Multicasty w sieciach

Podstawy

IP Multicast

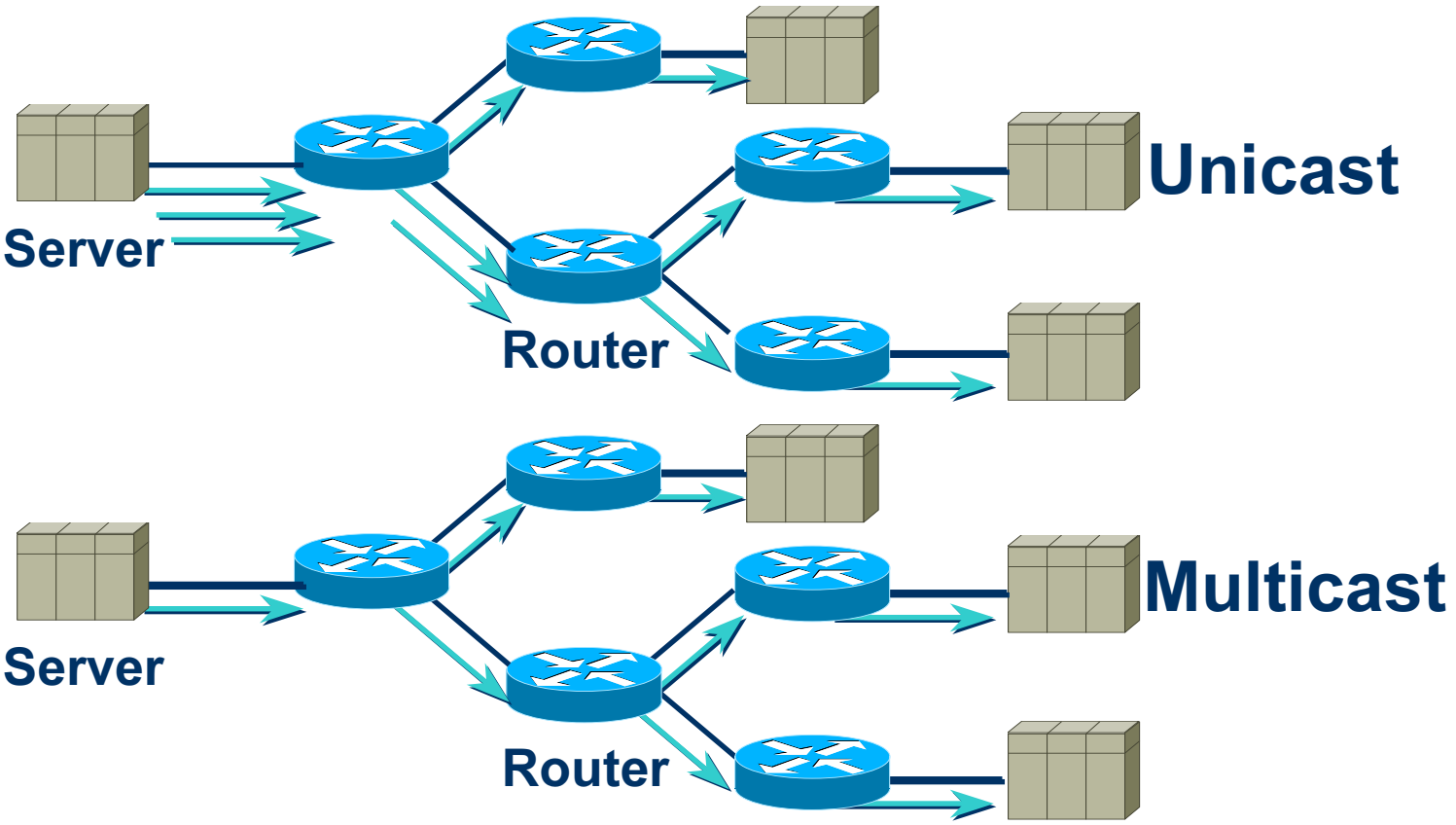
Agenda

- Kiedy stosować transmisję multicastową i dlaczego.
- Podstawowe koncepcje i aplikacje związane z technologiami multicastowymi
- Charakterystyka protokołów routingu IP multicast

Dlaczego IP multicast?

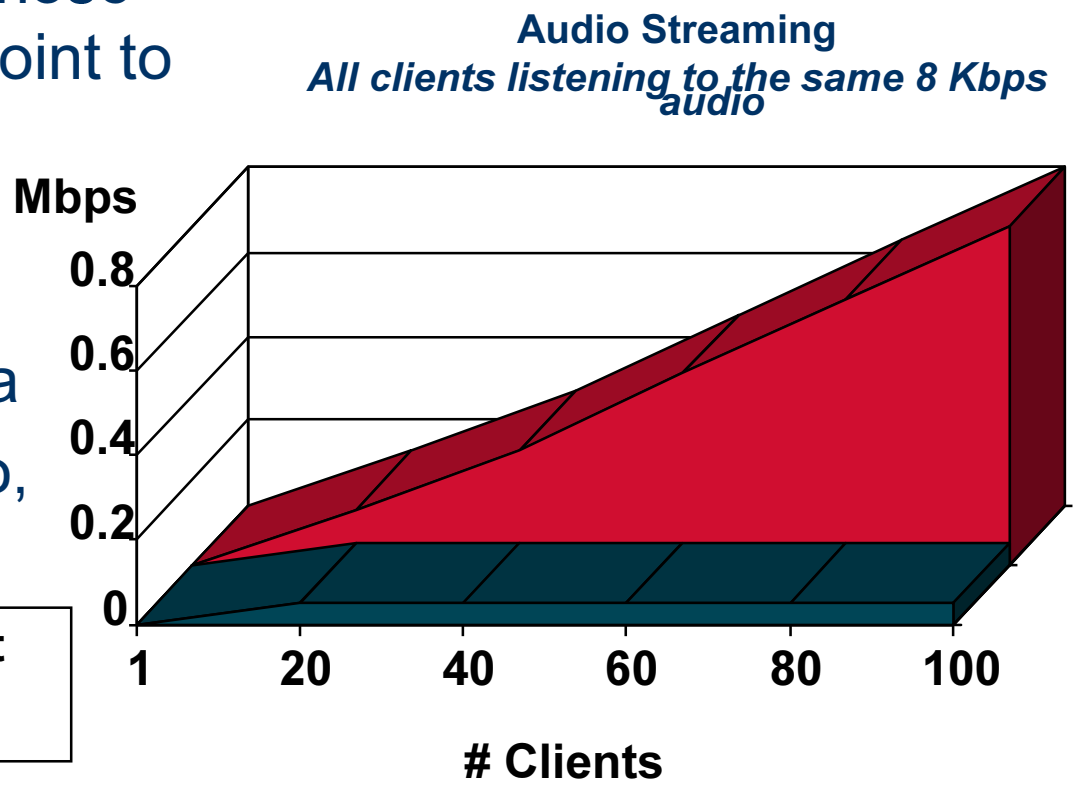
- Wysyłanie w tym samym czasie tych samych danych do wielu użytkowników
- Lepsze wykorzystanie pasma
- Mniejsze wymagania dla serwerów i routerów

Porównanie transmisji unicast i multiast



Zalety transmisji multicastowej

- Zwiększona wydajność dla aplikacji typu point to multipoint
- Optymalizacja wykorzystania dostępnego pasma
- Konferencje (audio, video)



Wady transmisji multicastowej

- Bazują na przesyłaniu pakietów UDP
- Dostarczanie pakietów **Best Effort**.
- Brak zapobiegania przeciążenia w sieci
- Możliwość duplikacji pakietów
- Brak możliwości kontroli kolejności nadchodzących pakietów

Aplikacje

Live TV and Radio Broadcast
to the Desktop

Corporate Broadcasts

Distance Learning

Multicast File Transfer
Data and File Replication





Training

Video Conferencing

Video-On-Demand

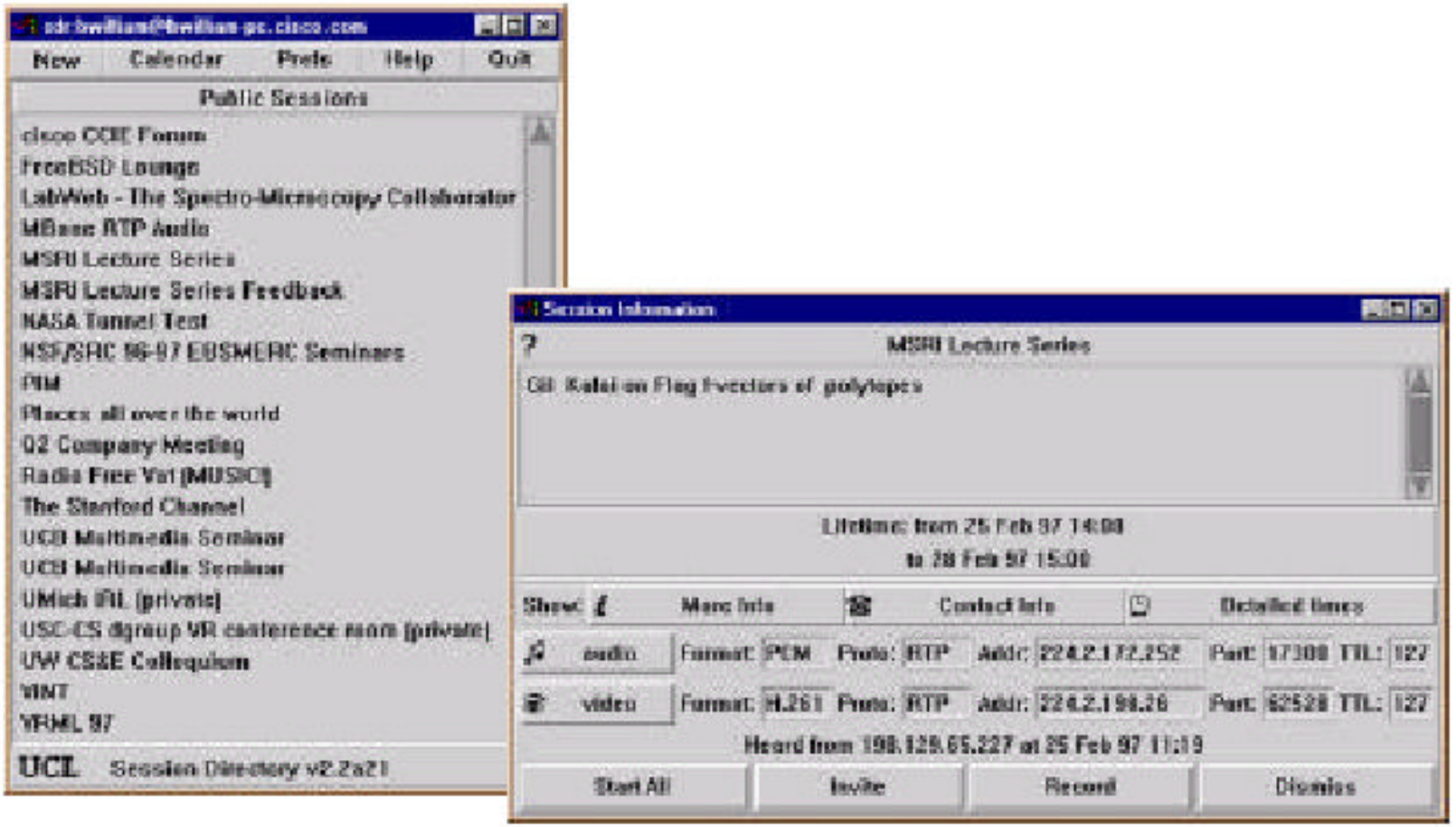
Whiteboard/Collaboration

Real-Time Data Delivery—Financial

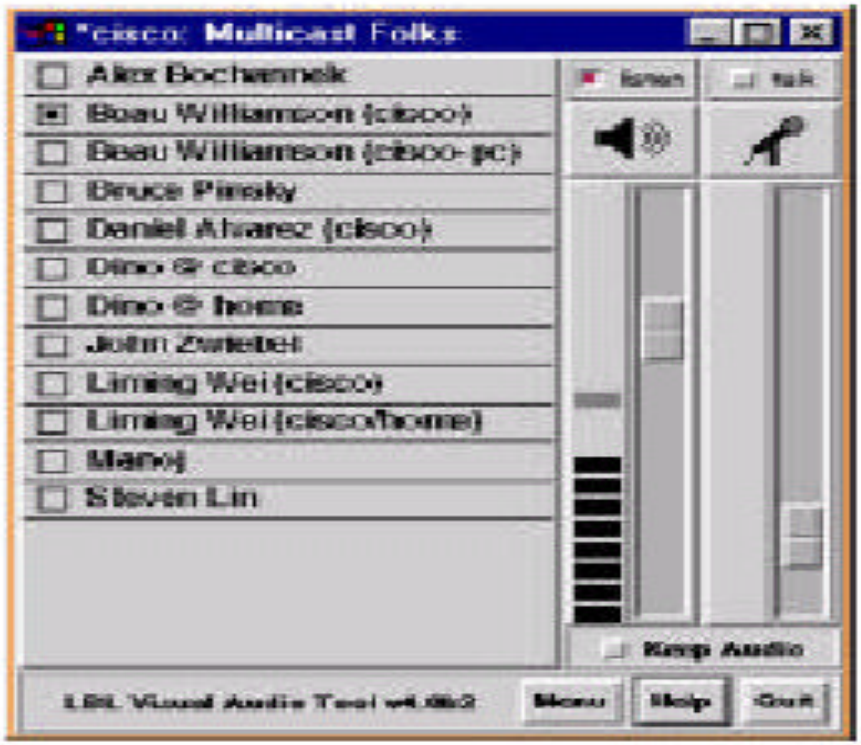
Aplikacje

- sdr – session directory
  - ✓ tworzy listędostępnych sesji multicastowych
  - ✓ uruchamia odpowiednie aplikacje multicastowe
- vat – audio konferencje
  - ✓ obsługuje formaty PCM, DVI, GSM, LPC4
- vic – video konferencje
  - ✓ obsługuje format H.261
- wb – white board
  - ✓ narzędzie do wspólnego rysowania
- Viewers (Cisco IP/TV)
  - ✓ przeglądarka obsługujące wiele formatów min. mp3, mpeg
  - ✓ współpracuje z content managerem (Cisco IP/TV server) lub nasłuchuje sesji rozgłaszanych przez SAP

sdr – Session Directory

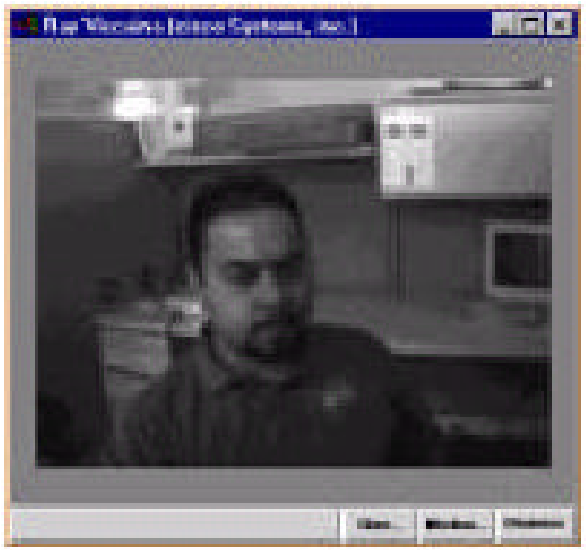


vat – Audio Conferencing



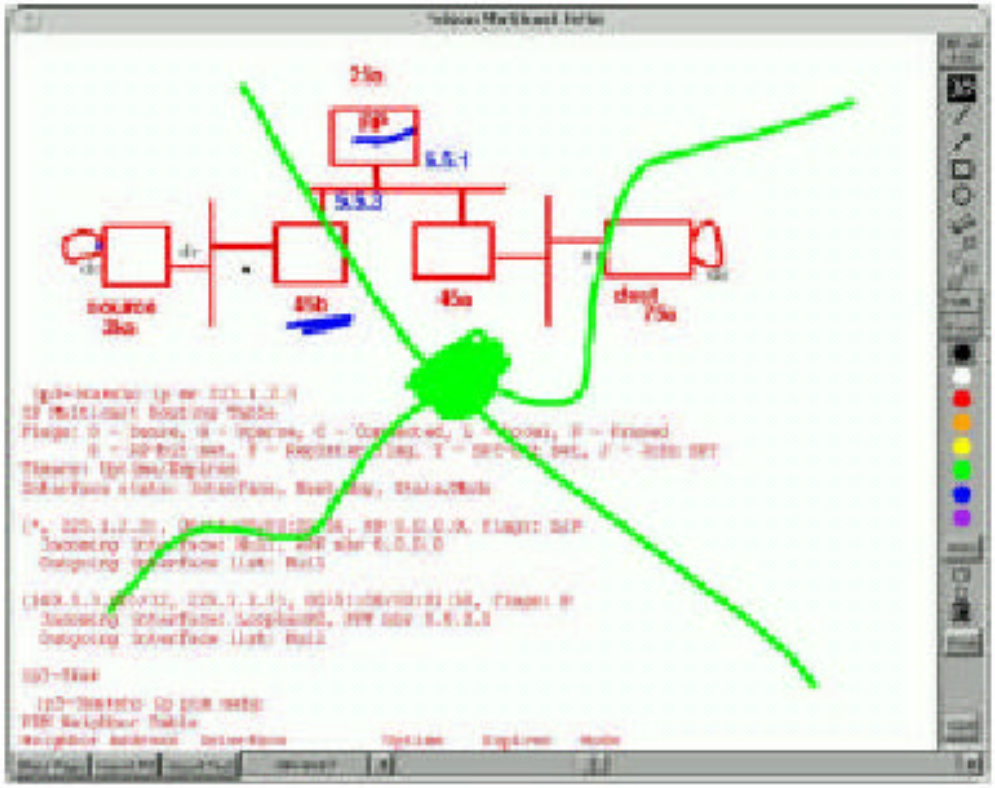


vic – Video Conferencing





wb – White Board





Model usług IP Multicast

- RFC 1112 (Host Ext. for Multicast Support)
- Każda grupa multicastowa określana jest przez adres IP klasy D
- Członkowie grupy mogą być zlokalizowani w dowolnym miejscu w sieci Internet
- Server (nadawca) nie musi być członkiem grupy
- Routery używają protokołu routingu IP multicast aby odpowiednio kierować ruch do poszczególnych grup

Model usług IP Multicast

- Adres IP grupy multicastowej
- Klasa D – ustawione 3 bardziej znaczace bity (224.0.0.0)
- Zakres adresów: 224.0.0.0 do 239.255.255.255
- Adresy zarezerwowane przez IANA
- Zakres zarezerwowany 224.0.0.0 – 224.0.0.255 np.:
  - 224.0.0.1 wszystkie systemy multicastowe w podsieci
  - 224.0.0.2 wszystkie routery w podsieci
- Adresy tymczasowe, przydzielane dynamicznie
- Zasięg globalny: 224.0.1.0 – 238.255.255.255
- Zasięg ograniczony: 239.0.0.0 – 239.255.255.255
  - Site-local: 239.253.0.0/16
  - Organization-local: 239.192.0.0/14

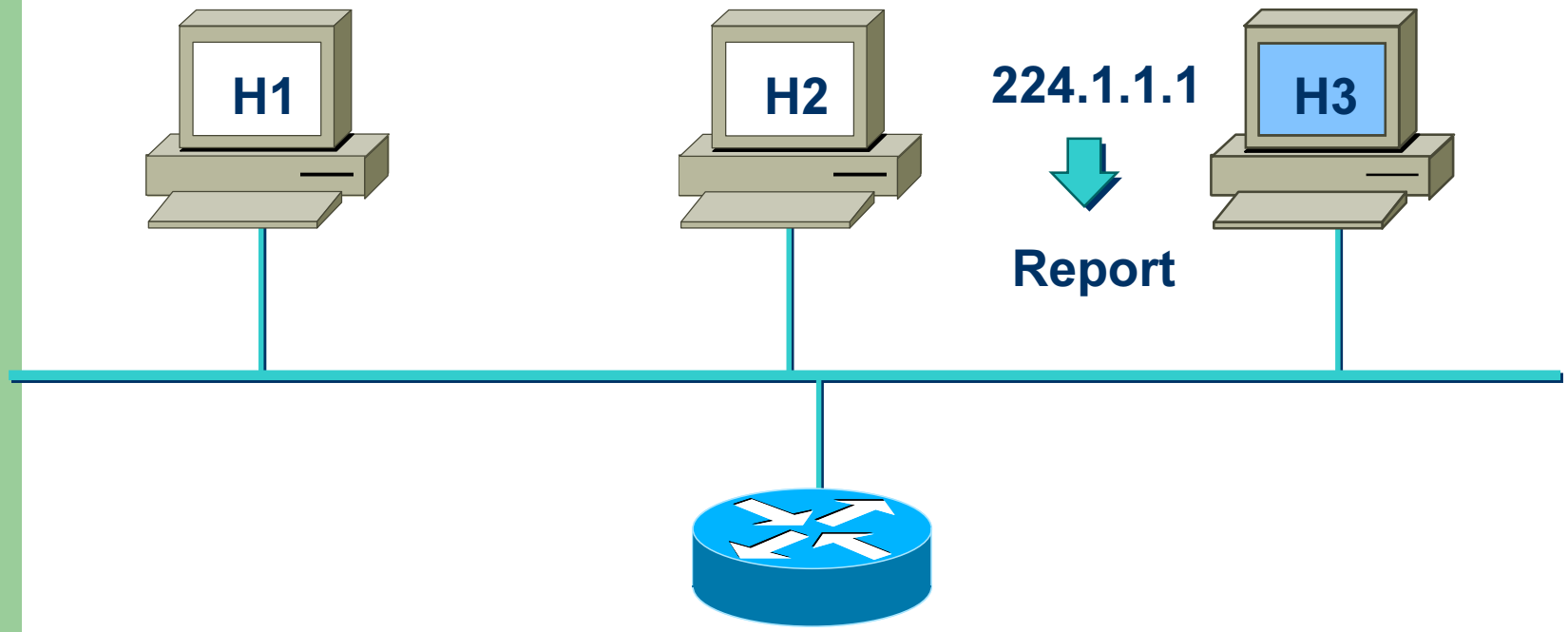
Adresacja IP multicast

- Dynamiczny przydział adresów grupowych
  - przy wykorzystaniu aplikacji SDR
    - Sesje/grupy rozgłaszane przy pomocy adresów „well-known”
    - Kolizje adresów wykrywane w czasie tworzenia sesji
    - Słabo skalowalne
  - w przyszłości przy wykorzystaniu Multicast Address Set-Claim (MASC)
    - Hierarchiczny, dynamiczny schemat przydziału adresów
    - Bardzo skomplikowana architektura

Adresacja IP multicast

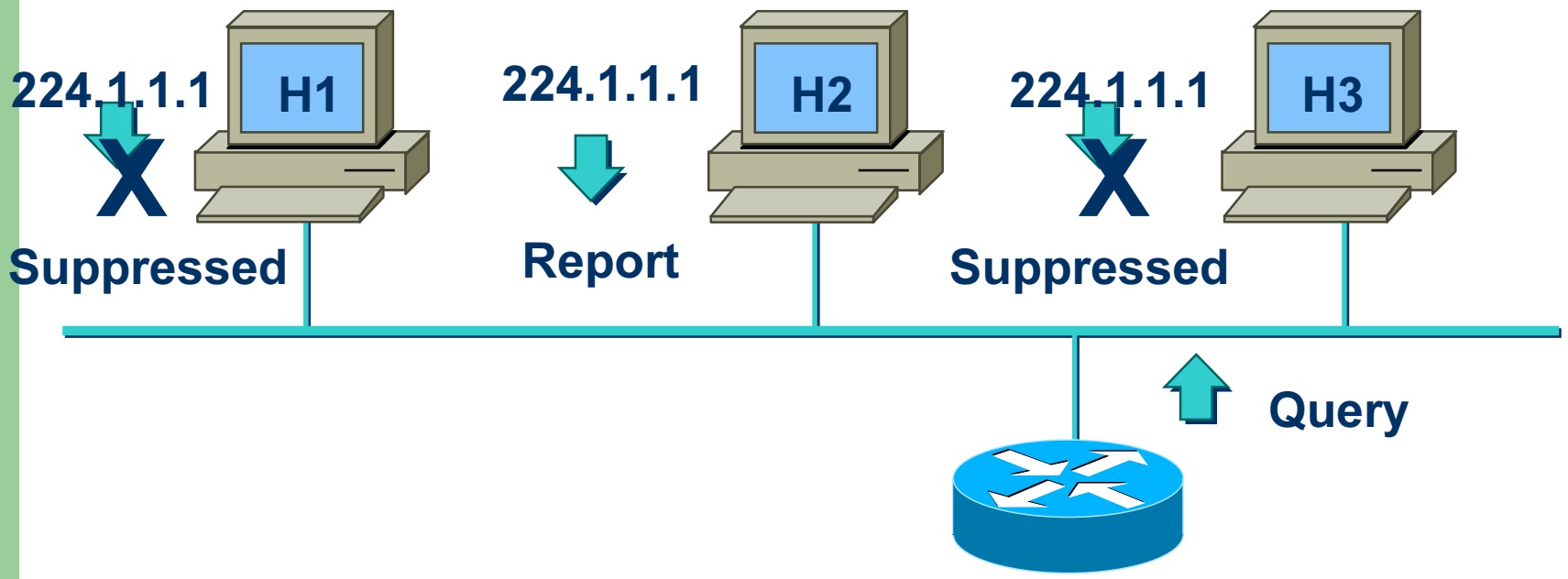
- Statyczny przydział adresów
  - Tymczasowa metoda, mająca na celu zaspokojenie wysokiego zapotrzebowania na adresy grupowe
  - Zakres adresów: 233.0.0.0 – 233.255.255.255
    - Numer AS umieszczany w dwóch środkowych oktetach adresu IP
    - Ostatni oktet określa grupę multicastową w danym AS
  - Zdefiniowany w drafcie IETF
    - „draft-ietf-mboned-glop-addressing-xx.txt”

Sygnalizacja IGMPv1



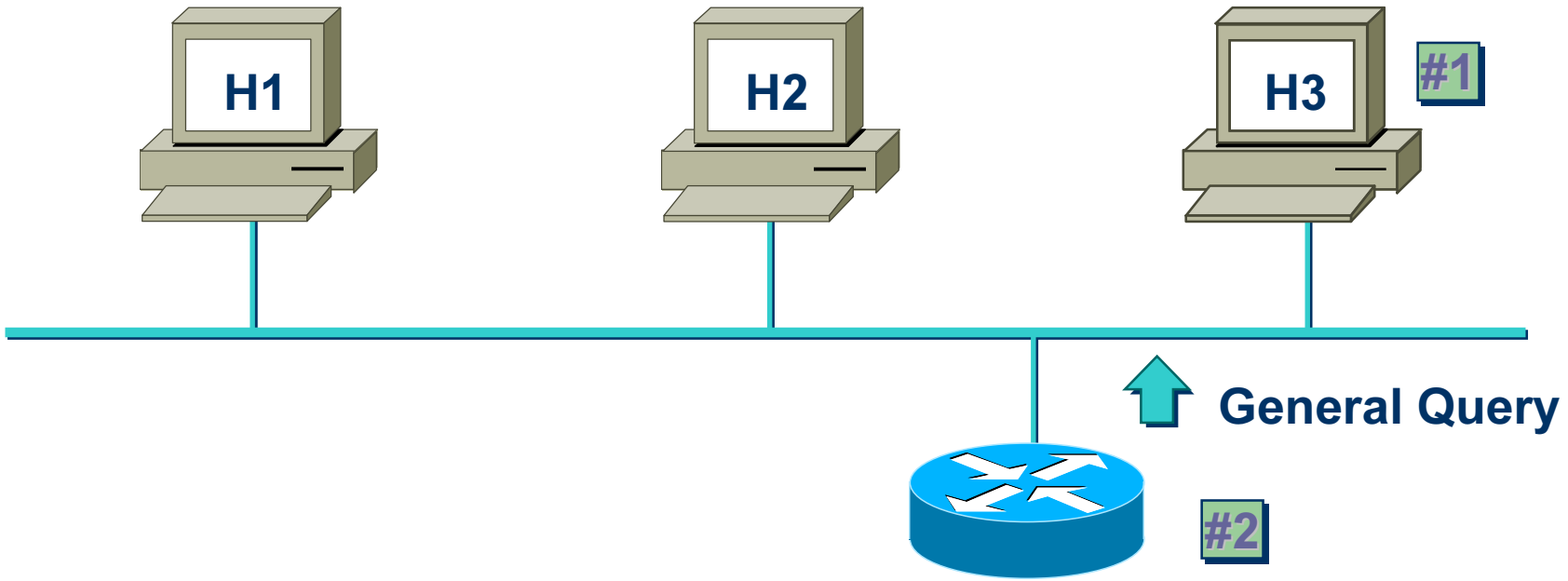
- Host wysyła żądanie IGMP dołączenia do grupy

Sygnalizacja IGMPv1



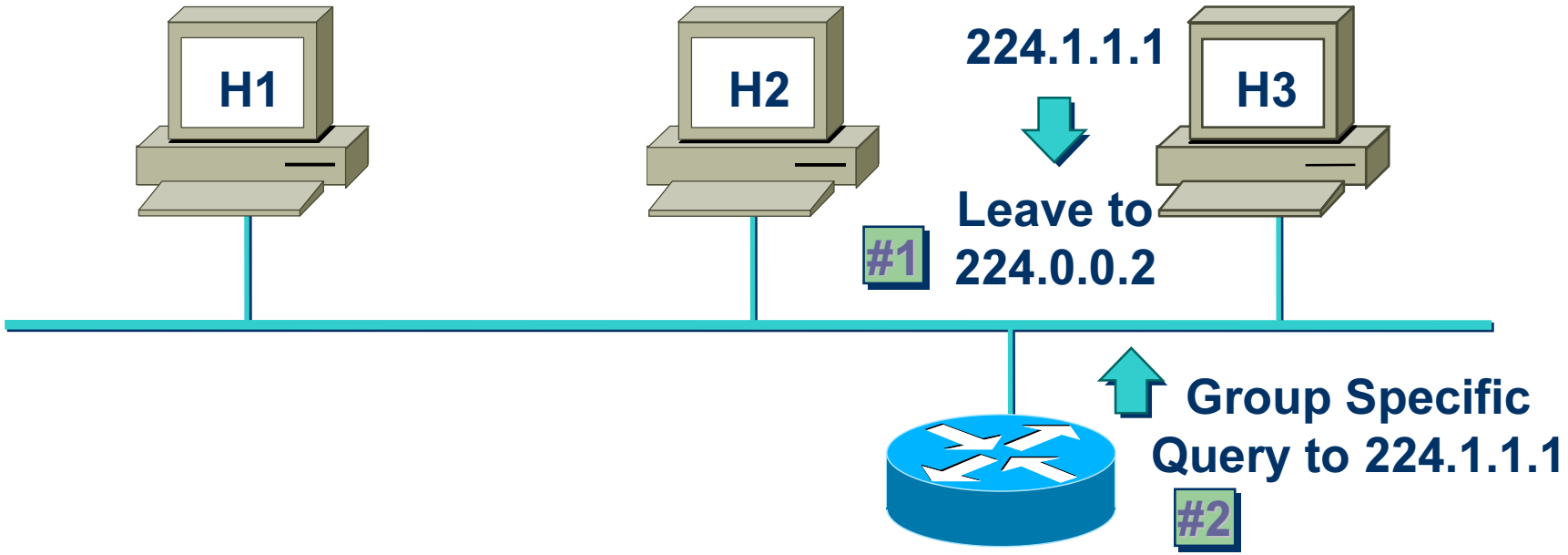
- Router okresowo wysyła pytania do hostów (224.0.0.1)
- Na pytania odpowiada jeden z hostów

Opuszczanie grupy (IGMPv1)



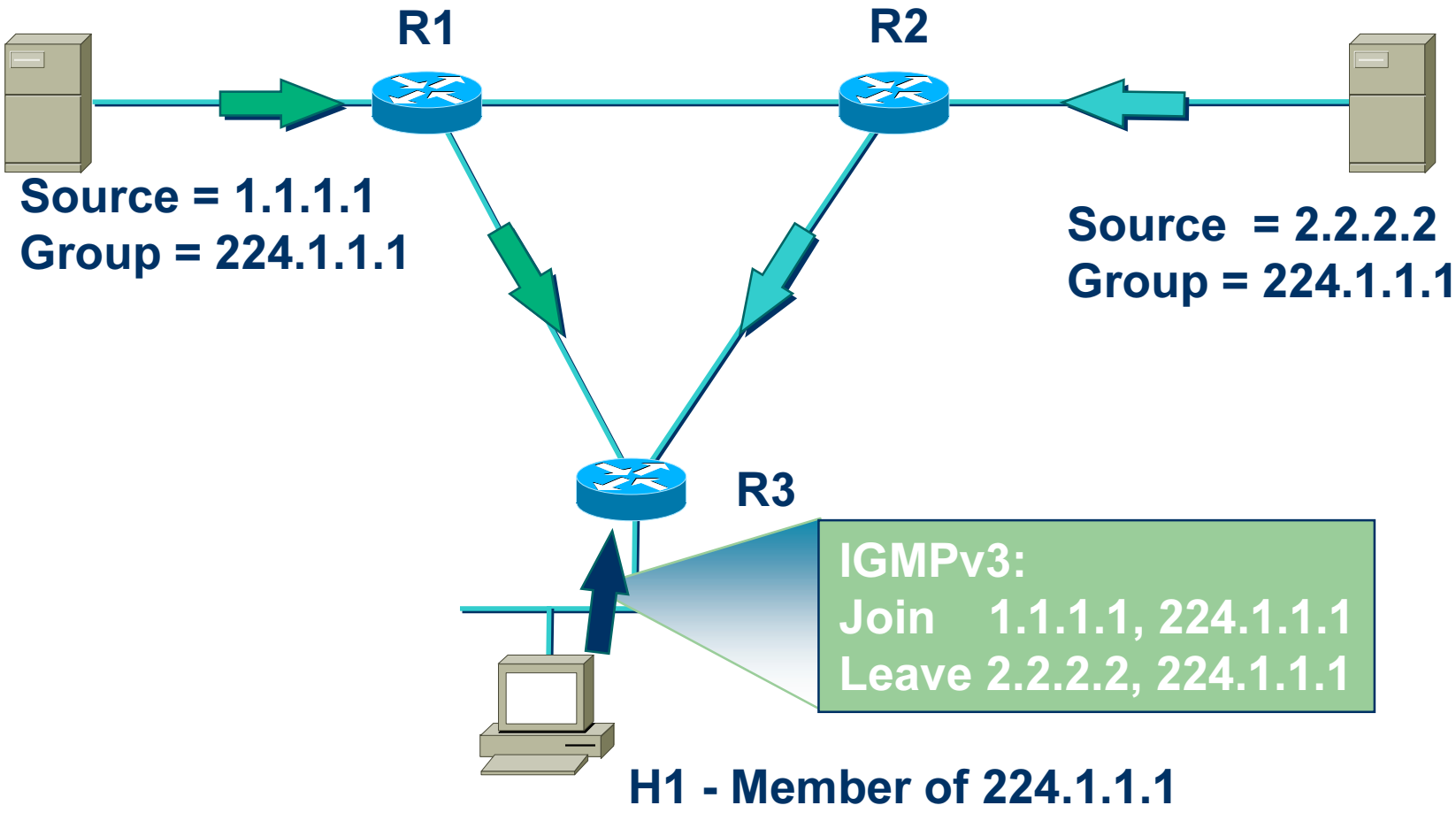
- Host opuszcza grupę
- Router wysyła 3 pytania o odbiorców grupy (60s)
- Nie otrzymuje odpowiedzi IGMP
- Timeout

Opuszczanie grupy (IGMPv2)



- Host wysyła wiadomość „Leave” (224.0.02)
- Router wysyła pytanie o pozostałych członków grupy 224.1.1.1
- Nie otrzymuje odpowiedzi IGMP w ciągu 3 sekund
- Timeout

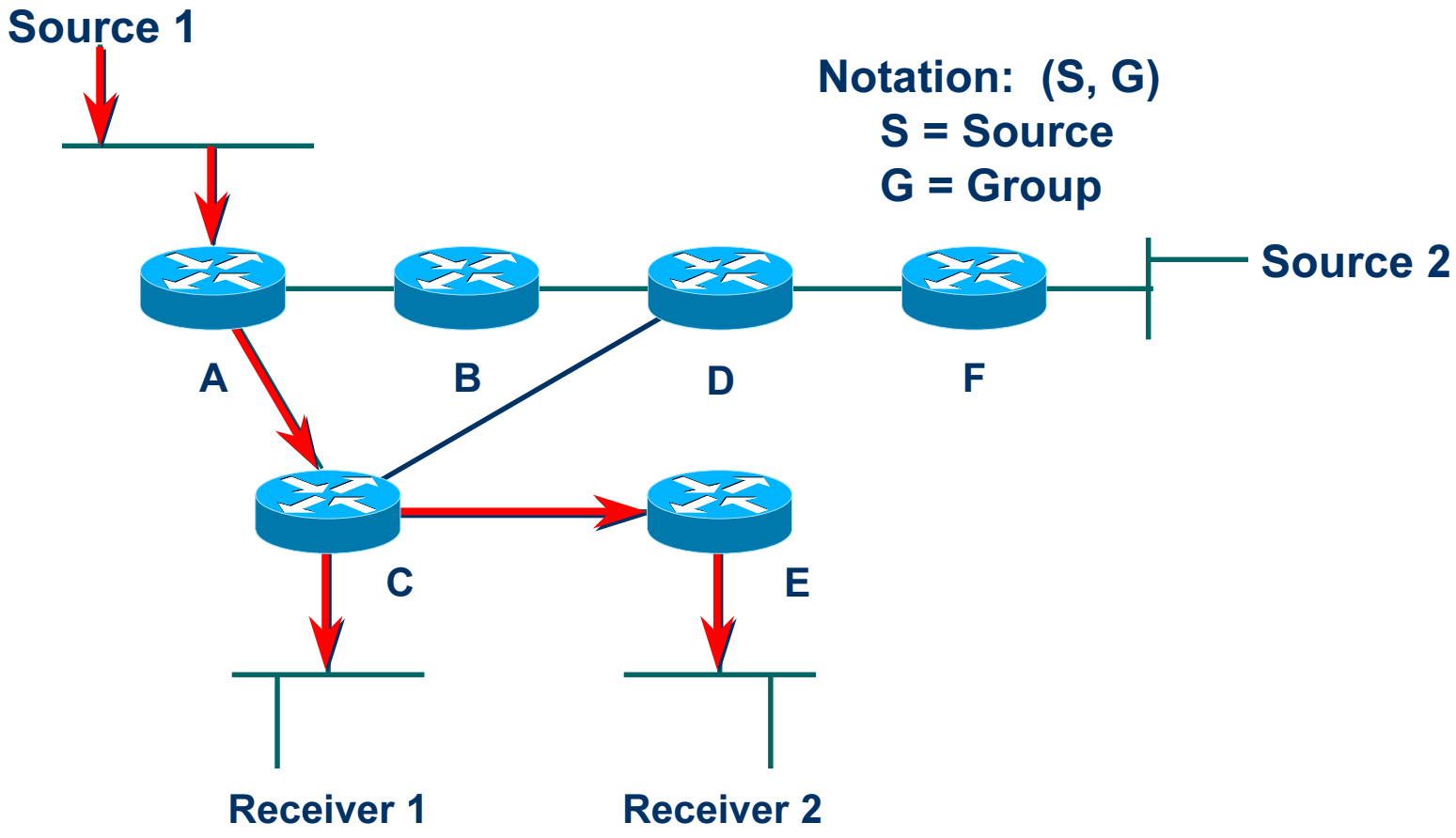
Sygnalizacja IGMPv3



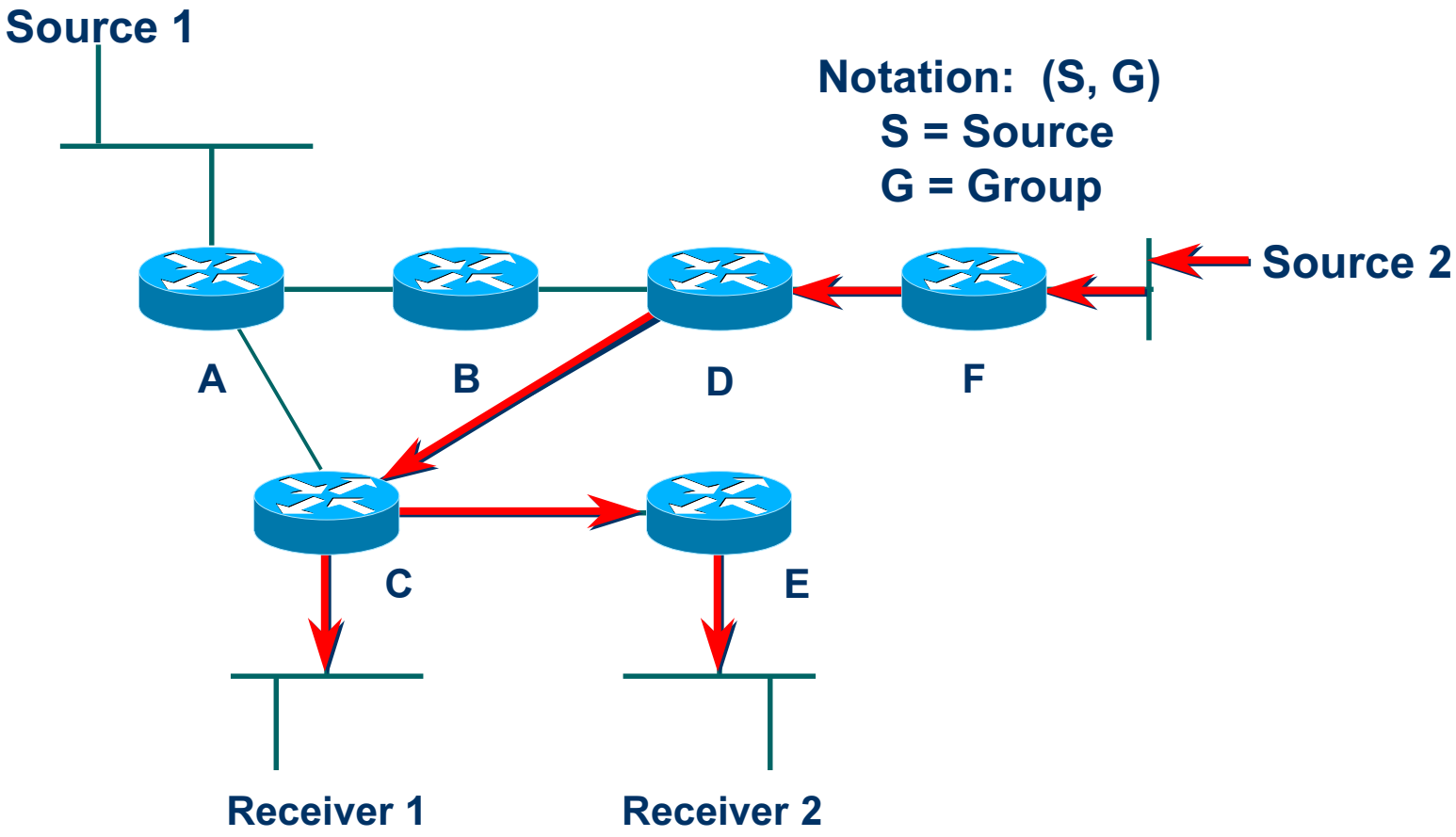
Podstawy protokołu multicastowego

- Drzewa dystrybucyjne (MDT)
  - Określają ścieżkę po której transmitowane są pakiety od źródła do odbiorców
- Przekazywanie pakietów
  - Używanie adresu źródłowego do podejmowania decyzji o przesyłaniu pakietów
- Typy protokołów multicastowych
  - Dense mode
  - Sparse mode

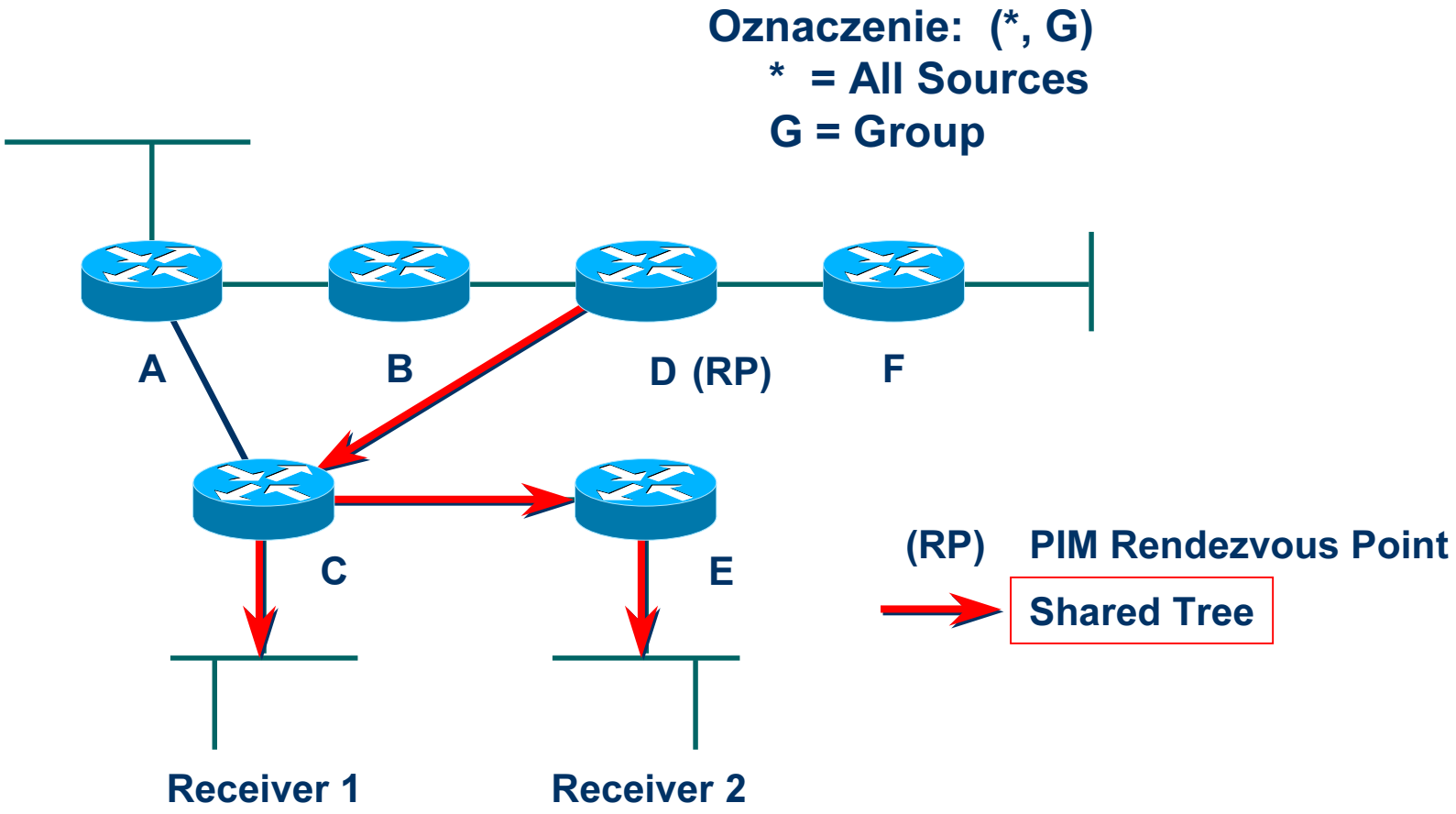
Drzewa dystrybucyjne
(Shortest Path/Source Distribution Tree)



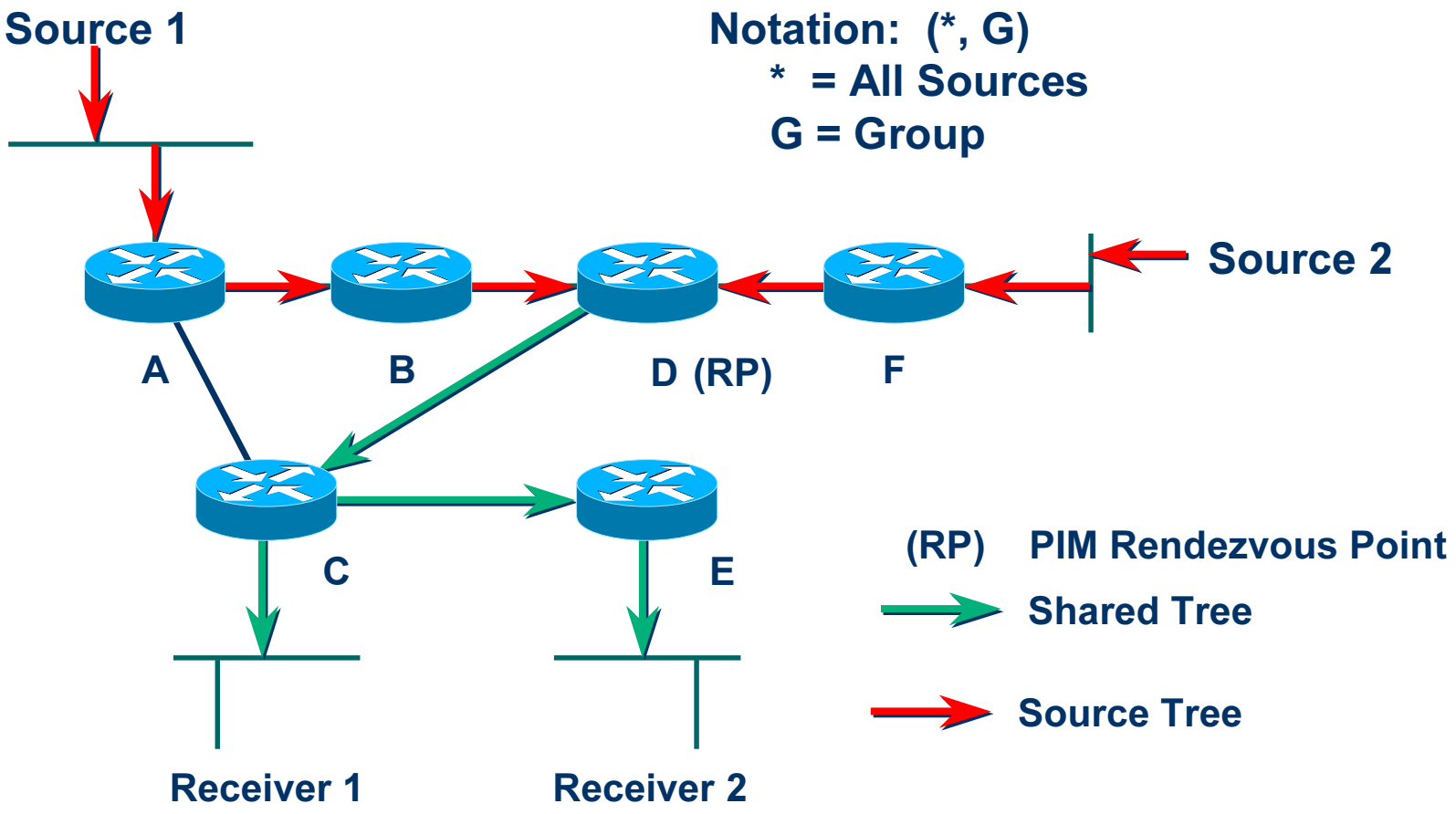
Drzewa dystrybucyjne
(Shortest Path/Source Distribution Tree)



Drzewa dystrybucyjne
(Shared Distribution Tree)



Drzewa dystrybucyjne
(Shared Distribution Tree)



Drzewa dystrybucyjne
Charakterystyka

- Source / Shortest Path Trees:
  - Większe zużycie pamięci routera ze względu na informacje (S,G)
  - Możliwość wyboru optymalnej ścieżki od źródła do odbiorców co minimalizuje opóźnienia
- Shared trees
  - Mniejsze zużycie pamięci routera (*.G)
  - Ścieżki mogą być nieoptymalne co może spowodować powstawanie dodatkowych opóźnień

Drzewa dystrybucyjne
Sposób powstawania

- PIM
  - Używa istniejącej tablicy routingu oraz mechanizmu Join/Prune/Graft
- DVMRP
  - Używa tablicy routingu DVMRP oraz mechanizmu Poison-Reverse
- MOSPF
  - Używa rozszerzonego mechanizmu link-state protokołu OSPF
- CBT
  - Używa istniejącej tablicy routingu i mechanizmu Join/Prune/Graft

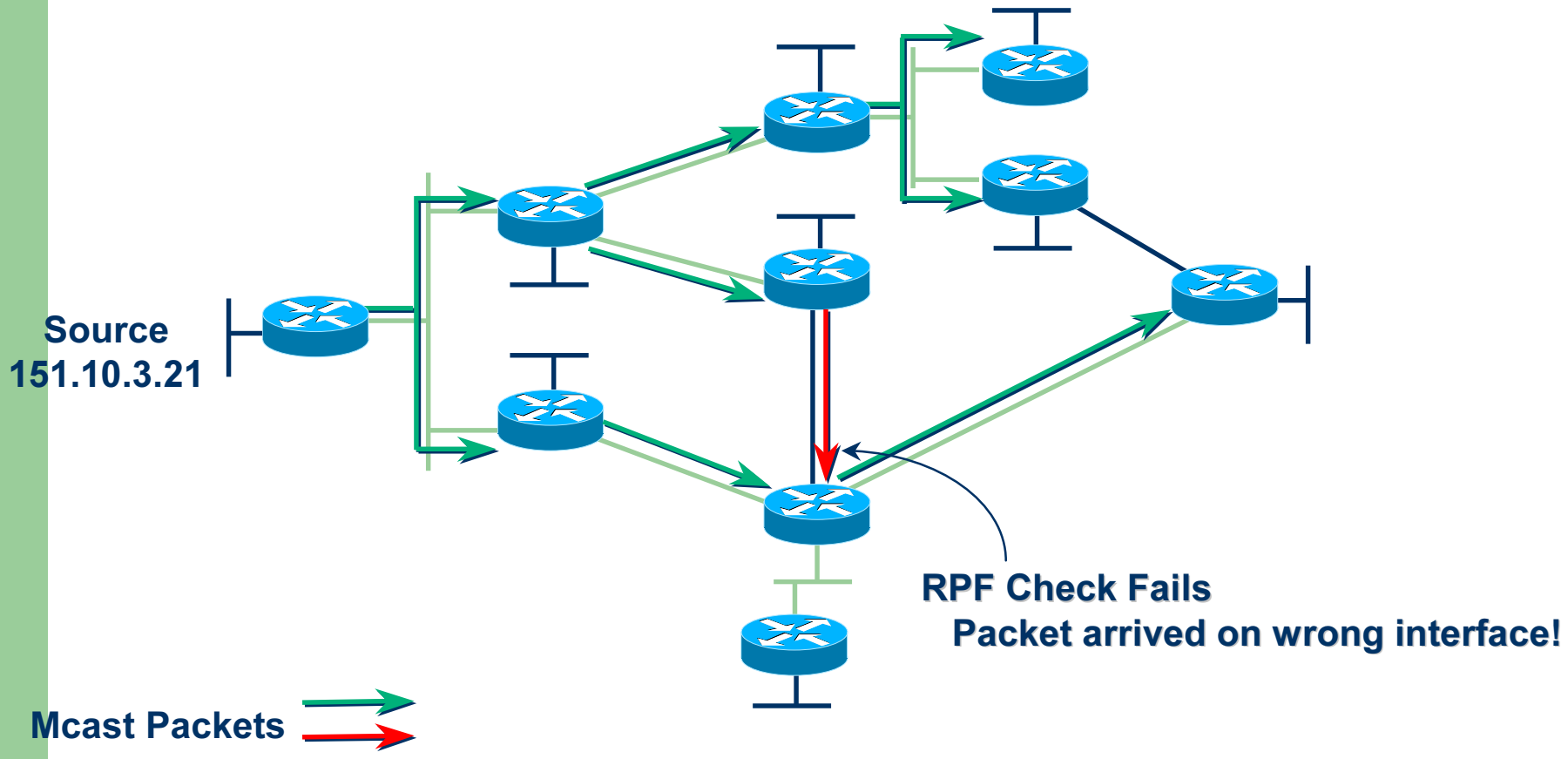
Przekazywanie pakietów multicastowych

- Unicast - Do kogo pakiet ma być wysłany?
- Multicast - Skąd pakiet został przysłany?
- Mechanizm Reverse Path Forwarding (RPF)
  - Broadcast: przesyła pakiety na wszystkie interfejsy poza tym z którego pakiet nadszedł
  - Prune: eliminuje te gałęzie drzewa w których nie ma członków danej grupy multicastowej
  - Selektywne przekazywanie: wymaga własnego zintegrowanego protokołu unicastowego

Reverse Path Forwarding

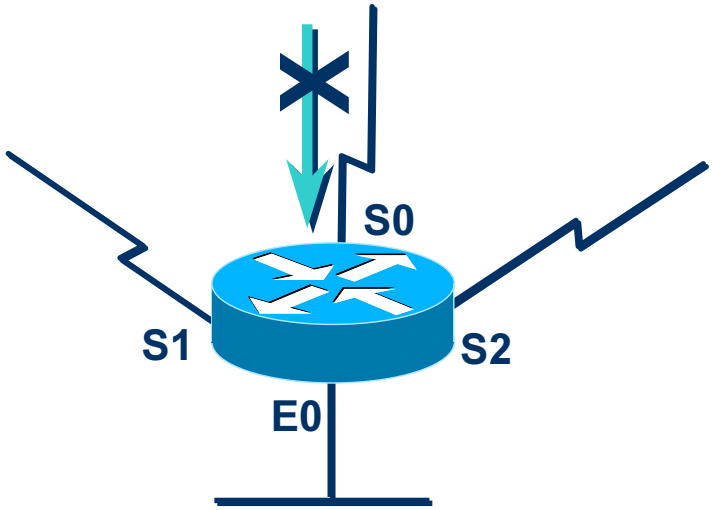
- Router przekazuje dalej tylko ten pakiet multicastowy, który nadszedł z interfejsu w kierunku źródła.
  - Sprawdzana jest tablica routingu pod względem adresu źródłowego w pakiecie
  - Jeżeli datagram nadszedł z interfejsu określonego w tabeli routingu (wskazanego przez adres źródłowy) wówczas może zostać przekazany dalej na interfejsy znajdujące się na liście interfejsów wychodzących
  - W przeciwnym wypadku pakiet nie jest przekazywany
  - Pakiety nigdy nie są przekazywane na interfejs z którego nadeszły

Reverse Path Forwarding



Reverse Path Forwarding

Multicast Packet from
Source 151.10.3.21



RPF Check Fails!

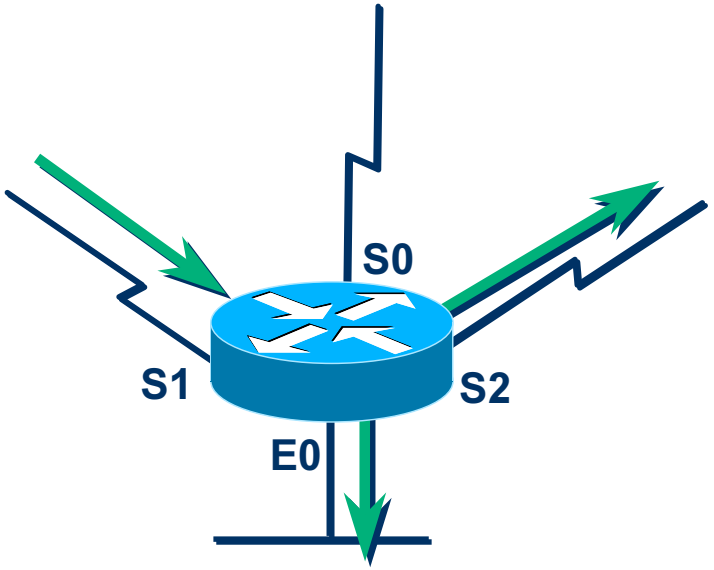
Unicast Route Table	
Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Packet Arrived on Wrong Interface!

Discard Packet!

Reverse Path Forwarding

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)

Typy protokołów multicastowych

- Dense-mode
  - Używa modelu „Push” (zakłada, że na interfejsach są zawsze odbiorcy, chyba, że nadejdzie informacja „Prune”)
  - Ruch początkowo jest transmitowany na wszystkie interfejsy
  - Jeżeli nadejdzie informacja „Prune” pakiety nie są na niego przekazywane
- Sparse-mode
  - Używa modelu „Pull” (zakłada, że na interfejsach nie ma odbiorców dopóki nie nadejdzie wyraźne żądanie)
  - Pakiety transmitowane są tylko na te interfejsy z których nadeszła wiadomość „Join”

Przegląd multicastowych protokołów routingu

- Interdomain
  - DVMRP
    - Distance Vector Multicast Routing Protocol
  - MOSPF
    - Multicast OSPF
  - PIM-DM
    - Protocol Independent Multicast dense mode
  - PIM-SM
    - Protocol Independent Multicast sparse mode
  - CBT
    - Core Based Tree

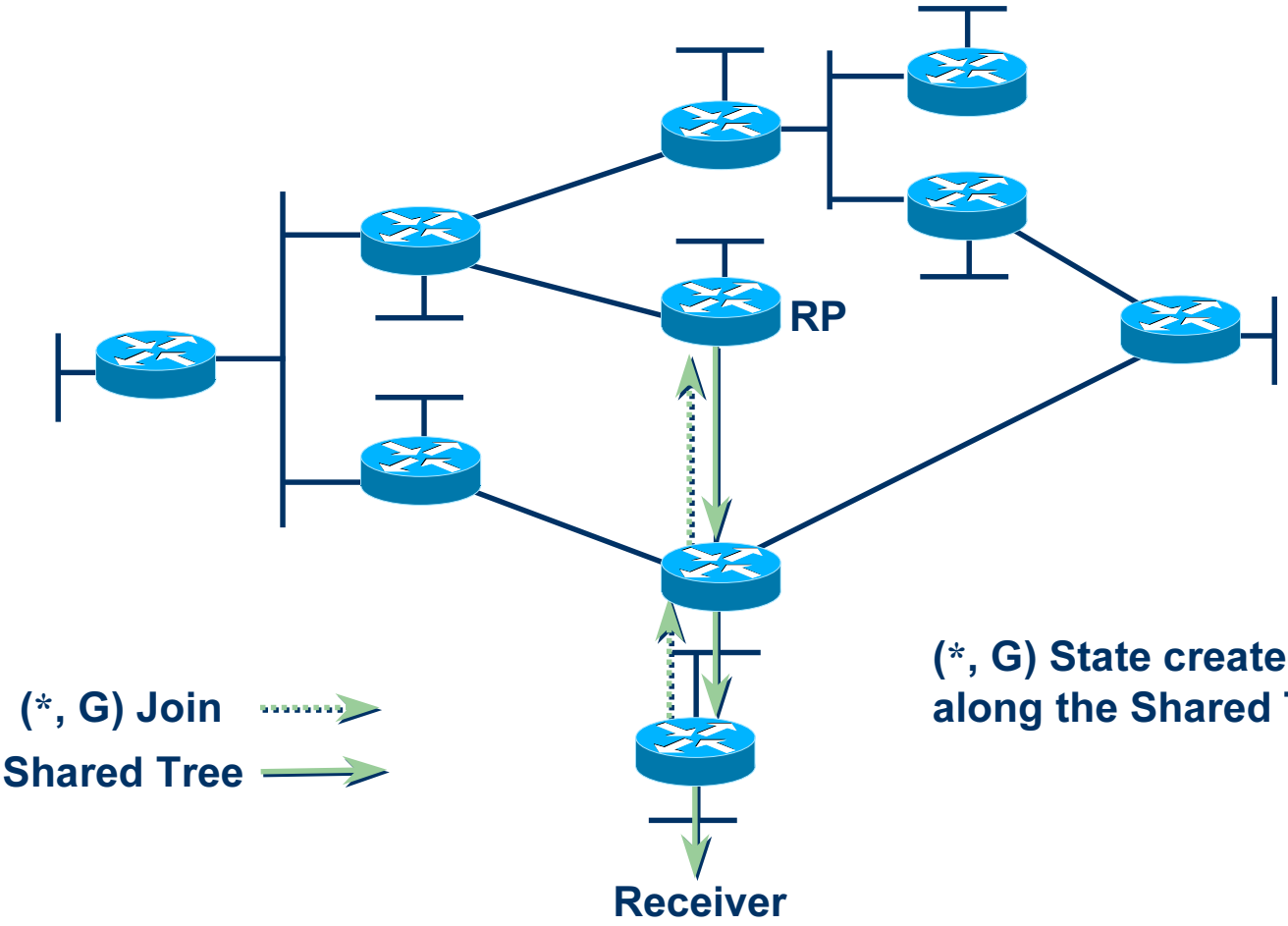
Przegląd multicastowych protokołów routingu

- Intradomain
  - MSDP
    - Multicast Source Discovery Protocol
  - MBGP
    - Multicast BGP
  - BGMP
    - Border Gateway Multicast Protocol
- SSM
  - Source Specific Multicast
  - Wymaga:
    - PIM-SM
    - IGMPv3

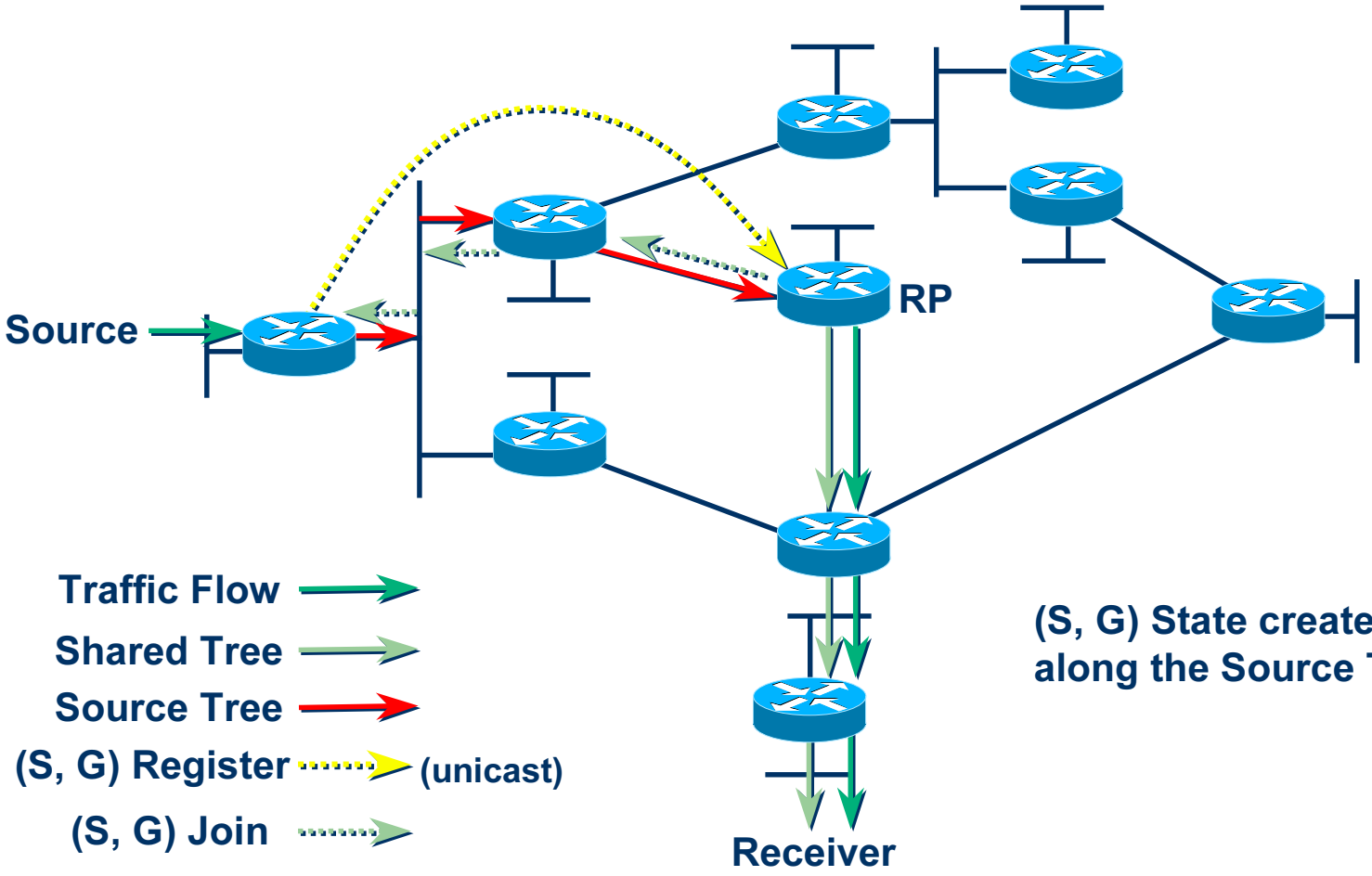
PIM-SM
(RFC 2362)

- Stosuje Source Trees i Shared Trees
  - Zakłada, że żaden host nie chce odbierać pakietów multicastowych do póki nie prześle wyraźnego żądania
- Używa Rendezvous Point (RP)
  - Rejestruje aktywne źródła
  - Odbiorcy dołączani są do drzewa Shared Tree w którym rdzeniem jest RP
- Cechy:
  - Skalowalność
  - Możliwość stosowania w sieciach o różnym rozkładzie odbiorców (sparse lub dense)

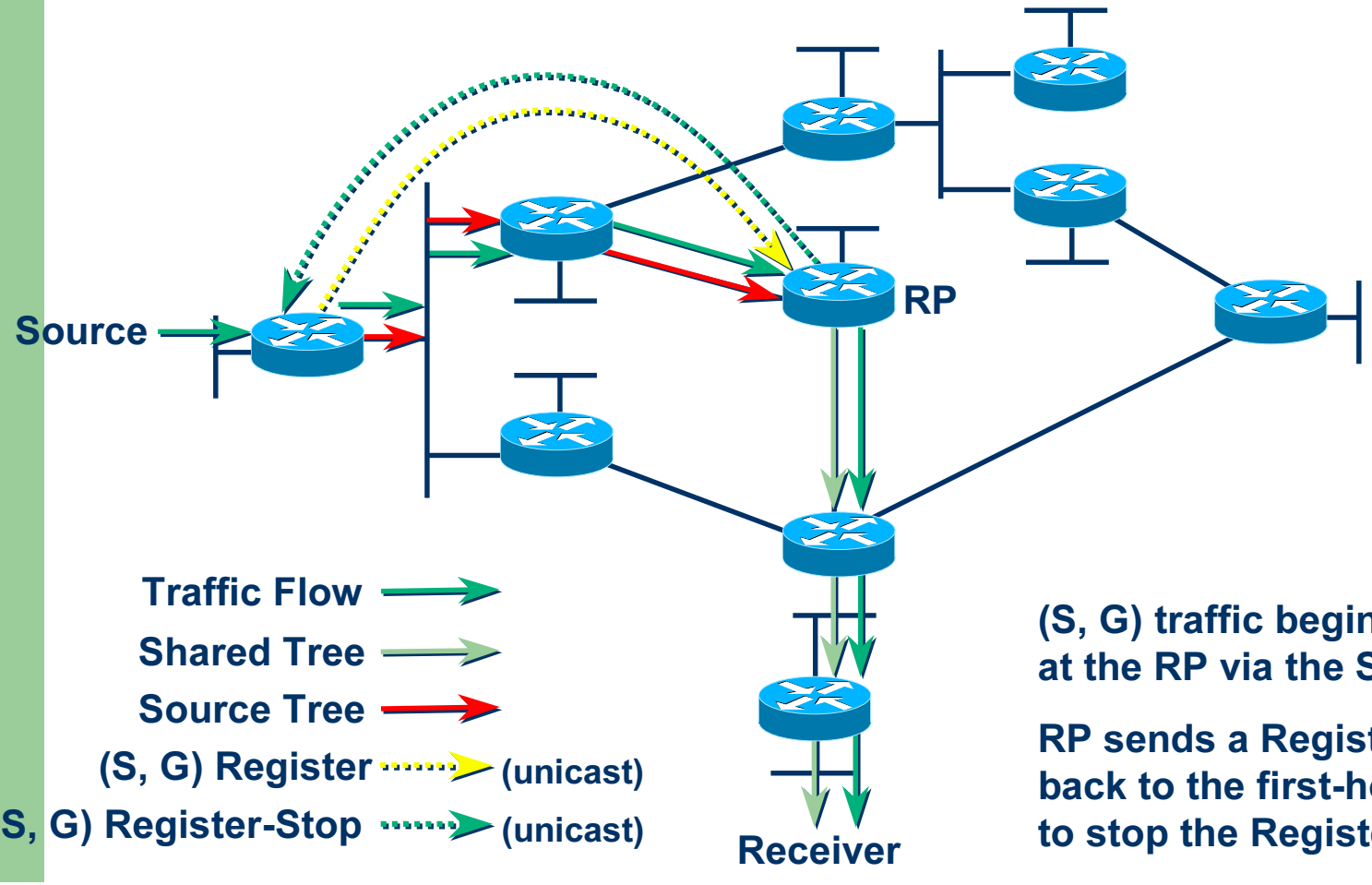
PIM-SM Shared Tree



PIM-SM Rejestracja Źródła



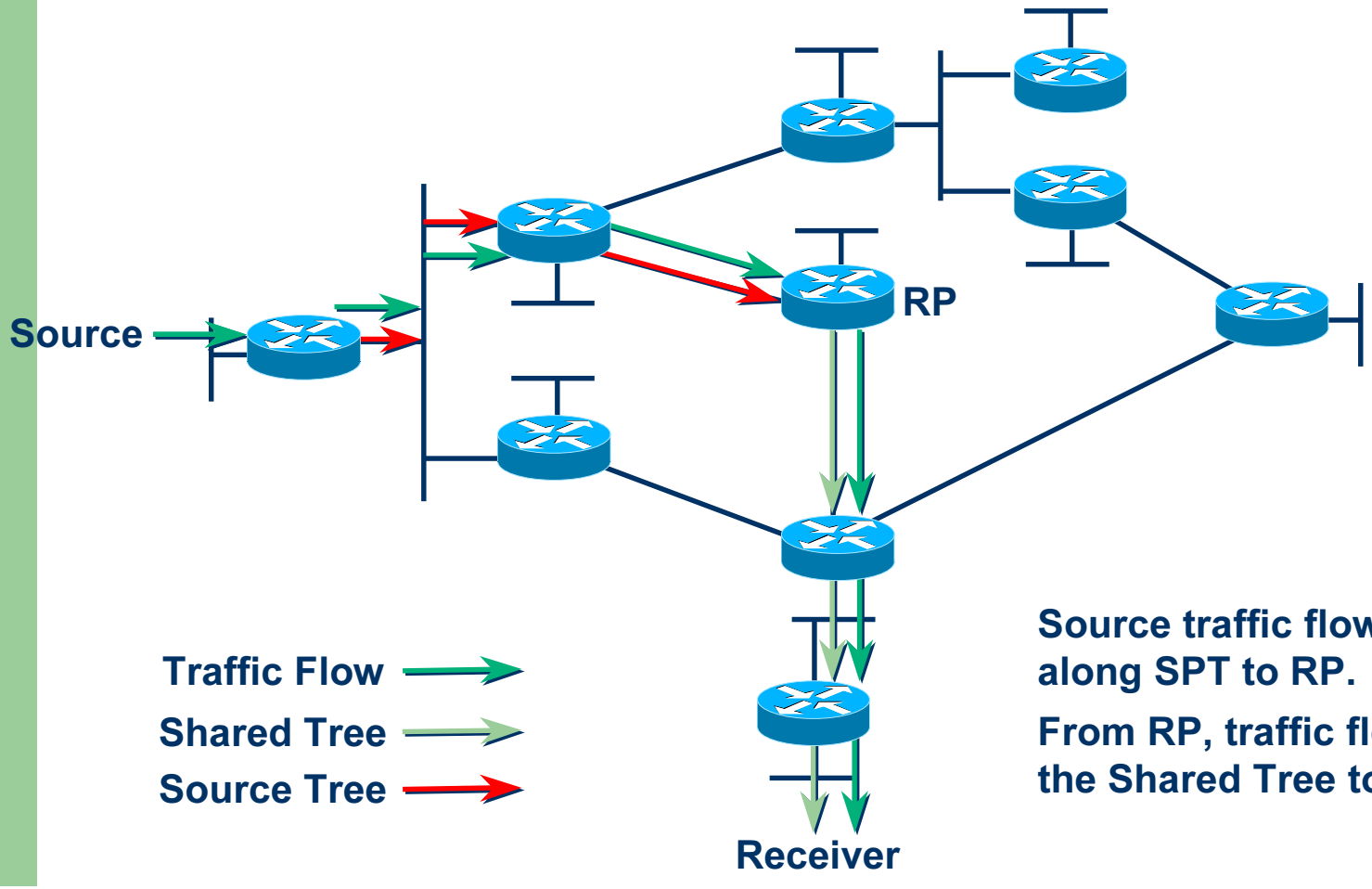
PIM-SM Rejestracja Źródła



(S, G) traffic begins arriving at the RP via the Source tree.

RP sends a Register-Stop back to the first-hop router to stop the Register process.

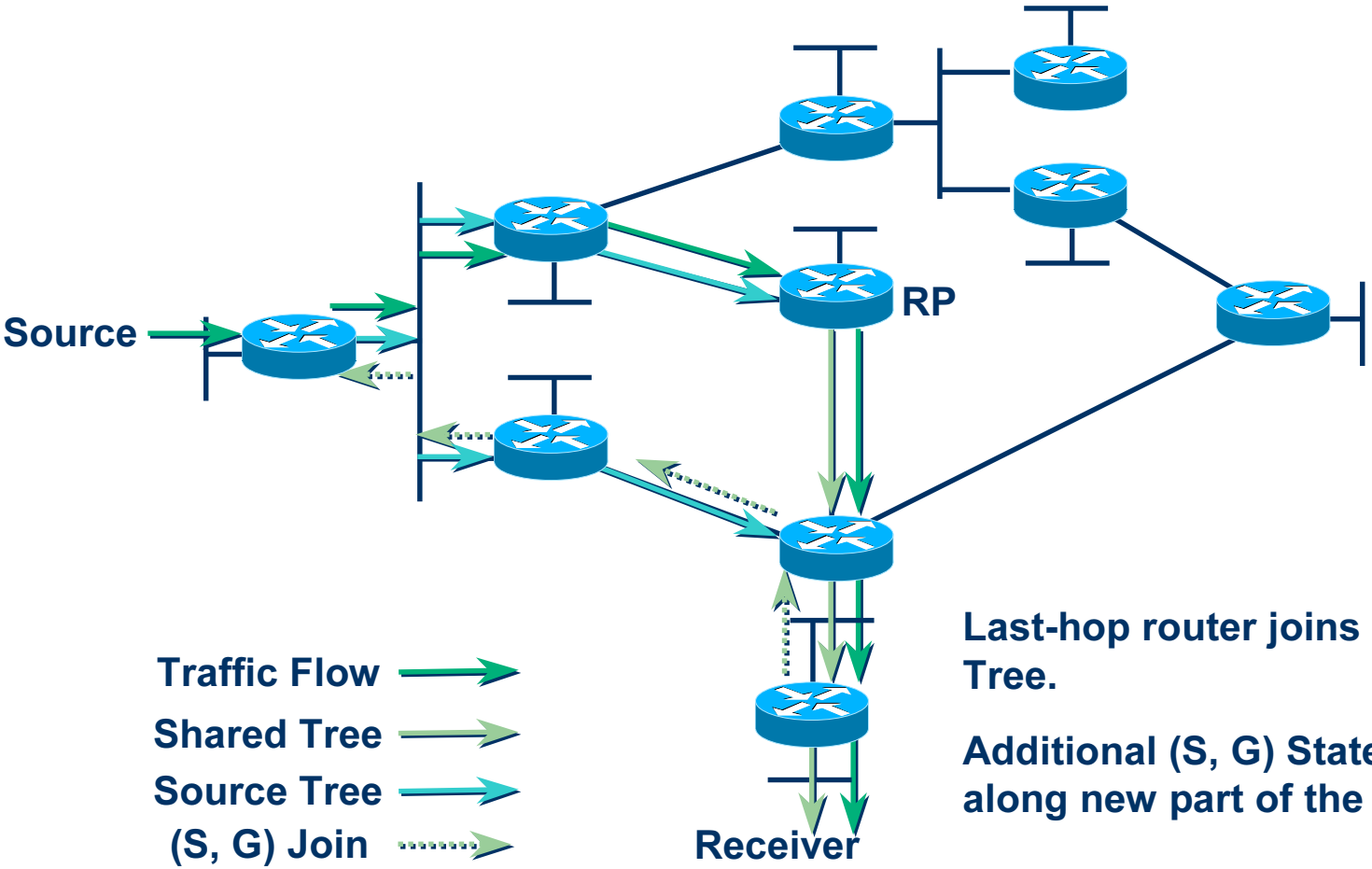
PIM-SM Rejestracja Źródła



Source traffic flows natively along SPT to RP.

From RP, traffic flows down the Shared Tree to Receivers.

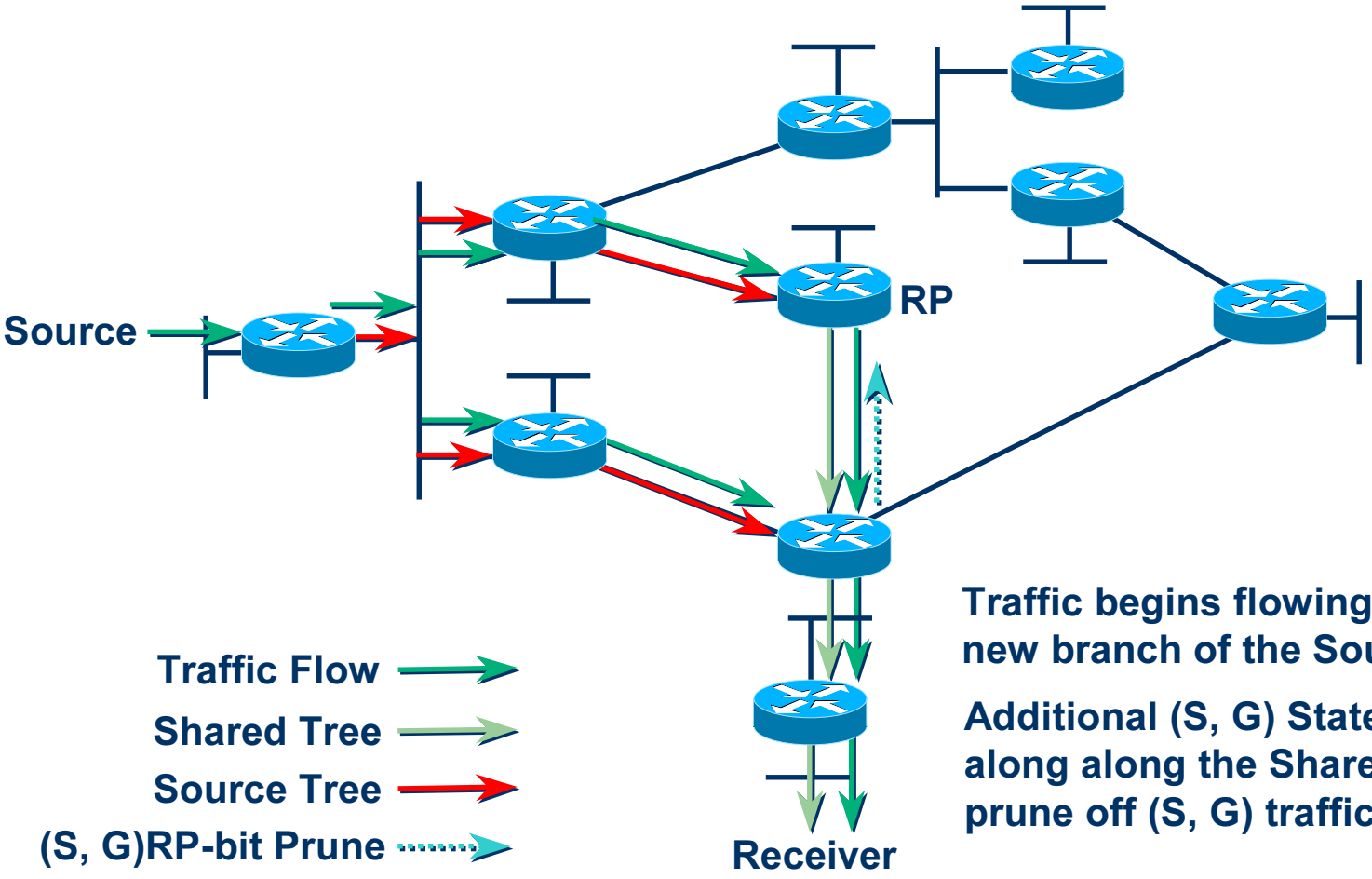
PIM-SM Przełączanie ścieżek



Last-hop router joins the Source Tree.

Additional (S, G) State is created along new part of the Source Tree.

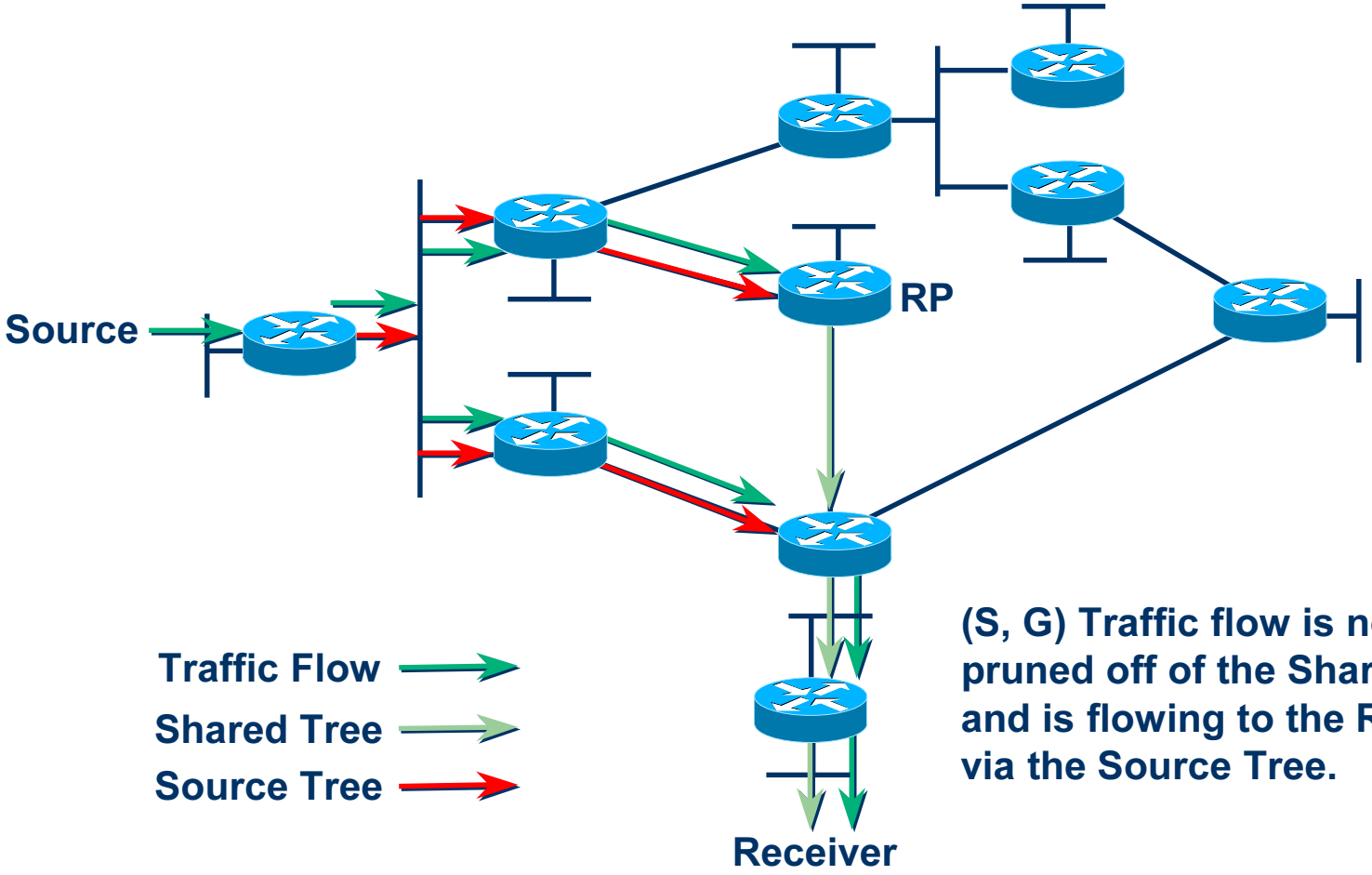
PIM-SM Przełączanie ścieżek



Traffic begins flowing down the new branch of the Source Tree.

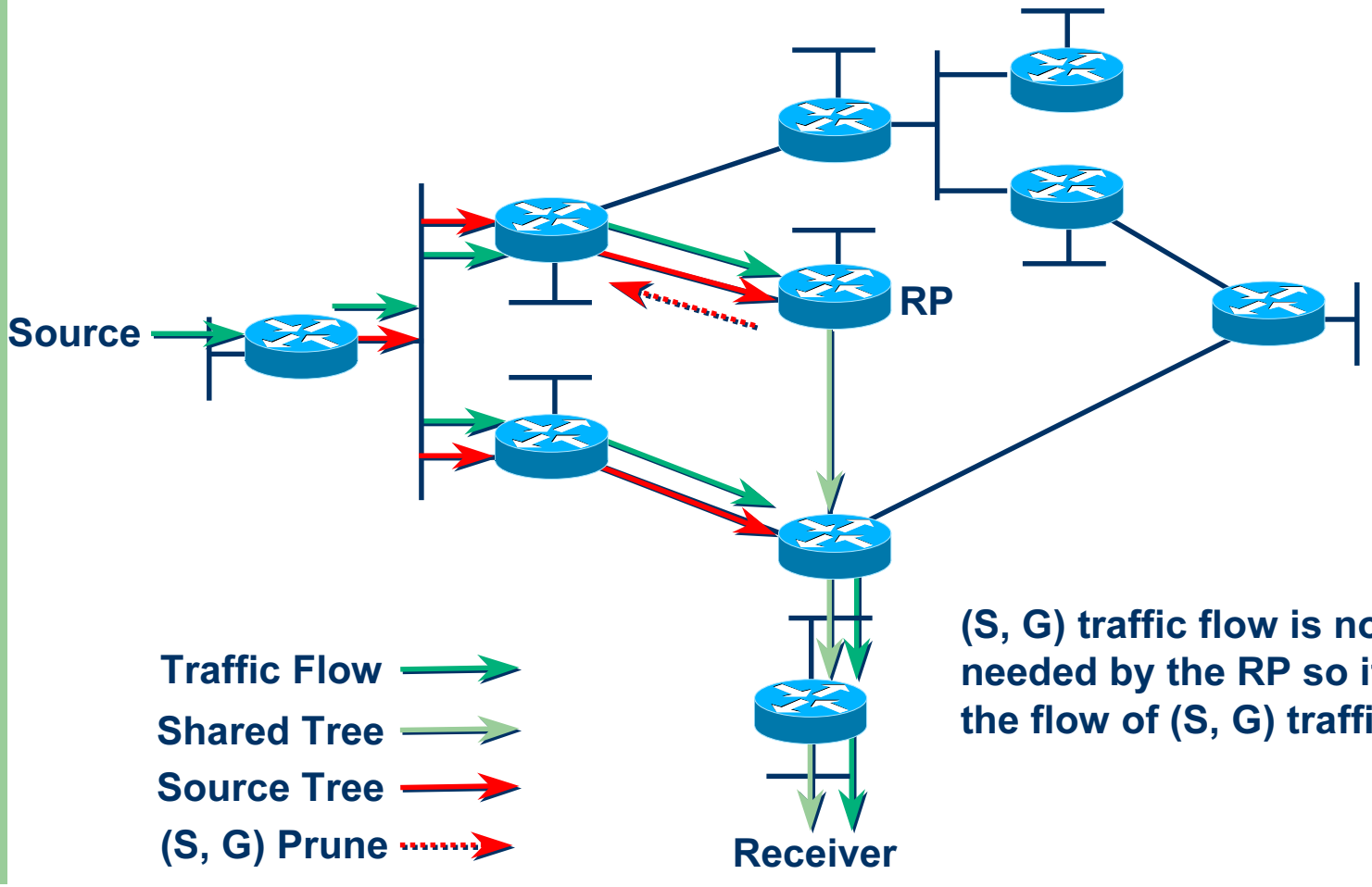
Additional (S, G) State is created along along the Shared Tree to prune off (S, G) traffic.

PIM-SM Przełączanie ścieżek

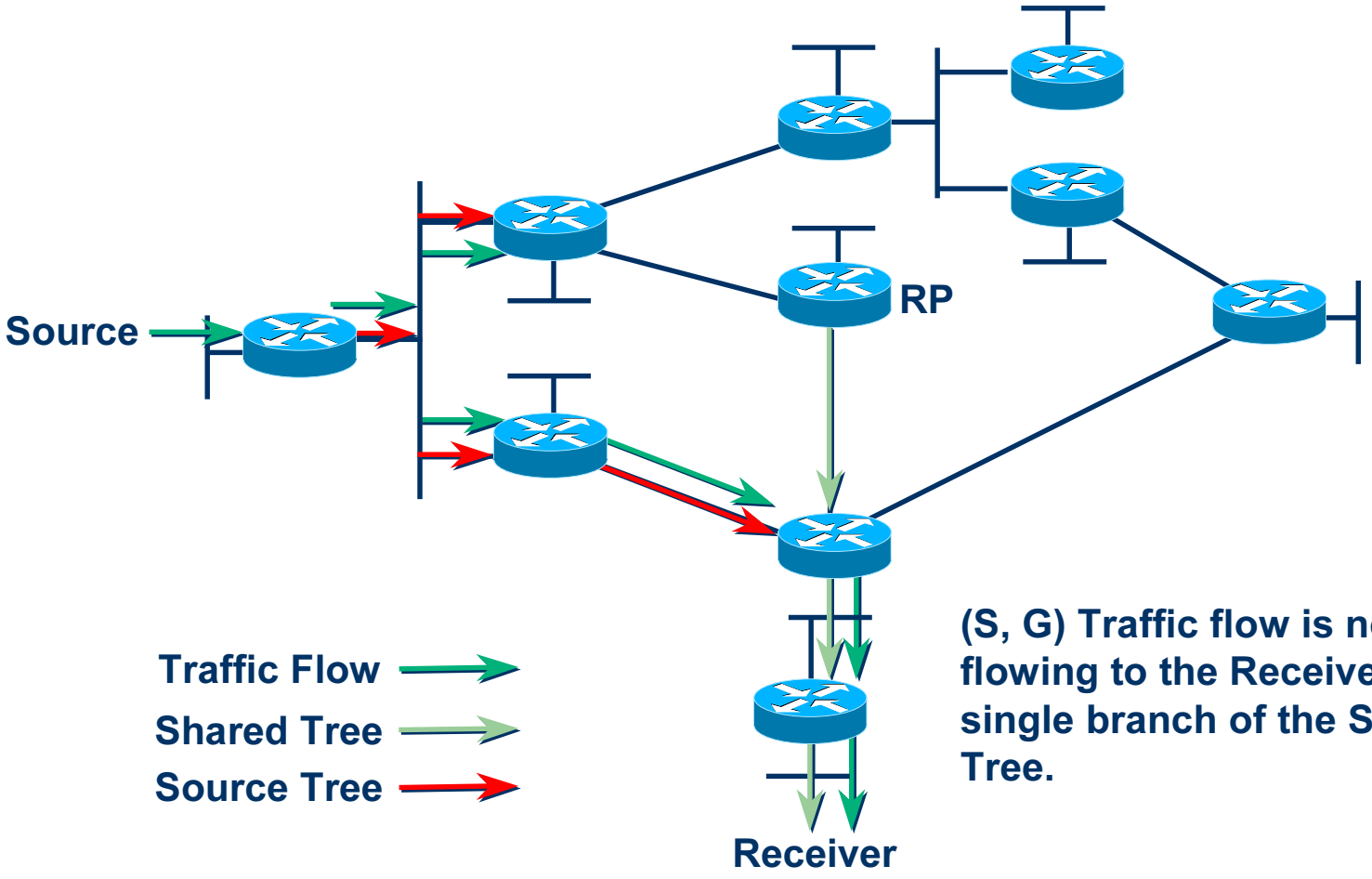


(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver via the Source Tree.

PIM-SM Przełączanie ścieżek



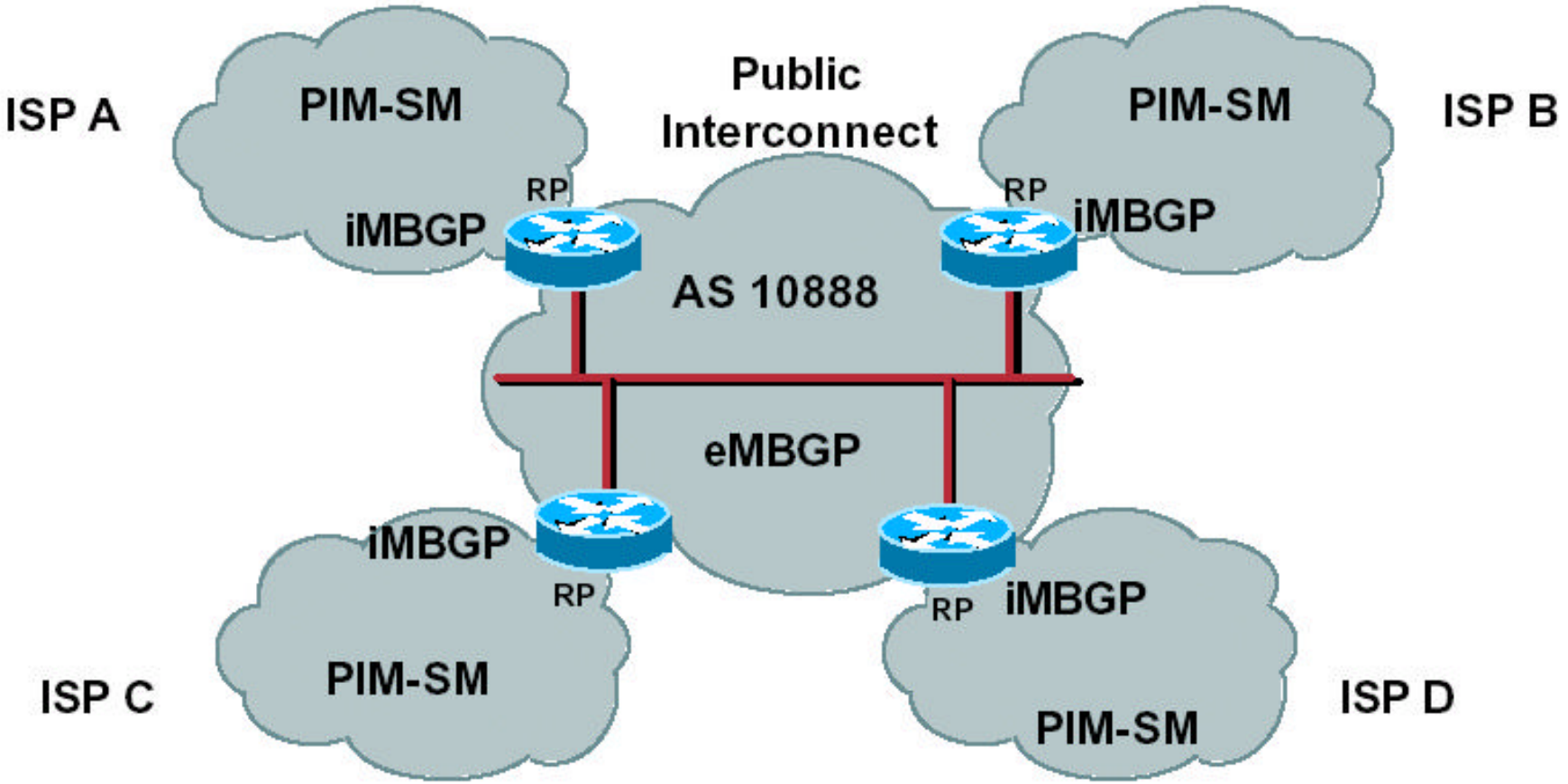
PIM-SM Przełączanie ścieżek



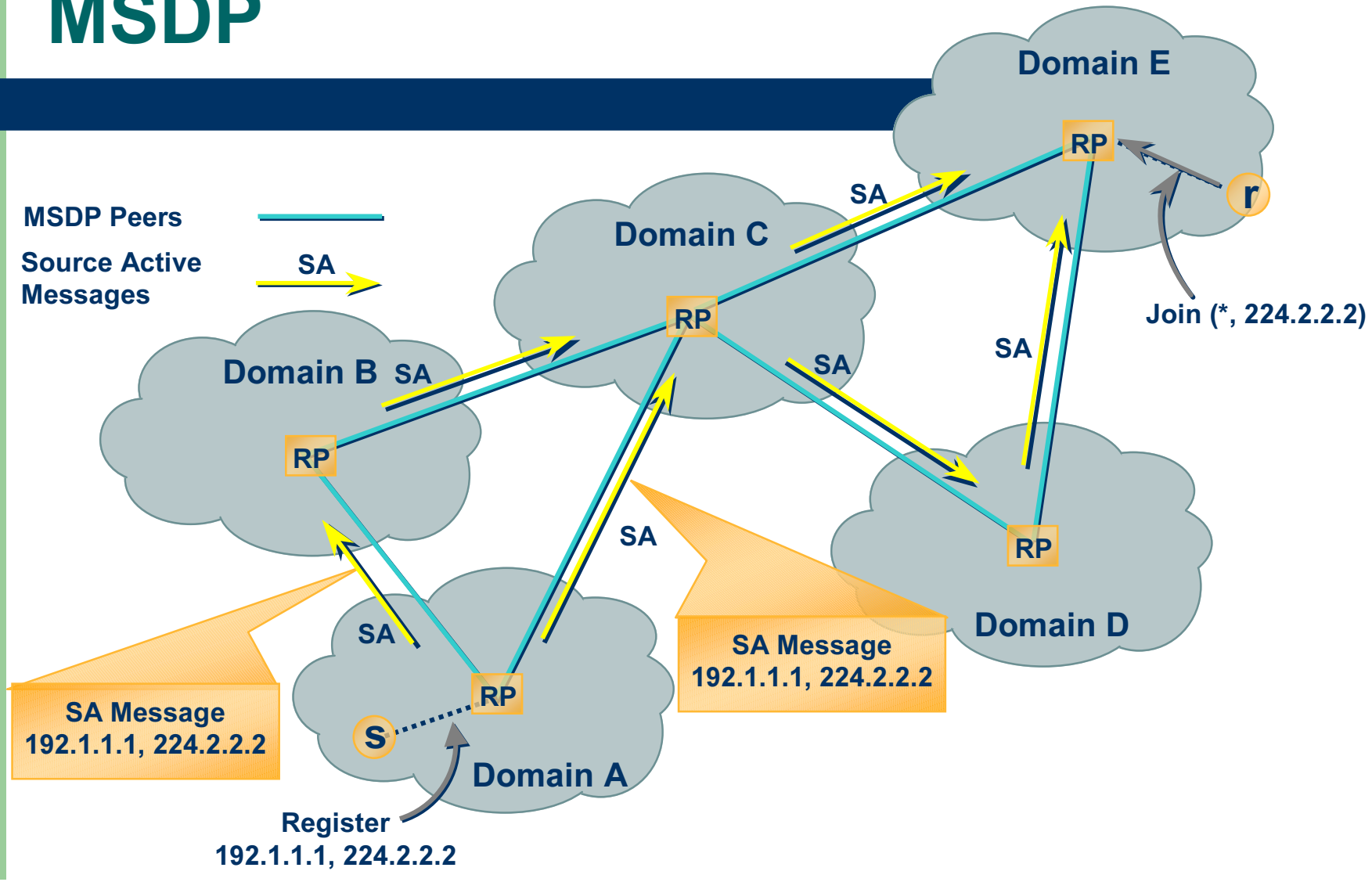
MSDP
(Multicast Source Discovery Protocol)

- Przesyłanie informacji o aktywnych źródłach między domenami PIM-SM
- Współpracuje wyłącznie z PIM-SM

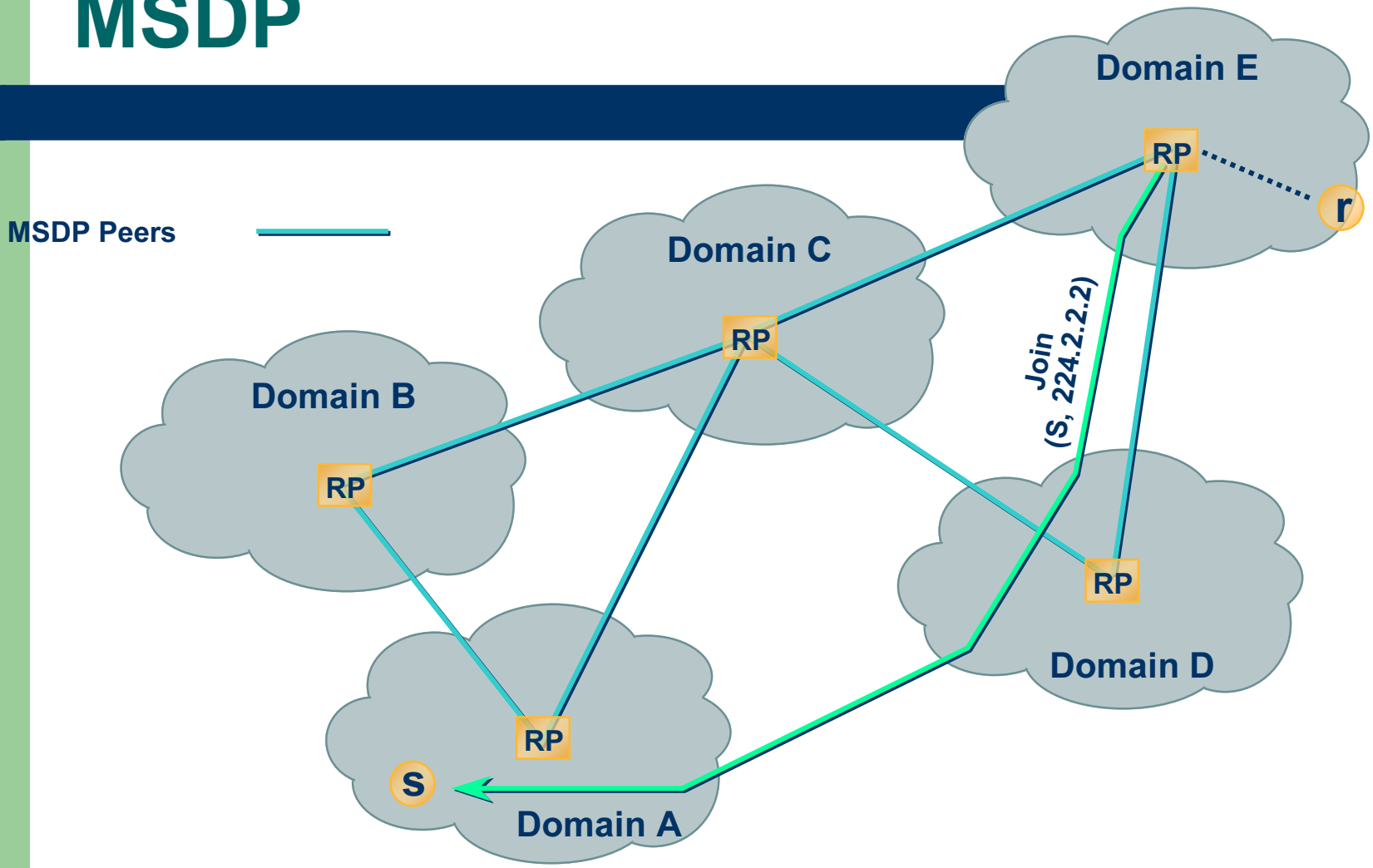
MSDP/MBGP


ISP A
PIM-SM
iMBGP
RP
Public
Interconnect
AS 10888
eMBGP
RP
PIM-SM
iMBGP
ISP B
iMBGP
RP
PIM-SM
ISP C
RP
iMBGP
PIM-SM
ISP D

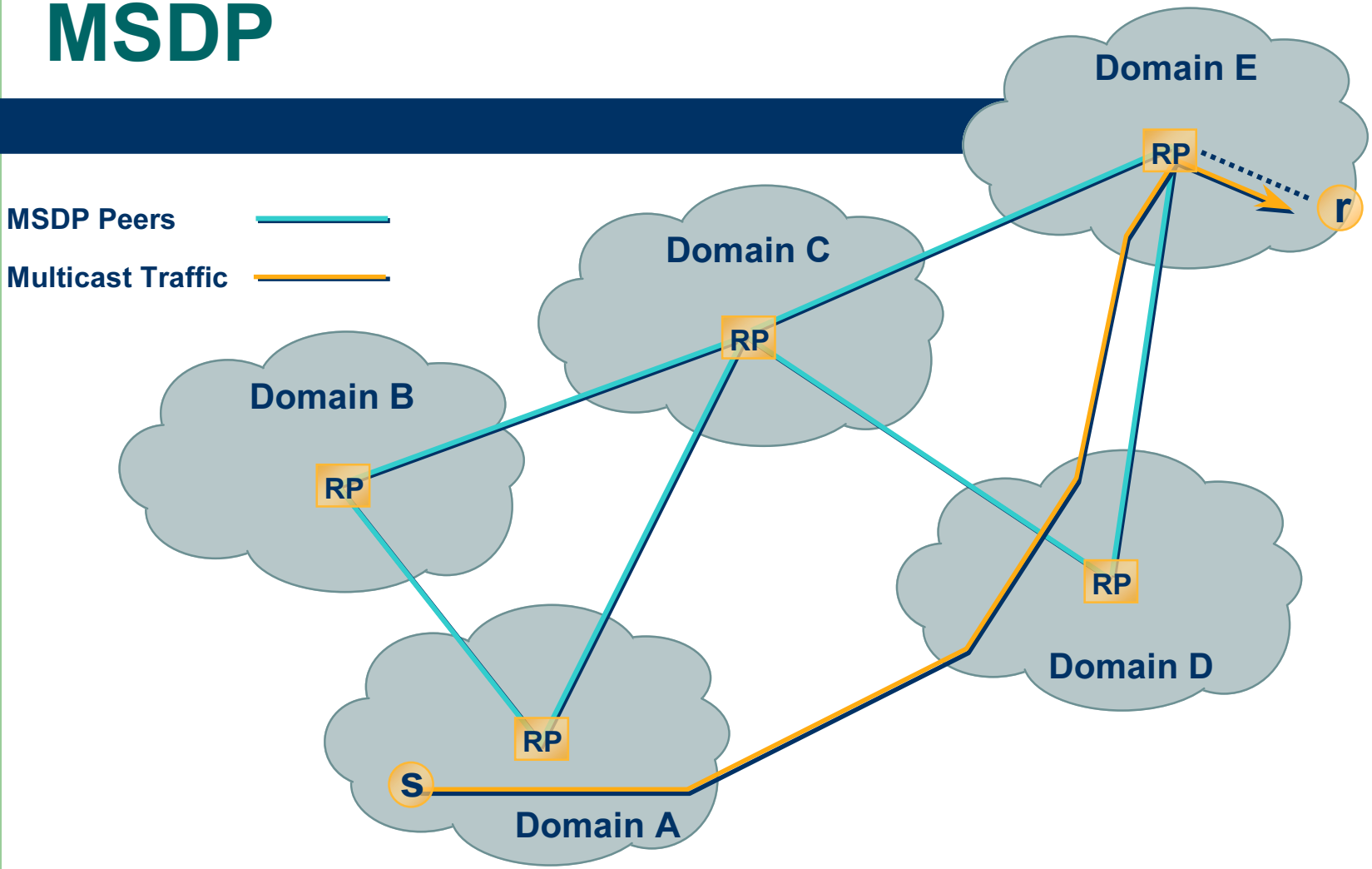
MSDP


Domain E
RP
r
Join (*, 224.2.2.2)
MSDP Peers
Source Active Messages
SA
Domain C
SA
RP
SA
Domain B
SA
SA
RP
RP
SA
SA
Domain D
SA Message
192.1.1.1, 224.2.2.2
SA Message
192.1.1.1, 224.2.2.2
RP
S
Domain A
Register
192.1.1.1, 224.2.2.2

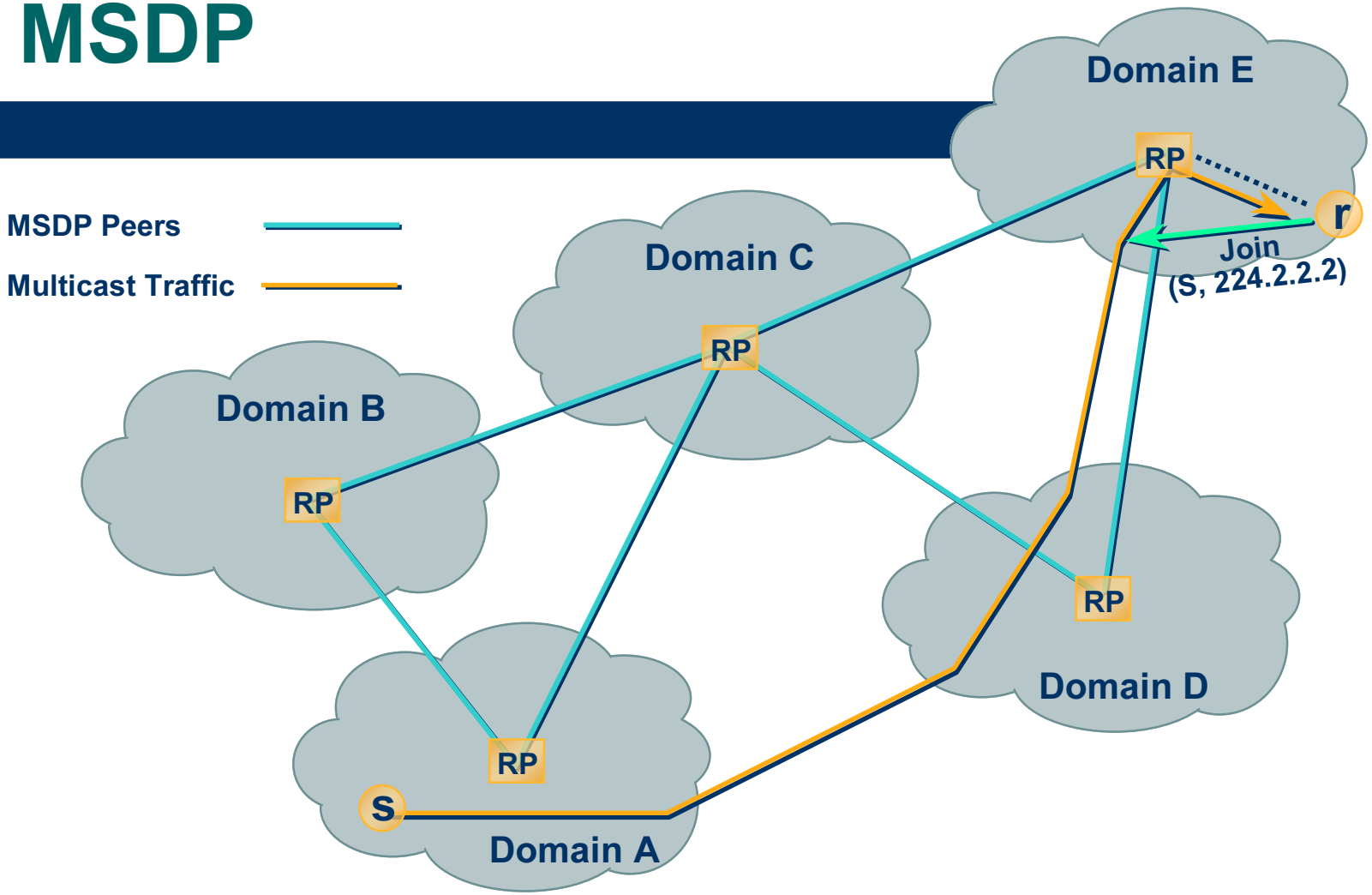
MSDP


MSDP Peers
Domain E
RP
r
Domain C
RP
Join
(S, 224.2.2.2)
Domain B
RP
RP
Domain D
RP
S
Domain A

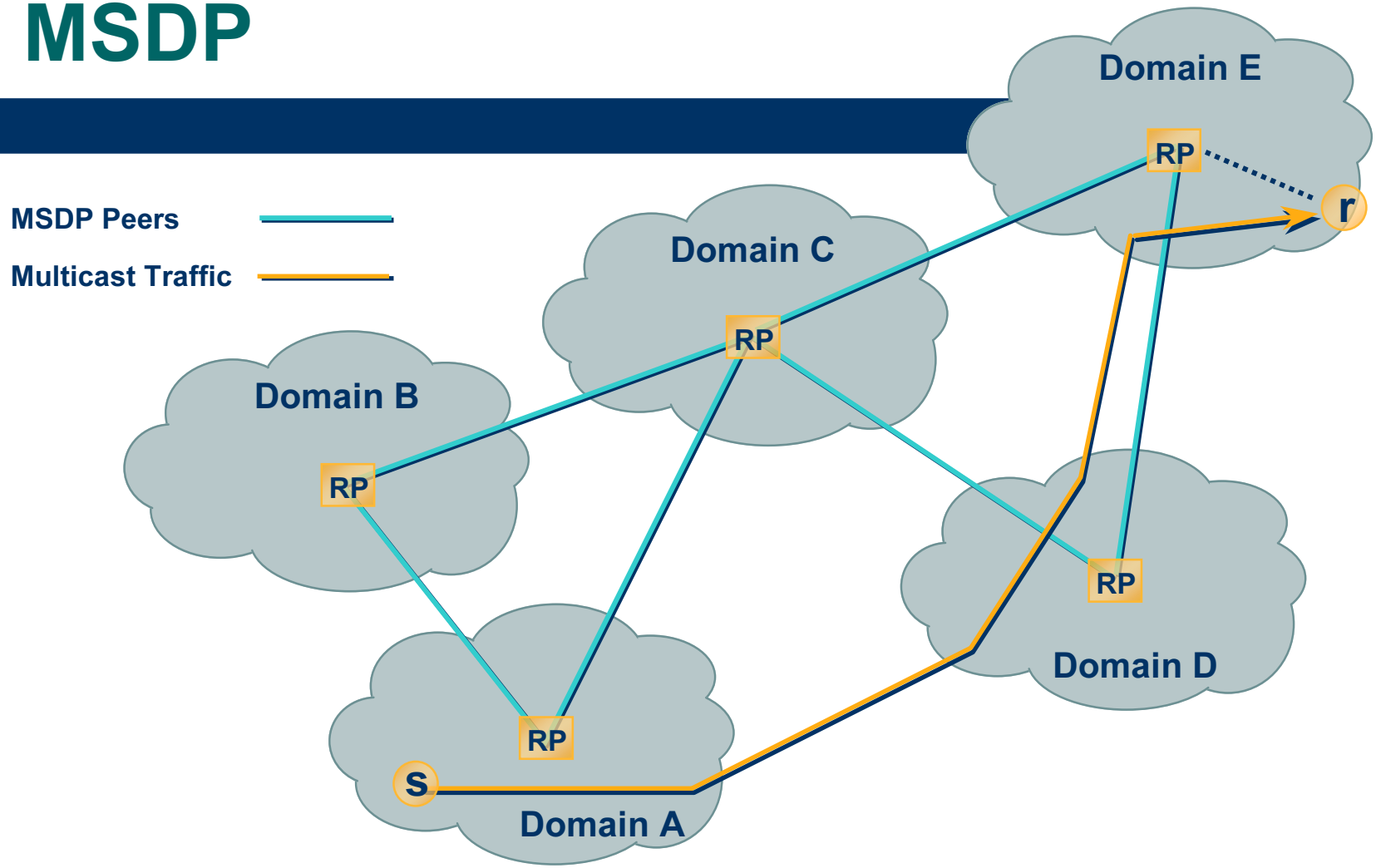
MSDP


MSDP Peers
Multicast Traffic
Domain E
RP
r
Domain C
RP
Domain B
RP
RP
Domain D
RP
S
Domain A

MSDP


MSDP Peers
Multicast Traffic
Domain E
RP
r
Join
(S, 224.2.2.2)
Domain C
RP
Domain B
RP
RP
Domain D
RP
S
Domain A

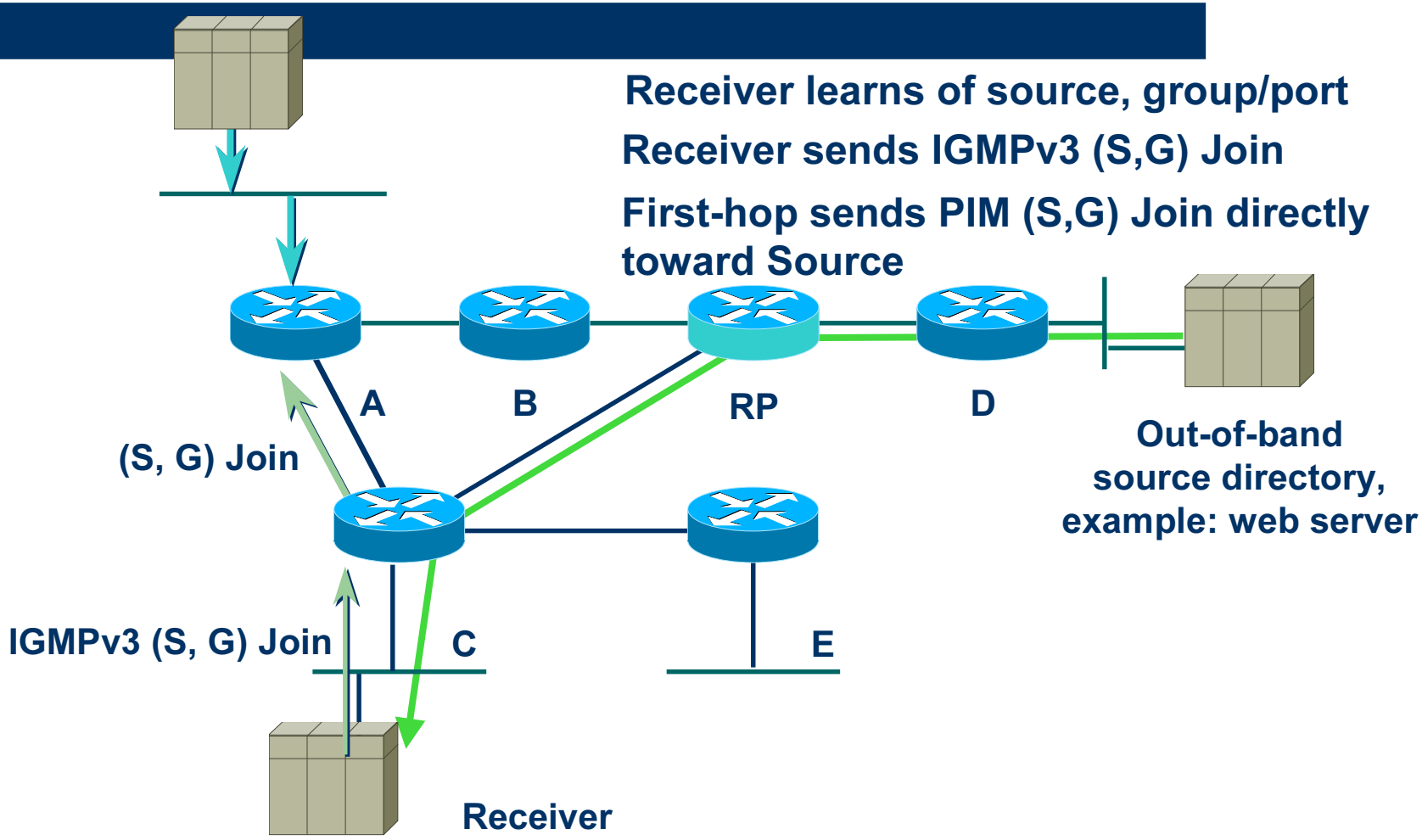
MSDP


MSDP Peers
Multicast Traffic
Domain E
RP
r
Domain C
RP
Domain B
RP
RP
Domain D
RP
S
Domain A

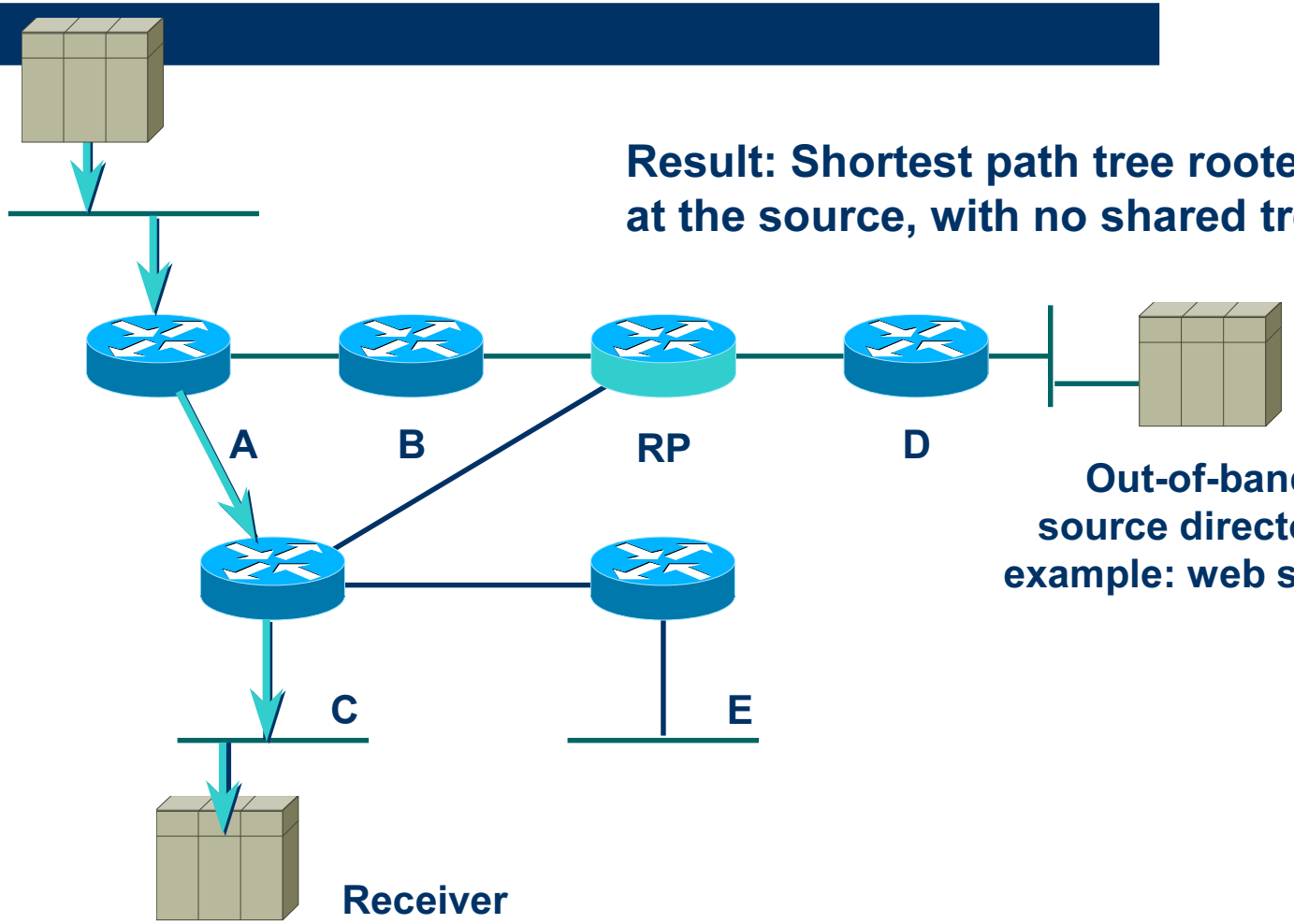
Source Specific Multicast

- Umożliwia przekazywanie pakietów po najkrótszej ścieżce bez konieczności budowania drzewa „shared tree”
- Eliminuje konieczność znajdowania aktywnych źródeł za pomocą MSDP
- Zarezerwowane adresy 232/8

SSM


Receiver learns of source, group/port
Receiver sends IGMPv3 (S,G) Join
First-hop sends PIM (S,G) Join directly toward Source
A
B
RP
D
(S, G) Join
Out-of-band source directory, example: web server
IGMPv3 (S, G) Join
C
E
Receiver

SSM



Result: Shortest path tree rooted at the source, with no shared tree.

Out-of-band source directory, example: web server