

Introduzione a Usenet



Software per Usenet	696
Organizzazione generale del sistema di news di Usenet	697
Selezione dei gruppi e feed	698
Gestione degli articoli	700
Protocollo NNTP	702
Messaggi di controllo	702
Distribuzioni	703
Organizzazione del software per la gestione delle news	704
Collocazione locale degli articoli	704
Gruppi amministrativi	704
File amministrativi	705
Distribuzione degli articoli	706
Utente specifico	706
Riferimenti	707

Usenet è una sorta di rete astratta il cui scopo è quello di gestire l'organizzazione e la diffusione di un sistema pubblico di messaggi, ovvero di *articoli*. Per «rete astratta» si intende qualcosa che va oltre i confini di una rete avente una sua tecnologia particolare: anche se oggi la rete più diffusa è Internet, Usenet non è necessariamente qualcosa che riguardi esclusivamente questo tipo di supporto.¹²

L'idea alla base della nascita di Usenet è stata quella di riuscire a realizzare un sistema automatico di diffusione di articoli tra un gruppo

di elaboratori, in modo da permettere agli utenti di questi di leggerli e di poter aggiungere i propri. Inizialmente il meccanismo attuato era molto semplice: a intervalli regolari (magari solo una volta al giorno) ogni nodo impacchettava gli articoli di cui disponeva e li spediva ai suoi nodi corrispondenti, i quali provvedevano poi a selezionare quelli che non avevano già ricevuto da altri (eliminando così gli articoli doppi).

L'evoluzione di Usenet ha portato all'introduzione di tecniche di diffusione più dinamiche, soprattutto attraverso l'introduzione del protocollo NNTP all'interno di Internet. Tuttavia, la complessità della storia di questo sistema di messaggi ha portato ad altrettanta difficoltà nella configurazione e nell'amministrazione del software relativo.

Software per Usenet

«

Il primo software utilizzato per realizzare questo meccanismo di diffusione di articoli ha una sua storia importante: «A», «B» (Bnews), «C» (C News) e INN (InterNet News). All'interno di questa sequenza, nel 1986, ovvero subito prima che apparisse C News, si inserisce lo standard NNTP (*Network news transfer protocol*), il cui scopo è quello di dare un protocollo per il trasferimento degli articoli di Usenet all'interno di Internet (attraverso il TCP), ricalcando l'esperienza di SMTP (il protocollo per i messaggi di posta elettronica).

L'introduzione del protocollo NNTP, documentata dall'RFC 977, coincideva con la realizzazione del «demone NNTP», ovvero ciò che adesso è conosciuto come la «realizzazione di riferimento» (*reference implementation*). Questo demone si inserisce in pratica come

un'interfaccia rispetto a sistemi di gestione delle news come Bnews e C News.

Da questa situazione un po' confusa emerge in particolare InterNet News, ovvero INN, che oltre a gestire il protocollo NNTP è in grado di amministrare anche lo strato sottostante che in altre circostanze sarebbe di competenza di Bnews o di C News.

Organizzazione generale del sistema di news di Usenet

Inizialmente si accennava al fatto che Usenet sia una sorta di rete sulle reti: in pratica è formata dall'insieme di nodi che offrono un servizio di pubblicazione e diffusione di articoli. Ogni nodo riceve da altri nodi gli articoli che gli interessano e probabilmente provvede a sua volta a fornirli ad altri aggiungendovi i propri. Questo meccanismo è il feed, con il quale si alimenta il flusso degli articoli di Usenet.

Le news di Usenet sono organizzate in **gruppi** (ovvero gruppi di discussione o *newsgroup*), all'interno dei quali gli utenti possono leggere gli articoli e spedirne dei nuovi (*post*). Questi gruppi di discussione sono denominati in modo gerarchico, utilizzando una notazione simile a quella dei nomi a dominio di Internet, senza avere però alcuna relazione con questi. L'idea che sta alla base della gerarchia dei gruppi di discussione di Usenet è la stessa delle directory in un file system. Per esempio, potrebbe esistere il gruppo '**grano.farina**', ma anche il gruppo '**grano.farina.farro**', così come '**grano.sementi**'; eventualmente potrebbe mancare il gruppo '**grano**' puro e semplice. Evidentemente, i nomi usati nella

definizione della gerarchia danno già un'idea degli argomenti per i quali sono stati fatti.

Come si intuisce, la gerarchia si sviluppa da sinistra a destra, esattamente come nella notazione che si usa per rappresentare i percorsi all'interno di un file system, utilizzando però il punto per separarne i vari elementi.

Abituati oggi con Internet, può risultare difficile la comprensione del meccanismo che sta alla base di Usenet. Ci si potrebbe domandare come mai non sia possibile contattare questo o quel gruppo di discussione raggiungendo semplicemente il nodo da cui ha origine. Il fatto è che non c'è un'origine vera e propria per un gruppo; tutto quanto è frutto di una cooperazione tra i vari siti che offrono accesso a Usenet e per raggiungere un gruppo occorre avere accesso presso uno di questi siti (uno che gestisca il gruppo a cui si è interessati).

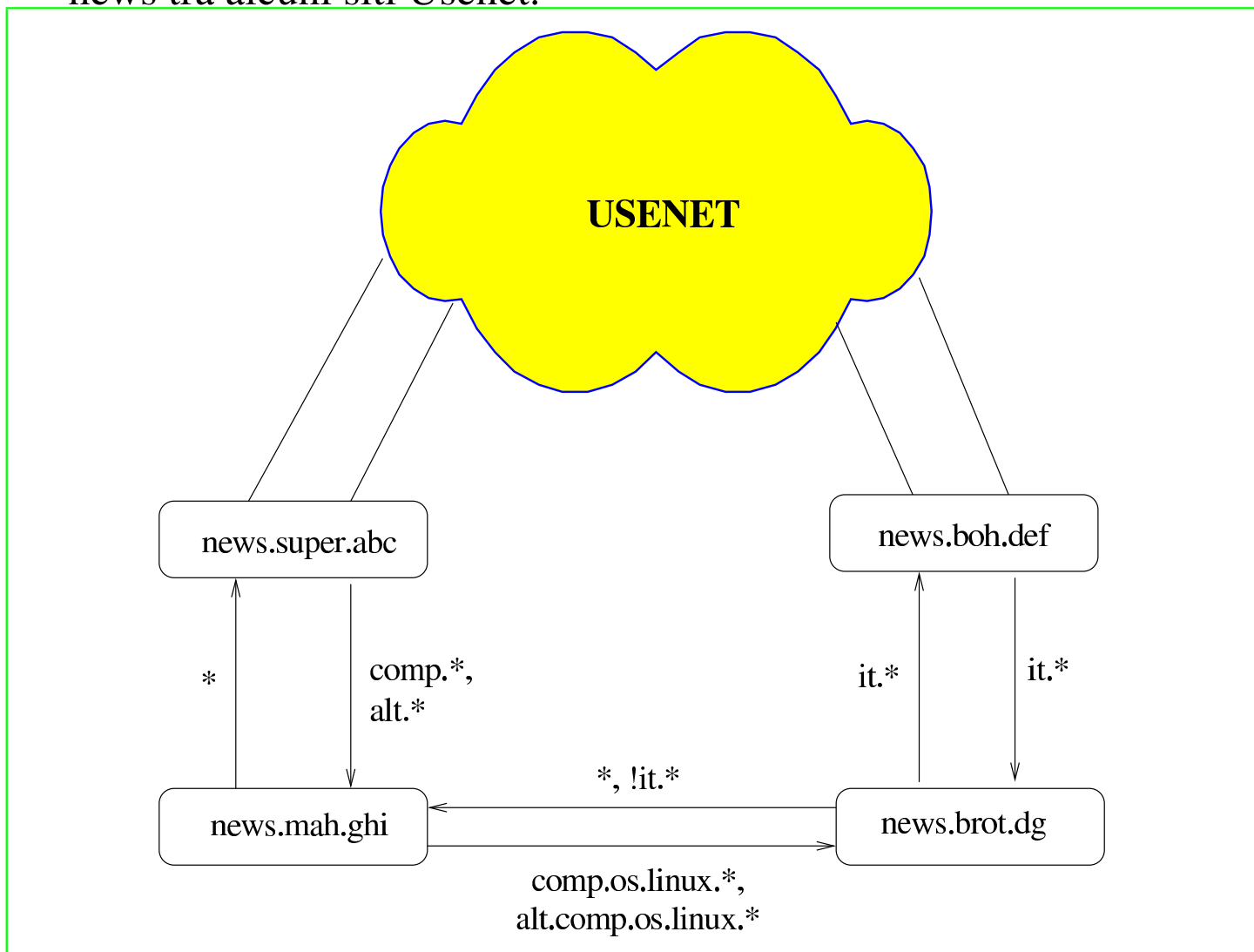
La quantità e la varietà dei gruppi esistenti è tale per cui è diventato impossibile gestire tutti i gruppi esistenti in un solo «sito Usenet», soprattutto a causa del volume di traffico che si genererebbe nella rete. In questo senso, l'amministrazione di un servizio del genere comporta due problemi fondamentali: decidere i gruppi che si vogliono gestire e trovare i siti Usenet con cui comunicare per potervi partecipare.

Selezione dei gruppi e feed

«

Il feed è il meccanismo alla base del funzionamento di Usenet. In pratica è ciò che permette la diffusione degli articoli dei vari gruppi di discussione. Volendo predisporre un server di news, ovvero un sito Usenet, occorre ottenere il feed dei gruppi a cui si è interessati, filtrando ciò che non interessa.

Figura u54.1. Esempio di come potrebbe funzionare il flusso di news tra alcuni siti Usenet.



Nell'esempio di figura u54.1 vengono mostrati dei nodi con nomi a dominio tipici di Internet, i quali collaborano al sistema di Usenet. In particolare, '**news.super.abc**' è inteso come un sito Usenet che dispone di una grande quantità di gruppi; da questo '**news.mah.ghi**' attinge i gruppi '**comp.***' e '**alt.***'. Inoltre, da un'altra parte dell'universo di Usenet appare '**news.boh.def**' che tra gli altri dispone dei gruppi '**it.***'; questi vengono forniti in particolare a *news.brot.dg*. Ma a *news.brot.dg* non basta perché vuole avere anche i gruppi '**comp.os.linux.***' e '**alt.comp.os.linux.***': questi

li ottiene da **'news.mah.ghi'**.

Tuttavia, non è sufficiente ottenere la copia degli articoli di questo o quel gruppo, bisogna considerare che ogni sito Usenet può aggiungere i propri e questi dovrebbero diffondersi su tutti gli altri siti Usenet che gestiscono il gruppo corrispondente; inoltre, ogni nodo potrebbe aggiungere i propri gruppi locali, che potrebbero o meno essere accolti anche da altri.³

Il feed che si vede indicato nella figura ha due direzioni: una per i gruppi da ricevere da un nodo e l'altra per i gruppi da inviare allo stesso. Per esempio, **'news.mah.ghi'** riceve solo i gruppi **'comp.*'** e **'alt.*'** da **'news.super.abc'**, ma gli restituisce indietro tutto; precisamente, restituisce tutto quello che **'news.super.abc'** è disposto ad accettare. In questo modo, gli articoli spediti per mezzo di **'news.mah.ghi'** prendono la loro strada attraverso Usenet. Invece, il caso di *news.brot.dg* è più complesso, perché da una parte ha un feed che gli permette di accedere ai gruppi **'it.*'**, dall'altra ne ha un altro per alcuni gruppi **'comp'** e **'alt'**. Dallo schema della figura si intende che questo mantenga un feed con **'news.boh.def'** esclusivamente per i gruppi **'it.*'**; dall'altra parte, nel collegamento con **'news.mah.ghi'** possono essere mandati tutti gli articoli che semplicemente non appartengono ai gruppi **'it.*'**, poi è compito di **'news.mah.ghi'** escludere quello che non gli riguarda.

Gestione degli articoli

«

Data l'organizzazione di Usenet in cui gli articoli si distribuiscono attraverso percorsi non prevedibili, è necessario che ogni sito sia in grado di gestire intelligentemente ciò che lo riguarda. Per esempio occorre evitare di accettare articoli che sono già stati ricevuti in

qualche modo; gestendo una quantità elevata di gruppi occorre dare una scadenza alla loro conservazione, eliminandoli periodicamente; inoltre è opportuno evitare di diffondere articoli attraverso nodi che sono già stati attraversati da questi.

Per ottenere questo risultato, il software che si utilizza per gestire le news deve avere un sistema di registrazione del traffico che lo riguarda, conservando anche le informazioni sugli articoli scaduti che nel frattempo sono stati cancellati localmente. Per distinguere gli articoli occorre un modo preciso e «sicuro»; questo si ottiene attraverso una stringa di identificazione, generata in qualche modo da chi riceve per la prima volta l'articolo e inserita nell'intestazione del messaggio attraverso il campo '**Message-ID:**', come nell'esempio seguente:

```
Message-ID: <36F130C4.2E242945@brot.dg>
```

La scadenza di un articolo può essere indicata anche da chi lo scrive, attraverso il campo '**Expires:**'; il sistema di gestione locale delle news può stabilire una scadenza predefinita e anche una scadenza massima, senza rispettare necessariamente quanto richiesto dall'articolo stesso.

Ogni messaggio porta con sé anche l'informazione del percorso dei vari nodi di Usenet che ha attraversato nel campo '**Path:**'. La forma è quella del cosiddetto *bang path*, perché si tratta di una stringa in cui si utilizza il punto esclamativo per separare i vari nomi. Per esempio,

```
Path: roggen.brot.dg!dinkel.brot.dg
```

rappresenta il transito da *dinkel.brot.dg* a *roggen.brot.dg*. Evidentemente, conoscendo i suoi passaggi, si deve evitare che questo articolo sia ritrasmesso ai nodi che l'hanno già visto transitare. Quando ciò

accade, quello è il punto in cui quella copia particolare dell'articolo si ferma.

Come nella posta elettronica, un articolo può essere diretto a più di un gruppo. In tal caso, i sistemi di gestione delle news che gestiscono tutti o alcuni di questi gruppi, dovrebbero riprodurre una sola copia «fisica» dell'articolo, mantenendo dei riferimenti all'interno di tutti i gruppi in cui questo è diretto. In pratica, nei sistemi Unix questo si ottiene con la tecnica dei collegamenti (fisici o simbolici che siano).

Protocollo NNTP

«

Prima di inserirsi in Internet, Usenet non aveva alcun bisogno del protocollo NNTP, ma adesso, con questo si semplificano molte cose. Di solito, un servente NNTP è anche un servente di news, dal quale un programma cliente, come può esserlo un navigatore grafico comune, può prelevare gli articoli e spedirne dei nuovi. Tuttavia, il protocollo NNTP fornisce anche un modo in più per ottenere il feed tra i vari siti di Usenet.

Messaggi di controllo

«

Più o meno come accade con le liste attraverso la posta elettronica, è possibile controllare in qualche modo il servizio delle news per mezzo di messaggi di controllo, contenenti campi particolari nell'intestazione. In questo caso si tratta del campo '**Control:**' e il comando più importante è '**cancel**', che dovrebbe permettere all'utente che ha spedito inizialmente un messaggio di ordinarne la cancellazione all'interno di tutta la rete Usenet. L'esempio seguente mostra in che modo potrebbe essere indicato questo comando:

```
Control: cancel <36F3CE29.F6176F5D@brot.dg>
```


Di solito un messaggio di questo tipo viene generato automaticamente dal programma utilizzato per accedere al servizio delle news, richiamando un comando particolare in base all'organizzazione del programma stesso. Tuttavia, è importante che il servente che lo riceve verifichi che si tratti verosimilmente dell'utente che aveva spedito originalmente l'articolo che adesso si richiede di cancellare.

La possibilità o meno di utilizzare i messaggi di controllo è regolata attraverso il file `/etc/news/control.ctl`, che è documentato nella pagina di manuale *control.ctl(5)*.

Distribuzioni

Gli articoli possono distinguersi, oltre che per gruppi, anche per *distribuzioni*. Si tratta del campo `'Distribution:'` che potrebbe apparire nell'intestazione di un messaggio. I nomi da attribuire alle distribuzioni possono servire a qualificare in modo alternativo gli articoli, per qualche scopo, per esempio per limitare la loro diffusione a un ambito locale o «regionale». Di sicuro si conoscono due distribuzioni comuni: `'world'` e `'local'`. In linea di massima si può dire che un articolo fatto per la distribuzione `'world'` dovrebbe essere lasciato diffondersi su tutti i siti Usenet, mentre un altro articolo etichettato per la distribuzione `'local'`, dovrebbe essere inteso per un uso locale riferito al sito Usenet attraverso cui è stato spedito.

```
Distribution: local
```

L'esempio mostra in che modo potrebbe essere composto il campo `'Distribution:'` quando viene utilizzato il nome `'local'`.

Di solito, l'utilizzatore normale non si cura di questo campo nell'intestazione dei suoi articoli e probabilmente non ha nemmeno il modo di aggiungerlo. In questo senso è poi il software utilizzato per

la gestione delle news che dovrebbe essere configurato in modo da attribuire un valore predefinito, eventualmente in base al gruppo in cui è stato spedito.

Organizzazione del software per la gestione delle news

«

Come già accennato, il software per la gestione di un sito Usenet segue un filone più o meno continuo, stabilendo implicitamente il legame tra Usenet e Unix. Il software più comune in questo momento è composto da C News, a cui si abbina normalmente un servizio NNTP, e da INN, che al contrario di C News può fare tutto da solo. Tra questi ci sono alcune affinità importanti la cui conoscenza può facilitare lo studio di ciò che si intende utilizzare effettivamente.

Collocazione locale degli articoli

«

Su un sistema GNU, gli articoli di un servizio di news sono collocati generalmente a partire dalla directory `/var/spool/news/`, con una struttura che ricalca quella del nome dei gruppi di discussione. Per esempio, il gruppo `comp.os.linux` dovrebbe trovarsi nella directory `/var/spool/news/comp/os/linux/`. All'interno di ogni directory riferita a un gruppo particolare vengono collocati gli articoli in forma di file, denominandoli in modo numerico: `'1'`, `'2'`,... Quel numero serve solo come riferimento, in base a quanto organizzato dal servizio locale di gestione delle news.

Quando si riceve un messaggio destinato a più gruppi, dei quali tutti o alcuni di questi sono gestiti localmente, quello che si ottiene è la riproduzione di questi attraverso dei collegamenti (generalmente dei collegamenti fisici).

Gruppi amministrativi

Il software di gestione delle news richiede normalmente la presenza di alcuni gruppi locali per uso amministrativo, che non vanno eliminati. Potrebbe trattarsi di:

'control'	utilizzato per conservare i messaggi di controllo, per esempio gli ordini di rimozione degli articoli;
'junk'	utilizzato per raccogliere gli articoli dei gruppi che localmente risultano mancanti;
'test'	utilizzato a livello diagnostico per poter sperimentare l'invio di articoli;
'to'	utilizzato per scopi vari dal software con cui si gestisce il servizio.

File amministrativi

I file amministrativi, come per esempio lo storico degli articoli già visti, dovrebbero trovarsi nella directory `/var/lib/news/`. In particolare, in questa directory si dovrebbe trovare il file `active`, contenente le informazioni necessarie a determinare quali siano i gruppi gestiti e i relativi intervalli di «numeri», ovvero dei nomi dei file rispettivi. In generale, dovrebbe essere sufficiente modificare questo file per creare o definire la gestione di un nuovo gruppo di discussione, tanto che di solito per avviare un servizio del genere si comincia prelevando il file `active` di un altro nodo Usenet e lo si modifica successivamente.

Distribuzione degli articoli

«

Il sistema di gestione delle news riceve in qualche modo gli articoli provenienti dai nodi corrispondenti, filtrando presumibilmente solo una parte dei gruppi e scartando i doppioni (ovvero ciò che in base allo storico risulta essere già stato visto localmente). Successivamente si deve occupare di dirigere una copia di questi articoli nel deposito locale, in modo da metterli a disposizione per la lettura a chi può averne accesso; inoltre si deve curare di far proseguire (in qualche modo) una copia di questi articoli ai nodi corrispondenti che non appaiano già nel percorso annotato nel campo **'Path:'**.

Nello stesso modo, si potrebbe definire l'accumulo di una copia degli articoli in un file di archivio, oppure anche verso un sistema di posta elettronica (probabilmente un indirizzo riferito a una lista).

Utente specifico

«

Generalmente, alla gestione del software per l'amministrazione delle news viene abbinato un utente di sistema, denominato **'news'**, al quale viene aggiunto normalmente anche il gruppo **'news'**. La directory personale di questo utente fittizio dovrebbe essere **'/var/spool/news/'**. Tutti i file amministrativi, compresi quelli di configurazione che si trovano sotto **'/etc/news/'**, assieme a tutto ciò che si trova a partire da **'/usr/lib/news/'**, dovrebbe appartenere all'utente (e al gruppo) **'news'**.

Tutte le volte che l'amministratore del sistema deve intervenire sulla gestione del software delle news dovrebbe avere l'accortezza di utilizzare l'identità e i privilegi dell'utente **'news'**, possibilmente attraverso il comando **'su'**, oppure deve fare attenzione a sistemare la proprietà dei file che crea.

Quando si utilizza il sistema Cron per eseguire delle elaborazioni periodiche, occorre preoccuparsi di questo fatto; eventualmente si può utilizzare il file crontab dell'utente **'news'**, oppure quello dell'utente **'root'**, badando però a sistemare l'identità dell'utente:

```
su - news -c "/usr/lib/news/bin/news.daily"
```

Riferimenti

- Marco d'Itri, *Usenet*
<http://www.linux.it/~md/usenet/>
- Brendan P. Kehoe, *Zen and The Art of the Internet*, 1992
http://www.cs.indiana.edu/docproject/zen/zen-1.0_toc.html
- Brian Kantor, Phil Lapsley, *RFC 977, Network News Transfer Protocol*, 1986
<http://www.ietf.org/rfc/rfc977.txt>
- Olaf Kirch, *NAG, The Linux Network Administrators' Guide*

¹ In altri termini, Usenet è stata il sistema BBS di Unix, prima attraverso il trasporto UUCP e poi estendendosi anche su Internet (oltre che su altri sistemi operativi).

² Il modo originale di scrivere il nome di questo sistema di messaggistica è USENET, utilizzando quindi solo le lettere maiuscole, ma successivamente si è diffusa la forma usata in questo documento, cioè con la sola iniziale maiuscola.

³ L'aggiunta di un gruppo di discussione al di fuori dell'ambito locale è una cosa che deve essere fatta di comune accordo, secondo una procedura stabilita.

