

画像情報特論 (11)

- その他の話題

- マルチキャスト
- CDN
- モビリティ

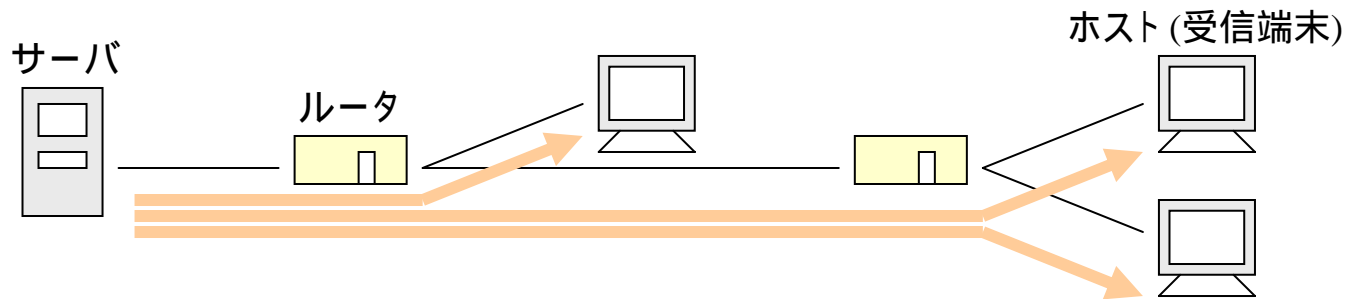
2002.07.02

電子情報通信学科 甲藤二郎

E-Mail: katto@katto.comm.waseda.ac.jp

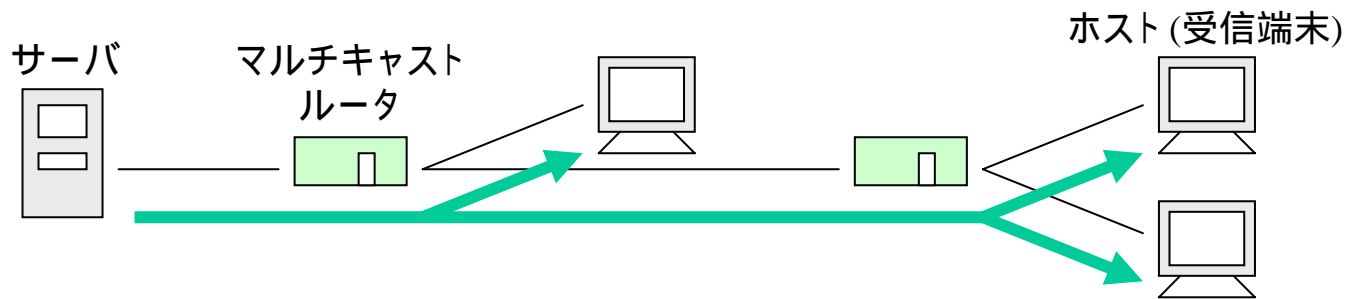
マルチキャスト

マルチキャスト



(a) ユニキャスト

• n 人の参加者 n 本のユニキャスト

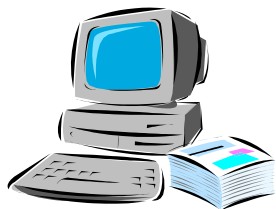


(b) マルチキャスト

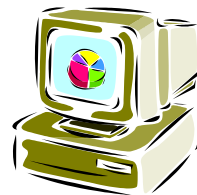
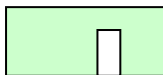
• n 人の参加者 1 本のマルチキャスト

IPマルチキャスト (1)

マルチキャスト
サーバ

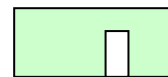


マルチキャスト
ルータ



マルチキャスト・ルーティング・
プロトコル

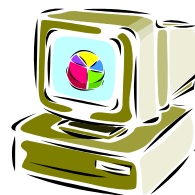
経路の確立・削除



マルチキャスト
ルータ

IGMP

Join/Leave



(S,G): マルチキャストグループ

S: 送信者アドレス

G: マルチキャストアドレス

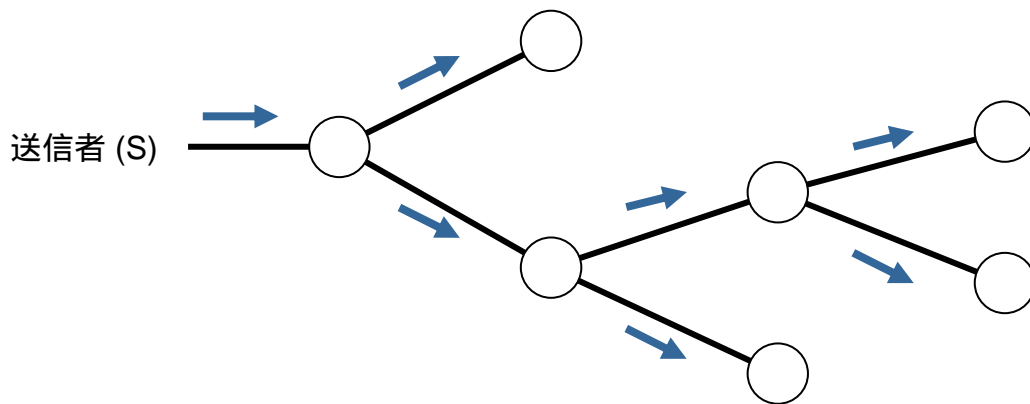
クラスDアドレス:

224.0.0.0 ~ 239.255.255.255

IPマルチキャスト (2)

• Shortest Path Tree と Shared Tree

Shortest Path Tree : (S, G)

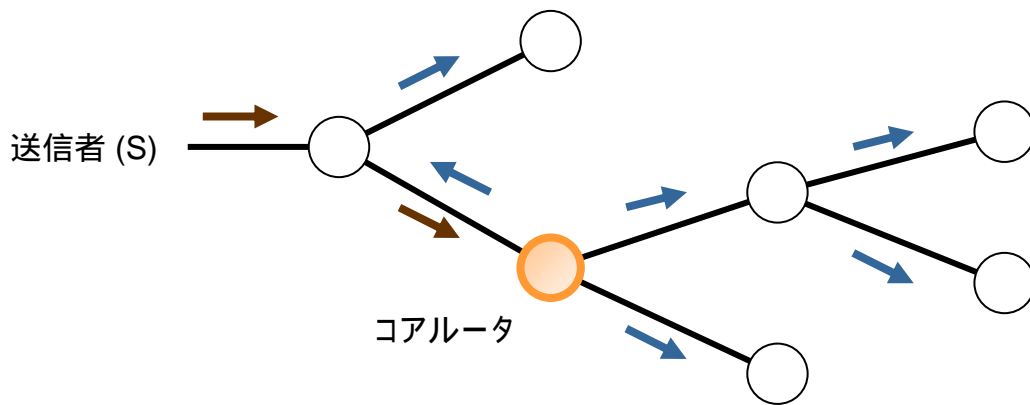


フラディング:

各ルータは、パケットを受信したインタフェース以外のすべてのインタフェースにパケット転送。(S,G) エントリによる経路管理。

下流のルータは、状況に応じて転送停止・再開要求を出し、経路を確定。

Shared Tree : (*, G)



コアルータ:

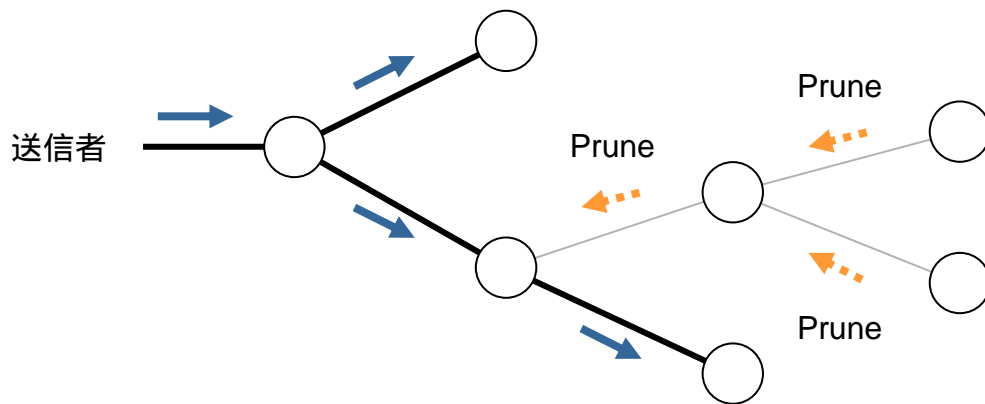
マルチキャストグループ毎に特定のコアルータにパケットをいったん集約。ここまでは、(S, G) エントリによる経路管理。

下流のルータは、必要に応じてコアルータに参加要求を出し、経路を確定。コアルータ以下は、(*, G) エントリによる経路管理。

IPマルチキャスト (3)

• DVMRP version 3

Prune メッセージ

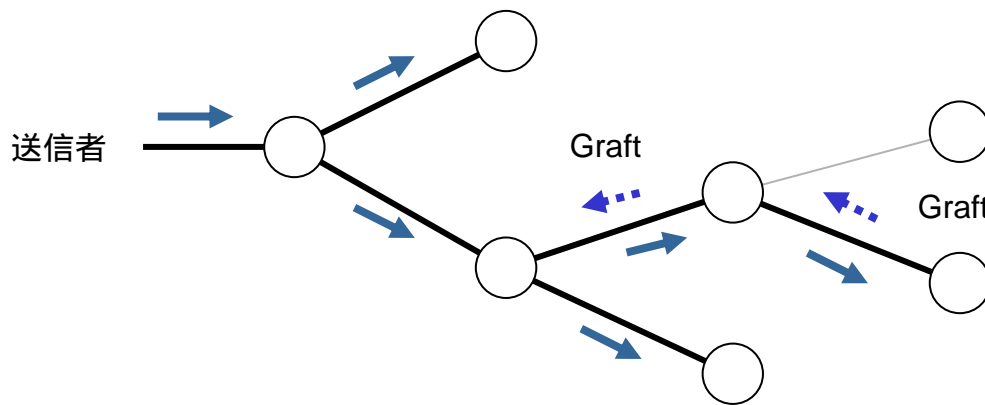


Prune (刈り取り):

下流にマルチキャストグループ参加者がいない場合、上流ルータにパケット配送停止を要求。

途中のルータ: (S, G) エントリ削除。

Graft メッセージ



Graft (接ぎ木):

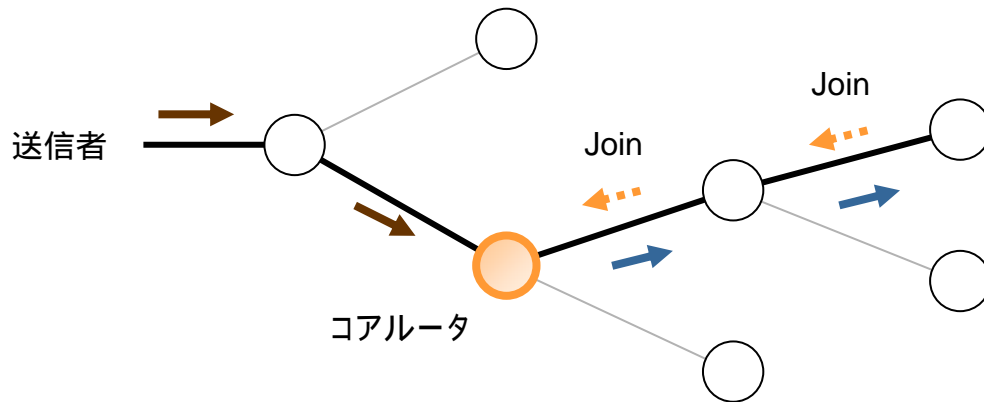
下流にマルチキャストグループ参加者が現れた場合、上流ルータにパケット配送再開を要求。

途中のルータ: (S, G) エントリ追加。

IPマルチキャスト (4)

• PIM-SM

Join メッセージ

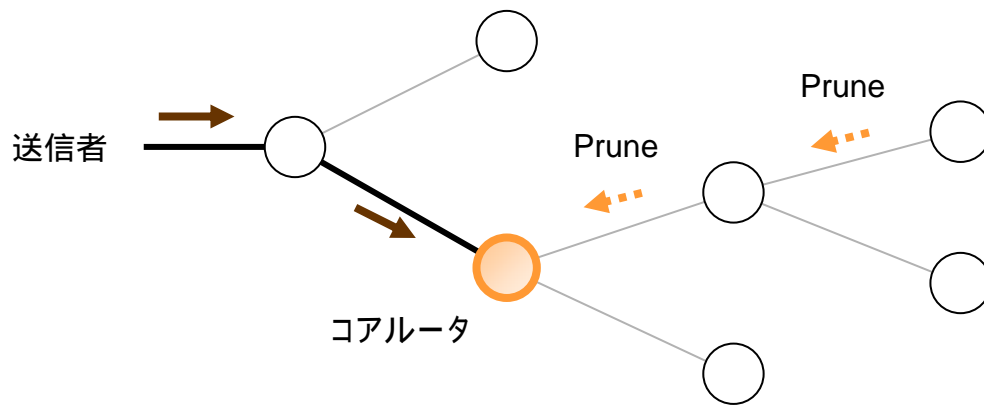


Join (参加):

下流にマルチキャストグループ参加者が現れた場合、上流ルータにパケット配送開始を要求。

途中のルータ: (*, G) エントリ追加。

Prune メッセージ



Prune (離脱):

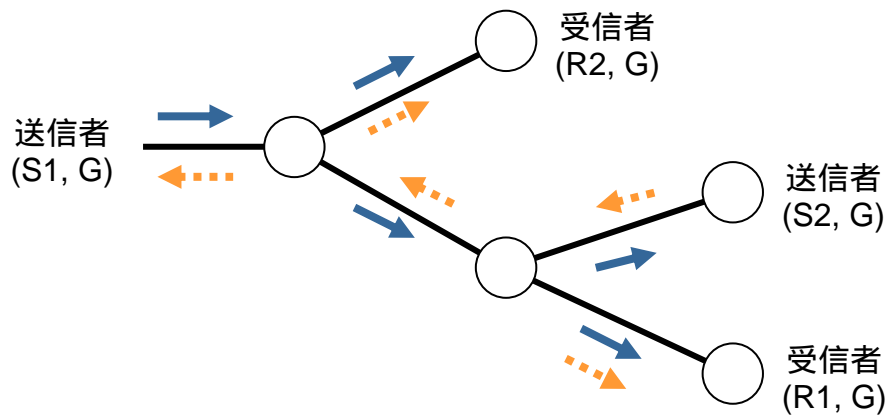
下流のマルチキャストグループ参加者が離脱した場合、上流ルータにパケット配送停止を要求

途中のルータ: (*, G) エントリ削除

IPマルチキャスト (5)

• SSM

Any Source



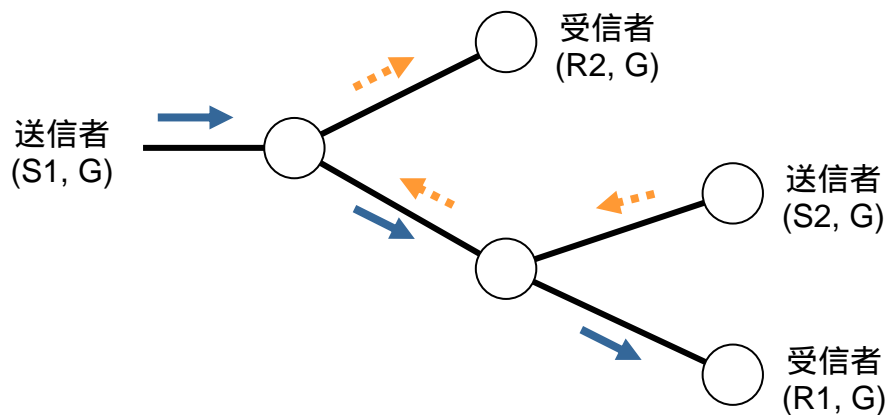
ASM (Any Source Multicast: 従来)

同じマルチキャストアドレス G を使用するセッションのすべての参加者にパケット配信

同じマルチキャストグループに複数の送信者が送信可能 (many-to-many)

多人数会議

Source Specific



SSM:

送信者によって限定される (S, G) セッション
参加者のみにパケット配信

送信者を一人に限定 (one-to-many)

インターネット放送

(232.0.0.0 ~ 232.255.255.255)

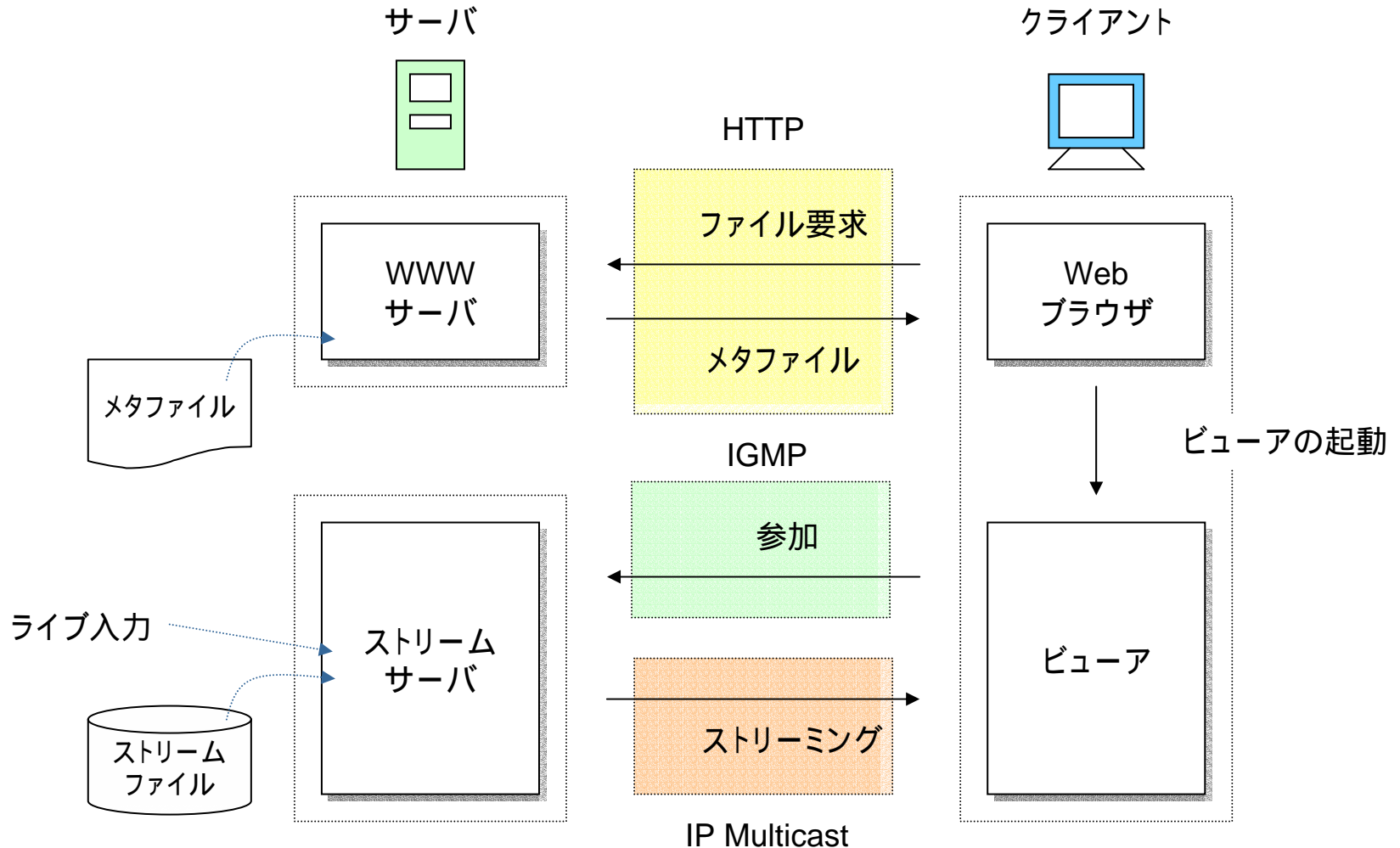
IPマルチキャスト (6)

• まとめ

プロトコル名	特徴	長所	短所
DVMRP	最小経路 (S, G) 送信者がパケットを投げると、フラッディングによって最小経路を確定、配信	最小経路	フラッディングによる不要なトラヒックの増加 拡張性
PIM-SM	送信者・コアルータ: 最小経路 (S, G) コアルータ・受信者: 共有経路 (*, G) 送信者がコアルータに「登録」すると、最小経路を確定 受信者がコアルータに「参加」すると、共有経路を確定、配信	フラッディングが不要 拡張性	共有経路が必ずしも最短経路にならない コアルータの決定方法 プロトコルが若干複雑 (最短経路と共有経路の動的切替え)
SSM	最小経路 (S, G) 受信者が送信者に subscribe すると最小経路を確定、配信	1 対多の放送型アプリケーション PIM-SM とのハイブリッド構成 (PIM-SSM)	1 対多に限定 IGMP v3 が必須

マルチキャスト放送 (1)

• (1) WWW による番組案内

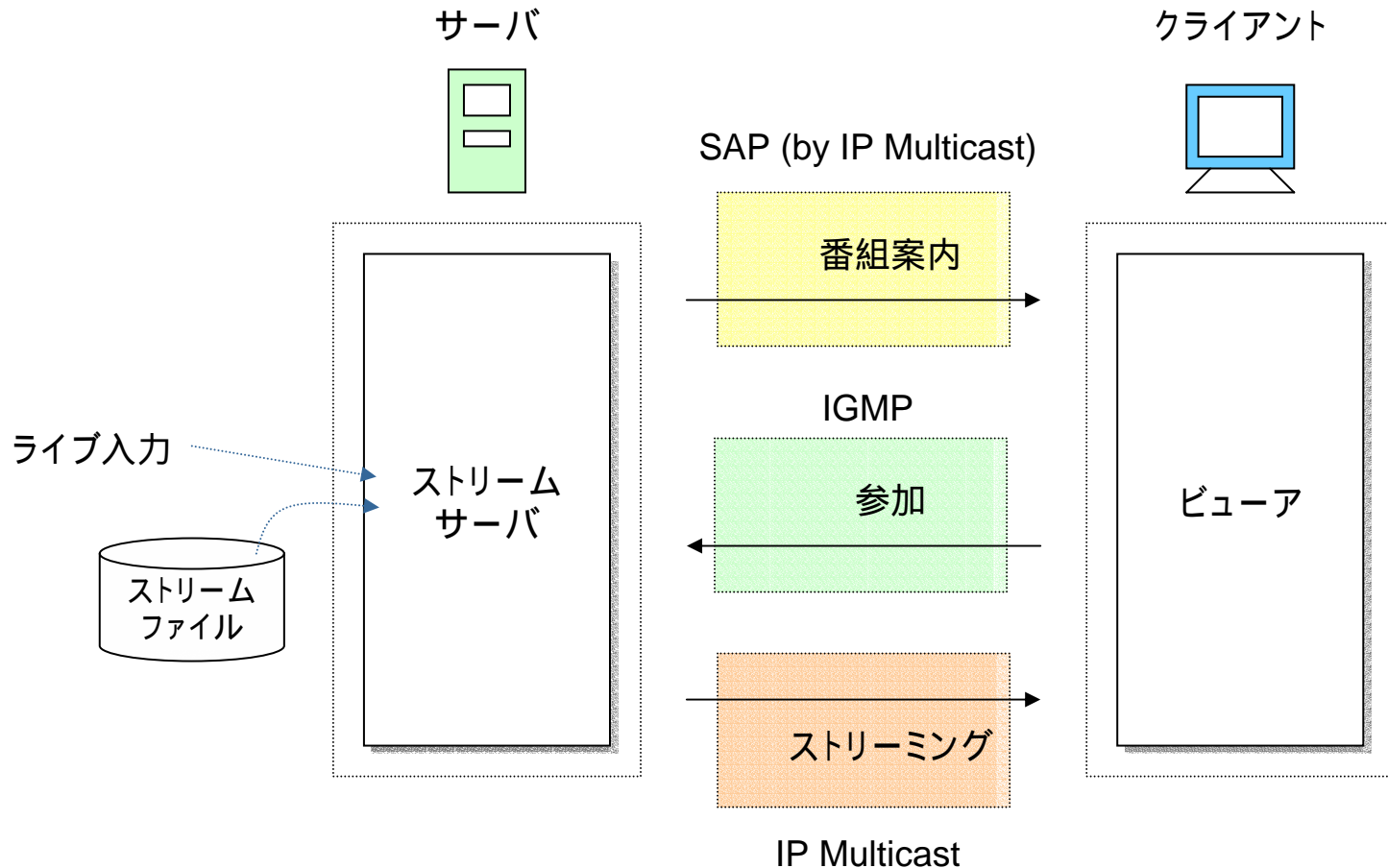


マルチキャスト放送 (2)

• (2) SAP による番組案内

SAP: Session Announcement Protocol

定期的に番組案内 (SDP) をマルチキャスト



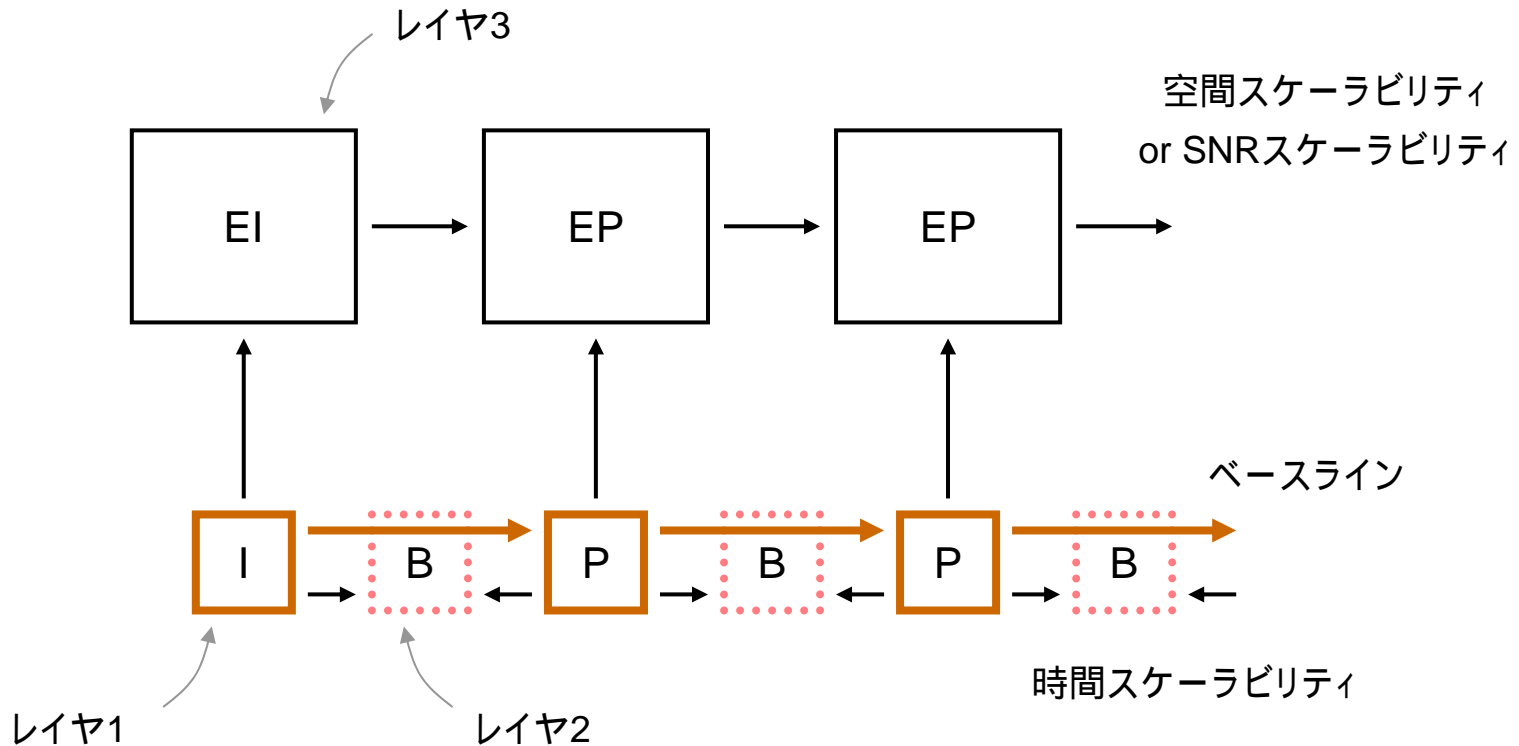
マルチキャスト放送の長所と短所

	ユニキャスト放送	マルチキャスト放送
長所	既存のシステムの変更が不要 クライアントの接続状況に合わせたふくそう制御が可能	トラフィックの削減 (原理的に冗長なパケットは発生しない)、およびサーバ負荷の削減
短所	クライアントの増加に伴うトラフィックの爆発、 ならびにサーバ負荷の増大 (線形増加)	マルチキャストルータの普及と各種設定 クライアント毎のふくそう制御が困難
課題		マルチキャストルーティングプロトコル ふくそう制御アルゴリズム



例：階層化マルチキャスト

スケーラブル符号化



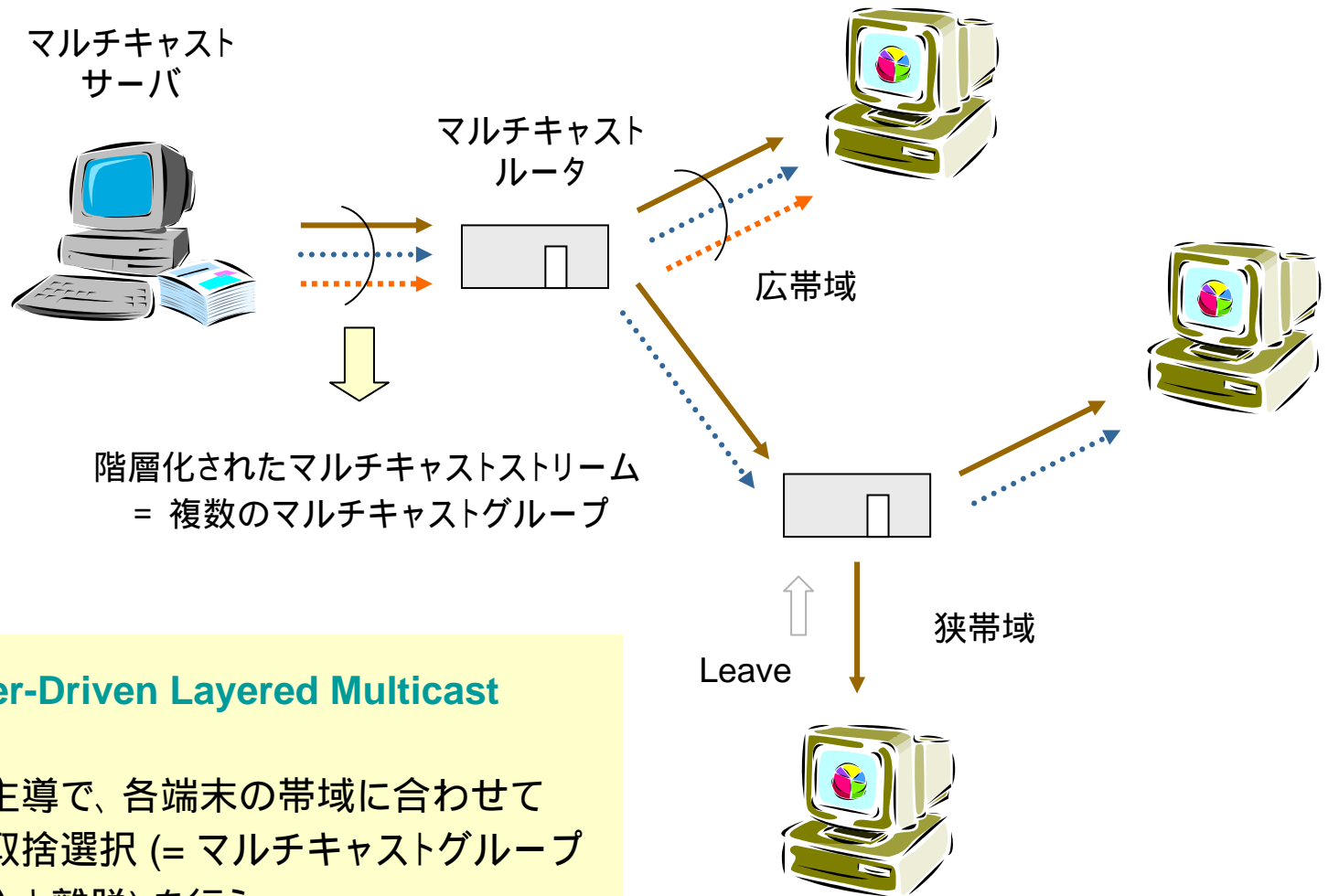
- 空間解像度の階層化: 空間スケラビリティ
- 時間解像度の階層化: 時間スケラビリティ
- SNRの階層化: SNRスケラビリティ

レイヤ1のみ: 低品質、低レート



すべてのレイヤ: 高品質、高レート

階層化マルチキャスト (1)



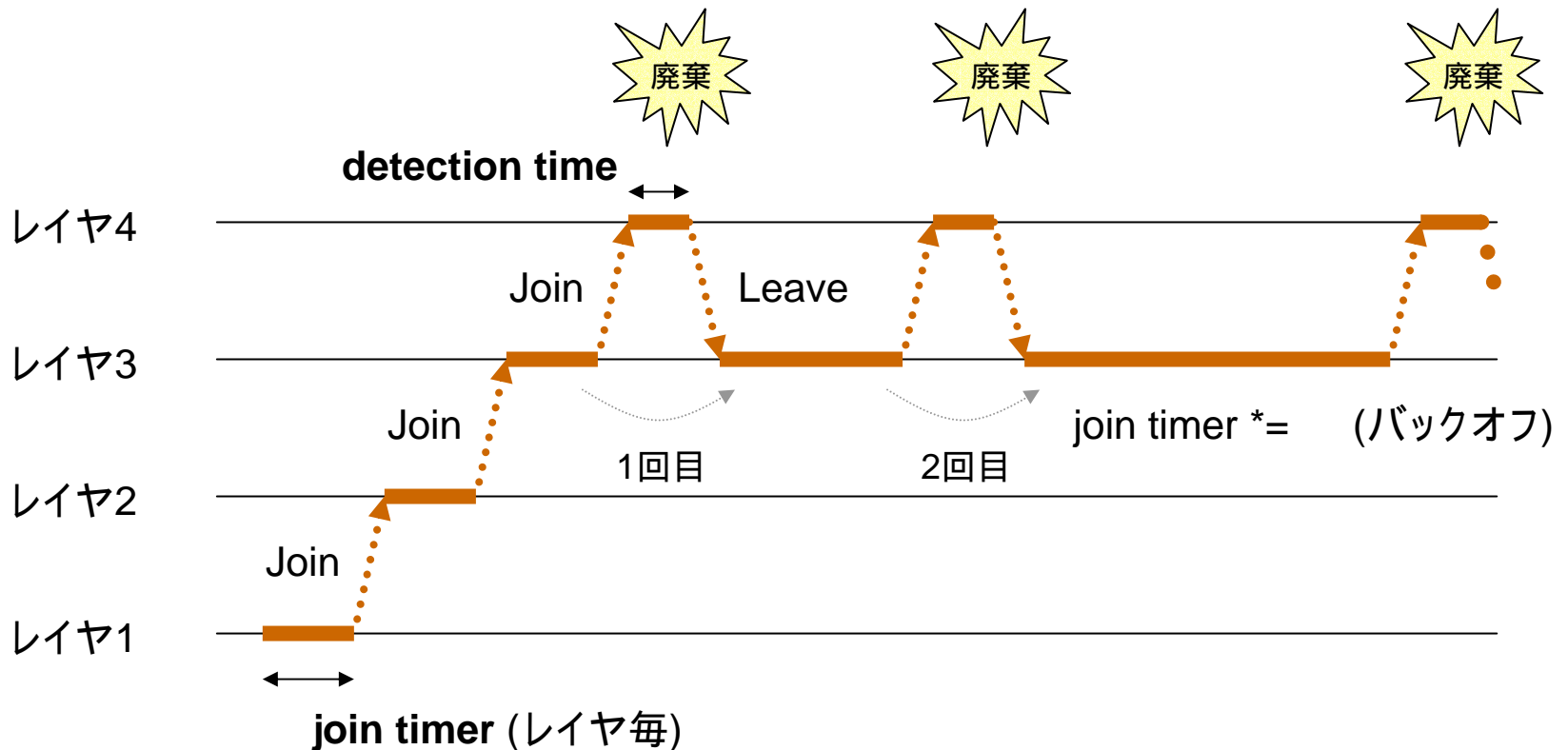
Receiver-Driven Layered Multicast

受信者主導で、各端末の帯域に合わせて階層の取捨選択 (= マルチキャストグループへの加入と離脱) を行う

階層化マルチキャスト (2)

• Join Experiment

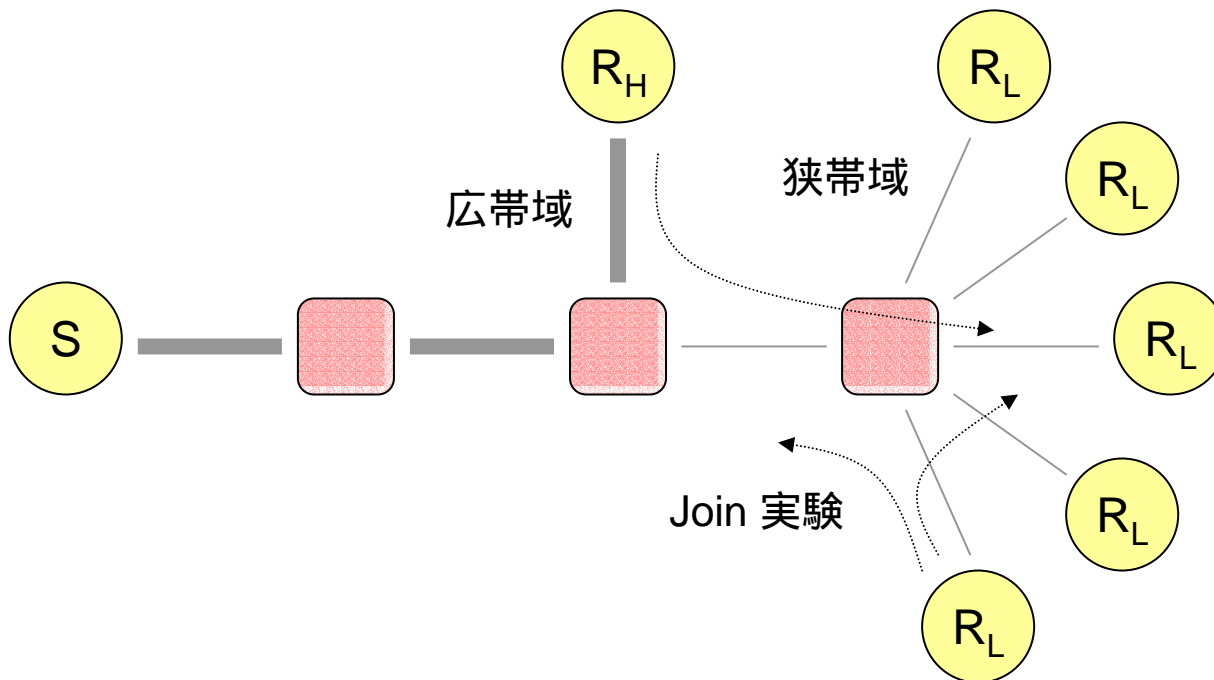
Join、Leave (ふくそう検出)、バックオフを繰り返し、レートを安定させる



階層化マルチキャスト (3)

• Shared Learning

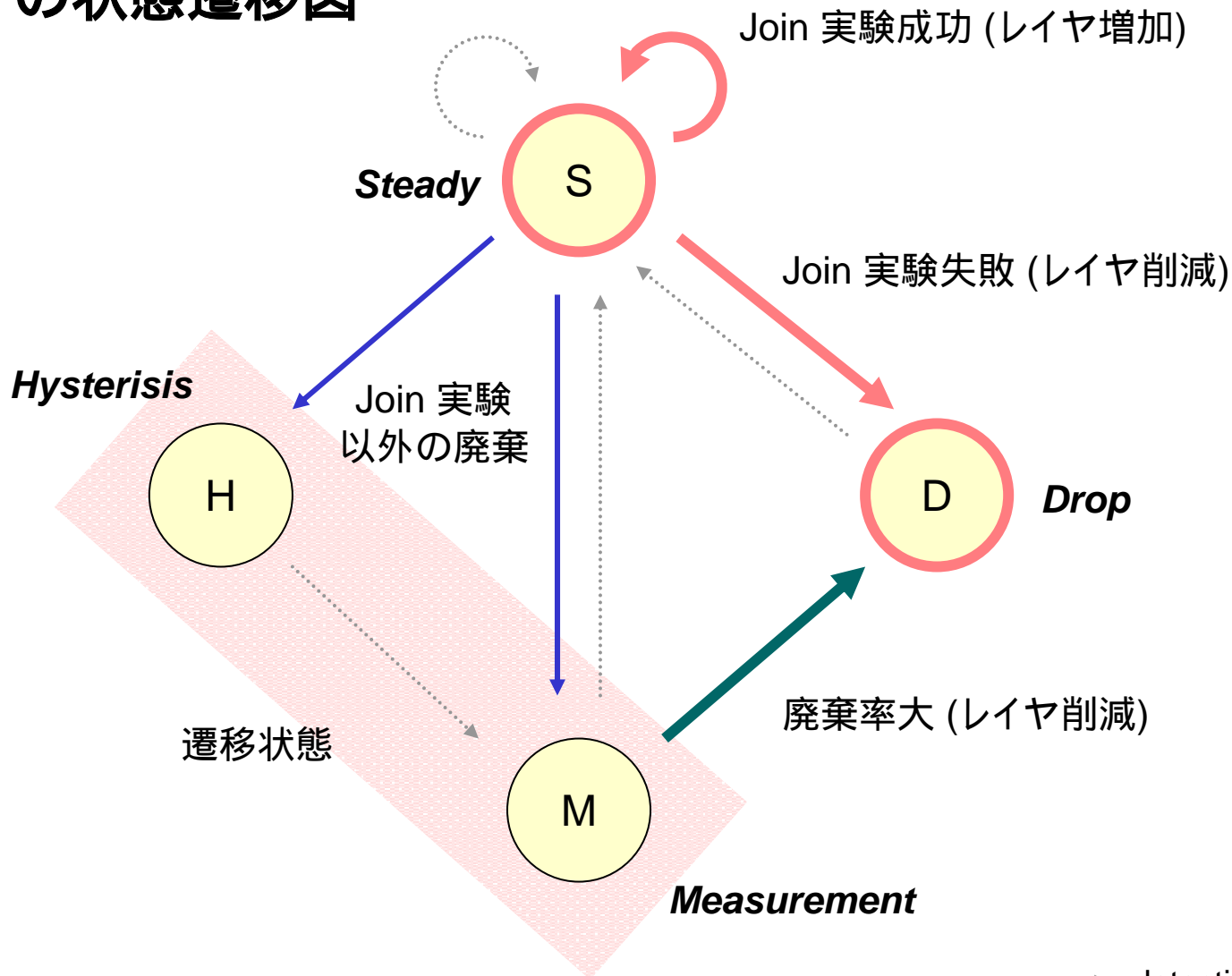
Join 実験の他の端末への通知



- 端末数の増加に伴う Join 実験の回数の増加を防ぐ
- 上流の広帯域 Join 実験と下流の狭帯域 Join 実験の結果の混同を防ぐ

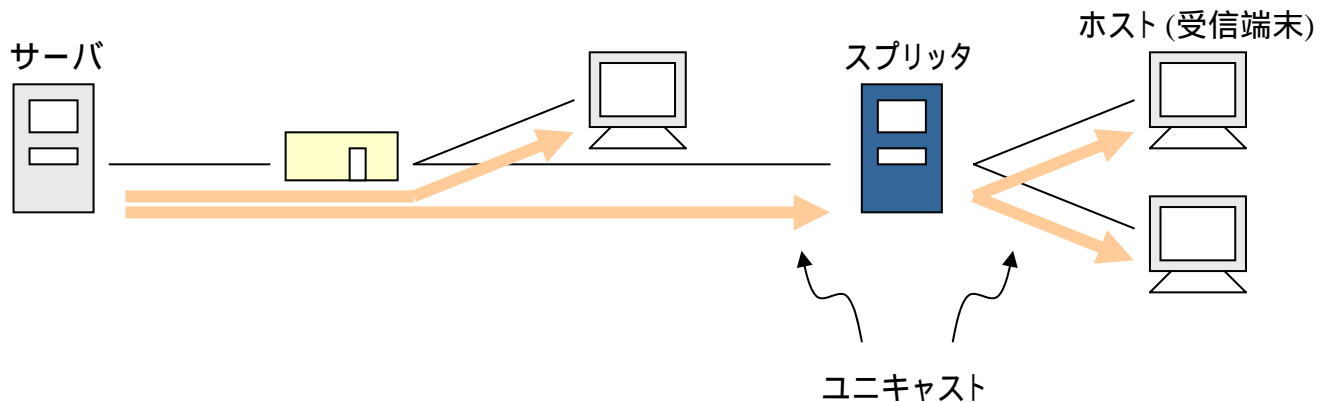
階層化マルチキャスト (4)

• RLM の状態遷移図

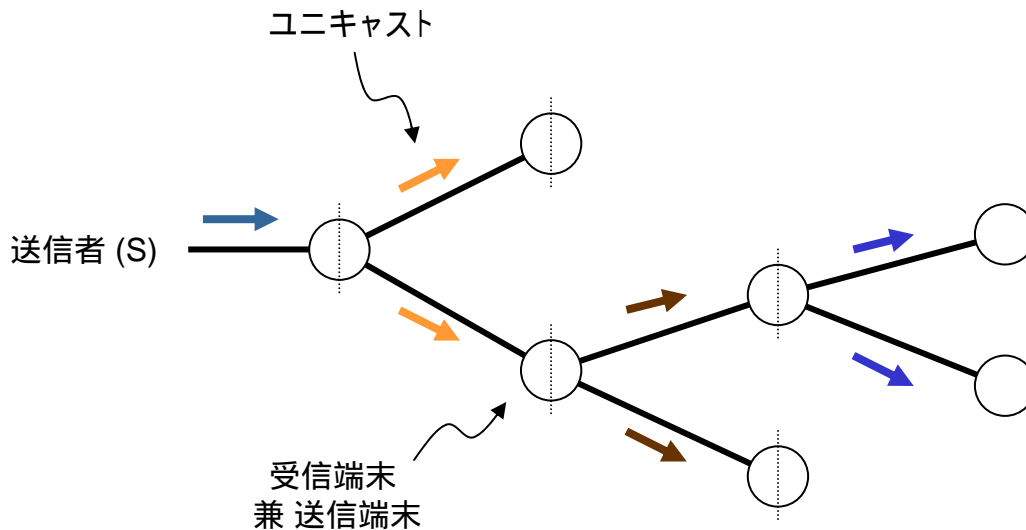


アプリケーション層マルチキャスト (1)

- スプリッタ

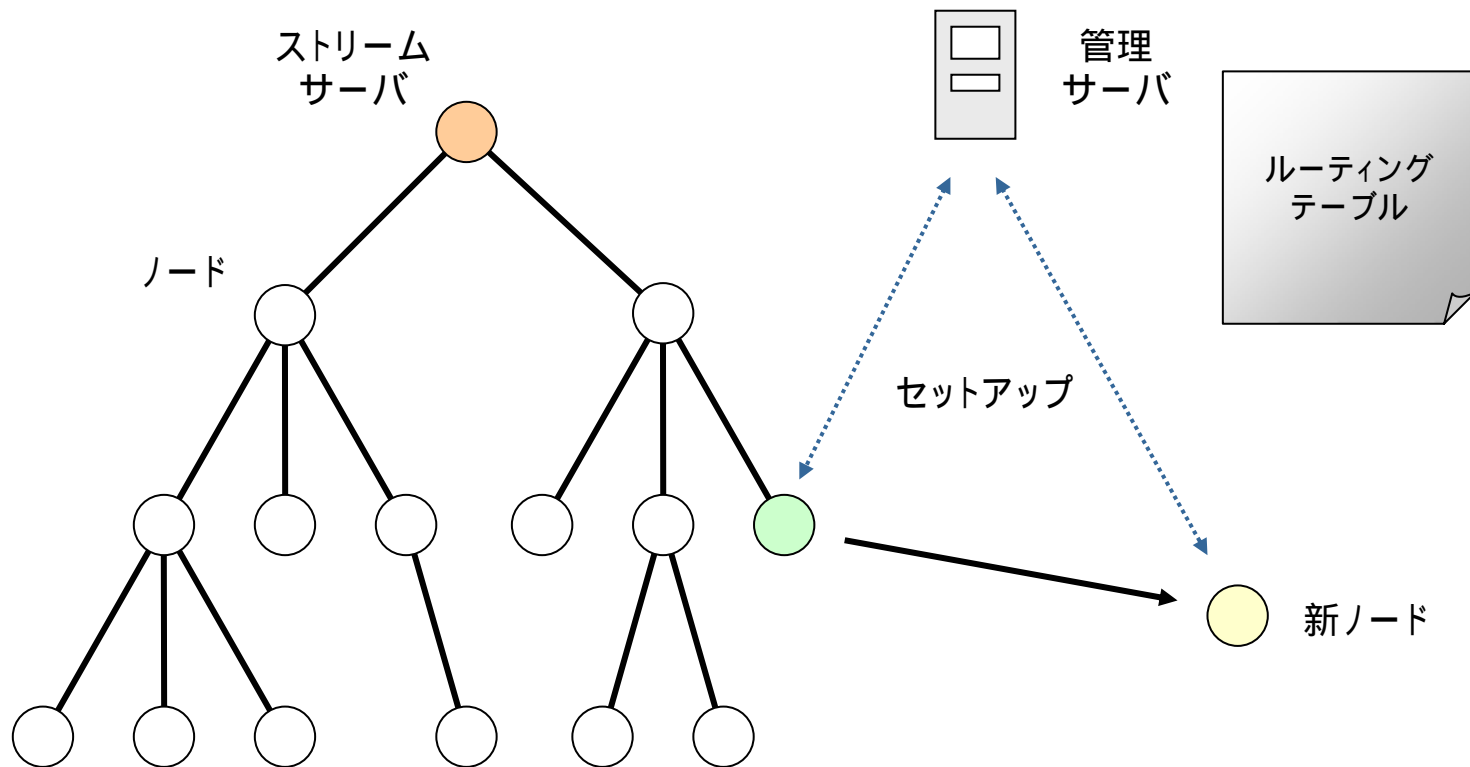


- P2P (Peer-to-Peer)



アプリケーション層マルチキャスト (2)

• P2Pマルチキャスト



長所: 簡単、既存ルータの変更不要

短所: 転送トラヒックの増加、経路の準最適性、管理サーバの負荷

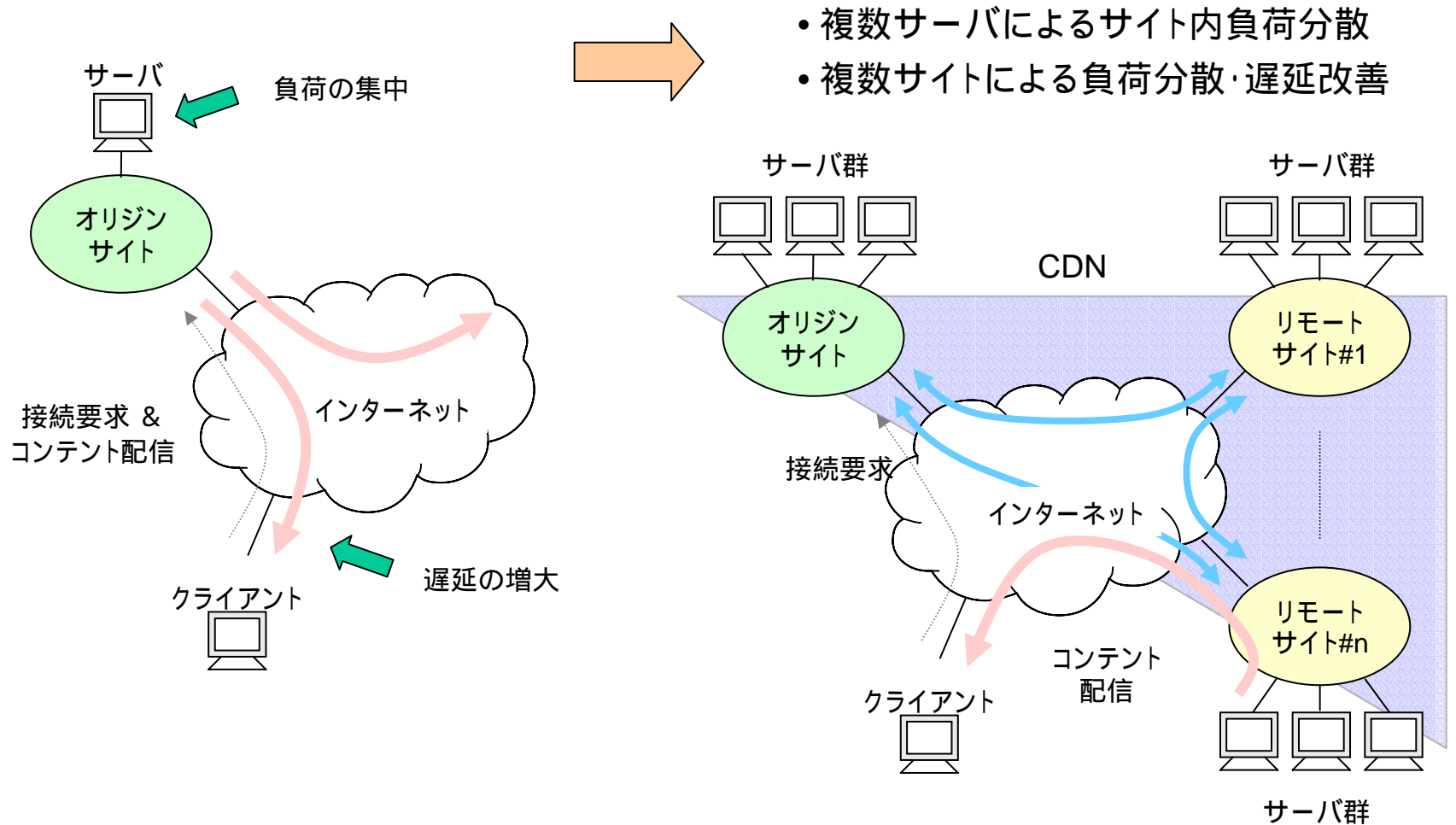
検討事項: ノードの追加と削除への対応、動的な経路変更、負荷分散

CDN

Content Delivery Network

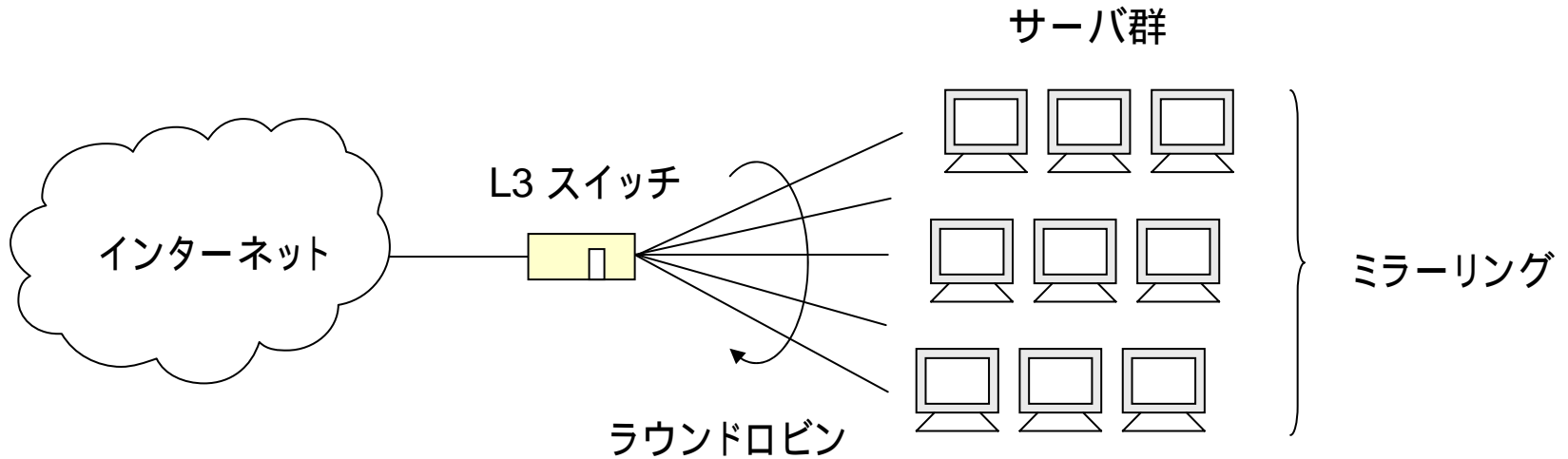
CDN

• サーバの負荷分散 & 転送遅延の改善



サイト内負荷分散 (1)

• L3 スイッチ



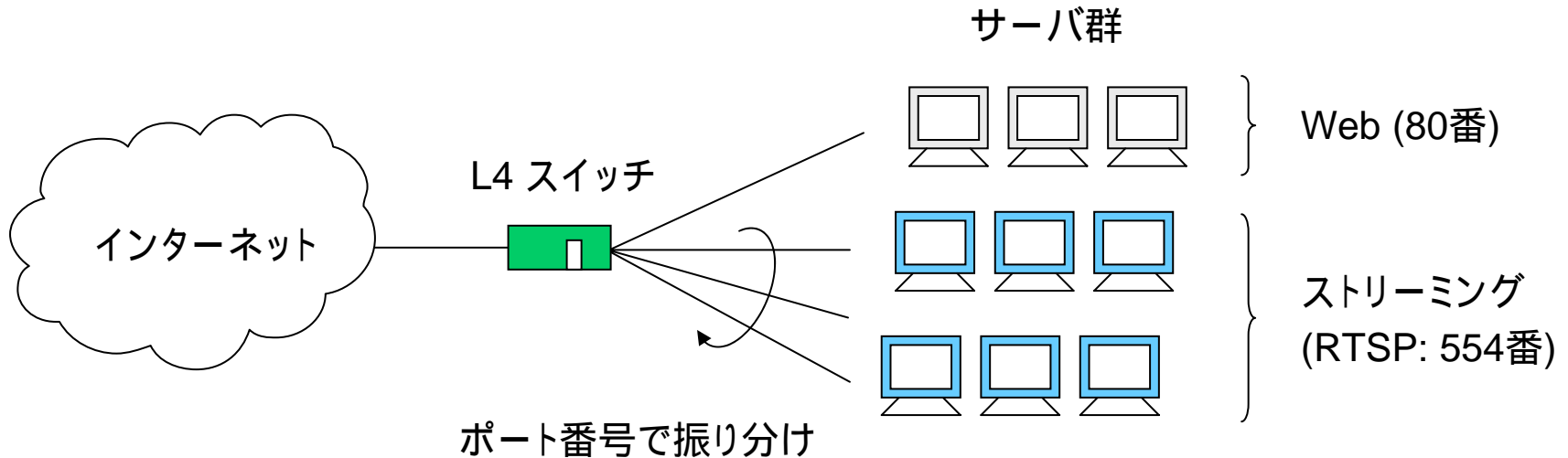
ミラーリングとラウンドロビンによる負荷分散:

長所: スイッチの負荷が軽い

短所: ミラーリングの効率が悪い (すべてのサーバが同じデータを持つ)

サイト内負荷分散 (2)

• L4 スイッチ



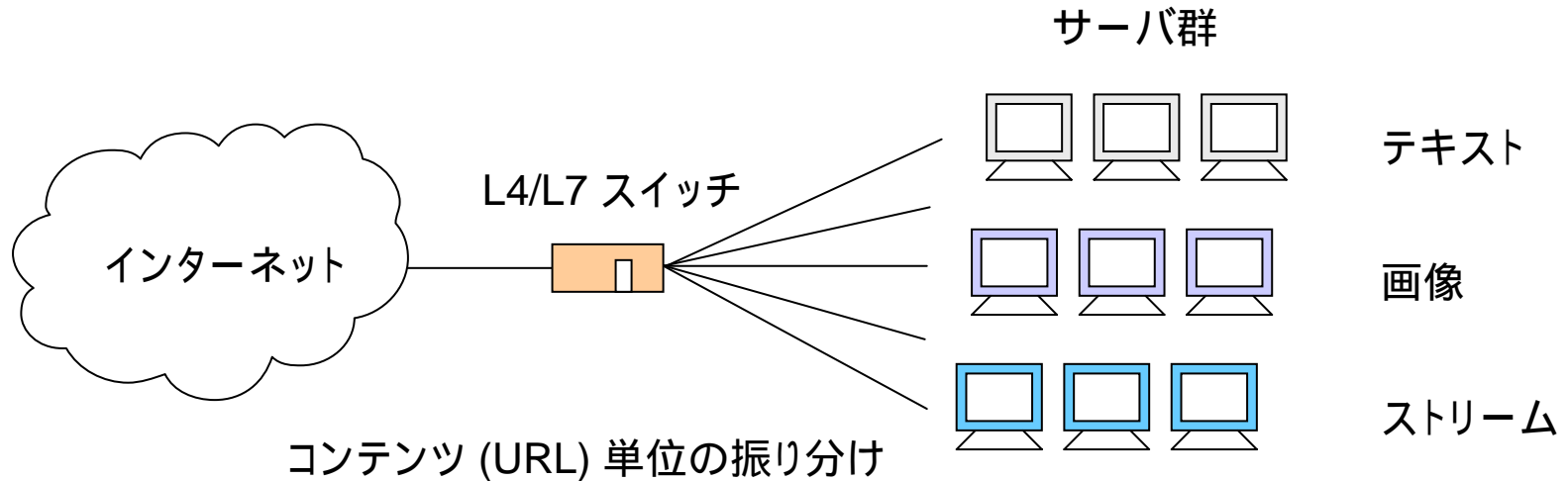
アプリケーション (ポート番号: L4情報) に応じた分散サーバ配置:

長所: アプリケーションに応じたきめこまかい負荷分散が可能

(短所: L3 スイッチよりはスイッチの負荷が大きい)

サイト内負荷分散 (3)

• L4/L7 スイッチ



コンテンツ (URL: L7情報) に応じた分散サーバ配置:

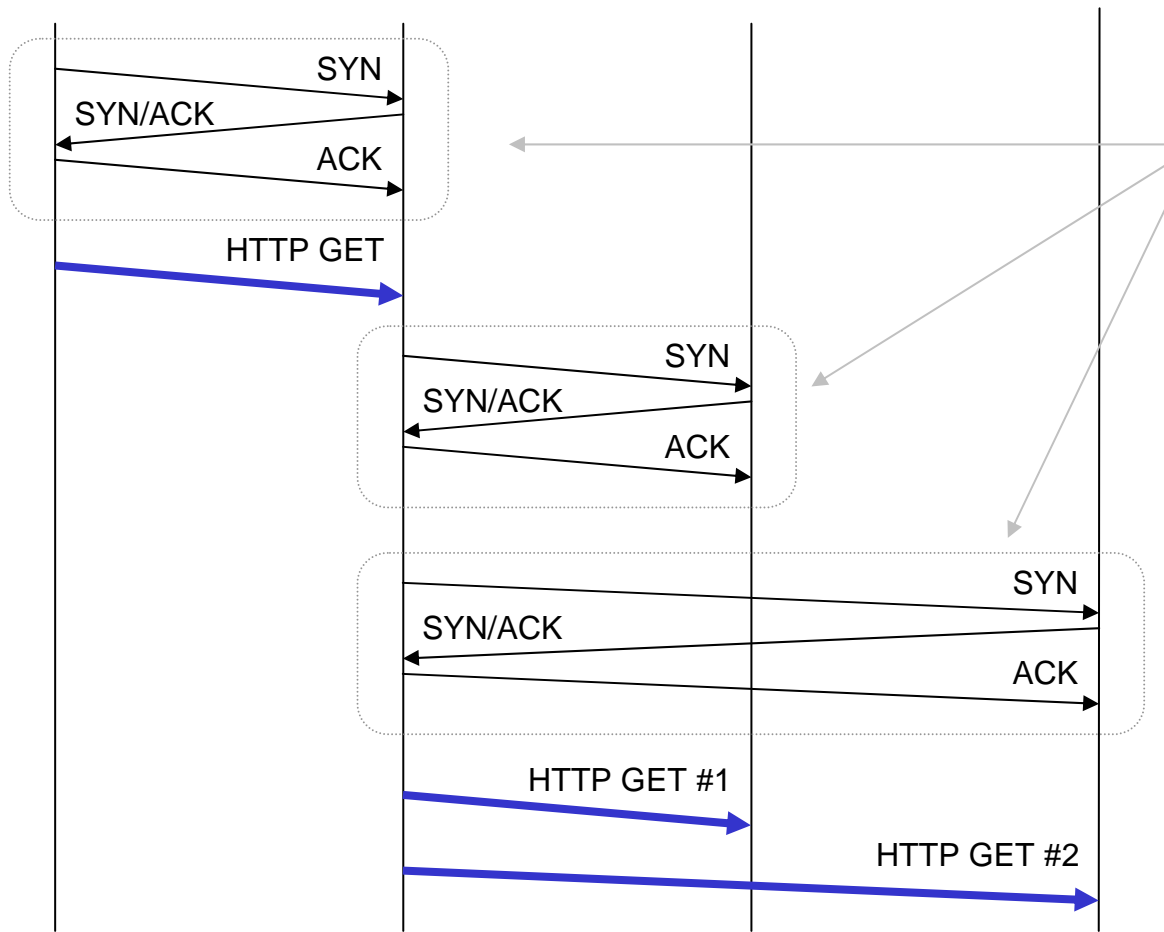
長所: コンテンツ単位のさらにきめこまかい負荷分散が可能

短所: スイッチの負荷が大きい

サイト内負荷分散 (4)

• Delayed Bound (1)

クライアント L4/L7スイッチ サーバ#1 サーバ#2

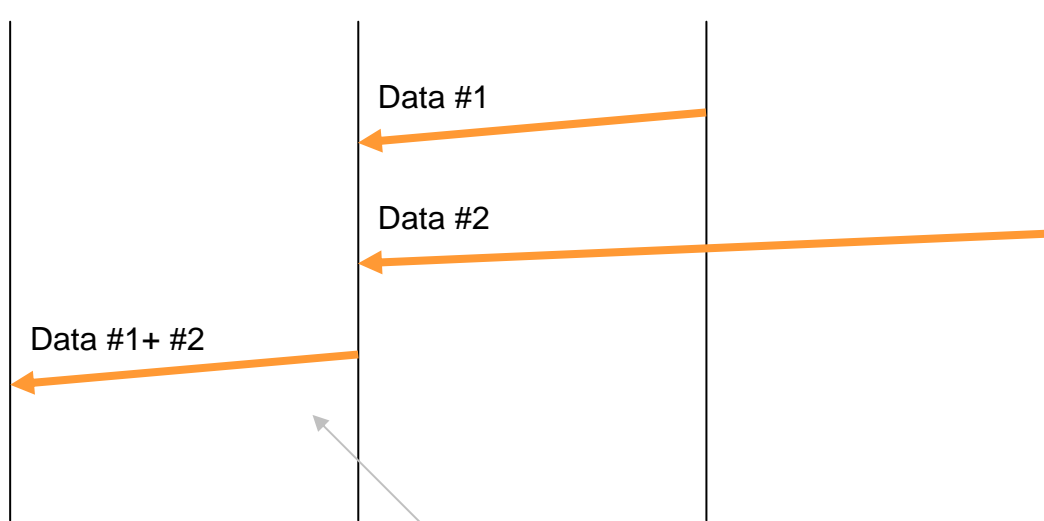


クライアント・スイッチ間、
スイッチ・サーバ間で
複数の TCP コネクション
を終端
= Delayed Bound

サイト内負荷分散 (5)

• Delayed Bound (2)

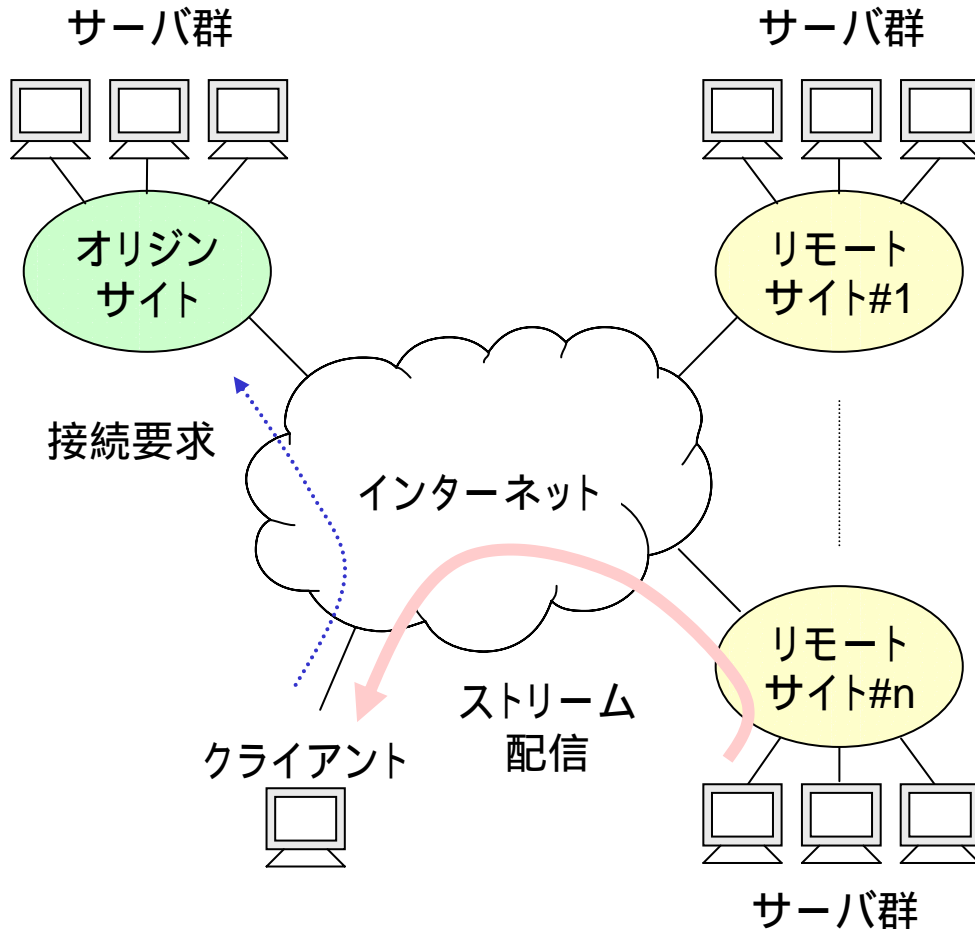
クライアント L4/L7スイッチ サーバ#1 サーバ#2



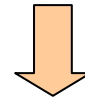
サーバ#1、サーバ#2
からのデータを集約
= Aggregate

サイト間負荷分散

• サイト間負荷分散 & 転送遅延の改善



複数サイト (サーバ群)
の分散配置



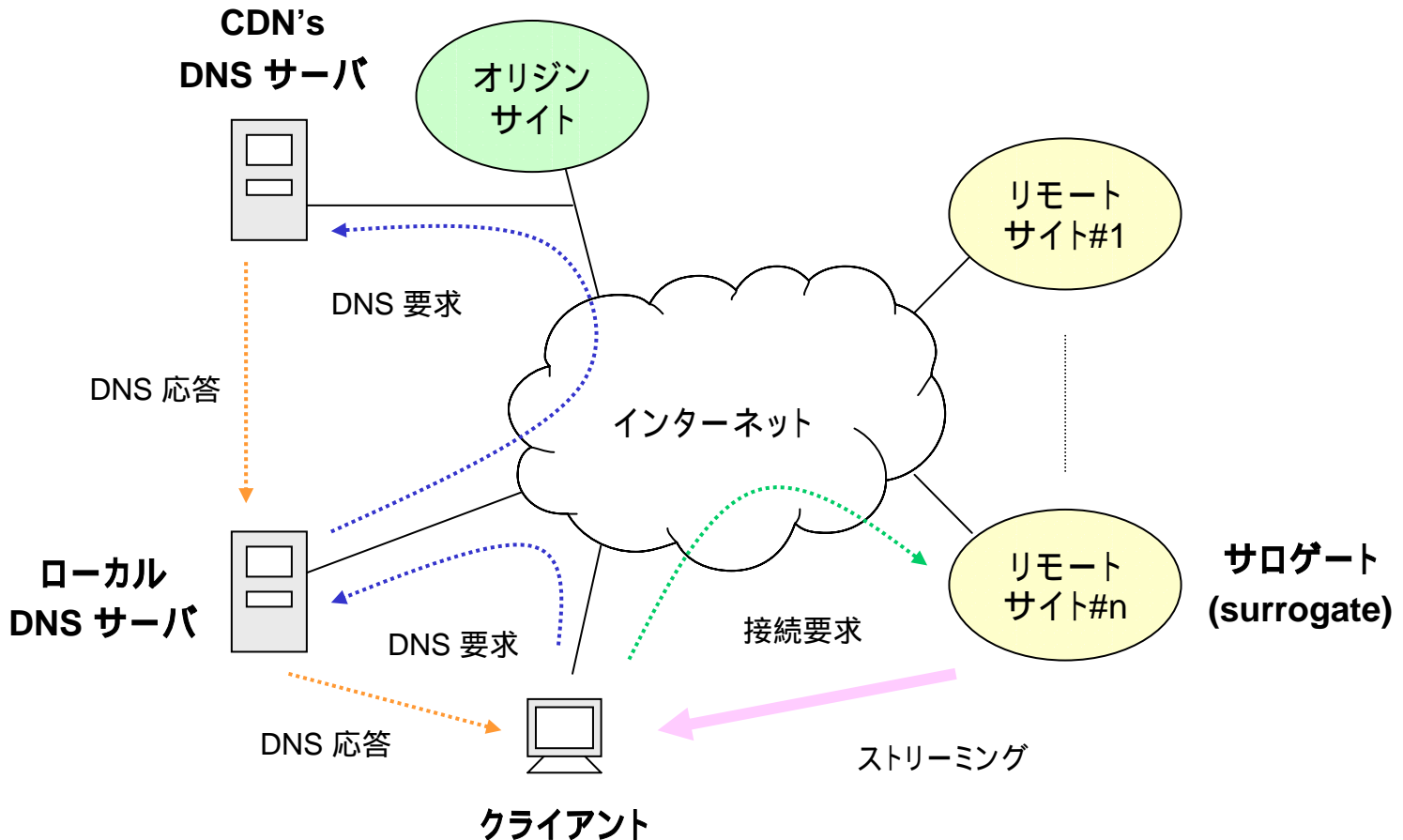
クライアントからの要求
に応じて、適切なサイト
を選択、誘導



サイト間負荷分散 &
転送遅延の改善

リクエストルーティング (1)

• DNS リダイレクション (1)



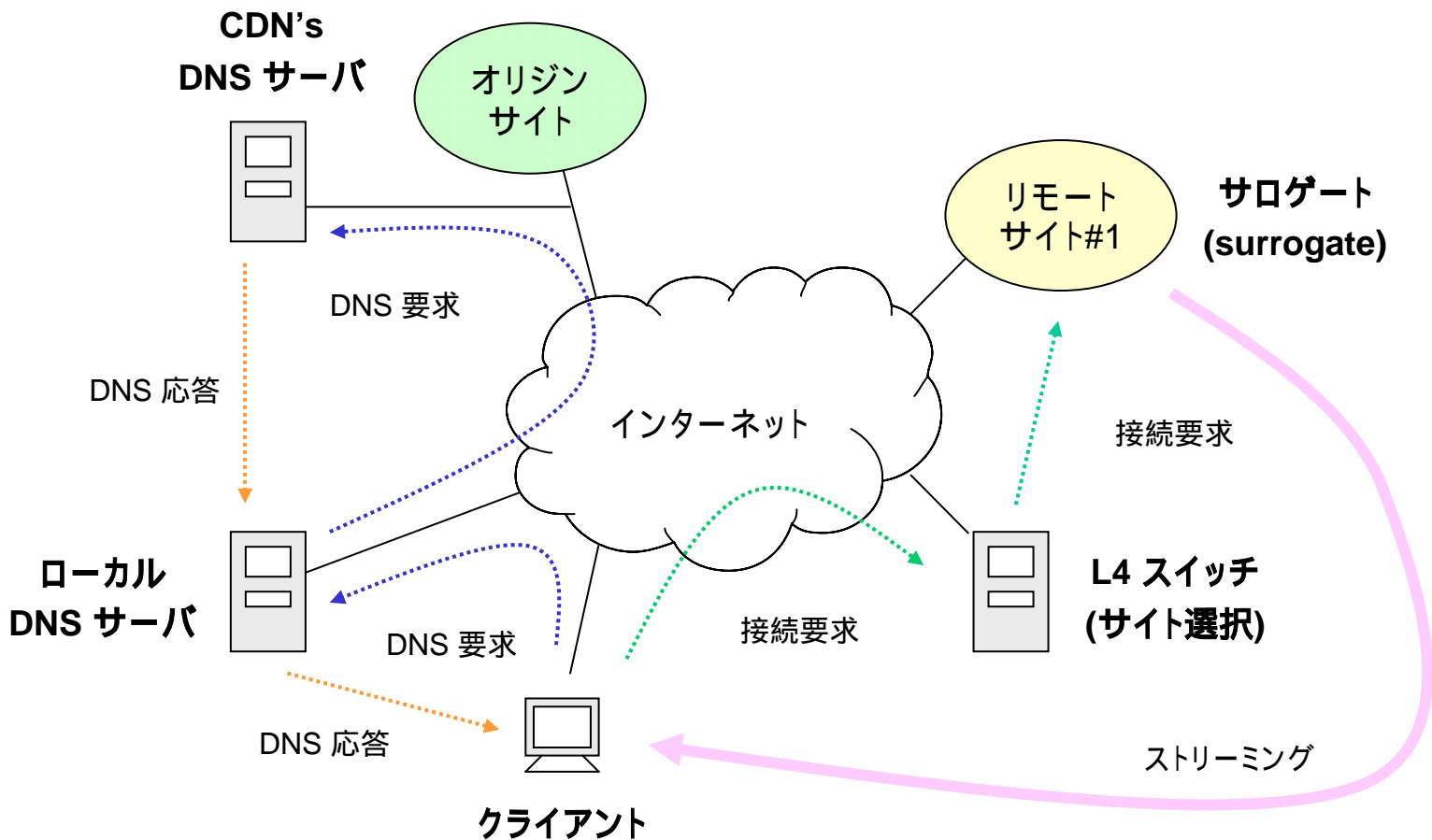
リクエストルーティング (2)

• DNS リダイレクション (2)

DNS リダイレクション	方式
Single Reply	CDN 内 DNS サーバが最適サロゲートを A レコード (IP アドレス) で返す方式 (例: stream.com 192.168.0.1)
Multiple Reply	CDN 内 DNS サーバが複数のサロゲート候補を A レコードで返し、ラウンドロビンでサロゲートを選択する方式 (例: stream.com 192.168.0.1, 192.168.0.2, 192.168.0.3 192.168.0.2)
NS Redirection	CDN 内 DNS サーバが、第三の DNS サーバに NS レコード (ネームサーバ) を返し、その DNS サーバが最適サロゲートを A レコードで返す方式 (例: stream.com server1.site1.stream.com 192.168.0.3)
CNAME Redirection	CDN 内 DNS サーバが、第三の DNS サーバに CNAME レコード (エイリアス) を返し、その DNS サーバが最適サロゲートを A レコードで返す方式 (例: stream.com site1.stream.com 192.168.0.4)
Object Encoding	DNS の名前にオブジェクトのタイプ等を埋め込んでしまい、それに応じてサロゲートの IP アドレスを振り分ける方式 (例: stream.com mpeg_content1.site1.stream.com 192.168.0.5)

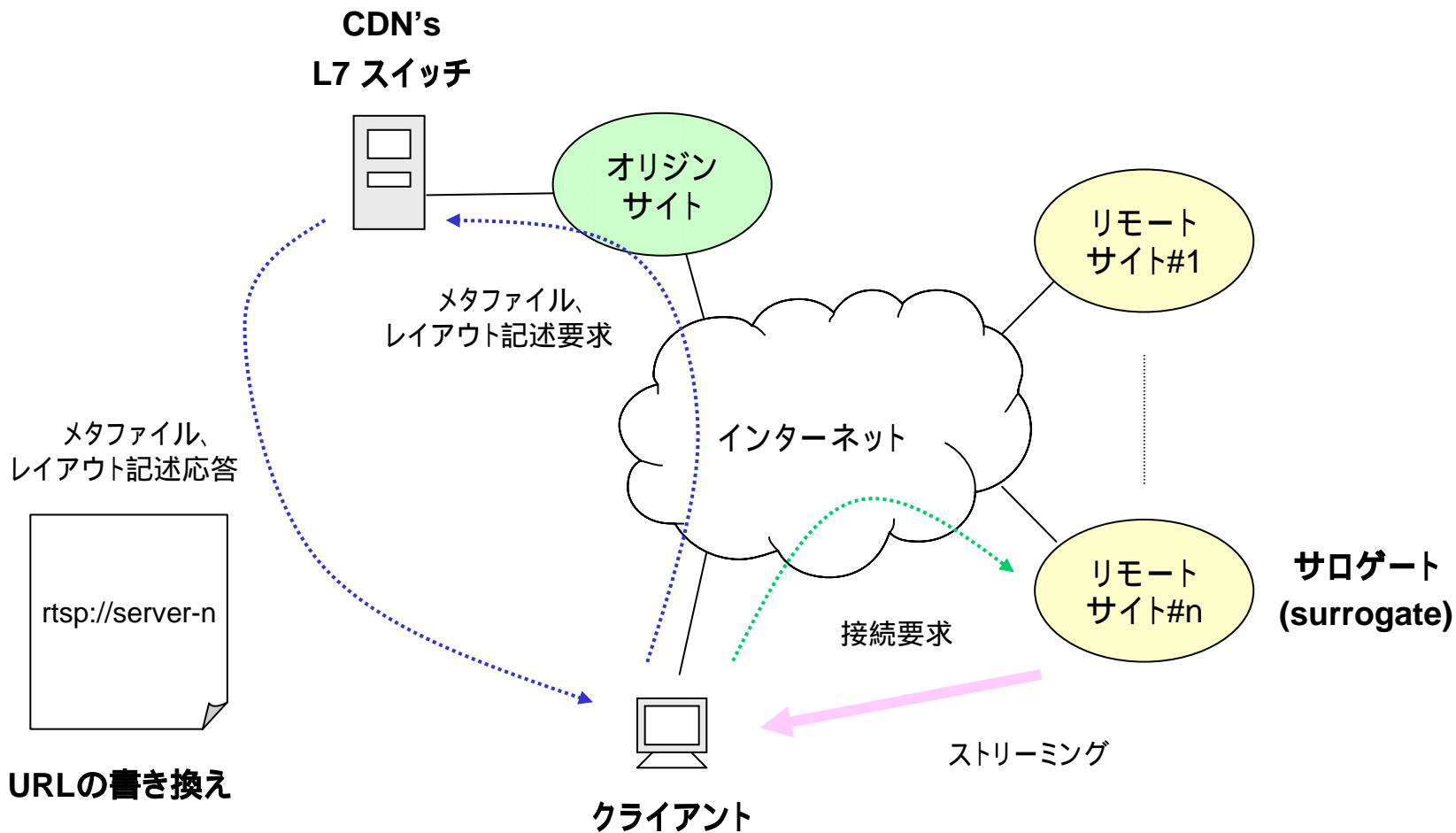
リクエストルーティング (3)

• DNS リダイレクション + L4 スイッチ



リクエストルーティング (4)

• URL リライティング (L7 スイッチ)



リクエストルーティング (5)

• URL リライティング (2)

URL リライティング	方式
Header Inspection (1)	RTSP 記述内に仮想的なサロゲートの URL を記述しておき、アクセスが来たら最適サロゲートへの 302 リダイレクションコードを返す (例) “302” Moved Temporarily
Header Inspection (2)	MIME ヘッダ内の Language、Cookie 等のフィールド情報に応じて、適切なサロゲートへのルーティングを行う (例) stream.com japanese.stream.com
Content Modification	クライアントからのリクエストに応じて、メタファイルやレイアウト記述ファイル内の URL フィールドを最適サロゲートの URL に書き換えて返す (例) rtsp://stream.com rtsp://site1.stream.com

リクエストルーティング (6)

• 最適サロゲートの推定方法

推定方法	方式
Proximity Measurement	クライアントに最も近いサロゲートの推定方法 (1) Active Probing : ping 等のプローブパケットの利用 (2) Passive Measurement : クライアントパケットのモニタリング 基準: 遅延、パケットロス、ホップ数、等 関連分野: インターネットの帯域測定技術
Surrogate Feedback	管理サーバとサロゲートの情報交換: エージェントを用いた Probing 基準: CPU 負荷、インターフェース負荷、コネクション数、等 関連分野: 負荷分散技術

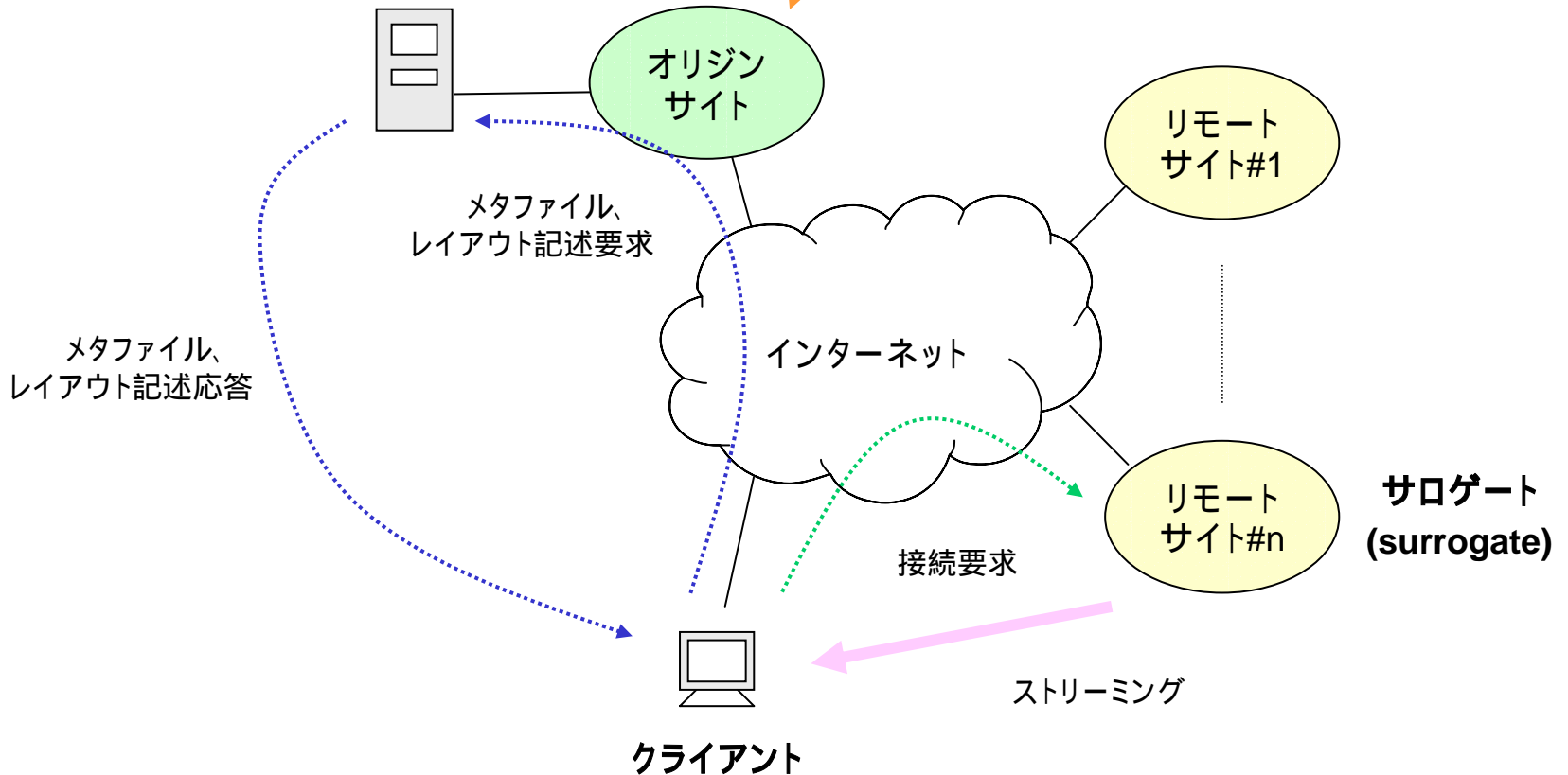
リクエストルーティング (7)

• Akamai FreeFlow

rtsp://x.akamai.net/y/stream.com/content1.rm

兼、課金情報

L7 スイッチ



リクエストルーティング (8)

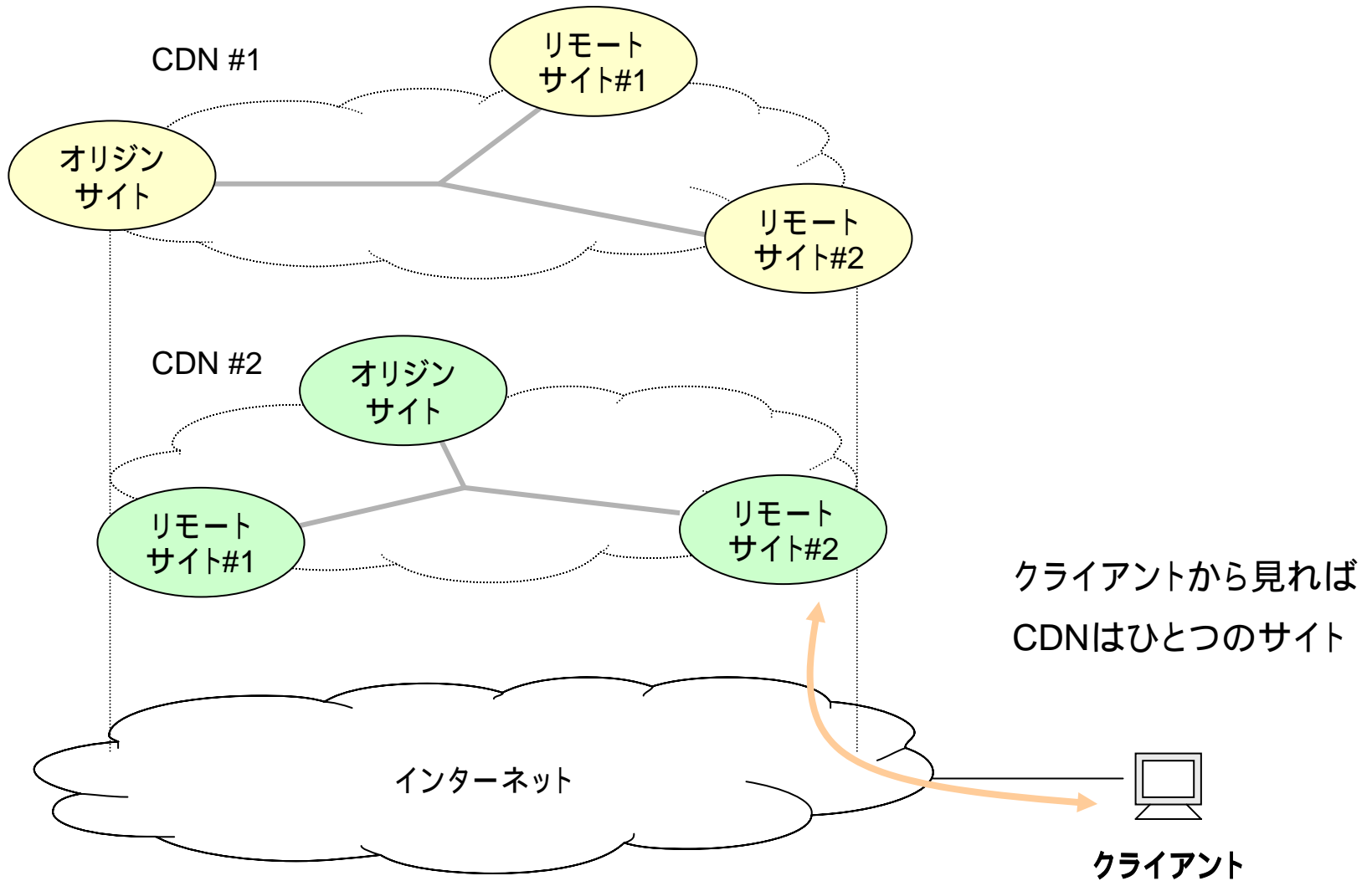
• CDN ベンダの例

CDNベンダ	コンテンツルーティング
Adero	DNSリダイレクション (full-site)
Akamai	ハイブリッド
Clearway	URLリライト
Digital Island	DNSリダイレクション (partial-site)
Fasttide	ハイブリッド
Mirror Image	DNSリダイレクション (full-site)
NetCaching	DNSリダイレクション (full-site)
Solidspeed	DNSリダイレクション (partial-site)
Speedera	DNSリダイレクション (partial-site)
Unitech Networks	DNSリダイレクション (full-site)

full-site: リモートサイトがオリジンサイトの完全ミラー

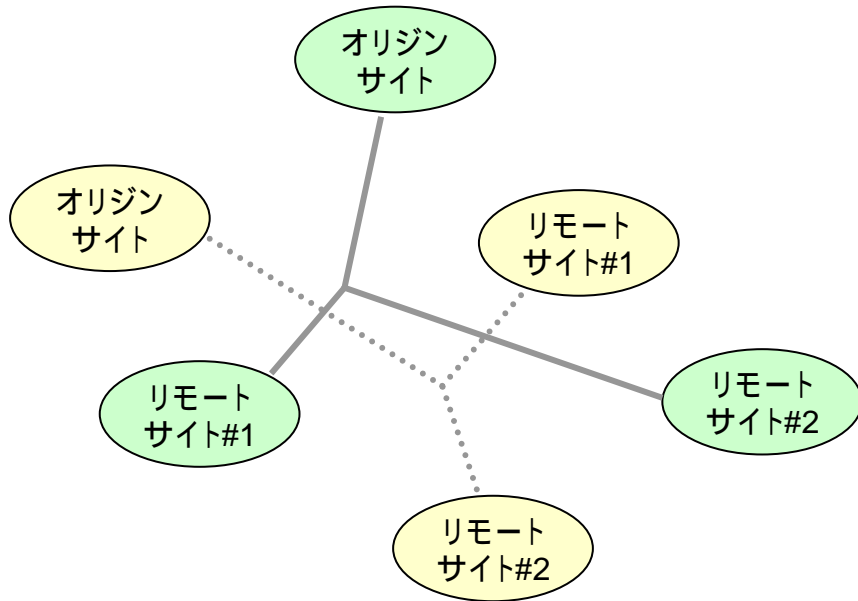
partial-site: リモートサイトがオリジンサイトの部分ミラー

オーバーレイネットワーク

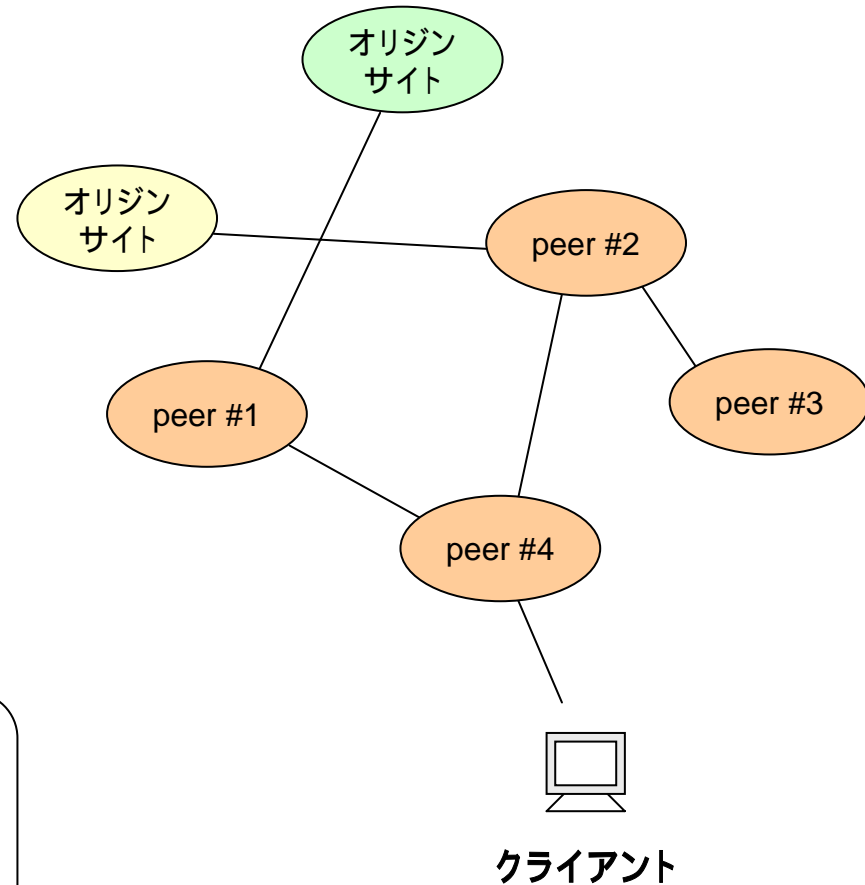


P2P CDN

複数のCDN



P2P CDN



新しい参加者をどの peer に接続するか
peer 間の経路をどのように更新するか

...

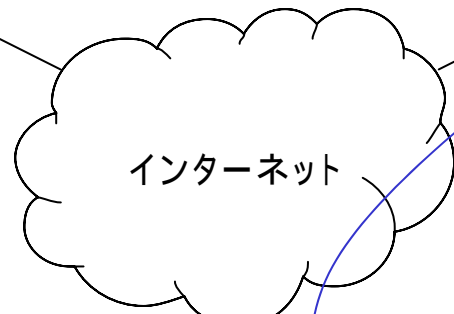
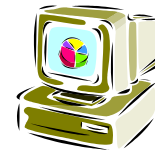
モビリティ

Mobile IP (1)

• Mobile IPv4 (制御)

CN (Corresponding Node)

HA (Home Agent)



Registration (CoA)

FA (Foreign Agent)



移動

MN (Mobile Node)

- ホームアドレス (MN: ホーム)
- 気付アドレス (CoA: 移動先)

Router Advertisement

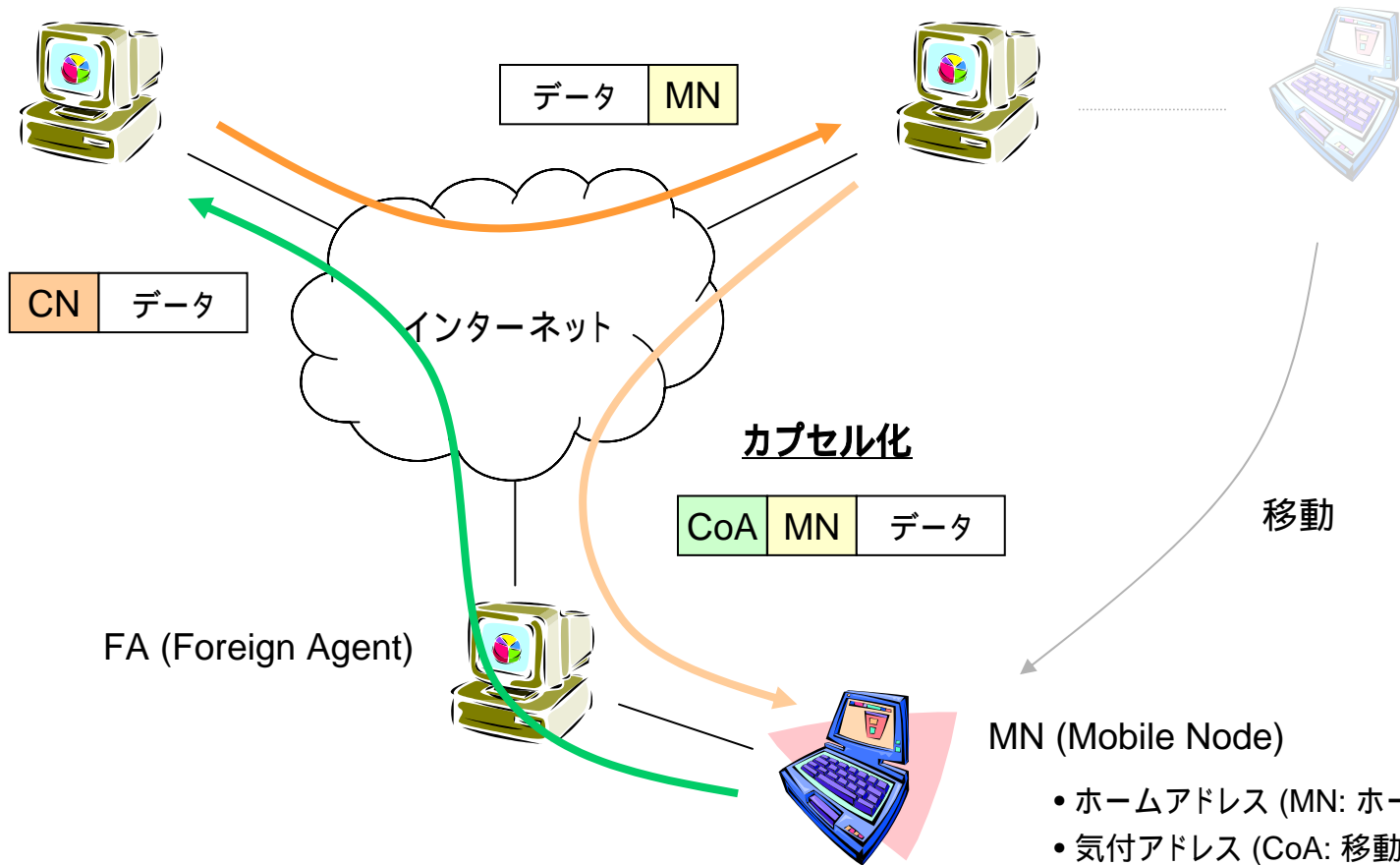


Mobile IP (2)

• Mobile IPv4 (データ)

CN (Corresponding Node)

HA (Home Agent)



MN (Mobile Node)

- ホームアドレス (MN: ホーム)
- 気付アドレス (CoA: 移動先)

Mobile IP (3)

• Mobile IPv4 (定義と手順)

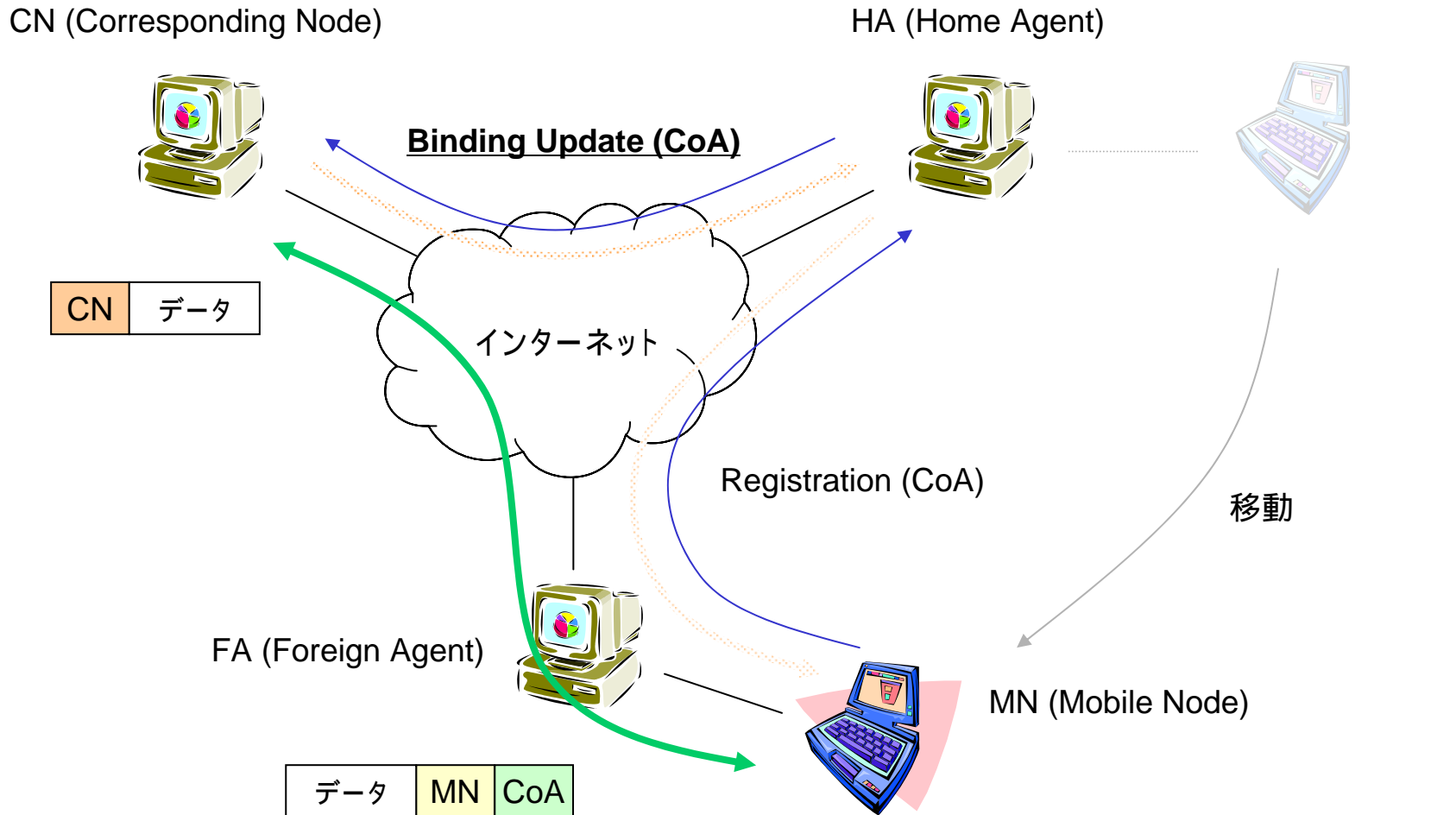
- 定義:
- MN (Mobile Node): 移動端末
 - CoA (Care of Address): 気付アドレス (共存気付と外部気付)
 - HA (Home Agent): 移動元エージェント
 - FA (Foreign Agent): 移動先エージェント
 - CN (Corresponding Node): 通信相手

共存気付アドレスの場合:

- MN が FA から CoA をもらう (**Discovery**: Advertisement, DHCP 等)。
- MN が HA に CoA を登録する (**Registration**)。
- CN からのパケットを HA が MN にカプセル化転送する (**Delivery**)。
- MN は、受信パケットのカプセル化をほどきデータを受信。
- MN は、送信元アドレスは MN のまま、CN に対してパケットを送信。

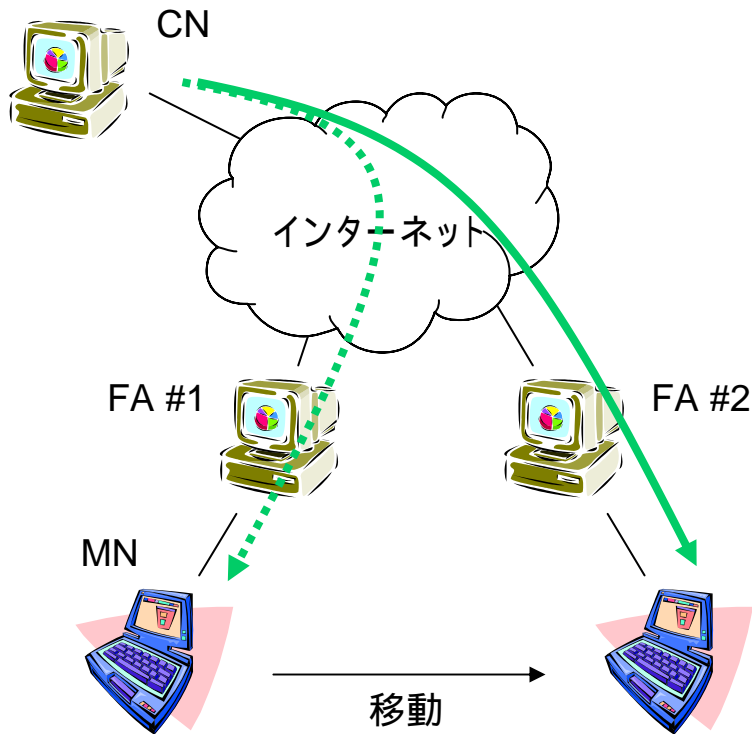
Mobile IP (4)

• Route Optimization (三角経路の回避オプション)



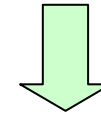
Mobile IP (5)

• Fast Handover (1)



課題:

- (1) シグナリング遅延の削減
- (2) パケットロスの削減

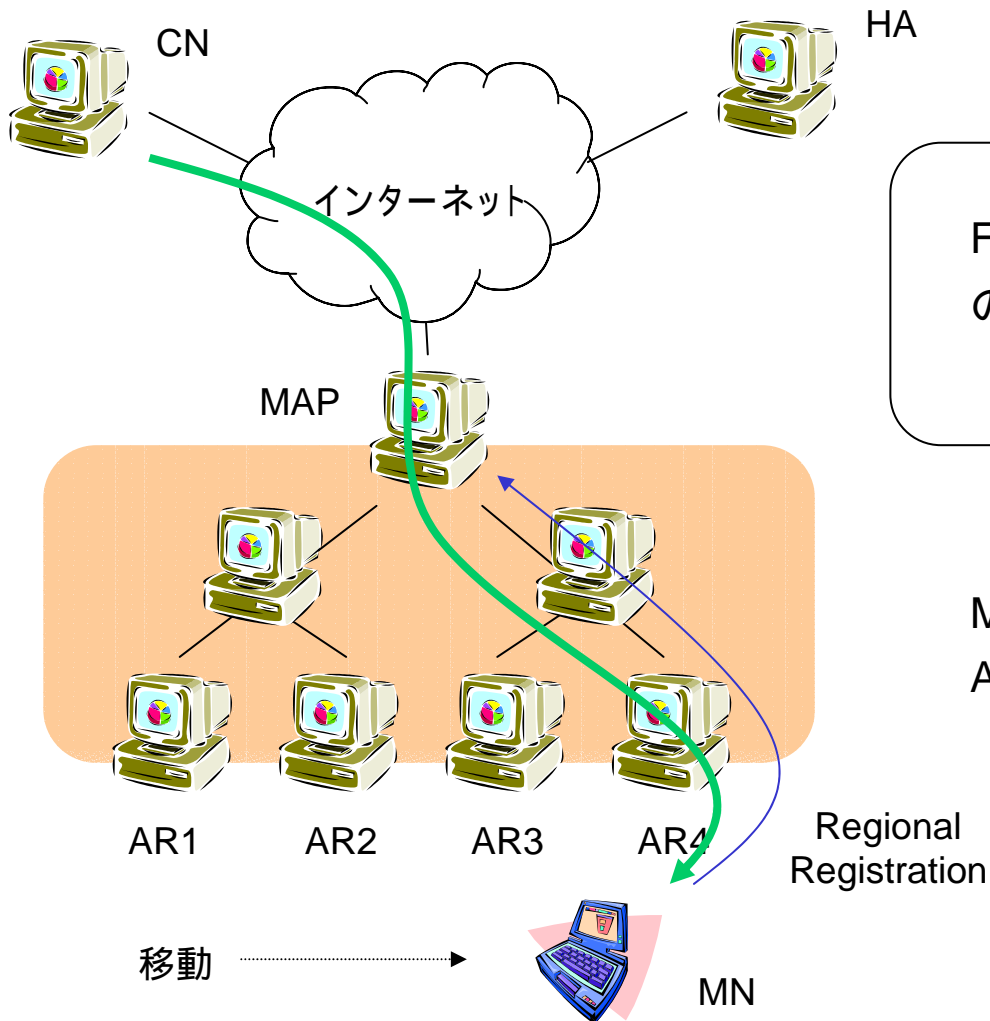


対策:

- (1) 階層化 MIP (HMIP)
- (2) バッファリング・トンネリング
- (3) バイキャストイング

Mobile IP (6)

• Fast Handover (2) 階層化 MIP



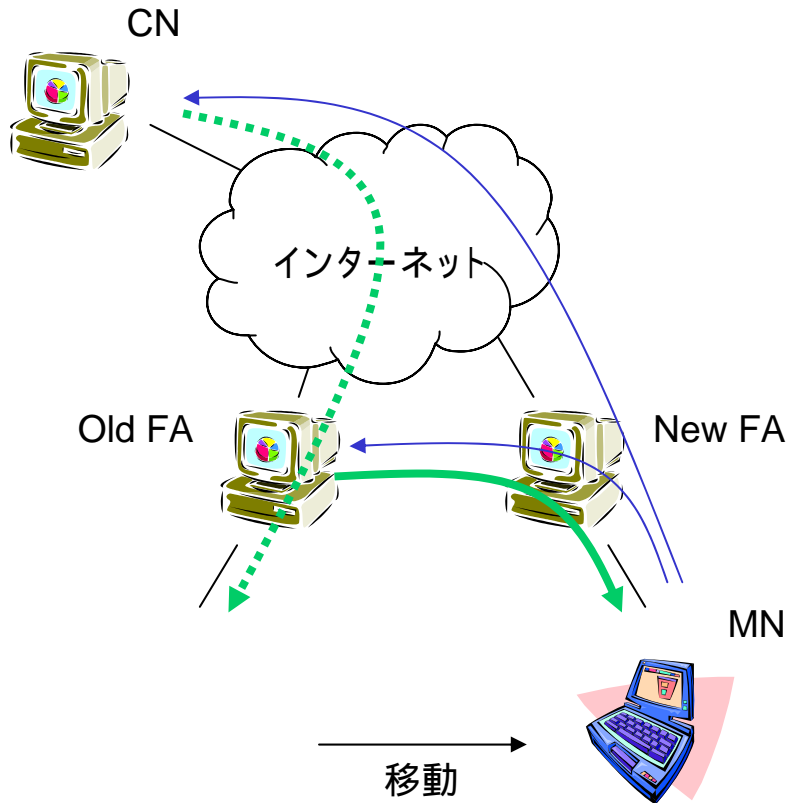
FA の階層化と、階層化ドメイン内のみ
のシグナリング
シグナリング遅延の削減

MAP: Mobility Anchor Point
AR: Access Router

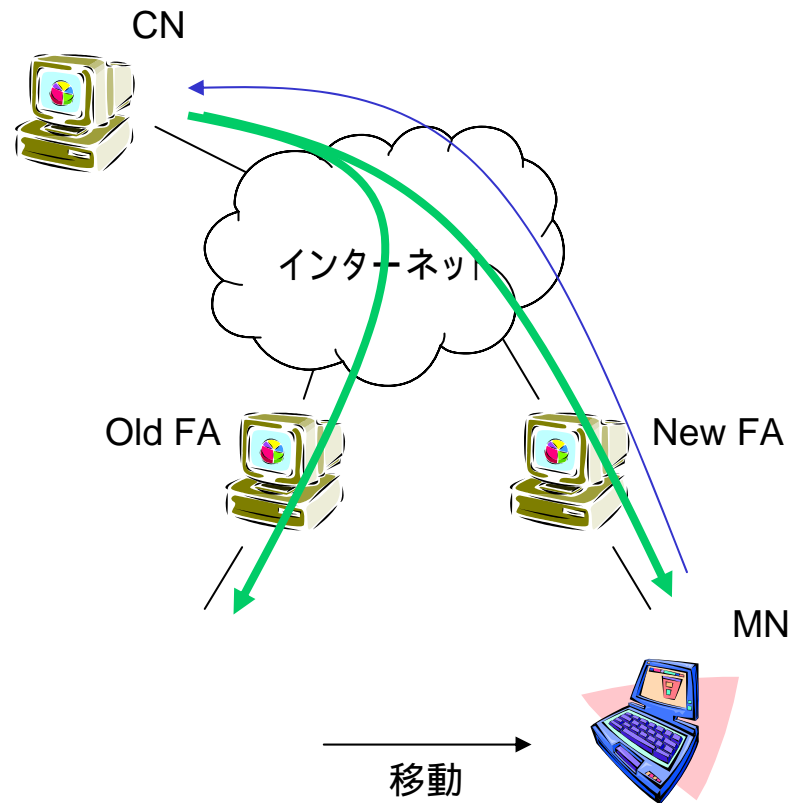
Mobile IP (7)

• Fast Handover (3) バッファリングとバイキャストイング

Buffering

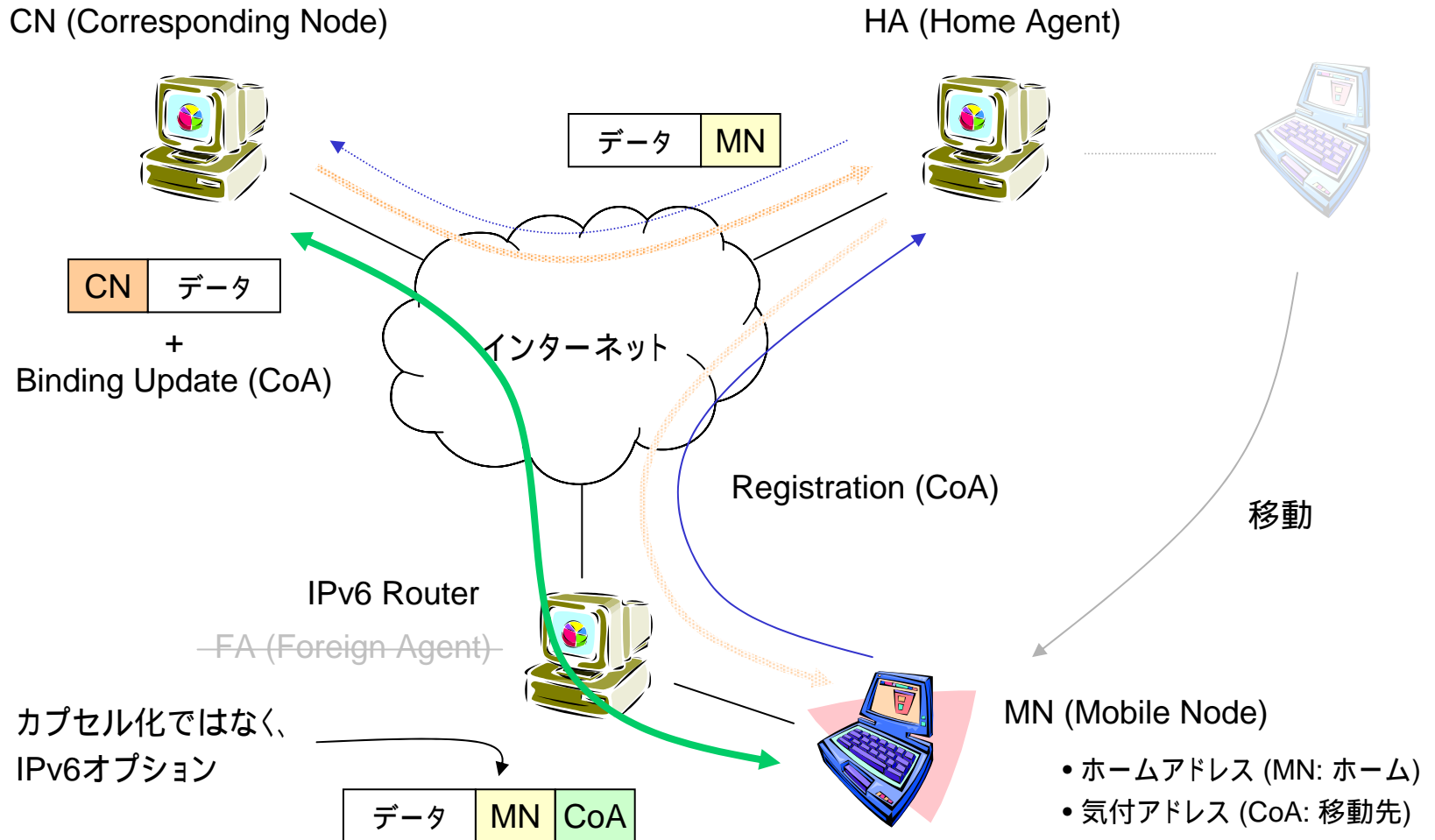


Bi-casting



Mobile IP (8)

• Mobile IPv6 (データ & 制御)



Mobile IP (9)

• Mobile IPv6 (定義と手順)

IPv4 との違い:

FA の廃止: IPv6 Stateless Address Autoconfiguration

Home Address Option: MN は発信元アドレスを CoA として送信

Destination Option: Binding Update をデータパケットに乗せられる

Route Optimization を IPv6 としてサポート

MIPv6 の手順:

MN が CoA を取得する (Stateless Address Autoconfiguration 等)。

MN が HA に CoA を登録する (MIP Registration)。

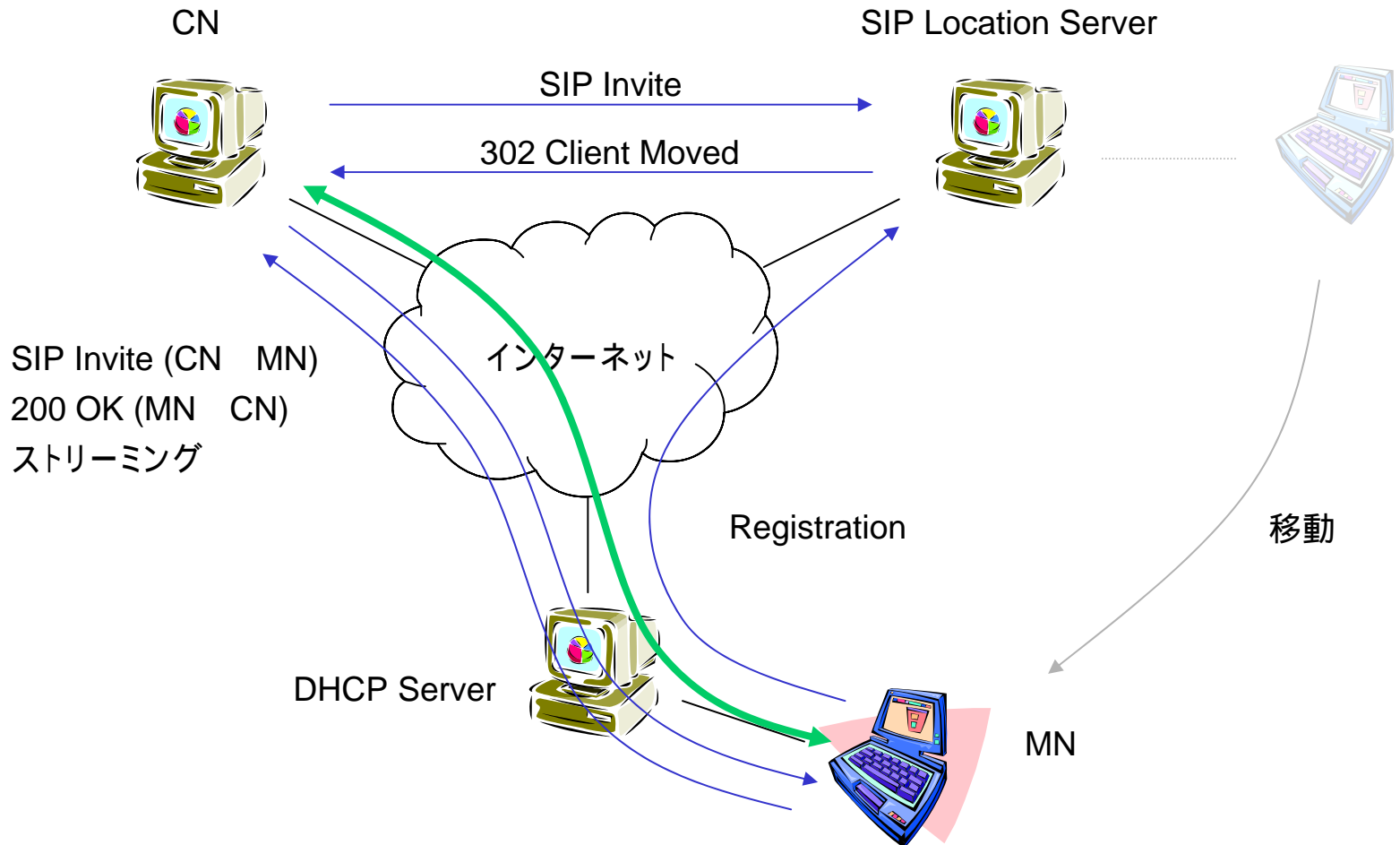
CN からのパケットが、HA からカプセル化されて MN に転送。

MN は、Binding Update を乗せて、CN にパケットを送信。

以降、MN と CN は、HA を介さずにパケットを送受信。

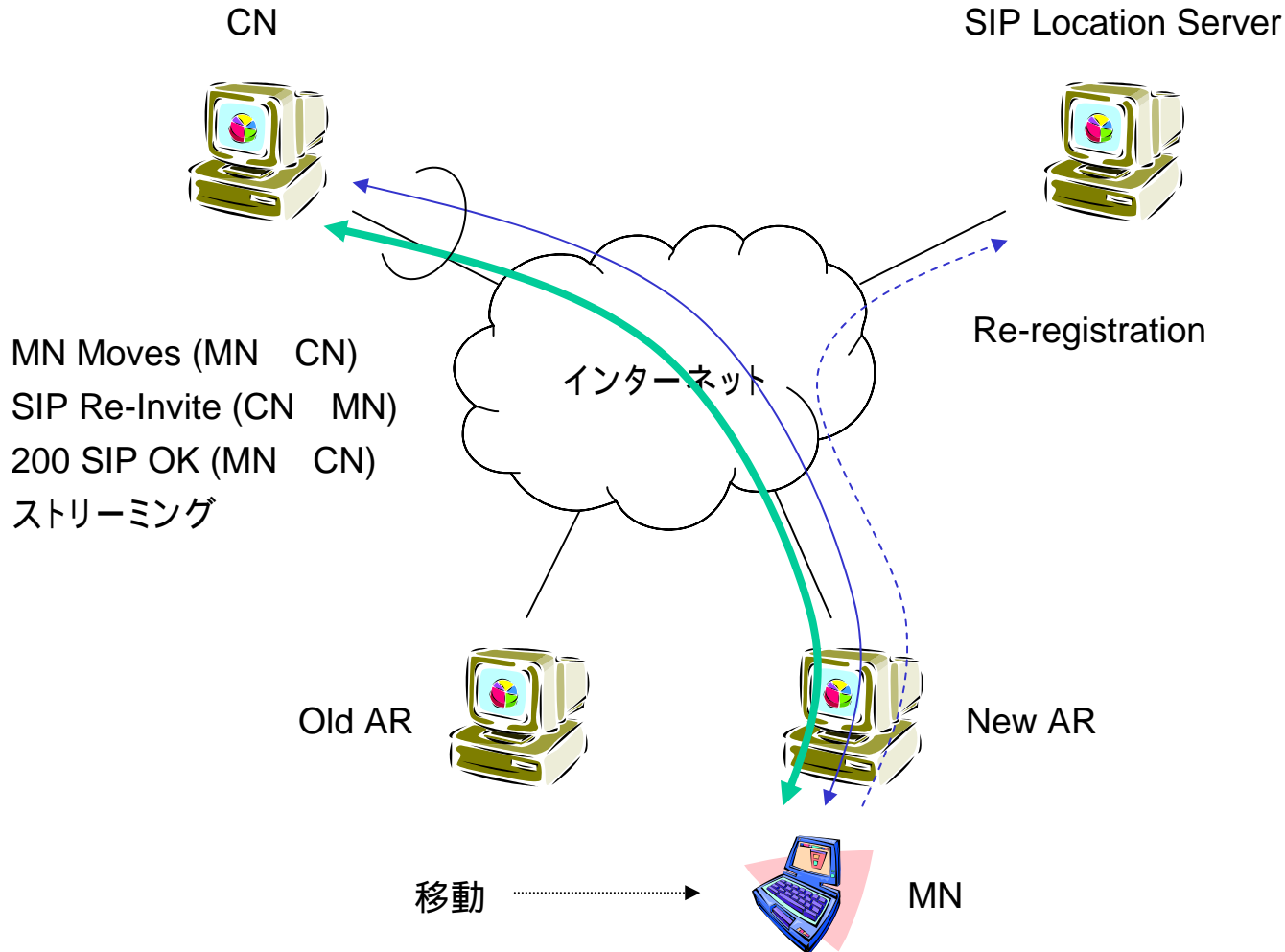
アプリケーション層モビリティ (1)

• SIP モビリティ (1) セッション前



アプリケーション層モビリティ (2)

• SIP モビリティ (2) セッション中



まとめ

まとめ

• 最近の傾向

L3/L4 技術 (ネットワーク・トランスポート層) と L7 技術 (アプリケーション層)

- マルチキャスト: IP マルチキャストと L7 マルチキャスト (P2P マルチキャスト)
- CDN: L3/L4 スイッチと L7 スイッチ、P2P CDN
- モビリティ: L3 モビリティ (Mobile IP) と L7 モビリティ (SIP Mobility)

L3/L4 技術

長所: 特性がよい、最適化が図れる

短所: 概して既存のシステムに変更が必要

L7 技術

長所: 既存のシステムに変更が不要、実装が容易

短所: L3/L4 技術に比べると特性が落ちる