パケットフィルタリングの iptables から BPF への移行妥当性調査

2030014 花家 研一

単一の Web アプリケーションが内包する プライベート IP アドレスの数は Docker を代表とするコンテナ技術と オープンソースのソフトウェアである Kubernetes を代表とするコンテナオーケストレーションシステムの普及により爆発的に増えた。Web 3 層アーキテクチャにおけるアプリケーション層には複数のノードがロードバランサーに紐付いており、ノード内には複数のコンテナが乗り、コンテナはノードを跨いで構築された仮想ネットワーク空間にて CIDR によって区切られた IP レンジの IP アドレスが割り当てられた仮想ネットワークインターフェースを介してノードを跨いだコンテナ間の通信を行っている。巨大なアプリケーションともなるとコンテナの数は何万という数となり、その分の プライベート IP アドレスをコンテナオーケストレーションシステムは管理している。

このような背景から物理ノード内には Linux のカーネル機能である cgroups によって分断された仮想ネットワーク空間が存在しており、ネットワーク脅威防止のためのパケットフィルタリングは物理ノード間のみならずコンテナ間においても実装が求められ、いくつかのコンテナネットワークインターフェース（CNI） はコンテナ間のパケットフィルタリングに対応している。一方で Linux におけるパケットフィルタリングの技術は古くから iptables が用いられてきたが、昨今のクラスタ事情のようにフィルタリングのルールが何万という膨大になるケースにおいてはパフォーマンスのボトルネックとなることがわかった。

そこで代替技術として活用され始めたのが Berkeley Packet Filter（BPF）によるパケットフィルタリングである。BPF 自体はパケットフィルタリング技術として古くから使われてきた技術ではあるが、昨今目覚ましい進化を遂げており、その汎用性、実装容易性からパケットフィルタリングのみならず様々な面で活用の場を広げている昨今注目されている技術の一つである。本研究では iptables と BPF を多角的に考察することでパケットフィルタリング技術の移行妥当性の調査を行う。