

# WELCOME



## OVERLAYモデルによる仮想ネットワークング技術 ～NVO3

日本アルカテル・ルーセント 鹿志村 康生  
yasuo.kashimura@alcatel-lucent.com

# NVO3 WG : NETWORK VIRTUALIZATION OVERLAYS (OVER L3)

## <NVO3の目指すソリューション>

既存のL2及びL3のネットワークの制限に縛られず、下記のマルチテナンシの要求事項を実現する。

- テナント間のTrafficの分離
- テナント間のアドレススペースの分離
- DCネットワーク内での自由なVMの配置及びMigration
- 数百万のVMや数十万の物理サーバーに対応可能なスケーリング
- Sub-secondでのVM Migration

<http://datatracker.ietf.org/wg/nvo3/charter/>

IETF 83 (2012 Mar) : NVO3 BOF

IETF 84 (2012 July) : NVO3 WG

# CURRENT TARGET OF NVO3

## <NVO3の現在のTarget>

- Problem statement

[draft-ietf-nvo3-overlay-problem-statement](#)

- Framework document

[draft-ietf-nvo3-framework](#)

- Control plane requirements documents

[draft-ietf-nvo3-nve-nva-cp-req](#)

- Data plane requirements documents

[draft-ietf-nvo3-dataplane-requirements](#)

- Operational requirements

[draft-ashwood-nvo3-operational-requirement](#)

- Gap analysis

[draft-gbclt-nvo3-gap-analysis](#)

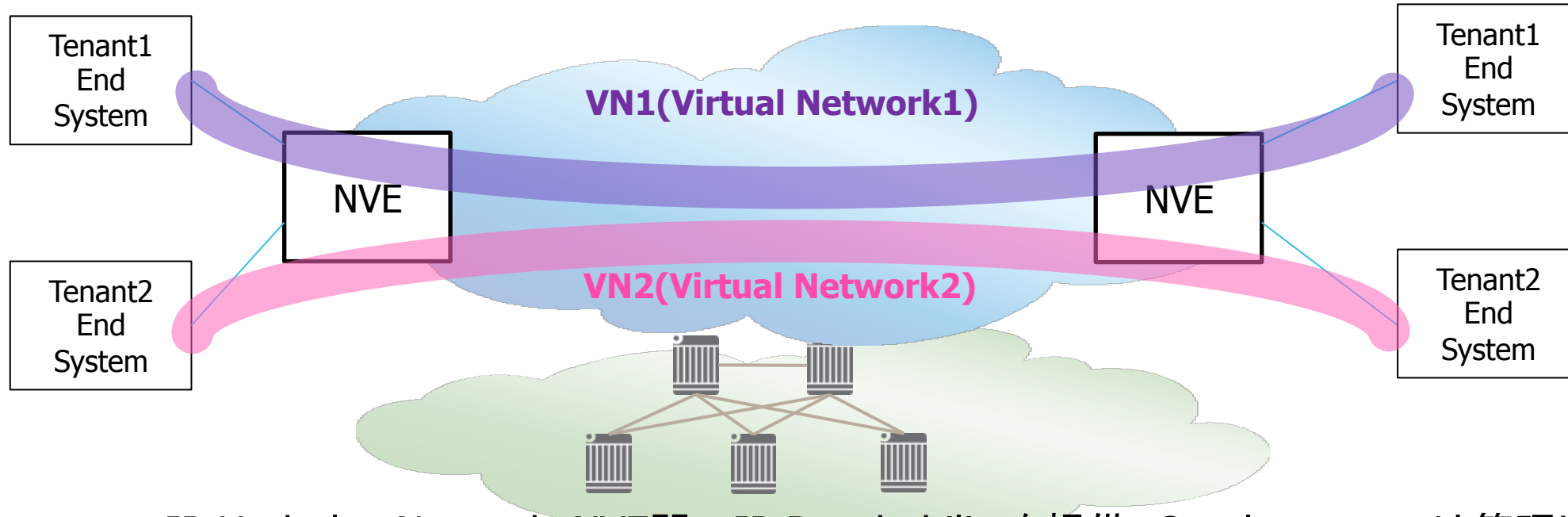
## <その他>

[draft-ietf-nvo3-use-case](#)

[draft-ietf-nvo3-vm-mobility-issues](#)

# NVO3 FRAMEWORK OVERVIEW (1)

Overlay Network:L3 overlayにより仮想ネットワーク機能を提供



IP Underlay Network:NVE間のIP Reachabilityを提供, Overlayのstateは管理しない。

NVE(Network Virtualization Edge) : OverlayによるL2/L3仮想ネットワーク機能を提供するエッジ

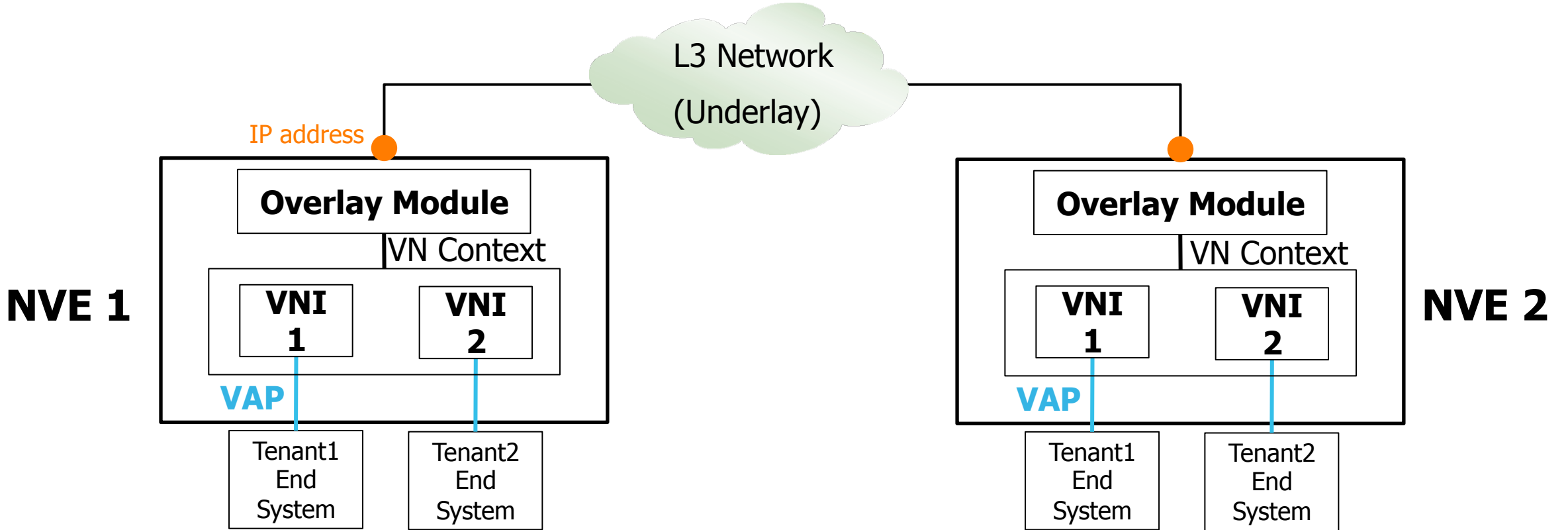
仮想化のためのIP Tunneling機能を持つ

VN(Virtual Network) : テナントシステムに対してL2/L3のネットワークサービスを提供する仮想ネットワーク

Closed User Group(CUG)

# NVO3 FRAMEWORK OVERVIEW (2)

NVE : Network Virtualization Edge



VNI (Virtual Network Instance): Virtual Networkの特定のインスタンス

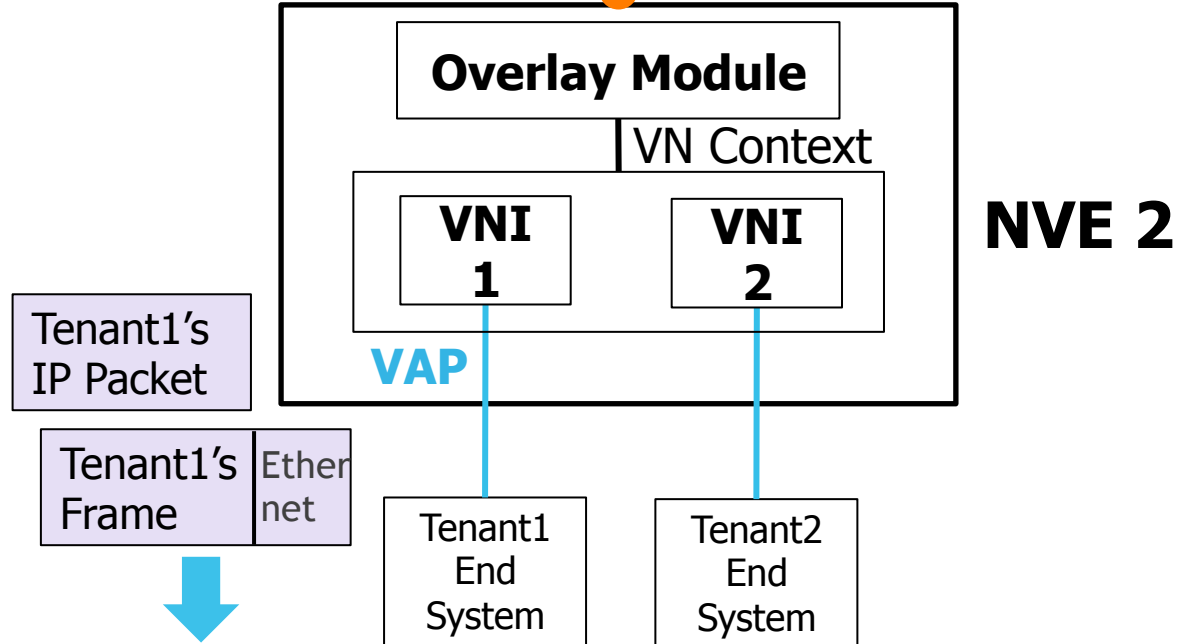
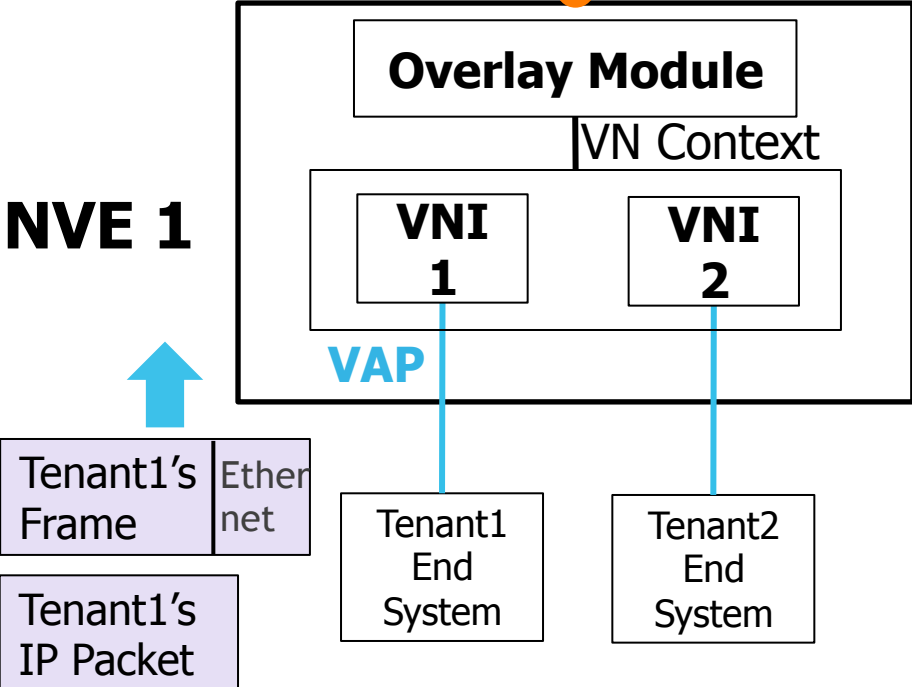
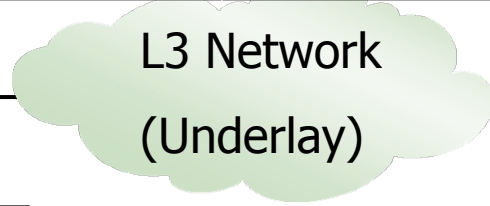
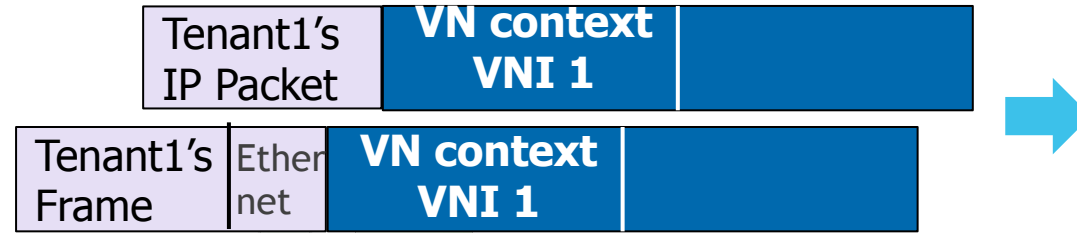
VAP (Virtual Access Point): テナントシステムを接続するためのポート(物理ポート/仮想ポート)

VN Context Identifier: VNの識別子

# NVO3 FRAMEWORK OVERVIEW (3)

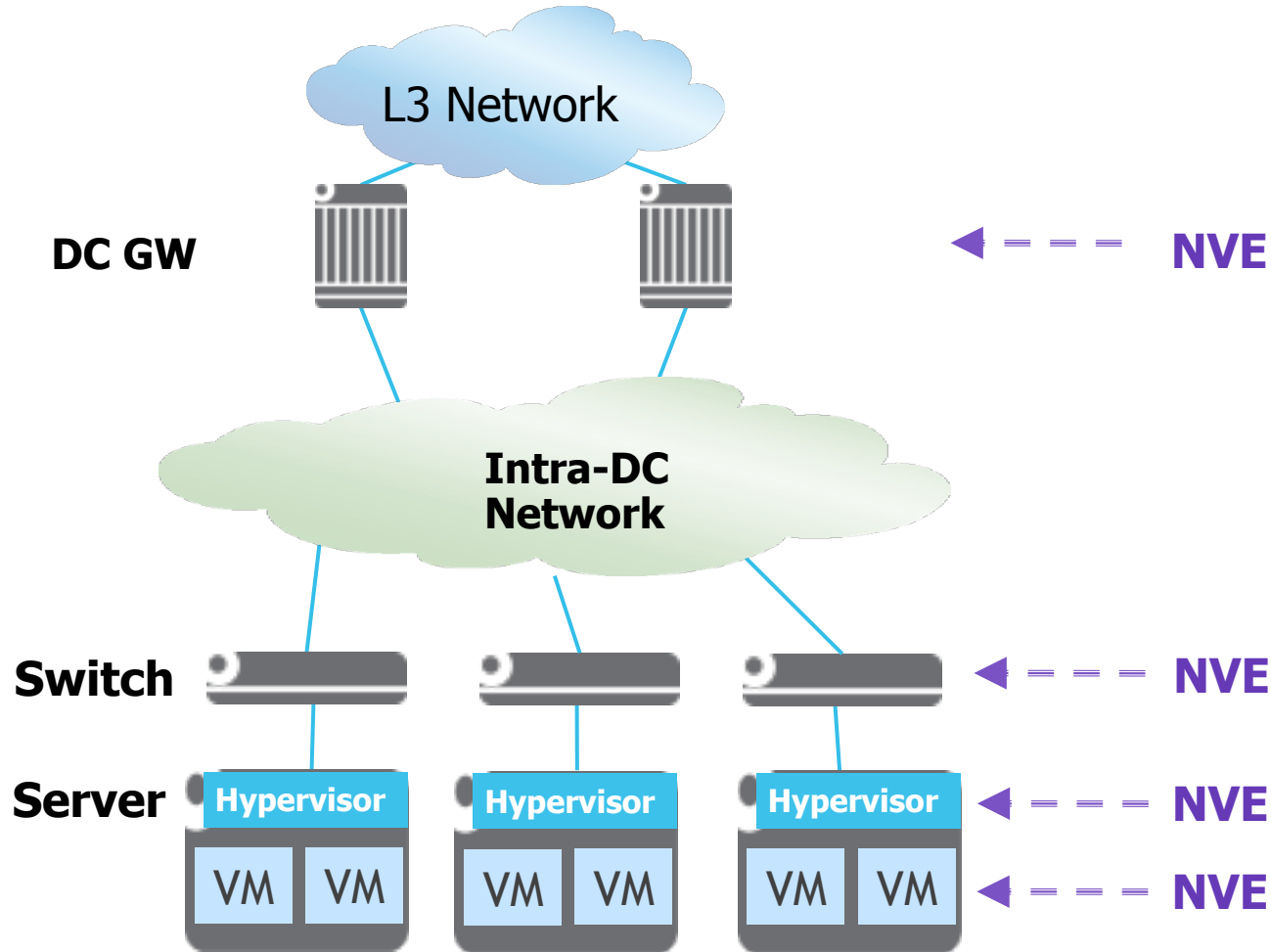
NVE : Network Virtualization Edge

L3 tunnel Header



# NVO3 FRAMEWORK OVERVIEW (4)

## NVE の機能配置



NVEの機能配備レイヤを決める際に考慮すべきこと

- 処理能力とメモリの要求
  - データパス(lookup/filtering/encap/decap)
  - Control plane(routing/signaling/OAM)
- FIB/RIBのサイズ
- Multicastのサポート
  - プロトコル/Replication
- Fragmentation
- QoS
- Resiliency

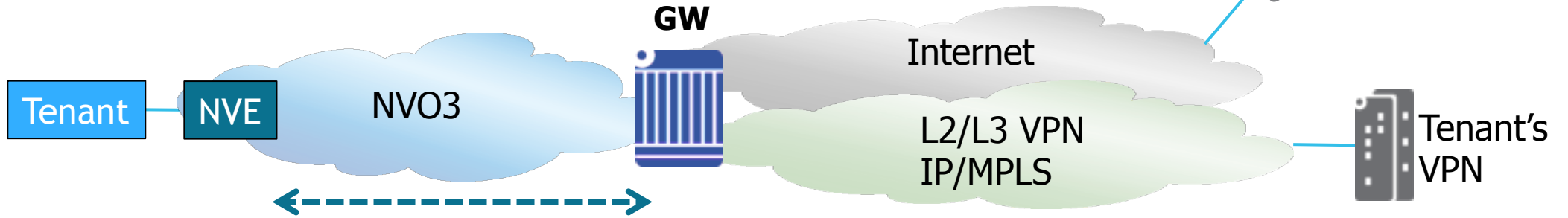


# NVO3 FRAMEWORK OVERVIEW (5)

←---→ NVO3 Tunnel

NVO3で求められる接続性

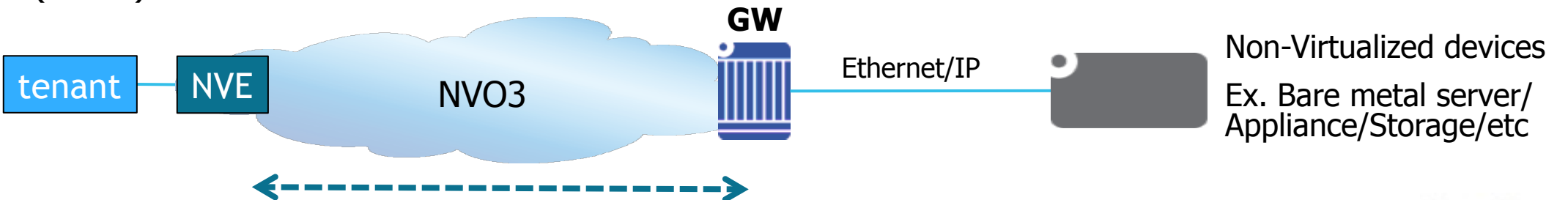
NVO3 <---> VPN/Internet



Inter-(NVO3)-DC

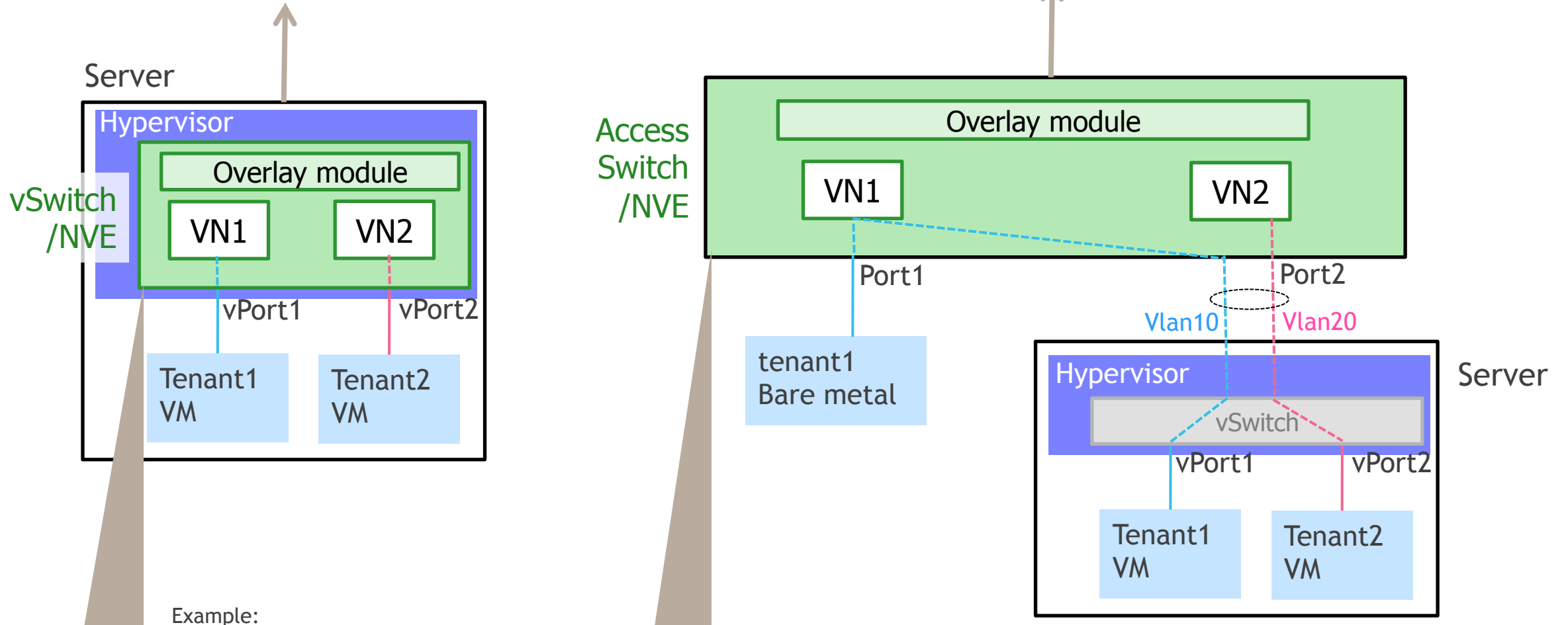


Intra-(NVO3)-DC



# TENANT SYSTEM <-> NVE

## VNEの配備例とVNの識別



Example:

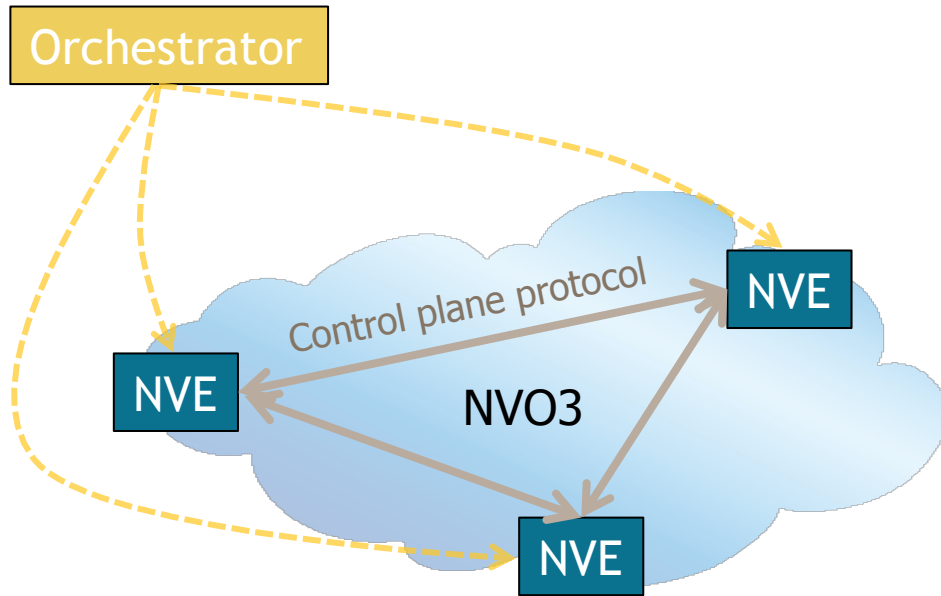
vPort/ Port(and Vlan)	Local Addr	VN Context	Dest Addr	Encap Type	Local- NVE-Addr	Remote- NVE-Addr
--------------------------	------------	------------	-----------	------------	--------------------	---------------------

# コントロールプレーンのREQUIREMENTS

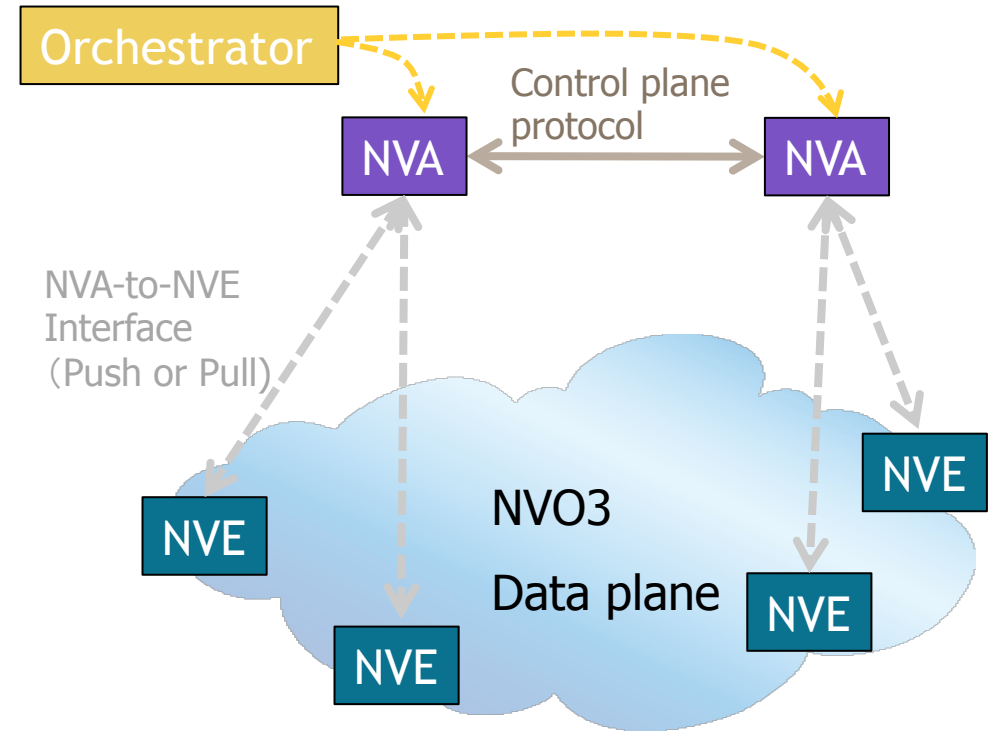
- StoreするStateの量は最小限に
  - 自分で必要なStateのみをしておく
- 必要なStateを迅速に取得出来ること
- 不要なStateを迅速に検知しRemove出来ること
- 処理のOverheadは最小限とすること
- 高いスケール性を持つこと
  - 数十万のNVE, 数百万のVMに対応でき、高スケールでもProtocol overheadが極端に大きくなるならない
- インプリの複雑さを最小限にすること
- 機能拡張が可能であり、旧VersionとのInteroperabilityを持つ
- シンプルなプロトコル設定

# コントロールプレーンの機能配備

NVEにコントロールプレーンを実装

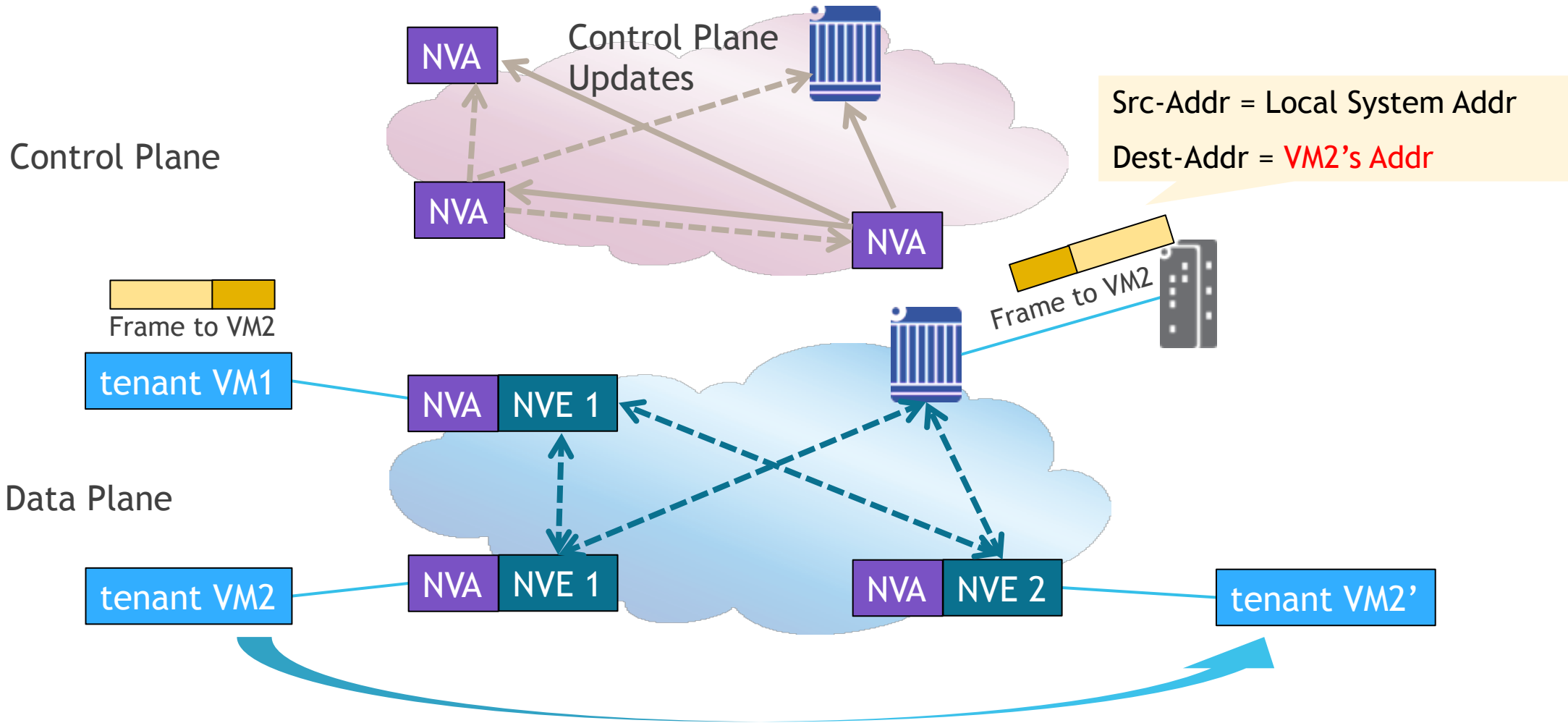


コントロールプレーンを分離



NVA (Network Virtualization Authority):  
ReachabilityやForwarding infoをNVEに提供する  
エンティティ

# VM MOBILITYへの対応



L2: MACアドレスの継続性、L2 Table更新

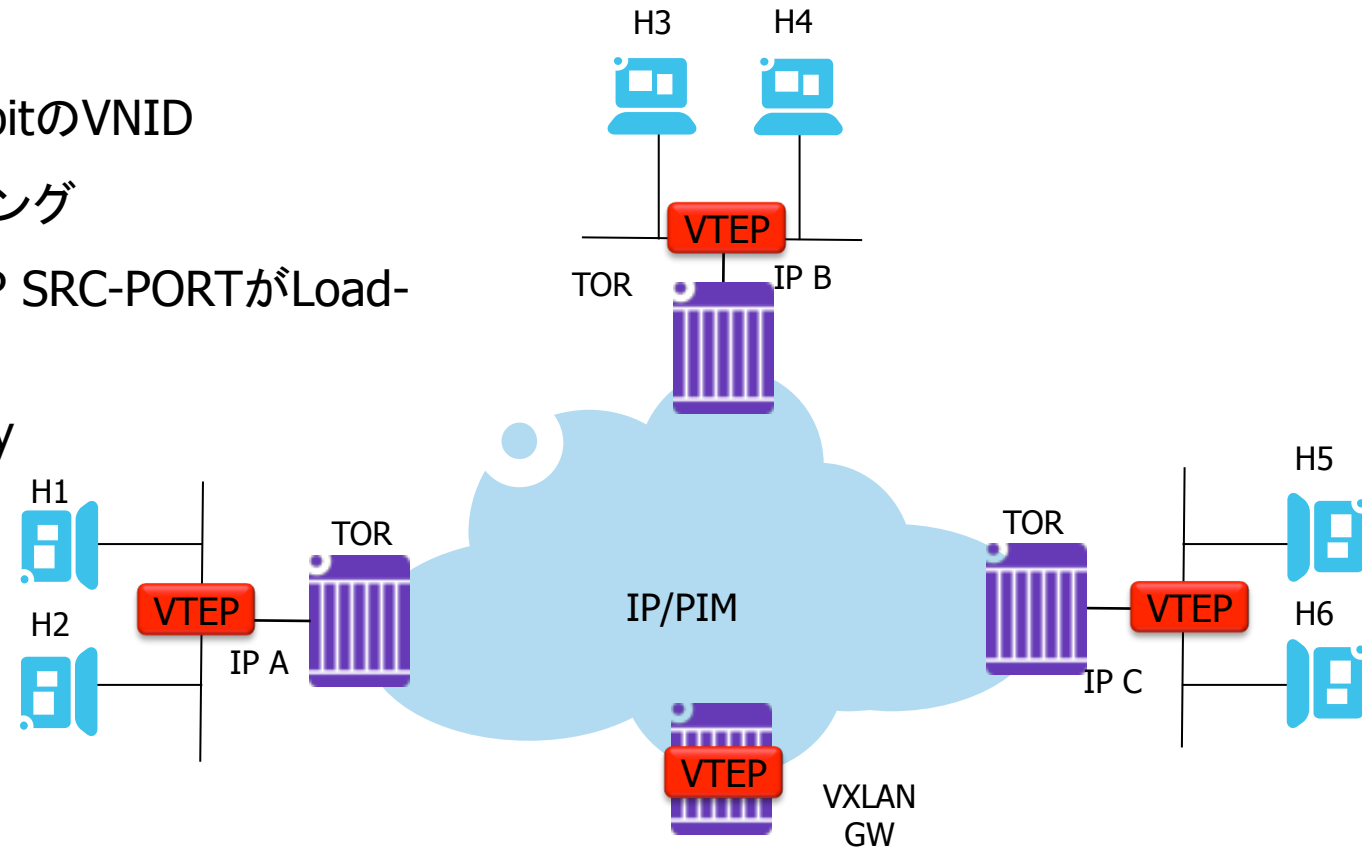
L3: IPアドレスの継続性、L3 Table更新、ARP

# 現在候補としてあがっているTECHNOLOGY

- NVGRE : draft-sridharan-virtualization-nvgre
- VxLAN : draft-mahalingam-dutt-dcops-vxlan
- L2VPN : VPLS : RFC 4761, 4762  
EVPN : draft-ietf-l2vpn-evpn
- L3VPN : VPLS : RFC 4365

# VXLAN: VIRTUAL EXTENSIBLE LAN OVERVIEW

- IP上でOverlayによりLANサービスをEmulate、24-bitのVNID
- データプレーン/コントロールプレーンのMACラーニング
- Ethernet over UDP over IPのEncapsulation、UDP SRC-PORTがLoad-balance時のhashingに使用可能
- BGP-EVPN又はSDN ControllerでのAuto-discovery
- Multicast(PIM)を使用したMulticastの最適化
- レジリエンシ:
  - コア: ECMP
  - アクセス: Active/Active or Active/Standby
- 通常DC内で使用されるがWAN側へも拡張可能
- Alternative proposed by Microsoft: NVGRE



VTEP: VxLAN Tunnel End point

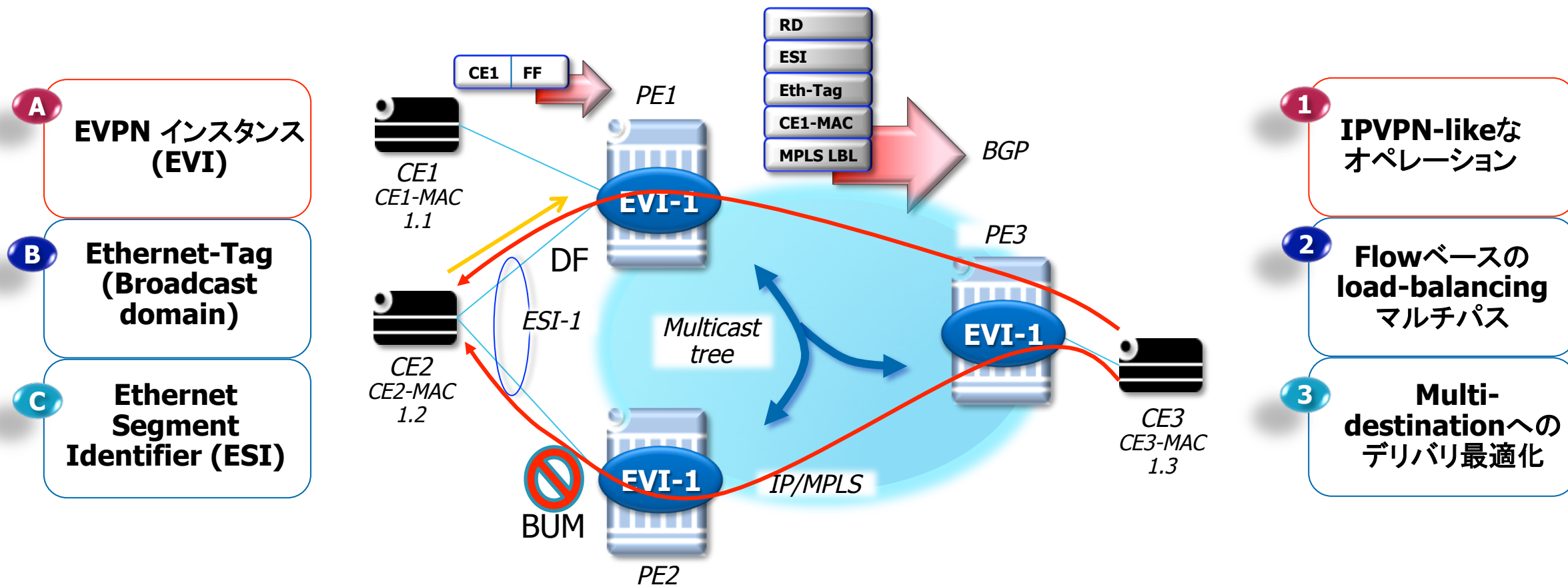
<http://tools.ietf.org/html/draft-mahalingam-dutt-dcops-vxlan-04>

# E-VPN: OVERVIEW

- 広く使用されているEthernet L2VPNサービスで、下記の新たなRequirementsに対応する
  - All-active forwardingでのMulti-homing
  - Multi-destination frameのデリバリの最適化
  - 簡易なService Provisioning
  - IP-VPN likeなオペレーション、IPサービスのサポート
- E-VPNは、これまでのデータプレーンでのMAC学習ではなく、MP-BGPコントロールプレーンでの学習により、これらの新たなRequirementに対応する
- WAN/DC環境どちらにも適用可能
- MPLS 又は IP データプレーン (VXLAN/NVGRE/MPLSoGRE/etc)



# E-VPNの3つのKey conceptとそれにより可能となること



E-VPNによりL2サービスについてもIP-VPN-Likeなオペレーションが可能となり、  
 またFlowベースのマルチパス/LB、複数対向へのデリバリ、高速なコンバージェンスが実現される

# TECHNOLOGY COMPARISON

	PBB-VPLS	E-VPN	PBB-EVPN	VXLAN/NVGRE
Data-Plane	IP or MPLS	MPLS or VXLAN/ NVGRE	MPLS	UDP-IP
Control-Plane	"MP-BGP support"	MP-BGP	MP-BGP	MP-BGP/ISIS/ OSPF/PIM
Multipath	No	Yes	Yes	Yes
Access	Active/Standby	Active/Active	Active/Active	Active/Active
Scale	16M + 16M	16M	16M	16M
Positioning	WAN	DC/WAN	WAN	DC/WAN
Other	L2 only	L2 + L3	L2 only	L2, Multicast dependency

NVO3 WGでは、[draft-gbclt-nvo3-gap-analysis](#)で、これからAnalysisを行っていくところ。

まだこんな→感じ

Supported Approaches	NVGRE	VxLAN	VPLS	EVPN	L3VPN
Data Plane Learning					
- - -	- - -	- - -	- - -	- - -	- - -
Explicit Signaling					

# NVO3 今後のRoadmap

Done - Problem Statement submitted for IESG review

August 2013 - Framework document submitted for IESG review

Dec 2013 - Data plane requirements submitted for IESG review

Feb 2014 - Operational Requirements submitted for IESG review

Feb 2014 - Control plane requirements submitted for IESG review

Mar 2014 - Gap Analysis submitted for IESG review

Apr 2014 - Recharter or close Working Group

# まとめ

## <Overlay方式のAdvantage>

- Unicast Tunneling Stateの管理はエッジ(NVE)が行うため、UnderlayのノードはTenantのStateを気にする必要が無い。(Multicastの配信をUnderlay multicast protocolで行う場合は、Multicast stateを管理する必要あり)
- Tunnelingによりtenant addressをUnderlayから隠すことが可能であり、Underlayではtenant stateを管理する必要が無い
- tenant毎のアドレスのSeparation
- 高スケールへの対応

## <Challenges>

- Overlay networkとUnderlay networkの管理手法
  - Network path, Resource, Performance, etc
- Overlay networkのFirewallやNAT deviceへの対応
- Multicast scalability(underlayでのmulticastを使用する場合)
- Hash-baseのLoad-balanceへの対応

[www.alcatel-lucent.com](http://www.alcatel-lucent.com)