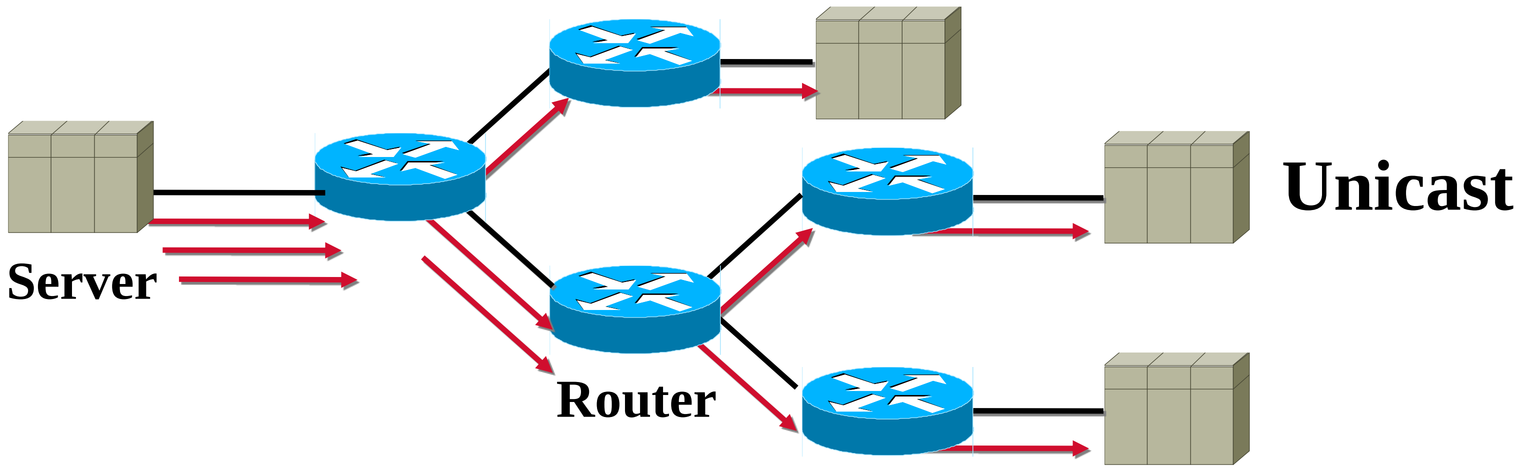
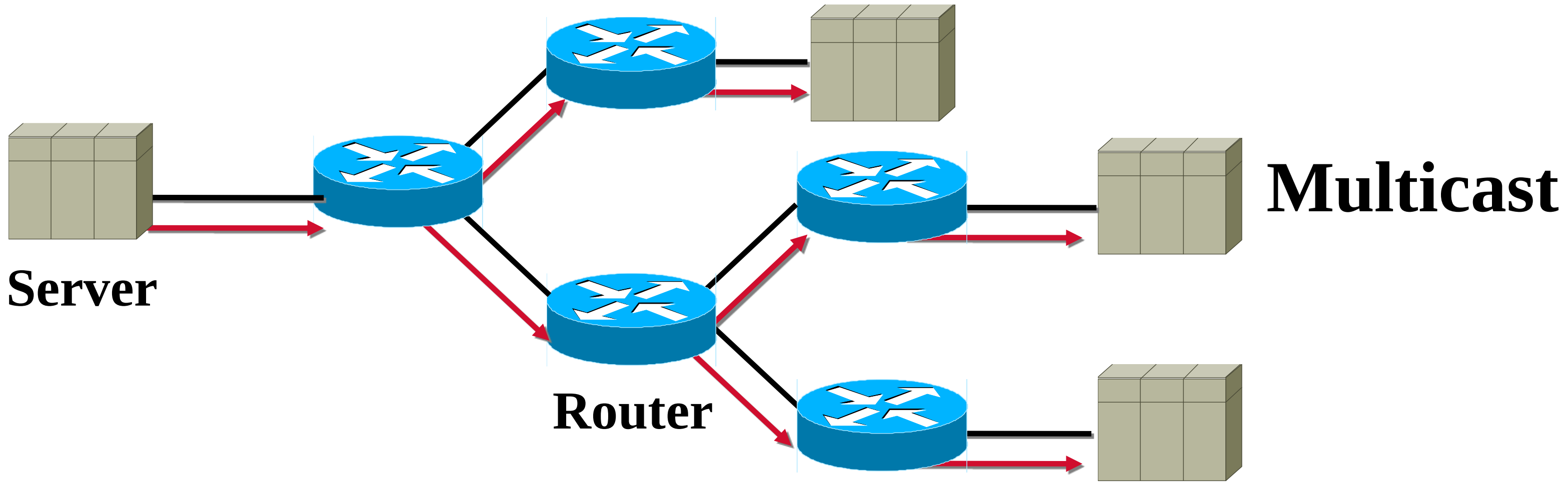
IP Multicast

Introduction to Multicast

- Why multicast?
  - When sending same data to multiple receivers
  - Better bandwidth utilization
  - Lesser host/router processing
  - Receivers' addresses unknown
- Applications
  - Video/audio conferencing
  - Resource discovery/service advertisement
  - Stock distribution

What is Multicasting?





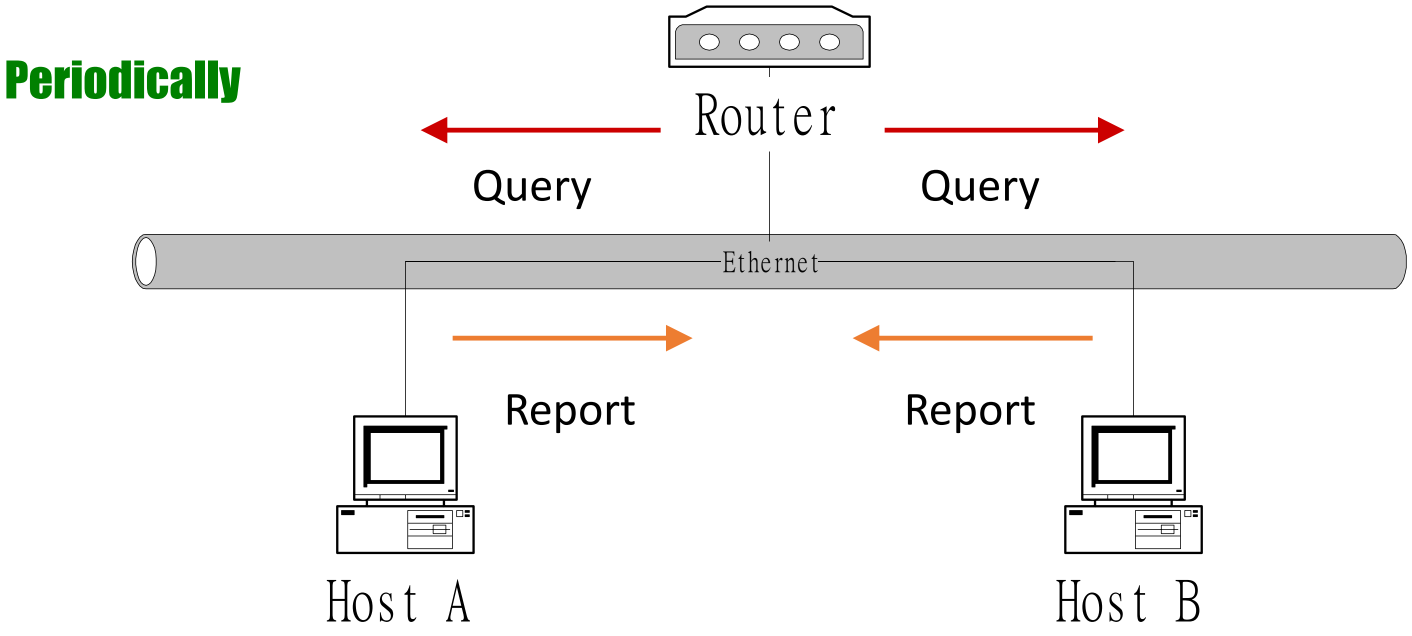
- In 1995 the first mcast network was born: MBone
- MBONE
  - Internet’s Multicast Backbone
  - More than 2800 subnets in over 25 country by 1996 March
  - A semi-permanent IP multicast testbed
  - Use IP tunneling to send data
  - Multicast routing protocol: DVMRP, MOSPF, PIM

Related protocols

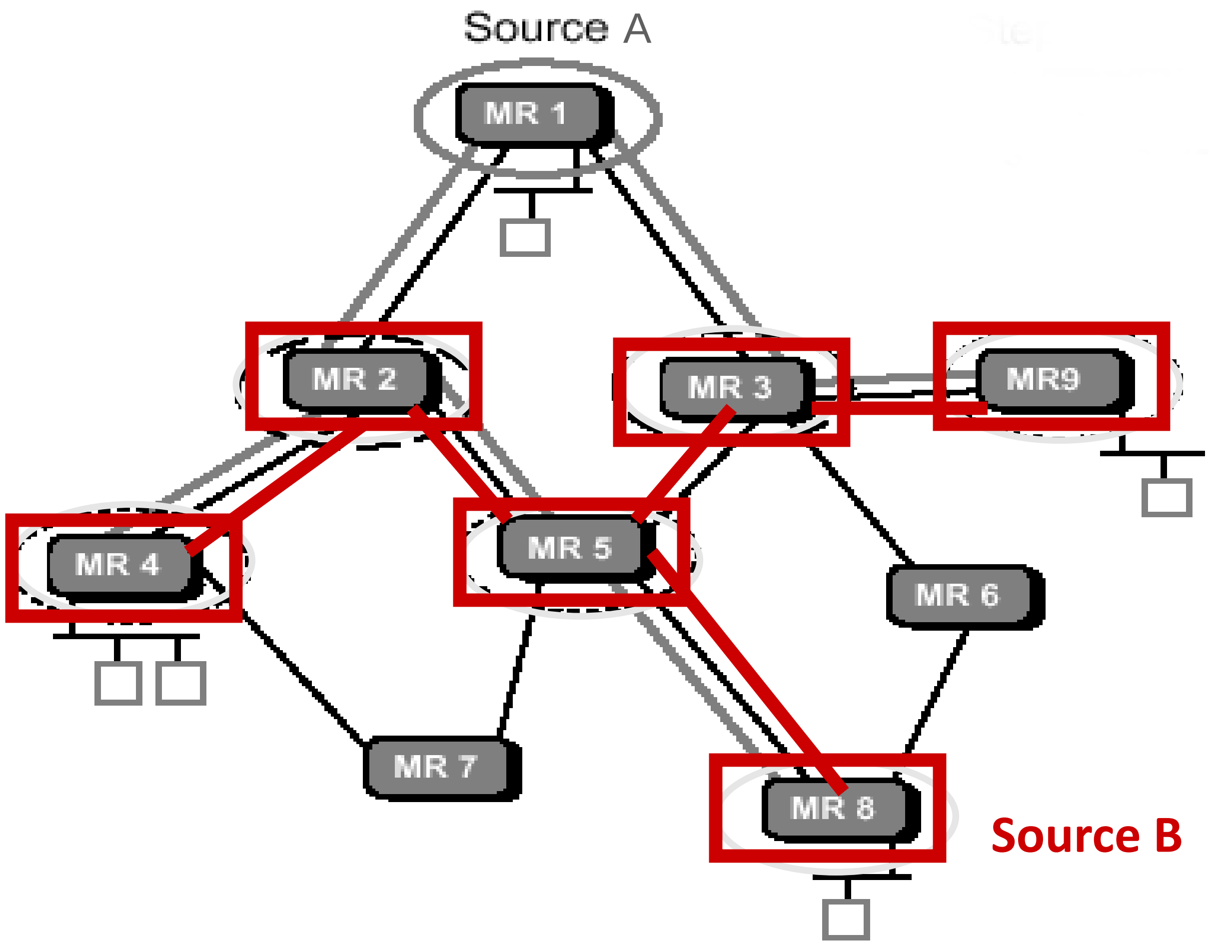
- Group management protocol:
  - IGMP(Internet Group Management Protocol)
- Multicast routing protocol:
  - Source-based tree
    - Multicast tree per (S,G)
    - DVMRP, MOSPF, PIM-DM
  - Shared tree
    - Multicast tree per (*,G)
    - CBT, PIM-SM

IGMP

- The multicast router use IGMP to obtain the group-membership information of the LAN.



Source-based tree

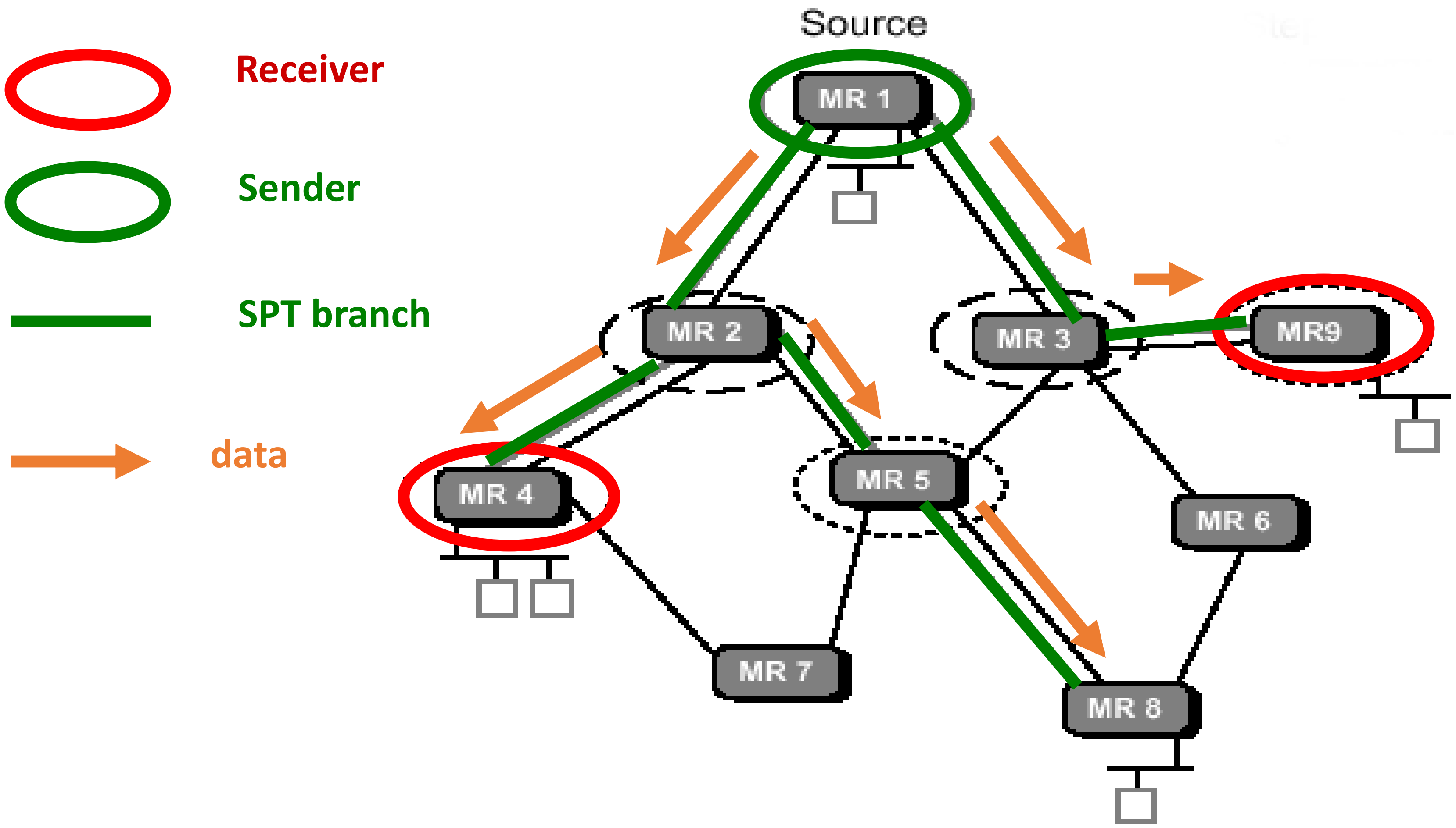


- Source A and B use the same multicast address to send data.
- Intermediate routers use (S,G) to forward data.
- Intermediate routers create forward entry while receiving data.

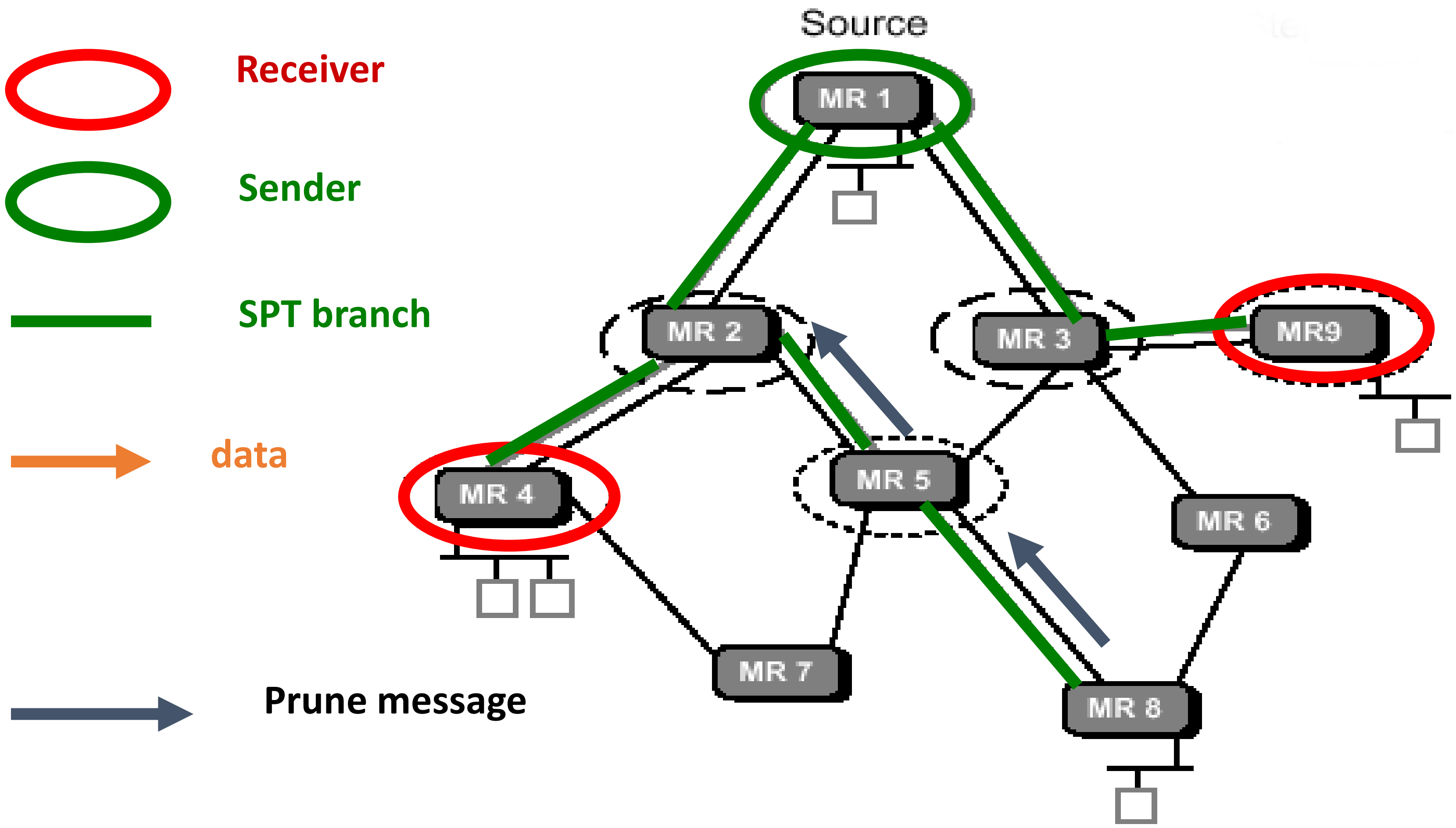
DVMRP

- Distance vector multicast routing protocol
- Rely on distance vector based unicast routing protocol(RIP)
- For each (S,G), grow a shortest-path tree rooted on the source
- Use distance vector routing algorithm to support reverse path multicasting(RPM)
- DVMRP subnetworks was interconnected through the unicast Internet infrastructure with tunnels
- Flood and Prune technology
  - Very successful in academic circles

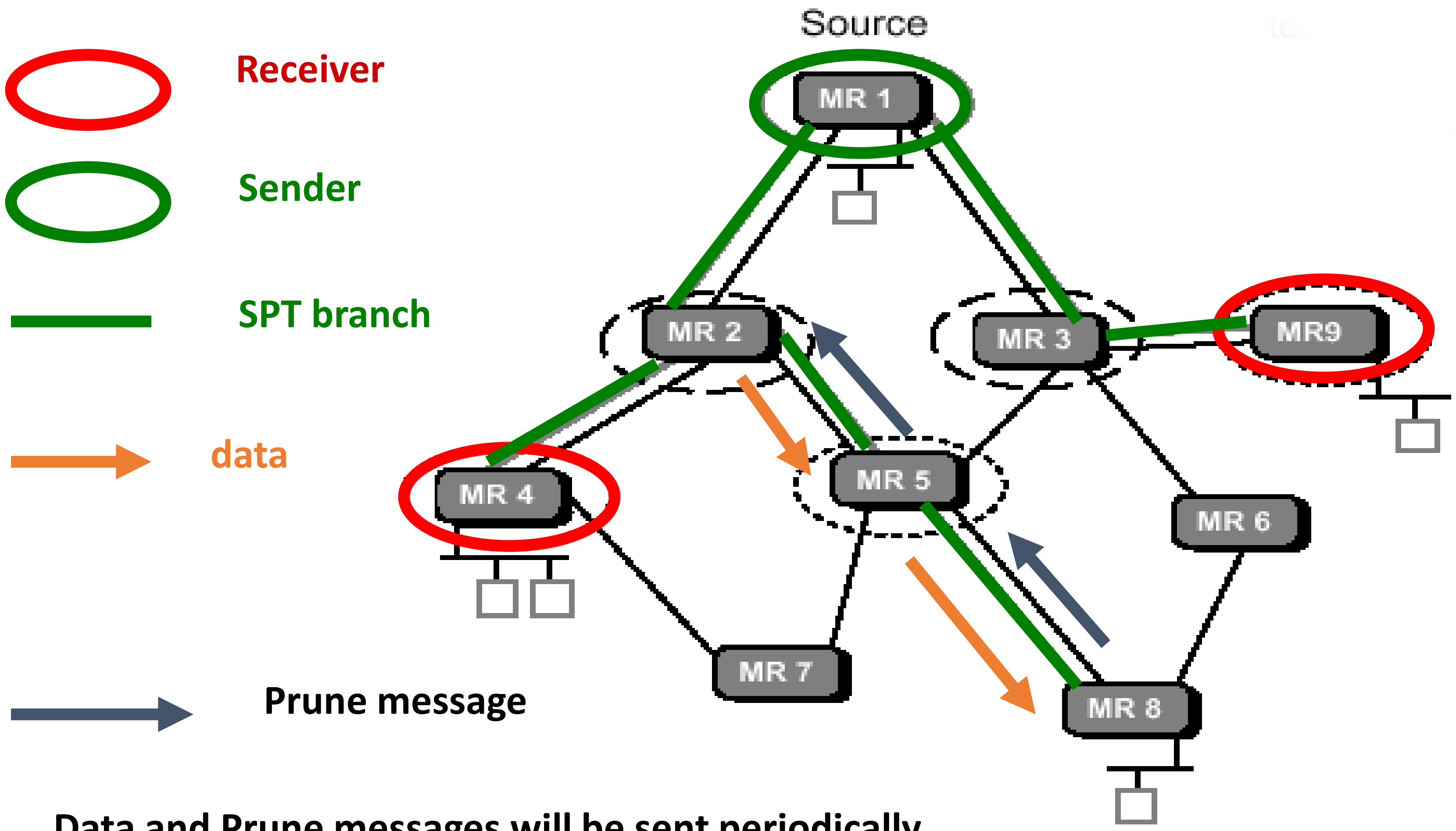
DVMRP


Receiver
Sender
SPT branch
data
Source
MR 1
MR 2
MR 3
MR9
MR 4
MR 5
MR 6
MR 7
MR 8

DVMRP

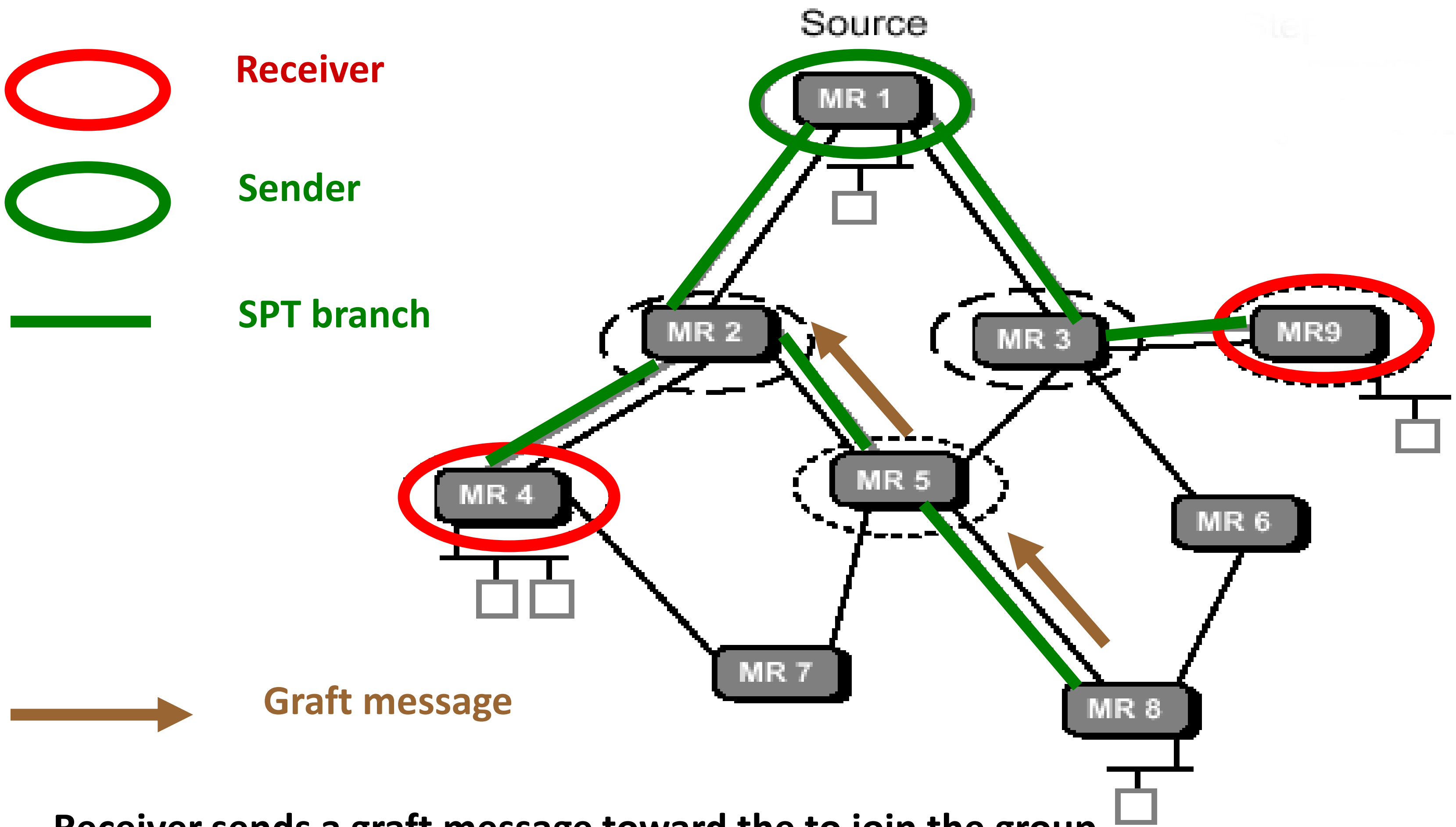

Receiver
Sender
SPT branch
data
Prune message
Source
MR 1
MR 2
MR 3
MR9
MR 4
MR 5
MR 6
MR 7
MR 8

DVMRP



. Data and Prune messages will be sent periodically.

DVMRP



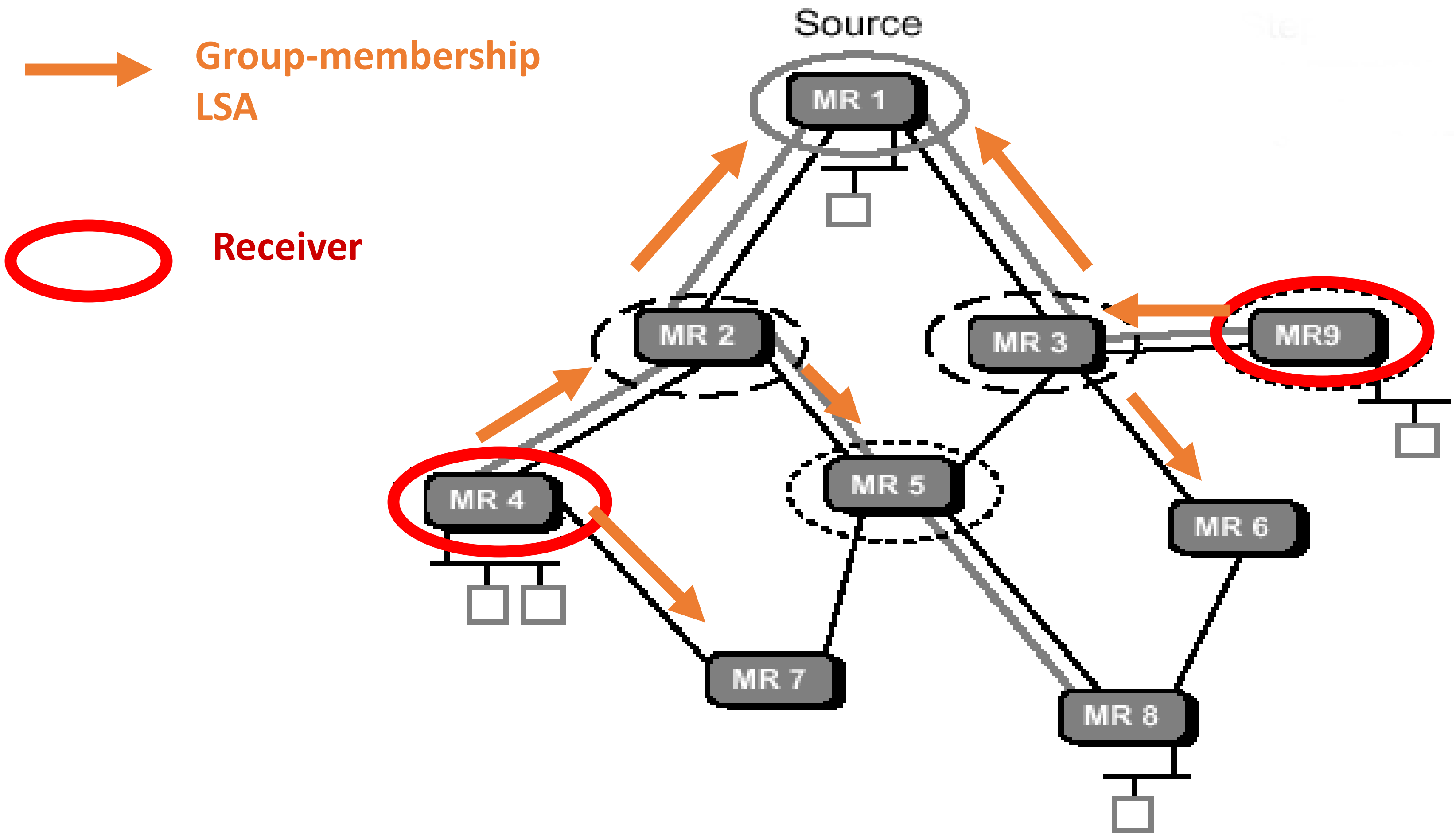
- Receiver sends a graft message toward the to join the group quickly.

- Problem DVMRP can’t scale to Internet sizes
  - Distance vector-based routing protocol
  - Periodic updates
    - Full table refresh every 60 seconds
  - Table sizes
    - Internet > 40,000 prefixes at that moment
  - Scalability
    - Too many tunnels, hop-count till 32 hops, etc

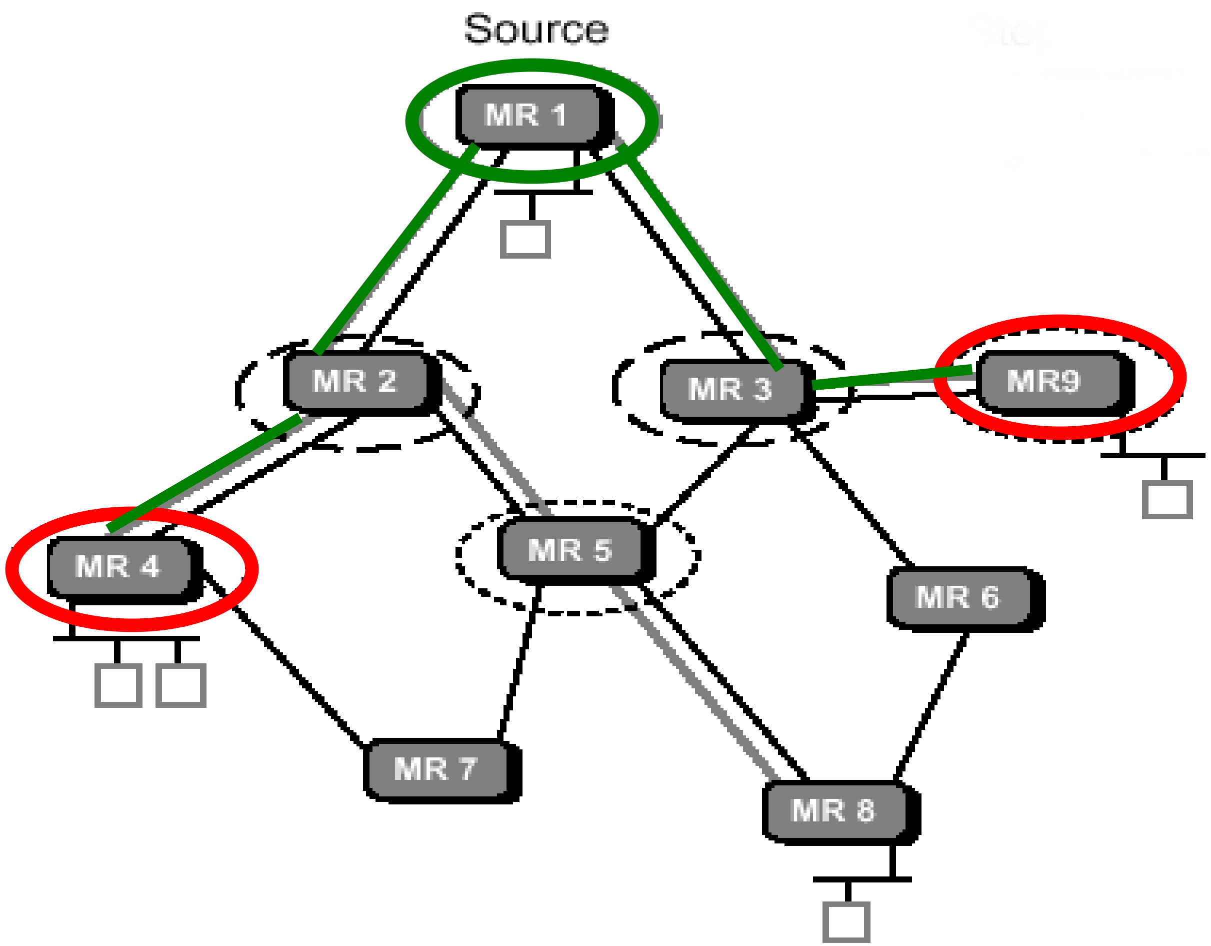
MOSPF

- Multicast open shortest path first routing protocol
- Rely on link state based unicast routing protocol(OSPF)
- For each (S,G), grow a shortest-path tree rooted on the sources
- Use Dijkstra’s algorithm and group-membership LSAs to compute shortest path tree

MOSPF


Group-membership LSA
Receiver
Source
MR 1
MR 2
MR 3
MR9
MR 4
MR 5
MR 6
MR 7
MR 8

MOSPF



- MOSPF creates tree without prune messages

In 1997, a native protocol is developed, Protocol Independent Multicast

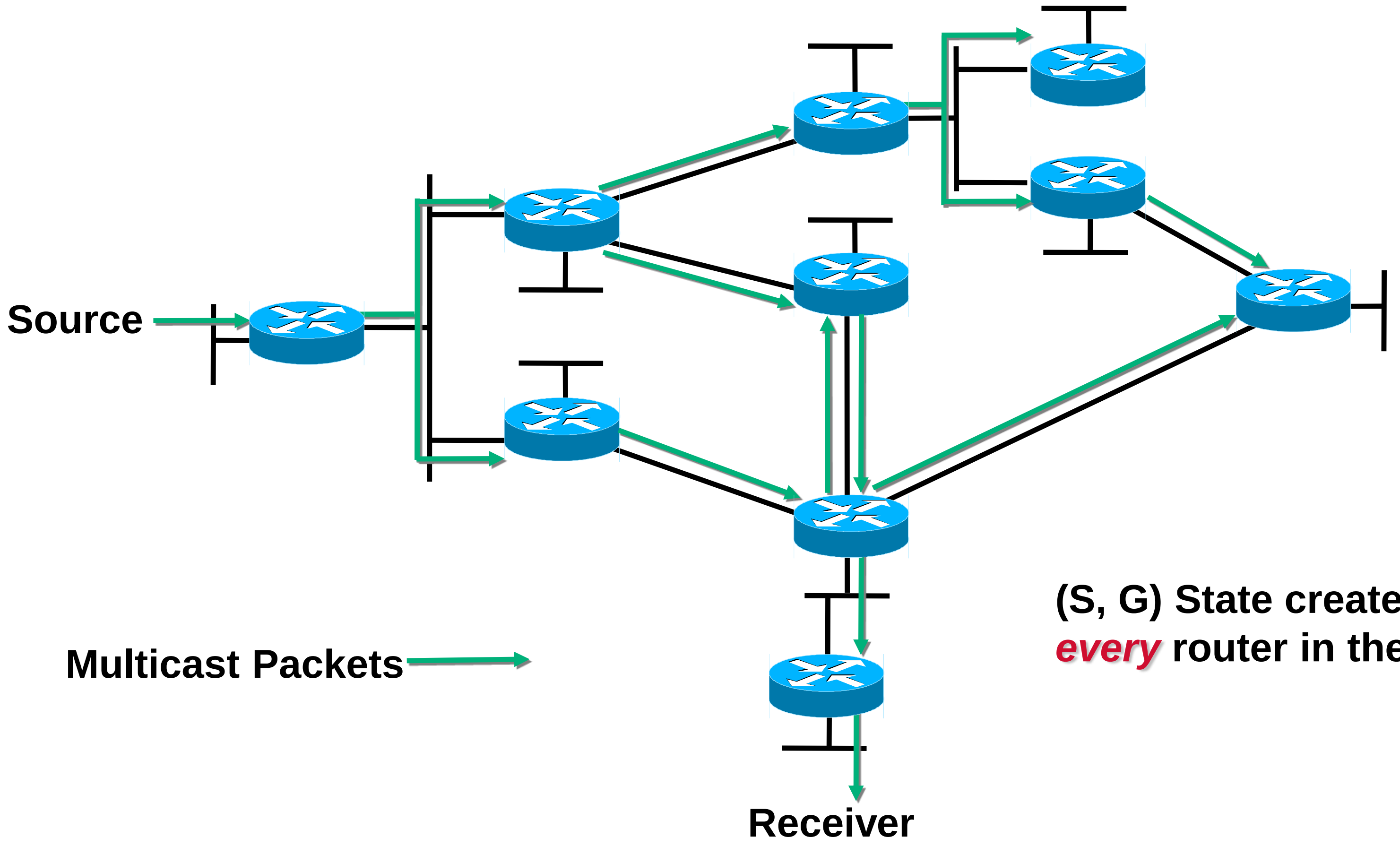
- PIM Dense mode
  - Flood and Prune behavior very inefficient
    - Can cause problems in certain network topologies
  - Creates (S, G) state in EVERY router
    - Even when there are no receivers for the traffic
  - Complex Assert mechanism
    - To determine which router in a LAN will forward the traffic
  - No support for shared trees

(S,G) notation

- For every multicast source there must be two pieces of information: the source IP address, S, and the group address, G.
  - This is generally expressed as (S,G).
  - Also commonly used is (*,G) - every source for a particular group.
  - The router creates a table with the entries (*,G),(S,G).

PIM Dense Mode Overview

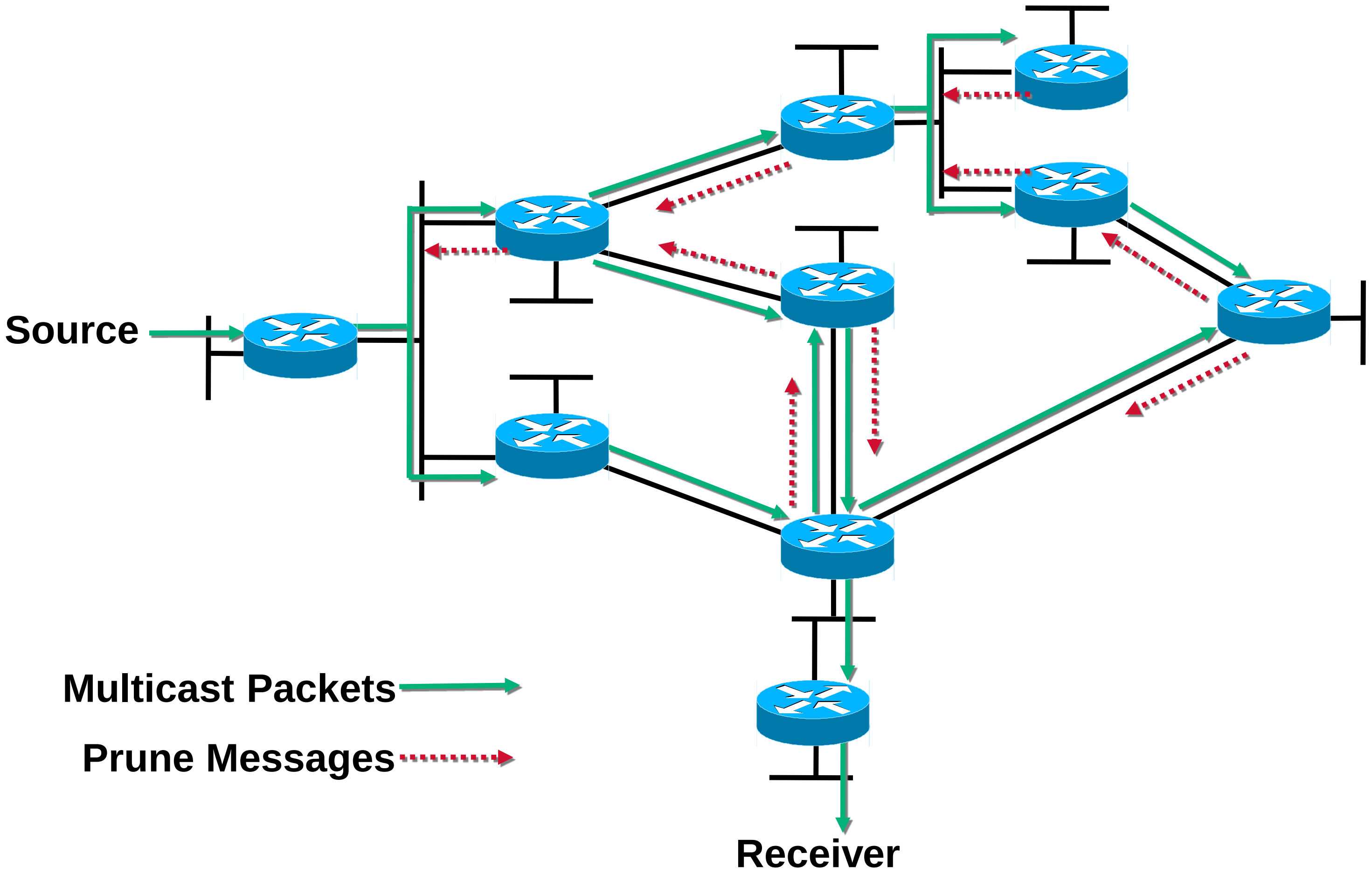
Initial Flooding



(S, G) State created in
every router in the network!

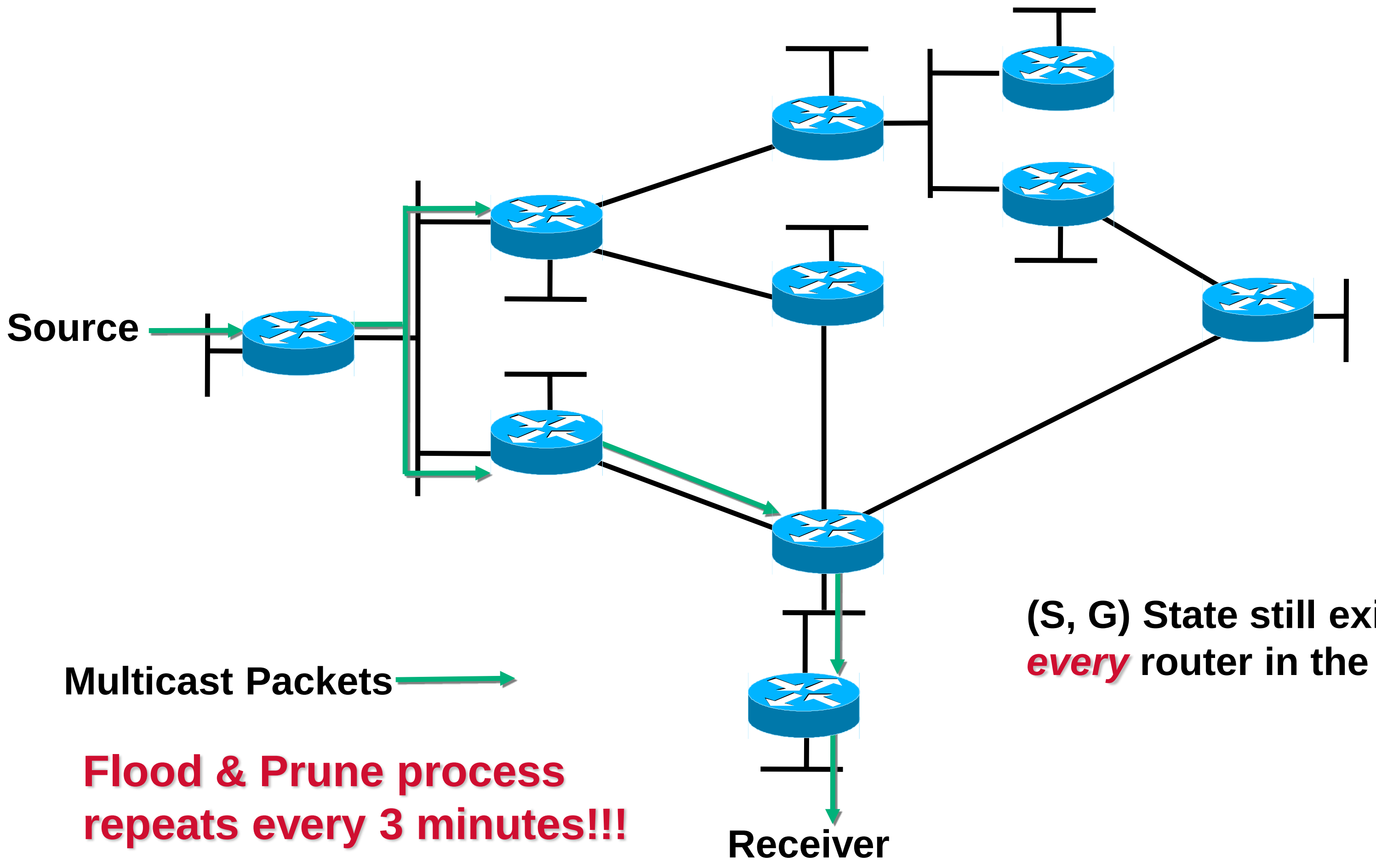
- Initial Flooding
    - In this example, multicast traffic being sent by the source is flooded throughout the entire network.
    - As each router receives the multicast traffic via its Reverse Path forwarding (RPF) interface (the interface in the direction of the source), it forwards the multicast traffic to all of its PIM-DM neighbors.
    - Note that this results in some traffic arriving via a non-RPF interface such as the case of the two routers in the center of the drawing. (Packets arriving via the non-RPF interface are discarded.) These non-RPF flows are normal for the initial flooding of data and will be corrected by the normal PIM-DM pruning mechanism.

Pruning Unwanted Traffic



- Pruning unwanted traffic
    - In the example above, PIM Prunes (denoted by the dashed arrows) are sent to stop the flow of unwanted traffic.
    - Prunes are sent on the RPF interface when the router has no downstream members that need the multicast traffic.
    - Prunes are also sent on non-RPF interfaces to shutoff the flow of multicast traffic that is arriving via the wrong interface (i.e. traffic arriving via an interface that is not in the shortest path to the source.)
        - An example of this can be seen at the second router from the receiver near the center of the drawing. Multicast traffic is arriving via a non-RPF interface from the router above (in the center of the network) which results in a Prune message.

Results After Pruning



(S, G) State still exists in *every* router in the network!

Flood & Prune process repeats every 3 minutes!!!

- Results after Pruning
  - In the final drawing in example shown above, multicast traffic has been pruned off of all links except where it is necessary. This results in a Shortest Path Tree (SPT) being built from the Source to the Receiver.
  - Even though the flow of multicast traffic is no longer reaching most of the routers in the network, (S, G) state still remains in *ALL* routers in the network. This (S, G) state will remain until the source stops transmitting.
  - In PIM-DM, Prunes expire after three minutes. This causes the multicast traffic to be re-flooded to all routers just as was done in the “Initial Flooding” drawing. This periodic (every 3 minutes) “Flood and Prune” behavior is normal and must be taken into account when the network is designed to use PIM-DM.

PIM Sparse mode

- Must configure a Rendezvous Point (RP)
  - Statically (on every Router)
  - Using Auto-RP(Cisco) or Bootstrap Router (BSR) (i.e. Routers learn RP automatically)
- Very efficient
  - Uses Explicit Join model
  - Traffic only flows to where it’s needed
  - Router state only created along flow paths
- Scales better than dense mode
  - Works for both sparsely or densely populated networks