

# RFC 2328 : OSPF Version 2

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 11 Mai 2008

Date de publication du RFC : Avril 1998

<http://www.bortzmeyer.org/2328.html>

---

OSPF est le protocole de routage interne de l'IETF. Si RIP est encore utilisé à certains endroits, si des protocoles non-standard, spécifiques à un fabricant de routeurs sont parfois utilisés sur des sites où l'intérêt des normes ouvertes n'a pas été compris, si son concurrent ISO 10589, IS-IS a connu quelques déploiements, OSPF est de loin le plus répandu. Ce RFC est la spécification normative d'OSPF version 2, la version pour IPv4.

Par définition, les IGP, comme OSPF, sont utilisés uniquement à l'intérieur d'un système autonome. Il est donc toujours difficile de produire des statistiques fiables directes, rien ne permettant de dire, vu de l'extérieur, quel IGP utilise une organisation. Mais il n'y a pas de doute qu'OSPF est de loin le plus répandu. Son histoire a parfois été cahotique, et vient d'une grande discussion à l'IETF pour créer un IGP moderne, discussion très chaude qui avait opposé les partisans des protocoles à vecteur de distance, comme RIP aux protocoles à état des liens, comme OSPF. (Christian Huitema, dans son excellent livre « Et Dieu créa l'Internet » décrit ce débat. Le premier article sur un protocole à états des liaisons dans l'ancien Arpanet est l'excellent "*An Overview of the New Routing Algorithm for the ARPANET*" <[http://zoo.cs.yale.edu/classes/cs433/readings/arpanet\\_routing.pdf](http://zoo.cs.yale.edu/classes/cs433/readings/arpanet_routing.pdf)> en 1979.)

OSPF a finalement été normalisé dans le RFC 1131<sup>1</sup> en octobre 1989, qui connaîtra plusieurs successeurs, les plus récents étant le RFC 2178, et son successeur, notre RFC 2328. L'annexe G décrit les différences avec le RFC 2178, aucune n'étant vitale (il s'agit bien du même protocole).

Notre RFC 2328 est un des plus longs jamais publiés par l'IETF, avec presque deux cent pages. Il n'est pas spécialement pédagogique, son but étant de permettre le développement d'implémentations compatibles du protocole. Il ne cherche pas à aider l'administrateur réseau, ni à informer le curieux de passage. La table des matières est déroutante, conçue selon une logique qui a un sens, mais pas

---

1. Pour voir le RFC de numéro NNN, <http://www.ietf.org/rfc/rfcNNN.txt>, par exemple <http://www.ietf.org/rfc/rfc1131.txt>

---

celui du nouveau venu. Pour une explication de plus haut niveau (mais pas forcément facile) d'OSPF, on peut consulter le livre écrit par l'auteur du RFC <<http://www.bortzmeyer.org/ospf-moy.html>>. Pour une présentation plus orientée vers l'ingénieur en charge du réseau, on peut citer mon cours OSPF <<http://www.bortzmeyer.org/deux-cours-routage.html>>. Mais, de toute façon, OSPF est compliqué, il ne faut pas se faire d'illusions.

Quels sont les principes de base d'OSPF ? Comme décrit dans le résumé du RFC et dans la section 1, OSPF est un protocole à **état des liens** ("*link state*") et ce terme se retrouve partout dans le RFC). Chaque routeur diffuse par inondation ("*flooding*") l'état de ses interfaces réseaux et tout routeur connaît donc l'état de tout le réseau, tous les routeurs ont la même information. À partir de celle-ci, chaque routeur fait tourner l'algorithme SPF pour trouver les routes à suivre. Le fait que tous les routeurs aient la même information facilite le calcul de routes sans boucles.

Le routage proprement dit se fait ensuite uniquement sur les en-têtes IP, sans encapsulation ultérieure.

Comme le système autonome a en général des connexions avec d'autres systèmes autonomes, certains des routeurs apprennent également des routes par d'autres protocoles comme BGP (RFC 4271) et OSPF permet d'importer ces routes dans la base de données, comme « routes externes ».

OSPF dispose d'un mécanisme hiérarchique, les **zones** ("*areas*"), qui permet de limiter la diffusion de l'information, rendant ainsi ce protocole moins sensible aux changements d'échelle. La règle « tous les routeurs ont exactement la même base de données des états des liens » doit donc être nuancée : elle n'est valable qu'à l'intérieur d'une seule zone.

Tout le vocabulaire d'OSPF est rassemblé dans la section 1.2. Comme souvent avec les documents informatiques en anglais, il n'y a pas une seule traduction en français qui fasse autorité. J'ai repris ici le vocabulaire de mon cours <<http://www.bortzmeyer.org/deux-cours-routage.html>>.

Vue la taille du protocole, il est impossible de citer toutes les sections du RFC. Seules celles qui me semblent les plus importantes sont listées ici.

La section 2 du RFC détaille la base des états de liens, le cœur d'OSPF. Sa section 2.1 précise comment sont représentés les routeurs et les réseaux. Ces derniers peuvent être point à point, à diffusion comme Ethernet, ou bien NBMA ("*Non-Broadcast Multiple Access*", qui font l'objet de la section 2.1.1 et pour lesquels les choses sont assez complexes) comme le relais de trames. La représentation du lien dépend du type de réseau auquel le routeur est connecté.

La section 2.1.2 illustre tout ceci avec un exemple de base de données pour un réseau petit mais relativement complexe. La section 2.2 montre ensuite le résultat du SPF sur ce réseau.

La 2.3 explique la modélisation des informations externes, obtenues depuis un autre protocole. Elles sont de deux types, le type 1 pour les informations venant du système autonome (et donc dont le coût peut être comparé aux coûts OSPF) et le type 2 pour les informations extérieures, non seulement à OSPF, mais également au système autonome, et pour lesquelles le coût est toujours supérieur au coût OSPF.

La section 4 décrit le fonctionnement du protocole : chaque routeur OSPF, au démarrage, rassemble les informations sur ses propres interfaces réseau, puis apprend l'existence de ses voisins, et forme une adjacence sur chaque réseau. Il échange ensuite l'information sur l'état des liens avec le routeur adjacent (rappelons que tous les voisins ne sont pas adjacents, seuls le DR - "*Designated Router*" - et le BDR - "*Backup Designated Router*" le sont).

Cette information sur les états des liens est transmise sous forme d'unités nommées LSA ("*Link State Advertisement*"), qui sont transmises par inondation (procédure décrite en détail en section 13), jusqu'à ce que tous les routeurs de la zone aient la même information.

Plus concrète, la section 4.3 normalise le format des paquets OSPF. Celui-ci fonctionne directement sur IP, avec le numéro de protocole 89. Les champs du protocole sont de taille fixe, pour faciliter la mise en œuvre. Les détails du format sont renvoyés à l'annexe A. Le traitement des paquets figure dans la section 8.

La section 5 est dédiée aux différentes structures de données que doit manipuler le routeur OSPF. Par exemple, le "*Router ID*" est un entier de 32 bits qui identifie de manière unique un routeur du système autonome. Cet entier est en général représenté avec la même syntaxe qu'une adresse IPv4 (même pour la version d'OSPF conçue pour IPv6). D'autres structures de données sont décrites dans les sections 9 (les interfaces réseaux) ou 10 (les routeurs voisins).

La section 7 explique le protocole par lequel les routeurs OSPF d'un même segment de réseau se trouvent et s'échangent des informations. Tous les routeurs du lien envoient des paquets de type `Hello` à une adresse de diffusion restreinte, `224.0.0.5` (les valeurs numériques nécessaires au protocole, comme cette adresse, sont dans l'annexe A). Ici, `tcpdump` montre ces paquets émis par le routeur de "*Router ID*" `192.134.7.252`, de la zone 0 (l'épine dorsale) :

```
22:29:04.906209 IP (tos 0xc0, ttl 1, id 22064, offset 0, flags [none], proto OSPF (89), length 64) 192.134.7.252
  Router-ID: 192.134.7.252, Backbone Area, Authentication Type: none (0)
  Options: [External]
  Hello Timer: 10s, Dead Timer 40s, Mask: 255.255.255.240, Priority: 1
  Designated Router 192.134.7.252
```

Sur un réseau à diffusion comme Ethernet, un routeur est élu routeur principal ("*Designated Router*", DR, section 7.3 mais l'algorithme d'élection figure, lui, dans la section 9.4). Les adjacences ne se forment qu'avec lui et tous les échanges OSPF passent par lui (il existe aussi un routeur de secours, le BDR, "*Backup Designated Router*", section 7.4).

La section 12 présente les paquets LSA ("*Link State Advertisement*", ou annonces de l'état des liens), par lesquels les routeurs OSPF se tiennent au courant de leur situation. La section est très longue car il y a beaucoup d'informations possibles à transmettre !

Une fois que les routeurs ont toute l'information nécessaire, ils peuvent faire tourner l'algorithme SPF de Dijkstra et calculer ainsi les routes, selon la procédure que décrit la section 16. (RIP utilisait Bellman-Ford.)

D'autres sections couvrent des parties plus « exotiques » du protocole, moins nécessaires en première approche.

La section 3 est consacrée au découpage en zones. (Notez toutefois que seuls les plus grands réseaux OSPF en ont besoin. Dans un document plus pédagogique, les zones auraient été reléguées à la fin.) Les routeurs d'une zone ne connaissent pas la topologie interne aux autres zones, seulement le moyen de joindre un routeur ABR ("*Area Border Router*", section 3.3) attaché à cette zone. Les zones OSPF sont forcément attachées à une zone principale, numérotée 0 (section 3.1), qui sert d'épine dorsale au réseau.

La section 14 explique le mécanisme d'expiration des LSA dans la base (ils ne restent en effet pas éternellement).

La section 15 spécifie les liens virtuels qui permettent de connecter des parties physiquement disjointes d'une zone.

Les annexes B et C contiennent les paramètres d'OSPF, les valeurs numériques qui gouvernent le comportement du routeur. L'annexe B décrit les constantes, dont la valeur est fixe, comme `MaxAge`, la durée de vie maximale d'un LSA (fixée à une heure). L'annexe C contient les variables, que l'on peut modifier via la configuration du routeur (attention, il est souvent nécessaire que tous les routeurs du lien, ou bien tous les routeurs de la zone, soient d'accord sur la valeur). Par exemple, avec Quagga, on peut changer `Interface output cost` (le « coût » d'une interface réseau) en mettant dans le fichier `ospfd.conf` :

```
interface xxx0
! Ligne lente, son usage est découragé
ip ospf cost 10000
```

L'annexe D est consacrée à l'authentification. OSPF trouvant ses voisins par diffusion, un routeur méchant peut assez facilement s'insérer dans un réseau et diffuser de fausses informations. Sur un réseau partagé par plusieurs organisations, par exemple un point d'échange Internet, il n'est pas rare de voir passer des paquets OSPF, émis, espérons-le, suite à une erreur de configuration et pas par malveillance. Pour éviter que les adjacences ne se forment entre deux voisins, OSPF permet d'utiliser un secret partagé que doivent connaître les routeurs.

Il existe aujourd'hui deux mises en œuvre libres d'OSPF :

- Quagga
- Xorp <<http://www.xorp.org/>>

mais on trouve évidemment ce protocole dans tous les routeurs, des Cisco aux Juniper en passant par les Vyatta.