

Réseaux

Open Shortest Path First

1. Historique

2. Fonctionnement

3. Les aires

4. Les routeurs

5. Les messages

6. Adjacences

- OSPF a été formé au sein de l'IETF en 1987 pour remplacer RIP ;
- Il est inspiré du protocole ARPANET développé par BBN ;
- La première version est publiée en 1990 (**RFC 1131**) ;
- La deuxième version est publiée en 1991 (**RFC 1247**) ;
- l'Internet Engineering Steering Group (IESG) recommande OSPF pour Internet (**RFC 1371**).

- chaque routeur établit des relations d'adjacence grâce à des messages "Hello" ;
- chaque routeur propage la liste des réseaux auxquels il est connecté grâce à des messages "LSA" (**Link-State Advertisements**) ;
- Cette liste est propagée de proche en proche ;
- L'ensemble des "LSA" forme une base de données (*LSDB*) ;
- L'ensemble du réseau est découpé en aire ;
- Chaque routeur utilise l'algorithme de Dijkstra pour déterminer la route la plus courte (*SPF*) dans la *LSDB*;
- A chaque fois que la topologie change, les *LSA* sont propagés et les *SPF* calculés

- L'ensemble des routeurs est segmenté en groupes connexes appelés aires ;
- A la frontière de chaque aire on procède à des résumés.
- Chaque aire est distinguée par un nombre entier positif ou nul variant de 0 à 4 294 967 295. Ce nombre est parfois exprimé en notation décimale pointée, de la même manière qu'une adresse IP ;
- Chaque sous-réseau appartient à une seule aire;
- Il existe toujours une aire dorsale (backbone area), area 0 (ou 0.0.0.0) à laquelle toutes les autres aires sont connectées ;
- Les aires contenant des routeurs dialoguant avec un autre protocole de routage sont des ***Not-So-Stubby Area*** (NSSA)

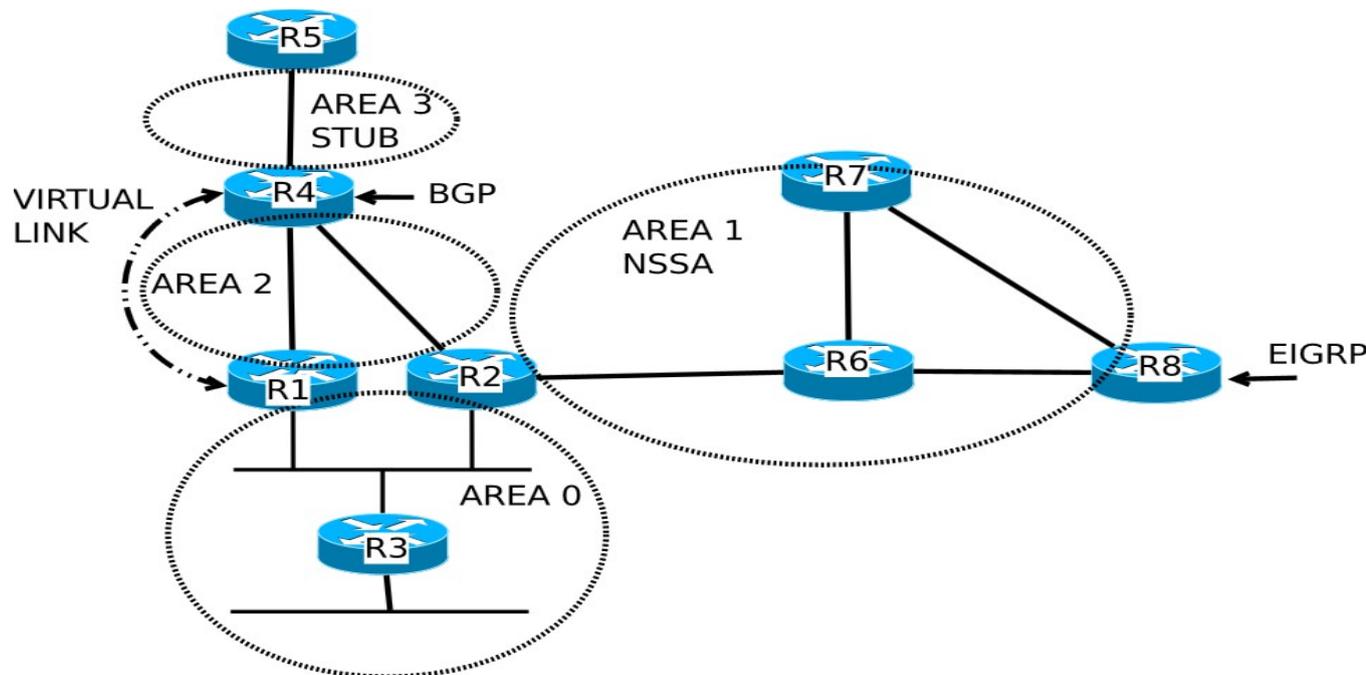
- Si des aires ne sont pas logiquement contiguës, les routeurs qui les constituent sont reliés aux routeurs du backbone entre grâce à des liens virtuels.
- Chaque routeur est identifié à l'aide un router-id unique qui est un nombre positif codé sur 32 bits et habituellement représenté sous la forme d'une adresse IP.
- La détermination du router-id a lieu uniquement à l'initialisation du processus OSPF et persiste ensuite, indépendamment de la reconfiguration ou du changement d'état des interfaces.

On distingue les types de routeurs suivants :

- Routeur interne : un routeur dont toutes les interfaces se trouvent dans la même aire ;
- Area Border Router (*ABR*) : un routeur qui dispose d'interfaces dans des aires différentes ;
- Autonomous System Border Router (*ASBR*) : un routeur qui injecte dans OSPF des routes qui proviennent d'autres protocoles de routage ou des routes statiques ;
- Routeur backbone : un routeur dont au moins une interface appartient à l'aire 0.

On distingue les types de routeurs suivants :

- Routeur interne : R5, R6;
- Area Border Router (ABR) : R1, R2, R4 ;
- Autonomous System Border Router (ASBR) : R4, R8 ;
- Routeur backbone : R1, R2, R3.

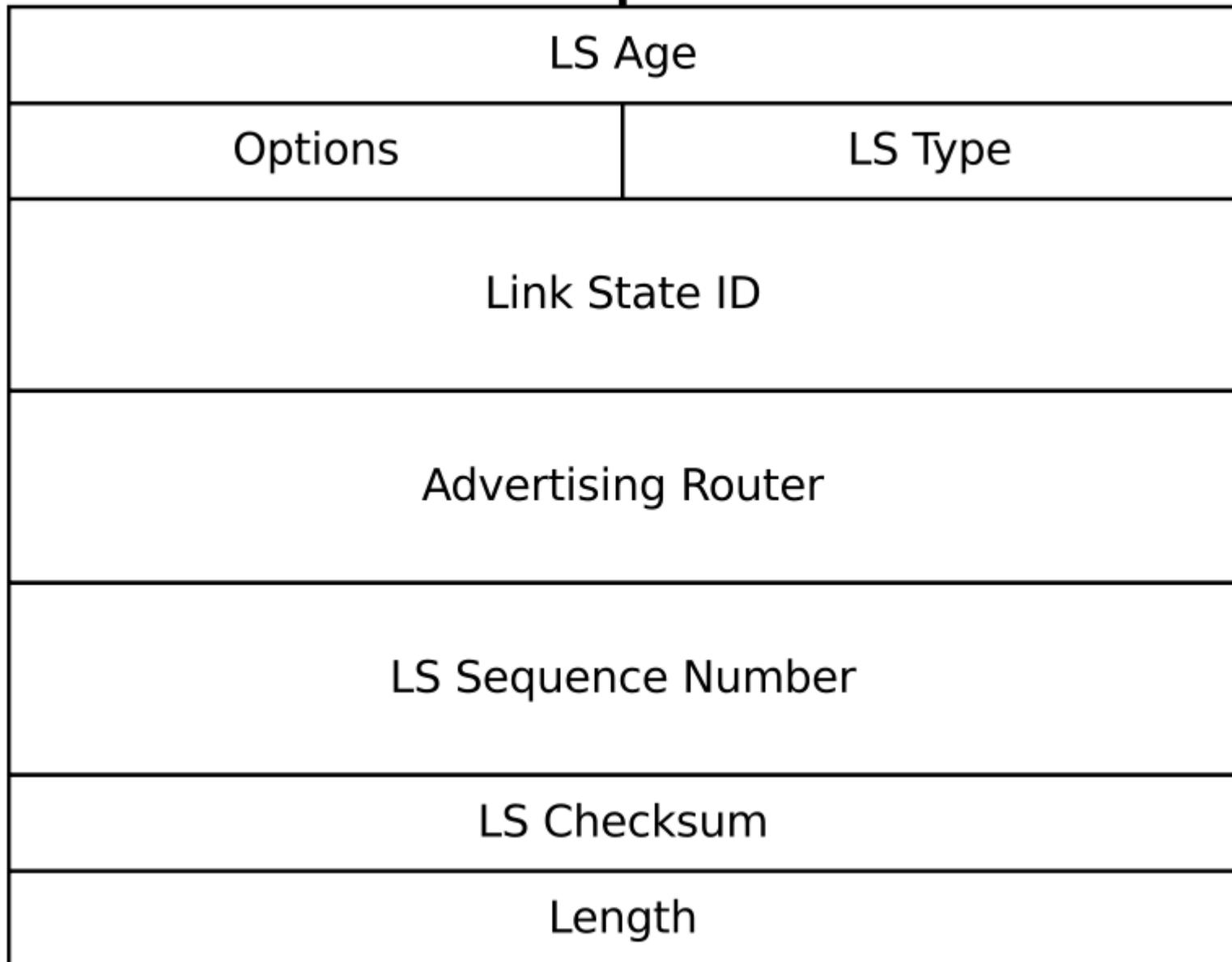


- OSPF utilise le numéro de protocole 89 d'IP ;
- Les messages ont un Time To Live (*TTL*) égale à 1 ;
- Le champ Type of Service (*ToS*) est fixé à 0 ;
- Il existe 5 types de paquets OSPF:
 - Hello (OSPF Pkt Type = 1) ;
 - Database Description (OSPF Pkt Type = 2) ;
 - Link State Request (OSPF Pkt Type = 3) ;
 - Link State Update (OSPF Pkt Type = 4) ;
 - Link State Acknowledgment (OSPF Pkt Type = 5).
- Le type 1 est utilisé pour l'établissement et le maintien des adjacences
- Les autres types sont utilisés pour la synchronisation de la LSDB

OSPF version	OSPF pkt type
OSPF length	
Source OSPF Router ID	
OSPF Area ID	
Packet Checksum	
Authentication type	
Authentication data	

Il existe 7 types de paquets LSA:

- Router (LSA Pkt Type = 1) ;
 - Network (LSA Pkt Type = 2) ;
 - Summary (LSA Pkt Type = 3) ;
 - Interarea summary (LSA Pkt Type = 4) ;
 - External (LSA Pkt Type = 5) ;
 - Multicast group membership (LSA Pkt Type = 6) ;
 - External NSSA (LSA Pkt Type = 7).
-
- Les types **1 et 2** sont utilisés par les **DR** pour les routes internes (intra-area) : type 1 = état des interfaces / type 2 = routeurs attachés
 - Les types **3 et 4** sont utilisés par les **ABR** pour les routes intra-area (IA) : type 3 = résumé de route / type 4 = route vers l'ASBR
 - Le type **5** est utilisé par les **ASBR** pour décrire des routes externes
 - Le type **7** est utilisé par l'**ASBR** d'un **NSSA**



Les routes se divisent en deux type:

- E1 → le coût d'accès à l'ASBR est ajouté à la métrique
- E2 → le coût de la métrique est fixe

Les LSA sont identifiés par :

- le routeur d'origine ;
- le type de LSA ;
- le LSA ID.

Le LSA ID varie en fonction du type de LSA :

LSA Type	LSA ID
1	Le router-id de l'émetteur
2	L'adresse IP de l'interface du DR
3	L'adresse du réseau de destination
4	Le router-id de l' ASBR
5	L'adresse du réseau externe

Le LSA de type 1 se subdivise en quatre sous-type dépendant du type de réseau décrit :

Sous-type	Description	Link ID	Link data
1	point-to-point	Router-id du voisin	Adresse IP de l'interface ou MIB-II ifindex si l'interface n'est pas numérotée
2	Lien vers un réseau de transit	Adresse de l'interface du DR	Adresse IP de l'interface
3	Lien vers un réseau de stub	Masque réseau	Masque réseau
4	Lien virtuel	Adresse IP de l'interface	Adresse IP de l'interface

- Tous les LSA sont accompagnés d'une somme de contrôle ;
- Elle est vérifiée, enregistrée dans la LSDB et transmise aux voisins sans modification ;
- Elle est revérifiée à intervalles réguliers pour s'assurer que la LSDB n'est pas corrompue.

- Le champ **LS Sequence number** s'incrémente de 1 à chaque changement ;
- le premier numéro de séquence est 0x80000001 (InitialSequenceNumber, $-2^{31}+1$) jusqu'à 0x7fffffff (2^{31}) ;
- Chaque version supérieur du LSA remplace les versions précédentes ;
- En l'absence de changement les LSA sont rafraîchis au bout du **LSRefresh Time** (environ 30 minutes) ;
- Un LSA qui atteint **MaxAge** (environ 1h) est supprimé ;
- Le routeur d'origine du LSA peut forcé le non rafraîchissement grâce au bit **DoNotAge** (placé à 1 dans les options).
- Le routeur d'origine du LSA peut le faire disparaître de la LSDB en diffusant un **LS Age = MaxAge**.

- Aire 0 : aire backbone ;
- stub area : aire vers laquelle ne sont pas propagées les LSA de type 5 (routes externes) ;
- totally stub area : aucun LSA de type 3, 4, 5 ou 7 n'y est propagée, à l'exception d'une route par défaut ;
- not-so-stubby area (NSSA) : type de stub area qui permet l'injection de routes externes via un LSA de type 7. Le type 7 sera converti en type 5 quand il sera transmis en dehors de l'aire ;
- totally NSSA : NSSA sans LSA 3 et 4 à l'exception d'une route par défaut.
- aire normale : toute les autres aires.
- Il n'est pas possible de créer des liens virtuels à travers une stub area, ni des ASBR internes à une stub area

- Point-to-point : le sous réseau correspond à un lien point à point ;
- Broadcast multiaccess : le sous-réseau peut comporter plus de deux routeurs qui peuvent tous communiquer entre eux grâce à une adresse broadcast ;
- Point-to-multipoint : le sous-réseau est constitué d'un routeur central et d'autres routeurs qui ne communiquent pas entre eux. OSPF voit le réseau comme une collection de liens point-à-point ;
- NBMA (Non-broadcast multiaccess) : le sous-réseau est constitué de routeurs qui peuvent communiquer entre eux mais sans adresse broadcast. Dans ce type de réseau, OSPF émule le type broadcast en répliquant les LSA à tous les voisins adjacents ;

	Point-to-point	Broadcast	Point-to-multipoint (non-broadcast)	Point-to-multipoint (broadcast)	NBMA
DR / BDR	non	oui	non	non	oui
Hello / dead	10s / 40s	10s / 40s	30s / 120s	30s / 120s	30s / 120s
Découverte des voisins	oui	oui	non	oui	non
Suporté par OSPF v.2 (RFC 2328)	oui	oui	oui	Cisco uniquement	oui

- Dans un sous-réseau, les routeurs qui se découvrent grâce au protocole hello sont appelés voisins ;
- Dans les réseaux de type point-à-multipoint et NBMA, les voisins sont configurés explicitement. Pour autant que les paramètres des voisins soient compatibles, ils tentent alors de former une relation d'adjacence ;
- L'établissement de l'adjacence est requise avant l'échange d'information de routage ;

Etat	Description
Down	aucune information n'a été reçue sur ce segment
Attempt	sur les réseaux NBMA, indique qu'aucune information récente n'a été reçue du voisin configuré
Init	un paquet hello a été reçu
2-Way	<ul style="list-style-type: none">• un paquet hello a été reçu et celui-ci contient son propre router-id, ce qui montre qu'il existe une communication bidirectionnelle ;• L'élection du DR et du BDR a lieu dans cet état ;• La décision de former une adjacence est prise au terme de cet état.
ExStart	<ul style="list-style-type: none">• les routeurs tentent d'établir les numéro de séquence initiaux qui seront utilisés dans les paquets d'échange d'information ;• Pour cet échange, un des routeurs deviendra le routeur primaire et l'autre le secondaire.
Exchange	les routeurs envoient leur LSDB par des paquets database description (DBD)
Loading	l'échange des LSDB se finalise, les routeurs réclament les LSA dont ils ont besoin
Full	la LSDB est synchronisée et l'adjacence est établie

- Une adjacence est nécessaire pour qu'il est partage des informations de routage ;
- Dans un réseau de broadcast la charge serait excessive $\sum_{k=1}^n k = \frac{n(n+1)}{2}$
- On choisit un DR qui réceptionne toutes les informations de route et les retransmet ;
- OSPF utilise du multicast pour communiquer :
 - 224.0.0.5 (AllSPFRouters) utilisé par le DR dialoguer avec les autres routeurs du segment ;
 - 224.0.0.6 (AllDRouters) utilisé par tous les routeurs pour envoyer les LSA vers le DR et le BDR.
- Seul le DR génère le network LSA (type 2) correspondant au sous-réseau :
 - Le LSID = l'adresse IP du DR dans le sous-réseau.
 - Ce LSA liste les routeurs attachés au sous-réseau.

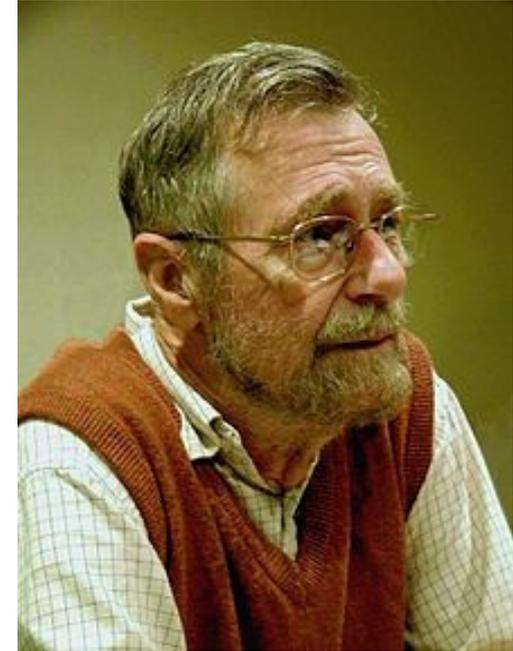
- Les messages hello sont envoyés à intervalle régulier sur les interfaces où OSPF est actif.
- Sur les liens :
 - point-à-point et les liens broadcast, ils sont diffusés sur l'adresse multicast 224.0.0.5 (AllSPFRouters) ;
 - sans broadcast, ils sont envoyés vers l'adresse IP unicast du voisin.
- Les informations suivantes sont contenues dans le paquet hello :
 - L'adresse du DR et du BDR ;
 - la priorité ;
 - le type de réseau ;
 - la liste des routeurs OSPF connus du sous-réseau ;
 - le masque de sous-réseau ;
 - les périodes hello et dead ;
 - les options.
- Le numéro d'aire est inclus dans l'en-tête paquet OSPF.
- Une adjacence ne se forme pas si des paramètres ne sont pas compatibles (n° d'aire, type d'aire (stub ou non), périodes hello/dead, authentification).

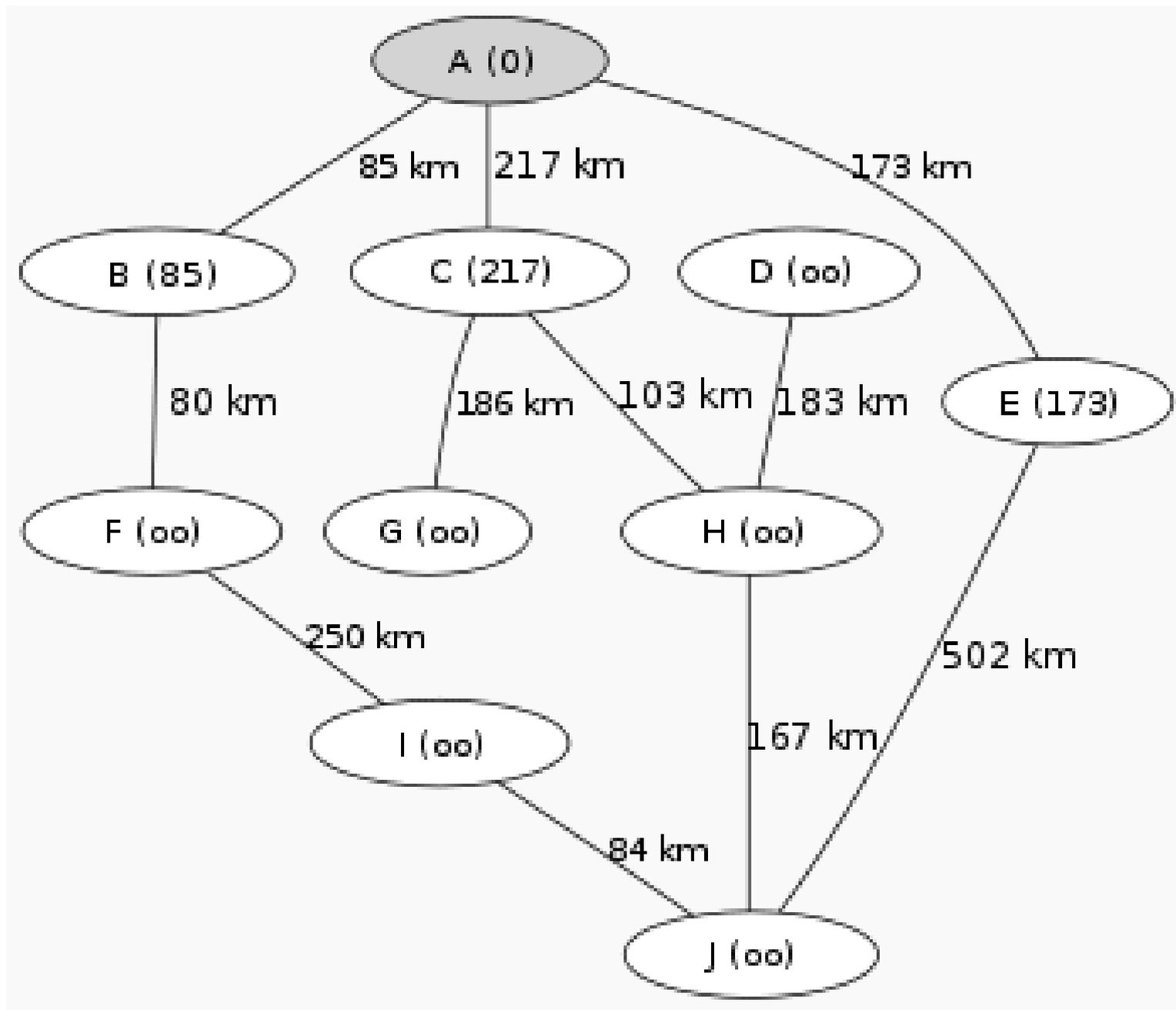
- Chaque routeur OSPF possède une **priority** comprise entre 0 et 255. Si la priorité est configurée à 0, le routeur ne participe pas à l'élection.
- Les autres routeurs qui sont dans l'état 2-Way sont éligibles.
- L'élection commence par le BDR :
 - s'il existe plusieurs routeurs qui indiquent être le BDR dans leurs paquets hello, celui qui a la priorité la plus haute sera retenu comme BDR ;
 - s'il existe plusieurs routeurs avec la priorité la plus haute, alors celui dont le router-ID est le plus élevé est retenu comme BDR ;
 - s'il n'existe qu'un routeur qui indique être le BDR, ce choix persiste ;
 - si la liste des candidats est vide, alors le BDR est le routeur éligible non-DR dont la priorité est la plus élevée ;
 - s'il en existe plusieurs, celui dont le router-id le plus élevé est retenu.
- On procède de même avec le DR, s'il n'y a pas de candidat DR, alors le BDR est promu en DR et on recommence l'élection du BDR comme au paragraphe précédent.
- Les autres routeurs sont adjacents au DR et au BDR. En cas de défaillance du DR, le BDR devient DR et on procède à une nouvelle élection du BDR.

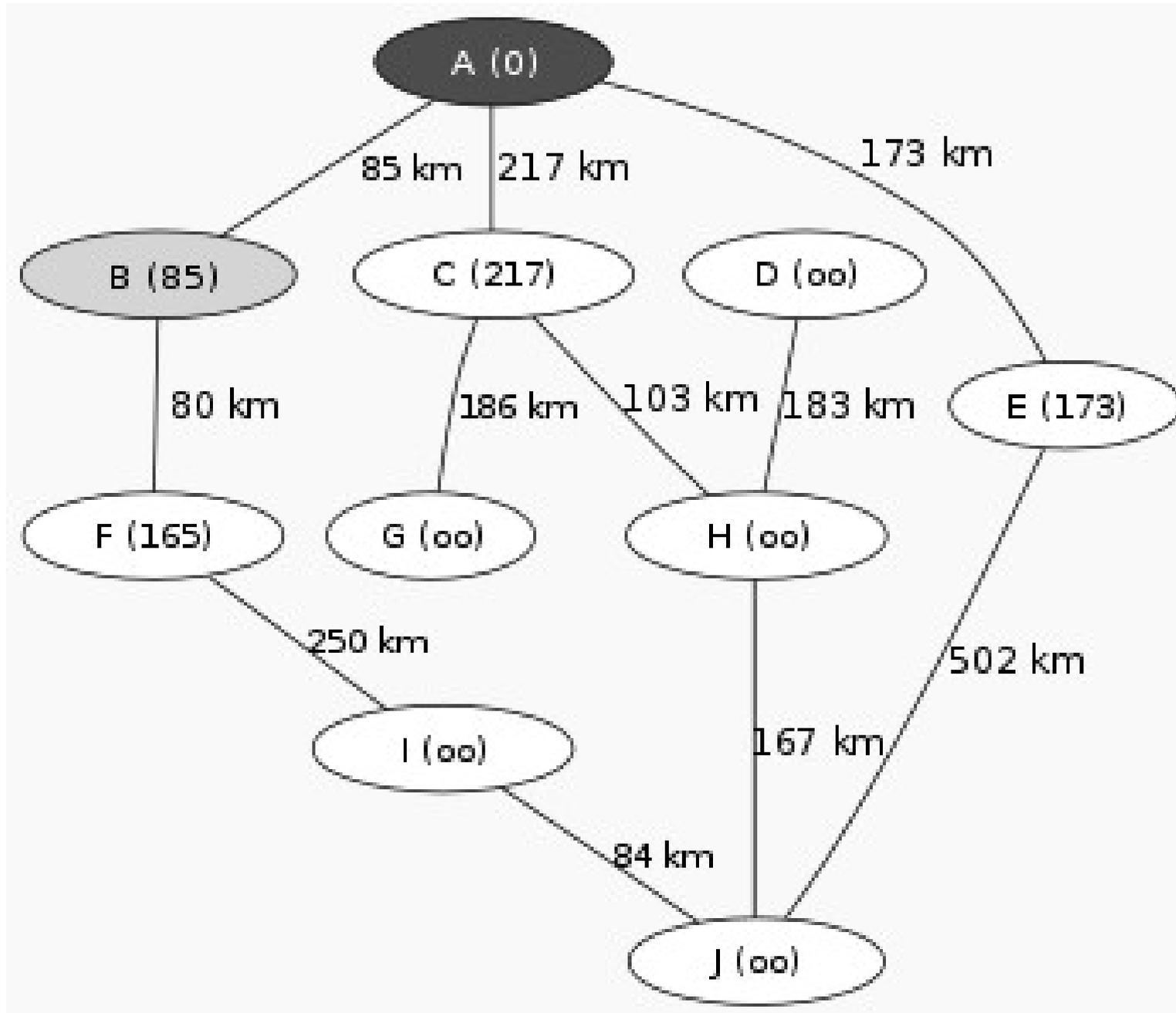
- Une fois l'adjacence établie, les routeurs vont déterminer un routeur primaire et un secondaire ainsi qu'un numéro de séquence initial ;
- Le routeur primaire va envoyer des paquets Database Description (DBD) au secondaire, ceux-ci consistent en une liste des en-têtes de LSA (mais sans les données du LSA proprement dites) ;
- le secondaire va noter les LSA qu'il ne possède pas ou dont les numéros de séquence sont plus élevés que celui dans sa LSDB et va les réclamer ensuite au primaire avec des paquets Link State Requests ;
- Le routeur primaire lui répond avec des paquets Link State Update qui contiennent les LSA demandés ;
- Le secondaire envoie ensuite les paquets Link State Update qui correspondent aux LSA dont le primaire ne dispose pas ou qui sont plus à jour chez le secondaire ;
- Une fois la synchronisation terminée, l'adjacence bascule dans l'état Full et les LSA sont diffusés normalement.

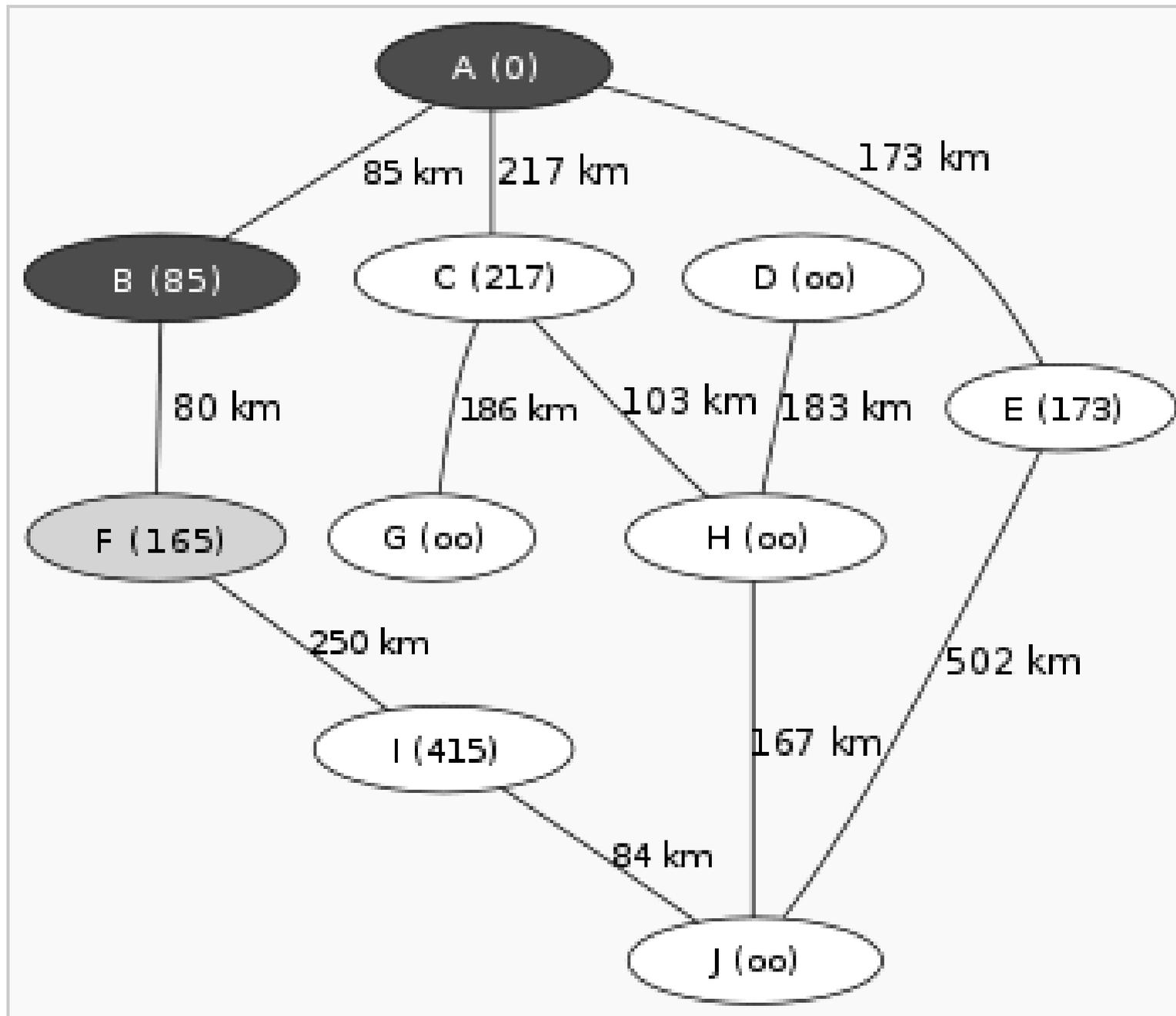
- OSPF utilise une métrique numérique, basée sur un coût additif qui peut varier de 1 à 65535 ;
- La spécification ne donne pas de signification particulière à cette métrique, seulement additionner les coûts de liens successifs doit déterminer le coût ;
- Cisco utilise une valeur par défaut du coût d'un lien qui vaut 10^8 /bande passante du lien en bit/s. Par exemple, un lien de 10 Mbit/s aura par exemple un coût de 10.
- On peut fixer arbitrairement le coût de chaque lien, ou bien fixer une bande passante de référence supérieure à celle par défaut.
- Indépendamment de la métrique, les types de routes suivants sont préférés dans cet ordre :
 - routes intra-area ;
 - routes inter-area ;
 - routes externes E1;
 - routes externes E2.
- OSPF est capable de répartir la charge sur plusieurs liens, pour autant que la métrique soit exactement identique pour chaque destination.

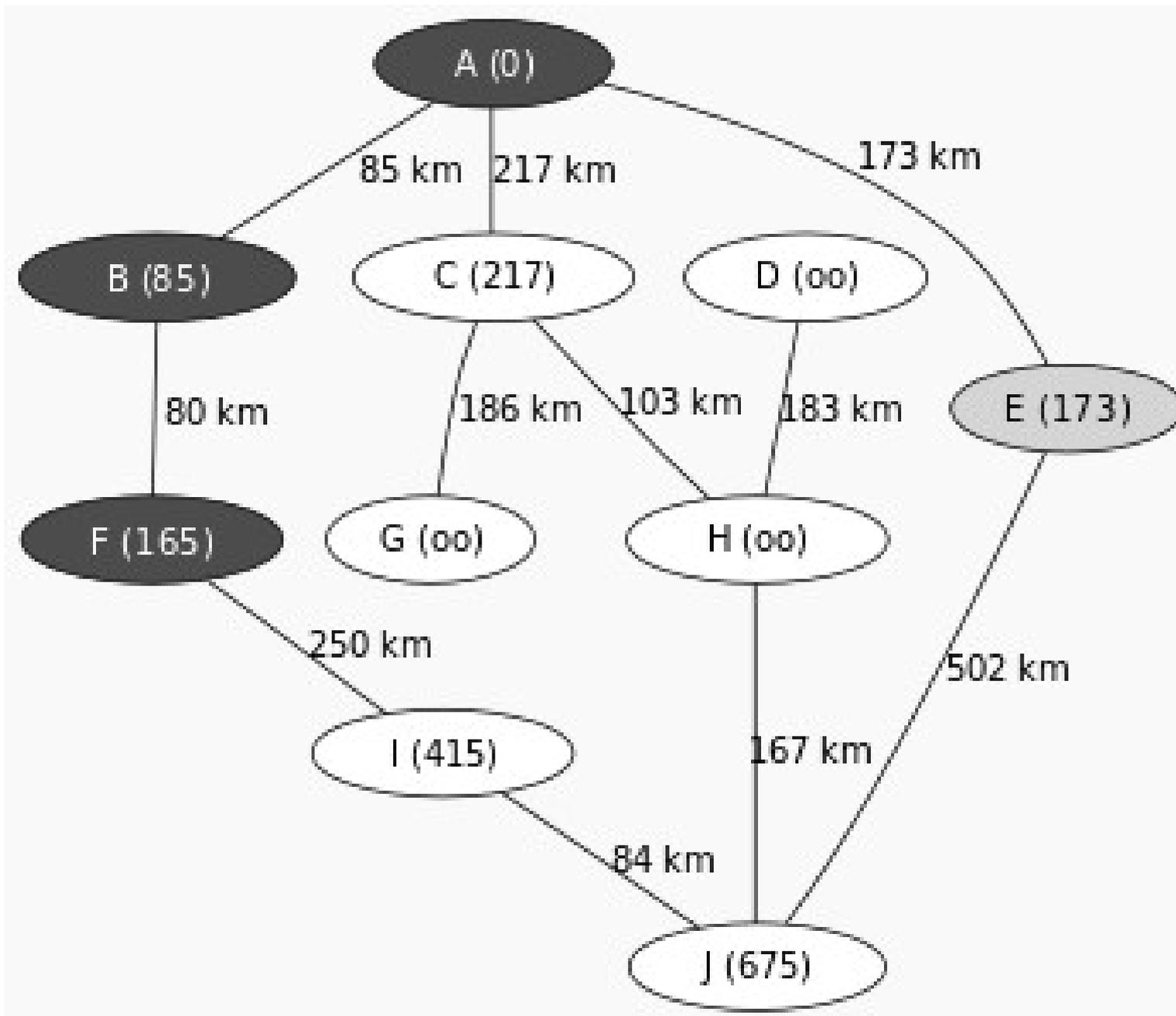
- OSPF utilise l'algorithme de Dijkstra pour déterminer le meilleur chemin à prendre. On le nomme aussi algorithme SPF (Shortest Path First) ou algorithme du plus court chemin d'abord.
- Il a été formulé par **Edsger Dijkstra**.
- OSPF déclenche ses mises à jour à chaque changement dans la topologie du réseau, ce qui permet de réduire le temps de convergence.
- À partir d'une mise à jour, un routeur met en place une base de données topologique permettant le calcul de l'accessibilité aux réseaux grâce au calcul d'un arbre de la topologie dont le routeur est la racine.

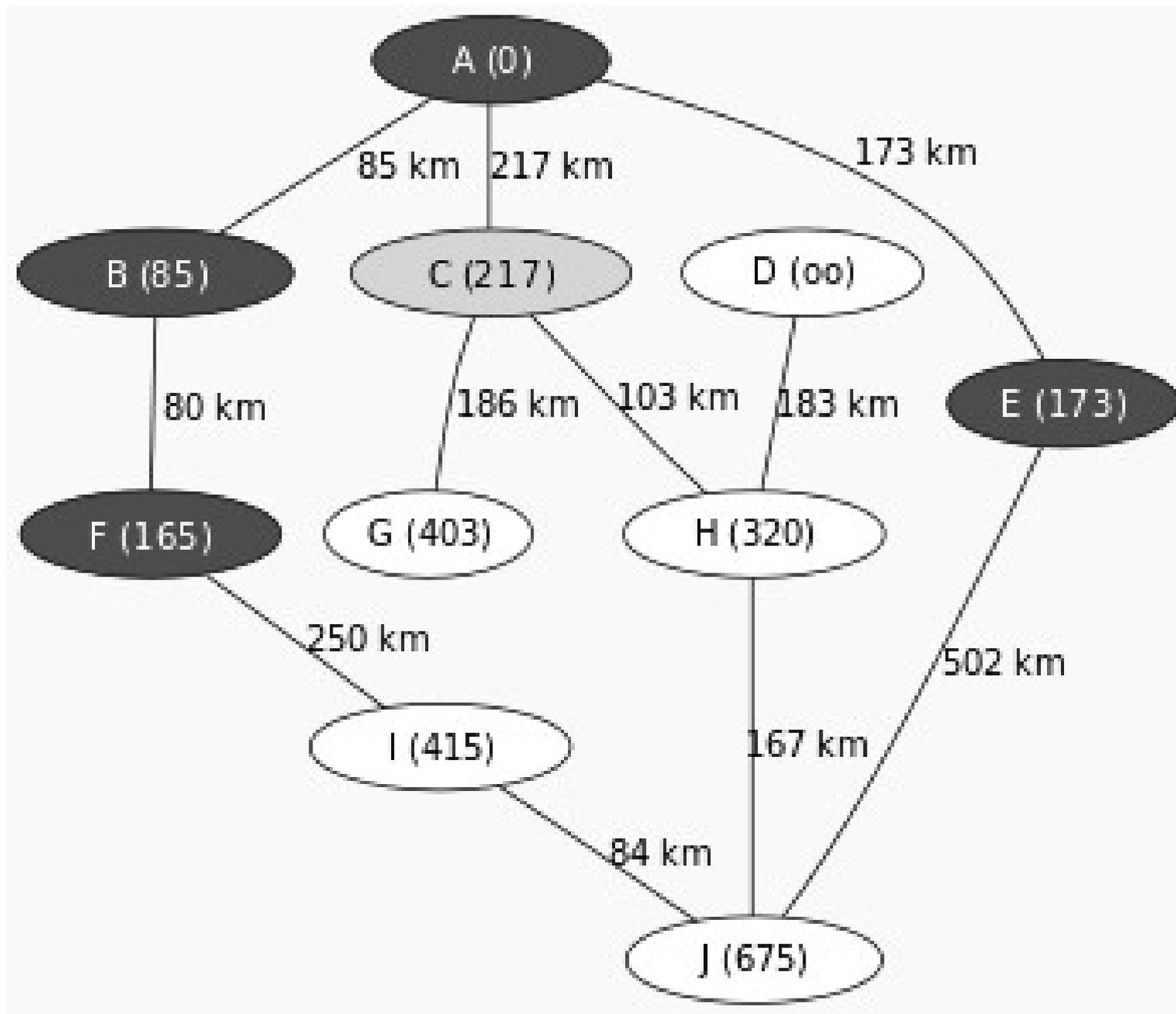


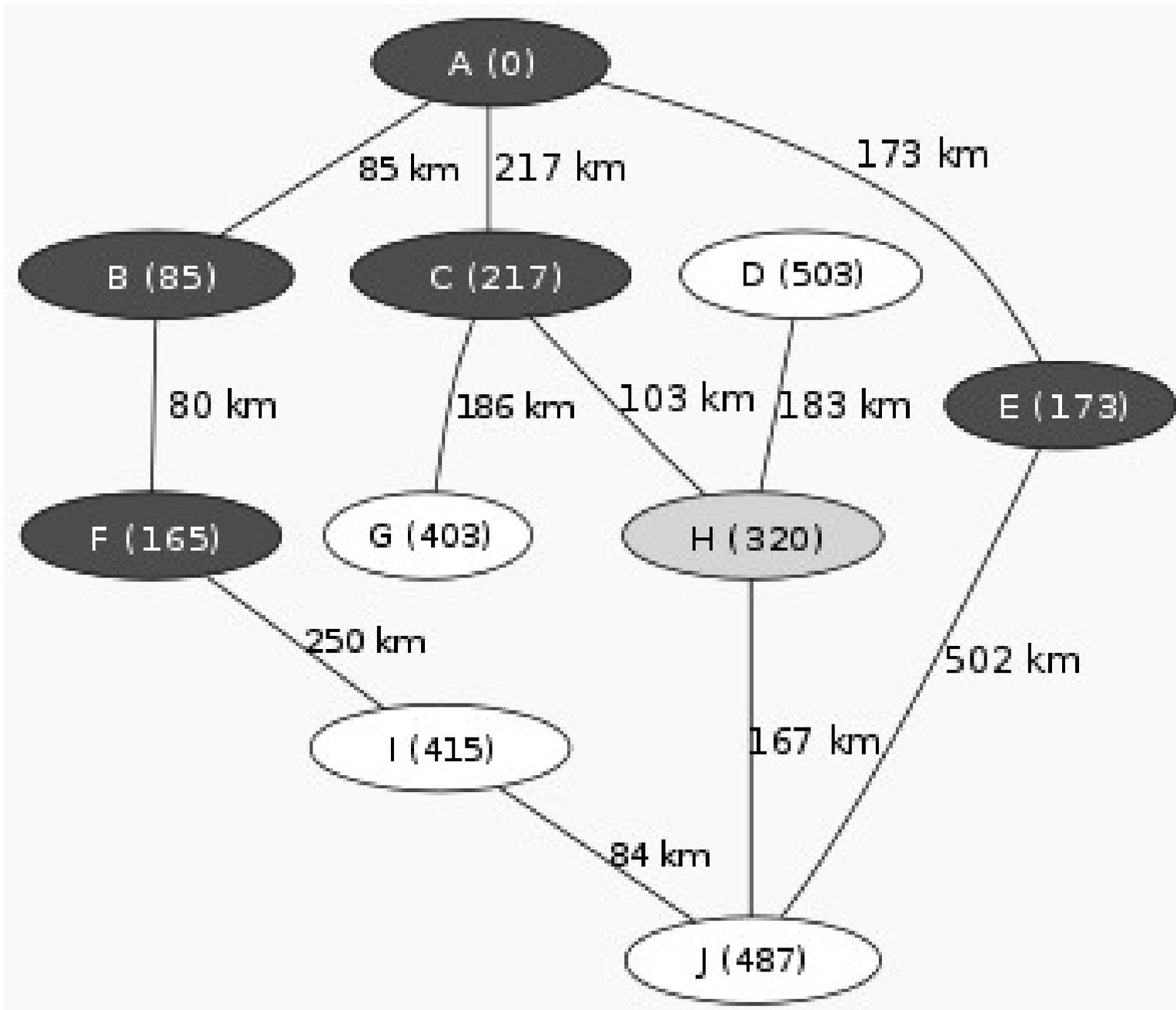












- C'est un standard IETF qui fait l'objet du RFC 2328, il fait donc l'objet d'implémentations par de nombreux vendeurs et ne pose pas de problème d'interopérabilité ;
- Il est de ce fait particulièrement populaire, son temps de convergence est particulièrement rapide, il intègre la notion de taille de masque variable (VLSM), indispensable à la gestion des réseaux sans classe actuels ;
- Il est économe en bande passante : en régime, seuls de courts messages hello sont envoyés, et en cas de changement de topologie, seuls les LSA modifiés sont envoyés aux voisins. Chaque routeur retransmet cependant l'ensemble de ses LSA à ses voisins toutes les trente minutes.

- Comme chaque routeur dispose de la totalité de la base de données de liens, tous doivent disposer de la capacité mémoire suffisante pour la stocker ;
- OSPF est sensible au phénomène de bagottement ou flapping, la capacité CPU joue un rôle dans le calcul SPF et donc la vitesse de convergence,
- On estime à 1000 le nombre de routeurs max sur un réseau OSPF ;
- La configuration d'OSPF est plus complexe, principalement si le réseau est segmenté en aires ;
- Le concept de backbone area peut limiter les topologies possibles et ne permet pas la répartition de la charge sur plusieurs liens de métrique différente, comme EIGRP peut le faire ;
- OSPFv2 est spécifique à IP. Pour d'autres protocoles il faut OSPFv3.