# 経路制御(2) – ドメイン間の経路制御

情報ネットワーク論I

門林　雄基

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

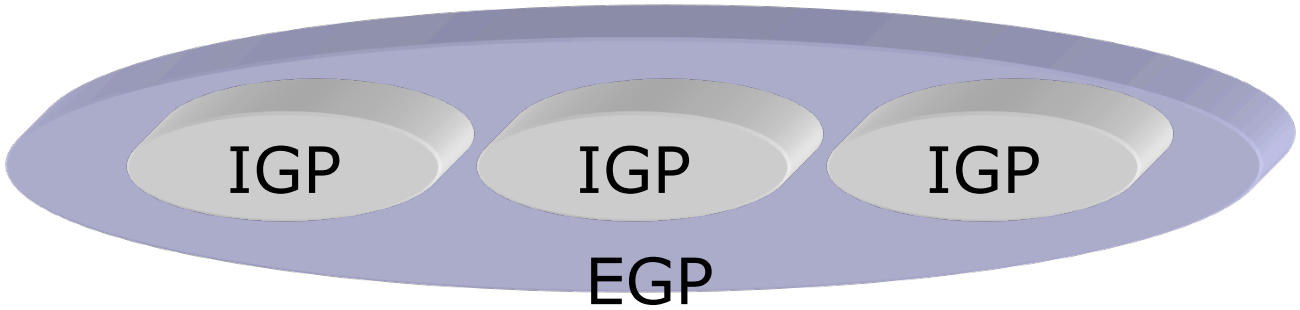
# アウトライン

- 距離ベクトル型経路制御
  Distance vector routing
- リンク状態型経路制御
  Link state routing
- IGP と EGP
  - ドメイン内経路制御、ドメイン間経路制御
- **パスベクトル型経路制御 (Path vector routing)**
- BGP: Border Gateway Protocol
- 経路の集約

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

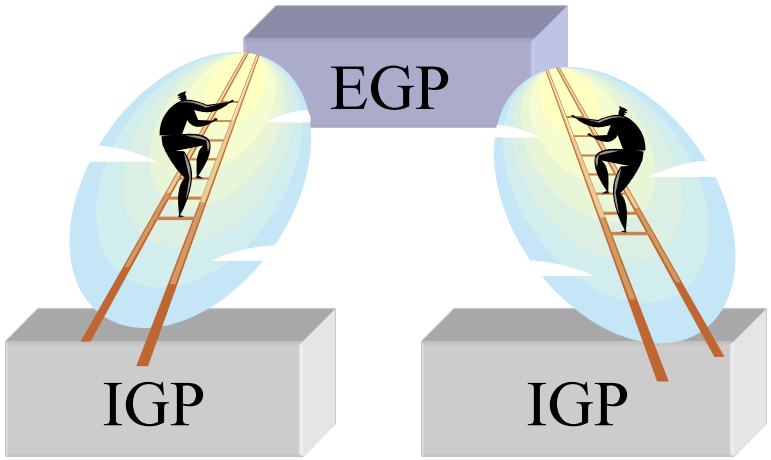
# 経路制御の階層化

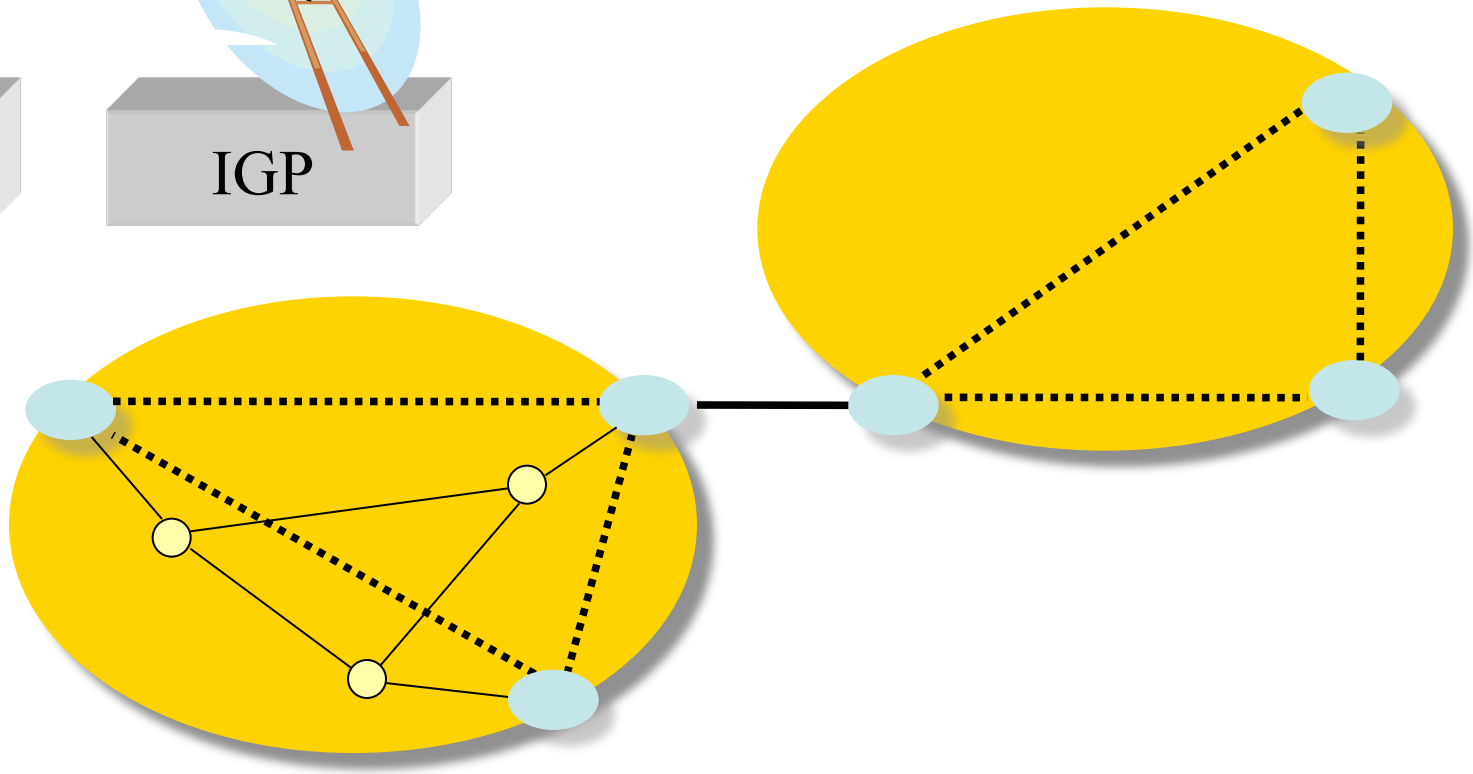
- ルーティングドメイン (routing domain)
  - 境界を定義
  - 障害の分離、経路の集約（後述）などの効用
- ドメイン内／ドメイン間経路制御プロトコルの区別
  - IGP (Interior Gateway Protocol)
  - EGP (Exterior Gateway Protocol)




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路制御の階層化 – 2階層での経路制御



©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路制御の階層化 – IGP と EGP

- IGP
  - RIP-2, RIPng: distance vector routnig
  - OSPF, IS-IS: link state routing
  - 各リンク・ルータの状態をできるだけ早く伝播させることに主眼
- EGP
  - BGP4, BGP4+: path vector routing
  - Internet 全体の経路制御の安定に主眼

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

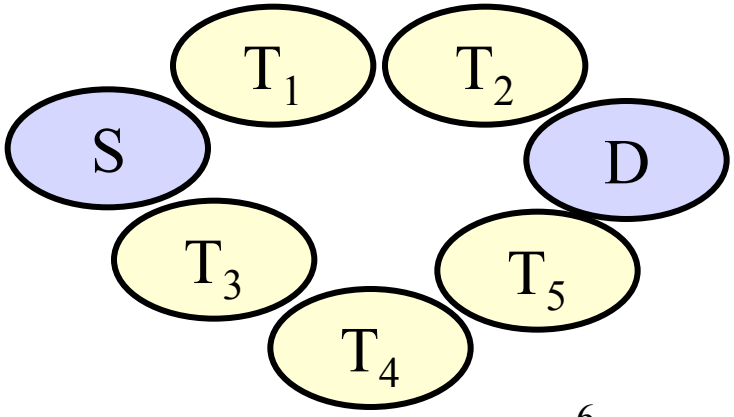
# パスベクトル型経路制御

- Bellman-Ford algorithm からの派生形
- 距離ベクトル型で交換する情報: (prefix, metric)
- パスベクトル型で交換する情報: (prefix, path, attributes)

- 距離だけでなくパス情報を経路情報に付与
  → ループの起きない経路制御が可能
  パスベクトル長の短い経路を優先

Sからみたへの経路の表現

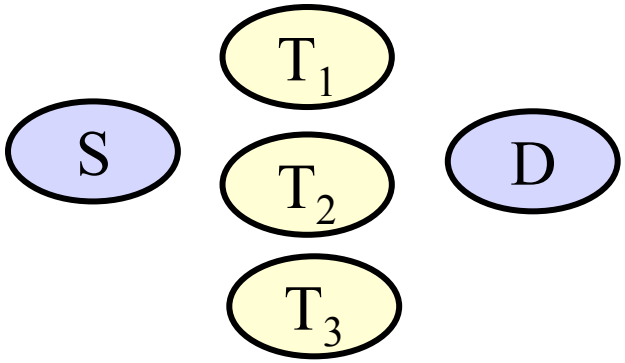
- (D, { $T_1$, $T_2$ }, ...)
- (D, { $T_3$, $T_4$, $T_5$ }, ...)




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# パスベクトル型経路制御：その背景

- 複数の代替経路
  - 事業者同士の密なつながり
  - どの経路を優先すべきか？
    - コスト、契約、負荷等による制約



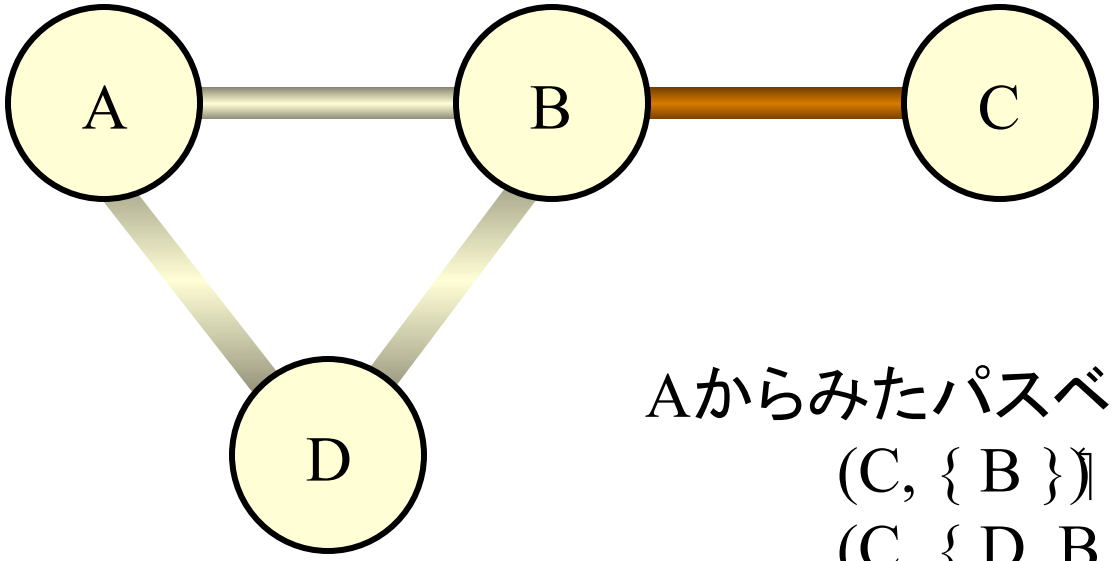

- 経路制御におけるポリシ
  - 経由する事業者の意向などを符号化
  - 代替経路からの経路選択方針を記述

  ➡ スカラー値に縮退できない

- ループのコスト
  - 過渡状態からの収束時間 ~ RTT

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# パスベクトル型経路制御におけるループ回避




Aからみたパスベクトル:

(C, { B })

(C, { D, B })

- リンクB-Cが落ちた場合、Bがパスベクトルを取り消し
- 自らが含まれるパスベクトルを棄却
- ループを回避

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# Q&A

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGP
# Border Gateway Protocol

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGP: Border Gateway Protocol

RFC 4271, 4760

- アルゴリズム
  - パスベクトル型
- トランスポート
  - トランスポート層プロトコルとしてTCPを利用
  - 再送、確認応答はTCPに委ねる

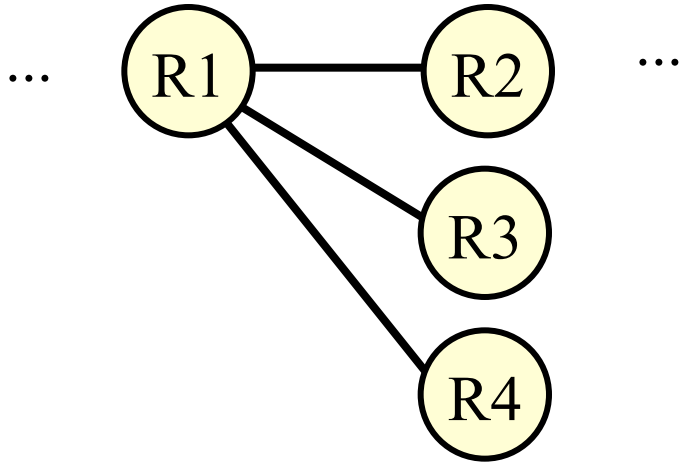
- **隣接関係と状態遷移**

- 経路情報
- トポロジ

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

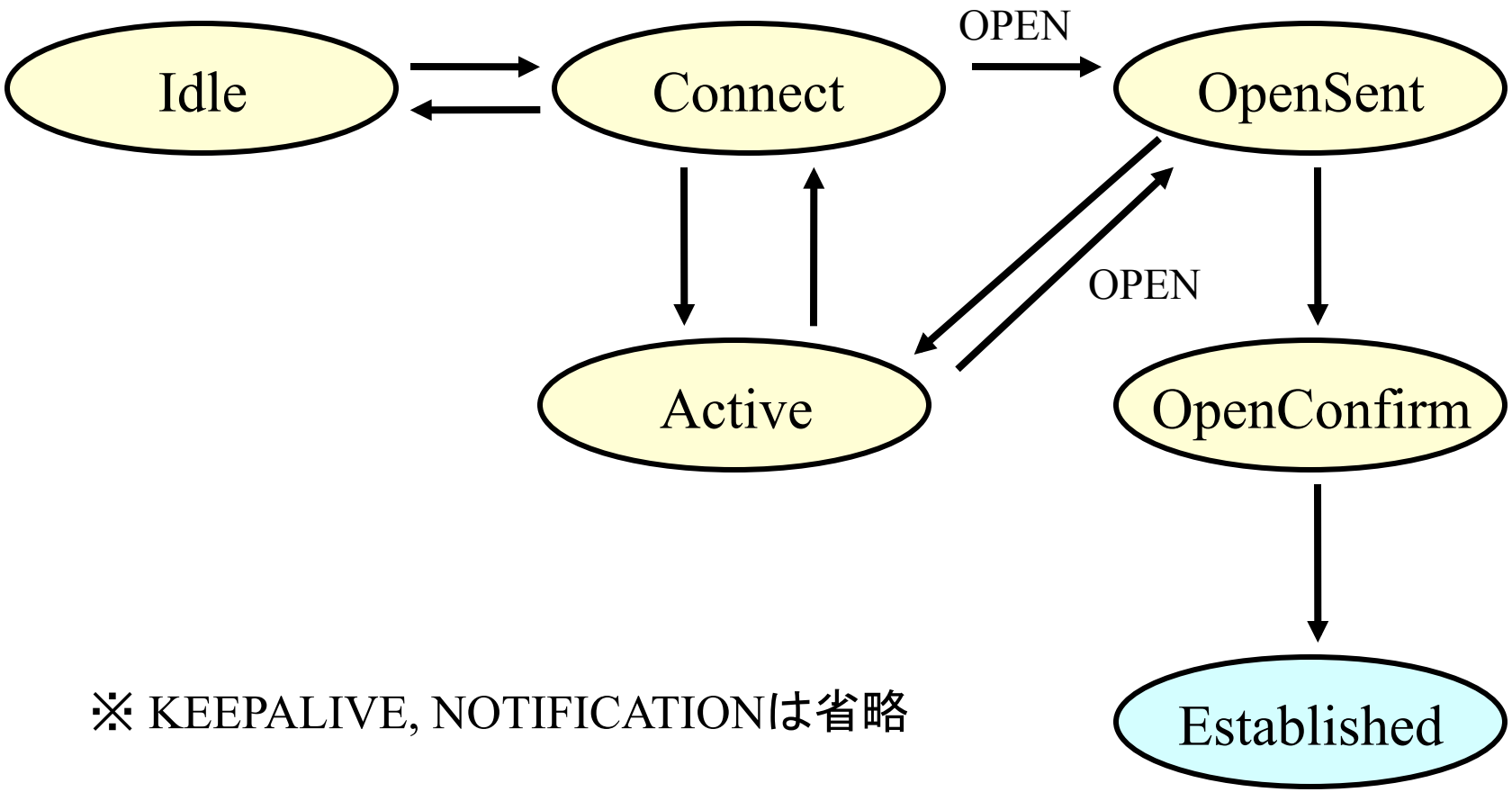
# BGPの隣接関係

- 隣接関係を運用者が明示的に指示
- 指示なき場合は隣接関係とならない
- なぜか？
  - C.f. OSPF ではパラメータが合えば隣接関係に




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGPの状態遷移: 隣接関係が確立するまで




※ KEEPALIVE, NOTIFICATIONは省略

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGP: Border Gateway Protocol

- アルゴリズム
- トランスポート
- 隣接関係と状態遷移

- **経路情報**
  - **効率的なパスベクトルの表現**

- トポロジ

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGPにおけるパスベクトルの表現

- AS (Autonomous System) を番号で表現
  - AS: 一貫したポリシを持ったルーティングドメイン
- (prefix, AS-path, attributes) をまとめて符号化
  - (AS-path, attributes, { prefix1, prefix2, ... } )
  - 通信量を削減

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGP におけるパスベクトルの例

Looking glass で
自分で見てみよう

```
show ip bgp 163.221.0.0
BGP routing table entry for 163.221.0.0/16, version 30149334
Paths: (6 available, best #3)
  Not advertised to any peer
  6461 2516 2500 2500
    208.185.175.169 from 208.185.175.169 (216.200.254.220)
      Origin IGP, metric 0, localpref 400, valid, external
  6461 2516 2500 2500, (received-only)
    208.185.175.169 from 208.185.175.169 (216.200.254.220)
      Origin IGP, metric 0, localpref 100, valid, external
  2500, (received & used)
    203.181.70.232 (metric 2) from 203.181.70.232 (203.181.70.227)
      Origin IGP, metric 505, localpref 400, valid, internal, best
      Originator: 203.181.70.227, Cluster list: 0.0.0.16
  2500, (received & used)
    203.181.70.233 (metric 2) from 203.181.70.233 (203.181.70.227)
      Origin IGP, metric 505, localpref 400, valid, internal
      Originator: 203.181.70.227, Cluster list: 0.0.0.16
  2500
    202.249.2.1 from 202.249.2.1 (203.178.136.4)
      Origin IGP, metric 1010, localpref 400, valid, external
  2500, (received-only)
```

# Q&A

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

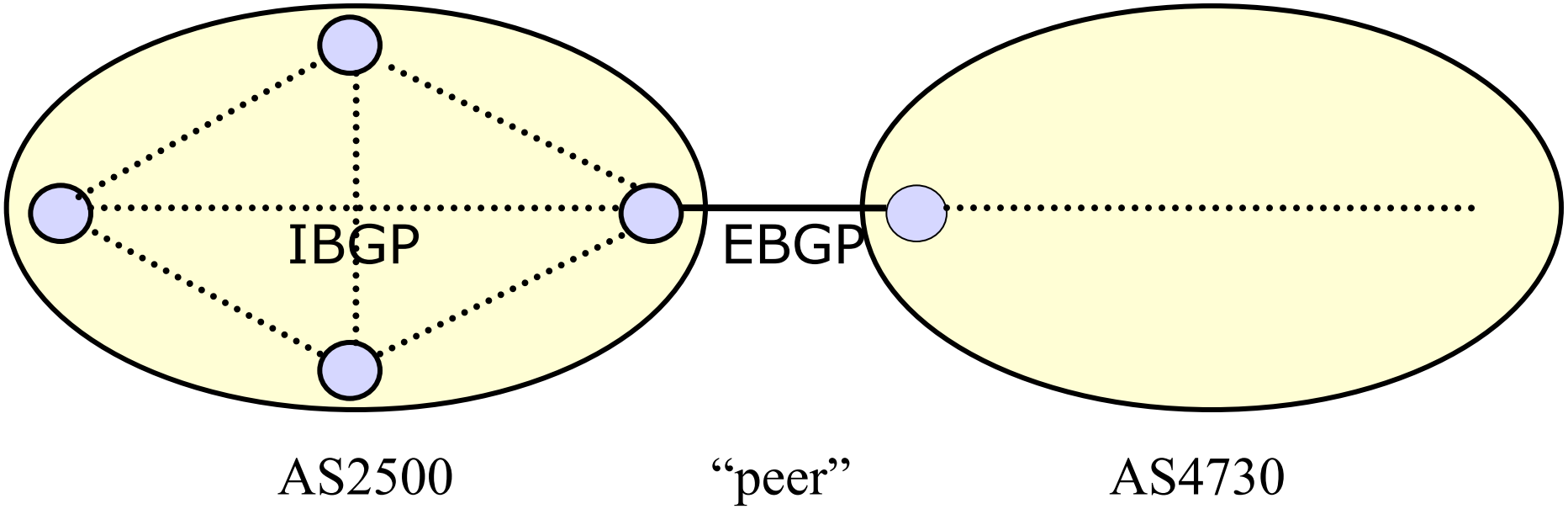
# BGP: Border Gateway Protocol

- アルゴリズム
- トランスポート
- 隣接関係と状態遷移
- 経路情報

- **トポロジ**

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGPのトポロジ: IBGP と EBGP




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# IBGP, EBGP におけるトポロジ制約

- IBGP
  - 同一AS内のすべてのIBGPルータとメッシュ状に peer を確立しなければならない
  - 物理的に隣接する必要はない
  - すべてのIBGPルータが同じ情報を持つ
- EBGP
  - 隣接ASのEBGPルータと物理的に隣接しなければならない（原則として）
  - 隣接ASと同じ情報を持つ必要はない

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# Q&A

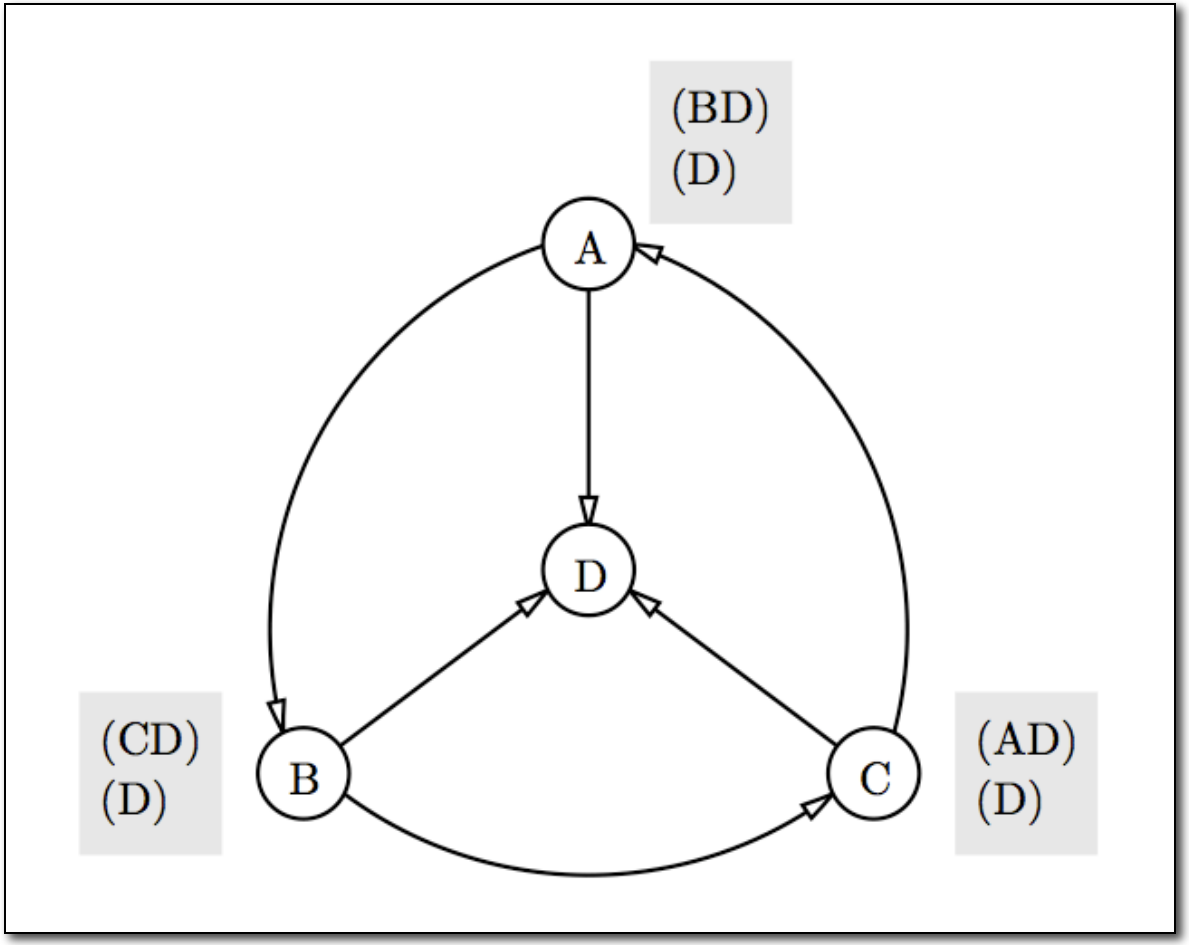
©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGPにおけるポリシ

- 経路制御におけるポリシ
  - 経由する事業者の意向などを符号化
  - 代替経路からの経路選択方針を記述
- 他事業者に伝えるポリシ
  - MED
- 事業者内でのポリシ
  - Local preference, admin distance, route map, etc.

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# BGPにおけるポリシの限界: dispute wheel




箱は各ノードでの
パスの優先順位

Dに辿り着けるか？

N. Feamster et al., "Implications of Autonomy for the Expressiveness of Policy Routing", SIGCOMM'05.

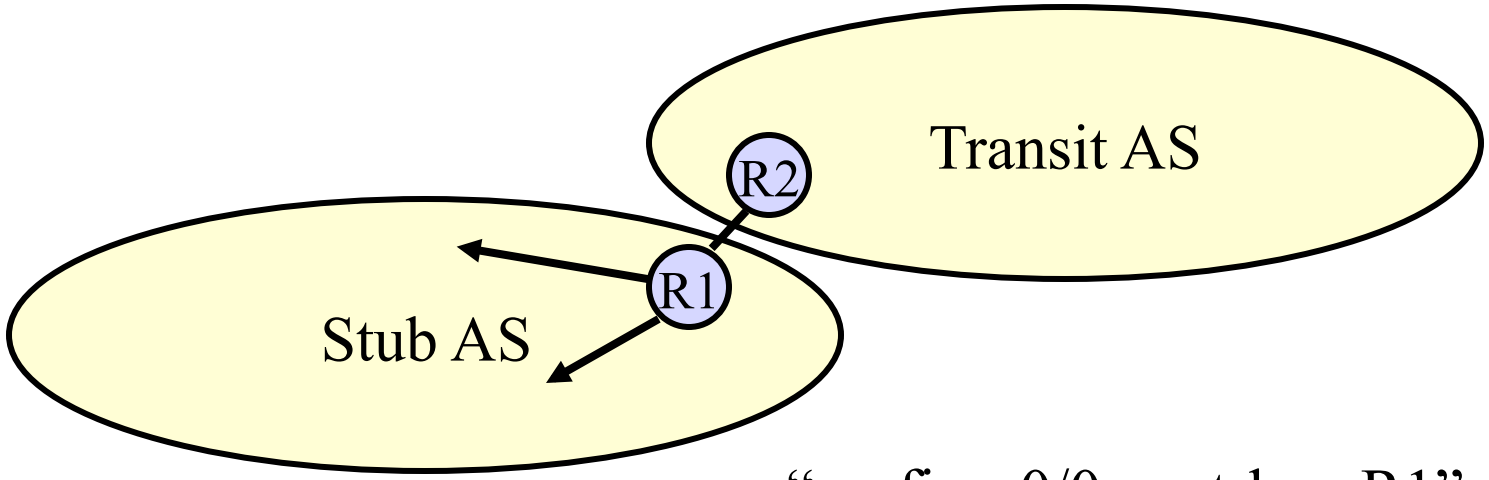
©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路の削減と集約
# Reduction and aggregation of routes

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路の削減: Default route

- 0.0.0.0/0 (IPv4), ::0/0 (IPv6)
- 最長一致 → 経路検索の最後にマッチ
- 経路の隠蔽、経路数の削減に効果

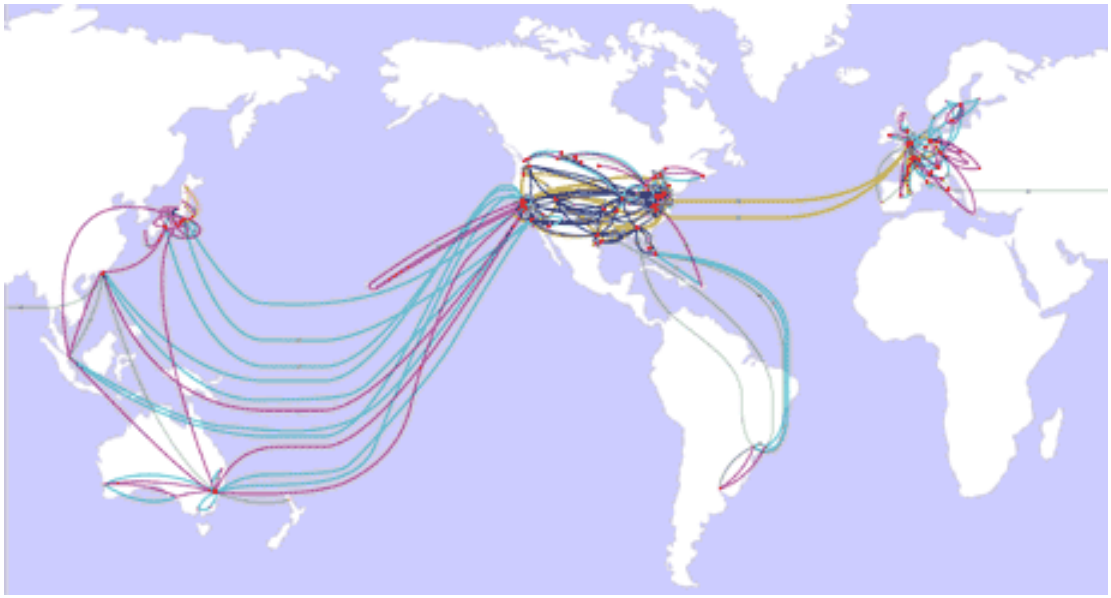



“prefix ::0/0 next-hop R1”

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路の削減ができない状況とは

- Default-free
  - Default routeがない状態
- Tier-1 ISPs, North America backbone

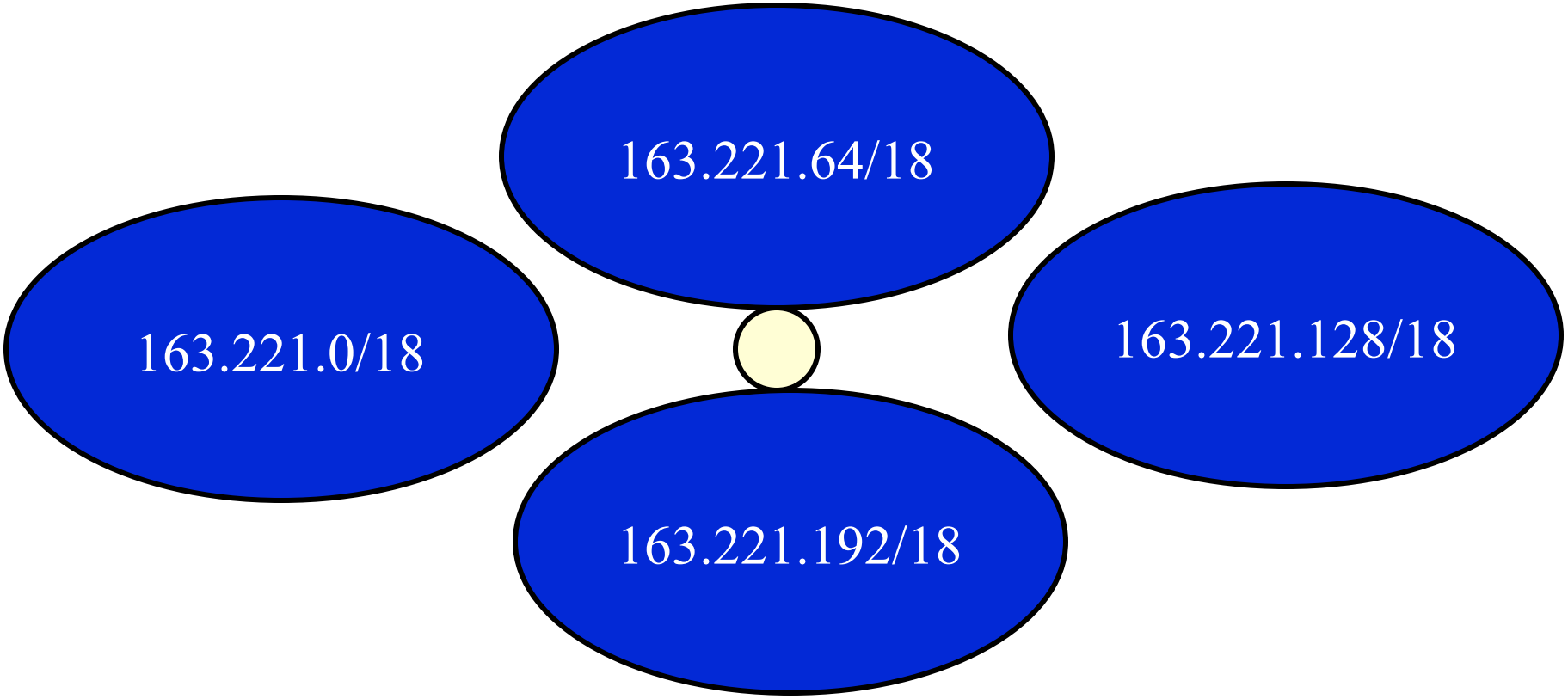


(source: UUNet network maps, www.uu.net)

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路の集約 (aggregation)

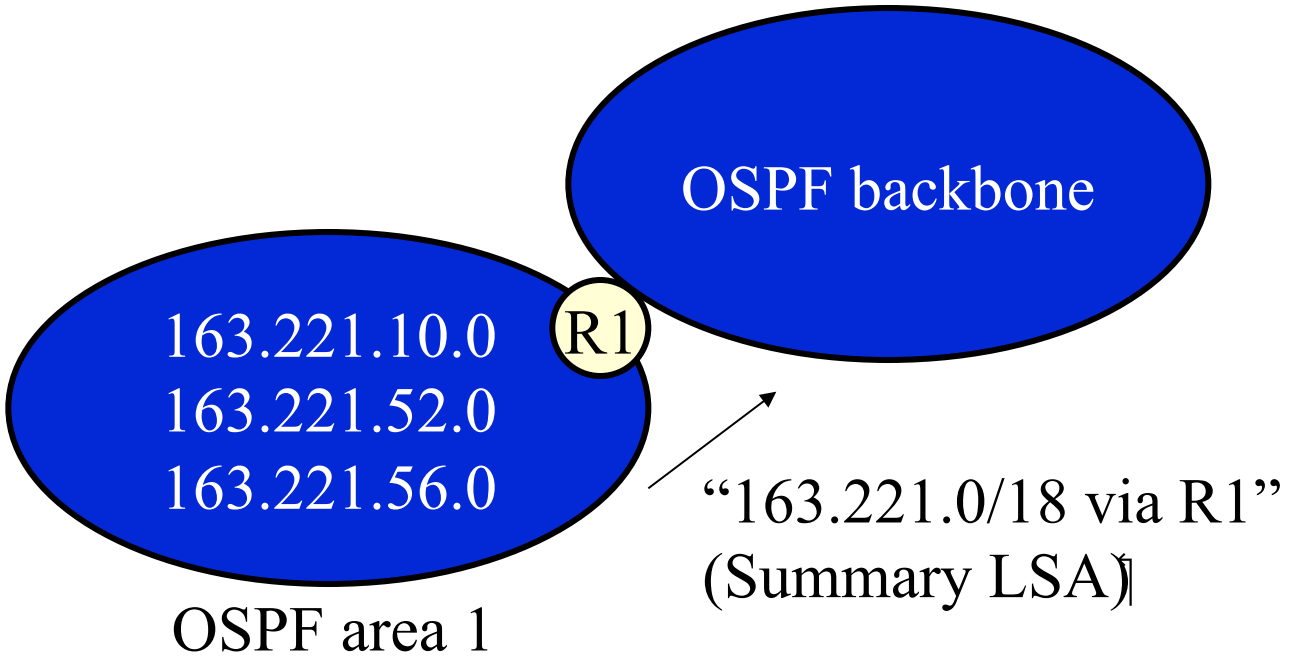
- 163.221.10.0/24 と 163.221.11.0/24 は 163.221.10.0/23 で表現できる




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

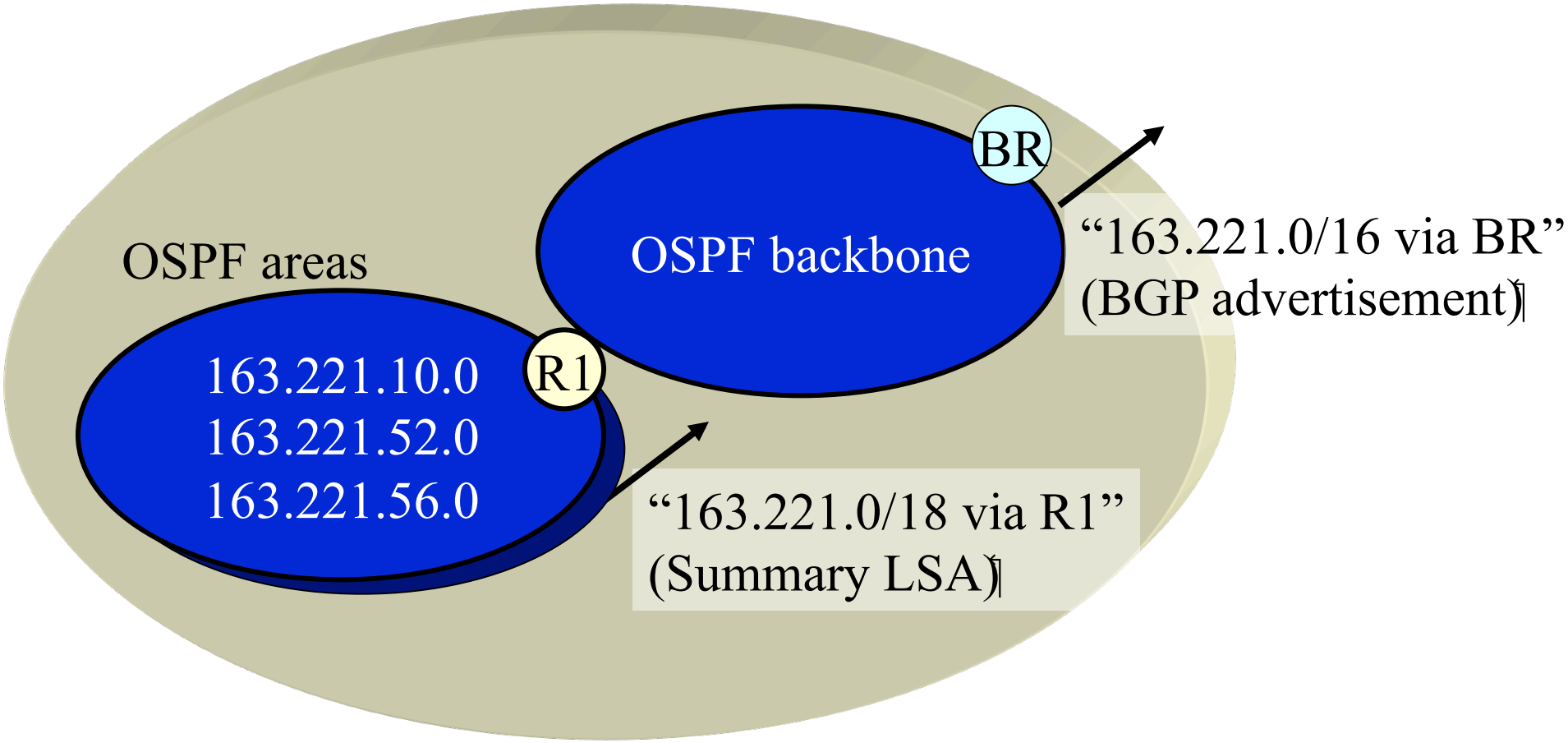
# 境界における経路の集約

- 経路の集約はルーティングドメインの境界で




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 階層的な経路の集約




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 経路の集約における問題点

- 集約できるかどうかはアドレス割り当て次第
  - 計画的なアドレス割り当てが重要
- 奈良先端がn研究科になると予想できたか？
  - あるISPが急成長すると予想できたか？

- → prefix renumbering
  - 集約できるアドレスへ付け替え
  - 技術開発がおこなわれている
- 経路集約は手作業でおこなう必要がある
  - インターネット全体で250,000経路
  - BGP table growth trends - Telstra

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# Q&A

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

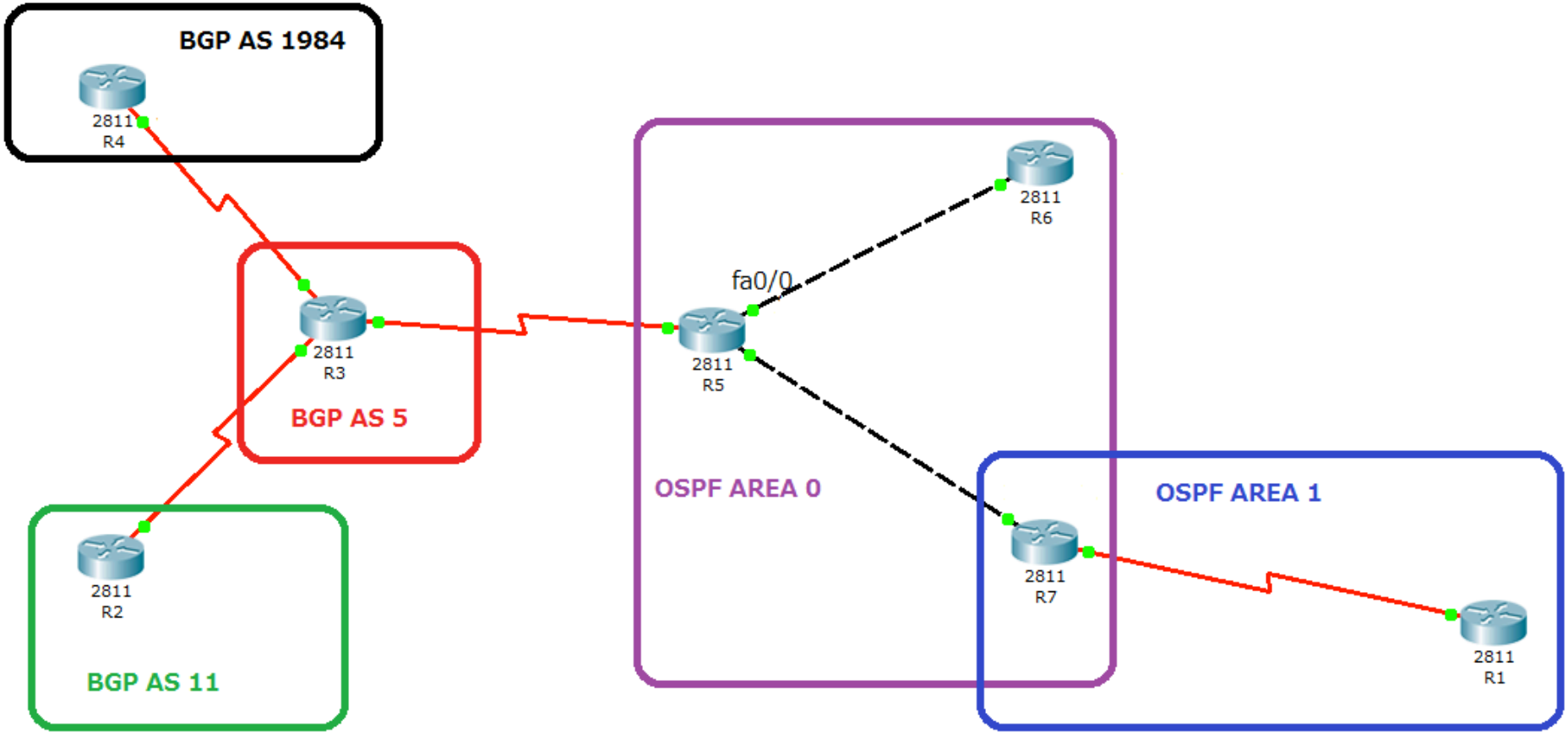
# まとめ

- 階層化経路制御の考え方
  - IGP, EGP
- パスベクトル型経路制御
  - Loop-free, policy-aware
- BGP
  - 状態遷移、経路情報、トポロジ
  - ポリシの適用限界
- 経路集約
  - 考え方、問題点

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 課題

- 目標: BGPとOSPFがどのように動いているのかを理解する




©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 課題

- 作業:
  - PacketTracerファイルを以下のリンクからダウンロード
    - http://iplab.naist.jp/class/infoN/2012/materials/BGP-OSPF.pkt
  - 前ページの図のようにBGPを設定
  - 前ページの図のようにOSPFを設定
    - ヒント:最初はAREA 0を設定しよう

＊IP等の基本的な設定はしてあります

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.

# 課題

- 提出方法
  - 2012/5/16 (水) 17:00 (JST)
  - PacketTracerファイル名を以下のように指定

    BGP-OSPF_YourNameYourStudentID

    (例: BGP-OSPF_DoudouFall1234567.pkt)

  宛先:
    - network1-2012_at_ is.naist.jp
- 質問はメール(上と同じ)で受け付けます

©2012 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.