Routing for Integrated Services

New routing requirements

- Multiparty communication:
  - conferencing (audio, video, whiteboard)
  - remote teaching
  - multi-user games
  - networked entertainment – “live broadcasts”
  - (distributed simulations)
  - (software distribution)
  - (news distribution)
- Support for QoS in routing

Questions

- How can we support multiparty communication?
- How can we provide QoS support in routing?

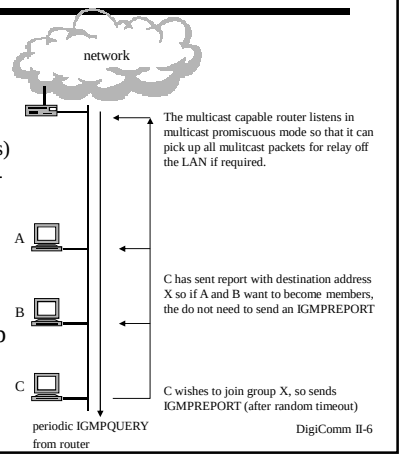
Many-to-many communication: IP multicast

Group communication using IP

- Many-to-many:
  - many senders and receivers
  - **host group** or **multicast group**
- One transmission, many receivers
- Optimise transmissions:
  - e.g. reduce duplication
- Class D IP address:
  - 224.0.0.0 - 239.255.255.255
  - **not** a single host interface
  - some addresses reserved
- Applications:
  - conferencing
  - software update/distribution
  - news distribution
  - mutli-player games
  - distributed simulations
- Network support:
  - LAN
  - WAN (Internet routers)
  - scoped transmission: IP TTL header field

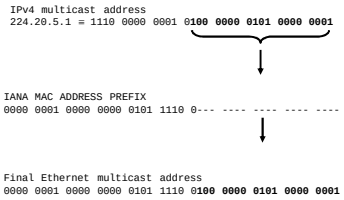
IP multicast and IGMP

- Features of IP multicast:
  - group of hosts
  - Class D address
  - leaf nodes (hosts) and intermediate nodes (routers)
  - dynamic membership, leaf-initiated join
  - non-group member can send to group
  - multicast capable routers
  - local delivery mechanism
- IGMP: group membership control

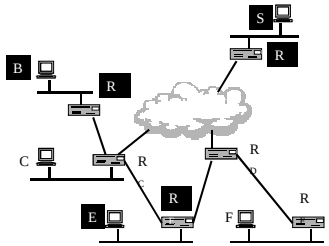


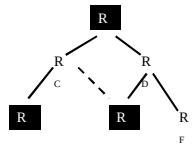
Multicast: LAN

- Need to translate to MAC address
- Algorithmic resolution:
  - quick, easy, distributed
- MAC address format:
  - IANA MAC address allocation
  - last 23-bits of Class D
  - not 1-1 mapping
- Host filtering required at IP layer



Multicast routing [1]

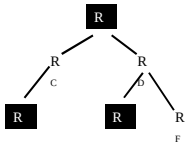




- First refinement
  - **reverse path broadcast (RPB)**
  - duplication

- Starting point: **flood**
  - creates looping

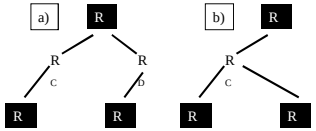
Multicast routing [2]



- Second refinement
  - eliminate duplicates
  - need routing information

- Distance vector:
  - need next hop information
  - (or use **poisoned reverse**)
- Link state:
  - construction of all SP trees for all nodes possible
  - "tie-break" rules required

Multicast routing [3]



- Third refinement:
  - **pruning**
  - need to refresh tree – **soft-state**
  - **reverse path multicasting (RPM)**
- RPM:
  - used in many multicast protocols
  - per-sender, per-group state

- Networks with no group members pruned from tree
- Must somehow allow tree to re-grow
- Soft-state:
  - timeout – re-flood
  - downstream nodes prune again
- Explicit **graft**:
  - downstream nodes join tree

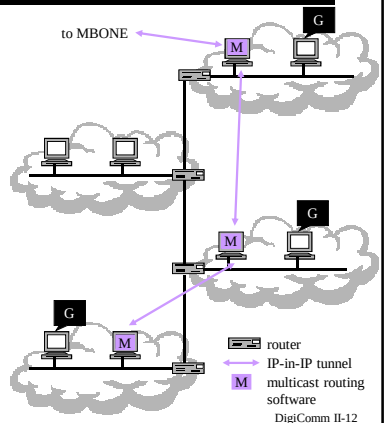
DVMRP and the MBONE

- DVMRP:
  - RPM
  - used on MBONE
- MBONE:
  - virtual overlay network
  - distance vector routing

MBONE Visualisation Tools
http://www.caida.org/Tools/Manta/
http://www.caida.org/Tools/Otter/Mbone/

MBONE configuration

- Routers not multicast aware:
  - use virtual network
- Multicast islands:
  - connected by virtual links
  - can not use normal routing info – use multicast hops
- IP tunnelling:
  - software runs on a host
  - *ad hoc* topology
- Use TTL for scope:
  - TTL expiry: **silent discard**
  - **administrative scope** possible



MOSPF

- Link-state algorithm
- RPM
- Intended for larger networks
- Soft-state:
  - router advertisement sent on group join
  - tree evaluated as routing update for a group arrives
- Still suffers from scaling problems:
  - a lot of state-required at each router
  - per-group, per-link information required

CBT

- Core router(s):
  - core distribution point for group
- Leaf sends IGMP request
- Local router sends *join request* to core
- *Join request* routed to core via normal unicast
- ✓ Intermediate routers note only incoming i/f and outgoing i/f per group
- ✓ Explicit join and leave:
  - no pruning
  - no flooding
- ✗ Distribution tree may be sub-optimal
- ✗ Core is bottleneck and single-point-of-failure:
  - additional core maybe possible
- Careful core placement required

PIM

- PIM:
  - can use any unicast routing protocol info
  - two modes: **dense mode** and **sparse mode**
- Dense mode:
  - RPM
  - flood-and-prune with explicit join
- Sparse mode:
  - similar to CBT
  - core (rendezvous point) or shortest-path possible
  - rendezvous point sends keep-alive
  - explicit graft to tree

Multicast address management

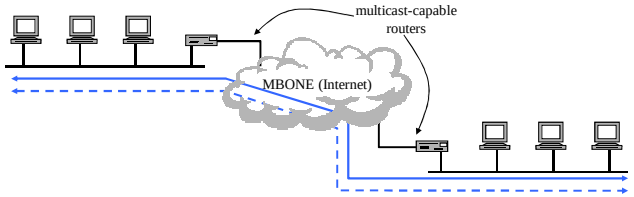
- Some addresses are reserved:
  - 224.0.0.1 all systems on this sub-net
    224.0.0.2 all routers on this sub-net
    224.0.0.4 all DVMRP routers
    (plus many others)
- No central control as in unicast addresses
- Others generated pseudo-randomly:
  - 28-bit multicast ID (last 28 bits of Class D address)

Multimedia conferencing [1]

- Multimedia applications:
  - voice - *RAT*
  - video - *VIC*
  - text - *NTE*
  - whiteboard - *WBD*
- Support:
  - session directory - *SDR*
  - gateway - *UTG*
- All use **IP multicast**:
  - local – direct
  - wide area – MBONE
- RTP/RTCP
- IP multicast:
  - 224.2.0.0 - 224.2.255.255
  - different address per application per session
- Scoping:
  - IP TTL header field:
    16 local (site)
    47 UK
    63 Europe
    127 world
  - administrative

Multimedia conferencing [2]

- Two multicast channels per application per session:
  - RTCP and RTCP
- Stand-alone - *ad hoc*:
  - individual applications
- Advertised conference:
  - SDR
  - configuration information

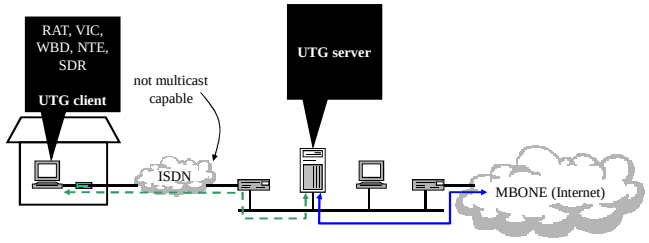


Multimedia conferencing [3]

- Inter-flow synchronisation:
  - e.g. audio-video (lip-synch)
  - RTP/RCTP time-stamps
  - e.g. *RAT+VIC*: synch to *RAT* flow
- Inter-application communication:
  - conference bus
  - local communication (e.g. pipes)
- Heterogeneity:
  - data rates
  - (QoS)
- Gateway:
  - **transcoding**
  - multicast-to-unicast
  - supports dial-up users via BR-ISDN
  - (similar to H.323 Gatekeeper)

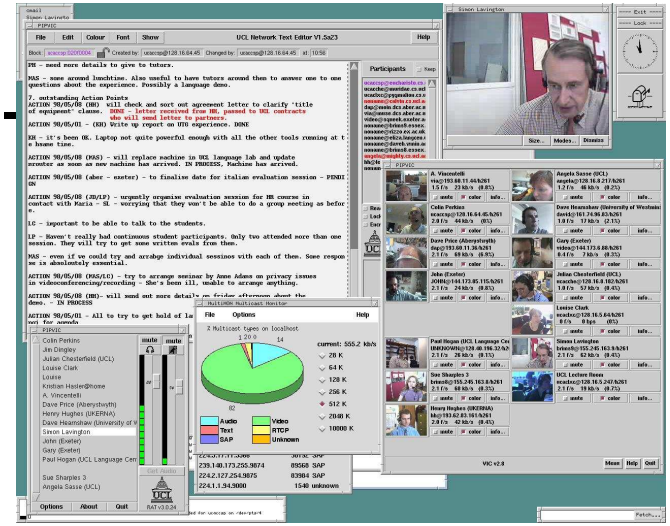
Multimedia conferencing [4]

- *UTG* server:
  - performs transcoding and relay
  - *UTG* clients register with server
- Dial-up users:
  - unicast to UTG client
  - local multicast at remote (client) host



Multimedia conferencing [5]

- *RAT*:
  - packet audio: time-slices
  - numerous audio coding schemes
  - redundant audio for repair
  - unicast or multicast
  - data-rate configurable
- *VIC*:
  - packet-video: frames
  - numerous video coding schemes
  - unicast or multicast
  - data-rate configurable



Multicast conferencing [7]

- Floor control:
  - who speaks?
  - chairman control?
  - distributed control?
- Loose control:
  - one person speaks, grabs channel
- Strict control:
  - application specific, e.g.: lecture
- Resource reservation:
  - not supported on the MBONE(!)
  - ~500Kb/s per conference (using video)
- Per-flow reservation:
  - audio only
  - video only
  - audio and video

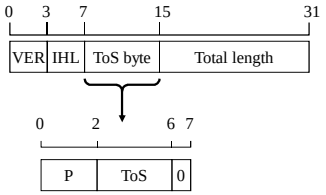
QoS-based routing

What is QoS-based routing?

- Traditional routing:
  - destination address chooses path/route
  - routers have one “optimal” path to destination
  - routing metrics are single values
- QoS routing:
  - multiple paths possible
  - alternative paths have different QoS properties
  - routing updates include QoS parameter information
  - use destination address, source address, ToS, etc.
- RSVP/INTSERV/DIFFSERV:
  - signalling may still be required

IPv4 ToS byte

- IPv4 header – ToS byte:
  - 3-bit precedence, P
  - 4-bit ToS
- Precedence:
  - 000: lowest
  - 111: highest
- ToS – flags:
  - 1xxx: minimise delay
  - x1xx: maximise throughput
  - xx1x: maximise reliability
  - xxx1: minimise cost (£)
  - 0000: “normal” service



- Not widely used:
  - no global agreement
  - (some use in Intranets)
- RFC1349 – now historic:
  - superseded by DIFFSERV
  - not compatible with ECN

Multi-metric routing

- Use multiple metrics:
  - minimum delay path
  - maximum throughput path
  - maximum reliability path
  - minimum cost path
- Example – OSPF:
  - QoS parameters passed in link-state packets
  - ToS byte used in IPv4
  - multiple executions of shortest-path algorithm
- Sequential filtering:
  - filter paths using metrics
- Granularity of QoS:
  - can be per-flow, but requires much state in routers
- Router overhead:
  - more per packet processing
  - larger router updates
  - more state at routers
  - possibility of instability during routing updates

Route pinning and path pinning

- Dynamic routing:
  - path change → QoS change
- Keep route fixed for flow?

Route pinning

- Ensure that route is fixed while packet forwarding in progress
- Disrupts normal routing behaviour
- May cause congestion conditions

Path pinning

- Allow route to change:
  - existing flows remain on fixed path
  - new flows use new route
- Allow different paths for different flows:
  - pin separate flows to separate paths
- Inconsistency:
  - could affect stability if flow is long lived
- (Use of RSVP?)

MPLS

- Multi-protocol label switching:
  - fast forwarding
  - IETF WG
- MPLS is an enabling technology:
  - claimed to help scaling
  - claimed to increase performance
  - forwarding still distinct from routing
- Intended for use on NBMA networks:
  - e.g. ATM, frame-relay
- Many supporters:
  - e.g. Cisco
- Many cynics:
  - introduces much more complexity into routers
  - more state required at routers
  - (non)-interaction with routing protocol operation may cause instability
  - may not work very well at high speeds
  - other IP-level mechanisms exist

Intra-domain routing

- Can use agreed single/multiple metrics
- Allow autonomy in domains to remain
- Should indicate disruptions to QoS along a path
- Must accommodate best-effort traffic:
  - no modification to existing, best-effort applications
- Optionally support multicast:
  - allow receiver heterogeneity and shared reservations
- Still a research issue

Inter-domain

- **Must be scaleable**
- QoS-routing should not be highly dynamic:
  - few router updates, relatively small amounts of information
  - may have to rely on traffic engineering and capacity planning
- Must not constrain intra-domain routing mechanisms
- Allow QoS information aggregation
- Optionally support multicast

QoS-based routing for multicast

- Reliable multicast:
  - retransmissions from sender does not scale
  - research issue
- QoS for multicast:
  - need to support widely/sparsely dispersed groups
  - dynamic membership changes
  - must scale across domains (across AS boundaries)
  - should allow heterogeneity in group
  - support for shared reservations
  - research issue

Summary

- Many-to-many communication:
  - IP multicast
  - DVMRP, MOSPF, CBT, PIM
  - conferencing example
- QoS-based routing:
  - multi-metric
  - route/path pinning
  - intra-domain and inter-domain
  - QoS-based routing for multicast