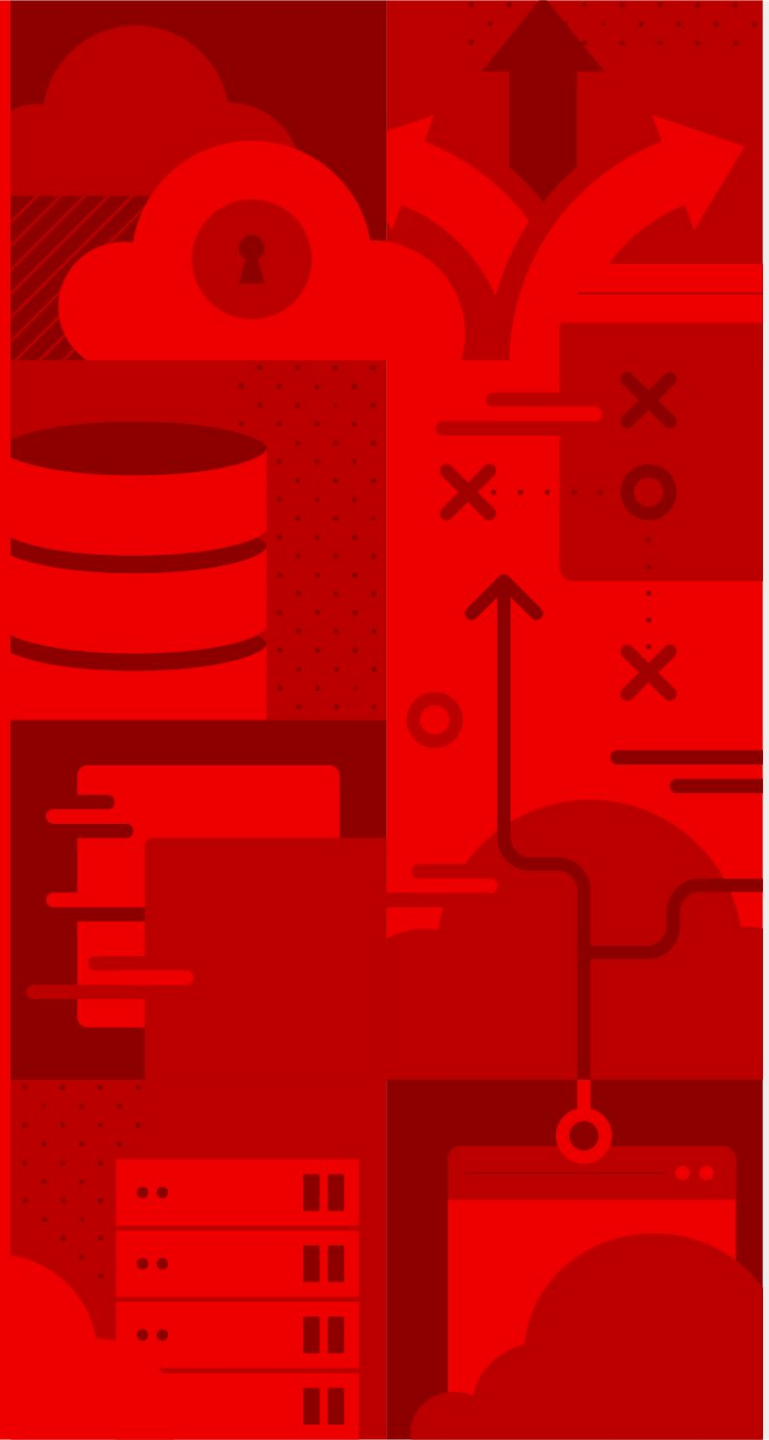


RED HAT CLOUD PLATFORMS BUSINESS UNIT

OVNのご紹介

v1.0

2019-11-01

Manabu Ori
@orimanabu

Red Hat

注意

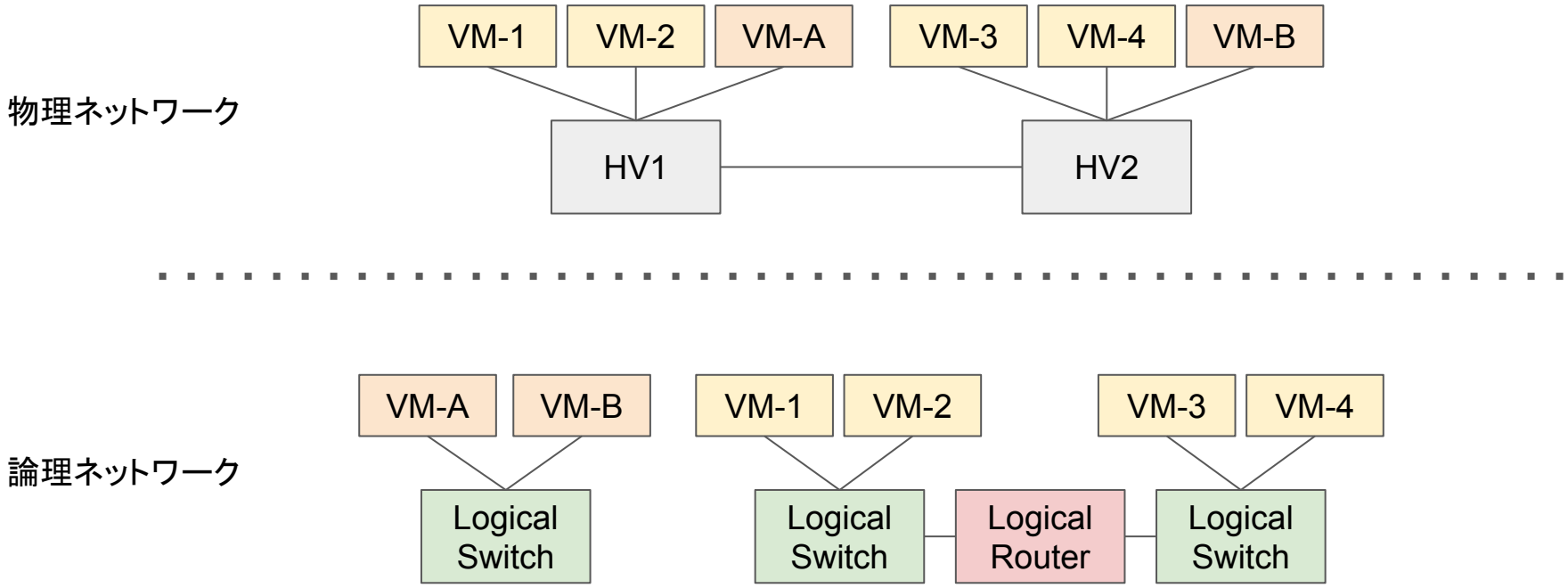
- この資料は2019年11月1日時点の情報を元に作成しました

Red Hat

OVNとは

OVN (Open Virtual Network) とは

- 複数ハイパーバイザ上のOVSにまたがった仮想ネットワークを作る仕組み
- OVS (Open vSwitch) のサブプロジェクトとして、2015年に始動
  - 最初のリリース: 27 Sep 2016 (OVS v2.6)
  - OpenStack Neutron Plugin (networking-ovn) の最初のリリース: 06 Oct 2016 (Newton)
  - OVS v2.11からリポジトリが分離 https://github.com/ovn-org
- オーバーレイネットワークを論理ネットワークとして抽象化



Red Hat

OVNの特徴

- データベース操作によるコンフィギュレーション
- Logical Flowによる設定
  - 物理ネットワーク(OVS)と仮想ネットワークを分離
  - だいたいOpenFlowと同じ気分
    - フローテーブルのパイプライン、フローのmatchとaction
- ハイパーバイザ間のカプセリングはGeneve,STT
- 分散L2, L3処理
- NAT、DHCP、ロードバランサのネイティブ実装
- L2, L3ゲートウェイ
- 他のCMS (Cloud Management System) と連携することを想定したデザイン
  - OpenStack, Kubernetes, Docker, Mesos, oVirt, ...

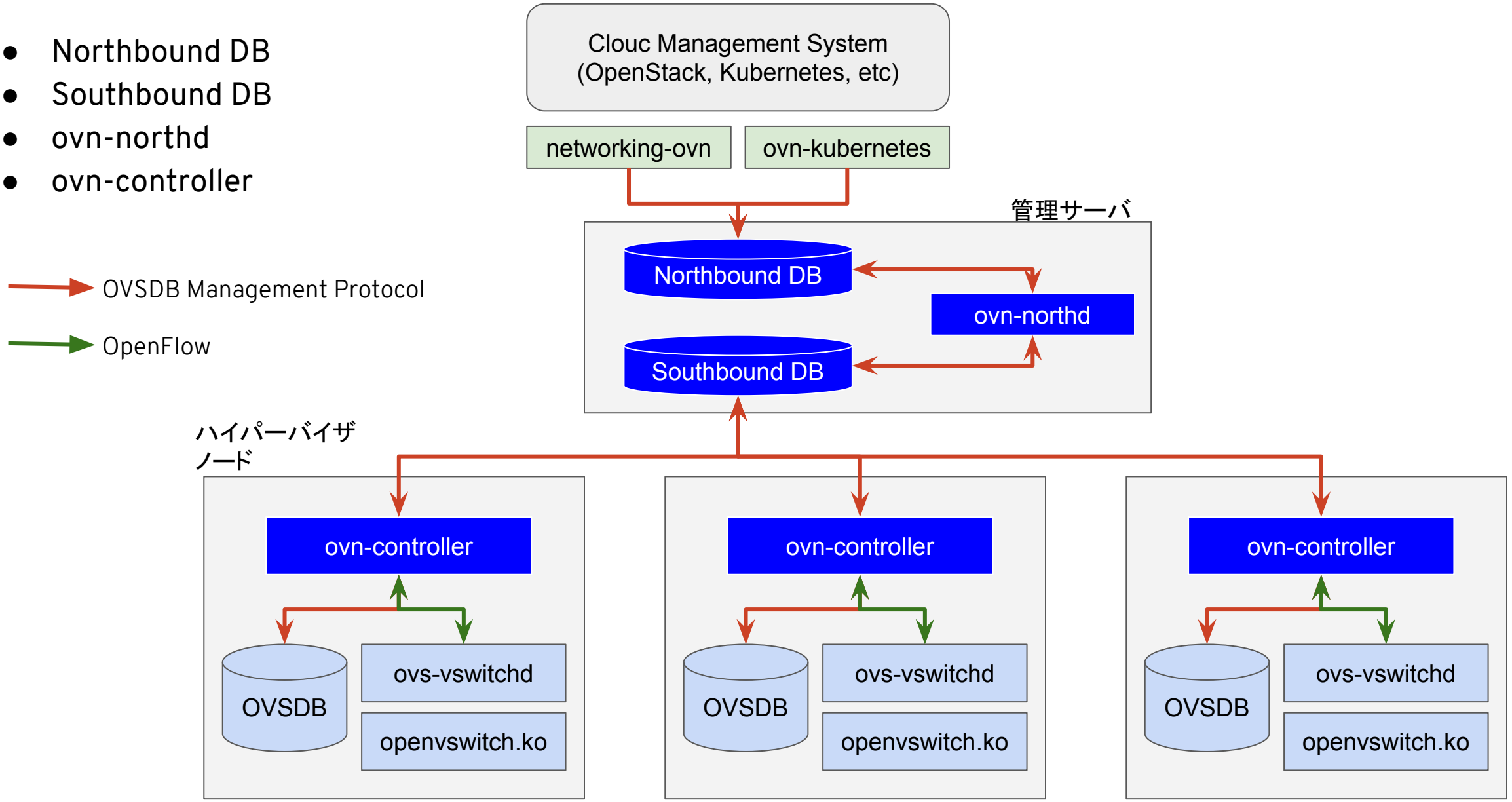
	OVS	OVN
対象	1台のホスト内の仮想スイッチ	複数のホストにまたがる仮想ネットワーク
設定	OpenFlow + OVSDB	Logical Flow + OVSDB

Open vSwitch (OVS) の課題

- OVSは超強力、だけどOpenFlowでSDN環境を構築するのは大変
  - 「現時点では、低レベルのフローロジックを直接作り込む必要があるなど、導入の敷居はあまり低くありません」
    - 技術文書 OpenFlowの概要, VA Linux Systems Japan
  - 「プログラミング言語に例えるとアセンブラ、もしくは標準ライブラリがないC言語」
    - マスタリングTCP/IP OpenFlow編, オーム社
- OVSは超強力、だから
  - OVSネイティブな機能を活用するとより効率的に処理できるはず
  - 現状はOVS, Network Namespace, iptables, etcを組み合わせて様々な機能を実現している
- 仮想化/コンテナ基盤のソフトウェア製品それぞれでOpenFlowの作り込みをするのはつらい
  - OpenStack
  - Kubernetes
  - oVirt, ...

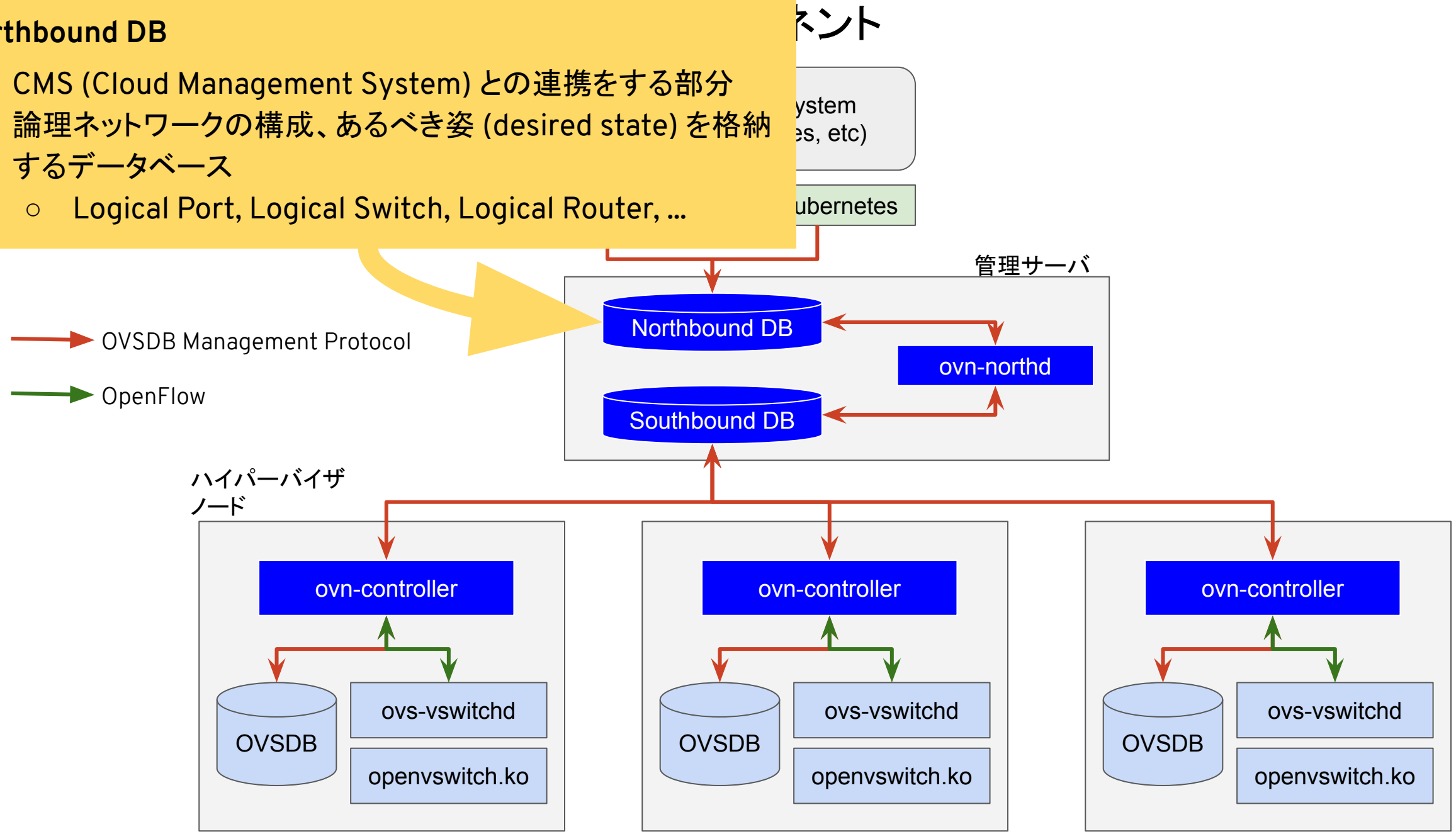
OVNのコンポーネント

- Northbound DB
- Southbound DB
- ovn-northd
- ovn-controller



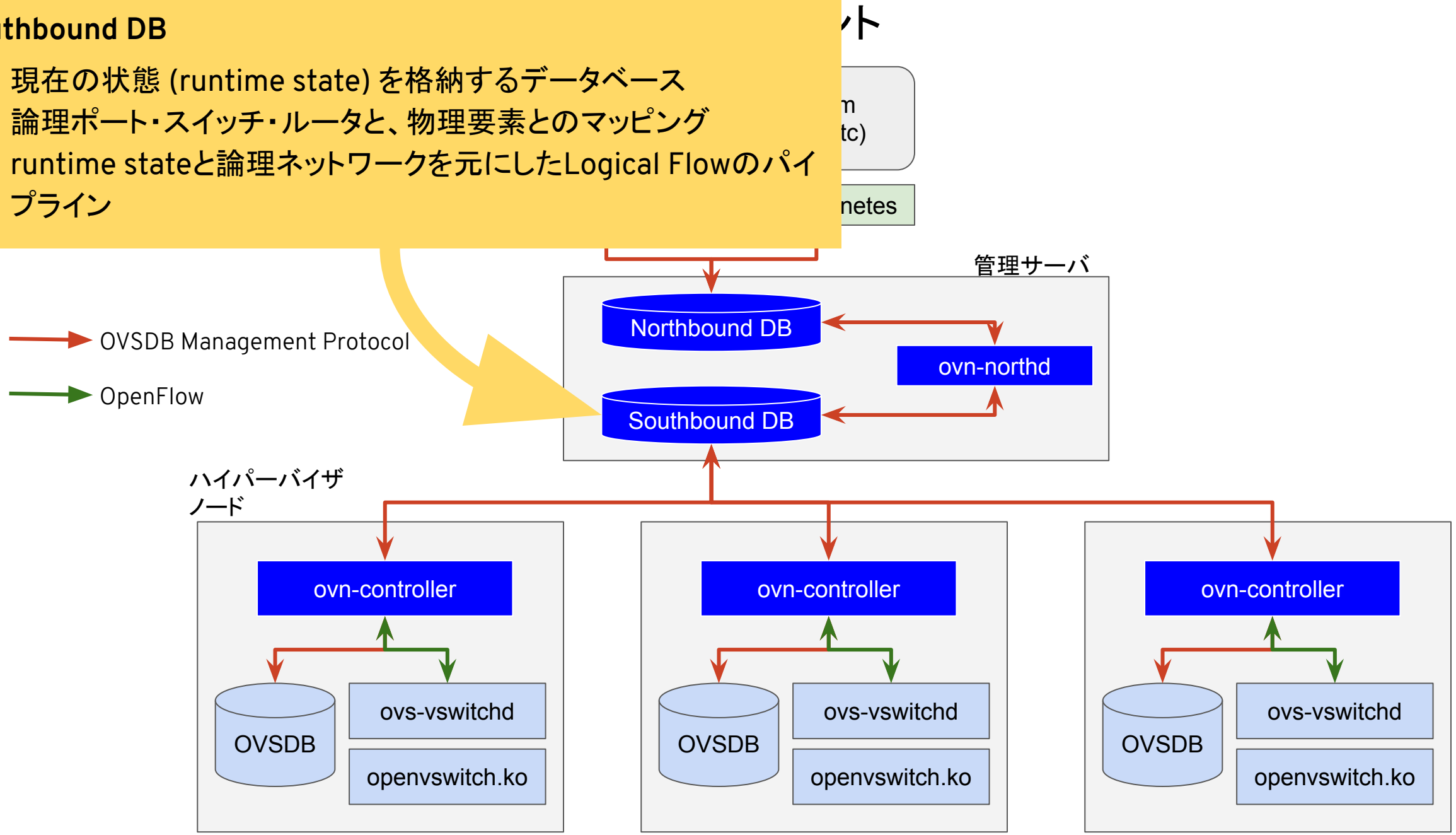
Northbound DB

- CMS (Cloud Management System) との連携をする部分
- 論理ネットワークの構成、あるべき姿 (desired state) を格納するデータベース
  - Logical Port, Logical Switch, Logical Router, ...



Southbound DB

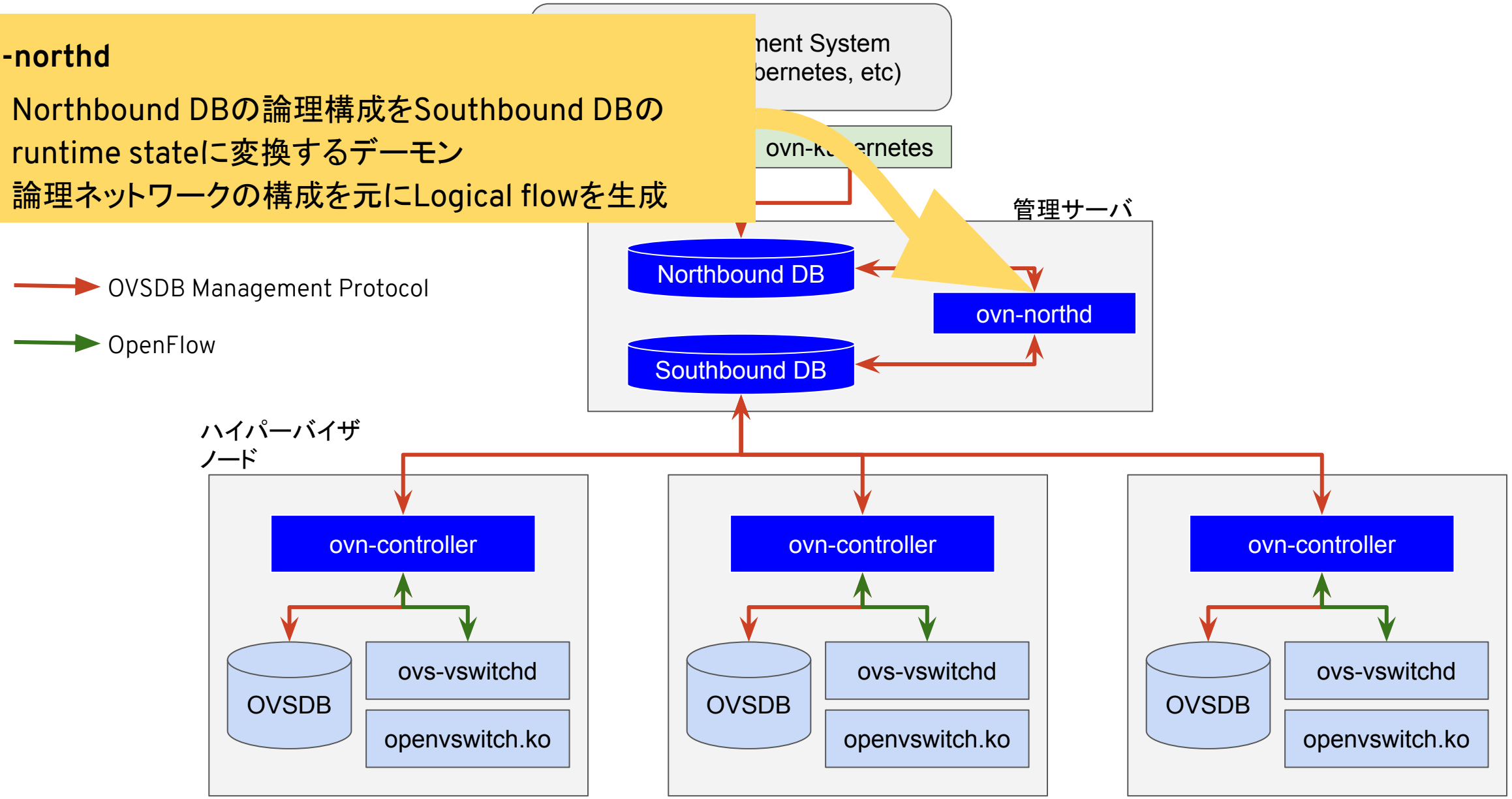
- 現在の状態 (runtime state) を格納するデータベース
- 論理ポート・スイッチ・ルータと、物理要素とのマッピング
- runtime stateと論理ネットワークを元にしたLogical Flowのパイプライン



OVNのコンポーネント

ovn-northd

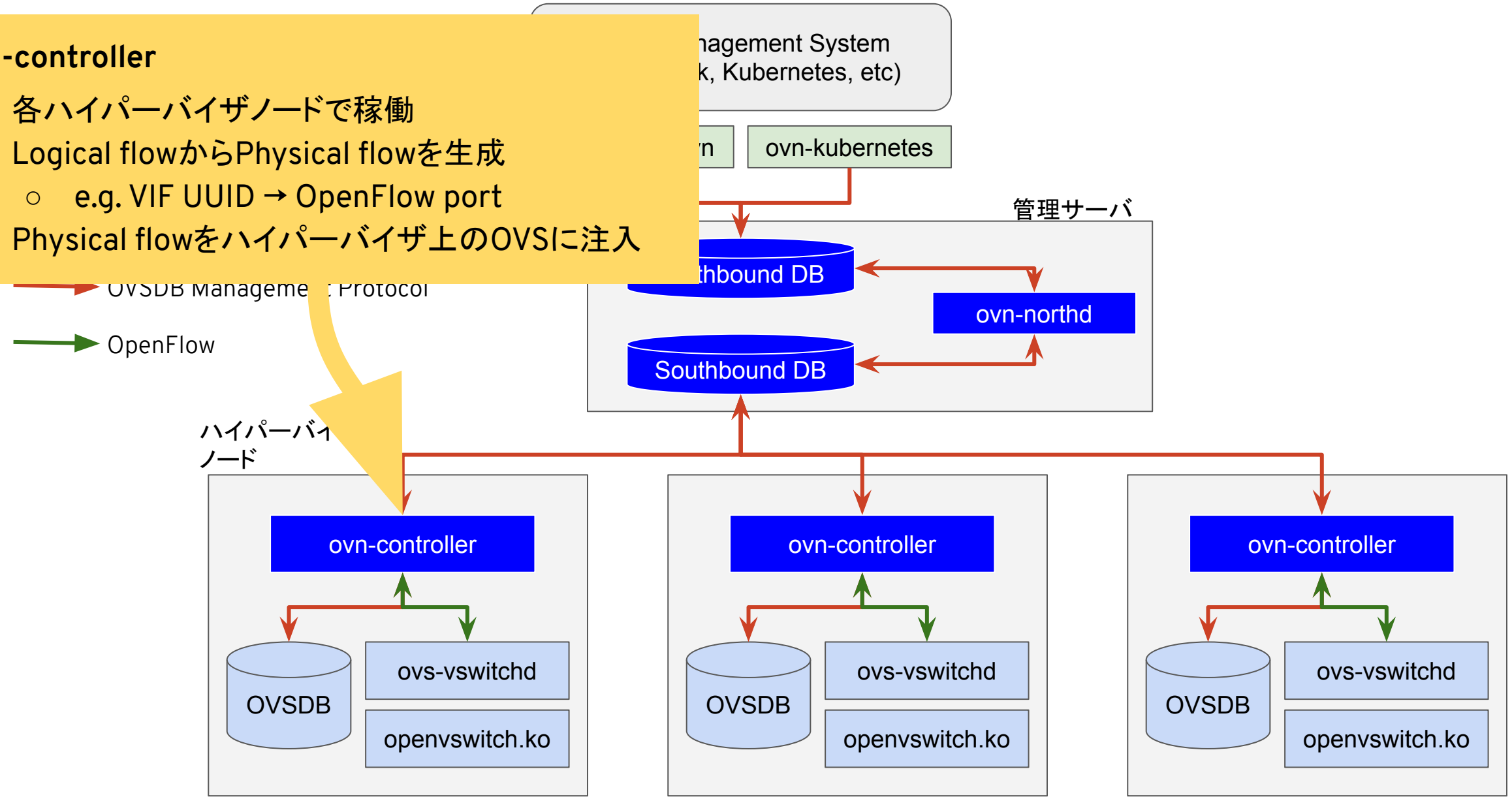
- Northbound DBの論理構成をSouthbound DBのruntime stateに変換するデーモン
- 論理ネットワークの構成を元にLogical flowを生成



OVNのコンポーネント

ovn-controller

- 各ハイパーバイザノードで稼働
- Logical flowからPhysical flowを生成
  - e.g. VIF UUID → OpenFlow port
- Physical flowをハイパーバイザ上のOVSに注入



OVNのコンポーネント

- Northbound DB
  - CMS (Cloud Management System) との連携をする部分
  - 論理ネットワークの構成、あるべき姿 (desired state) を格納するデータベース
    - Logical Port, Logical Switch, Logical Router, ...
- Southbound DB
  - 現在の状態 (runtime state) を格納するデータベース
  - 論理ポート・スイッチ・ルータと、物理要素とのマッピング
  - runtime stateと論理ネットワークを元にしたLogical Flowのパイプライン
- ovn-northd
  - Northbound DBの論理構成をSouthbound DBのruntime stateに変換するデーモン
  - 論理ネットワークの構成を元にLogical flowを生成
- ovn-controller
  - 各ハイパーバイザノードで稼働
  - Logical flowからPhysical flowを生成
    - e.g. VIF UUID → OpenFlow port
  - Physical flowをハイパーバイザ上のOVSに注入

Logical Table Flow Structure - Logical Switch Datapaths

Ingress

Table	Flow category
0	Admission Control and Ingress Port Security - L2
1	Ingress Port Security - IP
2	Ingress Port Security - Neighbor discovery
3	from-lport Pre-ACLs
4	Pre-LB
5	Pre-stateful
6	from-lport ACLs
7	from-lport QoS marking
8	from-lport QoS meter
9	LB
10	Stateful
11	ARP/ND responder

Table	Flow category
12	DHCP option processing
13	DHCP responses
14	DNS Lookup
15	DNS Responses
16	Destination Lookup

Egress

Table	Flow category
0	Pre-LB
1	to-lport Pre-ACLs
2	Pre-stateful
3	LB
4	to-lport ACLs
5	to-lport QoS marking
6	to-lport QoS meter
7	Stateful
8	Egress Port Security - IP
9	Egress Port Security - L2

Logical Table Flow Structure - Logical Router Datapaths

Ingress

Table	Flow category
0	L2 Admission Control
1	IP Input
2	DEFRAG
3	UNSNAT
4	DNAT
5	IPv6 ND RA option processing
6	IPv6 ND RA responder
7	IP Routing
8	ARP/ND Resolution
9	Gateway Redirect
10	ARP Request

Egress

Table	Flow category
0	UNDNAT
1	SNAT
2	Egress Loopback
3	Delivery

OVNの手動構成

- OVSDBの操作
  - ovsdb-tool
  - ovsdb-client

- Logical Switchの作成
  - `ovn-nbctl lswitch-add SWITCH_NAME`
- Logical Portの作成
  - `ovn-nbctl lport-add SWITCH_NAME PORT_NAME`
- Logical PortにMACアドレスを設定
  - `ovn-nbctl lport-set-address PORT_NAME MAC_ADDRESS`
- Logical PortとPhysical Portの紐付け
  - `ovs-vsctl add-port BRIDGE INTERFACE -- set Interface INTERFACE external_ids:iface-id=PORT_NAME`

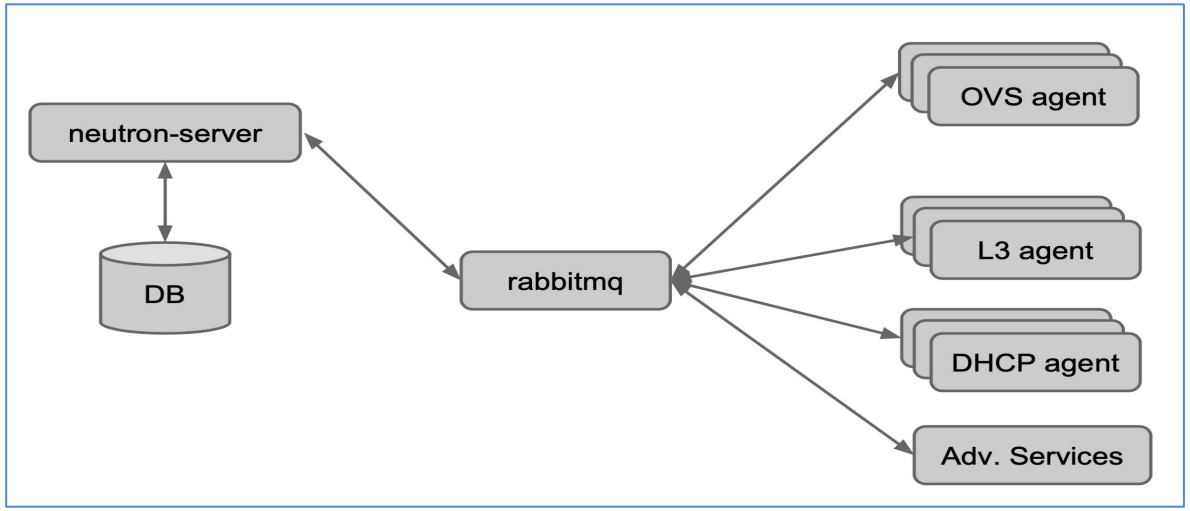
  ↓
- OpenStack, Kubernetes等と連携するときは、この辺りはNeutron ML2 driver/CNI Pluginがやってくれます

OpenStack Integration

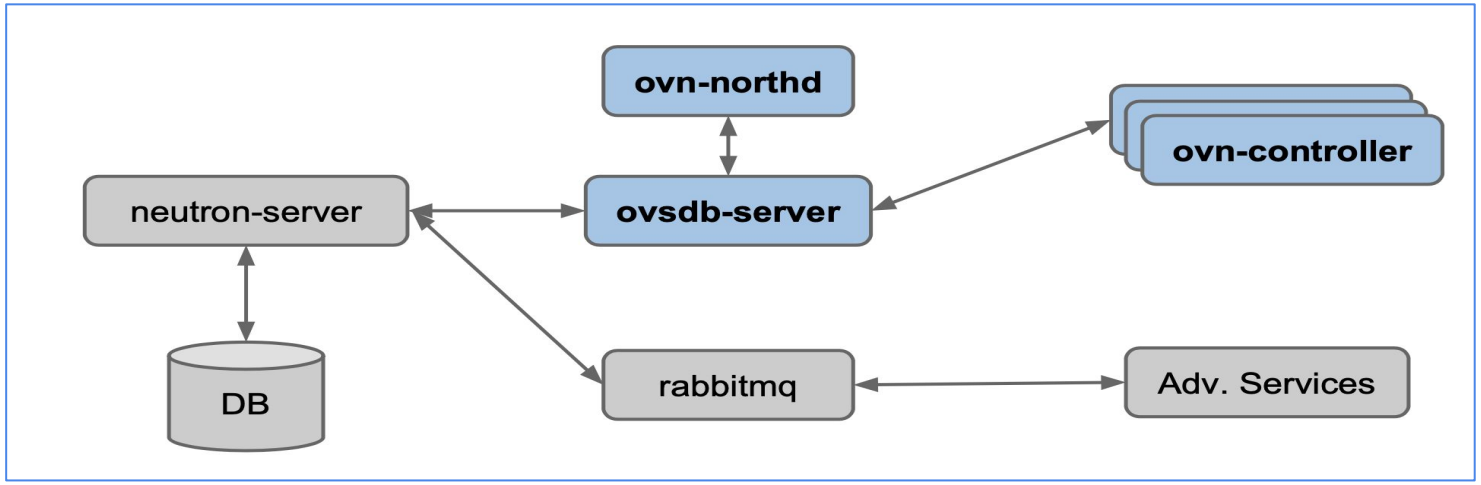
OpenStackとの連携

- Neutron ML2 driver: networking-ovn

ML2/OVS



ML2/OVN



Red Hat

NeutronとOVNの構成要素のマッピング

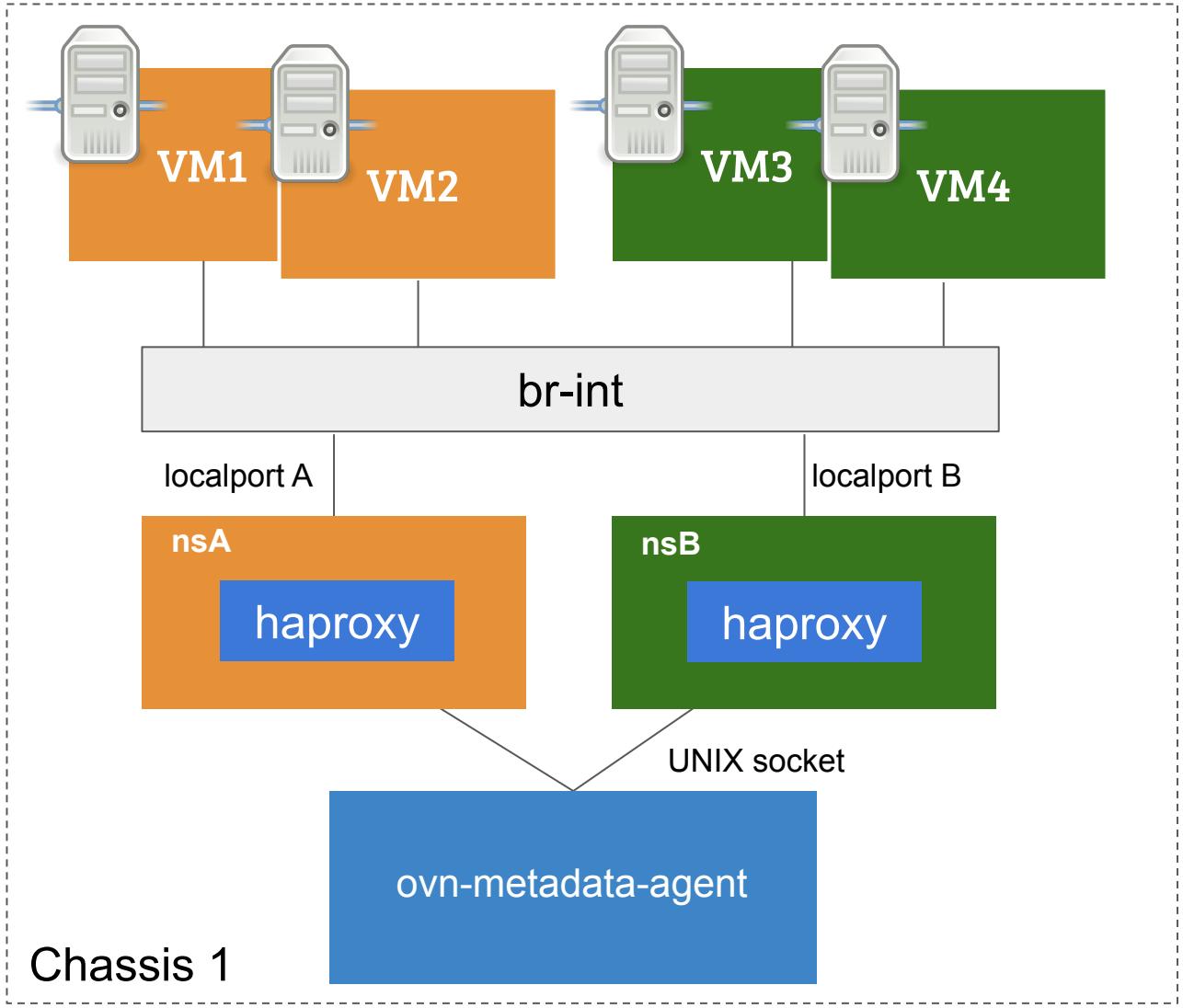
NEUTRON	OVN
router	logical router + gateway_chassis (scheduling)
network	logical switch + dhcp_options
port	logical switch port (+ logical router port)
security group	Port_Group + ACL + Address_Set
floating ip	NAT (dnat_snat entry type)
(in octavia WIP!)	Load_Balancer

networking-ovnの特徴

- L2
  - ARP responderの機能
- L3
  - OVNでIPv4/IPv6ルーティングのネイティブサポート
    - L3 agentは必要ない
  - 分散ルータ
  - namespaceを渡る必要がないので効率的
- Security Group
  - カーネルのconntrackモジュールをOVSから直接利用
  - Neutronの firewall_driver = openvswitch と同じ動き
- DHCP
  - ovn-controllerがDHCPの機能を持つ
    - dhcp agentは必要ない
    - dnsmasqがたくさん地獄にならない
  - シンプルなユースケースのみ想定

networking-ovnの特徴

- Metadata
  - 今の実装では namespace + haproxy
  - metadata-agentとneutron-serverとの通信は不要
- Octavia
  - OVNのOctavia driver開発中
  - Amphora VMが必要なくなる



Red Hat

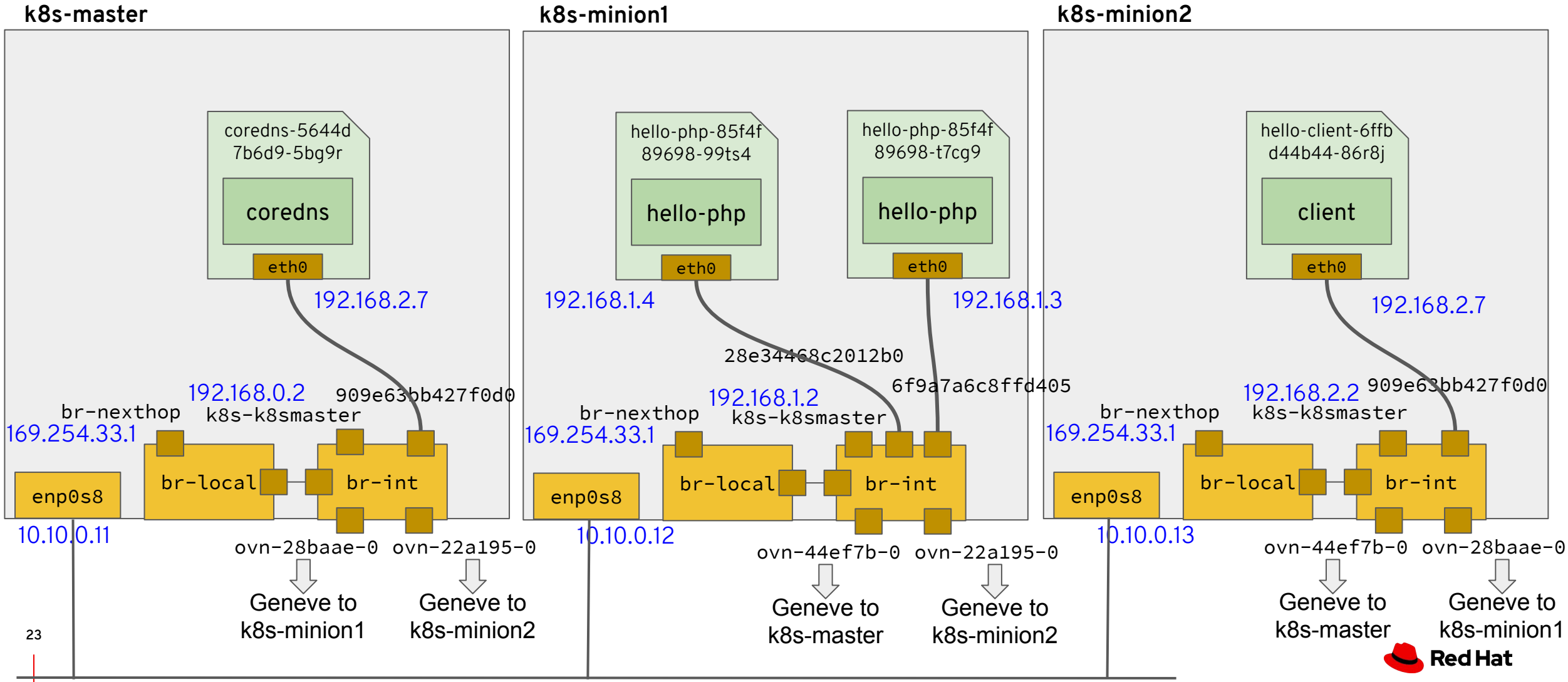
Kubernetes Integration

Kubernetesとの連携

- OVN用のCNIプラグイン: ovn-kubernetes https://github.com/ovn-org/ovn-kubernetes

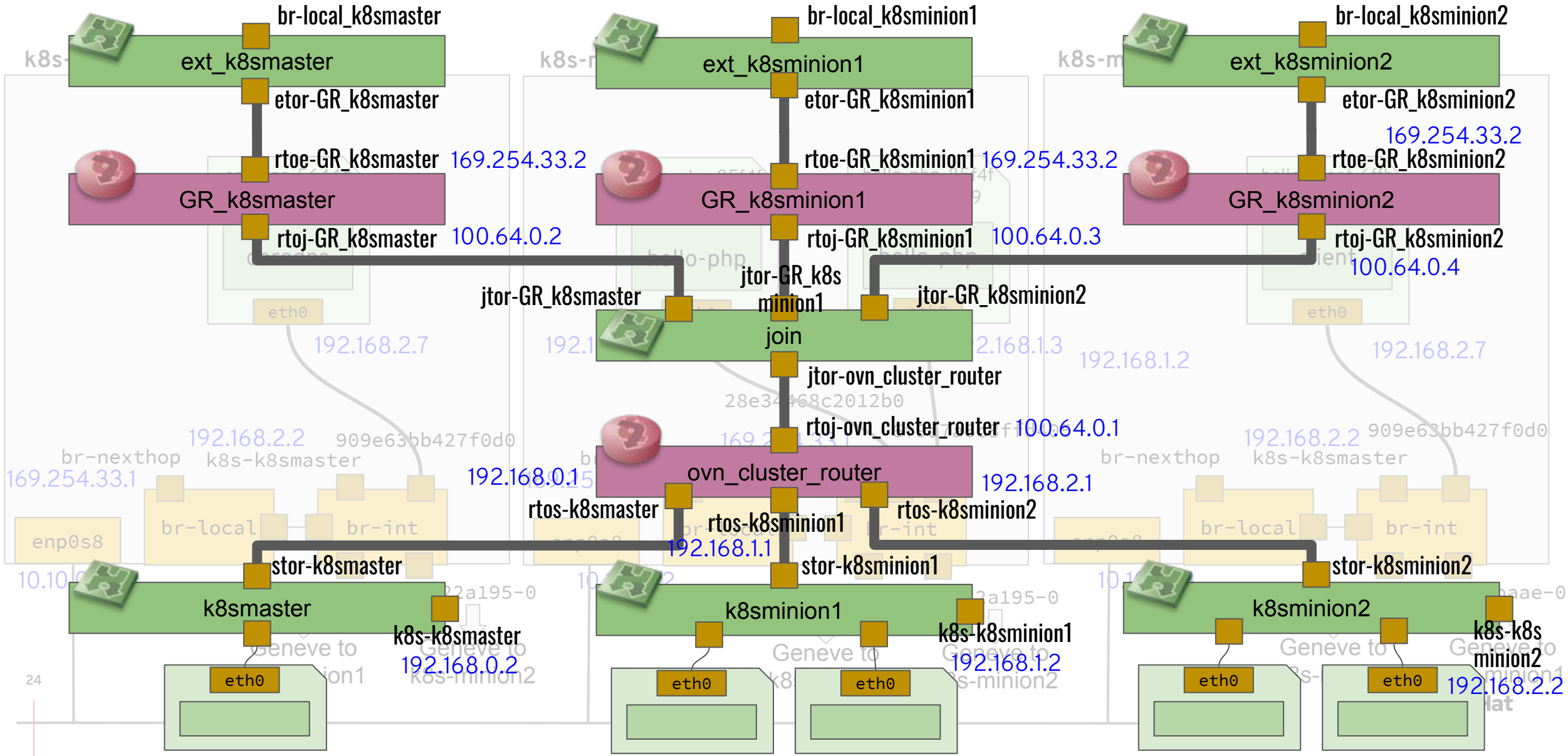
- 他のCNIプラグインとの主な違い
  - Serviceオブジェクトは基本的にOVSの機能で実現している
    - Service → PodのDNAT
    - Service → 複数Podのロードバランス
  - Network Policyの制御はOVSで実現
  - その他はだいたい従来のCNIプラグインと同じ
    - 内部DNSは今のところCoreDNSを使う前提になっている...ように見える

ovn-kubernetes 物理構成

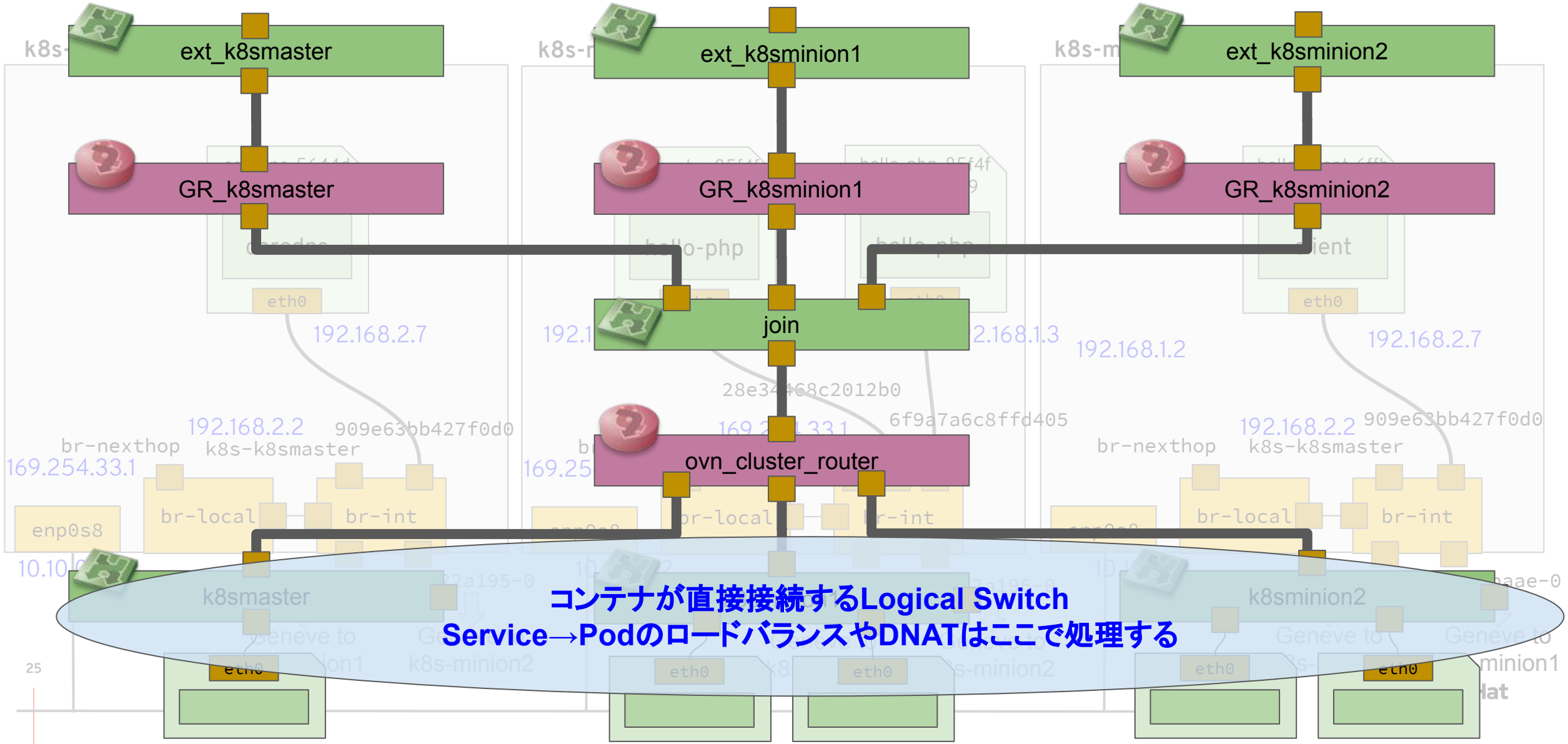


Red Hat

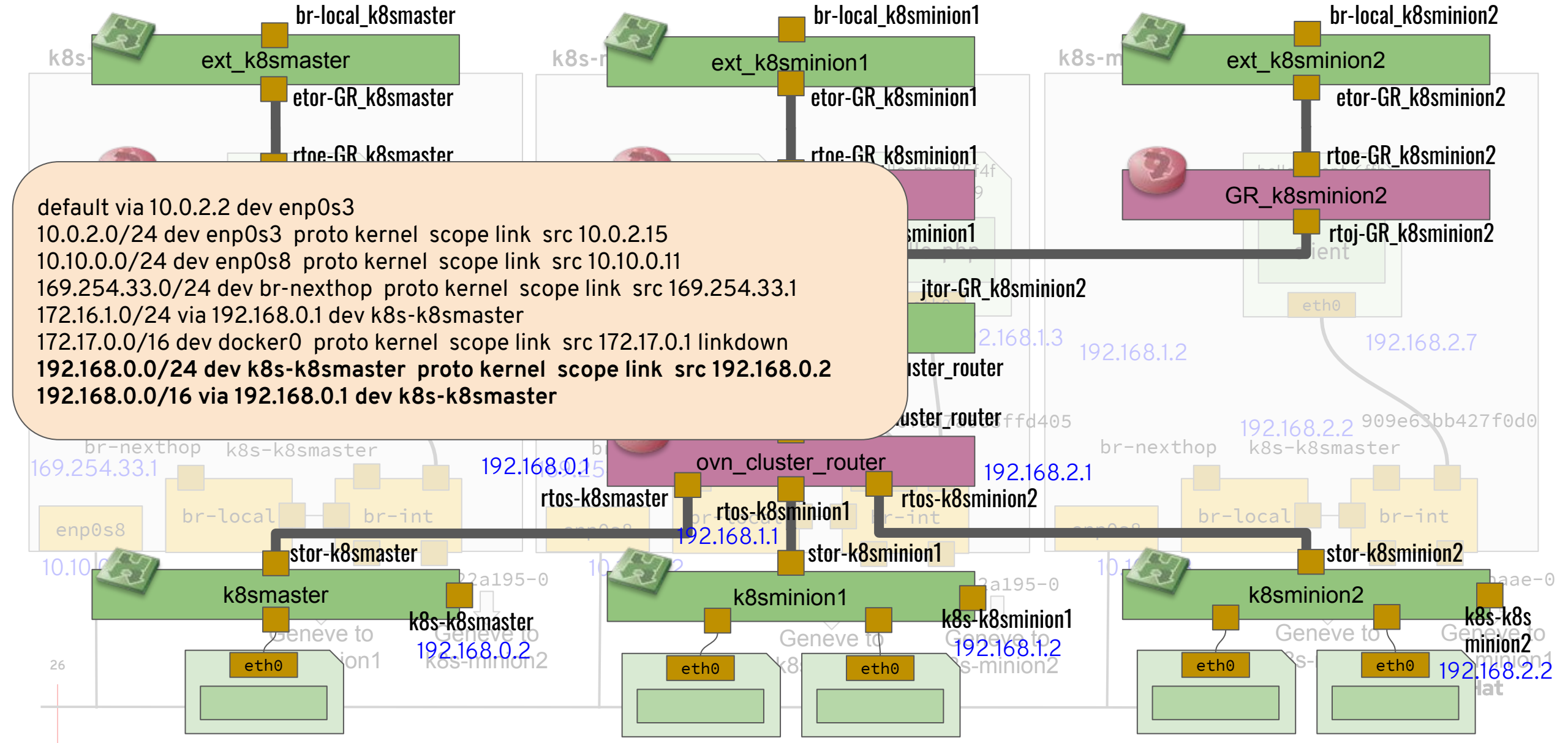
ovn-kubernetes 論理ネットワーク



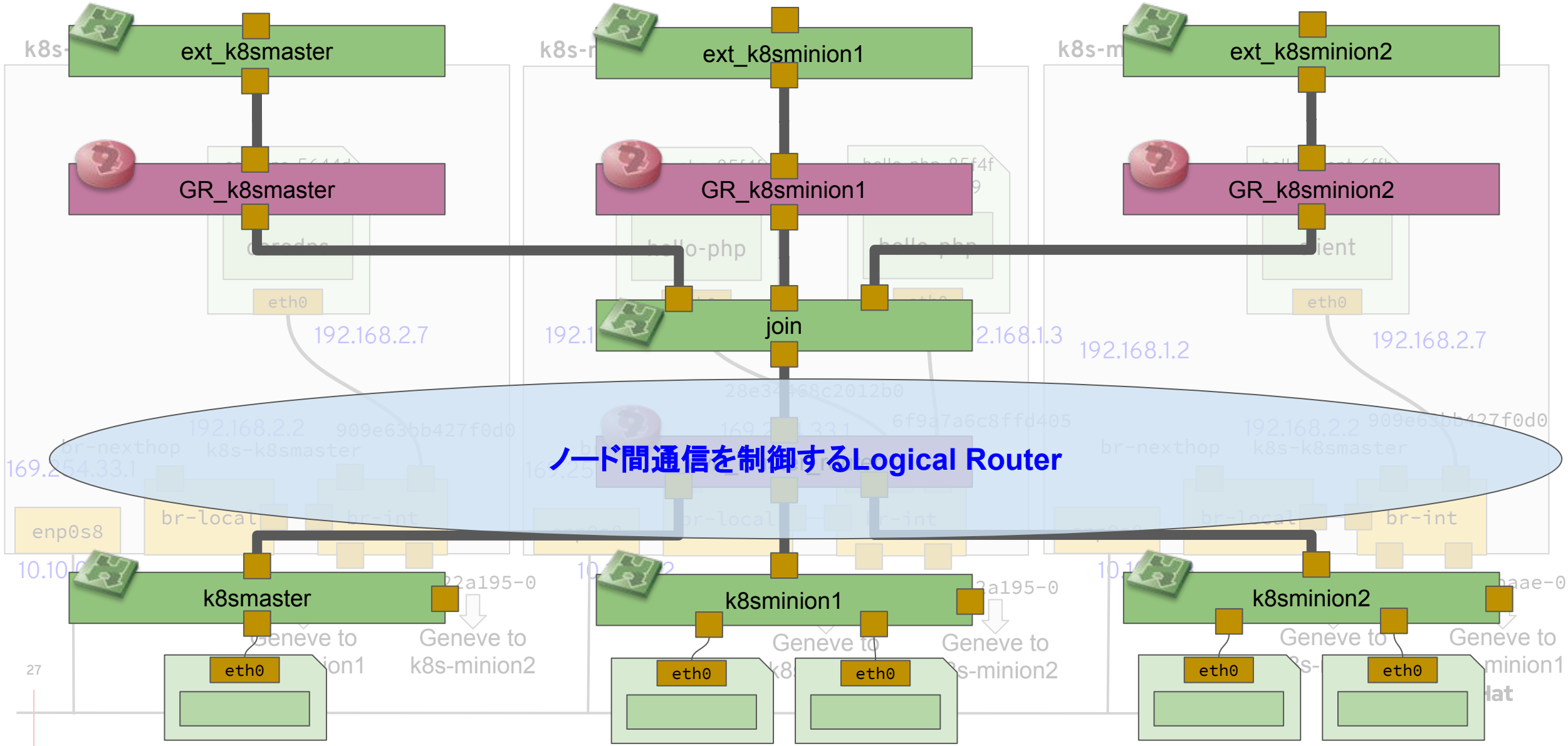
ovn-kubernetes 論理ネットワーク



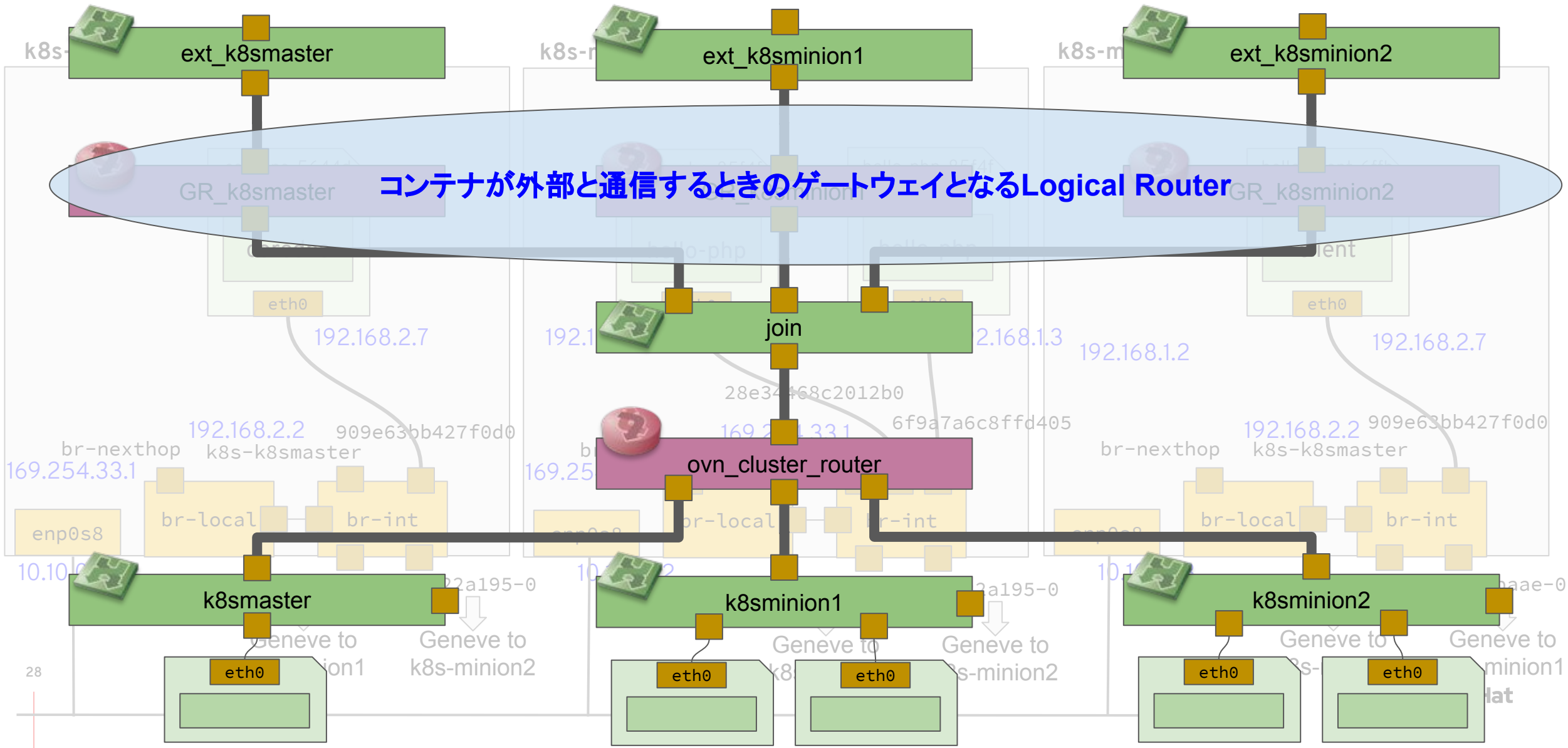
ovn-kubernetes 論理ネットワーク



ovn-kubernetes 論理ネットワーク

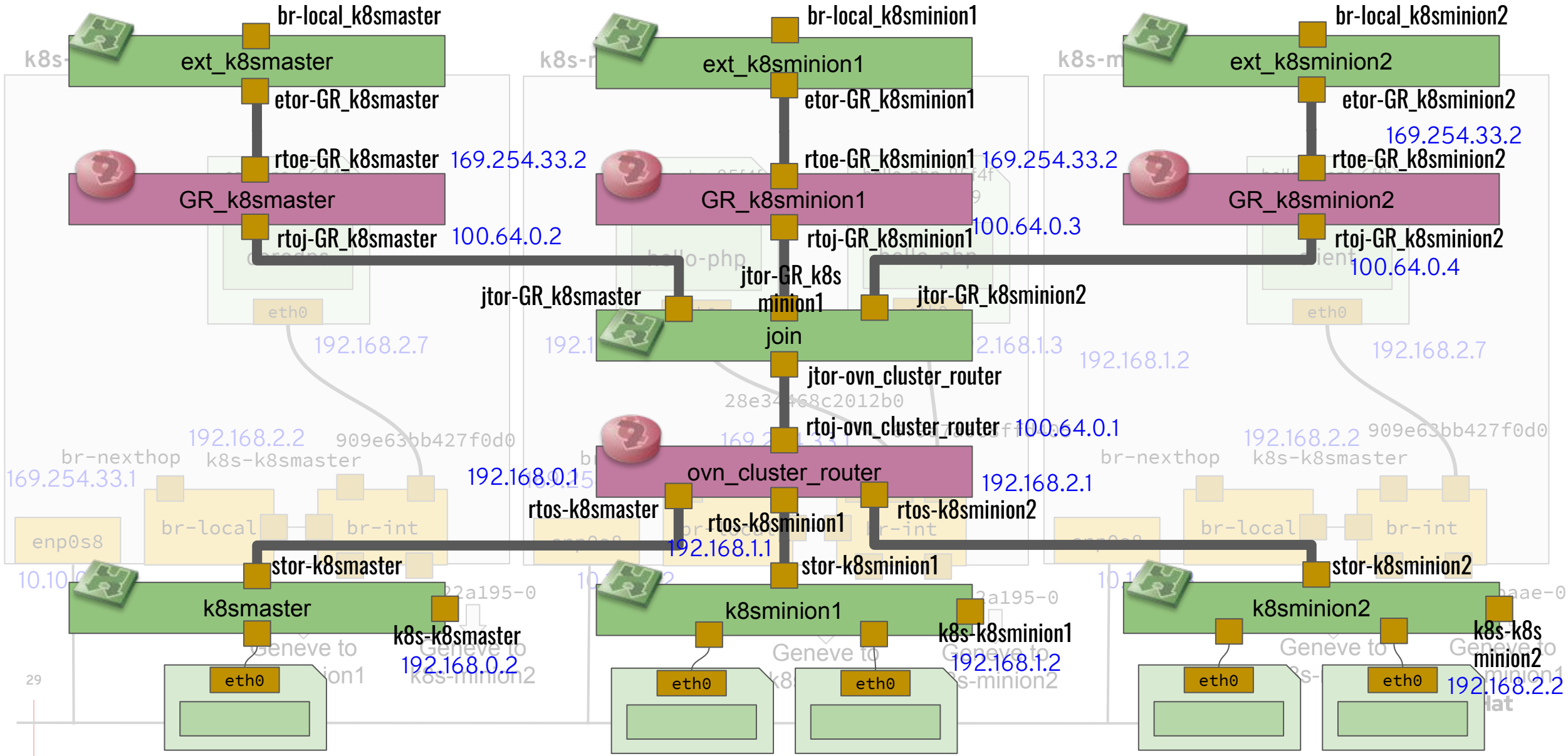


ovn-kubernetes 論理ネットワーク

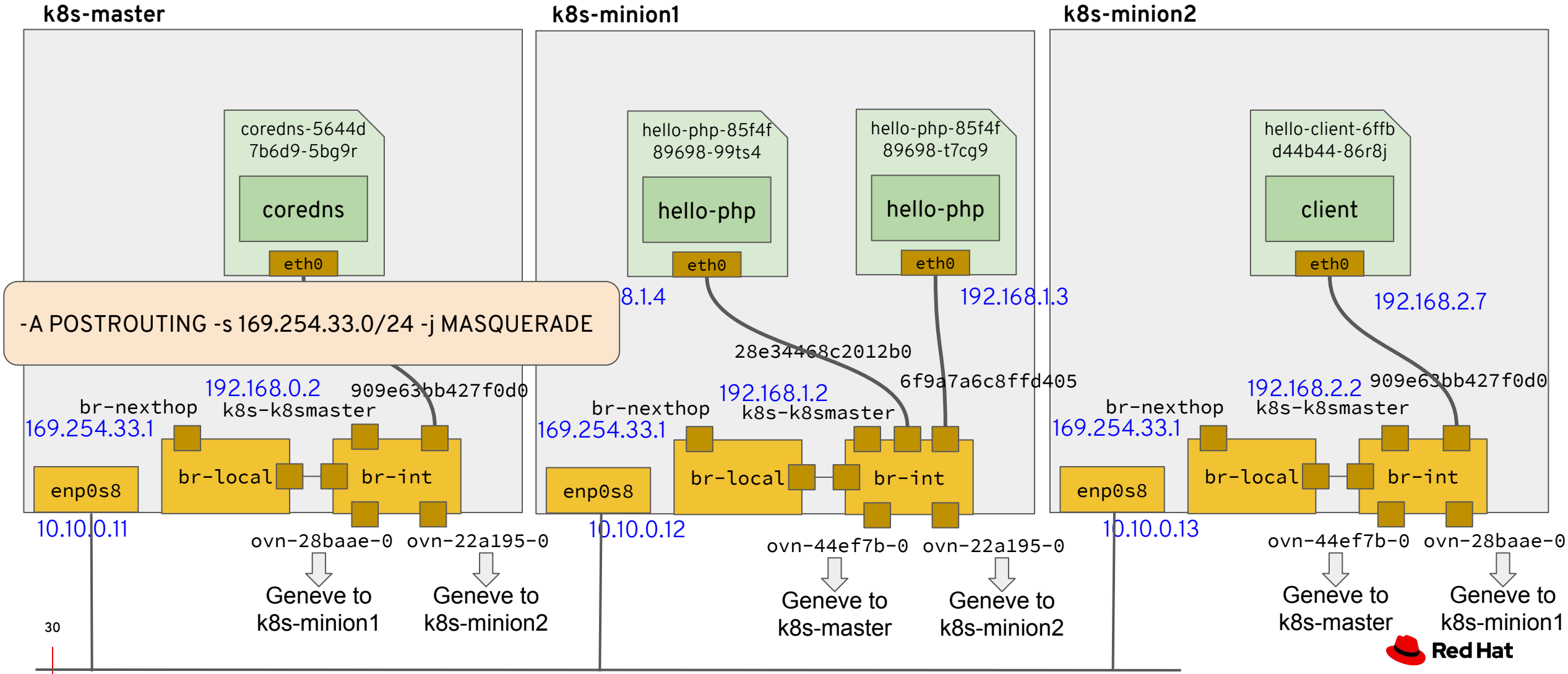


-A POSTROUTING -s 169.254.33.0/24 -j MASQUERADE

netes 論理ネットワーク



ovn-kubernetes 物理構成



Red Hat

OVNの今後

- Multi master OVSDB Server Clustring
- スケーラビリティ改善 (特にOVSDB)
- BPF/DPDK Datapath
- Service Function Chaining

- Red Hatの製品
  - Red Hat OpenStack Platform 15 (Stain)
    - OpenStackの製品版
    - OVNがデフォルトのNeutron ML2 ドライバ
  - Red Hat OpenShift Container Platform 4.2
    - Kubernetesの製品版
    - 4.2でTech Preview、次かその次くらいで正式サポート→デフォルトのCNIプラグイン
  - Red Hat Virtualization
    - 4.2以降でOVNサポート

参考文献

- ovn-architecture(7) http://www.openvswitch.org/support/dist-docs/ovn-architecture.7.txt
- ovn-nb(5) http://www.openvswitch.org/support/dist-docs/ovn-nb.5.txt
- ovn-sb(5) http://www.openvswitch.org/support/dist-docs/ovn-sb.5.txt
- ovn-northd(8) http://www.openvswitch.org/support/dist-docs/ovn-northd.8.txt
- ovn-controller(8) http://www.openvswitch.org/support/dist-docs/ovn-controller.8.txt
- OVSConの資料 http://www.openvswitch.org/support/ovscon2019/

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

linkedin.com/company/Red-Hat

youtube.com/user/RedHatAPAC

facebook.com/RedHatAPAC

twitter.com/Red_Hat_APAC

Red Hat