

Proposition de présentation courte

Journées Cloud 2018

Adrien Lebre, Alexandre van Kempen

1 Localisation de réplicas dans les infrastructures Fog prenant en compte la topologie physique

Dans cette présentation, nous nous intéresserons aux architectures Fog, et plus particulièrement aux problématiques liées au stockage des données dans ce type d'architecture. Déployer un système de stockage sur de telles infrastructures présente de nombreux avantages tels que garantir un accès aux données avec une très faible latence, confiner le trafic réseau, garantir l'accès aux données malgré une partition réseau, etc... Lorsqu'un objet est accédé depuis un certain site, il est généralement mis en cache afin de garantir une faible latence pour les futures lectures. De ce fait, de multiples replicas d'un même objet peuvent coexister, chacun étant stocké dans les caches des différents sites. Le système de stockage doit donc être capable de localiser les différents réplicas d'un même objet afin de rediriger une requête vers le réplica le plus proche. Dans les systèmes existants, les informations relatives à la localisation des réplicas sont généralement stockées soit sur un serveur centralisé dans le cloud, soit sur une DHT (Distributed Hash Table). Cependant, ces deux solutions ne tiennent pas compte de la topologie réelle de l'infrastructure, ce qui peut engendrer de forts délais pour récupérer un objet. En effet, il peut être parfois plus lent d'obtenir l'information de localisation que l'objet lui même.

Afin de pallier aux problèmes inhérents à toute solution qui ne prend pas en compte la topologie réelle du réseau, nous proposons un protocole de gestion de localisation de réplicas qui tire partie de l'infrastructure réelle du système. Au cours de la présentation, nous rappellerons les enjeux liés au stockage d'objets dans une infrastructure Fog et détaillerons les problèmes des solutions existantes, notamment au niveau latence et tolérance aux partitions réseaux. Puis nous décrirons un algorithme qui permet de sélectionner l'ensemble minimum des sites devant être informés de la présence d'un nouveau réplica, afin que chaque requête soit servie le plus rapidement possible. Nous décrirons enfin les différentes couches logicielles permettant l'implémentation de notre protocole dans un système de stockage réel.