

# フェデレーション機能をもたない仮想化基盤の フェデレーション

2011-11-25 改訂

金田 泰, 垂井 俊明 (日立)

# はじめに

---

- 共同研究プロジェクトにおいてネットワーク仮想化基盤を開発し、発展させている。
  - ◆ NICT の仮想化ノード・プロジェクトにおいて仮想化基盤を開発した (東大, NTT, NEC, 富士通, 日立. 2008-2010).
  - ◆ NICT 委託研で新仮想化基盤とそのフェデレーション機能を開発中 (Project A & B: 上記にくわえて KDDI 研. 2011-2014).
  
- この委託研では、フェデレーション機能がない旧仮想化基盤をもとにして他の仮想化基盤とのフェデレーションを実現させる。
  - ◆【目標 0】複数の旧仮想化基盤のフェデレーション.
  - ◆【目標 1】旧 (+新) 仮想化基盤と G-lambda ベースの仮想化基盤とのフェデレーション.
    - G-lambda は KDDI 研, 産総研などによる Grid をむすぶネットワークの資源管理 (I/F) に関するプロジェクト.
  - ◆【目標 2】新仮想化基盤と ProtoGENI とのフェデレーション.
    - ProtoGENI は米国におけるネットワーク仮想化プロジェクト GENI で開発されているプロトタイプ (GENI = Global Environment for Network Innovations).

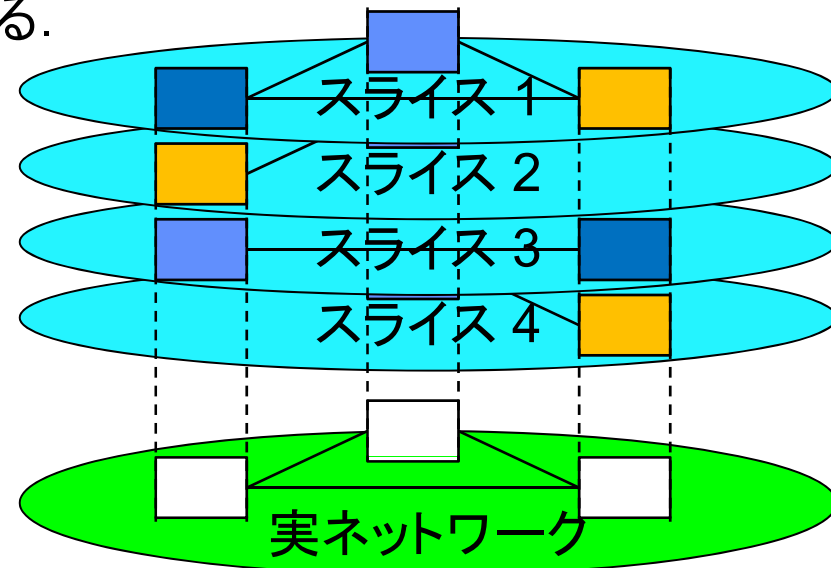
# 仮想化ノード・プロジェクトと仮想化基盤

## ■ 仮想化ノード・プロジェクト

- ◆ NICT が設定した場で東大および各社がネットワーク仮想化基盤 (以下, “旧仮想化基盤” とよぶ) を開発してきた.

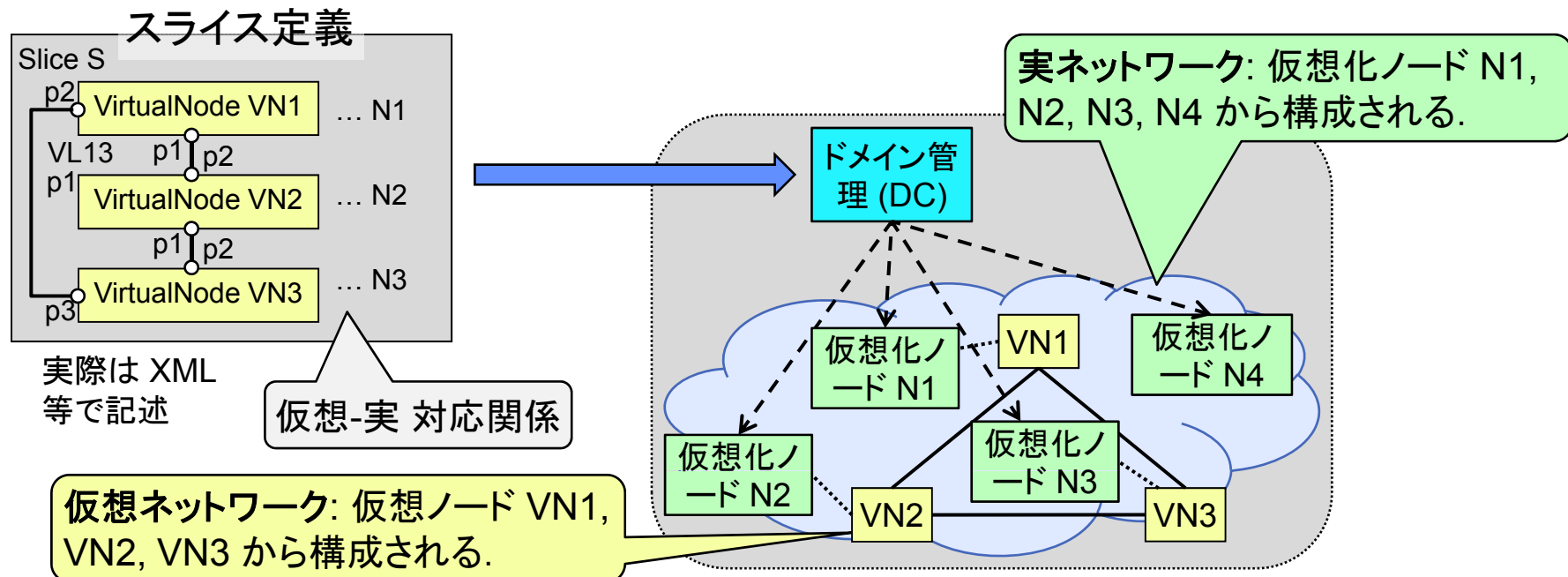
## ■ ネットワーク仮想化基盤

- ◆ 既存のインフラを利用して新世代ネットワークの研究ができるようにした.
- ◆ 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する.
- ◆ すでに研究開発用テストベッド・ネットワーク JGN-X に導入され, さらに (“新仮想化基盤” に) 発展しつつある.



# 仮想化基盤におけるスライスの生成・管理

- 仮想ネットワークの設計図 (構成図) を “スライス定義” として記述する.
  - ◆ 仮想ノード (node slivers), 仮想リンク (link slivers) とそれらの結合関係 (binds) が記述される.
  - ◆ 下図ではグラフとして表現しているが, 実際は XML を使用する.
- スライス定義をドメイン管理サーバ (DC) にあたえると仮想ネットワークが作られる. (DC = Domain Controller)



# 仮想化基盤間のフェデレーション

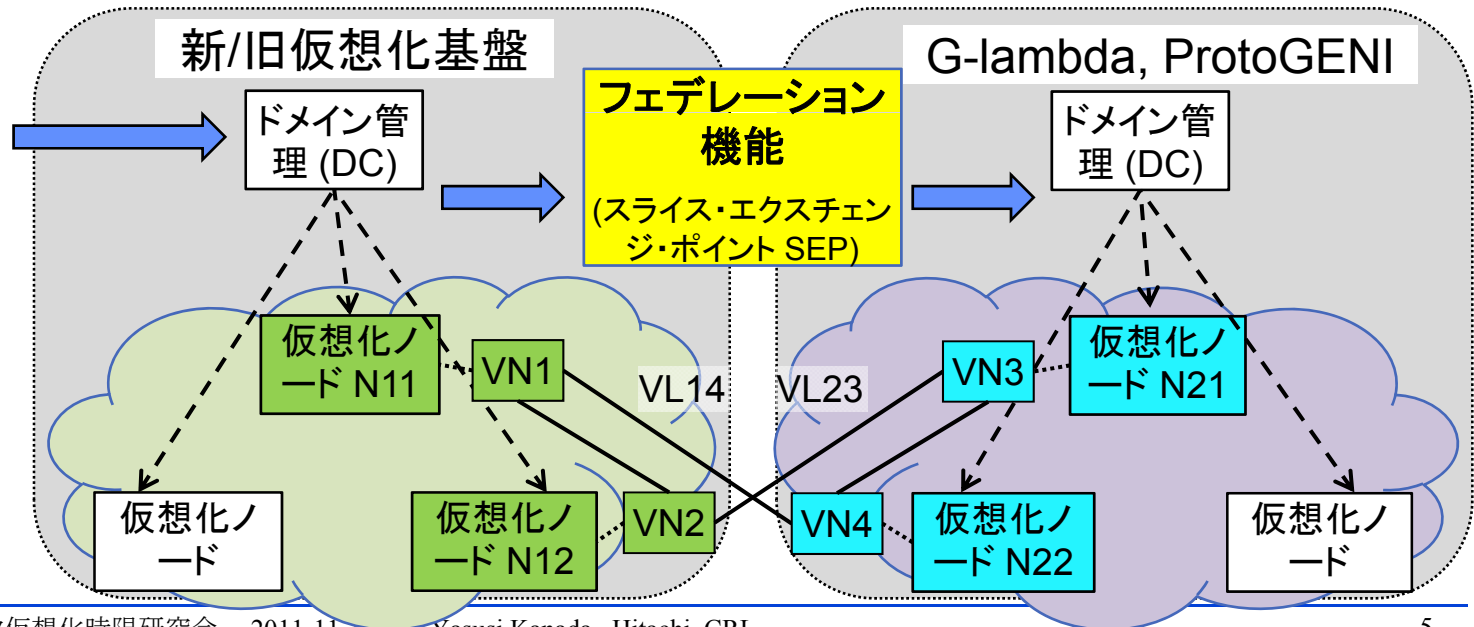
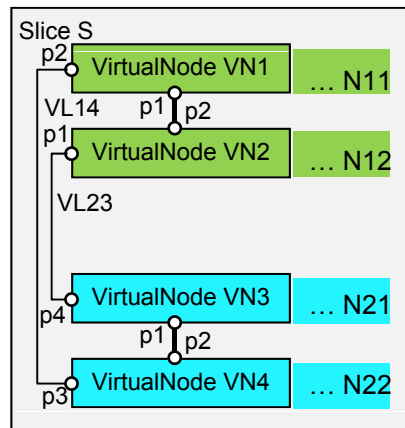
## ■ 仮想化基盤間のフェデレーション機能とは？

- ◆ 複数の仮想化基盤のうちの 1 個から、他の仮想化基盤にまたがる仮想ネットワークを生成・管理する機能.
- ◆ この発表では生成にフォーカスする.

## ■ KDDI 研と共同でフェデレーション機能を開発中

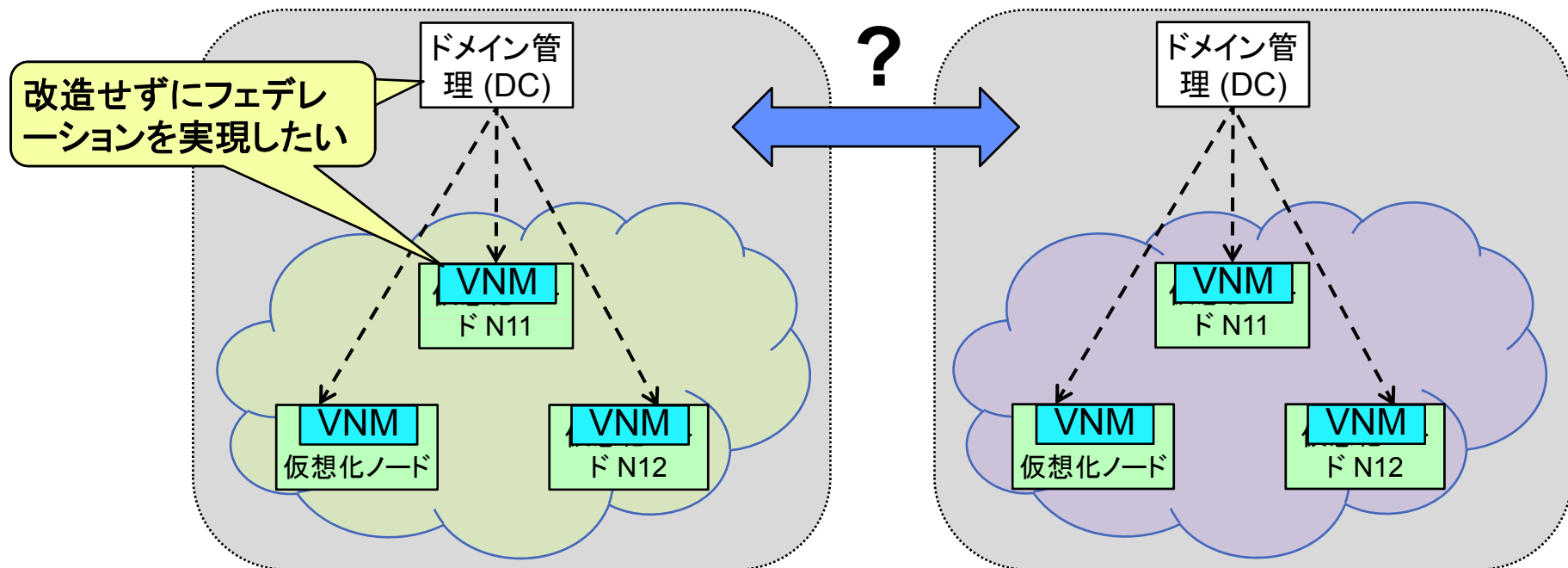
- ◆ 新/旧仮想化基盤側は日立担当.
- ◆ G-lambda, ProtoGENI 側は KDDI 研担当.
- ◆ フェデレーション・インターフェースは KDDI 研主導で共同開発.

### スライス定義



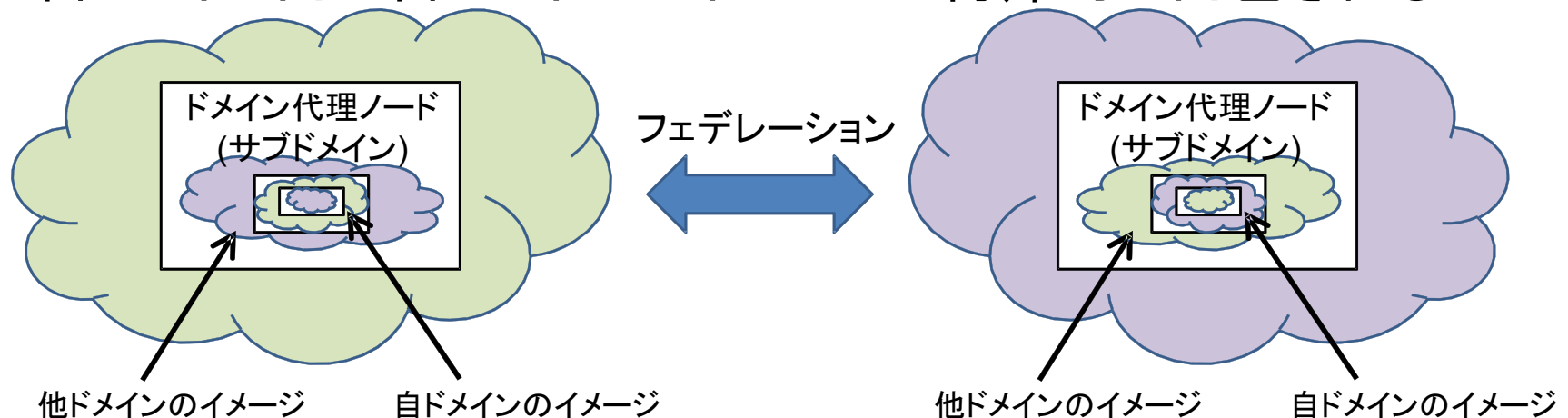
# 旧仮想化基盤のフェデレーション機能開発における目標

- フェデレーション機能をもたない仮想化基盤を改造せずにフェデレーションを実現する。
  - ◆ 改造によって発生しうるインパクトをなくす: バグの発生, 性能劣化, システム停止など.
- スケーラブルなフェデレーションを実現する。
  - ◆ 連携ドメイン数や他ドメインのノード数による性能劣化をなくす.



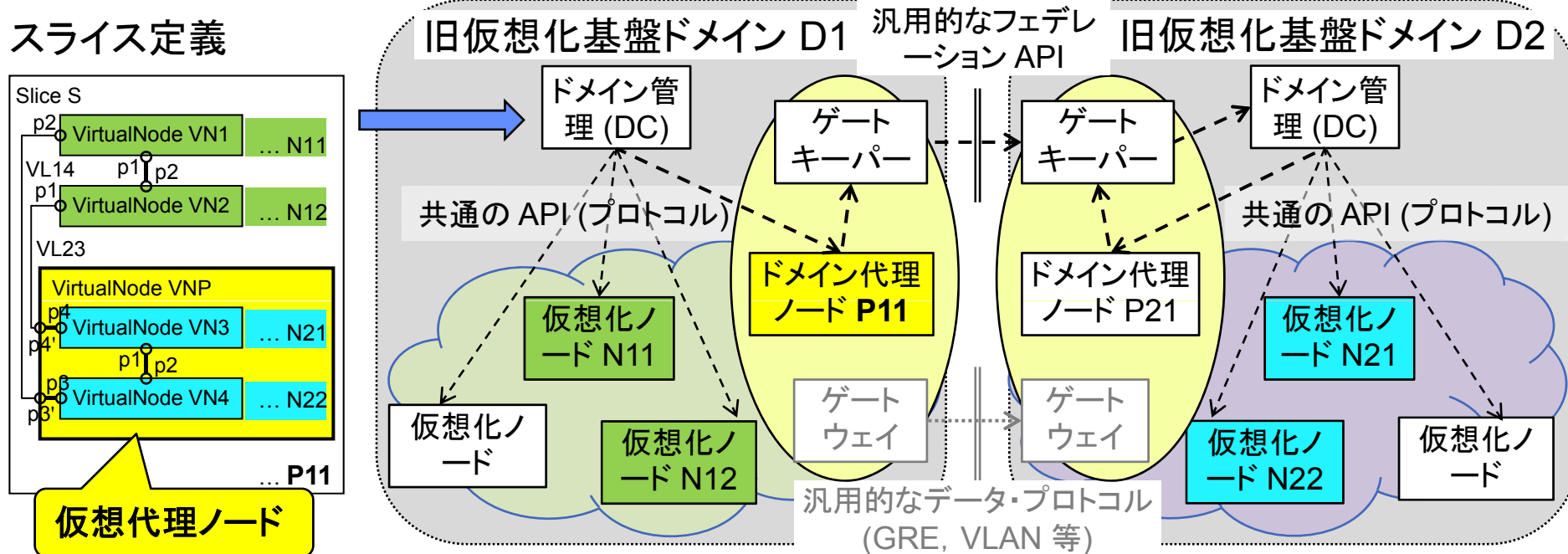
# “他ドメイン” という概念がないときのフェデレーション

- フェデレーション機能のない仮想化基盤には“他ドメイン” という概念がない — 存在するのは自ドメインだけ.
- スライス定義の“他ドメイン” 部分は自ドメインに属する — “他ドメイン” は自ドメインのサブドメインである.
- “ドメイン代理ノード” の概念
  - ◆ スライス定義においてサブドメインを表現する手段は仮想化ノードだけである — ドメイン代理ノードとして表現する.
  - ◆ ドメイン代理ノードは DC からは通常の仮想化ノードとしてみえる.
- 自ドメイン内に自ドメインのイメージが再帰的に内包される



# フェデレーション・アーキテクチャ: 管理情報の送信側

- ドメイン管理サーバ (DC) にあたえるスライス定義は, 他ドメイン部分を “ドメイン代理ノード” でくる。
  - ◆ DC はスライス定義のドメイン代理ノード内部分を参照しない (性能劣化しない).
- DC はスライス定義を実在するドメイン代理ノードに仮想化ノードと同様に 配布する (共通の API / プロトコル を使用する).
- ドメイン代理ノードは “ゲートキーパー” 経由でスライス定義を他ドメインに展開する.



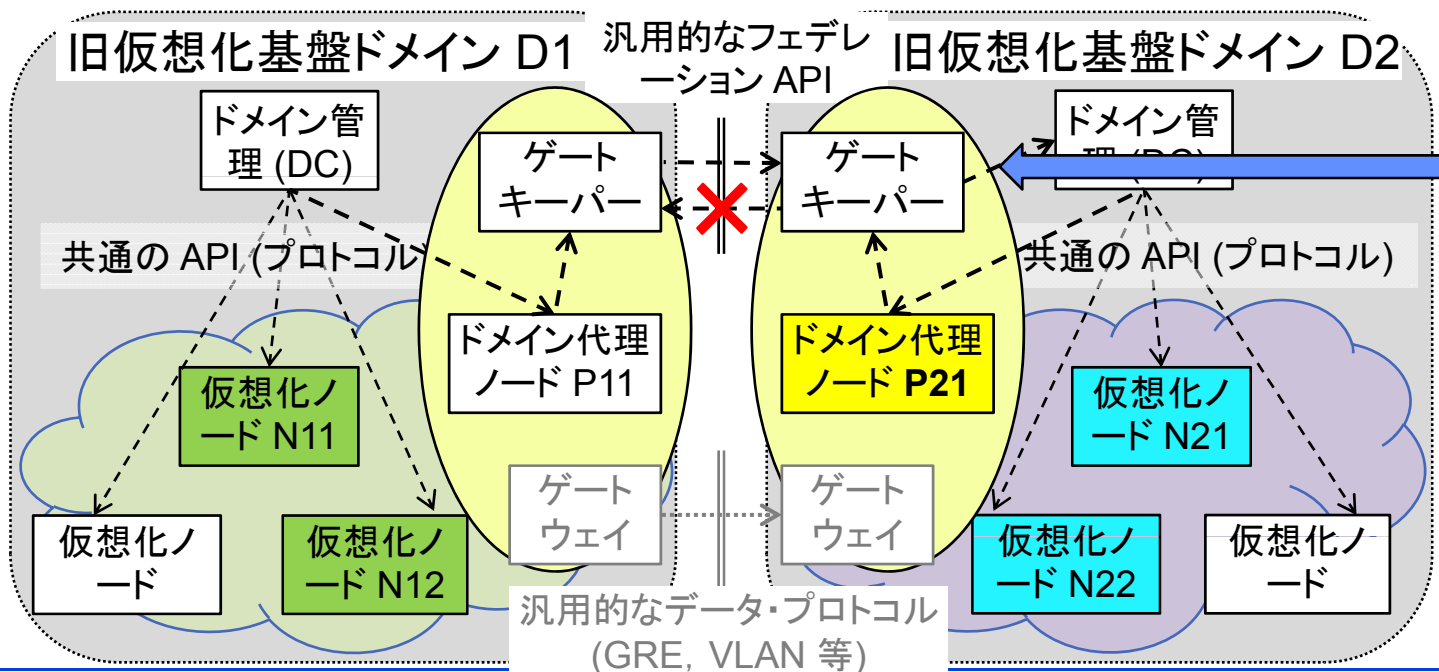
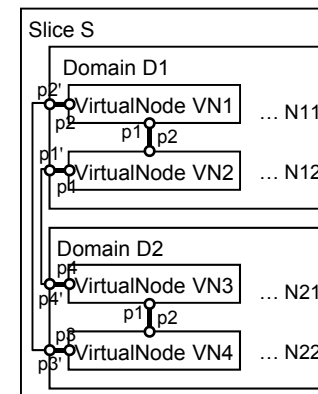


# フェデレーション・アーキテクチャ: 管理情報の受信側

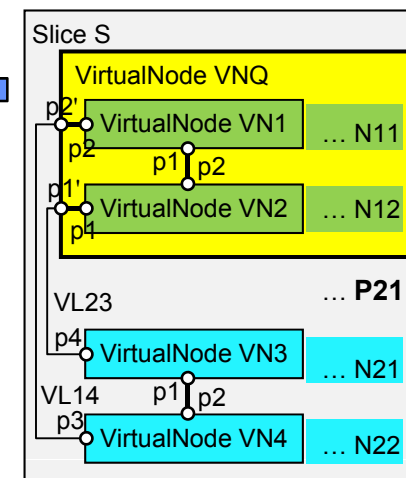
- 受信側の“ゲートキーパー”はスライス定義における相手ドメインのノードをドメイン代理ノードでくっつけて DC に転送する.

- ◆ 自ドメインのノードはくられていない. (フェデレーション API においては全ノードがドメインごとにくられている →)
- ◆ ドメイン内の情報をそのドメインと開発者だけが参照できるように暗号化することも可能.

- DC はスライス定義を実在するドメイン代理ノードに仮想化ノードと同様に 配布する.



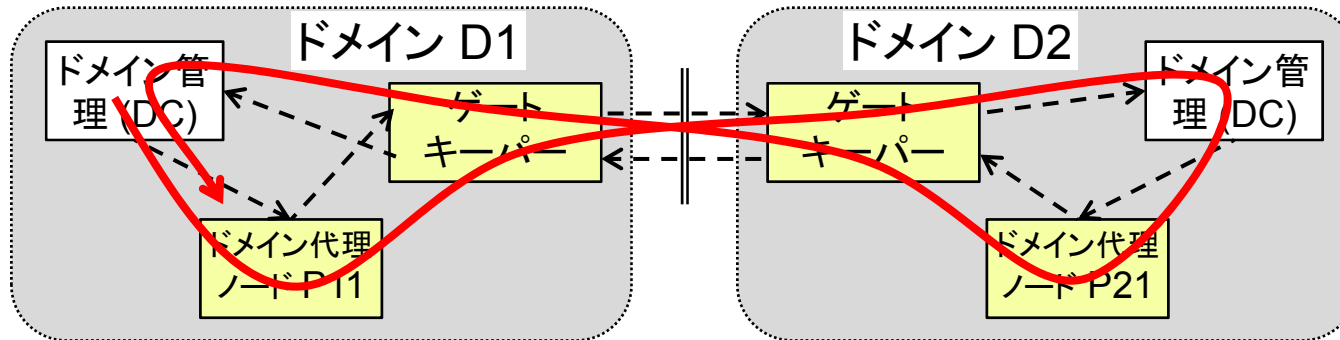
## スライス定義



# ドメイン間メッセージングの無限再帰とその防止法

## ■ ドメイン間メッセージはほうっておくと無限再帰する.

◆ドメインが3個以上のときはさまざまなパターンがある.



## ■ 無限再帰の防止法

◆ **基本:** メッセージを識別し, 処理済みのメッセージは転送/処理しない.

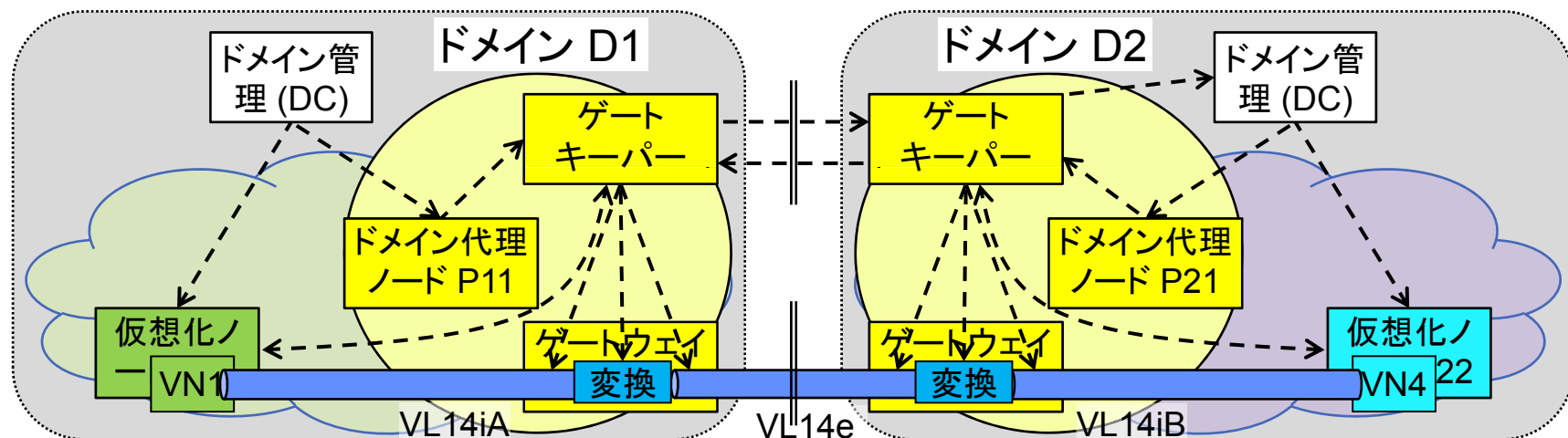
- なにを識別するか? — オーダやメッセージ中のスライスを識別.
- DC がスライス識別子をかきかえるときは要注意.

◆ **補足:** バグがあると基本の方法がうまく作用しない可能性があるため, つぎの方法で補足する.

- **マーキング:** 外部からのメッセージ (スライス情報) にマークをつける  
— 2ドメイン間のときだけ有効.
- **TTL (Time-To-Live):** 管理情報の投入時に TTL をつける  
— TTL をこえる数のドメインが数珠つなぎになっているとうまくいかない.

# ドメイン間リンクの生成

- ドメイン D1-D2 間の仮想リンクは 3 個のパスで実装される:  
ドメイン D1 内, ドメイン D1-D2 間, ドメイン D2 内.
- スライス定義に関するドメイン間要求-応答 のあいだに,  
ゲートキーパー - ゲートウェイ 間通信によって生成される.
  - ◆ ドメイン内は通常のドメイン内パス設定法によって生成される  
- 旧仮想化基盤においては GMPLS による.
- ゲートウェイにおいてプロトコル変換
  - ◆ 不要ならバイパスすることも可能 (たとえば, VLAN パスならゲートウェイへの MAC アドレス設定を省略できる).



## この方法の限界

---

### ■ 相手ドメインへのスリバー追加・削除が困難

- ◆ 旧仮想化基盤にあるスリバー追加・削除のコマンドでは実現できない。
- ◆ ノードスリバー変更というコマンドがあれば実現可能 (仮想代理ノードを変更する) だが, 現在は存在しない  
(来年度リンクスリバー変更は実装予定だが, ノードスリバー変更の実装予定はない).

## まとめ

---

- フェデレーション機能をもたない仮想化基盤のフェデレーションのためのアーキテクチャを提案した.
  - ◆ 他ドメインを代表する“ドメイン代理ノード”を使用する.
  - ◆ “ゲートキーパー”が設定管理情報の他ドメイン部分をドメイン代理ノードでくくったり, くくりをといたりする.
- 提案したフェデレーション・アーキテクチャでおこりうるメッセージングの無限再帰を停止させる方法を示した.
- ドメイン間仮想リンク生成の実装法を示した.
  - ◆ “ゲートキーパー”がゲートウェイに設定する.
  - ◆ ゲートウェイでプロトコル変換をおこなう.
  
- 謝辞: この研究発表は委託研究「新世代ネットワークを支えるネットワーク仮想化基盤技術の研究開発」課題イの成果をふくむ.