

Il protocollo BGP e le rotte Internet

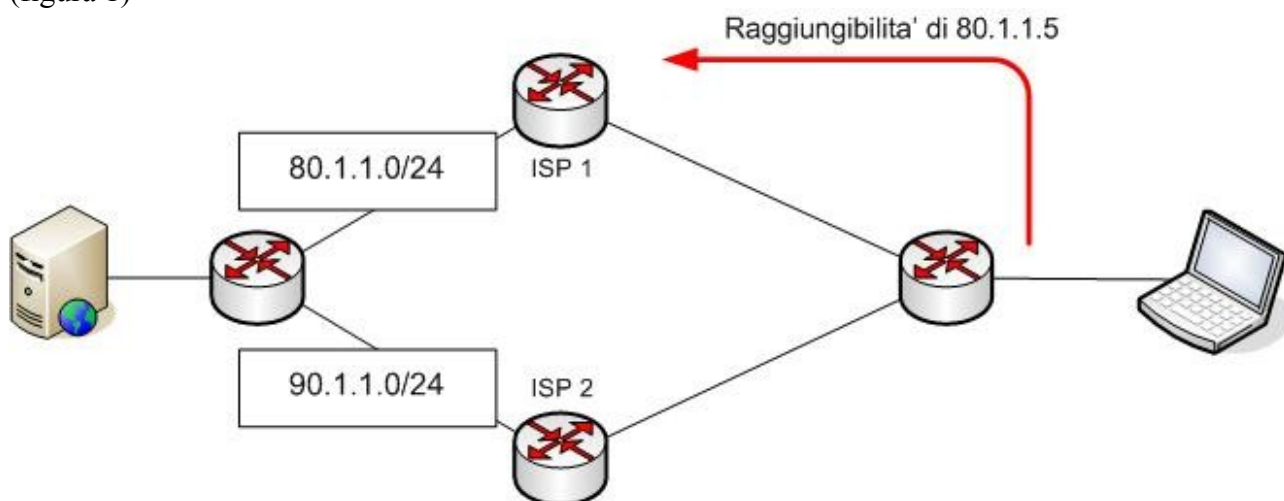
Introduzione al protocollo

Il protocollo BGP, Border Gateway Protocol è un protocollo di routing dinamico usato per connettere tra loro più router che appartengono a “sistemi autonomi” (AS) diversi. È quindi un protocollo di routing inter-AS, classificato come protocollo di “Exterior Gateway”, ovvero utilizzato per inviare le proprie route a organizzazioni esterne (e per ricevere le route di organizzazioni esterne). Differisce in questo dai protocolli di “Interior Gateway”, utilizzati per lo scambio di rotte all'interno della stessa organizzazione. (esempio: RIPv2, OSPF, ...)

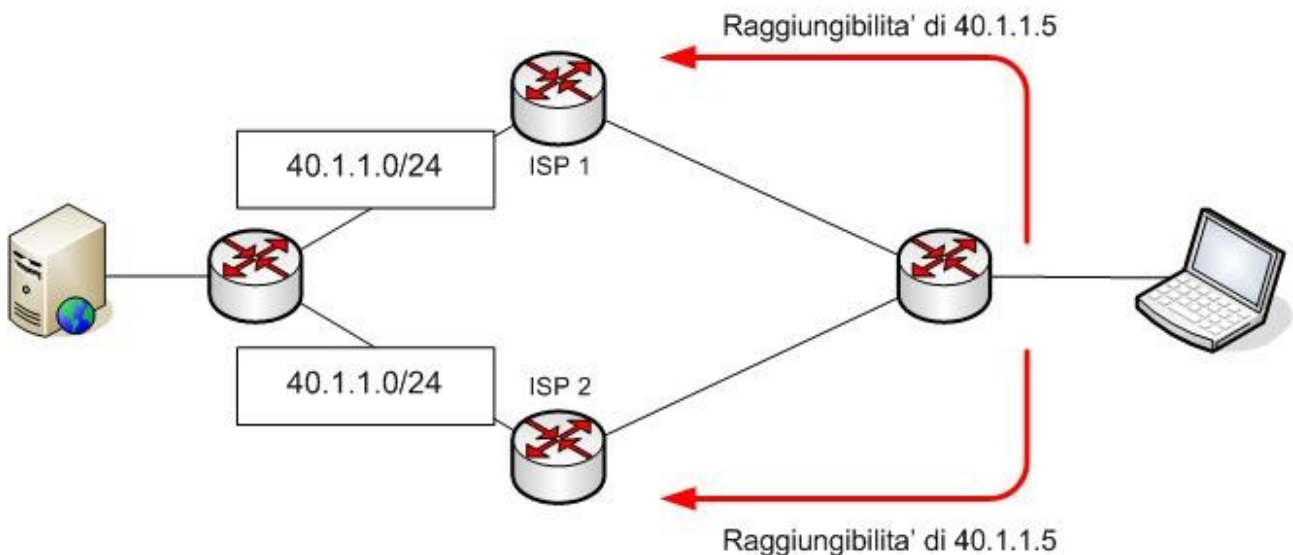
Il BGP è il protocollo di routing per antonomasia, sul quale è basato il funzionamento di internet. Esso funziona attraverso la gestione di una tabella di prefissi (reti IP), che forniscono informazioni sulla raggiungibilità delle diverse reti tra più sistemi autonomi. Si tratta di un protocollo a indicazione di percorso (path vector) che non usa metriche di carattere tecnico (ad esempio non considera le ampiezze di banda) ma prende le decisioni di instradamento basandosi su specifiche politiche.

La motivazione che spinge un'organizzazione all'utilizzo del protocollo BGP per scambiare le proprie route in internet è la necessità di avere una rete multihomed. Per rete multihomed si intende una rete che acquista transito internet da più provider di livello superiore (o anche più link con lo stesso provider) e che deve garantire la raggiungibilità dall'esterno dei propri servizi, tramite un set di indirizzi IP pubblici.

Infatti, in una situazione standard senza l'uso del protocollo BGP, ogni link verso internet sarà dotato della sua classe di indirizzamento IP, non condivisibile con altri link. Risulta evidente quindi che pubblicando un servizio (esempio: sito web) sull'indirizzo IP raggiungibile direttamente dallo specifico link, alla caduta del link stesso quel servizio risulterà irraggiungibile dalle altre reti. (figura 1)



Invece, utilizzando il protocollo BGP per propagare in internet la propria classe di indirizzi IP, gli indirizzi stessi saranno raggiungibili tramite qualsiasi link verso il quale siano state propagate le route. (figura 2)



Nel protocollo BGP gli AS adiacenti, detti peer, stabiliscono delle sessioni di scambio delle rotte tramite appropriata configurazione dei router. Le sessioni utilizzano il protocollo tcp per lo scambio di informazioni.

È importante notare che tutti i router all'interno di un AS che partecipano all'instradamento via BGP devono essere collegati secondo una topologia a maglie completamente connesse (full-mesh): ogni router deve essere peer di tutti gli altri. La cosa pone problematiche relative alla scalabilità, per questo sono implementabili soluzioni quali route-reflections e confederations. L'argomento non verrà qui trattato, in quanto necessario solo in AS di grandi dimensioni.

Come detto prima, BGP è un protocollo basato su path-vector, il che significa che il miglior percorso per raggiungere una determinata destinazione si basa principalmente sul numero di AS tra sorgente e destinazione.

Ad esempio, nella scelta tra i seguenti due percorsi, verrà scelto il primo (anche se con minore ampiezza di banda):

- AS100 → AS 200 → AS 300 → AS 400
- AS100 → AS 105 → AS 106 → AS 107 → AS 400

è tuttavia possibile utilizzare alcuni “tricks” per determinare il percorso preferito. Un esempio è il cosiddetto “AS prepending”, che consiste sostanzialmente nell'iniettare nel percorso AS più volte il proprio numero AS.

Tornando ai percorsi di prima, se il link numero 2 ha maggior capacità trasmissiva, posso “forzare” il transito attraverso di esso annunciando più volte il mio AS (400). Si noti quindi che il link 1 è più lungo del 2, quindi verrà utilizzato come seconda scelta.

- AS100 → AS 200 → AS 300 → AS 400 → AS 400 → AS 400
- AS100 → AS 105 → AS 106 → AS 107 → AS 400

È importante notare che proprio il fatto che il protocollo non badi ai canali trasmissivi imponga l'uso di canali quanto più simili tra loro.

Si aggiunge che il protocollo BGP è valido sia per il routing Ipv4 che per il routing Ipv6.

Cosa si può fare con BGP

Si va dalla semplice rete multihomed verso due provider alle reti che fanno peering con altri partner o reti di clienti, vuoi tramite un peering privato vuoi tramite un peering pubblico presso un IXP, Internet Exchange Point (es: MIX-IT).

Ovviamente, proprio grazie alla selezione dei path link vector, ogni cliente che deve raggiungere i servizi erogati utilizzerà il percorso con lunghezza minore.

Un esempio lampante:

- il provider A fornisce servizi internet e acquista transito internet da B e C;
- il provider A acquisisce il cliente Z, che si collega ad internet tramite D e ha bisogno di tempi di latenza bassissimi verso il servizio di A.

È evidente in questo caso che un peering diretto con D accorcerebbe notevolmente il percorso dei pacchetti tra A e Z.

Cosa non si può fare con BGP

Proprio per il fatto che il protocollo è link-vector, non si possono in alcun modo considerare i media fisici su cui transiteranno i pacchetti.

Criticità e requisiti

Per propagare in internet le proprie route via BGP è necessario essere Autonomous System riconosciuto (dal RIPE nel caso europeo) e avere una propria classe di indirizzi IP pubblici.

L'allocazione minore acquistabile è di una /24, ovvero 256 indirizzi IP, detta "PI" (Provider-Independent)

Tuttavia è molto difficile ottenere l'allocazione di una classe di questo tipo, in quanto bisogna rispondere a determinati requisiti difficili da completare.

È molto più facile ottenere allocazione di reti più grosse, come ad esempio una /22 (4 reti /24), dette "PA" (Provider-Allocatable). Questo perché nei vari router in internet si preferisce evitare di propagare route che hanno prefisso maggiore o uguale di 24, in quanto frammenterebbero troppo la routing table dei vari router.

È poi estremamente facile ottenere l'allocazione di una classe /32 di indirizzi Ipv6 (che è lungo 128 bit: si avrebbero quindi un numero enorme di sottoreti allocabili).

Da notare che è possibile propagare rotte Ipv6 solo se i transit provider supportano Ipv6 stesso.