Network Services, VU 2.0

Email (SMTP, POP3, IMAP)
News

Dipl.-Ing. Johann Oberleiter
Institute for Informationsystems, Distributed Systems Group

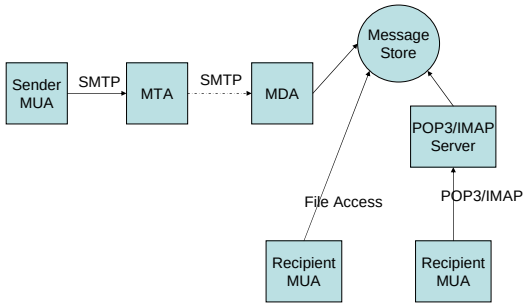
Agenda

- Mail general info
- SMTP (Simple Mail Transfer Protocol)
- POP3 (Post Office Protocol)
- IMAP4 (Internet Message Access Protocol)
- Mail vulnerabilities
- Usenet News / Network News Transfer Protocol

EMail Topics

- EMail Agents
  - Mail User Agent (MUA)
    - Client software used to compose, send, and retrieve email messages. Sends mail via MTAs.
    - Retrieves messages from mail store directly or POP3/IMAP
  - Mail Transfer Agent (MTA)
    - Server that receives and delivers mail.
    - Locally delivered mails handed off to MDA
  - Mail Delivery Agent (MDA)
    - Final delivery of messages for local recipients
    - Filtering or categorizing
    - Forwarding
- Postmaster
  - Mail Administrator

Email delivery



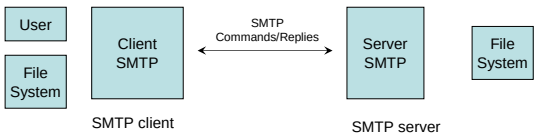
SMTP

- Simple Mail Transfer Protocol
- Started in 1982
- Mail consists
  - Envelope (RFC 2821 - originally 821)
    - Everything to accomplish transmission
  - Content (RFC 2822 – originally 822)

Main goals

- Transfer mail
- Reliably and Efficiently
- Requires a reliable ordered data stream channel (eg. TCP)
- Transfer between processes
  - Same network
  - Another network

SMTP Design

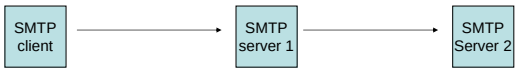


SMTP Client

- Main responsibilities
  - Transfer mail message to one or more SMTP servers
  - Report mail transfer failures

SMTP Server

- Final destination
- Intermediate Relay
  - Mail may pass through intermediate relay or gateway hosts
  - Server may assume role of SMTP client after receiving
- Gateway
  - Transport message with protocol other than SMTP
- Server responsibilities
  - Server accepts responsibility to deliver a message
  - Properly reports the failure to do so.
- Standard Port: TCP 25



Mail transmission

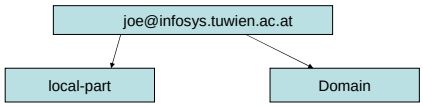
1. Establish Transmission channel
2. Initial Handshaking
3. Mail transaction
... Mail transaction
4. Shut down

Types of SMTP Systems

- Originator
  - Introduces Mail into a transport service environment such as the Internet
    - Eg . Mail Client
- Delivery
  - Receives Mail from transport service
  - Passes it to a mail user agent or stores it in message store (which is accessed by MUA)
- Relay
  - Receives mail from SMTP client and transmits it ,without modification to another SMTP server
- Gateway
  - Receives mail from a client in one transport environment and transmits it to a server system in another transport environment
  - Firewalls that rewrite addresses can be considered Gateway (even if SMTP is used on both sides)

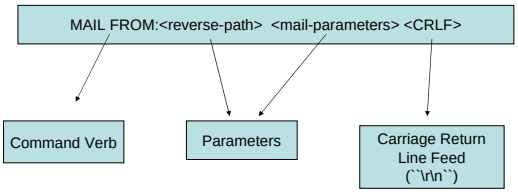
Mail Address

- Mail Address
  - Character string identifies user to whom mail will be sent
  - Location into which mail will be deposited



- Mailbox
  - Refers to the mail depository
- Address examples
  - muster.mann@infosys.tuwien.ac.at
- Alias
  - Names that refer to the same mailbox
    - m.mann@infosys.tuwien.ac.at
    - muster.mann@infosys.tuwien.ac.at

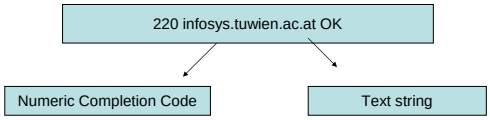
SMTP Commands



•Text commands (7-bit ASCII)
•Each command ends with CRLF
  •Will not be shown in the examples
  •in Telnet just press Enter

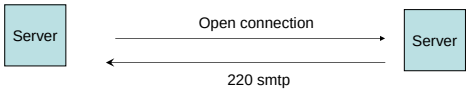
SMTP Reply

- Acknowledgment or Error Code
  - From receiver to sender
  - Response to a command
- About 20 numeric codes
  - 1xx : Positive Preliminary reply (only in extended SMTP)
  - 2xx : Positive Completion reply
  - 3xx : Positive Intermediate reply
  - 4xx : Transient Negative Completion reply
  - 5xx : Permanent Negative Completion reply



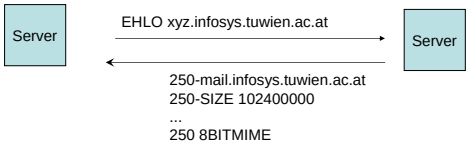
SMTP Procedures / 1

- Session Initiation
  - Client opens a connection to a server
  - Server responds with an opening message



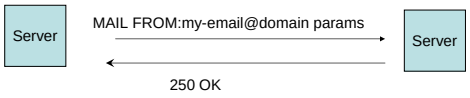
SMTP Procedures / 2

- Client Initiation
  - Client sends EHLO (=Extended Hello)
    - Old: HELO (Hello)
  - Client sends own domain as parameter
  - Responds with capabilities
  - Server resets variuos buffers



SMTP Procedures /3

- Start Mail Transaction
  - MAIL FROM: <reverse-path> <params>
  - <reverse-path> destination for errors
    - After server has accepted delivery
  - Example:
    - MAIL FROM:<xyz@myuser.xyz>

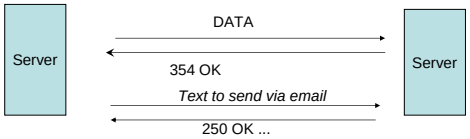


SMTP Procedures /4

- Set Mail Recipient
  - Multiple recipients allowed
  - RCPT TO:<forward-path>
  - <forward-path> destination (mailbox) for the e-mail
  - Example:
    - RCPT TO:<userx@blabla.com>

SMTP Procedures /5

- Send Data
  - DATA
    - Signals start of data transmission
    - Respone: 354 End data with <CRLF>.<CRLF>
  - Normal text can be sent
    - Internet message format
    - Ends with <CRLF>.<CRLF>



SMTP Example

220 smtp
EHLO <xyz@mydomain.com>
250-mail.mydomain.com
250 8BITMIME
MAIL FROM:<xyz@mydomain.com>
250 Ok
RCPT TO:<abc@otherdomain.com>
250 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
Testmail
Another Line
.
250 Ok: queued as 12345678AB
QUIT

Internet Message Format

- RFC 2822 (originally 821 / 822)
- Used for mail content
- Header
- Contents

Internet Message Format Header Fields / 1

- Originator Fields:
  - EMail Subject
    - Subject: Meine e-mail
  - From (authors)
    - From: addr1, addr2
  - Sender (mailbox of the sender)
    - Sender: my-mailbox@blabla.com
  - Reply-To
    - Reply-To: address-list

Internet Message Format Header Fields / 2

- Destination address fields
  - To (primary recipient)
    - To: you@blabla.com
  - Cc: (Carbon copy)
    - Cc: theboss@blabla.com
  - Bcc: (Blind carbon copy)
    - Send message to recpient without revealing those addresses to other recpients
    - SMTP server specific implementation strategies

Internet Message Format Header Fields / 3

- Identification Fields
  - Refers to previous messages
    - message-id
    - in-reply-to
    - references

Internet Message Format Header Fields / 4

- Informational Fields
  - Subject
    - Subject: Network Services
  - Comments
  - Keywords
- Resent Fields
  - For messages reintroduced by user
- ...

MIME / 1

- Multipurpose Internet Mail Extensions
  - RFC 2045-2049 + other RFCs
- Other media-types than text
  - Described with Content-Type and Subtype
    - Eg. image/jpeg, multipart/mixed
- Multiple formats in one file allowed
  - Attachments (Content-Type = `multipart`)
- Binary data encoded with 7bits
  - Base64 encoding
    - 64 Ascii characters used to represent binary data
    - `=` special processing
  - Other formats possible

MIME / 2

- Example

```
...
Subject: Sample Message
MIME-Version: 1.0
Content-type: multipart/mixed; boundary=``end``

Note to non-MIME conformant readers
--end

Mail part one
--end
Another part
--end--
```

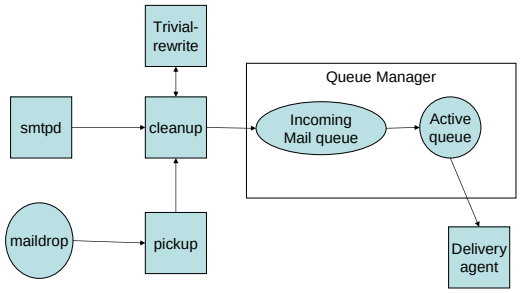
SMTP Server

- Sendmail
- Postfix
  - Used in Lab
  - Reliability
  - Security
  - Compatiblity with sendmail
    - Forwarding and aliasing
- Microsoft Exchange

Postfix / Architecture

- Several daemons
- master
  - Starts all others
- smtpd
  - Receives email from network
- pickup
  - Read mails from directory
- Cleanup
  - Sanity checks on message
- Trivial-rewrite
  - Completion of headers

Postfix / Architecture



Postfix / Queue Manager

- Core part of Postfix
- Administrates Queues
  - Goal to deliver email
- Queue Types
  - Incoming
    - First location of a message in Queue Manager
  - Active
    - Tries to forward to appropriat delivery agent
  - Deferred
    - Mails not delivered
    - Tried again
  - Corrupt
    - Contains damaged mails
  - Hold
    - User puts messages into hold queue
    - Won't be removed automatically

Postfix / Delivery Agents

- Local delivery agent
  - Users with shell acount on host
- Virtual delivery agent
  - For virtual mailbox addresses (no shell account)
- Smtp delivery agent
  - For relaying

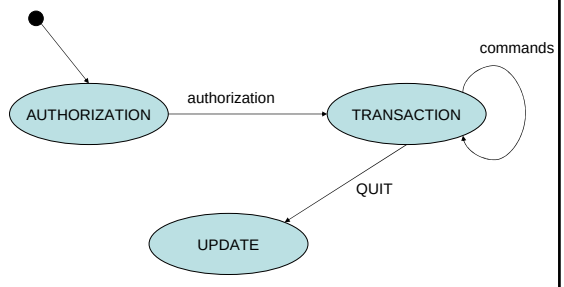
Message Storage Access

- SMTP not on all hosts feasable
- Other protocols to access maildrop
  - POP3 (RFC 1939)
  - IMAP4 (RFC 3501)
- Separate protocols
  - Primary idea
    - Use ISP SMTP server when offline
    - Use companies POP3/IMAP

POP3 / 1

- Primary mechanism
  - Download Mail
    - from server to client
  - Delete from Server
- POP3 server on TCP Port 110
- Commands similar to SMTP
  - Keyword & text-based
  - Multiline responses end with "."
  - All commands terminated with <CRLF>
  - Each message has a number
- POP3 session state-based
  - AUTHORIZATION state
    - Wait on authorization info
  - TRANSACTION state
  - UPDATE state
    - Removes mails from server maildrop

POP3 / State Machine



POP3 / 3

- USER & PASS (AUTHORIZATION)
  - Mailbox & Password
- APOP name digest (AUTHORIZATION)
  - Alternative to USER & PASS
  - Calculates shared secret based on server greeting (that must contain unique timestamp)
- STAT (TRANSACTION)
  - Status – information about maildrop
- LIST [msgNr] (TRANSACTION)
  - Scan listing for (all) messages
    - Message number & message size in octets (=bytes)
- RETR msgNr (TRANSACTION)
  - Retrieves the contents of a message
- DELE msgNr (TRANSACTION)
  - Marks messages as delete
- RSET (TRANSACTION)
- TOP msgNr n (TRANSACTION)
  - Retrieves header + first n lines of body of a message
  - Important for retrieving header
- QUIT (TRANSACTION)
  - POP3 server removes all messages marked as delete

POP3 / Telnet Trace

```
C:  <open connection>
S:          +OK POP3 mail.xyz.at server ready
C:          USER joe
S:          +OK User name accepted, password please
C:          PASS blabla
S:          +OK Mailbox open, 20 messages
C:          LIST 20
S:          +OK 20 2696
C:          TOP 20 1
S:          +OK Top of message follows
...
C:          RETR 20
S:          +OK 2696 octets
...
C:          DELE 20
S:          +OK message 20 deleted
C:          QUIT
S:          +OK Sayonara

C:  Connection to host lost
```

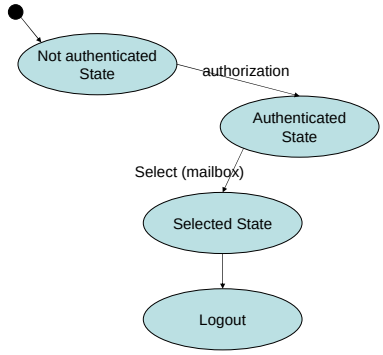
IMAP4 / 1

- IMAP4rev1
- More features than POP3
  - Operations for Mailboxes administration
  - Checking for new messages
  - Searching for messages
  - Message Flags
- IMAP4 server on TCP 143
- Keyword & text-based
  - All commands terminated with <CRLF>
  - Commands begin with unique identifier (tag)
  - Two different type of Responses
    - tagged response
    - Untagged response
  - Client may have to send continuation data
- Each message
  - unique identifier
    - MUST not change during a session
    - SHOULD not change between sessions
  - message sequence number
    - Relative position from 1 to number of messages in a mailbox
    - May be reassigned during a session

IMAP4 / 2

- Flags Message Attribute
  - 0-n named tokens associated with a message
  - Permant & Session-only flags
  - System flags = predefined
    - \Seen
    - \Answered
    - \Flagged (urgent/special attention)
    - \Deleted (marked as deleted)
    - \Draft (marked as draft)
    - \Recent (this IMAP session is first session notified about message)
  - Keywords
    - Not begin with "\"
    - Client may defined new keywords in the mailbox

IMAP4/3



IMAP4 / 4

- Server
  - may send data at any time
    - Even if client did not request this data
  - Server MUST send mailbox size updates automatically
  - Untagged response while no command in progress
  - Inactivity autologout time

IMAP4 / 5 – Client commands

- Any state
  - CAPABILITY
    - Requests listing of capabilities server supports
  - NOOP
    - No Operation
    - Prefered method to lookup new messages or status updates
  - LOGOUT
    - Server sends untagged BYE
    - Afterwards server sends tagged LOGOUT response

IMAP4 / 6 – Client commands

- Not authenticated
  - LOGIN
    - Plaintext password authentication (user name & password)
  - STARTTLS
    - Starts TLS/SSL negotation
    - On success all further commands under TLS layer
  - AUTHENTICATE
    - Indicates a SASL authentication mechanism to server
    - Server performs authentication protocol exchange to authenticate end identify client
    - May negotiate optional security layer for subsequent protocol interactions

IMAP4 / 7 – Client commands

- Authenticated State
  - SELECT mailbox
    - Selects a particular mailbox for subsequent requests
    - Only one mailbox can be selected in one connection
  - EXAMINE mailbox
    - Like SELECT, but read-only
  - APPEND mailbox messageData
    - Appends message to a mailbox
  - LIST refName mailboxName
    - Lists mailboxes relative to refName (eg. filePath)
  - Mailbox administration commands
    - CREATE,DELETE,RENAME
  - Subscription commands
    - SUBSCRIBE,UNSUBSCRIBE,LSUB

IMAP4 / 8 – Client commands

- Selected State
  - Based on currently selected mailbox
  - CLOSE & EXPUNGE
    - Removes all messages with \Deleted flag
    - Expunge sends untagged EXPUNGE response for each deleted message
  - SEARCH
    - Searches the mailbox for messages that match certain criteria (see RFC 3501 6.4.4)
  - FETCH
    - Retrieves data associated with a message (eg. Header, Body)
  - STORE
    - Alters data associated with a message

IMAP4 / 9 - Sample

```
C:  <opened connection>
S:  * OK [CAPABILITY IMAP4REV1 ...] mail.xyz.at
C:  A001 LOGIN joe mypasswd
S:  A001 OK [CAPABILITY IMAP4REV1 ...] User joe authenticated
C:  A002 SELECT mail/IEEE
S:  * 11 EXISTS
    * 0 RECENT
    * FLAGS (\Answered \Flagged \Deleted \Draft \Seen)
    * OK [UNSEEN 10] first unseen message in /home/joe/mail/IEEE
C:  A003 SEARCH ALL
S:  * SEARCH 1 2 3 4 5 6 7 8 9 10 11
    A003 OK SEARCH COMPLETED
C:  A004 FETCH 2:4 (BODY[HEADER])
S:  * 3 FETCH (BODY[HEADER]) {1085}
    ... mail messages ...
    A004 OK FETCH completed
C:  A005 LOGOUT
    * BYE mail.xyz.at IMAP4rev1 server terminating connection
    A006 OK LOGOUT completed
```

Message Disposition Notification

- RFC 3798
- Inform humans of the disposition of the message after successful delivery
- Additional message header field
  - "Disposition-Notification-To:"
- Sent as MIME message
- Problems:
  - Forgery (as regular emails)
  - Privacy
  - Non-Repudiation
  - Another way for Mail-bombing
- Better solution
  - Put message on Web server
  - Send secret URL via email
  - URL only accessible once

Procmail File

- Mail processing system
  - Sort incoming mail into folders
  - Preprocess email
  - Start programs upon arrival
- http://student.tuwien.ac.at/procmail/
- http://www.ii.com/internet/robots/procmail/qs/

Phishing

- Sending an email to a user claiming to be another sender
- Attempt to acquire private information from the user
  - Passwords
  - Pins
  - Credit Card Numbers
  - Bank Account Numbers
- Frequent attempt
  - HTML Links in HTML emails
    - www.amazon.com?
  - Link appears as www.amazon.com but links to 66.22.33.22
- Simple Solution
  - Don't use HTML emails

Spam

- Different meanings
  - Unsolicited Bulk Email
    - Massive number of recipients
    - Unsolicited!
    - Primarily Mass mails with commercial content (other Name: Unsolicited Commercial Email)
  - Fraud emails (Nigeria Connection)
    - Chain letter via email
  - Nonsense Postings in Internet forums (Trolling)

Spam - Principles

- Internet has a friendly nature
  - Email sent back to sender when receiver does not react/exist
  - Otherwise error message to postmaster
- Spam
  - Sends emails to huge number of potential recipients
  - Postmaster gets error message for non existent addresses
  - Removes these addresses from recipient list

Spam – Countermeasurements / 1

- Mask published email addresses
  - on Web pages
  - "email: joe at infosys dot infosys dot ac dot at"
    - Frequent pattern & rather weak (easily analyzable)
  - Better something like this:
    - "email: name@domain where name = joe and domain = infosys.tuwien.ac.at"
- Complain about spammer at the spammer's provider
  - Often same person
  - Provider in foreign country
  - Spammer is a client of the provider

Spam – Countermeasurements / 2

- Legal measurements
  - Accusing spammers
  - Possible for large companies
  - Only if spammer works in developed countries
  - Slow
  - First success stories
- Filtering based on Content and Format
  - In control of end-user
  - In control of end-user's provider
  - Today most successful
  - Does not fight Spam at the originator

Spam Filtering

- Scan on MTA
  - Good place for centralized checks
  - User specific settings cannot be used
- Scan on MDAs / Message store
  - Supports user specific configurations
  - Move Spam to particular mailbox
  - Spam verification done only after message received the system
  - Has to be installed & maintained on every system
- Problem – Different kind of users
  - Some don't want spam
  - Some want all emails
  - Legal problem of NOT delivering emails
    - Eg. German university

EMail Types

- HAM = Real-Negatives
  - Message is no SPAM
- SPAM = Real-Positives
  - Message is SPAM
- False-Positives
  - Message classified as SPAM but isn't
- False-Negatives
  - SPAM, not marked as SPAM
  - Goal of Spam Filtering is to minimize False-Negatives

Spam Assassin

- Rule-based Spam-Filter
  - Each rule applied on ingoing email
  - Each such test results in points
  - Each mail has total number of points
  - If points > predefined threshold -> Mark as SPAM
  - Reduce False-Negatives
    - Decrease Threshold; may lead to false positives
    - Better and More accurate rules
- Rule Types
  - Match Header or Message against text patterns for Spam
    - Detect Viagra (and variants)
  - Internet-Requests against Blacklists with IP-Sums or Checksums
  - Static Tests
  - Statistic & Self-Learning Tests

Spam Assassin

- Modifies email Header
  - X-Spam-Status YES/no
  - X-Spam-Checker-Version
  - X-Spam-Level
  - X-Spam-Flag

Spam Lists

- Lists contain sender
  - domain names
  - Email addresses
- Whitelists
  - Don't want email filtered
- Blacklists
  - Emails are Spam
    - Eg. DNSBL: emails sent or relayed from certain hosts are very likely Spam

Self-Learning Spam Filters

- Autowhitelisting
- Bayesian Filtering

Autowhitelisting

- Goal: Reduce false positives
- Reason behind:
  - People who send non-Spam messages won't start
- Reduction of test points for sender addresses on whitelist
- Example: each email sender has Total Spam Points stored in a database
- Correction of Points for current message
- Avg_Sender_Points = Total_Sender_Points / Count_Messages
- Points_New_Message += (Avg_Sender_Points – Point_New_Message) * CONV_FACTOR
- CONF_FACTOR $\in$ [0,1] (normally 0.5)

Bayesian Filtering

- Bayesian Theorem
  - The Probability of letting an event appear (Message A will be SPAM) when a certain test (Spam test result) is true is dependent on the probability of the event before the test result is known and the significance of the test.
- Practice
  - If some tests always evaluate to true when a message A is SPAM and never evaluate to true when a message is no SPAM then points will be adapted

Bayesian Filtering

- Requires training phase
  - Collection of messages that are definitively SPAM
  - Collection of messages that are definitively NO-SPAM
  - Finds token in messages based on these mesages
    - Words or word groups
  - Training based on mistakes
    - Classification only on false positives and false negatives
- Hints
  - Message must be from the same time frame
  - Better results with continous training

USNET News / Network News Transfer Protocol (NNTP)

- News articles stored in central database
  - Subscribers may select only interesting items
- Features
  - Indexing
  - Cross-referencing
  - Expiration of aged messages
- RFC 977 (1986)
- News articles
  - Primarily based on news article specification (RFC 850)
  - Need only be stored on one host
  - Subscribers on other hosts may read news articles using stream connections
  - Intermediate servers may be used
    - Mediating news reading requests
    - Caching of recently-retrieved news

Usenet News

- Advantages compared to mailing lists
  - More efficient than mailing lists when many people are recipients
    - Separate copies of the same mail to destination host
  - Maintenance of the list
    - Subscribers change jobs
    - Subscribers join/leave
    - Hosts coming in/out of service

NNTP

- Protocol for news articles
  - Distribution
  - Inquiry
  - Retrieval
  - Posting
- Reliable Stream (TCP)
  - Standard Port: TCP 119
- Similar to SMTP
- NNTP may support other formats than RFC 850

NNTP Commands

- ARTICLE message-id / nnn
  - Displays the header + body of an article
  - Other forms: HEAD or BODY or STAT
- GROUP newsgroupname
  - Select the newsgroup "newsgroupname"
- LIST
  - Returns a list of valid newsgroups
  - Other form NEWGROUPS lists new groups since a particular date & time
- NEWNEWS
  - List of message-ids of articles posted since date&time
- NEXT / LAST
  - Advances an internally maintained article pointer to the next article
- IHAVE message-id
  - Informs server that client has an article with given message-id
  - Server may request this article
- POST
  - Send a new article to the newsgroup
  - Server replies with 340 to notify clients to send the article

NNTP Sample

S: 200 xyz.at NNRP Service Ready (posting ok).
C: LIST
S: list1.public.abc 0001 0001 y
list2.public.abc 0001 0001 n
.
C: GROUP list1.public.abc
S: 211 1 1 list1.public.abc
C: HEAD
S: ... Delivers head of article
C: ARTICLE
S: ... Delivers complete article