



# **OSPF**

**(Open shortest path first)**

# Plan

- **Introducción al protocolo**
- **Visión de alto nivel**
- **Ruteo jerárquico**
- **Detalles del protocolo**
- **OSPF para IPv6**
- **Ejemplo de configuración**

# OSPF (v2)

- Protocolo de ruteo interior
- Estado de enlace (Link State)
- RFC 2328 (STD 54) 1998
- Versiones anteriores: RFC 2178 ,RFC 1583, RFC 1247
- Jerárquico
- Soporta balanceo de tráfico por múltiples caminos de igual peso

# Estado de enlace

- Idea: Construir el grafo de la red
- A partir del grafo, calcular mejores rutas a todos los destinos
- Paso 1: Descubrimiento de topología local
- Paso 2: Inundación confiable de la información a toda la red
- Paso 3: Construcción del “grafo” de la red y cálculo de los caminos más cortos
- Paso 4: Construcción de la tabla de ruteo

# Descubrimiento de topología local en OSPF

- El enrutador conoce (por configuración) las interfaces a las que está conectado y las direcciones correspondientes
- Descubre (protocolo Hello) a sus vecinos
- Tiene la métrica de utilizar cada interfaz
- Sabe a qué redes llega directamente
- Con esta información, genera la información que lo describe

# Base de datos de estado de enlace

- La información de estado de enlace se anuncia mediante lo que se conocen como LSA: Link State Advertisement (anuncios de estado de enlace)
  - Descripción del estado de una pieza de la red (un router o una red)
  - Router: estado de las interfaces y adyacencias
  - Red: IP/máscara y routers conectados
  - Más detalles más adelante
- Cada enrutador guarda una lista de todos los LSA recibidos (base de datos de estado de enlace)

# ¿Qué informa cada enrutador sobre sí mismo?

- Identificación del enrutador (router-id)
  - Identifica unívocamente al enrutador
  - Dirección IP, usualmente IP de la loopback de menor número (NO 127.0.0.1), de lo contrario, menor IP del enrutador
- Para todos los enlaces (interfaces):
  - Tipo de enlace
  - El costo (métrica) de usar la interfaz
  - A qué está conectado el enlace (otro enrutador, una subred, una red de tránsito...)
- Otros parámetros

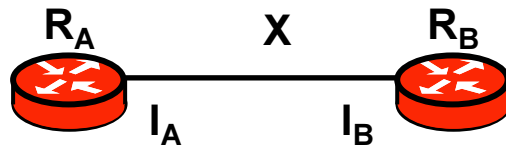
# Tipos de enlaces

- **Para OSPF:**
  - **Punto a Punto**
    - Ejemplo: línea serial, PVC ATM
  - **Broadcast**
    - Ej.: Ethernet
  - **NBMA (acceso múltiple sin broadcast)**
    - Ej.: ATM, Frame Relay (mallados)
  - **Punto – Multipunto**
    - Conectividad “Uno a muchos”
  - **Links virtuales**

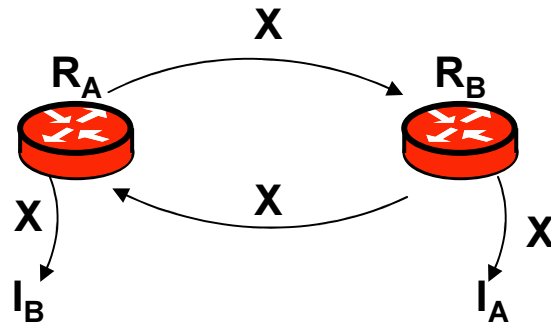


# Tipos de enlaces para OSPF: Punto a punto

## Enlace Punto a Punto



## Grafo correspondiente

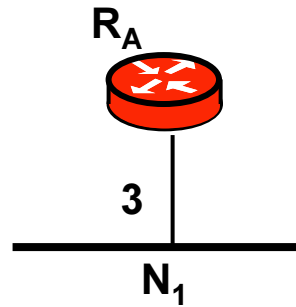


Hacia

|    |   | Desde |    |
|----|---|-------|----|
|    |   | RA    | RB |
| RA |   |       | X  |
| RB | X |       |    |
| IA |   |       | X  |
| IB | X |       |    |

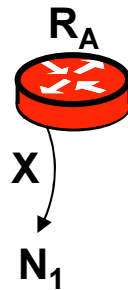
Las interfaces pueden no estar numeradas

# Redes “stub”

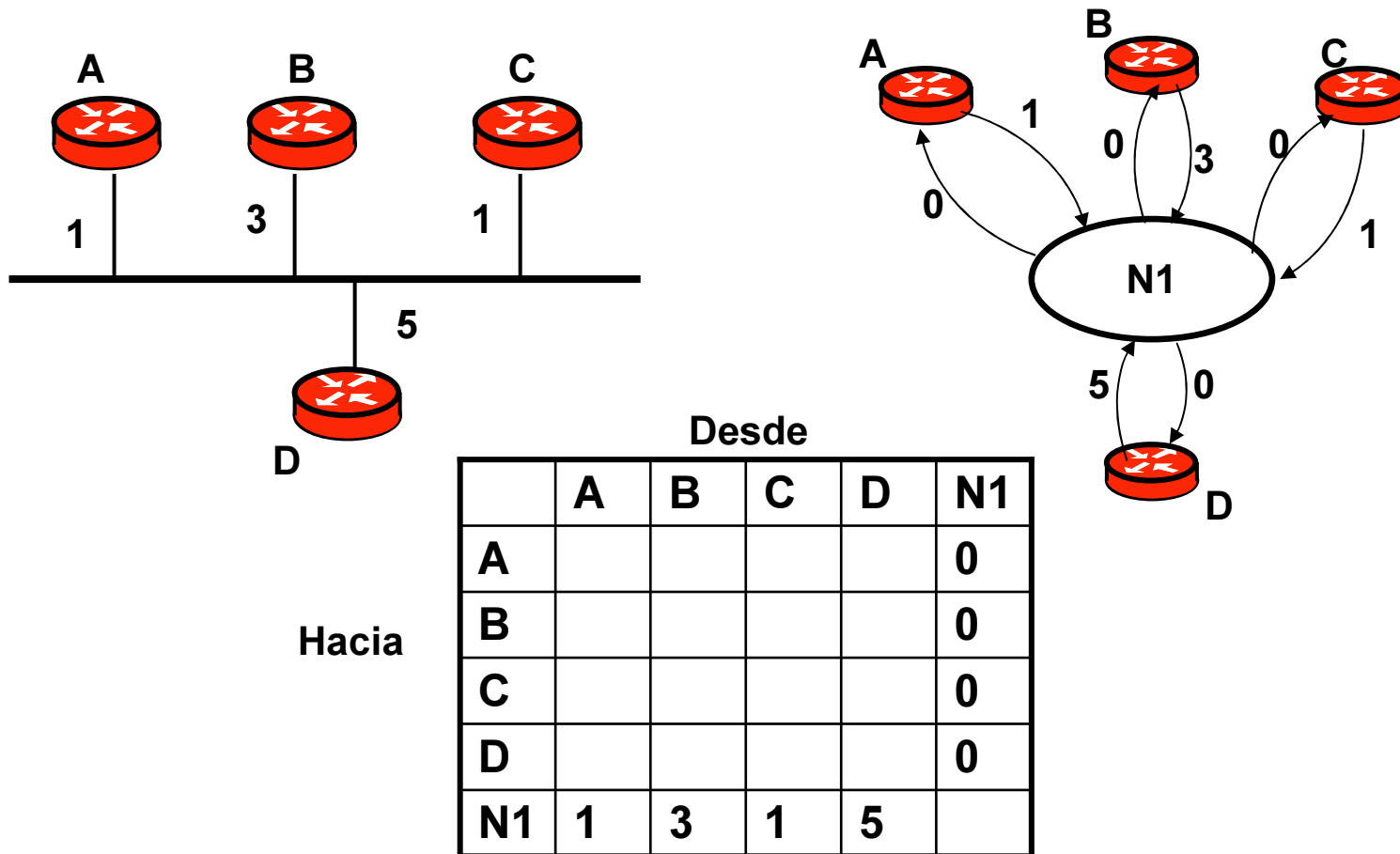


Hacia

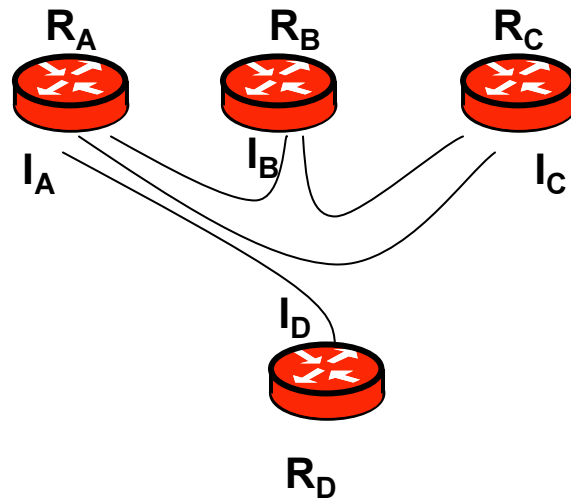
|       |     | Desde |       |
|-------|-----|-------|-------|
|       |     | $R_A$ | $N_1$ |
| $R_A$ |     |       |       |
| $N_1$ | $X$ |       |       |



# Redes multiacceso (broadcast o NBMA)



# Redes Punto-Multipunto



Hacia

Desde

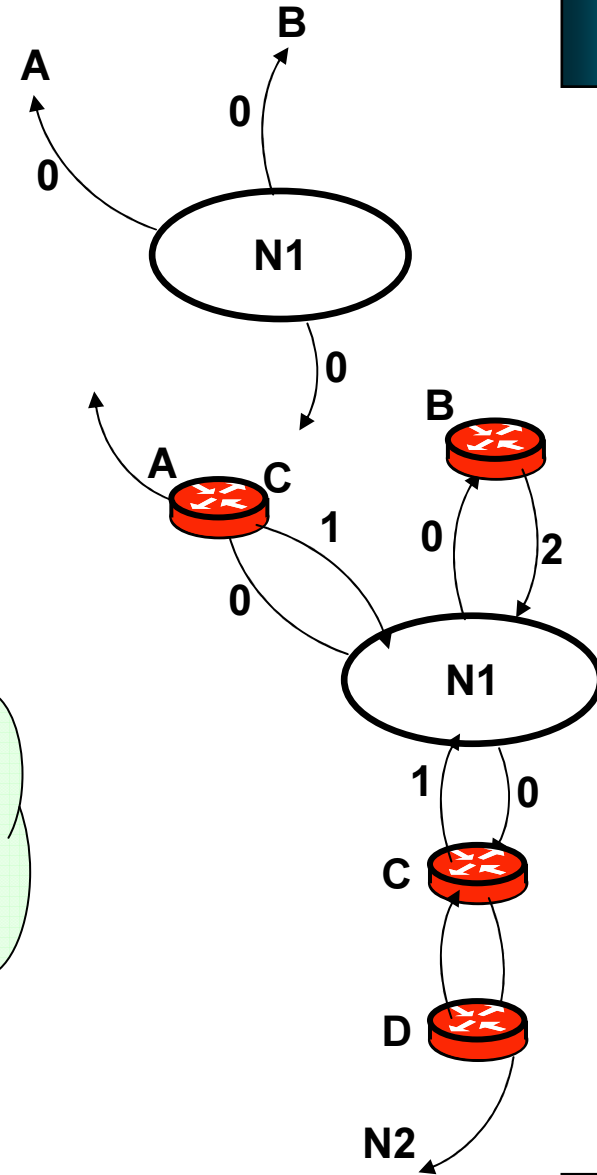
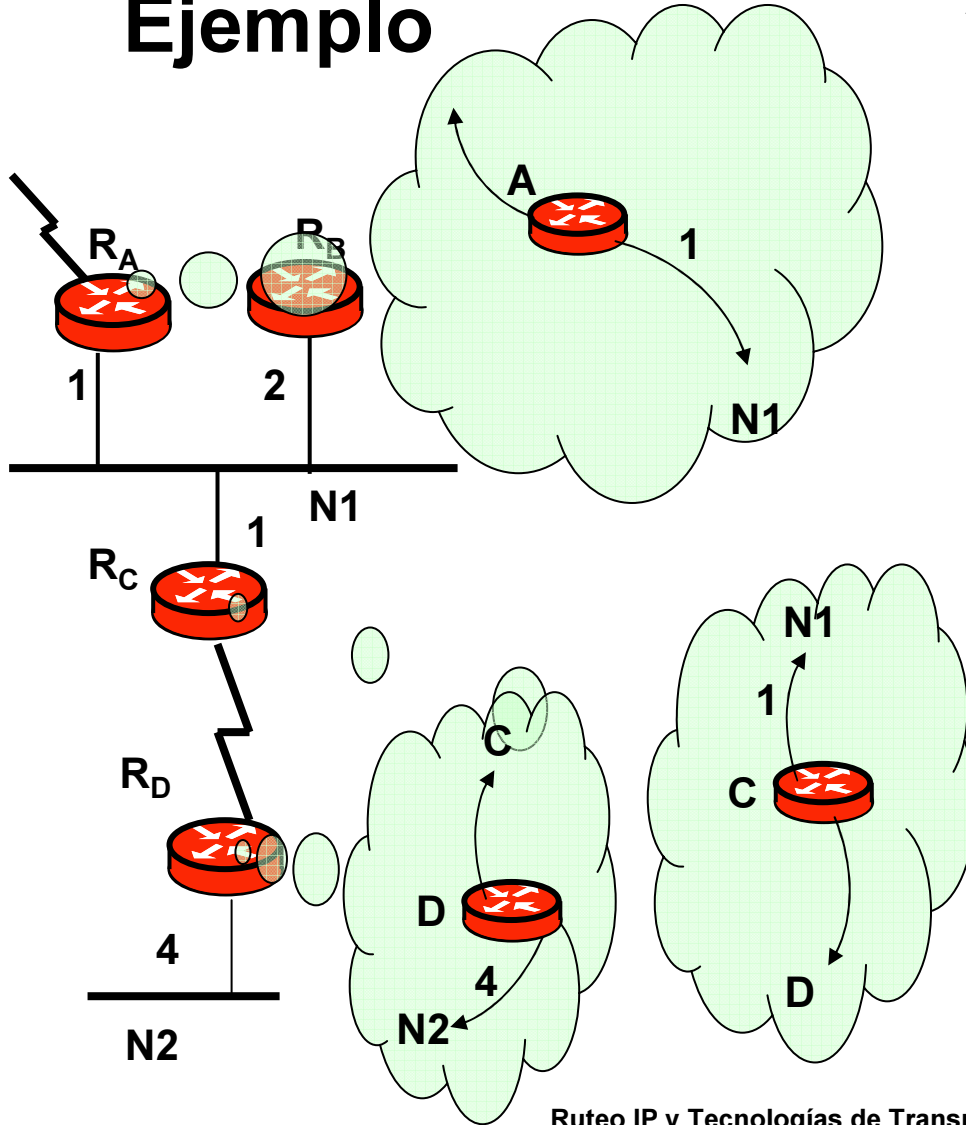
|                | R <sub>A</sub> | R <sub>B</sub> | R <sub>C</sub> | R <sub>D</sub> |
|----------------|----------------|----------------|----------------|----------------|
| R <sub>A</sub> |                | X              | X              | X              |
| R <sub>B</sub> | X              |                | X              |                |
| R <sub>C</sub> | X              | X              |                |                |
| R <sub>D</sub> | X              |                |                |                |
| I <sub>A</sub> | X              |                |                |                |
| I <sub>B</sub> |                | X              |                |                |
| I <sub>C</sub> |                |                | X              |                |
| I <sub>D</sub> |                |                |                | X              |

Se utilizan las redes punto-multipunto en general cuando no se tiene conectividad “todos con todos”

Ejemplo: múltiples PVC Frame-relay o ATM (no malla completa)

Se trata como múltiples links punto a punto

# Ejemplo



# Cálculo de la mejor ruta

- Se construye el grafo con las piezas que envía cada enrutador (LSA)
- Teniendo el grafo de la red, cada nodo aplica el algoritmo de Dijkstra
- Se obtiene un árbol, con raíz en el nodo, de los caminos de menor peso a todos los nodos de la red
- Del árbol se extraen las mejores rutas a cada destino
- En la tabla de enrutamiento se pone solamente el próximo salto

# Métrica en OSPF

- Valor sin dimensión, entero
- Se calcula utilizando 24 bits
- Infinito: 0xffffffff
- El costo de cada interfaz se representa con 16 bits
- Asignado administrativamente
- Por defecto, los equipos asignan un costo
  - Ej. Cisco (y otros):  $100.000.000/bw$
  - bw: ancho de banda nominal de la interfaz

# Ruteo Jerárquico

- Objetivos:
  - Simplificar la tabla de rutas
    - Mediante la agregación en la frontera entre “zonas”
  - Simplificar la topología que debe conocer cada enrutador
- Dificulta algunas aplicaciones (ingeniería de tráfico)



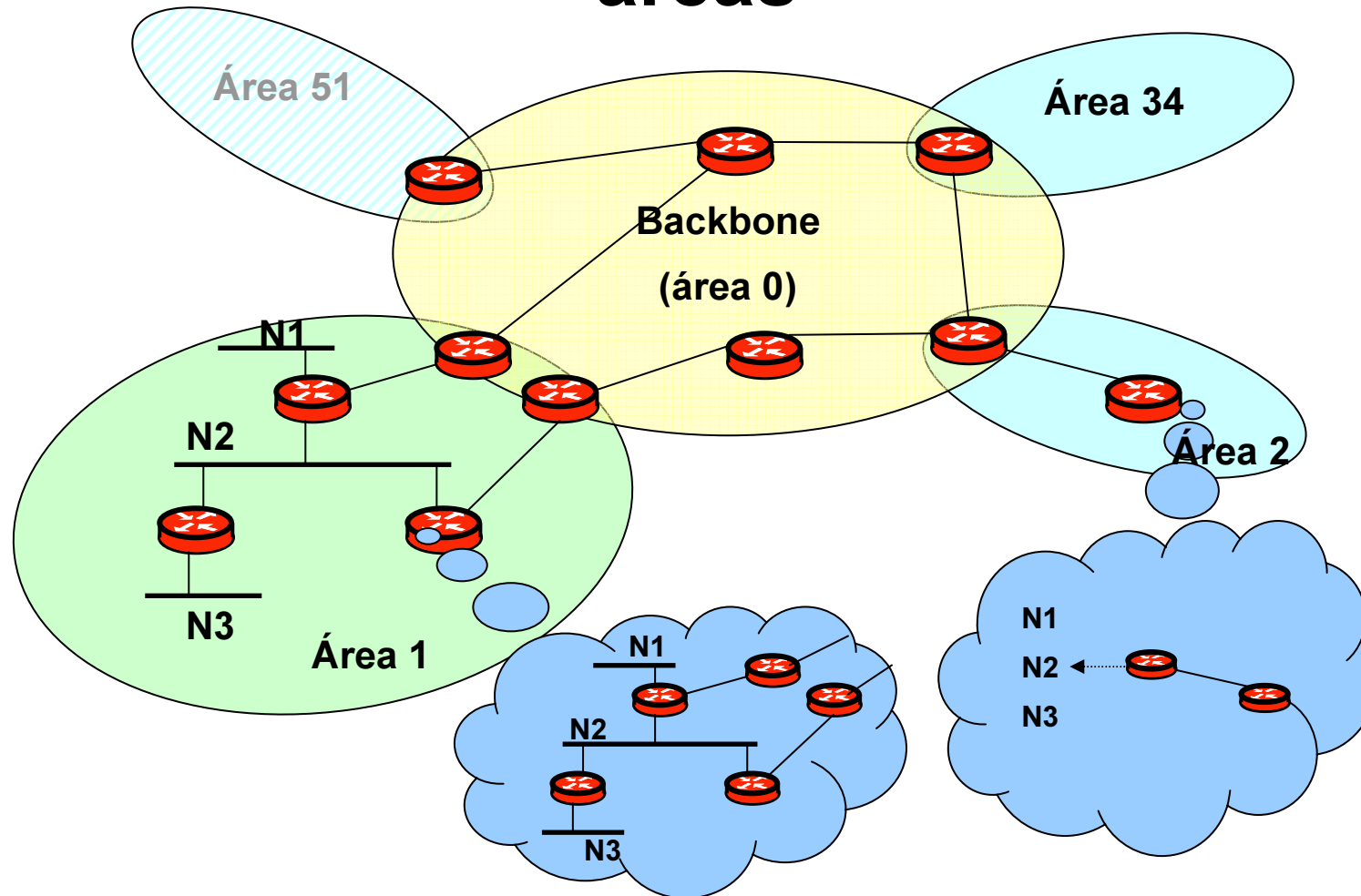
# Ruteo Jerárquico

- En OSPF, 2 niveles
- Basado en áreas
- Identificación de Área: número de 32 bits
- Área 0: backbone
- El tráfico entre áreas debe pasar por el backbone
- El backbone debe ser conexo (debe haber una única “area 0”)

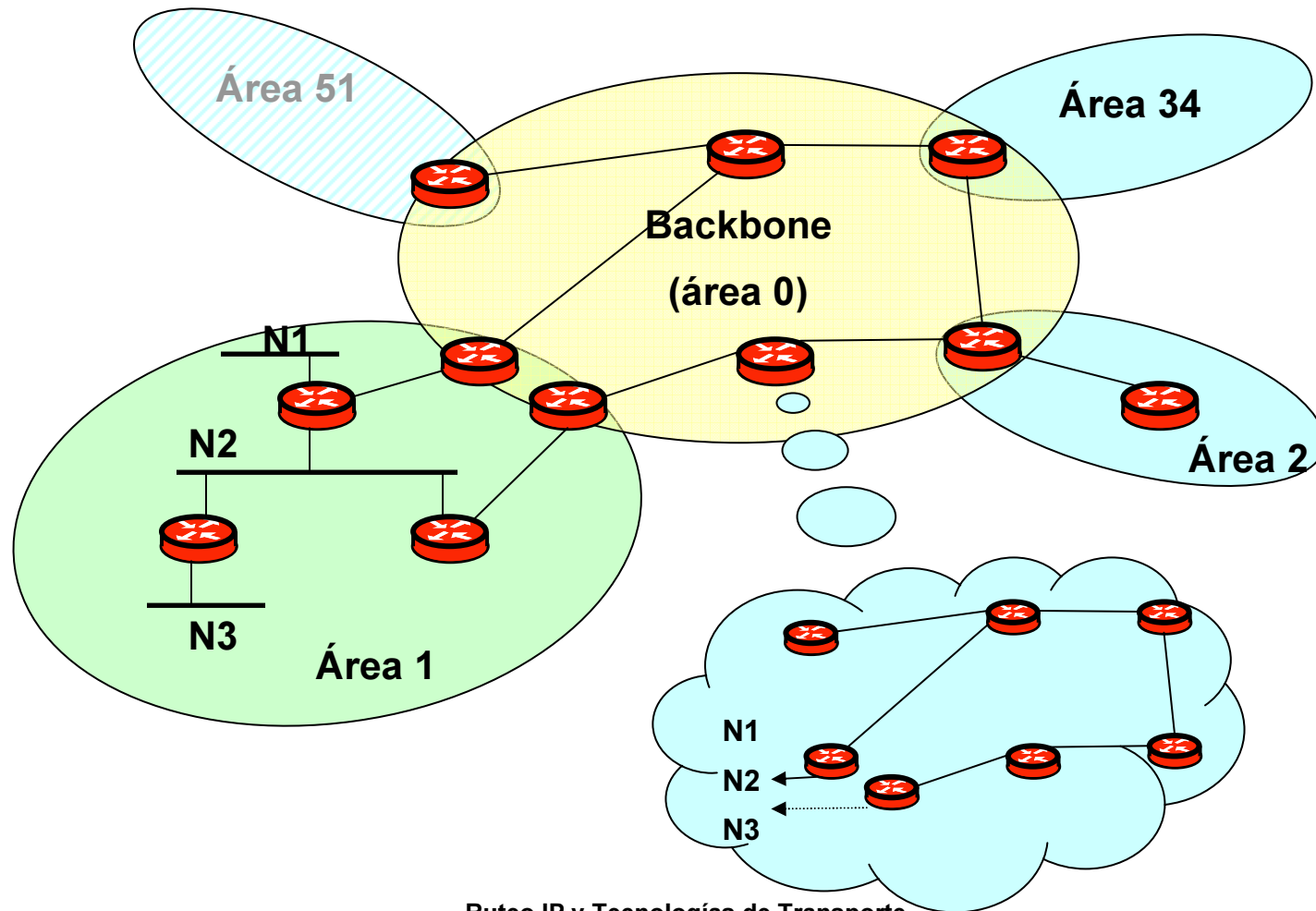
# Ruteo jerárquico

- Enrutadores en un área no conocen la topología externa al área
- La topología del área no se conoce fuera de ella
- Los enrutadores tienen una base de datos de estado de enlace por cada área a la que estén conectados
- Entre áreas solo se propaga las redes alcanzables y la métrica para alcanzarlas

# Visión topológica de las distintas áreas



# Visión topológica de las distintas áreas



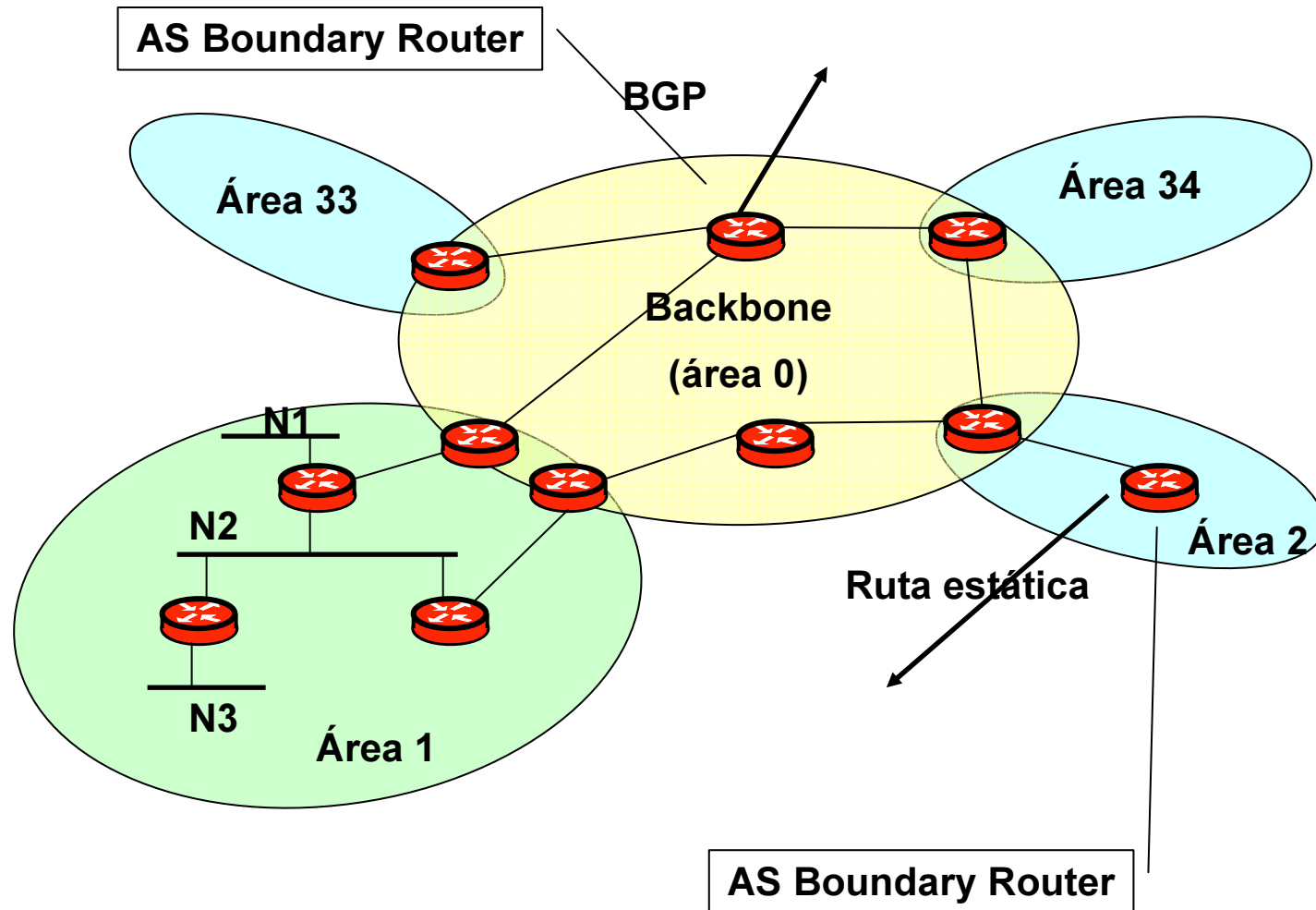
# Clasificación de enrutadores

- Enrutadores internos
  - Pertenece a una sola área
- Enrutadores de borde de área
  - Pertenece a más de un área
- Enrutadores de backbone
  - Todos los del área 0 (incluye los de borde de área)
- Enrutadores de borde de AS (sistema autónomo)
  - Inyectan en OSPF información proveniente de otros protocolos, por ejemplo BGP

# Información externa

- Proveniente de otros protocolos (o estática)
- Puede incluir ruta por defecto
- 2 tipos según la métrica:
  - tipo 1: métrica comparable con el costo de OSPF
  - tipo 2: métrica estrictamente mayor que cualquier métrica de OSPF
- Se inunda a todo el dominio OSPF

# Información externa



# Caminos inter áreas

- 3 tramos:
  - Del origen al enrutador de borde de área
  - En el backbone entre enrutadores de borde de área
  - Interno al área de destino



# ¿Por qué la comunicación entre áreas debe pasar por el backbone?

- Entre áreas no se intercambia información de topología
- Si permitiéramos comunicación entre áreas arbitrarias, podríamos tener los mismos problemas de conteo a infinito que con vector distancia
  - Al pasar por el área 0, forzamos una estructura de árbol sin loops

# Continuidad del backbone

- Si el backbone es discontinuo: áreas no podrán comunicarse
  - Las áreas distintas de “0” no hacen tránsito
- Solución: Links virtuales
  - Enlace virtual punto a punto a través de un área no backbone
- Área de tránsito: si un link virtual pasa por ella

# Agregación

- Puede configurarse (en el enrutador de borde de área) los rangos de IPs del área
- Se puede configurar que fuera del área, solo se propague la ruta correspondiente al rango configurado, y no las rutas específicas

# Paquetes OSPF

- OSPF funciona directamente sobre IP
  - Protocolo 89
- En aquellas redes que lo permiten, se utiliza multicast para el envío de mensajes
  - Redes con capacidad de broadcast (ej. Ethernet)
  - Redes punto a punto
  - Se evita configurar los vecinos
- En aquellas redes sin capacidad de multicast, deben configurarse los vecinos
- 2 direcciones de multicast reservadas para OSPF
  - 224.0.0.5 – AllSPFRouters
  - 224.0.0.6 – AllDRRouters

# Tipos de mensaje OSPF

- 5 tipos de paquete, con un encabezado común
  - Hello
  - Database Description
  - Link State Request
  - Link State Update
  - Link State Ack

# Encabezado OSPF

## Común a todos los paquetes

|                        |                       |                  |
|------------------------|-----------------------|------------------|
| Versión                | Tipo                  | Largo de paquete |
| Router Id              |                       |                  |
| Area Id                |                       |                  |
| Checksum               | Tipo de autenticación |                  |
| Datos de Autenticación |                       |                  |
| Datos de Autenticación |                       |                  |

# Autenticación OSPF

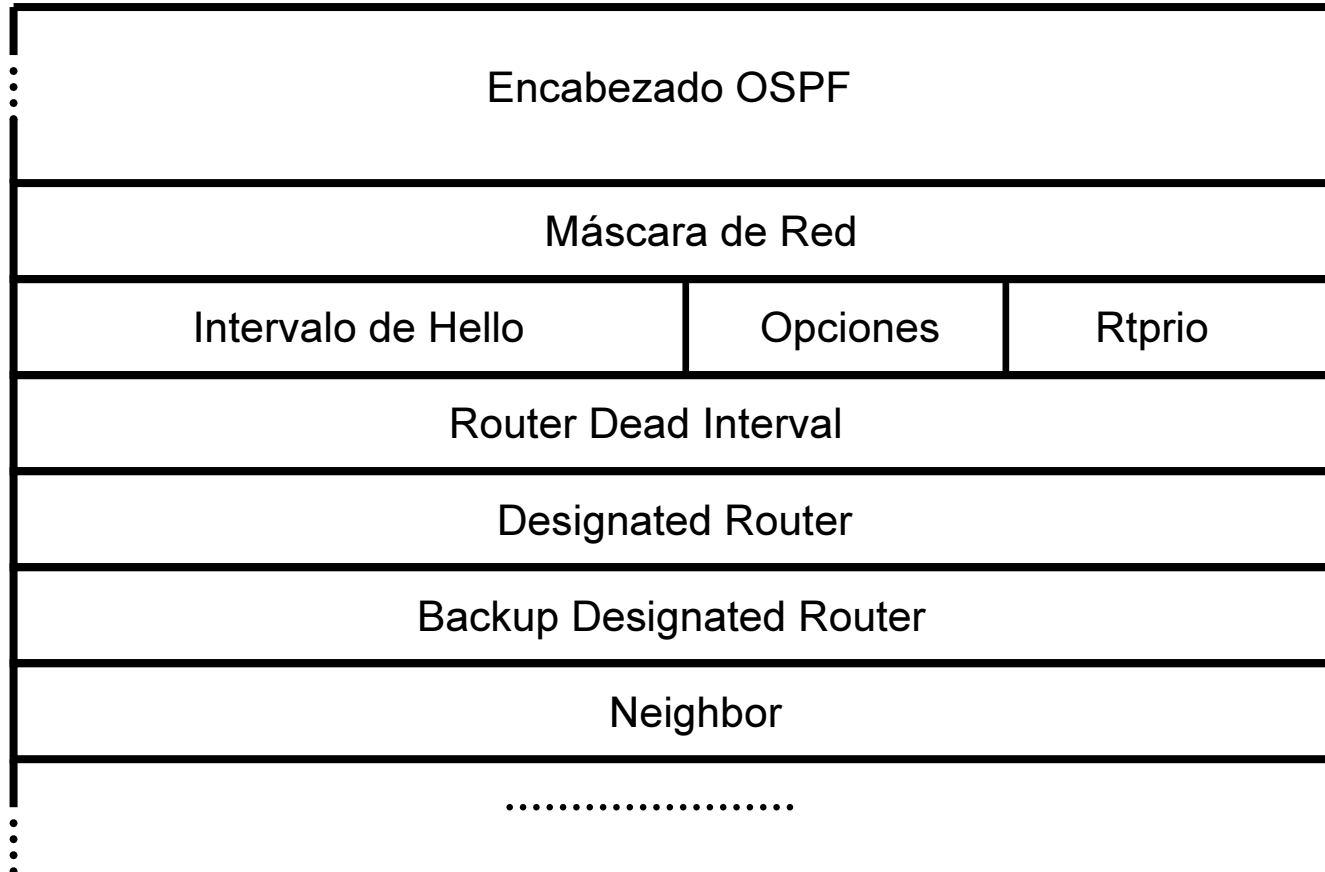
- 3 tipos definidos:
  - Null authentication (No autentifica!!!)
  - Simple password (password en claro!!!)
    - Se compara la password en el paquete con la configurada
  - Cryptographic authentication
    - Se envía un “message digest” (ej. MD5) del mensaje + un secreto compartido
    - En el campo de autenticación se describe la clave, largo del digest, y se agrega un número de secuencia (para evitar replay)

# Protocolo HELLO

- Envía paquetes de descubrimiento (HELLO) periódicamente (por defecto HelloInterval=10s)
- Descubrimiento de vecinos
- Anuncio de parámetros
- Establecimiento de comunicación bidireccional
- Generación de adyacencias
- En redes Broadcast y NBMA, elige DR
- Keepalive



# Paquete de Hello



# Campo de Opciones

- DC bit: soporte de OSPF en circuitos a demanda
- EA bit: Soporte de External Attributes (type 8) LSAs
- N bit: Soporte de NSSA LSAs
- P bit: ABR debería traducir LSAs tipo 7 en tipo 5
- MC: MOSPF (multicast)
- E: Soporte para AS External LSAs. Apagado en areas STUB
- T: Soporte de TOS

# HELLO (cont.)

- En redes Punto a punto y broadcast:
  - Enviado a 224.0.0.5 (AllSPFRouters)
- En redes sin capacidad de Multicast (NBMA): unicast (debe conocer los vecinos)
- Formación de Vecindades:
  - Solamente si los parámetros en el HELLO son iguales (máscara, RouterDeadInterval, Intervalo de Hello, área, Autenticación, Opciones)
  - (si no se forma una adyacencia, revisar parámetros!!!)

# Hello (cont.)

- Al recibir el mensaje hello, si aparezco como vecino, sé que la comunicación es bidireccional
- Adyacencia: comunicación punto a punto entre ciertos vecinos
- Se rompe si no recibo Hello en RouterDeadInterval (por defecto  $4 * \text{HelloInterval}$ )

# DR: Designated Router

- Solo en redes Broadcast y NBMA
- Se elige un router para propagar el LSA de la red de la LAN y minimizar anuncios
- Es adyacente de los demás routers de la LAN
- También se elige un backup D.R.
- DR y Backup DR mandan paquetes a 224.0.0.5 (AllSPFRouters)
- Los demás a 224.0.0.6 (AllDrRouters)

# Elección del D.R.

- Al ingresar un nuevo enrutador:
  - Si los paquetes hello recibidos contienen un D.R. y BDR, aceptarlos
  - Si no hay BDR elegido, router con mayor prioridad se convierte en BDR
    - Empate se define con el mayor Router ID
  - Si no hay DR, el BDR es promovido a DR
  - Se elige BDR nuevamente
- Estados posibles de una interfaz Broadcast: DR, BDR, DROther

# Deshabilitando elección del DR

- Hoy en día, es común tener enlaces ethernet “punto a punto”
  - Dos enrutadores conectados directamente con una capa de enlace Ethernet
- Elección del DR y generación de LSA de la red superfluos
- En estos casos (si el equipo lo permite) podemos configurar las interfaces como “punto a punto”
  - OJO, deben estar configurados igual en ambos extremos

# Formación de adyacencias

- Líneas punto a punto: los vecinos siempre se convierten en adyacentes
- Redes Broadcast/NBMA: todos los routers se convierten en adyacentes del D.R. y el Backup D.R.
- La información de estado de enlace se intercambia entre nodos adyacentes



# Sincronización inicial de la base de datos de estado de enlace

- Solo al inicializarse una adyacencia
- Relación maestro/esclavo (router con mayor router-id es maestro)
- Vecinos intercambian resúmenes de los LSA que disponen (Database Description Packets)
  - Age, Options, Type (of LSA)
  - Link State Id.
  - Advertising Router, Sequence Id.

# Sincronización (cont.)

- Determinan qué LSAs no tienen
- Solicitan los LSA que no conocen
  - Paquetes de LSA Request
- Son enviados en LSA Updates
- Deben ser reconocidos (LSA Ack)
  
- Una vez completado, la adyacencia está funcional

# Comparación de LSAs

- Entre distintos elementos de la red, comparo la identificación y el tipo
- Importante poder determinar si 2 LSA recibidos correspondientes al mismo elemento de la red corresponden a la misma información, o si uno de ellos tiene información más reciente
- Objetivo primordial: que todos los enrutadores tengan la misma visión de la topología de la red

# Comparación de LSAs: datos disponibles en cada LSA

- Número de secuencia: 32 bits con signo
  - Se incrementa con cada actualización
- Edad: en segundos, inicializada en 0
  - Entero sin signo (16 bits)
  - Se incrementa al reenviar y con paso del tiempo
- Edad máxima: MaxAge = 3600
- El LSA se descarta si pasa de MaxAge
  - Antes de descartar se inunda nuevamente

# Algoritmo de comparación de LSAs

- Número de secuencia mayor es más nuevo
- Mayor checksum es “más nuevo”
  - Desempate (\*)
- Si tiene AGE en MaxAge, es más nuevo
- Si las edades difieren en más de MaxAgeDiff, la menor es más nueva
- De lo contrario son iguales

# Comentarios

- (\*) Si recibimos 2 LSA que deberían ser iguales, con distinto checksum, uno de ellos “está mal”. Lo importante es que todos seleccionen el mismo
  - Quien lo origina se dará cuenta y lo generará nuevamente
- Considerar más nuevo un anuncio expirado, tiene como consecuencia que rápidamente todos los enrutadores descarten ese anuncio

# Inundación confiable de LSAs

- Se hace enlace a enlace
- Mediante paquetes de “Link State Update”
- Cada paquete de actualización puede contener varios LSAs
- Cada LSA debe ser reconocido
- Se pueden agrupar varios reconocimientos en un Link State Acknowledge Packet

# Inundación de LSAs (cont)

- Al recibir un LSA más nuevo que la copia disponible (o al generar uno nuevo)
- Se debe inundar por todas las interfaces (excepto por la que se recibió si no somos DR)
  - Se agrega a la lista de retransmisión para cada vecino (que no sepamos que lo tiene)
  - Se inunda en las interfaces con vecinos que no nos lo haya enviado



# Inundación (cont.)

- En interfaces Broadcast o NBMA:
  - Si somos DR -> se inunda a AllSPFRouters
  - Si no -> se inunda a AllDRouters
- Se retransmite hasta que veamos un ACK del vecino
  - Si el vecino nos lo retransmite se considera un ACK
- Todo LSA recibido debe ser reconocido

# Tipos de LSA (RFC 2328)

- 1. Router
- 2. Network
- 3. Network Summary
- 4. ASBR Summary
- 5. AS External



# Otros tipos de LSA

- 6. Group Membership
- 7. NSSA External
- 8. External Attributes
- 9. Opaque (link-local scope)
- 10. Opaque (area-local scope)
- 11. Opaque (AS scope)

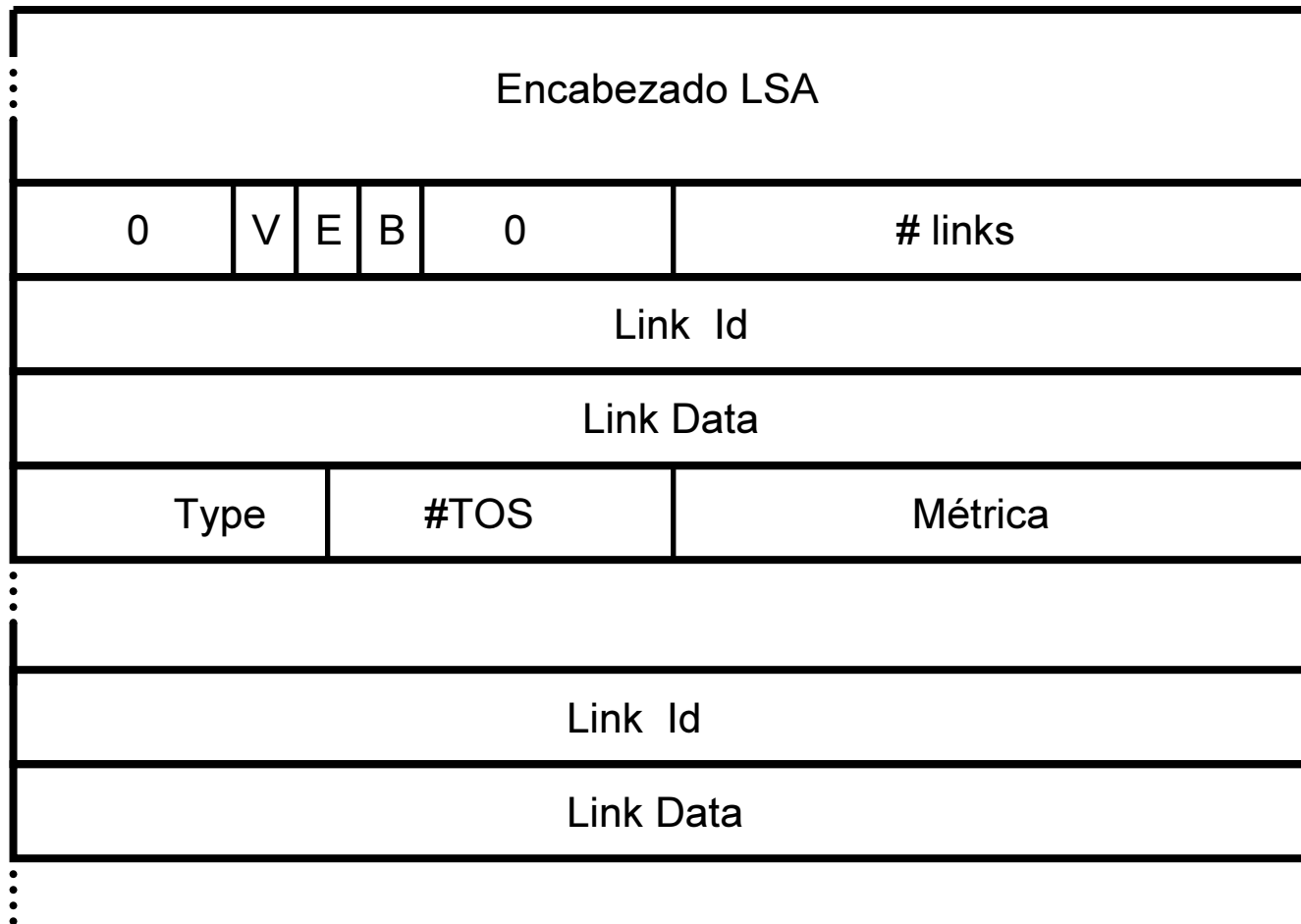
# Encabezado de un LSA

|                     |          |      |
|---------------------|----------|------|
| LS Age              | Opciones | Tipo |
| Link State Id       |          |      |
| Advertising Router  |          |      |
| Número de secuencia |          |      |
| LS Checksum         | Largo    |      |

# Router LSA

- Generado por cada router
- Incluye
  - Identificación del router (Link State Id)
  - Lista de todos los enlaces
  - Identificación de interfaz
  - Tipo de enlace
  - Métrica

# Router LSA



# Router LSA

- Link State Id (en el encabezado del LSA):  
Router Id
- V: router tiene links virtuales
- E: router es AS-boundary
- B: área border router

# Router LSA (cont)

- Para cada link:
  - Tipo
    - Punto a punto
    - Transit network
    - Stub network
    - Virtual
  - Link Id (a quién se conecta)
    - Si es a un router o red de tránsito ID
    - Caso contrario, red/máscara



# Router LSA (cont)

–Link data:

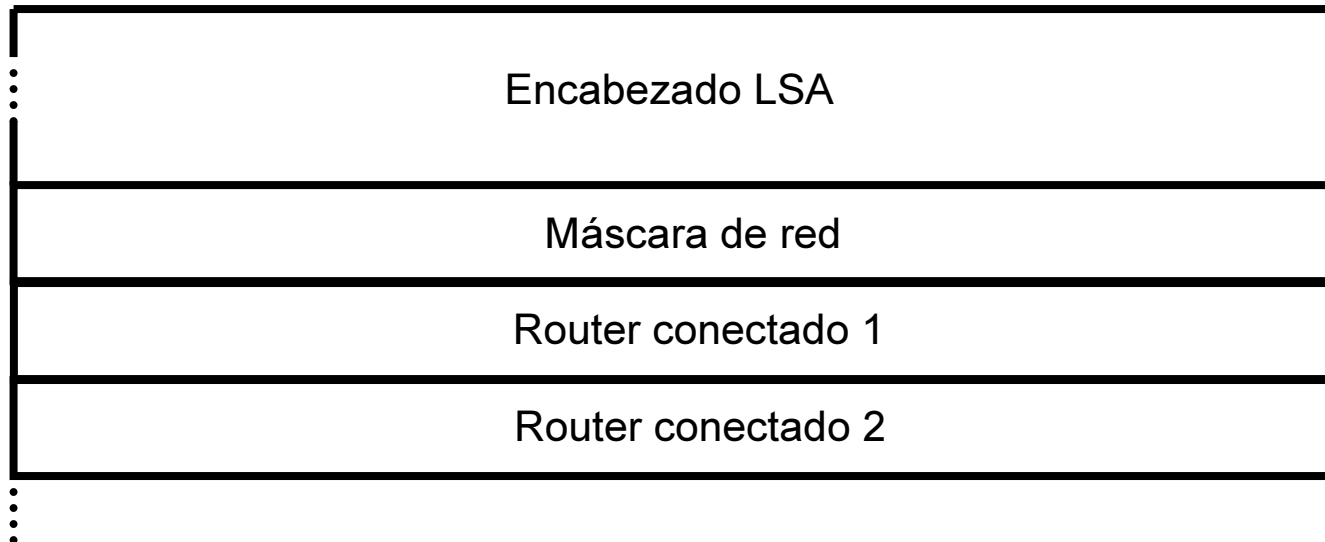
- para redes stub, la máscara
- para punto a punto no numeradas: identificador de la interfaz
- para el resto: IP del router en la interfaz

–Métrica

# Network LSA

- Solo para redes multiacceso
- Generado por el DR
- Representa la red multiacceso en el grafo
- Dirección de red y máscara
- Direcciones de los routers conectados

# Network LSA



- Router conectado: Router-id de cada uno de los routers adyacentes al DR

# Summary LSAs (Tipos 3 Y 4)

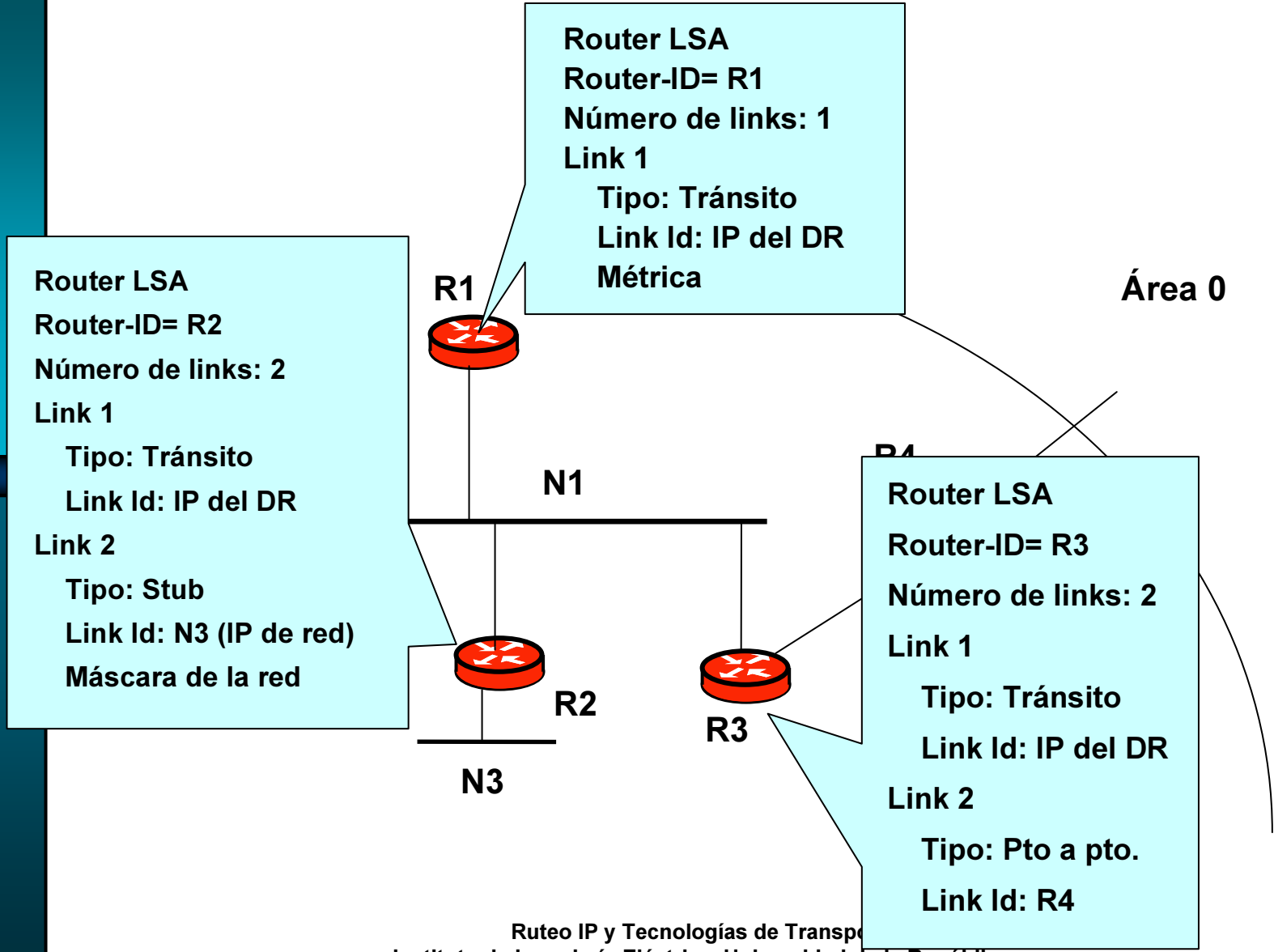
- Tipo 3: descripción de una red hacia otra área
- Link State Id: Dirección de red
- máscara de la red
- métrica
- Tipo 4: descripción de AS boundary router
- Link State Id: Router-id

# LSA tipo Summary

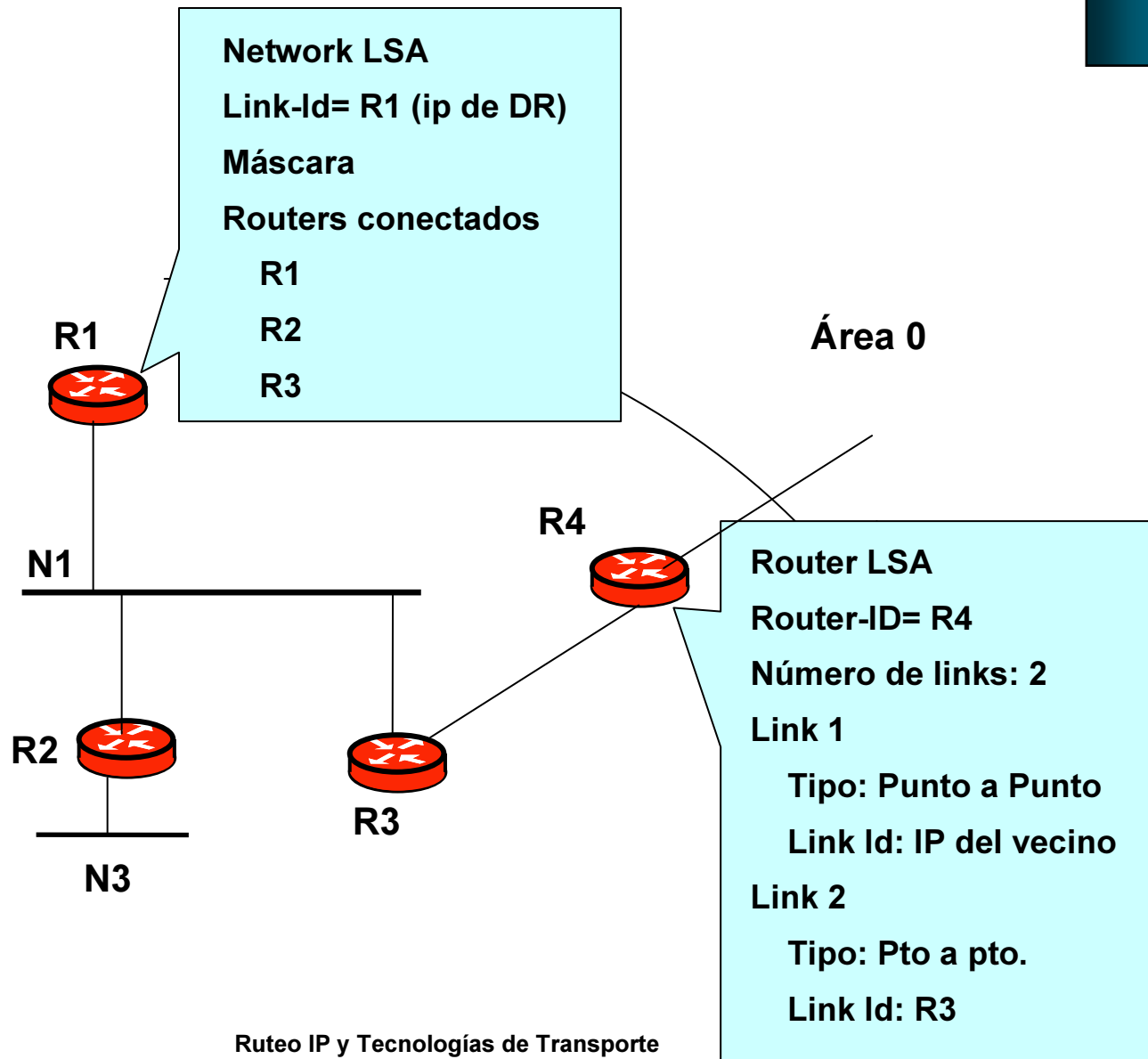
|                     |            |      |
|---------------------|------------|------|
| LS Age              | Opciones   | Tipo |
| Link State Id       |            |      |
| Advertising Router  |            |      |
| Número de secuencia |            |      |
| LS Checksum         | Largo      |      |
| Network Mask        |            |      |
| 0                   | Métrica    |      |
| TOS                 | TOS Metric |      |
| ⋮                   | ⋮          |      |

# AS-External-LSA (tipo 5)

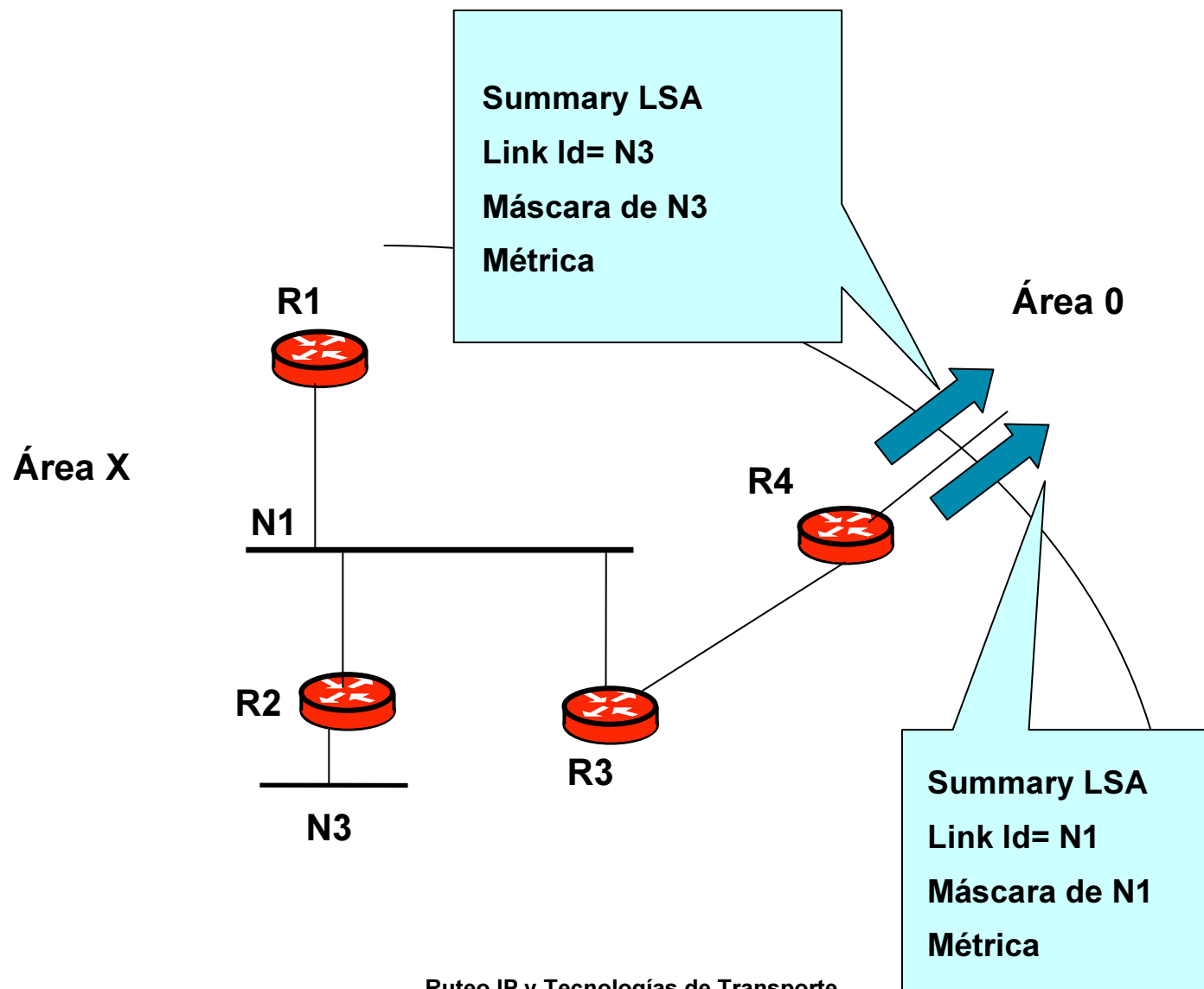
- Describe destinos externos al AS (externos a OSPF)
- Red
- Máscara
- Forwarding address (puede ser distinto del Advertising Router)
- Tipo de métrica externa
  - Tipo 1: comparable con OSPF
  - Tipo 2: estrictamente mayor que métrica OSPF



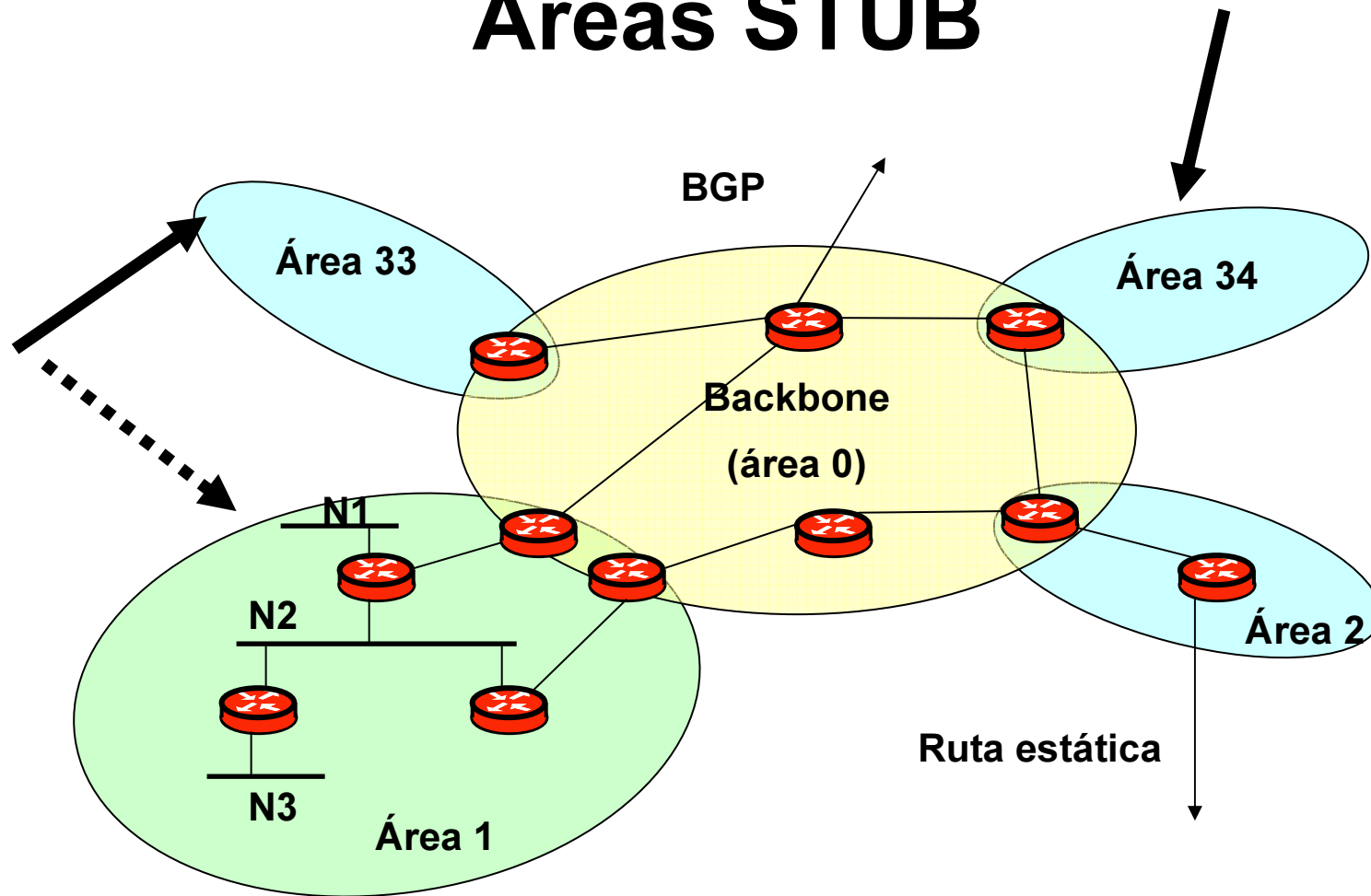
Área X







# Áreas STUB



# Áreas STUB

- “Simplificación” para áreas con “una” conexión al backbone
- Innecesario conocer información externa al AS y de otras áreas, si tenemos una sola salida del área
- Solo se publica una ruta por defecto hacia el área
- No puede haber links virtuales que pasen por un área stub
- No se permiten LSA tipo 5 (external)

# LSA tipo 7. NSSA External

- NSSA - Not So Stubby Area
- Objetivo: generar información externa en un área STUB
- Se utiliza un nuevo tipo de LSA (tipo 7), ya que el 5 está prohibido en áreas STUB
- Saliendo del área STUB, el enrutador de borde puede ser configurado para descartarlo o convertirlo a tipo 5

# OSPF v3 (OSPF para IPv6)

- RFC 5340 (Propuesta de standard, Julio 2008)
  - Antes RFC 2740 (Propuesta de standard)
- Mantiene el funcionamiento general
  - (inundación de LSAs, elección de DR, áreas,etc)
- Separa mejor la información de topología, de la información de direcciones
- La autenticación se deja a IP
- Se agregan LSAs para describir las subredes IPv6

# OSPF v3 (cont.)

- Corre “por enlace”, no por subred
- Router-id: 32 bits (número arbitrario)
- Área: 32 bits
- Los enrutadores se identifican siempre por el “router-id”
- El alcance de un LSA se codifica en el propio LSA. Link, Area o AS

## OSPF v3 (cont.)

- Para comunicación sobre un enlace, se utilizan direcciones link-local
- Router-LSA y Network-LSA no llevan información de direcciones
- En Network-LSAs y Link-LSAs el Link state ID es un identificador de interfaz del propio enrutador
- Los prefijos se expresan como prefijo/largo (y no máscara)

# OSPF v3 (cont.)

- Nuevos tipos de LSA:
  - Intra-Area-Prefix-LSA: lleva la información de prefijos
  - Link-LSA: informa a los vecinos las direcciones link-local del router en el link, y los prefijos asociados con el link



# Algunas Referencias

- Muy básico: Tanenbaum, 4ta. Edición
- RFC 2328
- [http://www.cisco.com/en/US/tech/tk365/tk480/tsd\\_technology\\_support\\_sub-protocol\\_home.html](http://www.cisco.com/en/US/tech/tk365/tk480/tsd_technology_support_sub-protocol_home.html)
- <http://www1.cs.columbia.edu/~ji/F02/>

# Ejemplo: Configuración Cisco

```
interface ethernet 0
```

```
  ip address 192.168.0.1 255.255.255.0
```

```
  ip ospf authentication-key 1234
```

```
Interface ethernet 1
```

```
  ip address 172.16.0.0 255.255.255.128
```

```
  ip ospf cost 20
```

```
router ospf 1
```

```
  network 192.168.0.0 0.0.0.255 area 0
```

```
  network 172.16.0.0 0.0.1.255 area 1
```

```
  network 10.0.0.0 0.0.255.255 area 10.255.255.1
```

```
  area 1 stub
```

```
  area 0 authentication
```

```
  redistribute rip metric 1 metric-type 2
```