

インターネットインフラ特論

10. ルーティング：マルチキャスト、経路表の縮約

太田昌孝

mohta@necom830.hpcl.titech.ac.jp

<ftp://chacha.hpcl.titech.ac.jp/infra10j.ppt>

マルチキャストとは？

- ネットワーク中でデータをコピーすることによる一対多、多対多の通信
 - ネットワークによる放送
- ネットワーク中に必用、エンドでは不可能
 - エンドでコピーするのはリフレクタ
- 資源予約と密接に関連
 - 混雑に応じた帯域の調整はできない
 - マルチキャストはルーティングテーブルを消費

ネットワーク

- 物流網
 - 郵便、宅配便、コンビニ
- 情報通信網
 - 出版網（書籍、新聞、レコード（CD）、映画）
 - 金融網
 - 電話網
 - 放送網
 - インターネット

出版網

- 同じ情報を大量に配布
- 情報流通は遅くていい
- 著作権法による保護
- いまのインターネットの好餌
 - 壊滅寸前

金融網

- お金のやりとりを管理
- 物流網でもあるが、今や、情報通信網としての面がはるかに大きい
- セキュリティー！！！！

電話網

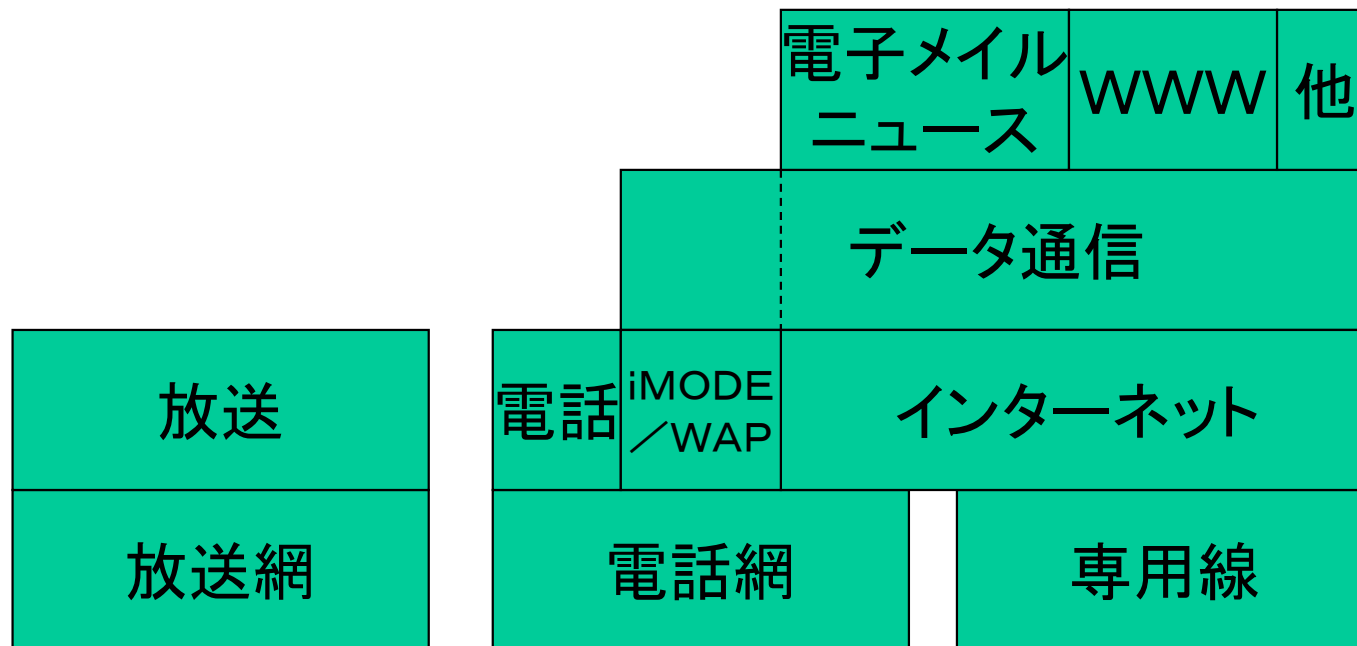
- 音声を実時間で伝送する網
 - 音声伝送の帯域を確保
 - 音声伝送の遅延を最小化(保証)
- 専用線事業も
 - あくまで音声伝送事業が主
- 遅くて高い
- 電電公社として保護、電気通信事業法で開放

放送網

- 音声、画像を実時間で多数に伝送する網
 - 伝送帯域を確保
 - 遅延を最小化
- 電波による広域一対多通信
 - ブロードキャスト／マルチキャスト
- 放送法による保護

放送	電話	データ通信
放送網	電話網	専用線

かつてのネットワーク



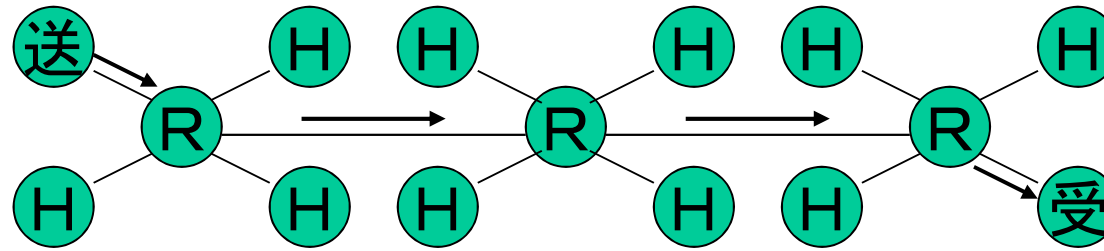
現在のネットワーク

放送	電話	電子メール ニュース	WWW	他
ストリーミング		データ通信(バッチ)		
インターネット				
専用線(含無線)				

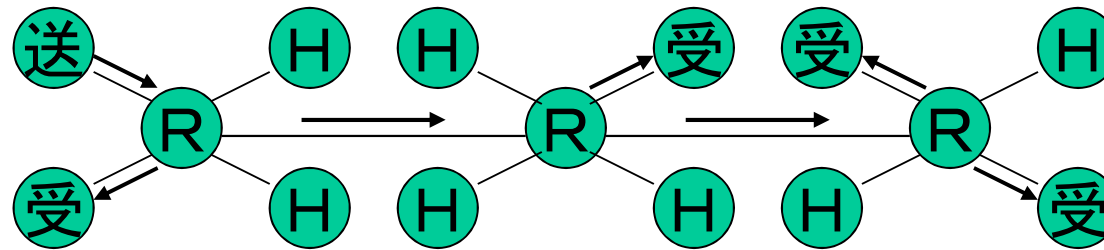
今後のネットワーク

マルチキャストと ブロードキャスト

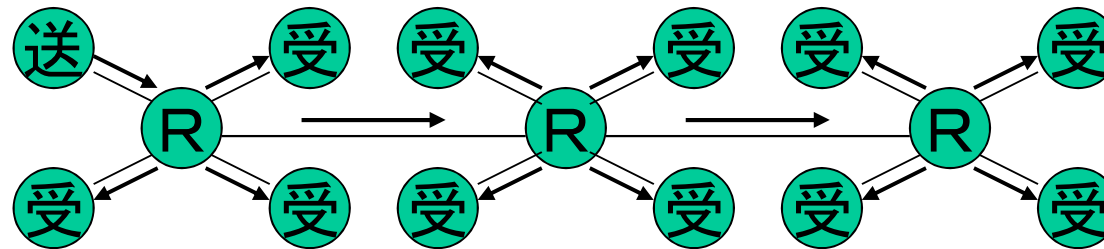
- ブロードキャスト
 - 全ホストに送信
 - インターネットでは非現実的
- マルチキャスト
 - 各グループのメンバーにだけ送信
 - メンバー数はいくら多くてもいい
 - メンバー管理は不可能、受信者の自己申告
 - グループはマルチキャストアドレスにより区別
 - 224. 0. 0. 0～239. 255. 255. 255



a) ユニキャスト

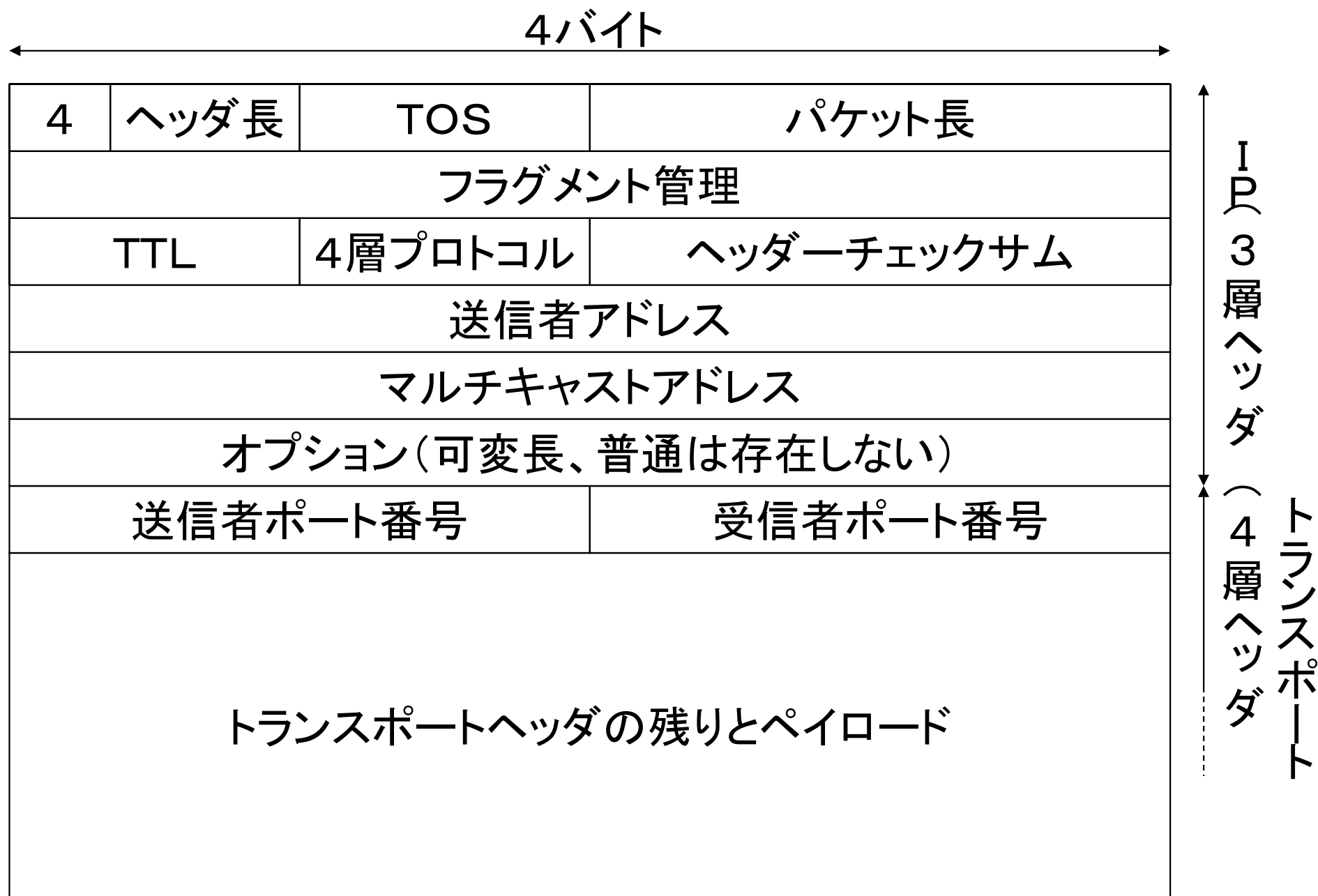


b) マルチキャスト



c) ブロードキャスト

●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者 →: データの流れ
 ユニキャスト、マルチキャスト、ブロードキャスト



IPv4マルチキャストパケットフォーマット

IGMPによるマルチキャスト(1)

- 受信側
 - ダイナミックに変動
 - IGMP(RFC988)により自己の存在を登録
 - IGMPはマルチキャストプロトコルと独立(?)
- 送信側
 - ダイナミックに変動
 - いきなりパケットを投げる
 - マルチキャストプロトコルと独立

IGMPによるマルチキャスト(2)

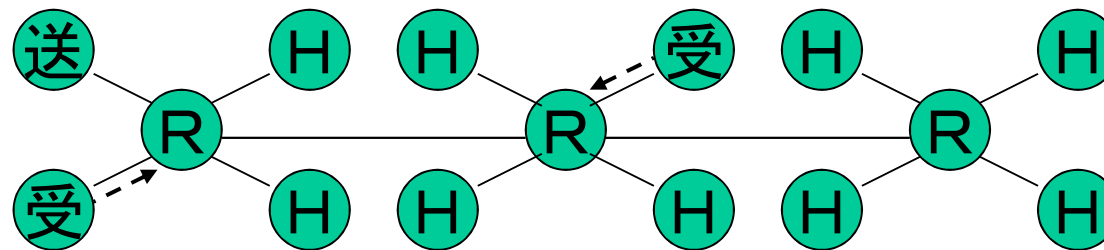
- ルータ
 - 特定のマルチキャストルーティングプロトコルを処理
 - それぞれのマルチキャストルーティングプロトコルに応じて
 - IGMPパケットを処理
 - 送信者のパケットを処理
 - エンドツーエンドモデル違反？

マルチキャストの関係者とエンド

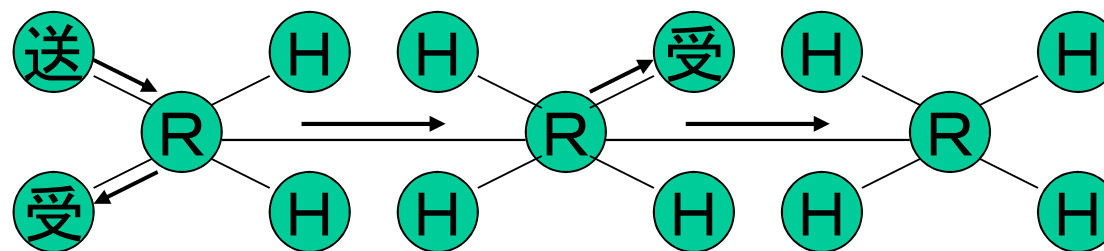
- 受信者 受信ホストが具現
- 送信者 送信ホストが具現
- グループ管理者を忘れてはいけない
 - 不在？
 - ISP？
 - 送信者が兼務(SSM)？
 - 管理ホストが具現！

マルチキャストルーティング プロトコル

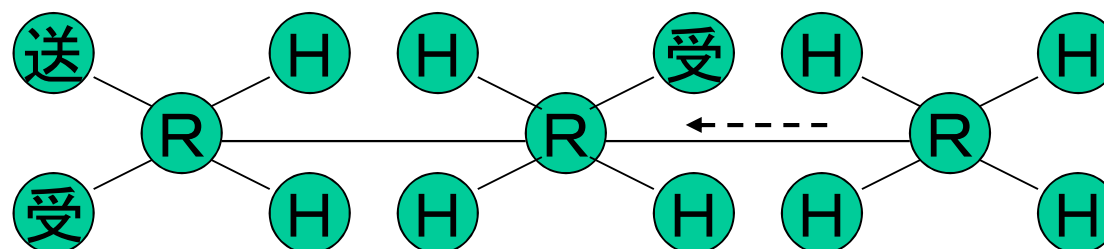
- DENSE
 - データをブロードキャストし不要部分を検出
 - DVMRP(RFC1075)、PIM-DM
- MOSPF(RFC1584)
 - 送信者、受信者の位置をブロードキャスト
- SPARSE
 - 中心をもち、データの流れを制御
 - CBT(RFC2189)、PIM-SM(RFC2362)



a) 登録



b) 初回のデータの流れ



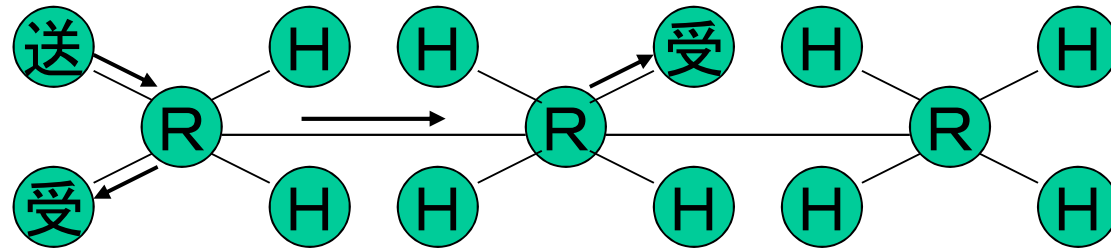
c) 末端の削除

●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者

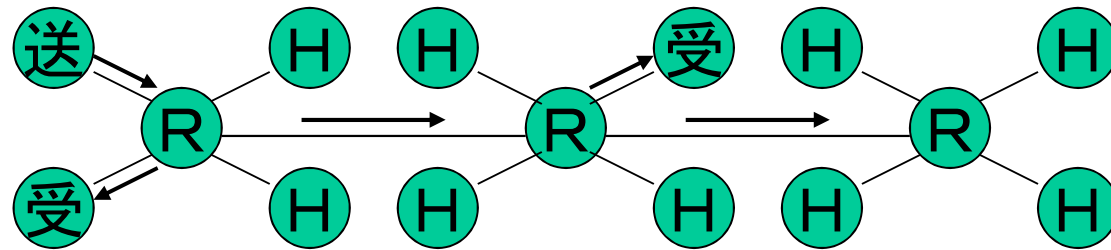
→: データの流れ

--->: 制御の流れ

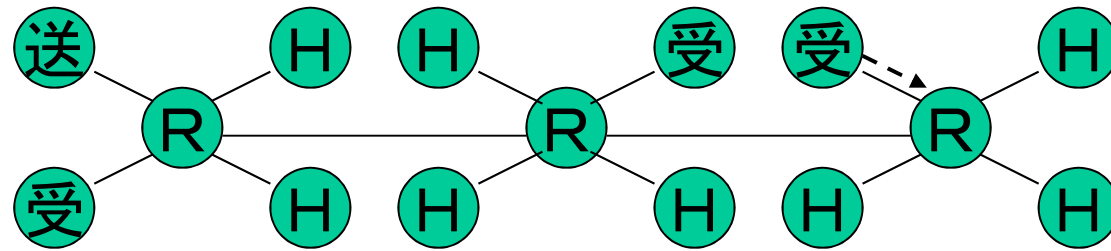
DVMRPの動作1



d) 末端削除後のデータの流れ



e) 一定時間経過後のデータの流れ



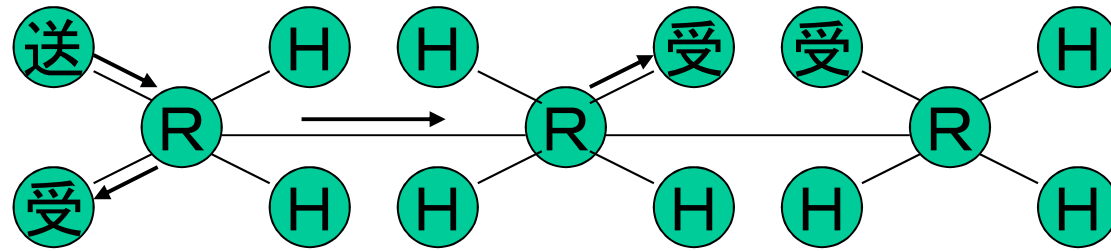
f) 新規受信者の出現

→: データの流れ

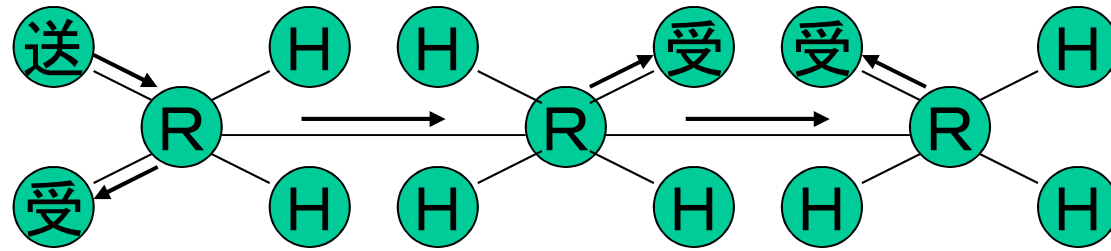
●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者

--->: 制御の流れ

DVMRPの動作2



g) 直後のデータの流れ

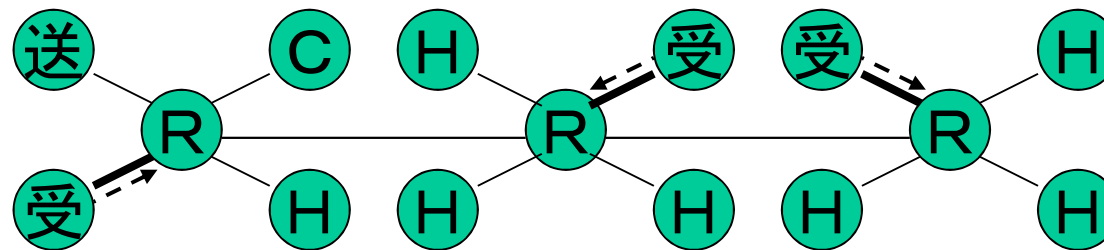


h) 一定時間経過後のデータの流れ

●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者 →: データの流れ
 --->: 制御の流れ
 DVMRPの動作3

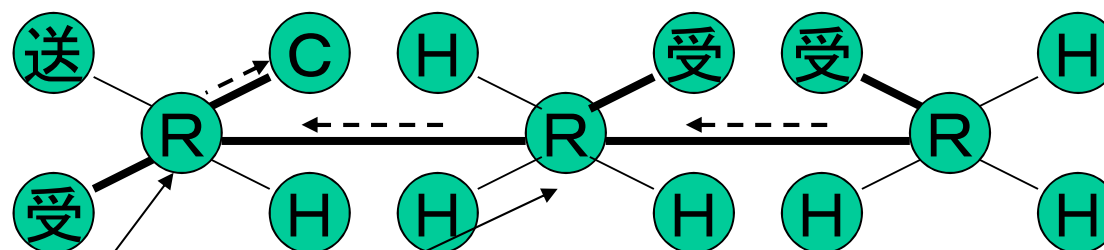
CBT (Core Based Tree)

- 受信者はCore(中心)に登録メッセージを送る
 - 複数の登録メッセージはマージ
 - Coreと受信者を含む双方向木を形成
- 送信者からのパケットはCoreに向け中継
 - 双方向木に到着したパケットは木に沿ってコピー



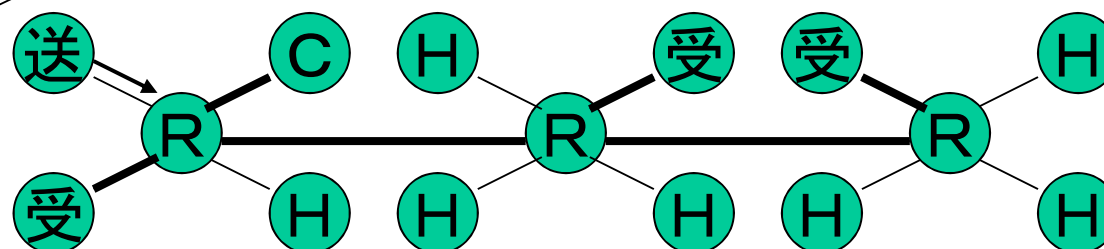
a) 受信者の登録

—: マルチキャスト木



b) 登録の伝播 (共有木の形成)

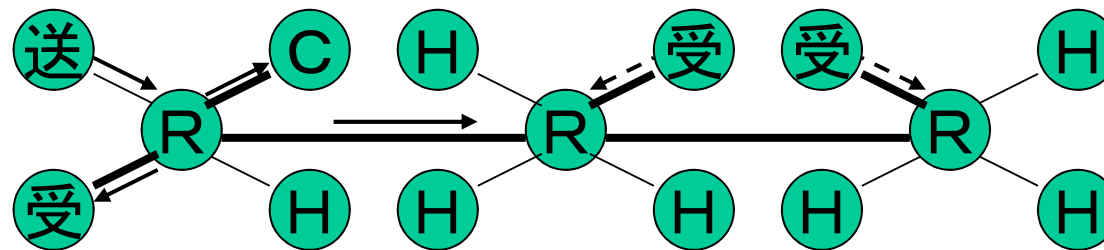
登録要求
のマーシ



c) Coreへの送信 (マルチキャスト) —: データの流れ

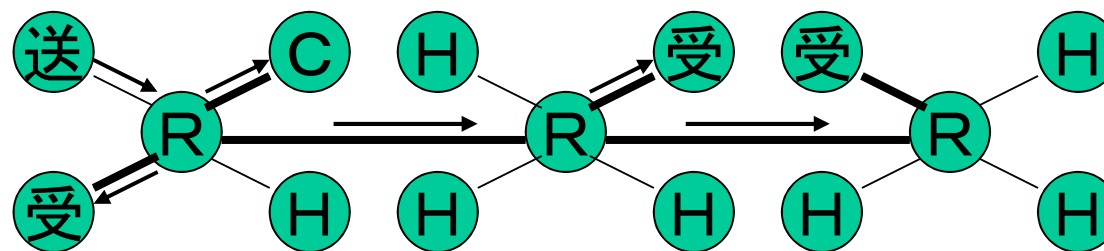
●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者 ---: 制御の流れ

CBTの動作1

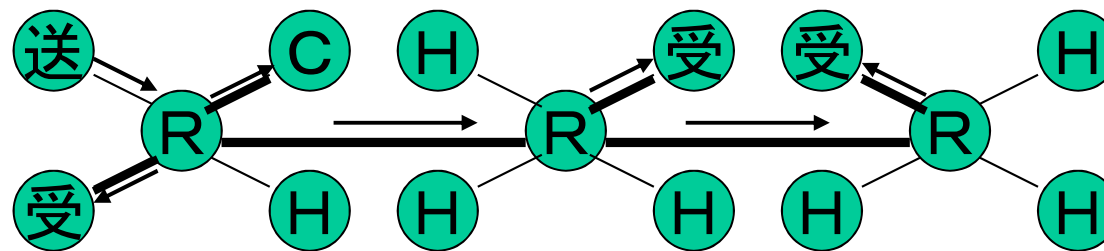


d) 共有木上の伝播1

—: マルチキャスト木



e) 共有木上の伝播2



f) 共有木上の伝播2

→: データの流れ

●H: ホスト

●R: ルータ

●送: 送信者

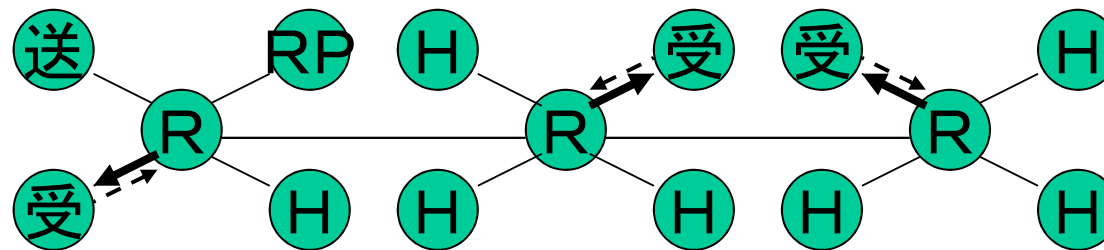
●受: 受信者

---: 制御の流れ

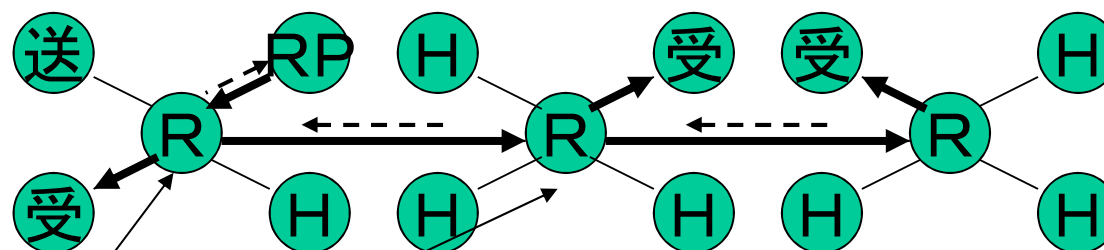
CBTの動作2

PIM (Protocol Independent Multicast) SM

- 受信者はRP (Rendez-vous Point) に登録メッセージを送る
 - 複数の登録メッセージはマージ
 - RPを根とする単方向木を形成
- 送信者(の隣のルータ)はRPにパケットをユニキャスト
 - RPに到着したパケットは木に沿ってコピー
- RPはグループ管理者を具現



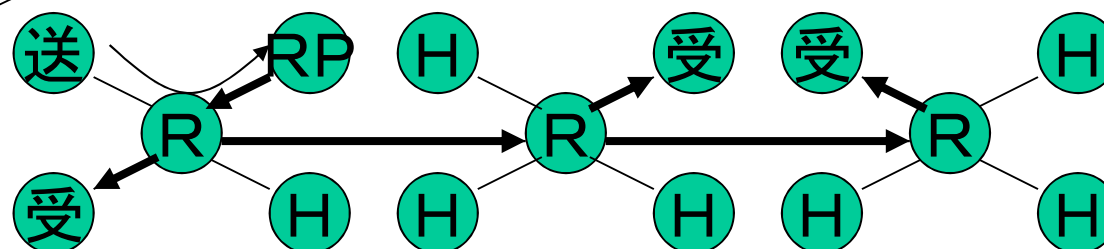
a) 受信者の登録



b) 登録の伝播

登録要求
のマーシ

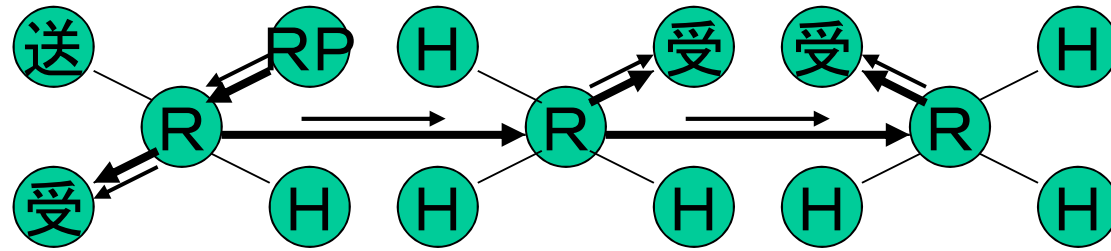
→: マルチキャスト木



c) RPへの送信 (ユニキャストネリング) →: データの流れ

●H: ホスト ●R: ルータ ●送: 送信者 ●受: 受信者 ---: 制御の流れ

PIMの動作1



d) RPによる中継(マルチキャスト)

H: ホスト R: ルータ 送: 送信者 受: 受信者
 →: マルチキャスト木 →: データの流れ
 --->: 制御の流れ
 PIMの動作2

COREとRP

- グループごとにどう決めるか？
 - ブロードキャスト？
 - マルチキャストアドレスとともにアプリケーション内の仕組みで配布？
 - ルータが不整合なメッセージをどうマージするか？
 - 多対多をあきらめて、送信者をRPに？
 - 静的マルチキャスト！
 - DNSの逆引きでCORE／RPを求める

インタードメインマルチキャスト

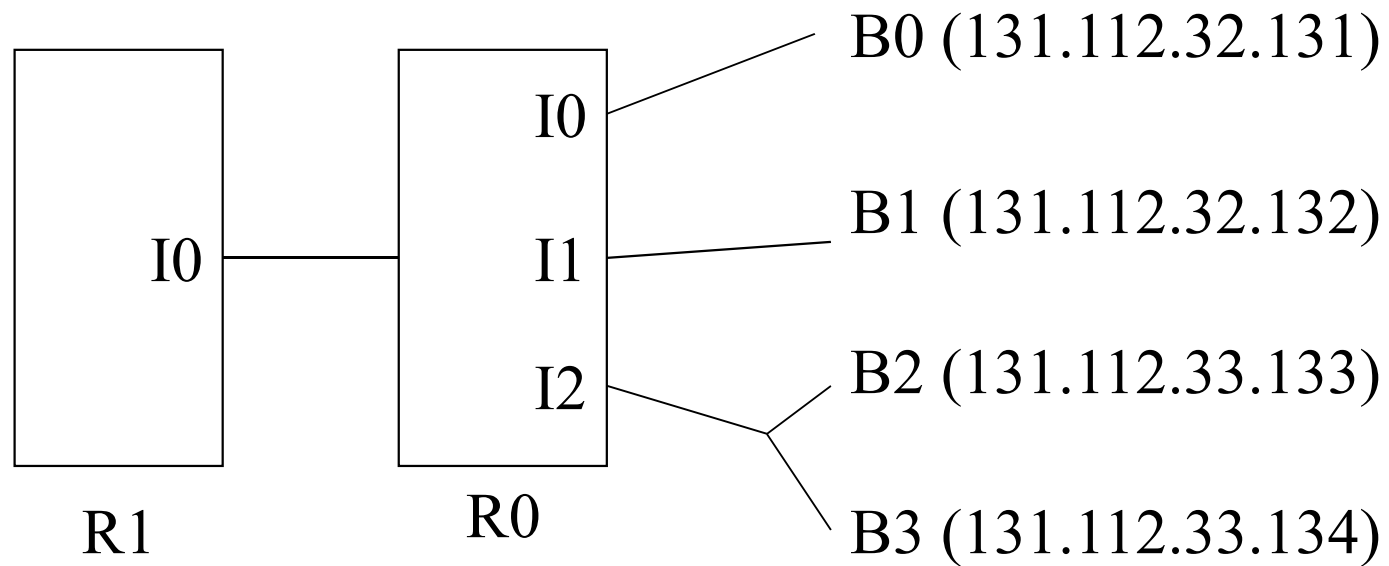
- 既存のマルチキャストプロトコルは、グループごとに別の経路表エントリが必用
 - 全インターネットではスケールしない？
- 既存のプロトコルは小さなドメインで利用
- ドメイン間では別のプロトコルを用意
 - BGMP (ドメイン間CBTのようなもの)
- ドメイン毎にマルチキャストアドレスをブロック化しドメイン間の経路表エントリを縮約

マルチキャスト経路表エントリの 縮約

- 不可能
- マルチキャストアドレスは位置をもたない
- グループが違えば(似たアドレスでも)受信者の分布は一般に似ても似つかない
- 中心をもつマルチキャスト(含BGMP)の
 - 送信者から中心までの(本質的にユニキャスト)経路は、もちろん縮約可能
 - 中心から受信者までの経路が縮約できない

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

経路の縮約が不可能な場合

- 経路が受信者アドレスだけで決まらない場合
 - QoSルーティング
- 受信者アドレスが場所と関係ない場合
 - マルチキャスト
- 地域のIPアドレスにまとまりがない場合
 - IPv4
 - ルーティングによるマルチホーミング

マルチキャスト経路表エントリが 縮約できないということは

- BGMPによるインタードメインマルチキャストは幻想
 - 静的CBTやPIMは、インタードメインでも動く
- インターネット幹線の経路表は大きい
- マルチキャストの各グループは、経路表エントリという有限の資源を占有する資源予約型通信
 - マルチキャストには従量制課金が必須

マルチキャストと帯域

- マルチキャストでは、受信者により混雑が
違う
 - 混雑制御をどこにあわせていいかわからない
- 送信側が速度を決める
 - 受信側はあわせるか、受信をあきらめる
 - 帯域(QoS)保証？

マルチキャストと課金

- 定額制課金のベストエフォートでは
 - ISPはマルチキャスト特別料金を取りたいが
 - 受信側にマルチキャストを使うメリットなし
 - ユニキャストのほうが楽
 - 送信側も、わざわざ使いたくない
- (QoS保証をとמונא)従量制課金だと
 - ISPはユニキャストで料金を取りたいが
 - 送受信側はマルチキャストで課金を減らしたい

マルチキャストのコスト例

8K県域放送

- 10Gbps県域網で考える(除アクセス)
 - ISP料金が一軒あたり月5000円、一軒2.5人とし、うち3割が県域網に使われると
 - 県域網コスト600円／(人・月)
 - 100Mbps占有マルチキャストは72円／(人・年)
 - 人口100万人の県で7200万円／年
- マルチキャストは
 - 電波放送網(2～4K)やCDNより安い
 - 100Gbps県域網ではさらに1／10

まとめ

- マルチキャストはネットワークの機能
 - エンドではできない
- IGMPはエンドツーエンド原理違反
- グループを管理するエンドシステムが重要
- マルチキャストでブロードキャストを利用しては話にならない
- マルチキャスト経路は縮約できない
 - マルチキャストは資源予約型通信