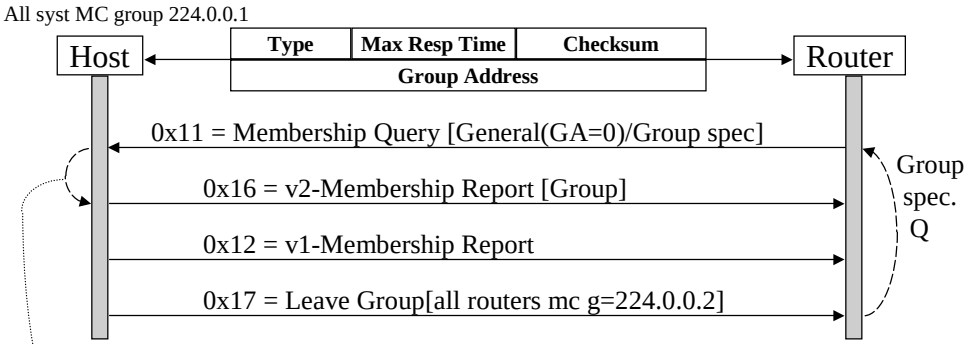
Multicast Protocols

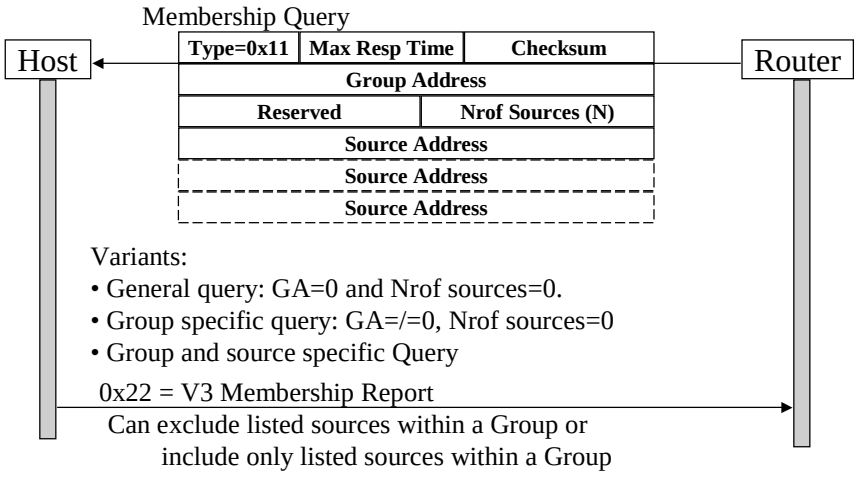
IGMP - IP Group Membership Protocol
DVMRP - DV Multicast Routing Protocol
MOSPF - Multicast OSPF
(see notes pages for some slides!)

IGMPv2 - Internet Group Management Protocol implements Group Membership



- IGMP runs directly on IP as protocol nr 2.
- TTL == 1 in all IGMP msges
- Host will wait random[0...Max Resp Time] prior to response and will supress its response if it sees another response to the same group

IGMPv3 adds selective reception from sources within a Group



Variants:
- General query: GA=0 and Nrof sources=0.
- Group specific query: GA=/=0, Nrof sources=0
- Group and source specific Query

0x22 = V3 Membership Report

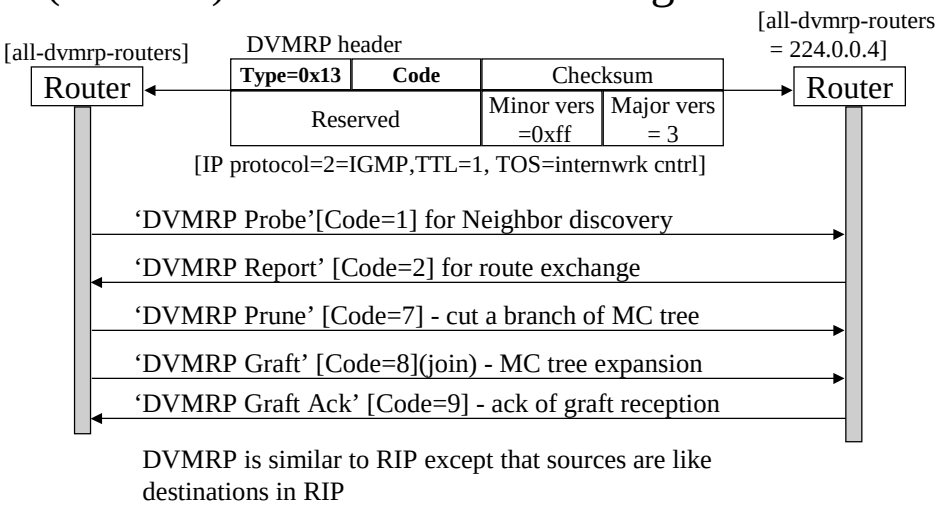
Can exclude listed sources within a Group or
include only listed sources within a Group

Experimental routing protocols have been developed for MBone - an overlay MC Internet

Shared tree	Source based trees	
	Bcast and Prune	Domainwide reports
PIM Sparse* Core Based tree*	DVMRP PIM Dense*	MOSPF

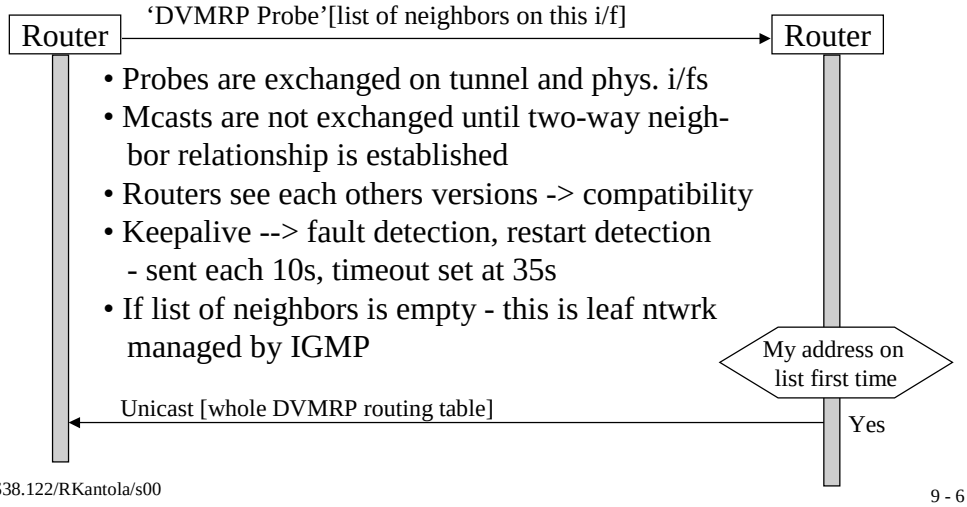
* Relies on Unicast routing protocol to locate MC-sources
Those that don't, can route MC on routes separate from unicast routes.

Distance Vector Multicast Routing Protocol (DVMRP) is used for MC routing in the MBone



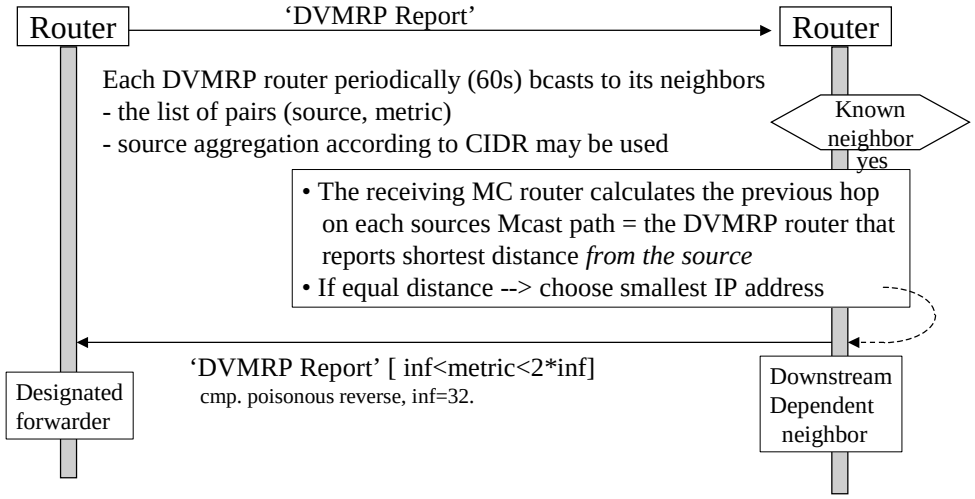
DVMRP is similar to RIP except that sources are like destinations in RIP

Probes are used for neighbor discovery

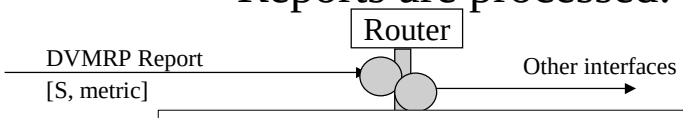


- Probes are exchanged on tunnel and phys. i/fs
- Mcasts are not exchanged until two-way neighbor relationship is established
- Routers see each others versions -> compatibility
- Keepalive --> fault detection, restart detection - sent each 10s, timeout set at 35s
- If list of neighbors is empty - this is leaf ntwrk managed by IGMP

Route reports are used to build the source based trees



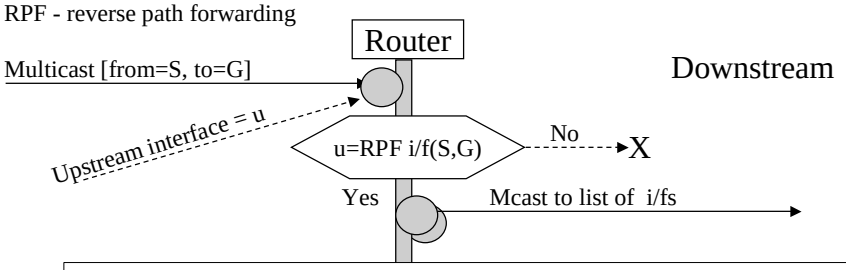
Reports are processed:




```
Adjusted metric=metric+interface cost
If  Metric<inf & Adjust metric≥inf
    Set adjusted metric to inf
If Route is new and Adj metric<inf
   Add route to RT
   Delete prune state of more general route
Elseif Route exists
   If Received metric < inf
      Check if Designated forwarder status for (S,G) changes
      If Adjusted metric > existing metric
         From same neighbor: update metric, Sch flash update for route
      Elseif Adj.metric < existing metric
         Update metric for the route
         If sender was different, update RT, schedule flash updates
      Elseif Adj.metric = existing metric: refresh route ...
   Elseif Received metric =inf ...
   Elseif Inf < Received Metric < 2 * inf ...
```

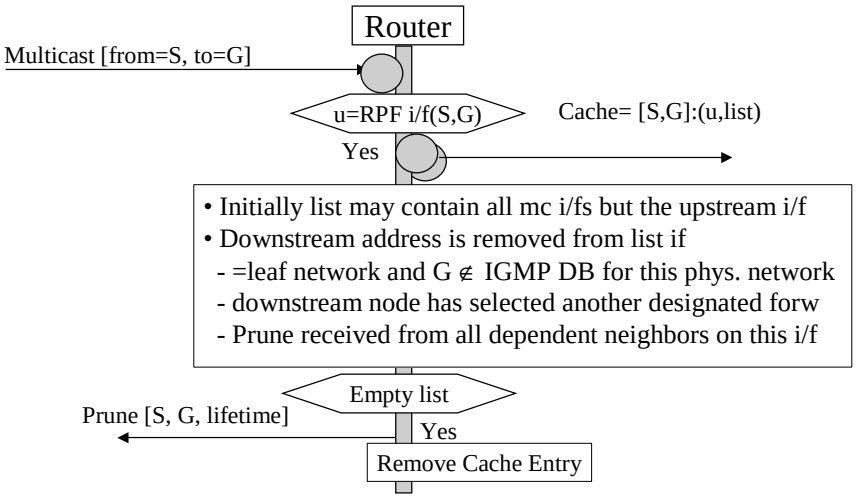
Multicast algorithm

RPF - reverse path forwarding



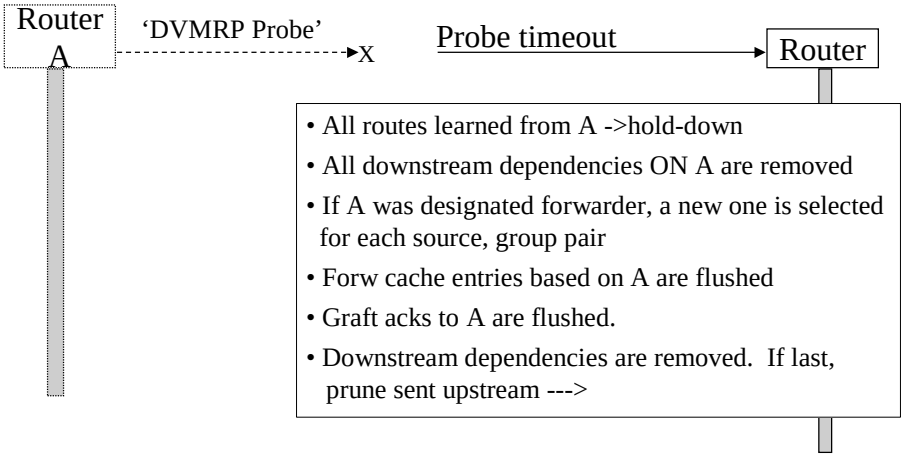
- At first mcast from RPF i/f a Forwarding Cache Entry [S,G]:(u,list...) is created using the DVMRP routing table
- List contains all downstream routers that have reported dependency on S
- Router is designated forwarder for downstream nodes
- If Designated forwarder becomes unreachable, Router assumes role of designated until it hears from a better candidate

List of dependent neighbors is used to minimise the MC tree



- Initially list may contain all mc i/fs but the upstream i/f
- Downstream address is removed from list if
  - =leaf network and G ∉ IGMP DB for this phys. network
  - downstream node has selected another designated forw
  - Prune received from all dependent neighbors on this i/f

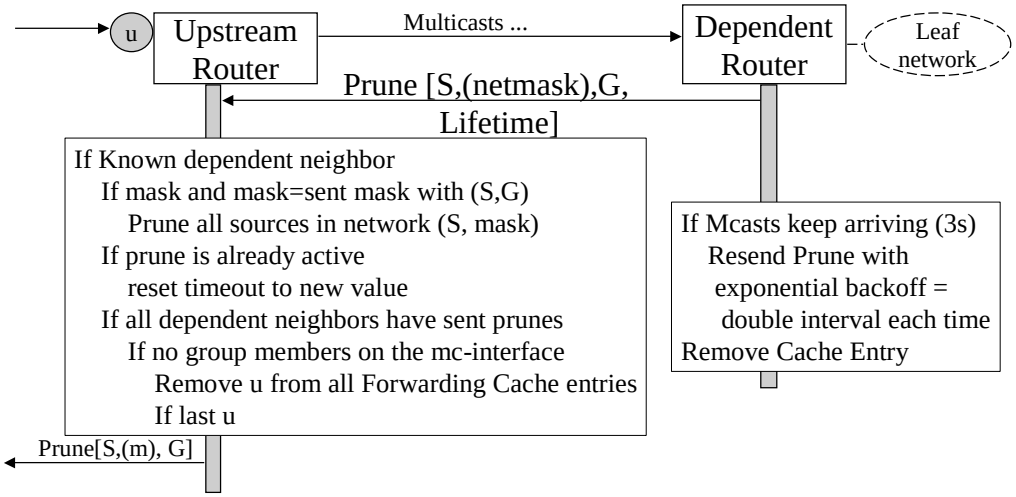
On Probe timeout Caches are flushed



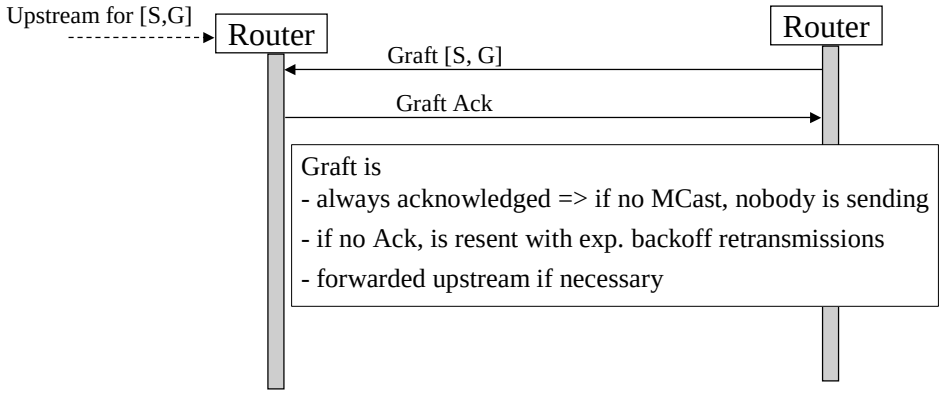
Route hold-down is a state prior to deleting the route

- Routes expire on Report timeout or when an infinite metric is received
- An alternate route (that in RIP caused temporary loops) may exist
- Routers continue to advertise the Route with inf metric for 2 report intervals - this is the hold-down period
- All Forw Cache entries for the Route are flushed
- During hold-down, the route may be taken back, if (<inf and = SAME) metric is received from SAME router

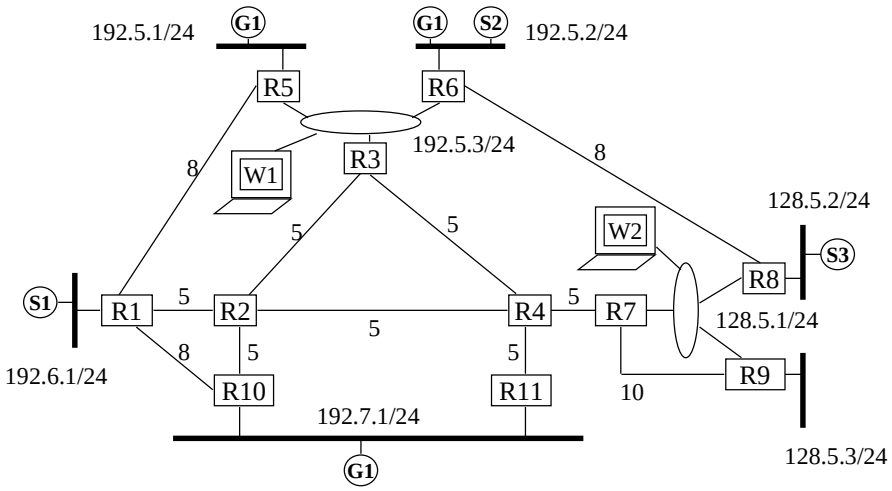
Prunes minimise the Mcast tree



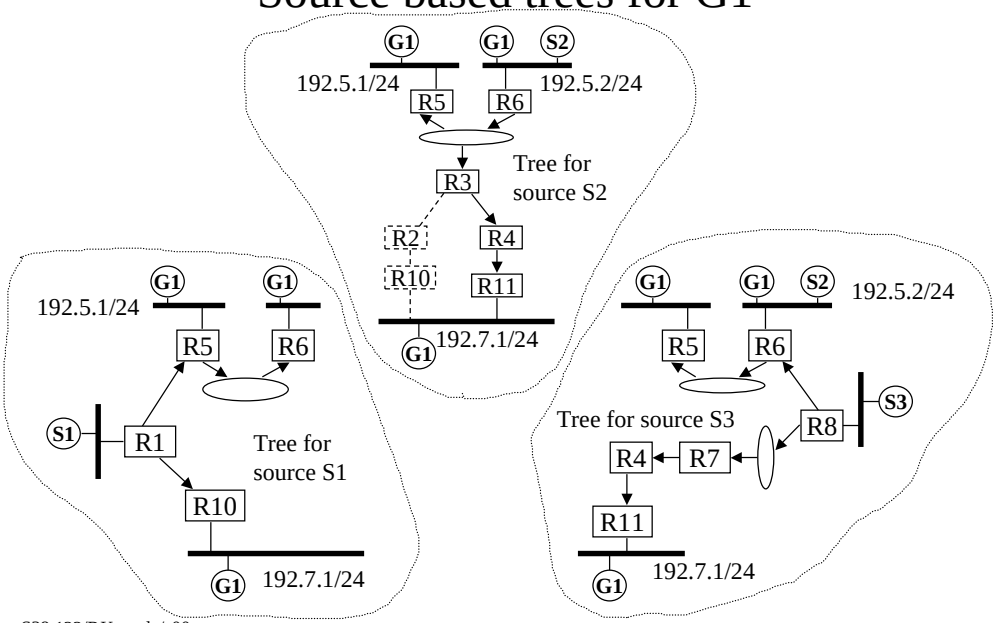
Grafts are used to grow the tree when a new member joins the Group



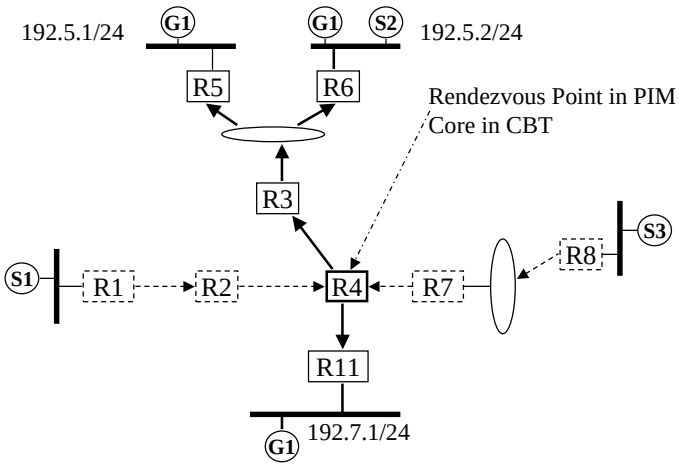
Multicast routing example



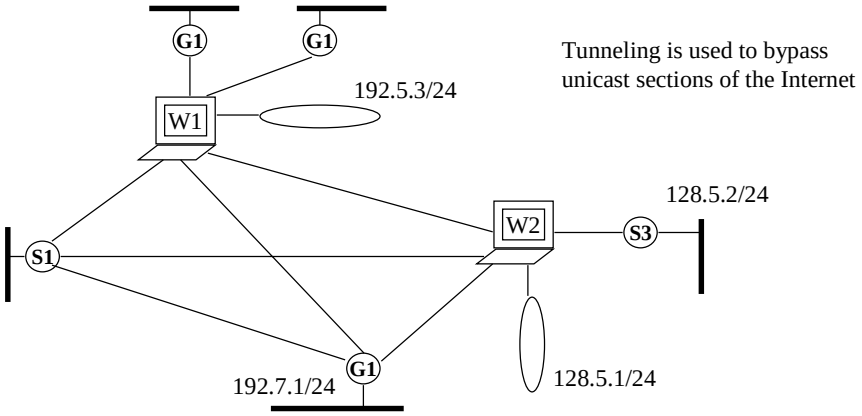
Source based trees for G1



Shared Multicast tree for G1



Mbone overlay is based on WSs running DVMRP



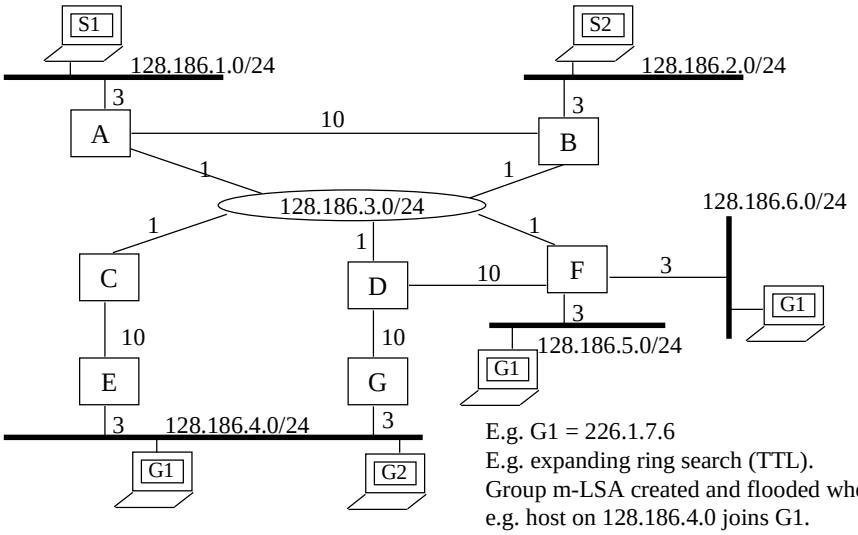
MOSPF (Multicast Extensions to OSPF)

- is an extension of OSPF, allowing Multicast to be introduced into an existing OSPF unicast routing domain
- unlike DVMRP, MOSPF is not susceptible to the normal convergence problems of Distance Vector algorithms.
- limits the extent of multicast traffic to group members, something e.g. DVRMP cannot always do. Restricting the extent of multicast datagrams is desirable for high-bandwidth multicast applications or limited-bandwidth network links (or both).

MOSPF can be deployed gracefully

- Introduces multicast routing by adding a new type of LSA to the OSPF link-state database and by adding calculations for the paths of multicast datagrams.
- The introduction of MOSPF to an OSPF routing can be gradual - MOSPF will automatically route IP multicast datagrams around those routers incapable of multicast routing, whereas unicast routing continues to function normally.
- MOSPF can be, and is in isolated places, deployed in the MBONE. A MOSPF domain can be attached to the edge of the MBONE, or can be used as a transit routing domain within the MBONE's DVMRP routing system.

An MOSPF Routing Domain

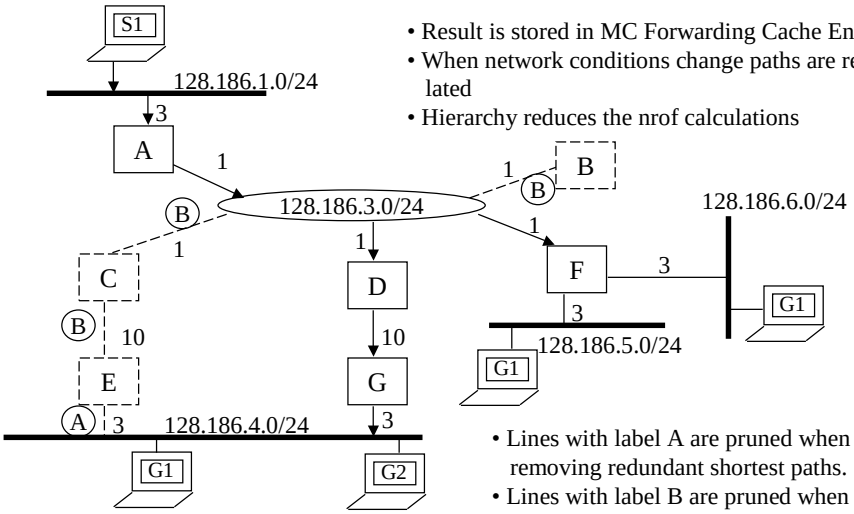


E.g. G1 = 226.1.7.6
E.g. expanding ring search (TTL).
Group m-LSA created and flooded when e.g. host on 128.186.4.0 joins G1.

Group-membership-LSA is created and flooded when an IP user joins an MC-group using IGMP

LS Age	0
Options / LS Type	E-bit. LS Type 6 (group-membership-LSA)
Link State ID	226.1.7.6 (group G1)
Advertising Router	128.186.4.1 (router E)
LS Sequence Number	0x80000001
LS Checksum	0x3da9
Length	28 bytes
Referenced LS Type	2 (network)
Referenced Link State ID	128.186.4.1 (128.186.4.0/24)

MOSPF calculates Shortest-path trees on demand



- Result is stored in MC Forwarding Cache Entry
- When network conditions change paths are recalculated
- Hierarchy reduces the nrof calculations

- Lines with label A are pruned when removing redundant shortest paths.
- Lines with label B are pruned when removing links that do not lead to G1

Forwarding Cache Entry stores MC path routing info

Source network, Group -->

Router or network for mcast reception	
List of Interfaces, mcasts must be sent	Metrics to nearest group member

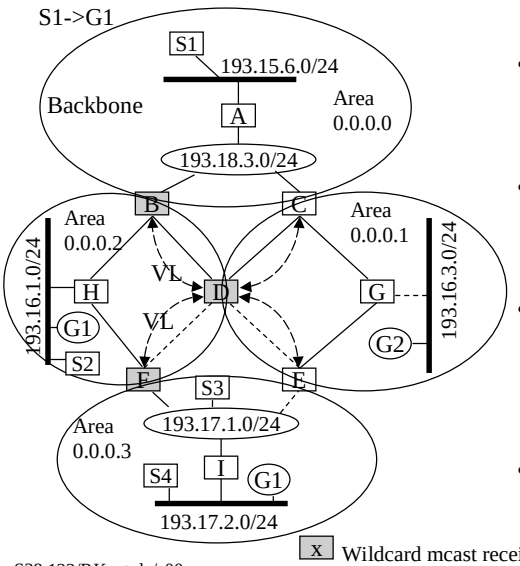
A Cache Entry may be deleted at any time -> will be recalculated on demand

Cache entries must be deleted, when changed LSAs are received
- Router-LSA, Network-LSA (on router or link failure or cost change) => delete all entries since can't tell which are affected
- Group-m-LSA: delete entries of that group
- Hierarchy => the farther away the change is the fewer cache entries are deleted

On demand route calculations use Dijkstra's SPF-algorithm

- Calculation is rooted on the source not the router as for unicast
- For a new mcast, every router performs the same calculation
- Stub networks do not appear in MOSPF calculation (e.g router F)
- Tiebreaks for equal cost routes - previous hop router that has highest address is chosen (e.g. G over E)

Two level hierarchy aggregates both sources and group addresses



- In aggregation some info is lost --> sometimes mcasts are sent needlessly: C->G:to G1
- Presense of sources is reported by summary-LSA with MC -bit set: F to H-> S3+S4 entry
- Area border router advertise Group-m-LSAs to bbone (B: G1, D,E,F:G1, C,D,E:G2) - no exact location
- Routers in non-bbone do not know location of group mmbrs

Summary of Multicast Protocols for the Internet

Shared tree	Source based trees	
	Bcast and Prune	Domainwide reports
PIM Sparse* Core Based tree*	DVMRP PIM Dense*	MOSPF

* Rely on Unicast routing protocol to locate MC-sources.

- Those that don't, can route MC on routes separate from unicast routes.
- For Shared tree protocols an additional step of finding the Core or Rendesvouz Point must be performed.
- Directories are useful on service management level.