

Algorithmes de routage IP

② routage intra-domaine: quelques généralités sur les protocoles RIP et OSPF

Filière F5 - Isima 2007-2008

Mickael Meulle

mickaelmeulle@gmail.com

Michael.meulle@orange-ftgroup.com

Plan

▶ 3. Le Protocole RIP

- ▶ 3.1 présentation: RIP est un protocole VD
- ▶ 3.2 les échanges de messages
- ▶ 3.3 l'effet rebond des protocoles à vecteur de distance
- ▶ 3.4 quelques remarques sur RIP

▶ 4. Le protocole OSPF

- ▶ 4.1 présentation: OSPF est un protocole LS
- ▶ 4.2 les échanges de messages
- ▶ 4.3 fonctionnement détaillé et remarques

3. Le protocole RIP

- 3.1 présentation: RIP est un protocole VD
- 3.2 les échanges de messages
- 3.3 l'effet rebond des protocoles à vecteur de distance
- 3.4 quelques remarques sur RIP

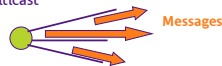
3.1 RIP

« Routing Information Protocol »

- Protocole intra-domaine : IGP (« Interior Gateway Protocol »)
 - Rfc 1058,2453
- Conçu à l'origine pour les LAN, mais son utilisation a été étendue aux réseaux distants
- Version 1.0 et 2.0
- Type « Ford-Fulkerson » ou encore « Bellman-Ford »
- A été très utilisé par le passé

Protocole à Vecteur de distance

- Inondation de messages
- Fonctionne en mode unicast ou multicast
- Itératif
- Distribué
- Non synchronisé



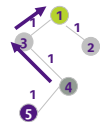
Algorithme de routage
Isma 2007-2008

DE

3.1 RIP est un protocole VD avec des métriques unitaires sur les liens

Une route est associée à une métrique

- le nombre de bonds (« hop count »)
- PAS de poids sur les liens par défaut
- On représente l'infini par « 16 »

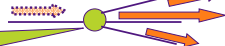


Les routeurs échangent leurs routes avec leurs voisins

Si l'information qui a provoqué la mise à jour provient d'un routeur adjacent, alors le routeur propage à tous ses voisins la mise à jour, sauf à :

Mise à jour d'une info

Message d'annonce (« update »)

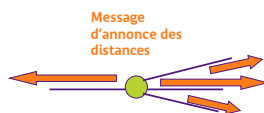


DE

3.1 protocole VD

Les routeurs annoncent leurs routes

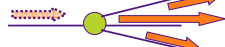
- périodiquement
- Lorsqu'une information a été mise à jour



Message d'annonce des distances

Mise à jour d'une info

Message d'annonce (« update »)

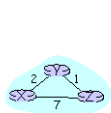


Algorithme de routage
Isma 2007-2008

DE

3.1 La comparaison de routes dans les protocoles VD: optimalité de Bellman

▶ Les chemins les plus courts sont choisis



		cost via	
d	e	X	Z
		2	∞
s	t	∞	7
		∞	∞

		cost via	
d	e	X	Z
		2	∞
s	t	∞	1
		∞	∞

$$D^X(Y,Z) = c(X,Z) + \min_w D^Z(Y,w)$$

$$= 7 + 1 = 8$$

$$D^X(Z,Y) = c(X,Y) + \min_w D^Y(Z,w)$$

$$= 2 + 1 = 3$$

3.1 les entrées de la table pour RIP

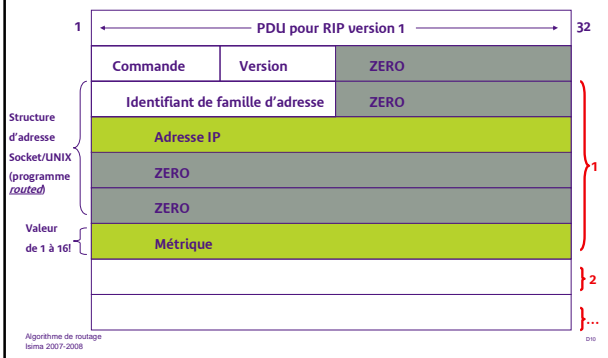
▶ Chaque entrée contient au moins

- L'adresse IPv4 de la destination
- Une métrique, qui représente le coût total de transport d'un datagramme de l'hôte à cette destination. Cette métrique est la somme des coûts associés aux réseaux qui seraient traversés pour arriver à la destination.
- L'adresse IPv4 du prochain routeur le long du chemin vers la destination. Si la destination est située sur l'un des réseaux directement connectés, cet élément n'est pas nécessaire.
- Un drapeau pour indiquer que l'information sur la route a changé récemment. Il sera référencé sous le nom de « drapeau de changement de route ».
- Différents temporisateurs associés à la route.

3. Le protocole RIP

- 3.1 présentation: RIP est un protocole VD
- 3.2 les échanges de messages
- 3.3 l'effet rebond des protocoles à vecteur de distance
- 3.4 quelques remarques sur RIP

3.2 Les messages du protocole RIP v1



3.2 Les échanges de messages du protocole RIP

- ▶ Les messages sont envoyés au dessus de UDP/IP port 520
 - ▶ Le port est < 1024
- ▶ Itératif
 - ▶ Continue la propagation de messages après une mise à jour
- ▶ Auto terminant
 - ▶ Pas de signal pour stopper
- ▶ Les messages
 - ▶ Request = demande de routes
 - Envoyées par diffusion
 - ▶ Response = annonces de routes
 - « Advertising »
 - ▶ Maximum 25 entrées par message

3.2 Les messages du protocole RIP

- ▶ Request
 - ▶ Requête de rien
 - Message vide
 - ▶ Requête de toute la table de routage d'un nœud
 - Message avec une seule entrée avec identifiant nul et une métrique de 16
 - Lorsque le routeur démarre sur le réseau
 - L'horizon partagée est inclus
 - ▶ Requête d'entrées spécifiques
 - Pour chaque destination dans le message, le nœud qui reçoit le message remplit ce dernier avec les distances qu'il utilise (éventuellement 16)
 - Est normalement utilisé pour l'administration (pas d'horizon partagée)
- ▶ Response
 - ▶ Réponse à une question
 - ▶ Mise à jour régulière (réponse non sollicitée)
 - ▶ Mise à jour déclenchée par un changement de route

3.2 Temporisations de RIP

► Chaque routeur reçoit des informations des routeurs qui lui sont adjacents

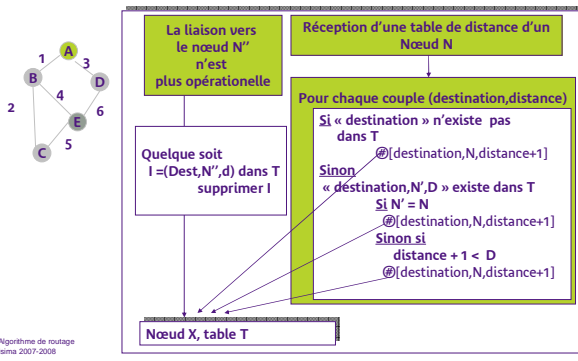
- Chaque routeur maintient l'état des routes transmises par ses voisins
- Des messages sont échangés périodiquement : « advertising »
 - Les messages sont appelés **update** (même si rien n'est mis à jour!)
 - Par défaut toutes les 30 secondes pour Cisco IOS
- Des routes qui ne sont pas maintenues actives par les voisins sont invalidées (inutilisables) puis supprimées
 - Un routeur qui ne transmet plus ses informations périodiquement est considéré comme mort et inutilisable dans le réseau
- Pour Cisco IOS
 - Après 180 secondes sans update du routeur R:
 - toutes les routes transmises par R sont marquées non utilisable
 - Après 240 secondes sans update du routeur R:
 - toutes les routes transmises par R sont supprimées de la table

3.2 Mises à jour de la table

► Mise à jour d'une information

- Perte de porteur dans les équipements de transmission sur le lien avec le nœud K
 - La distance vers toutes les destinations atteintes par K sont mises à l'infini
- Réception d'une route vers une destination inconnue
 - La destination est installée dans la table des distances
- Réception d'une mise à jour d'une route déjà connue
 - La distance est mise à jour
- Réception d'une route strictement meilleure qu'une route déjà connue vers la même destination
 - La distance et le next hop associés à la destination sont mis à jour

3.2 Mises à jour des distances

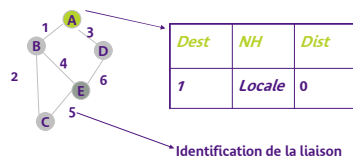


3.2 RIP: démarrage à froid

➤ Initialisation d'un nœud

- Chaque nœud a une connaissance minimale de la configuration
 - Son adresse, ses liaisons
- La table de distance est initialement vide

➤ Dans l'exemple, les liaisons ont un coût de 1



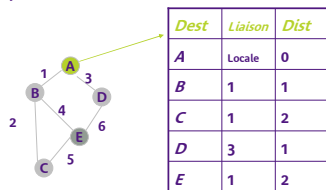
Algorithme de routage
Isma 2007-2008

D16

3.2 RIP: démarrage à froid (2)

➤ Diffusion du vecteur de distance sur toutes les liaisons locales à chaque nœud

➤ Après propagation des messages de voisins en voisins, on obtient pour A:



Algorithme de routage
Isma 2007-2008

D17

3. Le protocole RIP

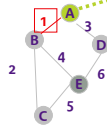
- 3.1 présentation: RIP est un protocole VD
- 3.2 les échanges de messages
- 3.3 l'effet rebond des protocoles à vecteur de distance**
- 3.4 quelques remarques sur RIP

Algorithme de routage
Isma 2007-2008

D18

3.3 VD: rupture de lien

Rupture du lien 1



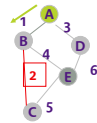
Dest	Liaison	Dist
A	Locale	0
B	3	3
C	3	3
D	3	1
E	3	2

Dest	Liaison	Dist
A	Locale	0
B	1	1->16
C	1	2->16
D	3	1
E	1	2->16

3.3 VD: rupture de lien et « effet rebond »

On considère désormais que la métrique de la liaison 5 est **10**

A			B		
Dest	LL	Dist	Dest	LL	Dist
C	1	2	C	2	1
C	1	2	C	2	16
C	1	2	C	1	2+1

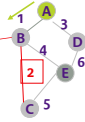


Rupture du lien 2

- B mets une métrique infinie pour la route vers C
- Avant que B envoie ses informations mises à jour à E et A, A envoie sa table de distance
 - B mets à jour son entrée concernant C

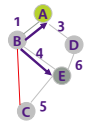
3.3 VD: rupture de lien et « effet rebond »

A			B			C			D			E		
Dest	LL	Dist	Dest	LL	Dist	Dest	LL	Dist	Dest	LL	Dist	Dest	LL	Dist
C	1	2	C	2	16	C	Loc	0	C	6	3	C	4	2
C	1	2	C	1	2+1	C	Lo	0	C	6	3	C	4	2
C	1	4	C	1	3	C	Lo	0	C			C	4	4

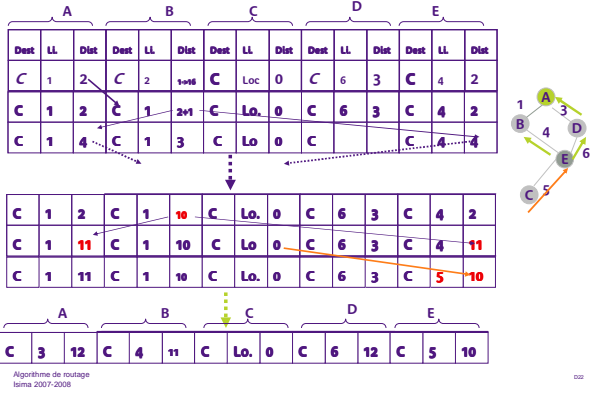


Rupture du lien 2

- B va envoyer les « fausses informations » à ses voisins A et E qui mettront à jour la route
 - A et E croient à une mise à jour de la route vers C (même NH)



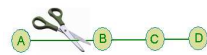
3.3 VD: convergence lente



3.3 « effet rebond » et convergence lente

- ▶ Lorsque la boucle apparaît entre A et B
 - Les paquets font du ping pong entre A et B jusqu'à l'expiration du TTL (IP)
 - La congestion peut entraîner des pertes de paquets
 - Si les paquets de routage se perdent, la convergence est encore plus lente!

- ▶ Effet « comptage jusqu'à l'infini »



Les routes vers A: (dist,NH)

B	C	D
16,A	2,B	3,C
3,C	2,B	3,C
3,C	4,B	3,C
5,C	4,B	5,C
5,C	6,B	5,C
...
16,C	16,B	16,C

Algorithme de routage
Ismaïl 2007-2008

3. Le protocole RIP

- 3.1 présentation: RIP est un protocole VD
- 3.2 les échanges de messages
- 3.3 l'effet rebond des protocoles à vecteur de distance
- 3.4 quelques remarques sur RIP

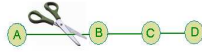
Algorithme de routage
Ismaïl 2007-2008

024

3.4 RIP tente d'être un bon protocole VD

▶ « Horizons partagés »

- un routeur C ne devrait pas envoyer des routes à un voisin B si cette route a été reçue de B



▶ « Mises à jour déclenchées »

- Les informations mises à jour sont immédiatement envoyées au lieu d'attendre l'envoi programmé

3.4 Adressage avec RIP

▶ On peut adresser

- des machines
- des réseaux
- une adresse spéciale indiquant une adresse par défaut

▶ lors du switching de paquets

- ⊕ regarde si la destination n'est pas dans la table
- ⊕ sinon regarde si la destination ne fait pas partie d'un réseau dans la table
- ⊕ sinon utilise la route par défaut

▶ L'évaluation du réseau peut être ambiguë

- dépend si le masque est connu ou non
 - Sinon le masque est identifié d'après les bits nuls dans l'adresse
- RIP v2 envoie toujours le masque

▶ Les routes sont synthétisées

- suivant la destination des messages RIP

3.4 RIP version 2

▶ Utilisation de multicast IP pour les messages

▶ Gestion des réseaux CIDR

- Sous réseaux, super-réseaux

▶ Transmission de messages

- Gestion aléatoire du déclenchement des émissions de mises à jours
 - 14 à 45 secondes.

3.4 les insuffisances de RIP

- ▶ **Métriques pas adéquates pour de « gros » réseaux**
 - ▶ 15 bonds maximum pour le plus long chemin
 - Si l'administrateur configure des coûts >1, c'est pire
- ▶ **Lenteur de convergence**
 - ▶ Le comptage à l'infini est l'unique solution pour résoudre des situations inhabituelles
 - ▶ Les phases transitoires sont instables et consomment de la bande passante
 - Problème pour les lignes bas débit
- ▶ **Requiert des métriques fixes**
 - ▶ Pas de QoS (métriques adaptatives au délai, à la bande passante,...)
- ▶ **Pas d'information concernant plusieurs routes**

4. Le protocole OSPF

- 4.1 Présentation
- 4.2 Les messages échangés
- 4.3 fonctionnement détaillé

4.1 un peu d'histoire sur OSPF

- ▶ **« Open Shortest path first »**
 - ▶ Algorithme de type SPF
 - Produit des recherches effectuées par ARPANet et surtout IETF
 - Groupe de travail formé en 1988
 - 1989 : version 1, rfc 1131
 - 1991 : version 2, rfc 1247
 - Revisions: rfc 1583, 2178, 2328
 - ▶ Seulement la version 2 est utilisée
 - Abus de langage commun: OSPFv2 = OSPF!
- ▶ **Protocole Intra-domaine**
 - ▶ IGP qui a été créé pour remédier aux problèmes de RIP, notamment en terme de temps de convergence et taille de réseaux
 - ▶ Protocole de type « Link State » (LS)

4.1 les concepts d'OSPF

▶ Link-state Database (LSDB)

- Structure de donnée fondamentale
- Une copie est maintenue par **tous les routeurs**
- Contient une description du graphe dirigé des routeurs du réseau (du système autonome souvent)
- Chaque lien vers un réseau ou vers un routeur est représenté par une entrée dans la LSDB **avec une métrique** (un coût) associée
 - La métrique ne doit pas forcément refléter le nombre de bonds comme dans RIP

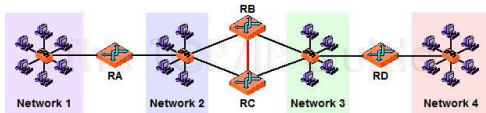
4.1 échange de la topologie du réseau

▶ Les routeurs envoient des paquets LSA ou LSP

- « Link State Advertisements/Packets »
- Chaque routeur envoie ce qu'il sait du réseau

▶ Les LSDB sont les mêmes dans chaque routeur

- Le processus d'envoi de messages LSA converge
 - à chaque mise à jour, des paquets LSA sont diffusés sur le réseau



4.1 calcul des routes

▶ Chaque routeur calcul un arbre des plus courts chemins

- Il obtient un chemin vers chaque autre routeur du réseau
- Si plusieurs chemins ont le coût minimum, le trafic est partagé sur les liens correspondants
 - **load-balancing**

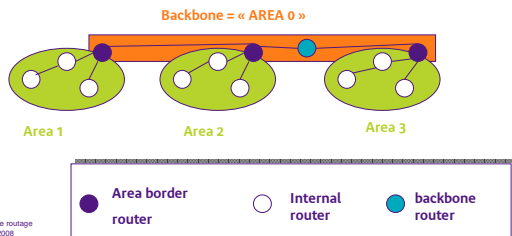
▶ L'arbre est recalculé au cours du temps

- Une mise à jour de la topologie va changer l'arbre des plus courts chemins préalablement calculé

4.1 hiérarchie

Des équipements adjacents sont regroupés en AREA

- Comme un sous-AS ou le routage OSPF est sans hiérarchie
- Hiérarchie à deux niveaux



4.1 hiérarchie

Internal router

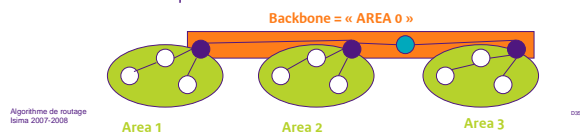
- Connecté seulement à des équipements d'une même area
- 1 LSBD

Area Border router

- Connecté à des équipements de plusieurs area
- Connecté au backbone
- n LSBD

Backbone router

- Area border router
- Et aussi uniquement connecté à des routeurs de backbone



4.1 hiérarchie

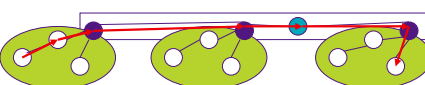
Les Internal routers n'ont besoin que des informations concernant leur area seulement

Les routeurs de backbone (au moins un routeur de bordure d'area par area) gardent des informations condensées concernant toutes les area

Communications dans la même area



Communications entre deux area

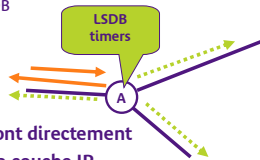


Algorithme de routage
Isma 2007-2008

4.2 Les messages OSPF

▶ Chaque routeur OSPF du réseau joue un rôle dans les échanges d'informations de routage

- Maintiens de leurs LSDB
- Gestion de timers
- envoie des messages
- Répond aux messages



▶ Les messages OSPF sont directement Construits au dessus de la couche IP

- '89' pour le champ *protocol* dans l'en-tête IP

4.2 Les messages OSPF

▶ 5 types de messages OSPF

- Hello
 - Pour la découverte des liens d'un routeur (« adjacencies »)
- Database Description
 - Description de la topologie ou d'une area d'un routeur vers un autre
 - La communication d'une grande LSDB requiert plusieurs messages envoyés en séquence d'un nœud **maître** vers un nœud **esclave**
- Link State Request
 - Requête de l'état de liens (spécifiés dans le message) d'un routeur depuis en provenance d'un autre routeur
- Link State Update
 - Envoyés en réponse à un message « Link State Request » ou broadcastés (multicast ou non) régulièrement par les nœuds
 - Le contenu sert à mettre à jour la LSDB du routeur qui reçoit le message
- Link state Acknowledgment
 - Accusé de réception d'un message « Link State Update »

4.2 format des messages OSPF

En-tête OSPF commun de 24 bits

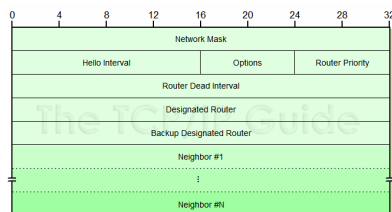
Field Name	Size (bytes)	Description
Version #	1	Version Number: 2.
Type	1	
Packet Length	2	Longueur du message OSPF avec les 24 bits de l'en-tête
Router ID	4	L'ID du routeur qui a généré ce message (généralement l'adresse de l'interface sortante)
Area ID	4	Area ID du routeur originaire du message
Checksum	2	similar to a standard IP checksum . (pas l'Authentication field).
AuthType	2	
Authentication	8	64-bit pour l'authentification (quand elle est utilisée)

4.2 format des messages Hello

Field Name	Size (bytes)	Description
Network Mask	4	Le masque du réseau envoyé par le routeur
Hello Interval	2	Hello Interval (nombre de secondes qu'un routeur attend pour envoyer un nouveau message hello)
Options	1	Indique les fonctionnalités OSPF supportées par le routeur
Rtr Pri	1	Router Priority (quand il y a élection d'un routeur de backup)
Router Dead Interval	4	Nombre de secondes de silence avant qu'un routeur soit considéré mort
Designated Router	4	Permet de désigner un routeur spécial. (vide si non utilisé)
Backup Designated Router	4	Permet de désigner un routeur de backup spécial. (vide si non utilisé)
Neighbors	Multiples of 4	L'adresse de chaque routeur dont les messages hello ont été reçus récemment

4.2 format des messages Hello

Type OSPF: 1



4.2 format des messages database description

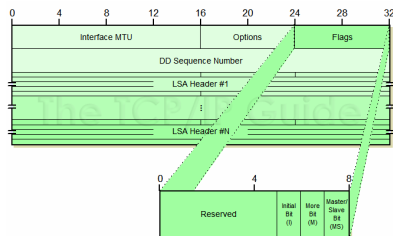
Field Name	Size (bytes)	Description
Interface MTU	2	Interface MTU : la taille du plus large message IP qui peut être envoyé sur l'interface du lien sans fragmentation
Options	1	Indique les fonctionnalités OSPF supportées par le routeur
Flags	1	
DD Sequence Number	4	Utilisé pour compter une séquence de messages Database Description (pour les remettre dans l'ordre)
LSA Headers	Variable	LSDB information

Flags: Special flags used to indicate information about the exchange of Database Description messages:

Subfield Name	Size (bytes)	Description
Reserved	5B (5 bits)	Reserved: Sent and received as zero.
I	1B (1 bit)	Init: Set to 1 to indicate that this is the first ("initial") in a sequence of Database Description messages.
M	1B (1 bit)	More: Set to 1 to indicate that "more" Database Description messages follow this one.
MS	1B (1 bit)	MS-Bit: Set to 1 if the router sending this message is the master in the communication, or 0 if it is the slave.

4.2 format des messages database description

Type OSPF: 2



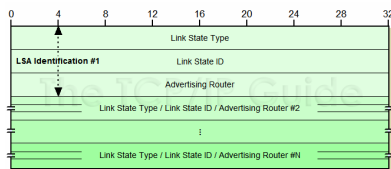
Algorithme de routage
Isma 2007-2008

D4

4.2 format des messages LSR

Type OSPF: 3

Field Name	Size (bytes)	Description
LS Type	4	Link State Type: type d'annonce LSA.
Link State ID	4	Link State ID: l'identifiant du LSA, généralement l'adresse IP du routeur ou du réseau relié
Advertising Router	4	Advertising Router: l'ID du routeur qui a créé le message LSA



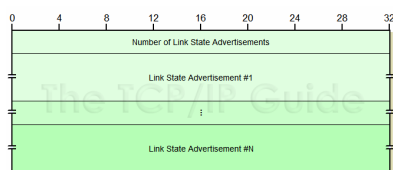
Algorithme de routage
Isma 2007-2008

D4

4.2 format des messages Link State Update

Type OSPF: 4

Field Name	Size (bytes)	Description
# LSAs	4	Number of LSAs: nombre de LSA
LSAs	Variable	LSAs: un ou plusieurs link-state advertisements .



Algorithme de routage
Isma 2007-2008

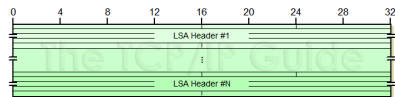
D4

4.2 format des messages Link State Acknowledgment

Type OSPF: 3

Field Name	Size (bytes)	Description
------------	--------------	-------------

LSA Headers Variable **LSA Headers:** contient l'en-tête des LSA acquies



Algorithme de routage
Isma 2007-2008

D46

4.2 format des messages Link State Advertisements

Subfield Name	Size (bytes)	Description
LS Age	2	LS Age: nombre de secondes depuis la création du message
Options	1	Indique les fonctionnalités OSPF supportées par le routeur
LS Type	1	
Link State ID	4	Link State ID: l'identifiant du LSA, généralement l'adresse IP du routeur ou du réseau relié
Advertising Router	4	l'ID du routeur qui a créé le message LSA
LS Sequence Number	4	Utilisé pour détecter les duplications
LS Checksum	2	Protection de la corruption des données
Length	2	Longueur du message (avec en-tête)

LS Type: Indicates the type of link this LSA describes:

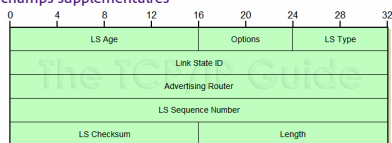
Value	Link Type	Description
1	Router-LSA	Link to a router
2	Network-LSA	Link to a network
3	Summary-LSA (IP Network)	When areas are used, summary information generated about a network
4	Summary-LSA (ASBR)	When areas are used, summary information about a link to an AS boundary router
5	AS-External-LSA	An external link outside the autonomous system

Algorithme de routage
Isma 2007-2008

D47

4.2 format des messages Link State Advertisements

- ▶ **LSA classiques**
 - ▶ Id router, métrique, border/internal
- ▶ **LSA pour réseaux**
 - ▶ Masque, infos sur machines dans le réseau
- ▶ **LSA résumé**
 - ▶ Métrique, résumé adresse, masque
- ▶ **External LSA**
 - ▶ Des champs supplémentaires



Algorithme de routage
Isma 2007-2008

D48

4.2 le flooding de paquets

➤ Dissémination fiable de l'information

⊕ **reliable flooding**

- Permet aux routeurs de mettre à jour leur LSDB

➤ Paquets LSP

- Échangés et transmis entre voisins pour simuler l'envoi N->N

– Envoi: chaque routeur envoie ses informations à tous ses voisins

– Transmission:

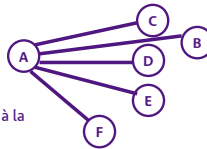
→ Si N est le plus récent numéro de séquence pour X:
Alors envoyer le message à voisins(A) \ {D}



4.2 Les LSA

➤ LSA généré par un nœud A

- ID du nœud (ici c'est l'ID de A)
- Liste des voisins de A avec le coût associé à la liaison
- Numéro de séquence (SEQNO)
- Age du paquet (« Time To Live »)



4.2 Réception des LSA

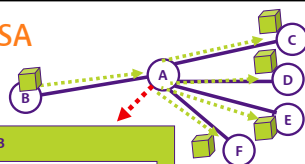
Rechercher enregistrement dans la LSDB

Si enregistrement absent

- ⊗ ajouter l'enregistrement dans la LSDB
- ⊗ diffuser le LSA aux autres voisins (**pas l'exp.**)

Sinon si enregistrement présent dans la LSDB
avec un SEQNUM inférieur à celui du paquet

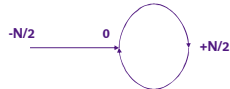
- ⊗ ajouter l'enregistrement dans la LSDB
- ⊗ diffuser le LSA aux autres voisins (**pas l'exp.**)



4.2 LSA et numéros de séquence

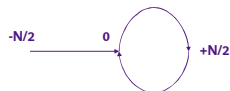
- ▶ Les informations sur la topologie sont diffusés
 - Après mise à jour d'une métrique ou de l'état d'un lien
 - périodiquement
- ▶ Les numéros de séquence sont incrémentés
 - Gestion habile des numéros
 - Lollipop Space

⊗ Si un routeur reboot, son numéro de séquence repart de $-N/2$



4.2 LSA et numéros de séquence

- ▶ $\text{NUMSEQ}(a) < \text{NUMSEQ}(b)$
 - Si $a < 0$ et $b > 0$
 - Si $a > 0$, $a < b$ et $b - a < N/4$
 - Si $a > 0$, $b > 0$, $a - b > N/4$
- ▶ Si un routeur Y reçoit un LSP de X avec un numéro de séquence plus vieux que dans la LSDB de Y
 - Y informe X de son numéro de séquence avant reboot



4.2 Messages Hello

- ▶ Boot de routeur
- ▶ Périodiquement
- ▶ Permet de découvrir les liaisons
 - Si nouvelle liaison:
 - Envoie de messages Database Description pour initialiser la LSDB des routeurs
- ▶ Désignation d'un routeur pour chaque interface
 - Suivant la plus haute priorité

4.2 pannes

▶ Messages Hello

▶ Destruction des vieux messages LSA

- ▶ l'âge d'un LSA est incrémenté après chaque bond
 - Destruction d'un LSA si son âge est supérieur à l'âge max
 - Rediffusion du LSA « âgé » pour provoquer la destruction dans les autres nœuds

4.3 OSPF est plus qu'un protocole LS

▶ Supporte l'authentification

- ▶ Meilleure sécurité

▶ Supporte les trois types d'adressage

- ▶ Classful
- ▶ subnetted classful
- ▶ classless (CIDR)

▶ Pour les grands réseaux

- ▶ On peut grouper les routeurs
- ▶ On peut organiser **hiérarchiquement** la topologie du réseau
- ▶ Permet de réduire le trafic des LSA

▶ Tout cela a un coût en **complexité**

- ▶ OSPF est plus difficile à administrer que RIP
