

IIJ

「ストリーミングシステム(II)」配信技術 IPマルチキャスト

アイアイジェイメディアコミュニケーションズ
藤井 直人
fujii@iij-mc.co.jp

Internet Week 2003
December 2, 2003



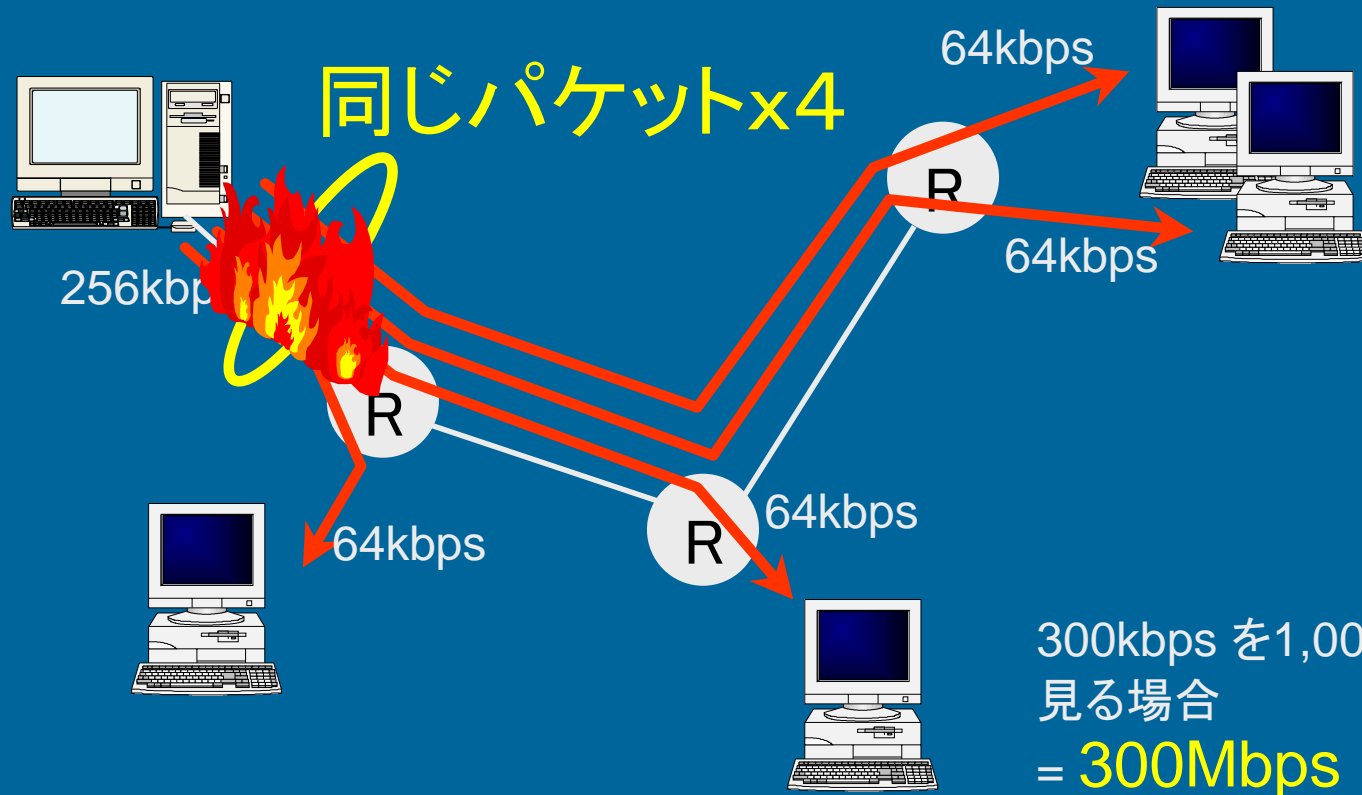
本日のアジェンダ

- IPマルチキャスト概略
- プロトコル詳細解説
- 現状の課題
- 最新トピック紹介
- 質問
- 休憩
- 後半へ続く



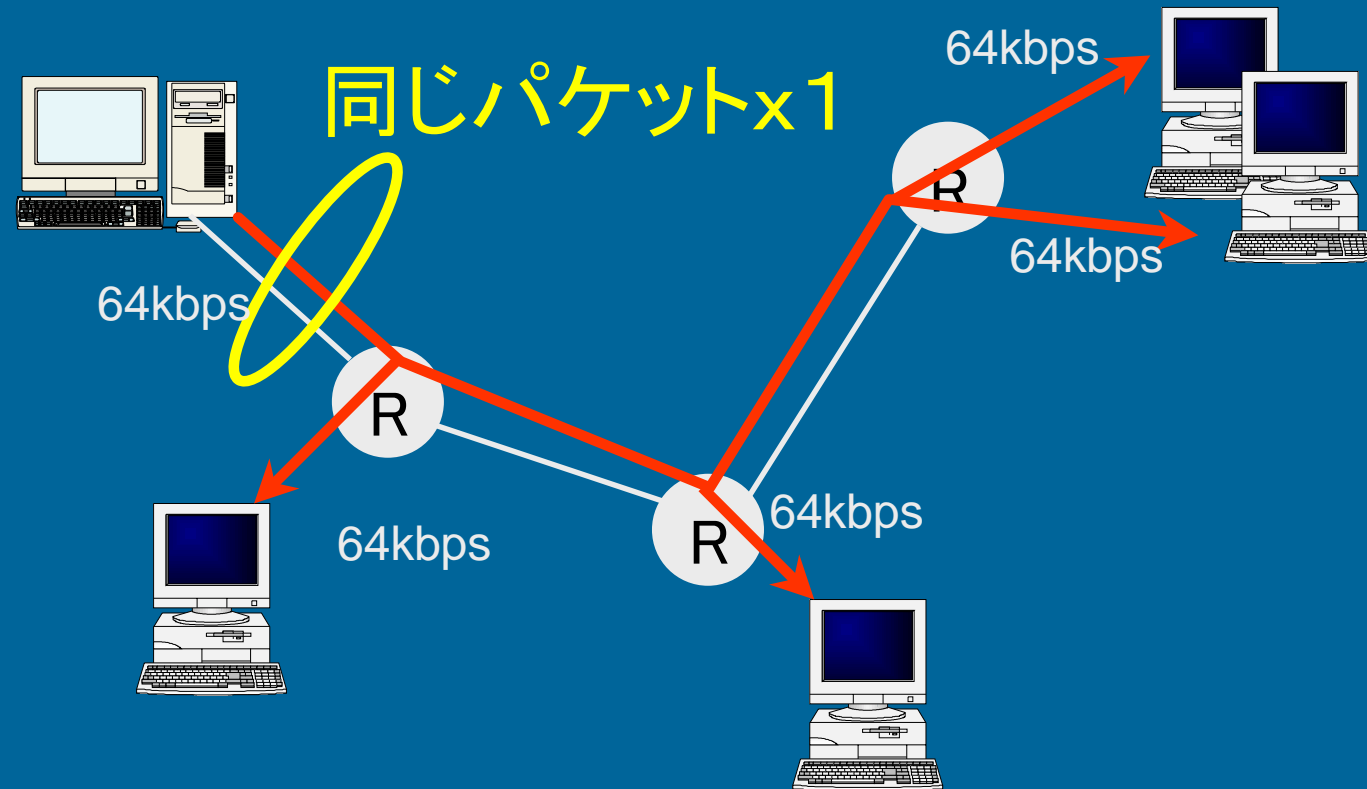
IPマルチキャスト概略

ユニキャストの場合



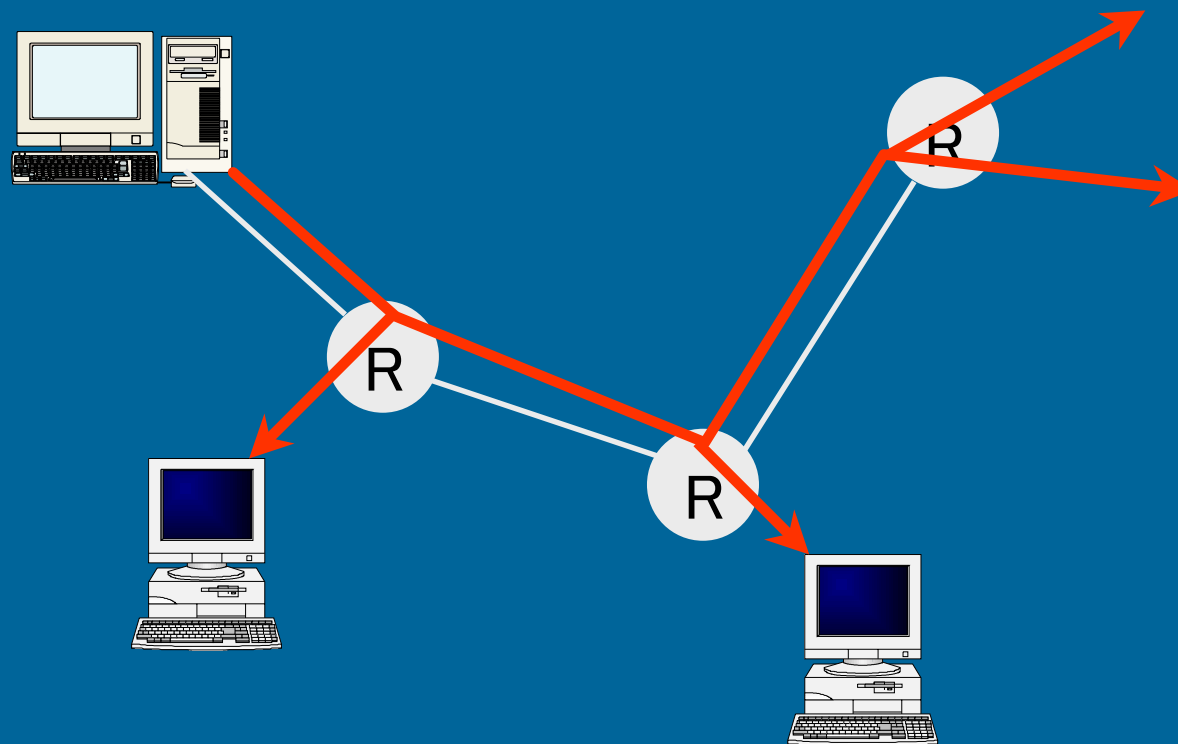
IPマルチキャスト概略(cont.)

マルチキャストの場合



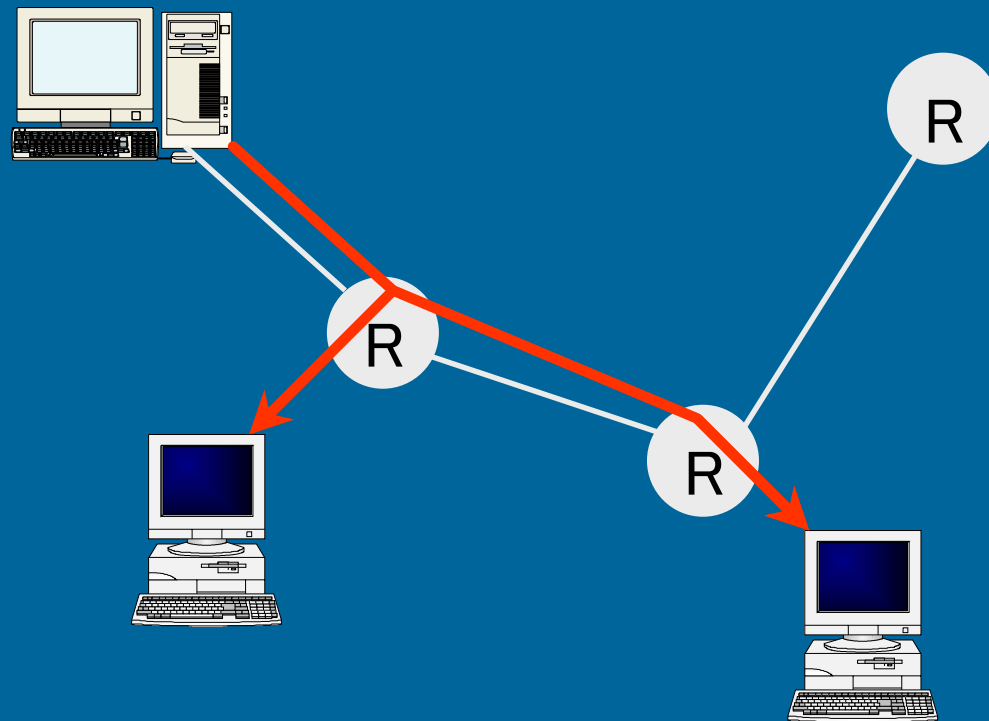
IPマルチキャスト概略(cont.)

マルチキャストの場合

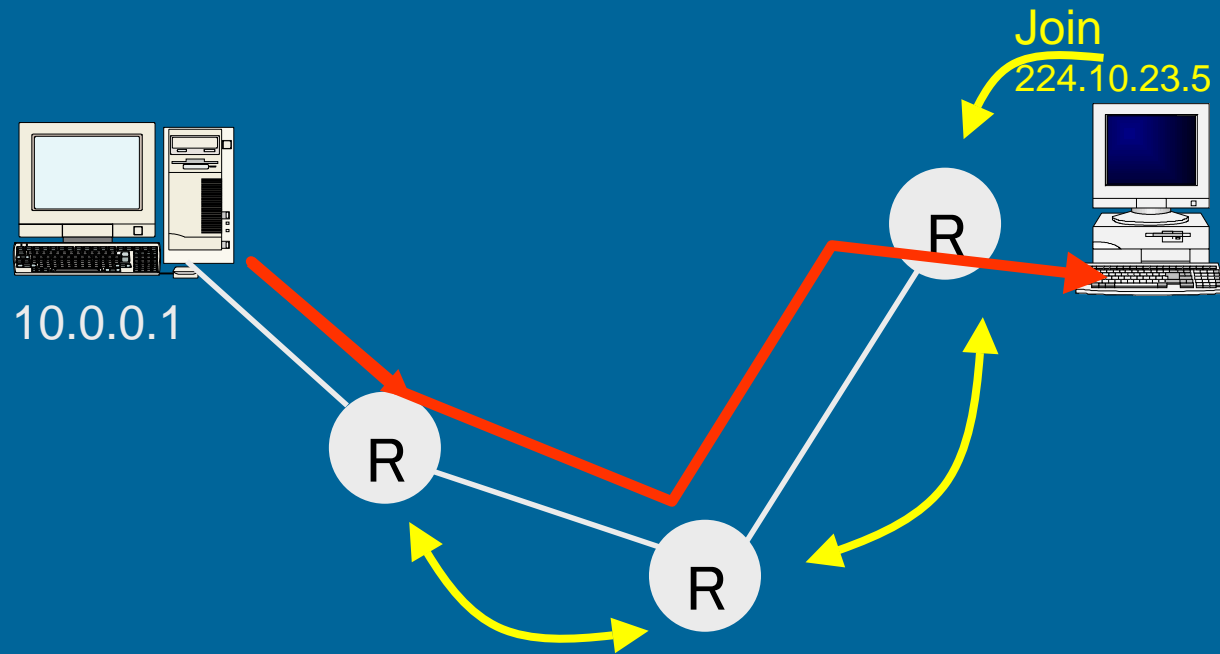


IPマルチキャスト概略(cont.)

マルチキャストの場合



マルチキャストの通信方法



| | src addr | dst addr |
|--|----------|-------------|
| | 10.0.0.1 | 224.10.23.5 |



マルチキャストアドレス1

- ClassD(224.0.0.0 - 239.255.255.255)を利用
- draft-ietf-mboned-rfc3171-update-01.txt
 - 224.0.0.0 - 224.0.0.255 (224.0.0/24) Local Network Control Block
 - 224.0.1.0 - 224.0.1.255 (224.0.1/24) Internetwork Control Block
 - 224.0.2.0 - 224.0.255.0 AD-HOC Block
 - 224.1.0.0 - 224.1.255.255 (224.1/16) ST Multicast Groups
 - 224.2.0.0 - 224.2.255.255 (224.2/16) SDP/SAP Block
 - 224.3.0.0 - 231.255.255.255 RESERVED
 - 232.0.0.0 - 232.255.255.255 (232/8) Source Specific Multicast Block
 - 233.0.0.0 - 233.255.255.255 (233/8) GLOP Block
 - 234.0.0.0 - 238.255.255.255 RESERVED
 - 239.0.0.0 - 239.255.255.255 (239/8) Administratively Scoped Block



マルチキャストアドレス2

- IANA
 - <http://www.iana.org/assignments/multicast-addresses>
- Local Network Control Block
 - 224.0.0.1 : ALL-SYSTEMS.MCAST.NET
 - 224.0.0.2 : ALL-ROUTERS.MCAST.NET
- Internetwork Control Block
 - 224.0.1.24 : microsoft-ds
 - 224.0.1.39 : CISCO-RP-ANNOUNCE
- AD-HOC Block
 - 歴史的経緯
 - 224.0.19.0 - 224.0.19.63 Walt Disney Company



アドレスアロケーション

- SAP(Session Announcement Protocol)
 - RFC2974
 - 予め決められたアドレスに対してアナウンスする
 - 224.2.127.254/9875 (global scope の場合)
 - 239.16.33.255 (239.16.32/24 の場合)
 - FF0X:0:0:0:0:0:2:7FFE (IPv6)
 - 新たにセッションを作る人は、しばらく受信した後に空いているアドレスを使用し、アナウンスする



セッション情報

- SDP(Session Description Protocol)
 - RFC2327
 - セッション情報を記述

| | |
|---|----------------------------|
| v=0 | プロトコル番号 |
| o=fujii 3142894548 3142894629 IN IP4 202.232.2.14 | |
| s=IIJ Test Channel | セッション名 |
| i=IIJ Test Channel from Osaka branch. | セッション情報 |
| u=http://help.iij4u.or.jp/Multicast/ | URL |
| e=<support@iij4u.or.jp> | メールアドレス |
| p=+81-3-5205-4444 | 電話番号 |
| t=3148678800 3151098000 | 有効期限 |
| m=audio 29748 RTP/AVP 0 | media name 0=PCM |
| c=IN IP4 239.253.128.81/31 | connection information |
| m=video 54210 RTP/AVP 31 | 31 = H.261 |
| c=IN IP4 239.253.128.44/31 | multicast address/init TTL |



GLOP addressing

- RFC3180
- 233/8 を各ASに static に割り当て
- 真中の16bit分にAS番号を割り当て、各ASは最後の8bit分を自由に使える
- AS2497 = 0x09c1 = 0x09 と 0xc1 = 9 と 193 = 233.9.193/24
- 計算CGI <http://gigapop.uoregon.edu/glop/>



マルチキャストアドレス3

- RFC2365 (Administratively Scoped IP Multicast)
- ユニキャストアドレスでいうところのプライベートアドレス
- 239.0.0.0 - 239.255.255.255 (239/8)
 - 239.255.0.0/16 Local scope (最小)
 - 239.192.0.0/14 organization-local scope



IPv6 マルチキャストアドレス

- RFC2373, RFC2375



Flags: 最下位bitが0 = 固定的に割り当て

最下位bitが1 = 動的に割り当て

- Scope:
- 1 node-local scope
 - 2 link-local scope
 - 5 site-local scope
 - 8 organization-local scope
 - E global scope

FF02:0:0:0:0:0:0:1 All Nodes Address (= IPv4のブロードキャストアドレス)

FF02:0:0:0:0:0:0:D All PIM Routers

FF0X:0:0:0:0:0:0:101 Network Time Protocol (NTP)

FF0X:0:0:0:0:0:0:108 SUN NIS+ Information Service



ユニキャストプレフィックスベースの IPv6 マルチキャストアドレス

- RFC3306



Flags: P=0 これまでの使い方

T=0 固定割り当てアドレス

T=1 動的割り当てアドレス

P=1 Network Prefix を使う

T=1 必ず1にすること

Plen: Prefix length

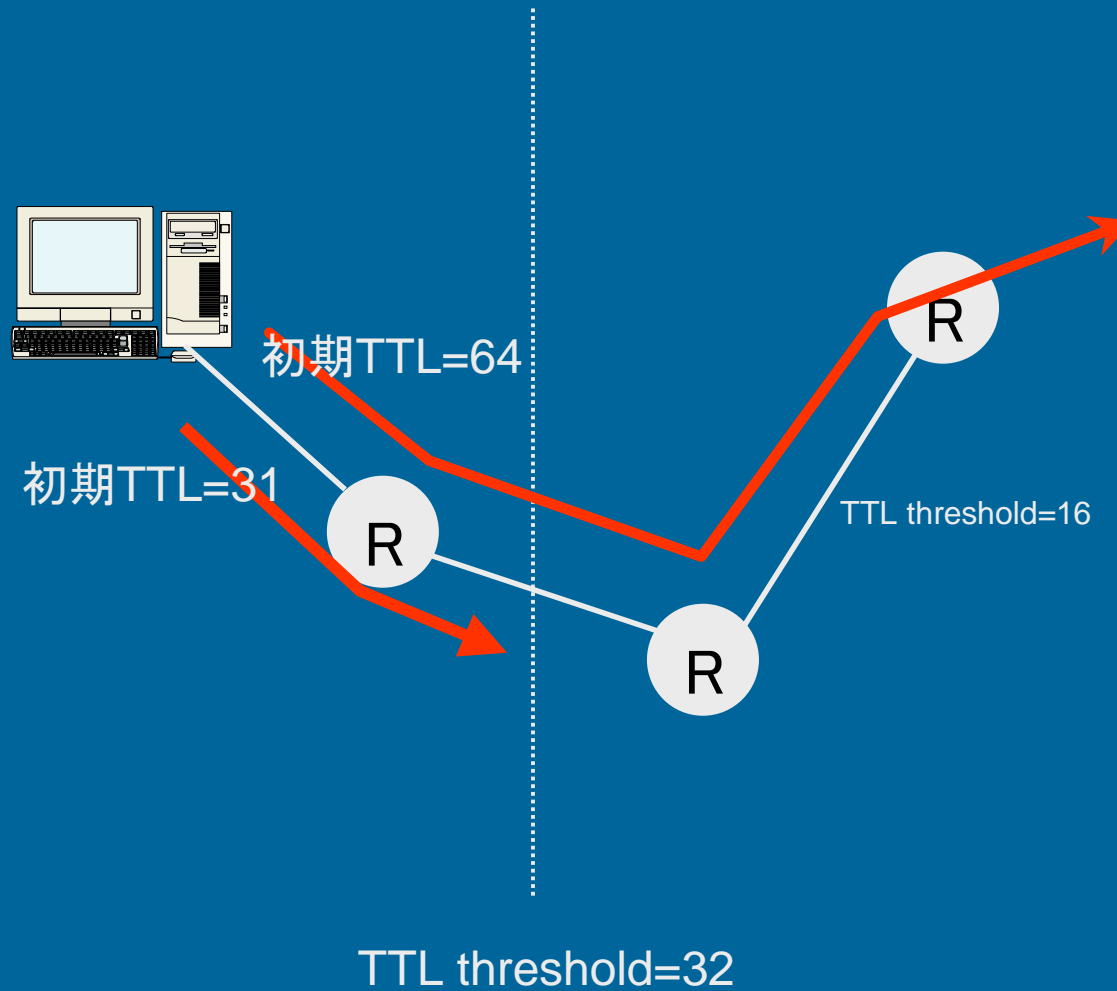
P=1 の場合のプレフィックスのビット数

Network Prefix:

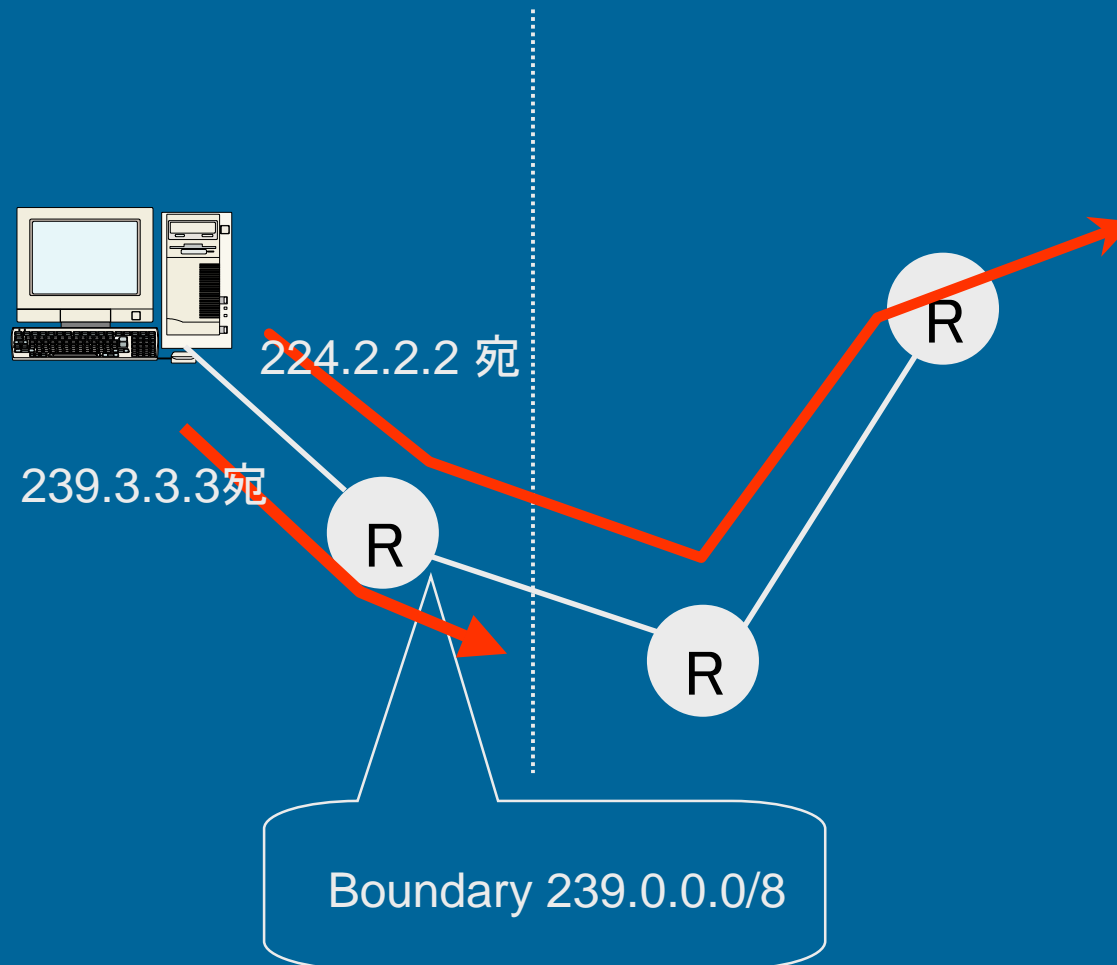
マルチキャストアドレスを所有しているユニキャストサブネットのネットワークプレフィックス



Scope と TTL



Scope と boundary

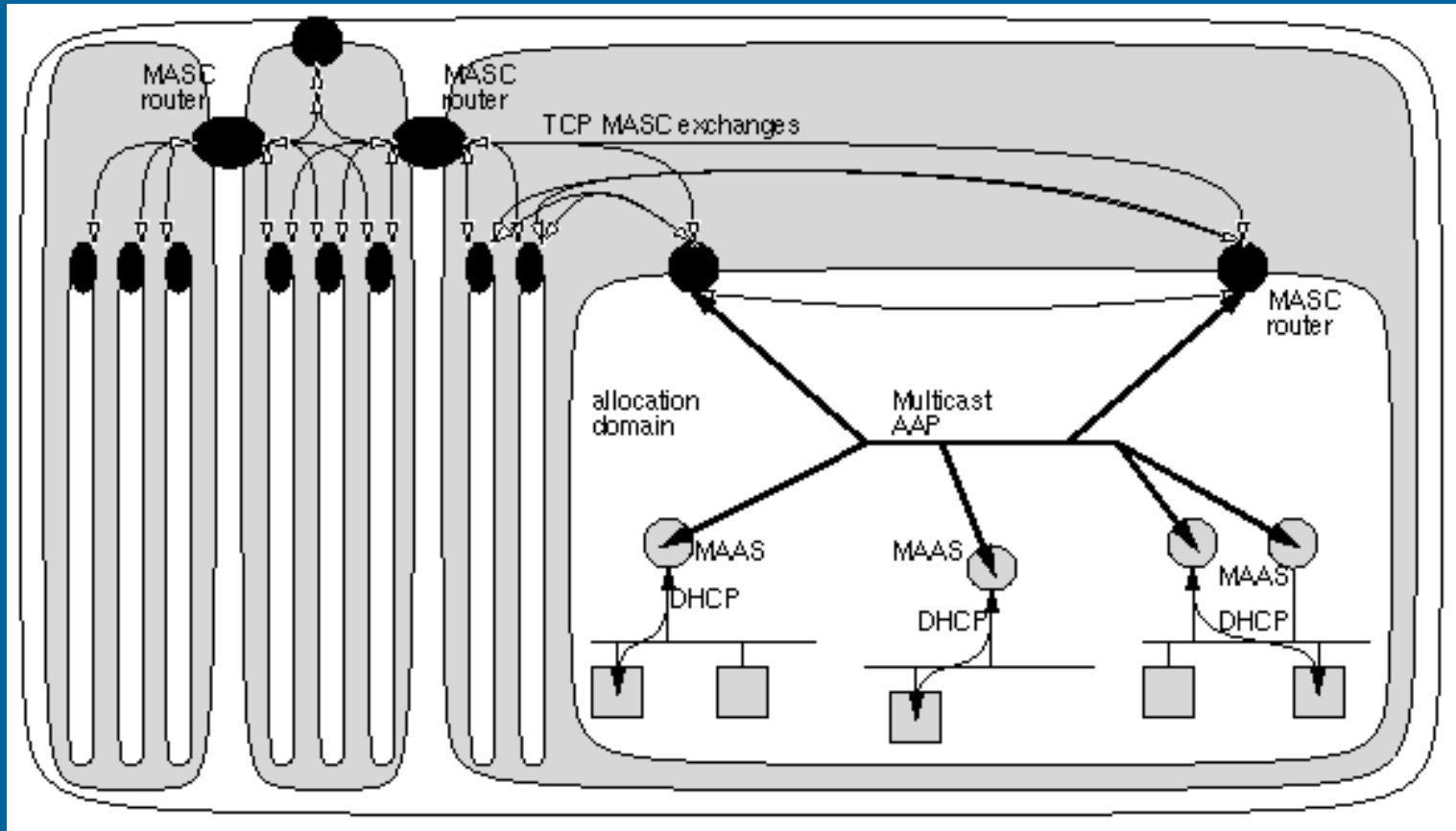


動的マルチキャストアドレス割り当て

- IETF MALLOC WG
 - RFC2771, RFC2908
 - マルチキャストアドレスの動的割り当て
 - 3階層からなる
 - ドメイン(AS)間: MASC
 - RFC2909
 - <http://netweb.usc.edu/masc/mascd/>
 - ドメイン内: AAP
 - draft-ietf-malloc-aap-04.txt
 - クライアントへ: MADCAP
 - RFC2730

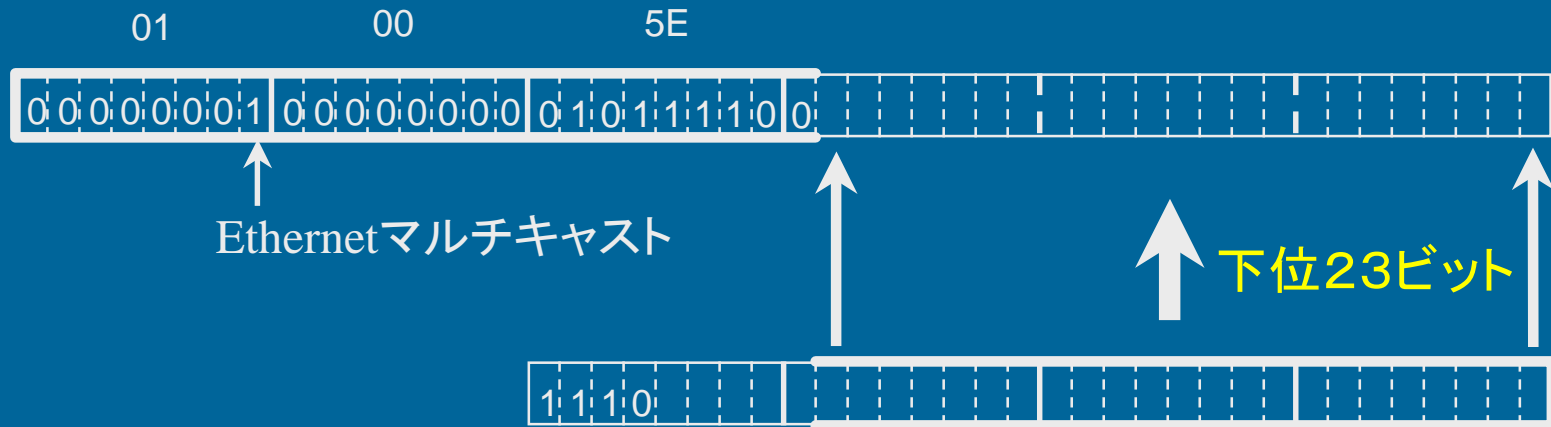


動的マルチキャストアドレス割り当て



イーサネットマルチキャスト

Ethernetアドレス(6オクテット)



IPマルチキャストアドレス(4オクテット)

$2^5=32$ 個分の IPマルチキャストアドレスが
一つの Ether マルチキャストアドレスにマップされる



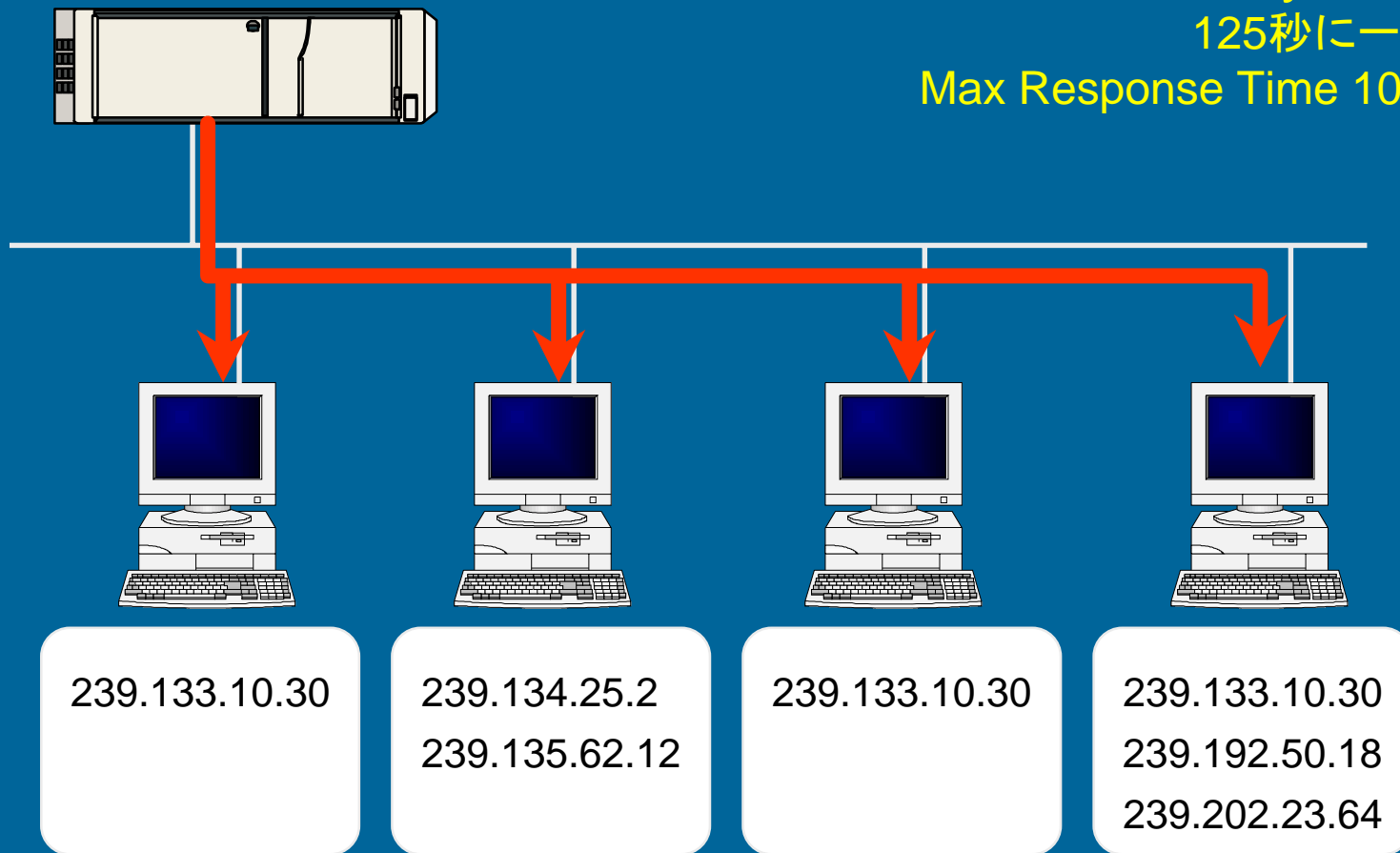
グループ管理プロトコル

- IGMP(Internet Group Membership Protocol)
- ローカルなサブネット上のホストがどんなグループに参加しているかを知るためのプロトコル
- IGMPv1 (RFC1112)
- IGMPv2 (RFC2236)
 - Leave Group
- IGMPv3 (RFC3376)
 - フィルターモード
- MLD(Multicast Listener Discovery)
 - RFC2710
 - ICMPv6 のサブセット



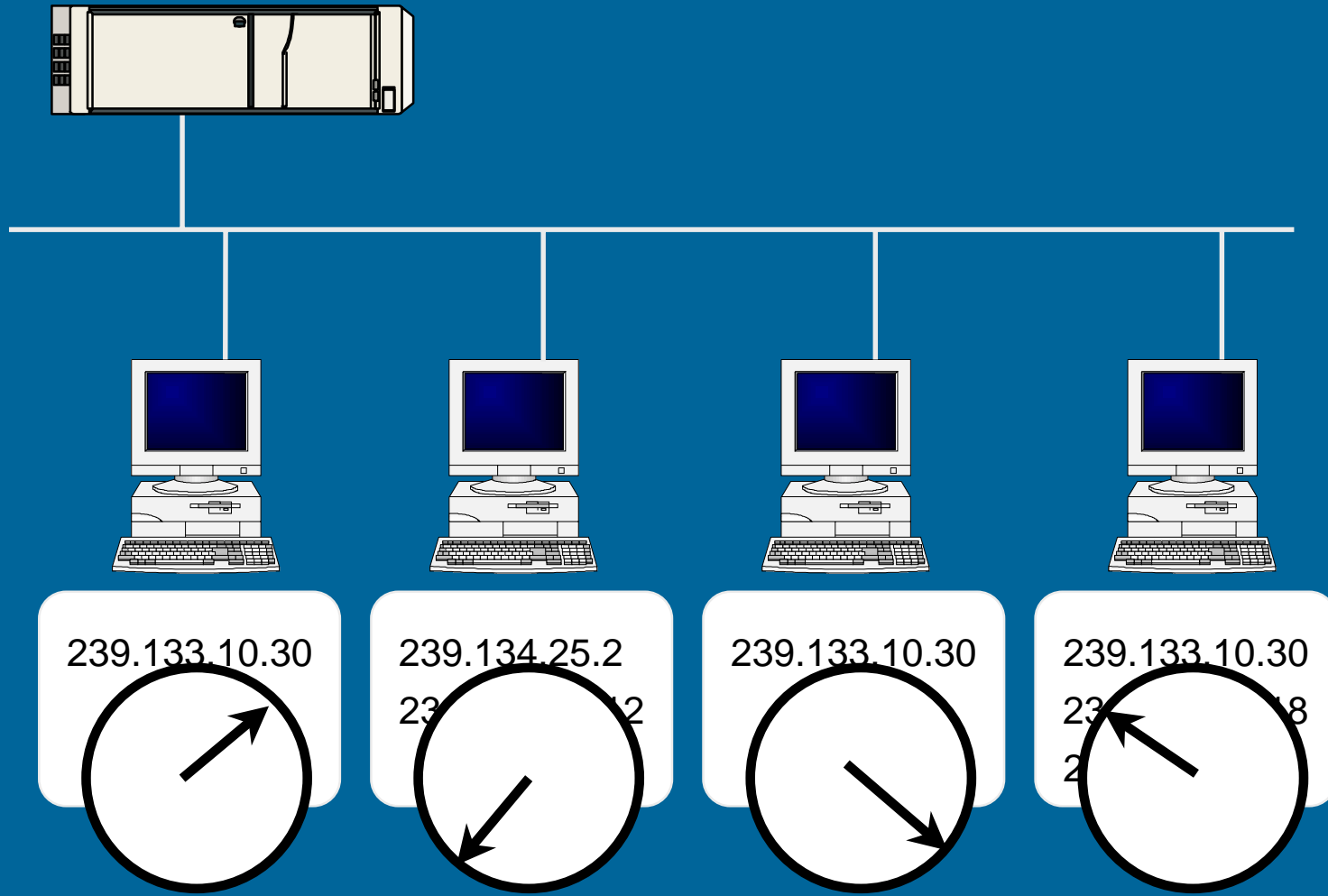
IGMPの仕組み

224.0.0.1 に対して IGMP Query 送信
 125秒に一度
 Max Response Time 10秒



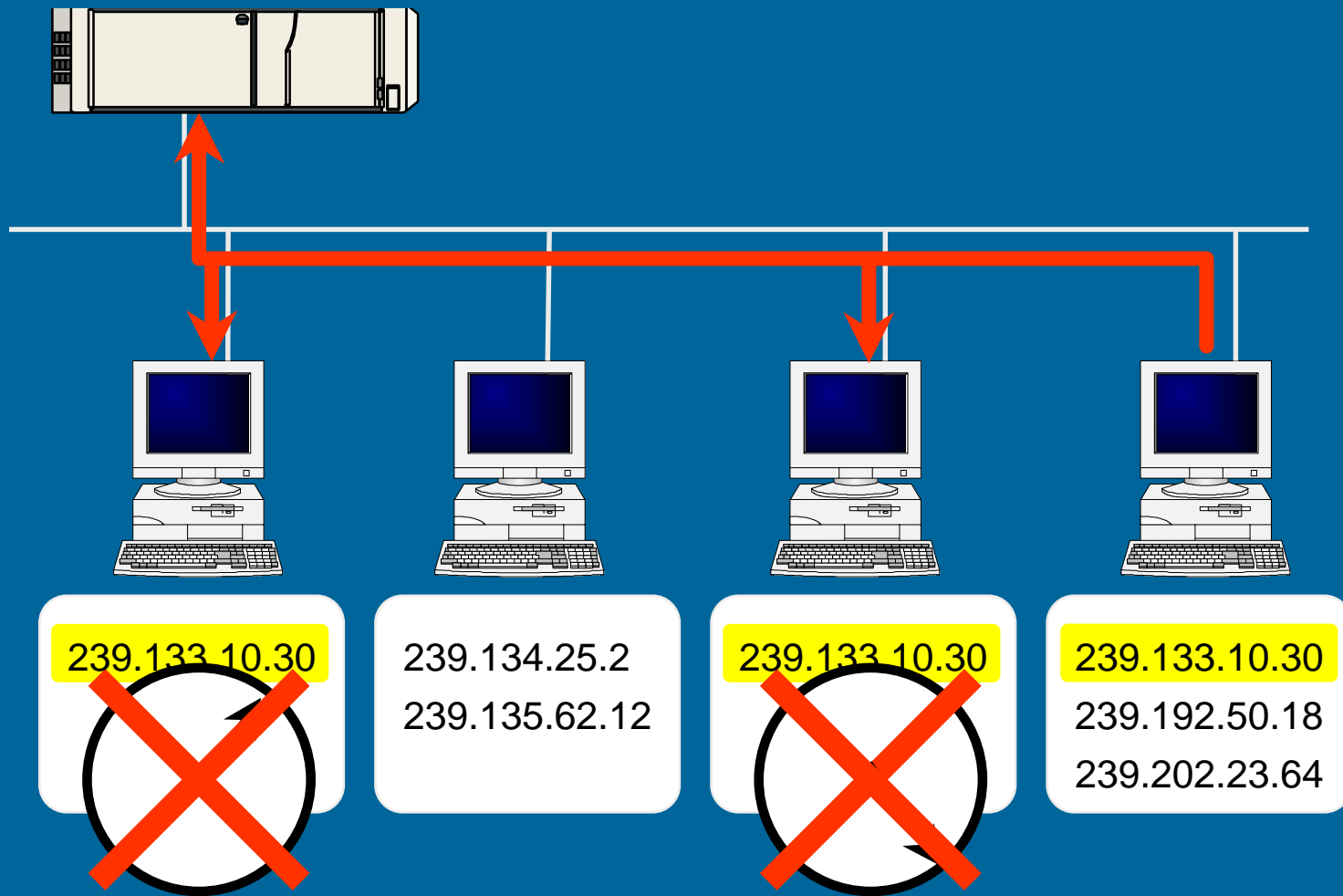
IGMPの仕組み(cont.)

返答タイマースタート



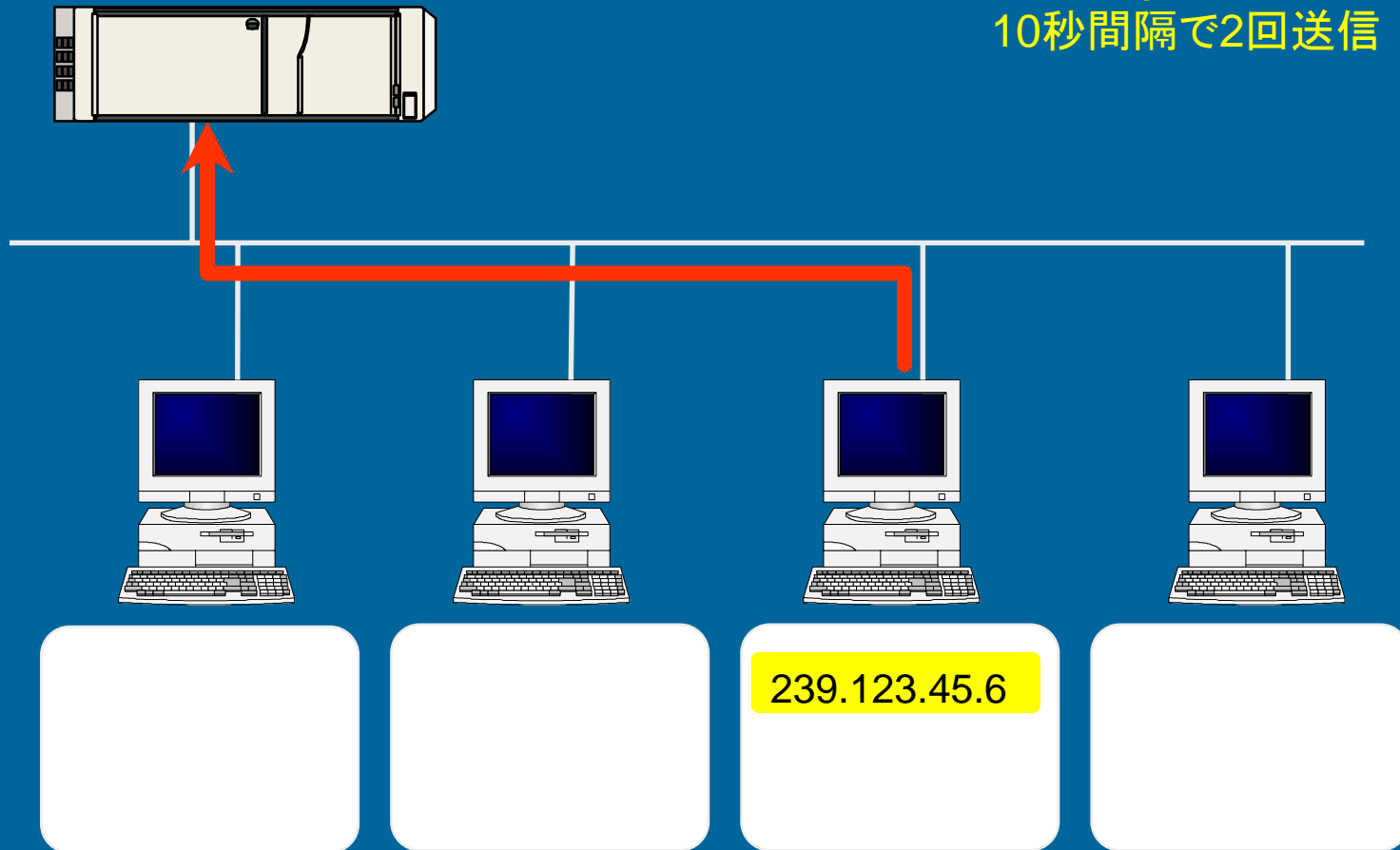
IGMPの仕組み(cont.)

239.133.10.30 に対する IGMP Report の送信を中止



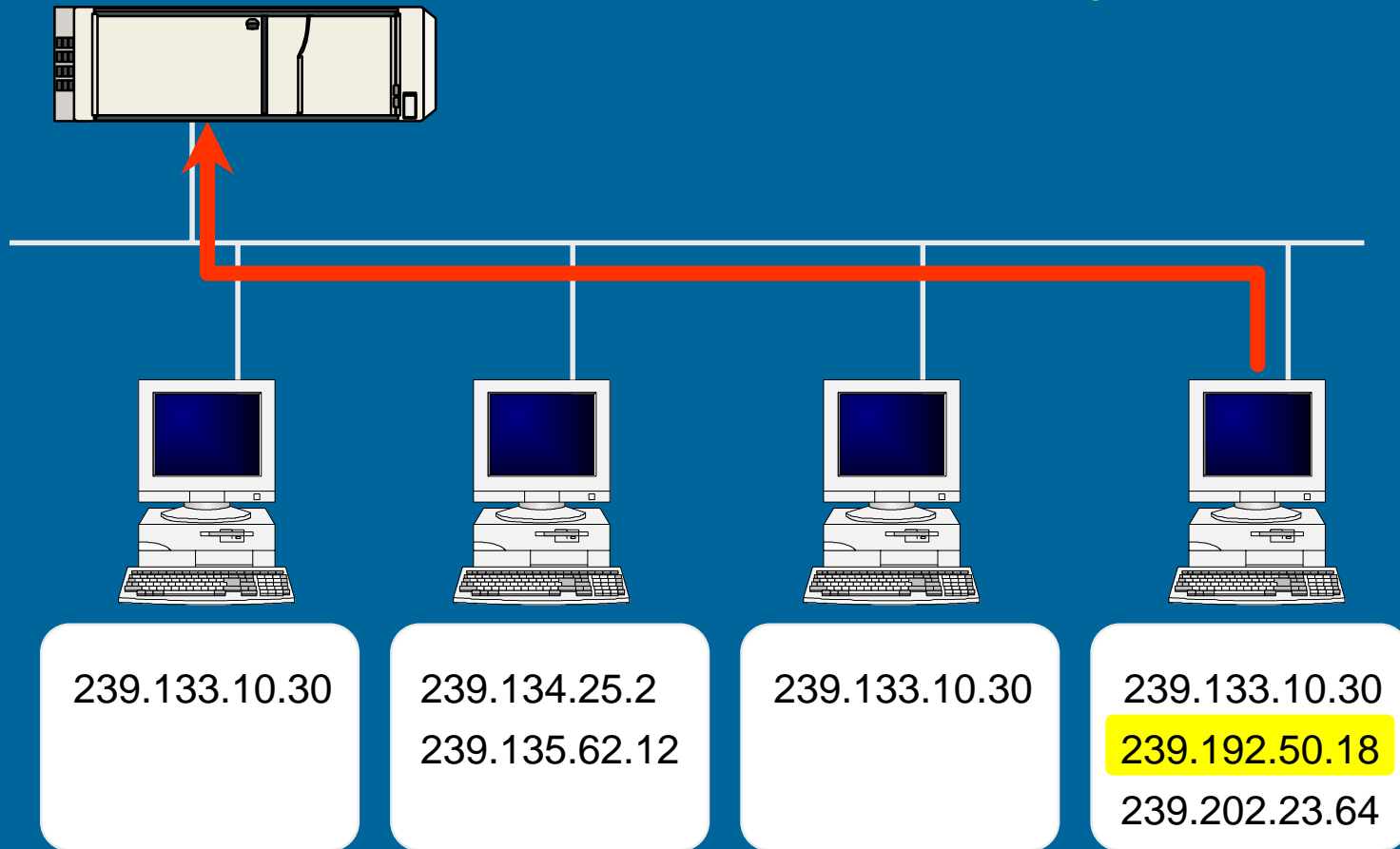
IGMPの仕組み (Membership Report)

239.123.45.6 に対して IGMP Report 送信
10秒間隔で2回送信



IGMPの仕組み (LeaveGroup)

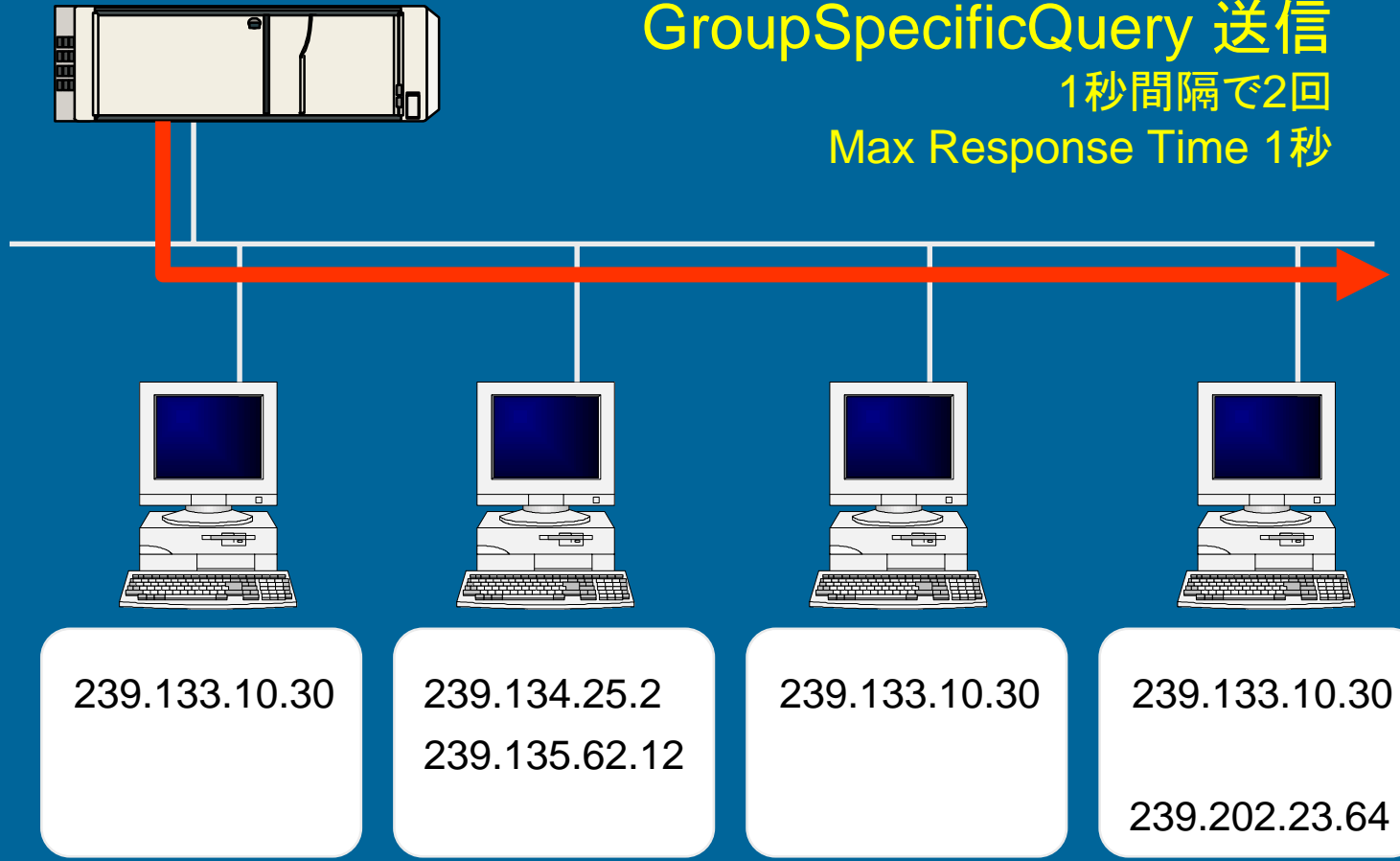
224.0.0.2 に対して LeaveGroup 送信



IGMPの仕組み

(LeaveGroup)

239.192.50.18 に対して
 GroupSpecificQuery 送信
 1秒間隔で2回
 Max Response Time 1秒



IGMPv3

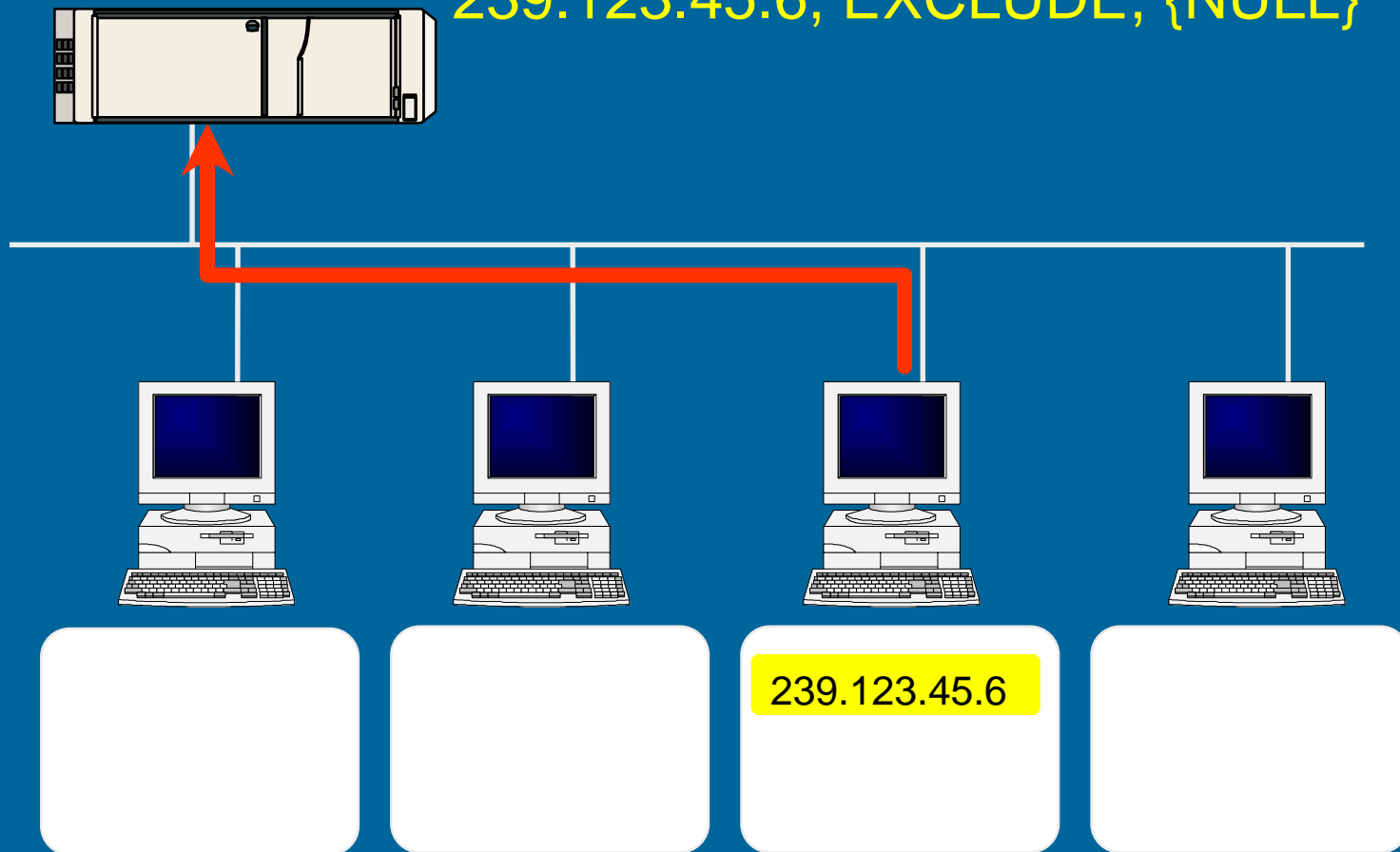
- フィルターモードという概念を追加
- IPMulticastListen (socket, interface, multicast-address, filter-mode, source-list)
- filter-mode:
 - INCLUDE: 以下のソースアドレスのパケットを欲しい
 - EXCLUDE: 以下のソースアドレスのパケットはいらない
- ソースアドレスは複数記述できる
- TTL値は1で
- TOS は 0xc0 (Precedence(優位)= 110
 - Internetwork Control ネットワーク間制御)



IGMPの仕組み

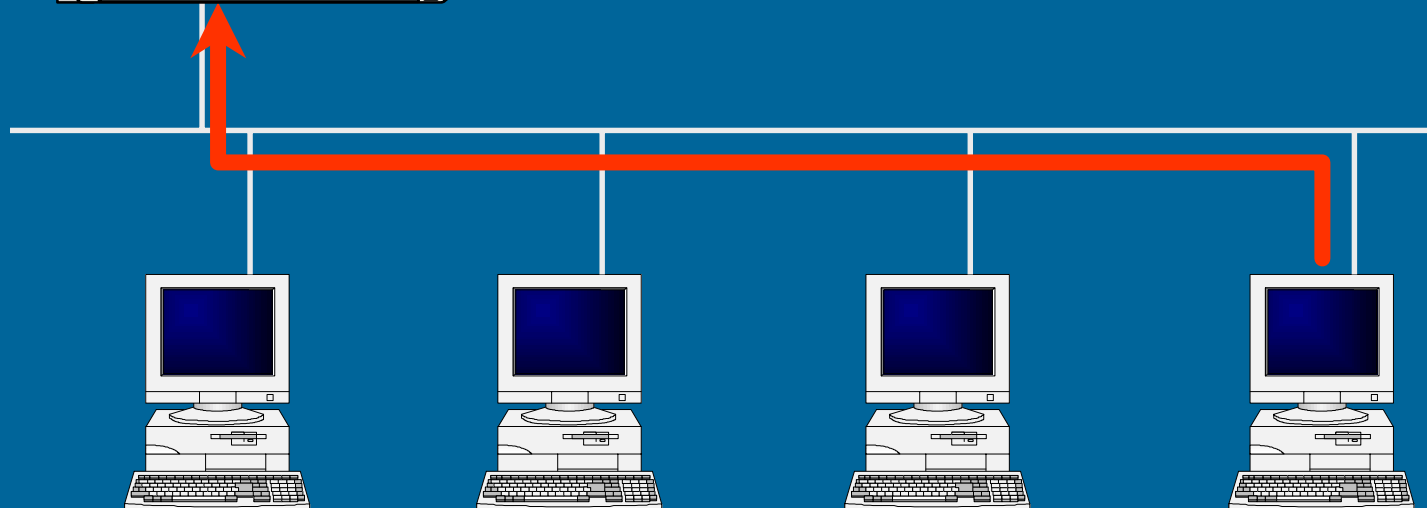
(IGMPv3 での単純な join message)

239.123.45.6, EXCLUDE, {NULL}



IGMPの仕組み (IGMPv3 での単純な leave message)

239.192.50.18, INCLUDE, {NULL}



239.133.10.30

239.134.25.2
239.135.62.12

239.133.10.30

239.133.10.30
239.192.50.18
239.202.23.64



IGMPv3 cont.

受信ホストの中で条件が重なったら

- EXCLUDE {a, b, c, d,}
- EXCLUDE { b, c, d, e} a と e は必要 (両方いないものだけ消す)
- INCLUDE { d, e, f} d は明示的に必要

= EXCLUDE { b, c}

- INCLUDE {a, b, c}
- INCLUDE {b, c, d}
- INCLUDE {e, f}

= INCLUDE {a, b, c, d, e, f} 全部合わせる



IGMPv3 Membership Query

- ルータからホストへ受信確認
- GroupAddr が 0 なら General Query (dest addr は 224.0.0.1)
- SrcAddr 指定なしなら Group-Specific Query (dest addr は Mcast addr)
- SrcAddr 指定ありなら Group-and-Source-Specific Query (dest addr は Mcast addr)



IGMPv3 Membership Report

- グループレコードを(複数)指定
- グループレコードはマルチキャストアドレスと(複数の)ソースアドレスからなる
- レコードのタイプは 3種類
 - Current-State Record: Query への返事に使う
 - MODE_IS_INCLUDE or MODE_IS_EXCLUDE
 - Filter-Mode-Change Record: 先に IPMulticastListen() で指定されているフィルタのモードが変更された
 - CHANGE_TO_INCLUDE_MODE or CHANGE_TO_EXCLUDE_MODE
 - Source-List-Change Record: フィルタのモードにアドレスの追加/削除を行なう
 - ALLOW_NEW_SOURCES or BLOCK_OLD_SOURCES



IGMPv3 Membership Report cont.

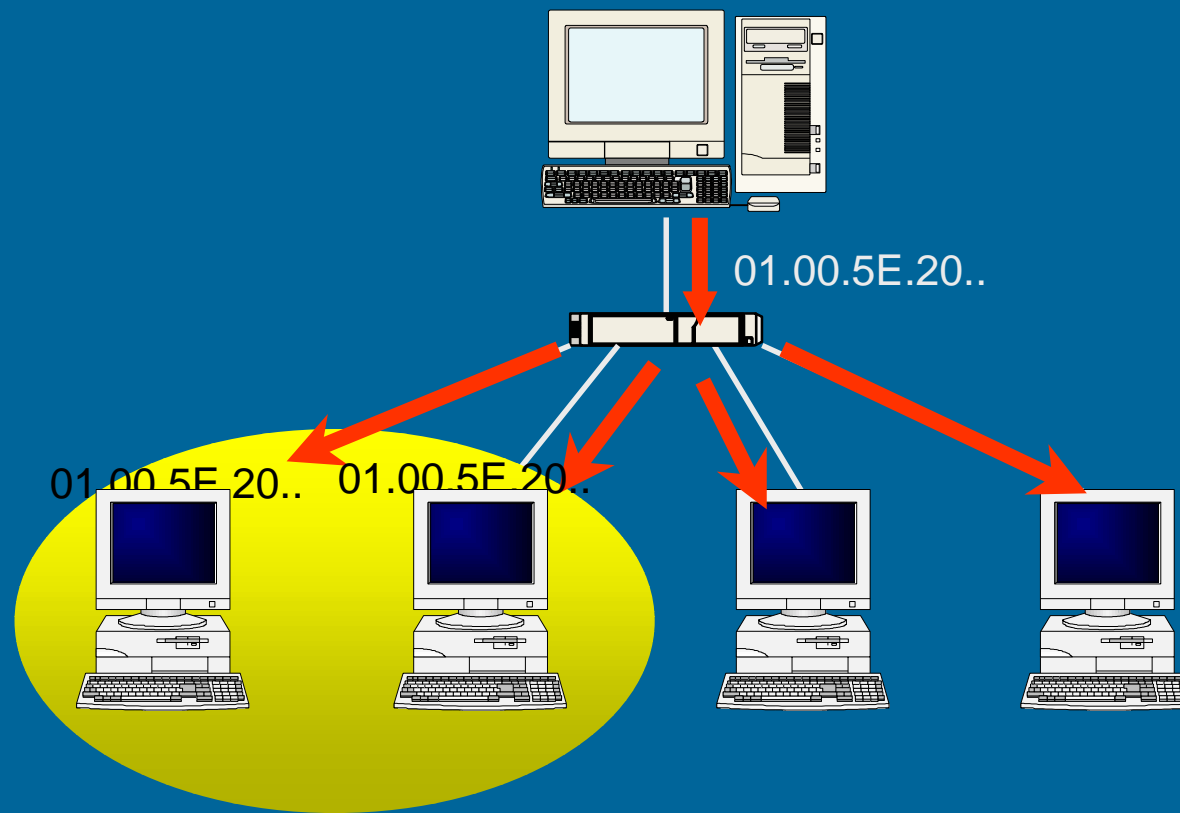
変化する例 (一つの I/F の中で変化が起こると, 右の report が発生する)

| Old State | New State | State-Change Record Sent |
|-------------|-------------|--------------------------|
| ----- | ----- | ----- |
| INCLUDE (A) | INCLUDE (B) | ALLOW (B-A), BLOCK (A-B) |
| EXCLUDE (A) | EXCLUDE (B) | ALLOW (A-B), BLOCK (B-A) |
| INCLUDE (A) | EXCLUDE (B) | TO_EX (B) |
| EXCLUDE (A) | INCLUDE (B) | TO_IN (B) |

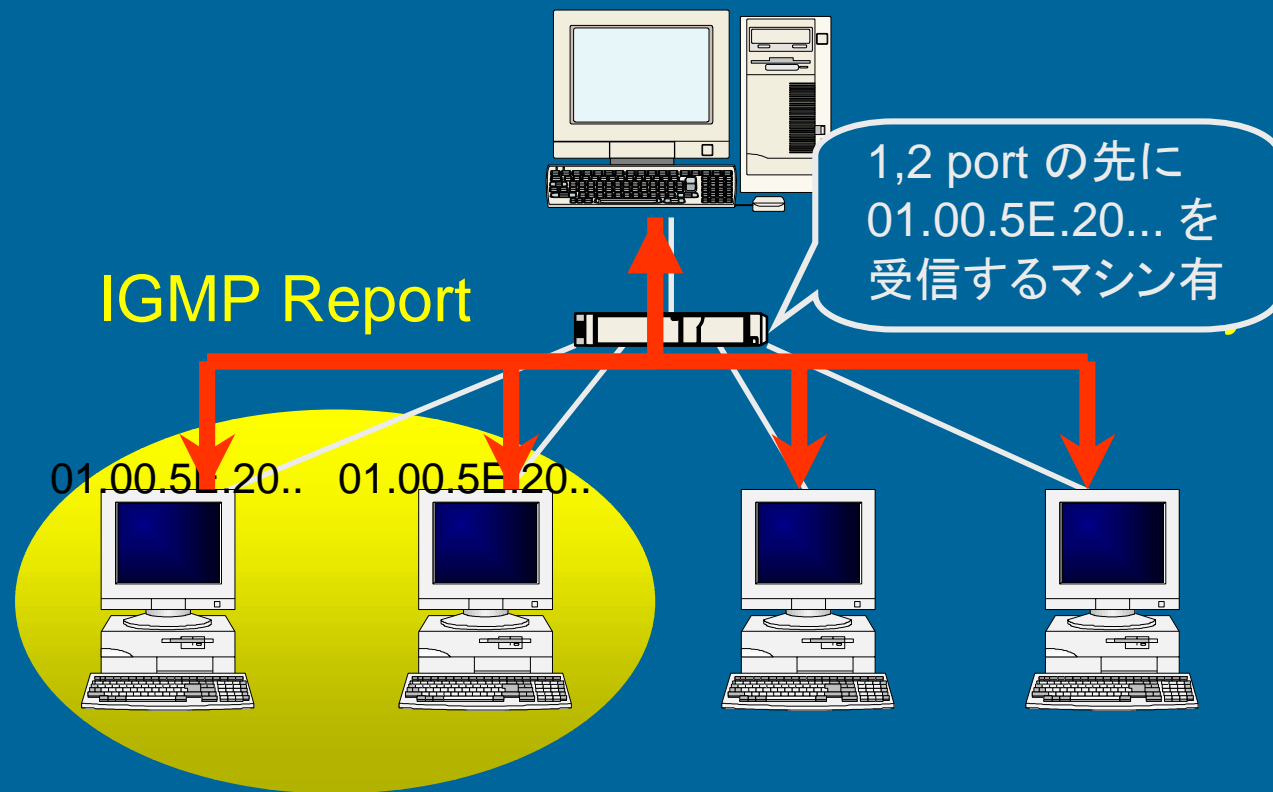
- membership report の dest addr は 224.0.0.22 を使う
 - (v1,v2 report 互換モードの場合は mcast addr を dest addr に)



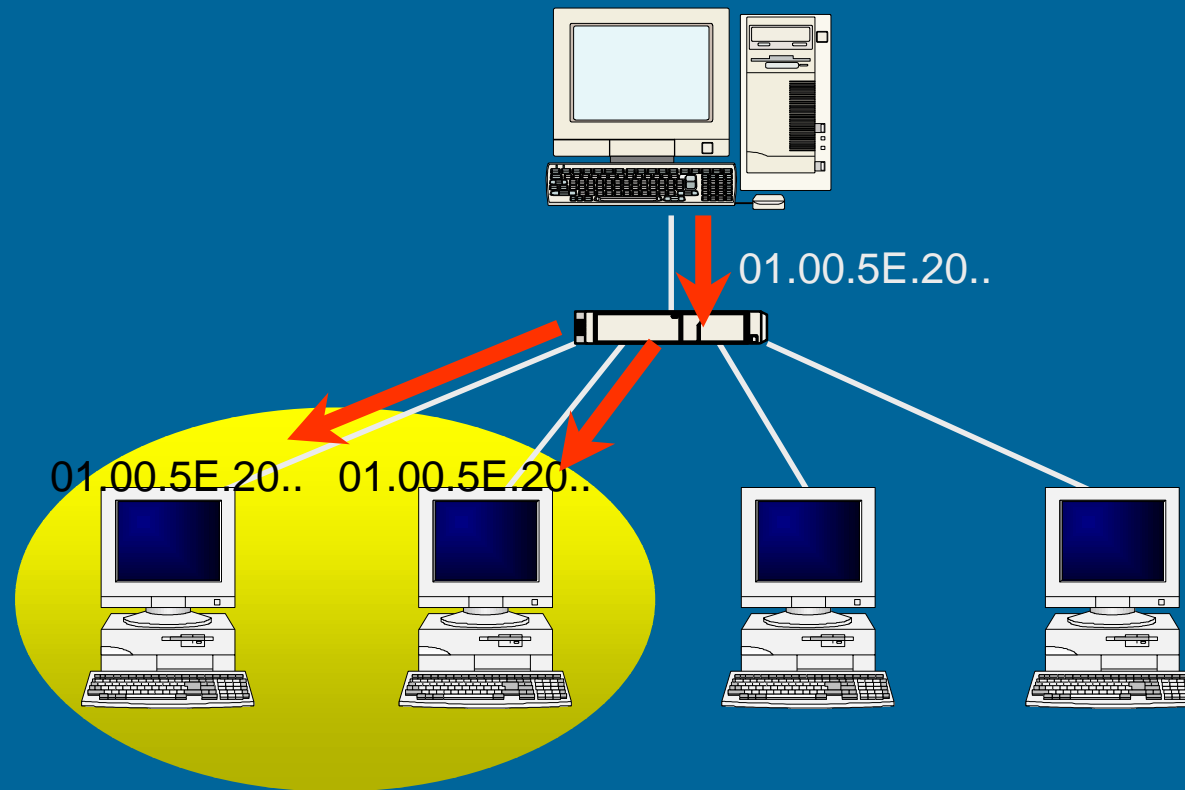
イーサネットスイッチ問題



イーサネットスイッチ問題(cont.)



イーサネットスイッチ問題(cont.)



スイッチの対応方法

- IGMP snooping
 - L3スイッチでないとCPUに負荷がかかる
 - IGMPパケットだけ選択して覗き見したいが、L2だと無理
 - L3スイッチだと価格が高くなる
 - 最近ではL2.5スイッチが登場
- CGMP(Cisco Group Management Protocol)
 - 最寄りのIGMP querierに教えてもらう
 - ベンダー依存のプロトコル
- IEEE 802.1 GMRP
 - GARP Multicast Registration Protocol
 - host(NIC)がswitchに対して受信希望グループ(L2)を申告

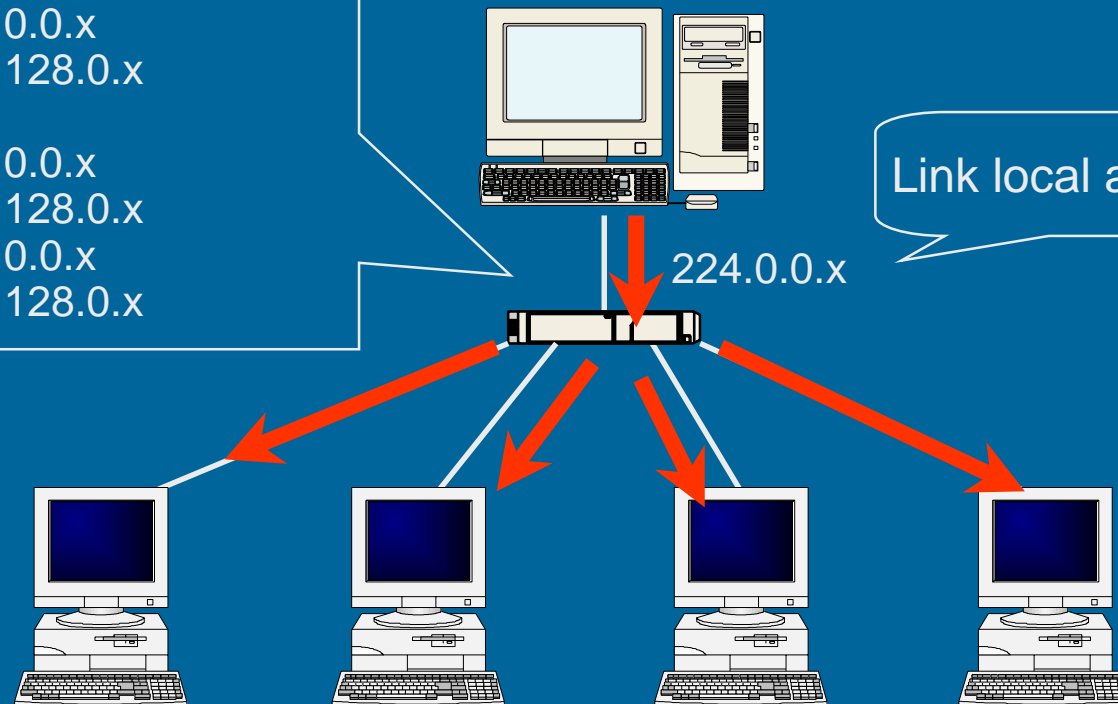


イーサネットスイッチ問題2

下位23bitのみマップ

224.0.0.x
 224.128.0.x
 225.0.0.x
 225.128.0.x
 :
 238.0.0.x
 238.128.0.x
 239.0.0.x
 239.128.0.x

⇒ 01:00:5E:00:00:x

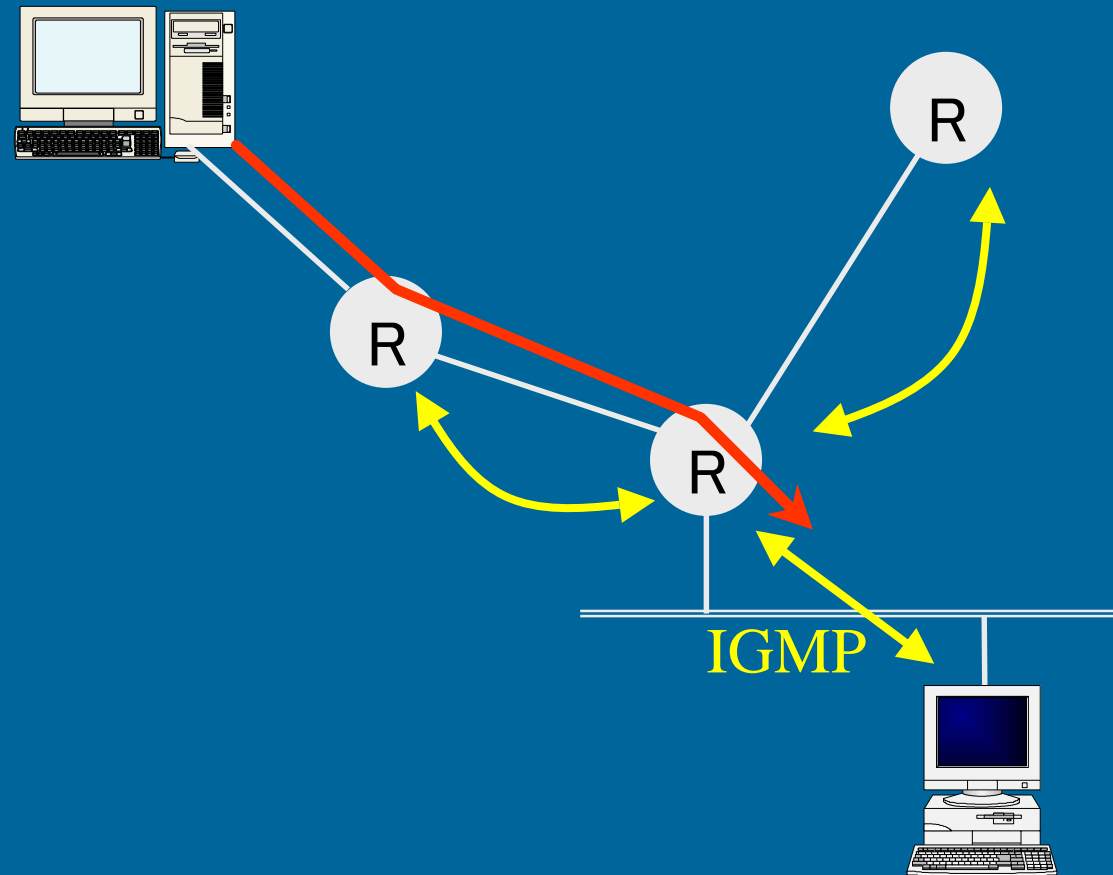


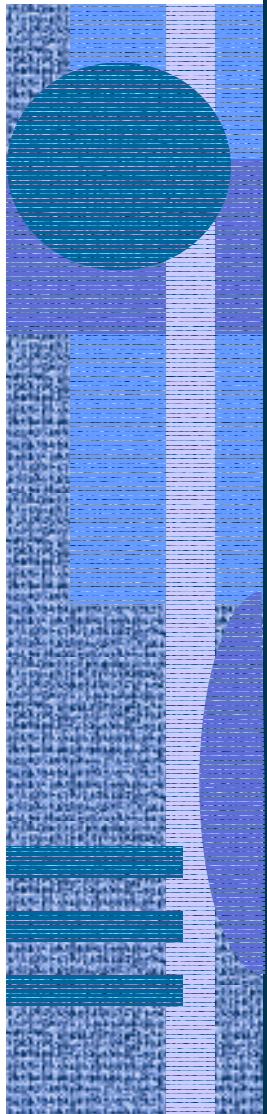
IGAP

- IGMP Authentication Protocol
- IGMP にユーザ認証の仕組みを追加する試み
- draft-hayashi-igap-03.txt
- NTT, Nortel Networks が提案
- IGMPパケットに UserAccount情報を入れておき、エッジルータがそれを使い RADIUSなどと話す
- リアルタイムに権限の更新が可能
- 詳細な視聴動向のログが取得可能



IPマルチキャストルーティングプロトコル



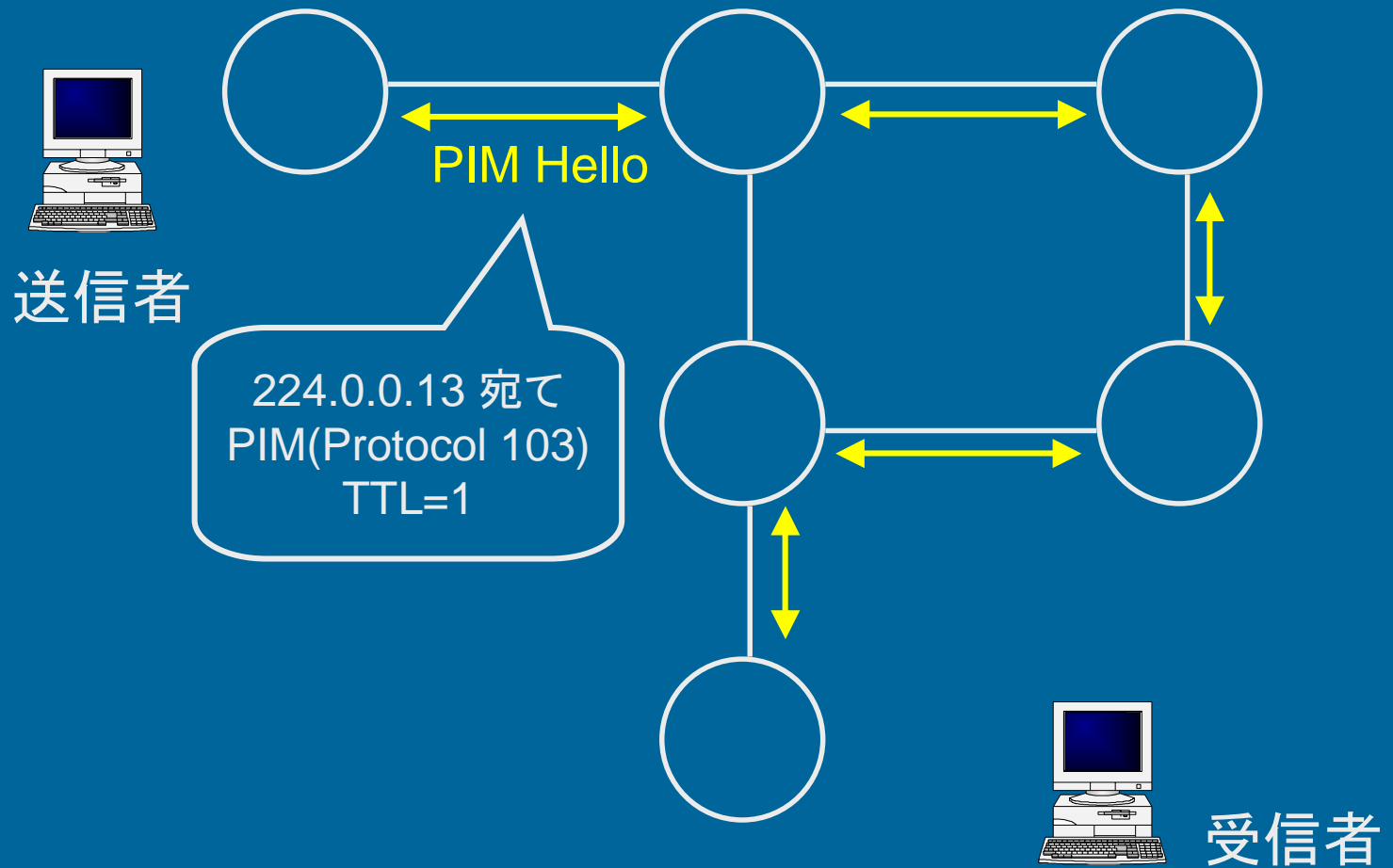


PIM

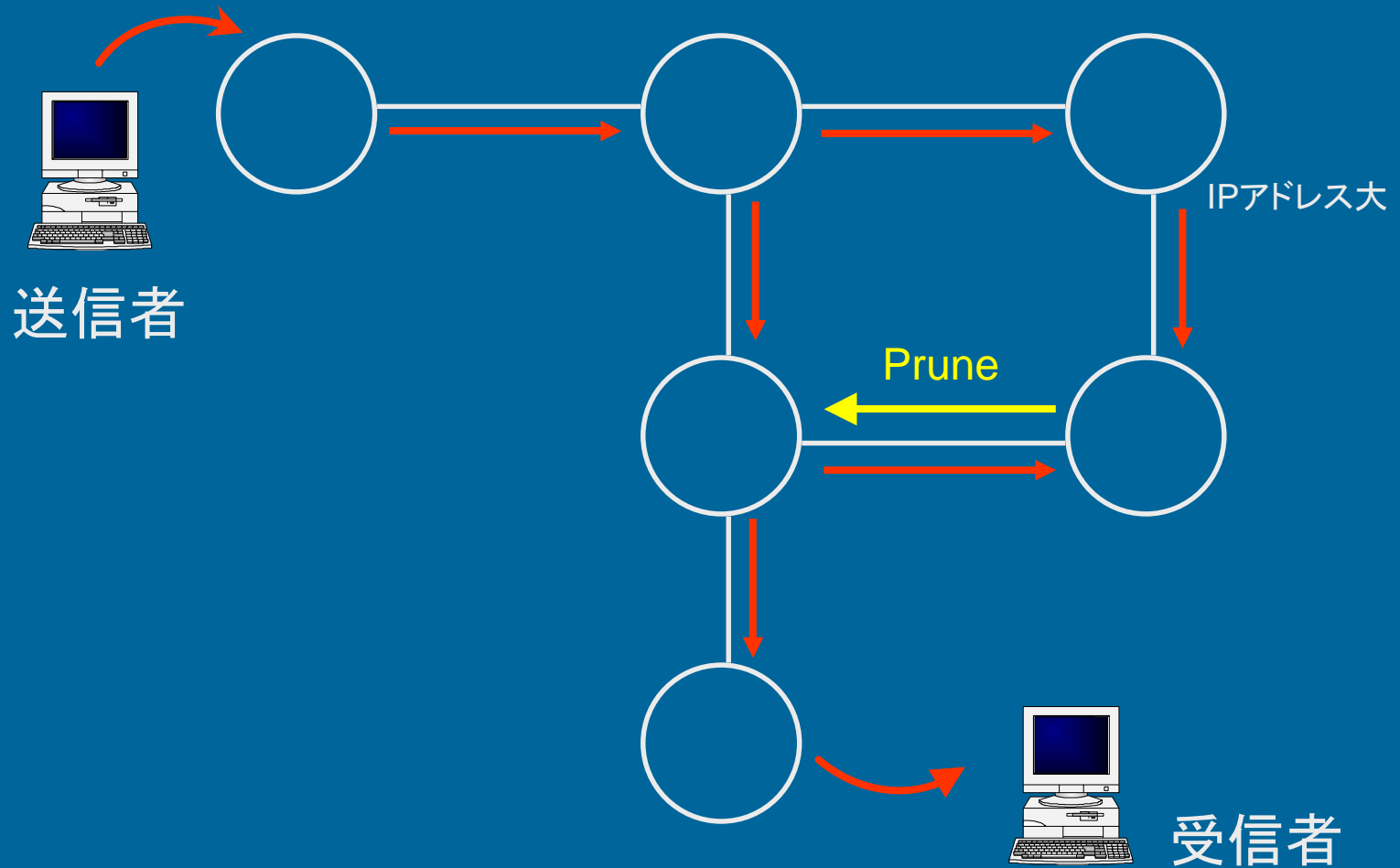
Protocol Independent Multicast

- unicast routing protocol に依存しない
- 密(Dense)モード
 - 狭い地域で、受信者が同じ番組を一斉に受信するような場合
 - 大学、企業(事業所)
 - flooding & pruning
- 疎(Sparse)モード
 - 広い地域で、受信者が離散しており、番組が沢山ある場合
 - ランデブーポイントを設定
 - 送信者はRPへ向けて送信
 - 受信者はRPへ明示的にもらいに行く

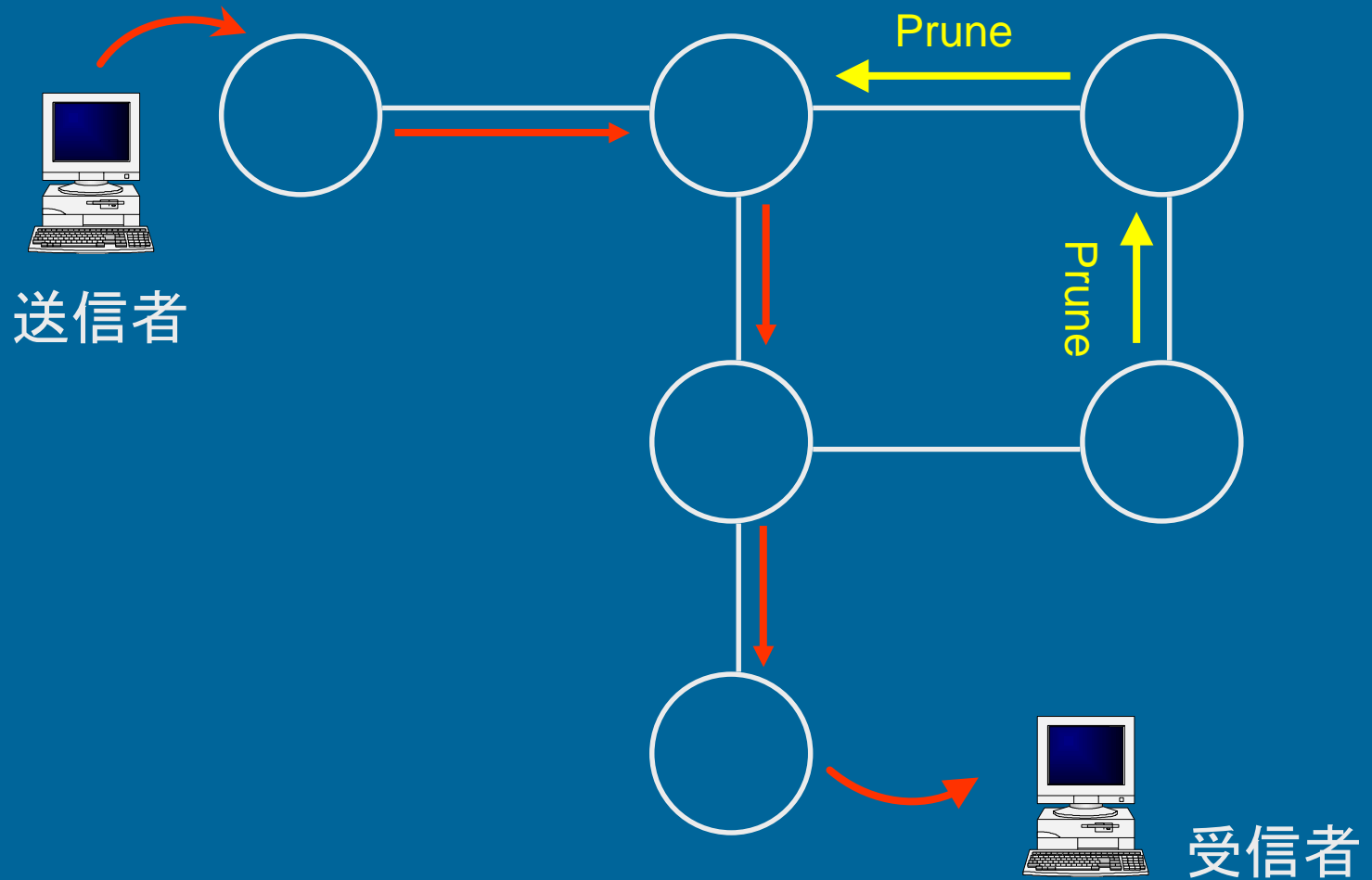
PIM Dense Mode



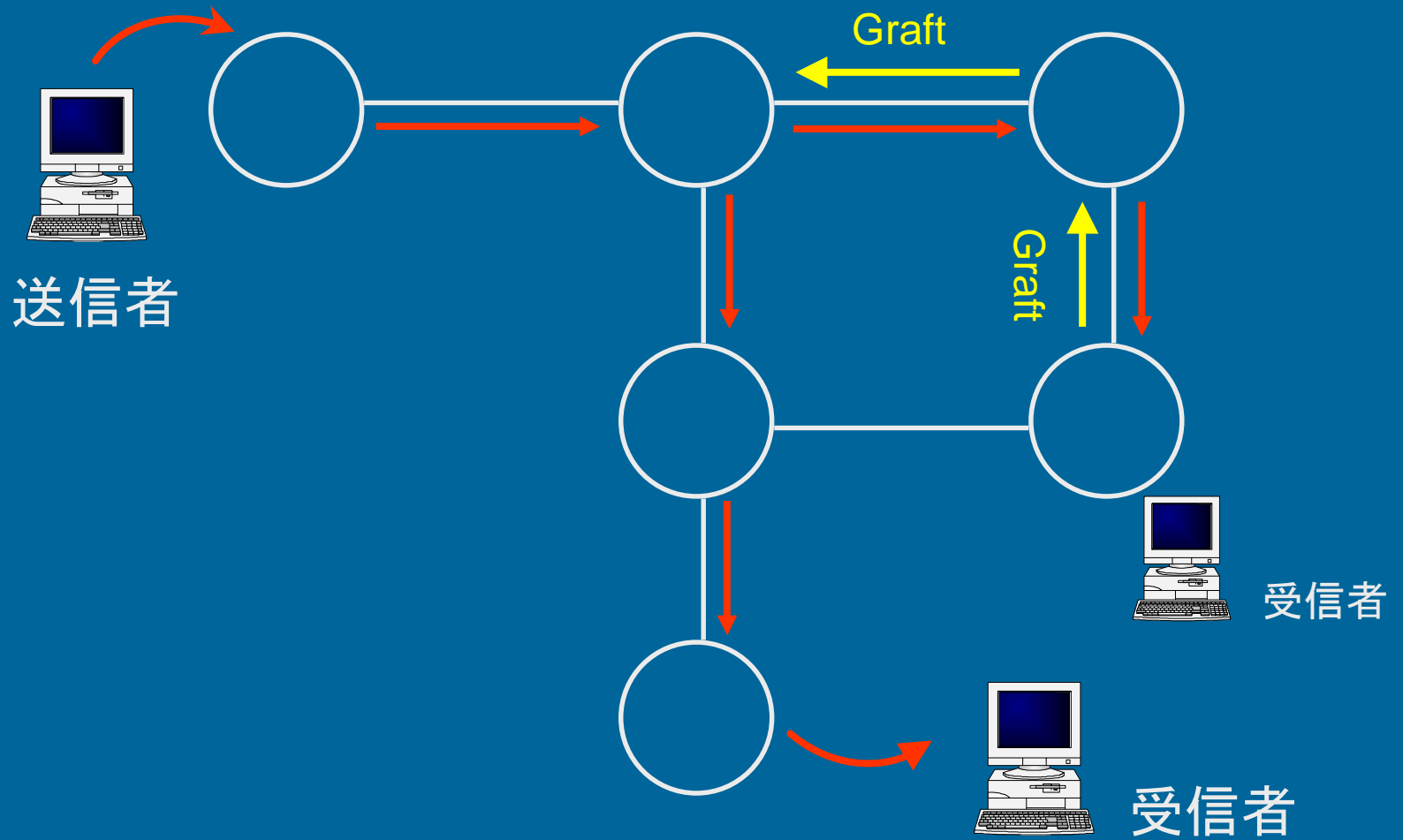
PIM Dense Mode



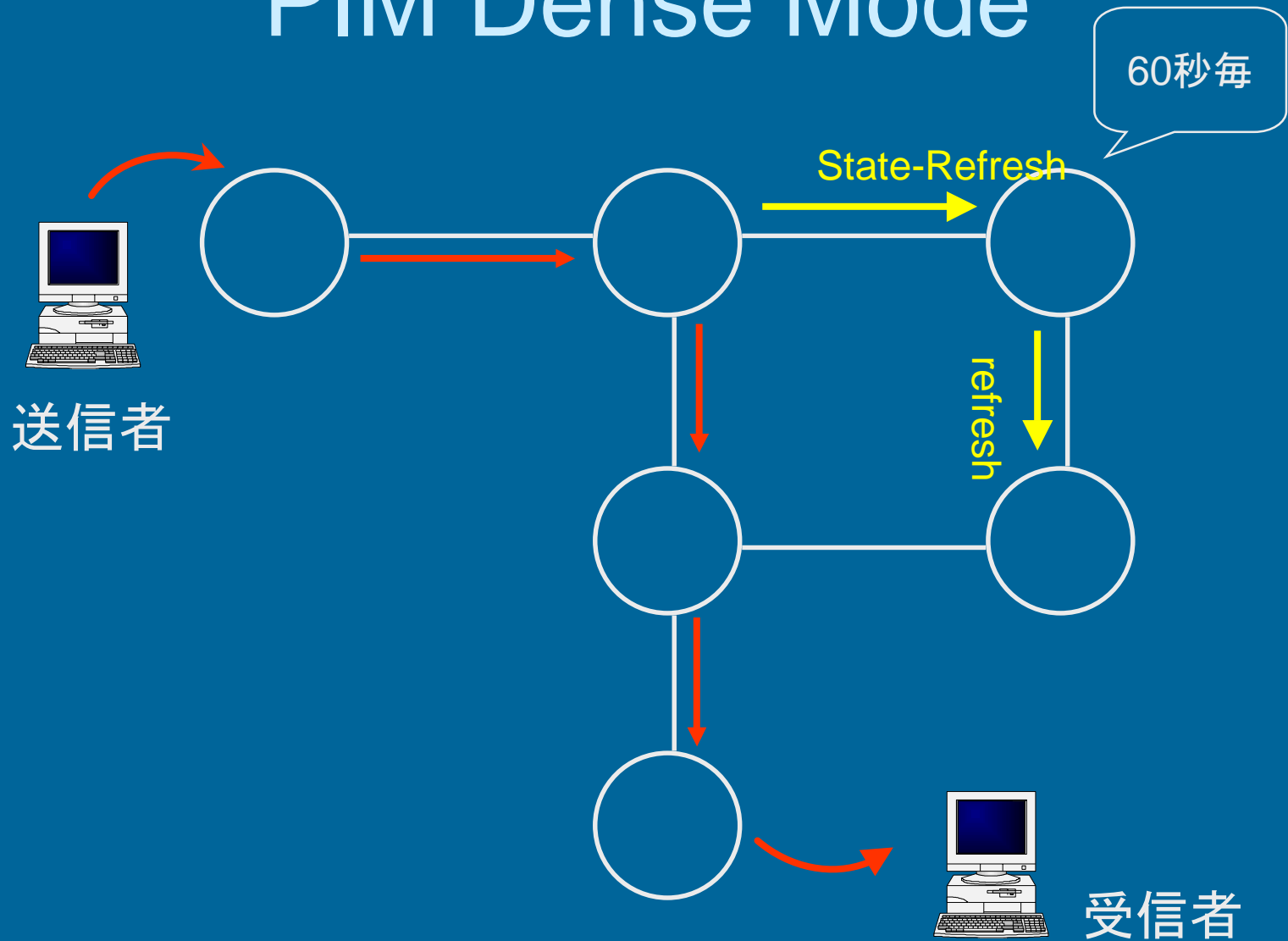
PIM Dense Mode

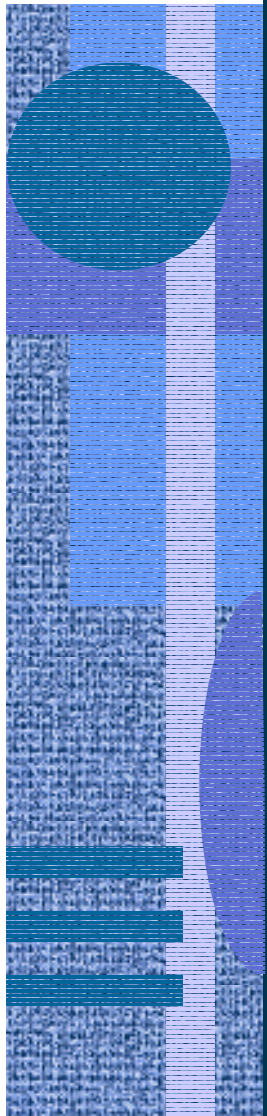


PIM Dense Mode



PIM Dense Mode

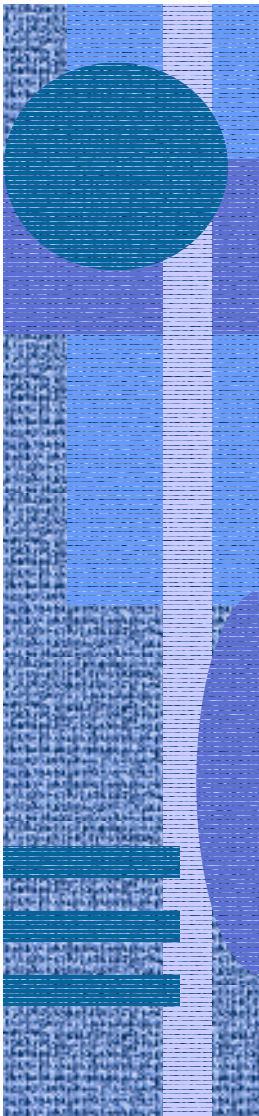




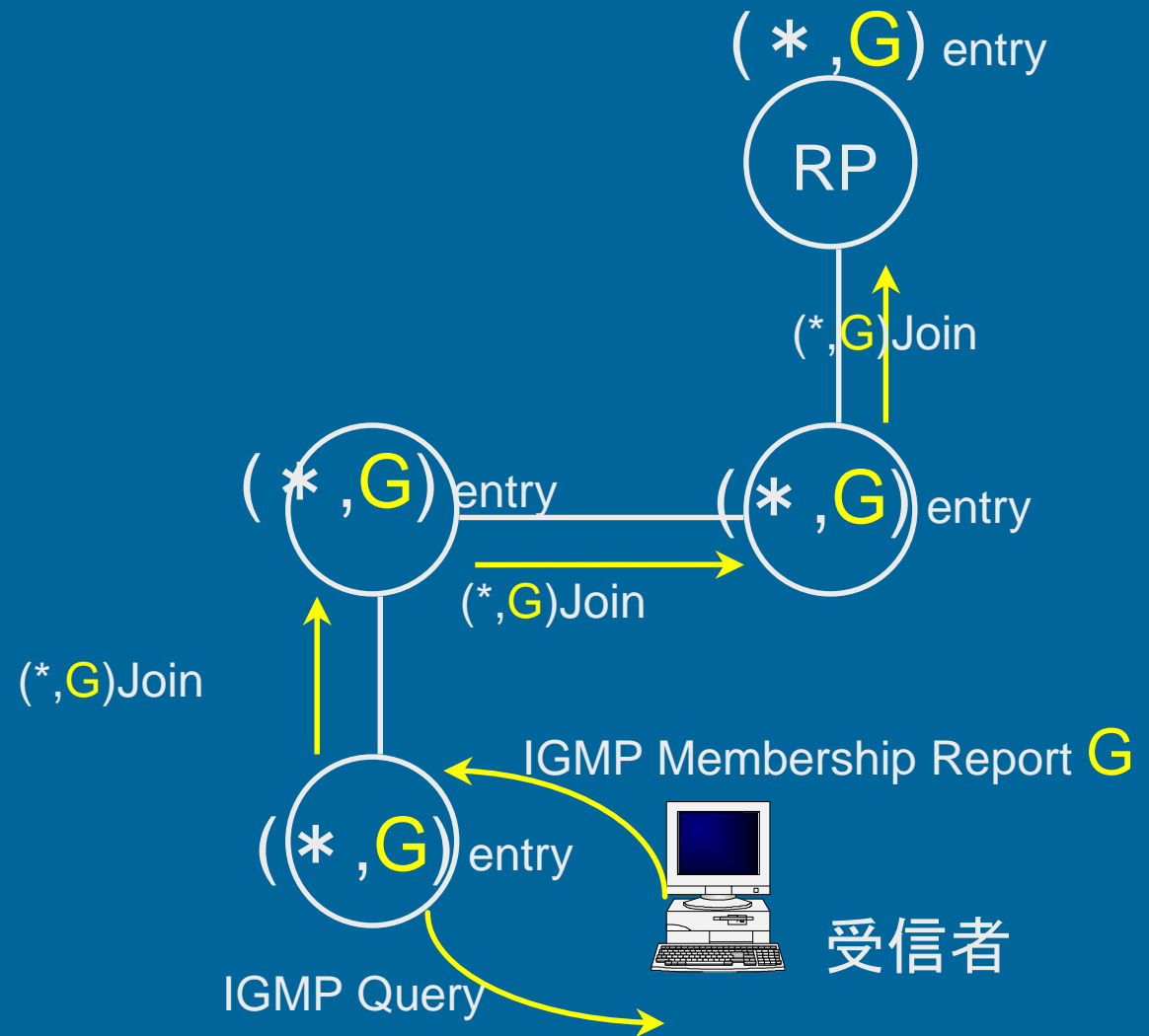
PIM

Protocol Independent Multicast

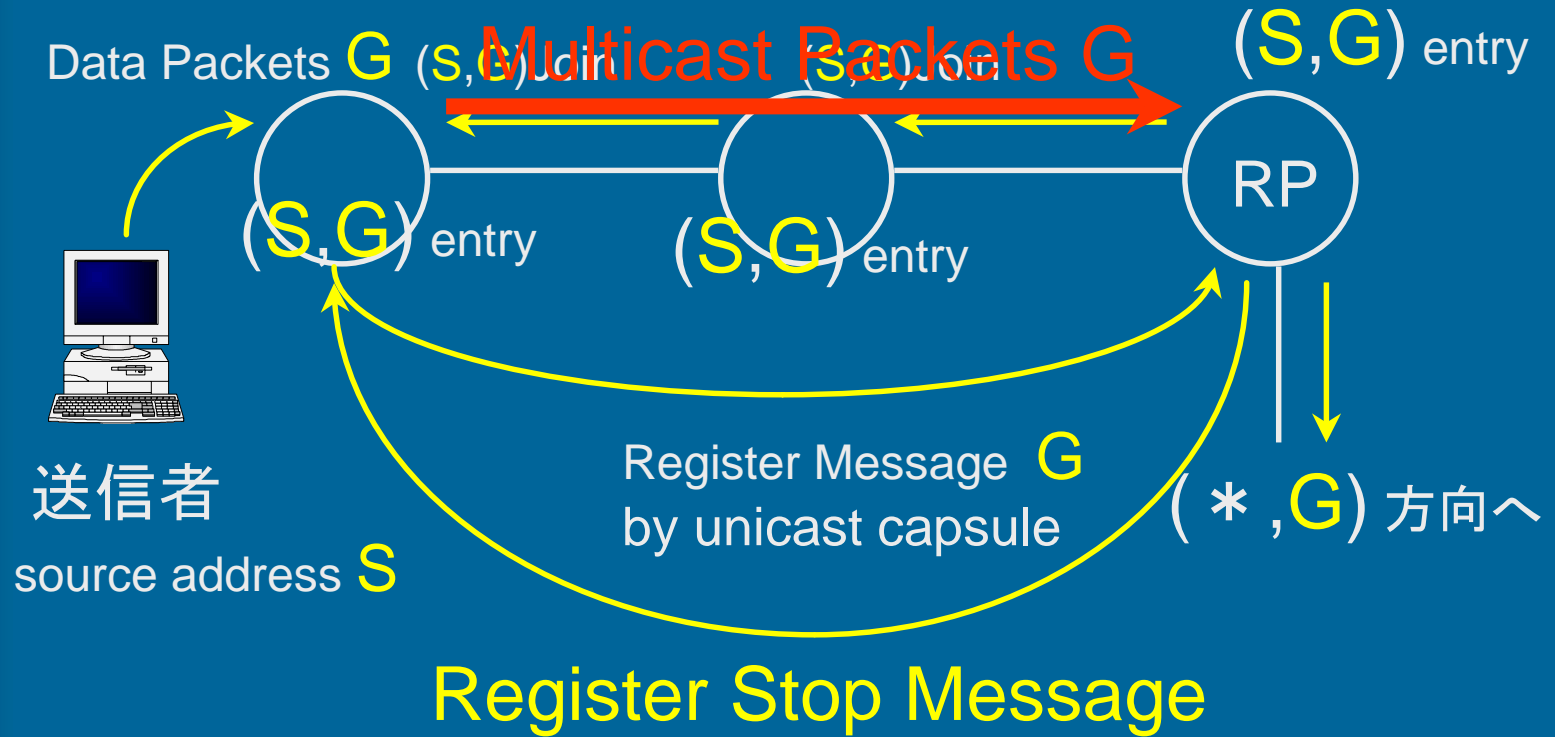
- unicast routing protocol に依存しない
- 密(Dense)モード
 - 狭い地域で、受信者が同じ番組を一斉に受信するような場合
 - 大学、企業(事業所)
 - flooding & pruning
- 疎(Sparse)モード
 - 広い地域で、受信者が離散しており、番組が沢山ある場合
 - ランデブーポイントを設定
 - 送信者はRPへ向けて送信
 - 受信者はRPへ明示的にもらいに行く



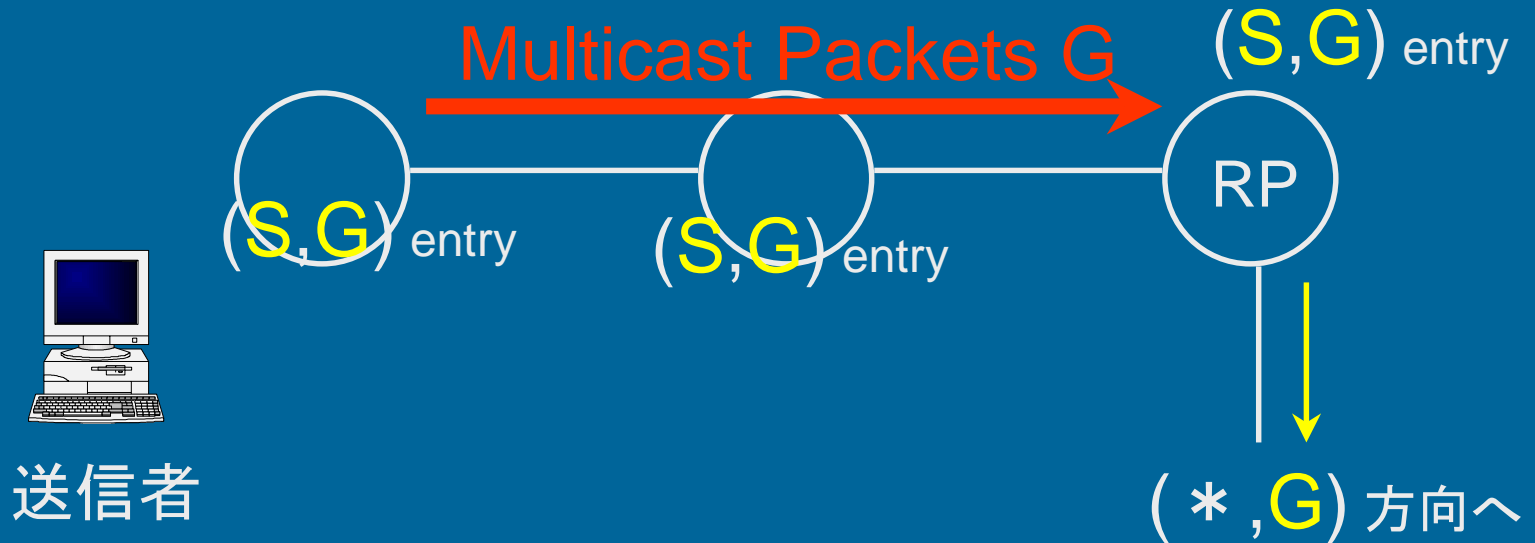
PIM Sparse Mode

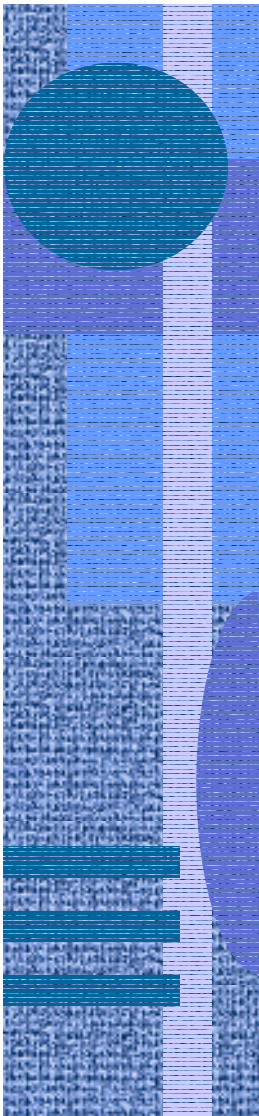


PIM Sparse Mode



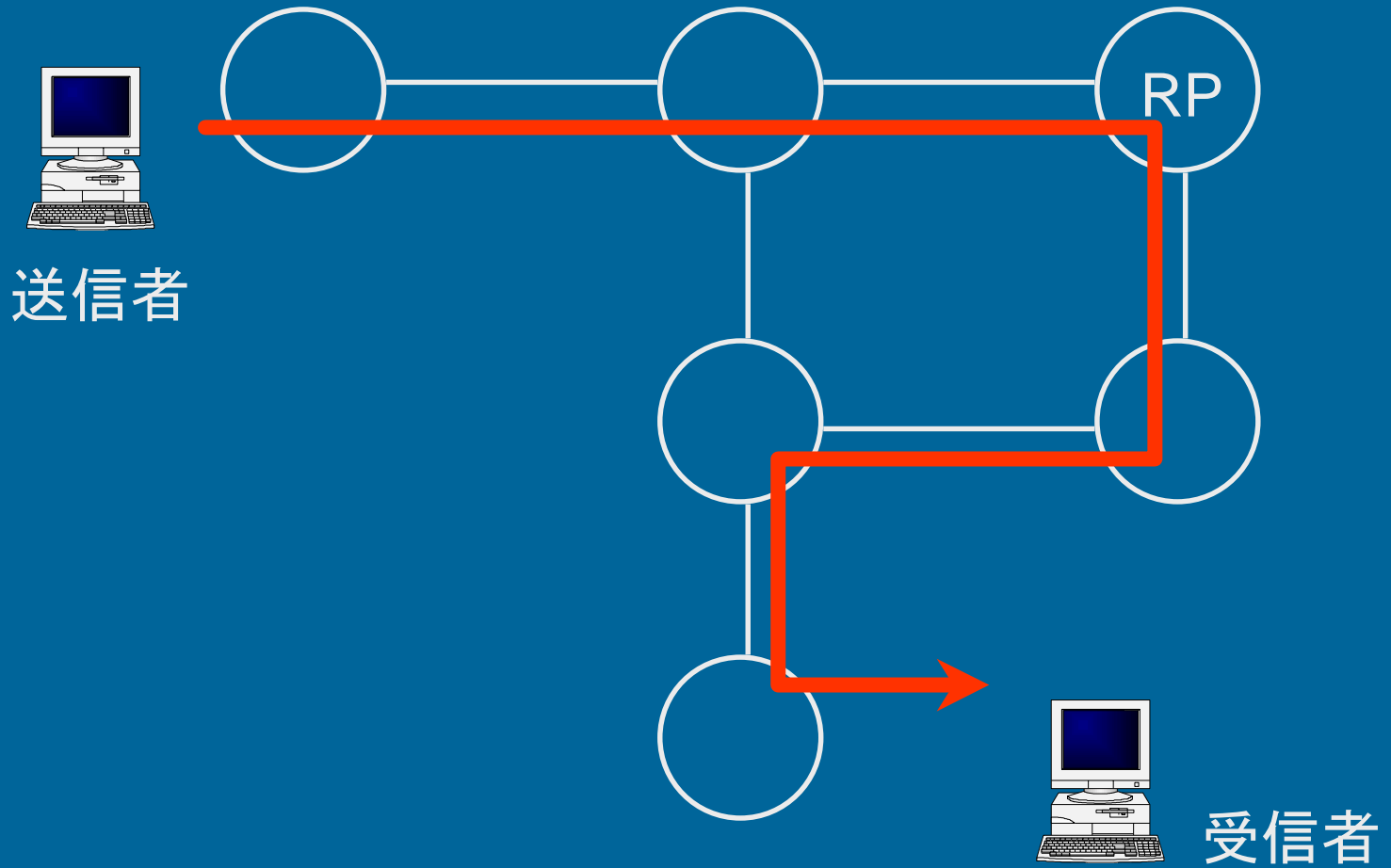
PIM Sparse Mode



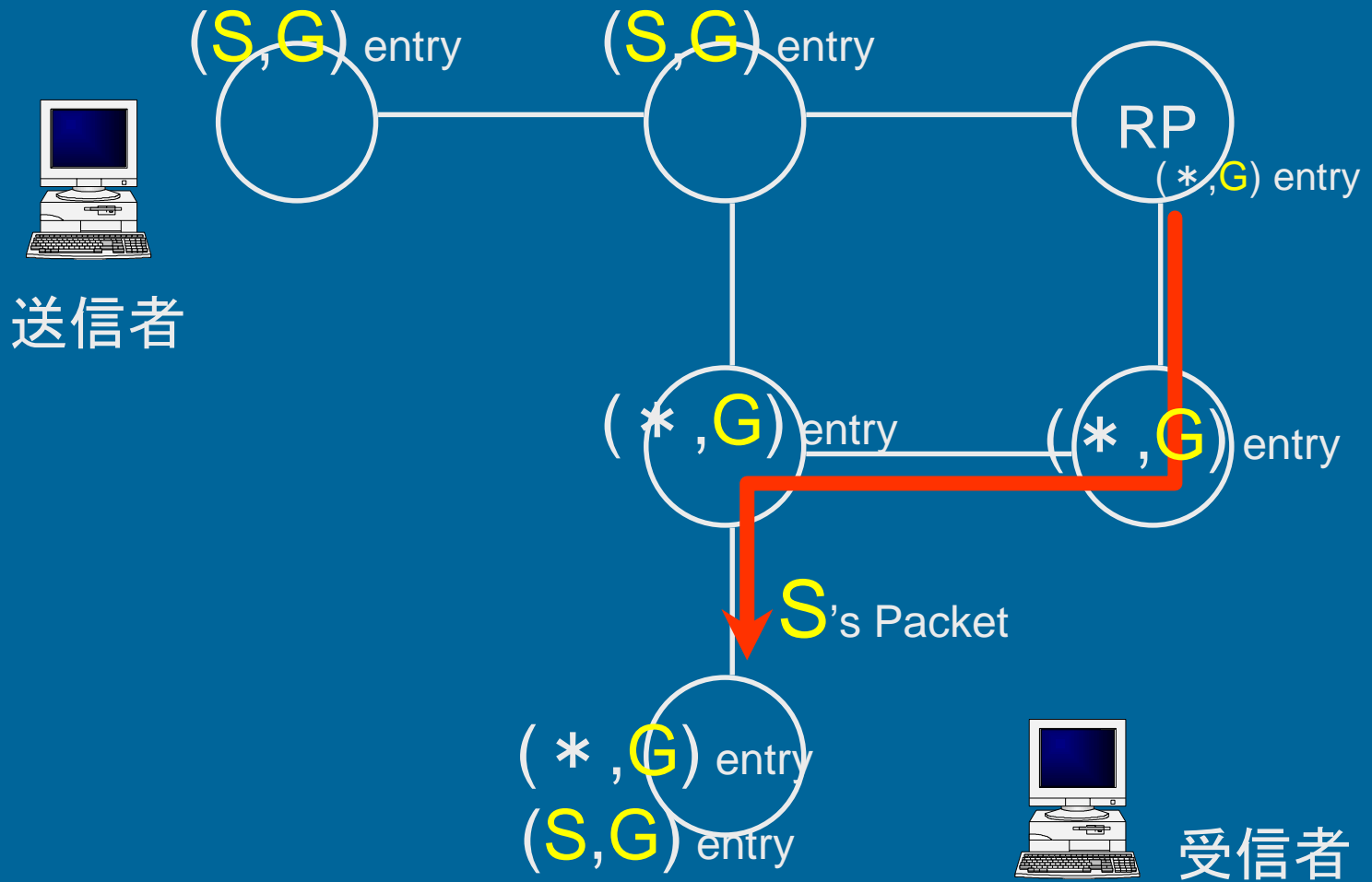


PIM Sparse Mode

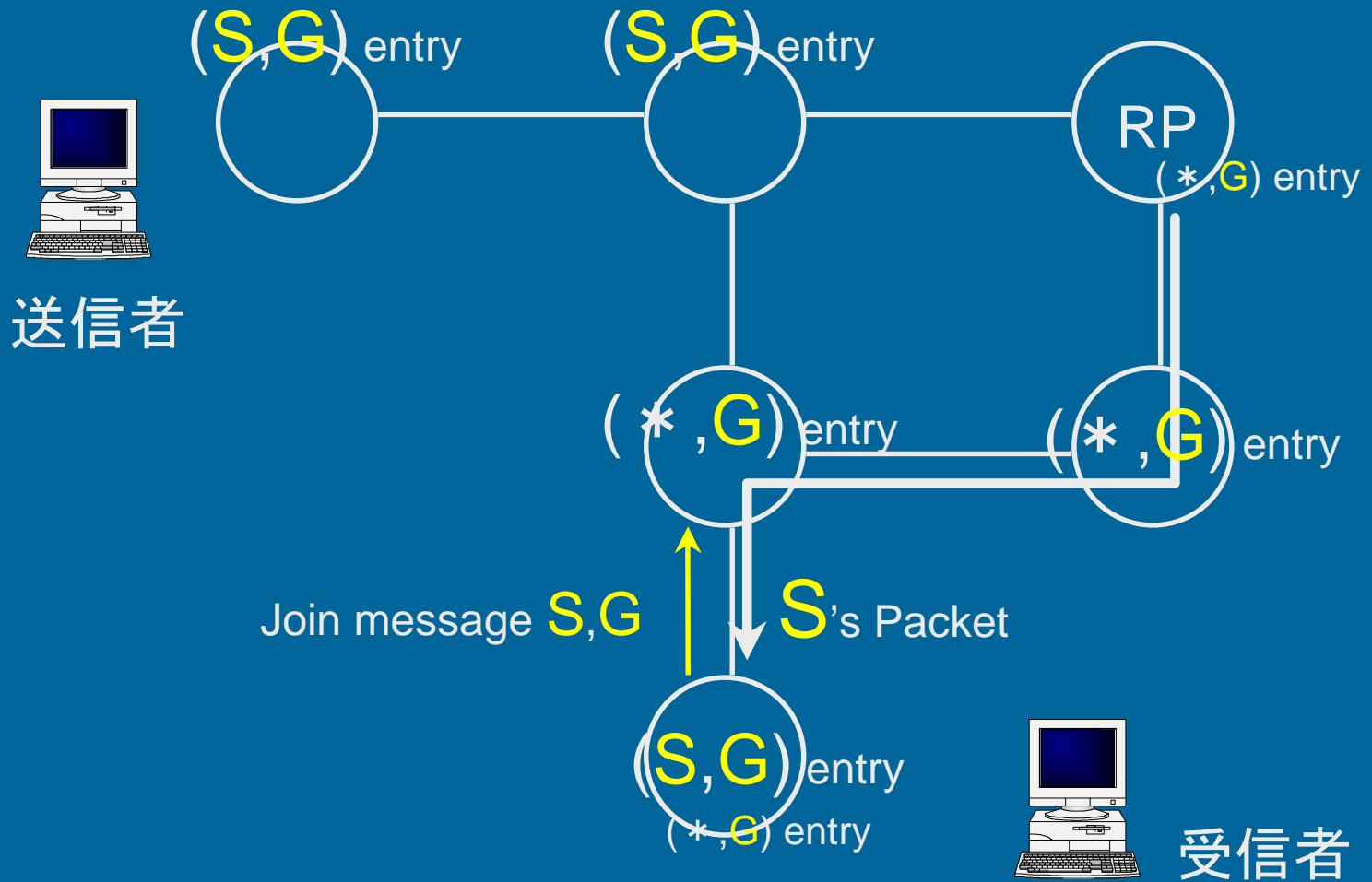
Multicast Packets G



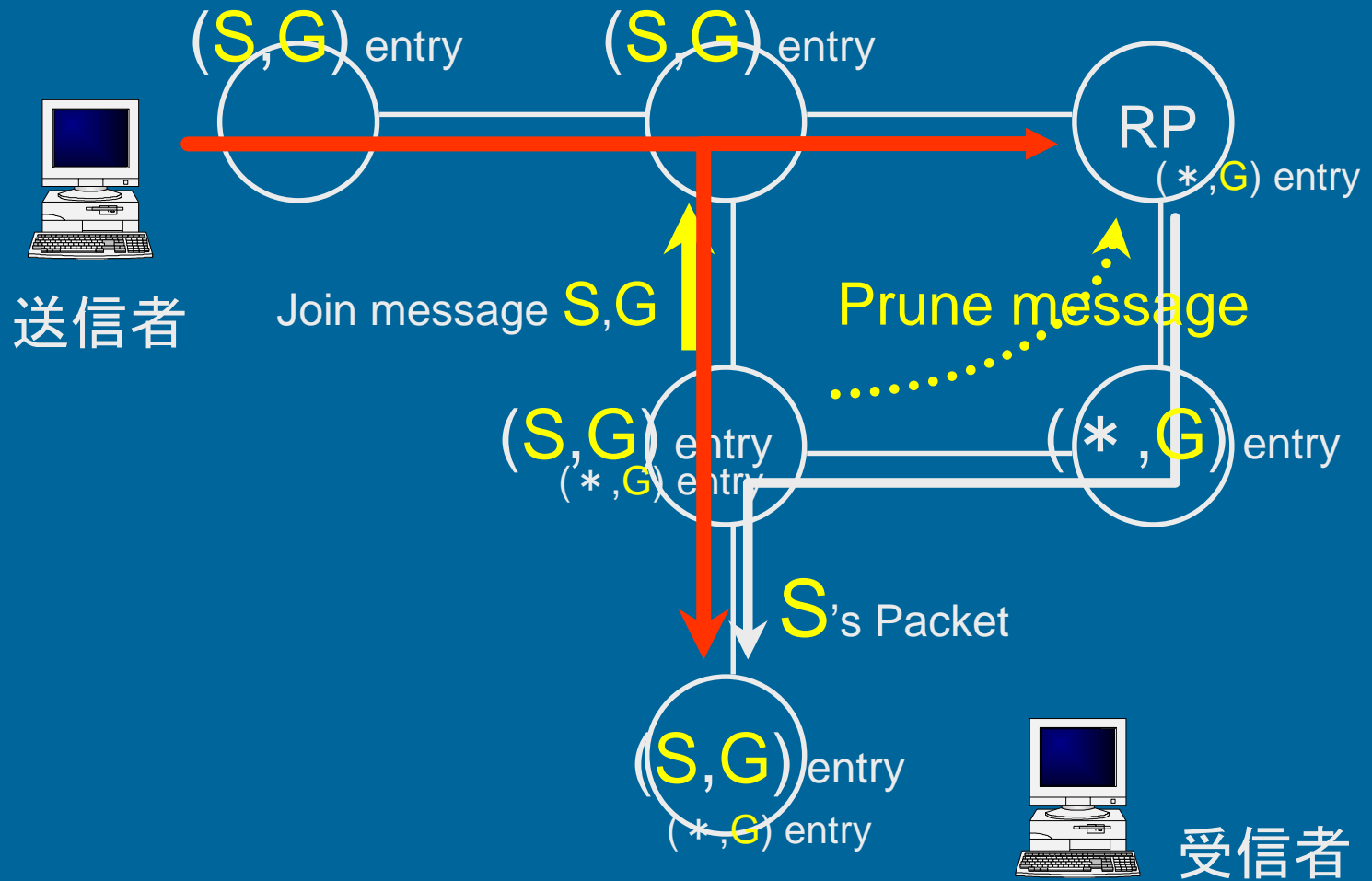
PIM Sparse Mode



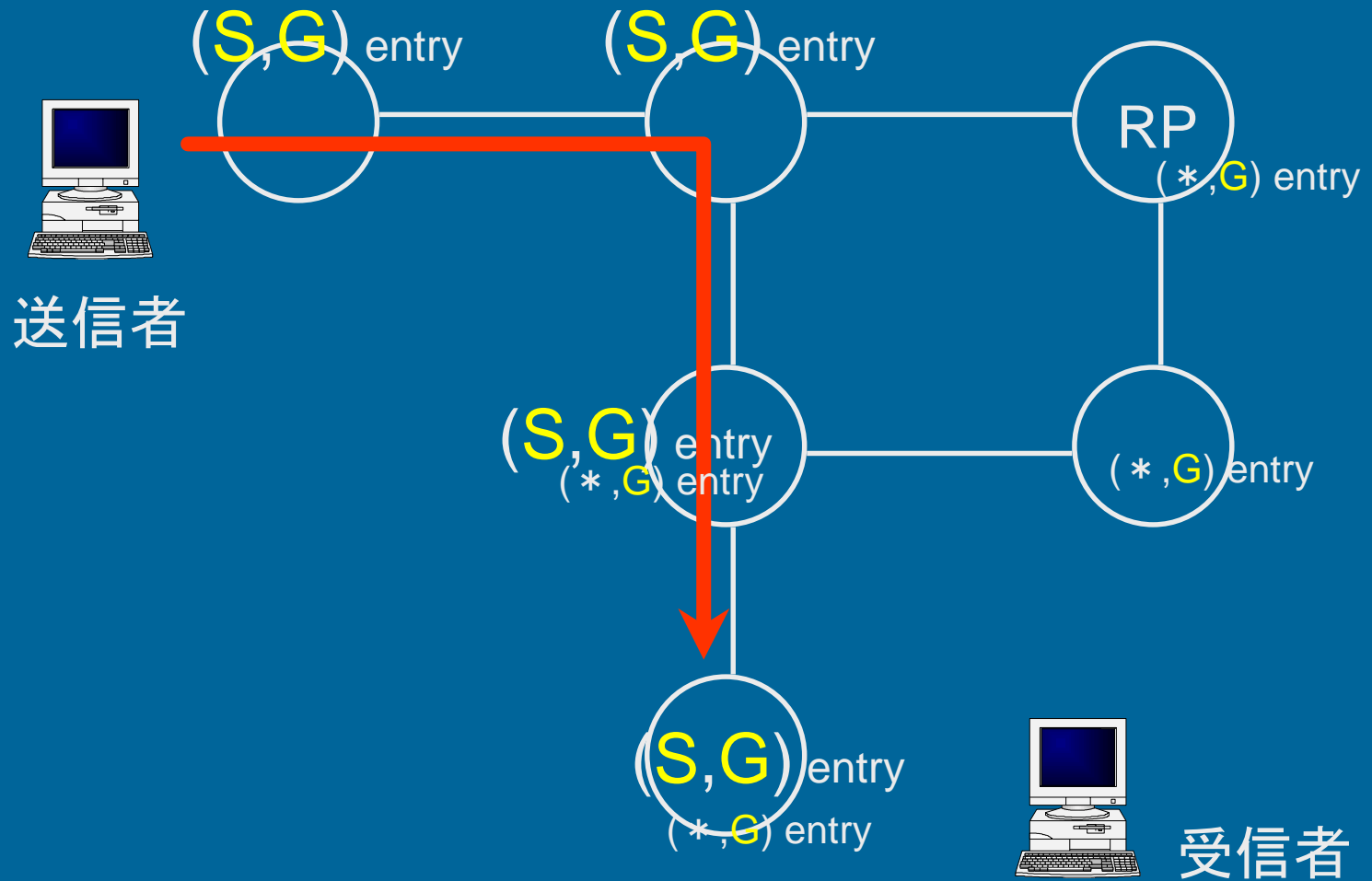
PIM Sparse Mode



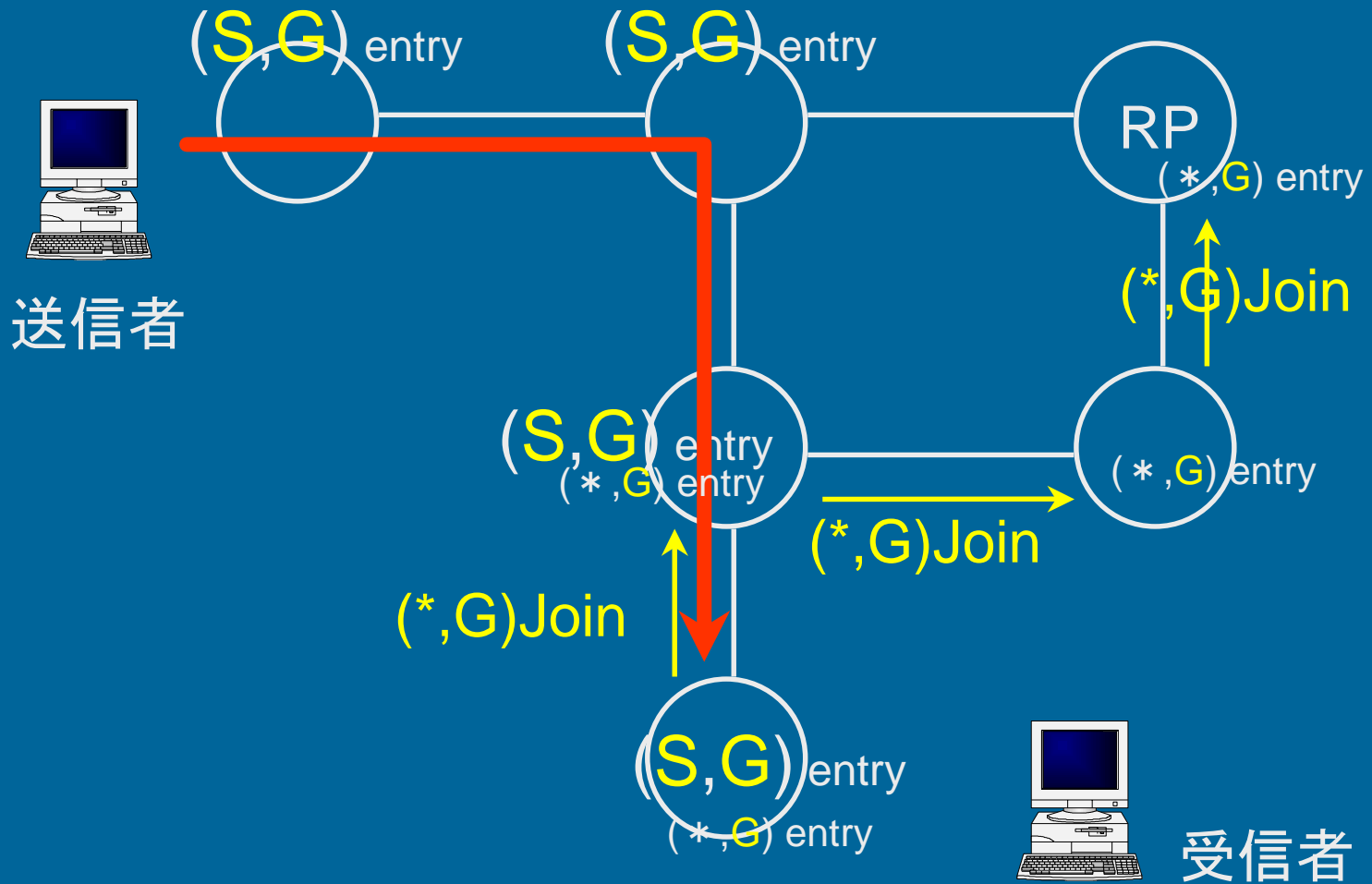
PIM Sparse Mode



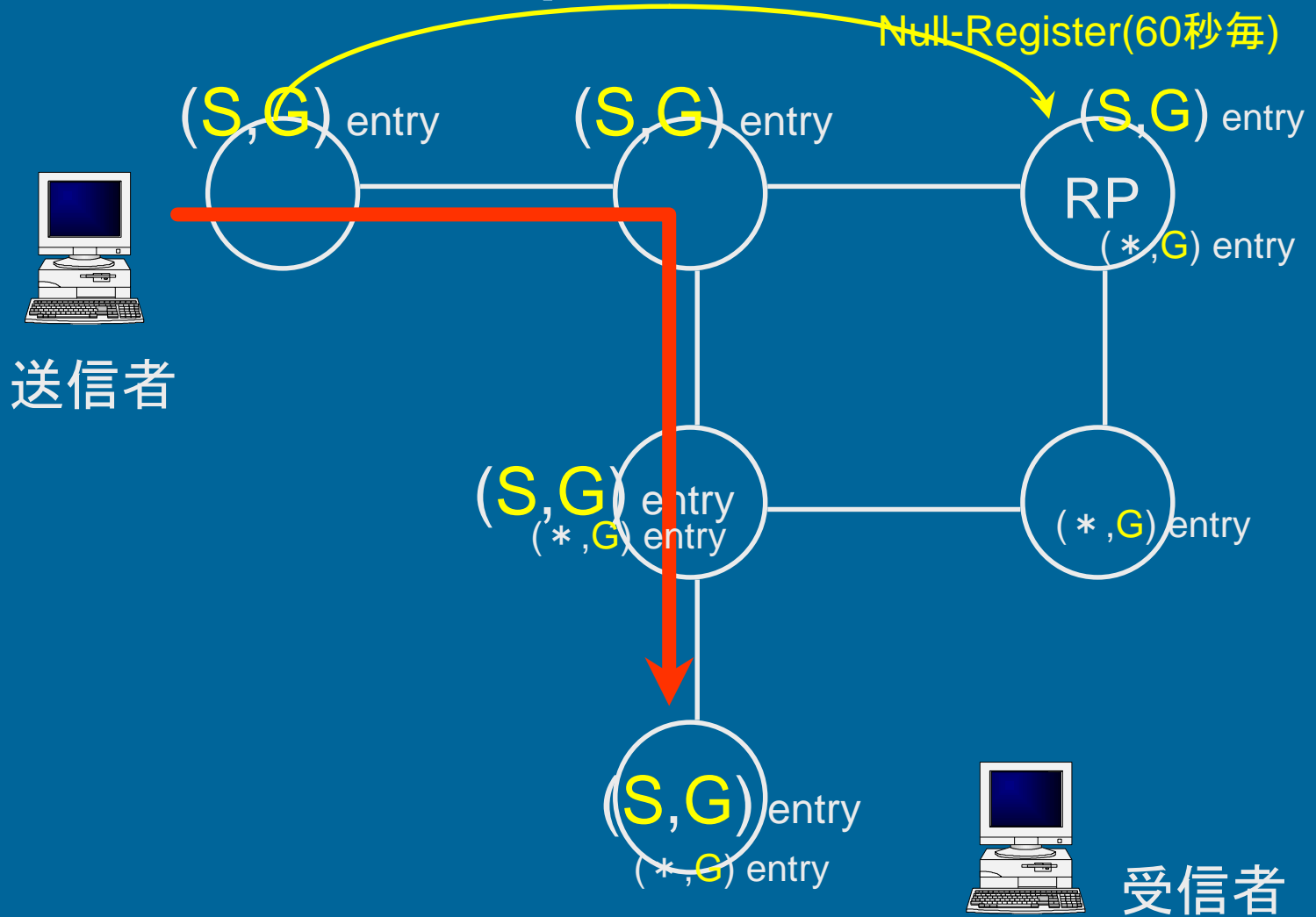
PIM Sparse Mode



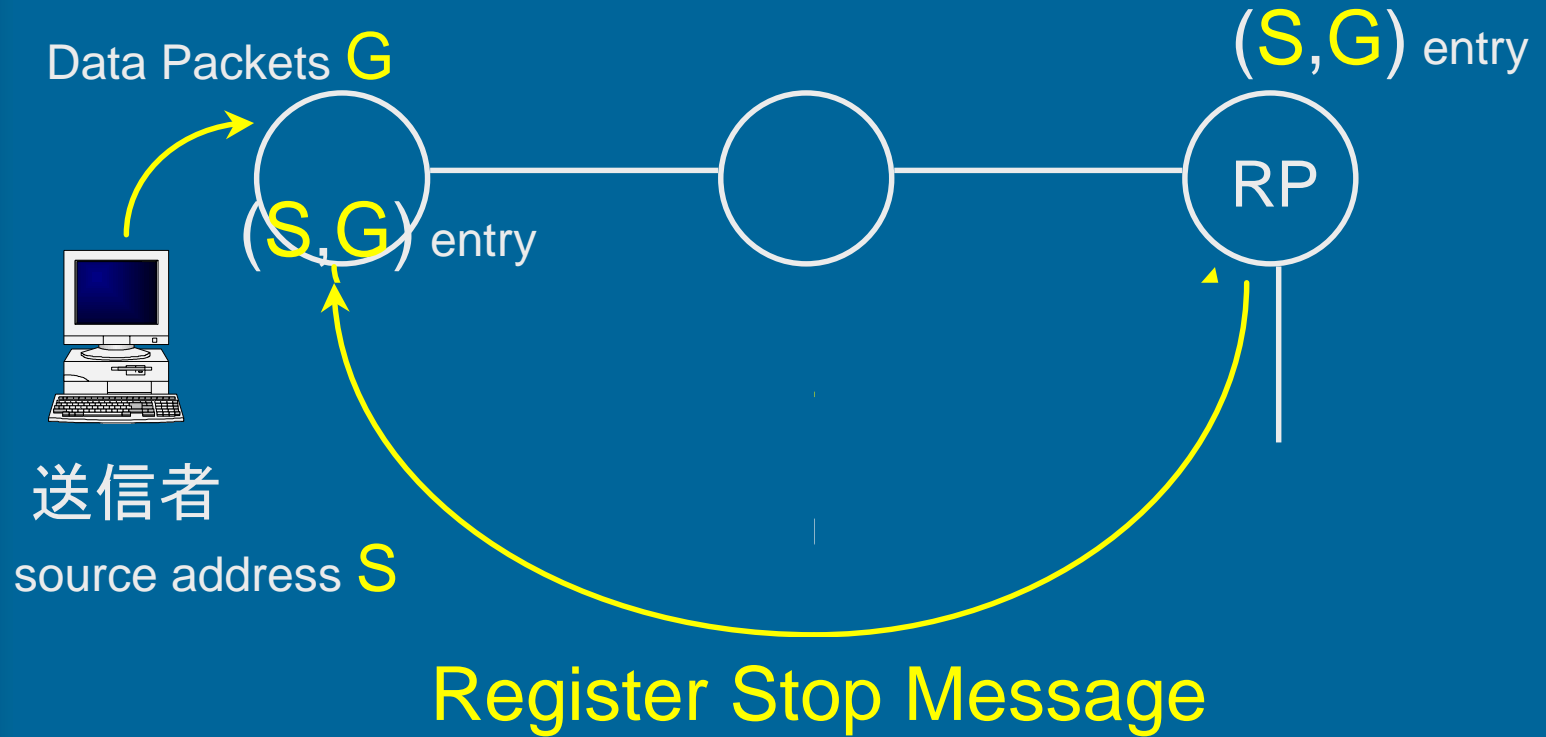
PIM Sparse Mode



PIM Sparse Mode



PIM Sparse Mode



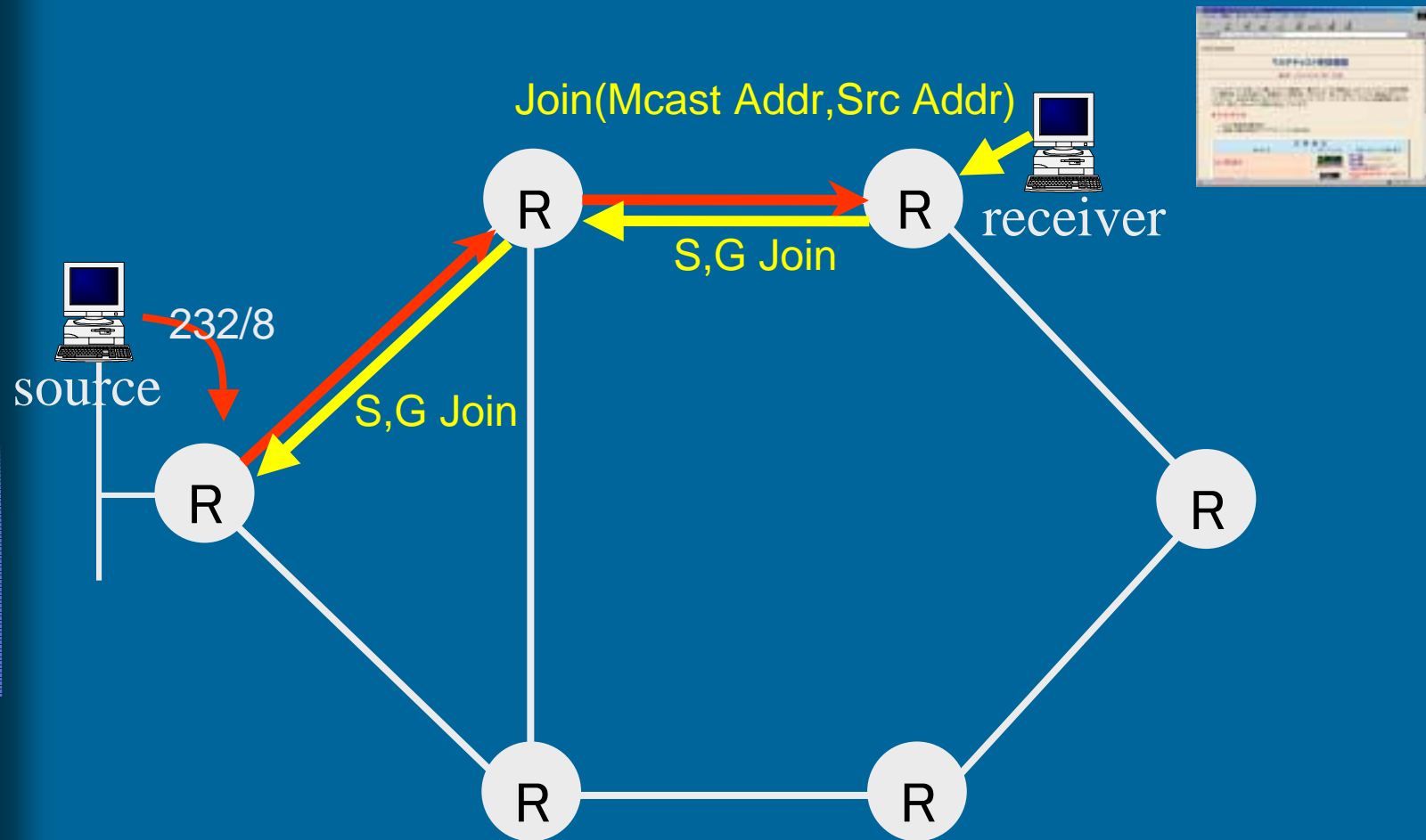
現状の課題

- ISP間のマルチキャスト相互接続
 - 各ASのRPの間をどう接続するか
 - MSDPはスケーラビリティに問題点
 - IPv6 では MSDP はサポートせず
 - SSM が一つの解
 - Embedding the Address of RP in IPv6 Multicast Address という方法も提案
- マルチキャストアドレスの重複問題
 - MASCの実装には時間がかかりそう、且つ複雑
 - これもSSMで解決



SSM

Source-Specific Multicast

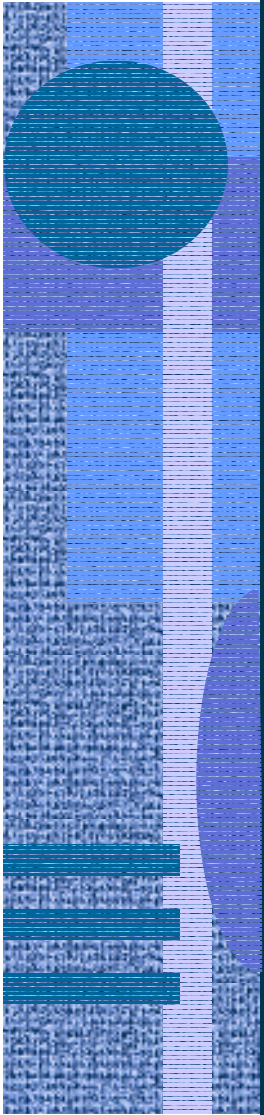


SSMの特徴

- 一対多の放送型用途を想定
- 232/8 のアドレス空間を使う(IPv6は FF3x::/96)
- 他のホストで同じマルチキャストアドレス使用可
- IGMPv3 のソースアドレス指定機能を使って実現
- クライアントOS側実装例
 - Sprint labs Linux patch
 - NetBSD,FreeBSD patch
 - MLDv2 もフランスで KAME に実装
 - Windows XP



IIJ



最近のトピック

Cisco社製ネットワーク機器の脆弱性について

- Cisco IOS Interface Blocked by IPv4 Packets
- access-list 101 permit tcp any any
- access-list 101 permit udp any any
- access-list 101 deny 53 any any
- access-list 101 deny 55 any any
- access-list 101 deny 77 any any
- **access-list 101 deny 103 any any ← pim**
- マルチキャストはトンネルで通しているサイトでも、RPが外にある場合には PIM register パケットがフィルタされてしまうので、受信はできるけど送信ができない状態になる

```
interface Tunnel0
ip unnumbered Ethernet0
ip pim sparse-mode
tunnel source Ethernet0
tunnel destination 192.168.10.1
```

```
ip route 192.168.20.1 255.255.255.255 Tunnel0
```



フレッツ網の試験サービス

- NTT東西が地域IP網内での「同報通信サービス」の試験提供を2003年7月31日より開始
- インターネット接続とは別のPPPアカウントを使用する、閉域のIPマルチキャスト網
- 1Gbpsのギガビットイーサを利用し、月額200万円



電気通信役務利用放送法にもどづく放送事業者1

- Yahoo!BBケーブルTV
 - IPv4 PIM-SM
 - MPEGをスクランブルして配信
 - MPEG2 2Mbps と 4Mbps
 - 受信はICカード入りのセットトップボックスで
 - ユーザーからの上りマルチキャストパケットはフィルタ



電気通信役務利用放送法にもどづく放送事業者2

- オンラインティーヴィ
 - ぷららネットワークスが構築したコンテンツ配信ネットワーク「4thMEDIA」利用
 - ジュピター・プログラミング、セコム、東北新社の3社
 - 沖電気製のSTB「OKI Streaming Player」を使用
 - MPEG4(1.5~2Mbps)とMPEG2(4Mbps)
 - 数分に一回CASの鍵を変更



電気通信役務利用放送法にもどづく放送事業者3

- KDDI 光プラス
 - マンションまでは 1Gbps
 - 建物内は VDSL
 - MPEG2 4Mbps (オンデマンドあり/CDN)
 - STB は松下電産 製
 - HomeGW(ルータ) は NEC アクセステクニカ(株) 製



情報交換ML

- ipmulticast@ijinet.or.jp
 - 参加条件なし
 - オープンな情報交換の場
 - 宣伝利用も歓迎
- <http://www.ijinet.or.jp/ipmulticast/maillinglist.html>

