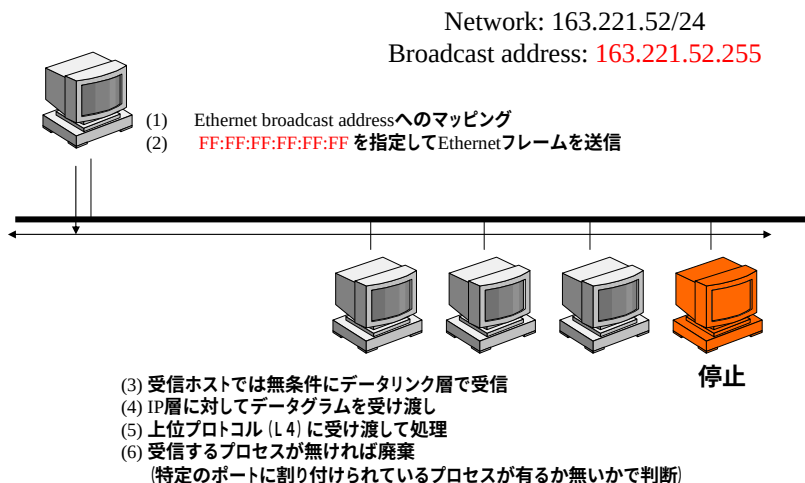


Unicast / Multicast / Broadcast / Anycast

通信モデル (1)

- Unicast
 - 1対1通信
 - インターネット上で行われる通信の大部分がUnicast
- Broadcast
 - 同報通信
 - 同一伝送メディア (データリンク) 内で稼動しているホスト全てに
対して送信
 - 同報通信機能ができるかどうかはデータリンクに依存
 - 同報通信機能がないデータリンクは沢山ある
 - 完全な同報を保証するものではない
 - 稼動していないホストは受け取れない
 - 送られたデータを処理するかどうかは受信ホストで稼動しているプロセスに依存

Broadcast



Broadcast (2)

- Ethernetは同報通信が可能なデータリンク
 - アドレスにFF:FF:FF:FF:FF:FFが指定されたフレームは、同一データリンク内のアクティブなホストは無条件に受信される
 - Ethernet が持つ機能
 - L2 switch による多段接続においても同報通信を実施
 - 同一データリンクの範囲にある
- 「1対不特定多数」通信
 - アクティブなホストだけが受信する
 - IP層は connectionless 型通信なので再送機能を持たない
- データグラム型通信しか対応できない
 - UDP (トランスポート層)
 - 「1対不特定多数」でのconnection oriented 型通信は定義が難しい
 - 幾つかの研究は存在しているのは確かだが....
 - Broadcast enabled TCP

Broadcast (3)

- データリンク層がbroadcastをサポートしていないときには、同報通信はできない

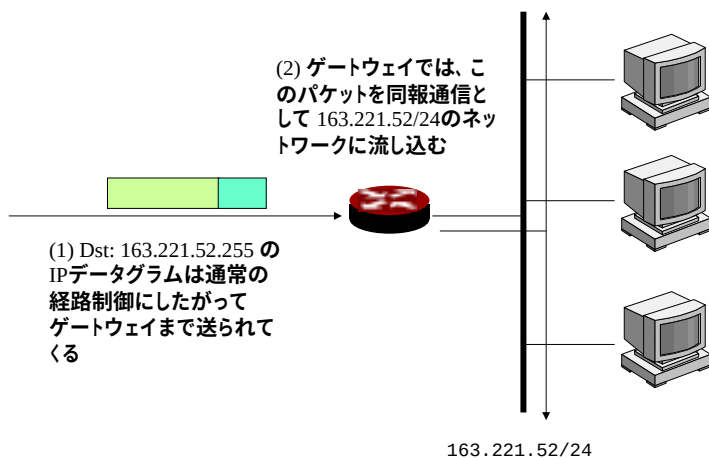
- ifconfig コマンドで確認できる

```
% ifconfig en0
en0: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
    inet 163.221.52.135 netmask 0xfffff000 broadcast 163.221.52.255
    ether 00:03:93:c9:31:52
    media: autoselect (100baseTX <half-duplex>) status active
```

Directed Broadcast

- 通常のIP broadcastはゲートウェイを越えない
 - 送信ホストが接続されているデータリンクの範囲で同報通信が実行される
 - 奈良先端大のキャンパスネットワーク 163.221/16 で、163.221.255.255 を指定したとしても、キャンパスネットワーク全体に同報通信されるような機能は持たない
- 他のネットワークから、宛先アドレスに IP broadcast addressが指定されているIPデータグラムを送った場合には何が起るのか？
 - Directed broadcast
 - 通常は、この機能は使えないようになっている
 - ところが、1990年代後半までは、「Directed broadcast 機能を使えるようにすることは良いこと」と多くの人が考えていた

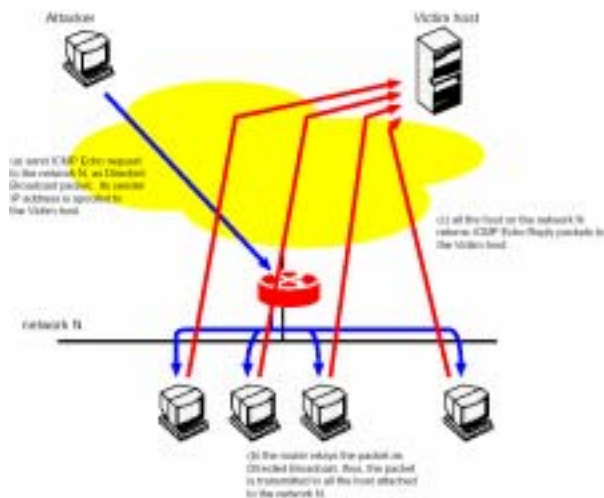
Directed Broadcast (2)



Smurf Attack

- ICMP ECHO request packet
 - Source addr: 133.1.2.3
 - Destination addr: 163.221.52.255
 - 163.221.52/24 のゲートウェイが directed broadcast 機能が有効になっていると.....
- 133.1.2.3 に対して、ICMP ECHO replay packet がネットワーク内のアクティブなホストから返される!
 - Input: 1 packet, output: multiple packets
 - (おお！増幅されている！)
 - Smurf amplifier

Smurf Attack (2)



Broadcastはコスト高

- IP broadcast は、同一データリンク内のアクティブなノード全てで受信され、**処理される**
 - データリンク層、IP層の処理が確実に駆動される
 - トランスポート層 (あるいはIP層の上位層) で、初めて受信したIPデータグラムが必要かどうかを判断して、受信するホストが無ければデータグラムを廃棄する
 - ブロードキャストによって、全てのアクティブなノードのOSが何らかの処理をしなければならない
 - Broadcast stormが発生するとネットワーク全体の性能が劇的に劣化する
 - Ethernet meltdown にも

通信モデル(2)

- Multicast

- グループ通信

- 多対多通信
 - グループは動的に形成される
 - グループを形成するための特別な手順が、IP層とは別に定義
 - 例えば、IGMP (Internet Group Management Protocol) を用いた動的なグループ設定のメカニズムなどがこれまで実装されてきた
 - 特別なIPアドレスを使う
 - 224.0.0.0 – 235.255.255.255
 - 特定のIPマルチキャストアドレスと、そこに参加しているノードのIPアドレスのリストが対応付けされる

Multicast

- Multicastはとても難しい

- L3ゲートウェイを越えてグループに属するノードにデータを配送する必要がある
 - あるグループに属するノードがどこに存在するのか?
 - どのようにして二重配信を防止するのか?
 - どのようにして配信を制御するのか?
 - 配送途中のゲートウェイの役割はなんであるのか?
 - インターネット全体での運用ポリシーは単一ではない
 - あるISPは IP multicast を許容するかもしれないが、別のISPでは限定された形での運用しかしていないかもしれない
 - IP multicast をサービス提供することは必須ではない
 - IP multicast を使える環境は (現時点では) 限られている

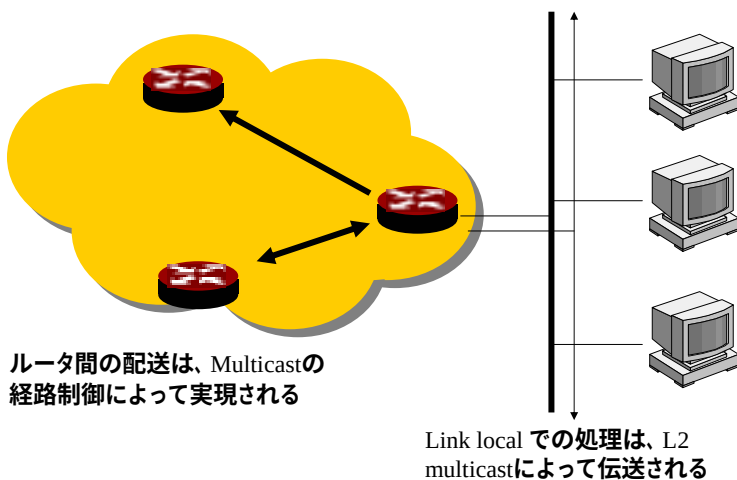
Multicast (2)

- Multicast を実現するために必要なパーツ
 - － グループ管理
 - 動的にメンバを管理するための機構
 - 特定のマルチキャストアドレスに対応付けられるホストグループの管理
 - 動的管理を行うために、これまでさまざまなプロトコルが開発されている
 - － 経路制御あるいは配信機構
 - “Mbone” と呼ばれる overlay network の構築
 - IP multicast に対応したゲートウェイ実装
 - Native IP multicast の運用

現在の運用

- AS内での運用が中心的
 - － AS: 同一の管理ポリシーに従うネットワーク群
 - 例えば、一つのISPが管理しているネットワーク
- L2とL3の連携
 - － L2の multicast 機能を積極的に利用
 - 一つの IP multicast group に対して、一つのL2 multicast address をマッピングする
 - IEEE 802 LAN に対するアドレスの作り方を標準化
 - － IP multicast address を埋め込んだ形で利用
 - － L3は経路制御によってカバー

IP multicast における L2 / L3 の役割分担



Address mapping (1)

- IP multicast address $\Rightarrow$ IEEE802 LAN multicast addr
 - IETF OUI: 0x00-00-5E
 - Organizationally-Unique Identifier
 - M: Unitcast/Multicast bit
 - L: Universal/Local bit
 - Manuf.: 各ベンダによって一意性が保証されている割り当て

	8	7	6	5	4	3	2	1
Byte 1	O						L	M
Byte 2	U							
Byte 3	I							
Byte 4								
Byte 5	Manuf.							
Byte 6								

Address Mapping (2)

- 上位24ビットは、IEEE OUI を
使い、Multicast bit をセット
 - L = 0, universal
 - M = 1, multicast
- 4オクテットの最上位ビットは0
とする
- 残り23ビットを、IP multicast
address から取り出して埋める
 - IP multicast address の下位
23ビットを使用する

	8	7	6	5	4	3	2	1
Byte 1	0	0	0	0	0	0	L	M
Byte 2	0	0	0	0	0	0	0	0
Byte 3	0	1	0	1	1	1	1	0
Byte 4	0							
Byte 5								
Byte 6								

L=0, M=1

Address Mapping (3)

Decimal:	239	147	6	99
Hex:	EF	93	06	63
Binary:	1110-1111	1001-0011	0000-0110	0110-0011

0000-0001	0000-0000	0101-1100	0	001-0011	0000-01110	-110-0011
-----------	-----------	-----------	---	----------	------------	-----------

IEEE802 LAN multicast: 0x01-00-05-13-06-63

IGMP

- 直近のルータに対して、IP multicast Groupに入っていることを宣言するためのプロトコルとして使用
 - L2 でのマルチキャスト配信をルータに依頼
 - IGMP join メッセージ
 - Join メッセージを受け取ったルータは、L3 でのマルチキャスト経路制御機構を使って、メッセージの配信を受ける

L3 multicast 制御

- 経路制御プロトコルに依存
 - 最近、PIM (Protocol Independent Multicast) を用いて配信をすることが多い
 - PIM-SM での運用が経験的には多いか？
 - その他の経路制御プロトコル
 - DVMRP (Distance Vector Multicast Routing Protocol)
 - Originally specified in RFC1075
 - MOSPF
 - Originally specified in RFC1584
- 効率の良い flooding を行うことを目標
 - 無駄な配送を避け、帯域を無駄遣いすることを避ける

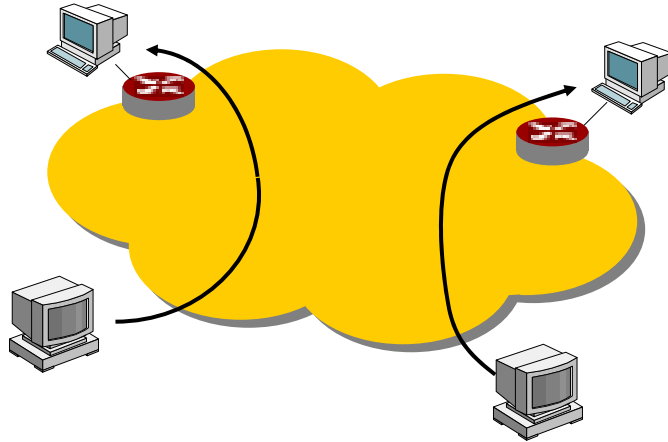
最近の傾向

- IP multicast は、もともと多対多通信として設計
 - 真の意味でグループ通信をするアプリケーションは少ない
 - 多くのマルチキャストは、1対多通信 (放送) を使うことが多い
 - 現在の IP multicast の設計 (ASM: Any Source Multicast) は处理的にも大変なので、SSM (Sender Specific Multicast) に特化した経路制御プロトコルが登場し始めている
 - PIM-SSM
 - Xcast

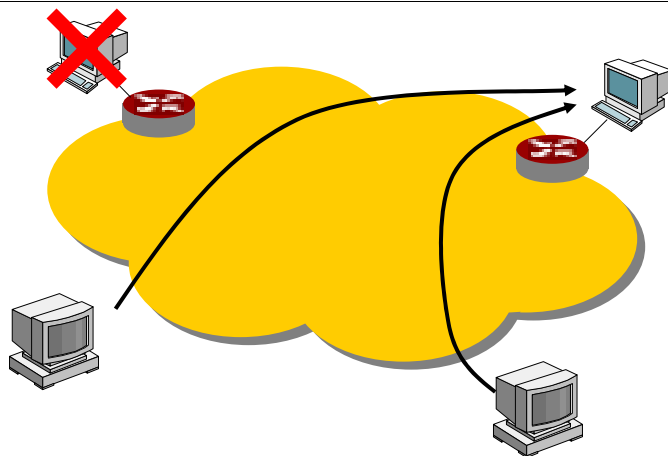
通信モデル (3)

- Anycast
 - インターネット内に、同一のIPアドレスを持つホストを複数設置し、経路制御的にも正しく通信できるようにする
 - 効能
 - 経路制御的に近いホストに到達して通信を行う
 - 一つのホストがダウンしても、経路制御機構により他のホストに対して通信が迂回される
 - ただしTCP (connection oriented L4) や、状態を持つアプリケーションの場合には、上位層でつじつま合わせに対応する必要がある
 - ある意味で Unicast の意味的な拡張

Anycast



Anycast



Anycast

- 現状で使われているサービスは限られている
 - F root DNS server
 - DNSのトップレベルのサーバ
 - 世界に13台しかない
 - www.root-servers.net
 - Operated by ISC
 - Geographical distribution using Anycast
 - Madrid
 - NYC
 - Hong Kong
 - Korea
 - サーバ側で状態を保持しない stateless service であれば使える可能性は高い