Configuration automatique de réseaux privés virtuels grâce à une table de hachage distribuée

Rémy Grünblatt – remy@grunblatt.org

Décembre 2022

Afin de déployer des infrastructures logicielles distribuées de manière sécurisée, il est d'usage de reposer sur l'utilisation de réseaux privés virtuels (VPN) sécurisant les communications entre les différentes parties de ces infrastructures, par exemple entre un serveur web « frontal » servant de terminaison TLS et des serveurs « dorsaux » pour différents services. Cette sécurisation est particulièrement bienvenue lorsque l'on considère des services géo-distribués, de type « edge », car les échanges utilisent très souvent Internet comme moyen d'interconnexion, avec tout ce que cela implique d'un point de vue de la sécurité. La configuration d'un VPN est donc bien souvent un prérequis à la mise en place d'architectures plus complexes.

Les tables de hachage distribuées (DHT) permettent de fournir des fonctionnalités analogues à celles d'une table de hachage, dans un système distribué, par exemple Internet. Parmi les propriétés intéressantes des DHT, on peut noter leur scalabilité (passage à l'échelle), leur décentralisation, et leur caractère « pair-à-pair » facilitant leur mise en place, car ne nécessitant pas de serveurs ou d'infrastructure déjà en place.

Le but de ce stage est d'explorer l'utilisation d'une DHT, par exemple la DHT « Mainline », basée sur Kademlia et notamment utilisée pour le protocole BitTorrent, afin de permettre à des nœuds possédant un accès à Internet et un secret partagé d'établir une connexion VPN entre eux (par exemple en utilisant le VPN wireguard), sans avoir initialement connaissance de leurs IPs respectives et de leur localisation respective, et sans se reposer sur une infrastructure existante mis à part celle d'Internet.

Ce stage pourra globalement être structuré en trois parties : une partie d'état de l'art, une partie implémentation, et une partie évaluation / analyse de la solution proposée. Il se déroulera au sein de Télécom SudParis et du laboratoire Samovar, à Évry et à Palaiseau (l'école Télécom SudParis est bilocalisée).

Note : Ce stage possède une composante technique et développement non négligeable. Une bonne familiarité avec un système d'exploitation type Linux et un langage de programmation (a minima Python, dans l'idéal Rust) est plus que bienvenue.