



Routing protocols

Corso di Reti di calcolatori
14 Gennaio 2013

Patrizio Bassi
patrizio.bassi@gmail.com

Che cos'è il routing

- Riferimento OSI
- Concetto di switch
- Concetto di routing
 - IP (Internet Protocol) è il protocollo di livello 3
 - IPv4 (32bit)
 - IPv6 (128 bit)

OSI model
7. Application layer NNTP · SIP · SSI · DNS · FTP · Gopher · HTTP · NFS · NTP · SMPP · SMTP · SNMP · Telnet · DHCP · Netconf · (more)
6. Presentation layer MIME · XDR
5. Session layer Named pipe · NetBIOS · SAP · PPTP · RTP · SOCKS · SPDY · TLS/SSL
4. Transport layer TCP · UDP · SCTP · DCCP · SPX
3. Network layer IP (IPv4 · IPv6) · ARP · ICMP · IPsec · IGMP · IPX · AppleTalk
2. Data link layer ATM · SDLC · HDLC · CSLIP · SLIP · GFP · PLIP · IEEE 802.2 · LLC · L2TP · IEEE 802.3 · Frame Relay · ITU-T G.hn DLL · PPP · X.25
1. Physical layer EIA/TIA-232 · EIA/TIA-449 · ITU-T V-Series · I.430 · I.431 · PDH · SONET/SDH · PON · OTN · DSL · IEEE 802.3 · IEEE 802.11 · IEEE 802.15 · IEEE 802.16 · IEEE 1394 · ITU-T G.hn PHY · USB · Bluetooth · RS-232 · RS-449

diffuso

Miliardi IP)

Che cos'è il routing (2) - Definizione

- **Instradamento** dei pacchetti verso la giusta destinazione
- La routing table si può aggiornare anche a mano in modo statico!
- Un protocollo di routing è un protocollo relativo al network layer che permette ai **router** di scambiarsi informazioni tra loro al fine di costruire delle **tabelle di routing**
 - **Sotto unix/linux: # route**
 - **Windows: c:\> route print**

Routing table

- IPv4 Tabella route

- =====

- Route attive:

Indirizzo rete	Mask	Gateway	Interfaccia	Metrica
0.0.0.0	0.0.0.0	172.16.92.100	172.16.92.106	281
127.0.0.0	255.0.0.0	On-link	127.0.0.1	306
127.0.0.1	255.255.255.255	On-link	127.0.0.1	306
127.255.255.255	255.255.255.255	On-link	127.0.0.1	306
172.16.92.0	255.255.255.0	On-link	172.16.92.106	281
172.16.92.106	255.255.255.255	On-link	172.16.92.106	281
172.16.92.255	255.255.255.255	On-link	172.16.92.106	281
224.0.0.0	240.0.0.0	On-link	127.0.0.1	306
224.0.0.0	240.0.0.0	On-link	172.16.92.106	281
255.255.255.255	255.255.255.255	On-link	127.0.0.1	306
255.255.255.255	255.255.255.255	On-link	172.16.92.106	281

- =====

- Route permanenti:

Indirizzo rete	Mask	Indir. gateway	Metrica
0.0.0.0	0.0.0.0	172.16.92.100	Predefinito

- =====

Autonomous System

- un sistema autonomo (Autonomous System, AS) è un gruppo di router e reti sotto il controllo di una singola e ben definita autorità amministrativa.
 - es. organismo governativo, rete ISP, grandissimi enti/società (Google, Microsoft, Berkley)
- Ogni AS che utilizza la rete pubblica deve essere registrato presso il rispettivo RIR (Regional Internet Registry): ARIN (us), RIPE (eu) e altri. Ciascun AS è inoltre identificato da un numero univoco a 16 bit o 32 bit.
- In pratica, ad ogni AS viene assegnato un numero di indirizzi IP pubblici, che possono essere distribuiti.
 - Fine ultimo: IP unici nel mondo

Protocolli di routing

- **Protocolli interni al sistema autonomo**

- Protocolli **Distance Vector**

- RIP Routing Information Protocol
 - IGRP Interior Gateway Routing Protocol

- Protocolli **Link State**

- IS-IS Intermediate System to Intermediate System
 - OSPF Open Shortest Path First

- Protocolli **ibridi**

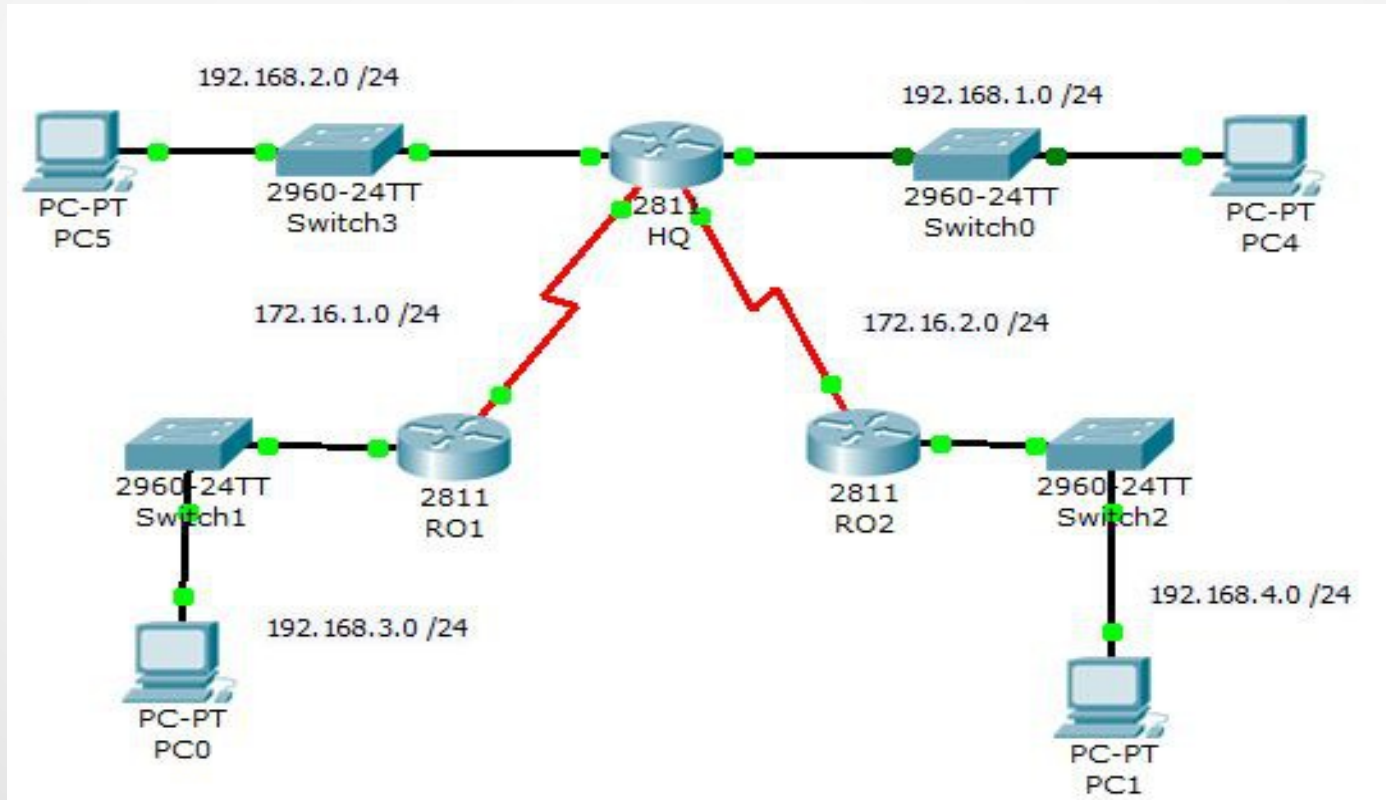
- EIGRP Enhanced Interior Gateway Routing Protocol

- **Protocolli esterni al sistema autonomo**

- EGP Exterior Gateway Protocol, obsoleto
 - BGP Border Gateway Protocol

Distance Vector

- Distance Vector (routing basato sul vettore delle distanze)
- Basato sull'algoritmo di Bellman-Ford



RIP (Routing Information Protocol) – 1

- Il numero massimo di hop consentito è 15 (16 significa infinito/invalido)
- Ogni router RIP trasmette di default, ogni 30 secondi, la propria tabella di routing completa a tutti i vicini direttamente collegati
 - Spreco banda e intervallo troppo largo
- RIP utilizza UDP come protocollo di trasporto, sulla porta riservata 520, IP destinazione 224.0.0.9
[VEDERE WIRESHARK]
 - Paradosso di utilizzare il livello 4 (UDP) come un'applicazione di livello 7
- Ogni router memorizza solo il primo passo dei percorsi che ha nella tabella.
 - A pubblica un percorso verso C, i vicini di A non possono sapere se sono stati inclusi da A nel percorso calcolato: possono formarsi cicli, risolti con la tecnica dello **split-horizon**
 - quando si interrompe un collegamento si può avere una situazione di count-to-infinity
 - rete "lineare" A-B-C-D-E-F e che si interrompa il collegamento tra A e B.
 - B sa che non può più raggiungere A con l'interfaccia "verso sinistra"
 - C (che è ancora inconsapevole della situazione) sta dichiarando di poter raggiungere A in due passi; B riterrà quindi di poter raggiungere A in tre passi tramite C.
 - Questa situazione può propagarsi sulla rete generando stime di distanza sempre maggiori.

RIP (Routing Information Protocol) – 2

- RIP TIMERS, valori di default
 - Update 30 sec: intervallo tra gli advertisement
 - Hold-Down 60/90 sec: ritardo di update nel caso di ripetuti cambi di link status (es. up/down)
 - Timeout 180 sec: attesa prima di dichiarare una rotta morta, viene refreshato ad ogni msg ricevuto
 - Flush 120 sec: attesa prima di cancellare una rotta dopo il timeout

RIP (Routing Information Protocol) – 3

- Ci sono 3 versioni di RIP: RIPv1, RIPv2 e RIPv6.

- **RIPv1**, definito da (RFC 1058), usa il routing "classful".

Gli aggiornamenti delle tabelle di routing non contengono la maschera di sottorete rendendo impossibile la creazione di sottoreti di dimensione diversa all'interno della stessa rete. Non viene supportata nessuna forma di autenticazione, lasciando RIPv1 vulnerabile ad attacchi;

- **RIPv2**, è stato sviluppato nel 1994 e definito da (RFC 2453), include il trasporto delle informazioni sulla maschera di sottorete.

Per garantire la sicurezza degli aggiornamenti sono disponibili 2 metodi: autenticazione semplice con testo in chiaro e MD5, (RFC 2082).

- **RIPv6**, (RFC 2080), è una estensione del protocollo originale RIPv1 per supportare IPv6;

Link State - Generico

- **Pacchetto Link State:**

- Contiene il router sorgente e tutti i nodi direttamente connessi
- ha un sequence number:
 - se il numero di sequenza è più recente si processa e inoltra a tutti
 - se uguale, si scarta
 - se meno recente si manda il più aggiornato al nodo mittente
- Ogni nodo memorizza i pacchetti ricevuti e costruisce una mappa completa e aggiornata della rete: il Link State Database (LSDB), che sarà uguale per tutti
- Ogni nodo esegue in maniera indipendente un algoritmo, generalmente una variante dell'Algoritmo di Dijkstra, per determinare il cammino minimo per raggiungere ogni nodo della rete ponendosi come radice dell'albero dei cammini minimi

- **Vantaggi:**

- Rapido nel convergere
- Evita loop
- Mappa completa
- Grande scalabilità con un numero di nodi

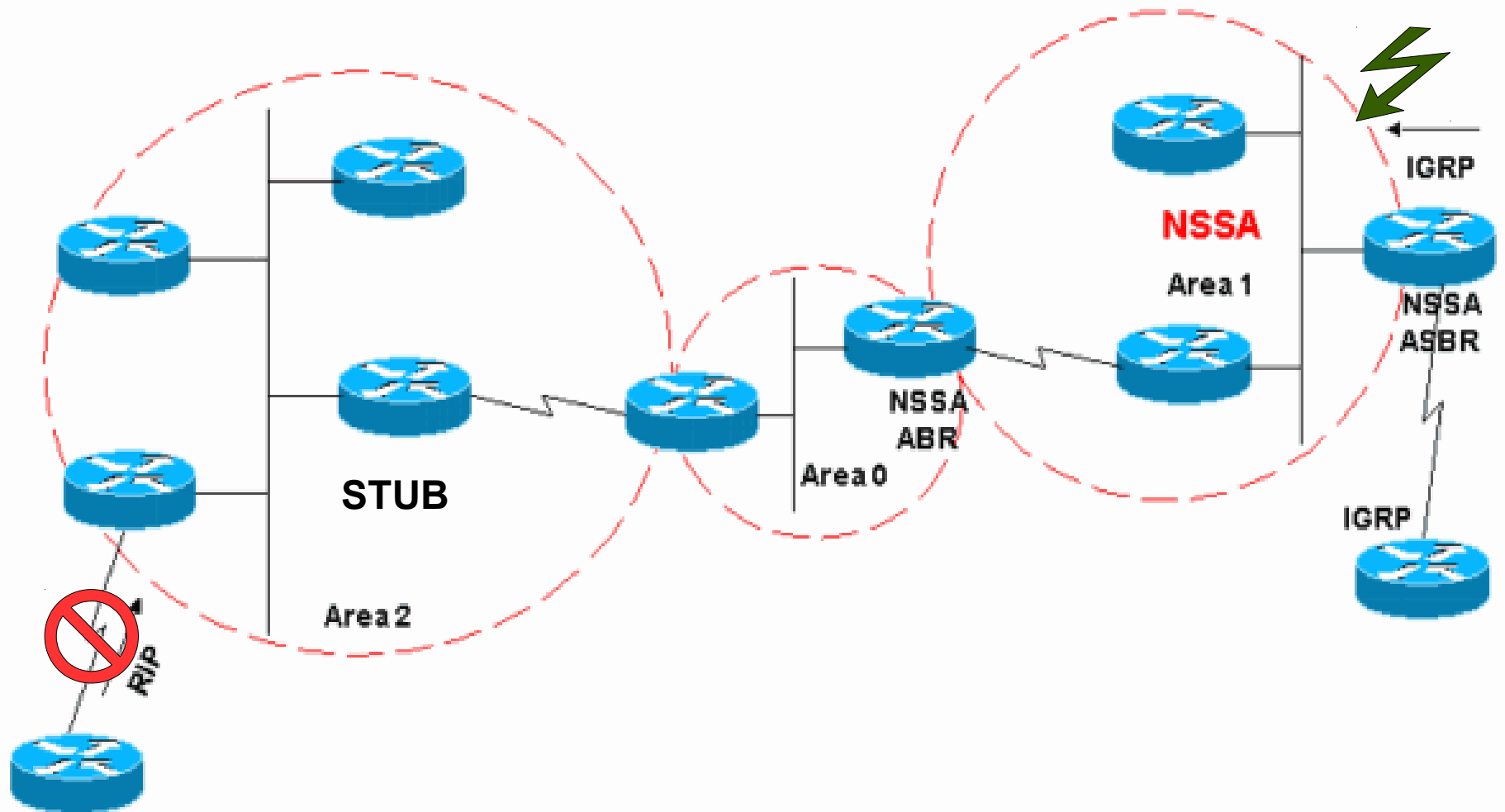
- **Svantaggi:**

- Richiede molta memoria ed elaborazione sui cambi per essere rapidi
- Complicato da implementare (e debuggare!)

OSPF - Open Shortest Path First (1)

- Ogni router OSPF è identificato in maniera univoca da un numero a 32bit, nella forma di un IP address
 - RouterID, sui sistemi unix legato alla loopback interface in modo da essere raggiungibile da qualunque interfaccia
- Gestisce reti broadcast, NBMA, PtP
- Ovviamente anche reti non classful
- Ha autenticazione: nulla, cleartext, md5
- Puro level3: usa multicast addresses 224.0.0.5 (all routers) e 224.0.0.6 (designated routers), ip protocol 89
- OSPFv2: IPv4
- OSPFv3: IPv6

OSPF (2)



OSPF – Tipi di Router (3)

- **Area Border Router (ABR)** è un router che connette una o più aree OSPF all'area di backbone. È membro di tutte le aree alle quali è connesso. Un ABR mantiene in memoria copie multiple del database link-state, uno per ciascuna area alla quale appartiene.
- **Autonomous System Boundary Router (ASBR)** è un router connesso a più di un Autonomous system (Internet) (AS), che scambia informazioni di routing con router in altri AS. Gli ASBR tipicamente utilizzano anche un protocollo di routing non-IGP, come il BGP. Un ASBR viene utilizzato per distribuire le rotte ricevute dagli altri AS attraverso il proprio AS.
- **Internal router (IR, router interno)** se ha solo adiacenze OSPF con router nella stessa area.
- **Backbone router (BR)** è un router con un'interfaccia verso l'area backbone. Un ABR è anche un BR, anche se non è necessario che sia vero l'inverso.

OSPF – Tipi di Router (4)

- **Designated router (DR)** è un router eletto tramite procedura di Hello dagli altri router appartenenti al suo dominio di broadcast.
- Elezione
 - Si parte con la priorità da 0 a 255 (max). Se la selezione di priorità su di un router OSPF è settata a 0, significa che tale router non potrà MAI divenire un DR o un BDR (Backup Designated Router).
 - Quando un DR si guasta e il BDR lo sostituisce, c'è una nuova elezione per chi farà da BDR.
 - Il router che manda i pacchetti Hello con la priorità highest.
 - Se due o più router si legano con la selezione di priorità più alta, vince il router che manda l'Hello con il più alto RID (Router ID).
 - SE un router OSPF con priorità più alta si mette in linea DOPO che l'elezione è avvenuta, non diventerà DR o BDR finché (almeno) il DR e il BDR si guastano.
 - E' l'unico che manda gli aggiornamenti di topologia a tutti in broadcast per ridurre l'uso di banda e risorse. Gli altri router mandano aggiornamenti al DR in caso di link change
 - Sul ptp non si elegge perchè non servirebbe
- **Backup Designated Router**
 - Un backup designated router (BDR, router di backup predefinito) è un router che diventa il principale se il router principale in uso ha un problema o si guasta. Il BDR è il router OSPF secondo per priorità nel momento dell'ultima selezione.

OSPF Packets (5) - Header

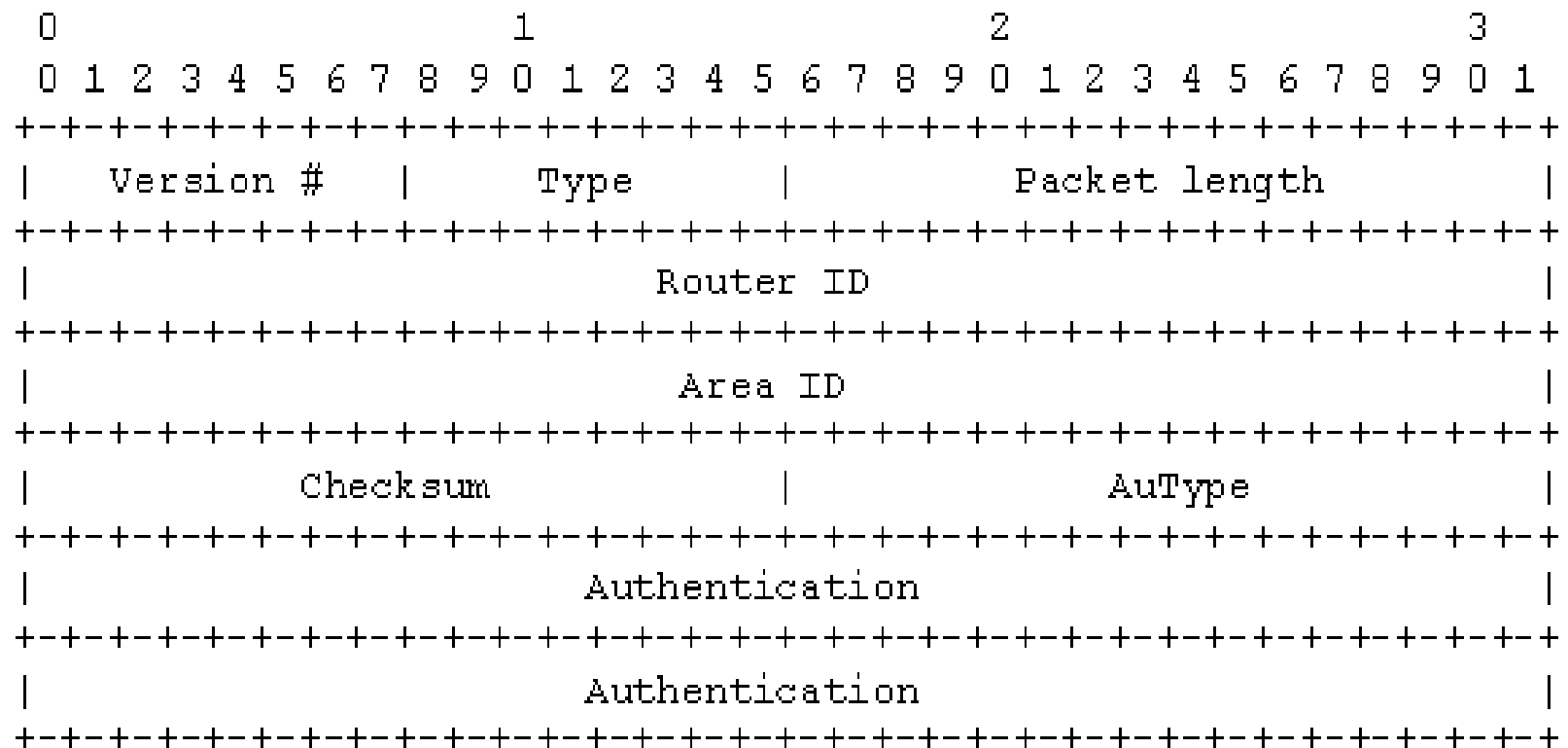


Figure 12

OSPF Packets (6) - Types

Type	Packet name	Protocol function
1	Hello	Discover/maintain neighbors
2	Database Description	Summarize database contents
3	Link State Request	Database download
4	Link State Update	Database update
5	Link State Ack	Flooding acknowledgment

Figure 13

OSPF Packets (7) - LSAs

- 11 types of link-state advertisements (LSAs), inside the Link State Update (type4) packets
- Type 1: router LSA — A router sends this to describe neighbors and its own interfaces.
- Type 2: network LSA — For broadcast networks only; this LSA is flooded by the DR and lists OSPF-speaking routers on the network.
- Type 3: network summary LSA — Sent by an ABR to advertise networks reachable through it. A stub area router will also use this for the default route.
- Type 4: ASBR-summary LSA — Sent by ABR, but only internally. This describes to the others how to get to the ASBR itself, and uses only internal metrics.
- Type 5: AS-external LSA — Used to describe external routes to internal areas. Can be used to advertise “this is the way to the Internet” (or some subset of).
- Type 6: Group summary — Used in multicast (MOSPF). Ignore this. Cisco Routers does not support LSA6.
- Type 7: NSSA area import.
- Type 8: External Attributes LSA — This is very rarely deployed and is used when BGP information is carried across OSPF AS.
- Type 9, 10, 11 are Opaque LSAs — They are designated for future specifically for application specific purposes.

Exterior Gateway Protocol – BGP (1)

- **Border Gateway Protocol (BGP)** è un protocollo di routing usato per connettere tra loro più router che appartengono a sistemi autonomi (Autonomous System, AS) distinti. Protocollo di routing inter-AS.
- Usa il metodo di **path vector**, metriche “politico/economiche”
- Routing indipendente dalle classi (Classless InterDomain Routing)
- Meccanismo di aggregazione degli instradamenti per diminuire le dimensioni delle relative tabelle.

BGP (2)

- IBGP: Nella topologia più semplice (e ridondante) ogni router di un singolo AS dovrebbe connettersi in “full mesh” con tutti gli altri.
- Questo causa un evidente problema di scalabilità: il numero delle connessioni (link) cresce col quadrato del numero dei router
- Per ridurre le problematiche BGP ha due opzioni:
 - route reflectors (RFC 4456): simile ai DR OSPF
 - confederations (RFC 5065): gruppi di AS raggruppati su Internet
- Scambio di pacchetti che contengono NLRI (Network Layer Reachability Information)

BGP (3)

- AS adiacenti, peer, si connettono manualmente con una sessione TCP sulla porta 179.
- Nella OPEN della connessione si avvisa della possibilità di estensioni (es. portare informazioni extra da altri protocolli)
- L'iniziatore della sessione BGP (speaker) invia periodicamente (per default ogni 60 secondi) dei messaggi keepalive da 19 byte per mantenere attiva la connessione.
- Peering manuale, contrattuale

BGP (4) – Incidenti - Politica

- Google offline per quasi 5% della popolazione mondiale
 - <http://blog.cloudflare.com/why-google-went-offline-today-and-a-bit-about>
 -
- Telecom Italia depeering ([PDF](#))

Implementazioni

- GPL Quagga → <http://www.nongnu.org/quagga/>

DOMANDE?