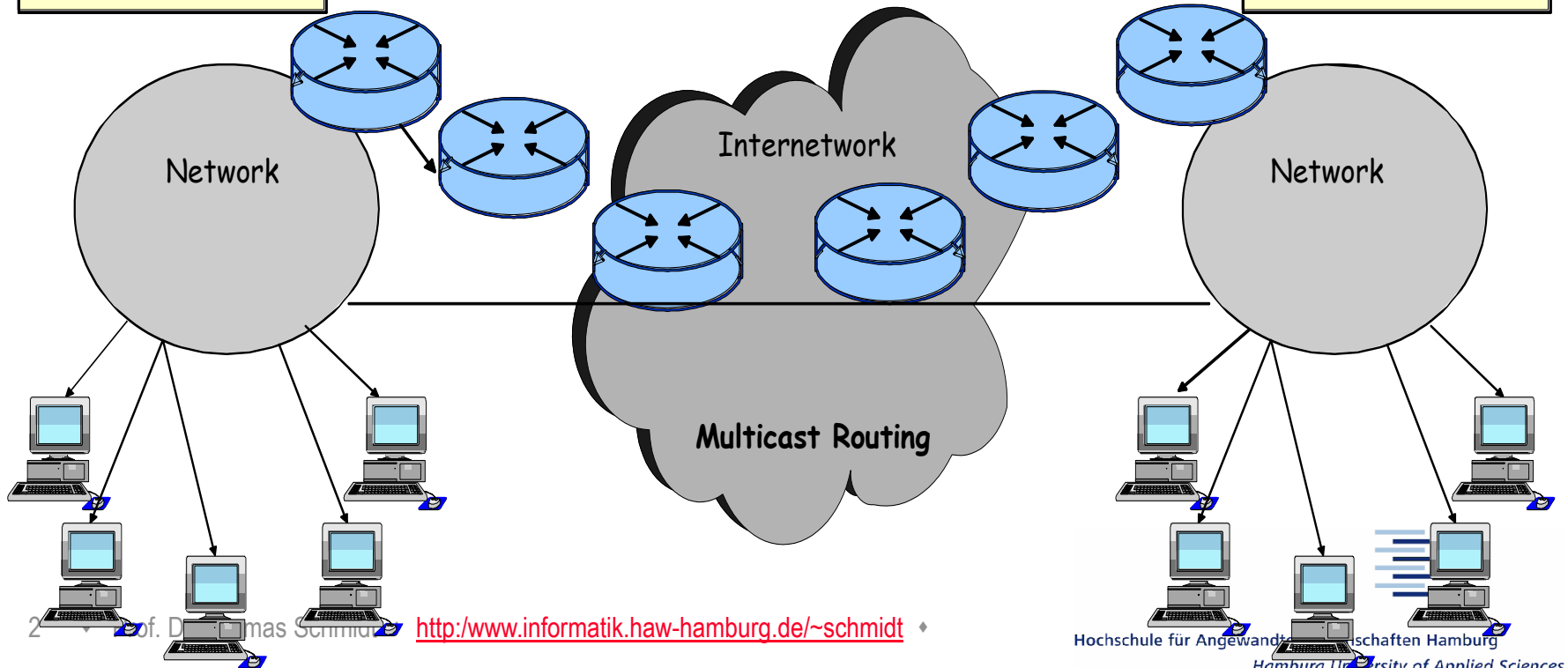
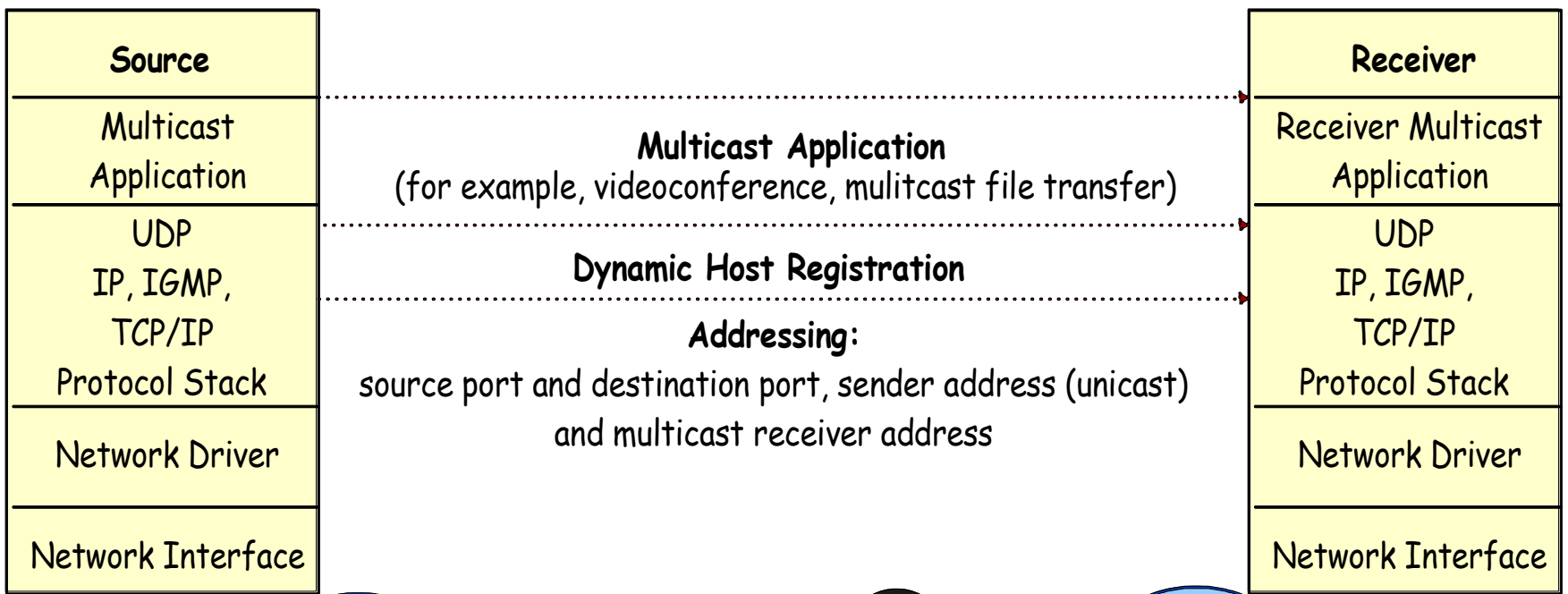


# Multicast Routing

- The Problem of IP Mcast Routing
- Routing Algorithms
- ASM Routing Protocols
- SSM Routing
- Properties of Multicast Distribution Trees
- Efficiency versus Deployment Complexity

Some graphics originate (in part) from cisco





# Multicast Routing

## Unicast IP-Routing

- ▶ Guides IP-Datagrams stepwise to *one* receiver
  - ▶ Routing decision on *where* to forward packet to
  - ▶ Solely based on *destination* address
  - ▶ Adapts to Router topology, *never* to IP-Packets
- ⇒ Multicast turns Routing upside down



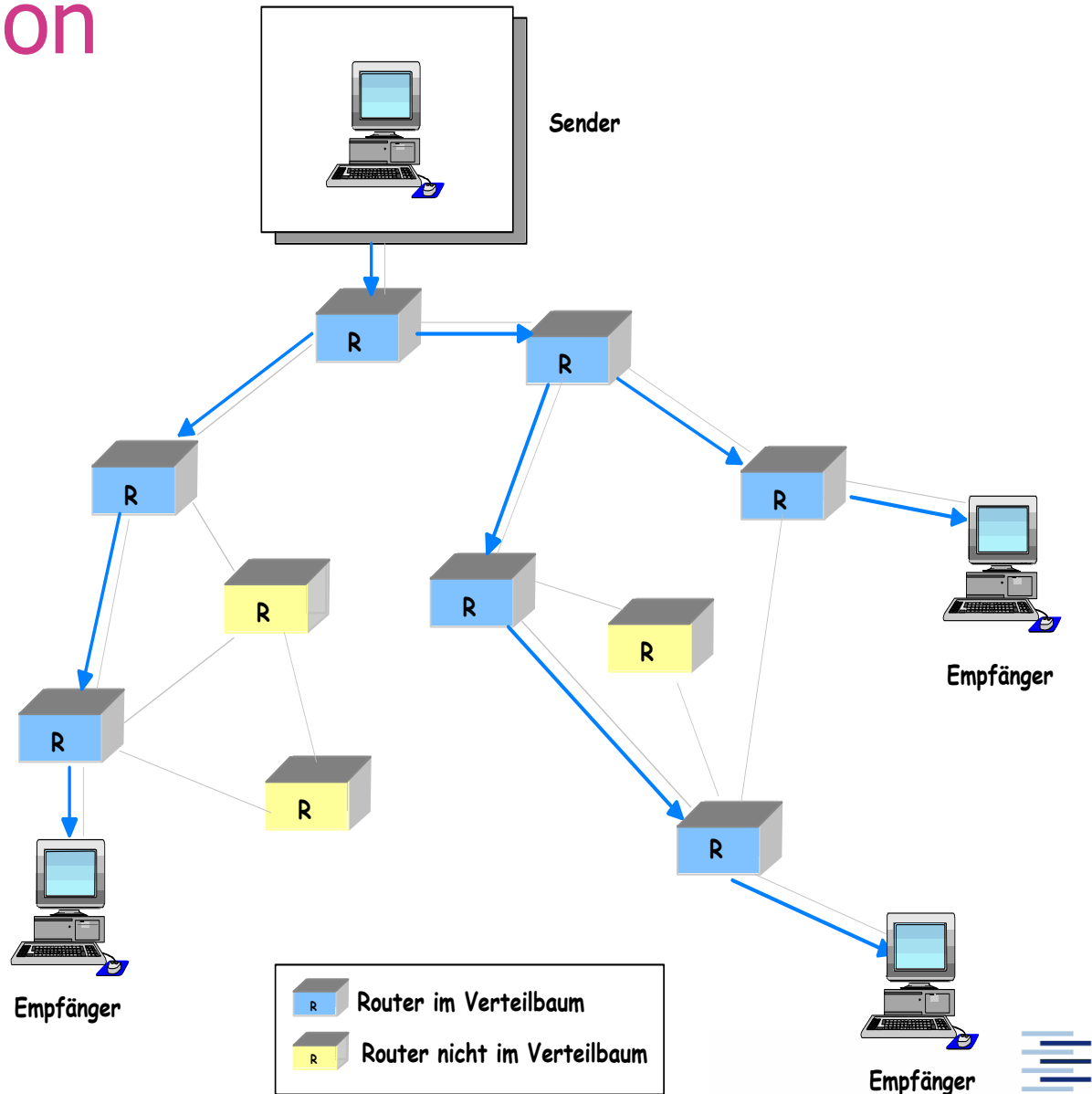
# Multicast Routing (2)

IP Multicast - Routing is receiver initiated:

- Guides mcast-Datagrams according to a distribution tree
- Duplicates Datagrams
- Based on *Source* address
- Changes according to group dynamics
- Uses ‚Reverse‘ Paths



# Distribution Tree



# Receiver Initiated Routing

- ▶ Group initiation by sender results in distribution tree
- ▶ Two types of distribution trees:
  - ▶ **Source Specific Tree** originating at sender  $(S,G)$  or
  - ▶ **Shared Tree** originating at Rendezvous Point  $(*,G)$   
(serving a group of senders)
- ▶ Calculation of Routing Information stimulated by receiver
  - ▶ A receiver adds/removes branches to/from distribution tree
- ▶ Unicast routing tables usable (requires symmetric routing!)
- ▶ Forwarding Algorithm: Reverse Path Forwarding



# Reverse Path Forwarding (RPF)

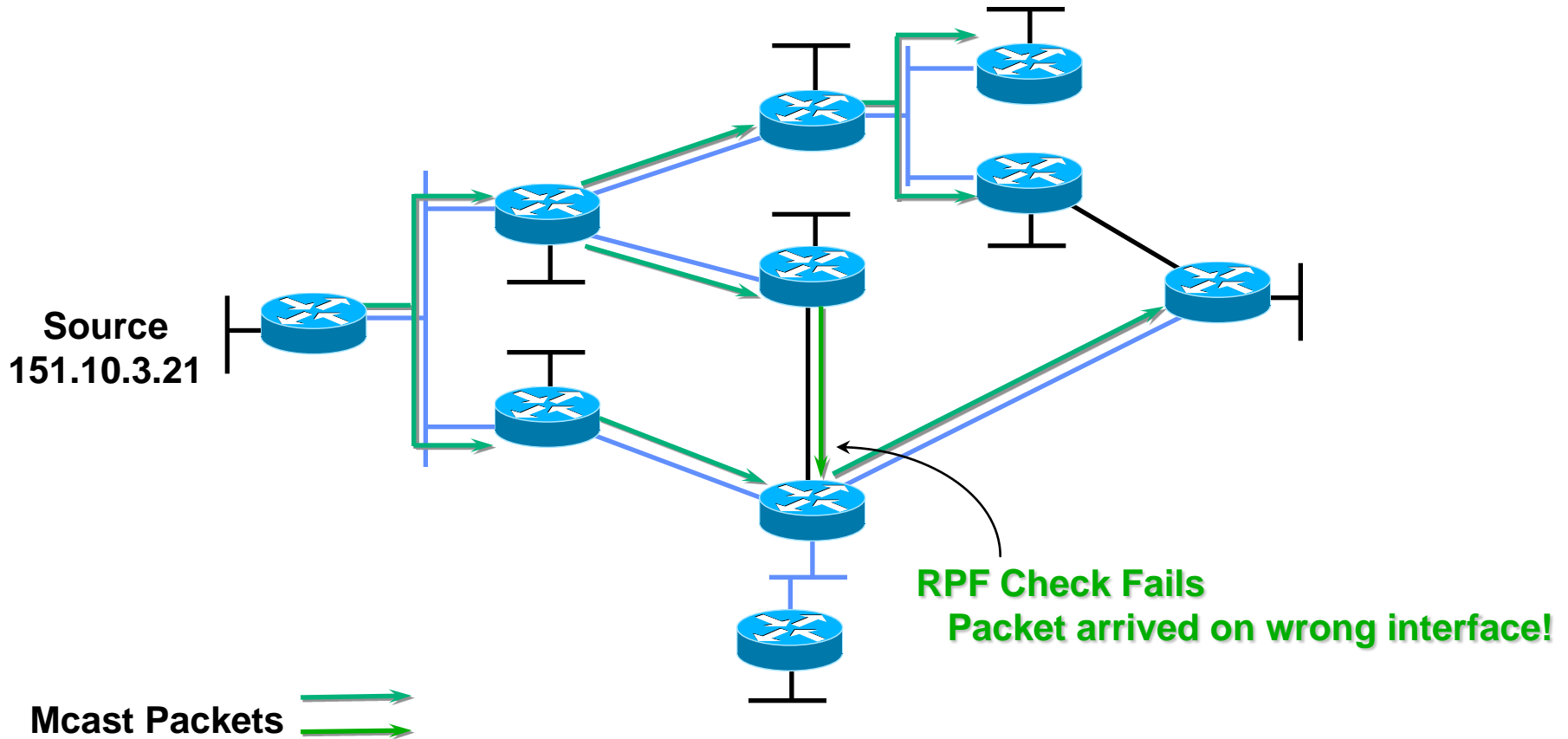
A Router forwards a packet only, if it was received on the proper route to source.

RPF Check:

- ▶ active routing table searched for *source*-address
- ▶ Packet transmitted, if received on the interface foreseen as source address destination
- ▶ Packet discarded otherwise



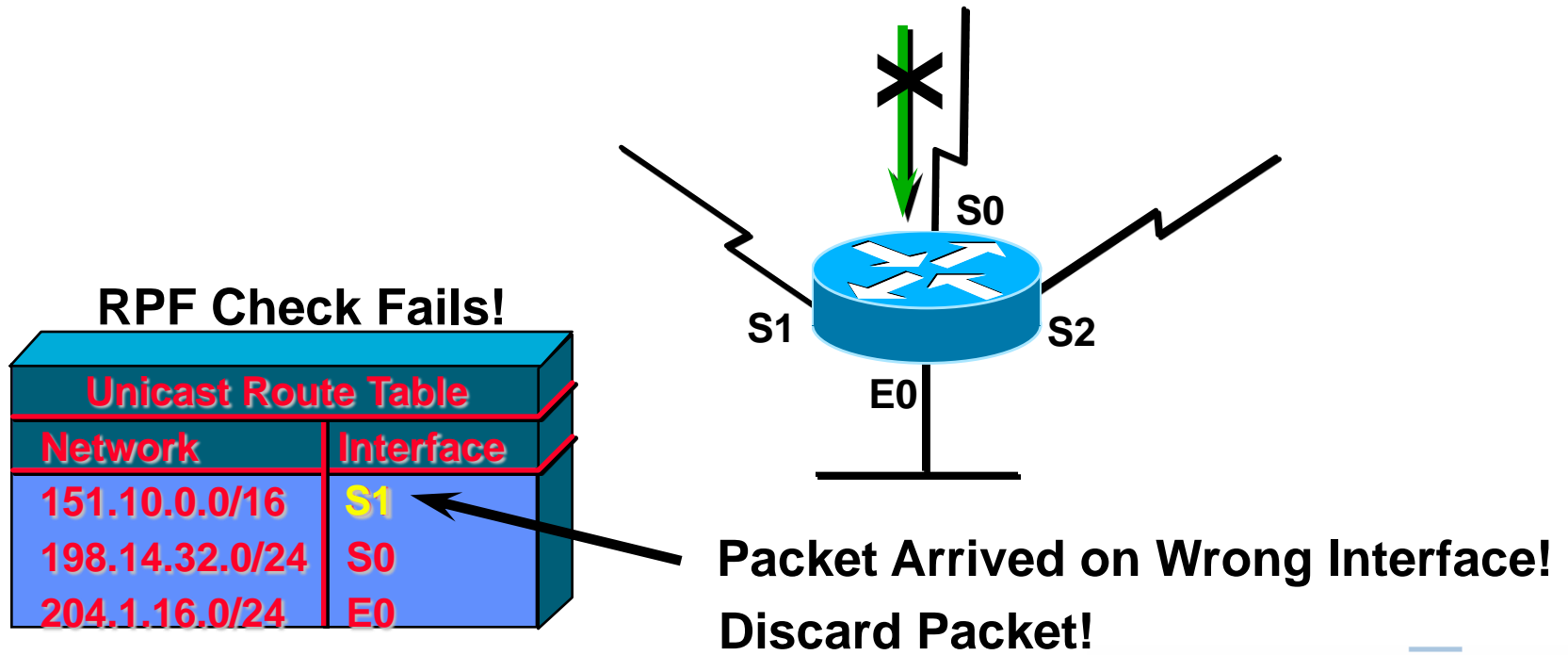
# RPF Check





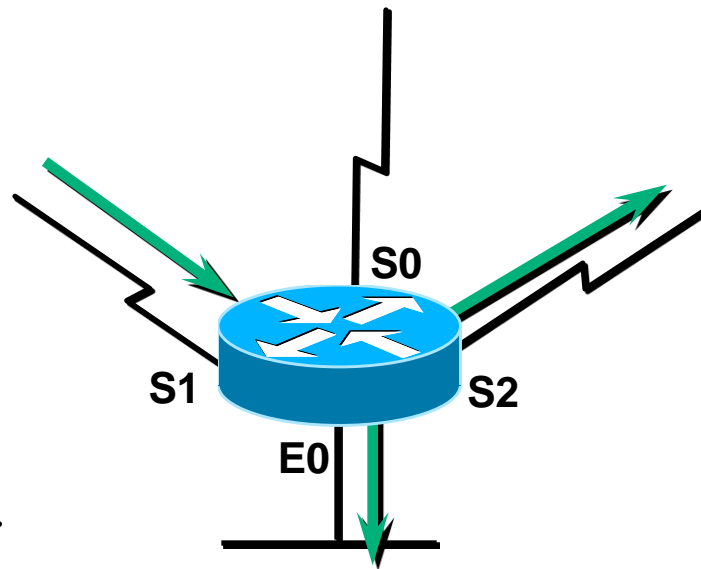
# RPF Check: Failure

Multicast Packet from  
Source 151.10.3.21



# RPF Check: Success

Multicast Packet from  
Source 151.10.3.21



RPF Check Succeeds!

Unicast Route Table	
Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Packet Arrived on Correct Interface!  
Forward out all outgoing interfaces.  
(i. e. down the distribution tree)



# Any Source Multicast (ASM)

How to construct distribution tree to reach all receivers?

Two classes of algorithms:

## Dense Mode

- Push Model
- Flooding and Pruning

## Sparse Mode

- Pull Model
- Directional traffic only
- Rendezvous Points



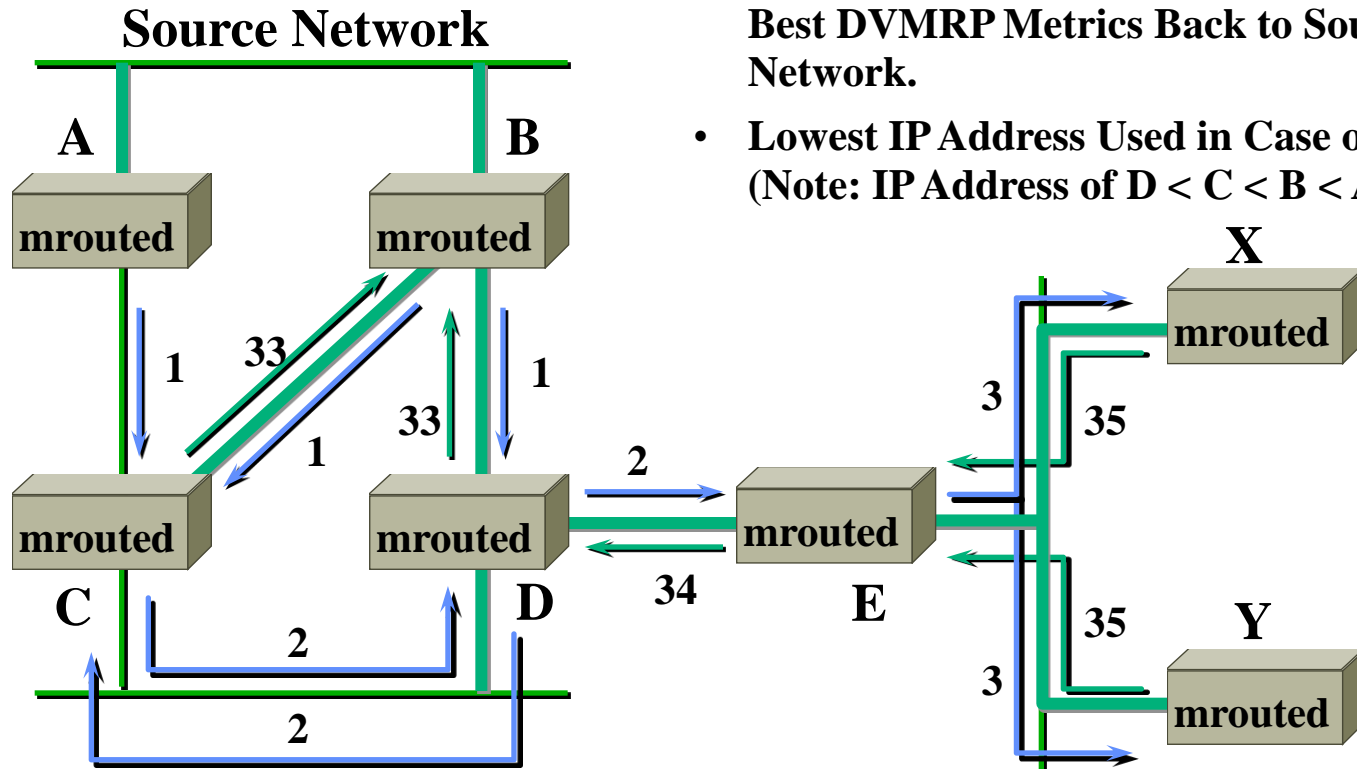
# Distance Vector Multicast Routing Protocol (DVMRP)

- ▶ Oldest IP mcast Routing Protocol (v1: RFC 1075, Deering)
- ▶ Destination based Distance Vector Protocol
- ▶ Dense Mode Protocol
- ▶ Generates source-specific shortest path trees
- ▶ Currently V3 allows for Mcast tunnelling
- ▶ Operates on RIP bases (as Unicast Routing Protocol)
- ▶ Transmits Subnetmasks
- ▶  $\infty$  = 32 Hops, sometimes 16



# DVMRP Distribution Tree: Construction

- **Truncated Broadcast Trees Are Built using Best DVMRP Metrics Back to Source Network.**
- **Lowest IP Address Used in Case of a Tie.**  
(Note: IP Address of D < C < B < A)



Route for source network of metric "n"



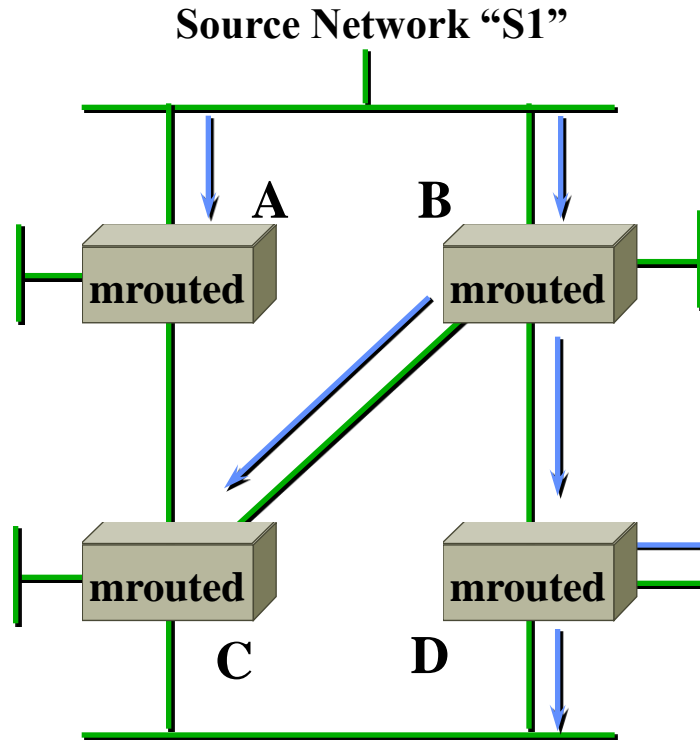
Poison reverse (metric + infinity) sent to upstream "parent" router.  
Router depends on "parent" to receive traffic for this source.



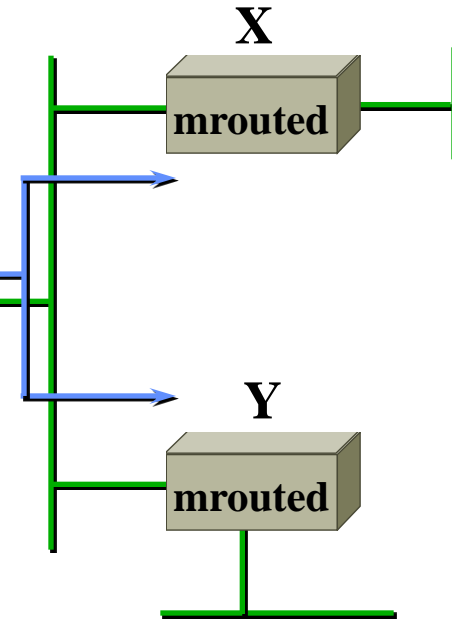
Resulting Truncated Broadcast Tree for Source Network



# DVMRP Distribution Tree



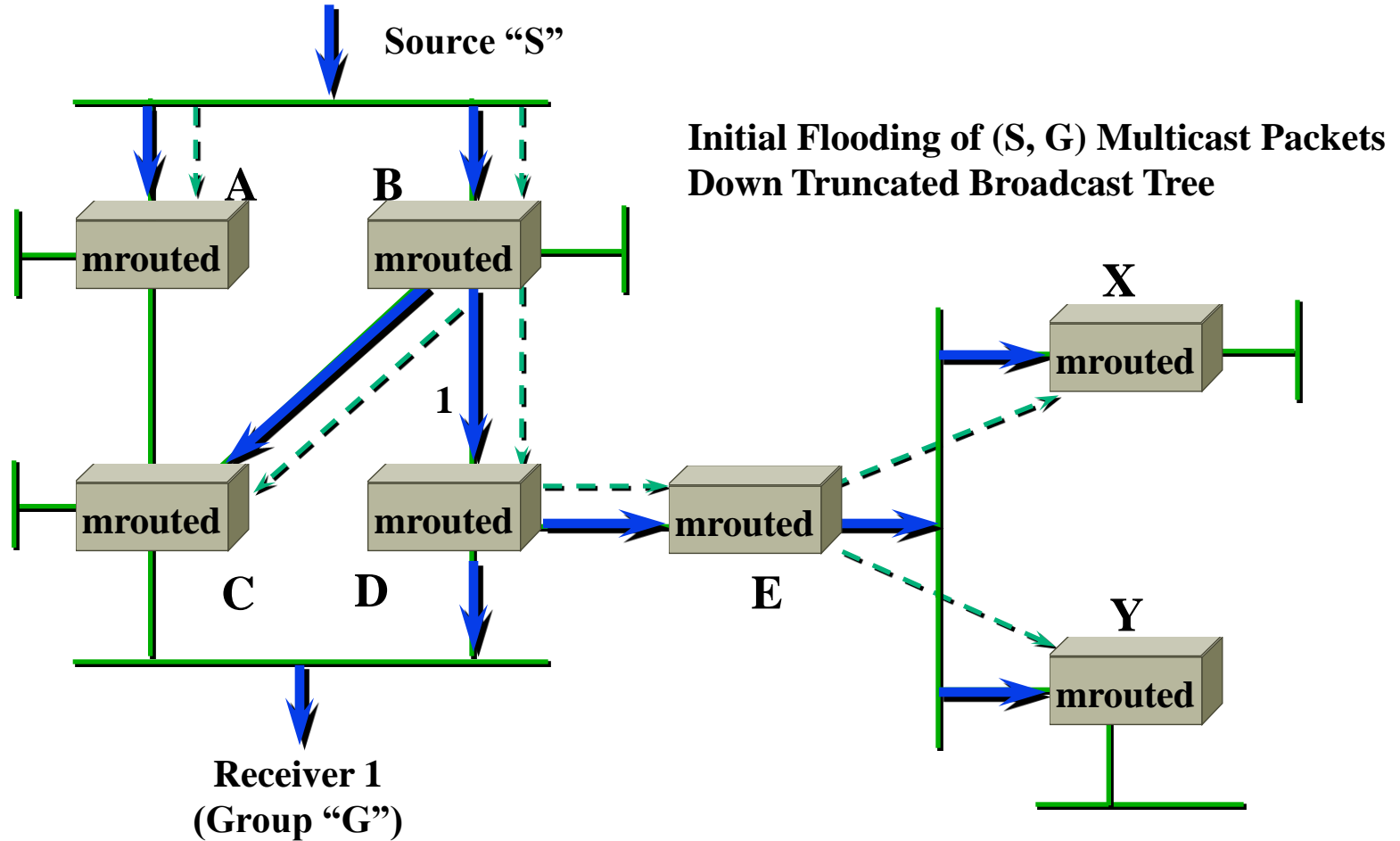
Resulting Truncated Broadcast Tree for Source Network "S1"



 S1 Source Tree



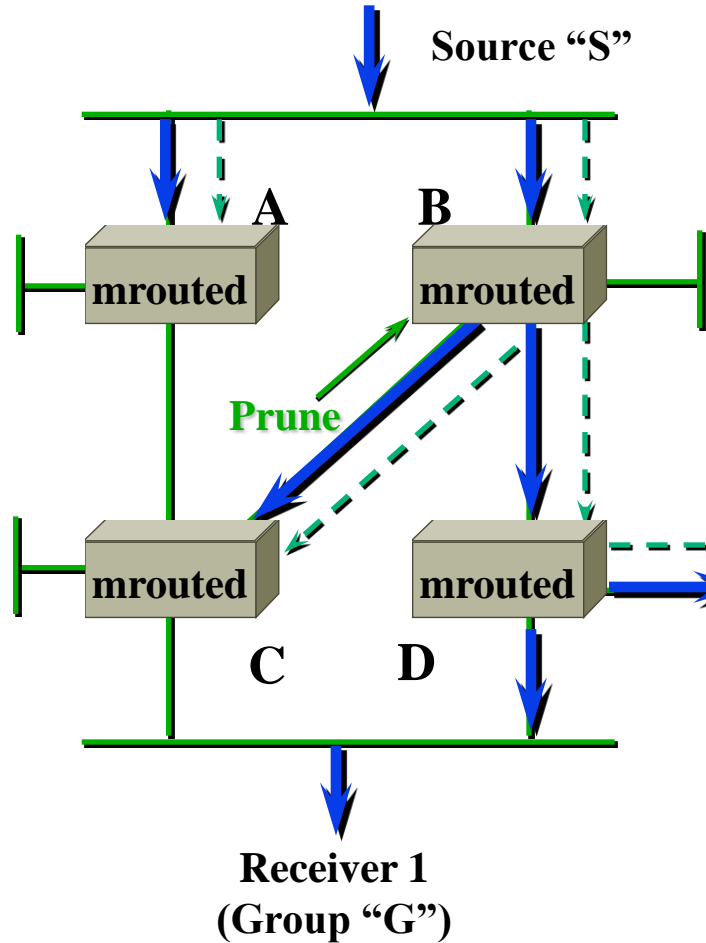
# DVMRP Flood & Prune



 Truncated Broadcast Tree based on DVMRP route metrics  
 (S, G) Multicast Packet Flow



# DVMRP Flood & Prune



Routers C is a Leaf Node so it sends an "(S, G) Prune" Message

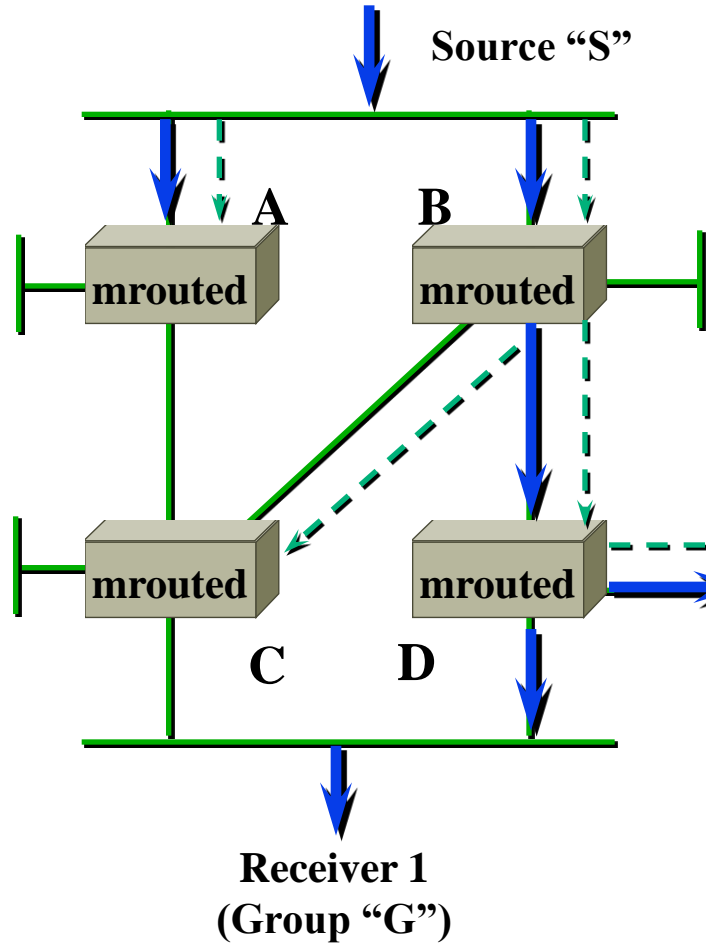
Router B Prunes interface.

 Truncated Broadcast Tree based on DVMRP route metrics  
 (S, G) Multicast Packet Flow





# DVMRP Flood & Prune



Routers X, and Y are also Leaf Nodes so they send "Prune (S, G)" Messages

Router E prunes interface.

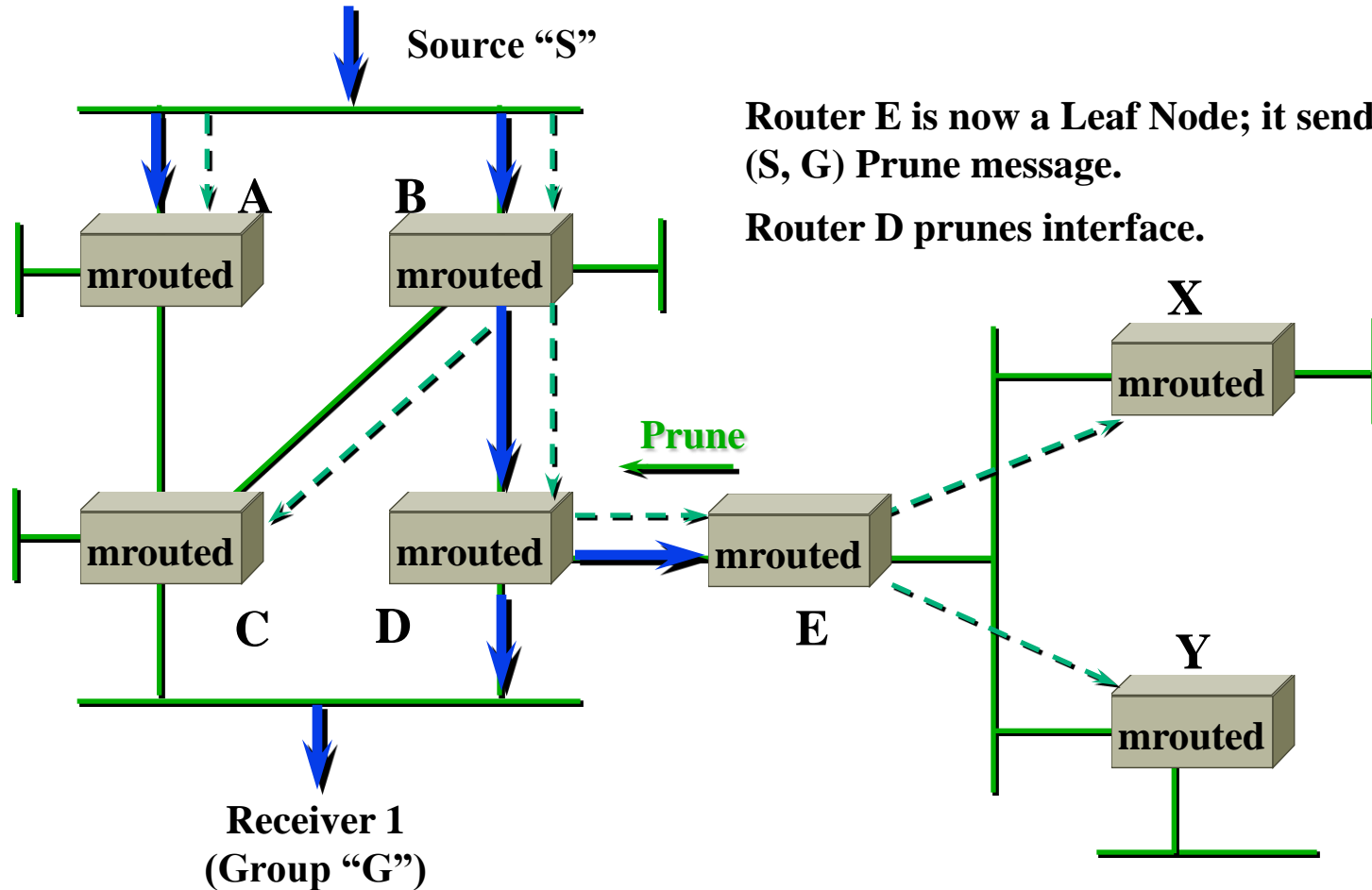
Prune

Prune

- - - - - Truncated Broadcast Tree based on DVMRP route metrics  
—————> (S, G) Multicast Packet Flow



# DVMRP Flood & Prune

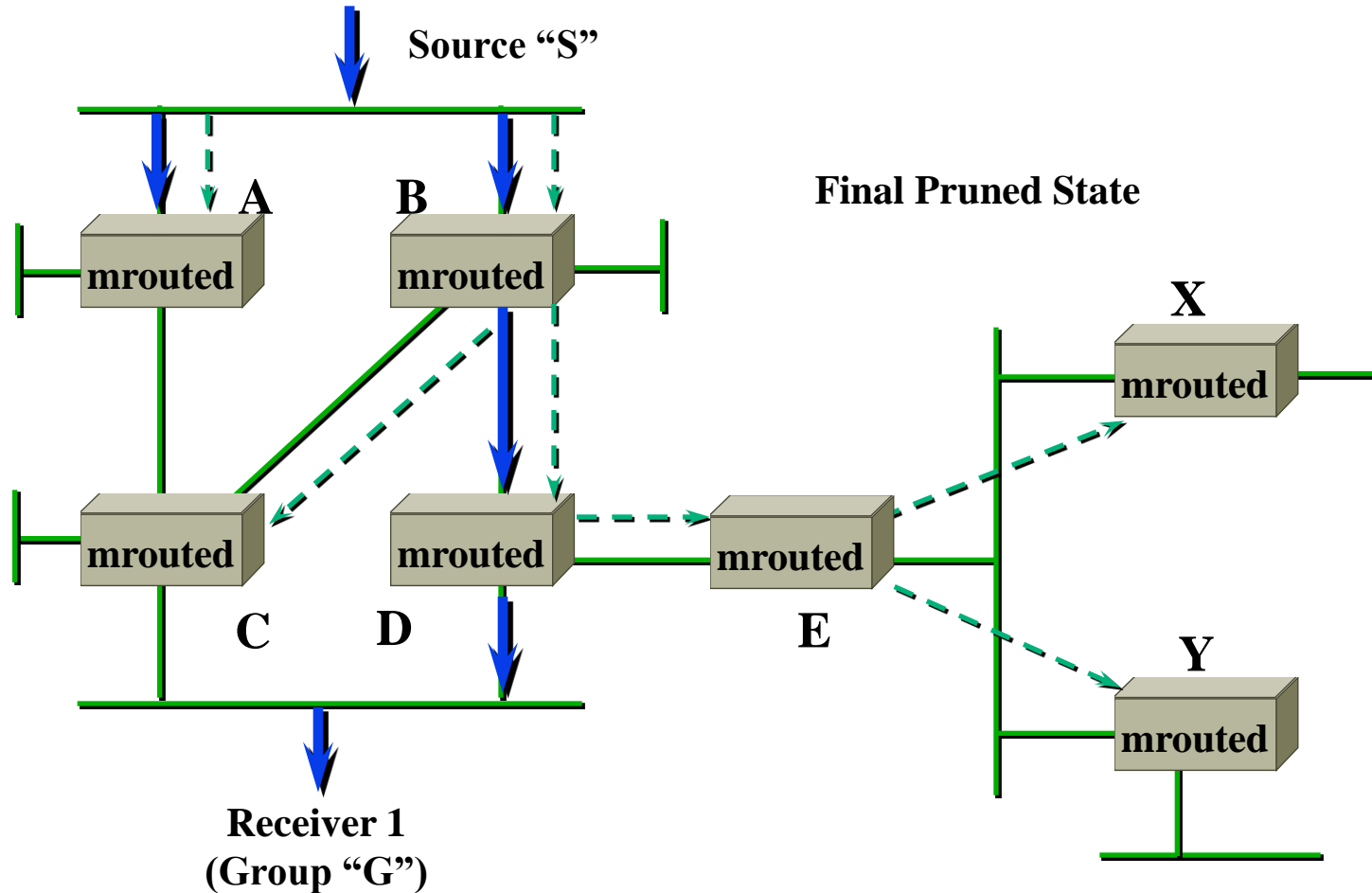


Router E is now a Leaf Node; it sends an (S, G) Prune message.  
Router D prunes interface.

- - - - - Truncated Broadcast Tree based on DVMRP route metrics  
—————> (S, G) Multicast Packet Flow



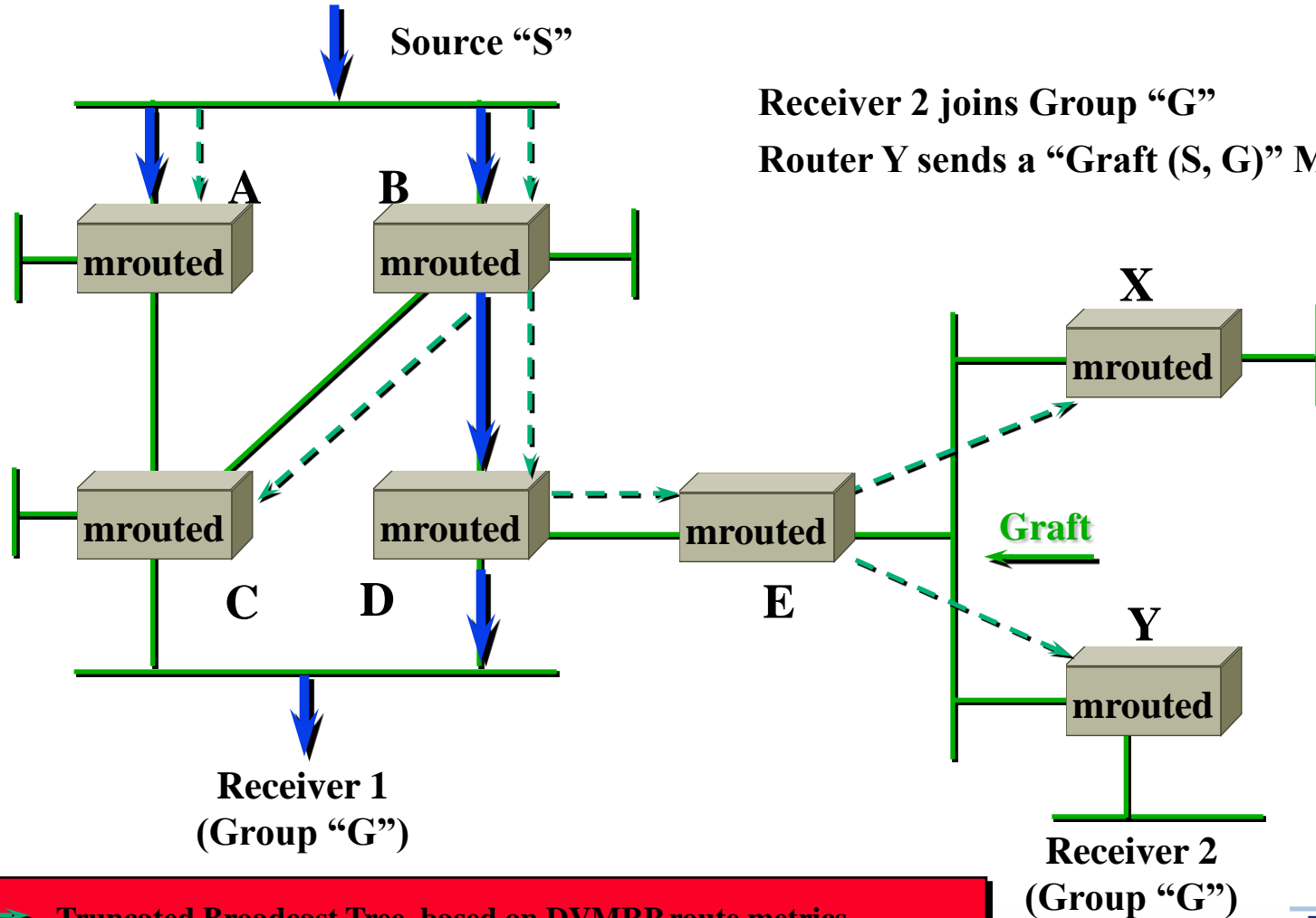
# DVMRP Flood & Prune



**--->** Truncated Broadcast Tree based on DVMRP route metrics  
**→** (S, G) Multicast Packet Flow



# New Receiver in DVMRP: Grafting

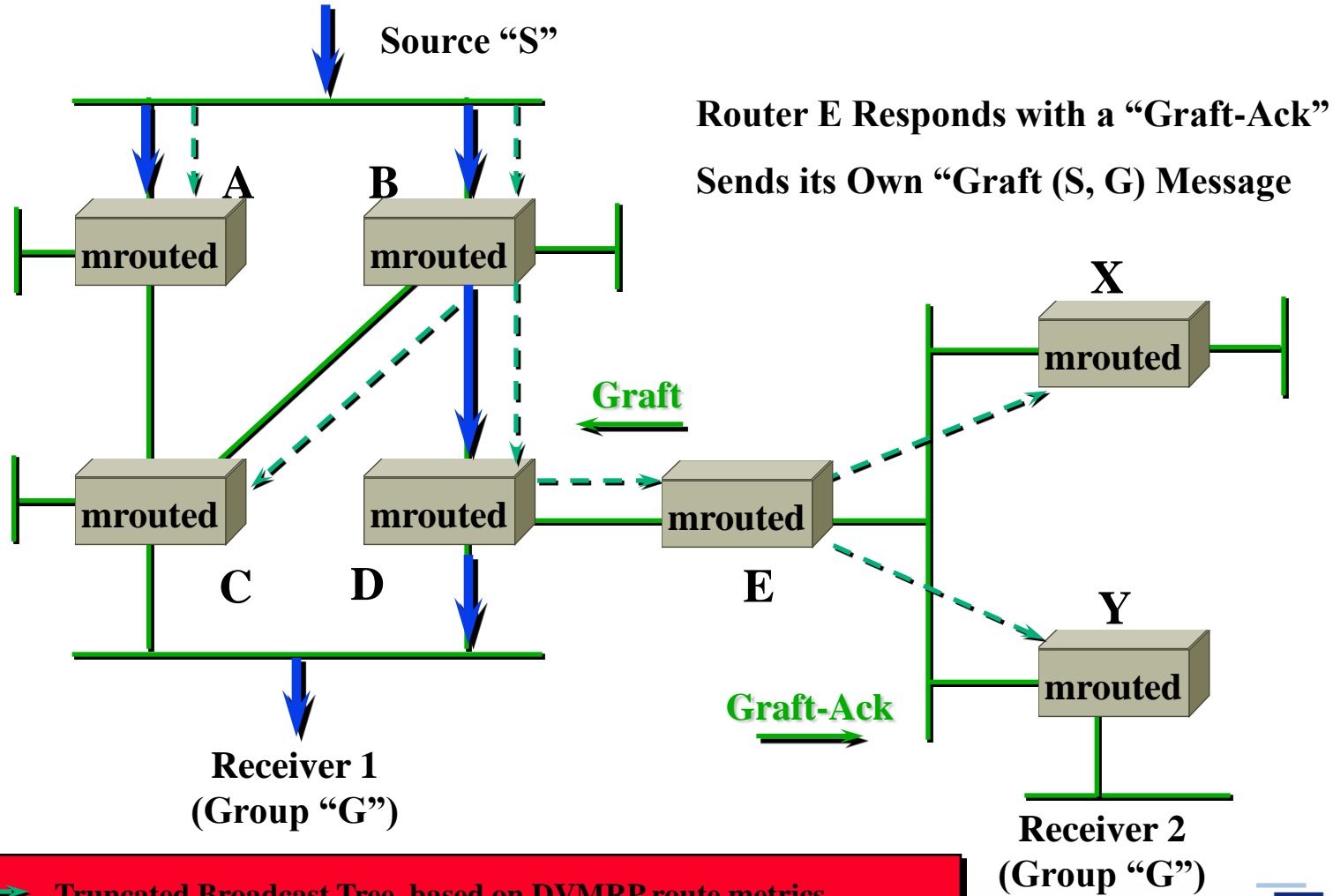


Receiver 2 joins Group "G"  
Router Y sends a "Graft (S, G)" Message

- - - - - Truncated Broadcast Tree based on DVMRP route metrics  
→ (S, G) Multicast Packet Flow



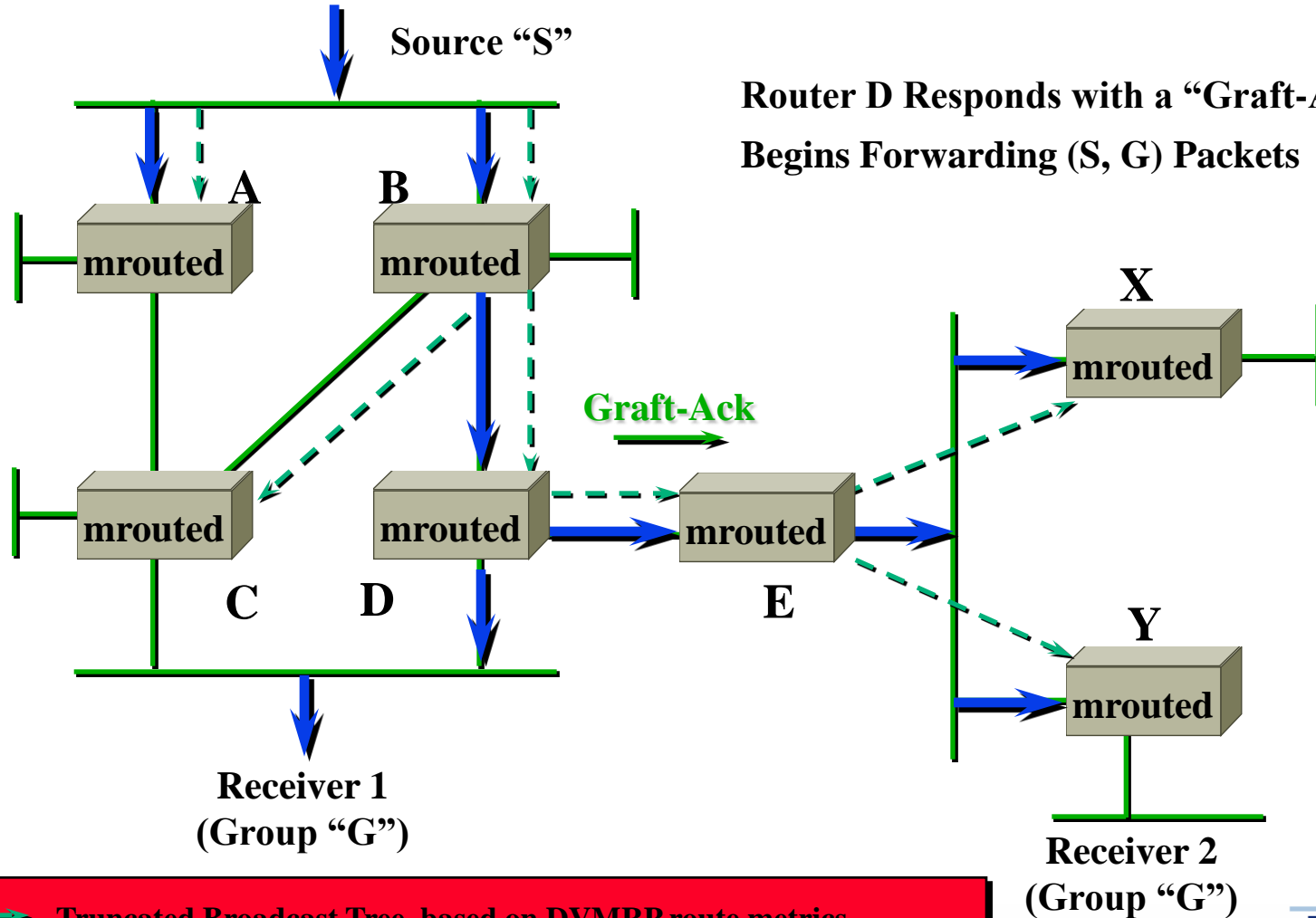
# DVMRP Grafting



- - - - - Truncated Broadcast Tree based on DVMRP route metrics  
—————> (S, G) Multicast Packet Flow



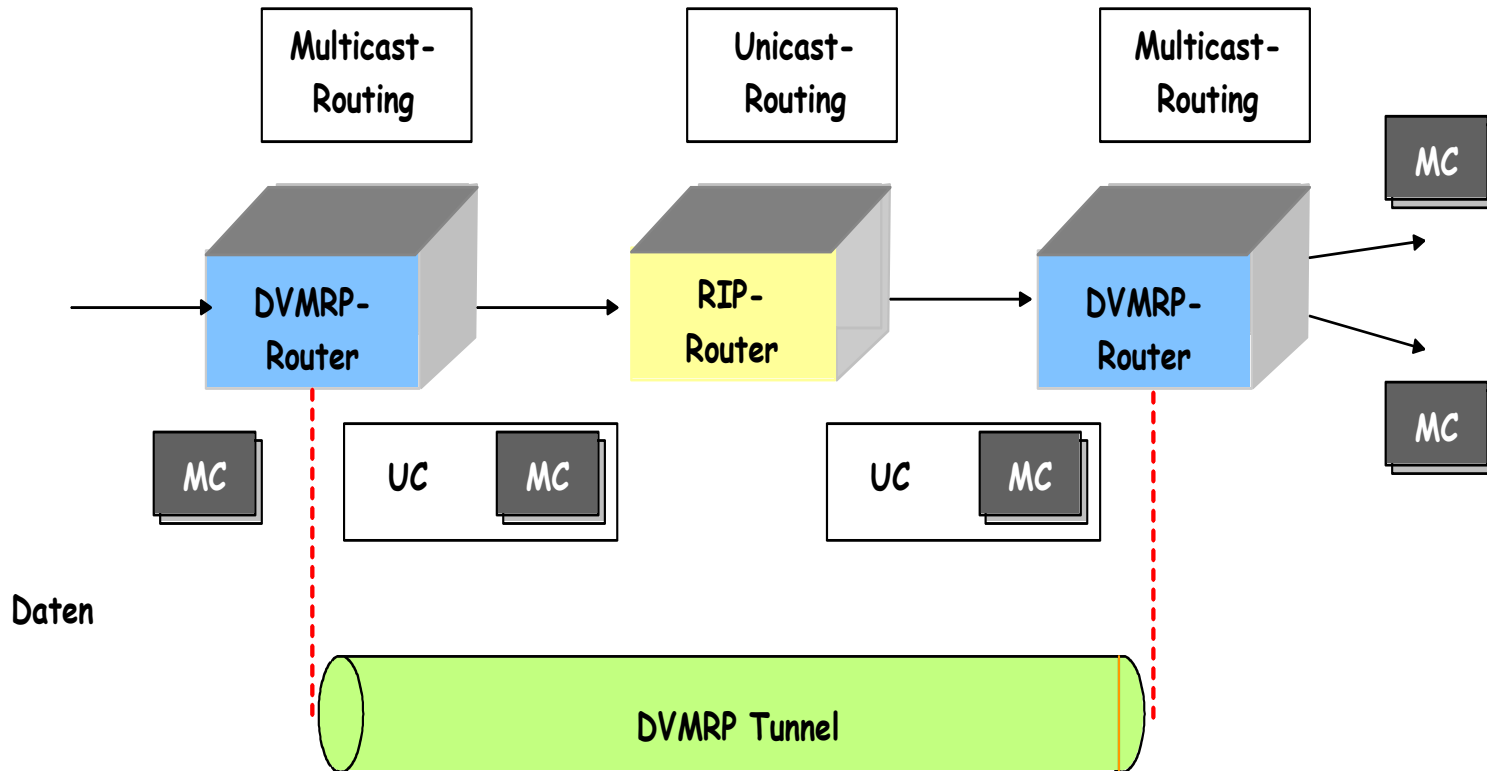
# DVMRP Grafting



 Truncated Broadcast Tree based on DVMRP route metrics  
 (S, G) Multicast Packet Flow



# DVMRP Tunnelling



**UC: Unicast**  
**MC: Multicast**

# Multicast Open Shortest Path First (MOSPF)

- ▶ Extends OSPF for Multicast Routing
- ▶ Destination based link state protocol (dense mode)
- ▶ Distribution of link states (OSPF)
- ▶ Group member link states flooded
- ▶ Every router learns a complete topology and calculates shortest path tree
- ▶ **MOSPF corresponds to OSPF-Unicast-Routing**





# Protocol Independent Multicast (PIM)

- ▶ *Protocol independence:*

  - works with all underlying Unicast Routing Protocols

- ▶ Dense und Sparse Mode PIM (RFC ..., current RFC 4601 08/'06)

- ▶ Dense Mode PIM floods & prunes (as DVMRP)

- ▶ Sparse Mode PIM uses Rendezvous Points (RP)

  - ▶ Constructs a shared distribution tree centred at RP

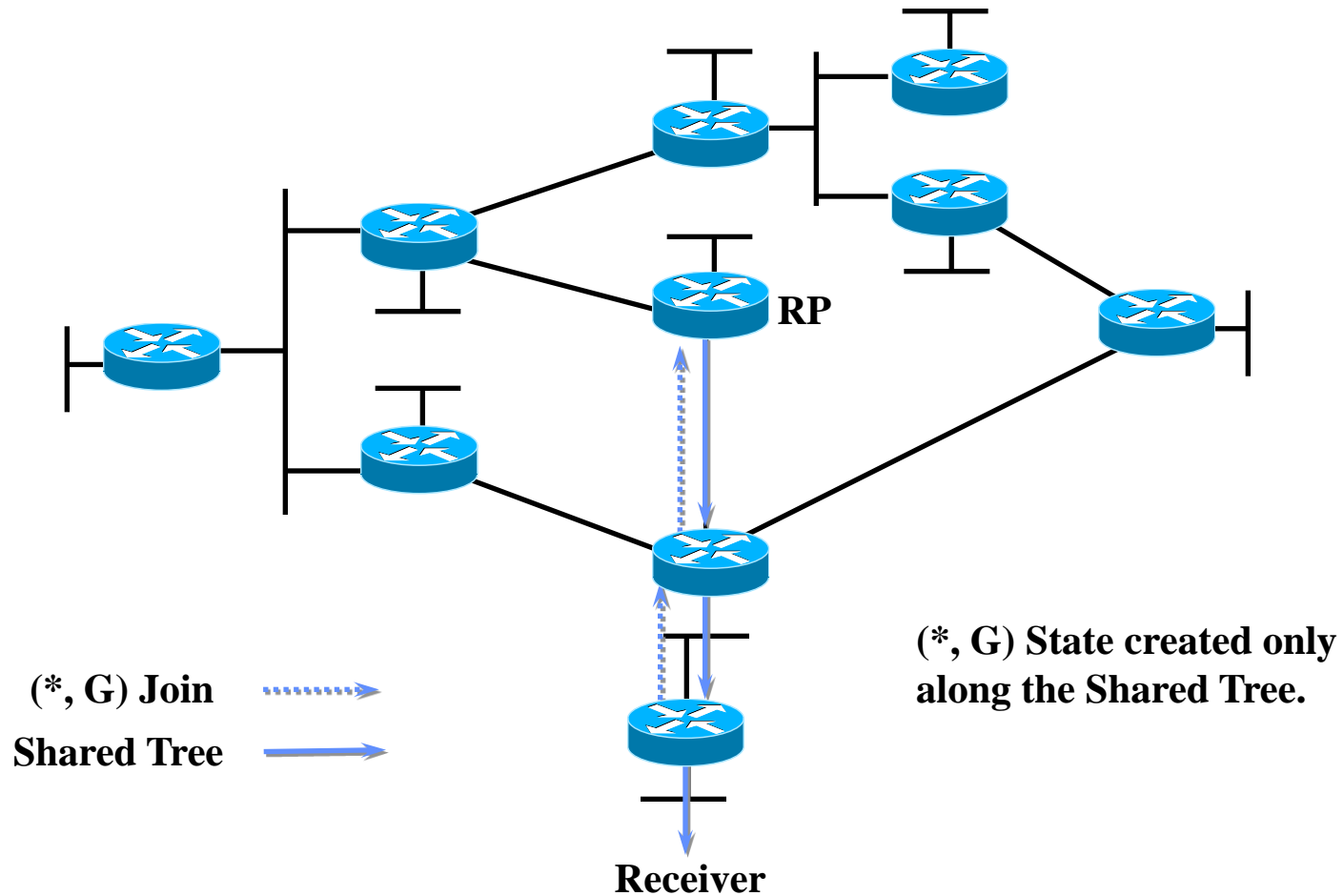
  - ▶ Efficient for widely distributed groups

  - ▶ Favoured for wide area networks – problem:  
inter-RP signalling

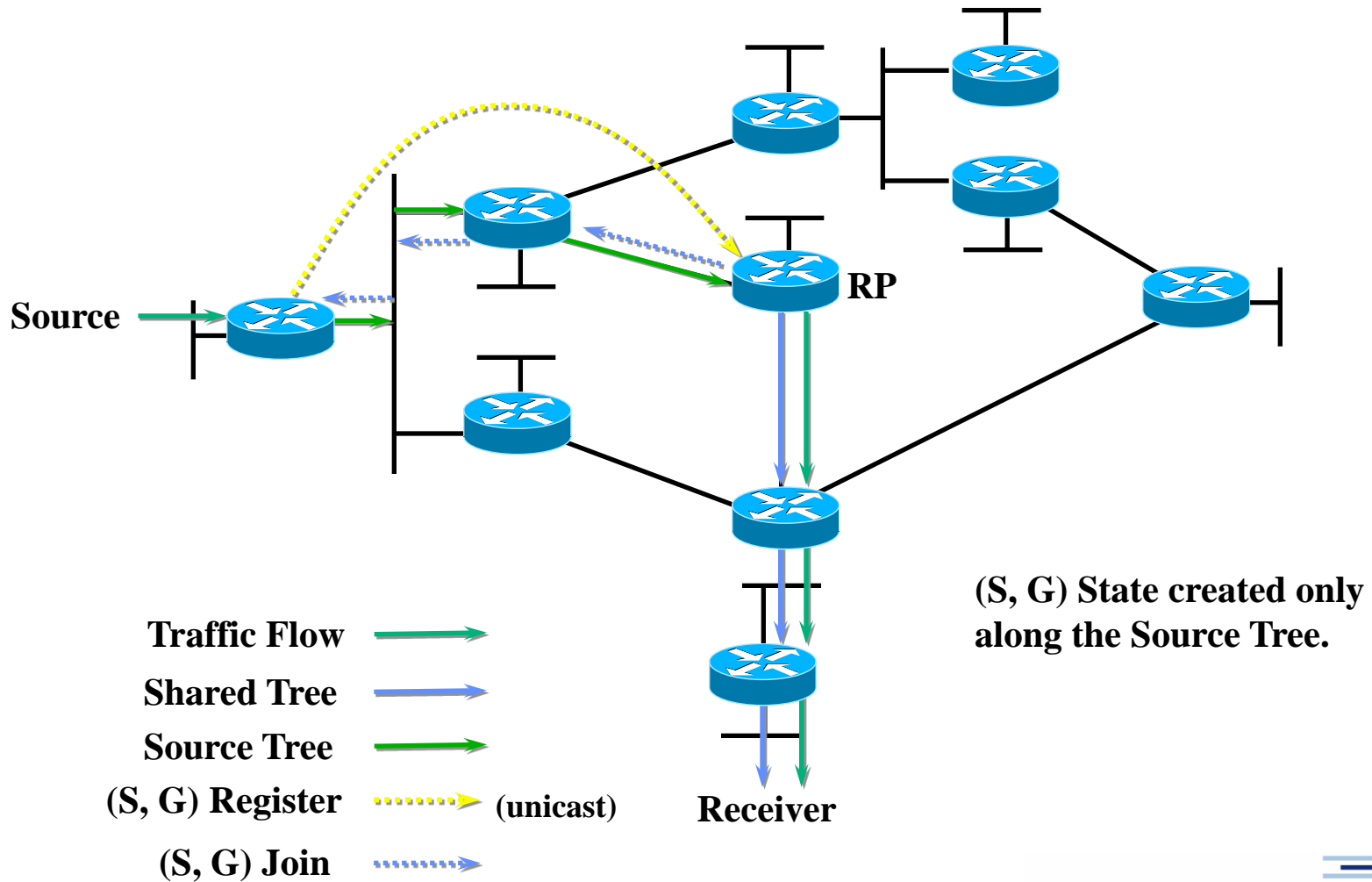
  - ▶ Now widely implemented



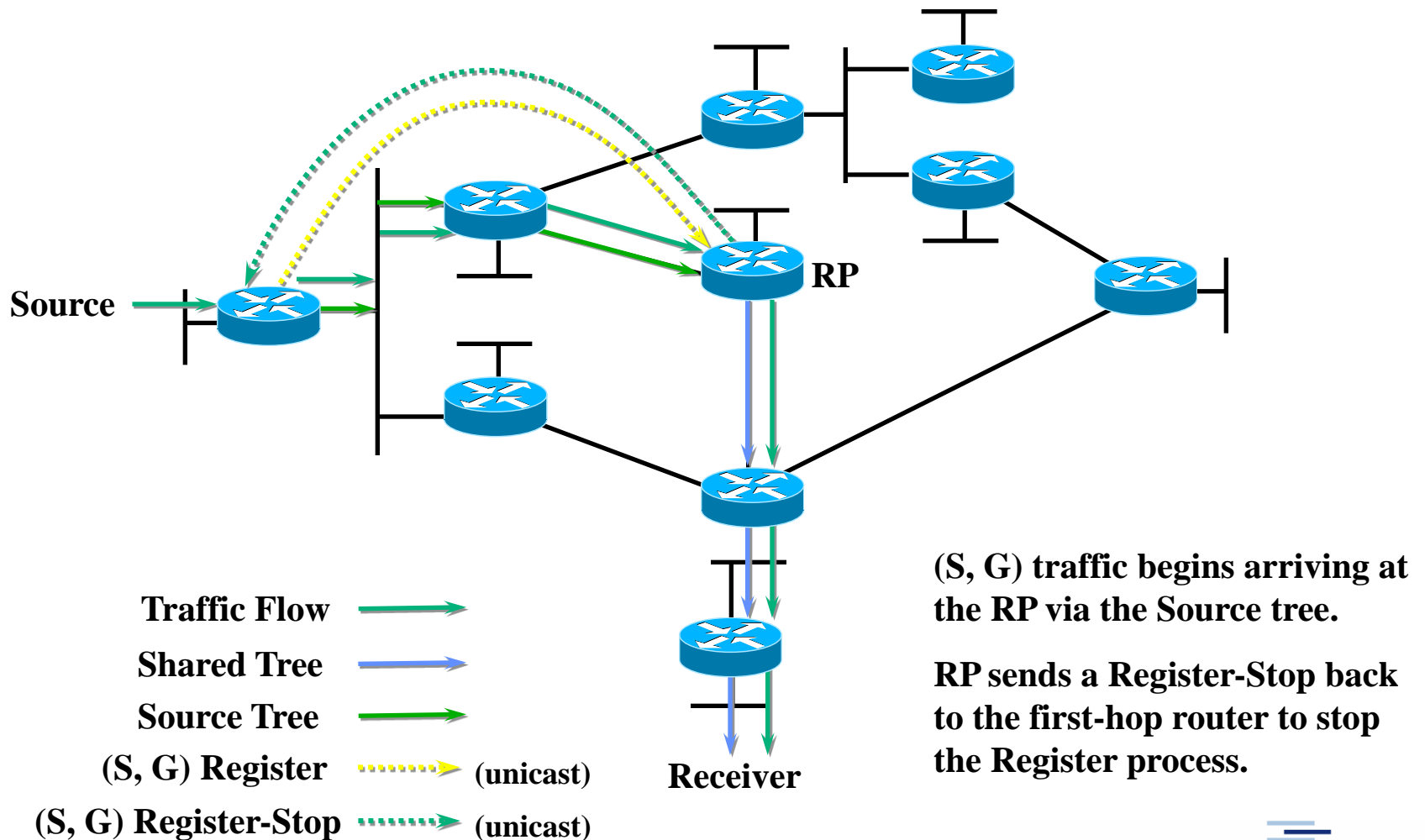
# PIM SM Tree Joins



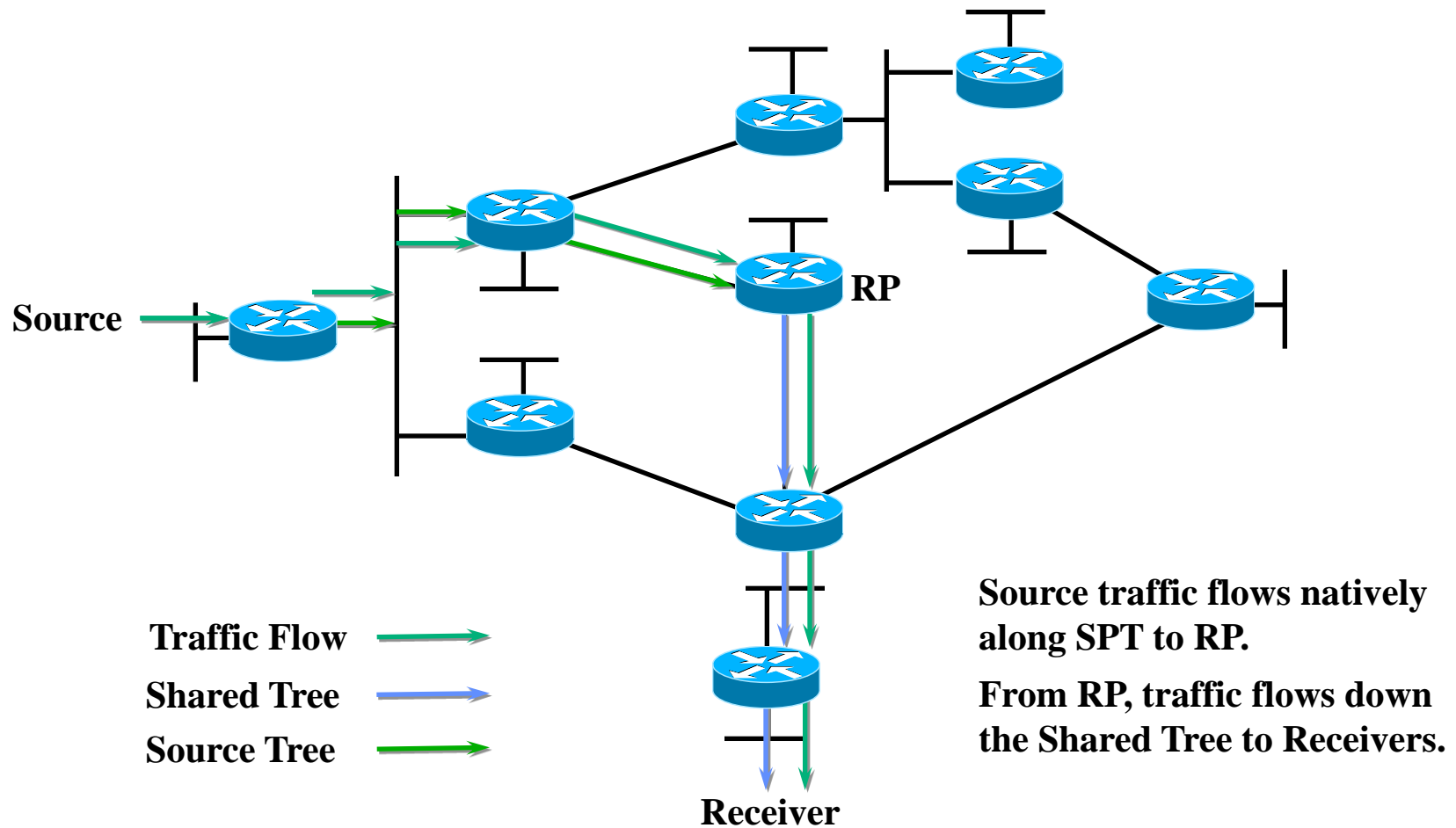
# PIM SM Sender Registration



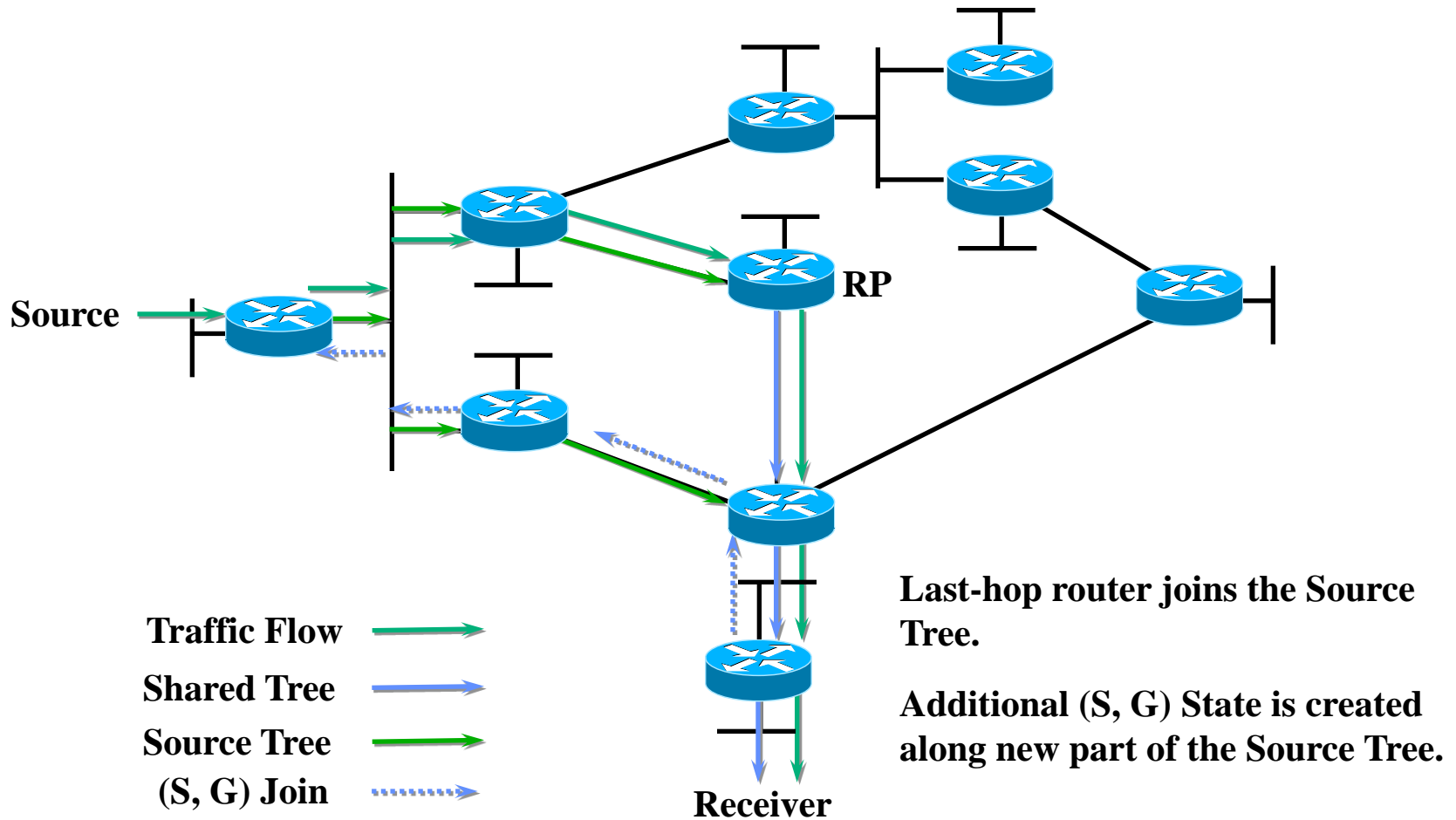
# PIM SM Sender Registration



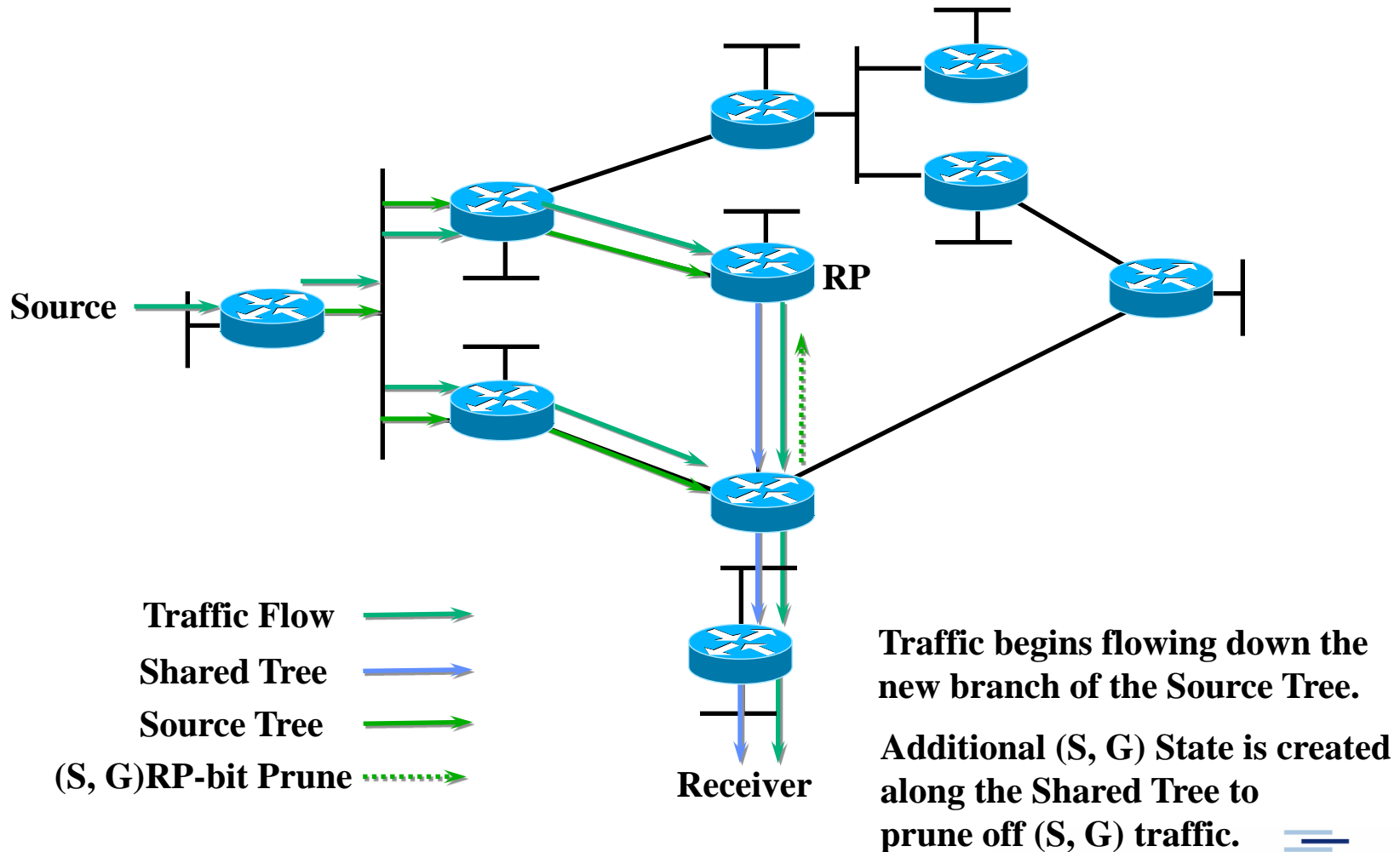
# PIM SM Sender Registration



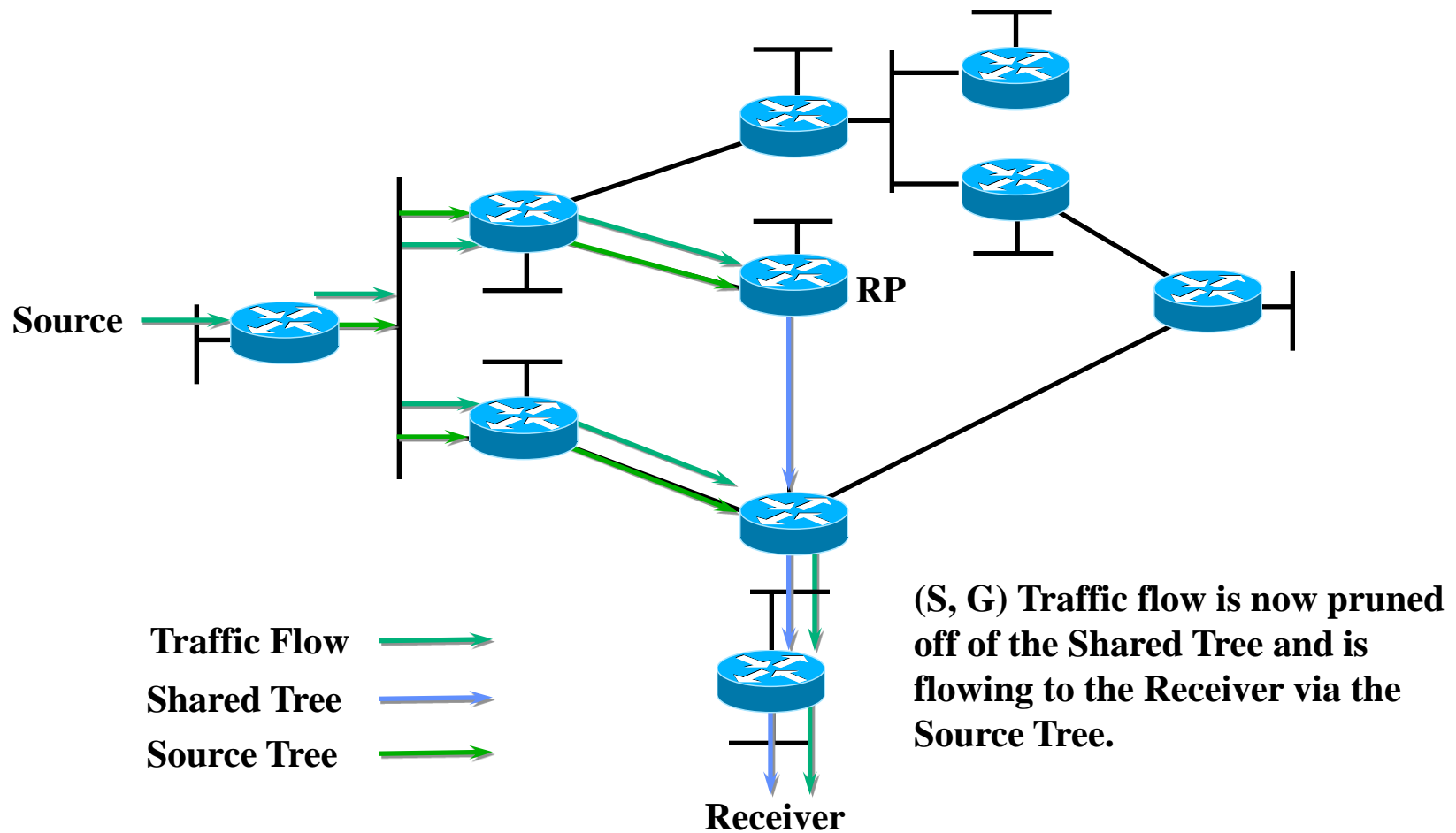
# PIM SM Short Cut



# PIM SM Short Cut

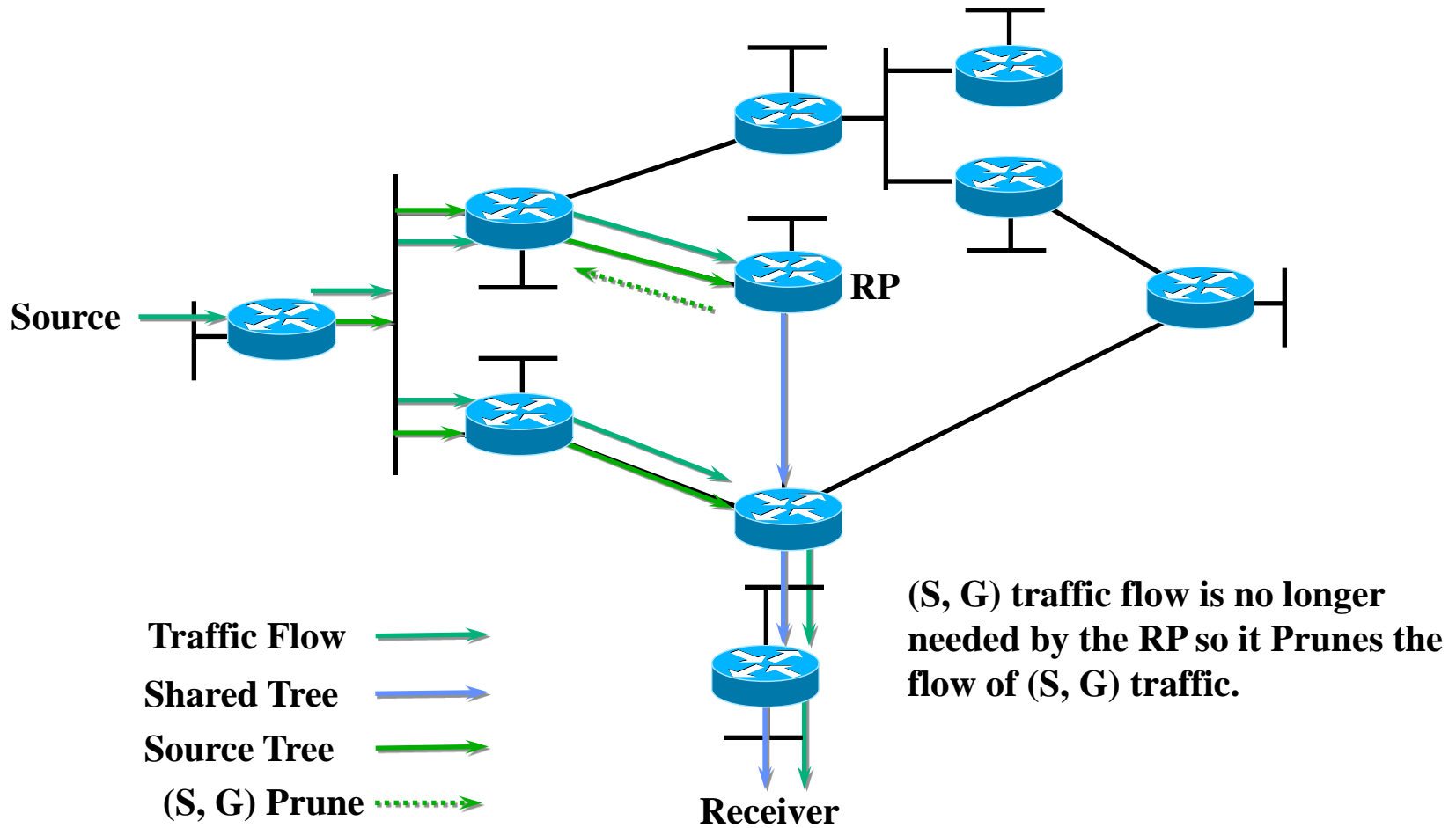


# PIM SM Short Cut

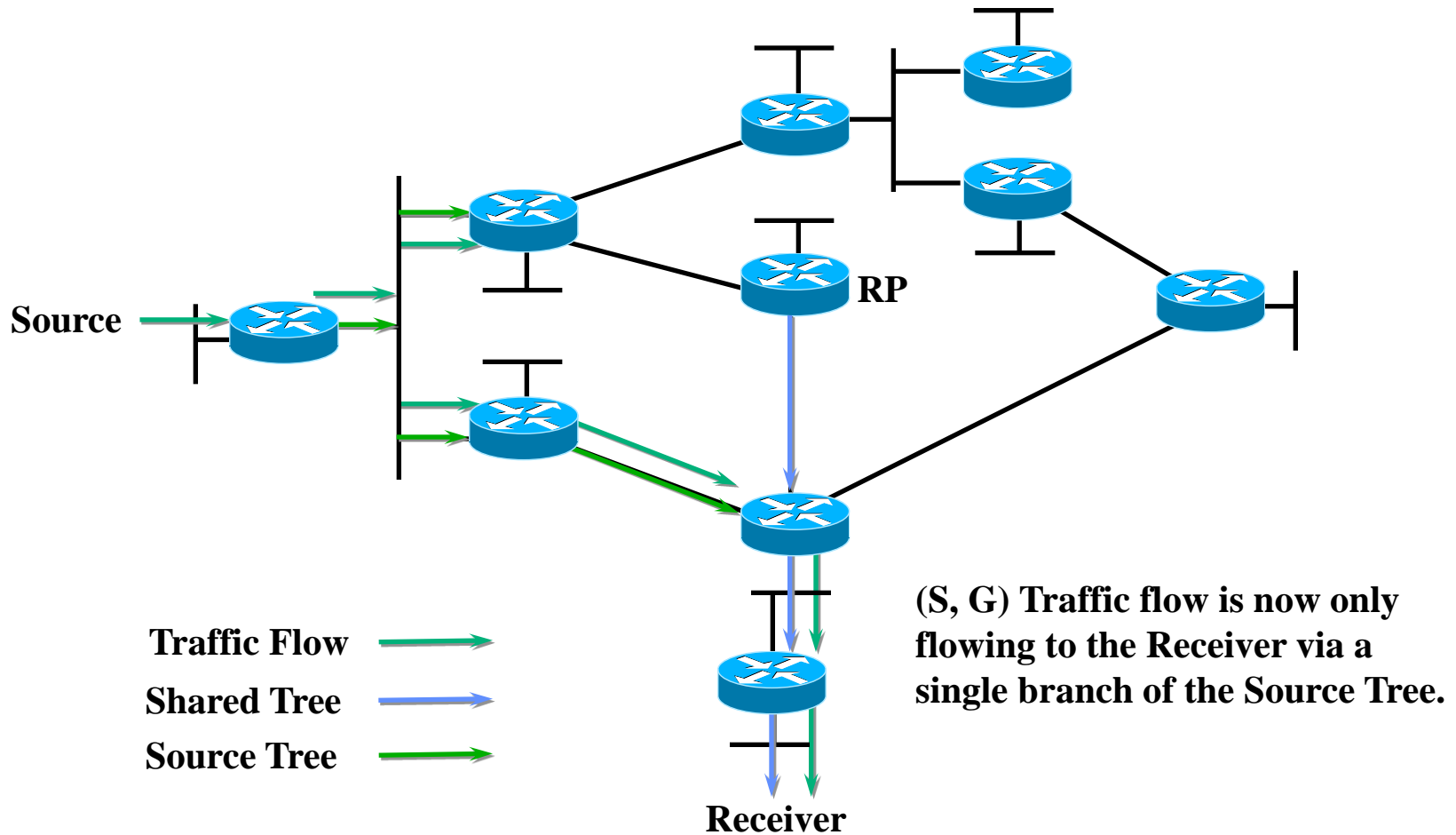




# PIM SM Short Cut



# PIM SM Short Cut



# Bidirectional PIM

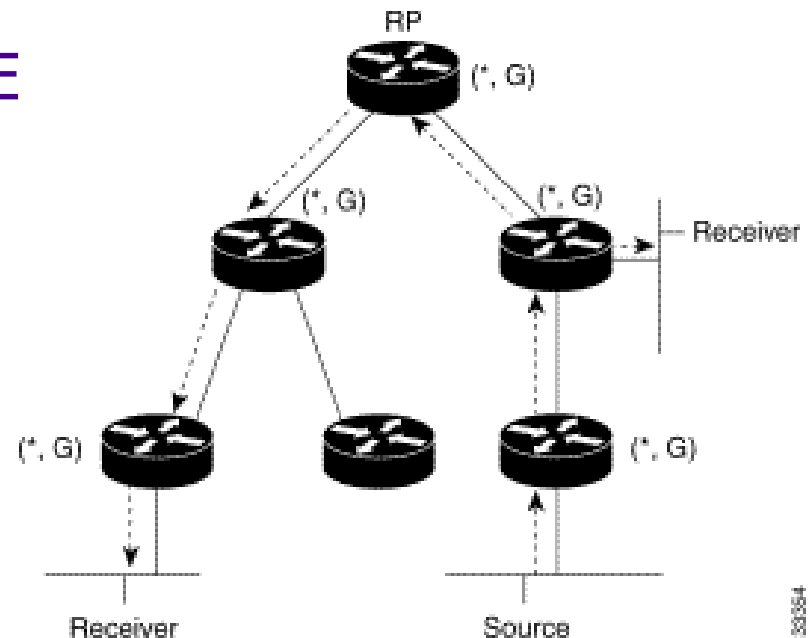
Mark Handley et. al.: RFC 5015

- ▶ Intra-domain protocol
- ▶ Selects (per Group) a “virtual” rendezvous point address (RPAs) – this may be an unused address on the rendezvous point link (RPL)
- ▶ Generates (RPA) a **shared tree** of designated forwarders (DFs): One router per link with best route to RPA
- ▶ Explores a domain by per group shared forwarding states: "NoInfo" or "Include"
- ▶ Decouples state management from data plane



# Bidirectional PIM (2)

- ▶ Trees have RPA as virtual root, branch on RPL
- ▶ Group specific states are propagated by JOIN/PRUNE messages towards RPA
- ▶ Shared trees are operated bidirectionally
- ▶ Sources always forward upstream even without on-link receivers



38/38



# Source Specific Multicast - SSM

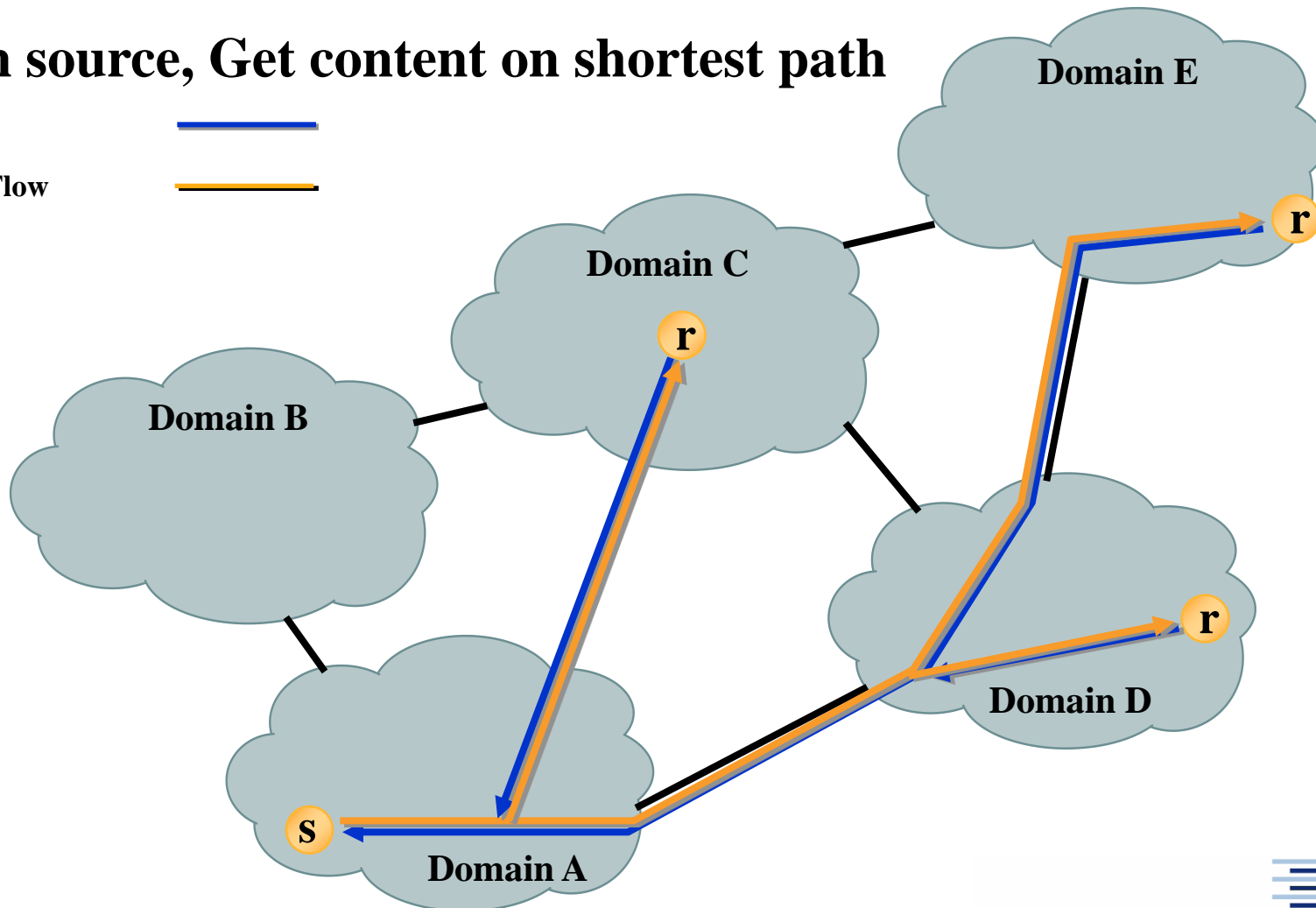
- ▶ Recently released (RFC 3569, RFC 4607 08/'06)
- ▶ Assumes source address known at receiver
  - ▶ Allows for source selection
  - ▶ Source discovery offline or via MSDP
- ▶ Receiver subscribes to (S,G) using IGMPv3/MLDv2
  - ▶ No state aggregation on shared trees
- ▶ Routing: PIM-SSM, a subset of PIM-SM (in RFC 4601)
  - ▶ Obsoletes rendezvous points & flooding
- ▶ Simpler, well suited for single source media broadcast or interdomain apps



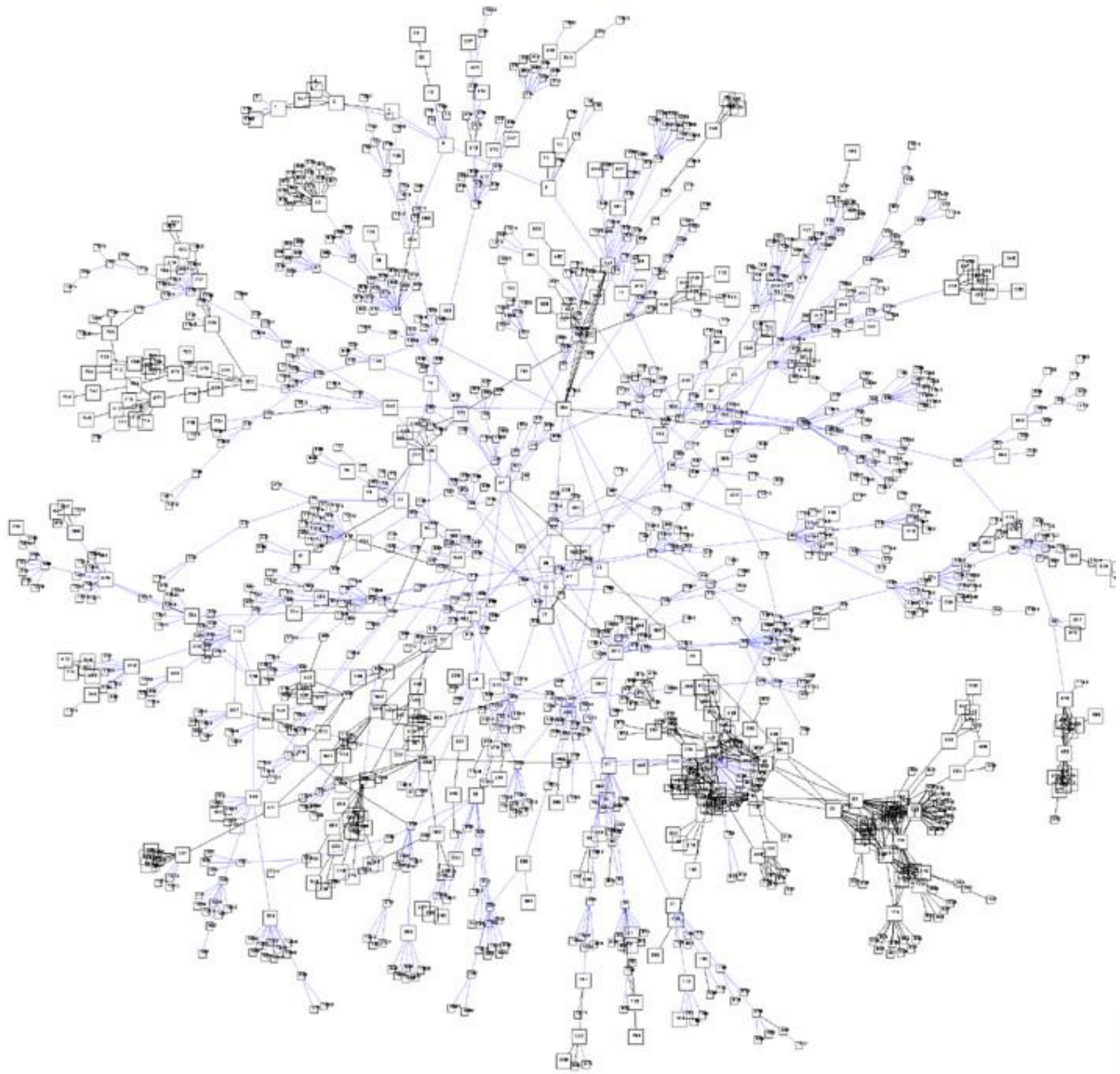
# SSM Routing

**Join source, Get content on shortest path**

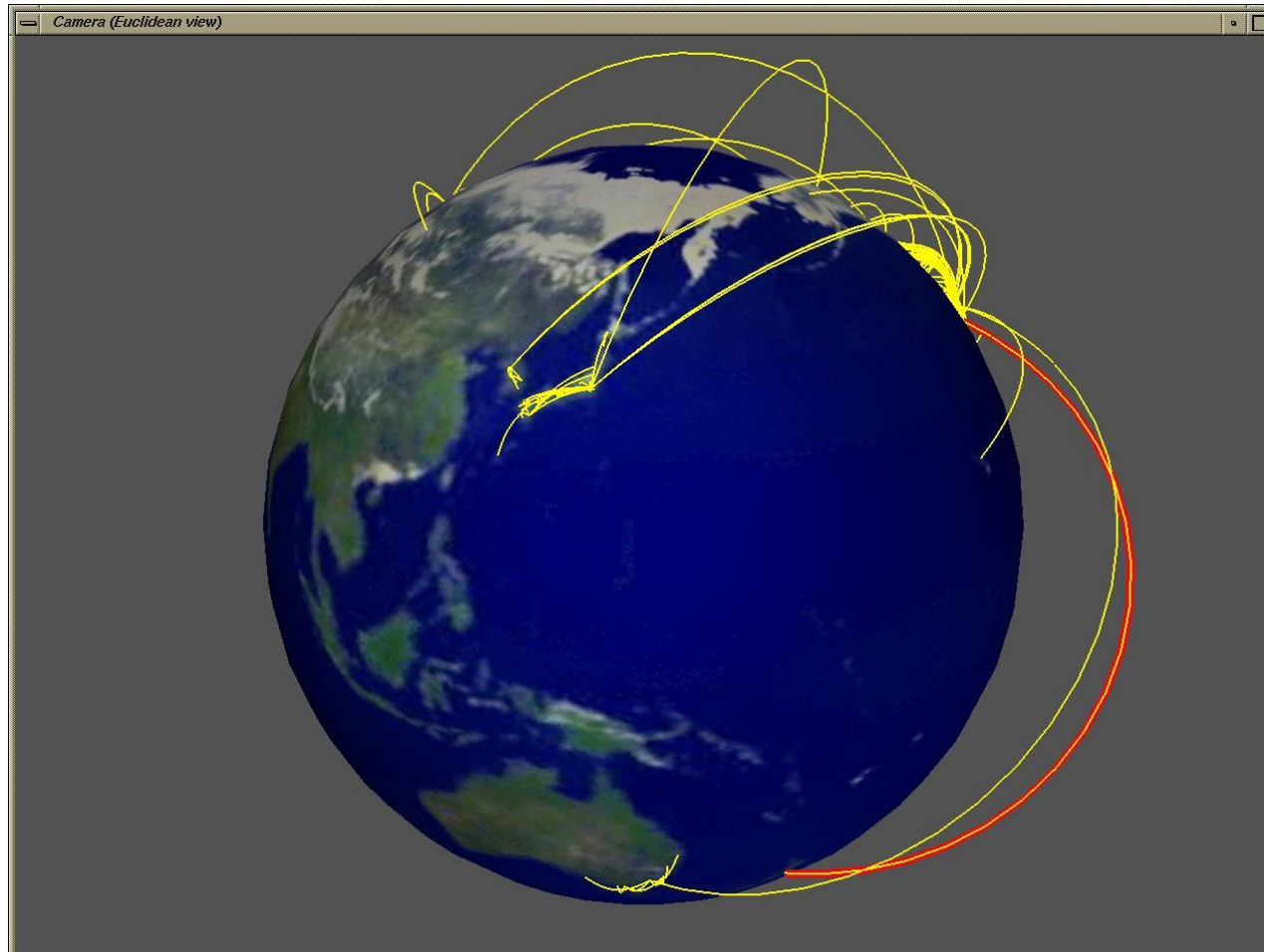
Join   
Data Flow 



# MBone



# Visualisation of Multicast Group



```
DST DEC3800-1-FDDI-0.LOSANGELES.MCI.NET LOS ANGELES CA 34.05 -118.24  
SRC IHUG.CO.NZ WELLINGTON NEWZEALAND -41.00 175.50
```

Image & Video by Tamara Munzer, Univ. of British Columbia





# Efficiency of Multicast

- ▶ For  $m$  receivers
  - ▶  $L_M(m)$  : Number of links in multicast SPT
  - ▶  $\langle L_U \rangle$  : Average # of unicast hops between uniformly chosen end nodes, then clearly

$$L_M(m) < m * \langle L_U \rangle$$

- ▶ Empirical Scaling Law (Chuang and Sirbu 1998/2001):

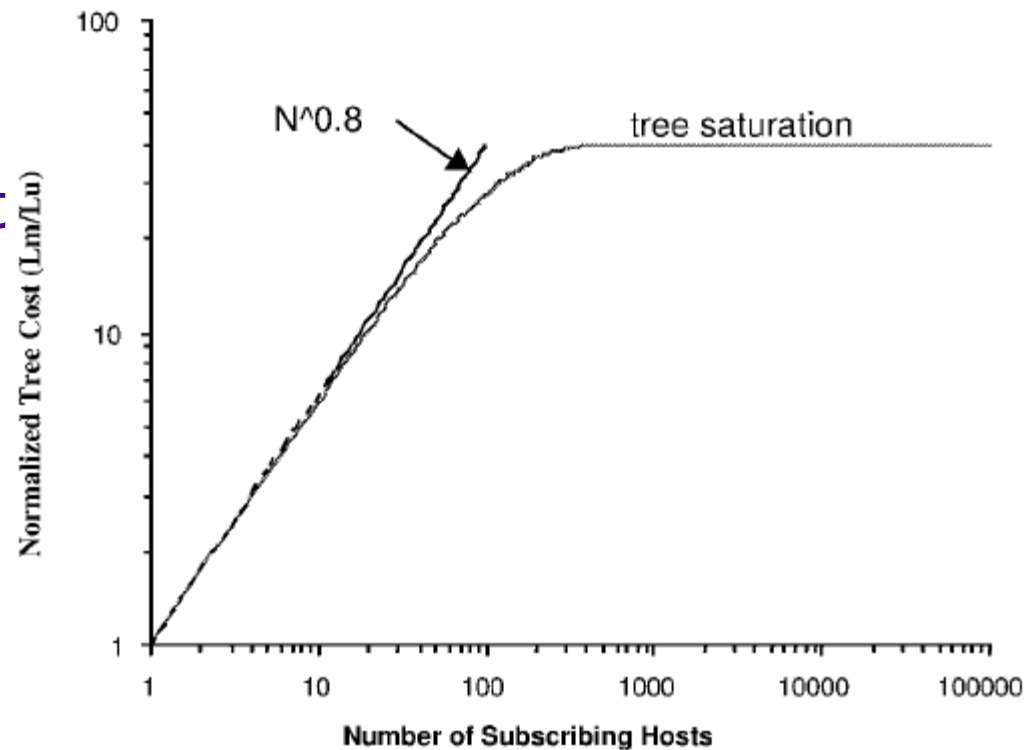
$$L_M(m) \approx \langle L_U \rangle * m^{0.8}$$

- ▶ This means: multicast shortest path trees are of self-similar nature with many nodes of small, but few of higher degrees
- ▶ Trees are shaped rather tall than wide



# Chuang and Sirbu Scaling Law

- ▶ Empirical measurement on Internet & generated topologies
- ▶ Exponent found to be topology-independent
- ▶ Saturation due to exhaustive network exploration



Graphic from Chuang Sirbu (2001)



# Efficiency of Multicast (2)

- Van Mieghem et al. (2001) proved that the Chuang and Sirbu scaling law cannot hold in general, but can be reasonably well approximated by

$$L_M(m) \approx \langle L_U \rangle m^k, \quad k = k(N) = \frac{\text{var}[L_U(N)]}{\langle L_U(N) \rangle}$$

where  $N$  is the number of core nodes of the underlying network and  $m \ll N$

- For the current Internet size ( $N \approx 250.000$  core nodes) and moderate receiver numbers  $m \ll N$ :

$$k \approx 0.8.$$



# Properties of Shortest Path Trees

- ▶ Assume:  $m$  multicast receivers are uniformly chosen out of  $N$  network nodes\*, then

*If the link weights are iid., exponentially with mean 1, the Shortest Path Tree is a **Uniform Recursive Tree***

- ▶ URTs are well studied self-similar trees
- ▶ Relevant quantities can be derived analytically:  
Average hopcount, path weights, stability ...
- ▶ Allows to answer á priori deployment questions, e.g. cost efficiency of multicast ...

\* This assumption has been theoretically and empirically justified, cf. Van Mieghem 2006

# Markov Discovery : Uniform Recursive Trees

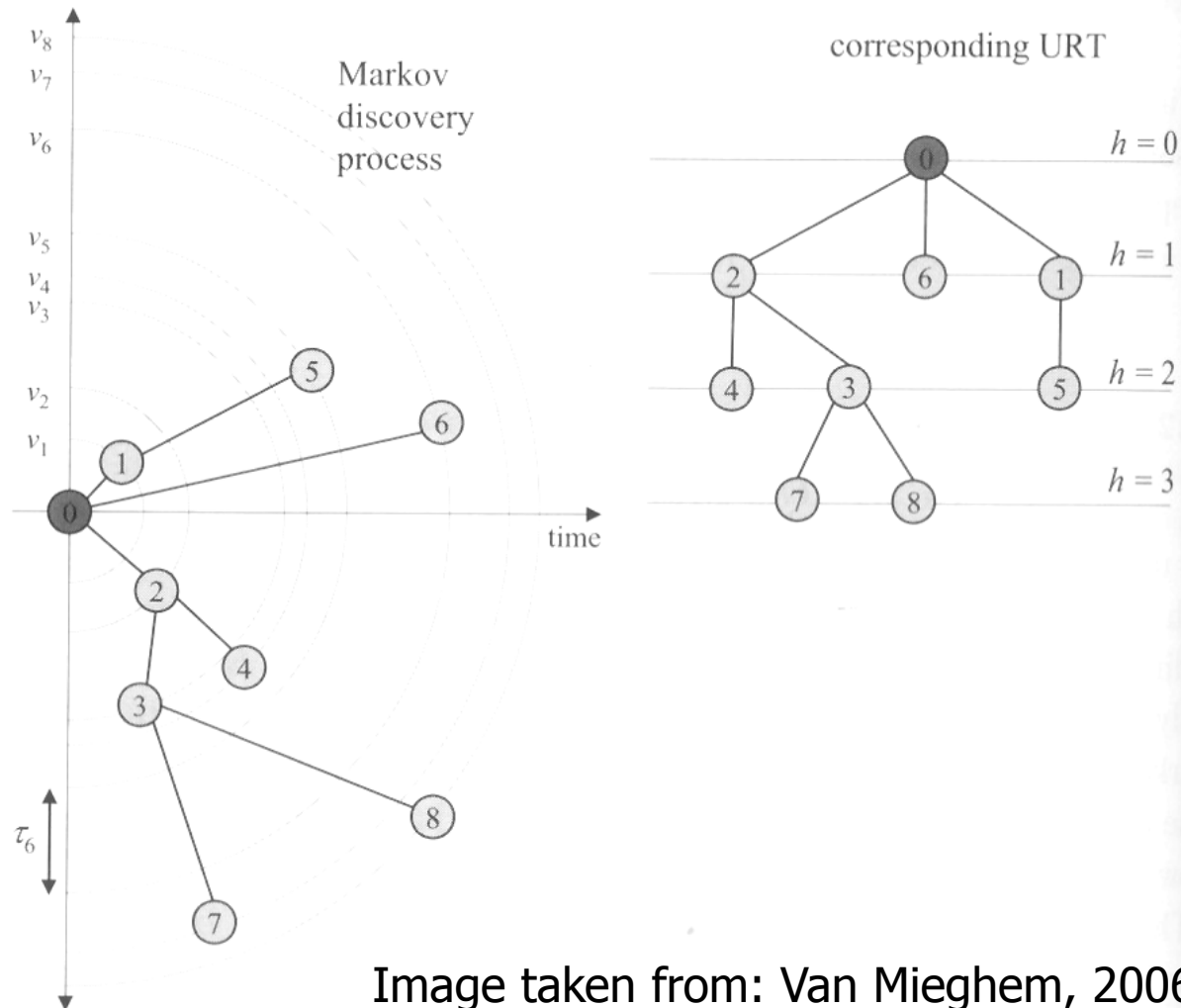


Image taken from: Van Mieghem, 2006



# IP Mcast Deployment Issues

- ▶ Complexity versus Performance Efficiency
  - ▶ IP Multicast most efficient, but burdens infrastructure
- ▶ Provider Costs
  - ▶ Provisioning of knowledge, router capabilities & maintenance, Interdomain mcast routing problem
- ▶ Security
  - ▶ ASM simplifies DDoS-attacks
- ▶ End-to-End Design Violation?
  - ▶ Service complexity objects implementation at lower layer



# QoS for Multicasting

## Resource ReSerVation Protocol (RSVP)

- RFC 2205
- Destination oriented Reservations
  - Sender pushes periodically PATH messages
  - Receiver answers with RESV packets
  - Router interpret these along the paths
- Soft-State-Concept: States time out
- Sender remains unsynchronised



# Further Reading

- ▶ R. Wittmann, M. Zitterbart: *Multicast Communication*, Morgan Kaufmann, 2001
- ▶ [www.ipmulticast.com](http://www.ipmulticast.com)
- ▶ [www.rfc-editor.org](http://www.rfc-editor.org)
- ▶ <ftp://ftpeng.cisco.com/ipmulticast.html>
- ▶ J. Chuang and M. Sirbu: *Pricing Multicast Communication: A Cost-Based Approach*, Telecommunication Systems 17(3), 281 – 297, 2001.
- ▶ P. Van Mieghem: *Performance Analysis of Communication Networks and Systems*, Cambridge University Press, Cambridge, 2006.
- ▶ P. Van Mieghem, G. Hooghiemstra and R. van der Hofstad: *On the Efficiency of Multicast*, IEEE/ACM Trans. Netw. 9(6), pp. 719-732, 2001.

