[招待講演]

ネットワーク仮想化基盤における仮想ノード間とスライス-外部ネットワーク間の接続機能

金田 泰, 白石 圭 (日立)
中尾 彰宏 (東大 / NICT)

はじめに

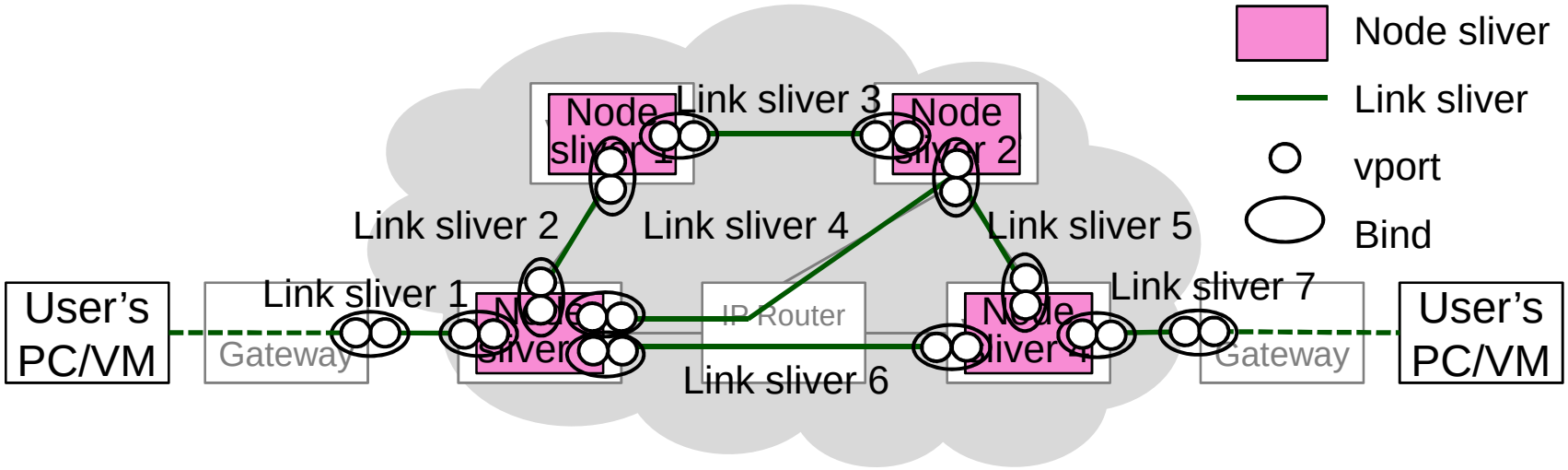
- **VNode プロジェクトで仮想ネットワークのノード間やノードと外部ネットワークをつなぐ部分を担当してきた.**
  - NICT の場で東大, NTT, 富士通, NEC, 日立の各社が 2009-2010 年に**共同研究**し, 2011 年以降は NICT 委託研として継続している.
  - 既存のインフラを利用して新世代ネットワークを研究するための**ネットワーク仮想化基盤**とくにそのための物理ノード VNode を開発している.
  - ひとつの物理ネットワーク上で独立かつ自由に設計された**複数の仮想ネットワークが同時に動作**する環境を実現 している.
- **VNode は JGN-X にも導入されたが, これまで利用者視点での紹介が十分でなかった.**
- **この発表では, VNode 上に生成される仮想化基盤内とその外部への “接続” について紹介する.**
  - “接続” が利用者からどのように参照されるか?
  - “接続” が VNode においてどのように実現されるか?
  - 日立開発部分に限定された内容ではない.

仮想化基盤における役割

- **仮想ネットワークの生成と運用にかかわる役割は，つぎの 3 者にわけられている.**
  - ◆運用者 (オペレータ)
    - ポータル (Web インターフェース) を使用して実ネットワークを管理.
    - 開発者は運用者がシステムに登録.
  - ◆開発者 (デベロッパ) — “利用者” 1
    - ポータルを使用して仮想ネットワークを生成し，一般ユーザを管理.
    - 仮想ネットワーク上で一般ユーザにサービスを提供.
  - ◆一般ユーザ — “利用者” 2
    - 仮想ネットワークを使用.
    - 端末の設定と端末に関する情報は一般ユーザがポータルに登録.
- **この報告においては開発者の視点を中心にする.**

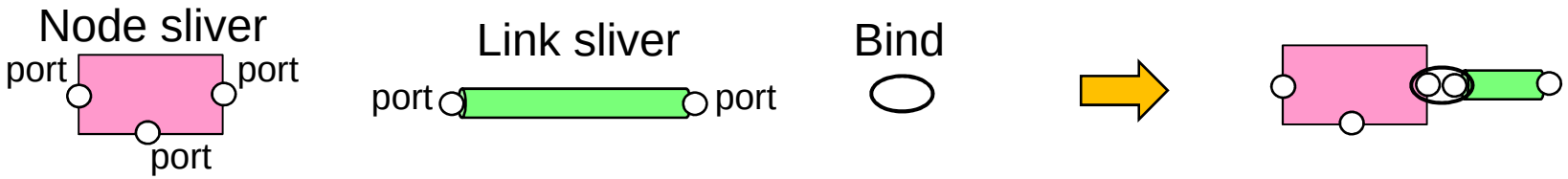
スライス

- **スライス (slice) は仮想化基盤上につくられる抽象化・構造化された仮想ネットワーク.**
  - ◆ PlanetLab, GENI 等の “スライス” とちがって抽象化・仮想化されている.
- **開発者はスライスを定義 (記述) してポータルから投入する.**
  - ◆ XML にもとづくスライス定義言語で記述する.



スライスの構成要素

- **スライスの構造はつぎの 3 つできまる.**
  (1) ノードスリバー，(2) リンクスリバー，(3) 結合



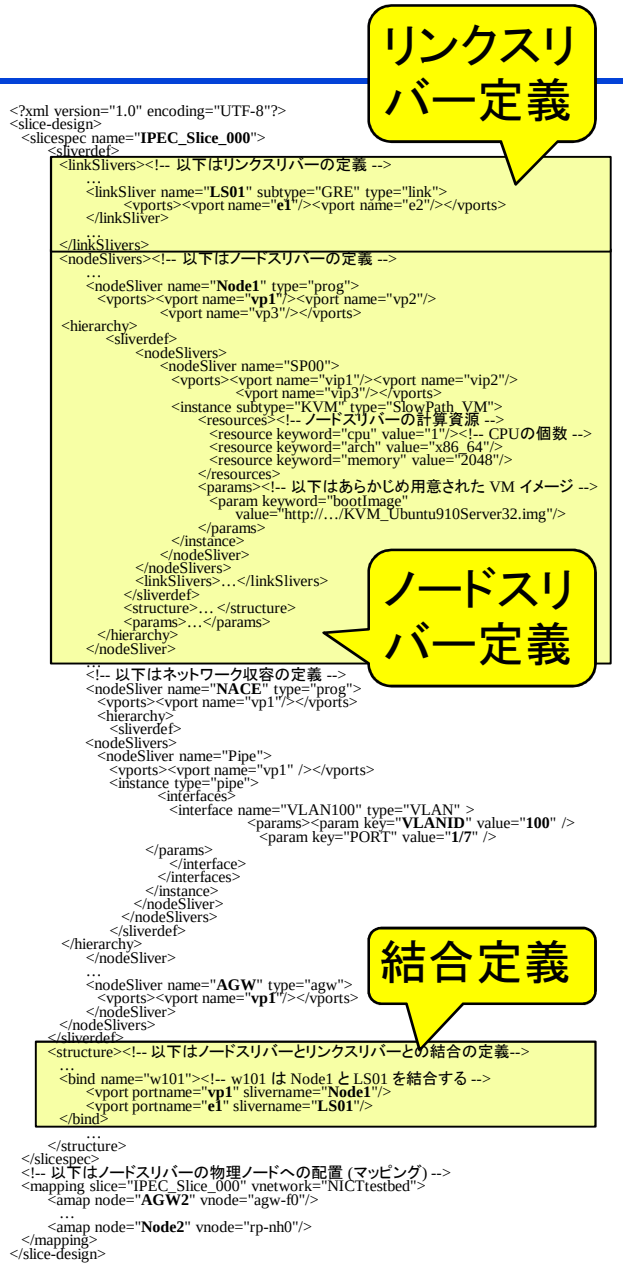
  - ◆ 結合がノードスリバーとリンクスリバーのポートをむすびつける.

- **ノードスリバー (Node sliver, 仮想ノード)**
  - ◆ プロトコル処理やノード制御などを実行するのに使用する.
  - ◆ 初期状態ではなんの機能ももたず，プログラムや Linux の機能を使用しないかぎりはパケットはスイッチングもルーティングもされない.
  - ◆ ノードスリバーの内部構造や動作はスライス定義とはべつにあたえる.

- **リンクスリバー (Link sliver, 仮想リンク)**
  - ◆ ノードスリバー間を結合する 2 層 (L2) の仮想リンクであり，単純にその一方の端点から入力されたパケットを他方の端点につたえる.
  - ◆ ことなる物理ノード内にあるノードスリバーを結合する.

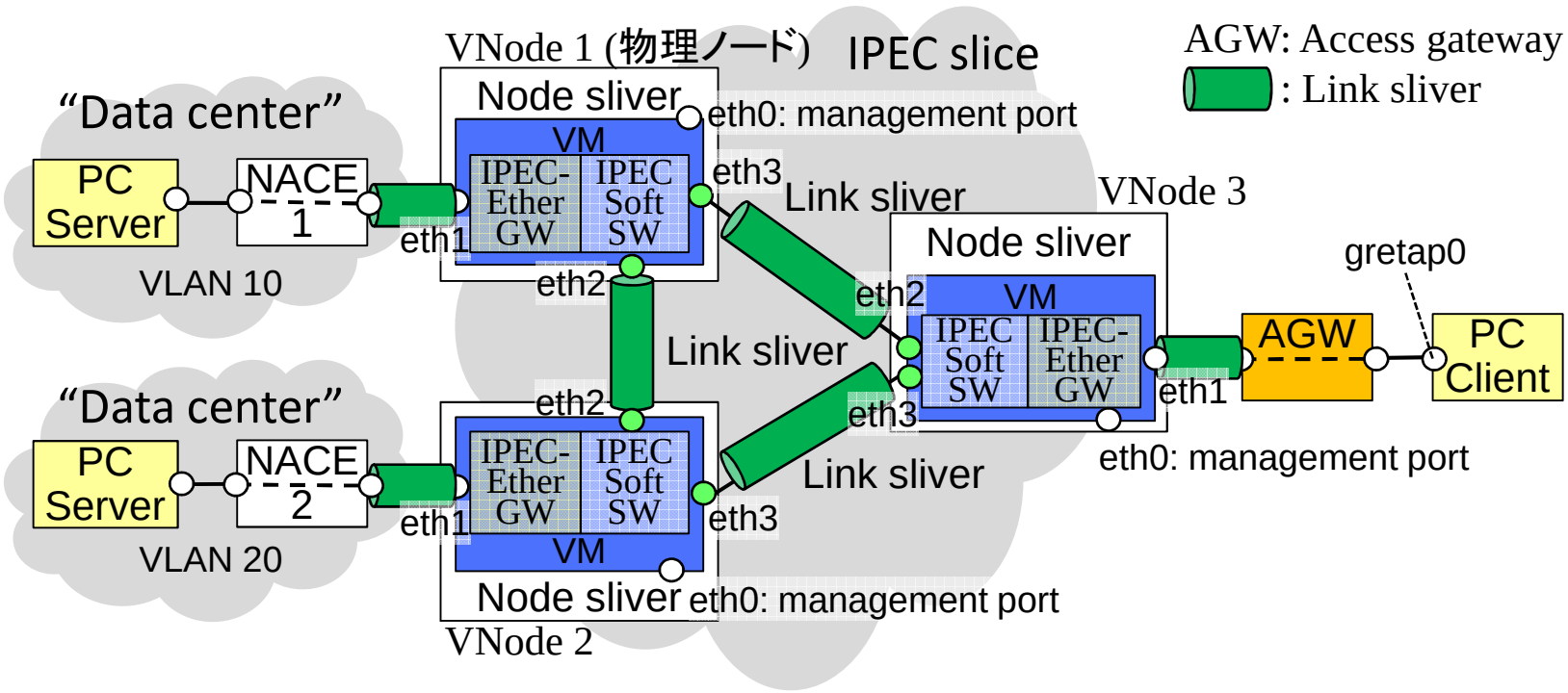
スライス開発手順

- **Step 1: 開発者は XML でスライス定義を記述する.**
  - ノードスリバーとリンクスリバーを独立に記述し，それらの結合を別途指定.
  - 各ノードスリバーはスローパス (VM) やファストパス (ネットワーク・プロセッサ) をふくむ.
- **Step 2: ノードスリバー要素への指示をスライス定義とはべつにあたえる.**
  - スローパスの場合: 開発者が必要なら VM イメージを用意し，VM にログインして各種の設定，プログラムのロード・実行などをおこなう.
  - ファストパスの場合: スライス定義において指定されたプログラムがロードされ実行される.



「Step 1: スライス定義の記述」と基盤内接続

- **開発者はスライス定義をポータル (Web I/F) から投入して生成する.**
  - スライス生成時にノードスリバーどうしが接続され，ノードスリバーが外部ネットワークや端末へのゲートウェイ (AGW) にも接続される.
  - VM 内部の “接続” はこの時点ではつくられない.

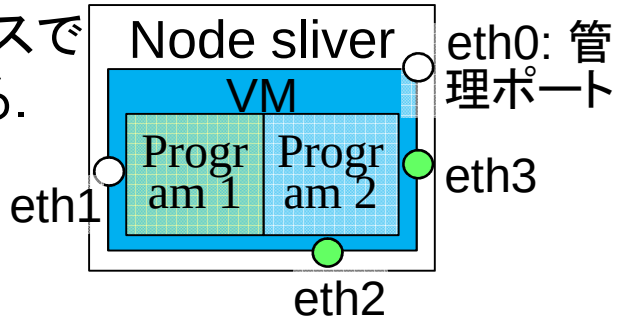


「Step 2: ノードスリバー要素への指示」と基盤内接続

■ ノードスリバーと外部との接続の開発者からのみえかた

◆管理用インターフェース

- 各 VM の eth0 は管理用のインターフェースであり，開発者用のネットワークに接続される．
- eth0 には自動的に IP アドレスがつけられるが，開発者はその値をポータルで知ることができる．



◆ノードスリバーとリンクスリバーのインターフェース

- 開発者にとってのリンクスリバー出入口は eth1, eth2, … である．
- 開発者はスライス定義上のポートと eth1, eth2, … との対応を指定しないので，生成してから対応を知る必要がある．

◆ノードスリバー上でのプロトコル

- IP/Ethernet を使用すると Linux の “遺産” が使用できるので，プログラムは容易になる．

「Step 2: ノードスリバー要素への指示」と基盤内接続 (つづき)

■ ノードスリバー要素への “接続” などの指示

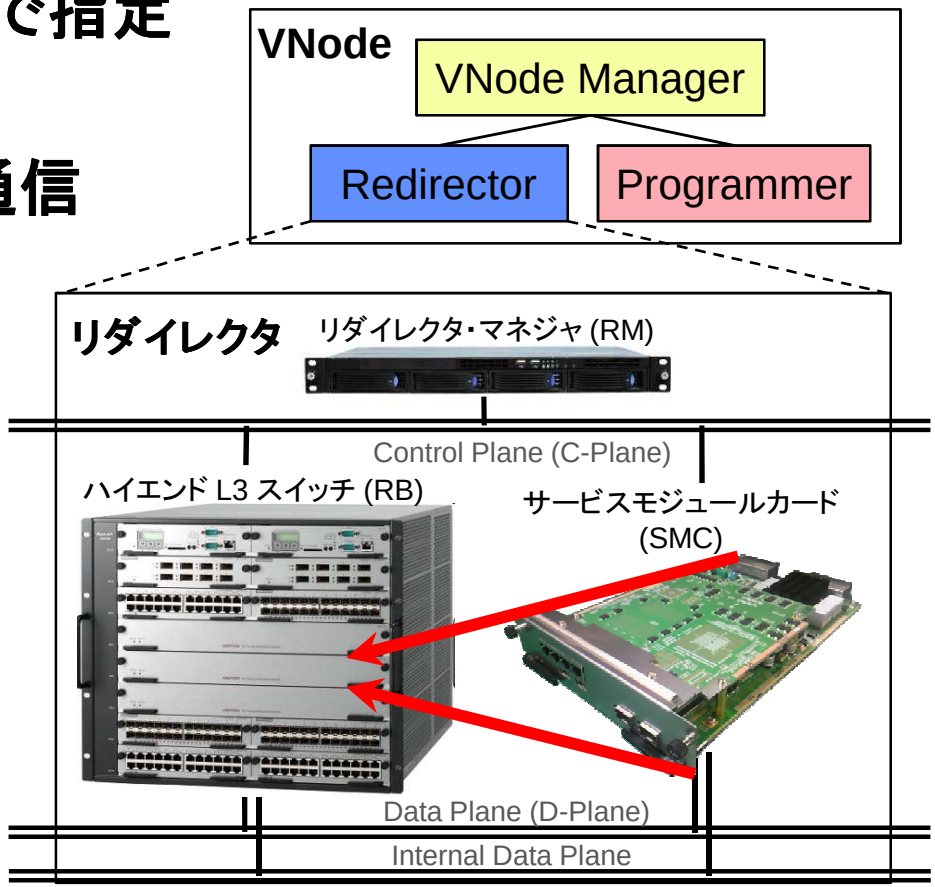
- ◆ 開発者は管理用インターフェース eth0 から ssh 等でログインする.
- ◆ リンクスリバーと eth1, eth2, … の対応をしらべる (導通ツール等使用).
- ◆ eth1, eth2 などに必要な設定をする.
  - IP 通信をおこなうときは，つぎのようなコマンドを使用してIPアドレスを設定:
    ifconfig eth1 192.168.1.1 (ただし，VM イメージで設定されていれば必要ない.)
- ◆ eth0 経由でプログラムをロードし実行する.
  (VM イメージにプログラムがふくまれていればロードする必要はない)

■ 補足: 非 Ethernet 通信について

- ◆ eth1, eth2, … などという名称のインターフェースを使用するが，パケット･フォーマットは自由 (Ethernet パケットをつかわなければならないわけではない).
- ◆ 自由なパケット･フォーマットをあつかうため，VM 上のプログラムで promiscuous mode をつかう.
  - これによってパケットの最初のビットから操作できる.
- ◆ 現在の実装では GRE トンネルによって自由なフォーマットのパケット通信を実現している.

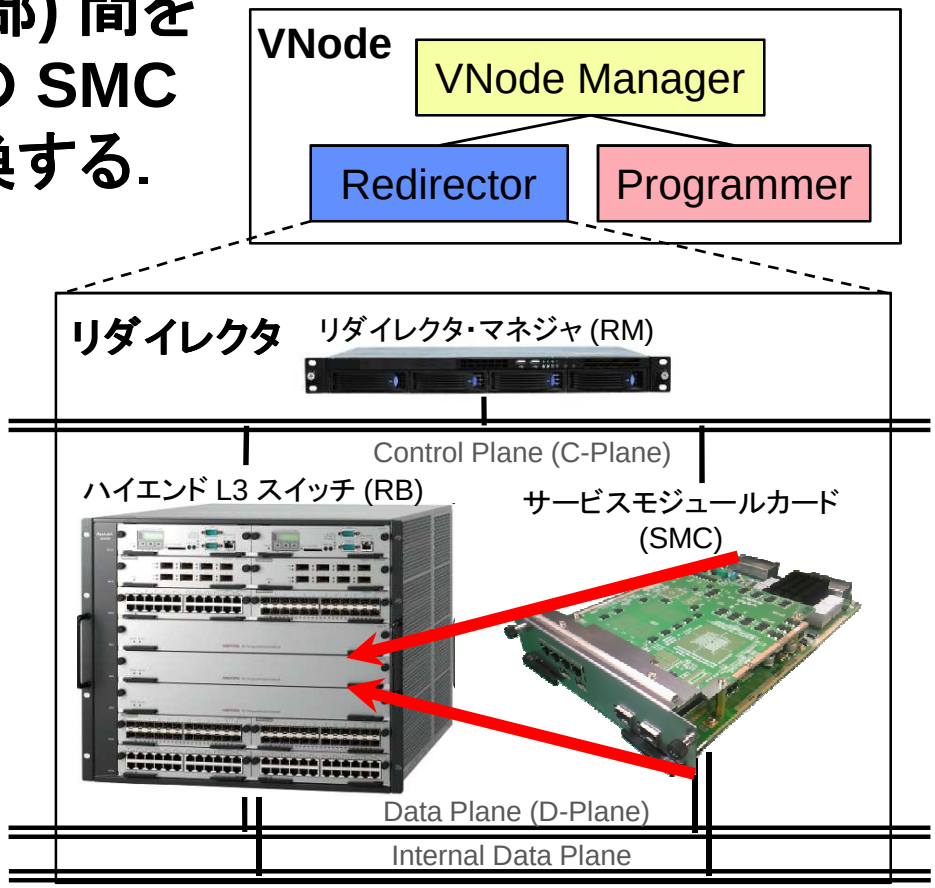
VNode と基盤内接続の実現

- VNode の構成要素「リダイレクタ」が中心となって，基盤内(ノードスリバー間) 接続を実現
- プログラマがノードスリバーを実現: プログラマ内ではリンクスリバー入口が eth1, eth2, … で指定される.
- VNode 内部は Ethernet で通信する: MAC アドレス対 (MACP, MACR) がリンクスリバーに対応.
- リダイレクタがリンクスリバーを実現: VNode 外部は GRE/IP で通信する: 3 つ組 (IPVN1, IPVN2, GREkey) がリンクスリバーに対応.



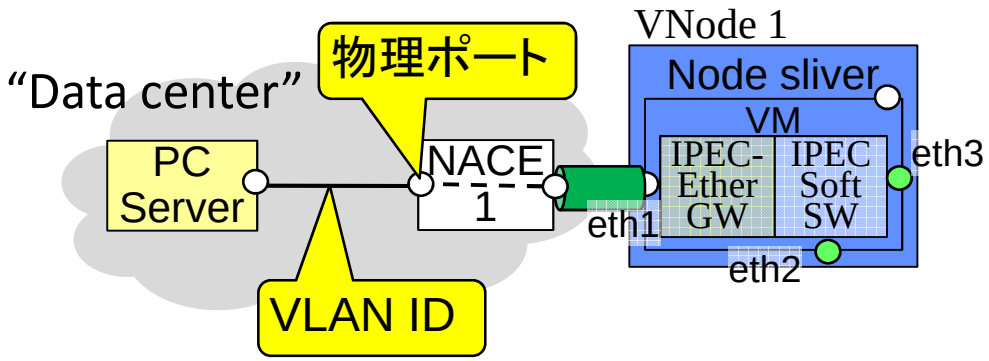
VNode と基盤内接続の実現 (つづき)

- プログラマとリダイレクタとを独立させ，VNode 内外のデータ表現も独立にきめられるようにすることで，VNode の多様な発展 (進化) を可能にしている.
- これらのデータ表現 (内部-外部) 間をネットワーク･プロセッサ搭載の SMC によって最高 10 Gbps で変換する.



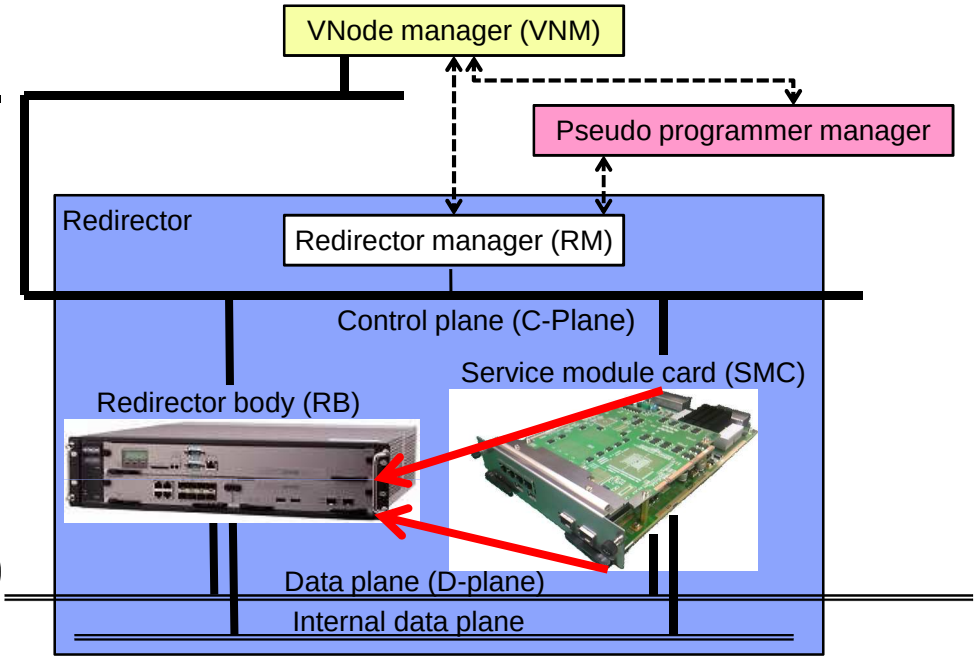
開発者による外部ネットワークとの接続

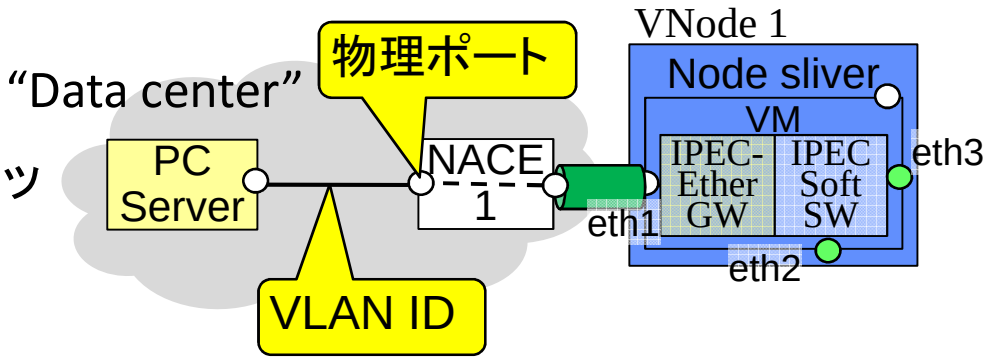
- 外部ネットワークと接続するとき，スライス定義上で通常は物理ポートと VLAN ID とを指定する.
- 開発者がネットワーク収容装置 (NACE, NC) から外部ネットワークへの配線をする (または既存の配線にしたがってそれらをスライス定義に指定する).
- 補足: 現在，外部ネットワークは Ethernet にかぎられている.



外部ネットワーク収容装置と外部接続の実現

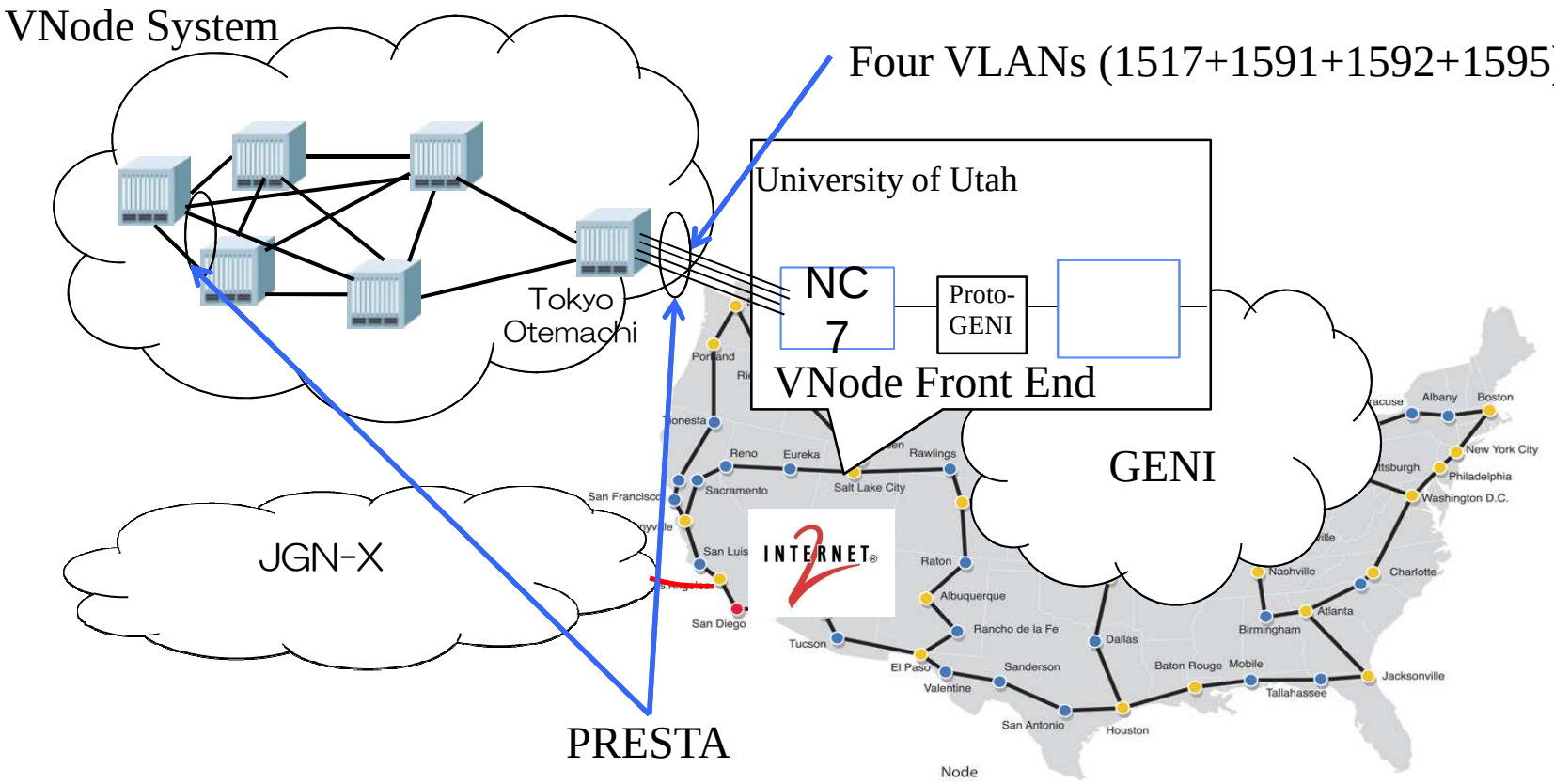
- **NACE を使用して外部ネットワークをスライスに接続する.**
- **NACE の構成: NACE はゲートウェイの一種だが，構造上は VNode にちかい.**
  - ◆ VNode マネジャ，リダイレクタ，疑似プログラマで構成される.
- **接続: 指定した NACE (RB) の物理ポートに，指定した VLAN ID で外部ネットワークを接続する.**
  - ◆ SMC を使用して最高 10 Gbps でパケット･フォーマット変換する.





GEC 15 におけるスライス間接続とデモ

- **10/23-25 の GEC 15 では，ユタ大においた NACE を使用してProtoGENI のスライスを (外部ネットワークとして) JGN-X 上のスライスと接続して，双方がデモをおこなう予定．**
  - ◆ GEC 15 = 15th GENI Engineering Conference @ Houston, TX



まとめ

- **仮想化基盤においては，開発者が XML によって記述したスライス定義をポータルから投入することでスライスが生成できる.**
- **開発者がどのように接続先を参照・指定し，それがどのように VNode と NACE で変換され通信が実現されるかを説明した.**
  - **ノードスリバー間**: 開発者はスライス定義でノードスリバー外側，インターフェース名によって内側の接続を指定する. VNode の内外でことなるデータ表現を SMC が高速変換する.
  - **外部ネットワーク**: NACE を使用すれば開発者がスライス定義上で指定した物理ポートと VLAN ID によって外部ネットワークと接続することができ，高速通信できる.

付録: スライスと実ネットワークとのマッピング

- **ノードスリバーと VNode (物理ノード) とのマッピングをスライス定義に記述することができる.**
  - ◆ 記述しなければ自動的にマップされる.