

# Algoritmi di routing e ICMP

## 1 Algoritmi di routing

Gli **algoritmi di routing** determinano come instradare ogni datagramma IP dal mittente al destinatario, in base all'indirizzo di destinazione contenuto nel datagramma stesso. Un algoritmo di routing può essere di tipo:

- **Senza tabella:**
  - **random:** ogni nodo (router) inoltra i datagrammi ricevuti a un nodo successivo scelto casualmente;
  - **flooding:** ogni nodo inoltra i datagrammi a tutti i nodi vicini, “inondando” la rete;
  - **source routing:** il percorso che un datagramma deve compiere è scelto direttamente dalla sorgente del datagramma, cioè dal mittente.
- **Gerarchico.**
- **Con tabella:** ogni router mantiene una tabella che consulta per determinare come inoltrare i datagrammi. Un algoritmo con tabella può poi essere:
  - **statico:** la tabella rimane statica;
  - **dinamico:** la tabella viene aggiornata dinamicamente.

Gli algoritmi dinamici si distinguono ulteriormente in **distance vector** e **link state**.

In Internet si usano algoritmi con tabella di tipo dinamico, tra cui principalmente RIP, OSPF e BGP, che verranno illustrati in seguito.

## 2 Tabella di routing

Come appena detto, quando si usa un algoritmo di routing con tabella, ogni router consulta la propria tabella di routing per determinare il *next hop*, cioè per scegliere verso quale porta di uscita (e dunque verso quale nodo successivo) inoltrare ciascun datagramma ricevuto.

Una tabella di routing associa ogni rete di destinazione conosciuta (che può essere direttamente connessa al router, oppure remota, cioè raggiungibile tramite altri router) alle porte di uscita che permettono di raggiungerla, memorizzando anche i costi di ciascun percorso, che possono essere calcolati secondo varie *metriche*. Nella tabella deve sempre essere presente anche una *default route*, che indica la porta di uscita su cui instradare i datagrammi indirizzati a reti che il router non conosce. Un esempio di tabella per un router con porte di uscita chiamate L1, L2, ecc. potrebbe allora essere:

Indirizzo rete destinazione	Costo	Next hop
193.204.59.0/24	...	L1
...	...	...
Default route	...	L2

### 2.1 Metriche per il costo dei percorsi

Per stabilire il costo dei percorsi di routing, e quindi poter scegliere il percorso migliore per un datagramma, esistono varie metriche, che possono essere usate singolarmente o in combinazione. Alcune delle principali sono:

- *bandwidth*: la banda (capacità) dei link sul percorso;
- *delay* (ritardo): il tempo che un datagramma impiega ad attraversare il percorso;
- *load*: il carico di lavoro degli elementi della rete (ad esempio router e link) che compongono il percorso;
- *reliability*: l'affidabilità (in genere il tasso di errore) dei link;
- *hop count*: il numero di salti che un datagramma deve compiere (cioè di router che esso deve attraversare) per giungere a destinazione;
- *cost*: un valore arbitrario, basato ad esempio sul costo economico della banda di un link.

### 3 Instradamento gerarchico in Internet

Nella rete Internet, un insieme di router amministrati dallo stesso gestore è definito **Autonomous System (AS)**.

All'interno e all'esterno degli Autonomous System vengono usati algoritmi di routing differenti:

- si dicono **intra Autonomous System** gli algoritmi usati all'interno di un singolo Autonomous System, come ad esempio RIP e OSPF;
- l'algoritmo BGP, usato invece per il routing tra diversi Autonomous System, è detto **inter Autonomous System**.

### 4 Distance vector

In un algoritmo di routing di tipo distance vector, ogni router invia ai propri vicini, periodicamente e/o quando c'è un cambiamento topologico, la propria tabella di routing, che prende il nome di **distance vector**.

Ogni router *non* conosce l'intera topologia di rete, e sceglie la strada migliore per ciascun datagramma da inoltrare in base solo ai distance vector che ha ricevuto fino a quel momento. La ricerca del percorso migliore si basa sull'**algoritmo di Bellman-Ford**.

#### 4.1 Routing loop e counting to infinity

Siccome i router che usano un algoritmo distance vector non conoscono l'intera topologia della rete, e in particolare non sanno quali siano esattamente i nodi che compongono un determinato percorso, può succedere, solitamente in seguito a cambiamenti topologici, che si formino dei **routing loop**, cioè percorsi di routing circolari, nei quali i datagrammi continuano a girare (finché il loro TTL non raggiunge 0), senza mai arrivare a destinazione.

Un altro problema correlato è il "**counting to infinity**": in sintesi, quando una rete diventa non più raggiungibile, tale informazione non si propaga immediatamente a tutti i router, e ciò causa la formazione di percorsi di routing circolari, i cui costi nelle tabelle di routing aumentano progressivamente all'infinito.

Esistono varie contromisure per evitare i loop e il counting to infinity:

- definire un valore massimo di costo, per evitare che questo possa aumentare all'infinito;
- **split horizon**: le informazioni su un percorso non vengono mai inviate "all'indietro", cioè verso la porta dalla quale tali informazioni sono state ricevute;

- **split horizon with poisoned reverse**: se la porta da cui viene appreso un percorso è la stessa che costituisce il next hop del percorso, il percorso viene annunciato su tale porta con un costo “infinito”;
- **trigger update**: gli aggiornamenti delle tabelle di routing sono inviati non solo allo scadere di un timer prefissato, ma anche a ogni cambiamento topologico;
- una rete che risulta irraggiungibile è considerata tale, ovvero la sua entry viene eliminata dalle tabelle di routing, solo allo scadere di un **hold-down timer**.

## 5 RIP

**RIP, Routing Information Protocol**, è un protocollo di routing di tipo *distance vector*, adatto a reti di piccole dimensioni. Esso usa come metrica l'*hop count*, con un valore massimo di 15, mentre 16 indica una rete irraggiungibile (per evitare il counting to infinity).

Ciascun router esegue un *routing update* (cioè comunica lo stato della propria tabella di routing agli altri router direttamente raggiungibili) ogni 30 s, e a ogni cambiamento topologico. L'*hold-down timer* è pari a 120 s.

Esistono due versioni di questo protocollo: RIPv1 e RIPv2.

### 5.1 Messaggi RIP

Per lo scambio di messaggi di routing RIP si usa il protocollo di trasporto UDP.

Ogni messaggio contiene fino a 25 descrittori di percorso, ciascuno dei quali indica sostanzialmente l'indirizzo IP di destinazione e l'*hop count* del percorso. Nel RIPv2 vengono aggiunti ai descrittori di percorso anche:

- la *subnet mask* associata all'indirizzo di destinazione;
- la possibilità di specificare un *next hop* diverso dal router che ha mandato il messaggio;
- un campo *route tag* per distinguere percorsi “interni” (gestiti da RIP) ed “esterni” (ottenuti da altri protocolli di routing).

Un'altra caratteristica introdotta da RIPv2 è la possibilità di autenticare i messaggi di routing.

## 6 OSPF

**OSPF, Open Shortest Path First**, è un protocollo di routing di tipo *link state*, cioè nel quale ogni nodo conosce l'intera topologia di rete. Esso calcola il costo dei percorsi in base a una combinazione di metriche, (ad esempio bandwidth, reliability, delay, ecc.), e determina il percorso migliore usando l'**algoritmo di Dijkstra**. Oggi OSPF è più usato rispetto a RIP, nonostante abbia lo svantaggio di una maggiore complessità.

Un Autonomous System in cui si usa OSPF viene suddiviso in **aree**, e OSPF opera in modo sostanzialmente indipendente in ciascuna di queste aree. Così, limitando la dimensione delle aree e i messaggi tra di esse, OSPF risulta altamente scalabile, cioè in grado di gestire reti molto grandi.

In base alla loro posizione all'interno della suddivisione in aree, i router assumono tre ruoli diversi:

- quelli aventi solo interfacce nella stessa area sono detti **internal router**, e si occupano solo dello stato dell'instradamento all'interno della singola area;
- quelli tra due aree sono detti **area border router**, e devono avere almeno un'interfaccia verso la cosiddetta **backbone area**, un'area "speciale" tramite la quale vengono messe in comunicazione aree diverse di uno stesso Autonomous System;
- quelli situati al confine dell'intero Autonomous System (cioè situati in un'area, ma aventi una o più interfacce verso Autonomous System diversi) sono detti **boundary router**.

### 6.1 Link State Advertisement

Periodicamente, e a ogni cambiamento, ciascun nodo invia lo stato dei suoi collegamenti, sotto forma di degli annunci chiamati **Link State Advertisement (LSA)**. Esistono cinque tipi di LSA (router LSA, network LSA, summary LSA, Autonomous System boundary router LSA e Autonomous System external LSA), ciascuno dei quali contiene informazioni diverse; ad esempio, i *summary LSA* sono generati dagli area border router per comunicare, "riassumere" ai router interni di un'area la raggiungibilità di altre aree.

Ciascun router aggiunge alla propria base di dati i nuovi LSA che riceve, e rimuove invece gli LSA "vecchi" (in genere, un LSA è considerato vecchio dopo un'ora).

L'invio degli LSA avviene tramite degli appositi pacchetti, chiamati **Link State Packet (LSP)**, ciascuno dei quali può contenere più LSA. Gli LSP sono direttamente incapsulati nei pacchetti IP (senza l'uso di un protocollo di trasporto, a differenza ad esempio del RIP, che usava UDP), e supportano vari meccanismi di autenticazione. Ci sono tre tipi di LSP, che hanno funzioni diverse:

- *link state request*: è una richiesta di invio di uno o più LSA;

- *link state update*: contiene un insieme di LSA;
- *link state acknowledgement*: conferma l'avvenuta ricezione di un insieme di LSA.

## 7 BGP

Il protocollo di routing usato tra router di frontiera di diversi Autonomous System, per indicare la raggiungibilità di altri Autonomous System, è **BGP**, **Border Gateway Protocol**. Esso è un algoritmo di tipo **path vector**: i router si scambiano tabelle di routing che contengono, per ogni rete raggiungibile, la sequenza di Autonomous System da attraversare per arrivarci. Ogni Autonomous System ha un identificativo univoco.

I messaggi BGP vengono scambiati usando a livello di trasporto il protocollo TCP. Quando un router di frontiera di un Autonomous System è disponibile allo scambio di tabelle di routing, invia dei messaggi di *open*; questa richiesta di scambio può essere accettata o rifiutata dagli altri router di frontiera, sulla base degli accordi stipulati.

## 8 ICMP

**ICMP**, **Internet Control Message Protocol**, è utilizzato per segnalare informazioni di controllo (come ad esempio l'occorrenza di una situazione di errore) relative al livello rete.

I messaggi ICMP sono trasportati direttamente come payload dei datagrammi IP (senza un protocollo di trasporto). Essi possono essere di vari tipi, e il tipo di un messaggio è indicato da due campi numerici presenti al suo interno: *ICMP type* (tipo) e *code* (codice).

Alcuni esempi di applicazione di ICMP sono gli strumenti *ping* e *traceroute*:

- Ping permette di determinare se un host è raggiungibile, inviando a esso un messaggio ICMP di tipo *echo request* e aspettando che questo risponda con un'*echo reply*.
- Traceroute permette di individuare i singoli hop del percorso verso un host, inviando a tale host una sequenza di normali datagrammi IP con  $TTL = 1, 2, 3, \dots$ . Il TTL di ciascun datagramma della sequenza "scade" (raggiunge zero) all'hop successivo del percorso, e ogni router presso cui scade il TTL rimanda al mittente un messaggio ICMP *TTL expired* (che contiene informazioni su tale router).