Hierarchical Distance-Vector Multicast Routing for MBone

Presented
by
Nitin Deshpande
Darpan Bhuva

Outline

- Introduction
- Current MBone Scenario
- Hierarchical DVMRP
- Protocol Evaluation
- Other Issues
- Conclusion

Introduction

What is MBone?

- MBone is Multicast Internet Backbone
- Established in 1992 with 40 subnets in 4 countries
- Interconnected set of routers and subnets that provide IP multicast delivery in Internet

Distance Vector Multicast Routing Protocol (DVMRP)

- MBone routers run a protocol to decide where to forward IP multicast packets
- Routers treat MBone topology as a single flat routing domain
- Entry for every subnet in the MBone
- Problem of additional processing resources and memory
- If nothing is done the MBone will collapse

- Solution lies in using Hierarchical Distance-Vector Multicast Routing for the MBone
- Use two-level hierarchy in which the MBone is divided into regions and the regions contain subnets
- The routing protocol in each region maintains topological information only for its own region, not for other regions

- The intra-region multicast routing may be accomplished by any number protocols, including DVMRP
- Inter-region protocol maintains information only about interconnection of regions and not about any internal topologies
- Inter-region routing protocol uses a modified version of DVMRP that computes multicast routes among regions rather than among subnets

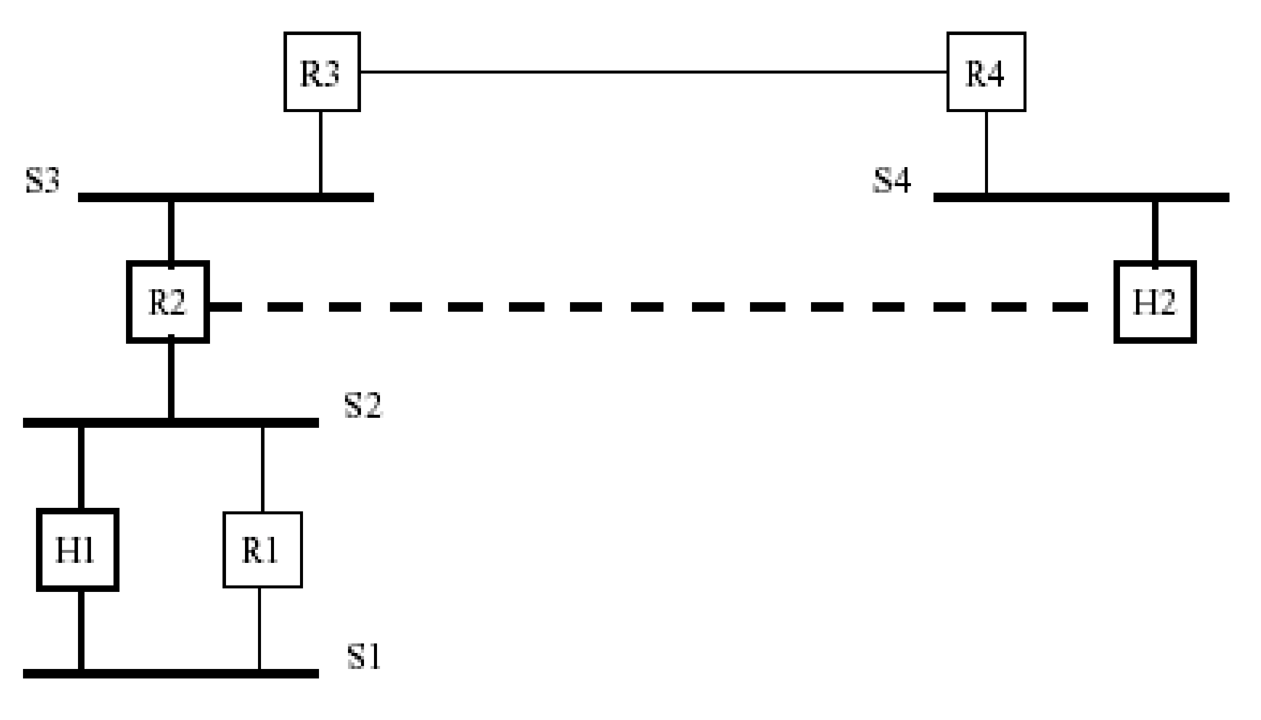
Advantages of using Hierarchical Routing

- Partitioning of regions allows different multicast routing protocols to be used in different regions
- Topological changes such as link or router failures are isolated to that particular affected region
- Limit on the maximum diameter of topology imposed by some protocols can be relaxed

Current MBone Structure and Routing.

- Very few of the Internet routers support multicast routing
- Most of the MBone routing is done by general purpose routers which run multicast routing software

MBone Components.



Description of Hierarchical DVMRP.

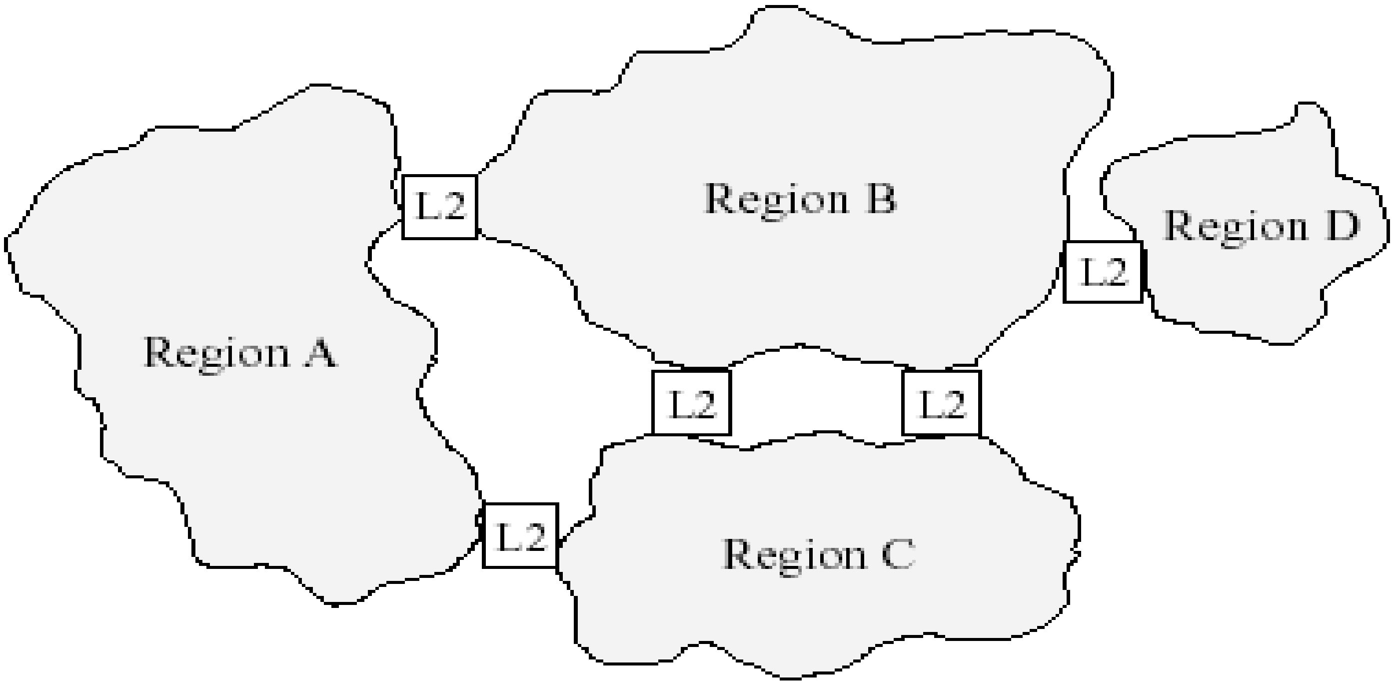
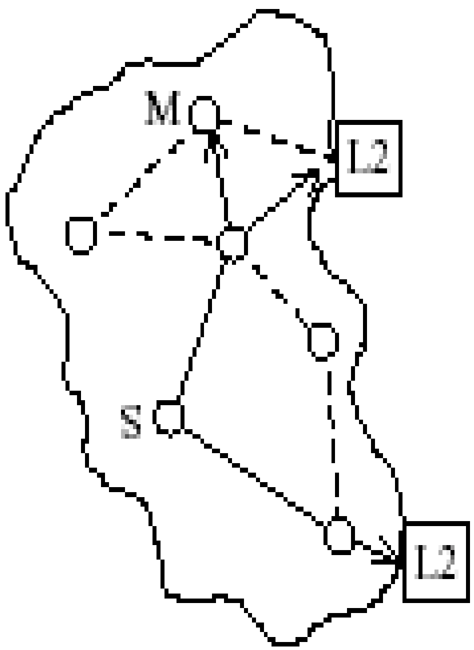


Fig: Regions interconnected by L2 routers

- MBone is divided into multiple, non-overlapping regions, each region being assigned a unique region identifier
- The multicast routers internal to a region run a L1 multicast protocol for forwarding traffic within the region and the boundary routers run a L2 protocol for forwarding inter-region traffic
- The boundary L2 routers also include L1 functionality

Phase 1.Routing in Originating Region.

- A multicast packet that originates within a region has to be forwarded to all the routers in that region which are the members of destination multicast group
- The packets are also forwarded to all the boundary routers attached to that region
- The DVMRP uses the “Broadcast and Prune” method



S	G	data

S - Source of the multicast datagram

G - Multicast group destination

Fig.5(a)

Fig.5(b)

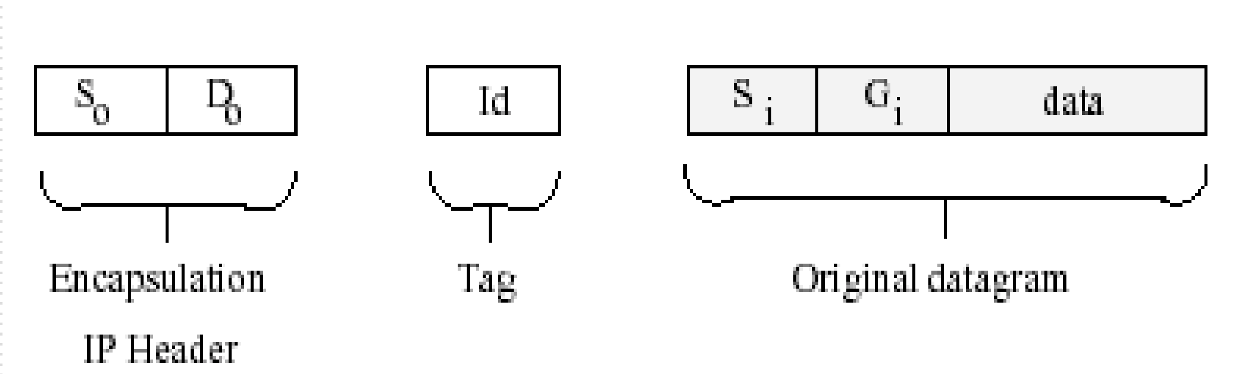
Phase 2. Routing Between Regions.

- Each region is given a unique identifier called *Region-Id*
- Each region is logical equivalent of a subnet and the region Id's are equivalent to subnet addresses
- Each L2 router also acquires information on all IP subnet addresses of the region to which it is directly attached
- Group L2 routers of directly attached regions called All_Boundary_Routers(ABR)

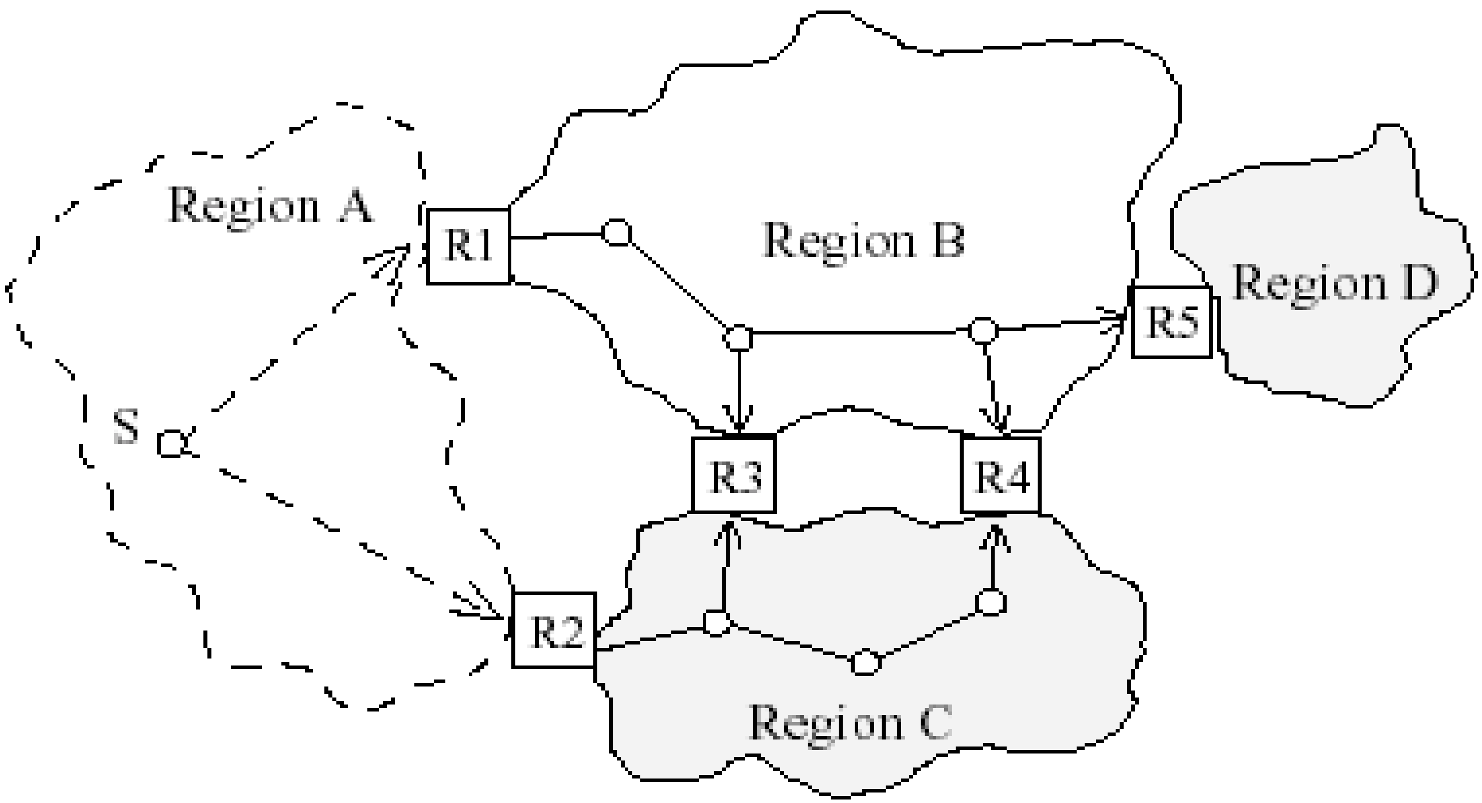
- Upon the arrival of a multicast packet from one from its connected regions, an L2 router performs the following operations :

  a) It checks if the source of the packet is one of the subnets in the region from which the packet arrived. If the check fails the packet is ignored

  b) The packet is tagged with the region Id representing the region from which the packet originated c)



If the L2 router decides to forward the packet then it encapsulates the packet as shown in the Figure above and sends it. The inner IP header contains the original source address (Si) and the destination group address (Gi). The Tag field contains the originating Region-Id. The outer encapsulation header contains the address of the sending boundary router (So) and the destination ABR group address (Go)



Routing of packets between regions

- On receiving an encapsulated packet of this type an L2 router performs the following operations :

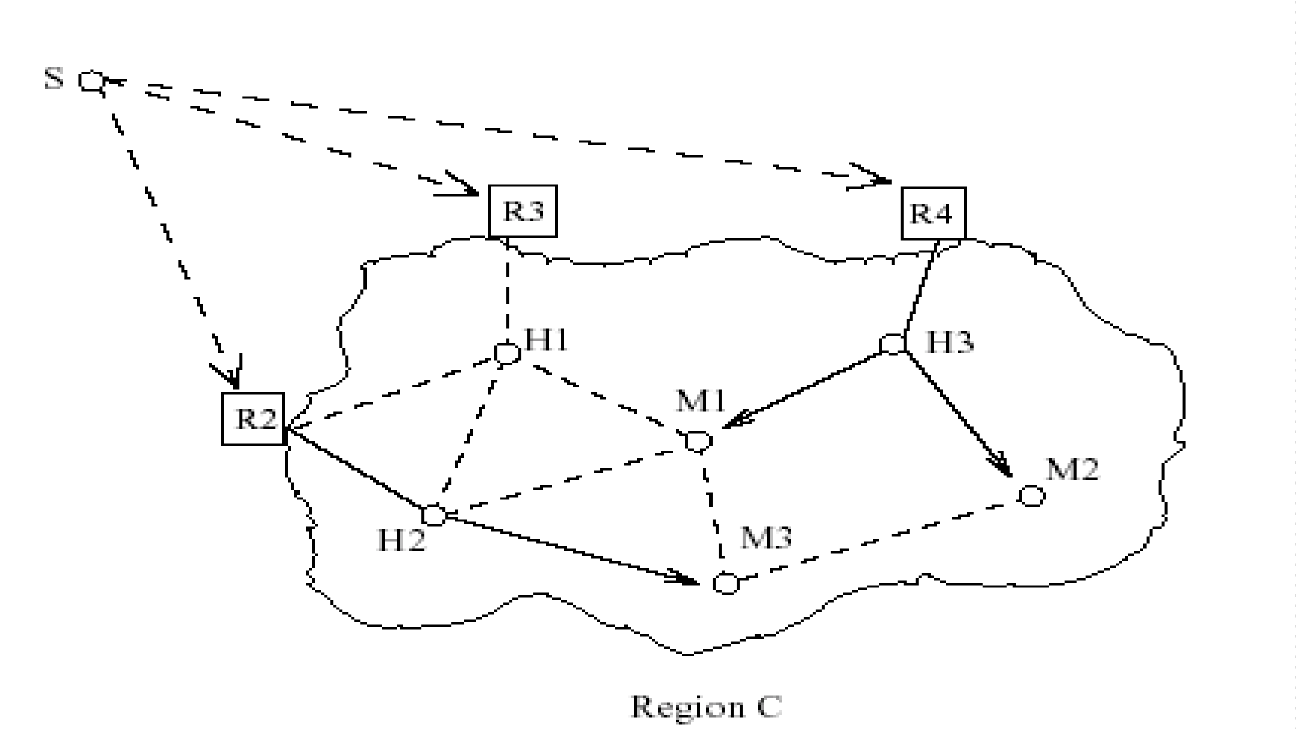
  a) The encapsulation is stripped off and a check is performed to see if the packet is arrived via the shortest path. If the check fails the packet is ignored

  b) The downstream regions, to which the packets are to be forwarded, are determined by using the region-Id and the group address of the original packet

- If any of the connected regions have members of the destination multicast group, a copy of the original packet is injected into the region
- The membership information within a region is obtained by having the L1 routers within the region send *Region Membership Reports* (RMRs) to all of the region's boundary routers via the ABR group. Also whenever the group membership status changes the L1 routers send an RMR message to the boundary routers

Phase 3.Routing in Destination Regions.

- Each L2 router advertises a default route with certain pre-configured metrics into each region
- It uses this default route when there is no entry for the source subnet
- Each L1 router upon receiving these default routing updates, determines the interface corresponding to the nearest L2 router using the default metric advertised by each L2 router



Routing packets in the destination Region C.

Protocol Evaluation

- This approach reduces overhead but at the same time has some drawbacks
- Packet Duplication over Internal links
- Effect on Fair Queuing and Resource Reservation Algorithms

Packet Duplication over Internal Links

- Problem of packet arriving at the destination over encapsulated as well as original form
- Delivery path R4-H3-M1 for external (decapsulated) packets to reach M1
- Packet R4 receiving encapsulated packets from R2 via path R2-H1-M1-H3-R4

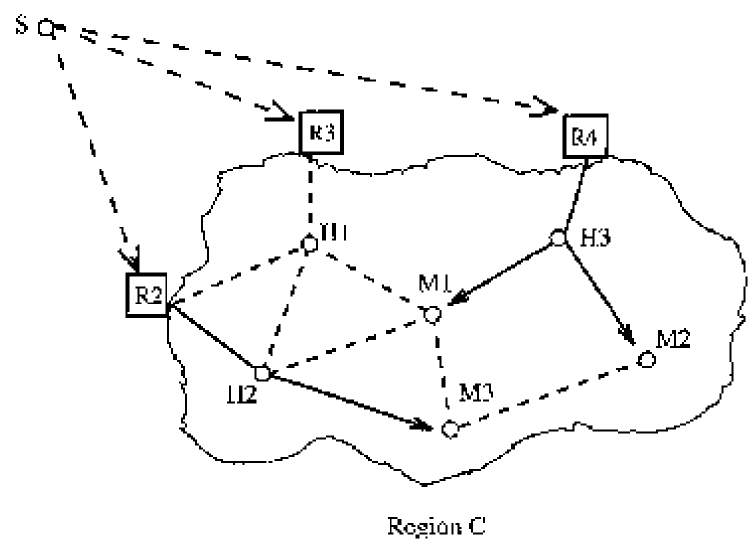


Figure 8: Routing packets in the destination Region C.

Contd.

- Route H3-M1 traversed twice by the same packet

Overcome by

- Router decapsulation

  L1 routers to act like “pseudo L2” routers, by checking if there are any group members corresponding to each packets multicast address in any of the subnets

Contd

- Host decapsulation

Group packets could be forwarded like regular multicast packets i.e. L1 router stripping the encapsulation and processing it instead of an L2 router

Effect on Fair Queuing & Resource Reservation Algorithms

- Fair Queuing and other Resource reservation algorithms associate source and destination address of the packet to bandwidth and other constraints
- Packets forwarded at Level 2 have encapsulation headers having source and destination addresses that are different from the original one there by causing error in decision making process

Other Issues

Enabling Multiple level Hierarchy

By clustering various regions into *super-regions* interconnected by L3 routers and assigning identifiers to them. Reduces table sizes at the router but increases overhead as more encapsulation is required.

A *“CIDR-type”* approach in issuing Region-Id’s

This approach is more beneficial as no additional overhead is required.

Other Issues
Level 1-Level2 Interoperability

- DVMRP described here as the Level 2 protocol and any number of Level 1 protocol
- L1 routers generate RMR's to enable L2 routers to determine presence or absence of group members
- Protocols such as PIM and CBT do not maintain explicit routing information of their own, so they have to be *fixed* at each interface to the L2 routers to provide *filtered route* updates which only contain addresses and prefixes that are local to the region

Other Issues
Configuration Parameters

- Need to configure two metrics
  - Region-Id.
  - Default route advertised by each L2 router.
- Possible to craft the default metrics such that a L1 router can forward packets from a particular L2 router only
- Desirable only if known that most of the traffic is arriving from which router

Other Issues
Avoiding Multiple levels of Encapsulation

- Since multicast not yet fully developed in the Internet there exist tunnels
- Need for two encapsulations, one for tunnel and the other for Phase 2 explained previously
- Problems of significant overhead
- Avoided by using the destination address of the tunnel instead of ABR

Other Issues
Boundary through links vs. routers

- Assumed so far that regional boundary fall across boundary routers

  What if regional boundary falls between two boundary routers?
- Can be achieved by treating a “*boundary link*” as a separate, degenerate region
- If there is no multicast in the *boundary link* region it need not have a Region-Id and its also its presence does to add to the size of L2 routing table

Other Issues
Boundary Leakage Issues

- Misconfigured boundary routers and links across regions can lead to multiple regions being accidentally configured as one
- If a routing message sent by an L2 router into one region appears into another region it signifies a misconfigured router. Also called boundary router causing traffic to leak through *backdoors*
- On detection of such a condition forwarding of packets by the router should be stopped to avoid looping

Conclusions

- Use of address independent region identifiers enables significant reduction in size of routing tables
- Deployment of such a strategy will reduce topological volatility that a router must handle and lax the current constraints on maximum diameter for the MBone
- Cannot be applied to the problem of unicast route scaling