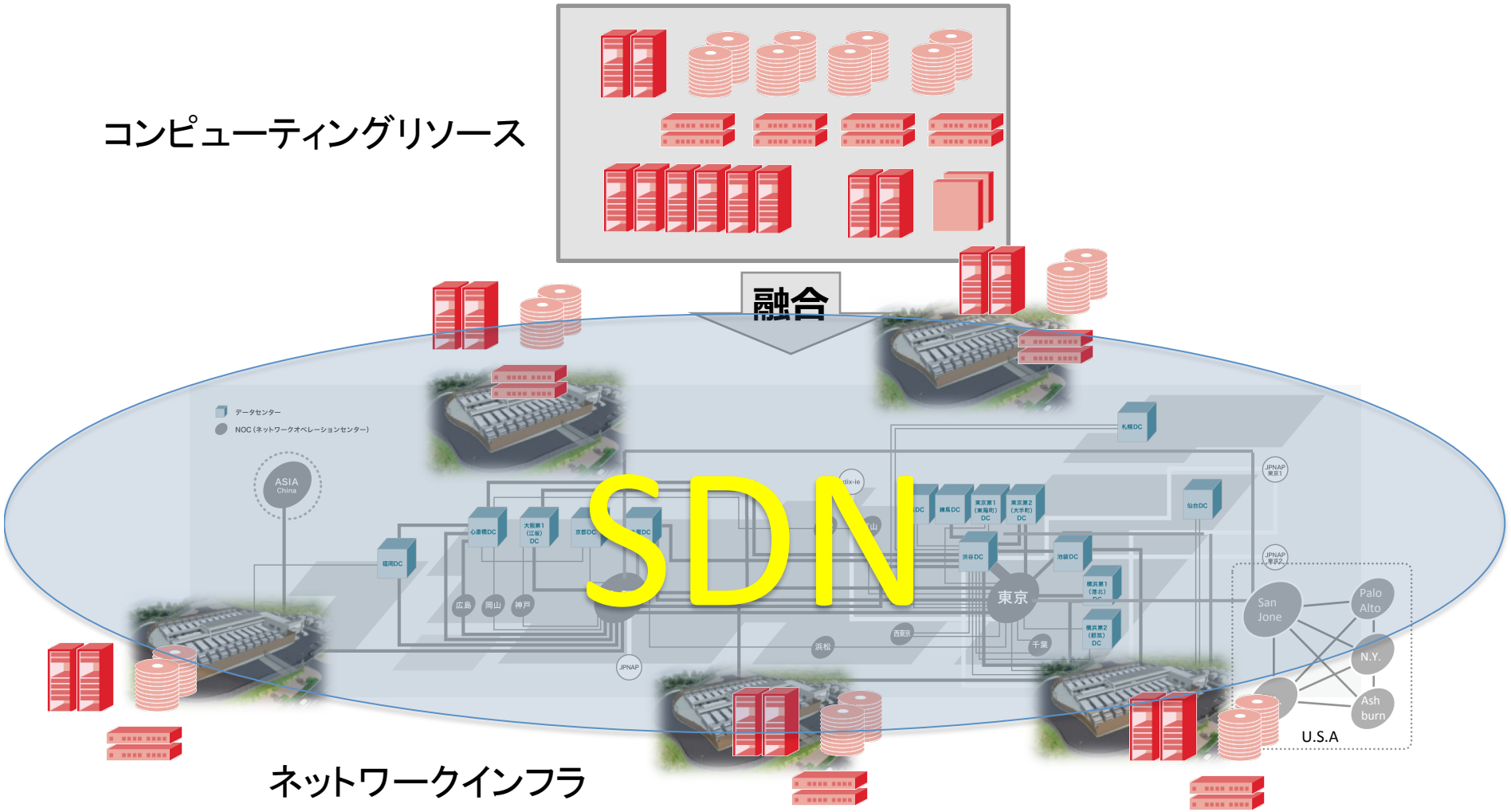
SDNの概要と
ストラトスフィアSDNプラットフォームのご紹介

株式会社ストラトスフィア
2013年5月

クラウド時代のインフラの変化

Stratosphere

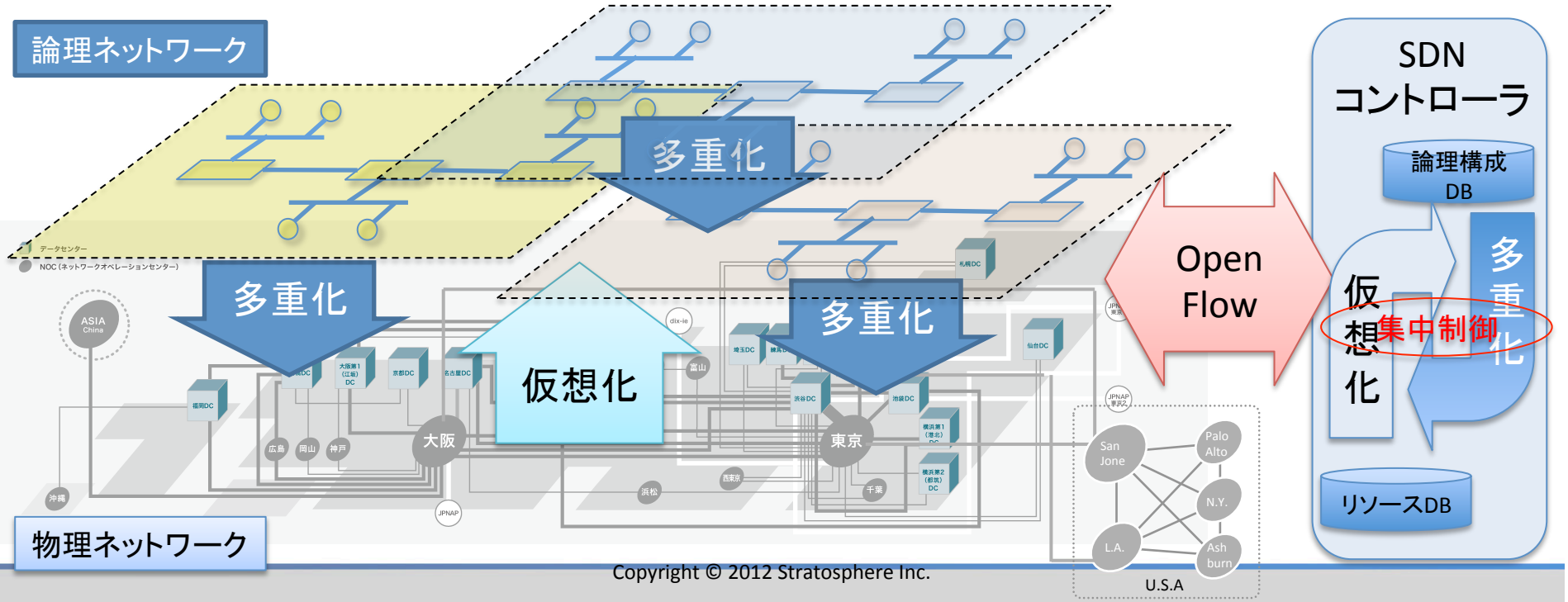


ネットワーク全体に分散されたコンピュータリソースの一体化・全体最適制御のために
ネットワーク仮想化およびSDNが必要

「ネットワーク仮想化」、「SDN」、「OpenFlow」

Stratosphere

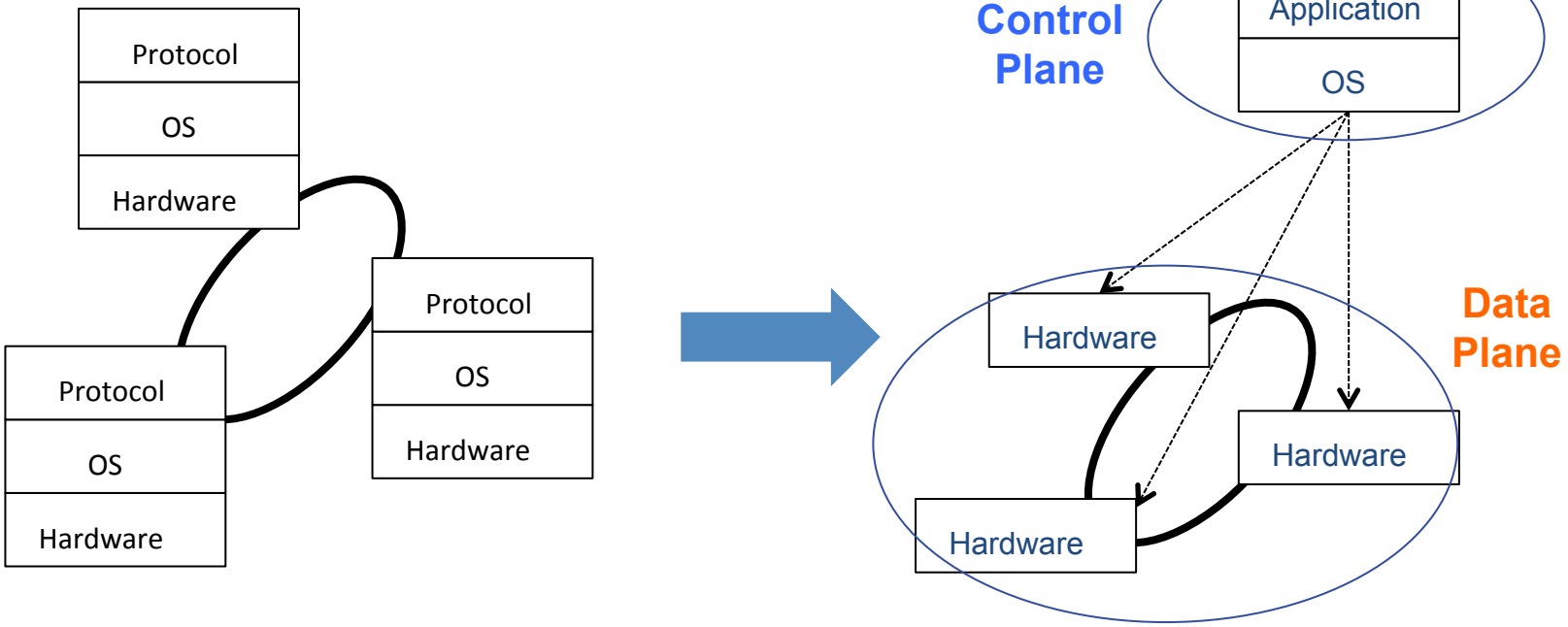
- ネットワーク仮想化とは
  - 一つの物理ネットワーク上で、多数の論理的なネットワークを互いに独立に動作させる技術
- SDN（Software-Defined Networking）とは
  - 物理ネットワークの仮想化、論理ネットワークの多重化、リソース共用管理・構成管理等をソフトェアにより集中制御しよう、というコンセプト
- OpenFlowとは
  - SDNコントローラが、ネットワーク機器の設定や動作を遠隔から制御するために用いる標準化されたプロトコルの一つ



Software Defined Networking (SDN)

Stratosphere

- Network Control Plane を物理トポロジから分離
- SDN自体は単なるコンセプト。特定のプロトコルや技術をさすものではない
  - 要素技術の 1つが OpenFlow

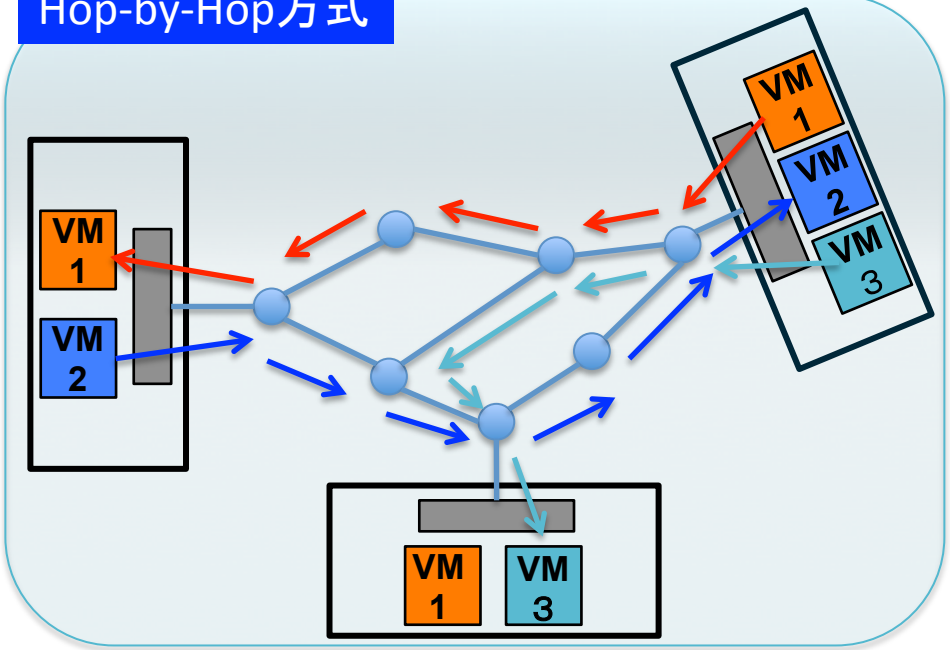


従来　　SDN

ネットワーク仮想化方式

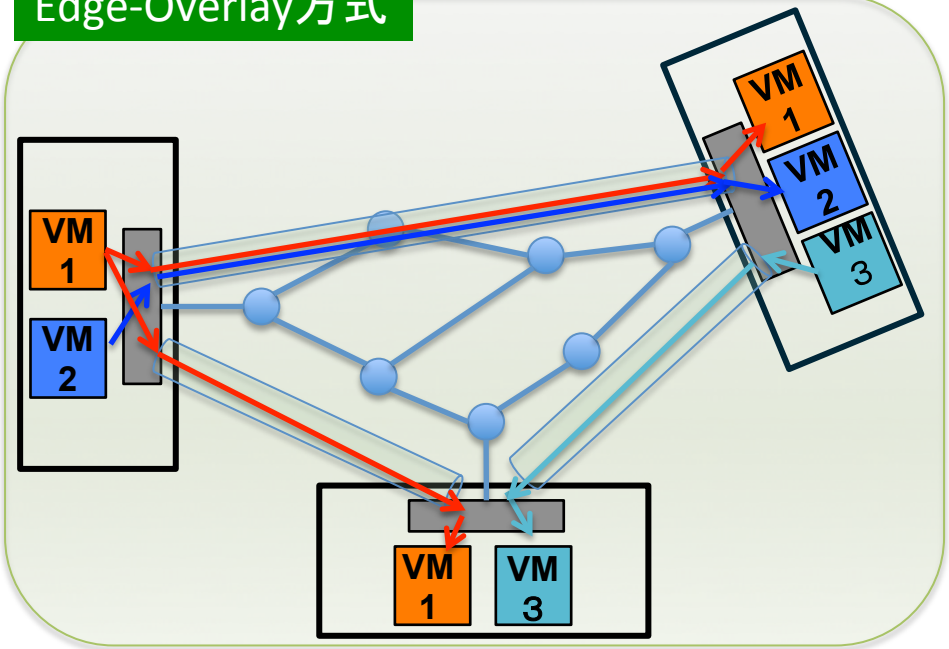
Stratosphere

Hop-by-Hop方式



特徴	ネットワークの各ノードに転送ルールを設定することで、パケットの流れを制御する方式
長所	各リンクに流れるトラフィックのきめ細かな制御が可能
短所	全てのノードに、一貫性のあるルール設定を行う必要があるため設定管理が大変

Edge-Overlay方式

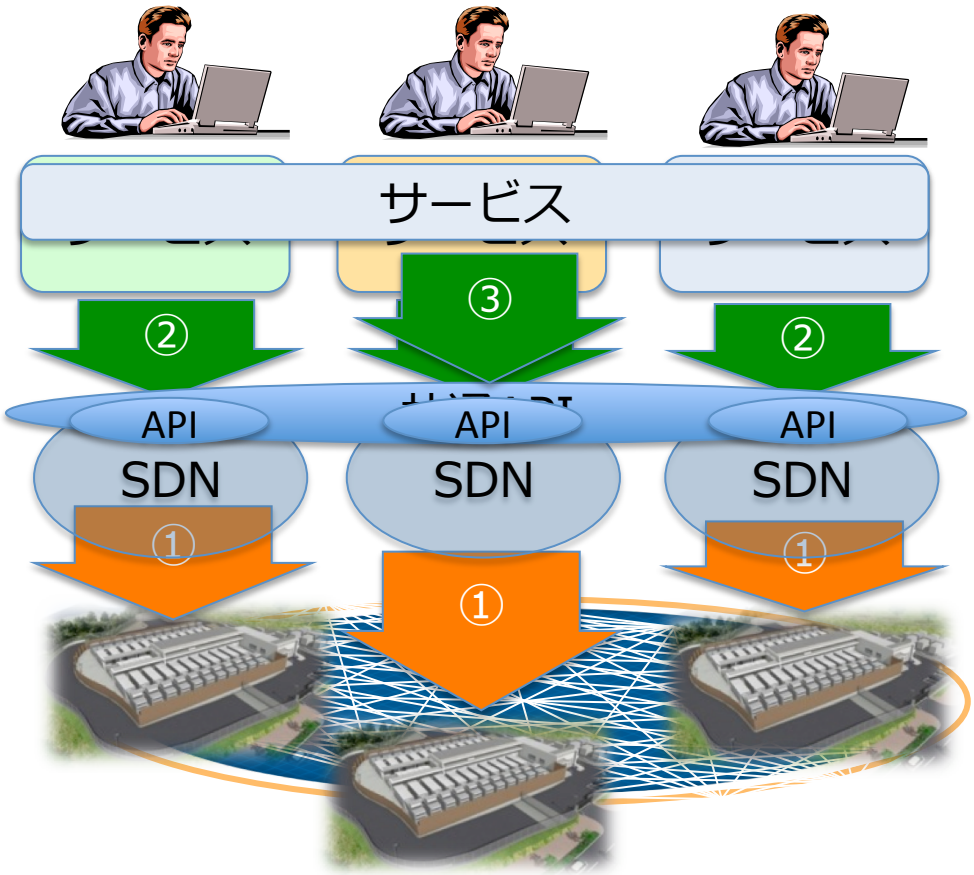


特徴	物理マシンをIPトンネリングで繋ぎ、トンネルを選ぶことで、パケットの転送先を制御する方式
長所	各マシンのテナントとトンネルの対応の制御のみで、柔軟に仮想ネットワークプレーンを構築可能
短所	IPネットワーク上でトンネルの中身を見ることができないため、きめ細かなトラフィック制御が難しい

実現可能な機能が違う・ハイブリッドなソリューションが必要

SDNのもたらす3つのValue

Stratosphere



広域ネットワークインフラ

1. インフラのソフトウェアによる制御
   a. 仮想NWの構成・制御
   b. 個々の機器の設定・制御
   c. 自動化・自律化

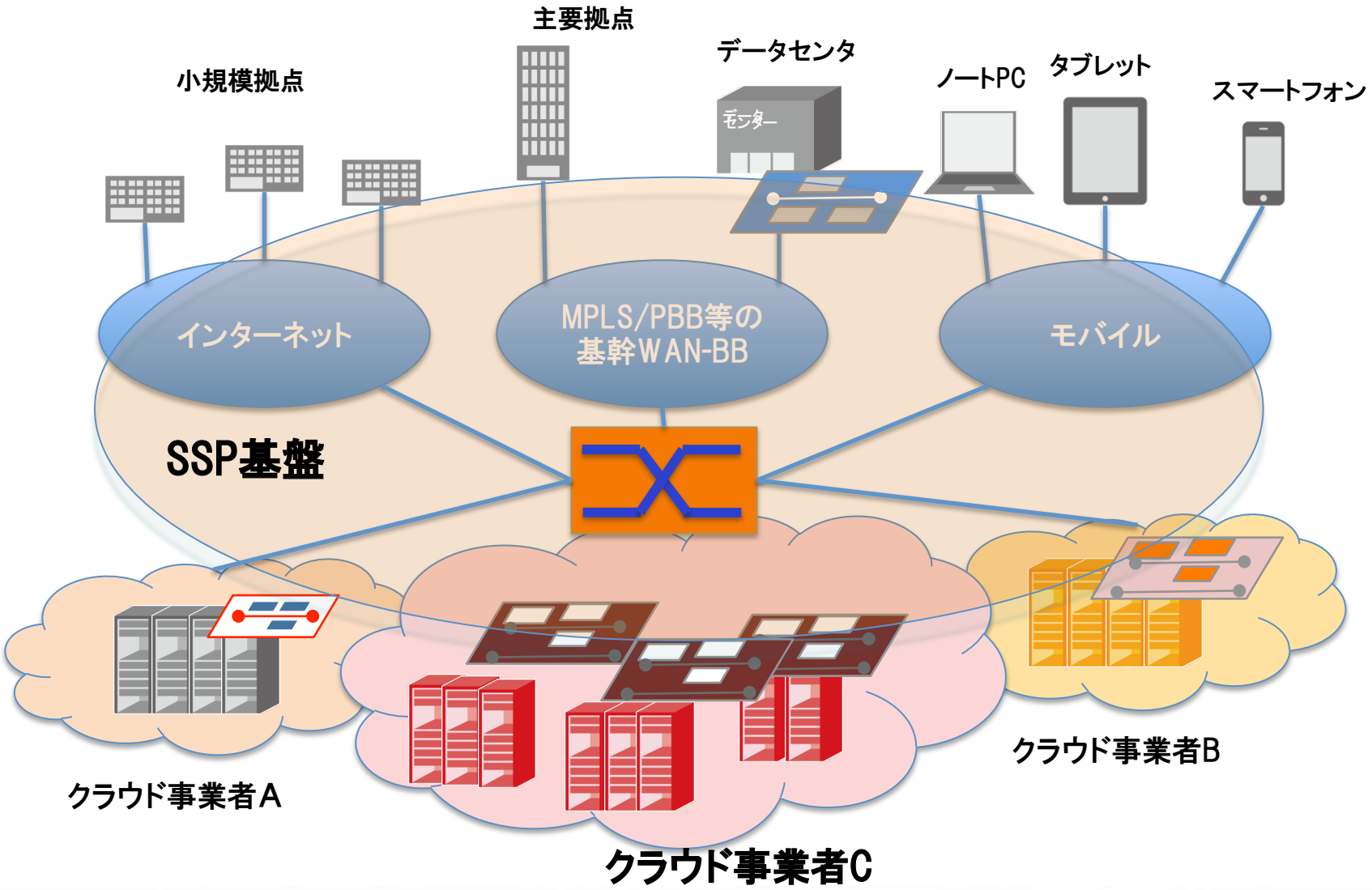
2. インフラのソフトウェアによるユーザへの解放
   a. ネットワークインフラ制御の機能をAPI化
   b. サービス構築をユーザに解放

3. インフラのソフトウェアによる統合・共通プラットフォーム化
   a. 複数インフラでAPIを共通化
   b. 共通APIを用いた広域サービス構築

Stratosphere SDN Platform(SSP)の方向性

Stratosphere

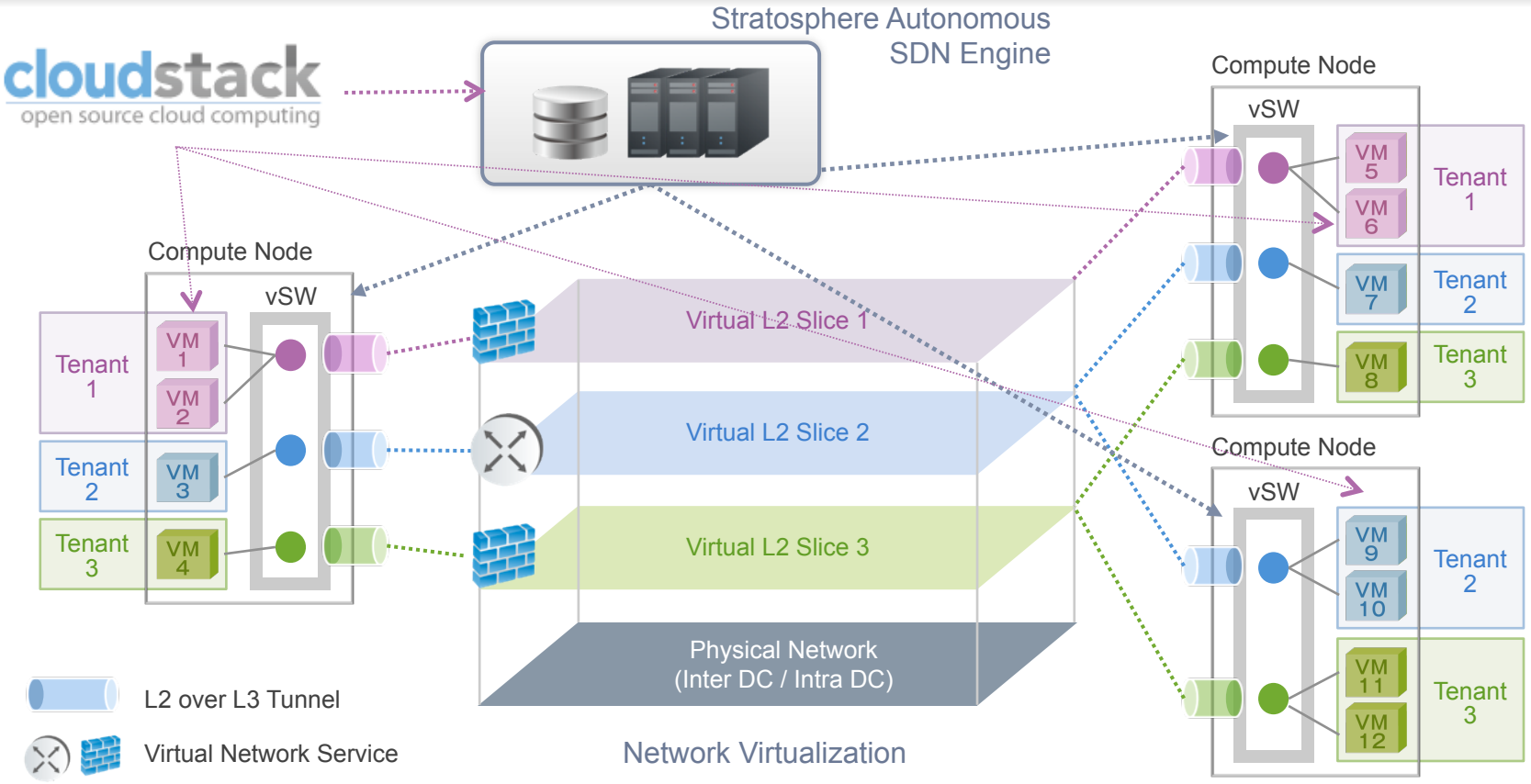
SSPは、マルチテナント、マルチサービスのための基盤提供を目指す(SD-VNO)



Cloud/データセンタソリューション

ストラトスフィア SDN プラットフォーム(SSP)

Stratosphere



■ エッジオーバーレイ型SDN

- Virtual Switch間のIPトンネルでVirtual L2 Sliceを実現
- VXLAN, STT, NVGREの各種オーバーレイプロトコルに対応
- KVM, VMWare, XEN, HyperV等のHyperVisorに対応(予定)

■ SDNエンジンによる仮想ネットワーク環境の統合制御

- 各テナントVirtual L2 Sliceの構成・管理
- 物理ネットワーク上の論理ネットワークの展開、稼働監視
- vRouter, vFirewall等の多様なネットワーク構成機能の組み込み

■ cloudstackと連動

トンネリングプロトコルの種別

Stratosphere

1.0で対応済み

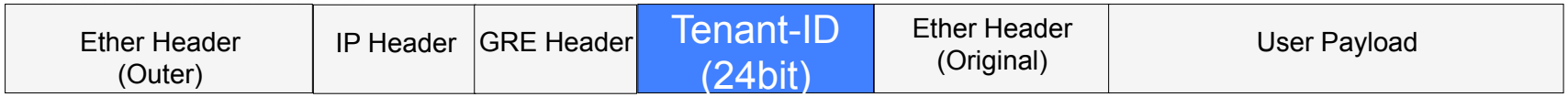
VXLAN

- VXLAN-IDを用い1600万テナントを一意に識別
- UDP ヘッダを用いたコア網内での負荷分散
- 実装例が増えており勢いのあるカプセル化方式

1.2以降

NVGRE

- Tenant-IDを用い1600万テナントを一意に識別
- 従来からあるGREを拡張したシンプルな実装
- Open vSwitch との親和性



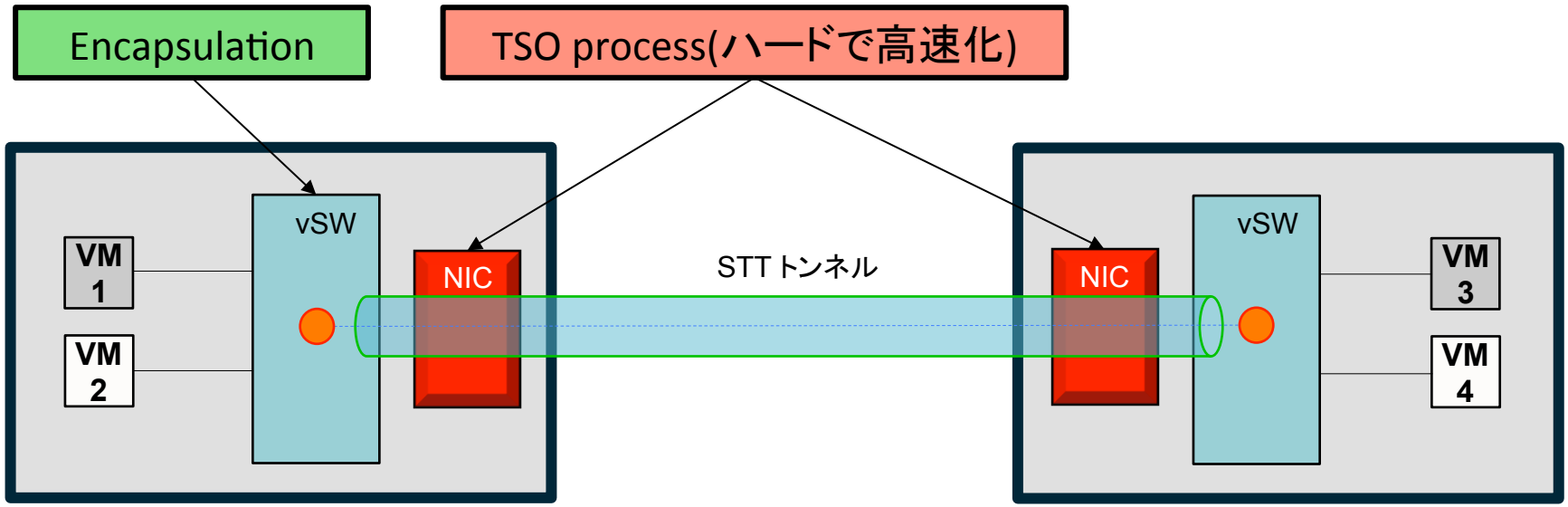
ハードウエア・アクセラレイティド・トンネル

Stratosphere

STT

もう一つのトンネリング技術…

- TCP encapsulation technology である。
- VXLAN/NVGRE の有する24bit-ID に対して、より大きなID (Context ID / (64bit) と呼称) を有し、様々なVirtual Network に関する属性情報を盛り込むことができる。
- NICの機能の一つであるTSOをレバレッジし、computing nodeのTCP segmentation によるCPUへの負荷を軽減することができる。

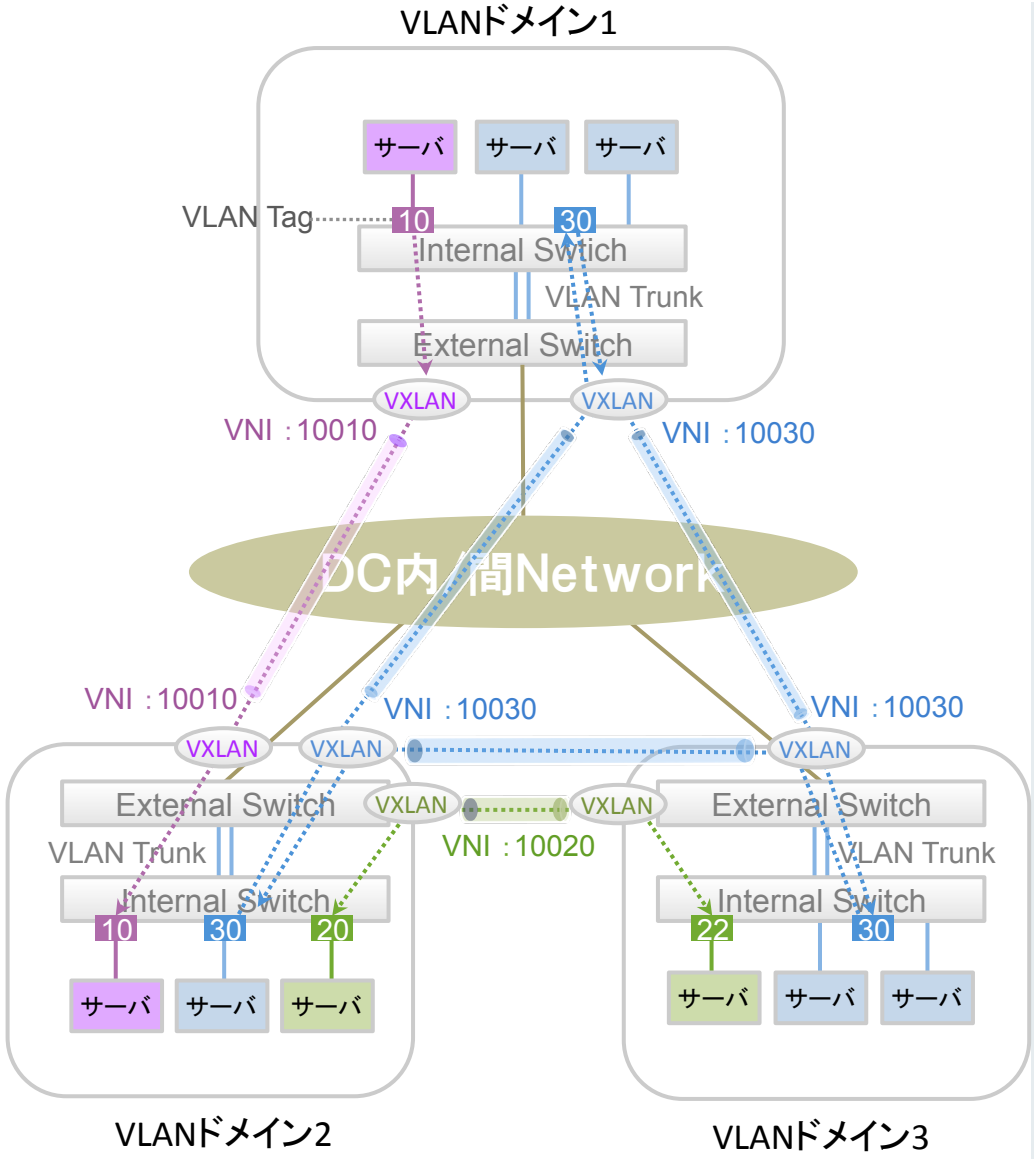


STT Reference Model

A Stateless Transport Tunneling Protocol for Network Virtualization (STT) draft-davie-stt-01

オーバレイネットワークアーキテクチャ

Stratosphere



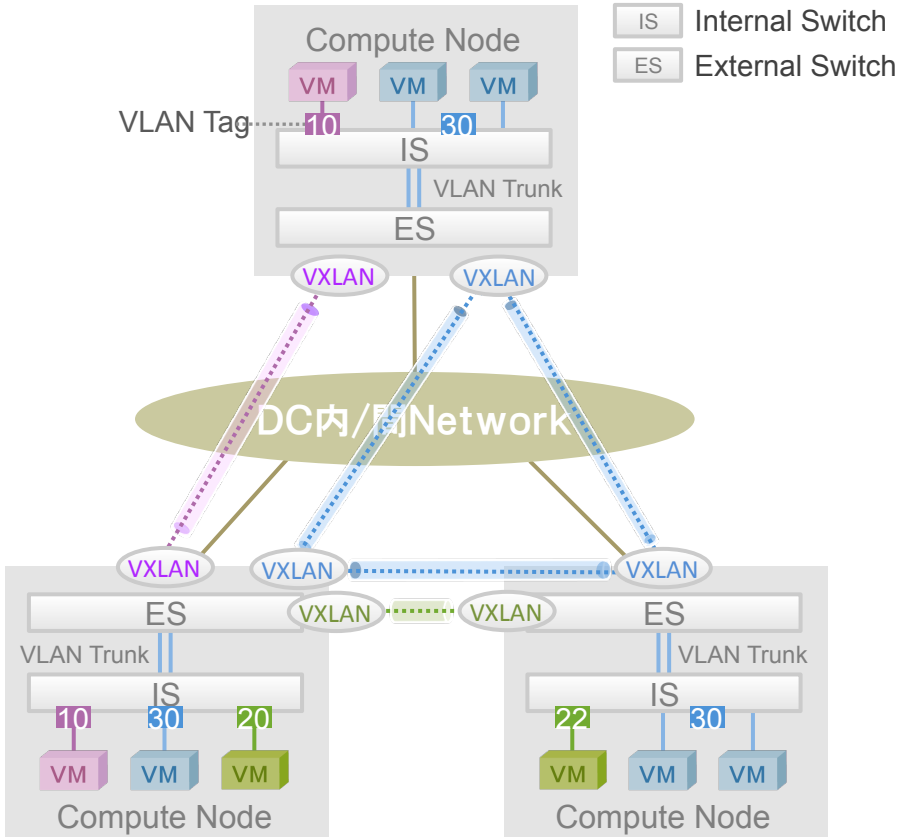
様々なネットワーク技術を活用したハイブッリッドアーキテクチャ

- VLANの活用
  - VLANドメイン内では、Internal Switchがサーバにユニークな VLANタグを付与しテナントを識別
  - VLANドメインは、1台の仮想化されたマシン内や1ラック内、1クラスタ内など、柔軟に構成可能
- 新しいオーバレイプロトコルで仮想NWを構成
  - External Switchで、VLANドメインを他のVLANドメインとIPオーバレイ方式で接続
  - VLANタグをVNI等のトンネルタグに変換することでID空間を拡張(VXLANの場合約1600万通りの仮想LANをサポート)
  - DC内/間ネットワークで、ホスト間のIPリーチャビリティが確保できさえすれば利用可能
- OpenFlowでの配送制御
  - External Switchでのパケット配送制御にFlow Ruleを活用
  - プロアクティブ制御でオーバヘッドを最小化
- 多様なL2プロトコルでDC内／間NWを制御(予定)
  - External Switchで物理ネットワークレイヤを制御
  - RSVP-TE/LDP、802.1ad/802.1ah、LLDP/TRILL

柔軟なモジュラー型システム構成

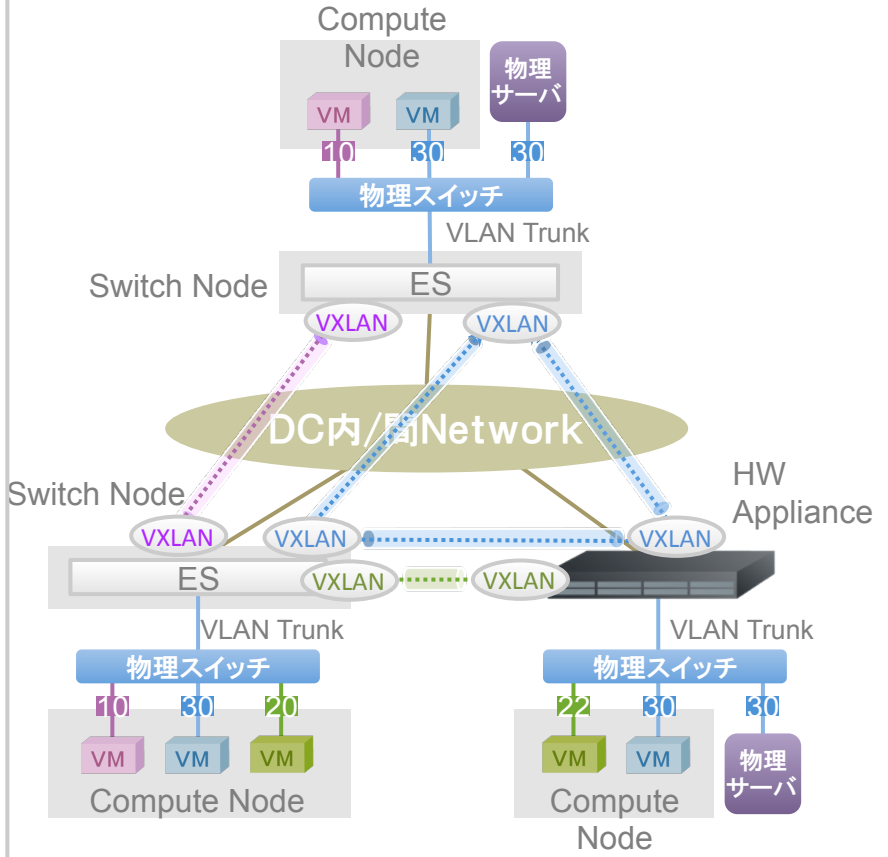
Stratosphere

ビルトインモード



- Compute Node(仮想化サーバ)上で全ての処理を実施
- 追加のネットワーク機器が不要
- 対応Hyper Visor: KVM, ESXi

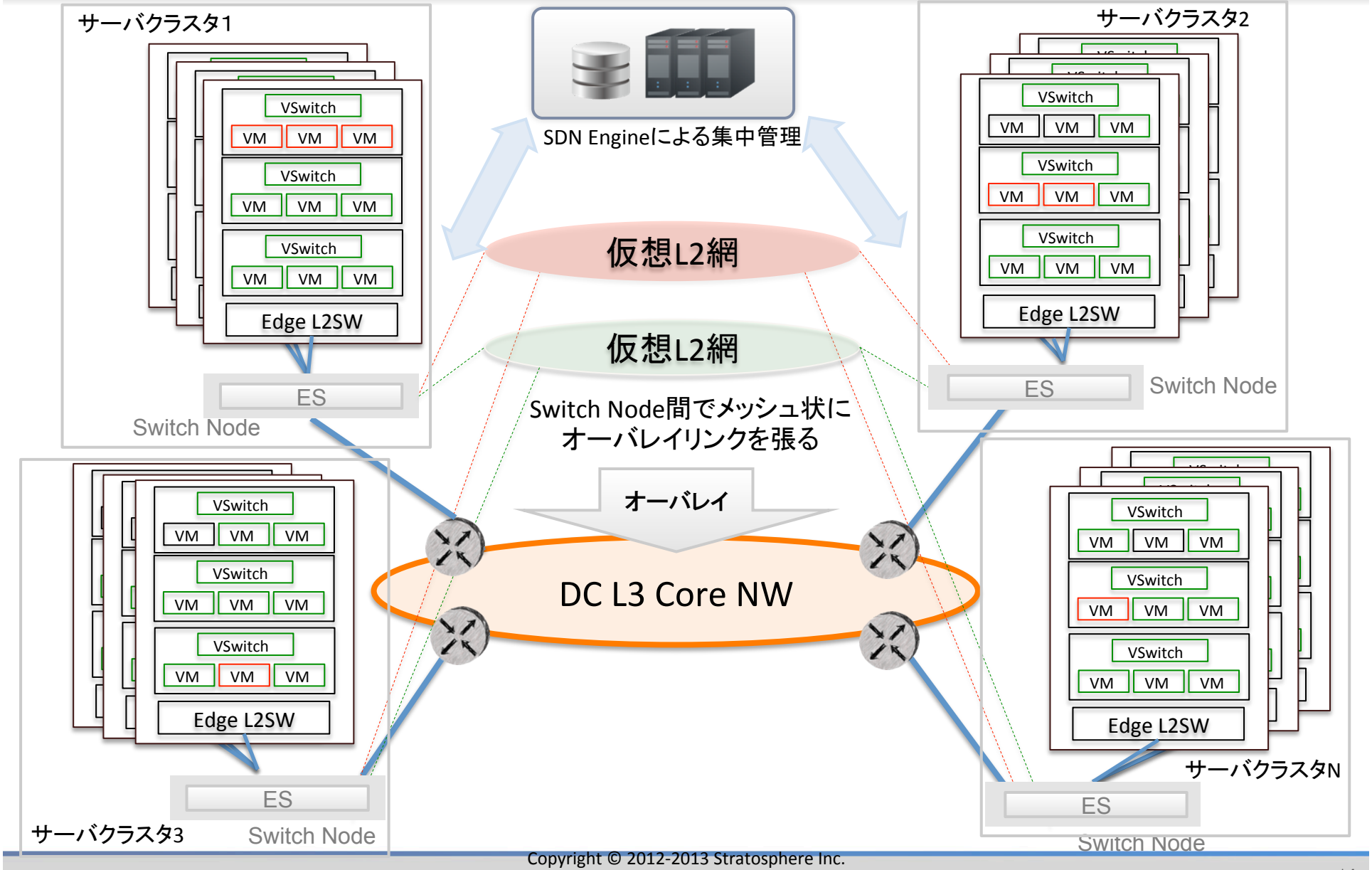
セパレートモード



- Compute NodeからSwitch Nodeを独立させ、スイッチ処理をオフロード
- 物理サーバの接続が可能
- HWアプライアンスによるハードウエア処理への拡張

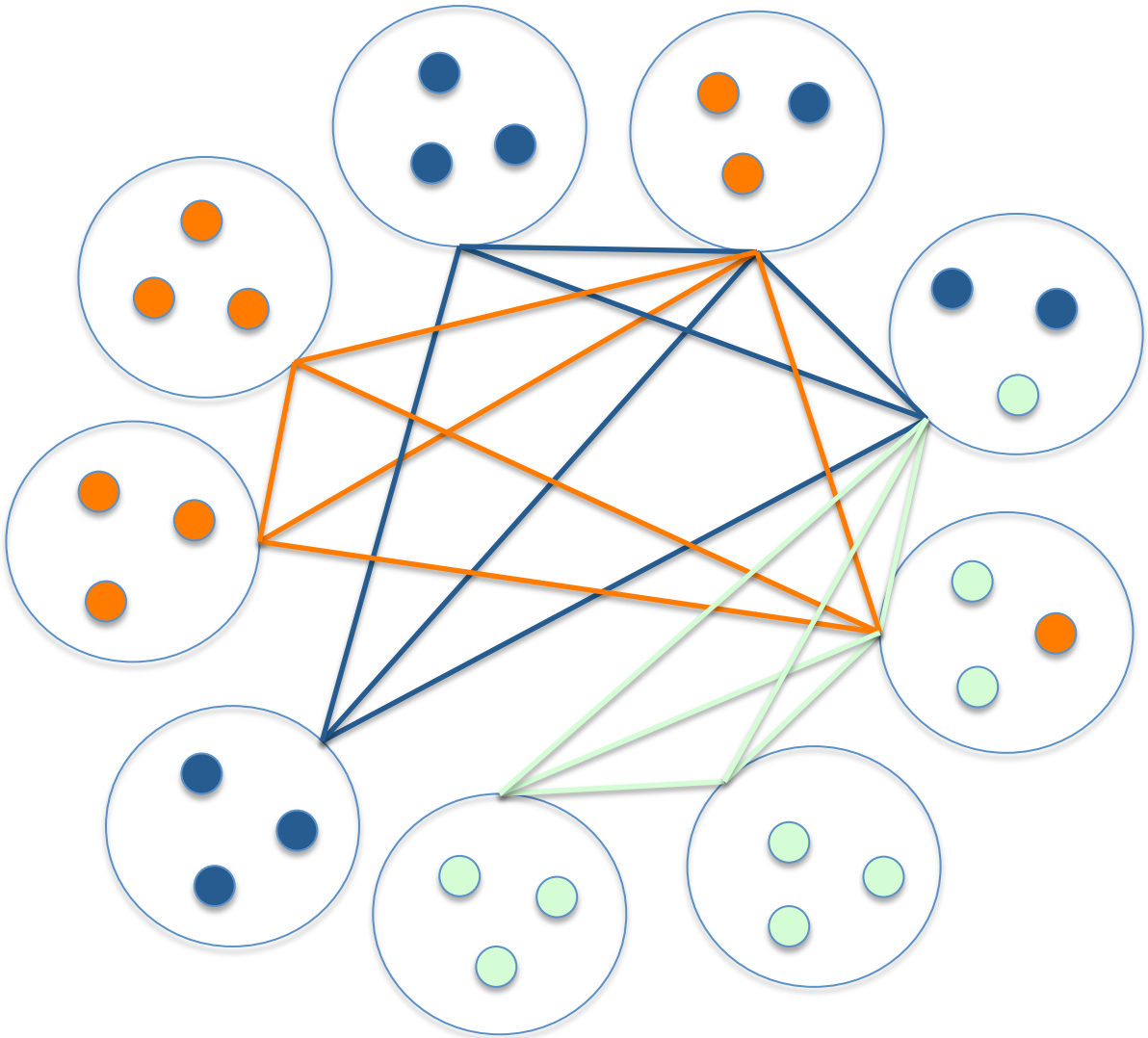
クラウド／DCへの適用（1クラスタ＝1VLANドメイン）

Stratosphere

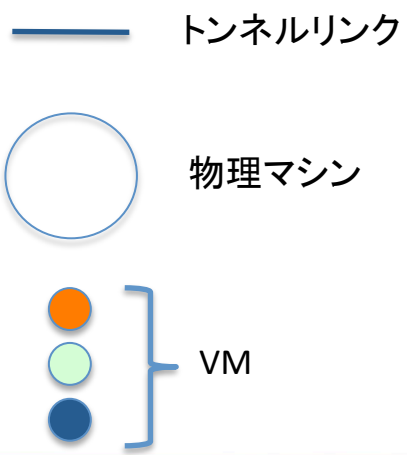


オーバレイメッシュの最適化

Stratosphere

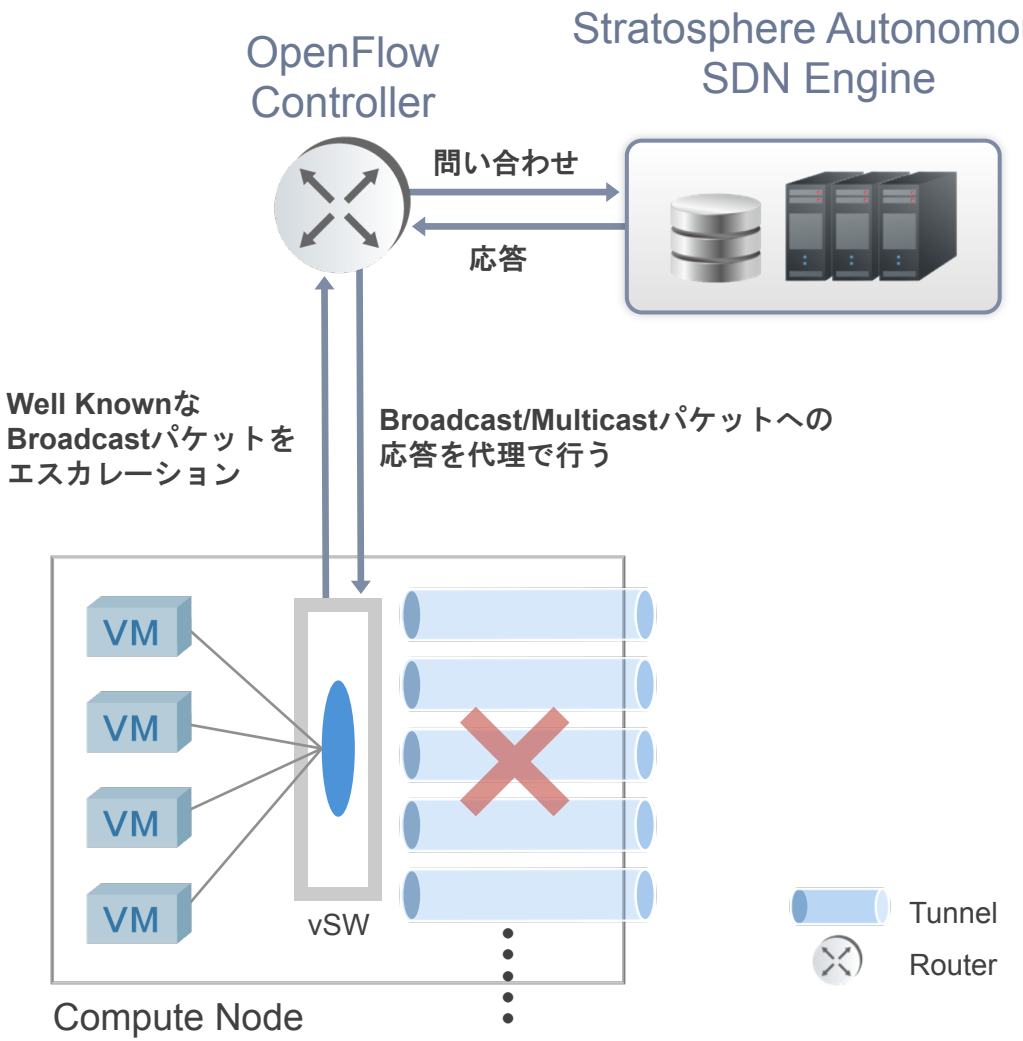


- 同一テナントのVMを収容する仮想化NodeやSwitch Node間を結ぶパーシャルメッシュ構成
- 図の例では、9ノード、27VM、3テナントで18リンク
- VMの配置とトンネルリンクの設定をコントローラで管理
- マイグレーション時に自動的にトンネルリンクを張り替える



Broadcastオフロード

Stratosphere

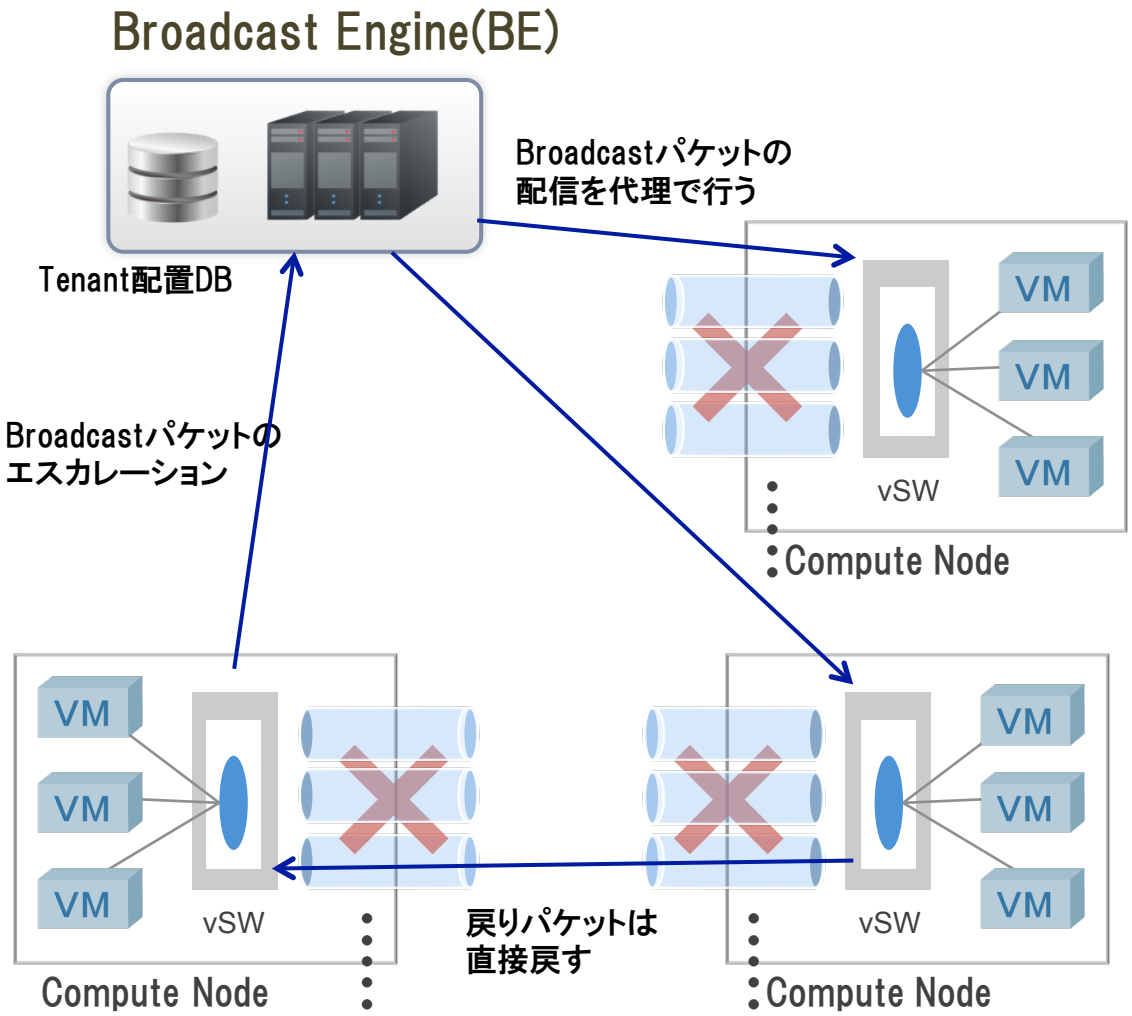


スケーラビリティと ハイパフォーマンスを追求

- ブロードキャスト/マルチキャストを利用するプロトコルは、IPトンネルにFloodingすることなく、コントローラにエスカレーションする
- SDNエンジンが保持する構成管理情報DBから必要な情報を取得し，クライアントに応答を返す

Broadcast Engine

Stratosphere

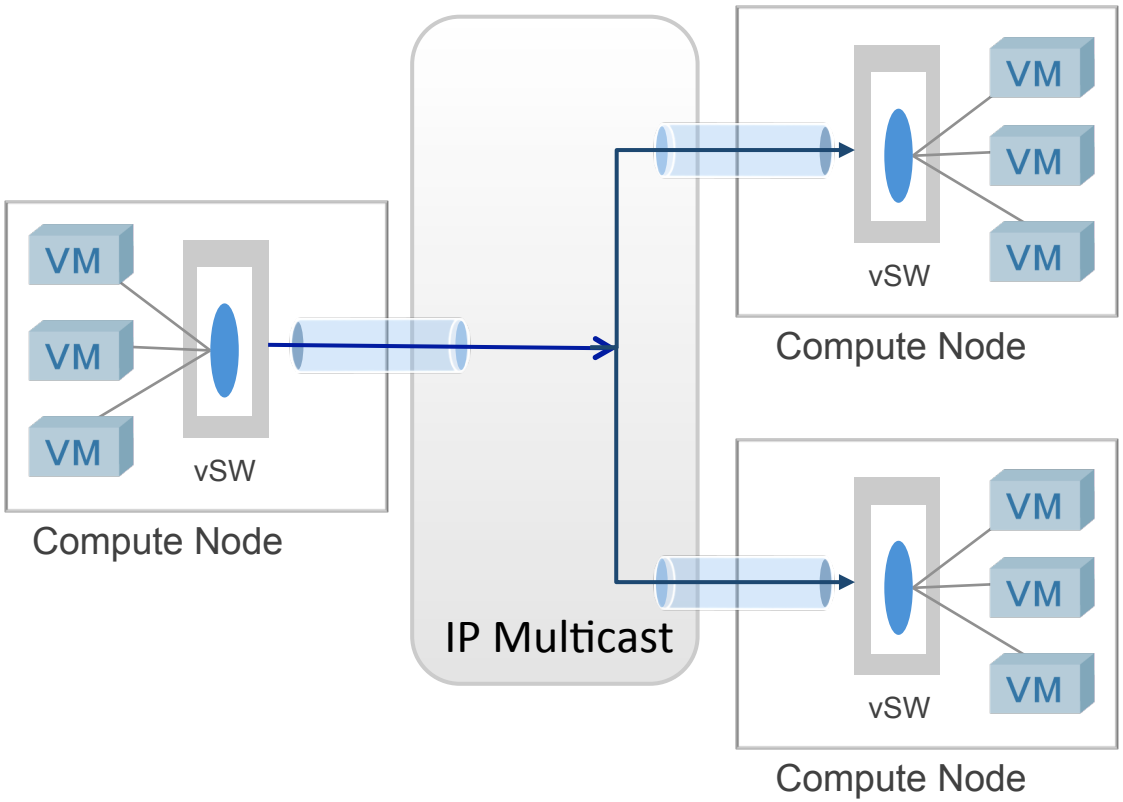


Broadcastパケットの配信専用のサーバ “Broadcast Engine”

- Full MeshからHub&Spokeへ
- Tenant VMの配置DBを持ち、必要なノードにのみパケットをコピーして送信する
- Broadcastパケットの送信元アドレスはオリジナルを保持
- 返信はBE経由ではなく、直接届けられる

VXLAN/IP Multicast

Stratosphere



Broadcastパケットを IP Multicastで配送

- IP Multicastを利用できる物理ネットワーク上で設定可能
- TEP(Tunnel End Point)をMulticast Group Addressに設定し送付
- GroupにjoinしているNodeにのみパケットがコピーされて配信される

WAN/LANへの拡張と
ネットワークサービスの基盤化

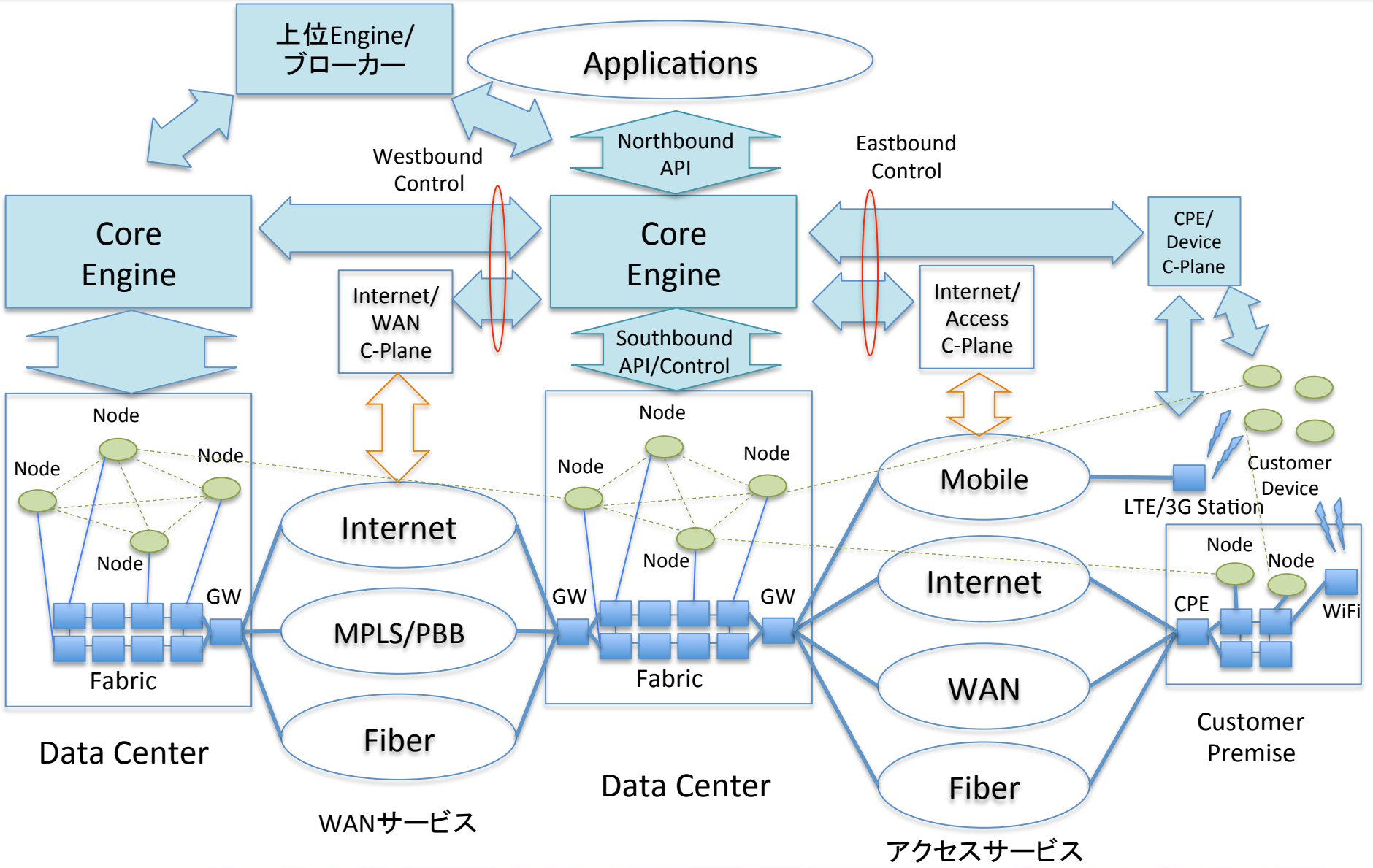
SSPによるネットワークサービスのための基盤の実現

Stratosphere

- Software Defined Managed Ethernet (SDME)
  - 広域かつ任意の地点間にオンデマンドでL2 Sliceを構成し、さまざまなエンティティーをLAN接続する機能を提供する
  - Cloudの基盤で利用する場合は、物理サーバ、Computation Node上のVMインスタンス、ストレージ等が基盤を利用するエンティティとなる
  - さまざまなインフラをまたがるL2 Sliceを構成可能で、インフラが提供する高度なネットワーク処理を利用可能とする
    - WAN, LANさまざまなインフラをまたぐ“スティッチング”
    - 階層化、二重化、プロテクションなど、キャリアインフラに求められる制御の実現
    - 複数回線/ポリシー制御も可能
    - ソフトウェア処理からハードウェア処理
- 上記何時でも利用したい時に利用したいネットワーク機能を速やかに利用出来る機能を『Network as a Service (NaaS)』として定義し、ソフトウェアから制御可能なサービスオペレーション(SD-VNO)基盤を構築する

SSPが対象とするSDNの領域

Stratosphere



広域ネットワークへの拡張

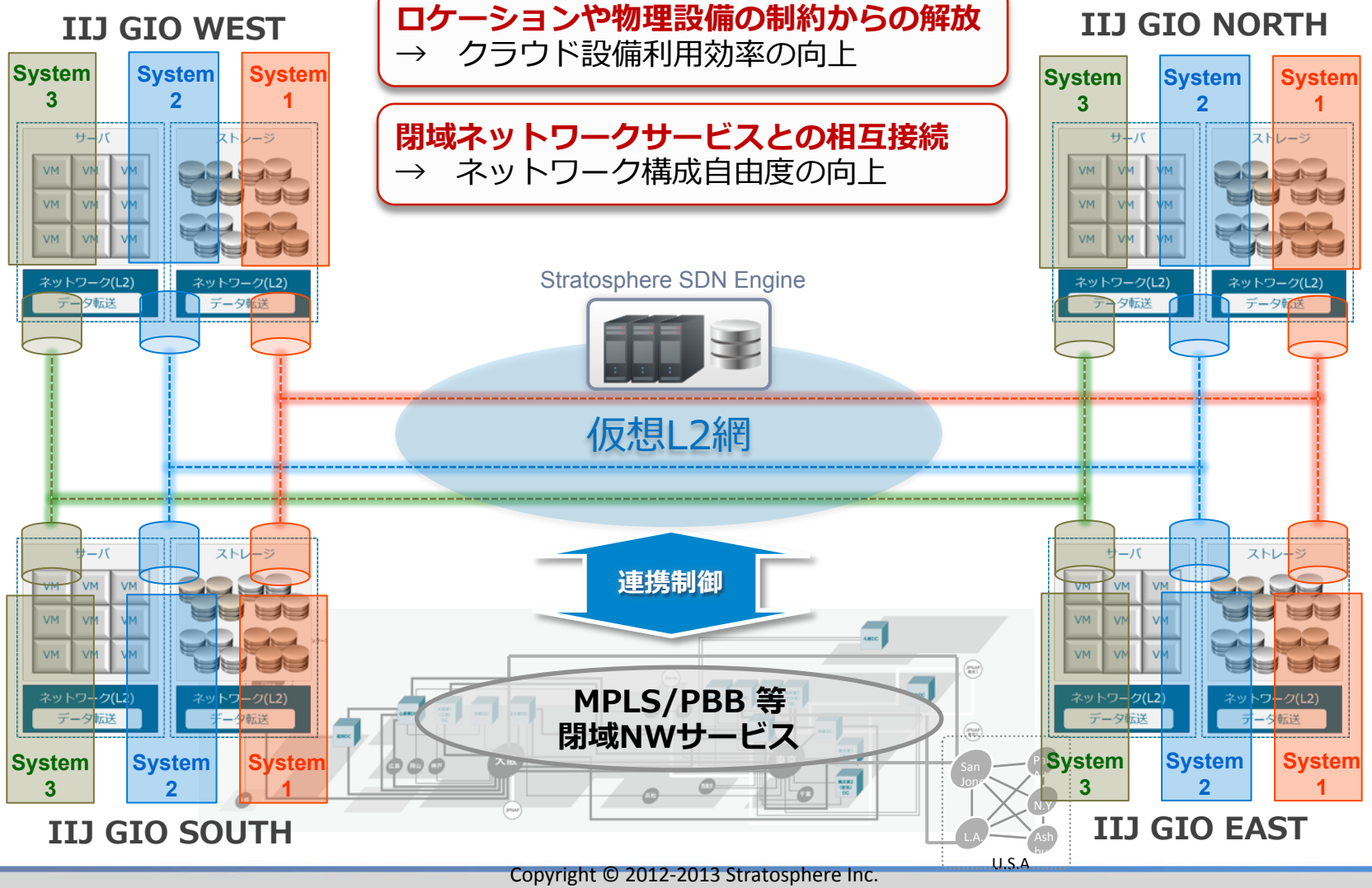
Stratosphere

複数DCをネットワークで繋いだ仮想DCへ

IIJ GIOでのサービス化は現在検討中

ロケーションや物理設備の制約からの解放
→　クラウド設備利用効率の向上

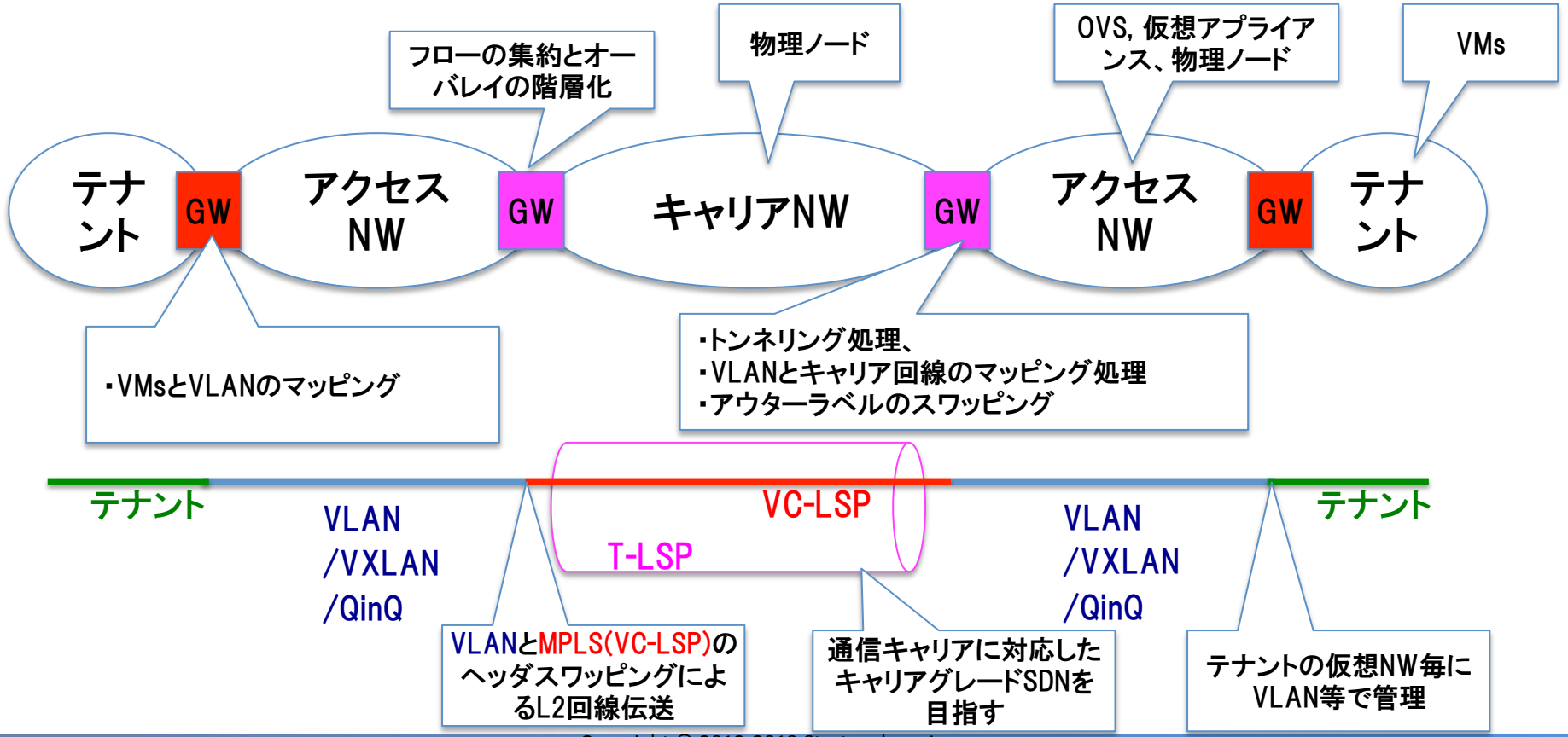
閉域ネットワークサービスとの相互接続
→　ネットワーク構成自由度の向上



仮想ネットワークとネットワークインフラの連携

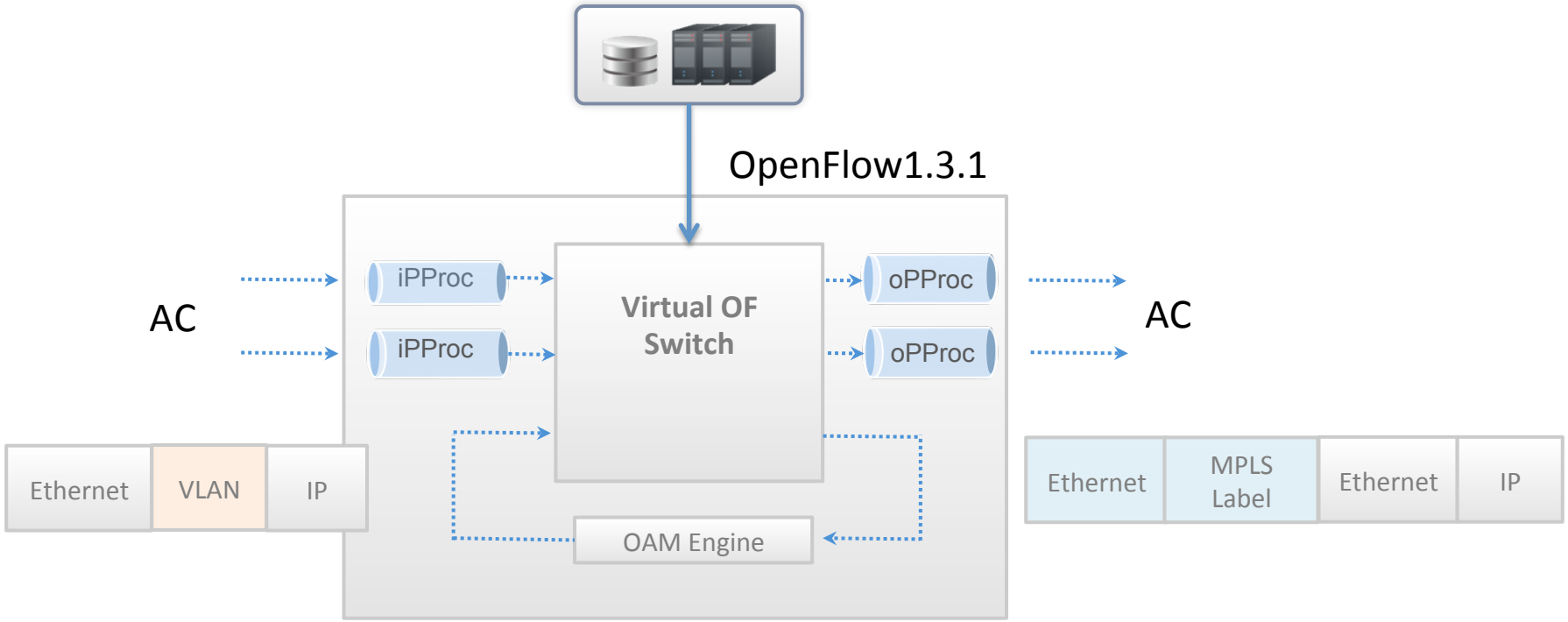
Stratosphere

- ネットワーク事業者向け仮想ネットワーク基盤を実現(SSP1.1から一部提供)
  - VM-VM間(3地点以上も可)に仮想L2 回線(面)を設定する
    - 回線のポリシー毎にグルーピング(VLAN等を活用)
    - 放路制御は、VXLAN等のIPオーバレイやMPLS LSP等で実装可能
  - オーバレイトラフィックの集約と階層化



WANへの拡張 – MPLS Forwarder

Stratosphere

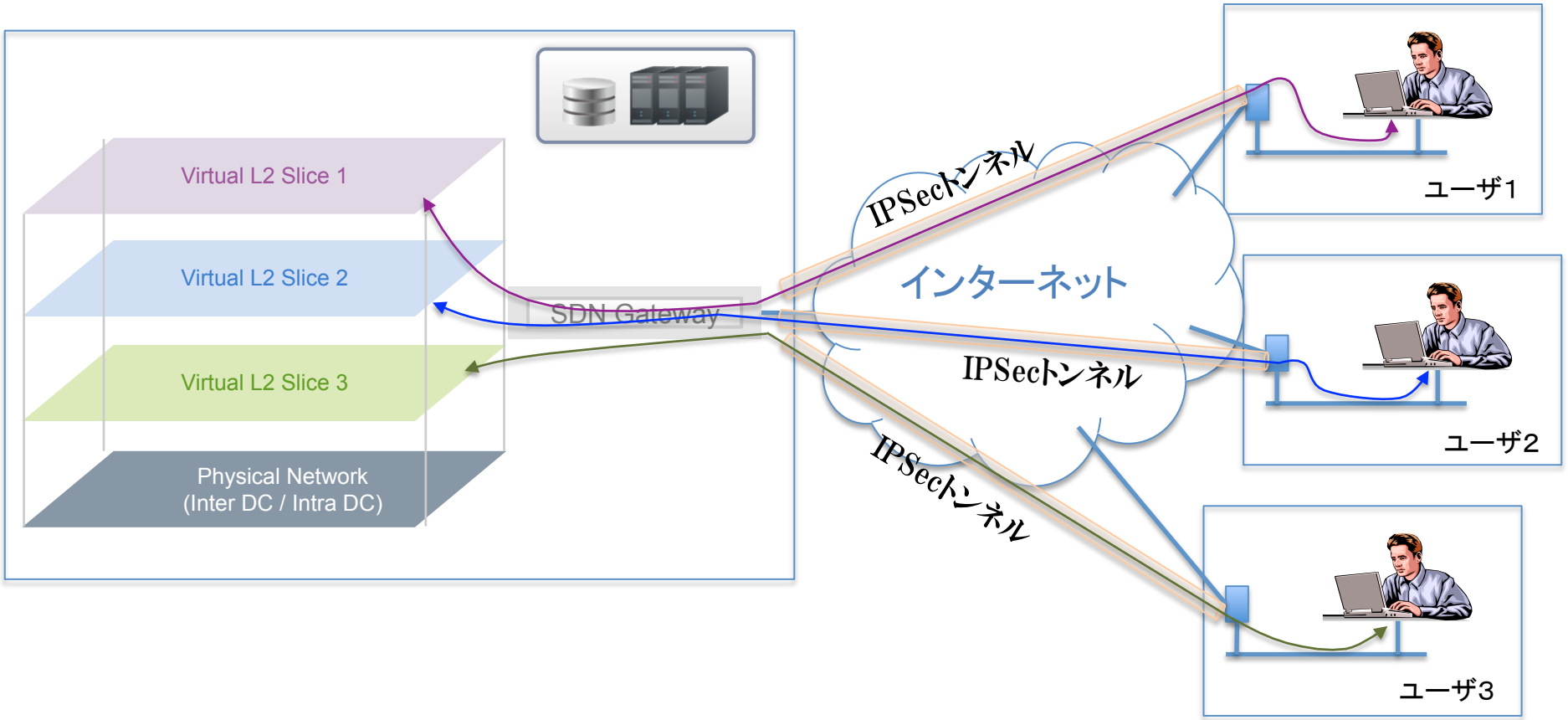


- Open Flow1.3対応により、MPLSラベルをPUSH/POP
- VLANタグをMPLSラベルに変換し、MPLS網上へ転送する
- MPLS OAM

- Draft-medved-pwe3-of-config-01
- iPProc: input Packet Processing Fuction
- oPProc: output Packet Processing Function

LANへの拡張 – SDN Gateway

Stratosphere



- DC側のSDN Gatewayとユーザサイト側のエッジGWとをIPSecトンネルにより接続し、ユーザのLANセグメントを仮想ネットワークと相互接続
- ユーザ側のLANは単一セグメントの転送、または、Tag VLANによる複数セグメントの転送が可能
- IPSecによりインターネット上にセキュアな通信路を設定
- 公開鍵暗号によるユーザサイト認証と暗号化

3層API ＝ Exosphere

Stratosphere

エンドユーザ

Tenant API

- テナントユーザ向け
- テナントネットワークの構築
- Amazon EC2ライク
- +拡張API

サービスプロバイダ

Provider API

- サービスプロバイダ向け
- プロバイダ向けリソースプールのバルク割当
- プロバイダ毎のオーバレイ網構築

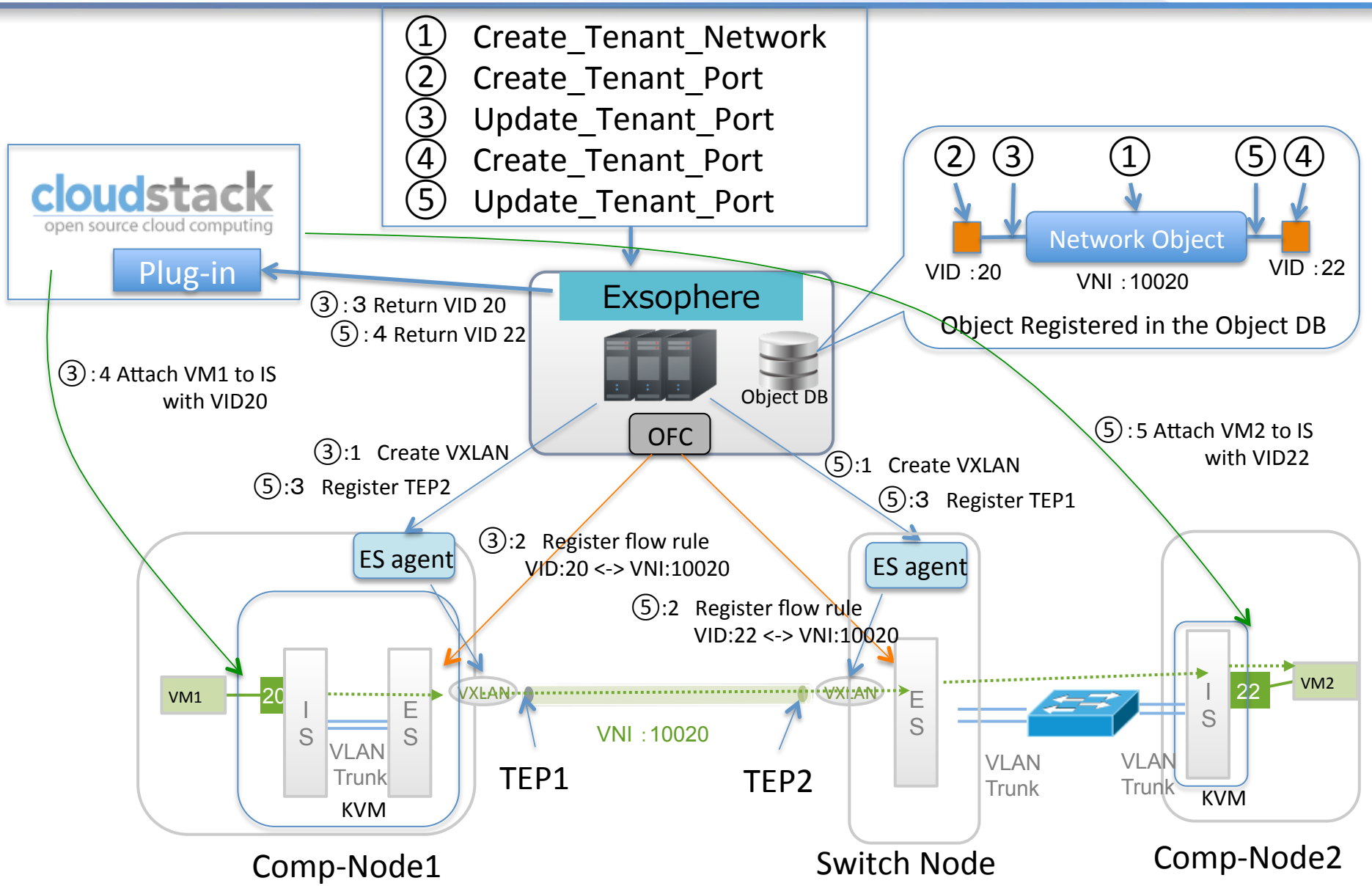
インフラオペレータ

Infrastructure API

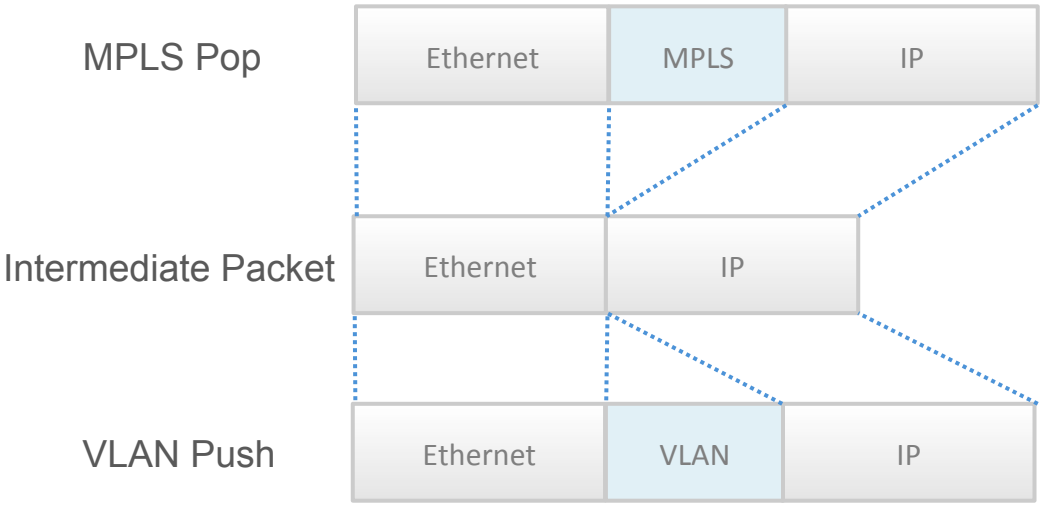
- クラウド/DC事業者向け
- DCネットワークの構築
- 物理ホスト／ネットワーク設定

APIを用いたオペレーション例1: SSP 基本設定

Stratosphere

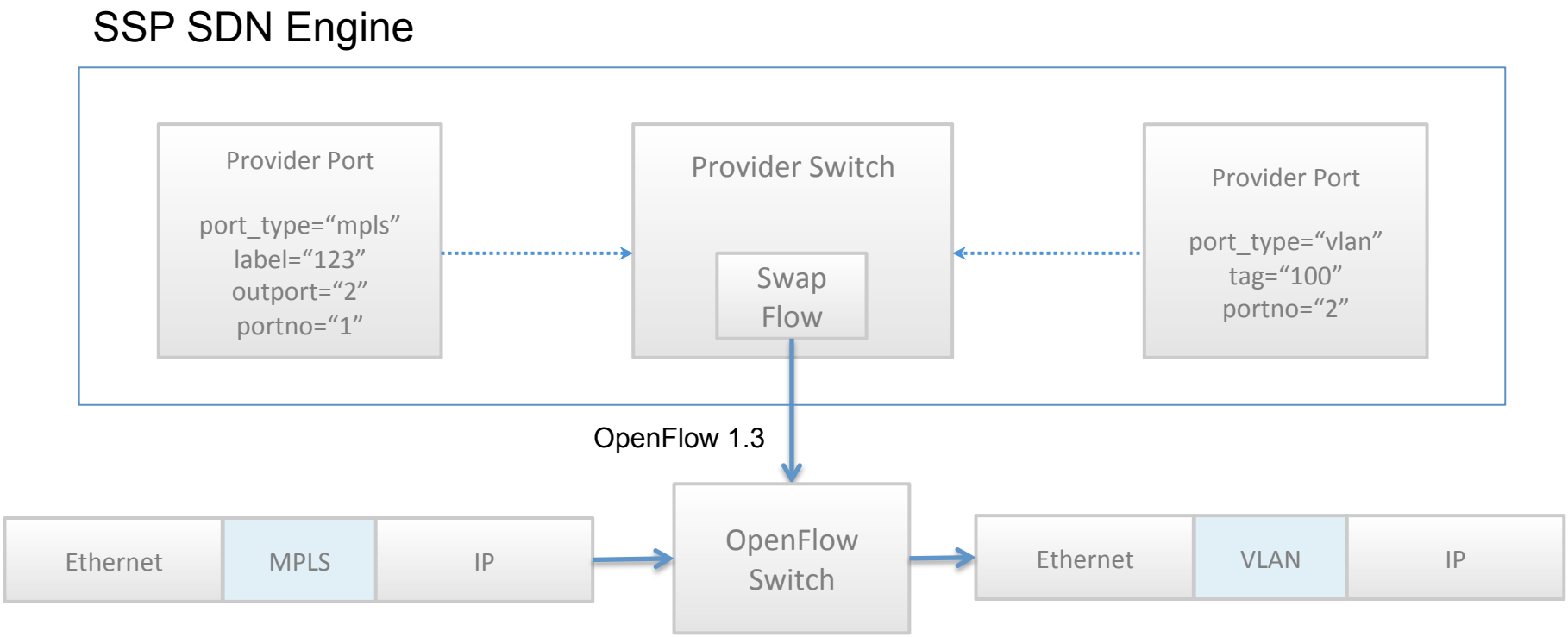


例2: MPLS/VLAN ラベルスワップ



- MPLS Popオペレーションで、L2ヘッダ直後のMPLSラベルを削除
- VLAN Pushオペレーションで新たなVLAN tagを挿入
- この2ステップで、MPLSとVLAN tagをスワップ可能

例2: MPLS/VLANスワップ設定



- MPLSラベル付きフレームを入力
- VLANフレームを出力
- MPLSとVLANをスワップするフロールールを自動生成

例2: 生成されたMPLS/VLANスワップFlow

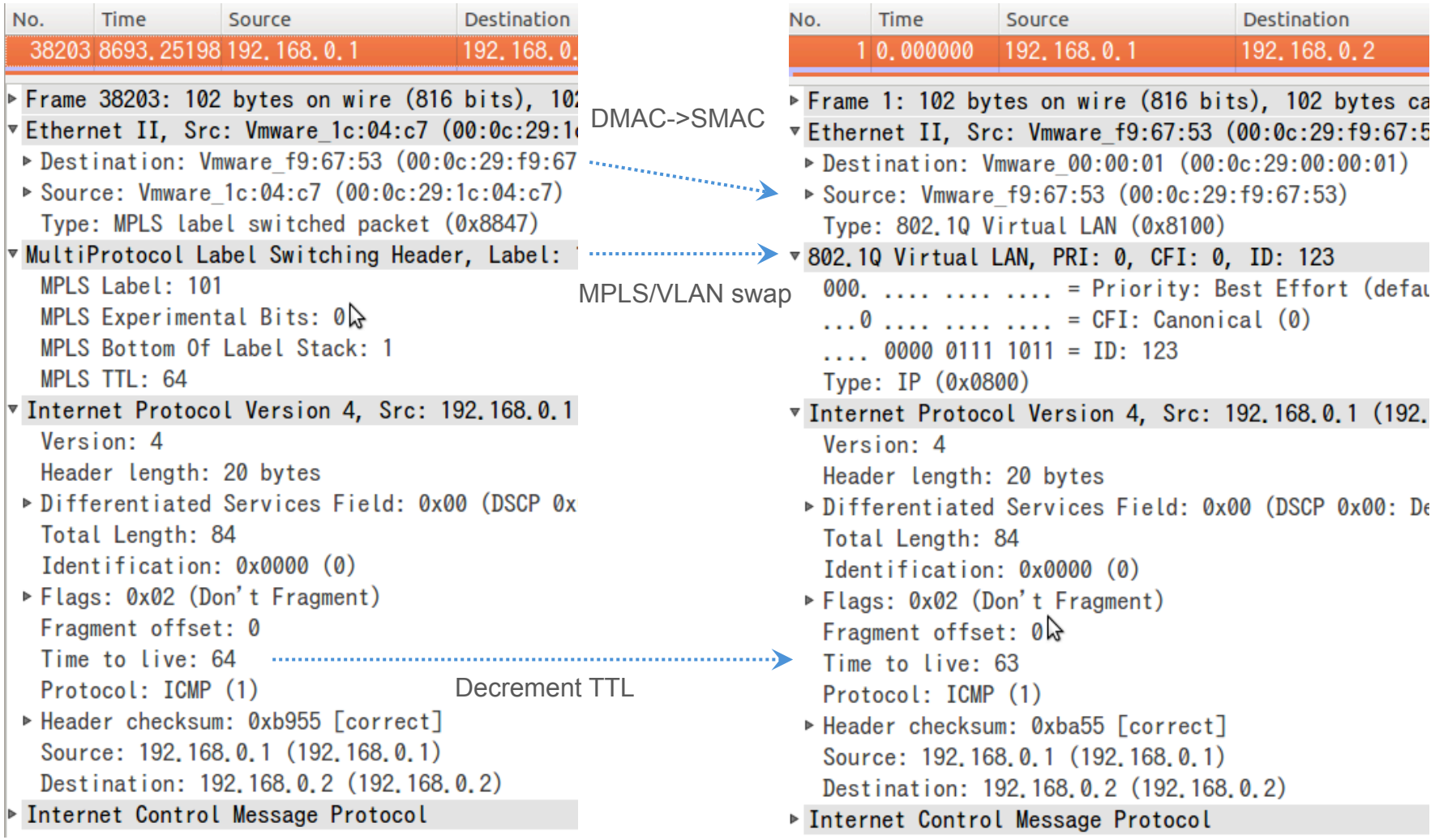
Stratosphere

- **MATCH(Port, Dest MAC, MPLS Label)**
- Match on Switch Input Port (OFPXMT_OFB_IN_PORT)
- Match on Ethernet DST (OFPXMT_OFB_ETH_DST)
- Match on Ethertype (OFPXMT_OFB_ETH_TYPE)
- Match on MPLS Label (OFPXMT_OFB_MPLS_LABEL)

- **ACTION**
- Pop the outer MPLS label (OFPAT_POP_MPLS)
- Decrement IP TTL (OFPAT_DEC_NW_TTL)
- Push a new VLAN tag (OFPAT_PUSH_VLAN)
- Set VLAN tag (OFPAT_SET_FIELD, OFPXMT_OFB_VLAN_VID)
- Set Ethernet SRC (OFPAT_SET_FIELD, OFPXMT_OFB_ETH_SRC)
- Set Ethernet DST (OFPAT_SET_FIELD, OFPXMT_OFB_ETH_DST)
- Output to switch port (OFPAT_OUTPUT)

例2: MPLS/VLANスワップ時の Packet Dump

Stratosphere



ユビキタスオフィスプラットフォームOmniSphere

2013年度8月リリースに向けて開発中

オフィスネットワークの課題

Stratosphere

根本的な問題：
組織変更やレイアウト変更のたびに繰り返されるネットワーク再設計と再構築

無線 LAN の課題

- AP 設置時には、電波干渉・電波強度を考慮する必要あり
  - ✓ マルチテナントビルの場合は調整は難しい
- ネットワーク構成を柔軟に変更できない
  - ✓ ネットワーク毎に AP (Access Point) 設置が必要
- 場所によって無線LANのSSIDを選ぶ必要がある
- 誰がアクセスしているかわからない
- ユーザへの帯域割当に不公平が生じることも
  - ✓ 特定端末が帯域を占有

有線 LAN の課題

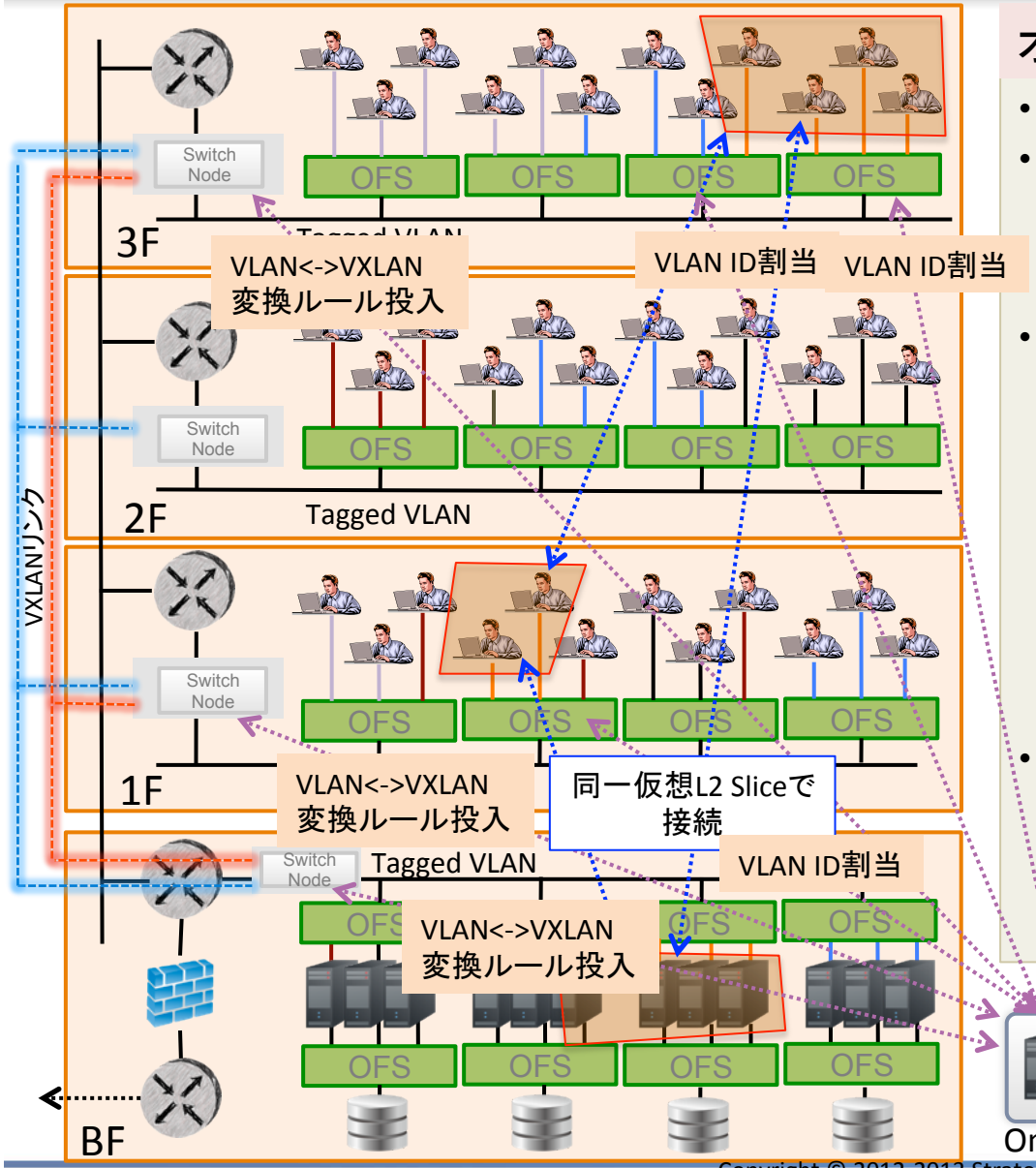
- フロアレイアウトとパッチポートの調査
- 組織変更に合わせた仮想ネットワーク再構成
  - ✓ VLAN による部署毎、プロジェクト毎のネットワーク分離が基本。各フロアスイッチの再設定が必要
  - ✓ 規模が大きくなれば設定工数増大。設定ミスを誘発
- 変更前後のネットワーク規模に合う適切なスイッチ（MACやVLAN学習数など）選定の問題
  - ✓ 場合によっては、取り替えも

共通の課題

- 例えばオフィスでは、自席や所属部署サーバにアクセスするためにPCのネットワーク設定を変更する必要がある
- 場所 (部屋、フロア、オフィス) を移動するとアドレスが変わる

オフィスネット管理SDNパッケージ　- OmniSphere

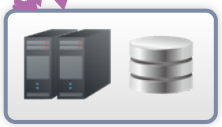
Stratosphere
Coming Soon!



オフィスビルネットワークの例

- 各フロアを1VLANドメインで構成
- フロア間の通信はSwitch Node経由
  - フロア間通信はL3になり、物理ネットワークの構築・運用の複雑度が軽減
  - L2/L3の境界は自由に設計可能
- OmniSphere Engine (OSE)が、ユーザの認証情報、人員配置情報等を集中管理
  - ユーザはどの（どこの）SWに接続しても、常に自分の所属ネットワークに接続される
    - MAC認証／パスワード認証で、自動的に繋いだポートのVLAN IDを設定
    - Switch Nodeで各フロアのVLAN IDとVXLAN IDを自動変換
    - サーバ室に部門別のDHCPサーバを設置しておけば、ユーザの端末には常に同じ設定が可能
    - 論理ネットワークが物理ネットワークから分離される
- OFSの設定はOSEがOpenFlowで行う
  - SW設定の簡略化
    - 基本的に同じ設定で全てのSWを管理可能
  - SW設定のロジックをOSEに集約
    - 設定変更時の変更箇所はひとつだけ

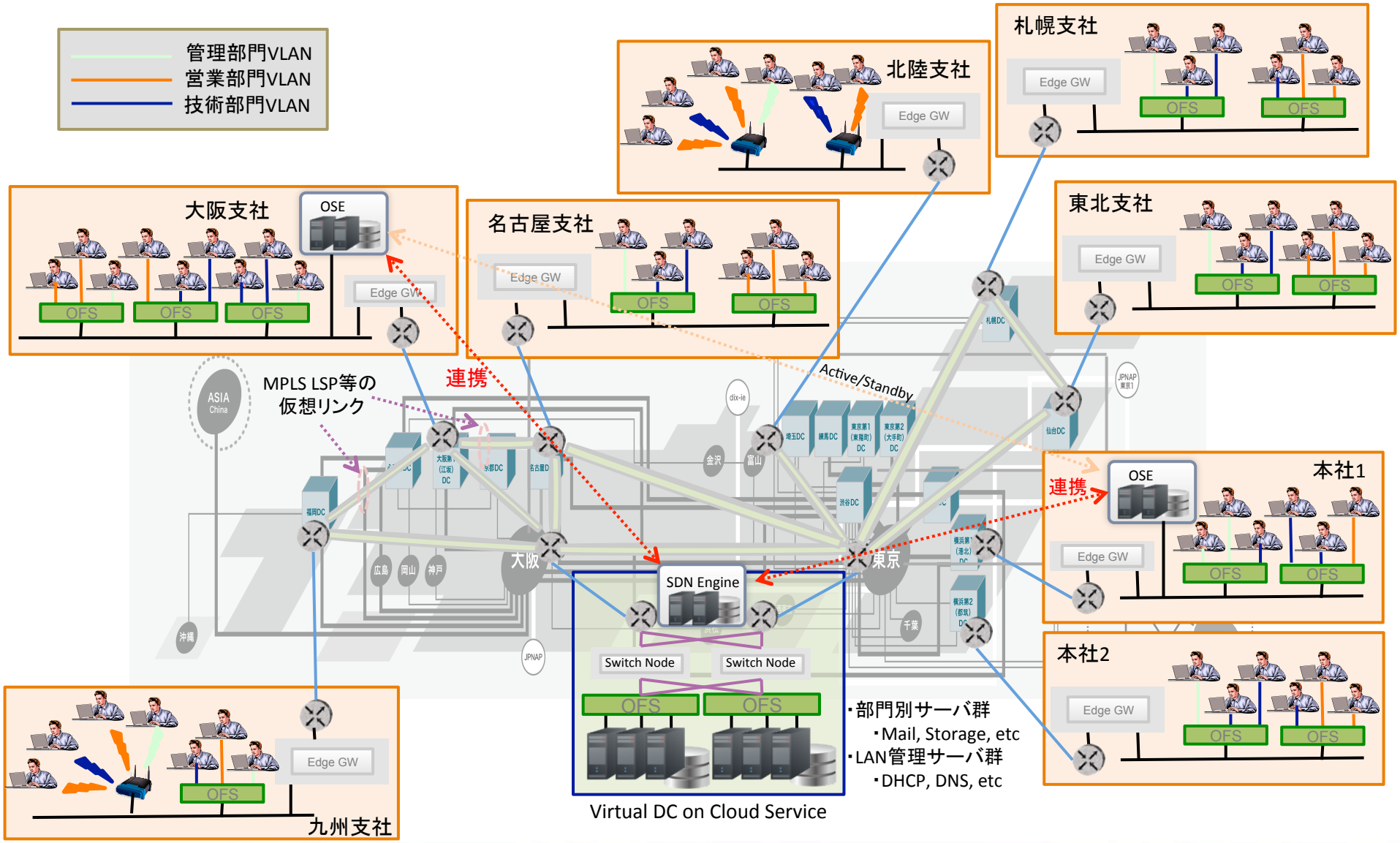
OpenFlow



人員配置情報
設定情報
認証情報
OmniSphere Engine(OSE)

Cloud + SDNによるVirtual IT

Stratosphere



まとめ

Stratosphere

- モジュラー型システム構成
  - ✓ IS, ES, SDN GWなど機能単位毎にソフトウェアをモジュール化
  - ✓ 各機能モジュールを組み合わせることで、柔軟なシステム構成を実現
- 既存ネットワーク技術との親和性
  - ✓ IS/ESの機能分離を明確にしたアーキテクチャーを採用し、端末やサーバ等の資源の分離をVLANで実施しているため、既存VLANネットワークを活用しやすく、導入障壁が少ない
  - ✓ ESの設置場所を、HV内、ToRの外側、クラスタの境界等と、ユーザニーズに合わせて個別に設定が可能
- ソフトウェア処理でも、ハードウェアリソース等を活用し高性能を追求
  - ✓ Proactive制御によりOpen Flow処理オーバヘッドを低減
  - ✓ モジュラー化により、機能毎のオフロードとスケールアウトを実現
- さまざまなネットワークシーンを想定したソリューションを用意
  - ✓ MPLSによる広域ネットワークへの接続を実現
  - ✓ IPSecによるインターネット経由でのユーザサイトの接続を実現
  - ✓ OmniSphereによる、ユビキタスオフィスネットワーク管理の実現