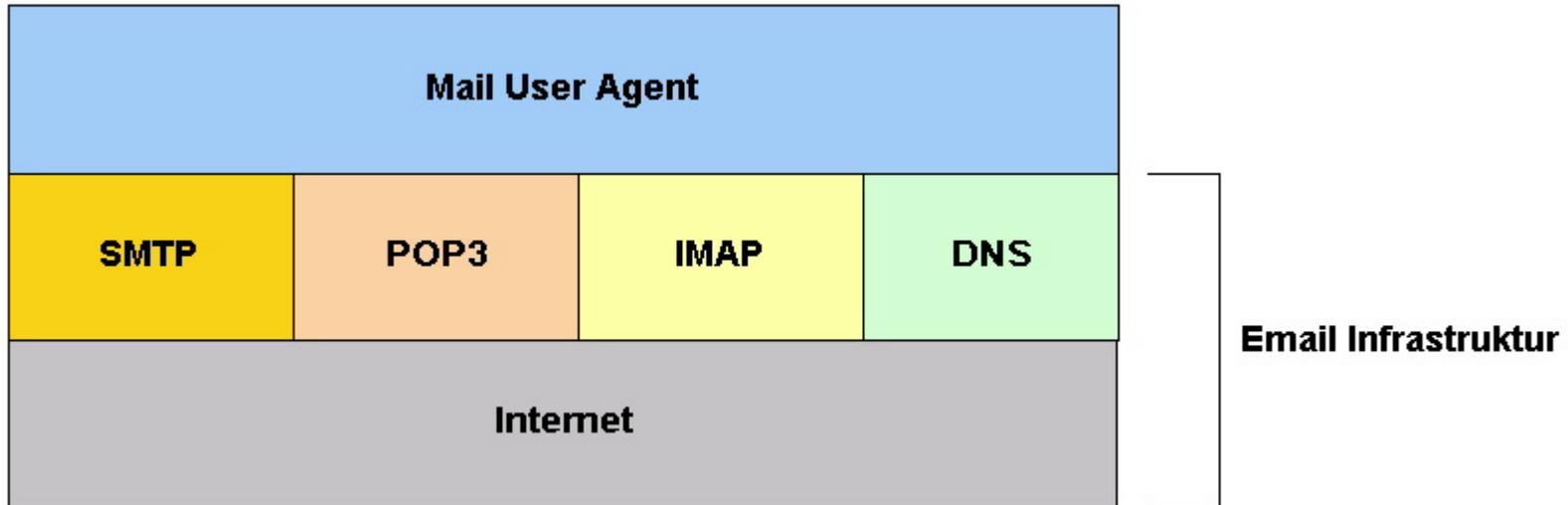
Übersicht

1. Email Grundlagen
2. SMTP Protokoll
3. POP3 Protokoll
4. IMAP Protokoll

Email Grundlagen

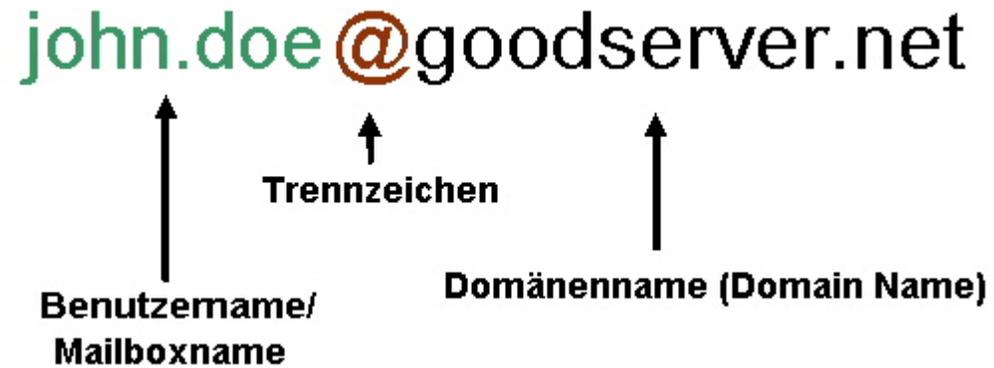
- Email System besteht aus
 - *Mail Transfer Agent (MTA)*
 - *Mail User Agent (MUA)*
- MUA ist die Schnittstelle zwischen Email System und Benutzer
- Seit 1995 immer ausgereifere User Agents
- MUA verwendet meist **POP3** oder **IMAP**
- MTA ist verantwortlich für die Übertragung der Nachrichten
- MTA sind meist **SMTP**-Server

Email Infrastruktur



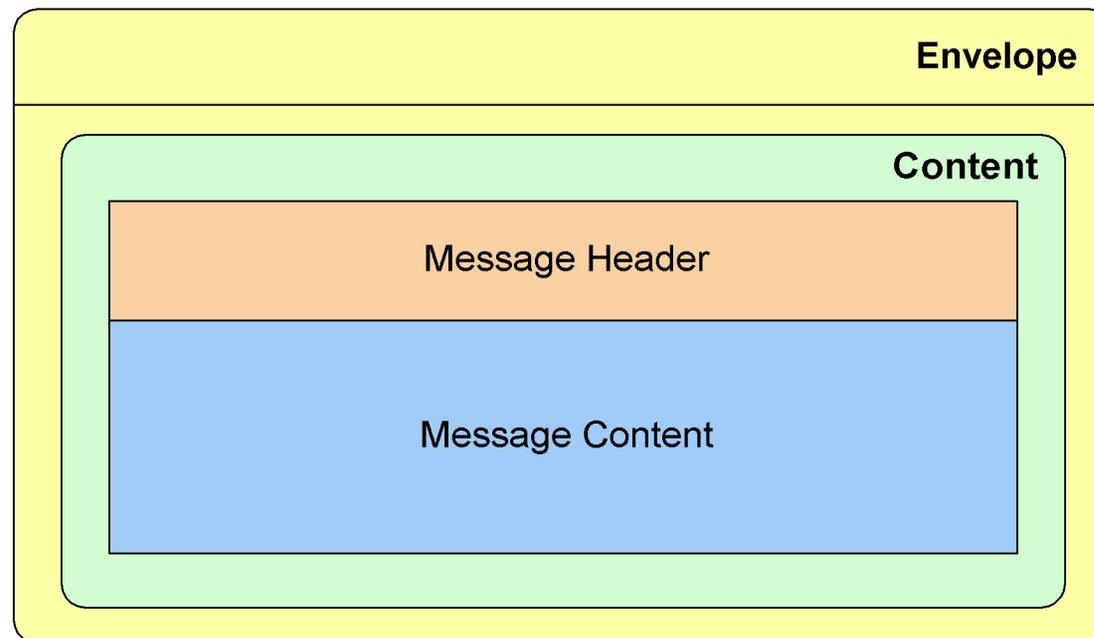
Internet Email Adressen

- Internet Email Adressen bestehen aus
 - einem Benutzernamen und
 - einem Domänen Namen (*Domain Name*)
- werden durch den sog. "Klammeraffen" (@) getrennt
- Domänen Namen werden verwendet um mittels **DNS** den SMTP Server des Empfängers zu ermitteln (MX - Einträge)



Das Format von Email Nachrichten

- Das Format von Internet Email Nachrichten ist definiert im RFC822 bzw. RFC2822 (RFC steht für Request For Comment und dient der Standardisierung von Internet Technologie)



- Email Nachrichten bestehen aus folg. Komponenten
 - Umschlag (Envelope)
 - Inhalt (Content)
 - * Nachrichtenkopf (Message Header)
 - * Nachrichteninhalte (Message Content)
- Der Umschlag enthält Informationen die für die Zustellung durch einen MTA notwendig sind (Sender, Empfänger, usw.)
- Der Inhalt umfasst die eigentliche Email Nachricht, die zu übertragen ist.
- Nachrichten bestehen aus Zeichen des US-ASCII Zeichensatzes (7 Bit)
- Der Nachrichtenkopf (oder Header) setzt sich aus mehreren Feldern der Form **Feldname: Feldkörper** zusammen

Feldname	Bezeichnung	Beschreibung
Date	origination date	Zeitpunkt, an dem die Nachricht vom Verfasser abgeschickt wurde
From	originator address	Email Adresse des Verfassers
Sender	sender address	Email Adresse des eigentlichen Absenders der Nachricht
Reply-To	reply-to address	Adresse(n) an die Antworten auf diese Nachricht gesendet werden sollen
To	primary recipients	Enthält die intendierten Empfänger dieser Nachricht
Cc	carbon copy	Enthält Adressen an welche Kopien dieser Nachricht gesendet werden sollen

Feldname	Bezeichnung	Beschreibung
Bcc	blind carbon copy	Enthält Adressen an welche Kopien dieser Nachricht gesendet werden sollen, die jedoch nicht sichtbar sind gegenüber den anderen Empfängern
Message-ID	message identification	Ist eine Zeichenkette (Nachrichten ID) welche beim sendenden Host erstellt wird und für jede Nachricht eindeutig ist
Subject	subject	Kurzbeschreibung des Inhalts der Nachricht
Comments	comments	Enthält weitere Kommentare zum Inhalt
Received	received	Kennzeichnet den Empfang einer Nachricht durch einen MTA

- Die Felder **FROM** und **DATE** müssen vorhanden sein
- Feldnamen die mit **X** beginnen sind für den lokalen Gebrauch reserviert

MIME

- Die Einschränkung von Mail Nachrichten auf 7 Bit ASCII birgt Probleme:
 - Dateien, die nicht auf Text basieren, können durch das Email System nicht übertragen werden. Das Versenden von Audio, Video oder ähnlichen Daten ist nicht möglich.
 - Das Senden von Mitteilungen in Sprachen, die kein Alphabet verwenden (Japanisch und Chinesisch) ist schwierig.
 - Sprachen mit Akzenten sind bei Email Nachrichten eingeschränkt .
- Diese Probleme führten zu einer Erweiterung der Email Definition die **Multipurpose Internet Mail Extensions (MIME)** genannt wird (siehe RFC1341, RFC1521). Aus Kompatibilitätsgründen baut MIME auf dem ursprünglichen Nachrichtformat auf.

- Die Erweiterungen die MIME einführt umfassen 5 neue Email Header-Felder. Diese neuen Felder sind:
 1. MIME-Version
 2. Content-Type
 3. Content-Transfer-Encoding
 4. Content-Id
 5. Content-Description

Typ	Subtyp	Beschreibung
text	plain	Normaler ASCII Text (unformatiert)
	html	HTML Text
	xml	XML Text
	richtext	Text im RichText Format
audio	basic	Audiodaten in einem PCM Standardformat (1 Kanal, 8 KHz)
	mpeg	Audiodaten im MPEG Format
video	mpeg	Videodaten im MPEG Format
image	gif	Bilder im GIF Format
	jpeg	Bilder im JPEG Format
	tiff	Bilder im Tagged Image File Format
message	rfc822	Nachricht nach RFC822
	partial	Teil einer Nachricht die zur Übertragung zerlegt wurde
	external-body	Die Nachricht muss von der angegebenen Adresse geholt werden
	html	Eine Nachricht im HTML Format

Typ	Subtyp	Beschreibung
application	zip	Daten die mit dem ZIP Verfahren komprimiert wurden
	octet-stream	Daten die zur weiteren Verarbeitung gespeichert werden sollten. Daten sind auf Byte Grenzen ausgerichtet
	pdf	Portable Data Format Dokument
	postscript	Daten im Adobe PostScript Format
multipart	mixed	Daten enthalten selbst wieder MIME Entitäten
	digest	Daten enthalten MIME Entitäten vom Typ message

Email Header

Ein Internet Email Header...

Received: from out10smtp6.ddc.dartmail.net [216.73.90.106] by elsnet.at (SMTPD32-5.01) id AC1B1E6900D4; Fri, 27 Dec 2002 20:38:03 +0200

Received: from core10pumper1 (unknown [10.21.8.70]) by out10smtp6.ddc.dartmail.net (Postfix) with SMTP id A16841FE617 for <praehauser@elsnet.at>; Fri, 27 Dec 2002 14:33:41 -0500 (EST)

Message-Id: <Kilauea69640-18088-50409678-1@flonetwork.com>

From: ExtremeTech Daily <daily@extreme.ziffdavis.com>

To: praehauser@elsnet.at

Subject: ExtremeTech Daily: Dawn of a Post-Digital Era In 2003?

MIME-Version: 1.0

Content-Type: text/html; charset="us-ascii"

Content-Transfer-Encoding: 7bit

Date: Fri, 27 Dec 2002 14:33:41 -0500 (EST)

X-RCPT-TO: <praehauser@elsnet.at>

X-UIDL: 341008044

Status: U

Das SMTP Protokoll

- SMTP steht für *Simple Mail Transfer Protocol* und wurde Anfang der 80er Jahre entwickelt
- Seine Aufgabe ist es text basierte Nachrichten an die Mailboxen oder Terminals anderer Benutzer zu übertragen
- Ursprünglich definiert in RFC821 (aktuelle Spez. ist RFC2821)
- SMTP ist offizieller Internet Standard (STD10)

Funktionsweise

- SMTP basiert auf einer Client-Server Architektur
- Heutzutage wird meist TCP (Transmission Control Protocol) als Kommunikationskanal für SMTP Übertragungen genutzt (weitere mögliche Medien wären z.B. Network Core Protocol oder X.25)
- Will ein SMTP-Client eine Nachricht übertragen baut er ein Verbindung zu einem SMTP-Server über Port 25 auf
- Kann der Server die Nachricht nicht direkt zustellen (d.h. ist die Mailbox des Empfängers nicht lokal) leitet er die Nachricht an einen anderen SMTP Server weiter. Er agiert als sogenannter *SMTP Relay*

- Ein *SMTP Gateway* ist im Prinzip ein SMTP Relay, allerdings befindet sich der Server an der Grenze zwischen zwei Netzen, die unterschiedliche Mail Protokolle verwenden.
- Die Interaktion zwischen Client und Server funktioniert mittels
 - SMTP Kommandos (SMTP Commands) und
 - SMTP Antworten (SMTP Replies)



- Ein SMTP Kommando besteht aus 4 ASCII Zeichen und kann Argumente haben

- Die Argumente (falls vorhanden) sind durch ein Leerzeichen vom Kommando getrennt
- Kommandos werden durch die ASCII Sequenz <CRLF> beendet; CR = Carriage Return (0x0d) und LF = Line Feed (0x0a)
- SMTP Antworten haben die Form: **xyz optionalerText**
 - **x** gibt die Art der Antwort an, z.B. positiv, negativ, unvollständig, usw.
 - **y** teilt die Antwort in verschiedene Kategorien ein
 - **z** gibt detaillierte Auskunft zu jeder Kategorie
- Der Client schickt ein Kommando an den Server und wartet auf eine Antwort. Solange diese nicht eingetroffen ist unternimmt der Client nichts (→ *lock step* Verfahren)

SMTP Mail Transaktion

- Ist der eigentliche Vorgang des Sendens einer Nachricht
- Wurde eine Verbindung zum SMTP-Server aufgebaut, kann eine Mail Transaktion eingeleitet werden
 - Der Client verbindet sich mit dem Server, indem er eine Transportverbindung aufbaut und danach das Kommando **HELO** sendet
 - Abgebaut werden SMTP-Verbindung durch das Kommando **QUIT**

- Ablauf einer Transaktion

1. Client sendet **MAIL** Kommando mit einem Argument der Form **FROM:<Reverse-Path>**. **<Reverse-Path>** ist eine Zeichenkette welche den Urheber der Nachricht identifiziert
2. Die Empfänger der Nachricht werden bekanntgegeben. Dies geschieht mit einem **RCPT** Kommando für jeden Empfänger (der entsprechende Empfänger wird als Argument mit angegeben)
3. Ist die Liste der Empfänger gesendet worden kann mit der Übertragung der eigentlichen Nachricht begonnen werden: hierzu sendet der Client das **DATA** Kommando.

Nun kann der Client mit dem Senden der eigentlichen Daten beginnen.

4. Das Ende der Daten wird dem SMTP-Server mit einer speziellen ASCII Sequenz aufgezeigt (**<CRLF>.<CRLF>**). Erkennt der Server diese Zeichenfolge antwortet er dem Client mit **250 OK** und sendet die Nachricht.

ESMTP

- SMTP Service Erweiterungen (Extended SMTP) im Jahr 1993 veröffentlicht
- ESMTP ermöglicht die dynamische Erweiterung des SMTP Protokolls
- ESMTP definiert also ein Framework für SMTP Protokollerweiterungen bestehend aus
 1. Ein neues SMTP Kommando (**EHLO**)
 2. Eine zentralen Datenbank in der alle öffentlichen Erweiterungen des SMTP Protokolls registriert werden
 3. Zusätzliche Parameter in den Kommandos **MAIL** und **RCPT**

SMTP Beispiel Sitzung

Ein Beispiel einer typischen SMTP Sitzung...

S:220 foo.com Simple Mail Transfer Service Ready
C:EHLO bar.com
S:250-foo.com greets bar.com
S:250-8BITMIME
S:250-SIZE
S:250-DSN
S:250 HELP
C:MAIL FROM:<Smith@bar.com>
S:250 OK
C:RCPT TO:<Jones@foo.com>
S:250 OK
C:RCPT TO:<Green@foo.com>
S:550 No such user here
C:RCPT TO:<Brown@foo.com>
S:250 OK
C:DATA
S:354 Start mail input; end with <CRLF>.<CRLF>
C:Blah blah blah...
C:...etc. etc. etc.
C:.
S:250 OK
C:QUIT
S:221 foo.com Service closing transmission channel

Das POP3 Protokoll

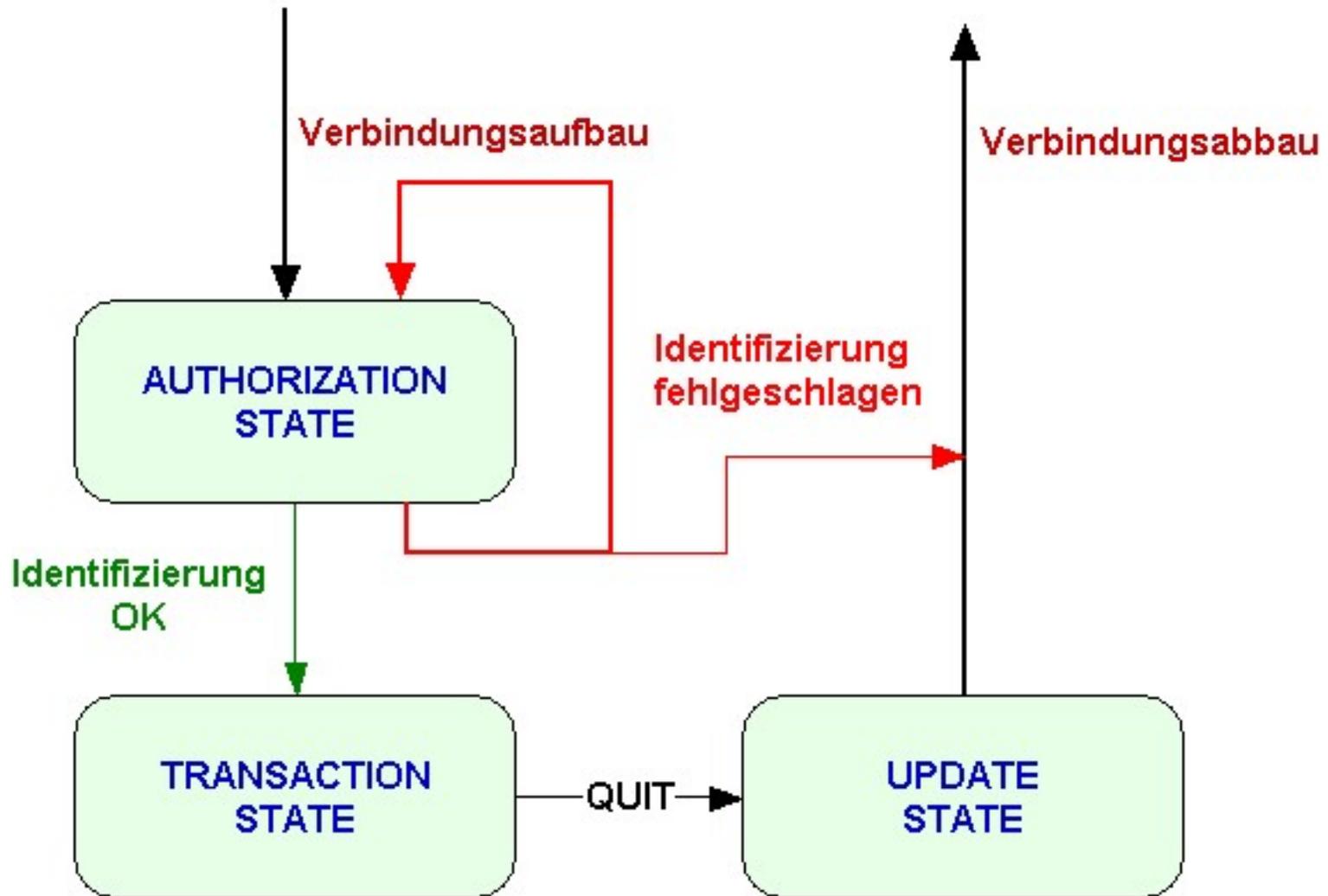
- Für Hosts die keine ständige Internetverbindung haben ist es nicht sinnvoll einen permanenten SMTP-Server zu betreiben
- Stattdessen speichert eine anderer MTA die Email (in sog. Mailboxes)
- Es wird ein Verfahren benötigt, um die gespeicherte Email abzurufen
- Diesem Zweck dient das *Post Office Protocol (POP)*
- Aktuelle Version ist POP3 (RFC1939)

- POP3 bietet folgende Funktionen
 - Authentifizierung
 - Empfang von Email Nachrichten
 - Löschen von Email Nachrichten auf dem Server

Funktionsweise

- POP3 wird beinahe ausschließlich über TCP Transportverbindungen verwendet
- Ein POP3 Server ist standardmäßig auf Port 110 zu finden
- Hinsichtlich der Interaktion von Client und Server ist POP3 ähnlich zu SMTP
- Auch POP verwendet das Kommando/Antwort Schema

- Eine POP3 Sitzung kann als Phasenmodell betrachtet werden, bestehend aus:
 1. AUTHORIZATION STATE
 - Client ist aufgefordert sich gegenüber dem Sever zu authentifizieren
 - Die zwei gängigsten Methoden sind:
 - * Die Kommandos **USER** und **PASS**
 - * Das Kommando **APOP**
 2. TRANSACTION STATE
 - Der Client wählt die Nachrichten aus, die er empfangen möchte und vermerkt jene Nachrichten die nicht mehr benötigt werden
 - Client beendet die Transaktionsphase mit dem Kommando **QUIT**
 3. UPDATE STATE
 - Alle markierten Nachrichten werden gelöscht
 - Die Kommunikationsverbindung wird abgebaut



Kommando	Beschreibung
LIST	Listet alle Nachrichten auf, die sich in der Mailbox befinden
RETR	Bewirkt, dass der Server die angegebene Nachricht an den Client sendet
DELE	Markiert die angegebene Nachricht zum Löschen
RSET	Setzt die Transaktionsphase zurück. Alle Löschmarkierungen werden verworfen
QUIT	Beendet die Transaktionsphase und beginnt die Aktualisierung

POP3 Beispiel Sitzung

Ein Beispiel einer typischen POP3 Sitzung...

S: +OK POP3 server ready <1896.697170952@dbc.mtview.ca.us>
C: APOP mrose c4c9334bac560ecc979e58001b3e22fb
S: +OK mrose's maildrop has 2 messages (320 octets)
C: STAT
S: +OK 2 320
C: LIST
S: +OK 2 messages (320 octets)
S: 1 120
S: 2 200
S: .
C: RETR 1
S: +OK 120 octets
S: <the POP3 server sends message 1>
S: .
C: DELE 1
S: +OK message 1 deleted
C: RETR 2
S: +OK 200 octets
S: <the POP3 server sends message 2>
S: .
C: DELE 2
S: +OK message 2 deleted
C: QUIT
S: +OK dewey POP3 server signing off (maildrop empty)

Das IMAP Protokoll

- **IMAP** ist die Abkürzung für **Internet Message Access Protocol**
- Es wurde 1986 an der **Stanford University** entwickelt und implementiert.
- Die heute gebräuchlichste Version ist IMAP4 (RFC 2060).
- IMAP ist ein Protokoll, das dem Client die Bearbeitung der Mails auf dem Server ermöglicht.
- Es wurde entwickelt, um Nachrichten nur nach Bedarf zu übermitteln: Der Anwender kann wählen, welche Daten er tatsächlich auf den eigenen Rechner übertragen will.

Einige weitere Merkmale des IMAP Protokolls sind:

- Die Email eines Benutzers können von diesem überall auf der Welt gelesen werden
- Benutzer können in ihren Mailboxen weitere Ordner anlegen und so eine strukturierte Ablage einrichten.
- Mehrere Personen können zum selben Zeitpunkt von unterschiedlichen PCs aus auf dasselbe Postfach zugreifen.
- Auf IMAP Server kann i. d. R. auch mittels POP3 zugegriffen werden. IMAP emuliert also POP3.
- Generell ist der Zugriff auf Email in einem IMAP-Ordner etwas langsamer als bei POP3 - da die Nachrichten ja nicht auf dem lokalen Rechner,

sondern auf einem Server gespeichert sind.

- Der Anwender kann den Nachrichtenstatus auf dem Server verändern, um so beispielsweise gelesene Mails wieder als ungelesen zu markieren und umgekehrt.
- Ein wesentliches Merkmal sind die direkten Suchoptionen auf dem Server: muss POP3 zunächst alle Nachrichten "blind" kopieren, erlaubt es IMAP vorab zu selektieren. Dies macht gerade bei der oftmals teuren und zumeist langsamen Datenabfrage über Mobilnetze mit PDAs oder Notebooks Sinn.

Vorteile zentraler Speicherung

- Email, die der Benutzer von einem Standort aus gelesen oder gelöscht hat, gilt auch auf allen anderen Rechnern, die er zur Bearbeitung seiner Email verwendet, als gelesen bzw. gelöscht.
- Wenn Sie eine Nachricht einmal gelesen oder in einem Folder abgelegt haben, finden Sie diese von jedem anderen Arbeitsplatz aus ebenso vor.
- Die Speicher von großen Servern sind meist zuverlässiger und sicherer als normale PC-Festplatten

- Ähnlich wie z.B. bei NetworkNews erhält der Client zuerst eine Nachrichtenübersicht, die rasch übertragen wird. Auf dem Rechner des Anwenders wird anschließend nur die Mail kopiert, die er aufgrund dieser Übersicht anfordert. Uninteressante Nachrichten oder solche mit umfangreichen Attachments können einfach überspringen werden.
- Die meisten IMAP fähigen MUA bieten die Möglichkeit, Attachments am Server zu belassen, aber dennoch die Nachrichten zu lesen.

Nachteile zentraler Speicherung

- Der Speicher-Bedarf auf dem Server ist bei IMAP viel größer als bei POP; daher ist der für den Benutzer verfügbare Speicherplatz meist limitiert.
- Wenn das zentrale Mailbox-Service ausfallen sollte oder nicht erreichbar ist, hat man während dieser Zeit auch auf alte Email Nachrichten keinen Zugriff.

IMAP Kommandos

Die wichtigsten IMAP Kommandos werden hier beschrieben.

Man unterscheidet grundsätzlich 3 Arten:

Housekeeping operations: AUTHENTICATE, LOGIN, LOGOUT,
CAPABILITY, NOOP

Mailbox operations: SELECT, EXAMINE, CHECK, CLOSE,
EXPUNGE, SEARCH, CREATE, DE-
LETE,
RENAME, LIST, LSUB, SUBSCRIBE,
UNSUBSCRIBE

Message operations: FETCH, PARTIAL, STORE, COPY,
APPEND

IMAP Beispiel Sitzung

Ein Beispiel einer typischen IMAP Sitzung...

```
S: * OK [Capability IMAP4 IMAP4REV1 LOGIN-REFERRALS AUTH= LOGIN
C: a001 LOGIN username password
S: a001OK LOGIN completed
C: a002 SELECT inbox
S: * 2 EXISTS
    * 2 RECENT
    ...
    a002 OK [READ-WRITE] SELECT completed
C: ...
C: a003 LOGOUT
S: a003 OK LOGOUT completed
```

Ressourcen

- Internet RFC Database (www.rfc-editor.org)
- Die Präsentations Unterlagen online:
http://www.cpn.2y.net/email_protocols/
- Ausführliche schriftliche Ausarbeitung ebenfalls unter:
http://www.cpn.2y.net/email_protocols/

— Ende —

Vielen Dank für Ihre Aufmerksamkeit !