Il routing in Internet
Interior Gateway Protocols

A.A. 2005/2006

Walter Cerroni

Routing Information Protocol (RIP)

- Protocollo **distance vector**, di implementazione vecchia (RFC 1058, Giugno 1988), discende dal protocollo di routing realizzato per la rete XNS di Xerox
- Ne esiste una **versione 2** più recente (RFC 2453)
- Molto diffuso in passato perché il codice di implementazione è liberamente disponibile
- Utilizzato praticamente solo su reti TCP/IP
- Utilizza due tipi di messaggi:
  - **REQUEST** serve per chiedere esplicitamente informazioni ai nodi vicini (ad es. all'avvio del nodo)
  - **RESPONSE** serve in generale per inviare informazioni di routing (cioè i distance vector)
- I messaggi RIP sono trasportati da UDP ed usano la porta 520 sia in trasmissione che in ricezione

2

RIP: la tabella di routing

- Ogni riga nella tabella contiene:
  - indirizzo di **destinazione**: è un indirizzo IP a 32 bit
    - della route di default se vale 0.0.0.0
    - di rete se Net-ID ≠ 0 e Host-ID = 0 (in base alla **classe**)
    - di sottorete se Subnet-ID ≠ 0 e Host-ID = 0
      - necessario conoscere la **netmask** → valido solo per reti direttamente collegate alle interfacce dei router
    - di host se Host-ID ≠ 0
  - **distanza** dalla destinazione (metrica)
    - in termini di hop-count (ogni link ha peso = 1)
    - la distanza massima (∞) per RIP è pari a **16**, al fine di limitare il conteggio all'infinito → adatto per reti relativamente piccole
  - **next-hop** sul percorso verso la destinazione
    - router vicino a cui inviare i datagrammi per la destinazione
  - due contatori: **timeout** e **garbage-collection timer**

3

RIP: invio delle informazioni

- Un **RESPONSE** con nuove informazioni di routing viene inviato:
  - periodicamente
  - come risposta ad una richiesta esplicita
  - quando una informazione di routing cambia (triggered update)
- Le informazioni periodiche sono inviate ogni 30 secondi, con uno scarto da 1 a 5 secondi, per evitare "tempeste" di aggiornamenti
- Se una route non viene aggiornata dopo 180 secondi (**timeout**), la sua distanza è posta all'infinito (si ipotizza una perdita di connettività)
- Dopo ulteriori 120 secondi (**garbage-collection timer**) la route viene eliminata del tutto dalla tabella

4

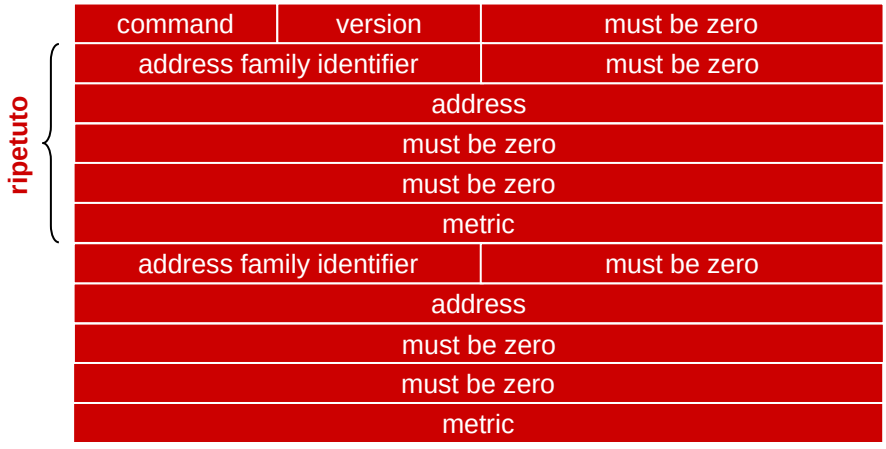
RIP: aggiornamento della tabella di routing

- Quando si riceve un RESPONSE viene inizialmente controllata la correttezza dei dati presenti nel distance vector ricevuto (indirizzi IP e metriche validi)
- Quindi si considerano solo le voci con distanze non infinite, che vengono incrementate di una unità e poi, per ogni destinazione:
  - se non esiste una entry corrispondente nella tabella, viene creata: la distanza è quella appena incrementata, il next-hop è il mittente del RESPONSE e si fa partire il timeout
  - se esiste già una entry ed ha distanza maggiore di quella indicata, la entry viene aggiornata e si fa ripartire il timeout
  - se esiste già una entry verso tale destinazione ed il next-hop è lo stesso che ha inviato il RESPONSE, la entry viene aggiornata se diversa dal valore precedente e, comunque, si fa ripartire il contatore del periodo di vita
  - in tutti gli altri casi non si fa nulla

5

RIP: formato dei pacchetti

- La struttura del pacchetto è basata su parole di 32 bit
- Il pacchetto può avere lunghezza variabile fino a 512 byte (max 25 entry)



6

RIP: significato dei campi

- I bit del pacchetto sono molto ridondanti rispetto alla quantità di informazioni da inviare (molti campi fissi con i bit tutti a zero)
  - inizialmente pensati per adattarsi ad altri protocolli
- **command**: distingue tra REQUEST (1) e RESPONSE (2)
- **version**: versione del RIP
- **address family identifier**: indica il tipo di indirizzo di rete utilizzato, vale 2 per IP
- **address**: identifica la destinazione per la quale viene data la distanza
- **metrica**: è la distanza dalla destinazione indicata

7

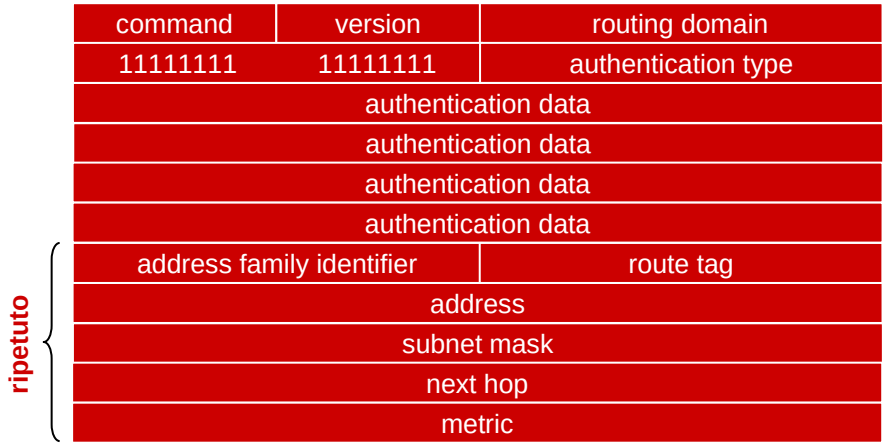
RIP: problematiche

- Fa uso di split horizon per cui le RESPONSE di interfacce diverse possono essere diverse
- Fa uso di triggered update: in questo caso non è necessario indicare nella RESPONSE tutte le entry della tabella ma solamente quelle appena modificate
- RIP non è un protocollo sicuro: chiunque trasmetta datagrammi dalla porta UDP 520 viene considerato come un router autorizzato
- Esempio di malfunzionamento indotto:
  - un router non autorizzato trasmette messaggi contenenti indicazione di una distanza 0 tra se stesso e tutti gli altri della rete
  - dopo qualche tempo tutti i percorsi ottimi convergono su questo router

8

RIP versione 2

- I miglioramenti introdotti riguardano soprattutto:
  - subnetting e CIDR
  - autenticazione



9

RIP versione 2

- Compatibilità verso il basso
  - RIP-1 ignora le entry con i campi riservati diversi da zero
- Possibilità di indicare sottoreti o indirizzamento CIDR
  - tramite il campo **subnet mask**
- Possibilità di **autenticare** chi invia i messaggi
- Possibilità di indicare il proprio AS e di scambiare informazioni con protocolli EGP
  - tramite i campi **route tag** e **routing domain**
- Possibilità di specificare un **next hop** più appropriato

- Comunque non adatto ad AS grandi
- Comunque ha problemi di convergenza
  - è pur sempre un distance vector

10

Open Shortest Path First (OSPF)

- Divenuto standard nella versione 2 (RFC 2328)
- Oggi è il più diffuso IGP
- Protocollo di tipo **link state**
  - invio di **Link State Advertisement** (LSA) a tutti gli altri router
- Incapsulato direttamente in IP
  - il valore del campo protocol dell'intestazione IP (89 per OSPF) serve a distinguere questi pacchetti da altri
- OSPF è stato progettato specificamente per:
  - semplificare il routing in reti grandi tramite la suddivisione in aree
  - gestire reti intrinsecamente diffusive (LAN IEEE 802, FDDI)
  - gestire reti intrinsecamente punto-punto (X.25, ATM, Frame Relay)
  - separare logicamente gli host dai router

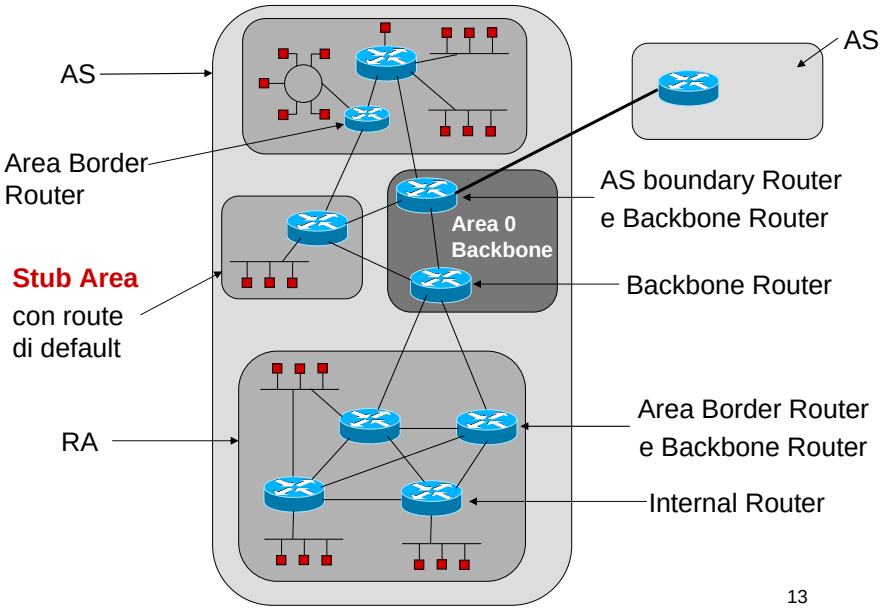
11

OSPF: aree di routing

- Un AS può essere suddiviso in porzioni dette **Routing Area** (RA) interconnesse da un **backbone** (Area 0)
  - ciascuna area risulta separata dalle altre per quanto riguarda lo scambio delle informazioni di routing e si comporta come un'entità indipendente (3°livello gerarchico di rout ing)
  - per interconnettere le aree vi devono essere router connessi a più aree e/o al backbone (almeno un router per area)
- Classificazione dei router secondo OSPF:
  - **Internal Router**: router interni a ciascuna area
  - **Area Border Router**: router che scambiano informazioni con altre aree
  - **Backbone Router**: router che si interfacciano con il backbone
  - **AS Boundary Router**: router che scambiano informazioni con altri AS usando un protocollo EGP

12

OSPF: aree di routing e tipologie di router



13

OSPF: ulteriori caratteristiche

- **Bilanciamento del carico**: se un router ha più percorsi di uguale lunghezza verso una certa destinazione, il carico viene ripartito equamente su di essi
- **Autenticazione**: per garantire maggiore sicurezza nello scambio delle informazioni di routing è prevista autenticazione con password ed uso di crittografia
- **Routing dipendente dal grado di servizio**: i router scelgono il percorso sul quale instradare un pacchetto sulla base dell'indirizzo e del campo Type of Service dell'intestazione IP, tenendo conto che percorsi diversi possono offrire diversi gradi di servizio

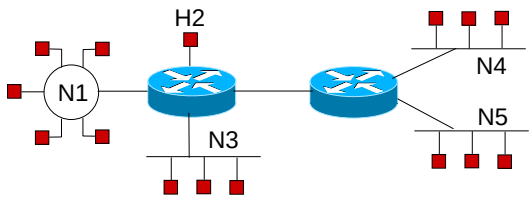
14

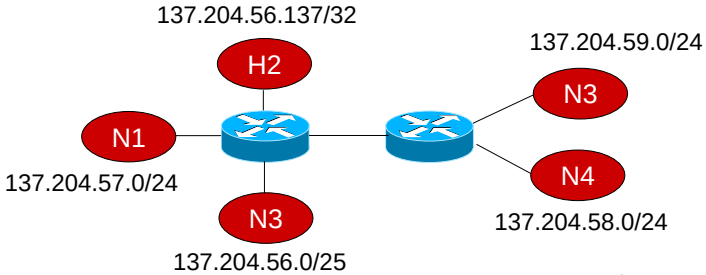
OSPF: host e router

- Nel modello OSPF i router sono i soli responsabili del routing
  - gli host sono solamente punti terminali da raggiungere
  - in teoria sarebbe necessario indicare ogni host nei grafi che rappresentano la rete (e nelle tabelle di routing)
- Se gli host di una rete IP sono connessi ad una LAN:
  - la singola rete IP viene vista come una sola entità raggiungibile in un colpo solo (identificata dall'indirizzo di rete)
  - vengono diffuse informazioni relative alla raggiungibilità dell'intera rete, non dei singoli host
- Se un singolo host è collegato direttamente ad un router:
  - è necessario indicarlo esplicitamente (tramite il suo l'indirizzo)

15

OSPF: rappresentazione di host e router





16

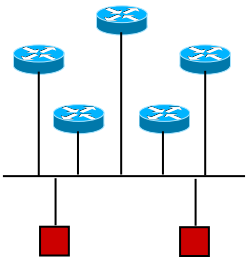
OSPF: tipologie di rete

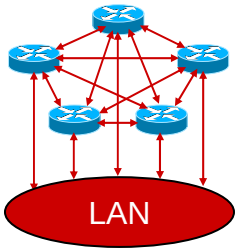
- OSPF è progettato per operare correttamente con reti:
  - **Point-to-Point**
  - **Broadcast Multi-Access** (diffusive: LAN, FDDI)
  - **Non-Broadcast Multi-Access** (NBMA: X.25, ATM, Frame Relay)

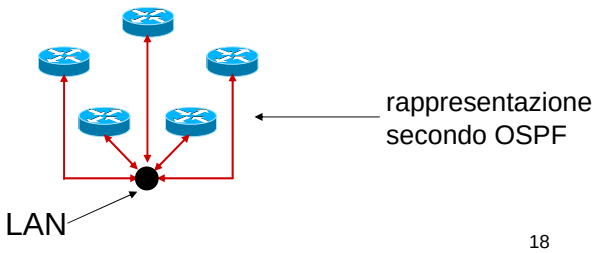
- In una **rete ad accesso multiplo** tutti gli N router connessi alla rete sono di fatto connessi con tutti gli altri
  - il numero di archi bidirezionali da inserire nel grafo è N(N–1)/2+N
  - il numero totale di LSA da trasmettere è N(N–1)
  - conviene adottare una **topologia a stella equivalente**, inserendo un nodo virtuale che rappresenta la rete
    → solo N archi bidirezionali

17

OSPF: rappresentazione di reti multi-accesso







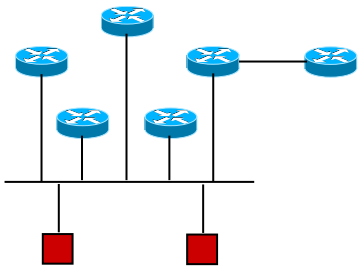
18

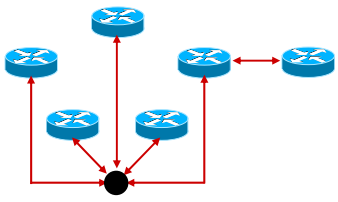
OSPF: vicinanza e adiacenza tra router

- **Vicini**: due router che sono connessi alla medesima rete e possono comunicare direttamente
  - punto-punto o punto-multipunto
- **Adiacenti**: due router che si scambiano informazioni di routing
- In una rete ad accesso multiplo risulta molto più efficiente eleggere un **Designated Router** (DR) fra gli N vicini
  - ogni router della LAN è adiacente solo al DR
  - lo scambio di informazioni di routing avviene solo tra router adiacenti (cioè DR fa da tramite)
  - inoltre il DR è l'unico a comunicare la raggiungibilità di router e host della LAN al mondo esterno
  - Per ragioni di affidabilità occorre avere anche un **Backup Designated Router** (BDR) adiacente a tutti i router locali

19

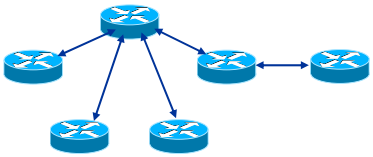
OSPF: vicinanza e adiacenza tra router





router vicini

DR



router adiacenti

20

OSPF: identificazione di router e priorità

- Ogni router di un AS utilizzante OSPF deve avere un identificativo univoco (**router ID**):
  - di default si prende l'indirizzo IP più alto fra quelli assegnati alle interfacce del router
  - si può assegnare manualmente un router ID ad ogni router configurando opportunamente l'interfaccia di loop-back
  - configurare l'interfaccia di loop-back è un modo più stabile e sicuro di assegnare il router ID perché questa interfaccia non viene mai disabilitata
- Ai singoli router di un'area possono essere associate delle priorità
  - utilizzate nell'elezione del DR
  - valore di priorità compreso tra 0 e 255 (8 bit)
  - di default tutti i router hanno priorità 0 (più bassa)

21

OSPF: elezione di DR e BDR

- Ciascun router nella rete ad accesso multiplo:
  - esamina la lista dei suoi vicini
  - elimina dalla lista tutti i router non eleggibili (ad esempio tutti quelli che hanno priorità nulla)
  - fra quelli rimasti seleziona il router avente la priorità maggiore
    - il più alto router ID in caso di uguale priorità
  - elegge il router selezionato a DR
  - rivede la tabella dei vicini e riseleziona gli eleggibili (a questo punto il router che è stato eletto DR non è più eleggibile)
  - seleziona ed elegge il BDR secondo le regole già adottate per il DR
  - termina la procedura una volta eletti DR e BDR

22

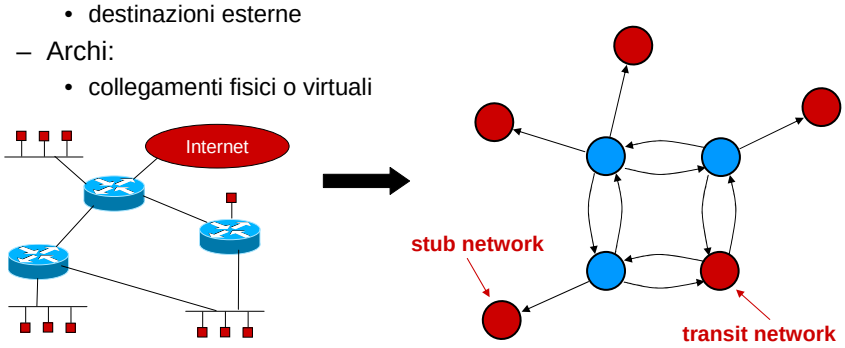
OSPF: tipi di LSA

- **router-LSA**
  - creati da ogni router e diffusi in una singola area
  - contengono le informazioni relative allo stato delle interfacce del router all'interno di quell'area
- **network-LSA**
  - creati dal DR di ogni rete ad accesso multiplo e diffusi in una singola area
  - contengono l'elenco dei router connessi a quella rete
- **summary-LSA**
  - creati dagli Area Border Router e diffusi in una singola area
  - contengono informazioni di routing per destinazioni appartenenti ad altre aree dello stesso AS
- **AS-boundary-router-summary-LSA**
  - creati dagli Area Border Router e diffusi a tutte le sue aree
  - contengono informazioni di routing per raggiungere gli AS-boundary router
- **AS-external-LSA**
  - creati dagli AS Boundary Router e diffusi a tutto l'AS
  - contengono informazioni di routing per destinazioni appartenenti ad altri AS (compresa la default route)

23

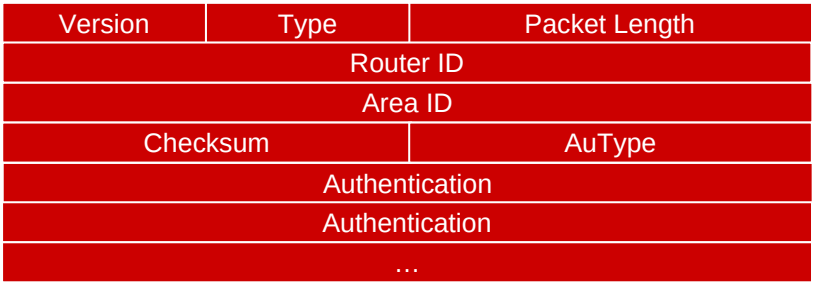
OSPF: Link State Database

- Il grafo orientato della rete sul quale ciascun router calcola lo **shortest path tree** è rappresentato dal **Link State Database** presente in ogni router
  - Nodi:
    - router
    - reti o host singoli
    - nodi virtuali delle topologie a stella equivalenti
    - destinazioni esterne
  - Archi:
    - collegamenti fisici o virtuali



OSPF: i protocolli

- OSPF invia messaggi utilizzando direttamente il protocollo IP (campo protocol = 89)
- Si compone di tre sottoprotocolli:
  - **hello, exchange, flooding**
- Tutti i messaggi hanno una intestazione comune
  - vengono aggiunte informazioni per il particolare scopo a cui il messaggio è destinato (tipo di pacchetto)



25

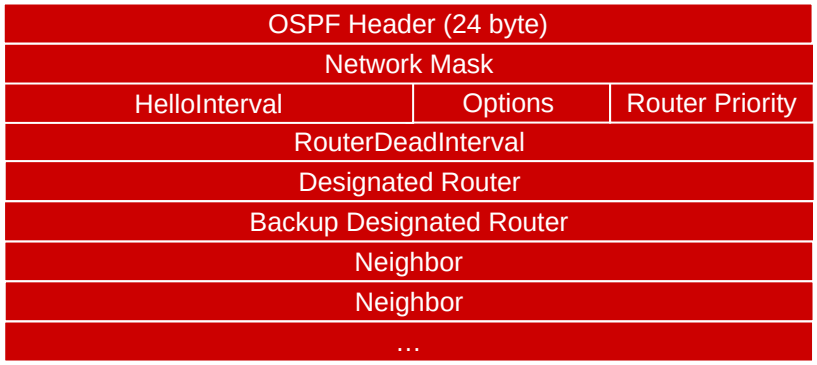
OSPF: intestazione comune

- **Version** indica la versione di OSPF (versione 2)
- **Type** indica il tipo di pacchetto
- **Packet Length** numero di byte del pacchetto
- **Router ID** indirizzo IP che identifica il router mittente
- **Area ID** identifica l'area di appartenenza
  - il numero 0.0.0.0 è l'area di backbone
- **Checksum** calcolata su tutto il pacchetto OSPF escludendo gli 8 byte del campo authentication
  - si utilizza l'algoritmo classico di IP
- **AuType** indica il tipo di autenticazione:
  - 0 nessuna autenticazione
  - 1 autenticazione semplice (password nel campo **authentication**)
  - 2 autenticazione crittografica (dati nel campo **authentication**)

26

OSPF: Hello protocol

- Unico tipo di pacchetto: **Hello** (Type = 1)
- Utilizzato per:
  - controllare l'operatività dei link
  - scoprire e mantenere relazioni fra vicini
  - eleggere DR e BDR



27

OSPF: Hello protocol

- I pacchetti HELLO sono inviati sulle interfacce periodicamente secondo quanto specificato dal parametro **HelloInterval**
  - si riescono così a scoprire i propri vicini
- Includono una lista di tutti i vicini (**Neighbor**) dai quali è stato ricevuto un pacchetto HELLO recente (cioè non più vecchio di **RouterDeadInterval**)
  - si riesce così a conoscere se per ciascun vicino è presente un collegamento bidirezionale e se esso è ancora attivo
- I campi **Router Priority**, **Designated Router** e **Backup Designated Router** sono utilizzati per l'elezione di DR e BDR
- **Network Mask** indica la maschera relativa all'interfaccia del router (l'indirizzo è nell'header IP)
- **Options** indica se si supportano funzionalità opzionali

28

OSPF: Exchange protocol

- Una volta stabilite le adiacenze, router adiacenti devono sincronizzare i rispettivi Link State Database
- La procedura di sincronizzazione è asimmetrica
  - si stabilisce chi è il master e chi lo slave
  - il master invia una serie di pacchetti **Database Description** (Type = 2) contenenti l'elenco dei LSA del proprio database
    - nell'elenco sono indicati il tipo di LSA, l'età, il router che lo ha generato e il numero di sequenza
    - non ci sono i dati relativi al LSA
  - lo slave risponde con l'elenco dei LSA del suo database
  - durante lo scambio ciascuno dei due router confronta le informazioni ottenute con quelle in proprio possesso
  - se nel proprio database ci sono dei LSA meno recenti rispetto all'altro, questi (e solo questi) vengono richiesti con un successivo pacchetto **Link State Request** (Type = 3)

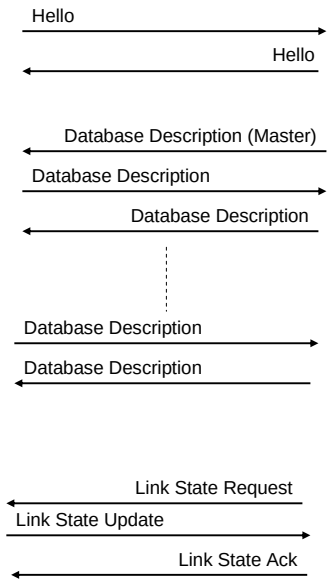
29

OSPF: Flooding protocol

- La diffusione dei LSA a tutti i router della rete avviene tramite l'invio di pacchetti **Link State Update** (Type = 4)
  - a fronte di un cambiamento nello stato di un collegamento
  - a fronte di una Link State Request
  - periodicamente (ogni 30 minuti)
- Si esegue in modalità flooding per fare in modo che tutti i router vedano gli aggiornamenti
  - flooding efficiente: si usano i numeri di sequenza dei LSA
- Si continua ad inviare lo stesso update finché non viene confermata la sua ricezione dai vicini tramite il pacchetto **Link State Acknowledgment** (Type = 5)
  - in questo modo si rende il flooding affidabile

30

OSPF: sincronizzazione e aggiornamento



Fase di **Hello**

I router scoprono l'esistenza reciproca

Fase di **Exchange**

Si sceglie il master e lo slave
Si confrontano i database

Fase di **Update**

Si inviano richieste di aggiornamento ai router adiacenti per aggiornare il database

31