16

L'ARCHITETTURA DI RETE TCP/IP

16.1 INTRODUZIONE

Nella prima metà degli anni '70, la Defence Advanced Research Project Agency (DARPA) dimostrò interesse per lo sviluppo di una rete a commutazione di pacchetto per l'interconnessione di calcolatori eterogenei, da utilizzarsi come mezzo di comunicazione tra le istituzioni di ricerca degli Stati Uniti. DARPA finanziò a tal scopo l'Università di Stanford e la BBN (Bolt, Beranek and Newman) affinché sviluppassero un insieme di protocolli di comunicazione.

Verso la fine degli anni '70, tale sforzo portò al completamento dell'*Internet Protocol Suite*, di cui i due protocolli più noti sono il TCP (*Transmission Control Protocol*) e l'IP (*Internet Protocol*).

Questi protocolli furono utilizzati da un gruppo di ricercatori per la rete ARPAnet e ottennero un elevato successo, anche perché posti sin dall'inizio nel dominio pubblico e quindi utilizzabili gratuitamente da tutti.

Il nome più accurato per l'architettura di rete rimane quello di Internet Protocol Suite, anche se comunemente si fa riferimento ad essa con la sigla TCP/IP o IP/TCP. Questo può portare ad alcune ambiguità: ad esempio è comune sentir parlare di NFS come un servizio basato su TCP/IP, anche se NFS non usa il protocollo TCP, ma un protocollo alternativo detto UDP appartenete all'Internet Protocol Suite. Visto l'uso estremamente comune della sigla TCP/IP, essa verrà adottata anche in questo libro in luogo del termine più corretto, quando non crei confusione.

TCP/IP è l'architettura adottata dalla rete Internet che, con le sue decine di milioni di calcolatori e il suo tasso di crescita del 5% al mese, è la più grande rete di calcolatori al mondo.

I protocolli appartenenti a questa architettura sono specificati tramite standard che si chiamano RFC (*Request For Comments*) facilmente reperibili sulla rete Internet. Famoso

è lo RFC 791 Internet Protocol, datato 1981, che specifica appunto il protocollo IP.

L'architettura TCP/IP ha dei componenti, quali l'IP, indubbiamente datati, ma non obsoleti: il grande successo di TCP/IP è quotidiano. Negli anni 1990, gli anni della maturità dell'ISO/OSI, l'unica architettura di rete che sembra interessare il mercato è quasi paradossalmente TCP/IP.

Anche gli enti di standardizzazione nazionali e internazionali hanno dovuto arrendersi davanti alla massiccia diffusione di TCP/IP e dargli la stessa dignità di ISO/OSI.

16.2 ARCHITETTURA

La figura 16.1 mostra l'architettura dell'Internet Protocol Suite e la paragona con il modello di riferimento ISO/OSI. Questa architettura permette l'esistenza di più pile di protocolli alternative tra loro ed ottimizzate per determinate applicazioni.

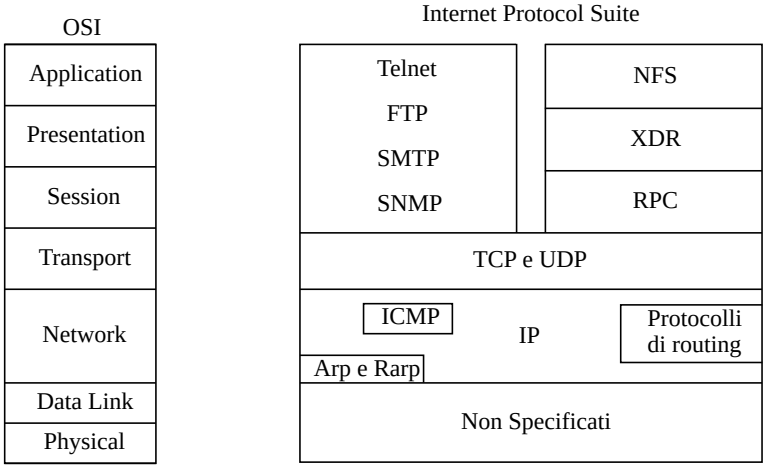

Fig. 16.1 - Internet Protocol Suite.

Esempi di possibili pile sono: telnet/TCP/IP/Ethernet, FTP/TCP/IP/Ethernet, SNMP/UDP/IP/FDDI e NFS/XDR/RPC/UDP/IP/Token-Ring.

16.3 SOTTO L'IP

L'architettura di rete TCP/IP non specifica i livelli 1 e 2 della rete, ma utilizza quelli normalmente disponibili e conformi agli standard. Ad esempio, nell'ambito

delle reti locali opera su Ethernet/IEEE802.3, Token-Ring e FDDI; nell'ambito delle reti geografiche su HDLC, PPP, SLIP, X.25, Frame Relay, SMDS e ATM.

Esistono anche realizzazioni per reti molto strane, spesso diffuse solo all'interno di certe comunità, ad esempio AIX.25, una rete packet switched dei radioamatori.

16.4 IL PROTOCOLLO IP

Il protocollo IP (*Internet Protocol*) è il protocollo principale del livello 3 (Network) dell'architettura TCP/IP. Si tratta di un protocollo semplice, di tipo datagram, non connesso (connectionless o CLNS), specificato in RFC 791. Insieme a TCP costituisce il nucleo originale e principale dell'Internet Protocol Suite.

IP si occupa di instradare i messaggi sulla rete, ma ha anche funzioni di frammentazione e riassemblaggio dei messaggi e di rilevazione (non correzione) degli errori.

Il formato dell'header del pacchetto IP è mostrato in figura 16.2. Esempi di pacchetti IP sono riportati in appendice B, paragrafo B.3.

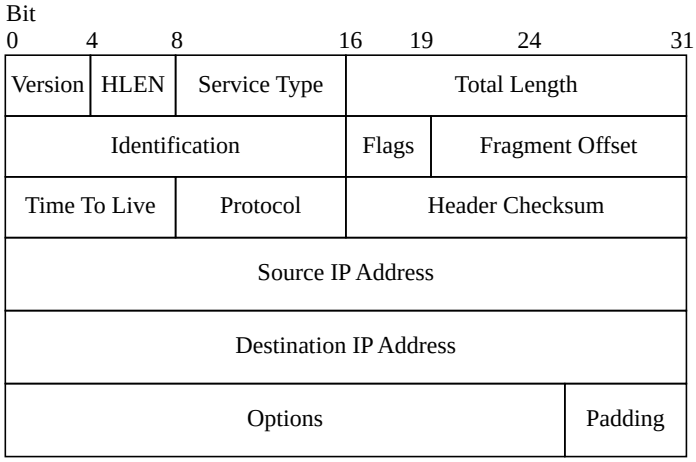

Fig. 16.2 - Header del pacchetto IP.

Il significato dei campi del pacchetto IP è il seguente:

- *Version*: è il numero di versione del protocollo IP che ha generato il pacchetto; attualmente questo campo vale sempre 4;
- *HLEN (Header LENgth)*: è la lunghezza dell'header IP, variabile in funzione del campo option, espressa come numero di parole da 32 bit;

- *service type*: specifica come un protocollo di livello superiore vuole che il pacchetto sia trattato; è possibile assegnare vari livelli di priorità utilizzando questo campo;
- *total length*: è la lunghezza del pacchetto IP (header più dati) in byte;
- *identification*: questo campo contiene un numero intero che identifica il pacchetto; è usato per permettere il riassemblaggio di un pacchetto frammentato;
- *flags*: specificano se un pacchetto può essere frammentato e se si tratta dell'ultimo frammento di un pacchetto;
- *fragment offset*: è l'offset del frammento in multipli di 8 byte;
- *time to live*: è un contatore che viene decrementato con il passaggio del tempo; quando il contatore arriva a zero il pacchetto viene scartato. Permette di eliminare i pacchetti che, a causa di un malfunzionamento, sono entrati in loop;
- *protocol*: identifica il protocollo di livello superiore contenuto nel campo dati del pacchetto. In appendice A, paragrafo A.7, sono riportati i codici dei protocolli che possono essere contenuti nel campo dati di IP;
- *header checksum*: è un campo utilizzato per controllare che l'header IP sia corretto;
- *source e destination address*: sono gli indirizzi IP di mittente e destinatario, entrambi su 32 bit;
- *option*: è un campo usato dall'IP per fornire varie opzioni, quali la sicurezza e il source routing, che può essere di tipo loose o strict.

L'header IP è seguito dal campo dati che contiene la PDU del protocollo di livello superiore.

16.5 INDIRIZZAMENTO IP

L'indirizzamento IP è parte integrante del processo di instradamento dei messaggi sulla rete. Gli indirizzi IP, che devono essere univoci sulla rete, sono lunghi 32 bit (quattro byte) e sono espressi scrivendo i valori decimali di ciascun byte separati dal carattere punto.

Esempi di indirizzi IP sono: `34.0.0.1`, `129.130.7.4` e `197.67.12.3`.

Agli indirizzi IP si associano per comodità uno o più nomi che possono essere definiti localmente in un file "hosts" che ha il seguente formato:

```
223.1.2.1  alpha
223.1.2.2  beta
223.1.2.3  gamma
223.1.2.4  delta    mycomputer
223.1.3.2  epsilon
223.1.4.2  iota
```

Questo approccio diviene impraticabile quando la rete IP cresce di dimensione e allora si preferisce utilizzare una base di dati distribuita per la gestione dei nomi (si veda il paragrafo 16.12.4).

Gli indirizzi IP comprendono due o tre parti. La prima parte indica l'indirizzo della rete (network), la seconda (se presente) quello della sottorete (subnet) e la terza quello dell'host.

Occorre subito evidenziare che non sono i nodi ad avere un indirizzo IP, bensì le interfacce. Quindi se un nodo ha tre interfacce, esso ha tre indirizzi IP. Poiché la maggior parte dei nodi ha una sola interfaccia, è comune parlare dell'indirizzo IP di un nodo. Questo tuttavia è senza dubbio sbagliato nel caso dei router che hanno, per definizione, più di una interfaccia.

Gli indirizzi IP sono assegnati da un'unica autorità e quindi sono garantiti univoci a livello mondiale*. Essi vengono assegnati a gruppi come dettagliato nel seguito.

Gli indirizzi IP sono suddivisi in cinque classi, come schematizzato in figura 16.3.

- Classe A. Sono concepiti per poche reti di dimensioni molto grandi. I bit che indicano la rete sono 7 e quelli che indicano l'host 24. Quindi si possono avere al massimo 128 reti di classe A, ciascuna con una dimensione massima di circa 16 milioni di indirizzi. Gli indirizzi di classe A sono riconoscibili in quanto il primo campo dell'indirizzo è compreso tra 0 e 127.
- Classe B. Sono concepiti per un numero medio reti di dimensioni medio-grandi. I bit che indicano la rete sono 14 e quelli che indicano l'host 16. Quindi si possono avere al massimo circa 16000 reti di classe B, ciascuna con una dimensione massima di circa 64000 indirizzi. Gli indirizzi di classe B sono riconoscibili in quanto il primo campo dell'indirizzo è compreso tra 128 e 191.
- Classe C. Sono concepiti per moltissime reti di dimensioni piccole. I bit che indicano la rete sono 21 e quelli che indicano l'host 8. Quindi si possono avere al massimo 2 milioni di reti di classe C, ciascuna con una dimensione massima di 256 indirizzi. Gli indirizzi di classe C sono riconoscibili in quanto il primo campo dell'indirizzo è compreso tra 192 e 223.

* In Italia esistono vari soggetti che possono assegnare indirizzi Internet. Chi fosse interessato può contattare, ad esempio, il GARR/NIS, c/o CNUCE/CNR, Via Santa Maria 36, 56126 Pisa, Tel. 050-593111, Fax 050-904052.

- Classe D. Sono riservati ad applicazioni di multicast secondo quanto descritto nel RFC 1112. In appendice A, paragrafo A.6, sono elencati gli indirizzi IP di multicast attualmente assegnati. Gli indirizzi di classe D sono riconoscibili in quanto il primo campo dell'indirizzo è compreso tra 224 e 239.
- Classe E. Questi indirizzi sono riservati per usi futuri. Gli indirizzi di classe E sono riconoscibili in quanto il primo campo dell'indirizzo è compreso tra 240 e 255.

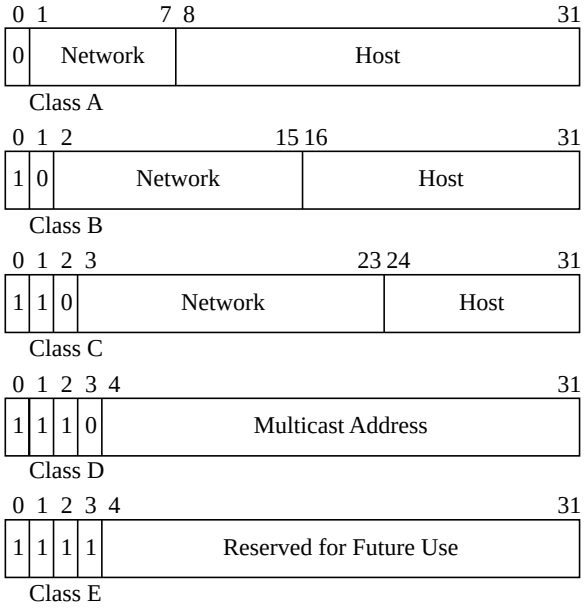

Fig. 16.3 - Classi di indirizzi.

In appendice A, paragrafo A.5, sono riportati gli indirizzi IP con significato particolare, ad esempio broadcast e loopback.

La parte host di un indiriirizzo di classe A, B o C può essere ulteriormente divisa in due parti: la subnet e l'host. La figura 16.4 mostra un indirizzo di classe B prima e dopo il subnetting. Il subnetting è discusso nel RFC 950.

L'ampiezza dei campi subnet e host può essere definita in modo molto flessibile tramite un parametro detto *netmask*. La netmask contiene bit a uno in corrispondenza dei campi network e subnet, e a zero in corrispondenza del campo host.

Ad esempio una netmask 11111111111111111111111100000000, più comunemente scritta come indirizzo IP 255.255.255.0 o in esadecimale ffffff00, indica che il campo host coincide con l'ultimo byte dell'indirizzo.

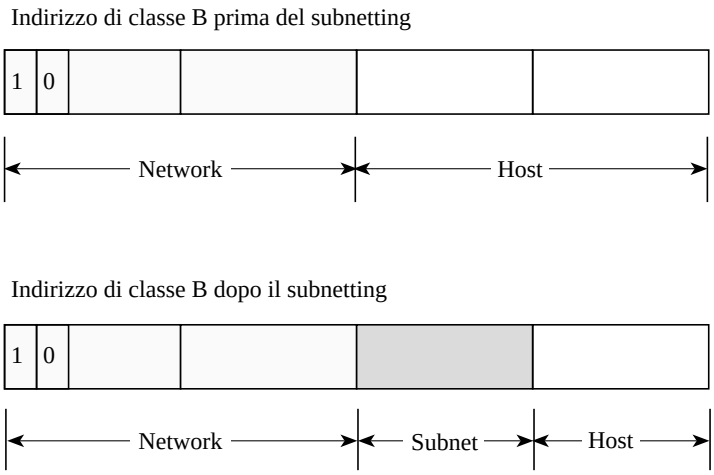

Fig. 16.4 - Subnetting.

All'interno di una network IP la netmask deve essere univoca, in quanto il partizionamento della network in subnet deve essere univoco. La netmask viene messa in AND bit a bit con gli indirizzi IP per estrarre la parte network e subnet. Tramite questo procedimento è possibile verificare se due indirizzi appartengono alla stessa subnet.

Ad esempio si supponga di aver una netmask 255.255.254.0 e i due indirizzi 128.155.4.77 e 128.155.5.75. Mettendo in AND bit a bit gli indirizzi con la netmask si ottiene in entrambi i casi 128.155.4.0 e quindi gli indirizzi appartengono alla stessa subnet. Un caso di due indirizzi simili ai precedenti, ma appartenti a subnet diverse è 128.155.5.75 e 128.155.6.77, in quanto i due AND rendono rispettivamente i valori 128.155.4.0 e 128.155.6.0.

L'importanza di comprendere se due indirizzi appartengono o no alla stessa subnet è fondamentale in quanto il primo livello di routing è implicito nella corrispondenza fissata in TCP/IP tra reti fisiche e subnet: *una rete fisica deve coincidere con una subnet IP*.

Questa situazione è illustrata in figura 16.5 dove sono mostrate più reti fisiche e le subnet IP associate. La rete è di classe A (primo byte uguale a 11) e ha una netmask 255.255.0.0.

Nell'esempio sono presenti le subnet 11.1 e 11.2. Si noti che il bridge, operando a livello 2 ed essendo trasparente al protocollo IP, collega reti Ethernet appartenenti alla stessa subnet 11.1, mentre il router collega reti Ethernet appartenenti a subnet diverse (11.1 e 11.2). A tal scopo il router ha due indirizzi IP diversi: uno appartenente alla subnet 11.2 (11.2.0.254) e uno appartenente alla subnet 11.1 (11.1.0.253).

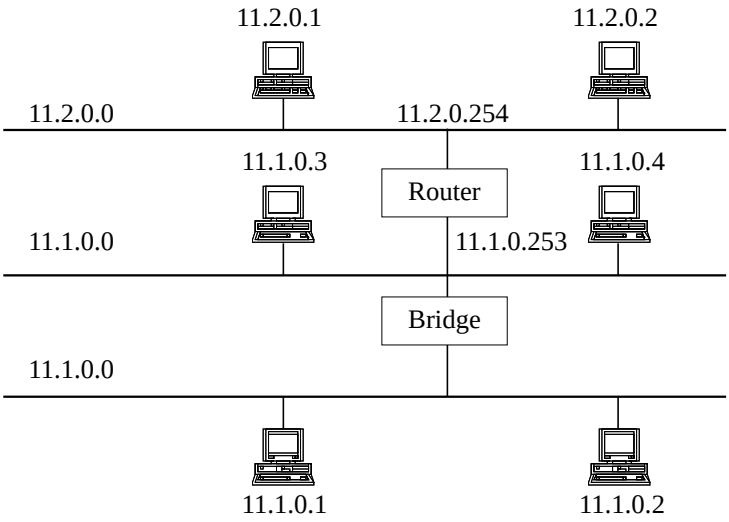

Fig. 16.5 - Indirizzi e reti fisiche.

Il bridge, essendo trasparente ai protocolli di livello superiore, non ha indirizzi IP.

La regola che impone una corrispondenza biunivoca tra subnet IP e reti fisiche è stata ultimamente leggermente rilassata per le LAN, dove è ammesso dalle implementazioni più recenti di TCP/IP che ad una LAN possano essere associate più subnet IP. Continua a non valere il viceversa.

Il concetto di subnet introduce un livello di gerarchia nelle reti TCP/IP. Il routing diventa un routing all'interno della subnet e tra subnet.

Il routing all'interno della subnet è banale in quanto la subnet coincide con una rete fisica che garantisce la raggiungibilità diretta delle stazioni ad essa collegate. L'unico problema che si può incontrare è quello di mappare gli indirizzi IP nei corrispondenti indirizzi di livello 2. Questo mappaggio è oggi quasi sempre gestito in modo automatico, tramite i protocolli ARP e RARP descritti nel seguito.

Il routing tra le subnet è gestito dagli IP router che originariamente erano stati definiti gateway. Tale definizione è infelice in quanto gli IP gateway sono quelli che OSI chiama router e i gateway OSI non hanno un corrispettivo nel mondo TCP/IP.

Nel seguito del capitolo si useranno i termini IP router e gateway come sinonimi.

Gli IP router effettuano l'instradamento sulla base di tabelle di instradamento che possono essere scritte manualmente dal gestore della rete o calcolate automaticamente tramite una serie di algoritmi di tipo distance vector o link state packet, descritti nel paragrafo 16.9.

Per comprendere meglio i concetti esposti sino a questo punto si consideri l'esempio di figura 16.6.

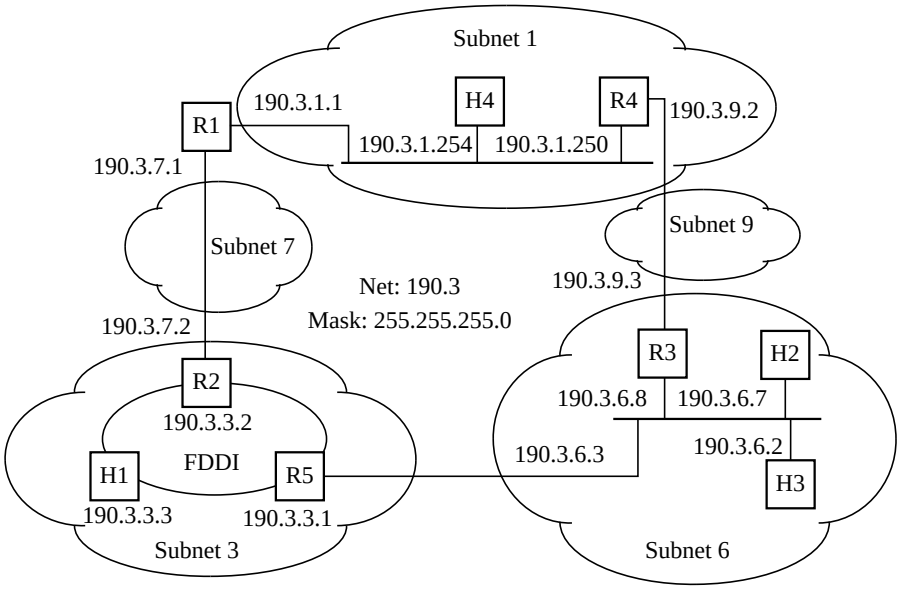

Fig. 16.6 - Esempio di indirizzamento IP.

In tale esempio la network 190.3 di classe B è stata partizionata in 256 subnet, grandi ciascuna 256 indirizzi. Nell'esempio sono utilizzate 5 subnet: la 190.3.1.0, la 190.3.3.0, la 190.3.6.0, la 190.3.7.0 e la 190.3.9.0.

Di queste cinque subnet, due corrispondono a reti fisiche di tipo Ethernet (la 190.3.6.0 e la 190.3.1.0), una ad una rete FDDI (la 190.3.3.0) e due a canali geografici di tipo punto-punto (la 190.3.7.0 e la 190.3.9.0).

Si noti che nel modello di gerarchia TCP/IP i router non fanno parte delle subnet, ma sono ad esse esterni. Inoltre non esiste il concetto di collegare direttamente due router: il collegamento avviene sempre tramite una subnet, eventualmente di tipo punto-punto e quindi formata da due soli indirizzi.

I router hanno tanti indirizzi quante sono le interfacce e quindi le subnet che collegano. Ad esempio il router R5 ha due indirizzi, uno associato alla rete FDDI (190.3.3.1) e l'altro associato alla rete Ethernet (190.3.6.3).

Sempre considerando il router R5, esso deve avere una tabella di instradamento che comprenda una entry per tutte le subnet cui il router non è direttamente collegato (in questo caso tre: la 190.3.1.0, la 190.3.7.0 e la 190.3.9.0).

La tabella di instradamento può essere simile a quella riportata in tabella 16.1.

Subnet di destinazione	Indirizzo del router cui inviare il pacchetto
190.3.1.0	190.3.3.2
190.3.7.0	190.3.3.2
190.3.9.0	190.3.6.8

Tab. 16.1 - Esempio di tabella di instradamento.

Si noti come tutti gli indirizzi della seconda colonna debbano appartenere a reti cui il router è direttamente connesso: nell'esempio appartengono infatti alle subnet 3 e 6 cui R5 è connesso.

La tabella di instradamento può essere creata manualmente con comandi del tipo:

```
route add 190.3.1.0 190.3.3.2
route add 190.3.7.0 190.3.3.2
route add 190.3.9.0 190.3.6.8
```

oppure calcolata dagli appositi algoritmi di routing descritti nel paragrafo 16.9.

Si noti che occorre anche definire per ogni host quale sia il suo router di default. Ad esempio, per l'host H4 si può definire che il router di default è R1, con un comando del tipo:

```
route add default 190.3.1.1
```

dato sul nodo H4 stesso.

Quando il nodo H4 deve trasmettere un pacchetto, per prima cosa verifica se il pacchetto è destinato ad un nodo appartenente alla sua stessa subnet. Se questo è il caso, la trasmissione può avvenire direttamente. In caso contrario invia il pacchetto al router di default (in questo caso R1). R1 instrada il messaggio a destinazione. Se durante tale operazione di instradamento R1 si trova a ritrasmettere il messaggio sulla stessa rete da cui lo ha ricevuto, ad esempio perché il messaggio è destinato alla subnet 9 e lo invia a R4, allora invia anche un messaggio di routing redirect al nodo mittente, in questo caso H4.

Il messaggio di routing redirect è inviato usando il protocollo ICMP che si appoggia su IP (si veda paragrafo 16.6).

Quando H4 ha deciso a quale indirizzo IP inviare il messaggio deve scoprire, se già non lo sa, qual è l'indirizzo di livello 2 (nel caso delle LAN l'indirizzo MAC) del destinatario. Per fare ciò utilizza il protocollo ARP descritto nel paragrafo 16.7.

16.6 IL PROTOCOLLO ICMP

Il protocollo Internet Control Message Protocol (ICMP) è stato progettato per riportare anomalie che accadono nel routing di pacchetti IP e verificare lo stato della rete. ICMP è specificato nel RFC 792.

La tabella 16.2 riporta i tipi di pacchetti ICMP.

Type Field	Message Type
0	Echo Reply
3	Destination Unreachable
4	Source Quence
5	Redirect
8	Echo Request
11	Time Exceeded for a Datagram
12	Parameter Problem on a Datagram
13	Timestamp Request
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply

Tab. 16.2 - Tipi di Messaggi ICMP.

I messaggi che riportano anomalie sono, ad esempio, *destination unreachable, time exceeded for a datagram* e *parameter problem on a datagram.*

I messaggi di verifica della raggiungibilità di un nodo sono *echo request* e *echo reply*.

Il messaggio *redirect* indica una condizione di stimolo ad un routing migliore, in quanto un router è stato attraversato inutilmente (ha dovuto ritrasmettere il messaggio sulla stessa rete da cui lo ha ricevuto).

Quando un host riceve un pacchetto di routing redirect tratta l'informazione in esso contenuta in modo simile a quella specificata da un comando di route add ed associa quindi un router diverso da quello di default a quella destinazione.

Gli ultimi messaggi ad essere stati introdotti nel protocollo ICMP sono *address mask request* e *address mask reply*, per permettere ad una interfaccia di scoprire automaticamente la netmask usata in quella network.

16.7 I PROTOCOLLI ARP E RARP

I protocolli *Address Resolution Protocol* (ARP) e *Reverse Address Resolution Protocol* (RARP) sono utilizzati per scoprire in modo automatico le corrispondenze tra gli indirizzi di livello 3 e gli indirizzi di livello 2 e viceversa. Questo è importante nelle LAN dove occorre creare una relazione tra gli indirizzi IP e gli indirizzi MAC.

I protocolli ARP e RARP sono specificati nel RFC 826.

Il protocollo ARP viene usato tutte le volte che una stazione collegata ad una LAN deve inviare un messaggio ad un nodo sulla stessa LAN di cui conosce unicamente l'indirizzo di livello 3.

Il protocollo RARP viene invece utilizzato dalle stazioni non dotate di memoria di massa (diskless) per scoprire il loro indirizzo IP in fase di bootstrap.

La figura 16.7 mostra la PDU di ARP/RARP.

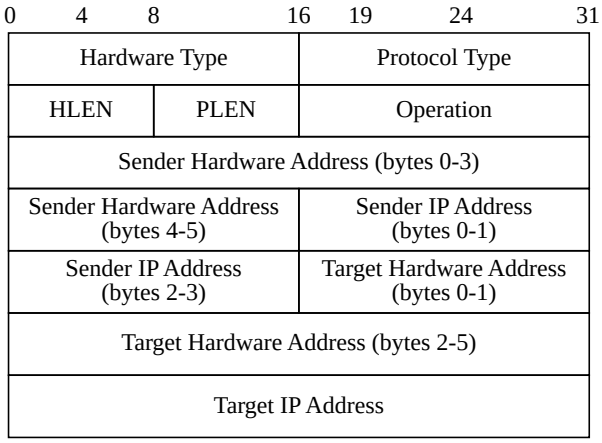

Fig. 16.7 - Il pacchetto ARP.

I protocolli ARP e RARP si appoggiano direttamente sulle reti locali e non su IP, come invece avviene nel caso di ICMP. Essi operano inviando le loro richieste in broadcast a tutte le stazioni della rete, anche quelle che non utilizzano TCP/IP.

La richiesta in broadcast di ARP contiene l'indirizzo IP del nodo di cui si vuole scoprire l'indirizzo di livello 2. Il nodo avente l'indirizzo IP specificato risponde alla richiesta fornendo il suo indirizzo di livello 2. Il protocollo RARP funziona in modo simile, ma è fornito un indirizzo di livello 2 e richiesto un indirizzo IP.

L'appendice A, paragrafo A.9, riporta i vari parametri del protocollo ARP, mentre un esempio di PDU ARP è riportata in appendice B, paragrafo B.3.3.

Per aumentare l'efficienza, i nodi mantengono in una cache locale le risposte ricevute alle richieste di ARP. Ad esempio, in figura 16.8 è mostrato il formato di ogni singola entry della ARP cache, come realizzata nel software Microsoft TCP/IP.

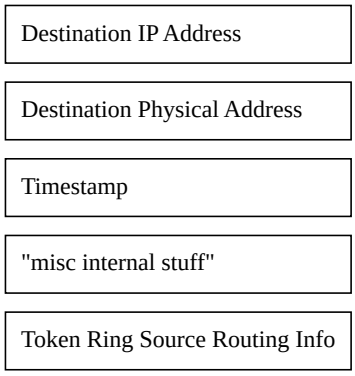

Fig. 16.8 - Una entry nella cache di ARP.

Il campo *timestamp* serve ad eliminare dalla cache le entry che sono più vecchie di 15 minuti.

Il contenuto della cache di ARP può essere visualizzato, in molte realizzazioni, tramite il comando:

```
arp -a
```

16.8 GLI AUTONOMOUS SYSTEM

Sino a questo punto il routing di TCP/IP è stato descritto come gerarchico su due livelli: un primo livello all'interno della subnet, implicito in quanto gestito dalla rete fisica; un secondo livello tra subnet gestito dagli IP router tramite tabelle di instradamento. Le subnet derivano dalla suddivisione di una network.

Le network sono ulteriormente raggruppate in *Autonomous System (AS)*, cioè in gruppi di network controllate e gestite da un'unica autorità.

Gli autonomus system sono identificati da un numero intero, univoco a livello mondiale, assegnato dalla stessa autorità che rilascia gli indirizzi Internet.

I router che instradano i messaggi all'interno dello stesso AS sono detti *interior router*, mentre quelli che instradano i messaggi anche tra AS diversi sono detti *exterior router*.

Un esempio di interconnessione di due AS è mostrato in figura 16.9.

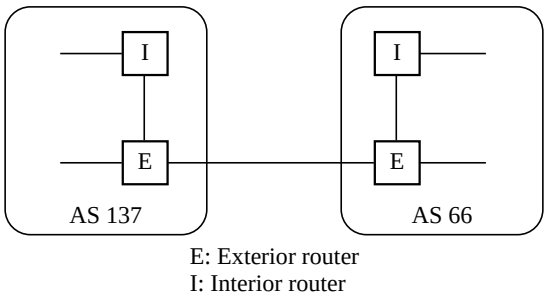

Fig. 16.9 - Exterior ed interior router.

Gli interior router possono scambiare informazioni di instradamento tramite un IGP (*Interior Gateway Protocol*), mentre gli exterior router utilizzano un EGP (*Exterior Gateway Protocol*).

Nei sottoparagrafi seguenti vengono illustrati i principali protocolli di routing di tipo EGP e IGP.

All'interno di un AS normalmente si usa lo stesso IGP su tutti i router.

16.9 I PROTOCOLLI DI ROUTING

L'architettura TCP/IP ha una varietà di protocolli di routing addirittura eccessiva. Nel seguito verranno discussi solo quelli che hanno maggior rilievo nella realtà italiana.

16.9.1 RIP

Il *Routing Information Protocol* (RIP) è un IGP originariamente progettato dalla Xerox per la sua rete XNS. È stato introdotto nell'architettura TCP/IP dall'Università di Berkeley nel 1982, definito come RFC 1058 nel 1988 e aggiornato con il RFC 1388 nel 1993.

RIP ha avuto una grandissima diffusione, soprattutto nelle implementazioni di reti di personal computer, ed è alla base di molti altri protocolli di routing: Novell, 3Com, Banyan, ecc.

RIP è un protocollo di tipo distance vector in cui ogni router invia il suo distance vector ai router adiacenti, ogni 30 secondi (si veda paragrafo 14.6). Le tabelle di instradamento memorizzano un solo cammino per ogni destinazione.

Il limite principale di RIP è che permette un numero massimo di hop pari a 15: ogni

destinazione più lontana di 15 hop viene considerata non raggiungibile.

Inoltre RIP ignora le velocità delle linee, non permette di definire costi o altre metriche, ma basa l'instradamento solo sulla minimizzazione del numero di hop. In caso di modifiche della topologia della rete, RIP è lento a convergere.

Per queste ragioni RIP può essere utilizzato solo in reti di piccole dimensioni.

Un esempio di PDU RIP è riportata in appendice B, paragrafo B.3.9.

16.9.2 IGRP

L'*Interior Gateway Routing Protocol* (IGRP) è un IGP sviluppato da Cisco System Inc. a metà degli anni 1980 per superare i limiti di RIP [5,6].

Si tratta anche in questo caso di un protocollo di tipo distance vector, ma con una metrica molto sofisticata. La scelta del cammino migliore è effettuata da IGRP combinando dei vettori di metriche contenenti: ritardo, banda, affidabilità, lunghezza massima del pacchetto e carico.

Inoltre IGRP permette il *multipath routing*, cioè la suddivisione del traffico tra più linee parallele. Il carico viene suddiviso in funzione delle metriche associate alle linee.

IGRP è nato come protocollo proprietario Cisco e sinora è stato reso disponibile solo sui router Cisco. A questo limite principale occorre aggiungere quelli generali, meno importanti, degli algoritmi distance vector descritti nel paragrafo 14.6.

16.9.3 OSPF

L'*Open Shortest Path First* (OSPF) è un IGP sviluppato appositamente per TCP/IP dall'IETF (*Internet Engineering Task Force*). Il gruppo di lavoro è stato costituito nel 1988 con lo scopo di realizzare un protocollo di tipo *link state packet* (si veda paragrafo 14.7) per TCP/IP.

OSPF è stato definito dal RFC 1247 nel 1991 e ridefinito dal RFC 1583 nel 1994.

OSPF ha il concetto di gerarchia. La radice della gerarchia è l'AS che può essere suddiviso in aree, ciascuna delle quali contiene un gruppo di reti contigue.

Il routing all'interno di un'area è detto intra-area, quello tra aree diverse inter-area. Ogni AS ha un'area detta di *backbone* ed identificata con 0.0.0.0 o più semplicemente 0. La backbone area (detta più semplicemente nel seguito backbone) può essere anche non contigua; in tal caso occorre configurare dei *virtual links* per garantire la coesione del backbone.

I router OSPF sono classificati secondo quattro categorie non mutuamente

esclusive:

- *Internal router*. Un router in cui tutte le network direttamente connesse appartengono alla stessa area. Questi router utilizzano una sola copia dell'algoritmo OSPF. I router che hanno solo interfacce sul backbone appartengono a questa categoria.
- *Area border router*. Un router che collega più aree. Questi router utilizzano più copie dell'algoritmo OSPF: una copia per ogni area direttamente connessa e una copia per il backbone. Gli area border router condensano le informazioni delle aree a loro collegate e le ridistribuiscono sul backbone. Il backbone ridistribuisce a sua volta queste informazioni alle altre aree.
- *Backbone router*. Un router che ha una interfaccia sul backbone. Questo include tutti i router che si collegano a più di un'area (area border router). I backbone router che hanno tutte le interfacce sul backbone sono considerati internal router.
- *AS boundary router*. Un router che scambia informazioni di routing con altri router appartenenti ad altri AS. Questa classificazione è ortogonale alle altre precedenti: un AS boundary router può essere un internal o area border router.

La figura 16.10 mostra un esempio di AS TCP/IP suddiviso in tre aree OSPF e connesso ad un altro AS.

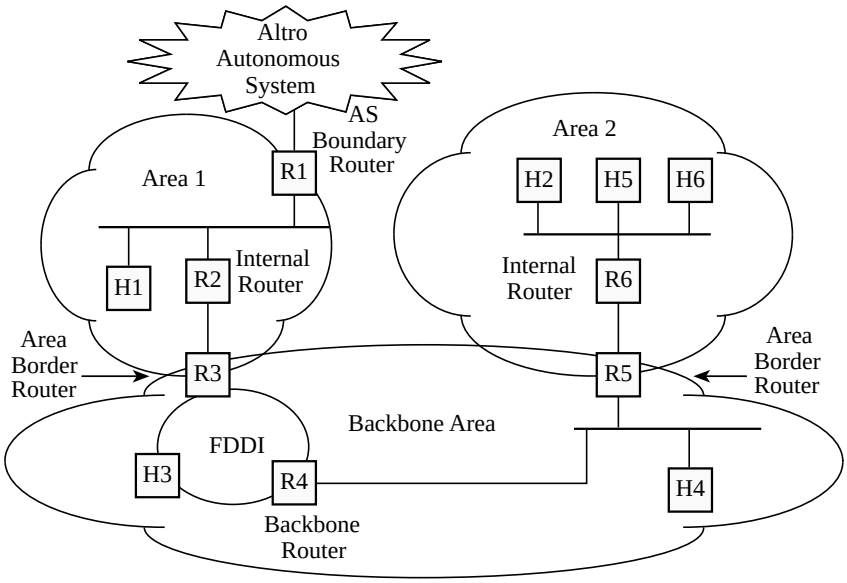

Fig. 16.10 - Esempio di utilizzo di OSPF.

OSPF è il protocollo più promettente per il routing di TCP/IP. Esso è infatti disponibile sui router di tutti i costruttori, è in grado di gestire reti di grosse dimensioni ed utilizza la tecnologia *Link State Packet* (LSP, paragrafo 14.7), che rappresenta lo stato dell'arte.

16.9.4 Integrated IS-IS

L'*integrated IS-IS*, detto anche *dual IS-IS*, è una versione del protocollo IS-IS (ISO 10589) che è in grado di ospitare informazioni di routing anche per protocolli diversi dall'OSI CLNS (ISO 8473). Per una discussione sui protocolli OSI si veda il capitolo 17.

Nell'ambito di un router multiprotocollo, questo approccio alla gestione di un solo protocollo di routing, comune a tutte le architetture di rete, si contrappone a quello più classico di avere un protocollo di routing per ogni architettura di rete e può portare ad una certa economia nell'utilizzo delle risorse di rete dei router.

16.9.5 EGP

L'Exterior Gateway Protocol è il primo EGP* ad essere stato ampiamente utilizzato all'interno della rete Internet. Specificato con RFC 904 nell'aprile 1984 è oggi ampiamente disponibile su tutti i router, anche se è ormai considerato un protocollo obsoleto e Internet lo sta sostituendo con il BGP (si veda il paragrafo 16.9.6).

EGP è simile ad un algoritmo distance vector, ma invece del concetto di costo specifica solo se la destinazione è raggiungibile oppure no. Questo ne impedisce il funzionamento su topologie magliate.

Esiste il concetto di un *core system* formato da una interconnessione di *core router* (figura 16.11).

EGP genera dei pacchetti di routing update che contengono informazioni di *network reachability*, cioè annunciano che certe reti sono raggiungibili attraverso certi router. I pacchetti di routing update sono inviati ai router vicini ad intervalli di tempo regolari e raggiungono tutti i router EGP. L'informazione in essi contenuta è utilizzata per costruire le tabelle di instradamento.

* Si noti che con il termine EGP si indica sia un generico protocollo di exterior routing, sia lo specifico protocollo oggetto del presente paragrafo.

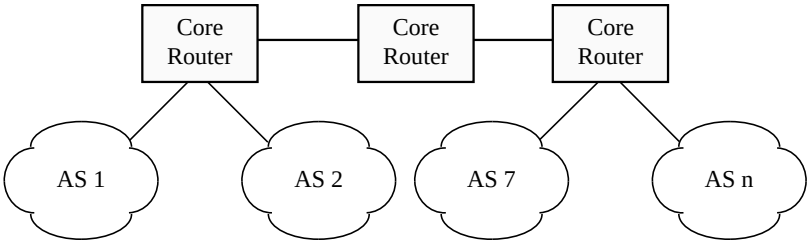

Fig. 16.11 - Esempio di utilizzo di EGP.

I limiti di EGP sono molti e gravi: EGP non ha una metrica associata alle linee e quindi basa le sue decisioni esclusivamente sulla raggiungibilità; EGP non ammette la presenza di magliature nella topologia, e tutti gli AS devono essere collegati in modo stellare ad un core system; i pacchetti di routing update possono essere molto grandi.

16.9.6 BGP

Il *Border Gateway Protocol* (BGP) è un exterior gateway protocol pensato per rimpiazzare il protocollo EGP ormai obsoleto. Il BGP è specificato per la prima volta dal RFC 1105 nel 1988, rispecificato come BGP-2 nel RFC 1163 nel 1990 e rispecificato ancora come BGP-3 nel RFC 1267 del 1991.

I router BGP comunicano tra loro utilizzando un livello di trasporto affidabile. Il BGP è un algoritmo di tipo distance vector, ma invece di trasmettere il costo di una destinazione, trasmette la sequenza di autonomous system da attraversare per raggiungere quella destinazione.

Ogni router calcola il suo instradamento preferito verso una data destinazione e lo comunica ai router BGP adiacenti tramite un distance vector. La politica con cui tale calcolo avviene è configurabile su ogni router BGP.

16.9.7 CIDR

Il *Classless Inter Domain Routing* (CIDR) è una modalità di propagazione dell'informazione di raggiungibilità (in gergo "annuncio") delle reti IP, che associa ad ogni indirizzo annunciato una netmask.

Il CIDR è specificato negli RFC 1517, 1518, 1519 e 1520.

Nei protocolli non CIDR la netmask viene derivata dalla classe dell'indirizzo. Con il CIDR questo non è vero ed indirizzi contigui possono essere propagati come fossero un indirizzo solo, operazione detta anche di clustering.

Ad esempio si supponga di volere annunciare le quattro seguenti reti di classe C: 199.9.4.0, 199.9.5.0, 199.9.6.0 e 199.9.7.0. Esse possono essere annuncate contemporaneamente tramite l'indirizzo 199.9.4.0 e la netmask 255.255.252.0.

Il CIDR riduce notevolmente la quantità di informazioni che devono essere propagate dagli EGP e anche dagli IGP e quindi ne aumenta l'efficienza. OSPF, Integrated IS-IS e la versione 4 del BGP realizzano il CIDR.

16.10 IL PROTOCOLLO TCP

Il TCP è un protocollo di transport di tipo connection-oriented che fornisce un servizio di tipo full-duplex (bidirezionale-contemporaneo), con acknowledge (conferma) e controllo di flusso.

Il TCP è utilizzato dalle applicazioni di rete che richiedono una trasmissione affidabile dell'informazione. Le applicazioni si connettono alle porte TCP e ad alcune applicazioni principali sono associate delle *well know port* cioè delle porte che hanno lo stesso numero su tutti i calcolatori (ad esempio all'applicazione telnet è associata la porta 23).

Il TCP segmenta e riassembla i dati secondo le sue necessità: ad esempio se un'applicazione fa cinque scritture su una porta TCP, l'applicazione destinataria può dover effettuare 10 letture per ottenere tutti i dati, oppure ottenerli tutti in una sola lettura.

Il TCP è un protocollo a sliding window (finestre) con meccanismi di time-out e ritrasmissione. La ricezione dei dati deve essere confermata dall'applicazione remota. La conferma può essere inserita in una PDU in transito nella direzione opposta, con una tecnica di piggybacking (si veda il paragrafo 13.2).

Come tutti i protocolli di tipo *sliding window*, TCP ha un massimo numero di dati in attesa di acknowledge. In TCP tale dimensione massima è specificata come numero di byte (window) e non come numero di segmenti TCP.

Il formato dell'header del pacchetto TCP è mostrato in figura 16.12, mentre un esempio di pacchetto TCP su IP è riportato in appendice B, paragrafo B.3.1.

I significati dei campi del pacchetto sono i seguenti:

- La *source port* e la *destination port* sono i numeri delle porte cui sono associati gli applicativi che usano la connessione TCP.

- Il *sequence number* è il numero di sequenza del primo byte del campo dati del messaggio. È utilizzato anche come identificatore della sliding window.
- Lo *acknowledge number* è il campo di acknowledge con tecnica di piggybacking della trasmissione nella direzione opposta. Contiene il numero di sequenza del primo byte che il mittente si aspetta di ricevere.
- Il campo *data offset* indica il numero di parole da 32 bit che compongono l'header TCP, variabile in funzione del campo option.
- Il campo *flag* contiene informazioni varie.
- Il campo *window* contiene la dimensione della receiving window del TCP mittente e quindi lo spazio disponibile nei buffer per il traffico entrante.
- Il campo *urgent pointer* punta al primo byte urgente nel pacchetto.

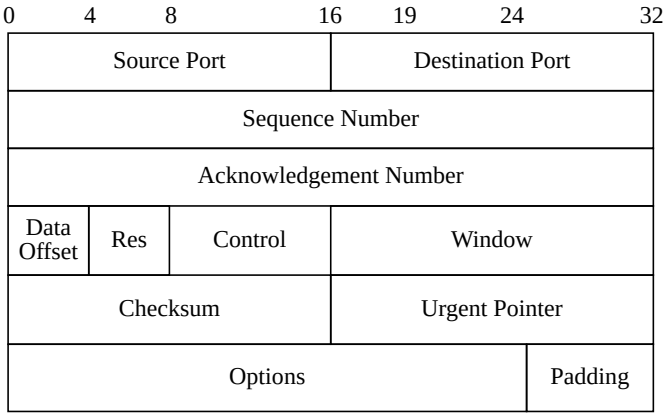

Fig. 16.12 - Header del pacchetto TCP.

16.11 IL PROTOCOLLO UDP

Lo *User Datagram Protocol* (UDP) è un protocollo di trasporto, alternativo a TCP, di tipo connectionless. UDP è un protocollo molto più semplice di TCP ed è utilizzato quando l'affidabilità di TCP non è richiesta.

La struttura del pacchetto UDP è mostrata in figura 16.13. I campi hanno significati simili a quelli di TCP, e la checksum è opzionale.

Un esempio di pacchetto UDP su IP è riportata in appendice B, paragrafo B.3.5.

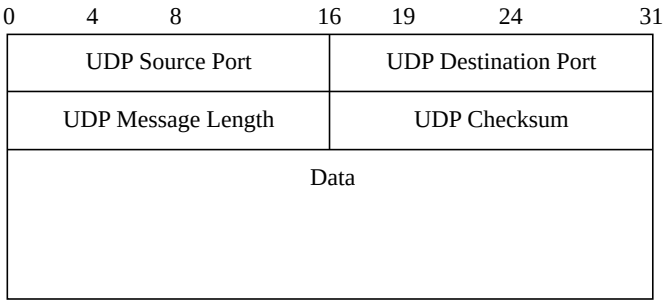

Fig. 16.13 - Il pacchetto UDP.

16.12 GLI APPLICATIVI

16.12.1 Telnet e rlogin

Telnet è un protocollo che permette ad un utente di collegarsi, tramite l'elaboratore locale, ad un qualsiasi altro elaboratore remoto connesso alla rete.

La connessione viene attivata facendo seguire al comando telnet il nome del calcolatore remoto o il suo indirizzo. Da quel momento in poi, tutti i caratteri battuti sulla tastiera sono inviati all'elaboratore remoto e le risposte da questo generate sono visualizzate sullo schermo locale. Il calcolatore locale è reso trasparente dal programma telnet e si opera come se si fosse direttamente connessi all'elaboratore remoto. Quando ci si scollega dall'elaboratore remoto, il programma telnet termina e ci si trova nuovamente a dialogare con il sistema operativo dell'elaboratore locale.

Normalmente il programma telnet include degli emulatori per i terminali più diffusi (es. Digital VT100 e IBM 3270).

Telnet è specificato dagli RFC 854 e 855 ed un esempio di PDU telnet su TCP/IP è riportato in appendice B, paragrafo B.3.1.

Alternativamente al telnet è possibile utilizzare il comando *rlogin* che ha funzionalità analoghe.

16.12.2 FTP, RCP e TFTP

Il *File Transfer Protocol* (FTP) è un applicativo che permette ad un utente collegato ad un elaboratore di trasferire file da e verso un altro elaboratore. La sicurezza è gestita chiedendo all'utente di fornire uno username e una password validi

sull'elaboratore remoto. FTP gestisce anche la conversione automatica di file di testo tra elaboratori con codifiche dei caratteri diverse. FTP è specificato nel RFC 959.

RCP è un applicativo simile a FTP in cui variano i meccanismi di gestione della sicurezza.

TFTP (*Trivial FTP*) è una versione semplificata di FTP usata normalmente per downline loading di software e specificata nel RFC 1350. Un esempio di PDU TFTP su TCP/IP è riportato in appendice B, paragrafo B.3.6.

16.12.3 SMTP

Il *Simple Mail Transfer Protocol* (SMTP) è probabilmente l'applicativo più importante del TCP/IP. Esso permette di inviare posta elettronica agli utenti della rete. Ogni utente è identificato dalla sintassi *Utente@Elaboratore* e non è richiesta alcuna autorizzazione per poter inviare un messaggio di posta elettronica. Il procedimento di invio avviene in batch, riprovando più volte sino a quando l'elaboratore remoto non diventa raggiungibile. L'utente remoto viene avvisato dell'arrivo di un nuovo messaggio.

I principali RFC che si occupano di posta elettronica sono lo RFC 821 e lo RFC 822.

16.12.4 DNS

Il *Domain Name Server* (DNS) è una base di dati distribuita e replicata per gestire principalmente la corrispondenza tra nomi e indirizzi IP. Un esempio di PDU DNS su TCP/IP è riportato in appendice B, paragrafo B.3.7.

Il DNS è specificato negli RFC 1035, 883 e 882.

16.12.5 BOOTP

Il *Boot Protocol* (BOOTP) è un protocollo per il bootstrap via rete di stazioni diskless. Un esempio di PDU BOOTP su TCP/IP è riportato in appendice B, paragrafo B.3.2.

Il BOOTP è specificato nel RFC 951.

16.12.6 ISODE

L'*ISO Development Environment* (ISODE) è un ambiente di sviluppo per applicativi OSI su reti TCP/IP. Un esempio di PDU ISODE su TCP/IP è riportato in appendice B, paragrafo B.3.12.

16.12.7 RSH, REXEC e RWHO

Le applicazioni *rsh* e *rexec* permettono di richiedere che un file di comandi o un programma eseguibile siano eseguiti su un elaboratore remoto invece che sull'elaboratore locale.

L'applicazione *rwho* permette di verificare quali utenti siano connessi da un elaboratore remoto. Un esempio di RWHO PDU su TCP/IP è riportato in appendice B, paragrafo B.3.4.

16.12.8 NFS e Netbios

Il *Network File System* (NFS) è un applicativo di sistema che permette a più elaboratori client di condividere un file system, messo a disposizione da un elaboratore server. Il tipo di network file system più noto è NFS proposto dalla SUN Microsystems ed adottato su tutti gli elaboratori con sistema operativo Unix.

SUN/NFS permette di avere molti server sulla rete e ad ogni elaboratore di fungere contemporaneamente da server e da client, per porzioni diversi del file system. Si appoggia su XDR (*eXternal Data Representation*), un pacchetto con scopi simili al livello Presentation OSI, e questo su RPC (*Remote Procedural Call*) e quindi su UDP e IP. Un esempio di PDU NFS è riportato in appendice B, paragrafo B.3.5.

SUN/NFS richiede una gestione coordinata della sicurezza degli elaboratori coinvolti nel file system distribuito che normalmente è realizzata con l'applicazione di sistema *Yellow Pages* (YP). Un esempio di PDU YP su TCP/IP è riportato in appendice B, paragrafo B.3.10.

Un altro tipo di file system distribuito molto utilizzato in ambito personal computer si basa su Netbios ed è trattato negli RFC 1001 e 1002. Un esempio di PDU Netbios su TCP/IP è riportato in appendice B, paragrafo B.3.11.

16.12.9 SNMP

Il *Simple Network Management Protocol* (SNMP) è un protocollo per la gestione degli apparati, basato su UDP/IP. SNMP è stato progettato per inviare dati sullo stato della rete provenienti dagli apparati ad un centro di gestione che li interpreti in modo opportuno. Con SNMP è anche possibile modificare alcuni parametri degli apparati di rete.

16.12.10 X-Window

X-Window è un software di rete client-server che permette ad un programma client di visualizzare dati grafici del display di un altro elaboratore che funge da server grafico.

Nato nell'ambito del progetto MIT Athena, X-window si è diffuso su tutti gli elaboratori e su tutti i protocolli, tra cui anche TCP/IP. Un esempio di PDU X-window su TCP/IP è riportato in appendice B, paragrafo B.3.8.

16.12.11 NIR

I *Network Information Retrieval* (NIR) sono servizi di tipo ipertestuale, distribuiti, che permettono di accedere ad un'ampia quantità di informazioni in modo semplice, usando un'interfaccia "user friendly" e ignorando dove l'informazione si trovi.

Nati per consentire l'utilizzo della rete Internet anche agli utenti meno esperti, hanno subito avuto un grosso successo e la loro diffusione è stata rapidissima.

Tra questi ricordiamo WAIS, gopher, WWW (Word Wide Web), netfind e X.500. Quest'ultimo è un applicativo nato nel mondo OSI, ma portato in quello Internet tramite ISODE (si veda il paragrafo 16.12.6).

16.12.12 Servizi Multicast

Sono gli ultimi ad essere stati sviluppati nel mondo Internet. Si tratta di servizi di audio e video conferenza basati su TCP/IP che affrontano problematiche nuove, quali quelle della multimedialità su rete. Tra questi ricordiamo Internet Talk Radio, IETF TV e Multimedia, Multiprotocol World.

Nell'ambito di Internet è stata definita una sottorete logica per fornire servizi di video e audio conferenza detta MBONE (*Multicast backBONE*).

BIBLIOGRAFIA

[1] A. Tanenbaum, "Computer Networks," Second Edition, Prentice-Hall.

[2] Douglas E. Comer, "Internetworking with TCP/IP", Volume 1, Second Edition, Prentice-Hall.

[3] L. Hedrick, "Introduction to the Internet Protocol", Rutgers University, New Jersey (USA), 3 July 1993.

[4] T. Socolofsky, C. Kale, "RFC 1180: A TCP/IP Tutorial", January 1991.

[5] Cisco Systems, "Internetworking Technology Overview", Codice documento DOC-ITO13 78-1070-01, 1993.

[6] Cisco Systems, "Router Products Configuration and Reference", Cisco Systems DOC-R9.1, Menlo Park CA (USA), September 92.

[7] M. K. Johnson, "Implementation Detail of the Microsoft LAN Manager TCP/IP Protocol", Microsoft Technical Note, Volume X, Number Y, March 1992.

[8] J. Postel, "RFC 768: User Datagram Protocol", 08/28/1980.

[9] J. Postel, "RFC 791, Internet Protocol", 09/01/1981.

[10] J. Postel, "RFC 792: Internet Control Message Protocol", 09/01/1981.

[11] J. Postel, "RFC 793: Transmission Control Protocol", 09/01/1981.

[12] J. Postel, "RFC 821: Simple Mail Transfer Protocol", 08/01/1982.

[13] D. Crocker, "RFC 822: Standard for the format of ARPA Internet text messages", 08/13/1982.

[14] D. Plummer, "RFC 826: Ethernet Address Resolution Protocol: Or converting network protocol addresses to 48.bit Ethernet address for transmission on Ethernet hardware", 11/01/1982.

[15] J. Postel, J. Reynolds, "RFC 854: Telnet Protocol specification".

[16] J. Postel, J. Reynolds, "RFC 855: Telnet option specifications".

[17] P. Mockapetris, "RFC 882: Domain names: Concepts and facilities", 11/01/1983.

[18] P. Mockapetris, "RFC 883: Domain names: Implementation specification", 11/01/1983.

[19] International Telegraph and Telephone Co, D. Mills, "RFC 904: Exterior Gateway Protocol formal specification", 04/01/1984.

[20] J. Mogul, J. Postel, "RFC 950: Internet standard subnetting procedure", 08/01/1985.

[21] W. Croft, J. Gilmore, "RFC 951: Bootstrap Protocol", 09/01/1985.

[22] J. Postel, J. Reynolds, "RFC 959: File Transfer Protocol", 10/01/1985.

[23] Defense Advanced Research Projects Agency, End-to-End Services Task Force, Internet Activities Board, NetBIOS Working Group, "RFC 1001: Protocol standard for a NetBIOS service on a TCP/UDP transport: Concepts and methods", 03/01/1987.

[24] Defense Advanced Research Projects Agency, End-to-End Services Task Force, Internet Activities Board, NetBIOS Working Group, "RFC 1002: Protocol standard for a NetBIOS service on a TCP/UDP transport: Detailed specifications", 03/01/1987.

[25] P. Mockapetris, "RFC 1035: Domain names - implementation and specification", 11/01/1987.

[26] C. Hedrick, "RFC 1058, RIP: Routing Information Protocol", 06/01/1988.

[27] S. Deering, "RFC 1112: Host extensions for IP multicasting", 08/01/1989.

[28] K. Lougheed, Y. Rekhter, "RFC 1267: A Border Gateway Protocol 3 (BGP-3)", 10/25/1991.

[29] K. Sollins, "RFC 1350: The TFTP protocol (revision 2)", 07/10/1992.

[30] G. Malkin, "RFC 1388: RIP Version 2 Carrying Additional Information", 01/06/1993.

[31] J. Moy, "RFC 1583: OSPF Version 2", 03/23/1994.

[32] Hinden, "RFC 1517: Applicability Statement for the Implementation of Classless Inter-Domain Routing (CIDR)", 09/24/1993.

[33] Rekhter, T. Li, "RFC 1518: An Architecture for IP Address Allocation with CIDR", 09/24/1993.

[34] Fuller, T. Li, J. Yu, K. Varadhan, "RFC 1519: Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", 09/24/1993.

[35] Rekhter, C. Topolcic, "RFC 1520: Exchanging Routing Information Across Provider Boundaries in the CIDR Environment", 09/24/1993.