



Complementi di Reti e Sistemi di Telecomunicazioni

5. Protocolli d'instradamento

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.1

Instradamento DVR e LSR

- Il confronto DV *Routing* (DVR) e LS *Routing* (LSR) è complesso, proviamo a distinguere diversi aspetti:
 - **Velocità di convergenza**
 - » In genere si tende a supporre che gli LSR convergano più rapidamente dei DVR, in pratica la velocità di convergenza dipende molto dalla topologia della rete e dalle caratteristiche del traffico.
 - **Volume di messaggi di controllo**
 - » LS: con N nodi e A archi richiedono lo scambio di $O(NA)$ messaggi per ciascun nodo.
 - » DV: deve solo scambiare i messaggi con i propri vicini.
 - **Robustezza**; se un nodo comincia a funzionare male:
 - » LS: il nodo propaga un costo sbagliato, ogni nodo si calcola separatamente la propria tabella
 - » DV: il nodo propaga un percorso sbagliato, ogni tabella viene calcolata facendo uso delle altre

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.2

Instradamento DVR e LSR

- I DVR non escludono la presenza di cicli a priori, ma con le opportune modifiche gli possono evitare efficacemente.
- Gli LSR, per contro, sono più complessi, devono fare uno sforzo significativo per mantenere i DB congruenti (generando anche un traffico di controllo più elevato) ed hanno *Distance Table* più grandi.
- Gli LSR possono usare più metriche diverse contemporaneamente.
- Gli LSR si prestano ad essere estesi per supportare con le stesse tabelle routing *unicast* e *multicast*

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.3

Instradamento Gerarchia

Ci sono due ragioni importanti per le quali nelle reti di una certa dimensione si tende ad usare meccanismi di instradamento gerarchici:

- **La scalabilità**
 - Per un numero di nodi elevato (WAN), indipendentemente dal tipo di algoritmo, la complessità dell'instradamento e la dimensione delle RT diventano comunque eccessive (oltre al traffico di segnalazione).
 - Per esempio, nel caso LS, con tanti archi quanti nodi, si ha una complessità di circa $O(N \log N)$ ed una RT con con dimensione $O(N)$, quindi

# nodi	RT	Calcoli
1000	1000	$O(3000)$
1.000.000	1.000.000	$O(6.000.000)$

- **L'autonomia amministrativa**

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.4

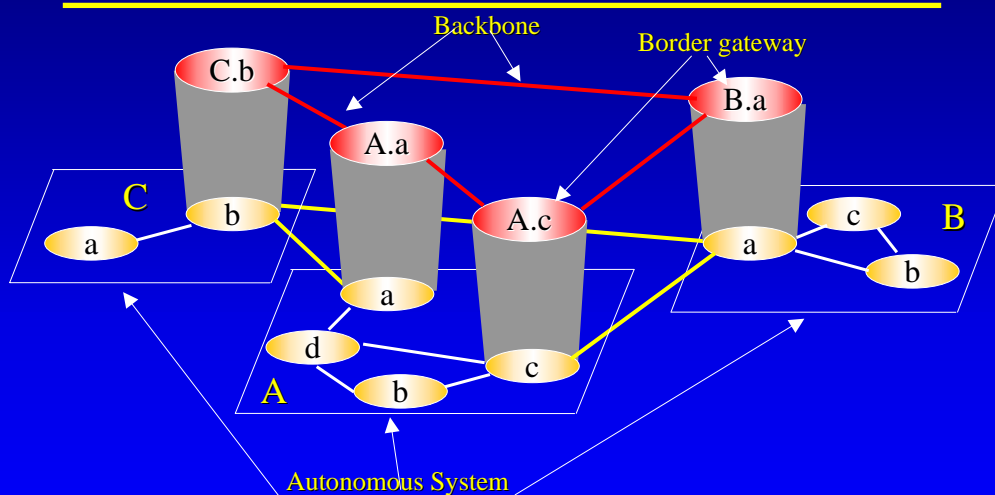
Instradamento Gerarchia

- Internet distingue tre livelli gerarchici principali:
 - Sottoreti o singoli domini di *broadcast*
 - » All'interno dei quali l'instradamento fa uso dell'ARP.
 - *Autonomous System (AS)*
In cui i protocolli di instradamento prendono il nome di *Interior Gateway Protocol* (IGP) e sono: RIP, IGRP e OSPF.
 - *Backbone*
In cui i protocolli di instradamento prendono il nome di *Exterior Gateway Protocol* (EGP) e sono: EGP e BGP.
- Ulteriori livelli possono essere inseriti tramite alcuni protocolli (OSPF), o sfruttando la "route aggregation".

Complementi di Reti e Sistemi R. Bolla, C. Nobile

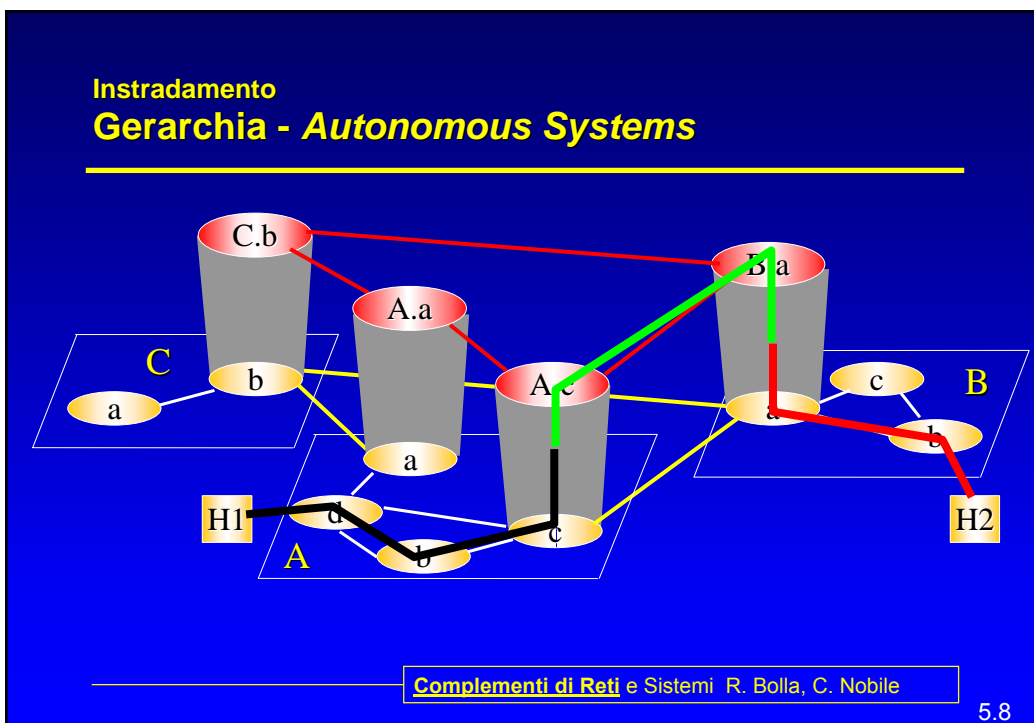
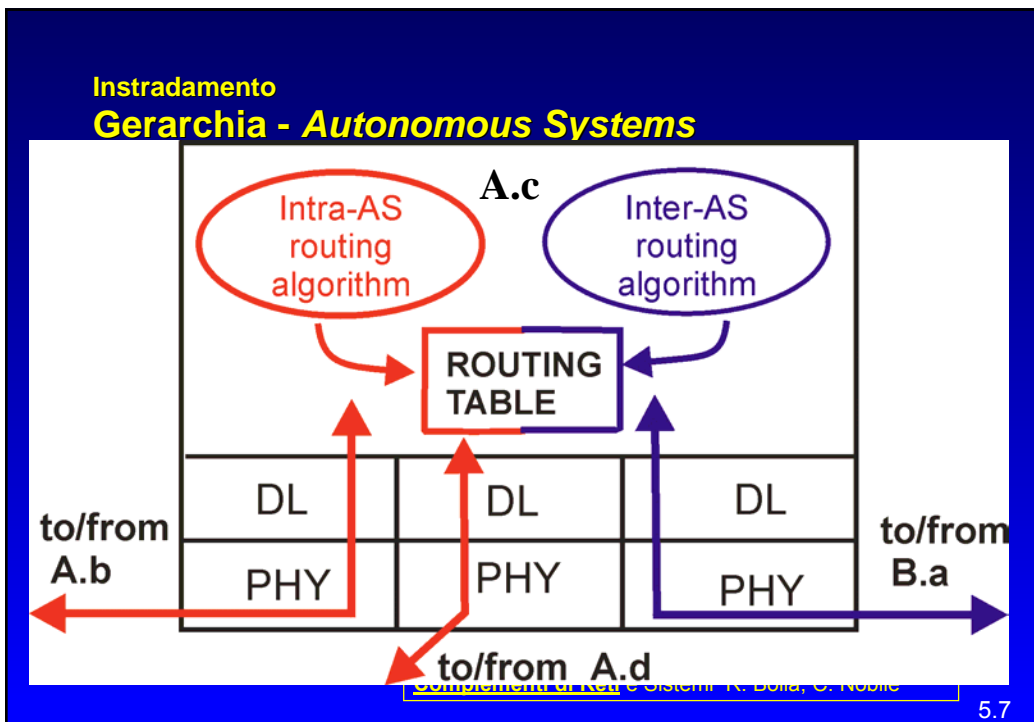
5.5

Instradamento Gerarchia - Autonomous Systems



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.6



Instradamento Gerarchia

- Il partizionamento è realizzato grazie alla gerarchizzazione degli indirizzi.
- L'obiettivo è avere relativamente "pochi" nodi per ogni livello (e area).
- Potenzialmente ogni livello può usare algoritmi diversi
- La gerarchia non è stretta, ossia il collegamento un area di un livello e il livello superiore può avvenire tramite più nodi (*Border Gateway, BG*)
- Ci sono dei router che partecipano all'instradamento di livelli differenti.
- Alcuni indirizzi possono non essere omogenei con lo spazio di indirizzamento dell'area/livello (questo diminuisce l'efficacia della gerarchia).

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.9

Instradamento Routing Information Protocol (RIP)

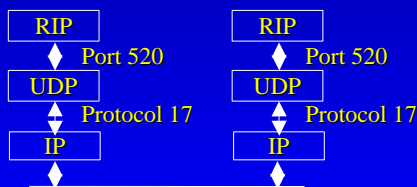
- E' un IGP originariamente progettato dalla Xerox per la propria rete, introdotto dall'Università di Berkley nella propria implementazione di TCP/IP (BSD)
- E' stato standardizzato con RFC 1058 nel 1988, la versione 2 è descritta nel RFC 1723.
- E' un DVR ed usa una metrica statica: il costo di un percorso è il numero di hop (ossia di linee) di cui è composto (ossia ogni linea a costo 1).
- Utilizza lo *split horizon with poisonous reverse*, e *i triggered update*.
- Aggiorna la RT (tramite *RIP response message* o *RIP advertisement*) ogni 30 s. e elimina ogni vettore non aggiornato per 180 s consecutivi (considerando la corrispondente linea non più disponibile).

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.10

Routing Informatio Protocol (RIP)

- Ha come valore massimo del costo 15, 16 corrisponde ad infinito. Quindi non permette reti con percorsi con più di 15 *router* attraversati.
- La limitazione di cui sopra è legata al fatto che per reti più grandi è troppo lento a convergere (non alla dimensione del campo costo).
- Ne esistono due versioni, la seconda (RIPv2) consente l'uso del CIDR, ossia la "route aggregation", e la "default route".
- Lo scambio di informazioni avviene attraverso un protocollo di livello 4 (UDP)



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.11

Instradamento Routing Information Protocol (RIP)

- Sollecito per un DV
 - Messaggio di *update* (anche su sollecito)
- Usato solo in RIPv2 per distinguere fra percorsi interni all'AS ed esterni.

Command	Version	Unused
Address Family ID		Route Tag
IP Address		
Subnet Mask		
Next Hop		
Metric		

- Usato per autenticazione
- Nel RIPv2 viene posto a FFFF e viene aggiunto successivamente un campo *password*

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.12

Interior Gateway Routing Protocol (IGRP)

- E' nuovamente un DVR, ma di tipo proprietario; infatti è stato sviluppato dalla CISCO verso la metà degli anni '80 ed è disponibile solo sui suoi prodotti.
- Usa una metrica dinamica e sofisticata (considera ritardo, banda, affidabilità, lunghezza del pacchetto ed il carico) in cui il costo della linea viene composto tramite una somma pesata, i cui pesi sono impostabili dal gestore.
- Permette la suddivisione del carico su più linee (multipercorso).
- Usa un meccanismo sofisticato per accelerare la convergenza ed evitare i cicli.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.13

Open Shortest Path First (OSPF)

- Nasce nel 1990 con l'RFC 1247 per sostituire il RIP
- E' un protocollo di tipo **Link State**
- Quindi ogni nodo costruisce al proprio interno la topologia di tutta la rete e invia in *flooding* i LSP, contenenti i costi dei *link* ad esso connessi.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.14

OSPF Storia

- 1987: istituito il Working Group OSPF
- 1989: 1° RFC (RFC 1131)
- 1990: prove di interoperabilità
- 1991: RFC1247 – OSPFv2
- 1992: raccomandazione come IGP
- 1993:
 - aggiornamento RFC1583
 - CIDR adottato in Internet
 - Supporto Multicast: M-OSPF

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.15

OSPF Storia

- 1994:
 - supporto interfaccia Multi-point
 - Crittografia per l'autenticazione
- 1997: RFC 2178 (OSPFv2)
- 1998:
 - RFC 2328 (OSPFv2)
 - RFC 2329 : OSPF Standardization Report

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.16

OSPF Obiettivi

- Superare i limiti del RIP
- Avere metriche più efficienti (migliori possibilità di descrivere la rete)
- Introdurre delle gerarchie (per maggior scalabilità)
- Separazione informazioni interne ed esterne (all'AS)
- Supporto subnetting variabile (CIDR)
- Sicurezza
- Routing dipendente anche dal TOS (In effetti il TOS è inutilizzato su IPv4 perché gli host non ne fanno uso e i router non li gestiscono ...)

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.17

OSPF Principi base: Caratteristiche

- **Open Shortest Path First (OSPF)**
- Classificato come IGP (interno all'AS)
- Algoritmo utilizzato di tipo *Link-state* (SPF)
- Progettato per IP supporta:
 - Subnetting
 - TOS (type of Services)
 - Multicast (M-OSPF)
 - Autenticazione
- Ottimizza traffico di segnalazione generato

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.18

OSPF

Principi base: Informazione

- Conoscenza stato locale (interfacce e vicini “attivi”).
- Invio in flooding dello stato locale.
- E quindi presenza di un data base con l'intera topologia dell'AS in ogni Router.
- Le Informazioni di instradamento provenienti dall'esterno di un AS sono gestite separatamente dalle informazioni di *Link-State* e sono trasmesse invariate all'interno.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.19

OSPF

Principi base: Algoritmo

- Instradamento di tipo **Link State (LS)** basato sull'indirizzo destinazione e ToS.
- Meccanismo comune a tutti i nodi per la costruzione ed aggiornamento di un database contenente il grafo della rete.
- Ogni router si considera *root* e calcola il proprio *Shortest Path Tree (SPT)*.
- Un SPT diverso per ogni livello di TOS.
- Possibile raggruppamento delle reti in aree.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.20

OSPF

Scelta tra LS o DV ?

- Scelta difficile che portò anche alla creazione di un Working group anche per un Open Distance Vector (i DV (come il RIP) hanno problemi di convergenza ma i LS sono più complessi da realizzare).
- La scelta LS è stata appoggiata dall'industria.
- ARPANET aveva sperimentato un protocollo LS con qualche problema.
- OSI sviluppa un protocollo LS (IS-IS) parallelamente a IETF.
- Si sceglie la tecnologia LS e la definizione di un protocollo ex-novo perché si ritiene troppo complesso e limitante modificare un protocollo esistente.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.21

OSPF

Scelta del tipo di segnalazione (cont.)

- Su livello 2 (seguendo le scelte dettate da OSI)
 - Problema MTU (Maximum Transfer Unit): la frammentazione e riassettaggio devono far parte del protocollo
- In pacchetti livello 3 (IP), vantaggi:
 - Utilizzo funzionalità proprie di tale livello (struttura pacchetto, ..)
 - Funziona su qualsiasi tipo di protocollo a livello 2 senza bisogno di progettare diverse forme di realizzazione.
- In pacchetti di protocolli di trasporto (UDP o TCP), vantaggi:
 - Disponibilità dell'interfaccia in modalità utente (IP è accessibile solitamente nella sola modalità super-user)
 - Checksum per verifica integrità (UDP)
 - Trasporto affidabile (TCP)

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.22

OSPF

Scelta del tipo di segnalazione

- Si è considerato IP la scelta “migliore”, assegnando il Type Protocol 89, facendo le seguenti considerazioni:
 - L'affidabilità può essere aggiunta tramite l'uso di meccanismi di “*flooding* affidabili”
 - Non permettere l'accesso da parte utenti è un limite che porta maggior sicurezza
 - Usando UDP (TCP) si avrebbe avuto 8 (20) byte di intestazione in più

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.23

OSPF

Elementi costitutivi

- Struttura del Data base contenente la topologia della rete.
- Struttura dei Link State Packet (LSP)
- Modalità di aggiornamento del Data base (segnalazione)
- Modalità di calcolo dell'instradamento
- Gerarchia interna (Aree)

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.24

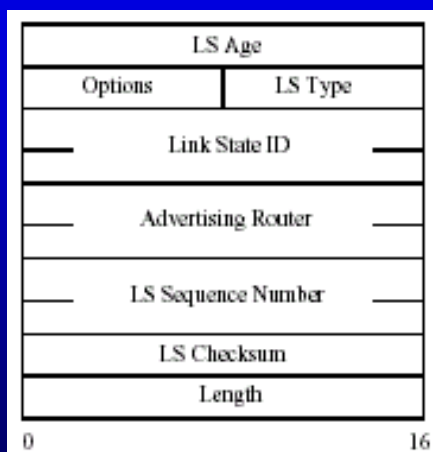
OSPF Data Base

- Tutti i router OSPF hanno una versione sincronizzata di un database composto da “Record Link State” (RLS).
- Gli RLS rappresentano la topologia e servono a calcolare lo *shortest path tree*.
- Gli RLS vengono aggiornati tramite “annunci”, o *Link State Advertisement* (LSA), provenienti da altri router, tutti aventi lo stesso formato.
- Il router che riceve un LSA modifica uno o più RLS del database.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.25

OSPF- Data Base Link State Record - Header



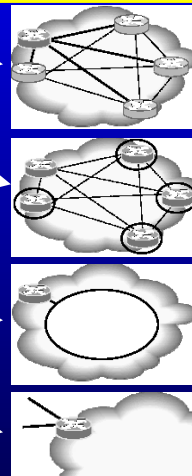
- LS Age: 16 bit – tempo trascorso dall’ultimo aggiornamento del record
- Options: da 1-6 non definiti
7 (bit E) = 1: link esterno
8 (bit T) = 1: supporto TOS
- LS Type: tipo di link annunciato – i parametri dei 5 tipi sono nel campo dati seguente l’header

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.26

OSPF- Data Base Link State Record - Header

- **Router link (tipo 1):** inviato da ogni router, il record contiene descrizione dei link (ID, TOS e metrica) che partono dal router.
- **Network link (tipo 2):** inviato dal solo router **designato**, contiene descrizione rete di transito con l'elenco dei router che hanno adiacenze con esso.
- **Summary di reti IP (tipo 3) o di router di bordo (tipo 4):** inviato dal solo router di bordo area (ABR), contiene netmask della sottorete IP ed elenco di TOS e metriche (routing gerarchico).
- **Link esterni (tipo 5):** inviato dal solo router di bordo area, contiene netmask di sottoreti IP esterne all'area ed elenco di TOS e metriche (routing gerarchico)



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.27

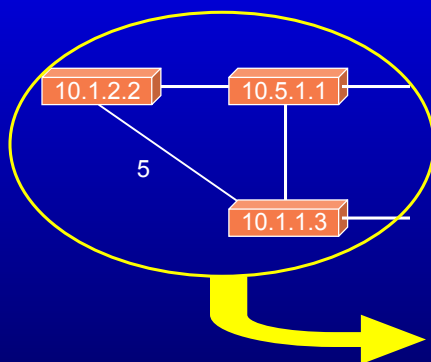
OSPF- Data Base Link State Record - Header

- **Link State ID:** indirizzo IP della destinazione connessa tramite il link annunciato
- **Advertising Router:** indirizzo IP del router che annuncia il link.
- **LS Sequence number:** numero progressivo dell'annuncio.
- **LS checksum:** codice di controllo (tipo IP) che protegge intestazione e dati

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.28

OSPF- Data Base Esempio di Link State Record



LS Age		0 seconds
Options	LS Type	E-bit,LS Type 1
Link State ID		10.1.1.1
Advertising Router		10.1.1.1
LS Sequence Number		0x80000006
LS Checksum		0x9b47
Length		60 bytes
Router Type	0	0 (ordinary)
# of links		3
Link ID		10.1.1.3
Link Data		Ifindex 2 (unnumbered link)
Link Type	#TOS Metrics	1(point to point), 0
Metric		5

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.29

Comunicazione tra router OSPF

- Gli LSA che OSPF fa scambiare tra i router sono inviati all'interno di pacchetti IP.
- Il protocollo OSPF è in realtà composto da tre “sottoprotocolli” con un *header* comune:
 - Protocollo HELLO (verifica dei link attivi).
 - Protocollo EXCHANGE (inizializzazione del DB)
 - Protocollo FLOODING (Aggiornamento dei DB)
- Per verificare se due router hanno le stesse informazioni (sono sincronizzati) si confrontano:
 - Il numero di LSA
 - La somma dei checksum

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.30

Comunicazione tra router OSPF

LS type	L.S. ID	Addr. Router	Check Sum	LS SN	LS Age
Router	10.5.1.1	10.5.1.1	0xda23	0x40000023	0
Router	10.5.2.1	10.5.2.1	0x67d1	0x40000005	5,312
Router	10.5.3.1	10.5.3.1	0x9021	0x40000018	6,234
Router	10.5.4.1	10.5.4.1	0xbcd1	0x40000003	0,426
Router	10.5.5.1	10.5.5.1	0x2254	0x40000026	3,234
Router	10.5.6.1	10.5.6.1	0xa12b	0x400000a2	1,453

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.31

Comunicazione tra router OSPF

- Esaminando il campo *Age* e *LS-SN* si evince quali sono le informazioni aggiornate più frequentemente, sono quelle:
 - con il tempo sempre basso (*Age*)
 - e il numero di sequenza (*LS-SN*) che varia frequentemente

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.32

OSPF Link State Packet

- Sono Pacchetti IP contenenti
 - *Protocol number 89*
 - TTL impostato a 1 (sono sempre “intercettati” dal protocollo OSPF del router direttamente collegato)
- Tre possibili indirizzi di destinazione (si usa il multicast se presente):
 - ALLDRouters (tutti i router “designati”)
 - ALLOSPFRouters (tutti i router OSPF)
 - IP del vicino (caso line punto-punto)
- La frammentazione è eseguita solo se indispensabile; per evitare la frammentazione si limita la dimensione a 1500 byte (per gli LSA).

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.33

Comunicazione tra router OSPF Il protocollo HELLO

- Usato per
 - controllare che i link siano operativi (devo ricevere un pacchetto di HELLO entro un *dead interval* per considerare attivo un link)
 - eleggere il router designato (LAN) e il suo backup, usando il valore di priorità del router.
- Invia pacchetti ogni n secondi ($n = \text{hello interval}$ (10 sec.) < *dead interval* (40 sec.))
- Negozia parametri (*hello interval* e *dead interval*)
- Il pacchetto di Hello include:
 - Indirizzo del router designato e del suo backup.
 - Indirizzi dei vicini da cui ha ricevuto un pacchetto di HELLO nell'ultimo *dead interval*

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.34

Comunicazione tra router OSPF Il protocollo EXCHANGE

- Usato per la sincronizzazione iniziale dei database dei router, che avviene fra una coppia di router adiacenti.
- Scambio di messaggi “Database Description” in modalità *master-slave* (chi invia per primo si elegge *master*, il ricevente è lo *slave*).
- Ogni messaggio è identificato da tre flag: I (*initialize*), M (*more*) e MS (*master/slave*) che regolano la procedura di sincronizzazione.
- Ogni messaggio include LS record del database del router, ed un *acknowledgment* dei pacchetti di EXCHANGE ricevuti nella direzione opposta.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

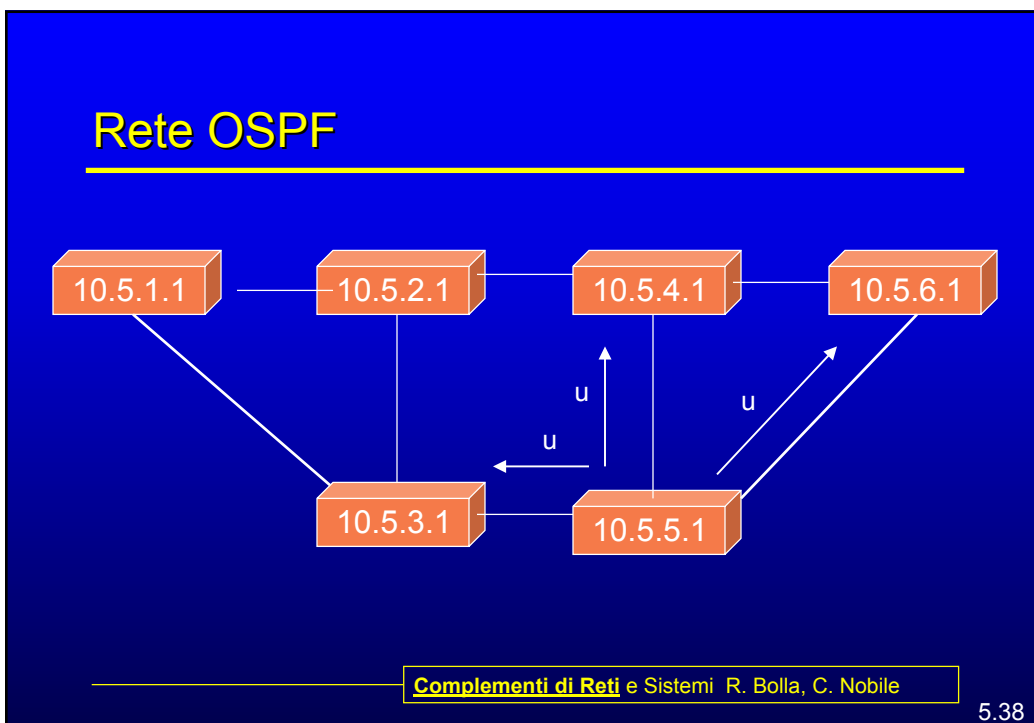
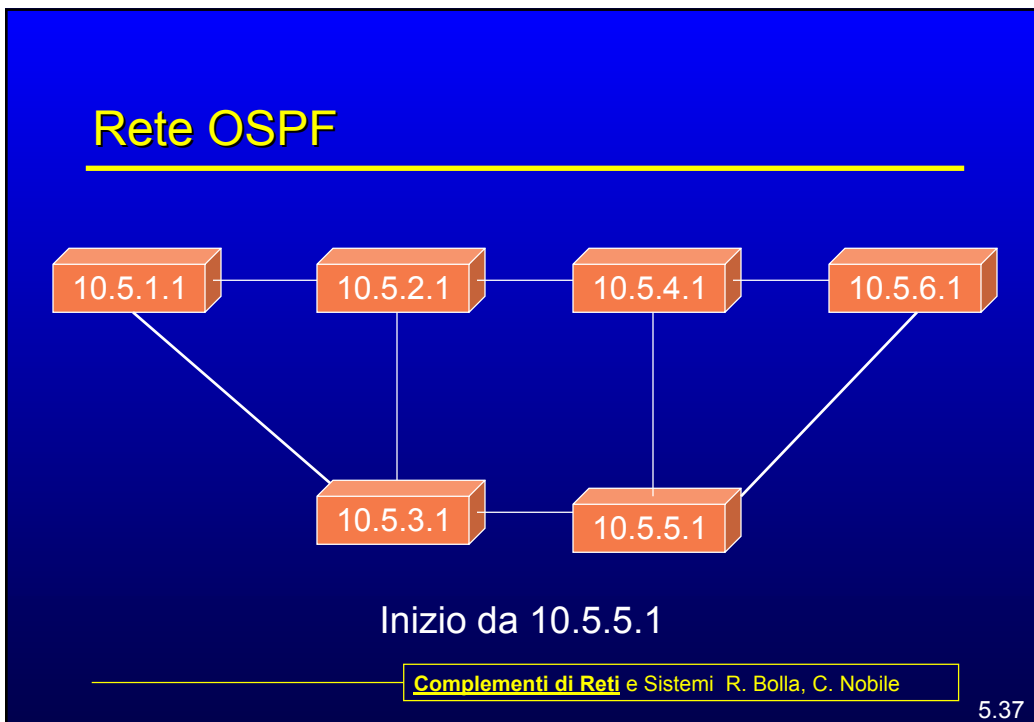
5.35

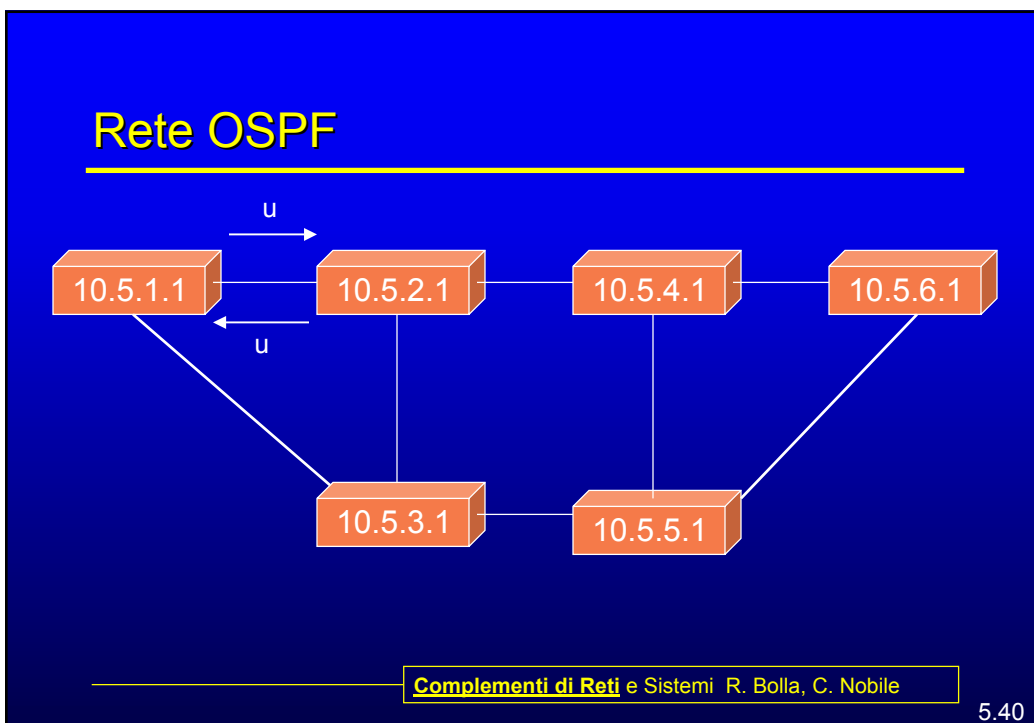
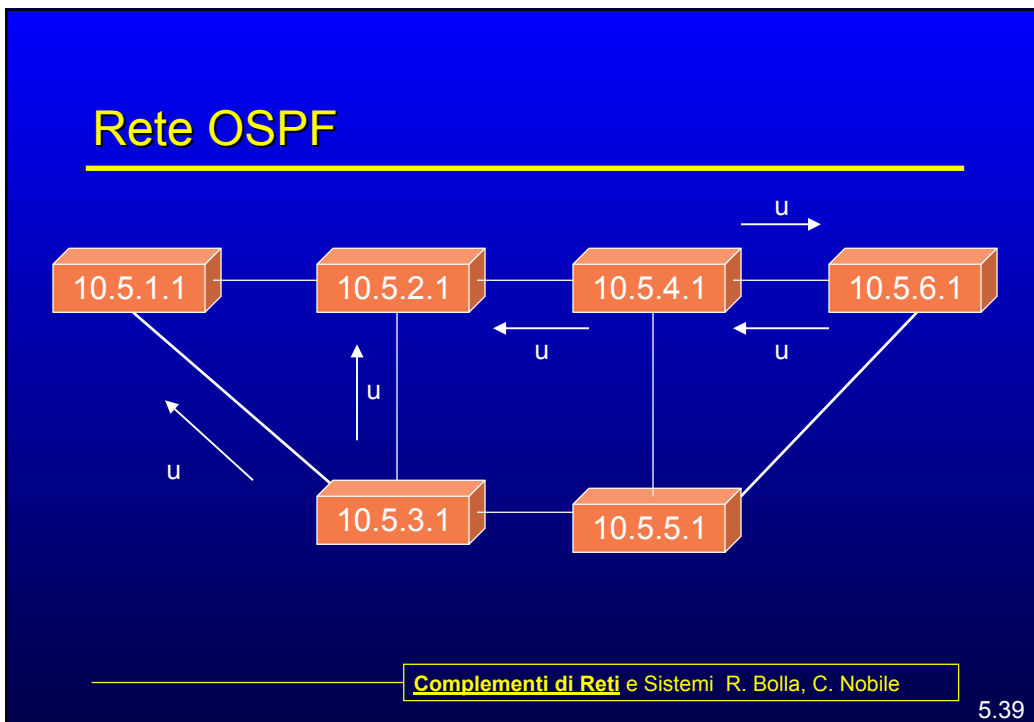
Comunicazione tra router OSPF Il protocollo FLOODING

- Usato per
 - annunciare cambi di stato di un link;
 - eseguire un rinnovo degli LSA (ogni 30 minuti)
- Ogni messaggio di *flooding* contiene un certo numero di LSA, e un numero di sequenza.
- Chi riceve il messaggio controlla se ha già ricevuto un messaggio con lo stesso numero di sequenza proveniente da quel router
 - In caso negativo, inoltra il messaggio su tutte le altre interfacce di rete
 - In caso affermativo, lo scarta.
- In ogni caso, invia un *acknowledgment* al mittente.
- Sulle reti locali i messaggi sono inviati in multicast.

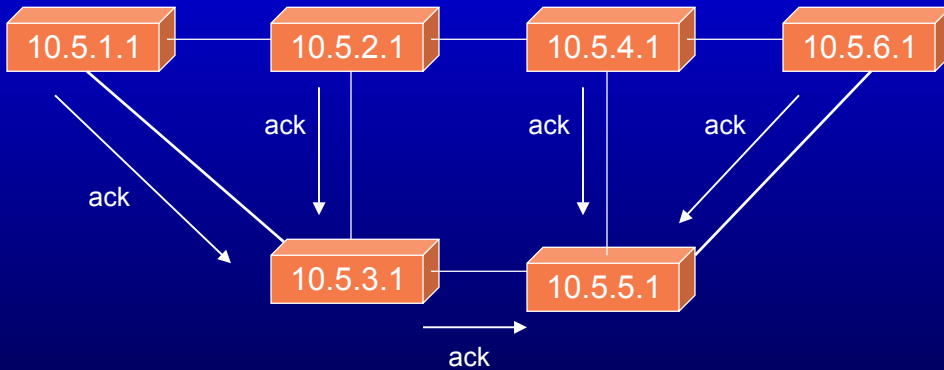
Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.36





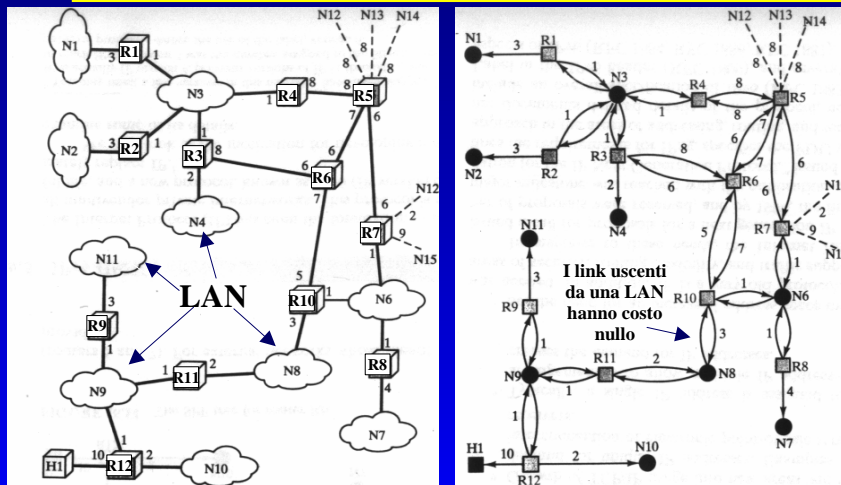
Rete OSPF



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.41

OSPF Calcolo dell'instradamento



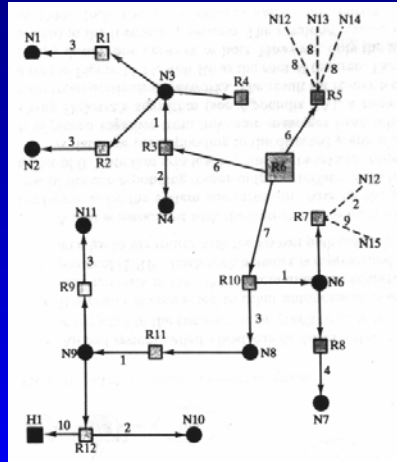
Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.42

OSPF Calcolo dell'instradamento

TABLE 16.5 Routing Table for RT6.

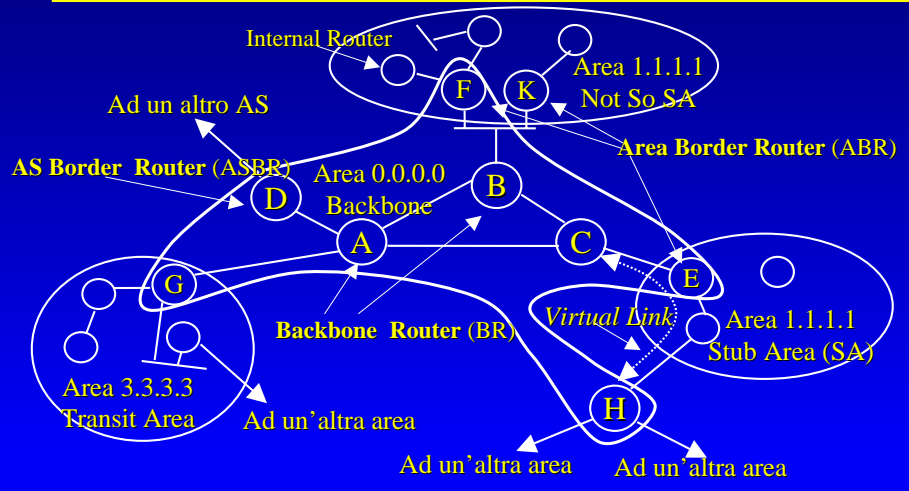
Destination	Next hop	Distance
N1	RT3	10
N2	RT3	10
N3	RT3	7
N4	RT3	8
N6	RT10	8
N7	RT10	12
N8	RT10	10
N9	RT10	11
N10	RT10	13
N11	RT10	14
H1	RT10	21
RT5	RT5	6
RT7	RT10	8
N12	RT10	10
N13	RT5	14
N14	RT5	14
N15	RT10	17



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.43

OSPF Gerarchia: Le aree



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.44

OSPF

Gerarchia: Le aree

- Le *Stub Area* non propagano informazioni interne o esterne ed accedono al *backbone* tramite un *router di default*.
- L'instradamento fra due aree viene realizzato in tre parti:
 - Il percorso nell'area sorgente fra la sorgente stessa ed un *Area Border Router*.
 - Il percorso fra i due ABR delle due aree tramite il *backbone*
 - Il percorso nell'area destinazione fra l'ABR che riceve il pacchetto dal *backbone* e la destinazione.
- In pratica si forza un instradamento a stella in cui il *backbone* rappresenta il centro stella.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.45

Instradamento

EGP- "EGP"

- Al più vecchio dei protocolli EGP è stato assegnato lo stesso nome che distingue la categoria: EGP.
- E' un protocollo di stile DV che però non propaga costi ma solo informazioni di raggiungibilità.
- Non è in grado di evitare cicli e quindi non può essere usato in topologie magliate ma solo ad albero.
- La sua struttura di riferimento è composta da "*Core Router*" (CR) collegati fra loro ad albero.
- Ogni AS può essere collegato ad un unico CR e quindi ogni CR fa da centro stella per un gruppo di AS

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.46

Instradamento

EGP - Border Gateway Protocol (BGP)

- E' il protocollo EGP relativamente recente, definito dal RFC 1654.
- La versione in uso attualmente è la 4 (BGP4).
- Permette la cooperazione fra *router* di AS diversi (chiamati *gateway*) per la realizzazione dell'instradamento fra AS.
- Per lo scambio di informazioni fra i nodi usa il TCP (porta 179).
- Non propaga vere e proprie metriche, ma lascia che la scelta dei percorsi venga determinata tramite "politiche" impostate dai singoli gestori.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.47

Instradamento

EGP - Border Gateway Protocol (BGP)

- Opera in tre passi:
 - Identificazione dei nodi adiacenti (*neighbor*)
 - Raggiungibilità dei nodi adiacenti
 - Raggiungibilità delle reti
- Utilizza un algoritmo DV, ed in particolare usa un *Path Vector*.
- Distingue tre tipi di reti
 - *Stub*: che hanno un'unica connessione con il *backbone* e non possono venir usate come transito
 - *Multiconnected*: che potenzialmente possono essere usate per transito (se lo permettono)
 - *Transit*: costruite per realizzare il transito.

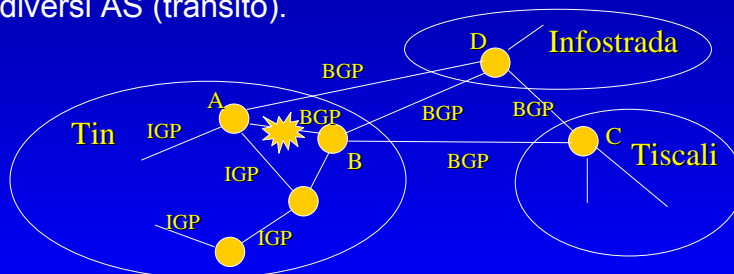
Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.48

Instradamento

EGP - Border Gateway Protocol (BGP)

- Anche se tutti i *router* all'interno di un AS sono fra loro cooperativi, non è detto che i nodi che interconnettono due AS si “fidino” l'uno dell'altro.
- Questo accade in quanto gli AS sono in genere controllati o posseduti da organizzazioni diverse e quindi l'instradamento deve dipendere anche dagli accordi fra i diversi AS (transito).



Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.49

Instradamento

EGP - Border Gateway Protocol (BGP)

- In quanto protocollo di tipo EGP, il BGP differisce dai protocolli IGP in diversi aspetti:
 - Politica: fra AS nella scelta del percorso la “politica” domina, ossia le scelte dipendono da più fattori che coinvolgono considerazioni strategiche, economiche e di sicurezza più che tecniche. Queste considerazioni sono specifiche di ogni AS, quindi la scelta è principalmente sotto il controllo amministrativo.
 - Scala: la scalabilità è molto importante perché le “reti” coinvolte sono generalmente grandi. Invece se un AS cresce eccessivamente lo si può sempre dividere in due.
 - Prestazioni: le prestazioni tecniche contano relativamente poco in un EGP.

Complementi di Reti e Sistemi R. Bolla, C. Nobile

5.50