Advanced Networks

(EENGM4211)

Rasheed Hussain

rasheed.hussain@bristol.ac.uk

Best contact : Microsoft Team

Slides courtesy: Prof. Reza Nejabati and Slides from Ross's Book

Unit Outline

1. Introduction
2. Internet Routing and Switching
3. IP Multicast
4. Networking for Realtime Applications
5. Routing in Wireless Networks
6. Quality of Service

Topics

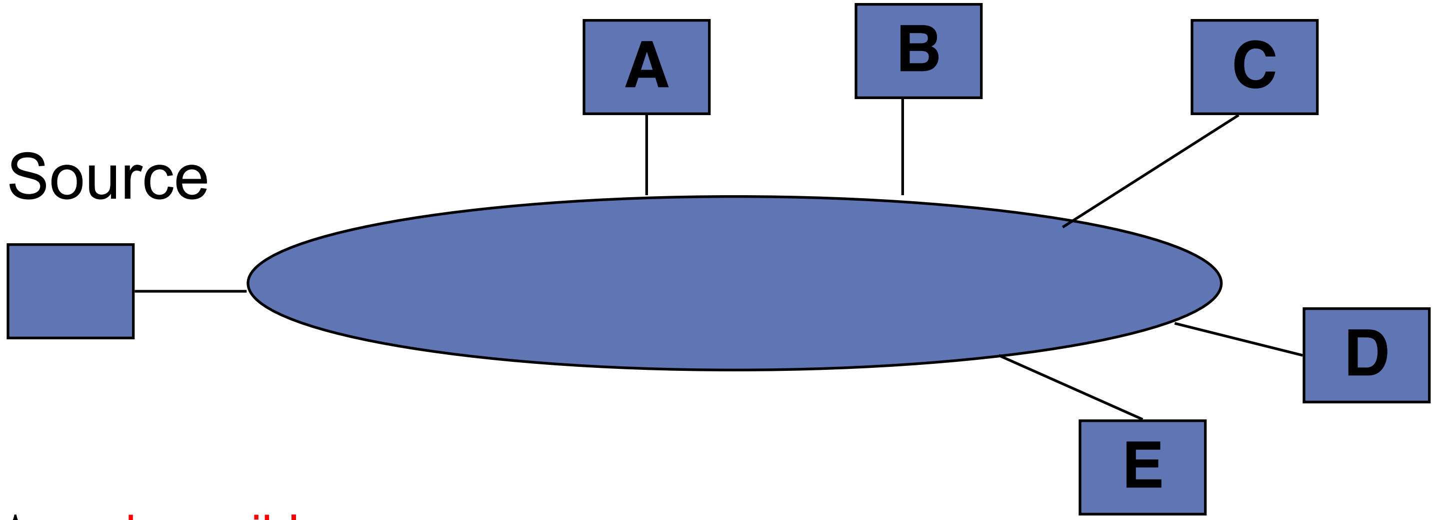
- What is multicast
  - why it is different
  - what applications use it
  - how is it implemented (layering)
- What does it mean for
  - addressing
  - routing
  - protocols
- Multicast Routing

ILOs

- Demonstrate knowledge of multicast
- Differentiate between unicast, multicast, and broadcast
- Understand and demonstrate the knowledge about multicast routing and strategies

Basic Multicast Routing Protocols (1)

- Problem:
  - Given a source and a set of destinations,
  - Route the same packet to at least (or exactly) this set of destinations



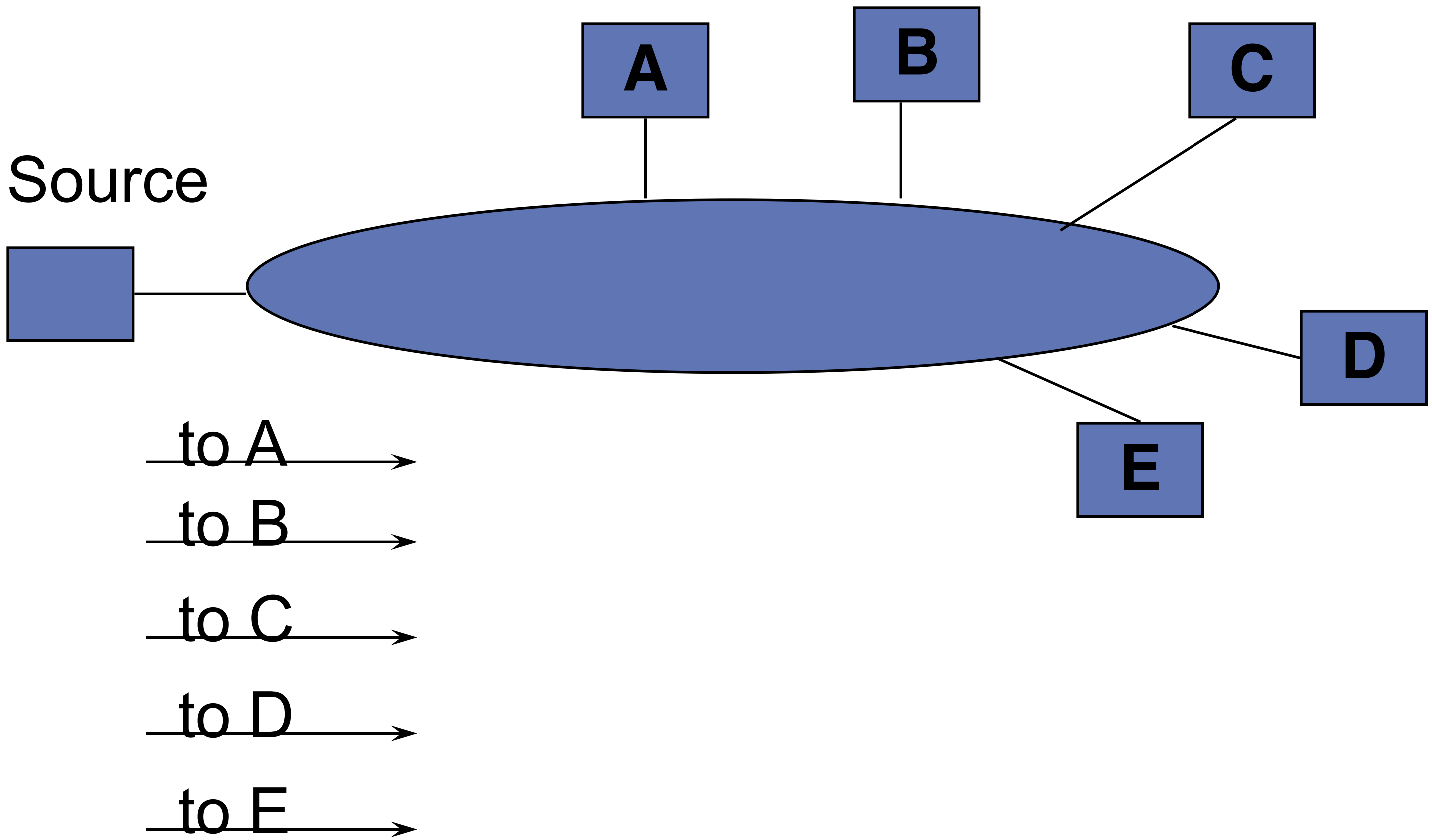
✦ and possibly:
- Optimise routes from source to receivers
- Maintain loop-free routes
- Distribute multicast load fairly/equally between all possible links (do not create hot spots)

Basic Multicast Routing Protocols (2)

- Multicast by Broadcast (flooding)
- How?
  - Filter above network layer
    - Natural in broadcast networks (satellite, bridged LANs)
    - Use flooding in packet switched networks
- Issues:
  - Bandwidth inefficient
  - Security concerns
    - Though these are endemic in the multicast open model

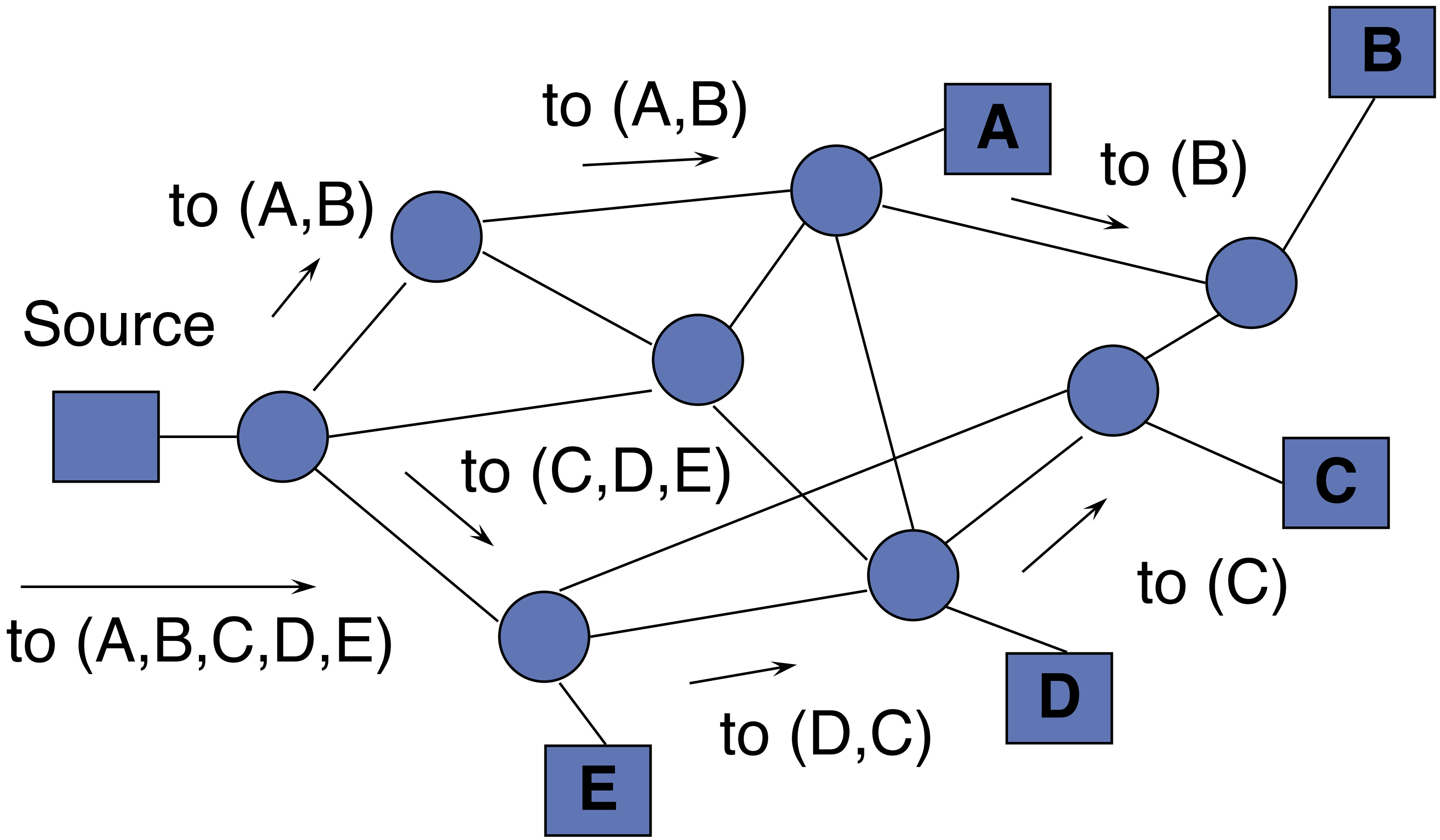
Basic Multicast Routing Protocols (3)

- Separately addressed packets



Basic Multicast Routing Protocols (4)

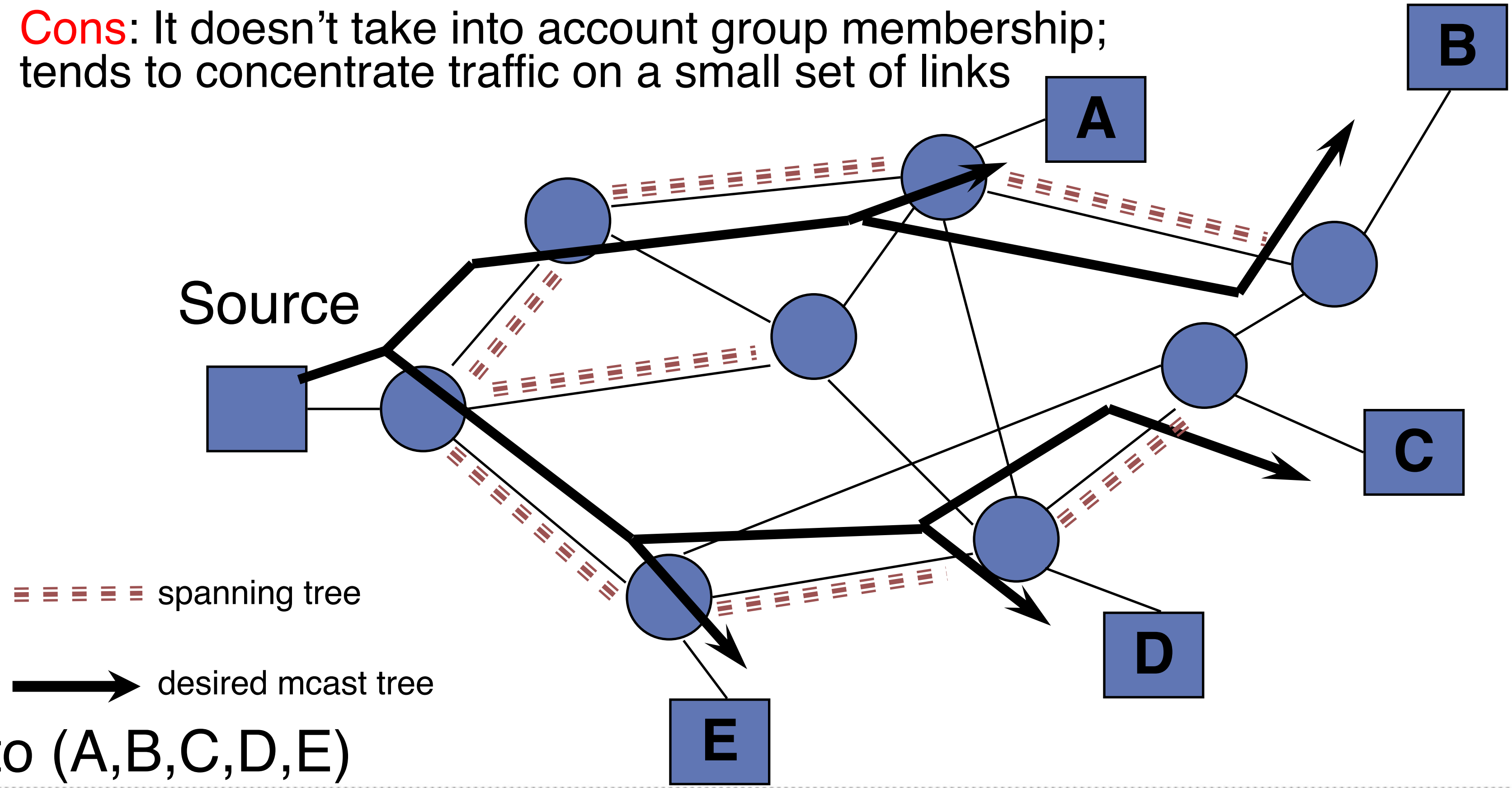
Multidestination Addressing



Basic Multicast Routing Protocols (5)

Spanning tree forwarding

- Pros: most efficient in minimising unnecessary traffic; robust; low memory requirement
- Cons: It doesn't take into account group membership; tends to concentrate traffic on a small set of links



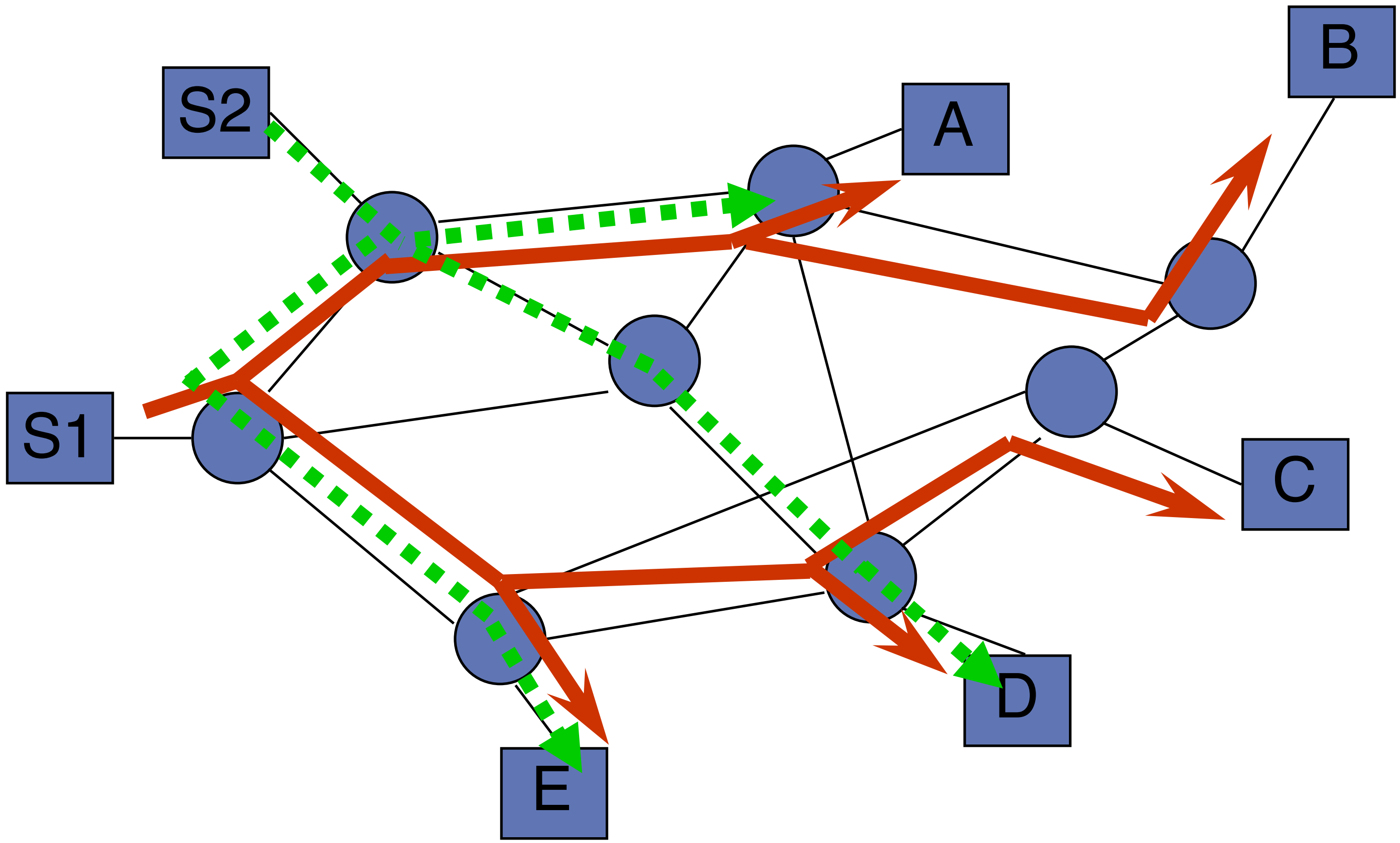
to (A,B,C,D,E)

Shared Tree vs. Source-Based Tree (1)

- Source-based tree
  - one per source/group pair (s, g)
  - good delay properties
  - overhead per source and group
- Shared tree
  - single tree for all sources (*, g)
  - per group overhead
  - higher delays
  - more traffic concentration
- Router information for multicast
  - In source-based tree
    - per-interface, per-group, per-source
  - In shared tree
    - per-interface, per-group

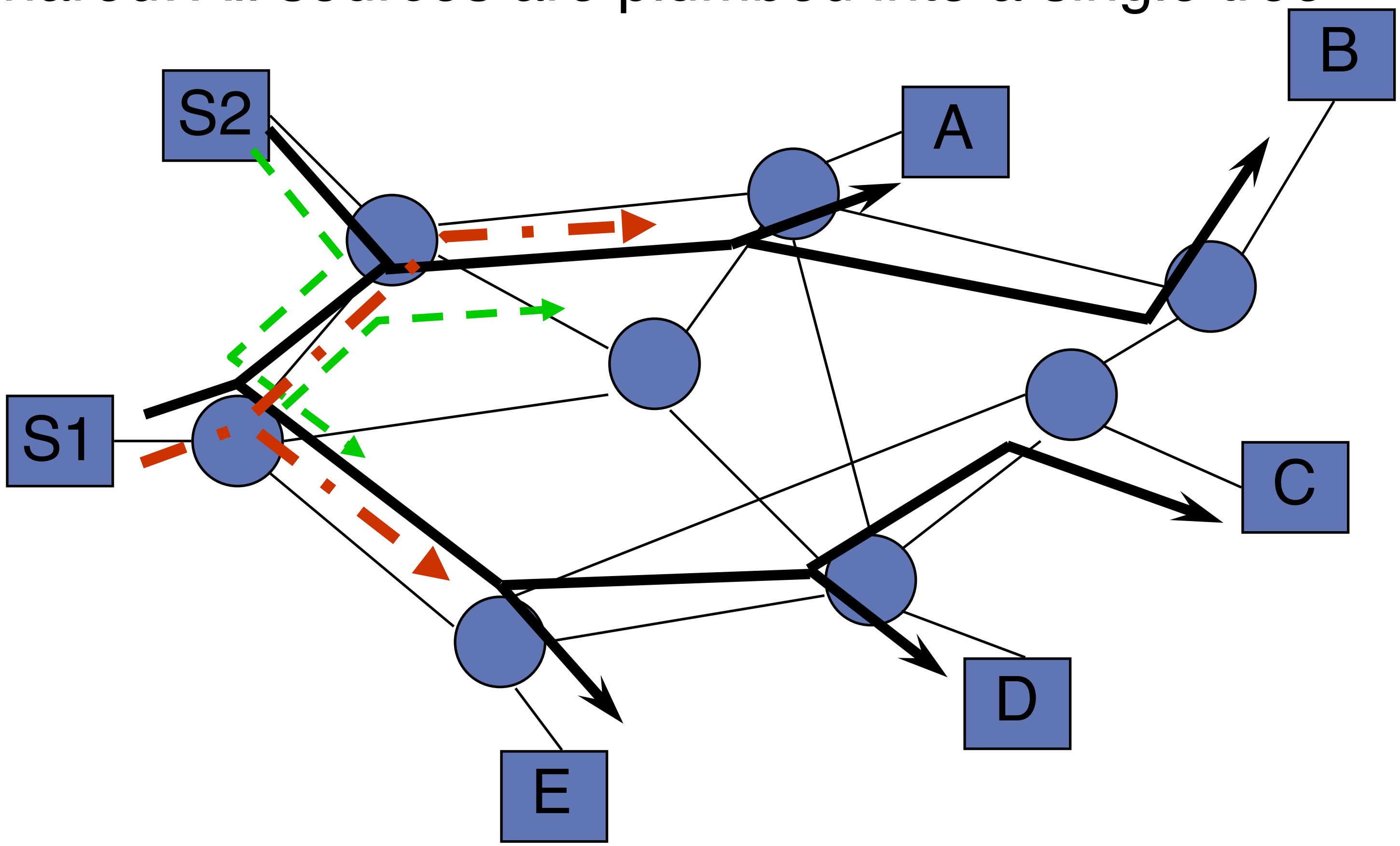
Shared Tree vs. Source-Based Tree (2)

Source-Based : All sources have their own tree



Shared Tree vs. Source-Based Tree (3)

Shared: All sources are plumbed into a single tree



Multicast Routing Algorithms

- Shortest Path Tree algorithms (source-based tree)
  - Compute a tree *rooted at the sender* and spanning all the receivers; the tree should be such that the distance between the sender and ***each*** receiver is minimum.
  - Minimises end-to-end delay.

- Minimum Cost Tree algorithms (shared tree)
  - Compute a tree spanning the sender and receivers such that *the overall cost of the tree* is minimal. The tree should not include any node that is not a member of the group.
  - Minimises overall cost of the multicast tree

- Constrained Tree algorithms
  - E.g. → compute the minimum cost tree that does not have any sender-receiver path that exceeds a delay bound ***d***.
  - Combines the two metrics: tree cost and delay.

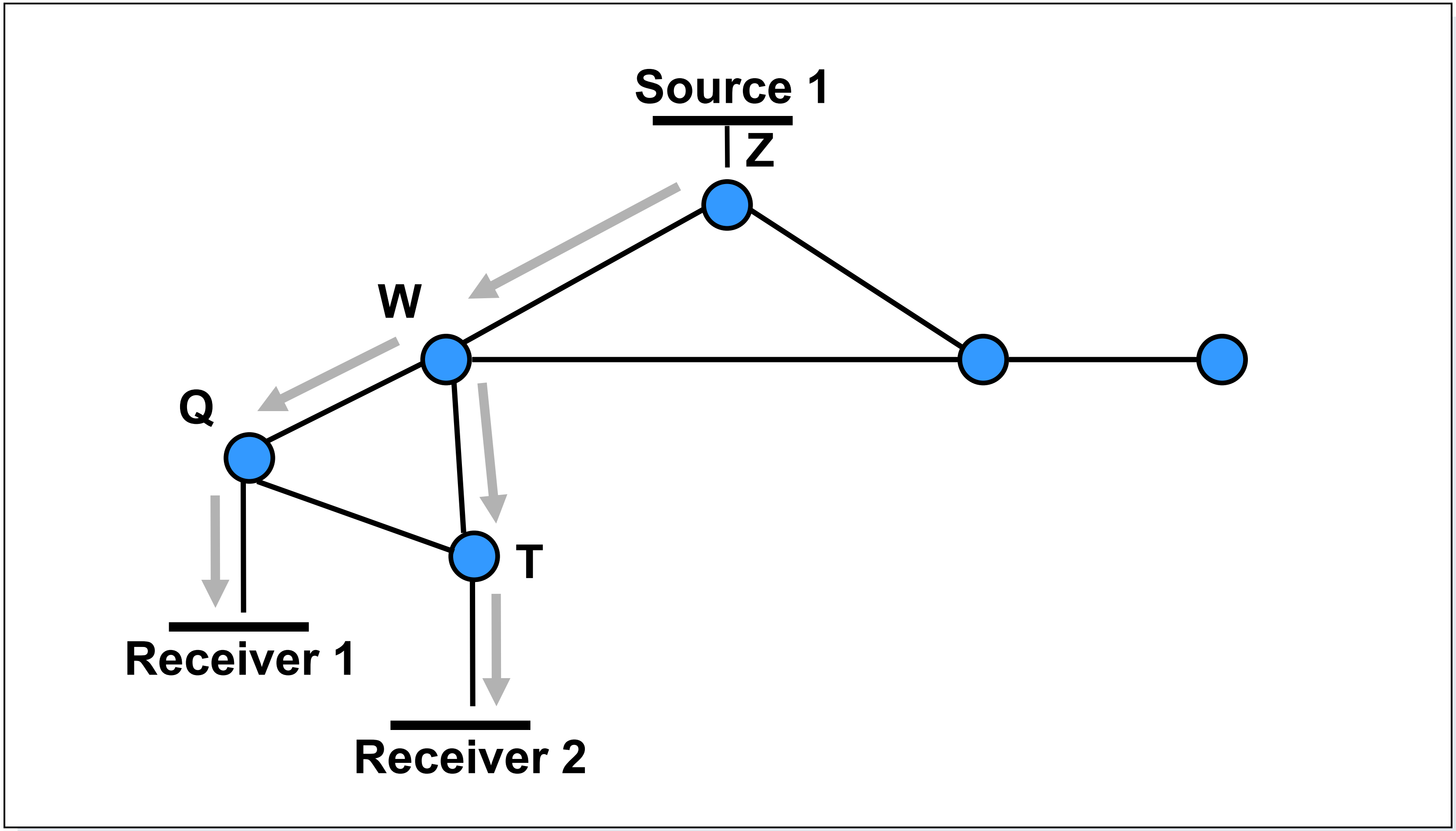
Routing Protocols

- Source-based Tree Protocols:
  - DVMRP
  - MOSPF
  - PIM-DM
- Shared-Tree Protocols
  - CBT
  - PIM-SM

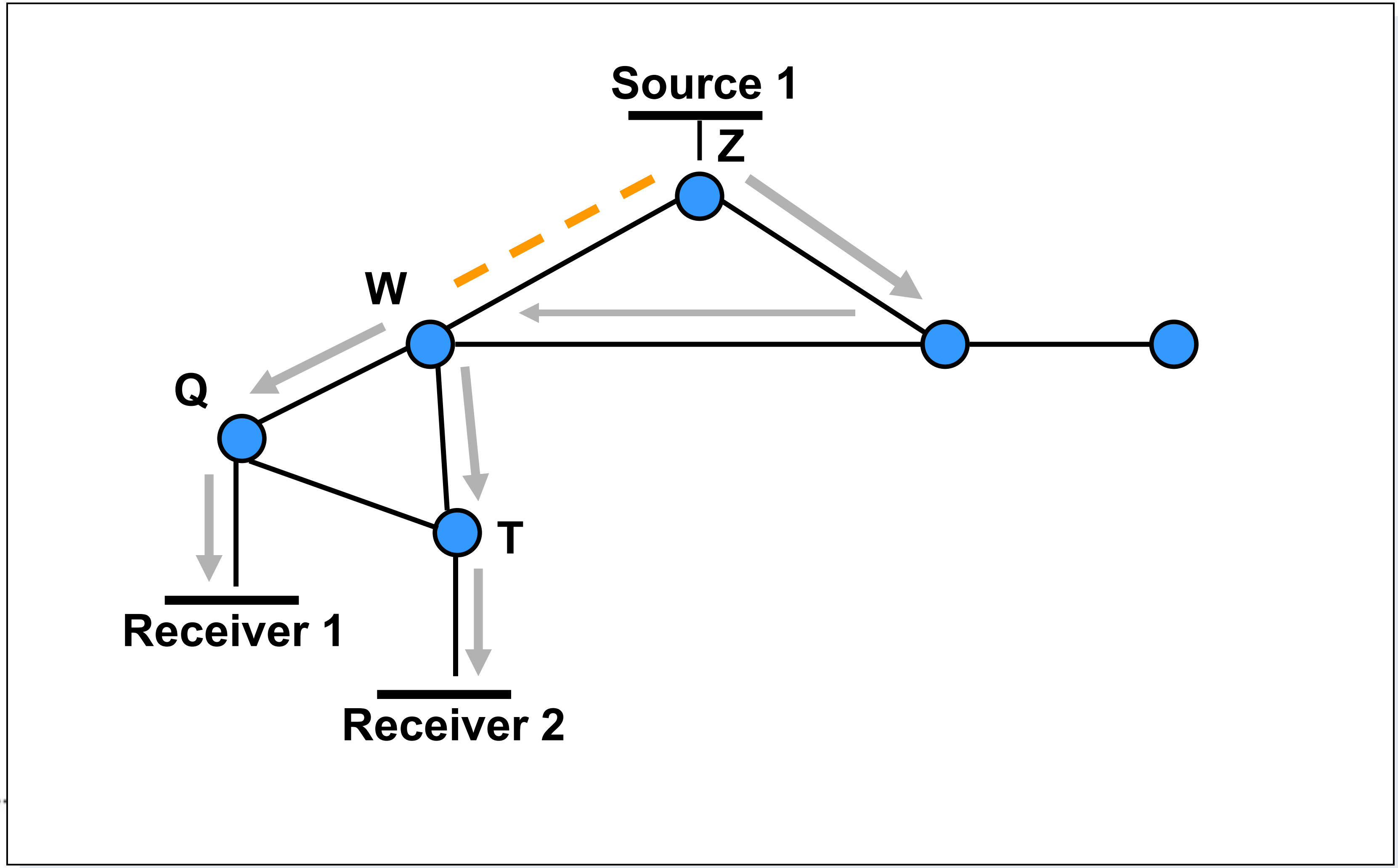
Multicast OSPF (MOSPF)

- Add-on to OSPF (Open Shortest-Path First, a link-state, intra-domain routing protocol)
- Multicast-capable routers utilise link state routing advertisements
- Link-state packets include multicast group addresses to which local members have joined
- Routing algorithm augmented to compute shortest-path distribution tree from a source to any set of destinations

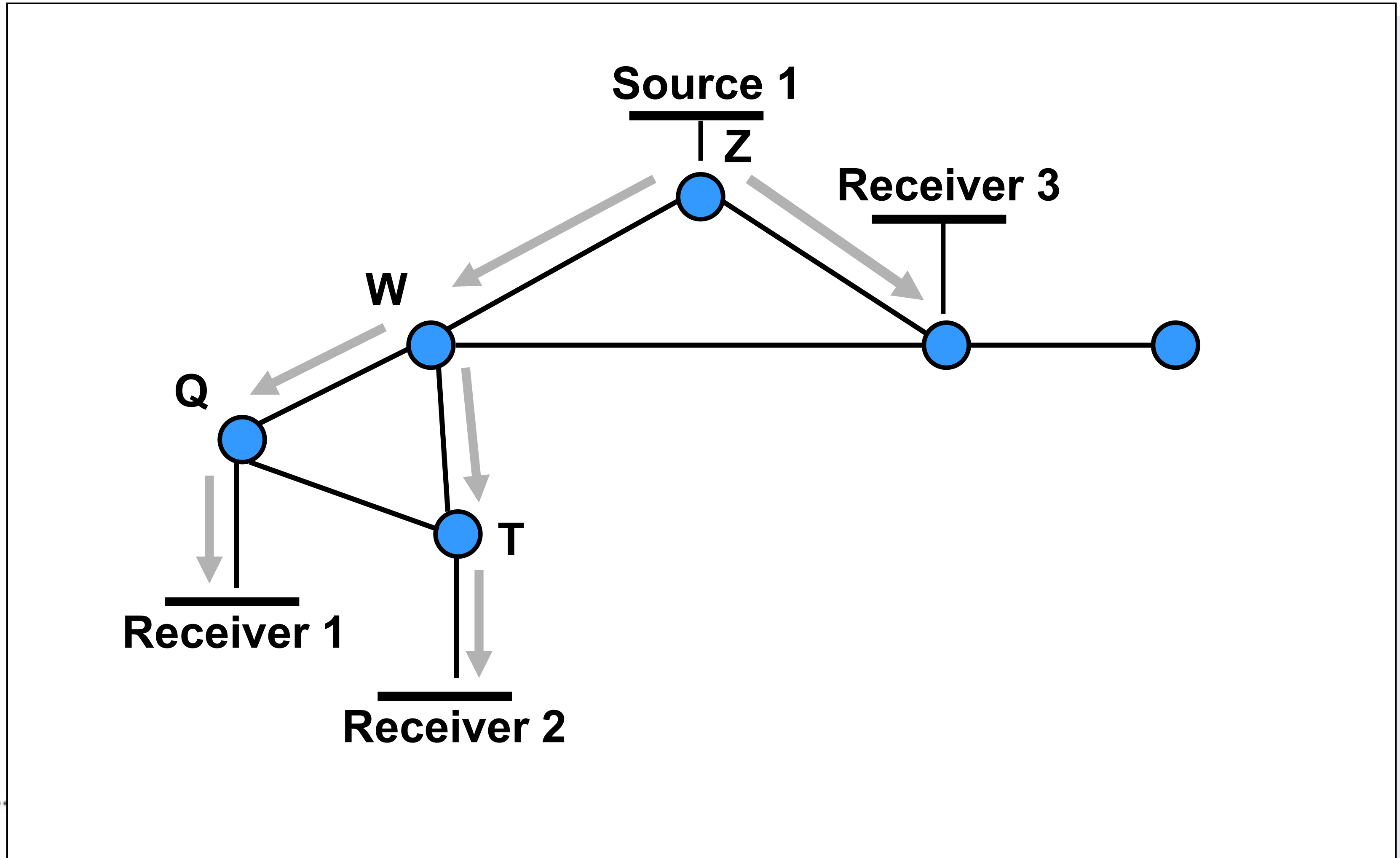
Example



Link Failure/Topology Change



Membership Change



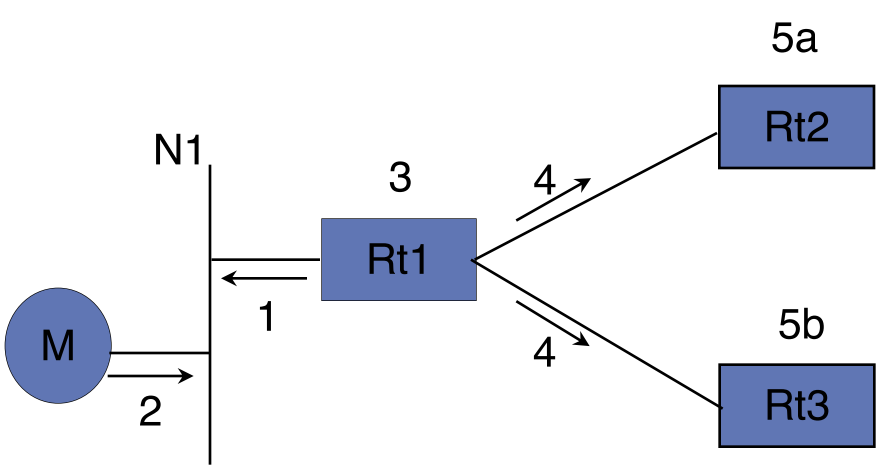
Slides from Srinivasan Seshan

Impact on Route Computation

- Can't pre-compute multicast trees for all possible sources
- Compute on demand when first packet from a source S to a group G arrives
- New link-state advertisement
  - May lead to addition or deletion of outgoing interfaces if it contains different group addresses
  - May lead to re-computation of entire tree if links are changed

Multicast OSPF (MOSPF)

- Link-State Multicast Routing
  - Multicast extensions to OSPF
- Routers maintain topology databases
- *Group-Membership-LSA* msgs broadcast by routers to advertise links with members
- Routers compute and cache pruned SPTs
- Pruning is done by timeout.



Joining multicast group

1. IGMP query
2. IGMP response
3. Create entry in local group DB {Rt1, (G,N1)}
4. Send group membership LSA (G, Rt1)
5. (a) Create entry in local group DB in Rt2 → {Rt2, (G, Rt1)}
   (b) Create entry in local group DB in Rt3 → {Rt3, (G, Rt1)}

Distance Vector Multicast Routing Protocol(1/4)

- Basis for DVMRP
  - DVMRP = Distance Vector Multicast Routing Protocol
- Main advantage: use of existing unicast routing infrastructure
- Broadcast-based protocol
- Group addressing used
- Routers/switches forward *based on source address* of multicast packet

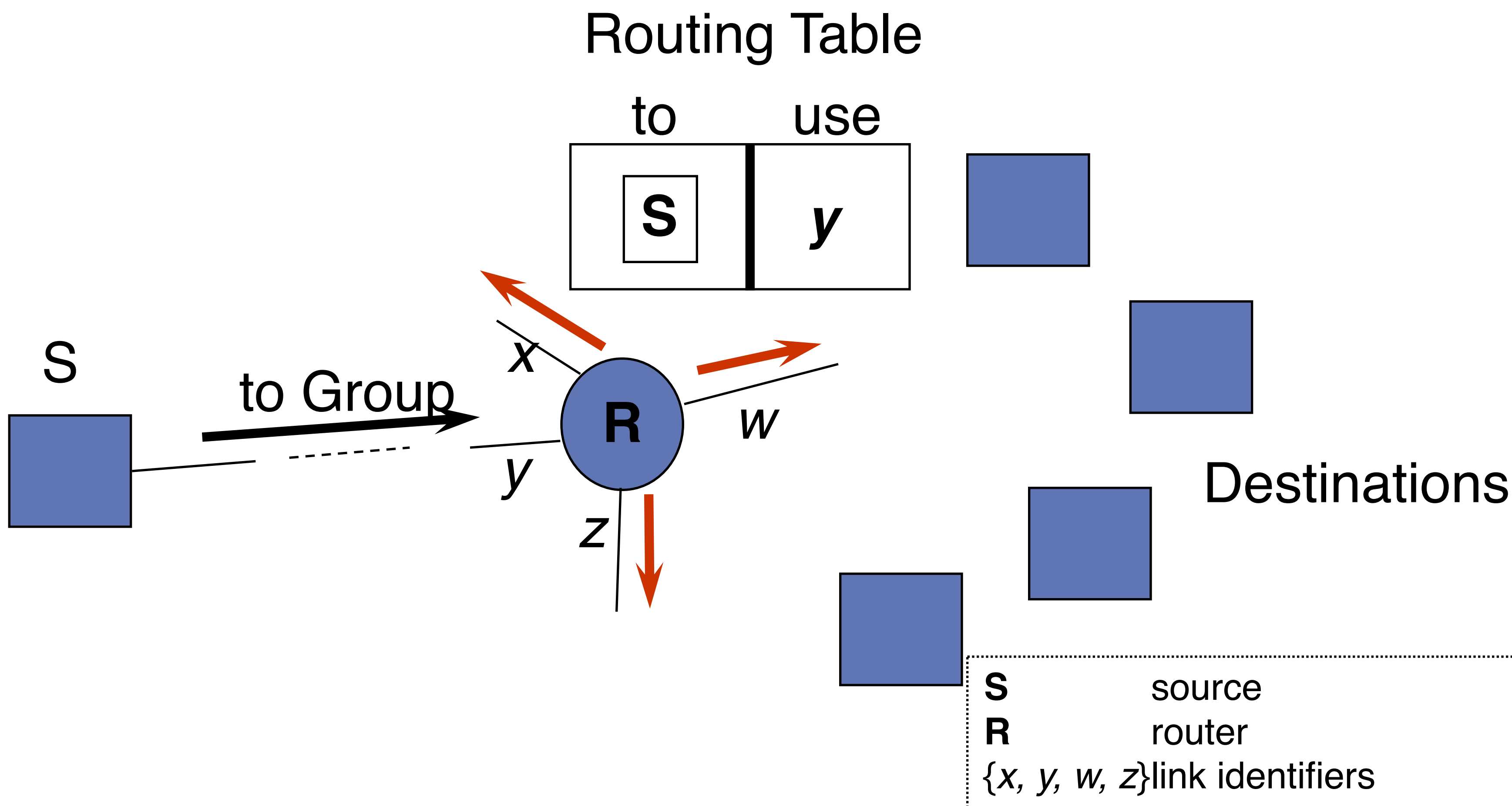
DVMRP
Reverse Path Forwarding (2/4)

- Uses shortest path tree from destinations to source (reverse tree)
- Unicast routing ensures it is the shortest
- The forwarding behavior of the routers is very simple:

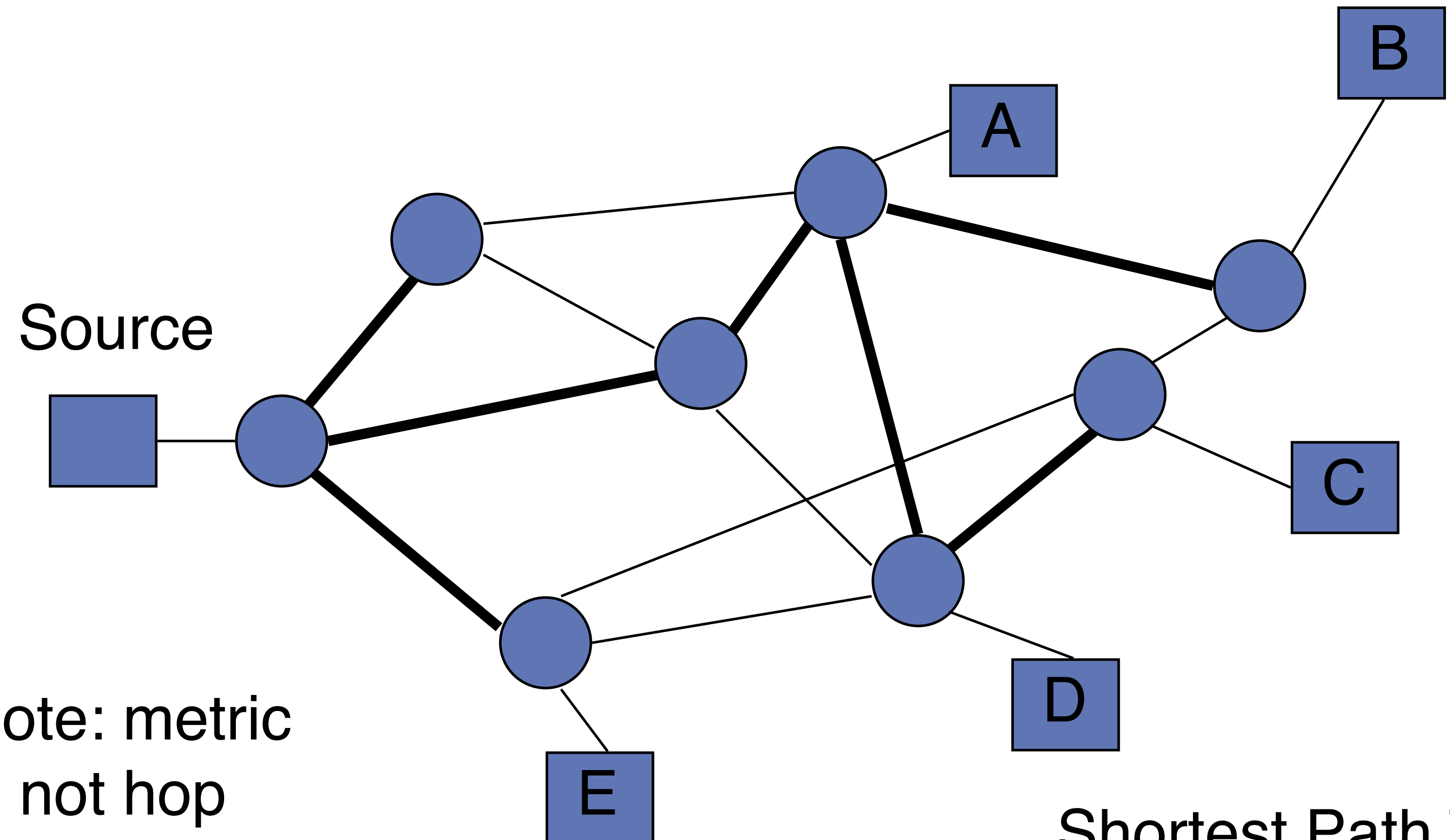
> ***if*** (mcast datagram received on incoming link on shortest path back to sender)
>
> ***then*** flood datagram onto all outgoing links
>
> ***else*** ignore datagram (drop it)

- Rule avoids flooding loops

DVMRP
Reverse Path Forwarding (3/4)



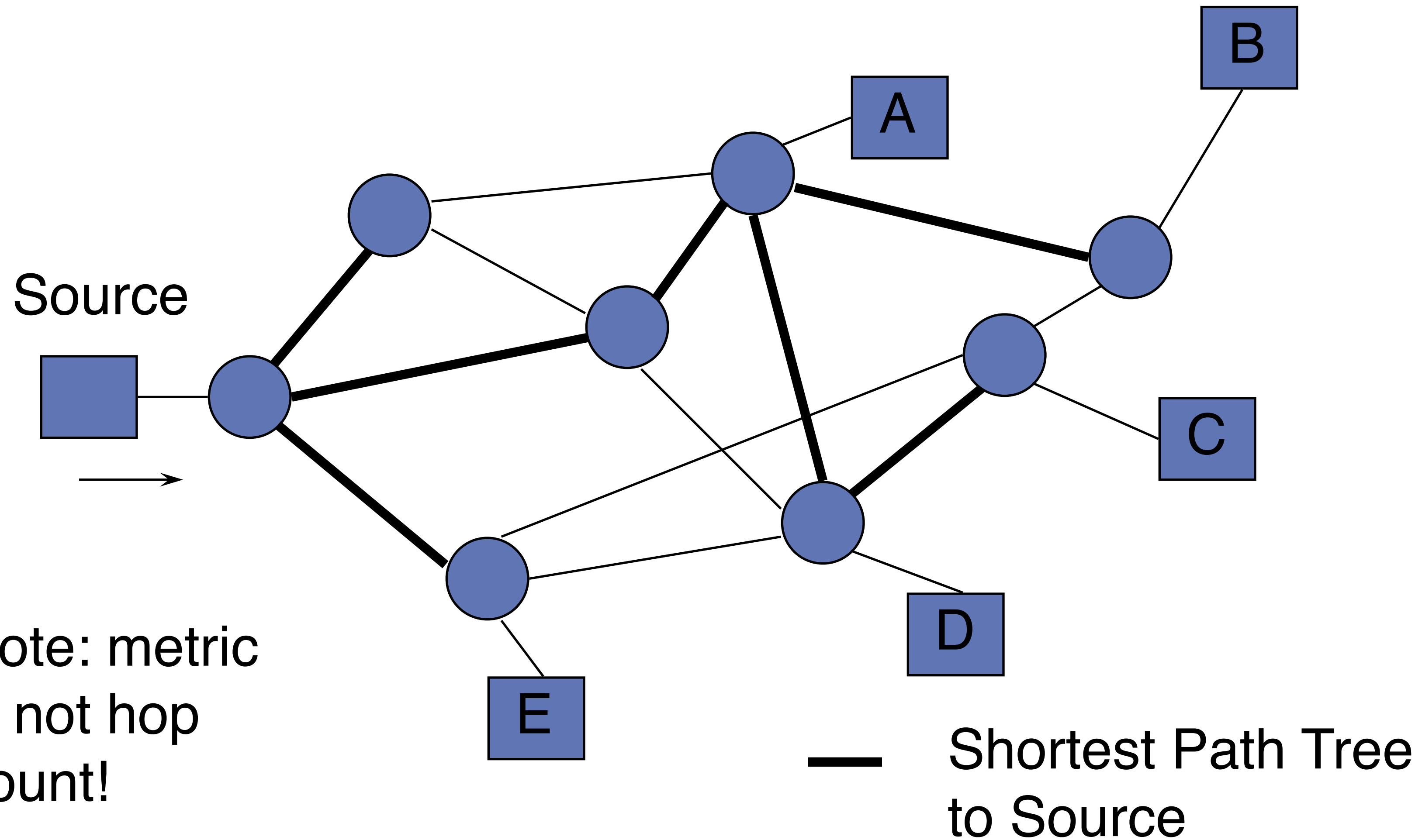
DVMRP Reverse Path Forwarding (4/4)



Note: metric is not hop count!

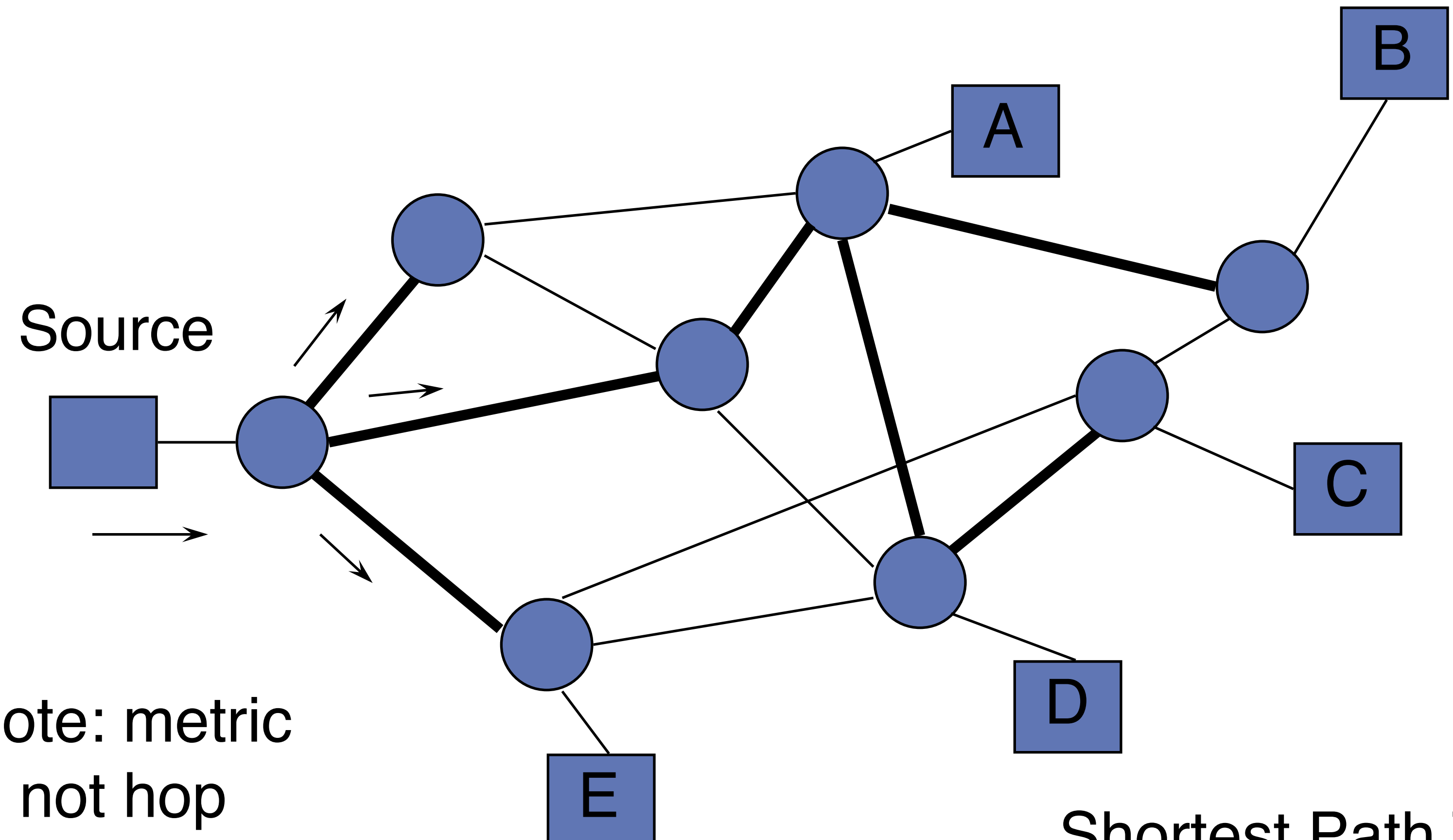
Shortest Path Tree to Source

DVMRP
Reverse Path Forwarding (4/4)



Note: metric is not hop count!

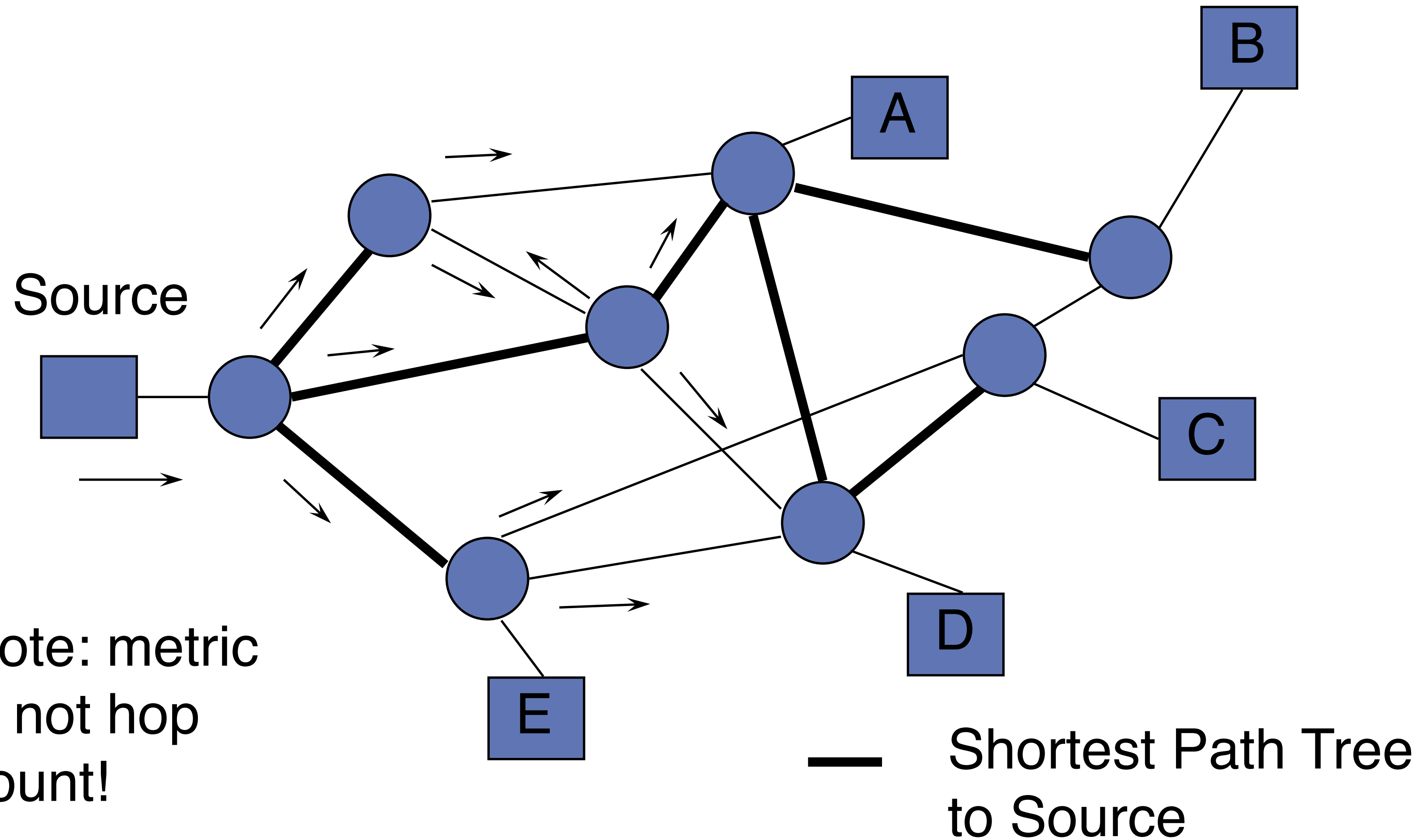
DVMRP
Reverse Path Forwarding (4/4)



Note: metric is not hop count!

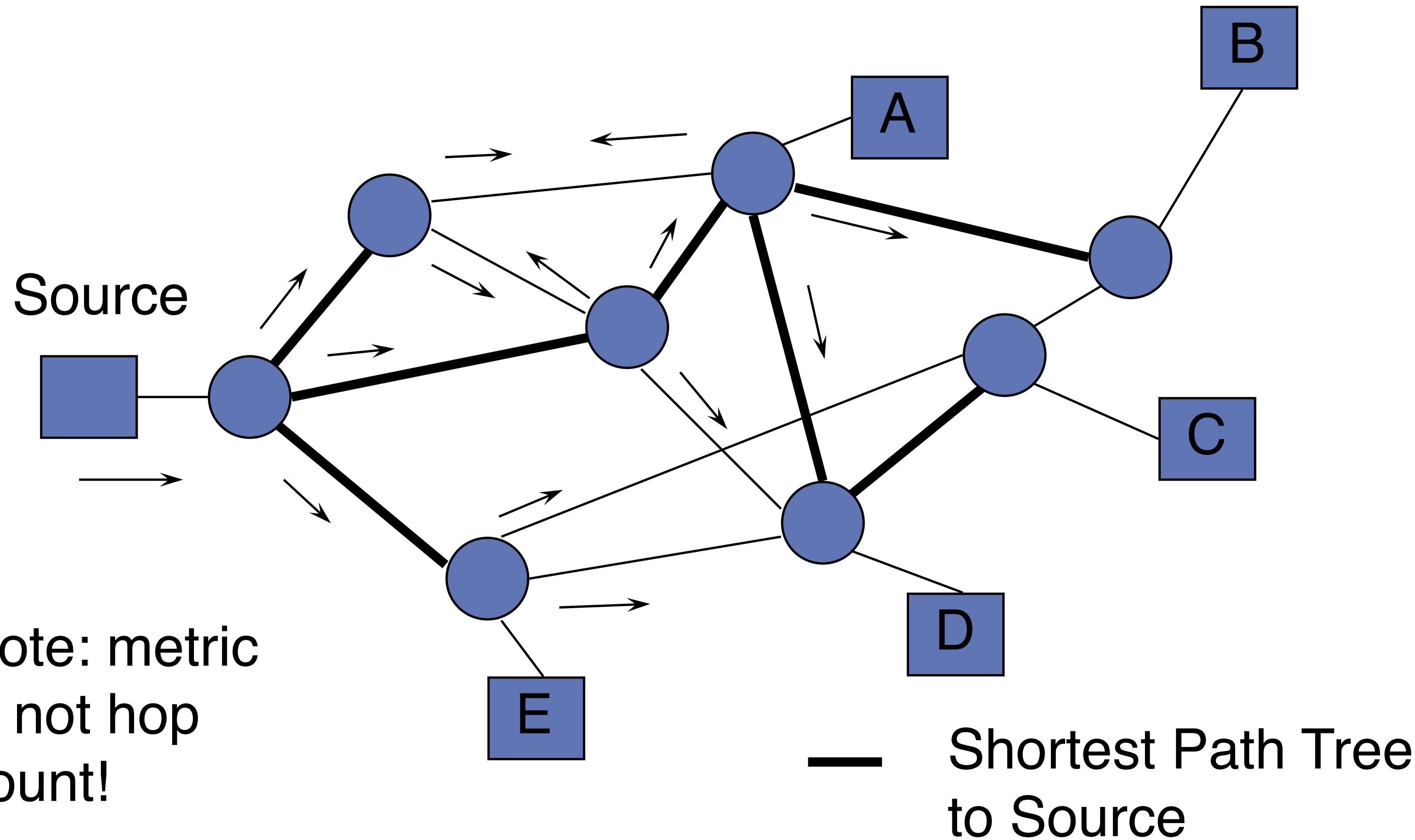
Shortest Path Tree to Source

DVMRP
Reverse Path Forwarding (4/4)



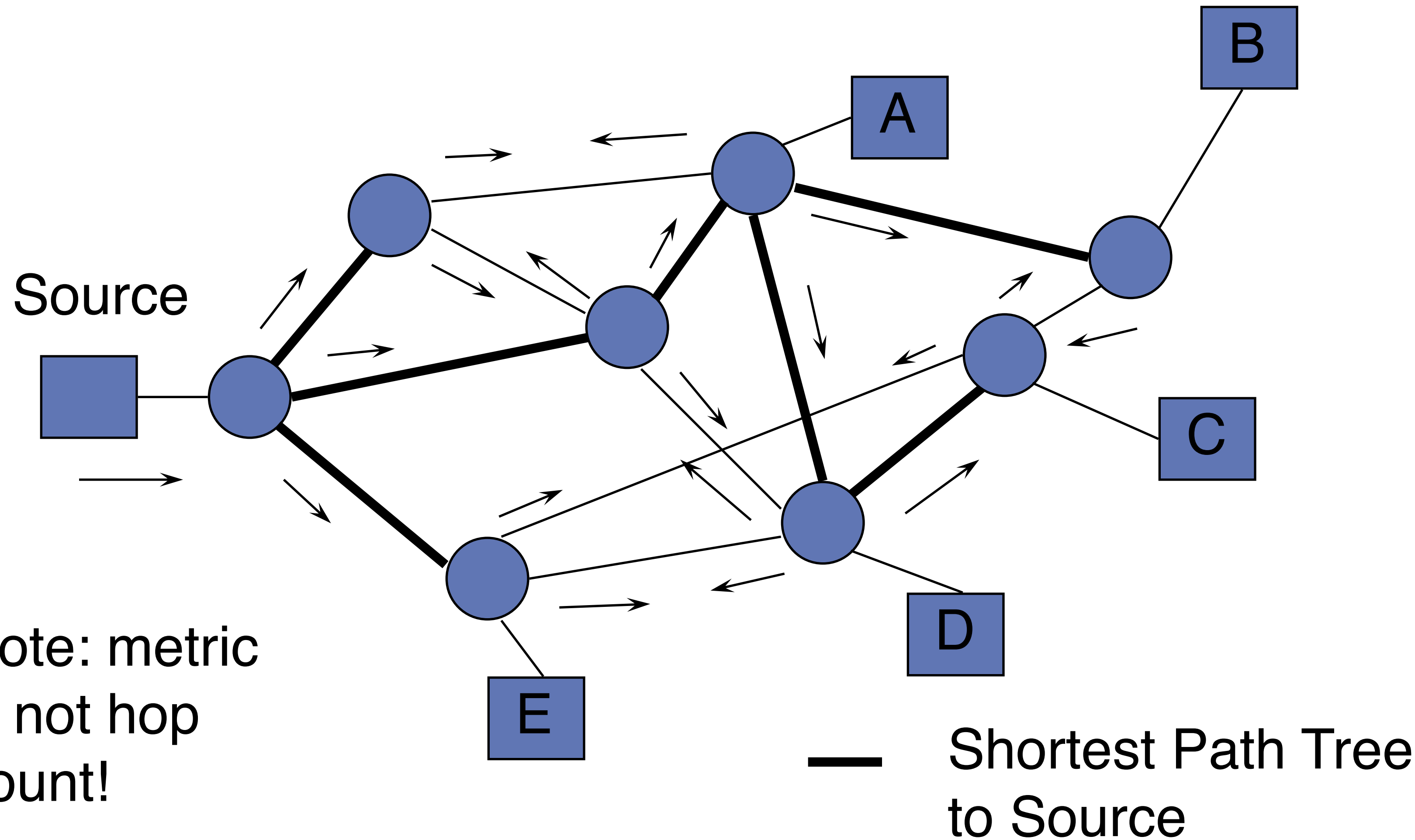
Note: metric is not hop count!

DVMRP
Reverse Path Forwarding (4/4)



Note: metric is not hop count!

DVMRP
Reverse Path Forwarding (4/4)



Note: metric is not hop count!

DVMRP

- DVMRP :
  - uses Distance Vector Routing packets for building the tree
  - prunes broadcast tree links that are not used (non-membership reports)
    - Prune messages are sent when a router receives a flooded packet addressed on a group that it does not have any member
    - allows for broadcast links

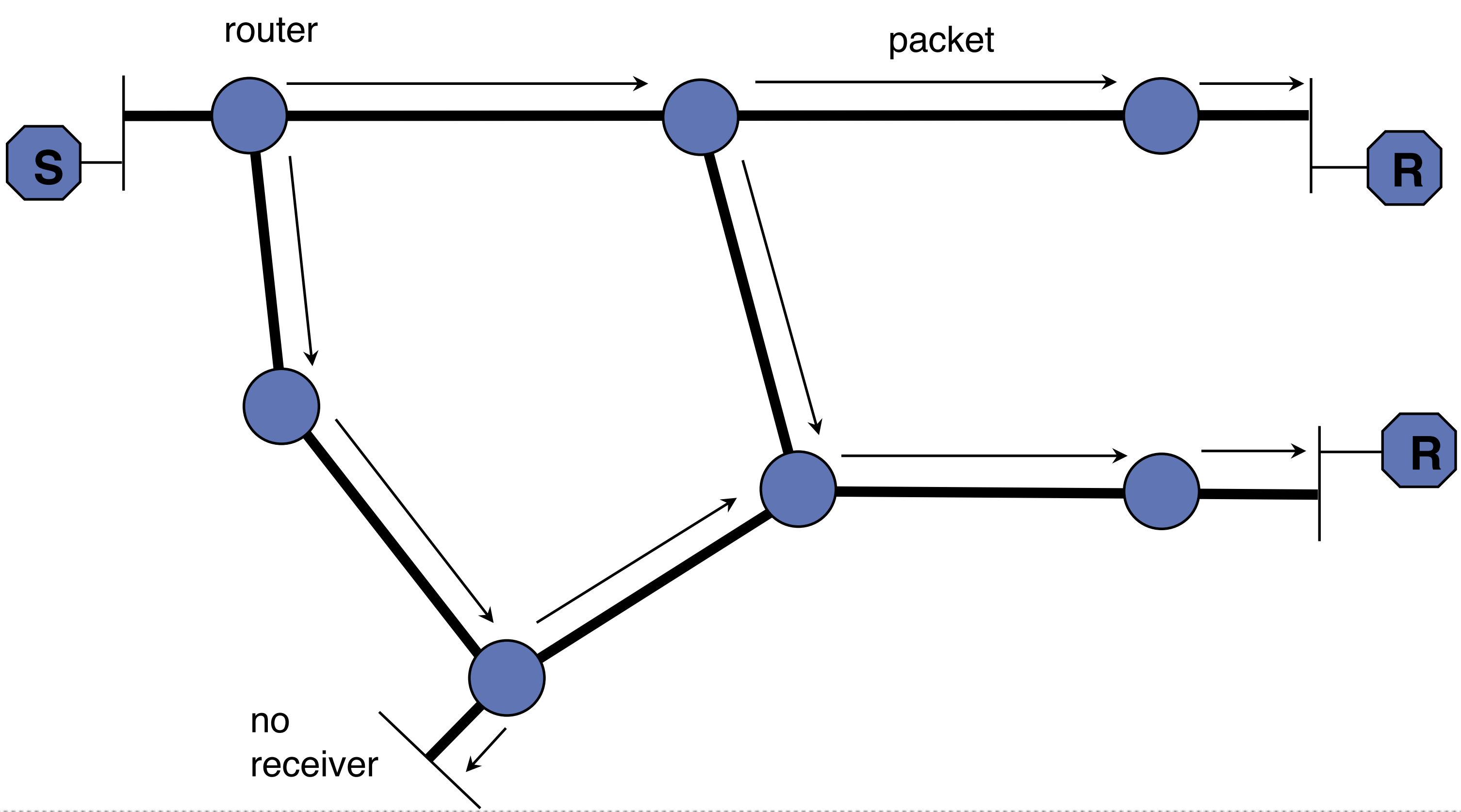
- It is similar to RIP, only used for multicast routing.
  - It is a soft-state protocol (state in the routers times out) i.e. the pruning and flooding should be repeated periodically.

DVMRP Steps :

1. Check incoming interface: discard if not on shortest path to source (reverse path forwarding)
2. Forward to all outgoing interfaces
3. Don’t forward if interface has been pruned
4. Prunes time out every minute
5. Routers may send grafts upstream
   - A graft is a subscription message sent upstream by a router to be added to the multicast group forwarding tree.

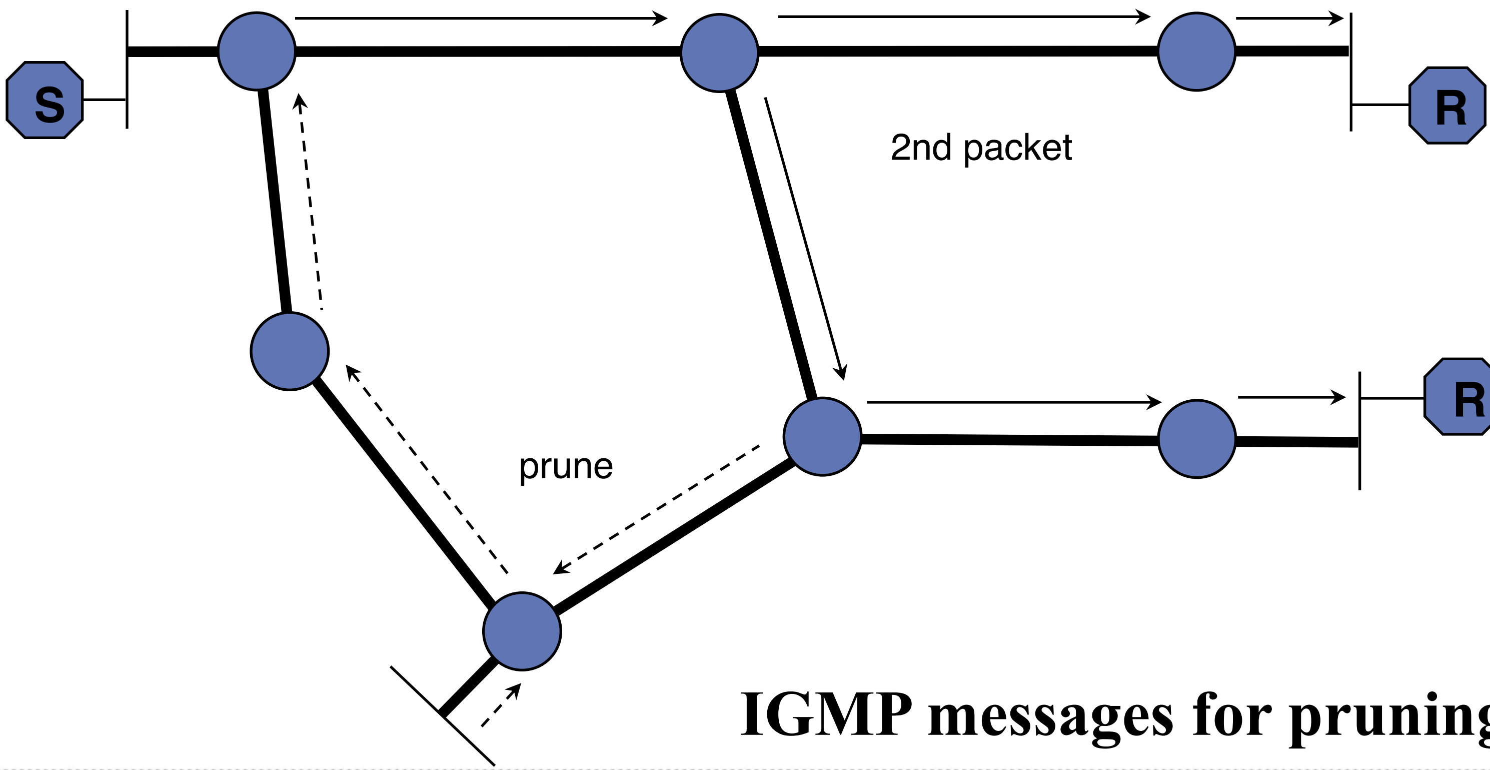
DVMRP

- Basic idea: Flood and Prune



DVMRP

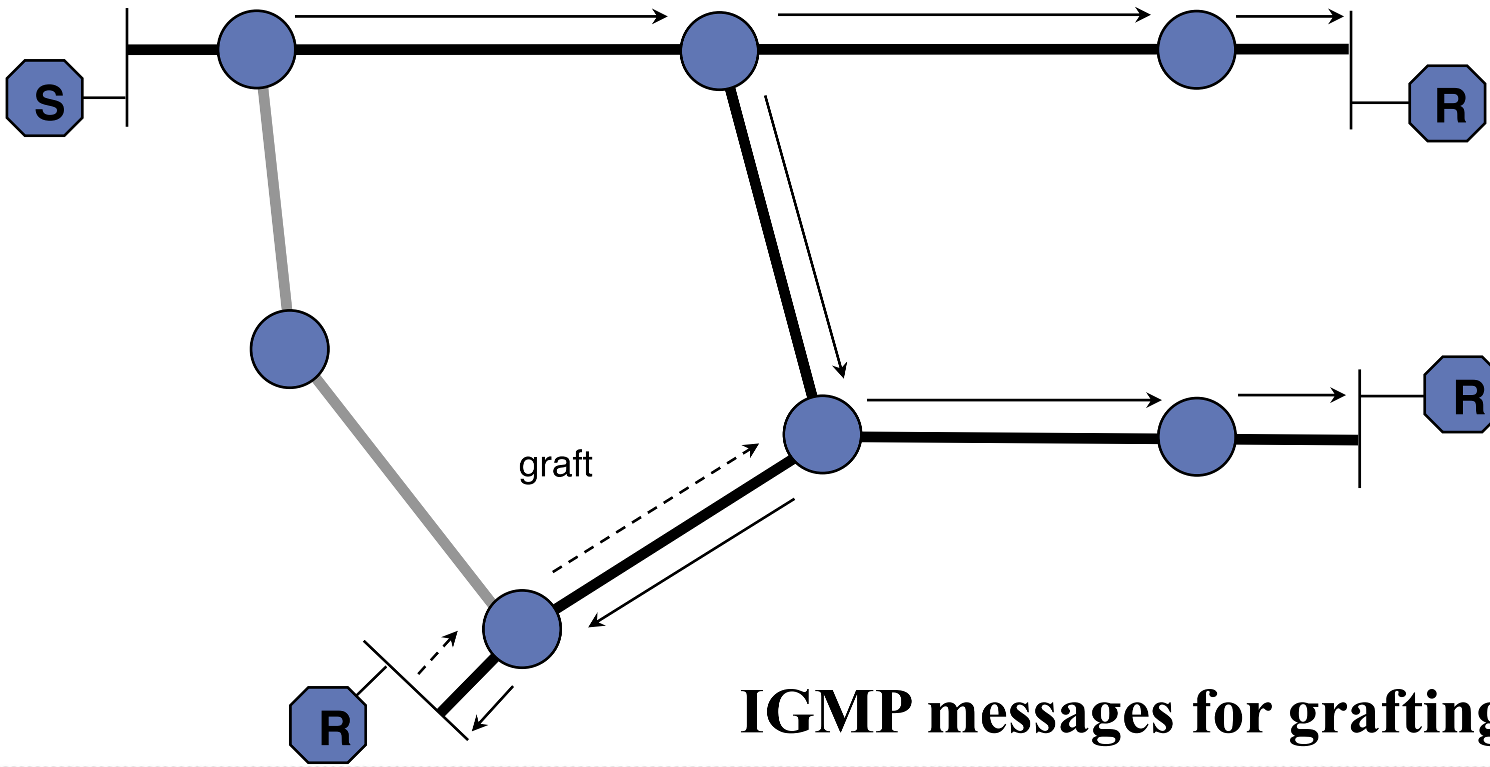
- Prune branches where no members and branches not on shortest paths



IGMP messages for pruning

DVMRP

- Add new user via grafting; departure via pruning



IGMP messages for grafting

Protocol Independent Multicast (PIM)

- DVMRP & MOSPF are good for dense group membership, but
  - Need shared/source-based tree flexibility
  - Independence from unicast routing would be good too.

- Two PIM modes:
  - PIM-DM : Protocol Independent Multicast - Dense Mode
  - PIM-SM: Protocol Independent Multicast-Sparse Mode

PIM Design Goals

- Cater for sparse mode regions
  - A region is classified as sparse if:
    - the number of networks with members <<< total number of networks in the region, and
    - group members are widely distributed, and
    - flood & prune mechanisms have high overheads (low efficiency).

- Low-latency (minimise end-to-end delay)

- Independence from the unicast routing protocol used. If used as inter-domain multicast routing protocol, it should interoperate with other domain multicast routing protocols, if deployed.

- Robustness: avoid single point of failure, and provide graceful adaptation to network topology changes.

- Scalability: overhead should be capped (percentage-wise) to the total bandwidth used by the group, independently of the size of the group.

University of BRISTOL

PIM-Dense Mode (PIM-DM)

- Independent from underlying unicast routing
- Approx. DVMRP – the main difference is independence from any specific unicast routing protocol
- Slight efficiency cost
- Contains protocol mechanisms to:
  - detect leaf routers
  - avoid packet duplicates

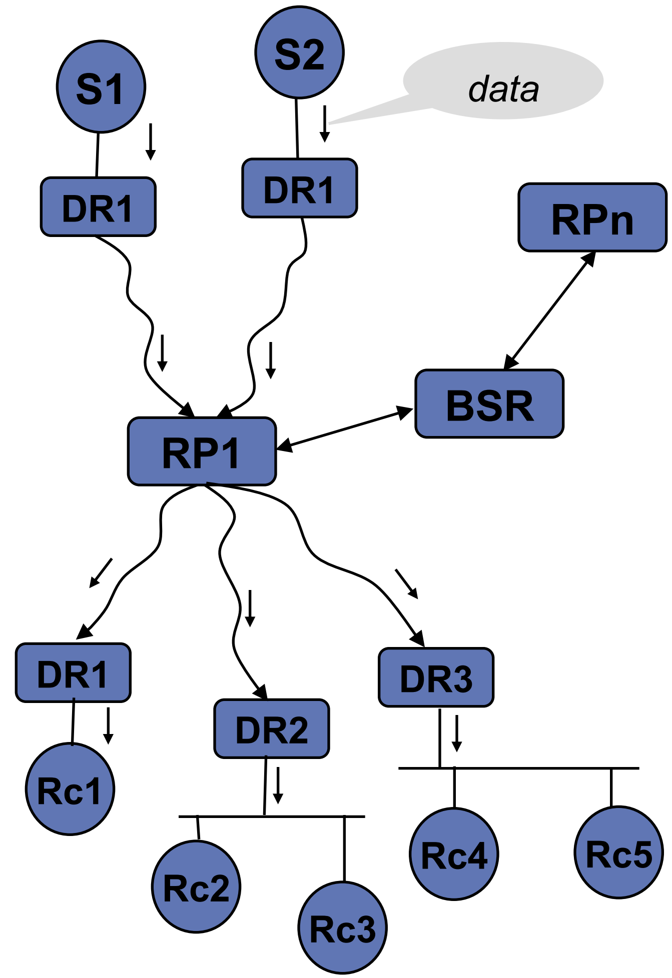
PIM Sparse Mode (PIM-SM) Components

Rendezvous Point - RP

- It is a router; one exists per each multicast group
- Rendezvous point :receivers meet sources

Designated Router - DR

- A router that is used by each sender/receiver in a multicast group:
- It is directly connected to member hosts
- It is used to create routing table entries and send JOIN/PRUNE messages to the RP



PIM Sparse Mode (PIM-SM) Components

Last Hop Router - LHR

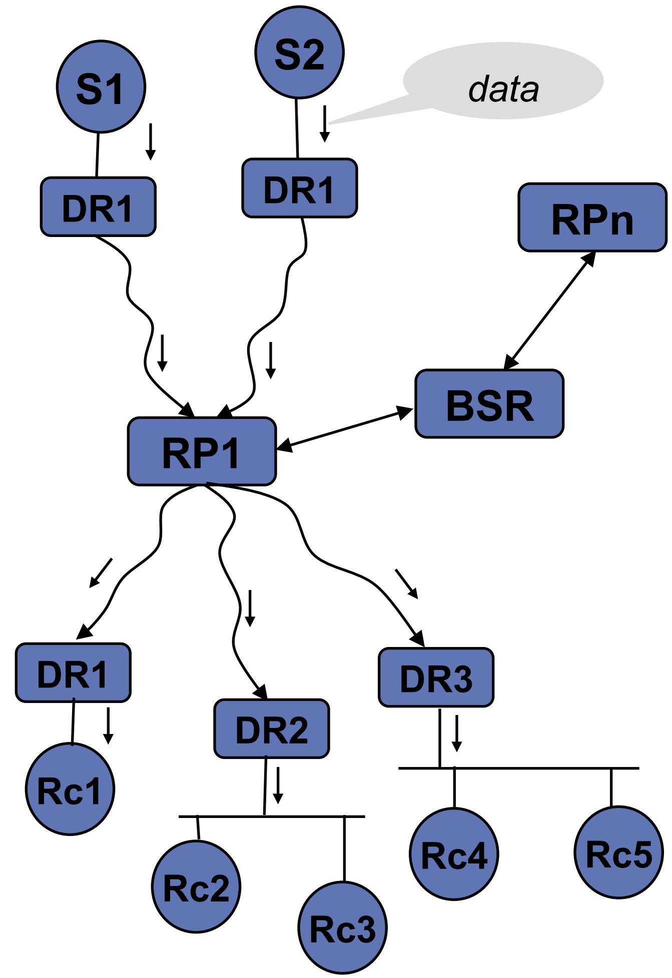
- A router directly connected to a receiver, that forwards multicast packets to it (often the same as DR, but not always the case).

Boot Strap Router - BSR

- The router that builds the set of RPs and distributes the set in a PIM domain.

Reception through RP connection ➔ shared tree

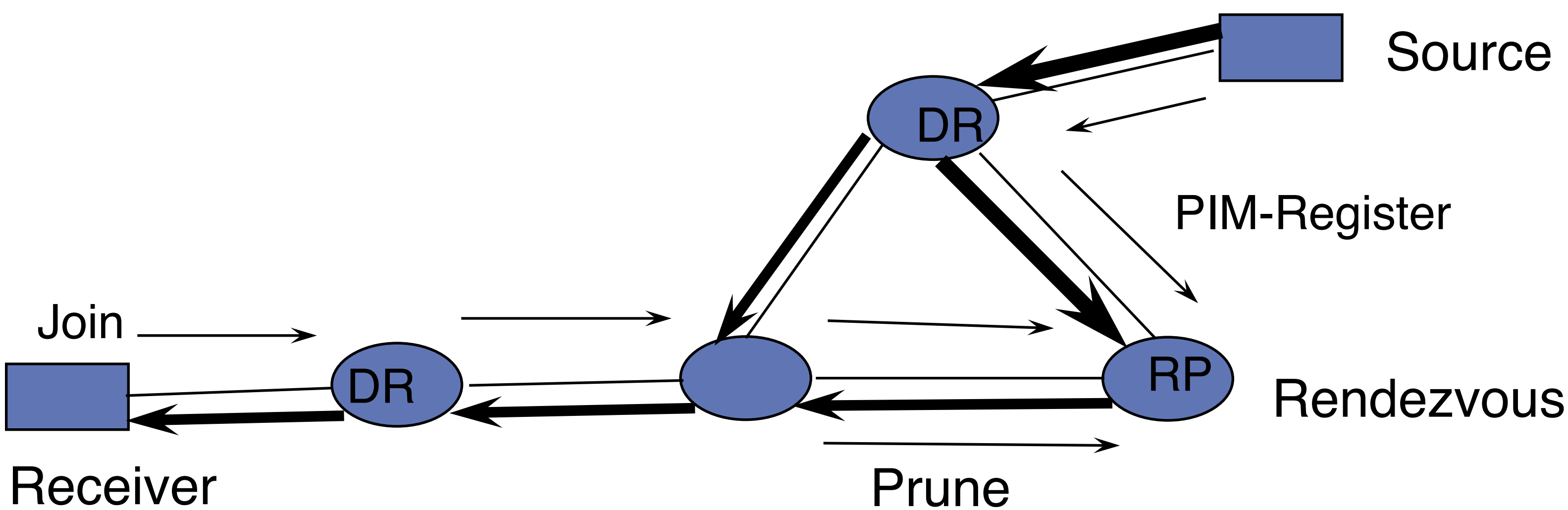
Establish path to source ➔ source-based tree



PIM - Sparse Mode

- **Receiver *Join***
  - Receiver informs DR (using IGMP)
  - DR computes RP for the group (it knows the multicast group and the RP set)
  - DR sends *PIM-Join/Prune* message towards RP.
  - Intermediate routers process the message, creating the multicast branch tree from RP to receiver.

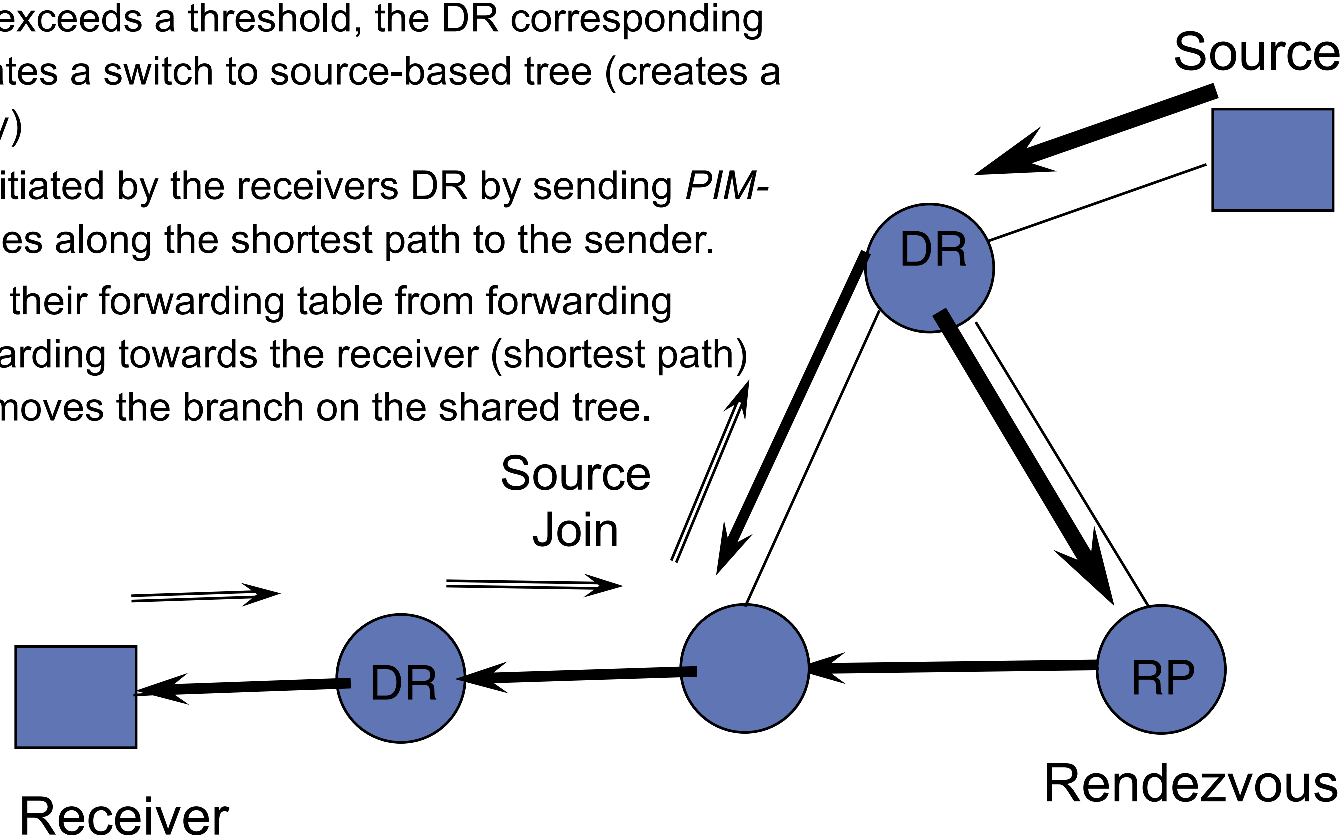
- **Source *Join***
  - Source uses IGMP to inform the DR about their interest on the mcast group.
  - DR encapsulates data from source in a *PIM-Register* packet and *unicasts* it to the RP.
  - RP decapsulates the packet and *multicasts* it along the shared tree rooted at RP.



PIM - Sparse Mode (4/4)

Switching from source-based to shortest-path tree

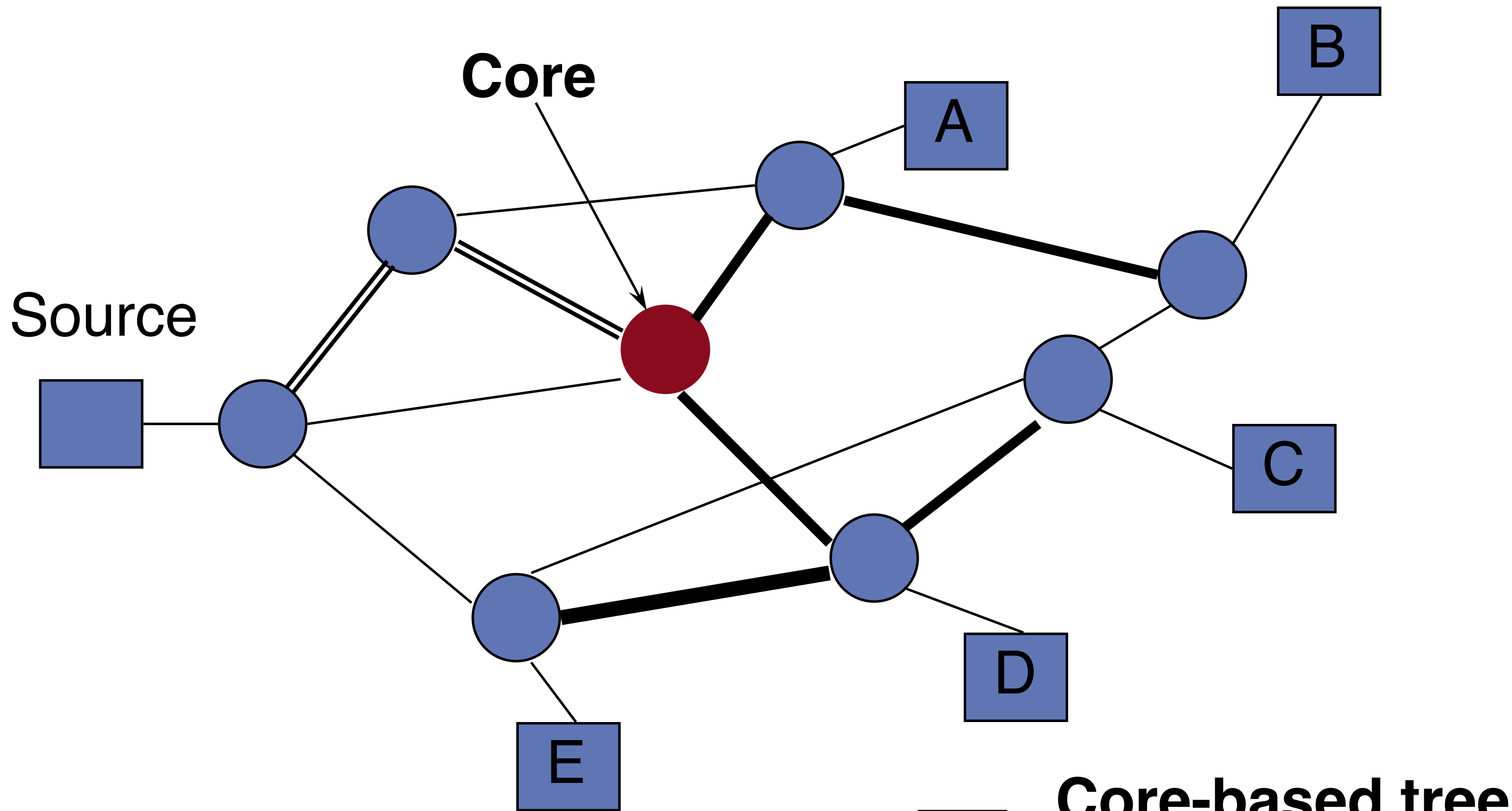
- PIM *always* starts by creating an RP-rooted shared tree.
- If sender data rate exceeds a threshold, the DR corresponding to the receiver initiates a switch to source-based tree (creates a specific (S, G) entry)
- A reverse path is initiated by the receivers DR by sending *PIM-Join/Prune* messages along the shortest path to the sender.
- Routers that switch their forwarding table from forwarding towards RP to forwarding towards the receiver (shortest path) notify RP, which removes the branch on the shared tree.



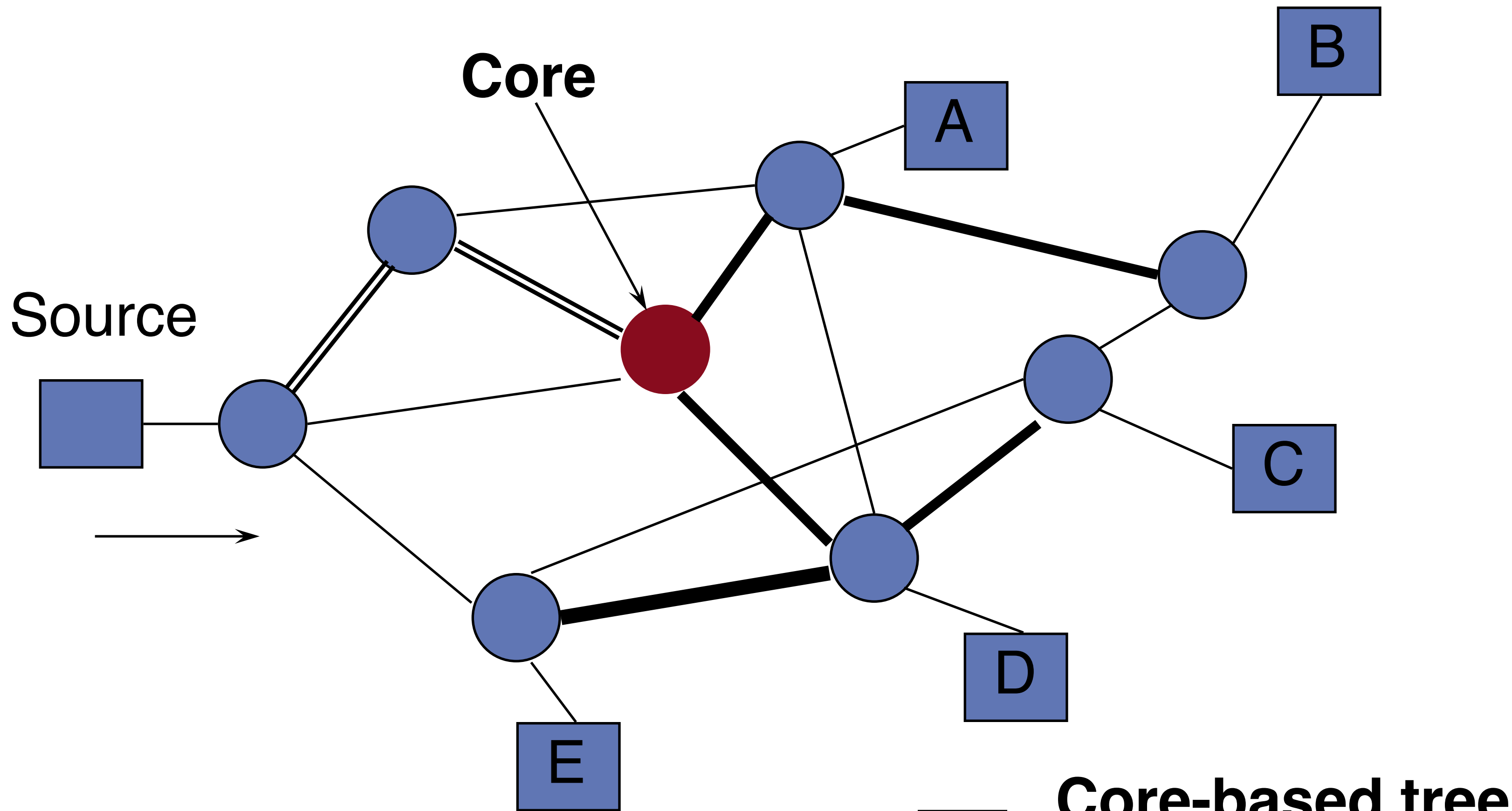
Core-Based Trees

- A shared-tree protocol
- One node on shared tree is Core
- Sender sends to core
- Core forwards over multicast tree

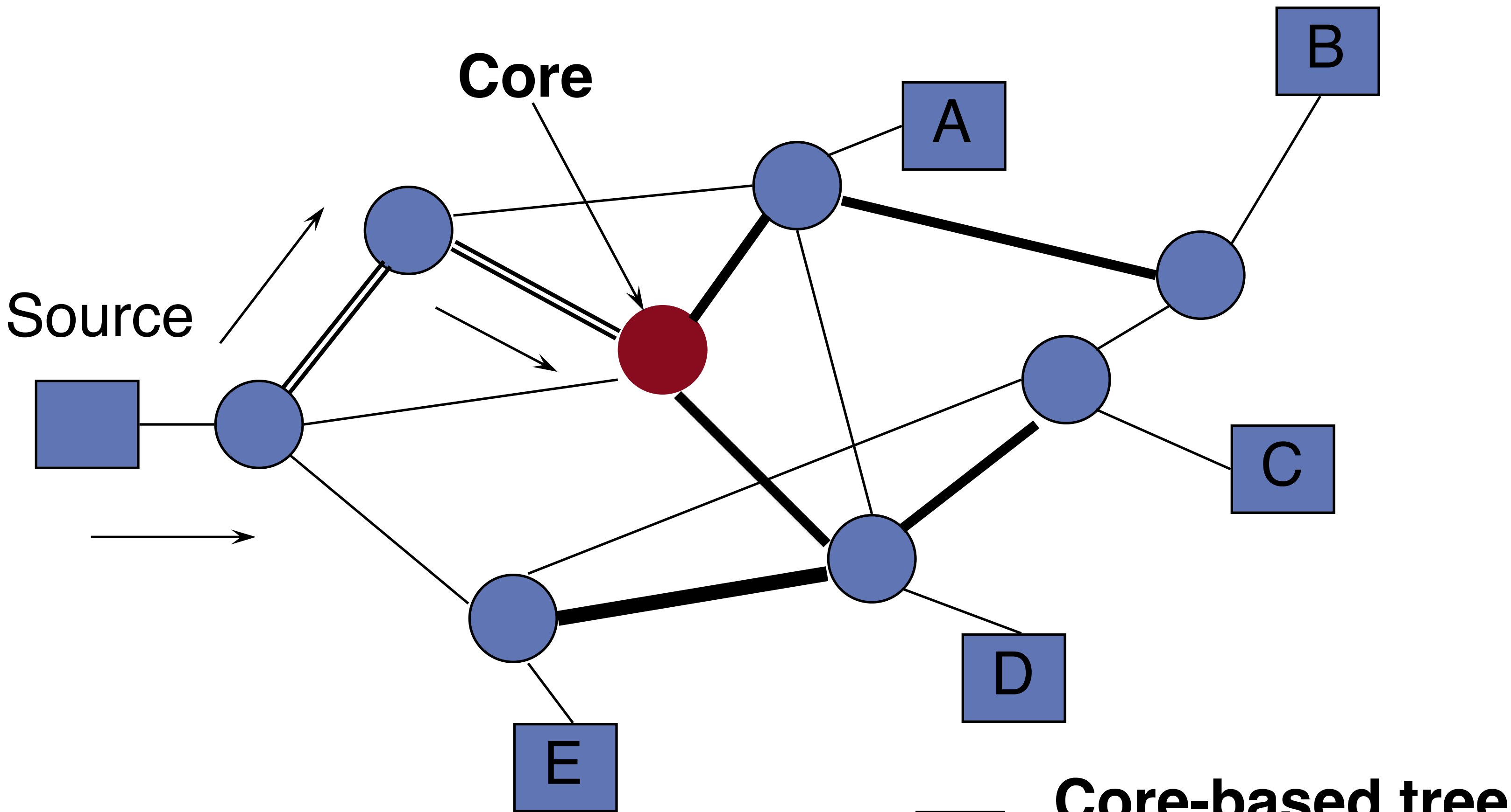
Core-Based Trees



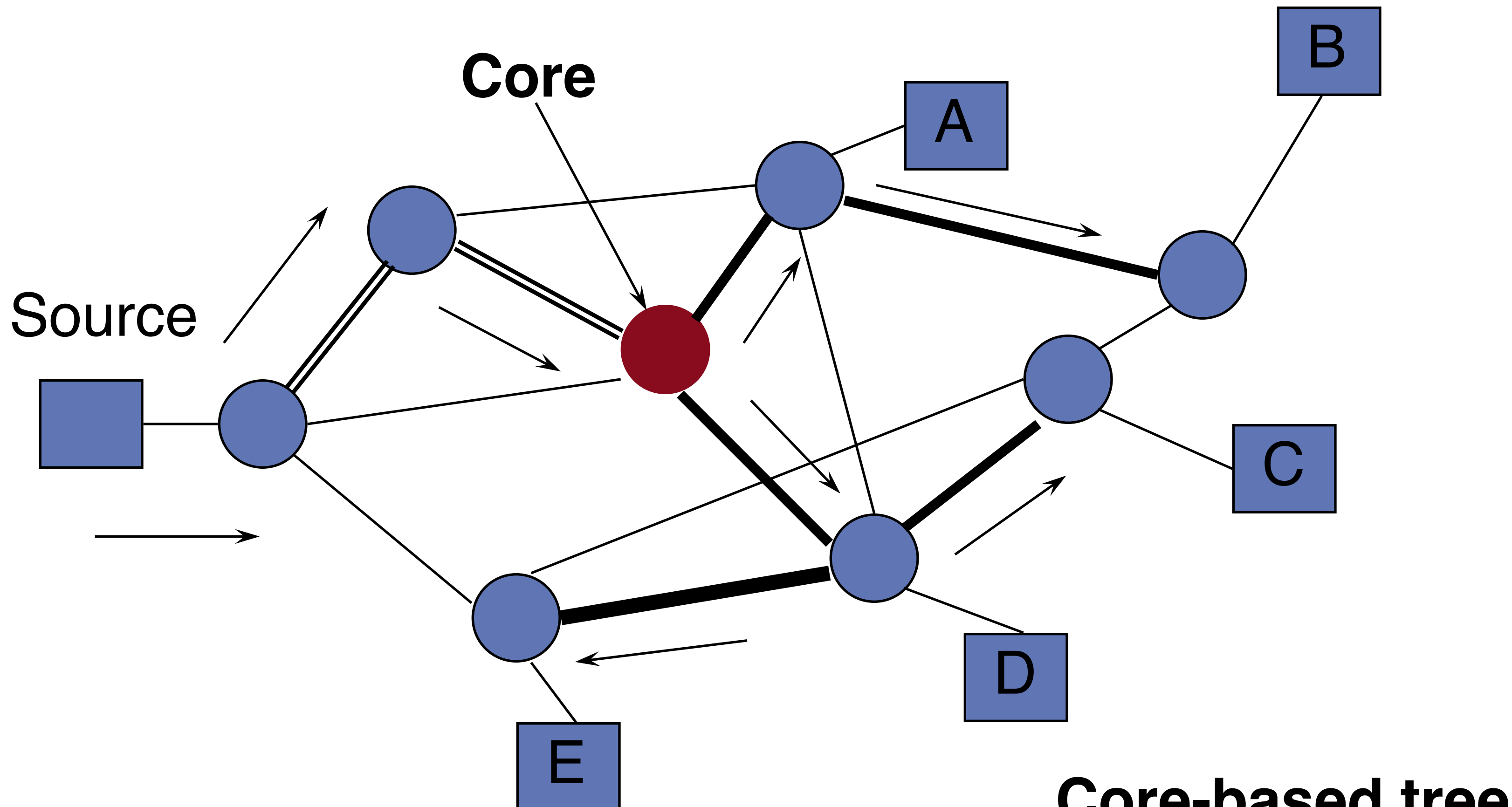
Core-Based Trees



Core-Based Trees



Core-Based Trees



PIM and CBT Issues

- Unidirectional vs. Bidirectional Shared Trees
- Core/RP placement and selection
- Multiple Cores/RPs (RP ::= Randezvous Point)
- Dynamic Cores/RPs

Conclusions

- Multicast routing protocols
  - structure and overhead
- Purpose and benefits of multicast
- Many approaches
- Only discussed the simple issues!

References

- PIM: S. E. Deering et al., “The PIM Architecture for Wide-Area Multicast Routing” – IEEE/ACM Transactions on Networking, Vol.4, No.2, pp 153-162, April 1996.
- MOSPF:
  - J. Moy, “Multicast routing extensions for OSPF”, Communications of the ACM, Vol.37, No.8, pp61-66, August 1994.
  - MOSPF: J. Moy, “Multicast Extensions to OSPF”, RFC 15984, March 1994.
- Multicasting book: S. Paul, “Multicasting on the Internet and its applications”, Kluwer Academic Publishers, 1998.

Next Up

1. Introduction
2. Internet Routing and Switching
3. IP Multicast
4. Networking for Realtime Applications
5. Routing in Wireless Networks
6. Quality of Service