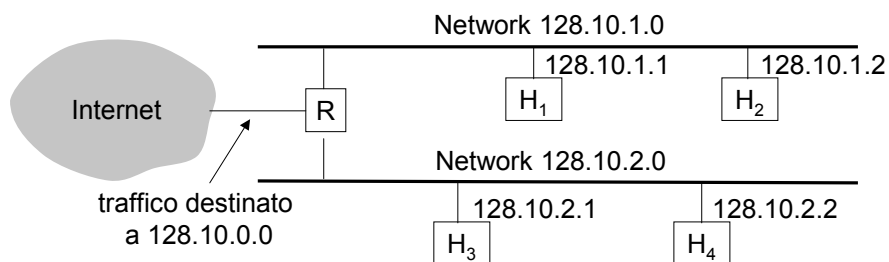


## Routing IP seconda parte

### Indirizzamento di subnet e supernet

- Schema di indirizzamento IP originale:
  - ad ogni rete fisica è assegnato un unico indirizzo di rete
  - ogni host appartenente alla rete ha come netid l'indirizzo di rete
- Esaurimento dello spazio di indirizzamento :
  - minimizzare il numero di indirizzi di rete: lo stesso netid può essere condiviso da più reti fisiche (**subnet addressing** per indirizzi di classe B)



## Indirizzamento di subnet e supernet (2)

- Il subnet addressing modifica l'interpretazione degli indirizzi IP:

- l'indirizzo IP è composto da una porzione di **rete** ed una **locale**

rete	locale	
rete	rete fisica	host

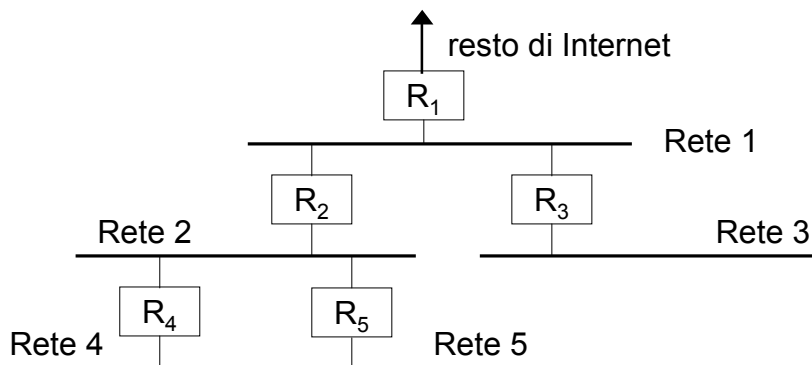
- Risultato: indirizzamento gerarchico → routing gerarchico

- Routing gerarchico: i router esterni usano i primi due byte dell'indirizzo IP per il routing, mentre il router della rete locale usa il terzo byte dell'indirizzo IP

## Indirizzamento di subnet e supernet (3)

- Esempio di rete con cinque reti fisiche suddivise in tre livelli:

- rete di classe B (16 bit per parte locale)
- 5 reti fisiche: occorrono 3 bit (essendo  $5 < 2^3 = 8$ ) per identificarle
- ad ognuna delle 5 reti fisiche è possibile attaccare  $2^{13} = 8192$  host



## Indirizzamento di subnet e supernet (4)

- Implementazione delle subnet usando le maschere:
  - **maschera di subnet** formata da 32 bit per ciascuna rete che usa il subnet addressing
  - nella maschera i bit settati ad 1 corrispondono alla parte di rete, quelli settati a 0 alla parte locale
  - bit della subnet = indirizzo IP  $\cap$  maschera (AND bit a bit)
- Esempio di rete con cinque reti fisiche suddivise in tre livelli:
  - maschera = 11111111 11111111 11100000 00000000

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.26

## Indirizzamento di subnet e supernet (5)

- Una tabella di instradamento convenzionale contiene voci della seguente forma:

(indirizzo di rete, indirizzo del next hop router)
- Con le sottoreti non è possibile individuare dal solo indirizzo IP quali bit corrispondono alla rete e quali all'host  
 $\Rightarrow$  l'algoritmo modificato utilizzato con le sottoreti riporta ulteriori informazioni nella tabella d'instradamento. Ogni voce della tabella contiene un campo aggiuntivo che specifica la maschera di sottorete usata per la rete di quella voce:

(maschera di sottorete, indirizzo di rete, indirizzo del next hop router )
- L'algoritmo modificato utilizza la maschera per estrarre i bit dell'indirizzo di destinazione da confrontare con la voce della tabella

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.27

## Indirizzamento di subnet e supernet (6)

- Esaurimento dello spazio di indirizzamento di classe B → approccio opposto al subnet addressing: una singola organizzazione può utilizzare più indirizzi di rete per la sua rete (**supernet addressing**)

- un blocco di indirizzi di classe C contiguo viene assegnato ad una singola organizzazione anziché un solo indirizzo di classe C
- usato dagli Internet Service Provider

- Problema: aumenta il numero di ingressi nella tabella di routing → si usa il meccanismo *Classless Inter-Domain Routing* (CIDR) in cui:

(network address, count)

- network address è il più piccolo indirizzo nel blocco
- count è il numero di blocchi di indirizzi di classe C contigui

## Sistemi Autonomi

- Cosa è un AS?
  - Un insieme di router sotto un'unica amministrazione
  - Usa un *interior gateway protocol (IGP)* e metriche comuni per instradare pacchetti all'interno dell'AS
  - Usa un *exterior gateway protocol (EGP)* per instradare pacchetti verso altri AS
- Gli AS possono usare molti IGP e metriche ma, per gli altri AS, appaiono come un solo AS

## Routing gerarchico

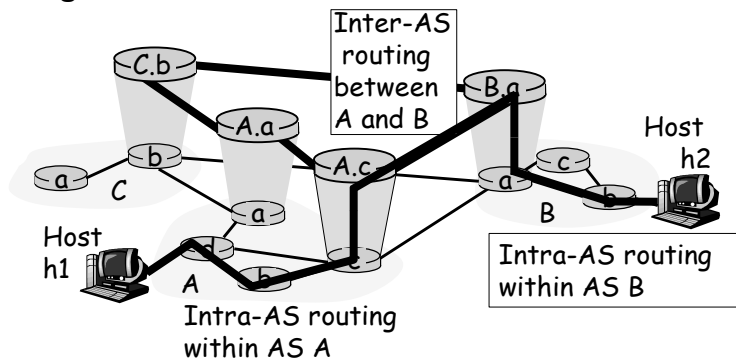
- L'insieme dei router non è omogeneo e tutti i router non eseguono lo stesso algoritmo per due ragioni:
  - **scalabilità**: all'aumentare del numero di router, l'overhead degli algoritmi di routing diviene proibitivo → occorre ridurre la complessità della computazione del cammino
  - **autonomia amministrativa**: un'organizzazione dovrebbe scegliere autonomamente come amministrare i propri router, rimanendo tuttavia in grado di connettere la propria rete alle reti esterne
- Soluzione: i router vengono aggregati in "regioni" o **sistemi autonomi (AS)**
  - I router appartenenti allo stesso AS eseguono lo stesso algoritmo di routing (LS o DV) detto *algoritmo di **routing interno al sistema autonomo***

## Routing gerarchico (2)

- Ogni router conosce i dettagli su come instradare i pacchetti verso destinazioni all'interno del proprio AS, ma non sa nulla sulla struttura interna degli altri AS.
- Crescendo il rapporto fra numero di AS e numero di router, si incrementa anche lo spazio risparmiato nelle tabelle di routing
- Questo però non è gratuito: i cammini hanno infatti una lunghezza maggiore

## Routing gerarchico (3)

- In un AS, alcuni router hanno la responsabilità del routing di pacchetti destinati all'esterno dell'AS
  - tali router sono detti **gateway router**
- I gateway devono sapere come effettuare il routing tra di loro
  - **routing esterno** al sistema autonomo



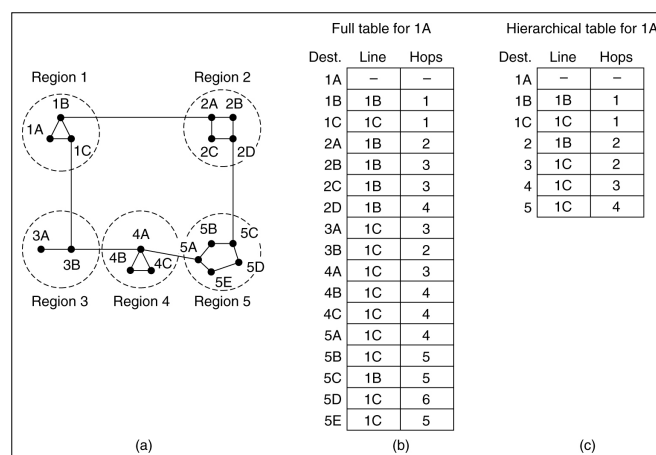
M. Colajanni, F.Mazzoni, 2003, Routing IP

5.32

## Routing gerarchico (4)

Gerarchia a 2 livelli con 5 regioni

- Una entry per ogni router dell'AS
- Una entry per ogni altro AS (ogni AS è visto come 1 router)



M. Colajanni, F.Mazzoni, 2003, Routing IP

5.33

## Routing in Internet

- Principali protocolli di routing **intra-AS**
  - **Routing Information Protocol (RIP)**
    - Algoritmo di routing distribuito (distance vector protocol)
  - **Open Shortest Path First (OSPF)**
    - Algoritmo di routing centralizzato (link state protocol)
    - successore di RIP
- Principale protocollo di routing **inter-AS**
  - **Border Gateway Protocol (BGP)**
    - Algoritmo di routing distribuito (distance vector protocol)
    - è lo standard *de facto* per il routing tra AS

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.34

## Routing Information Protocol (RIP)

- Uno dei primi protocolli di Internet per l'instradamento intra-AS
- Descritto negli RFC 1058 (ver. 1) e RFC 1723 (ver. 2)
- Protocollo distance-vector
- Costo metrico: conteggio degli hop  $\Rightarrow$  ciascun link ha costo 1
- Conteggio hop non dà sempre ottimi risultati: un percorso di 3 hop su una Ethernet può essere sostanzialmente più veloce di 2 hop su link satellitare (soluzione: si possono indicare più hop se la connessione è lenta)

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.35

## Routing Information Protocol (RIP) (2)

- Massimo costo di un percorso = 15  $\Rightarrow$  uso di RIP limitato agli AS che hanno un diametro  $< 15$  hop
- Tabelle di instradamento scambiate circa ogni 30 secondi usando il RIP response message o RIP advertisement
- Se un router non ha notizie da un suo vicino almeno una volta ogni 180 secondi  $\Rightarrow$  quel vicino è considerato non più raggiungibile (o il vicino è “morto” o il link ha cessato di funzionare)
- Un router può richiedere informazioni sul costo dei suoi vicini verso una data destinazione utilizzando il RIP request message

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.36

## Open Shortest Path First (OSPF)

- Il fatto di renderlo uno standard aperto, che chiunque può implementare senza pagare cessioni di licenza, ha incoraggiato molti fornitori a supportarlo.
- Include l'instradamento basato sul tipo di servizio. I gestori possono registrare più percorsi verso una destinazione data, uno per ciascuna priorità o tipo di servizio
- Fornisce il bilanciamento del carico: se un gestore specifica più instradamenti verso una specifica destinazione allo stesso costo, OSPF distribuisce il traffico in ugual modo lungo i percorsi

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.37



## Open Shortest Path First (OSPF) (2)

- Gestisce l'autenticazione dei router
- Gestisce le gerarchie all'interno del dominio di un singolo router
- Supporta varie metriche: distanza, ritardo, costo, ...
- Protocollo link state

M. Colajanni, F. Mazzone, 2003, Routing IP

5.38

## Open Shortest Path First (OSPF) (3) Intestazione

- Ogni messaggio comincia con un'intestazione fissa di 24 byte

0	8	16	31
VERSION (1)	TYPE	MESSAGE LENGHT	
SOURE ROUTER IP ADDRESS			
AREA ID			
CHECKSUM		AUTHENTICATION TYPE	
AUTHENTICATION (OTTETTI 0-3)			
AUTHENTICATION (OTTETTI 4-7)			

M. Colajanni, F. Mazzone, 2003, Routing IP

5.39

## Open Shortest Path First (OSPF) (4)

- Type prevede 5 tipi di messaggio
  1. Hello (usato per verificare la raggiungibilità)
  2. Descrizione del database (topologia)
  3. Richiesta dello stato del collegamento
  4. Aggiornamento dello stato del collegamento
  5. Conferma di ricezione dello stato di collegamento
- Source router IP address è l'indirizzo del mittente
- Area ID fornisce un numero di identificazione a 32 bit per l'area
- Authentication type specifica quale schema di autenticazione viene usato
  - 0  $\Rightarrow$  assenza di autenticazione
  - 1  $\Rightarrow$  uso di password

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.40

## Open Shortest Path First (OSPF) (5) Messaggio HELLO

0	8	16	24	31
OSPF HEADER WITH TYPE = 1				
NETWORK MASK				
DEAD TIMER		HELLO INTER	GWAY PRIO	
DESIGNATED ROUTER				
BACKUP DESIGNATED ROUTER				
NEIGHBOR <sub>1</sub> IP ADDRESS				
NEIGHBOR <sub>2</sub> IP ADDRESS				
...				
NEIGHBOR <sub>N</sub> IP ADDRESS				

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.41

## Open Shortest Path First (OSPF) (6)

- OSPF invia periodicamente messaggi HELLO su ogni collegamento per stabilire e verificare la raggiungibilità dei vicini
- Network mask contiene una maschera per la rete su cui il msg è stato inviato
- Dead timer fornisce il tempo in secondi trascorso il quale un vicino che non risponde è considerato inattivo
- Hello inter è il periodo, in secondi, che intercorre fra i msg HELLO
- Gway prio è la priorità del router mittente e viene usata nella scelta di un router di backup designato

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.42

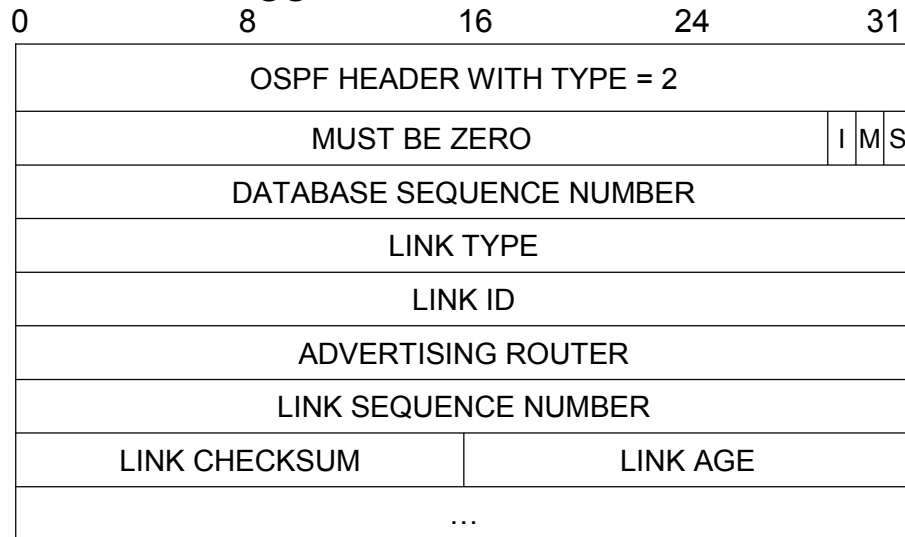
## Open Shortest Path First (OSPF) (7)

- Designated router e backup designated router contengono gli indirizzi IP del router e del router di backup designati per la rete su cui è inviato il msg
- Neighbor<sub>i</sub> IP address fornisce l'indirizzo IP di ogni vicino da cui il mittente ha recentemente ricevuto msg HELLO

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.43

## Open Shortest Path First (OSPF) (8) Messaggio di descrizione del DB



M. Colajanni, F. Mazzoni, 2003, Routing IP

5.44

## Open Shortest Path First (OSPF) (9)

- Per inizializzare il DB contenente le info sulla topologia della rete i router si scambiano msg di descrizione del DB OSPF
- Nello scambio uno funge da router principale (master), mentre l'altro è subordinato (slave)
- Poiché può essere molto esteso, il database della topologia va diviso in più msg usando i bit I e M: il bit I è posto a 1 nel msg iniziale; il bit M è posto a 1 se seguono altri msg. Il bit S indica se il msg è stato spedito da un master (1) o da uno slave (0)
- Il database sequence number numera in sequenza i msg in modo che il destinatario possa dire se ne manca uno

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.45

## Open Shortest Path First (OSPF) (10)

- I campi da link type a link age descrivono un collegamento nella topologia di rete e vengono ripetuti per ciascun collegamento
- Link type è il tipo di collegamento (del router, ad un router di confine, etc...)
- Link ID fornisce un'identificazione per il collegamento (ad es. l'IP di un router)
- Advertising router specifica l'indirizzo del router che annuncia il collegamento

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.46

## Open Shortest Path First (OSPF) (11)

- Link sequence number contiene un intero generato dal router per assicurare che i msg non vadano persi o siano ricevuti fuori ordine
- Link checksum fornisce un'ulteriore garanzia che le informazioni sul collegamento non sono state danneggiate
- Link age aiuta ad ordinare i msg, dando l'ora in secondi a partire da quando il collegamento è stato creato

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.47

## Open Shortest Path First (OSPF) (12)

0	16	31
OSPF HEADER WITH TYPE = 3		
LINK TYPE		
LINK ID		
ADVERTISING ROUTER		
...		

M. Colajanni, F.Mazzoni, 2003, Routing IP

5.48

## Open Shortest Path First (OSPF) (13)

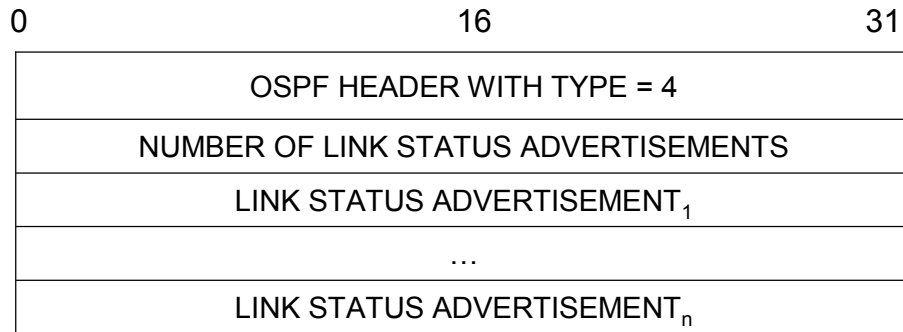
- Un router può scoprire che parti del suo database sono ormai superate. Per ottenere info aggiornate, manda un msg di richiesta dello stato dei collegamenti in cui sono elencate connessioni specifiche.
- Il vicino risponde con le info più aggiornate che ha su quei collegamenti
- I tre campi mostrati sono ripetuti per ogni link di cui si richiede lo stato

M. Colajanni, F.Mazzoni, 2003, Routing IP

5.49

## Open Shortest Path First (OSPF) (14)

Messaggio di aggiornamento dello stato dei collegamenti



M. Colajanni, F. Mazzoni, 2003, Routing IP

5.50

## Open Shortest Path First (OSPF) (15)

- Ogni aggiornamento consiste in un elenco di annunci (vedi slide precedente) e ogni annuncio ha un formato d'intestazione. I valori usati in ciascun campo sono gli stessi usati nel msg di descrizione del DB.

LINK AGE	LINK TYPE
LINK ID	
ADVERTISING ROUTER	
LINK SEQUENCE NUMBER	
LINK CHECKSUM	LENGHT

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.51

## Border Gateway Protocol (BGP): Scelte

- Link state o distance vector?
- Problemi di distance-vector:
  - Algoritmo di Bellman-Ford lento a convergere
- Problemi di link state:
  - Le metriche usate dai router non sono le stesse - cicli
  - Il database LS è troppo grande - tutta Internet
  - Può esporre le politiche agli altri AS

## BGP: Soluzione: Path Vectors

- Ogni aggiornamento della tabella di instradamento contiene tutto il cammino
- I cicli sono scoperti come segue:
  - Quando un AS riceve un cammino guarda se l'AS è già in quel cammino
    - Se sì, rifiuta il cammino
    - Se no, si aggiunge al cammino e (probabilmente) annuncia il cammino ad altri AS
- Vantaggio:
  - Le metriche sono locali - AS sceglie il cammino, il protocollo assicura che non vi siano cicli



## Border Gateway Protocol (BGP)

- Il BGP versione 4, definito nel RFC 1771 (vedi anche 1772 e 1773), è lo standard per il routing fra sistemi autonomi
- Assomiglia al protocollo distance vector, ma è meglio caratterizzato come un protocollo path vector, questo perché non propaga le info sui costi, ma le info sul percorso, come la sequenza di AS che si trovano sulla rotta verso un AS di destinazione
- Ciascun dominio può scegliere le sue rotte in accordo a qualsiasi criterio voglia

## Border Gateway Protocol (BGP) (2)

- Trasporto affidabile. Usa TCP per le comunicazioni
- Aggiornamenti incrementali. Non passa informazioni complete in ogni msg di aggiornamento. Solo la prima volta e poi diffonde i delta
- Supporta l'indirizzamento senza classi (CIDR)
- Autenticazione. Consente al destinatario di verificare l'identità del mittente

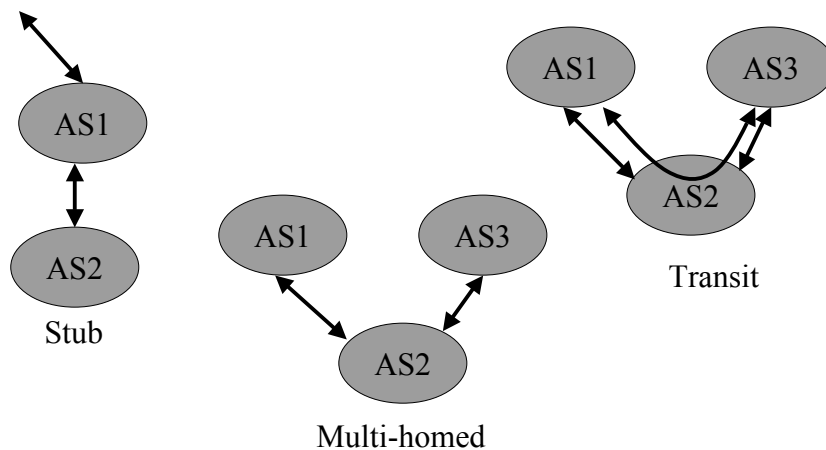
## Politiche con BGP

- BGP fornisce la capacità di sfruttare varie politiche
- Le politiche **non** fanno parte di BGP: sono fornite a BGP come informazioni di configurazione
- BGP fa rispettare le politiche scegliendo i cammini fra alternative multiple e controllando gli annunci verso gli altri AS

## Categorie di AS

- **Stub**: un AS che ha una sola connessione con un altro AS – trasporta solo traffico locale
- **Multi-homed**: un AS che ha connessioni verso più di un altro AS, ma che non trasporta traffico in transito
- **Transit**: un AS che ha connessioni verso più di un altro AS, e che trasporta sia traffico locale che in transito (sotto certe restrizioni date dalle politiche)

## Categorie di AS



M. Colajanni, F. Mazzoni, 2003, Routing IP

5.58

## Esempi di politiche di BGP

- Un AS multi-homed AS si rifiuta di fungere da transito
  - Annuncia limitazioni dei cammini
- Un AS multi-homed può diventare intermediario per alcuni AS
  - Annuncia i cammini solo a certi AS
  - Un AS può favorire o sfavorire certi AS per quanto riguarda il traffico in transito

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.59

## Border Gateway Protocol (BGP) (3)

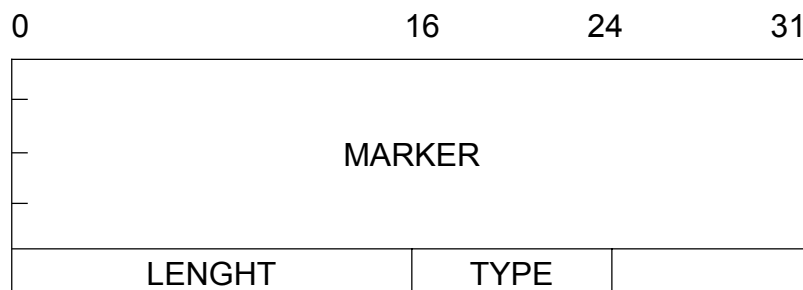
- Tre funzioni base
  1. Autenticazione e acquisizione del partner iniziale
  2. Invio info di raggiungibilità positive o negative
  3. Verifica che le connessioni funzionino correttamente
- Si definiscono 4 tipi di msg base in BGP-4
  1. OPEN inizializza la comunicazione
  2. UPDATE annuncia o ritratta gli instradamenti
  3. NOTIFICATION risponde ad un msg errato
  4. KEEPALIVE verifica attivamente la connessione paritaria

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.60

## Border Gateway Protocol (BGP) (4)

- Ciascun msg inizia con un'intestazione fissa, di 19 byte, che identifica il tipo di msg



M. Colajanni, F. Mazzoni, 2003, Routing IP

5.61

## Border Gateway Protocol (BGP) (5)

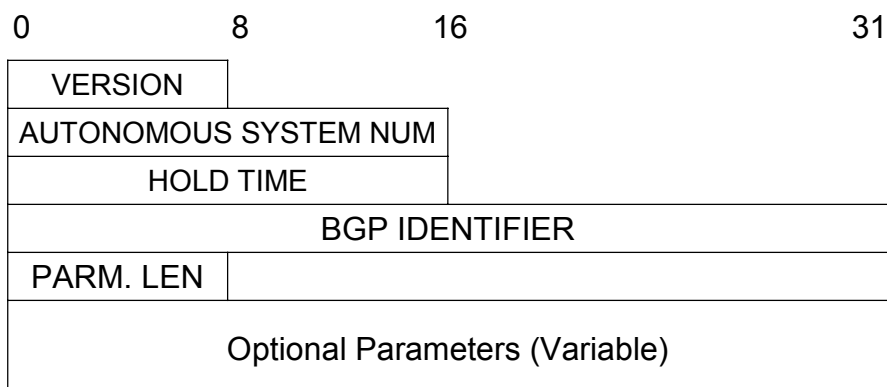
- Marker contiene un valore che entrambi i lati sono d'accordo di usare per contrassegnare l'inizio di un msg
- Marker è usato per la sincronizzazione: tutti i msg BGP viaggiano su TCP che non identifica il confine fra un msg e l'altro
- Length specifica la lunghezza totale del msg: la minima è 19 byte, la massima 4096 byte

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.62

## Border Gateway Protocol (BGP) (6)

- Formato del msg BGP OPEN inviato all'avvio.



M. Colajanni, F. Mazzoni, 2003, Routing IP

5.63

## Border Gateway Protocol (BGP) (7)

- Appena due partner BGP stabiliscono una connessione TCP, ciascuno invia un msg OPEN per dichiarare il suo numero di sistema autonomo e stabilire altri parametri operativi
- Hold time specifica il numero massimo di secondi che possono trascorrere fra la ricezione di due msg successivi
- BGP identifier è un intero che identifica univocamente il mittente (NB: è un indirizzo IP)
- Parm len indica la lunghezza in byte del campo facoltativo optional parameters

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.64

## Border Gateway Protocol (BGP) (8)

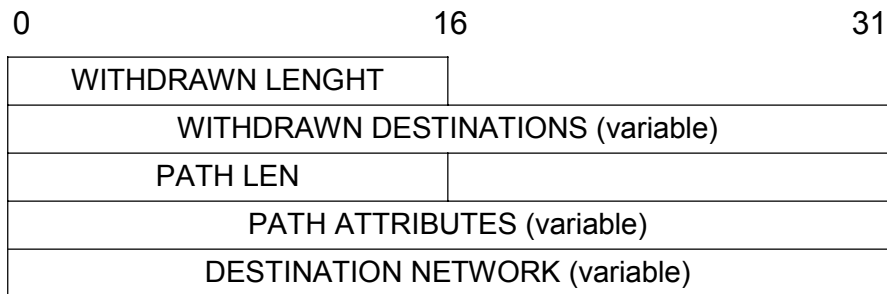
- Quando ci sono parametri, essi sono preceduti da un header di 2 byte, di cui il primo specifica il tipo di parametro e il secondo la lunghezza.
- Se non vi sono parametri il campo parm len è zero e il msg termina senza ulteriori dati
- Quando accetta un msg OPEN in entrata, un router BGP risponde inviando un msg KEEPALIVE (descritto in seguito)
- Entrambi i partner devono inviare un OPEN e ricevere un KEEPALIVE prima di poter scambiare informazioni di instradamento
- KEEPALIVE funziona come conferma di ricezione di OPEN

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.65

## Border Gateway Protocol (BGP) (9)

- Formato del msg BGP UPDATE.
- Le aree a dimensione variabile possono essere omesse



M. Colajanni, F.Mazzoni, 2003, Routing IP

5.66

## Border Gateway Protocol (BGP) (10)

- I partner usano i msg UPDATE per annunciare nuove destinazioni diventate raggiungibili o per annullare avvisi precedenti quando una destinazione è diventata irraggiungibile
- Ogni msg UPDATE è diviso in 2 parti: la prima elenca le destinazioni precedentemente annunciate che sono state eliminate e la seconda specifica nuove destinazioni divenute raggiungibili
- Withdrawn len specifica la dimensione di Withdrawn destinations
- Path len indica la dimensione degli attributi di instradamento

M. Colajanni, F.Mazzoni, 2003, Routing IP

5.67

## Border Gateway Protocol (BGP) (11)

- Il mittente può usare gli attributi di instradamento per specificare il salto successivo verso le destinazioni annunciate, un elenco di AS lungo il percorso verso le destinazioni e se le info di instradamento sono state apprese da un altro AS oppure derivate all'interno dell'AS del mittente
- Il mittente può specificare un cammino esatto verso la destinazione, elencando tutti gli AS da attraversare
- Permettono al destinatario di imporre limitazioni alle politiche di instradamento (es ctrl che non si passi per un AS non fidato)

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.68

## Border Gateway Protocol (BGP) (11)

- I partner BGP si scambiano periodicamente msg KEEPALIVE per verificare le connessioni di rete e per ctrl il reciproco funzionamento.
- Un msg KEEPALIVE consiste in un'intestazione standard senza ulteriori dati, quindi la sua dimensione totale è 19 byte
- Lo scambio periodico di msg è necessario perché BGP usa TCP che non include un meccanismo per verificare continuamente se il punto finale del collegamento è raggiungibile. TCP riporta però un errore se non può consegnare i dati

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.69



## Border Gateway Protocol (BGP) (12)

- I msg KEEPALIVE risparmiano larghezza di banda rispetto ad altri msg.
- Molti protocolli di instradamento si scambiano informazioni di routing anche per vedere se tutto funziona correttamente. A volte questo non è necessario e spreca banda.
- BGP separa la funzione di aggiornamento dei percorsi dalla verifica della connessione, consentendo di inviare spesso piccoli msg KEEPALIVE e riservare i più grandi msg UPDATE per le situazioni in cui le info di raggiungibilità cambiano

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.70

## Border Gateway Protocol (BGP) (13) Limitazioni dei protocolli di gateway esterni

- Un EGP (Exterior Gateway Protocol) non comunica e non interpreta le unità metriche di distanza, neanche se sono disponibili
- Poiché BGP diffonde solo info di raggiungibilità, il destinatario può implementare le limitazioni che desidera alle politiche d'instradamento, ma non può scegliere il percorso a minor costo. Il mittente deve quindi annunciare solo i percorsi che il traffico dovrebbe seguire

M. Colajanni, F. Mazzoni, 2003, Routing IP

5.71

## Border Gateway Protocol (BGP) (14)

- Formato del messaggio BGP NOTIFICATION

0	8	16	31
ERR CODE	ERR SUBCODE	DATA	

- È usato per il ctrl o quando si verifica un errore
- Gli errori sono permanenti; una volta che rileva un problema, BGP invia un msg di notifica e chiude la connessione TCP
- Per ogni possibile codice di errore, il campo subcode contiene un'ulteriore spiegazione

M. Colajanni, F.Mazzoni, 2003, Routing IP

5.72

## Border Gateway Protocol (BGP) (15)

- Valori possibili del campo ERR CODE

ERR CODE	Significato
1	Errore nell'intestazione del msg
2	Errore nel msg OPEN
3	Errore nel msg UPDATE
4	Timer di attesa scaduto
5	Errore della macchina a stati finiti
6	Fine (connessione terminata)

M. Colajanni, F.Mazzoni, 2003, Routing IP

5.73

## Border Gateway Protocol (BGP) (15)

- Valori possibili del campo ERR SUBCODE per errori nel mgs OPEN

ERR SUBCODE	Significato
1	Numero di versione non supportato
2	AS paritario non valido
3	Identificatore BGP non valido
4	Parametro facoltativo non supportato
5	Errore di autenticazione
6	Tempo di attesa non accettabile