評価方法

- 中間レポートと、期末レポート
- 出席はとらないが、、、
- 質問やコメントを義務付ける
  - 期中、講義に関する技術的な内容の質問やコメントを最低2回、授業中に行うこと
  - よい質問やコメントは、成績の加点対象
  - 質問者は、講義終了後に名前と学籍番号を申告のこと

インターネットインフラ特論
８．ルーティング：IGP、ポリシー、マルチホーミング、モビリティ

太田昌孝

mohta@necom830.hpcl.titech.ac.jp

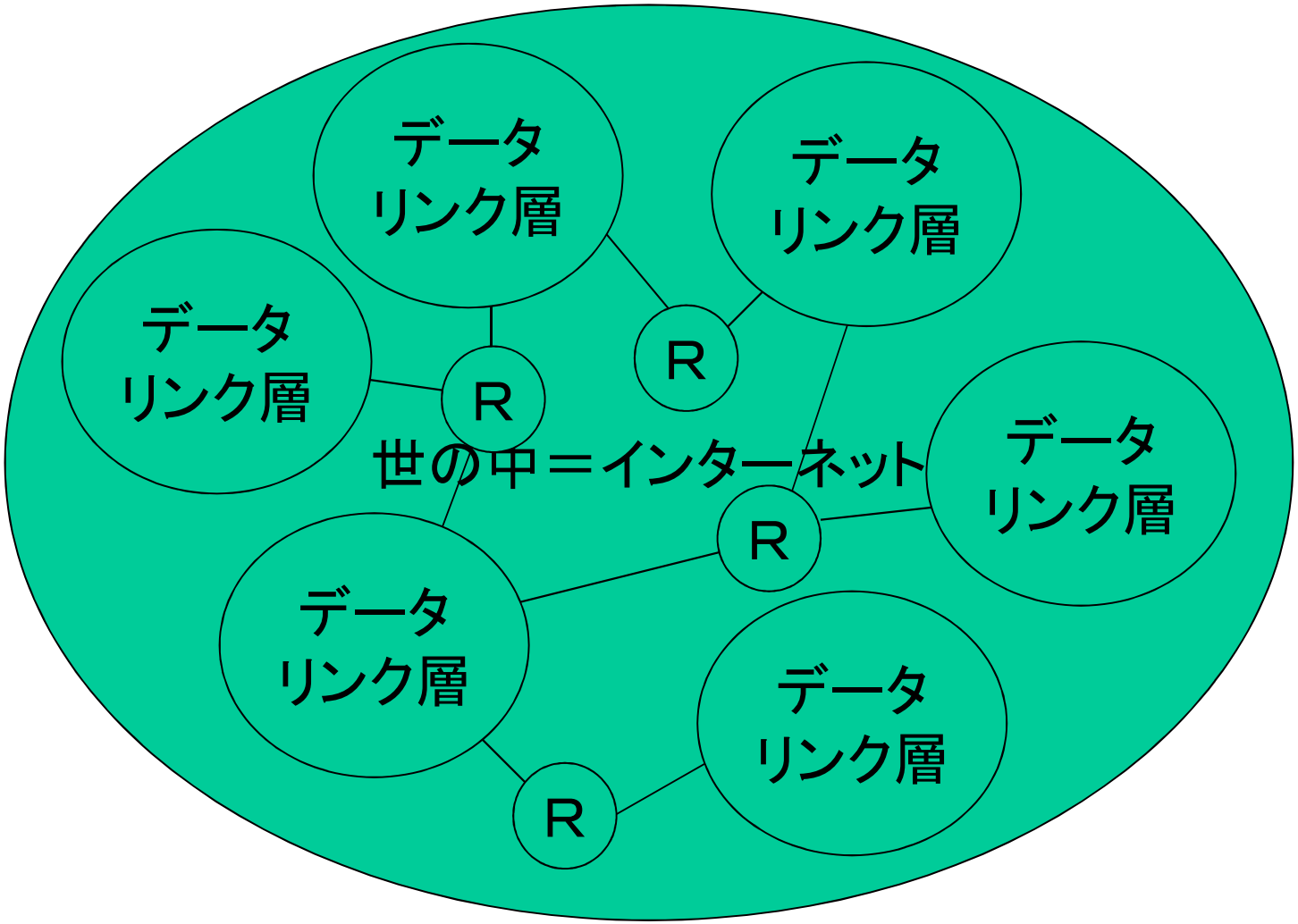
ftp://chacha.hpcl.titech.ac.jp/infra8.ppt

ルーティングとは？

- パケットを経路表に従って中継すること
  - インターネットの経路表は、ルーティングプロトコルにより自動生成
  - 障害部分は自動的に避ける
- ネットワークの知性である
  - エンドツーエンド原理違反？
  - 必用最低限はしょうがないが、、、

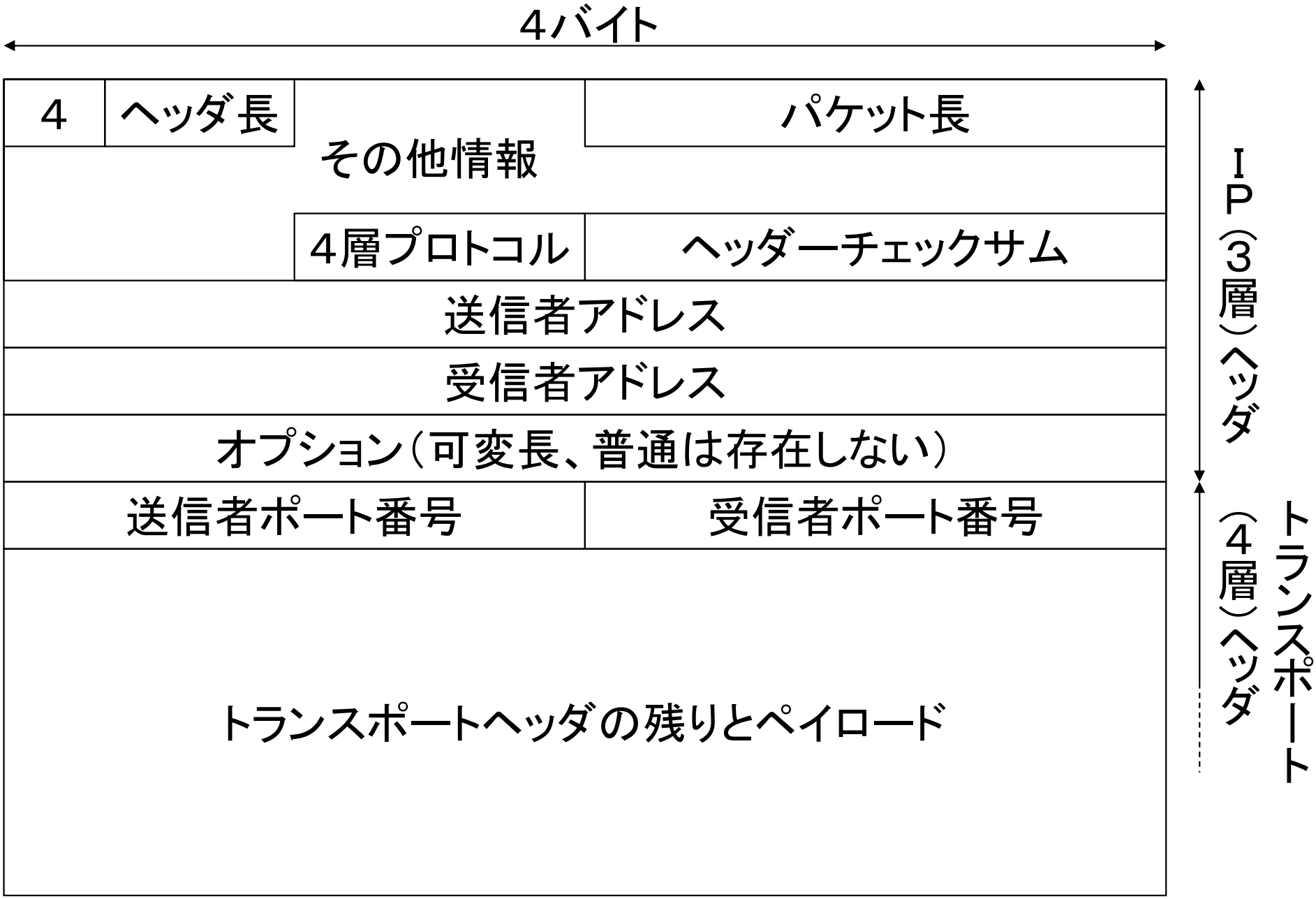
インターネットの構造

- CATENETモデル
  - 多数の小さな（機器の数が少ない）データリンク層をIP（Internet　Protocol）ルータで相互接続したもの



R　：ルータ

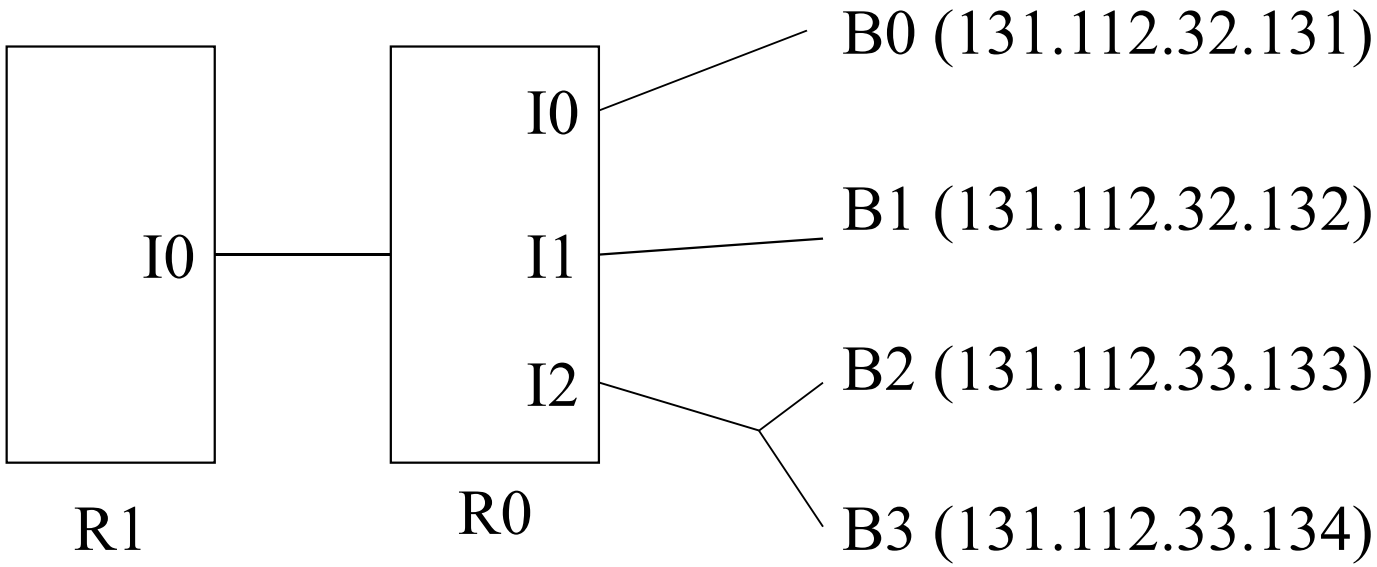
CATENETモデル



IPv4パケットフォーマット

経路表 （Routing Table）

- ルータは経路表の検索結果に基づきパケットを次のルータに送出
  - 経路表は受信者アドレスで引く
- ある地域のホストが似たIPアドレスを持っていれば、経路表のエントリは1つで済む
  - 経路の縮約（Route Aggregation）
  - 経路表はアドレス範囲ごとに1エントリ



R0の経路表

目的地	次
131.112.32.131	I0
131.112.32.132	I1
131.112.33.*	I2

R1の経路表

目的地	次
131.112.*	I0

経路表の縮約

インターネットはなぜ定額か？

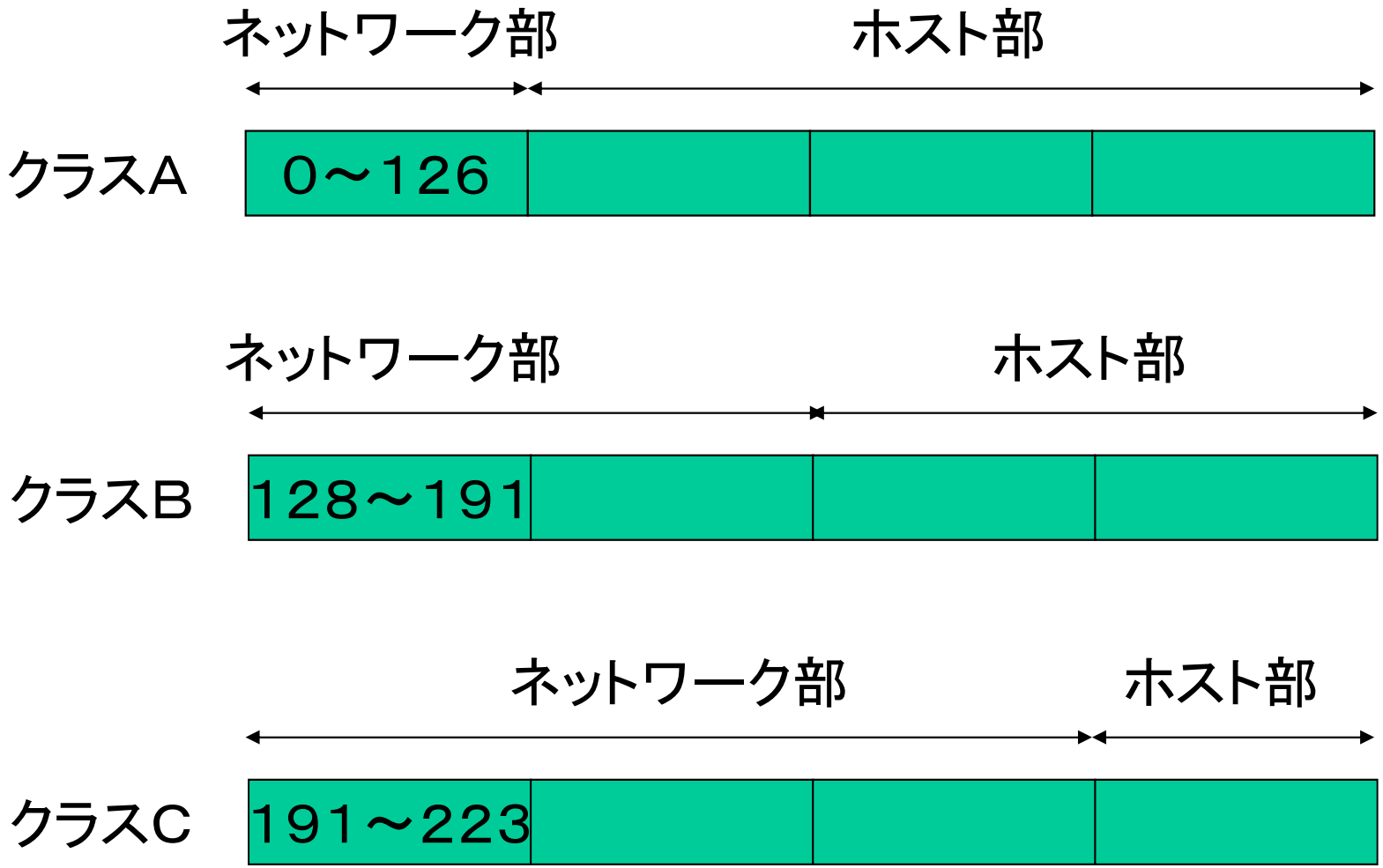
- 幹線が十分高速だから？
  - 幹線が遅い時代でも定額制であった
- 従量制課金は却って高くつくから？
  - 電話網では実際間接経費がほとんど
- 品質保証がないから？
  - 電話網でも品質は通話相手しだい

インターネットが定額なのは各通信が資源を占有しないから

- 事前の通信路設定（シグナリング）なし
- パケット単位に処理は完全に独立
  - 各パケットは行き先アドレスを持つ
  - 各ルータは行き先アドレスで経路表を引く
- 経路表のエントリは通信ごとには不要
  - 経路表エントリも有限な資源
  - 目的地ごとにすら不要
- 資源の占有には従量制課金が必要

クラス別ルーティング

- IPv4アドレスを5つのクラスに分類
  - クラスA、B、Cはユニキャストに
    - クラスDはマルチキャスト、Eはリザーブ
- ユニキャストIPアドレスを前半（ネットワーク部）、後半（ホスト部）にわけ、ネットワーク（データリンク）単位でルーティング
  - ホスト部全部1はネットワーク内ブロードキャストのアドレス
  - ホスト部全部0はネットワーク自体のアドレス



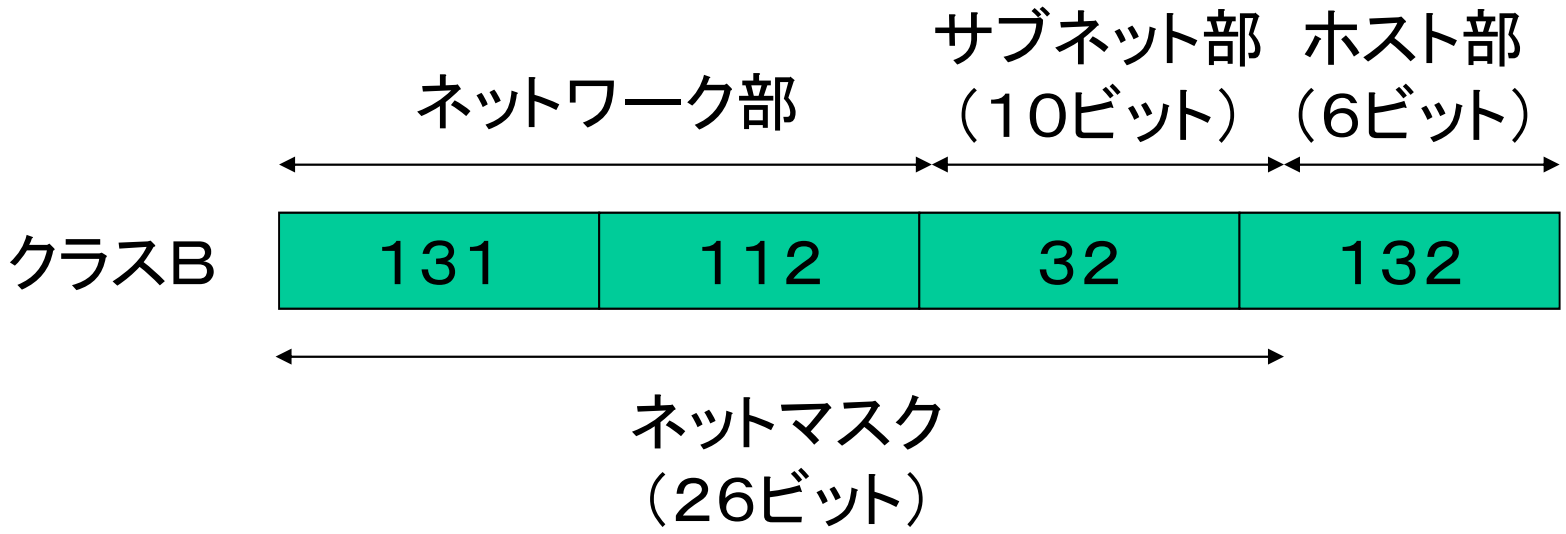
クラス別のIPアドレスの構造

クラス別ルーティングの問題点

- 個別データリンク内のホスト数は多くても数十程度
  - クラスCでも大きすぎる
    - ルート情報の不必要な増大
    - IPv4アドレスの不必要な消費
- IPv4アドレス構造の細分が必用
  - サブネット

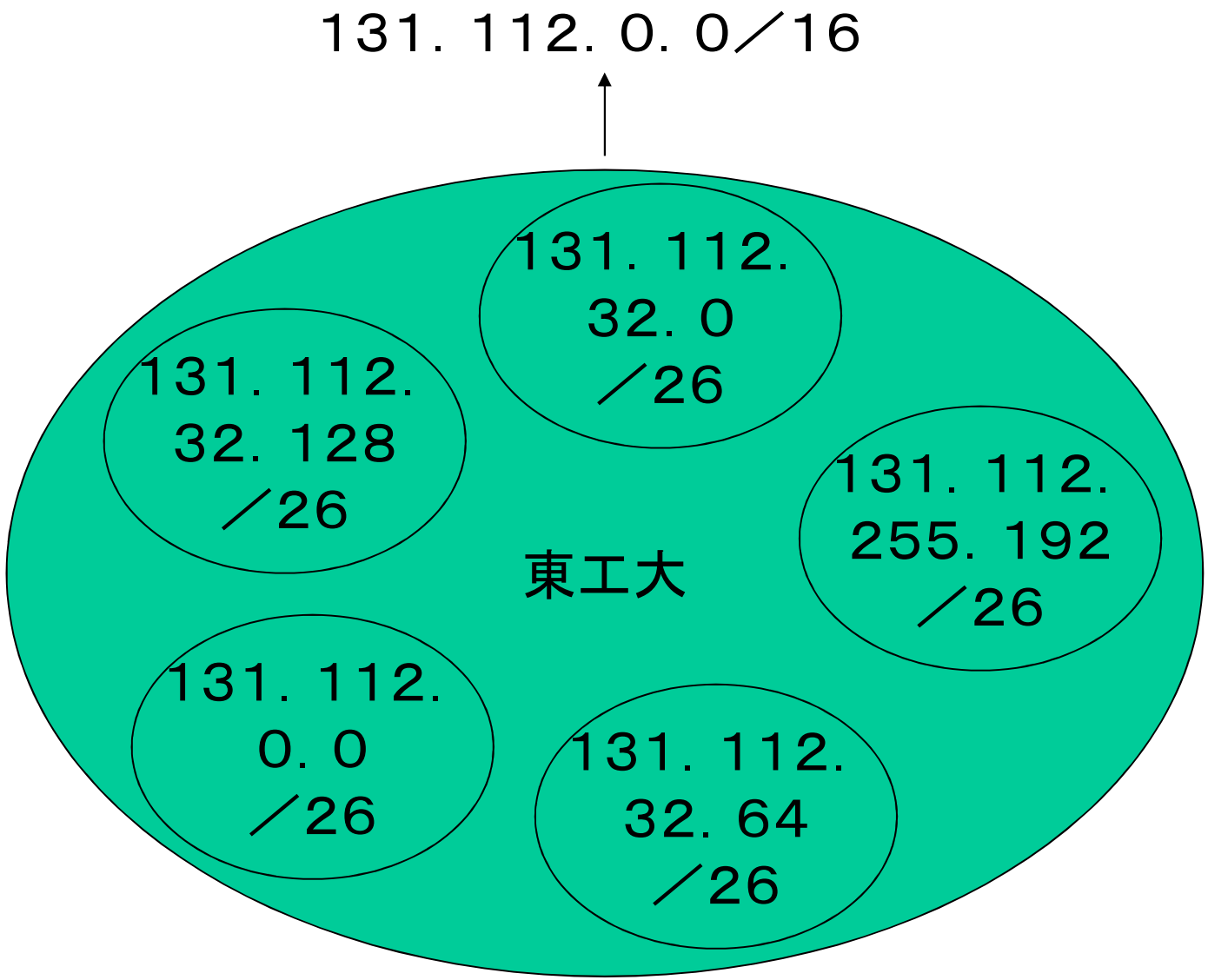
サブネット

- ユニキャストIPアドレスのホスト部を前半（サブネット部）、後半（ホスト部）にわける
- ネットワーク内ではサブネット（データリンク）単位でルーティング
  - １組織にクラスBアドレス1つでほぼ十分
- ネットワーク外ではネットワーク単位でルーティング
  - 外部にはルート情報は1つしかみえない



131．112．32．128／26

サブネット化IPアドレスの構造の例（東工大）

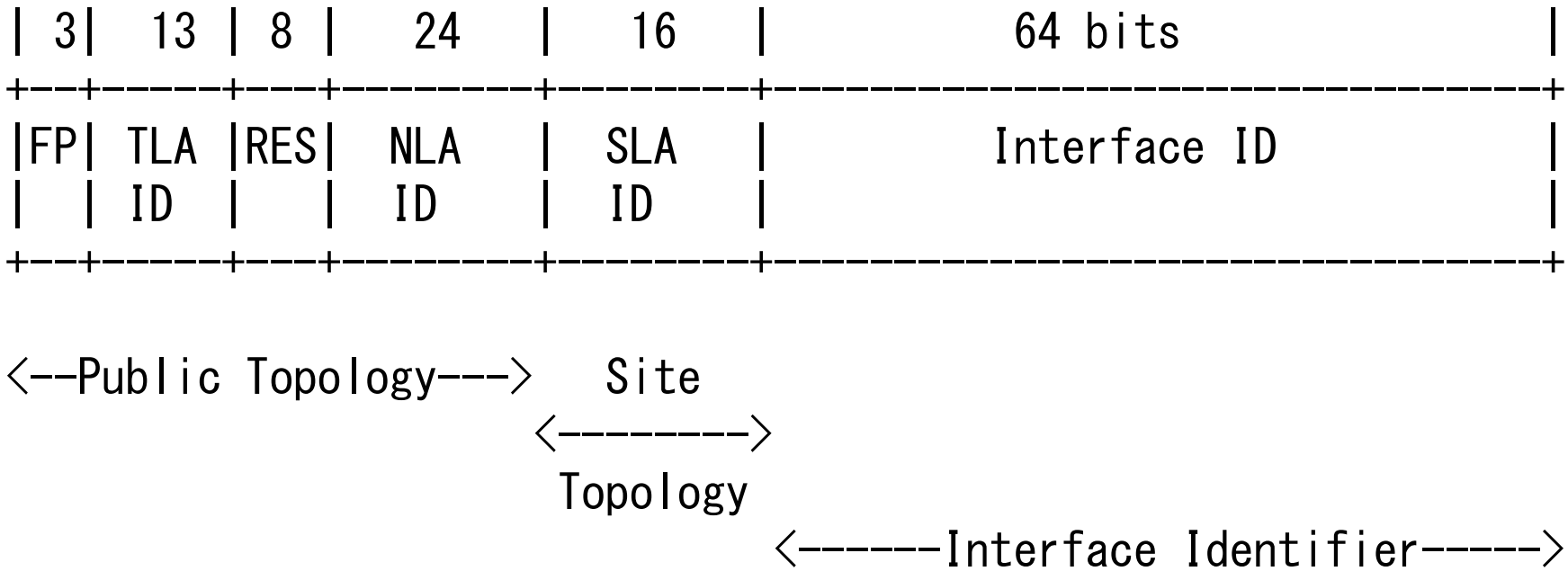

131. 112. 0. 0／16
131. 112.
32. 0
／26
131. 112.
32. 128
／26
東工大
131. 112.
255. 192
／26
131. 112.
0. 0
／26
131. 112.
32. 64
／26

CIDR （Classless InterDomain Routing）（RFC1519）

- クラスA、B、Cの区別を撤廃
  - ルーティングプロトコルはネットマスク長も運ぶ
- 階層的アドレス割り当ての例
  - ISPにアドレスを256個割り当て
    - ISP外部では／24でルーティング
  - ISPは各顧客にアドレスを8個づつ割り当て
    - ISP内部では32個の／29を個別にルーティング

IPv6アドレスのもともとの構造

- 強い階層構造
- ISPレベルで2段階
  - TLA （Top Level Aggregater）
  - NLA （Next Level Aggregater）
- 個別加入者（Subscriber）は65536個のサブネットをもてる
  - （Subscriber Level Aggregater）
- 各サブネット内は64ビットのアドレス


| 3| 13 | 8 | 24 | 16 | 64 bits |
|FP| TLA ID |RES| NLA ID | SLA ID | Interface ID |
<--Public Topology--->
Site
<-------->
Topology
<------Interface Identifier----->
IPv6アドレスの構造

ルーティングプロトコル

- 経路表自動作成のプロトコル
- DV（Distance　Vector）とLS（Link　State）の2方式に大別
- IGP（Interior　Gateway　Protocol）とEGP（External　Gateway　Protocol）に大別
- RIP（RFC1056）、OSPF（RFC2328）、BGP（RFC1771）、、、

DV方式

- ルータは自分に直結するネットワーク（サブネット）のルート情報を距離とともに発信
- ルート情報を受け取ったルータは、距離を増やして他のルータに中継
  - 同じネットワークへの複数のルート情報を受け取ったルータは、距離の小さいほうだけを他のルータに中継
    - 最短距離の分散計算

DVの特性

- 各ルータの計算量が少ない
  - 分散計算のため
- ルート情報が変化した場合の対応が遅い
  - ループがあると特に

LS方式

- ルータは自分に直結するネットワーク（サブネット）と他のネットワークとの接続状況を発信
- 接続情報を受け取ったルータは、そのまま他のルータに中継
- 各ルータは最短経路を個別に計算
  - エンドツーエンド原理に忠実

IGP

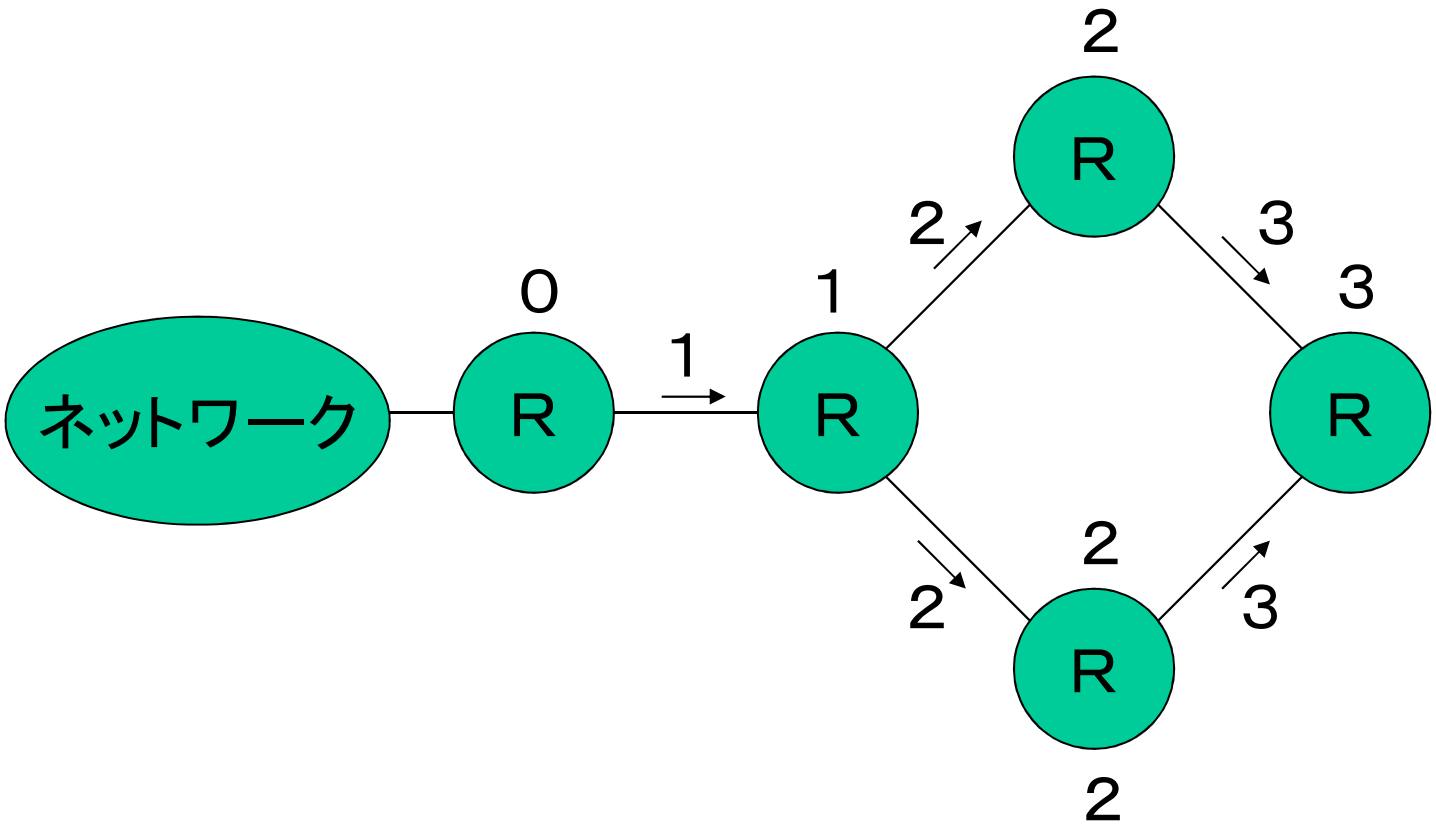
- お互いが完全に協力的な部分で利用
  - 組織内、ISP内等のルーティング
- 相手のトラフィックはよろこんで運ぶ
- 最短経路を選べばよい

EGP

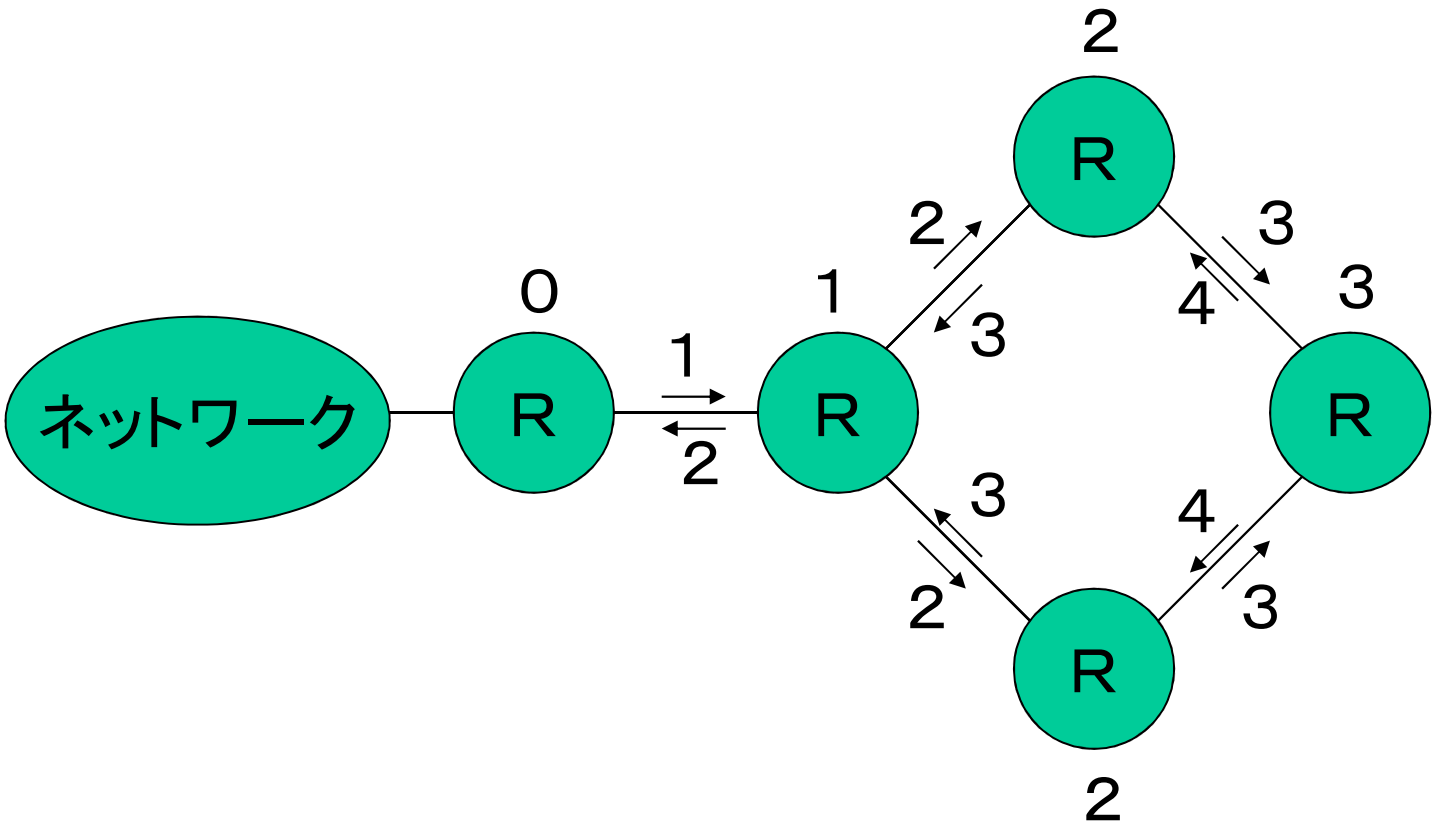
- 必ずしも利害が一致しない組織間のルーティング
- 相手のトラフィックを運ぶかどうかは交渉（バーター、お金等）しだい
- ポリシーによって経路を選択
  - といっても、あまり細かいことはやってられない

RIP （Routing Information Protocol）

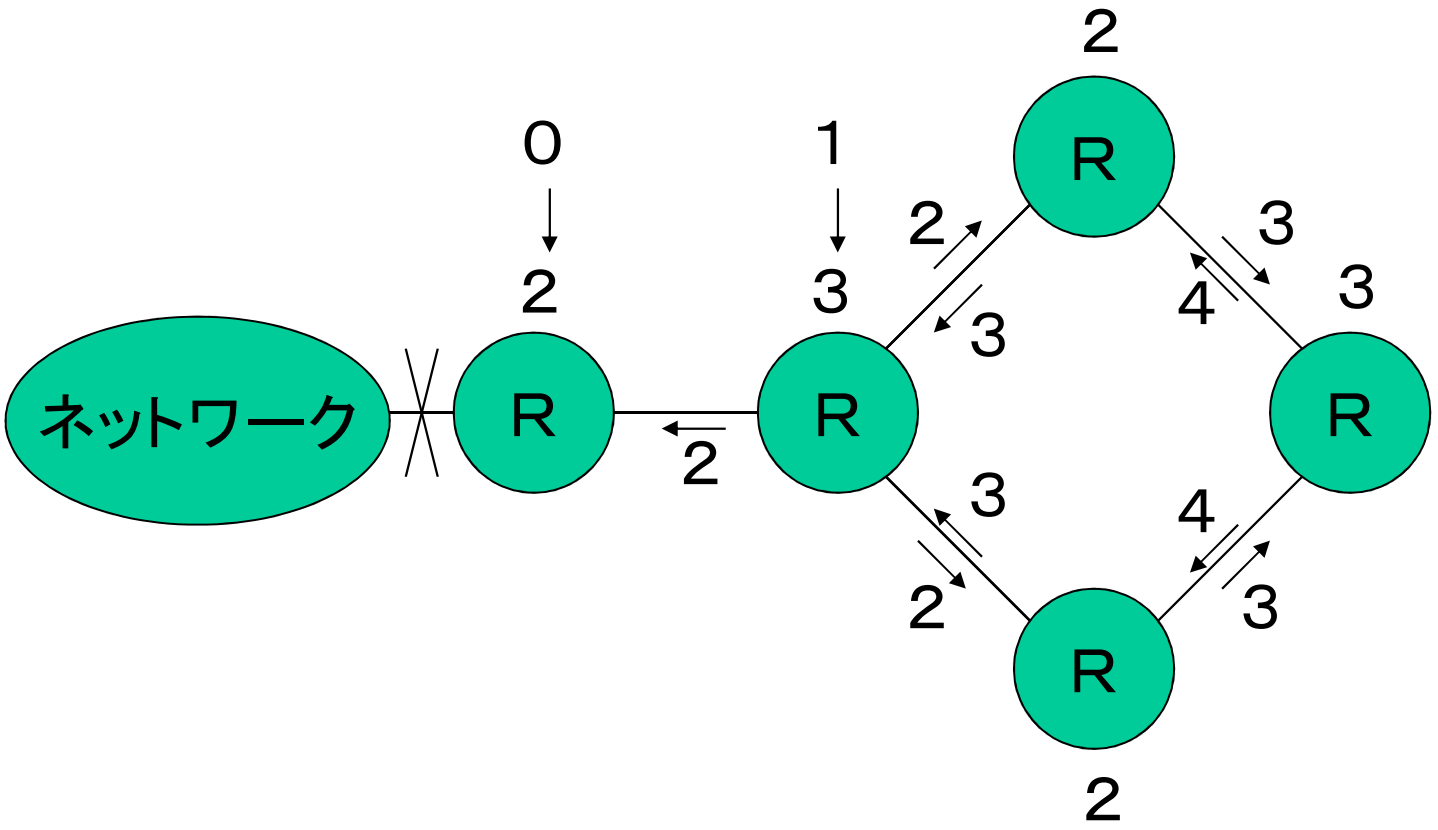
- DV型のIGP
- 古い
  - CIDR未対応
  - RIPv2（RFC2453）はCIDR対応
- 距離は0～15の整数（15は無限大）
  - 昔は15で全世界インターネットが覆えた



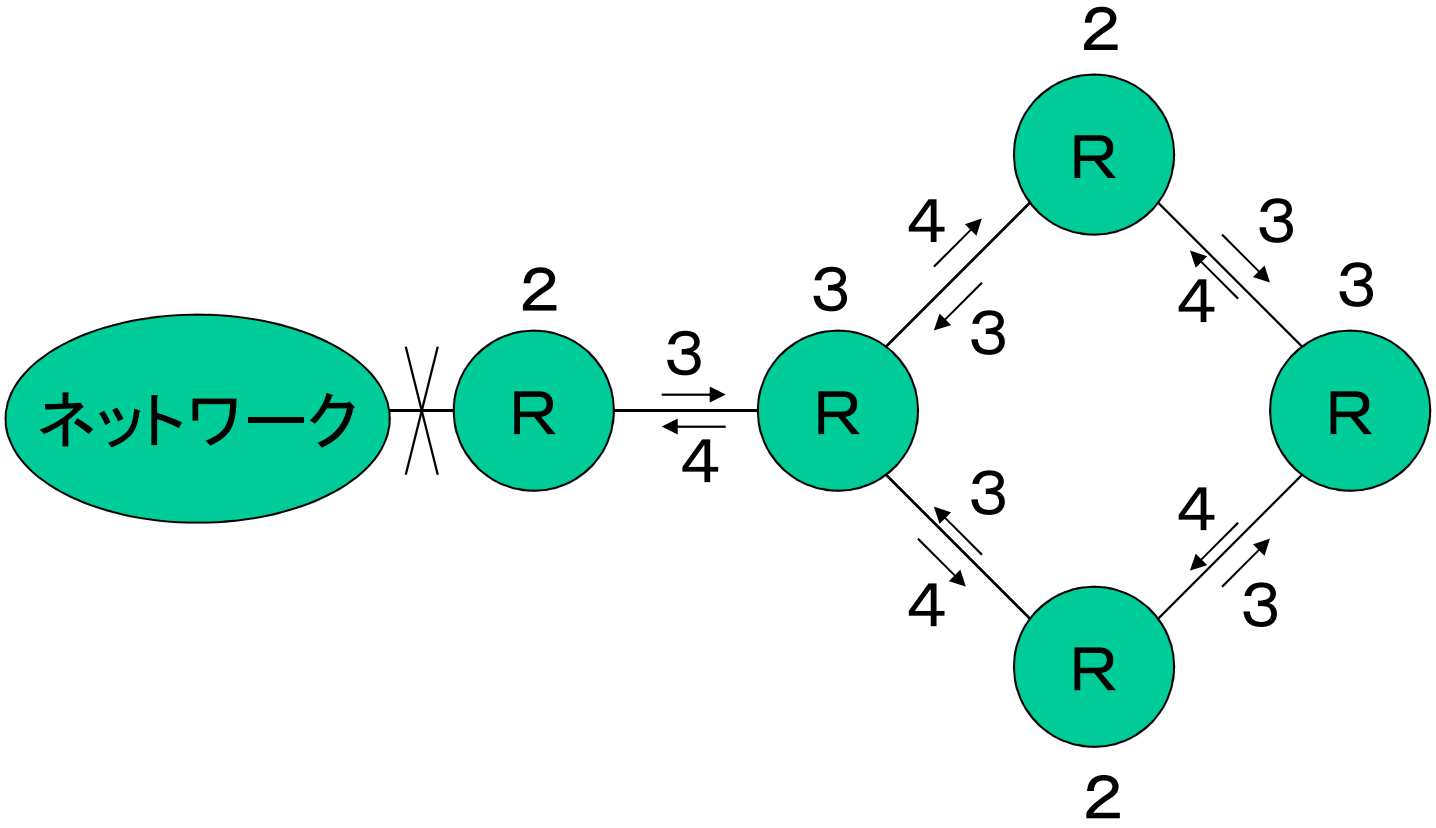
RIPの距離の伝搬



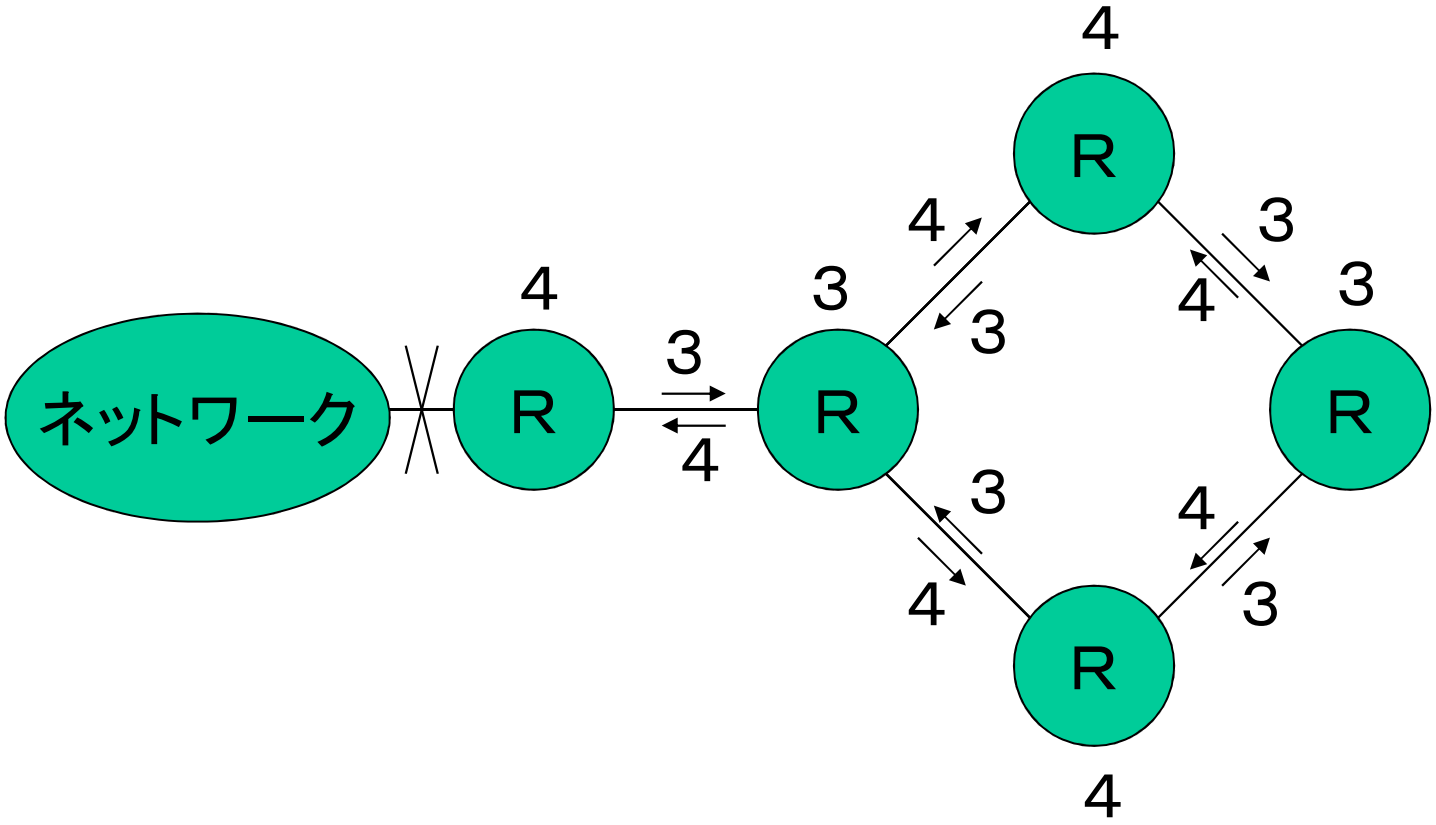
RIPの距離の伝搬



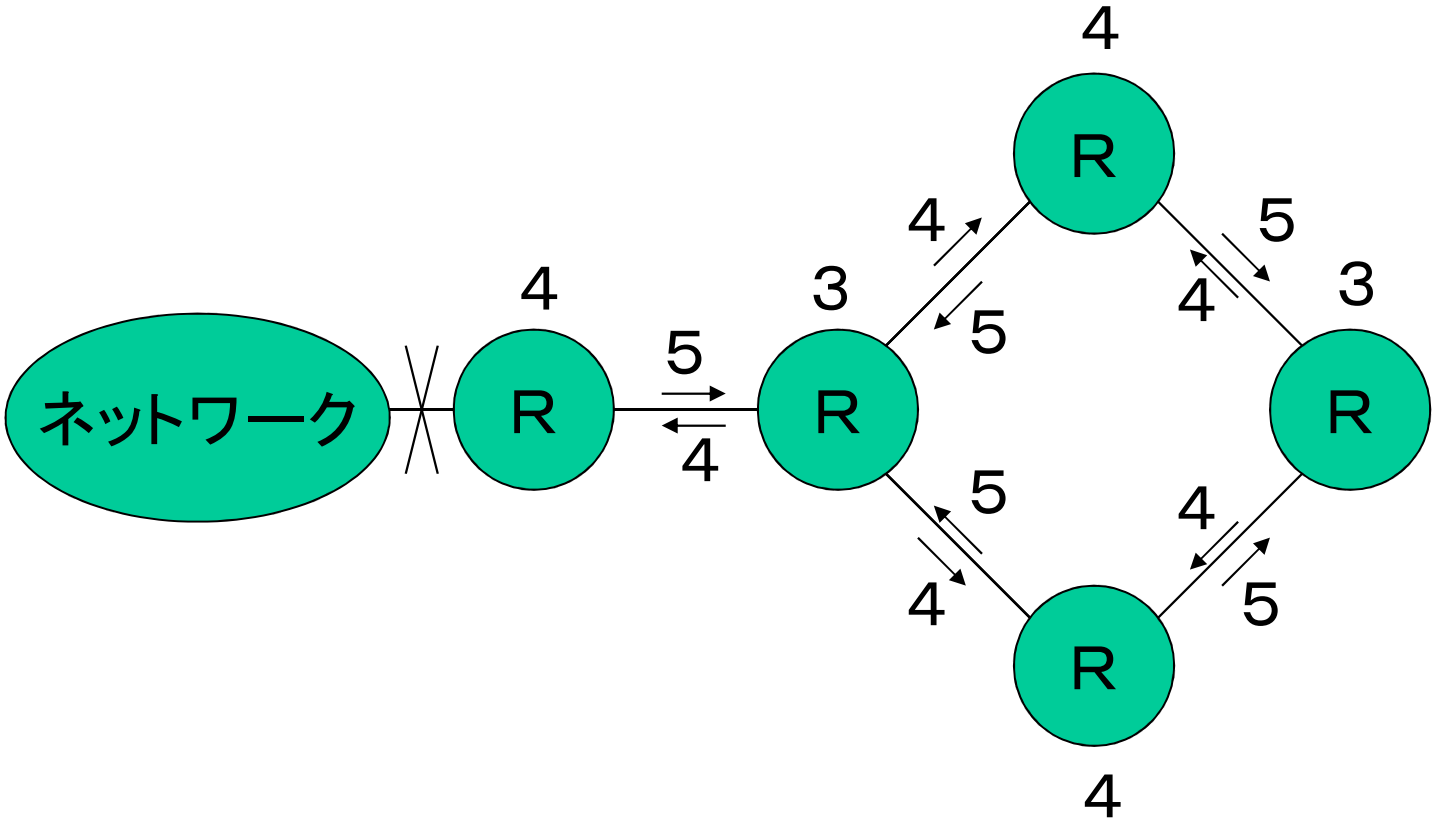
RIPの距離の伝搬



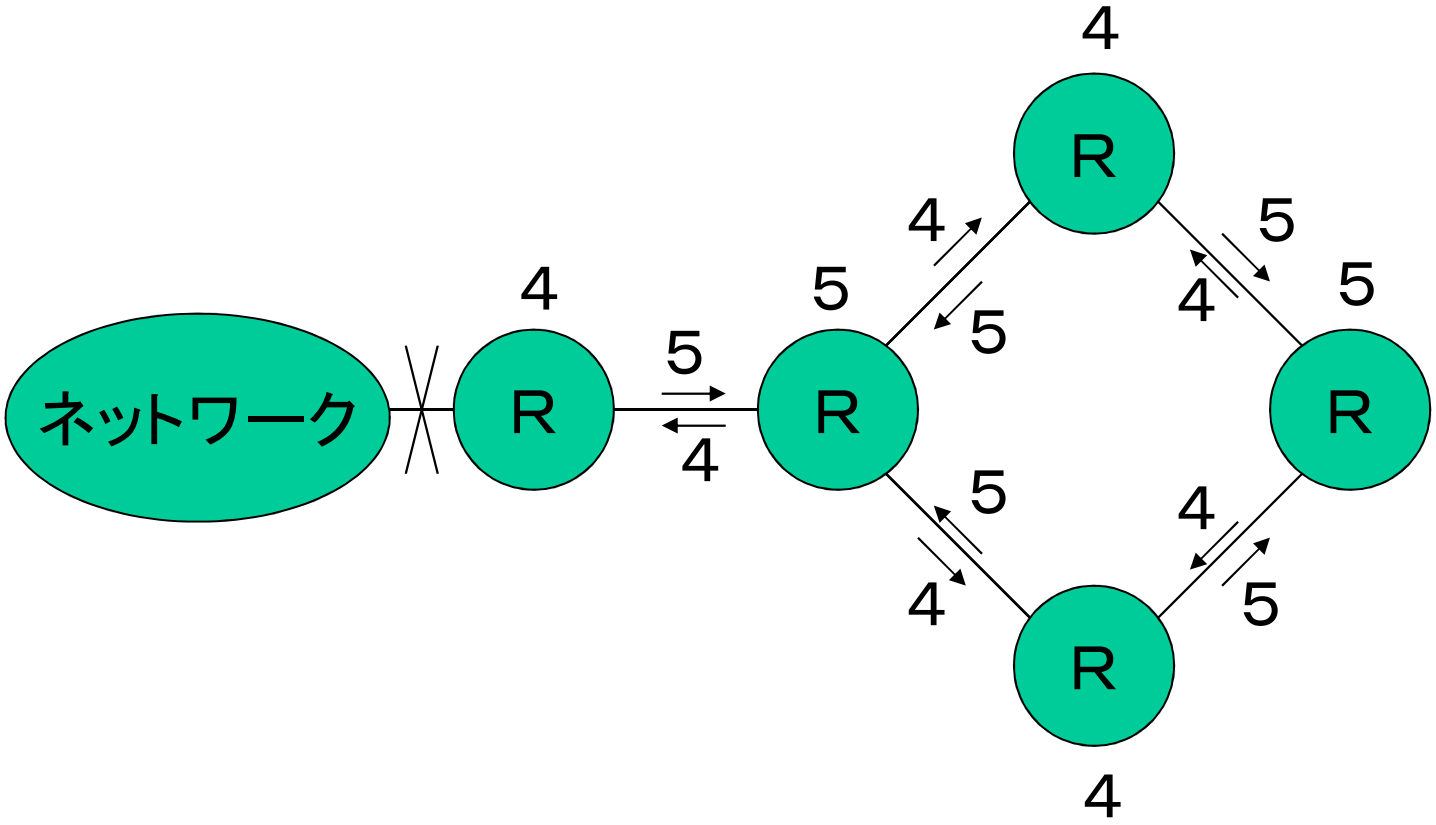
RIPの距離の伝搬



RIPの距離の伝搬



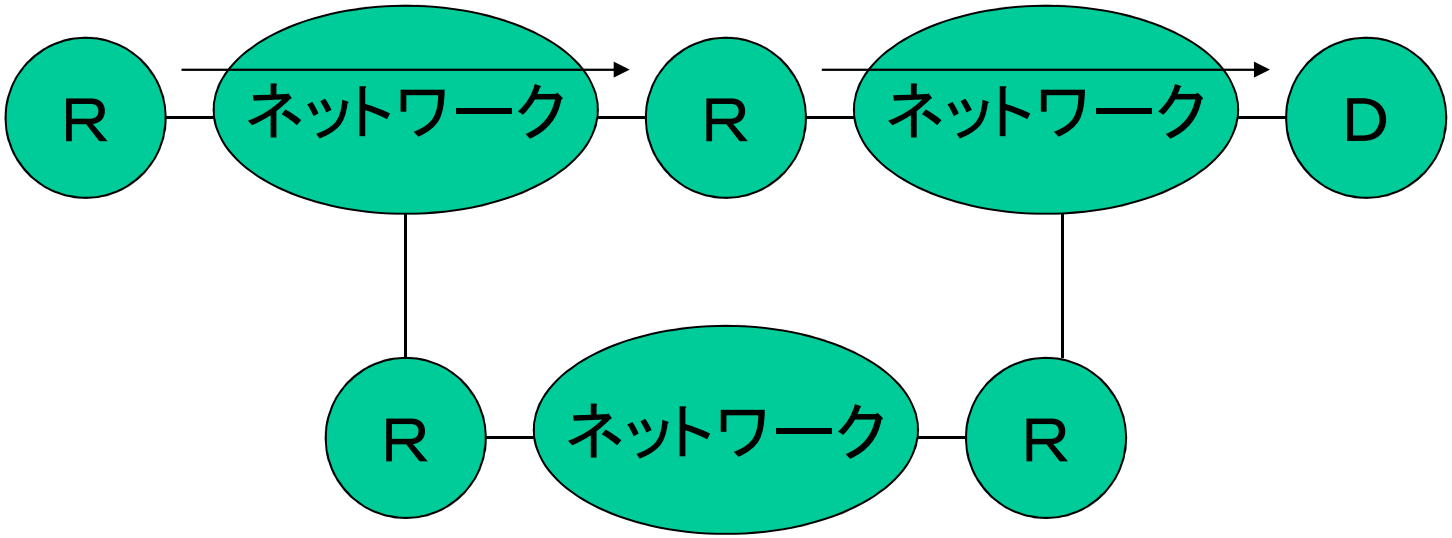
RIPの距離の伝搬



RIPの距離の伝搬

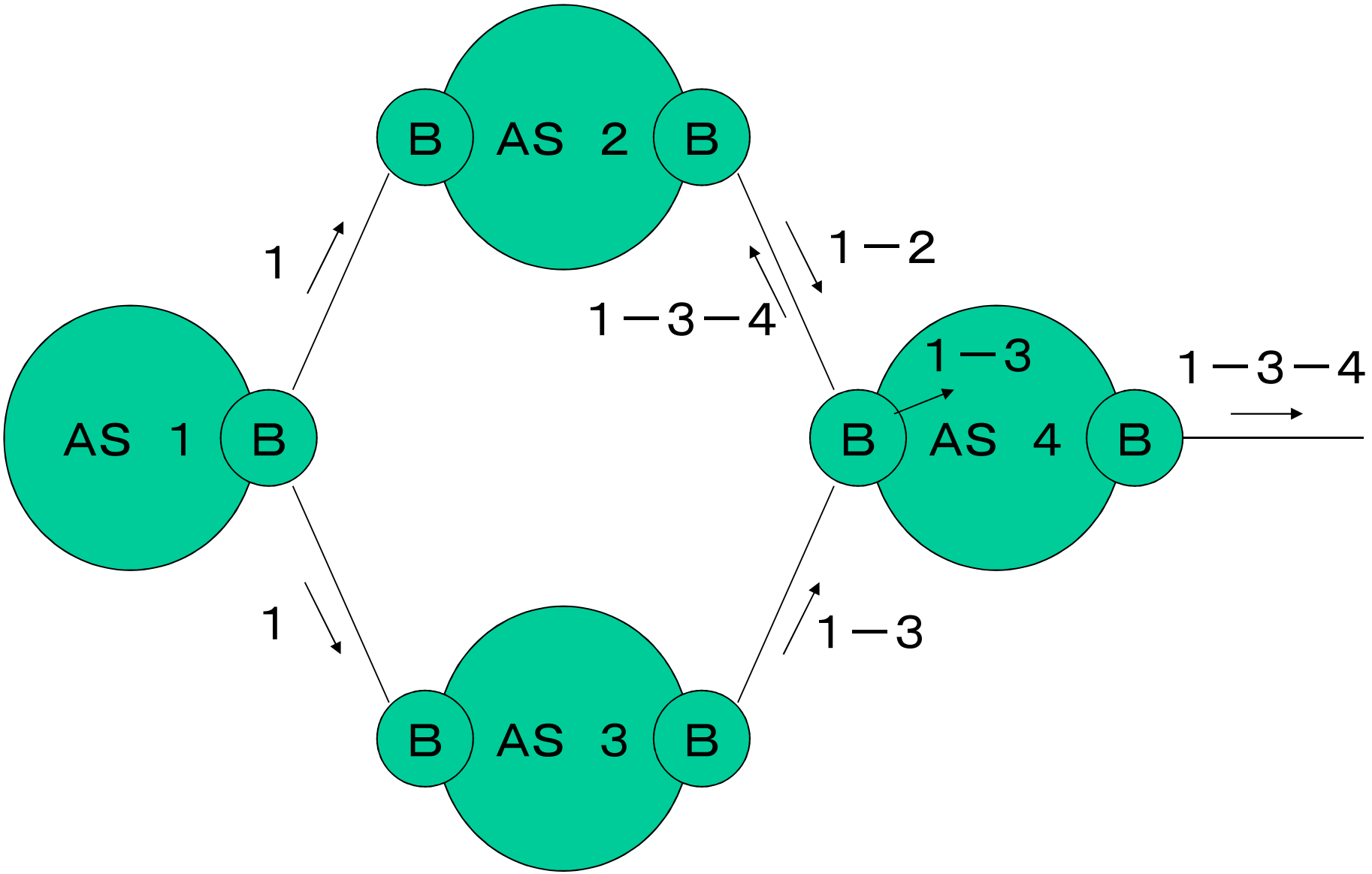
OSPF （Open Shortest Path First）

- LSのIGP
- 全ルータが同じ情報をもとに各目的地への最短経路を計算

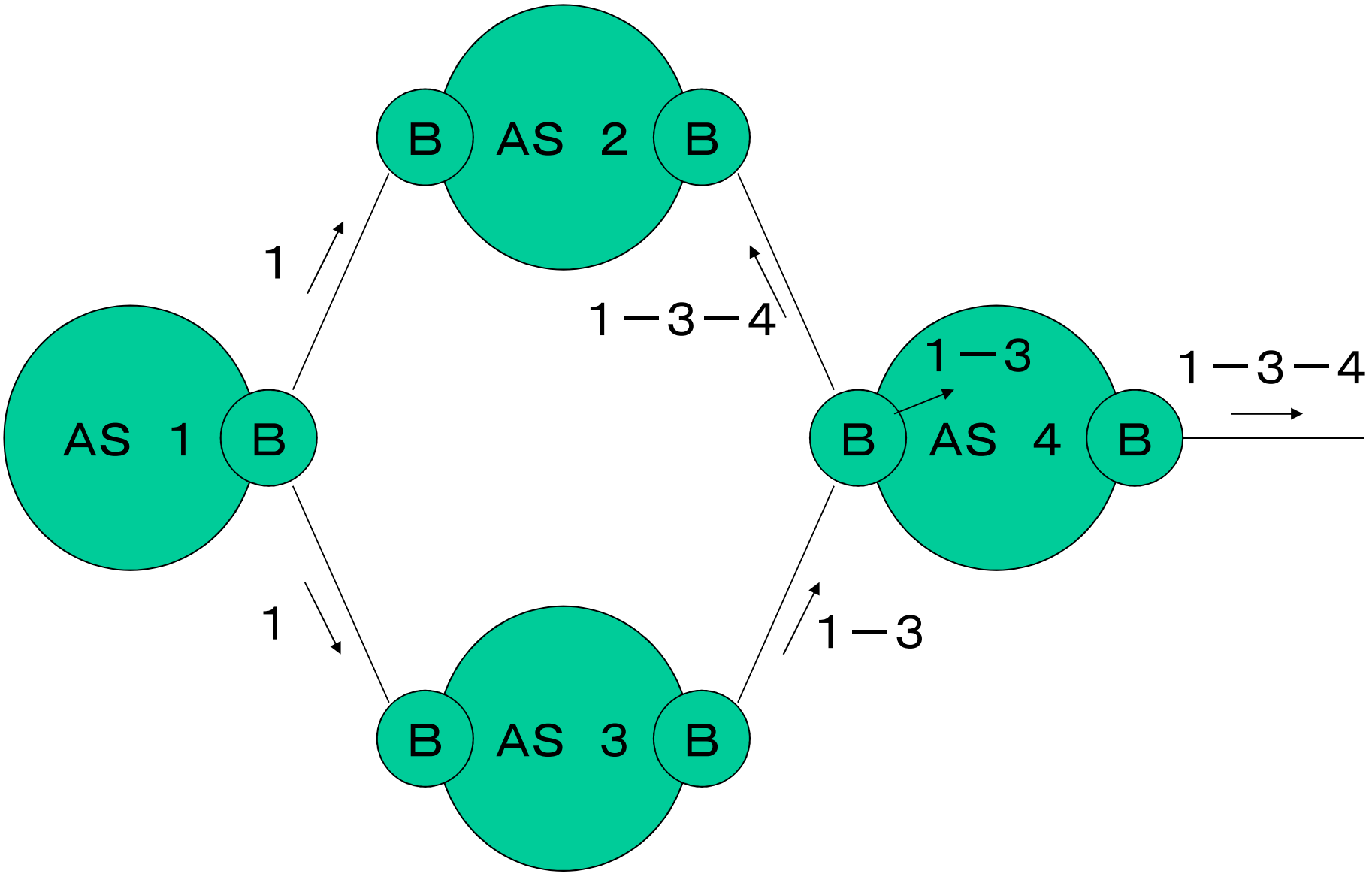


BGP（Border Gateway Protocol）

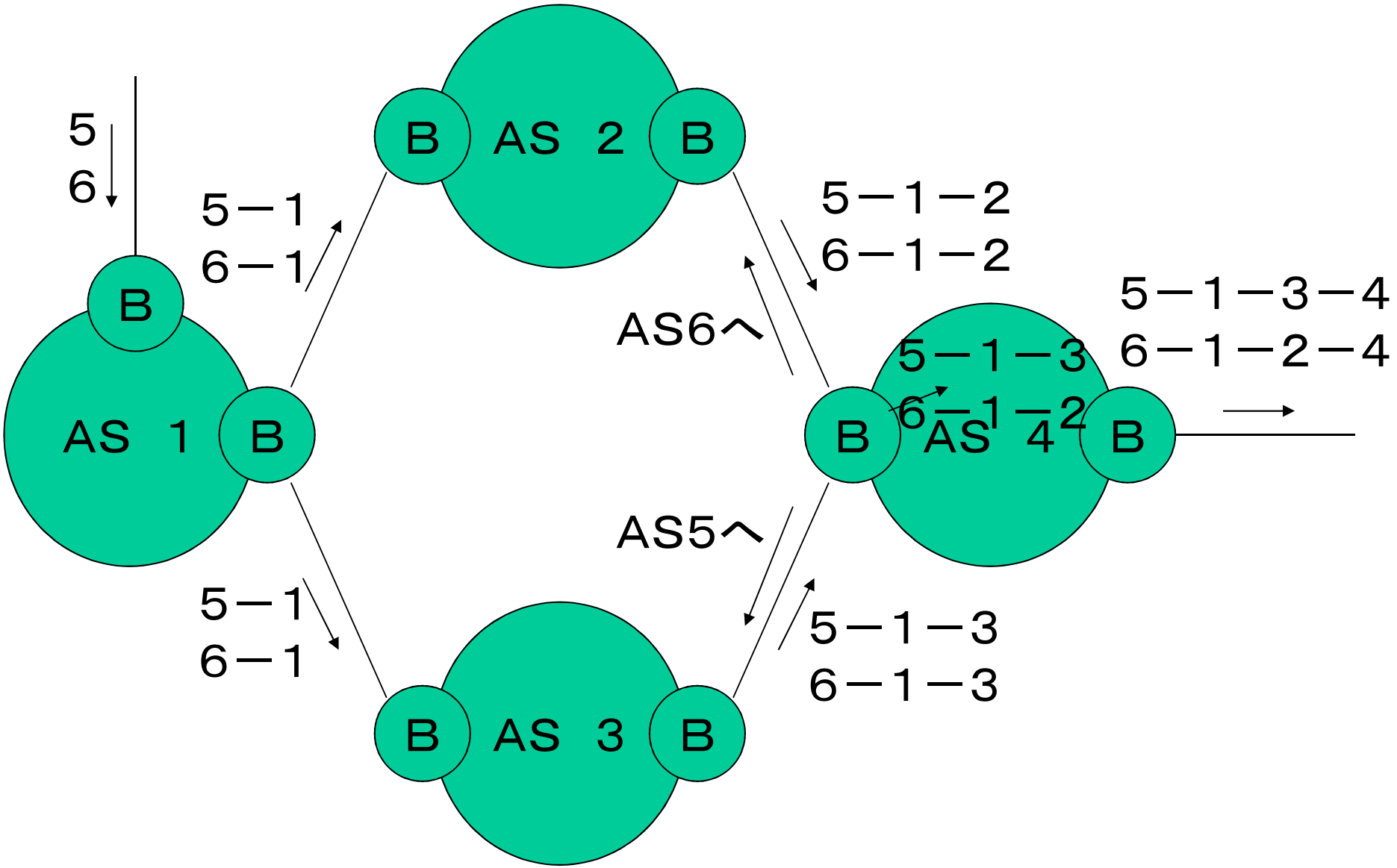
- DVのEGP
  - AS（Autonomous System）間で利用
- 距離ではなくASパスを利用
  - それまで経由したASの列
  - どのASパスを選択するかがポリシー
    - 実際にはASパスの短かさで決めることも多い
  - ASパスにはループがないので収束が速い？
- ボーダールータ間はTCPで通信
  - ASに属するアドレス範囲とASパスを交換



ASパスとポリシー（AS1はAS2よりAS3を好む）



ASパスとポリシー（AS2はAS1に協力しない）


5
6
B
AS 1
B
5－1
6－1
B
AS 2
B
5－1－2
6－1－2
AS6へ
5－1－3
6－1－2
B
AS 4
B
5－1－3－4
6－1－2－4
AS5へ
5－1－3
6－1－3
5－1
6－1
B
AS 3
B

ASパスと負荷分散

ルーティングレジストリ

- BGPの設定ミスは容易に全世界に波及
- どのアドレス範囲がどのASに属するかは、ルーティングレジストリとして管理
  - BGPの設定誤りをチェックできる
    - 誤った情報は伝搬しない
- あるアドレス範囲は一つのASにしか所属しない
  - エニキャストが難しい？

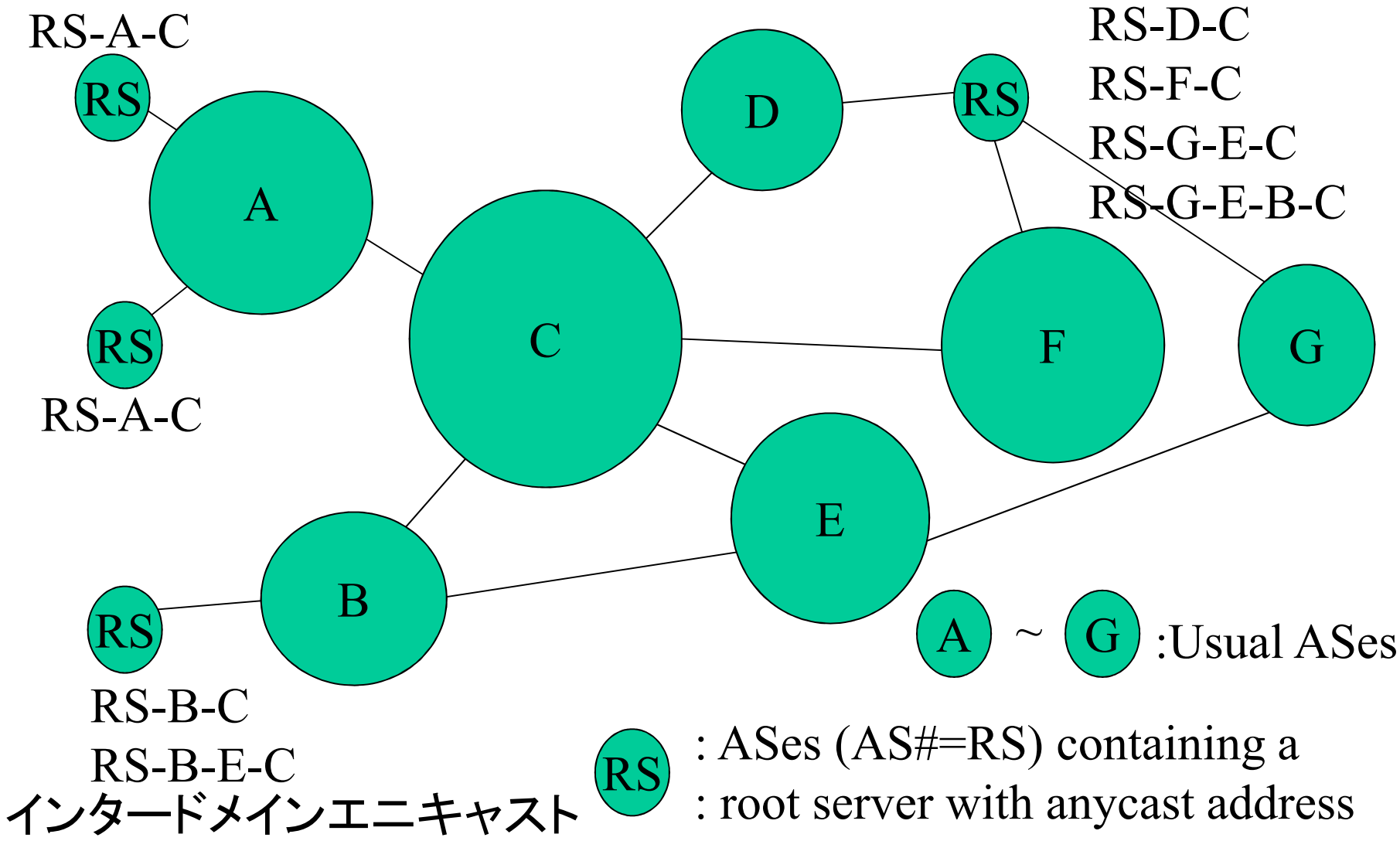
エニキャスト（anycast）

- １つのアドレスを複数のホストで共有
- エニキャストアドレス向きのパケットがどのホストで受け取られるかは、ルーティングプロトコルで決まる
  - IGPでは最寄り
  - EGPではポリシーにより選択可能
    - エニキャストASを使う
- エニキャストアドレスごとにルーティングテーブルを１個消費

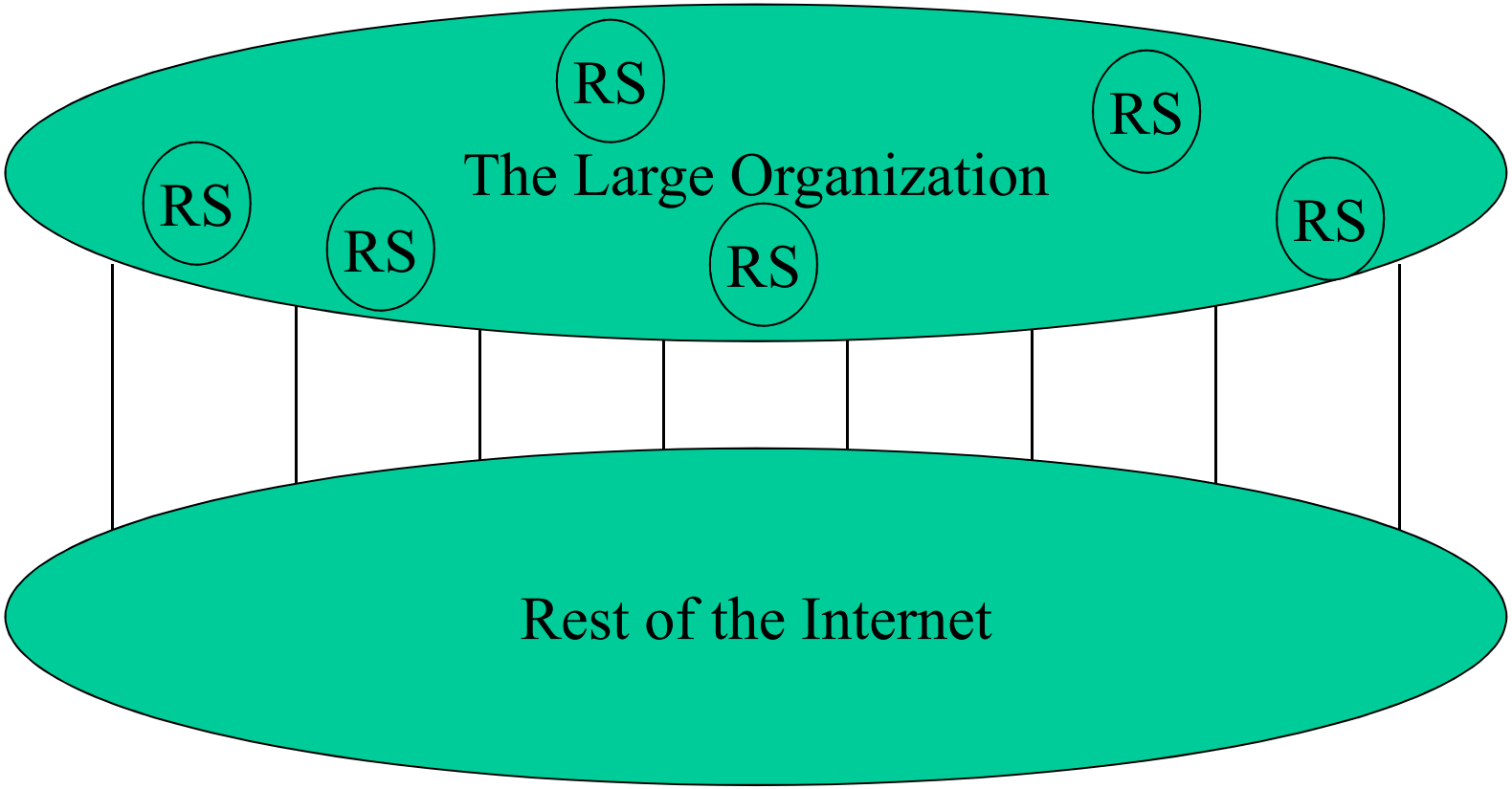
エニキャストの応用

- （DNSルート）サーバーの分散
- 緊急通信
- 位置情報を得る（後述）

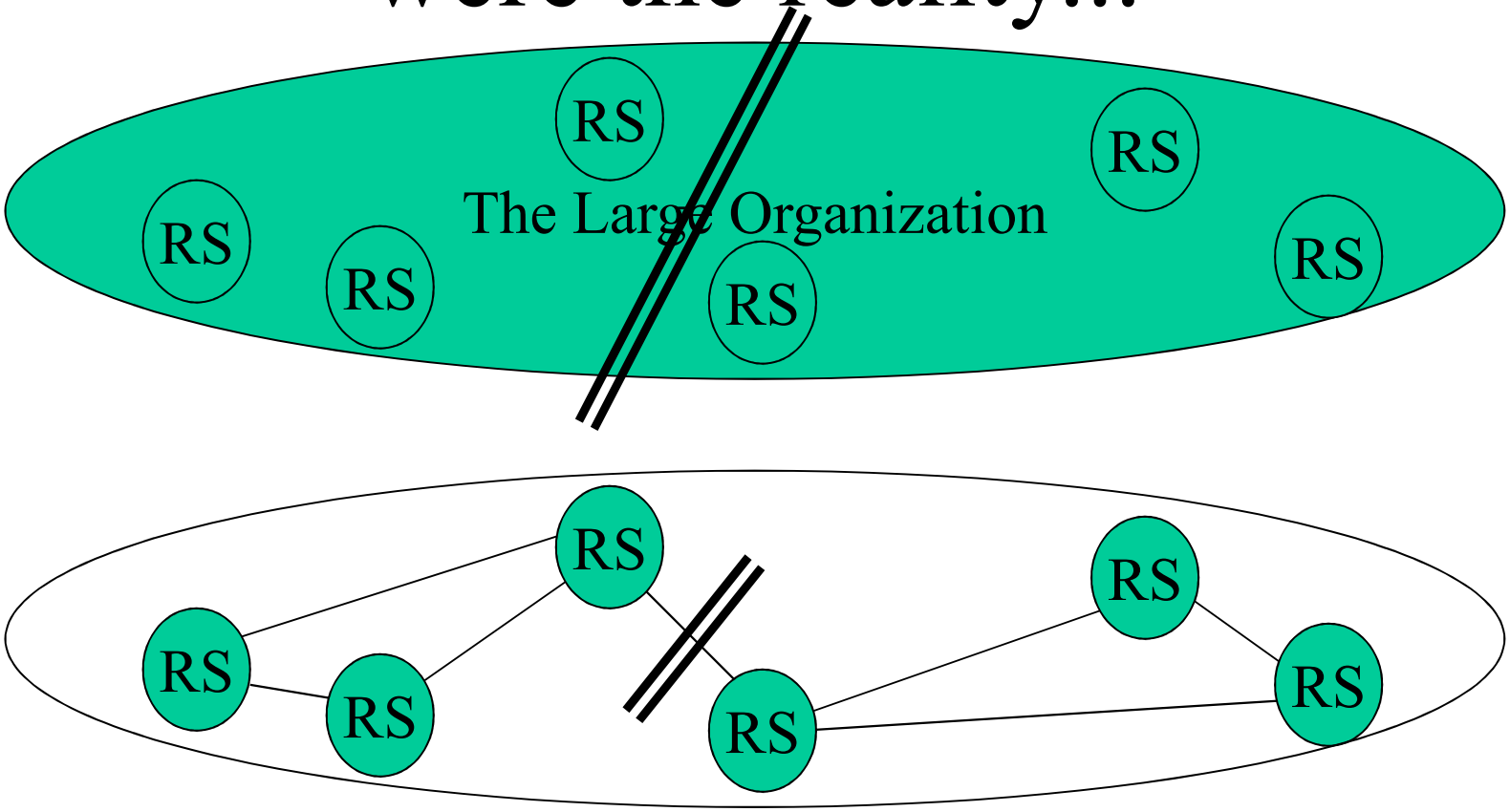
An Example (AS-pathes at C)



Another Approach: Intradomain Anycast with a Large AS



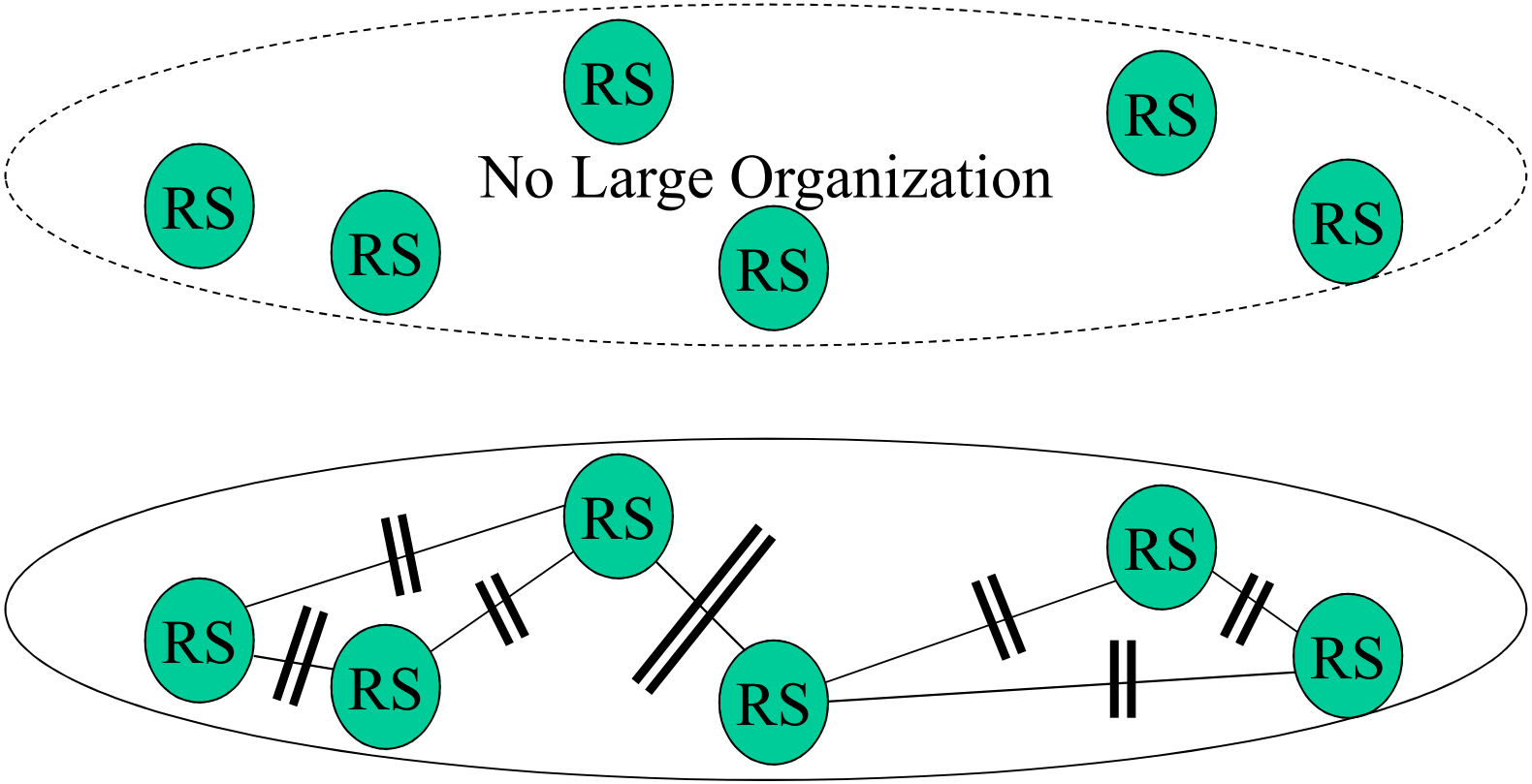
Even if the large organization(s) were the reality...



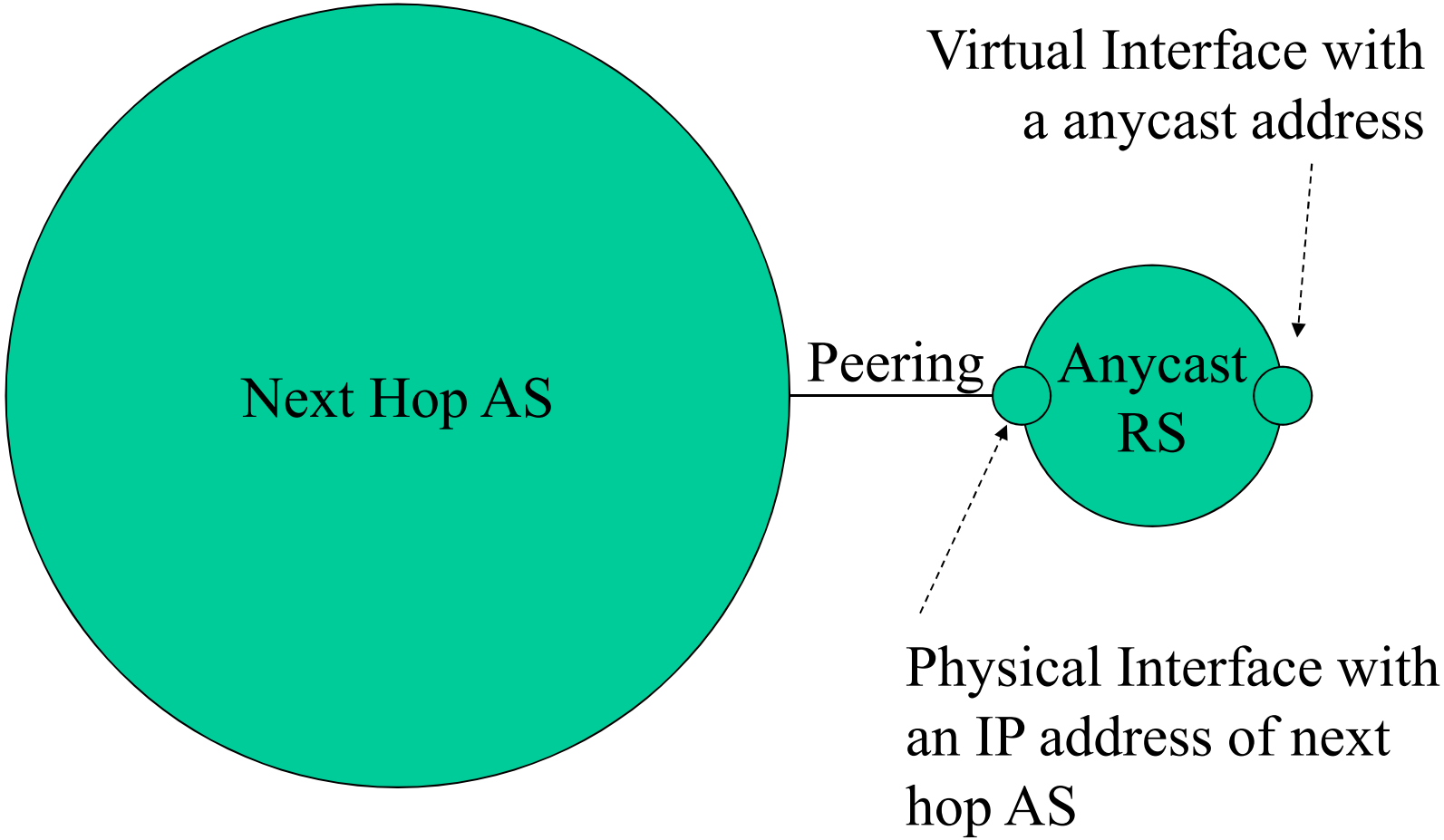
Can the geographically and topologically large organization have rich and robust internal connectivity?

Or, does the organization advertises a route to RSes only? But...

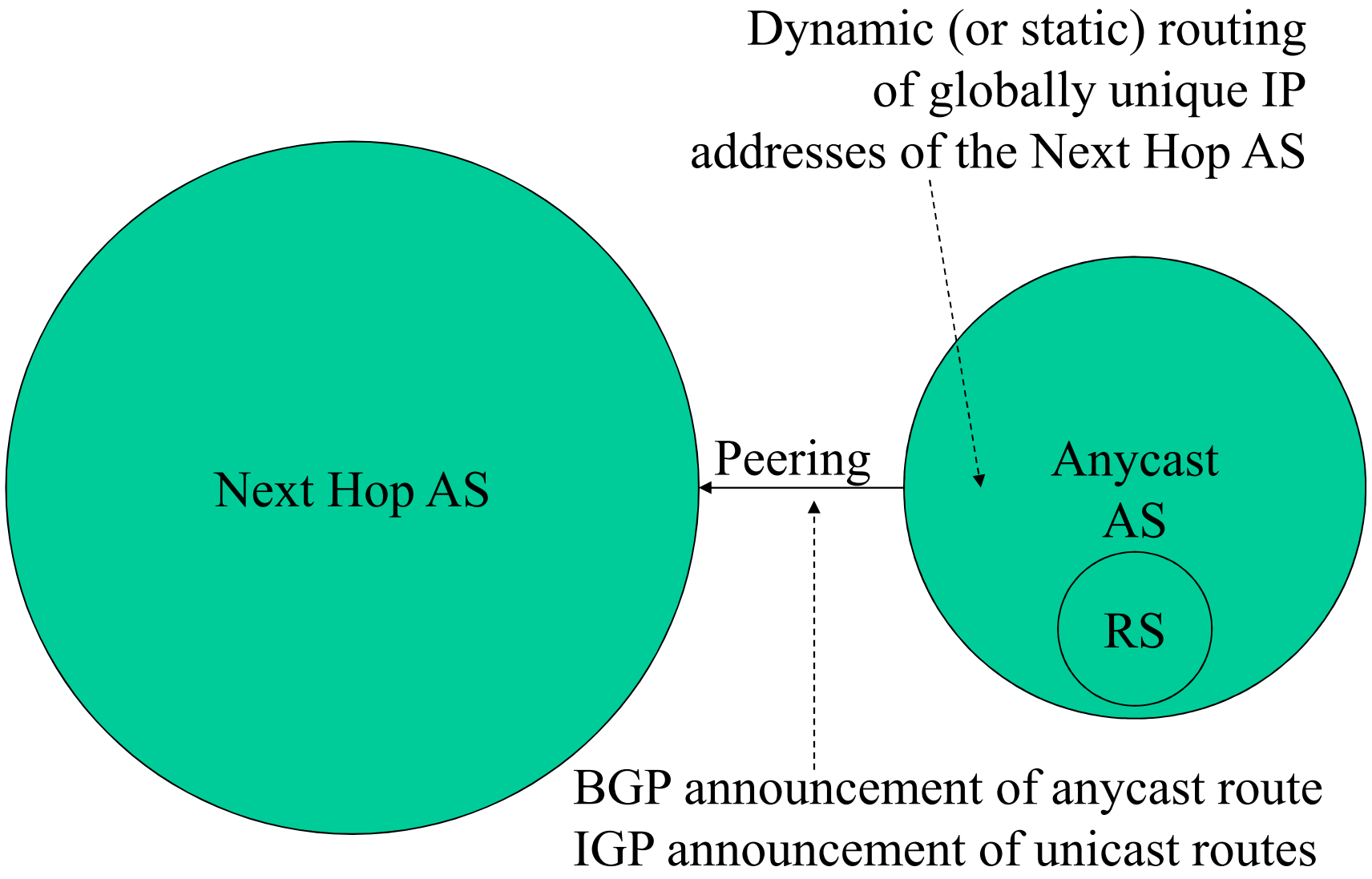
The Extreme Case



Anycast RS with Unique Addresses



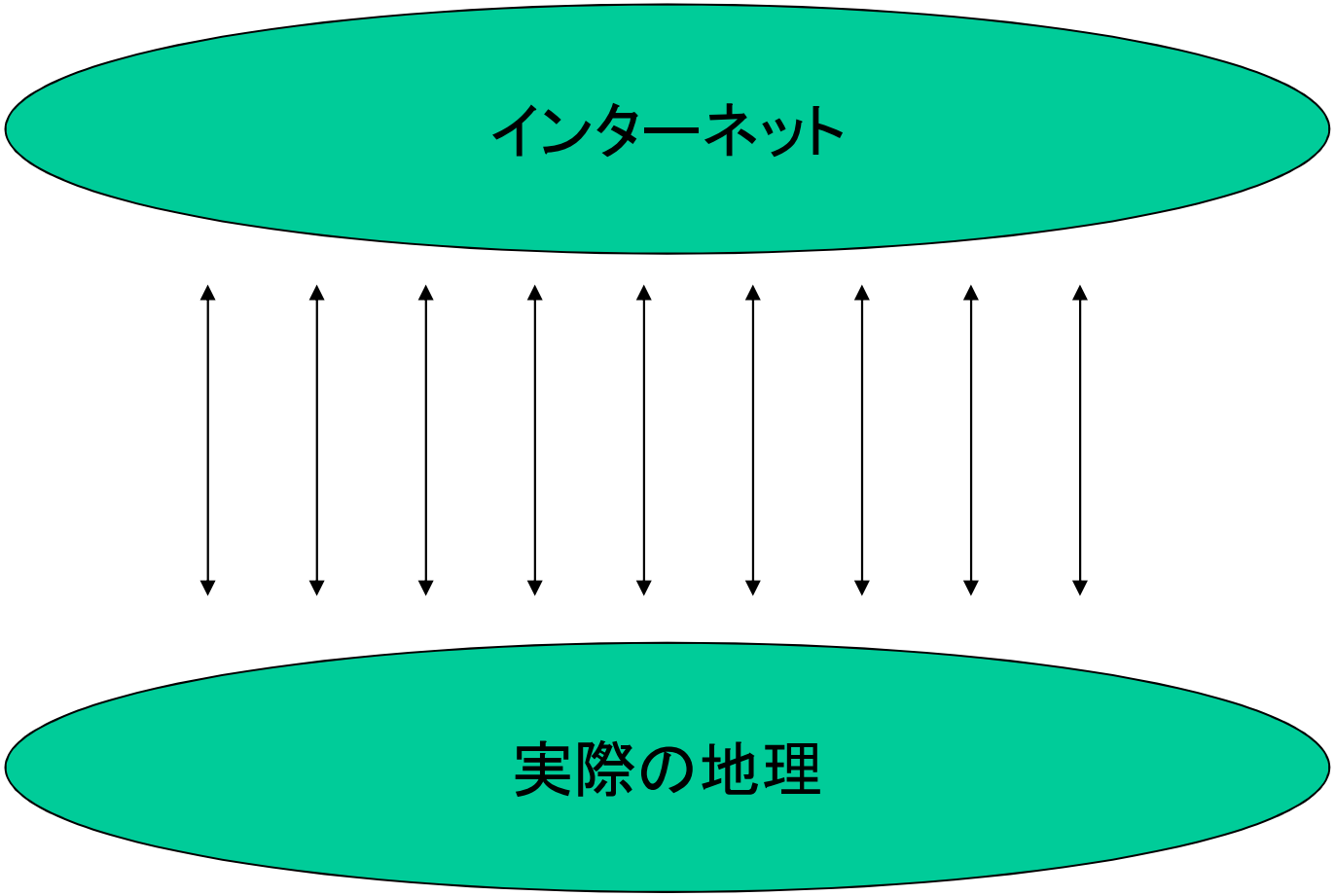
Anycast AS = Anycast Root Server

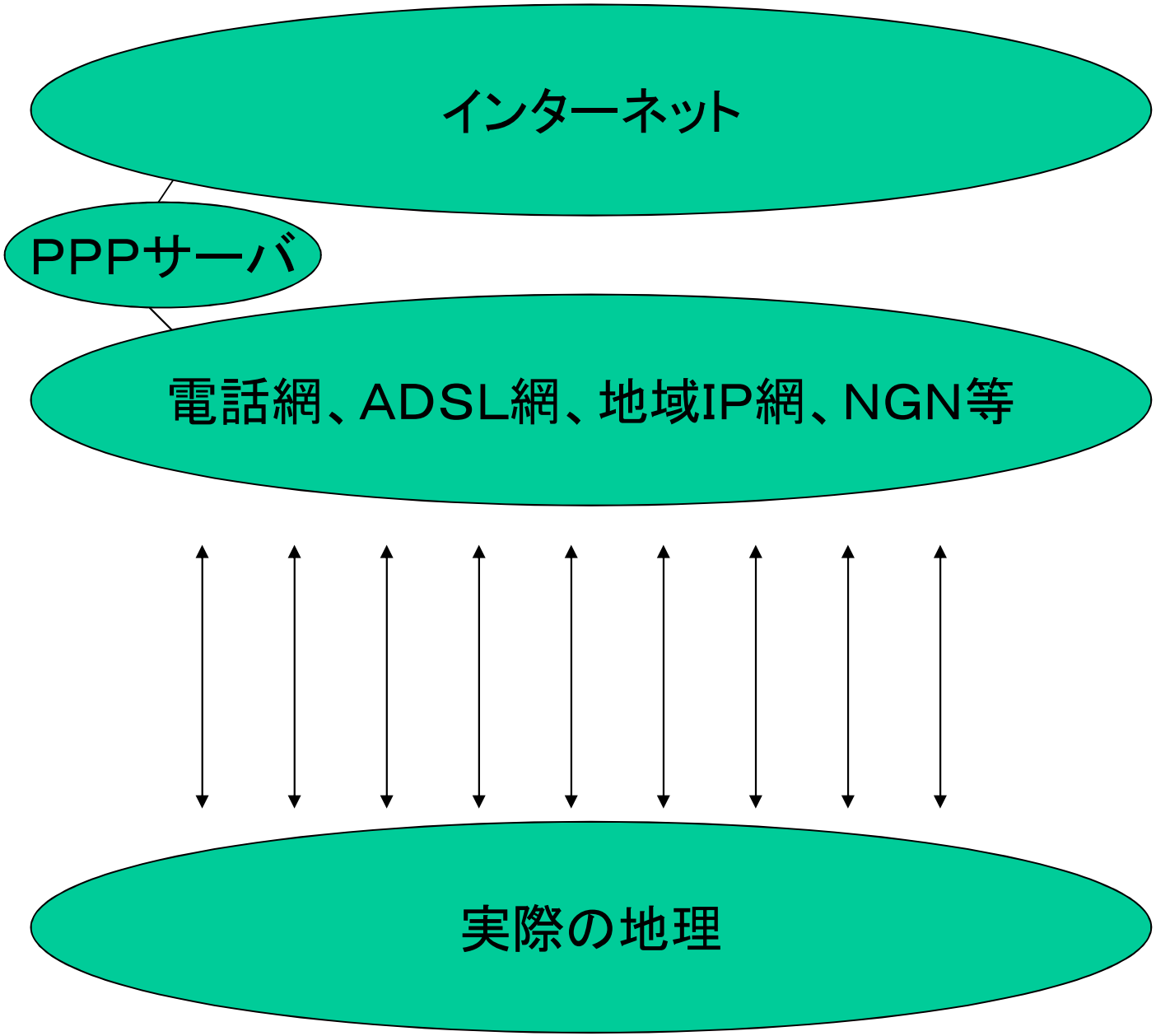


Anycast AS != Anycast Root Server

エニキャストと緊急通信

- 同じIPアドレスで、ある団体（消防、警察等）の自分の最寄の支所と交信したい
  - 電話網の１１０番や１１９番
- エニキャストを使えば、そのまま可能
  - 最寄にルーティングしてくれる
- インターネット常時接続なら
  - インターネットトポロジーと実際の地理は対応
  - ダイアルアップ（PPP）等ではうまくいかない


インターネット
実際の地理


インターネット
PPPサーバ
電話網、ADSL網、地域IP網、NGN等
実際の地理

デフォールトルート

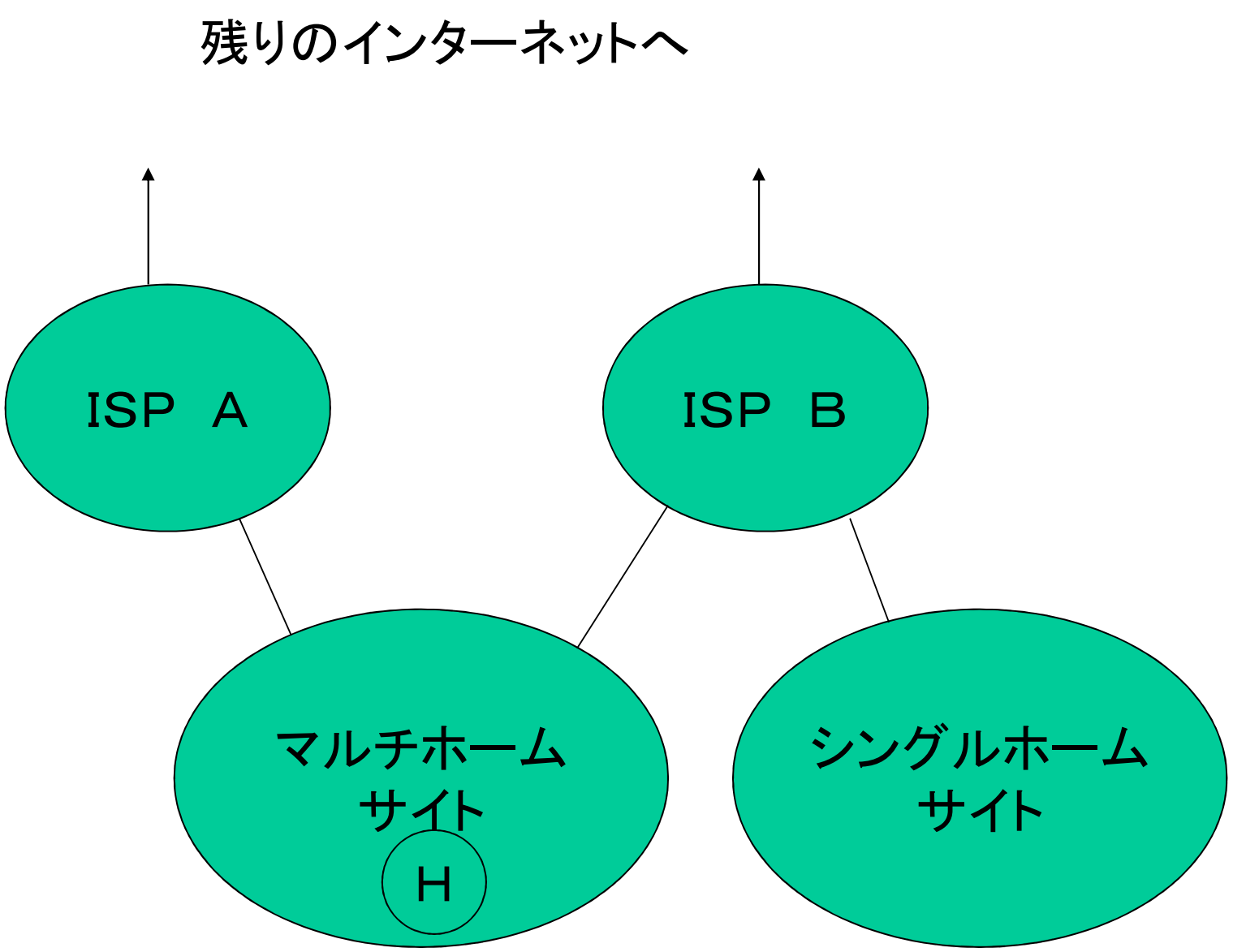
- インターネット全体の経路表
  - 現在は30万エントリ程度
  - ４～５年前は10万エントリ程度
    - マルチホーミングにより着実に増加
- 全ルータが全経路をもつ必用なし
  - 遅い通信路では全経路表は送れない
- 末端ではデフォールトルートを使う
  - あとはネットワークまかせ

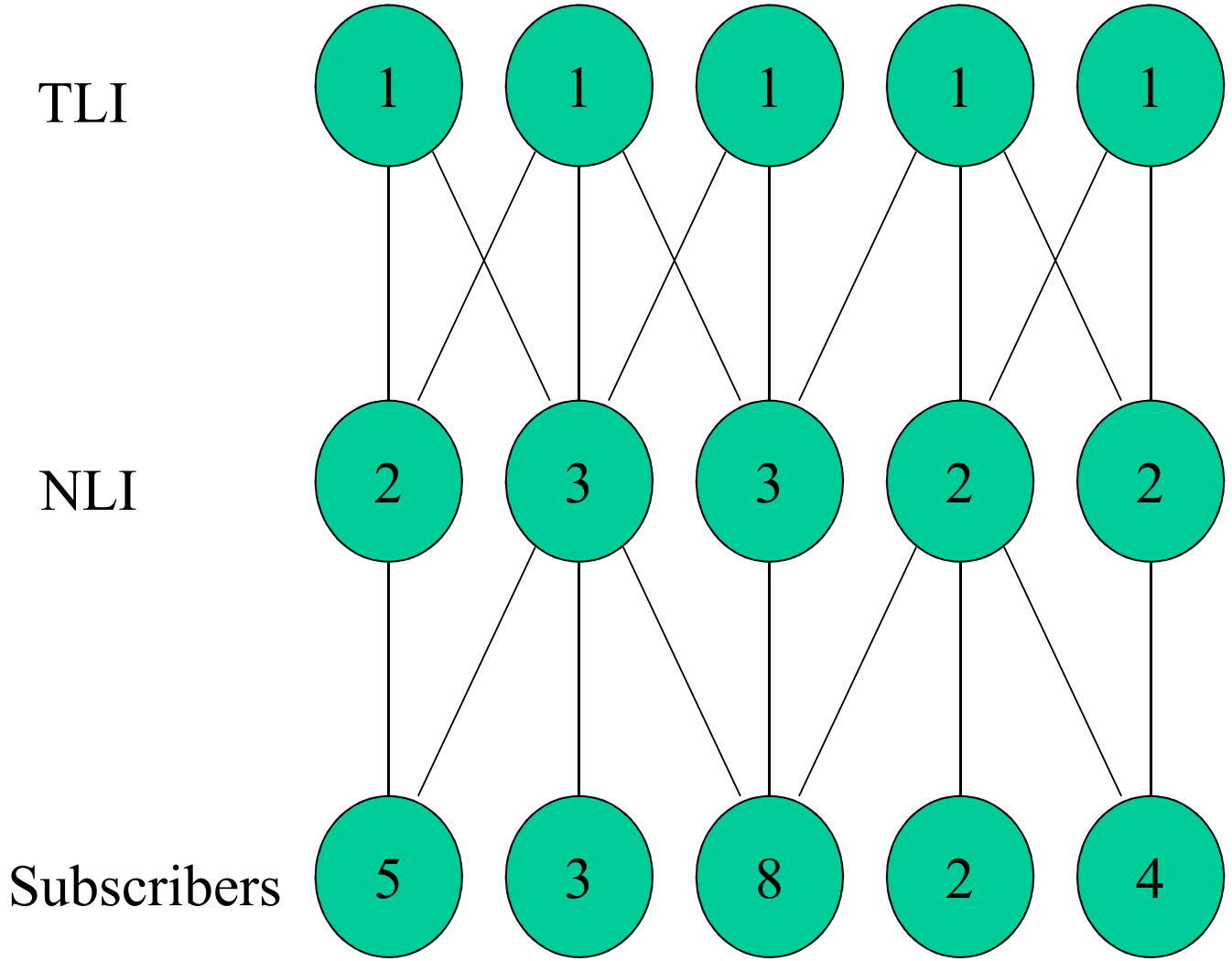
デフォールトルータ

- 経路表を得るにはIGPの理解が必用
- 普通のホストは経路表もつ必用なし？
  - ノードとルータの分離（IPv6）
  - ノードではIGPの設定が不要になる
- パケットはすべて手近のルータに送る
  - デフォールトルータ
- デフォールトルータは、別のルータを、ノードにICMP　redirectで紹介することも

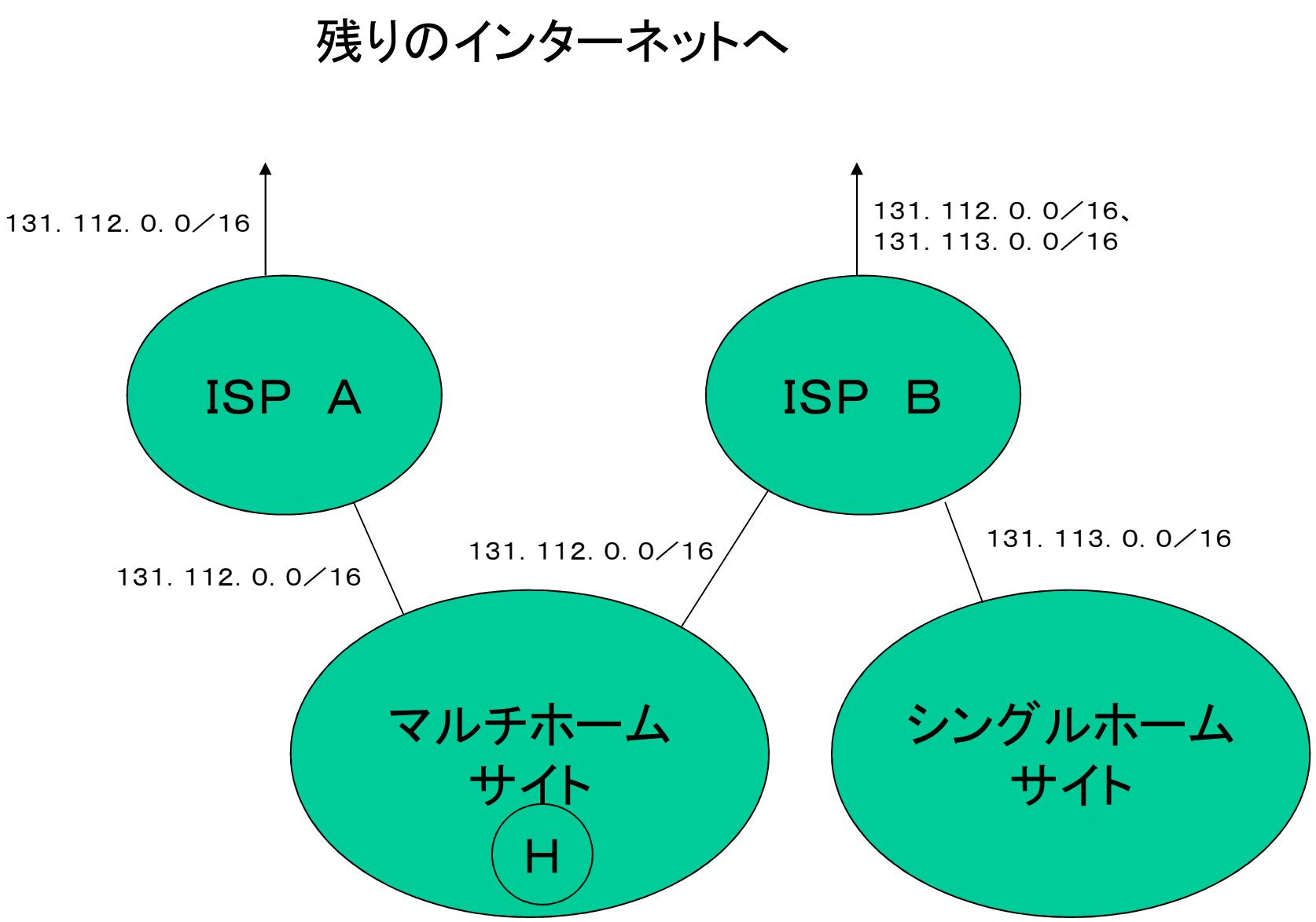
マルチホーミング

- 複数の上流ISPをもつ
  - どちらかがこけても大丈夫
- 信頼性のあるサービス（含ISP）には必須
  - IPv6ではNLISPは複数のTLISPに接続したいが、、、
- ルーティングによるマルチホーミングではISPはTLAしかもてない


残りのインターネットへ
ISP A
ISP B
マルチホーム
サイト
H
シングルホーム
サイト



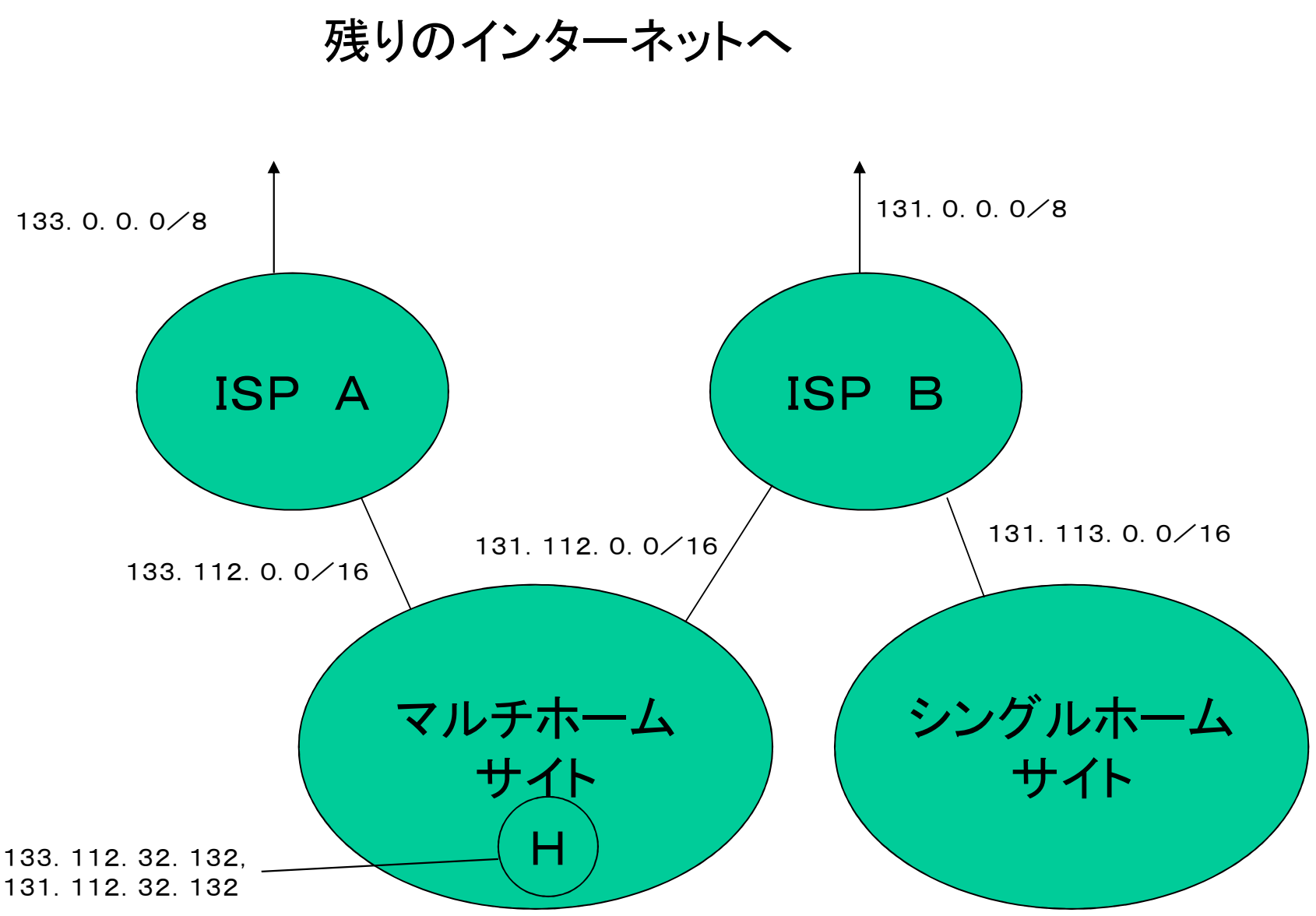
Number of Prefixes with E2E Multihoming



ルーティングによるマルチホーミング

エンドツーエンド
マルチホーミング

- ホストは複数のアドレスをもつ
- ホストは通信相手の複数のアドレスを自分で試す
  - どれかでつながれば通信は成立
  - 通信中にタイムアウトなどがおきれば、他のアドレスを試す
- ルーティングによるマルチホーミングは不要


残りのインターネットへ
133. 0. 0. 0／8
131. 0. 0. 0／8
ISP A
ISP B
133. 112. 0. 0／16
131. 112. 0. 0／16
131. 113. 0. 0／16
マルチホーム
サイト
H
シングルホーム
サイト
133. 112. 32. 132,
131. 112. 32. 132
エンドツーエンドマルチホーミング

マルチホーミングと
デフォールトルート

- マルチホームASでデフォールトルートを使ったらあまり意味がない
- エンドツーエンドマルチホーミングでは、最適な相手アドレスの選択にルート情報を参考にしたい
- 各ホストが全ルート情報をもってこそエンドツーエンド原理にかなう
  - IPv6では可能か？（駄目でした）

インターネットの今後

- 光ファイバーが中心
  - 圧倒的な速度（１心で１０Tbps）
- 無線も捨てがたい
  - 無線幹線網（一対多通信）
    - 衛星インターネット放送
      - 放送網のコンテンツがキラーアプリに
  - 無線アクセス網（配線不要）
    - 携帯インターネット
      - 電話網のコンテンツ（電話）がキラーアプリに

電波とインターネット

- 近距離（小電力）
  - 多数の基地局を設置
  - IPモビリティと組み合わせて携帯インターネットサービスを実現
- 長距離（大電力）
  - １対多通信には電波は最適
  - 衛星インターネットは速い？

携帯インターネット

- 携帯電話網は電話網でしかない
  - １２８バイト０．３円だと、６４Ｋｂｐｓで１秒２０円
- 高速低額常時接続定額の固定インターネットに無線機を接続すると
  - 高速低額常時接続定額の無線インターネット
  - 無線区間のセキュリティには一工夫必要
- IP　mobilityとの組み合わせで携帯インターネットに

携帯インターネット

- 無線インターネット＋IP　mobility
  - 無線による同一基地局内の自由な移動
  - IP　mobilityにより、基地局の移動後も同じIPアドレスが使え、TCP接続なども保たれる

無線インターネット

- 固定インターネットが大前提
  - 速くて安くて定額制の固定インターネットに速くて安い無線機を接続すると
    - 安くて定額制の無線インターネットが実現
  - 無線機の密度を十分高くすると
    - 速くて安くて定額制の無線インターネットが実現

無線インターネットの技術的課題

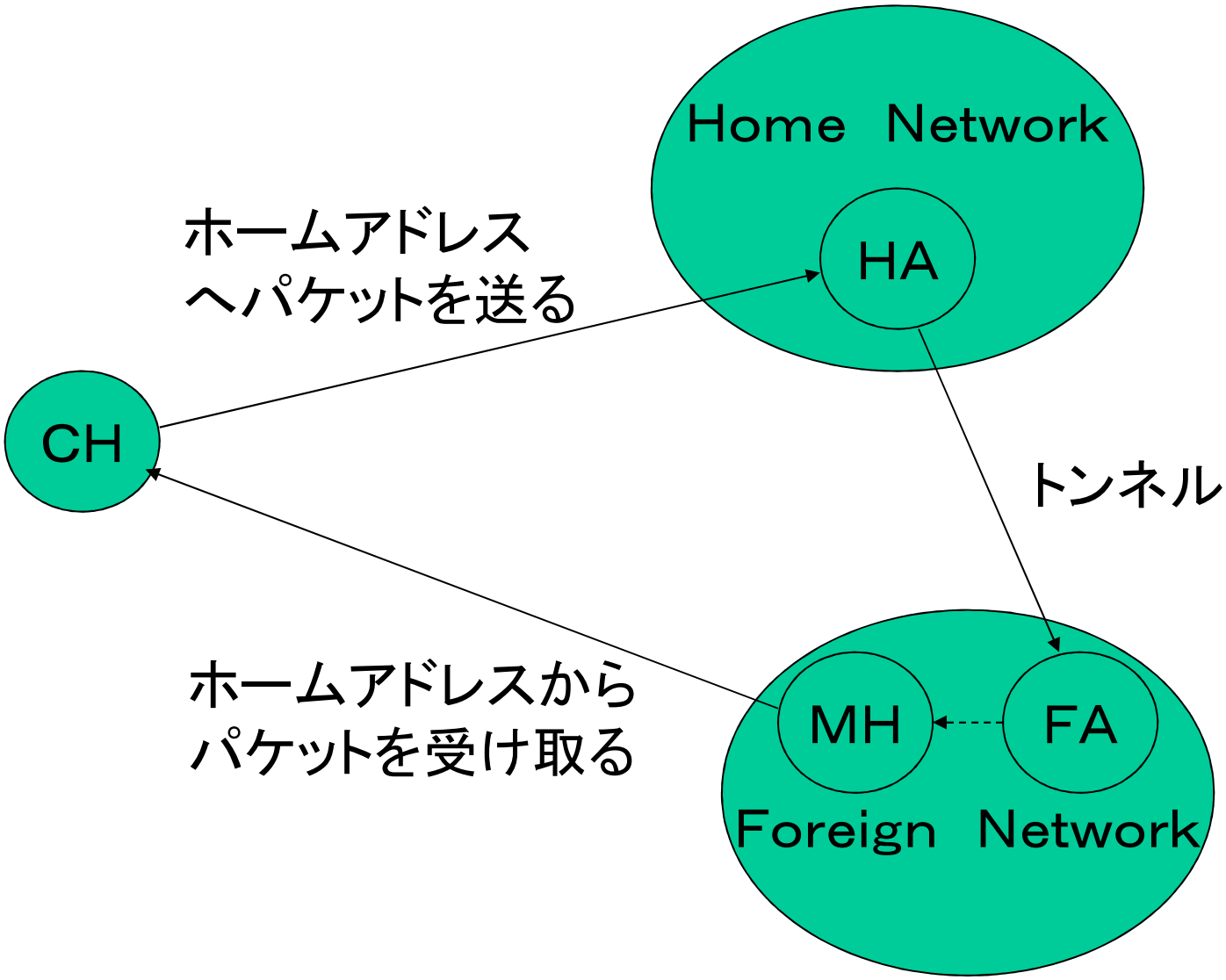
- 無線は不特定多数が使える
  - 認証
    - 誰もがいつでもどこでもインターネットをつかえるのはいいが
    - どこの誰なのか身元がわからないのは困る
      - 犯罪捜査
      - 課金
  - 暗号化
    - 本来はエンドツーエンド
    - 無線上でパスワードを入力するような場合に便利

無線インターネットの
セキュリティ

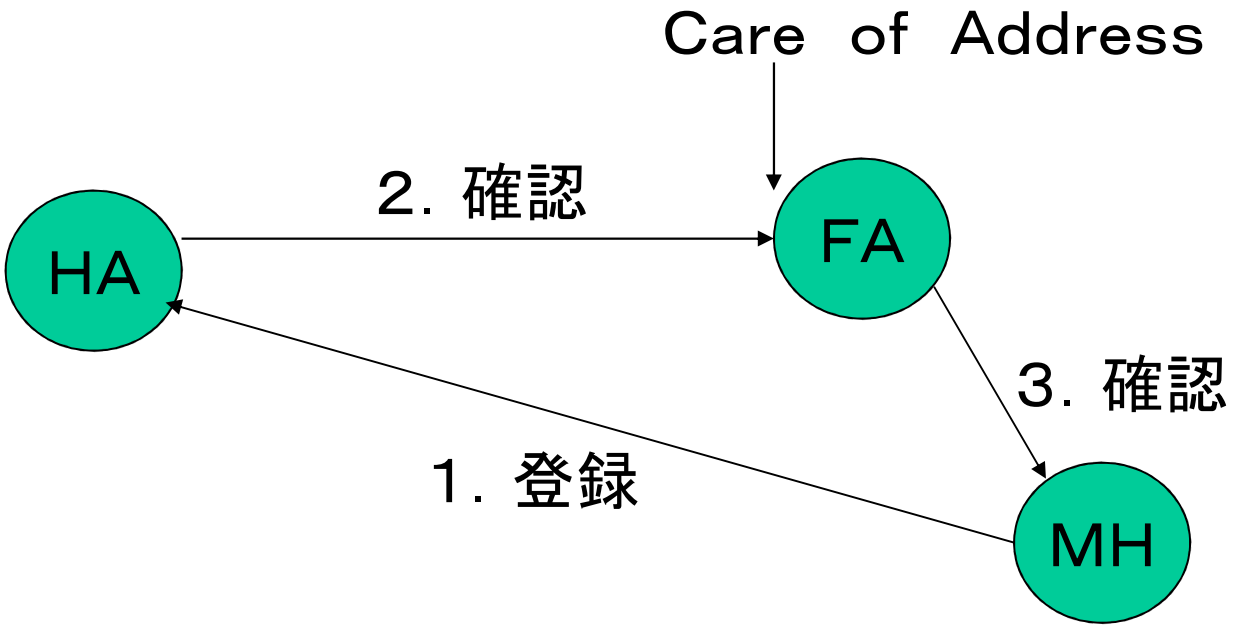
- RADIUSサーバでユーザごとの鍵を管理
- ユーザはセッション鍵を生成、自分の鍵で暗号化して無線基地局に送る
- 無線基地局はRADIUSサーバにセッション鍵の解読を依頼
- セッション鍵は認証と暗号化に利用可能

IP Mobility （RFC2002）

- 端末が移動しても、同じIPアドレスを使い続けたい
  - TCP接続なども維持
- ４つの要素
  - HA：（Home Agent）
  - FA（Foreign Agent）
  - MH（Mobile Host）
  - CH（Correspondent Host）、普通のホスト



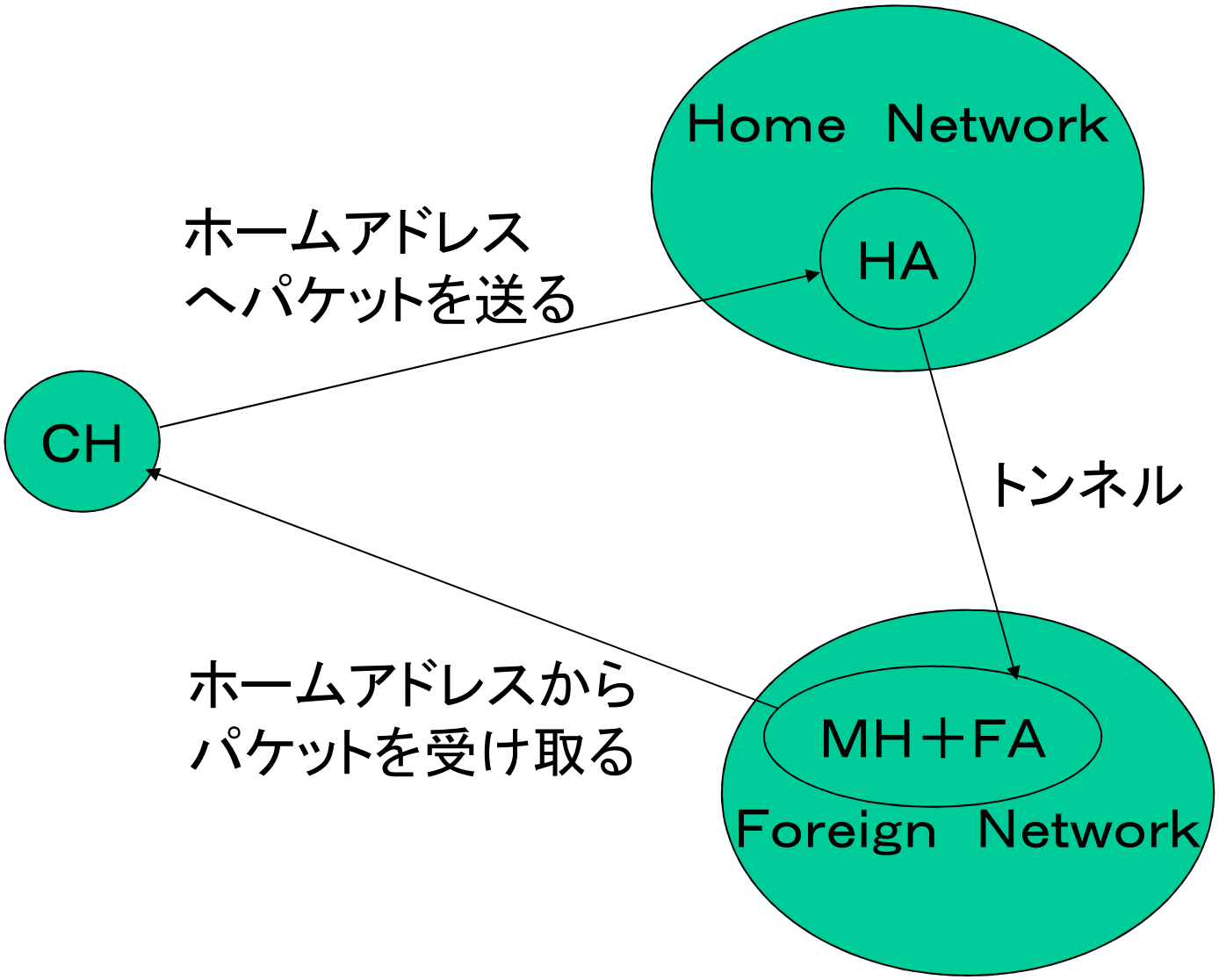
三角形型パケットのやりとり



位置(Care of Address)の登録

モビリティのエンド

- HA
  - FAの位置を把握
  - FAにパケットをフォワード
- MH
  - FAの位置をHAに登録
- FA？
  - HAとMHを仲介するネットワーク中の機器
  - エンドツーエンド原理違反


Home Network
HA
ホームアドレス
へパケットを送る
CH
トンネル
ホームアドレスから
パケットを受け取る
MH＋FA
Foreign Network

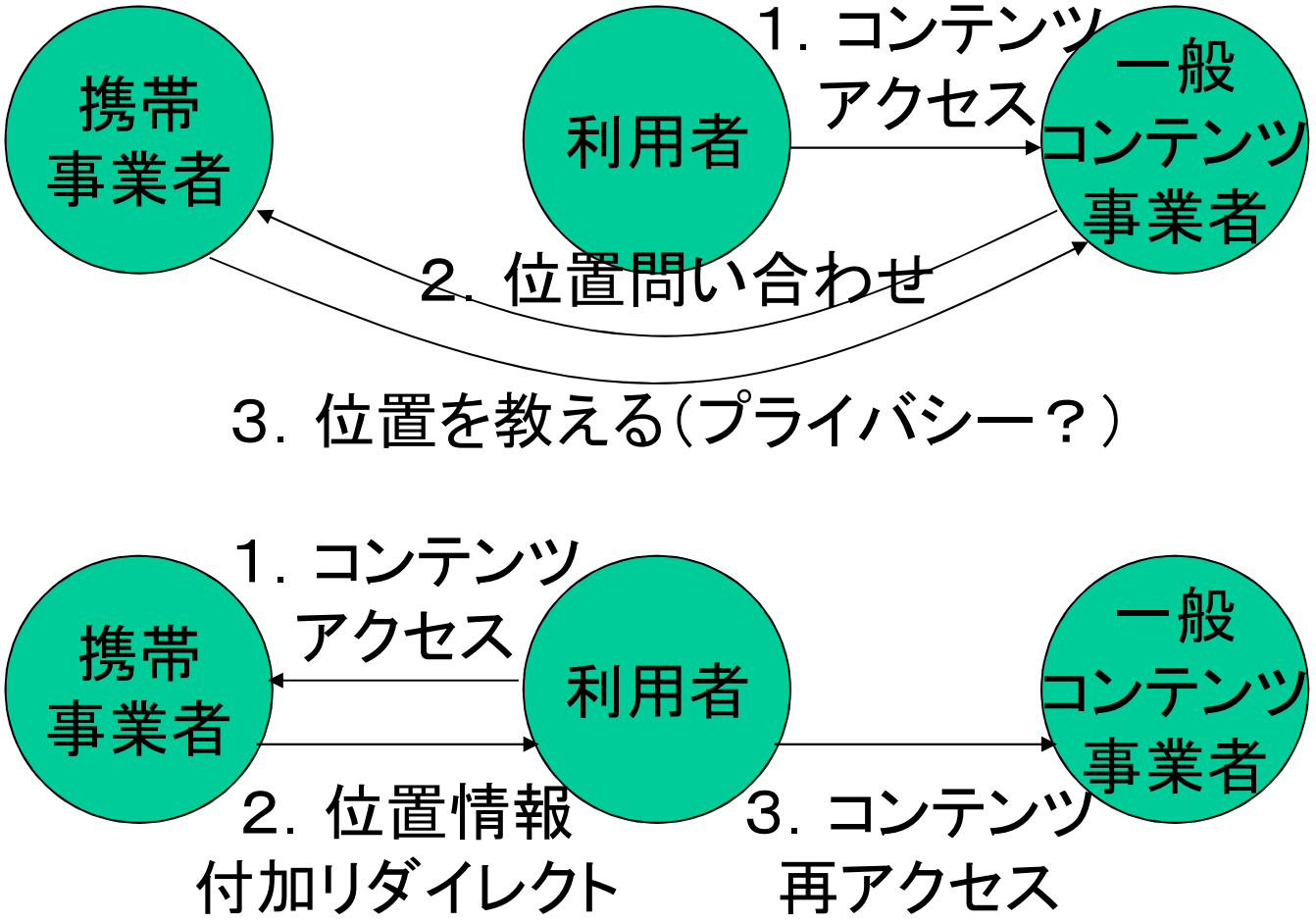
MHとFAの一体化

エニキャストによる位置依存情報の提供

- 携帯インターネットアクセスでは
  - 普通のインターネットコンテンツは全て利用可
  - さらに、端末の位置依存コンテンツの可能性
- 無線基地局でエニキャストアドレスを共有
  - 各基地局をサーバに
    - コンテンツは基地局ごとに異なる
  - エニキャストアドレスでコンテンツをアクセス
    - 最寄の基地局のコンテンツ（位置依存）を取得

位置依存情報と プライバシー

- 携帯事業者が利用者の位置を把握可能
  - 携帯事業者は位置依存コンテンツを提供可能
- 携帯事業者が利用者の位置を
  - 他の事業者に教えるとプライバシー侵害
    - 他の事業者は位置依存コンテンツの提供は困難
  - 利用者自身に教えるのは問題なし
    - 利用者がさらに他のコンテンツ事業者に教えるのも問題なし


携帯
事業者
利用者
1．コンテンツ
アクセス
一般
コンテンツ
事業者
2．位置問い合わせ
3．位置を教える（プライバシー？）
携帯
事業者
1．コンテンツ
アクセス
利用者
一般
コンテンツ
事業者
2．位置情報
付加リダイレクト
3．コンテンツ
再アクセス

まとめ（1）

- ルーティングプロトコルはネットワーク中の知性
  - DVは分散計算
  - LSは全ルータが全情報を知る
  - DVよりLSのほうが各ルータの自律性が強い
    - よりエンドツーエンドに適合
  - BGPも各ルータの自律性性が強い
- エニキャストは便利

まとめ（２）

- デフォールトルートはルータを特別視
  - エンドツーエンド原理違反
- マルチホーミングはルーティングプロトコル（ネットワーク中の知性）に頼らずに
  - エンドツーエンドマルチホーミング
- モビリティの問題点はセキュリティ（認証）
- モビリティのＦＡはネットワーク内部の機能
  - エンドツーエンド原理違反