RFC 8111 : Locator/ID Separation Protocol Delegated Database Tree (LISP-DDT)

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 27 mai 2017

Date de publication du RFC : Mai 2017

https://www.bortzmeyer.org/8111.html

Le protocole LISP (dont on ne rappelera jamais assez qu'il ne faut pas le confondre avec le langage de programmation du même nom) sépare les deux rôles de l'adresse IP, **identificateur** et **localisateur**. C'est très joli de séparer, cela permet plein de choses intéressantes, comme de lutter contre la croissance illimitée de la DFZ mais cela présente un défi : comment obtenir un localisateur quand on a un identificateur ? Dit autrement, « où est cette fichue machine que j'essaie de joindre ? » Ajouter une indirection, en informatique, oblige toujours à créer un **système de correspondance** permettant de passer par dessus le fossé qu'on vient juste de créer. LISP a plusieurs systèmes de correspondance possibles, tous expérimentaux, et ce nouveau DDT ("*Delegated Database Tree*") vient les rejoindre. C'est le système qui est le plus proche du DNS dans ses concepts. Comme je connais un peu le DNS, j'utiliserai souvent dans cet article des comparaisons avec le DNS.

Pour résumer DDT en un paragraphe : dans LISP (RFC 6830 [1]), l'identificateur se nomme EID ("*Endpoint Identifier*") et le localisateur RLOC ("*Routing Locator*"). Les EID ont une structure arborescente (leur forme syntaxique est celle d'adresses IP). Cet arbre est réparti sur plusieurs serveurs, les nœuds DDT. Un **nœud DDT** fait autorité pour un certain nombre de préfixes d'EID. Il délègue ensuite les sous-préfixes à d'autres nœuds DDT, ou bien à des "*Map Servers*" LISP (RFC 6833) quand on arrive en bas de l'arbre. (Une des différences avec le DNS est donc que les serveurs qui délèguent sont d'une nature distincte de ceux qui stockent les feuilles de l'arbre.)

LISP a une interface standard avec les serveurs qui font la résolution d'EID en RLOC, décrite dans le RFC 6833. En gros, le client envoie un message `Map-Request` et obtient une réponse `Map-Reply`, ou bien une délégation (`Map-Referral`) qu'il va devoir suivre en envoyant le `Map-Request` suivant au RLOC indiqué dans la délégation. Derrière cette interface, LISP ne spécifie pas comment les serveurs

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc6830.txt`

obtiennent l'information. Plusieurs propositions ont déjà été faites (comme ALT, dans le RFC 6836, ou NERD, dans le RFC 6837...), auxquelles s'ajoute celle de notre RFC. Un bon résumé est dans cette image <http://4.bp.blogspot.com/-LfHp8GH9QDY/T_1zUF8diQI/AAAAAAAAAAY/LDlm4WQkJyQ/s1600/ddt-packet-flow.png> (mais qui ne montre qu'un seul niveau de délégation, il peut y en avoir davantage.)

DDT vise avant tout le passage à l'échelle, d'où la structuration hiérarchique de l'information. La notion de **délégation** (d'un préfixe général à un sous-préfixe plus spécifique) est centrale dans DDT. Un client (le routeur LISP qui a un paquet destiné à un EID donné et qui cherche à quel RLOC le transmettre, ou bien un résolveur, un serveur spécialisé agissant pour le compte de ce routeur) va donc devoir suivre cette délégation, descendant l'arbre jusqu'à l'information souhaitée.

La délégation est composée, pour un préfixe délégué, d'un ensemble de RLOC (pas d'EID pour éviter des problèmes d'œuf et de poule) désignant les nœuds qui ont une information sur le préfixe délégué. (Ce sont donc l'équivalent des enregistrements NS du DNS, mais avec une indirection en moins, comme si la partie droite d'un enregistrement NS stockait directement une adresse IP.)

J'ai écrit jusque là que la clé d'accès à l'information (rôle tenu par le nom de domaine dans le DNS) était l'EID mais c'est en fait un peu plus compliqué : la clé est le XEID ("*eXtended EID*"), qui est composé de plusieurs valeurs, dont l'EID (section 4.1 de notre RFC).

Pour indiquer au résolveur qu'il doit transmettre sa requête à une autre machine, ce RFC crée un nouveau type de message LISP, `Map-Referral`, type 6 (cf. le registre IANA <https://www.iana.org/assignments/lisp-parameters/lisp-parameters.xml#lisp-packet-types>) détaillé en section 6, envoyé en réponse à un `Map-Request`, quand le nœud DDT ne connait pas la réponse. (Comme indiqué plus haut, c'est l'équivalent d'une réponse DNS avec uniquement une section Autorité contenant des enregistrements NS.)

Continuons un peu la terminologie (section 3 du RFC) :

— Un client DDT est une machine qui interroge les nœuds DDT (avec un `Map-Request`, cf. RFC 6833) et suit les `Map-Referral` jusqu'au résultat. C'est en général un résolveur ("*Map Resolver*", RFC 6833) mais cela peut être aussi un routeur LISP (ITR, "*Ingress Tunnel Router*").
— Un résolveur est serveur d'un côté, pour les routeurs qui envoient des `Map-Request`, et client DDT de l'autre, il envoie des requêtes DDT. Il gère un cache (une mémoire des réponses récentes). Le résolveur maintient également une liste des requêtes en cours, pas encore satisfaites.

La base des données des serveurs DDT est décrite en section 4. Elle est indexée par XEID. Un XEID est un EID (identificateur LISP) plus un AFI ("*Address Family Identifier*", 1 pour IPv4, 2 pour IPv6, etc), un identificateur d'instance (voir RFC 6830, section 5.5, et RFC 8060, section 4.1) qui sert à avoir plusieurs espaces d'adressage, et quelques autres paramètres, pas encore utilisés. Configurer un serveur DDT, c'est lui indiquer la liste de XEID qu'il doit connaitre, avec les RLOC des serveurs qui pourront répondre. Désolé, je n'ai pas de serveur DDT sous la main mais on peut trouver un exemple, dans la documentation de Cisco <http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_lisp/command/ip-lisp-cr-book/lisp-ddt-cfg-cmds.html>, où on délègue au "*Map Server*" de RLOC `10.1.1.1` :

```
router lisp
    ddt authoritative 2001:db8:eeee::/48
          delegate 10.1.1.1 eid-prefix 172.16.0.0/16
          delegate 10.1.1.1 eid-prefix 2001:db8:eeee::/48
```

Un autre exemple de délégation est l'actuelle liste des données dans la racine DDT <http://ddt-root.org/#root>.

Le DNS n'a qu'un type de serveurs faisant autorité, qu'ils soient essentiellement serveurs de délégation (ceux des TLD, par exemple) ou qu'ils soient serveurs « finaux » contenant les enregistrements autres que NS. Au contraire, LISP+DDT a deux types de serveurs, les nœuds DDT présentés dans ce RFC, qui ne font que de la délégation, et les traditionnels "*Map Servers*", qui stockent les correspondances entre EID et RLOC (entre identificateurs et localisateurs). Dit autrement, DDT ne sert pas à trouver la réponse à la question « quel est le RLOC pour cet EID », il sert uniquement à trouver le serveur qui pourra répondre à cette question.

Comme pour le DNS, il existe une racine, le nœud qui peut répondre (enfin, trouver une délégation) pour tout XEID. (Sur le Cisco cité plus haut, la directive `ddt root` permettra d'indiquer le RLOC des serveurs de la racine, voir aussi la section 7.3.1 de notre RFC.) Une racine expérimentale existe, vous trouverez ses RLOC en <http://ddt-root.org/>.

La section 5 de notre RFC décrit en détail la modification au message `Map-Request` que nécessite DDT. Ce message était normalisé par le RFC 6830 (section 6.1.2) et un seul ajout est fait : un bit qui était laissé vide sert désormais à indiquer que la requête ne vient pas directement d'un routeur LISP mais est passée par des nœuds DDT.

La section 6, elle, décrit un type de message nouveau, `Map-Referral`, qui contient les RLOC du nœud DDT qui pourra mieux répondre à la question. Cette réponse contient un code qui indique le résultat d'un `Map-Request`. Ce résultat peut être « positif » :

— `NODE-REFERRAL`, renvoi à un autre nœud DDT,
— `MS-REFERRAL`, renvoi à un "*Map Server*" (rappelez-vous que, contrairement au DNS, il y a une nette distinction entre nœud intermédiaire et "*Map Server*" final),
— `MS-ACK`, réponse positive d'un "*Map Server*".

Mais aussi des résultats « négatifs » :

— `MS-NOT-REGISTERED`, le "*Map Server*" ne connait pas cet EID,
— `DELEGATION-HOLE`, l'EID demandé tombe dans un trou (préfixe non-LISP dans un préfixe LISP),
— `NOT-AUTHORITATIVE`, le nœud DDT n'a pas été configuré pour ce préfixe.

Le fonctionnement global est décrit en détail dans la section 7 du RFC. À lire si on veut savoir exactement ce que doivent faire le "*Map Resolver*", le "*Map Server*", et le nouveau venu, le nœud DDT. La même description figure sous forme de pseudo-code dans la section 8. Par exemple, voici ce que doit faire un nœud DDT lorsqu'il reçoit un `Map-Request` (demande de résolution d'un EID en RLOC) :

```
if ( I am not authoritative ) {
    send map-referral NOT_AUTHORITATIVE with
     incomplete bit set and ttl 0
} else if ( delegation exists ) {
    if ( delegated map-servers ) {
        send map-referral MS_REFERRAL with
          ttl 'Default_DdtNode_Ttl'
    } else {
        send map-referral NODE_REFERRAL with
          ttl 'Default_DdtNode_Ttl'
    }
} else {
    if ( eid in site) {
        if ( site registered ) {
            forward map-request to etr
            if ( map-server peers configured ) {
                send map-referral MS_ACK with
                 ttl 'Default_Registered_Ttl'
            } else {
```

```
                send map-referral MS_ACK with
                 ttl 'Default_Registered_Ttl' and incomplete bit set
            }
        } else {
            if ( map-server peers configured ) {
                send map-referral MS_NOT_REGISTERED with
                 ttl 'Default_Configured_Not_Registered_Ttl'
            } else {
                send map-referral MS_NOT_REGISTERED with
                 ttl 'Default_Configured_Not_Registered_Ttl'
                 and incomplete bit set
            }
        }
    } else {
        send map-referral DELEGATION_HOLE with
         ttl 'Default_Negative_Referral_Ttl'
    }
}
```

Un exemple complet et détaillé figure dans la section 9, avec description de tous les messages envoyés.

Question sécurité, je vous renvoie à la section 10 du RFC. DDT dispose d'un mécanisme de signature des messages (l'équivalent de ce qu'est DNSSEC pour le DNS). La délégation inclut les clés publiques des nœuds à qui on délègue.

Il existe au moins deux mises en œuvre de DDT, une chez Cisco et l'autre chez OpenLisp <http://www.openlisp.org/>. (Le RFC ne sort que maintenant mais le protocole est déployé depuis des années.)