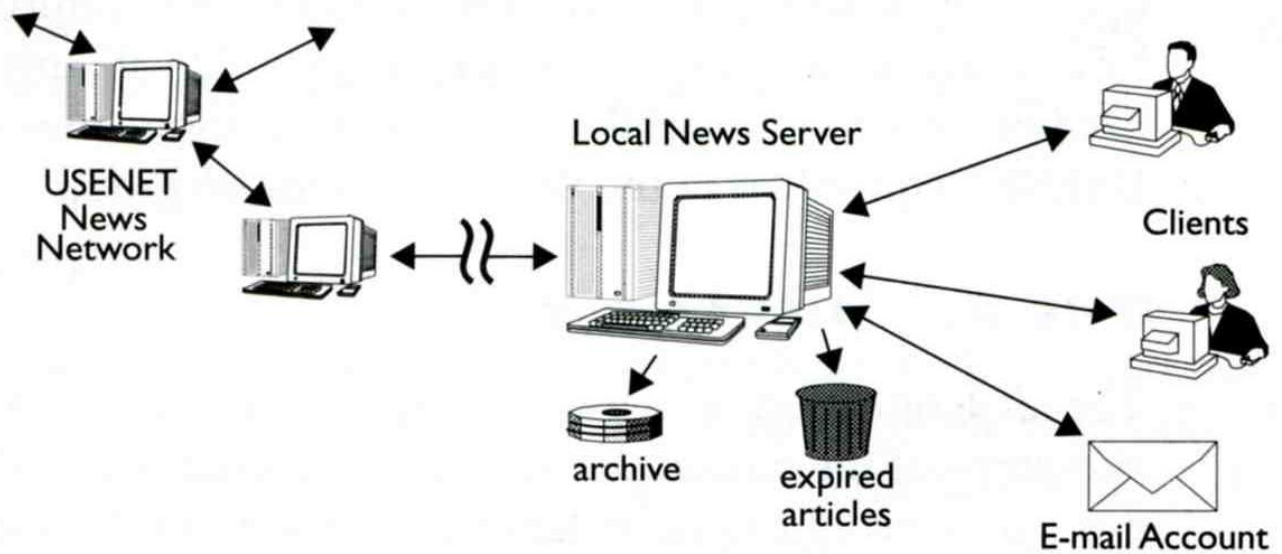


News

Usenet · NetNews



News outline

❑ Introduction

- History of netnews
- History of newsgroups
- News operations

❑ Transfer protocol

- NNTP – Network News Transfer Protocol

❑ News softwares

- Reader: tin
- Server: INN – InterNet News



Introduction to News

Introduction to news (1)

- ❑ Usenet is a worldwide distributed Internet discussion system.
 - The initial Internet community
 - Like bulletin board systems (BBS)
 - Hybrid between e-mail and web forums
 - Absence of a central server and dedicated administrator

- ❑ The format and transmission of Usenet articles is similar to that of Internet e-mail messages.

- ❑ Usenet has significant cultural importance in the networked world
 - FAQ, spam.

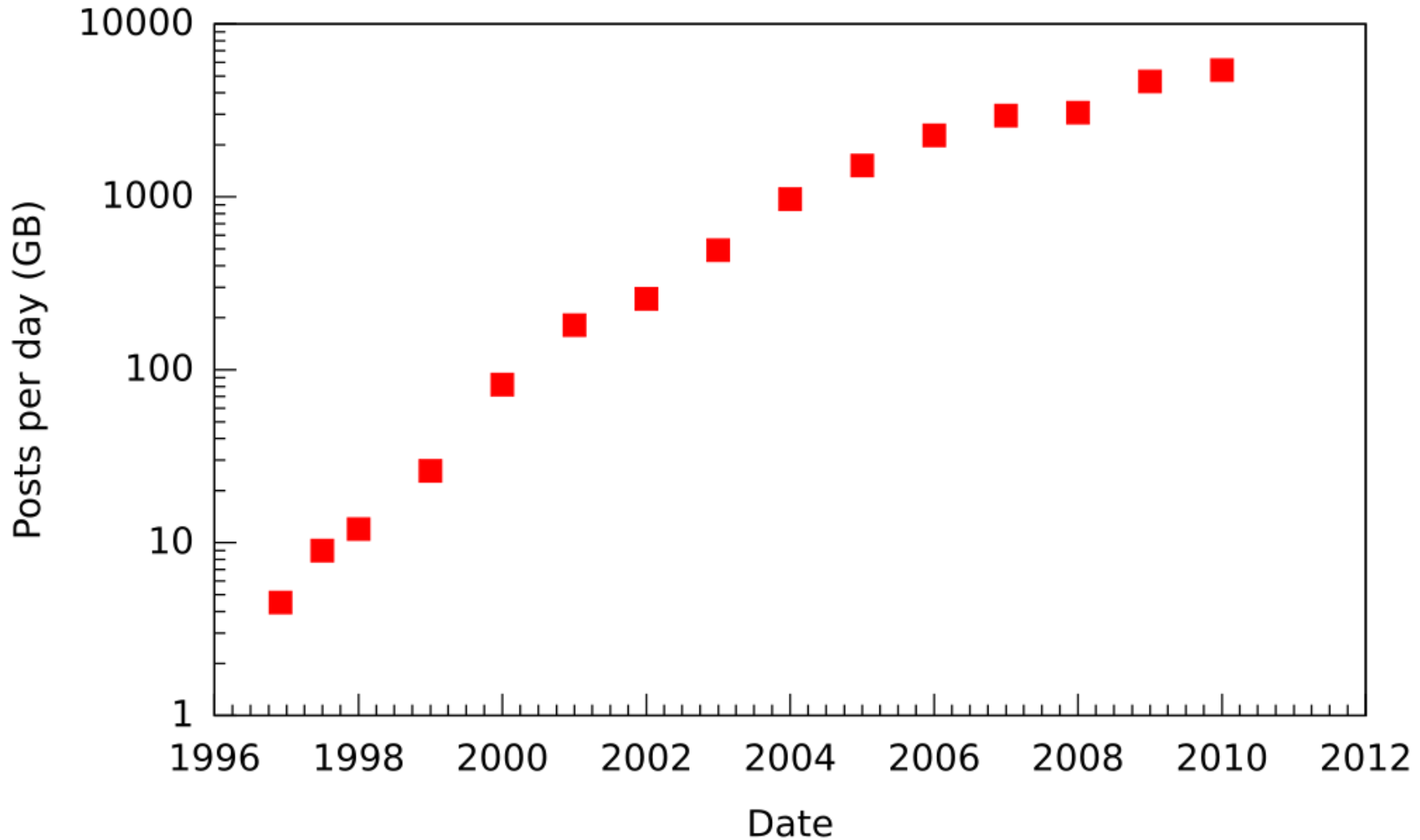
Introduction to news (2)

- Spamming: In 1994, advertisers began broadcasting messages to all newsgroups at once.
- Usenet has diminished in importance with respect to Internet forums, blogs, and mailing lists.
- Google made efforts to buy archive of Usenet newsgroups to attempt to create a complete one. Google has a web search interface and also allows newsgroup posting.
 - <http://groups.google.com>
- Usenet is a “must-have” resource?

Year	1996	1997	1998	1999	2000	2001	2002	2003
Daily Vol. (GB)	4.5	9	12	26	82	181	257	492
Year	2004	2005	2006	2007	2008	2009	2010	2011
Daily Vol. (TB)	1	1.52	2.27	2.95	3.07	4.65	5.42	7.52

Introduction to news (3)

Usenet Traffic Per Day (Source: altopia.com)



History (1)

❑ Origins

- Tom Truscott and Jim Ellis conceived the idea in 1979
- Steve Bellovin assisted with writing Bourne shell scripts
- The first “netnews” program
- UUCP as its transport protocol

❑ A News

- The public release of “netnews” in 1980
- Rewritten in C by Tom Truscott, Steve Daniel

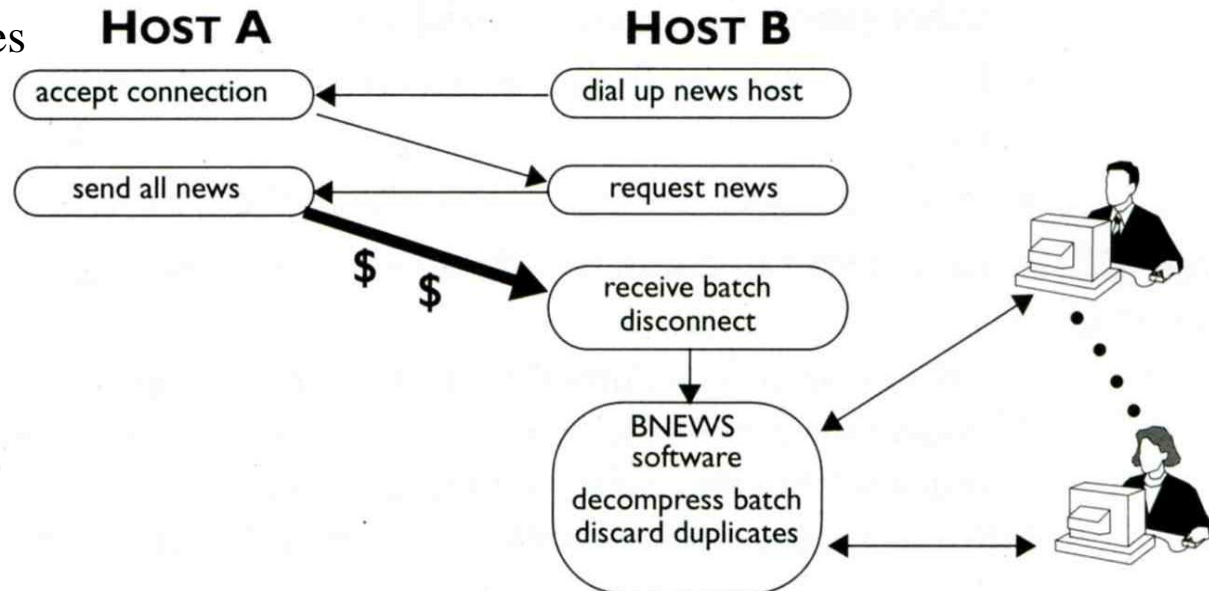
❑ B News

- Written by Mark Horton and Matt Glickman in 1981
- Development was passed on to Rick Adams in 1983 ~ 1989
- With a message format compatible with Internet mail
- Better performance for the grown news traffic

History (2)

□ UUCP

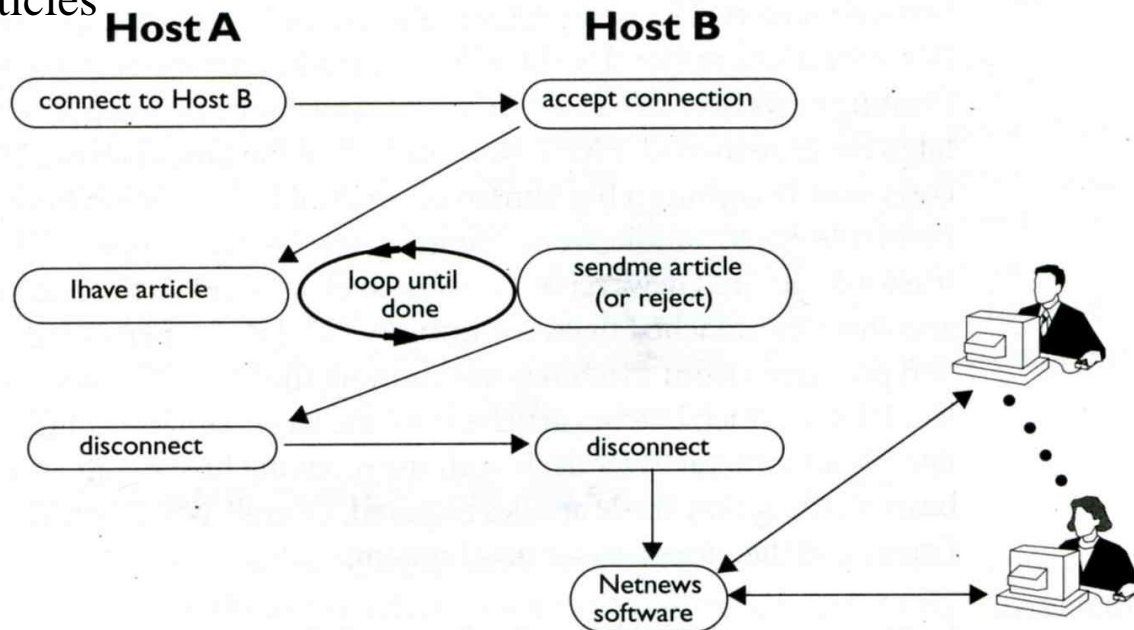
- Unix-to-Unix Copy Protocol
- Mostly commonly used over dial-up connections
- Replaced by the TCP/IP protocols SMTP for mail and NNTP for news
- Bang path are still use within the news network to record the nodes through which a message has passed
- Store-and-forward batch
- Duplicate articles



History (3)

□ NNTP

- Network News Transfer Protocol
- Introduced in 1985 to distribute netnews articles over TCP/IP network
- Specification is created by Phil Lapsley and Brian Kantor in 1986
- TCP port 119 for NNTP and TCP port 563 for NNTPS
- NNRP(Network News Reader Protocol)
- Sending required articles



History (4)

❑ The CNEWS version

- Developed by Henry Spencer and Geoff Collyer in 1987
- Rewrite the code of B News to result in a cleaner, faster program
- The creation of a news index function in 1992
- NNTP XOVER command
- Allow news readers to rapidly retrieve header and threading information

❑ INN (InterNet News)

- Released by Rich Salz in 1991
- The first news server package to fully integrate NNTP functionality
- NNTP and UUCP support
- The most commonly used news server package now
- Still actively developed by ISC(Internet Systems Consortium)

newsgroups (1)

- ❑ The articles are organized into topical categories called newsgroups.
- ❑ Newsgroups are organized into hierarchies, theoretically making it simpler to find related groups.

- ❑ A usenet newsgroup is a repository, a discussion group.
- ❑ Newsgroups
 - Types: text or binary.
 - Like the public message boards on BBS.
 - Over 100000 Usenet newsgroups



newsgroups (2)

❑ Great Renaming

- July 1986 ~ March 1987
- Before this
 - fr.* for groups gatewayed from ARPANET
 - mod.* for moderated discussions
 - net.* for unmoderated groups
- Primary reason: the difficulty of maintaining a list of all the existing groups.
- Alternative explanation:
 - European networks refuse to pay for some of the high volume and low content groups such as religion and racism.
 - The suggested category was talk.*

newsgroups (3)

❑ Great Renaming

- Initiated by Rick Adams
- Gene Spafford organized the “Backbone Cabal”
- Big 7 in 1987
 - comp.* for computer related discussions
 - news.* for discussions and announcements about netnews
 - rec.* for discussions of recreational activities
 - sci.* for science related discussions
 - soc.* for socializing and discussions of social issues
 - talk.* for discussions about various controversial topics
 - misc.* for miscellaneous discussions
- Big 8 in 1995
 - humanities.* for discussions of the humanities

newsgroups (4)

- ❑ After the Great Renaming, the creation and organization of new newsgroups became an automated voting process.
 - Post an RFD (Request for Discussion) into news.announce.newgroups
 - Discuss in news.groups.proposals
 - The proposal is formalized with a name, description, charter.
 - Big-8 management board will vote on whether to create the group.

- ❑ There was a great controversy over what newsgroups should be allowed.
 - Drugs, and sex were not allowed.

- ❑ Brain Reid suggested the alt.* hierarchy as an alternative to talk.*
 - Dissatisfied with mod.gourmand → rec.food.recipies
 - Created in May 1987.

newsgroups (5)

❑ alt.* hierarchy

- It is ‘alternative’ to the mainstream Big-7.
- The first alt.* newsgroups were alt.config, alt.gourmand, and alt.drugs
- Most newsgroups follow an informal procedure involving a public discussion in alt.config before created.
- Not confined to newsgroups of any specific subject or type.
- Because of the anarchistic nature with which the groups sprung up, some jokingly referred to ALT standing for “Anarchists, Lunatics and Terrorists”

❑ Two major sections of alt.*

- alt.sex.*: fearing the major news may be prevented from widely distributed.
- alt.binaries.*: often of extremely large size.

newsgroups (6)

- ❑ The language of preference in the “original” Usenet hierarchies, including alt.*, is English.
 - Other language hierarchies have later been created in parallel to the existing English ones.
 - de.*, fr.*, ...
 - tw.*, tw.bbs.* (innbbsd)

- ❑ Some access providers also create their own versions, prefixing the newsgroups names with their own name in a similar way.
 - Messages posted in there “private” newsgroups are generally not passed to the other providers or the Internet in general.

- ❑ Some users prefer to use “Usenet” to refer only to the Big-8 hierarchies, and to use “netnews” to incorporate the entire medium.

newsgroups (7)

❑ Binary newsgroups

- Binaries are posted in alt.binaries.*
- Usenet was originally created to distribute text content encoded in ASCII.
- Encoded binaries by uuencode, Base64, Quoted-Printable MIME, or yEnc.
- The binaries might be split into pieces by means of RAR and PAR.

❑ Moderated newsgroups

- A moderated newsgroup has one or more individuals as moderators.
- Messages submitted by readers are not distributed to Usenet, but are emailed to the moderators for approval, instead.
- Moderators receive submitted articles, review them, and inject approved articles. Such articles must bear the “Approved:” header line.

News operations (1)

❑ Articles and posts

- Binary content: split among multiple articles, reassembled into a single unit by the news reader
- Plain text: majority of the articles are responses to some other article.
 - Articles which can be traced to one single non-reply article is called a thread.

❑ Headers and overviews

- The overview is a list of the most frequently used headers, and additional information such as article sizes.

❑ newsfeeds

- When a user posts an article, it is initially only available on that user's news server. Each news server talks to one or more other servers(peers) and exchanges articles with them.

News operations (2)

❑ News Reader

- Read and post articles
- In early news implementations: news server and news reader were a single program suite.
- Mail clients or “communication suites” commonly have an integrated news reader.
 - However, these integrated clients are of low quality: incorrectly implementing Usenet protocols, standards and conventions.
- A user subscribes to a newsgroup, the news reader software keeps track of which articles that user has read.
- With the rise of the World Wide Web, web front-ends have become more common.
- Common text-based news reader: tin (chinese/tin 、 news/tin)

News operations (3)

□ News Server

- Continually increasing storage
- Network capacity requirements

- Completion: the ability of a server to successfully receive all traffic
- Retention: the amount of time articles are made available to readers
- Overall system performance

- Split into
 - Transit server: handle the basic articles exchange.
 - Numbering server: insert the “Xref:” header line into each article
 - Back end server: data storage
 - Front end server: a user would actually connect to. Only store overviews, and carry the heaviest CPU load

News operations (4)

□ Spools

- A disk storage for the body of an article
- Types:
 - One file per article
 - Oldest storage scheme
 - Performance relates to OS's ability to create, remove and locate files within a dictionary
 - Cyclical storage
 - Since 1990s
 - Articles are appended serially to large indexed container files
 - When the end of the file is reached, new articles are written at the beginning of the file, overwriting the oldest entries.
 - Traditional text manipulation tools are less well-suited to analyze the files
- Flexibility to retention and space used.

News operations (5)

❑ Storage management

- Article sizes
 - The maximum article size varies from site to site
- Retention
 - How long the server keeps articles
 - Text and binary

❑ Speed

- How quickly a server can deliver an article to the user

News operations (6)

❑ Control messages (RFC5537)

- newgroup messages
 - Create a new newsgroup
- rmggroup messages
 - Remove a newsgroup
- checkgroup messages
 - Lists all groups of a hierarchy
- cancel messages
 - Delete a article specified by Message-ID
 - In cases where unsuitable content has been posted, Usenet has support for automated removal of a posting from the whole network by creating a cancel message.

News operations (7)

❑ Spam filter

- Cancel message
 - 40% spams 、 40% cancel messages 、 20% useful information
 - There is no authentication in cancel messages
- Cleanfeed
 - Written by perl
 - Count by md5(body) 、 Ph/l(Posting-host/lines) 、 F/s/l(From/subject/lines)
 - Specific strings such as binary post 、 (mime) html post 、 MMF
- NoCeM
 - No see 'em
 - Cancel-Key = hash(password, Message-ID) 、 Cancel-Challenge
 - All notices must be public key signed to be accepted using PGP
 - innbbsd patch by leeym

Take a break – nonsense

- ❑ NCTU_Talk
- ❑ tw.bbs.comp.linux ↔ ptt
- ❑ tw.bbs.campus.nctu ↔ nctu.talk
- ❑ Group.NCTU.edu.tw



NNTP

RFC 3977

Interact with INN news server

NNTP (1)

□ Command and response

- Response codes
 - 1xx - Informative message
 - 2xx - Command completed OK
 - 3xx - Command OK so far; send the rest of it
 - 4xx - Command was syntactically correct but failed for some reason
 - 5xx - Command unknown, unsupported, unavailable, or syntax error

 - x0x - Connection, setup, and miscellaneous messages
 - x1x - Newsgroup selection
 - x2x - Article selection
 - x3x - Distribution functions
 - x4x - Posting
 - x8x - Reserved for authentication and privacy extensions
 - x9x - Reserved for private use (non-standard extensions)

NNTP (2)

□ NNTP vs. NNRP

- For your peers

```
liuyh@csnews$ telnet csnews2 433
Trying 140.113.235.121...
Connected to csnews2.cs.nctu.edu.tw.
Escape character is '^]'.
200 csnews2 InterNetNews server INN 2.4.6 ready
```

```
liuyh@nsd5$ telnet csnews2 433
Trying 140.113.235.121...
Connected to csnews2.
Escape character is '^]'.
502 Permission denied
```

- For your readers

```
liuyh@bsd5$ telnet csnews2 119
Trying 140.113.235.121...
Connected to csnews2.
Escape character is '^]'.
200 csnews2 InterNetNews NNRP server INN 2.4.6 ready (posting ok).
```

NNTP (3)

❑ LIST (find the newsgroups)

- AVTIVE

list active cs.*

```
215 Newsgroups in form "group high low flags".  
cs.computer-center 0000000509 0000000071 m  
cs.course.netadmin 0000000505 0000000131 y  
cs.course.object-oriented-programming.a 0000000284 0000000108 y
```

- NEWSGROUPS

list newsgroups cs.*

```
215 Descriptions in form "group description".  
cs.computer-center [系計中公告] (Moderated)  
cs.course.netadmin [課程] 計算機網路管理.  
cs.course.object-oriented-programming.a [課程] 物件導向程式設計 A 班
```

NNTP (4)

- ❑ LIST (overview format)
 - OVERVIEW.FMT

list overview.fmt

215 Order of fields in overview database.

Subject:

From:

Date:

Message-ID:

References:

Bytes:

Lines:

Xref:full

.

NNTP (5)

- ❑ GROUP (select a newsgroup)

```
group cs.course.netadmin  
211 359 131 505 cs.course.netadmin
```

- ❑ LISTGROUP (select a newsgroup)

```
listgroup cs.course.netadmin  
211 359 131 505 cs.course.netadmin  
131  
132  
134  
138  
141  
...  
504  
505  
.
```

NNTP (6)

- ❑ STAT [MessageID | Number]

```
stat
223 131 <f1sej8$2dev$1@csnews2.cs.nctu.edu.tw> status
stat 123
423 Bad article number
```

- ❑ NEXT
- ❑ LAST

```
stat 131
223 131 <f1sej8$2dev$1@csnews2.cs.nctu.edu.tw> status
stat 500
223 500 <CS-NetAdmin$A15S3QN9@bbs.cs.nctu.edu.tw> status
next
223 501 <CS-NetAdmin$A15S8UDD@bbs.cs.nctu.edu.tw> Article retrieved; \
request text separately.
last
223 500 <CS-NetAdmin$A15S3QN9@bbs.cs.nctu.edu.tw> Article retrieved; \
request text separately.
```


NNTP (7)

Subject:
From:
Date:
Message-ID:
References:
Bytes:
Lines:
Xref:full

- ❑ XOVER [range] (retrieve the overview)

xover 131-135

224 131-135 fields follow

```
131 [HW3] Sendmail sjwu <sjwu@cs.NCTU.edu.tw> Wed, 9 May 2007 \
    12:25:12 +0000 (UTC) <f1sej8$2dev$1@csnews2.cs.nctu.edu.tw> \
    70 Xref: csnews.cs.nctu.edu.tw cs.course.netadmin:131
```

132 ...

134 ...

.

- ❑ XHDR header [range | MessageID] (retrieve the specific header)

xhdr subject 131-135

221 subject matches follow (NOV)

131 [HW3] Sendmail

132 Re: [HW3] Sendmail

134 ...

.

xhdr newsgroups 131-135

221 newsgroups matches follow (art)

131 cs.course.netadmin

132 cs.course.netadmin

134 cs.course.netadmin

.

NNTP (8)

❑ ARTICLE [MessageID | Number]

article 499

220 499 <hpkcf\$283t\$1@csnews2.cs.nctu.edu.tw> article
Path: csnews2!csnews.cs.nctu.edu.tw!news.cs.nctu.edu.tw!not-for-mail
From: Yung-Hsiang Liu <liuyh@cs.NCTU.edu.tw>
Newsgroups: cs.course.netadmin
Subject: Re: 上課講解手release
Date: Thu, 8 Apr 2010 10:49:51 +0000 (UTC)
Organization: College of Computer Science, National Chiao Tung University
Lines: 1
Sender: liuyh@NCTUCSCC
Message-ID: <hpkcf\$283t\$1@csnews2.cs.nctu.edu.tw>
References: <CS-NetAdmin\$A15R07EB@bbs.cs.nctu.edu.tw> \
<hphpk6\$1sli\$1@csnews2.cs.nctu.edu.tw>
NNTP-Posting-Host: bsd5.cs.nctu.edu.tw
X-Trace: csnews2.cs.nctu.edu.tw 1270723791 73853 140.113.235.135 ...
X-Complaints-To: usenet@csnews2.cs.nctu.edu.tw
NNTP-Posting-Date: Thu, 8 Apr 2010 10:49:51 +0000 (UTC)
User-Agent: tin/1.8.3-20070201 ("Scotasay") ...
Xref: csnews.cs.nctu.edu.tw cs.course.netadmin:499

上課講解手檔案已放到網頁上。

NNTP (9)

- ❑ HEAD [MessageID | Number]
- ❑ BODY [MessageID | Number]

head 499

```
221 499 <hpkccf$283t$1@csnews2.cs.nctu.edu.tw> head
Path: csnews2!csnews.cs.nctu.edu.tw!news.cs.nctu.edu.tw!not-for-mail
From: Yung-Hsiang Liu <liuyh@cs.NCTU.edu.tw>
Newsgroups: cs.course.netadmin
Subject: Re: 上課講解手release
Date: Thu, 8 Apr 2010 10:49:51 +0000 (UTC)
Organization: College of Computer Science, National Chiao Tung University
Lines: 1
Sender: liuyh@NCTUCSCC
Message-ID: <hpkccf$283t$1@csnews2.cs.nctu.edu.tw>
```

...

```
Xref: csnews.cs.nctu.edu.tw cs.course.netadmin:499
```

.

body 499

```
222 499 <hpkccf$283t$1@csnews2.cs.nctu.edu.tw> body
上課講解手檔案已放到網頁上。
```

.

NNTP (10)

- ❑ POST (posting an article used by news clients)

```
post
340 Ok, recommended ID <hs8pdf$162r$1@csnews2.cs.nctu.edu.tw>
From: liuyh@cs.nctu.edu.tw
Newsgroups: cs.test
Subject: test telnet NNTP post

ha~ test
.
240 Article posted <hs8pdf$162r$1@csnews2.cs.nctu.edu.tw>
```

- 2xx - Command completed OK
 - 3xx - Command OK so far; send the rest of it
 - x4x - Posting
-
- RFC 5536

NNTP (11)

- ❑ IHAVE (exchange articles with peers)

```
ihave <hs8pdf$162r$1@csnews2.cs.nctu.edu.tw>  
435 Duplicate  
ihave <fjaowiejfmavoewifjaefarwvoijesfr@bsd5.cs.nctu.edu.tw>  
335  
Path: bsd5.cs.nctu.edu.tw!liuyh  
From: liuyh@cs.nctu.edu.tw  
Newsgroups: cs.test  
Subject: post from ihave  
Date: 10 May 2010 19:28:01 GMT  
Message-ID: <fjaowiejfmavoewifjaefarwvoijesfr@bsd5.cs.nctu.edu.tw>  
  
test  
.  
235
```

➤ x3x - Distribution functions

- RFC 5536

NNTP (12)

- ❑ NEWGROUPS [YY]YYMMDD hhmss

```
newgroups 100401 000000  
231 New newsgroups follow.  
alt.binaries.movies.hdtv.russian 0 1 y  
cs.u99 8 1 y  
sci.physics.acoustics 47 1 y  
alt.tv.channels.syfy 9 1 y  
alt.comp.lang.lua 2 2 y  
alt.fr.agence.notation.alteferienne 25 1 y  
.
```

- ❑ NEWNEWS newsgroups [YY]YYMMDD hhmss

```
newnews cs.course.netadmin 100222 000000  
230 New news follows  
<CS-NetAdmin$A15ONVSO@bbs.cs.nctu.edu.tw>  
<CS-NetAdmin$A15OVOGR@bbs.cs.nctu.edu.tw>  
...  
<CS-NetAdmin$A15UCP60@bbs.cs.nctu.edu.tw>  
<CS-NetAdmin$A15UDBNR@bbs.cs.nctu.edu.tw>  
.
```



tin

A Usenet news reader

Install (1)

- ❑ Two versions in ports
 - news/tin (default: US-ASCII)
 - chinese/tin (default: Big5)
- ❑ In RFC 3977
 - This specification extends NNTP from US-ASCII [ANSI1986] to UTF-8 [RFC3629]
 - Implementations **MUST NOT** use any other encoding.
 - The character set of article bodies **SHOULD** be indicated in the article headers
 - The newsgroup description **SHOULD** be in US-ASCII or UTF-8
 - 8-bit encodings other than UTF-8 **SHOULD NOT** be used
 - But...
 - IMHO, what is the hope?

Install (2)

- ❑ Back to install tin
 - 'make install' is ok
 - TIN_EDITOR
- ❑ Environment Variables
 - VISUAL
 - EDITOR
 - TINRC='-r'
 - NNTPSERVER=csnews2.cs.nctu.edu.tw

Start

- ❑ First execution
 - Generate ~/.tin/ and ~/.newsrc

- ❑ Connect

```
tin 1.8.3 release 20070201 ("Scotasay") [UNIX] (c) Copyright 1991-2006 Iain Lea.  
csnews.cs.nctu.edu.tw InterNetNews NMRP server INN 2.4.6 ready (posting ok).  
Reading keymap file...  
Reading input history file...  
Reading groups from active file.../  
Reading groups from newsrc file...-  
Reading global attributes file...  
Reading attributes file...  
Reading newsgroups file.../  
Reading newsrc file...
```

- ❑ Edit ~/.tin/tinrc

```
# Charset used for MIME (Content-Type) header in postings.  
mm_network_charset=Big5
```

- ❑ Second execution
 - If you use UTF-8 environment
 - Screen: encoding big5 utf8
 - env LC_ALL=zh_TW.Big5 tin (This is for post)

Search and subscribe newsgroups

- ❑ Show some default subscribed newsgroups

```

Group Selection (news 6)                                You have mail
M 1 3 news.announce.newusers Explanatory postings for new users. (M
M 2 14 news.newusers.questions Q & A for new users of Usenet. (Modera
3 166 misc.test For testing of network software. Very
M 4 misc.test.moderated Testing of posting to moderated groups
M 5 72 news.announce.newgroups Calls for newgroups & announcements of
M 6 542 news.answers Repository for periodic USENET article

```

- ❑ y (yank): list all newsgroups
- ❑ /?: search forwards/backwards for a group name
- ❑ s: subscribe to chosen group
- ❑ u: unsubscribe from chosen group

```

Group Selection (news 2)                                You have mail
1 484 cs.course.netadmin [課程] 計算機網路管理.
2 1442 cs.course.sysadmin [課程] 計算機系統管理.

```

Read and post

- ❑ Just like what you do in BBS
 - h: help
- ❑ Browse in a newsgroup
 - /?: search for articles by Subject line forwards/backwards
 - aA: search for articles by author forwards/backwards
 - r: toggle display of all/unread articles
 - <TAB>: display next unread article
 - w: post an article to current group
- ❑ Read an article
 - /?: search forwards/backwards within this article
 - f: post followup to current article
 - D: cancel (delete) or supersede (overwrite) current article
 - r: reply through mail to author
 - ^H: show article in raw-mode (including all headers)

newsrc

- ❑ After you quit tin or save configuration
 - ~/.newsrc will keep
 - newsgroups you subscribed/unsubscribed
 - articles you have read

```
liuyh@NASA$ cat ~/.newsrc
news.announce.newusers! 1-4466
news.newusers.questions! 1-675699
misc.test! 1-546806,546808,546823,546885,546908
misc.test.moderated! 1-646
news.announce.newgroups! 1-4441
news.answers! 1-230198
cs.course.netadmin: 1-505
cs.course.sysadmin: 1-644,648,652,656-658,661-664,668-669,687-
688,694,698,703,707,799,831,868,903,913-914,923,930-
932,968,983,1023,1054,1115,1131,1142-1145,1164,1186-
1187,1190,1192,1252,1323,1337,1433,1444,1451,1475,1489,1512,152
0,1523,1536,1559,1572,1614,1619,1647,1687,1757-1758,1766-
1768,1772,1781,1795,1818-
1819,1906,1918,1988,1997,2012,2028,2057,2059,2142
```

A crosspost example

❑ Before posting

```
From: 資工系 <help@cs.nctu.edu.tw>  
Newsgroups: cs.u95,cs.u96,cs.u97,cs.u98,cs.g97,cs.g98,cs.computer-center,cs.cour  
se.netadmin,cs.course.sysadmin,cs.phd,cs.talk,nctu.talk  
Subject: 徵資工系系務自動化系統開發 (大學部+研究生) 助教  
Followup-To: cs.talk
```

徵資工系系務自動化系統開發 (大學部+研究生) 助教

❑ After posting

```
Path: csnews.cs.nctu.edu.tw!news.cs.nctu.edu.tw!not-for-mail  
From: 資工系 <help@cs.nctu.edu.tw>  
Newsgroups: cs.u95,cs.u96,cs.u97,cs.u98,cs.g97,cs.g98,cs.computer-  
center,cs.course.netadmin,cs.course.sysadmin,cs.phd,cs.talk,nctu.talk  
Subject: 徵資工系系務自動化系統開發 (大學部+研究生) 助教  
Followup-To: cs.talk  
Xref: csnews.cs.nctu.edu.tw cs.u95:2369 cs.u96:1625 cs.u97:3571  
cs.u98:1490 cs.g97:317 cs.g98:85 cs.computer-center:486  
cs.course.netadmin:491 cs.course.sysadmin:2142 cs.phd:816  
cs.talk:2518 nctu.talk:1472787
```