Advanced Internet and IoT Technologies

- From IP Multicast to Information Centric Networking-

Prof. Dr. Thomas Schmidt

http://inet.haw-hamburg.de | t.schmidt@haw-hamburg.de

Agenda

- Motivation
  - Content Distribution to Groups of Receivers
- IP Multicast
  - Host Group Model
  - Multicast Addressing
  - Group Membership Management
- Multicast Routing
  - Routing Algorithms
  - ASM Routing Protocols
  - SSM Routing
- Information Centric Networking
  - Motivation
  - ICN Approaches
  - Routing & Forwarding
  - ICN in the IoT

Motivation

The Internet today has a major purpose in content delivery

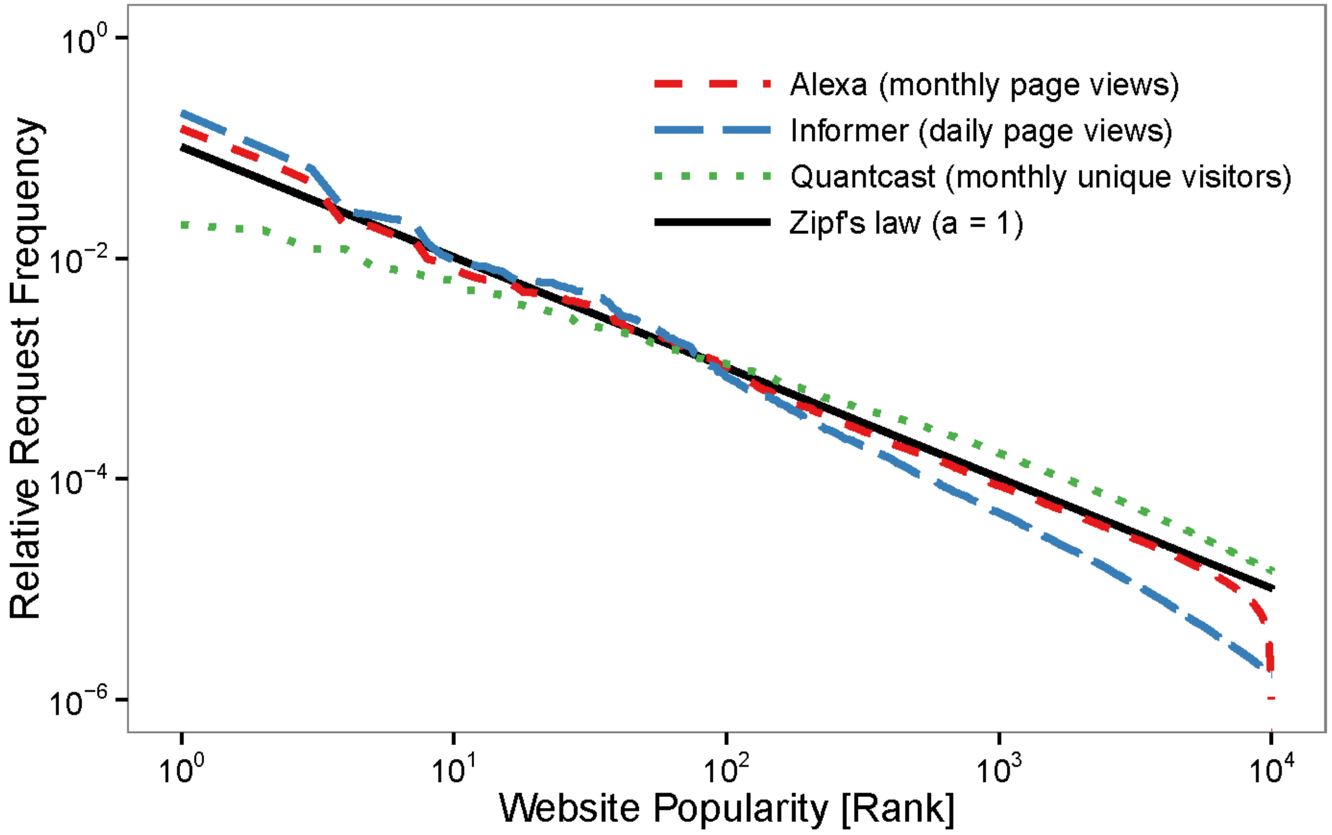
Content popularity is sharply peaked, i.e.,

- Many consumers request the same content
- Few publishers dominate Internet traffic
- Content dissemination assisted by CDNs

Many Internet applications are inherently for groups

- Chats, calls
- Games, infotainment
- Social networks

Content Popularity Distribution



Why to Talk in Groups?

Many use cases on the Internet:

- Multimedia Content Distribution
- Broadcasting Offers (IPTV)
- Time-sensitive Data (Stock Prices)
- Collaboration, Gaming
- Rendezvous and Coordination Services

⇒ Scalable Communication Paths needed to Distribute Data in Parallel

IP Multicasting

Method for Transferring IP Datagrams to Host-Groups

Initially: RFC 1112 (S. Deering & D. Cheriton, 1989)

Addresses a host group by *one* group address

Two kinds of multicast:

- Any Source Multicast (ASM)
- Source Specific Multicast (SSM)

Client Protocol for registration (IGMP/MLD)

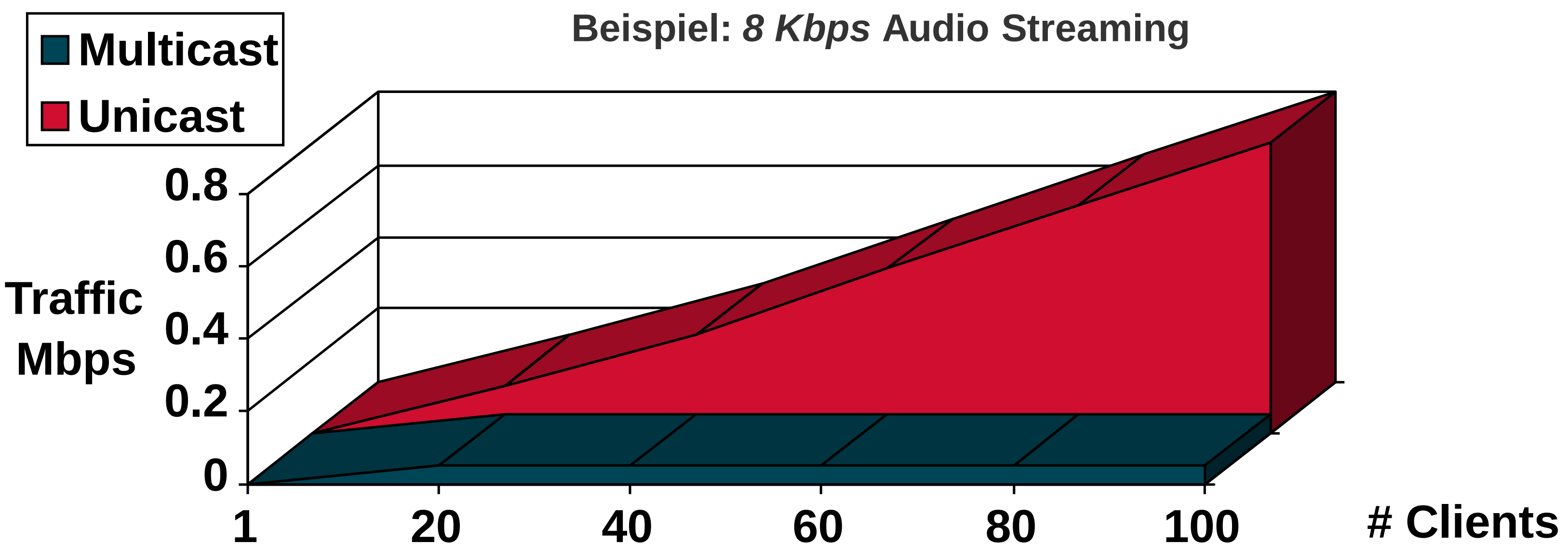
Routing throughout the Internet (Multicast Routing)

Address translation to Layer 2

Properties of IP Multicasting

Prevents redundant network traffic

Reduces network and server load



Multicast Addressing

Denote delocalized group identifiers

IPv4 Multicast Group addresses

- 224.0.0.0–239.255.255.255
- Class “D” Address Space
- Special SSM block: 232.*.*.*

IPv6: scoped multicast addresses

- FF00::/8
- Special SSM block: FF3x::/32

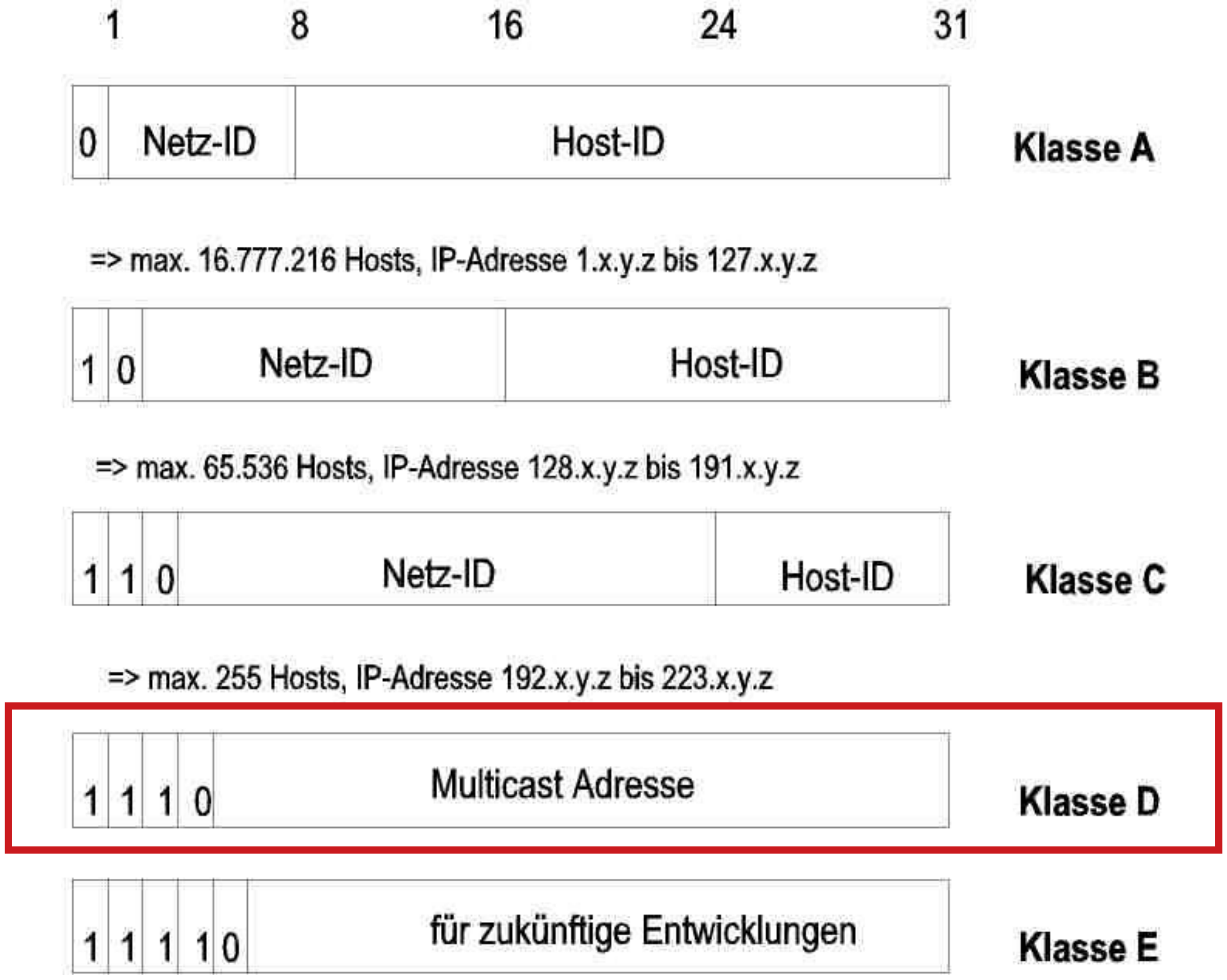
Permanent Addresses assigned by IANA

- RFC 1700: Assigned Addresses
- *“http://www.iana.org/assignments/multicast-addresses ”* lists reserved addresses

Dynamic Addresses

- independent of local IP-address space (IPv4)
- Unicast based Multicast addresses (IPv6)

IPv4 Address Class



Reserved Multicast Addresses

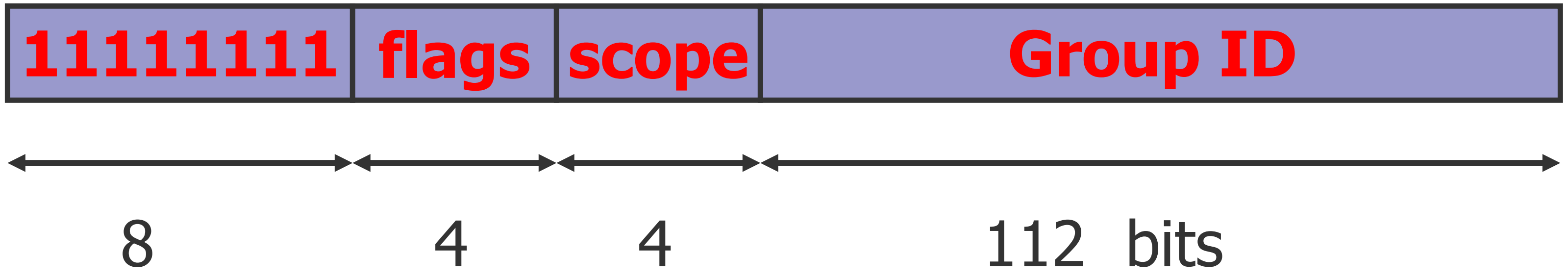
Permanent IP Multicast Group Addresses

- 224.0.0.0–224.0.0.255
- Examples:
  - 224.0.0.1 All Systems of Subnet
  - 224.0.0.2 All Routers of Subnet
  - 224.0.0.4 All DVMRP Router
  - 224.0.0.5 All OSPF Router
  - 224.0.0.9 All RIP(v2) Router
  - 224.0.0.13 All PIMv2 Router
  - 224.0.1.1 NTP
  - 224.0.1.9 Multicast Transport Protocol (MTP)

TTL – Standards

- TTL = 1: This Subnet
- TTL = 15: This Site
- TTL = 63: This Region
- TTL = 127: This Internet

IPv6 Multicast Addresses

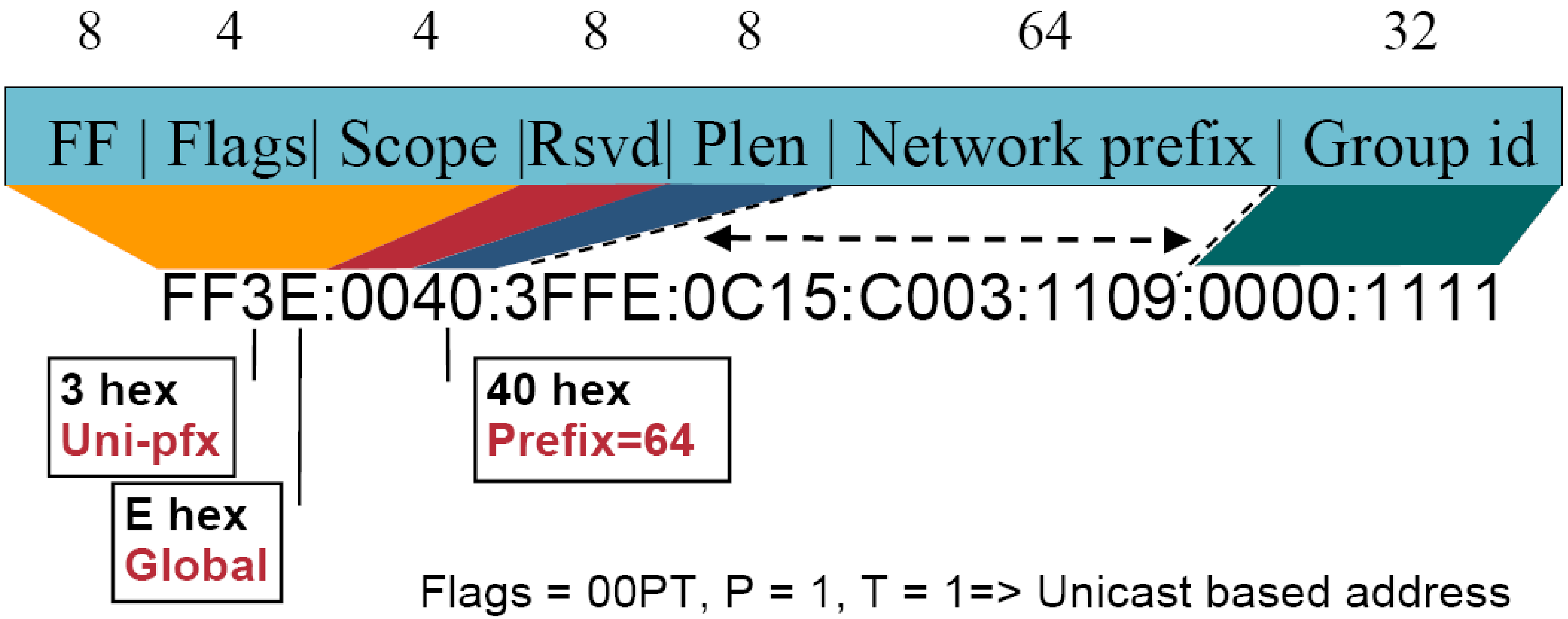


Flag field: lower bit indicates permanent (=0) respectively transient (=1) group, rest is reserved (==0)

Scope field:
- 1 - node local
- 2 - link-local
- 5 - site-local
- 8 - organisation local
- B - community-local (deprecated)
- E - global (other values reserved)

IPv6 Unicast Based Multicast Addresses (RFC 3306)

- **Solves the old IPv4 address assignment problem:**
  How can I get global IPv4 multicast addresses (GLOB, ..)
- **In IPv6, if you own an IPv6 unicast address prefix you implicitly own an RFC3306 IPv6 multicast address prefix:**



Internet Group Management

Client Protocol to initiate, preserve and terminate group membership

Local Router collect and monitor information

IPv4: Internet Group Management Protocol (IGMP)

- IGMP v1 RFC 1112
- IGMP v2 RFC 2236 – implemented almost everywhere
- IGMP v3 RFC 3376 – implemented in most OSes

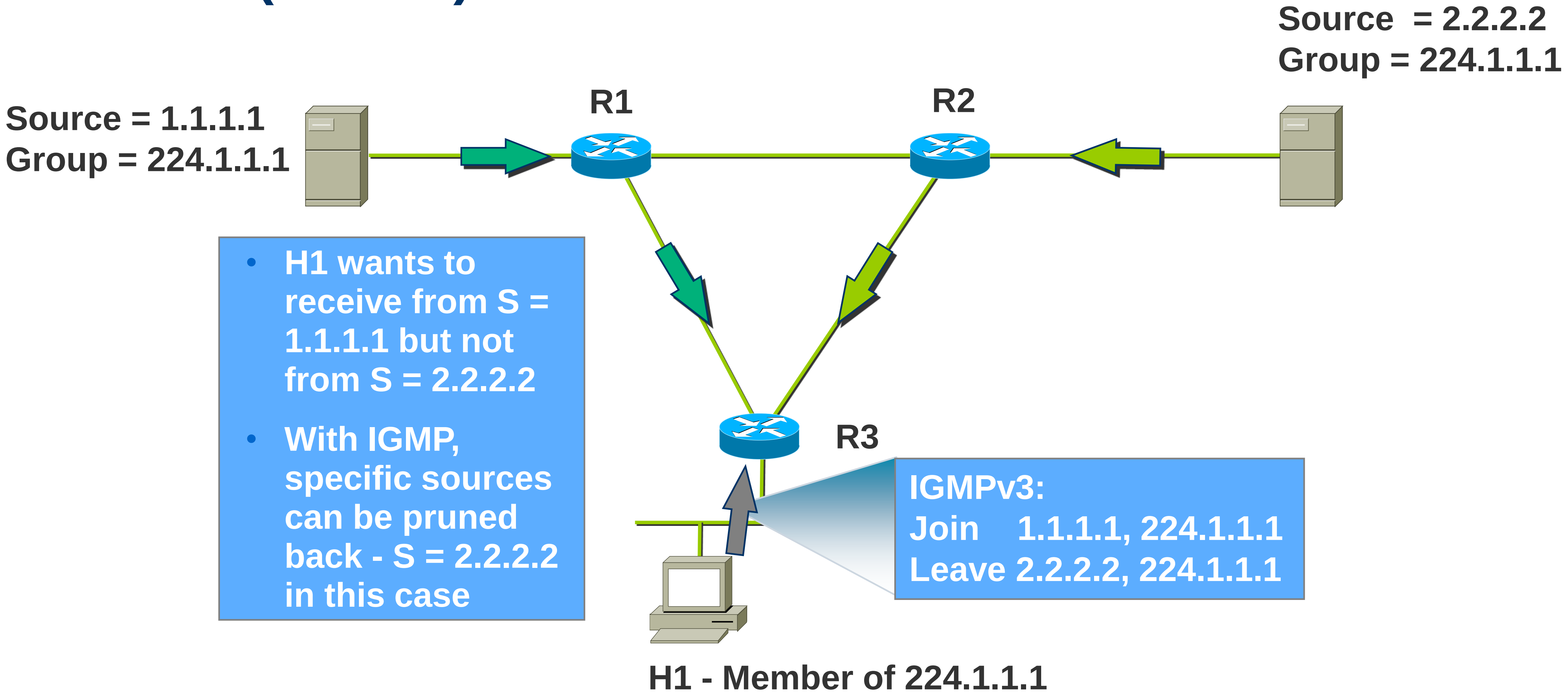
IPv6: Multicast Listener Discovery Protocol (MLD)

- MLDv1 (RFC 2710) – analogue to IGMPv2
- MLDv2 (RFC 3810) – starting from IGMPv3

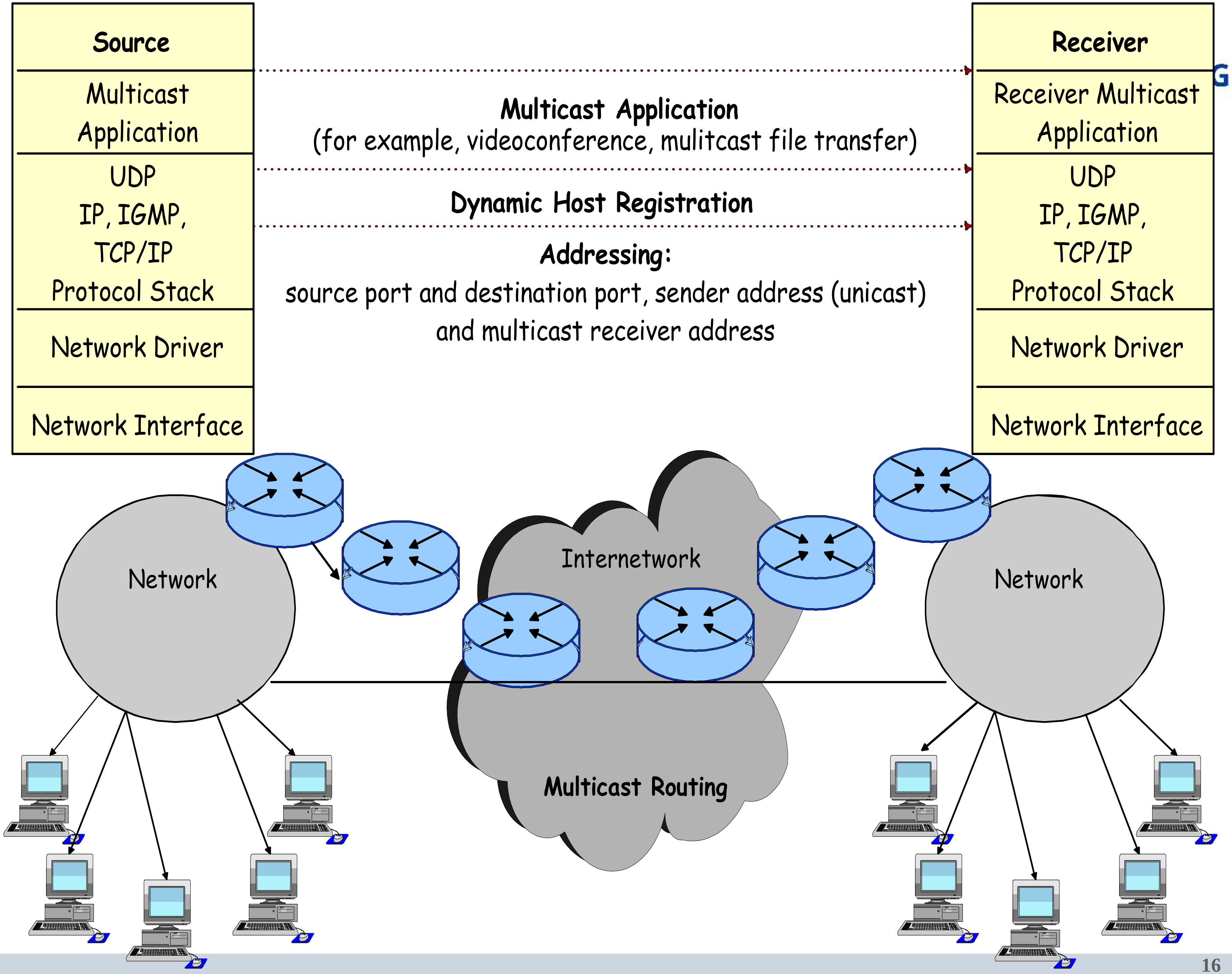
SSM Specialities: RFC 4604

IGMPv3 (MLDv2)

- H1 wants to receive from S = 1.1.1.1 but not from S = 2.2.2.2
- With IGMP, specific sources can be pruned back - S = 2.2.2.2 in this case



The Internet Multicast System



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

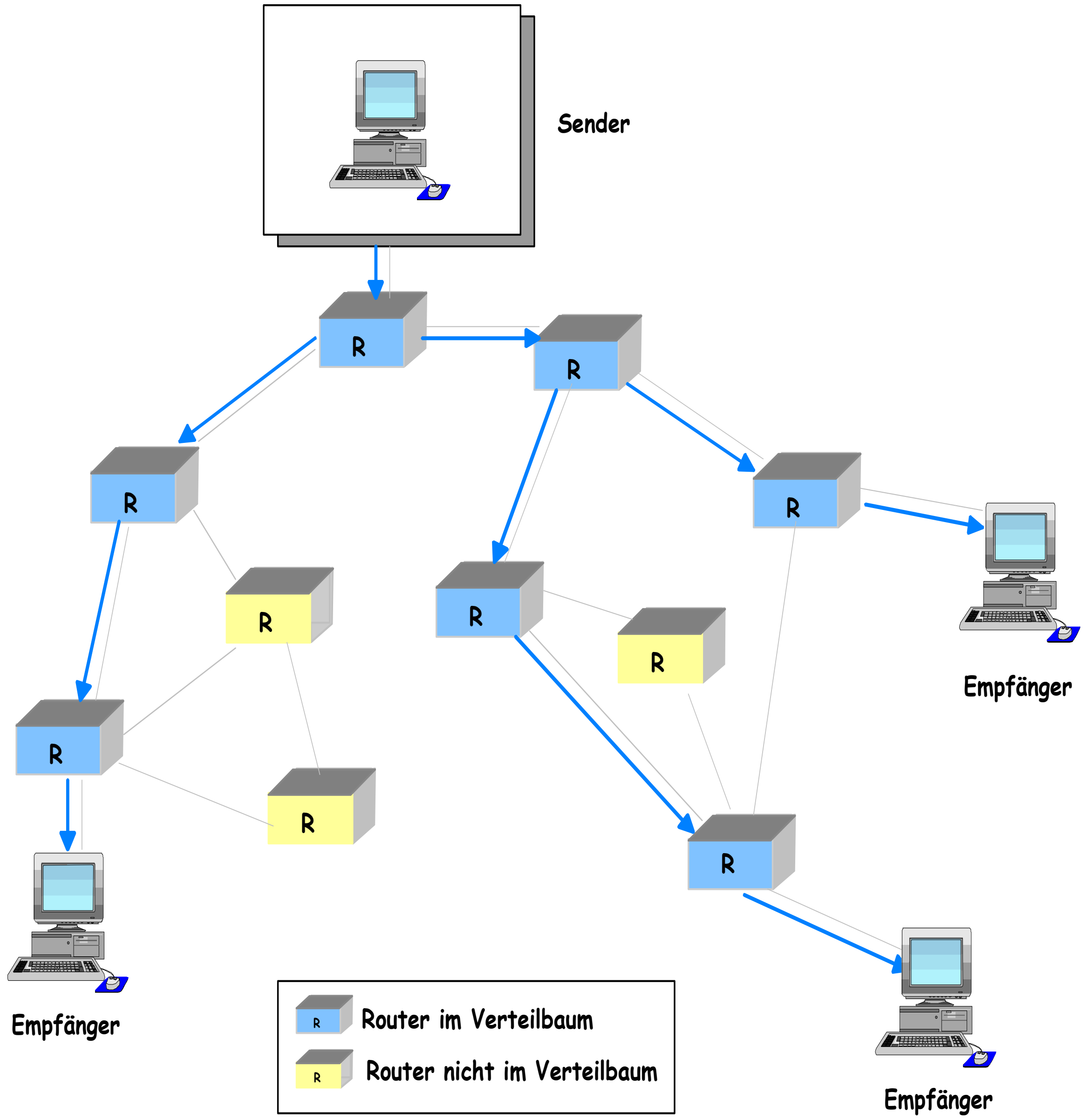
Multicast Routing (2)

IP Multicast is a publish-subscribe approach

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

Multicast 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.

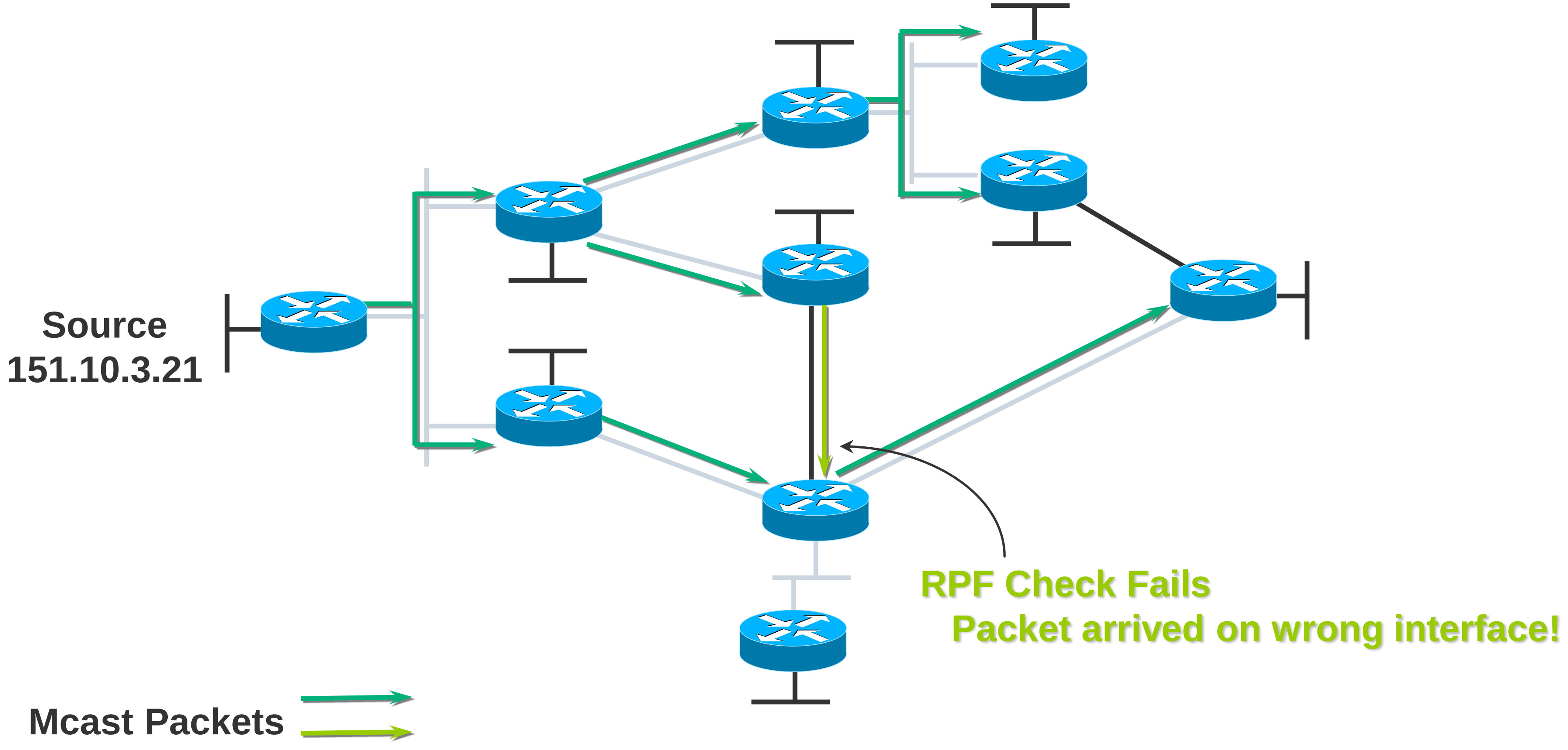
RPF Check:

Active routing table searched for *source*-address

Packet transmitted, if received on the interface foreseen as destination for source address

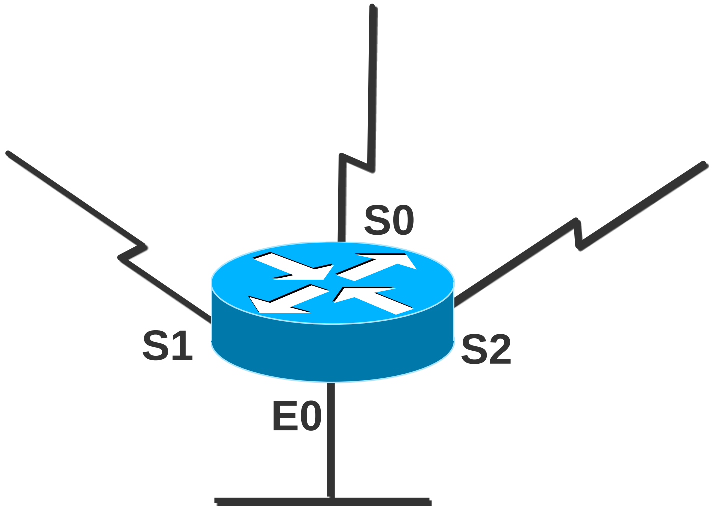
Packet discarded otherwise

RPF Check



Source
151.10.3.21
RPF Check Fails
Packet arrived on wrong interface!
Mcast Packets

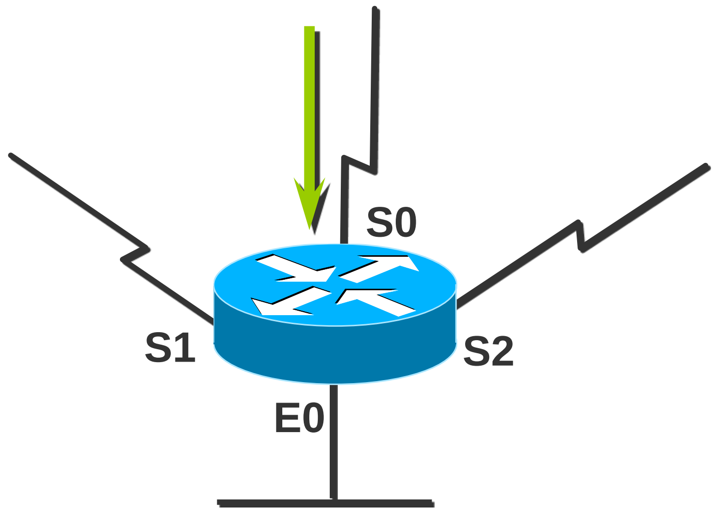
RPF Check: Failure



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

RPF Check: Failure

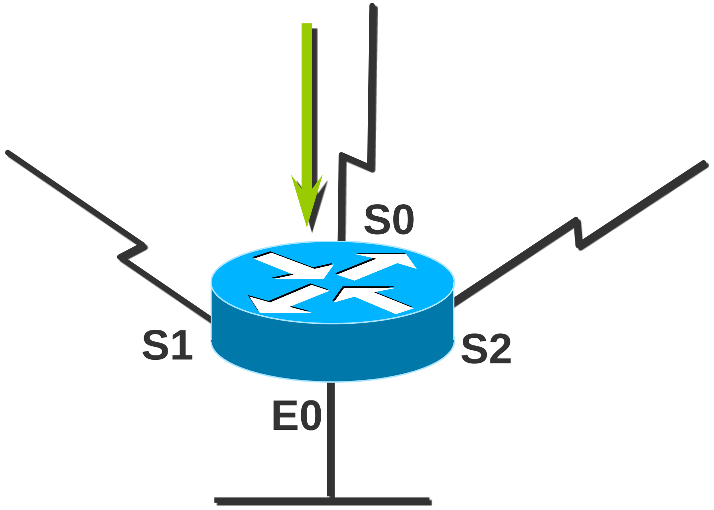
Multicast Packet from
Source 151.10.3.21



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

RPF Check: Failure

Multicast Packet from Source 151.10.3.21



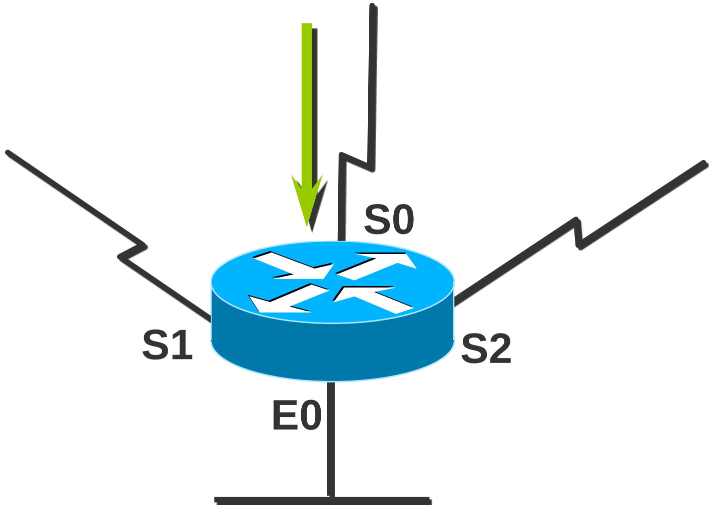
Unicast Route Table

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!

RPF Check: Failure

Multicast Packet from
Source 151.10.3.21



RPF Check Fails!

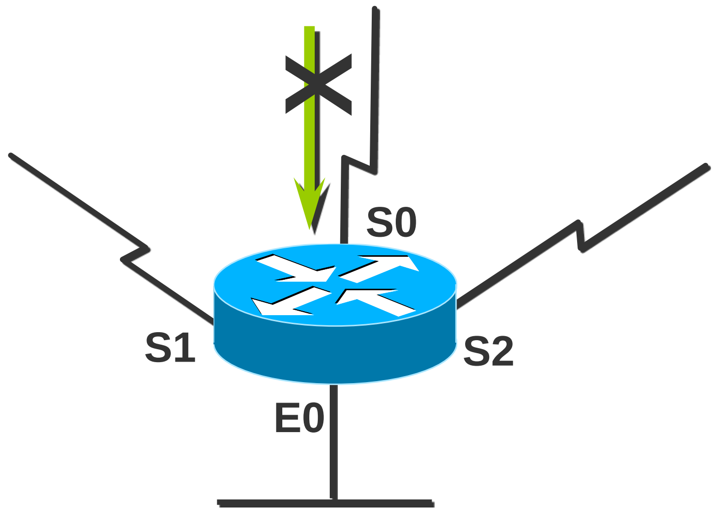
Unicast Route Table

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!

RPF Check: Failure

Multicast Packet from
Source 151.10.3.21



RPF Check Fails!

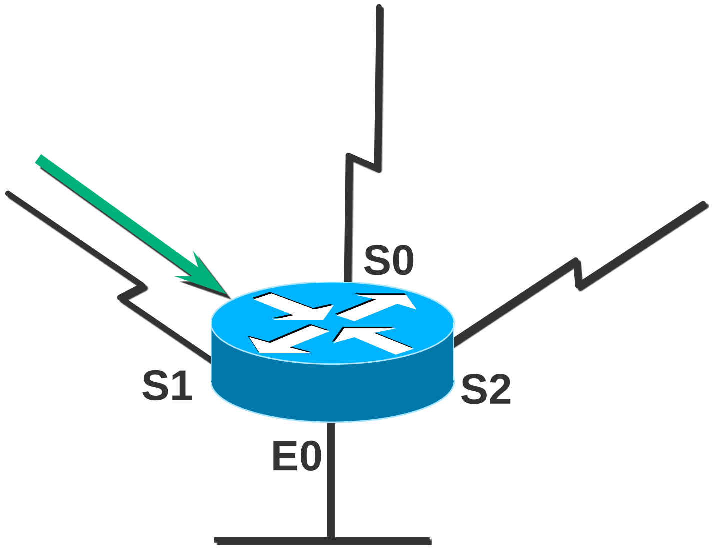
Unicast Route Table	
Network	**Interface**
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Packet Arrived on Wrong Interface!

Discard Packet!

RPF Check: Success

Multicast Packet from
Source 151.10.3.21

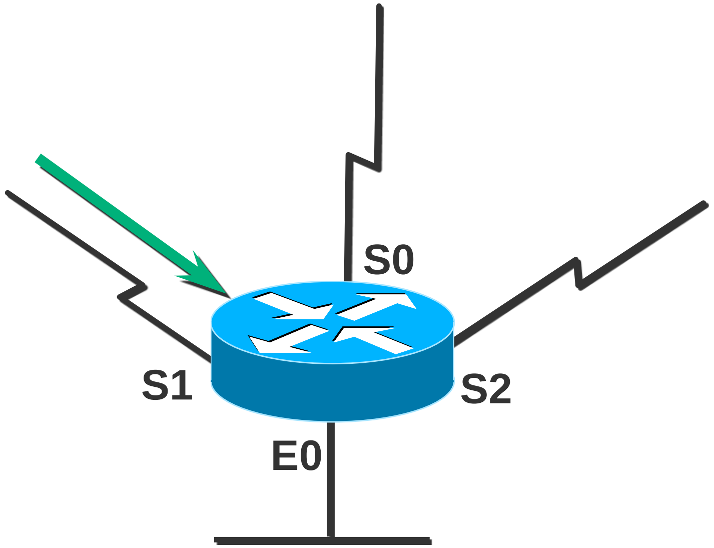


Unicast Route Table

Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

RPF Check: Success

Multicast Packet from
Source 151.10.3.21



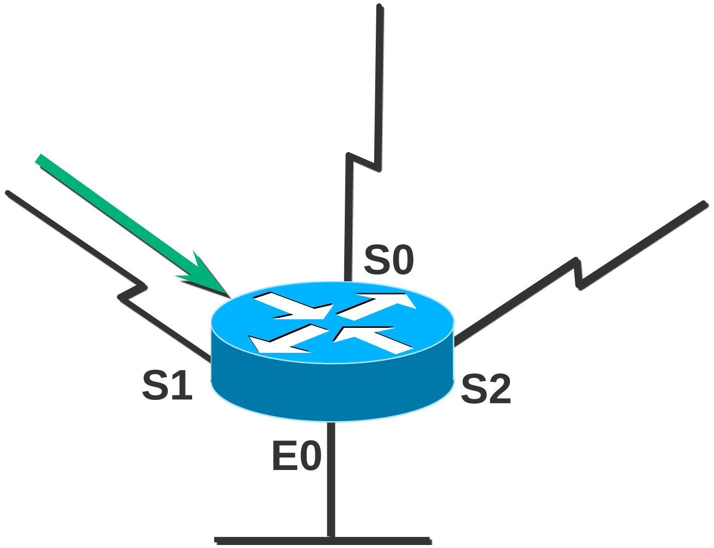
Unicast Route Table

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!

RPF Check: Success

Multicast Packet from
Source 151.10.3.21



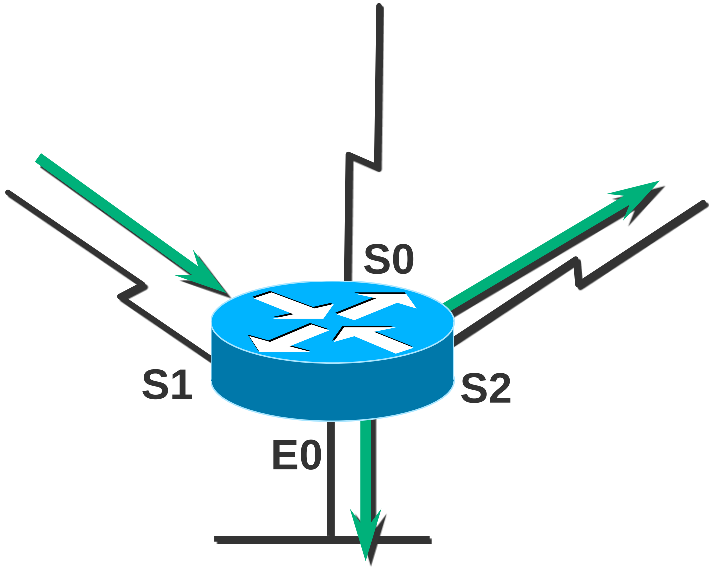
RPF Check Succeeds!

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!

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 trees to reach all receivers?

Link-state (MOSPF)

- Augment links with forwarding state
- Flood link state

Dense Mode (DVMRP, PIM-DM)

- Push traffic
- Flooding and pruning

Sparse Mode (PIM-SM, BIDIR-PIM)

- Pull traffic
- Directional traffic only
- Rendezvous Points

Protocol Independent Multicast Sparse Mode (PIM-SM)

Protocol independence:

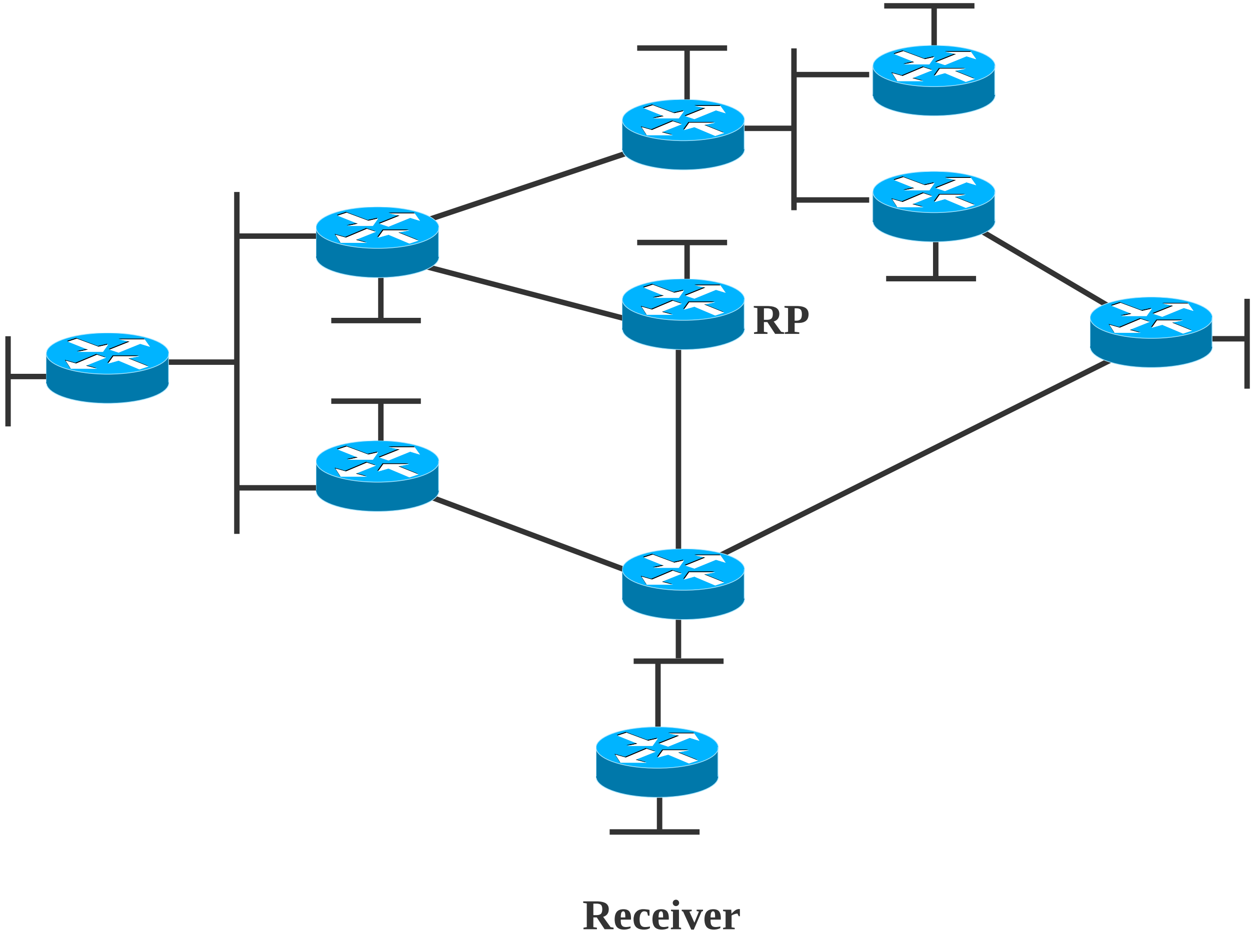
works with all underlying Unicast Routing Protocols

Long history of standards (RFCs 2326 … 4601 … 7761)

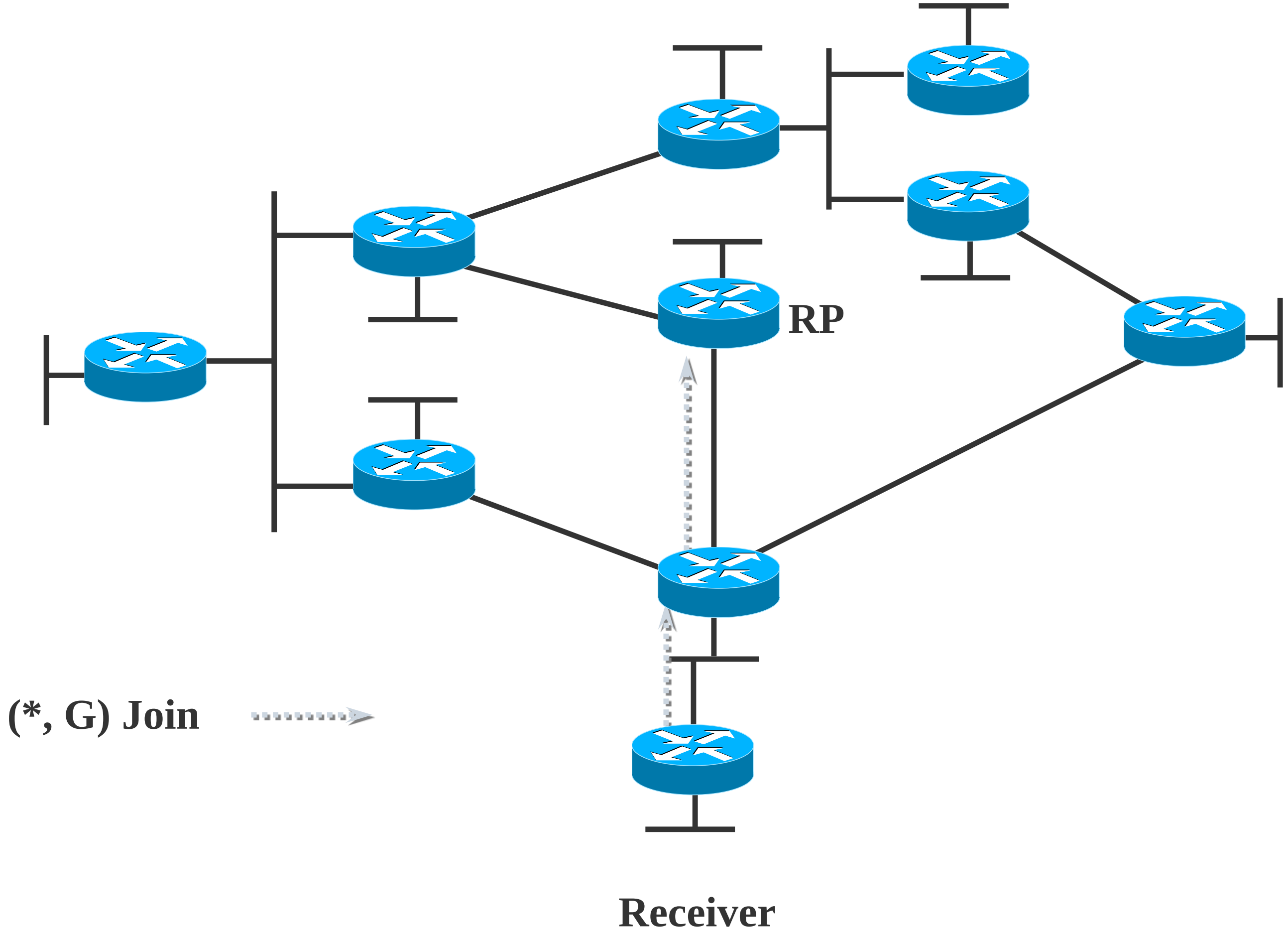
Sparse Mode PIM uses Rendezvous Points (RP)

- Constructs a shared distribution tree centred at RP
- Efficient for widely distributed groups
- Favoured for wide area networks
  - problem: inter-RP signalling
- Widely implemented

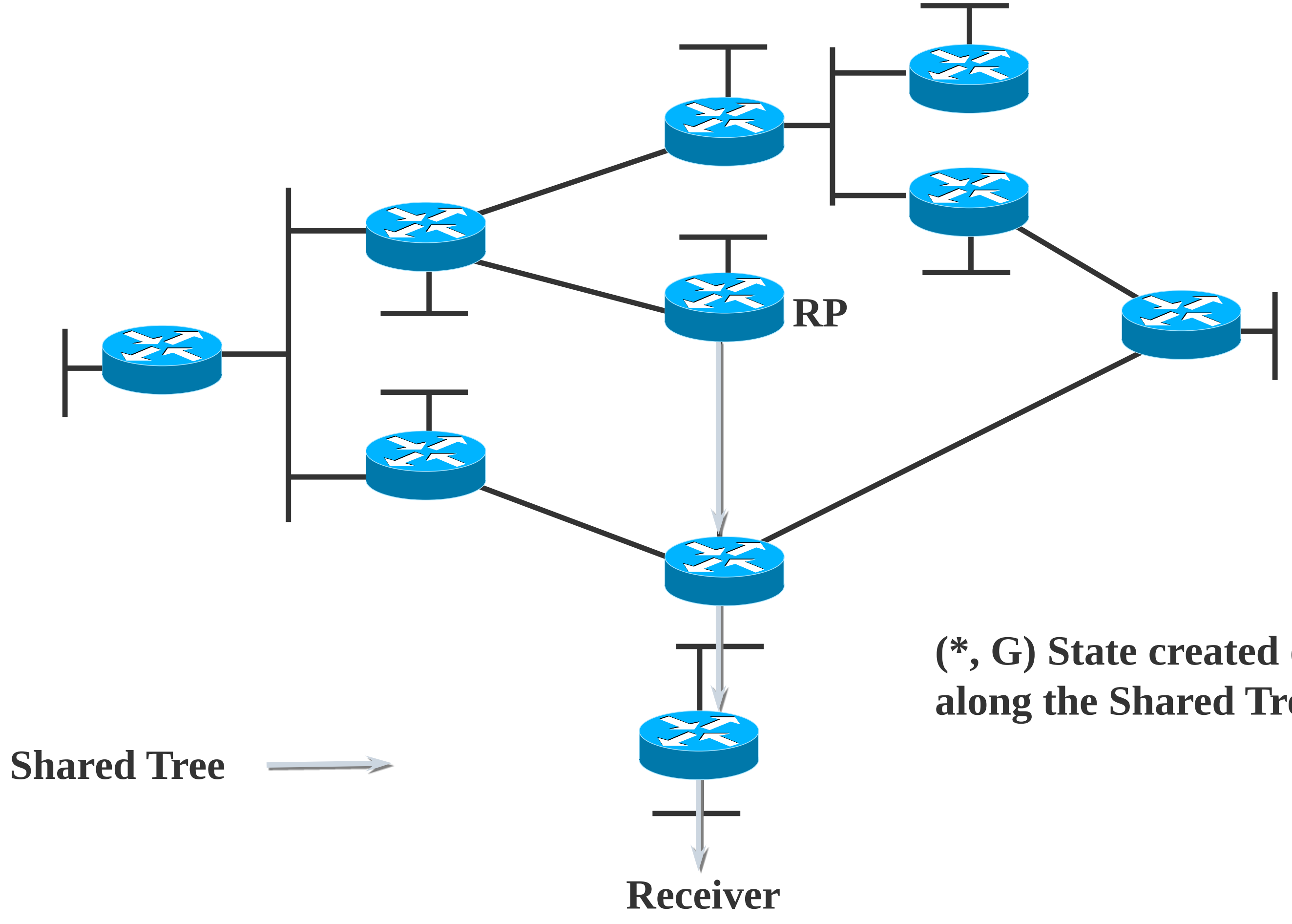
PIM SM Tree Joins



PIM SM Tree Joins

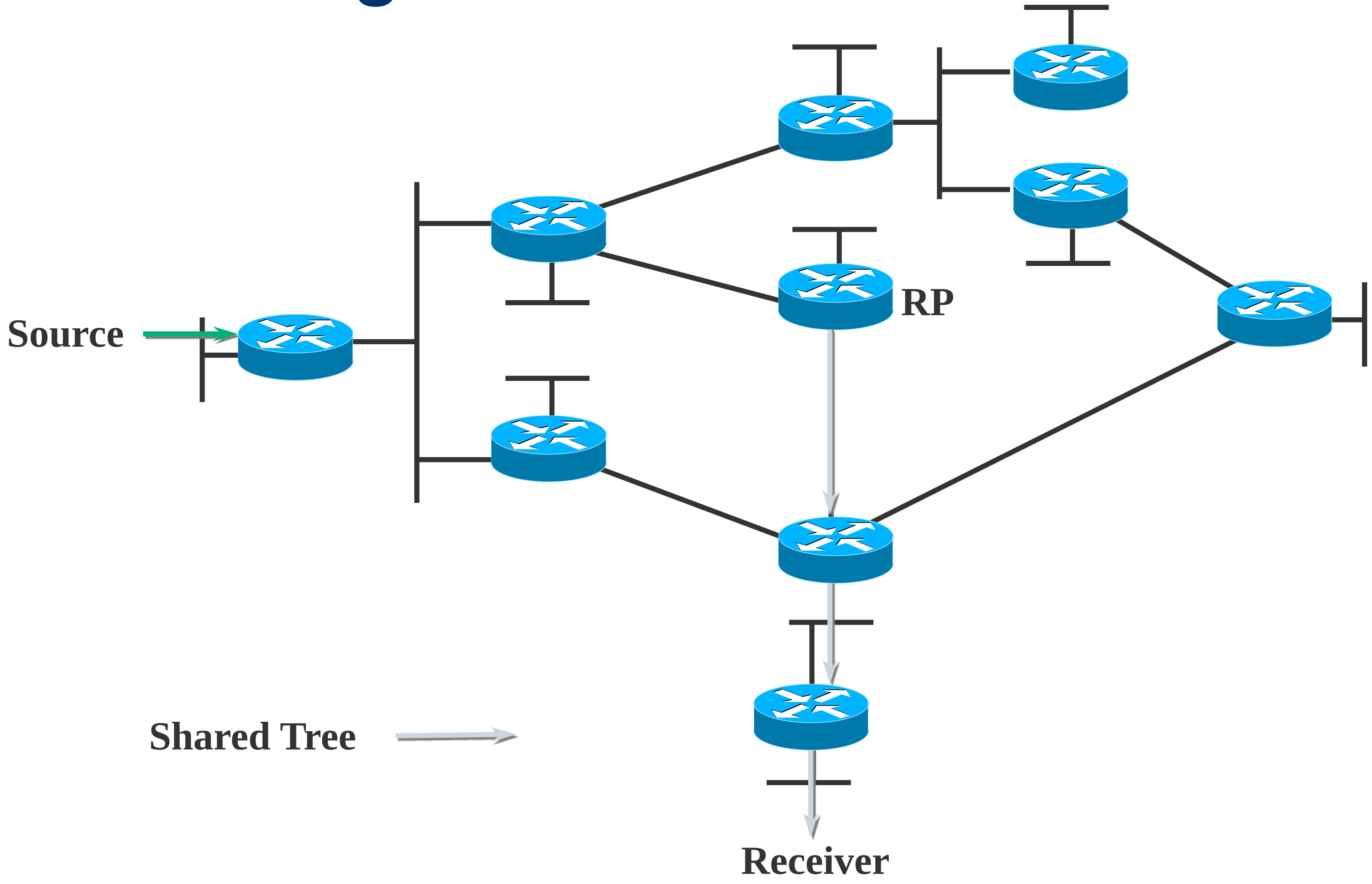


PIM SM Tree Joins

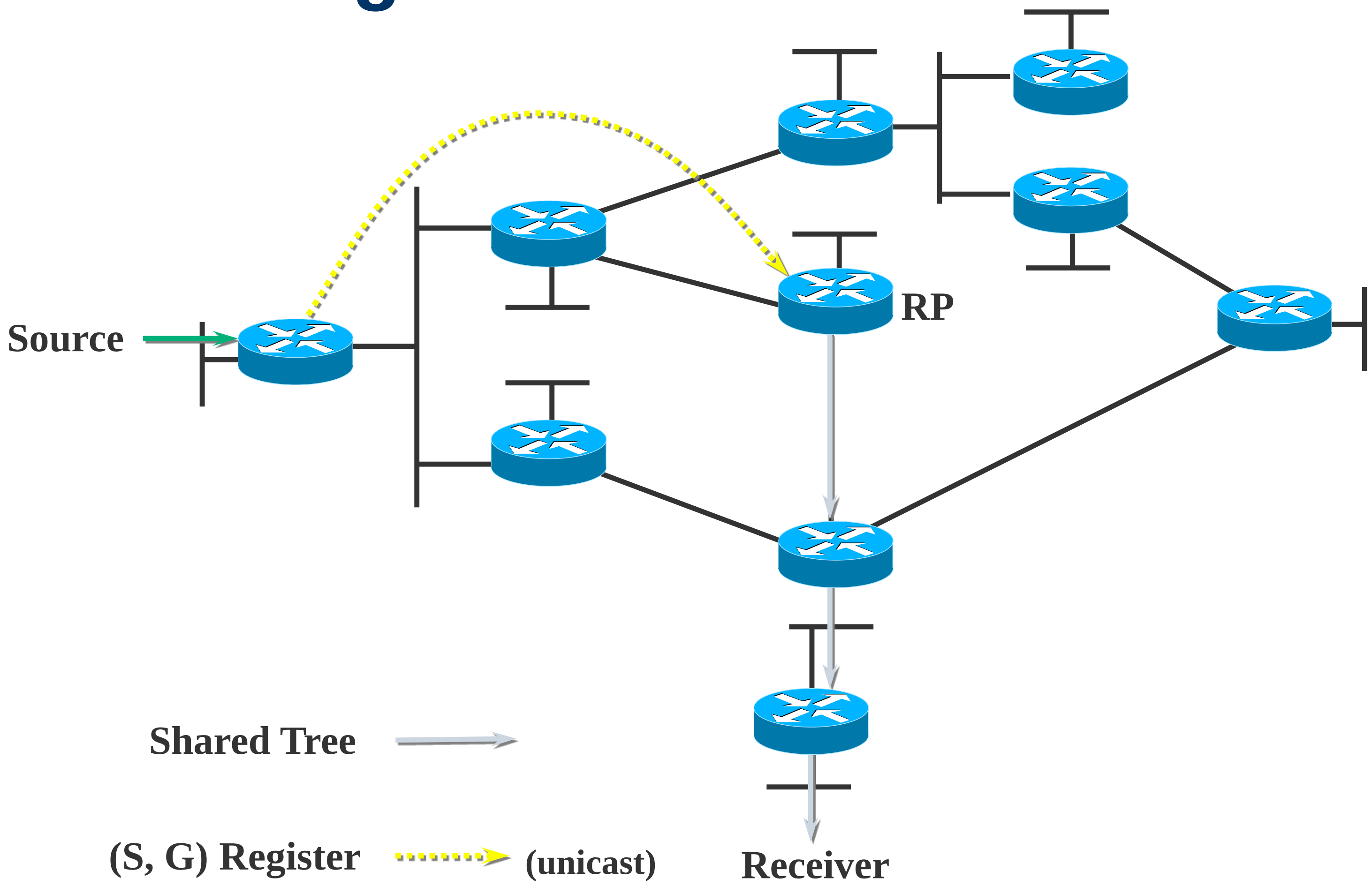


(*, G) State created only along the Shared Tree.

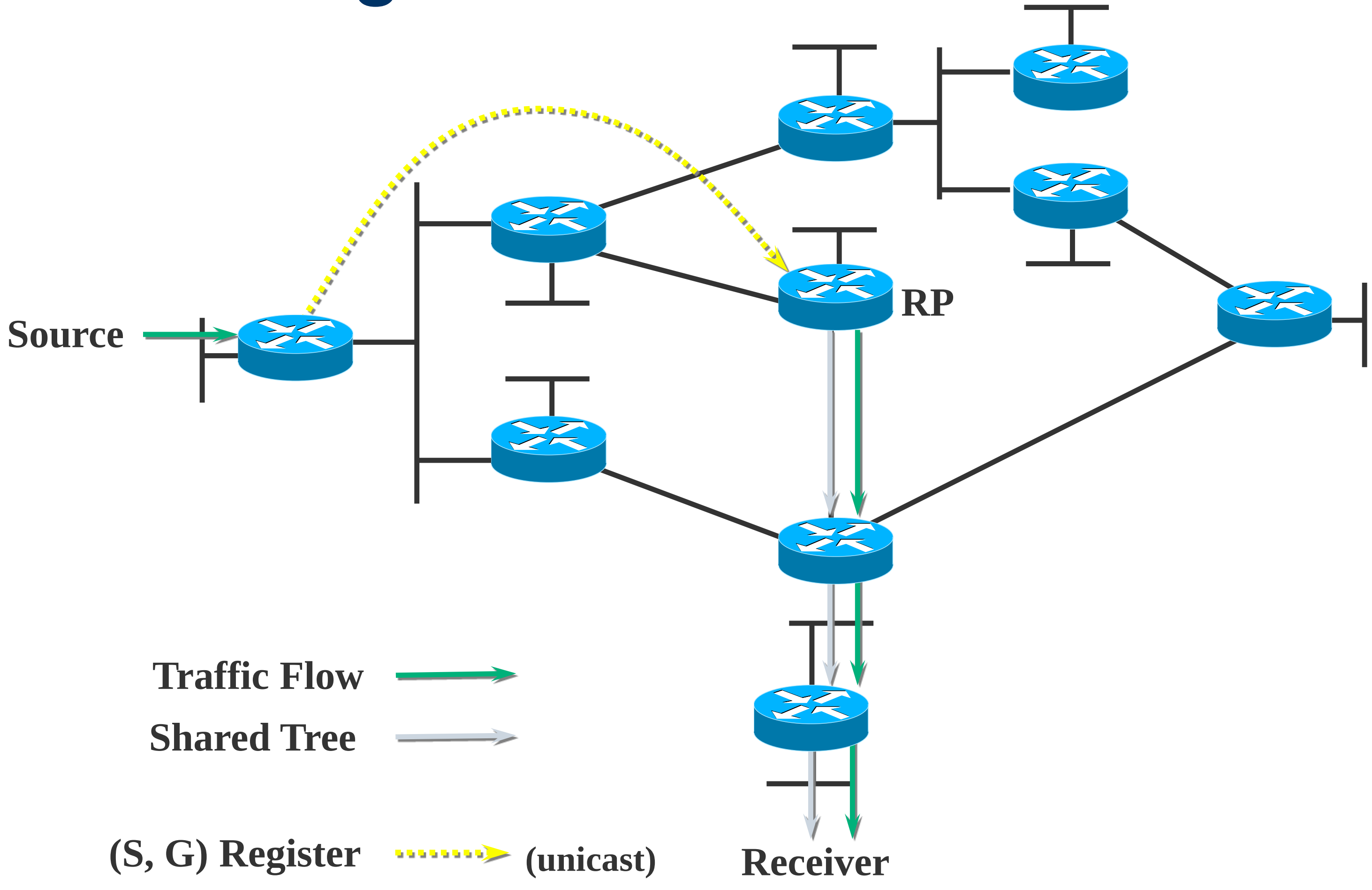
PIM SM Sender Registration



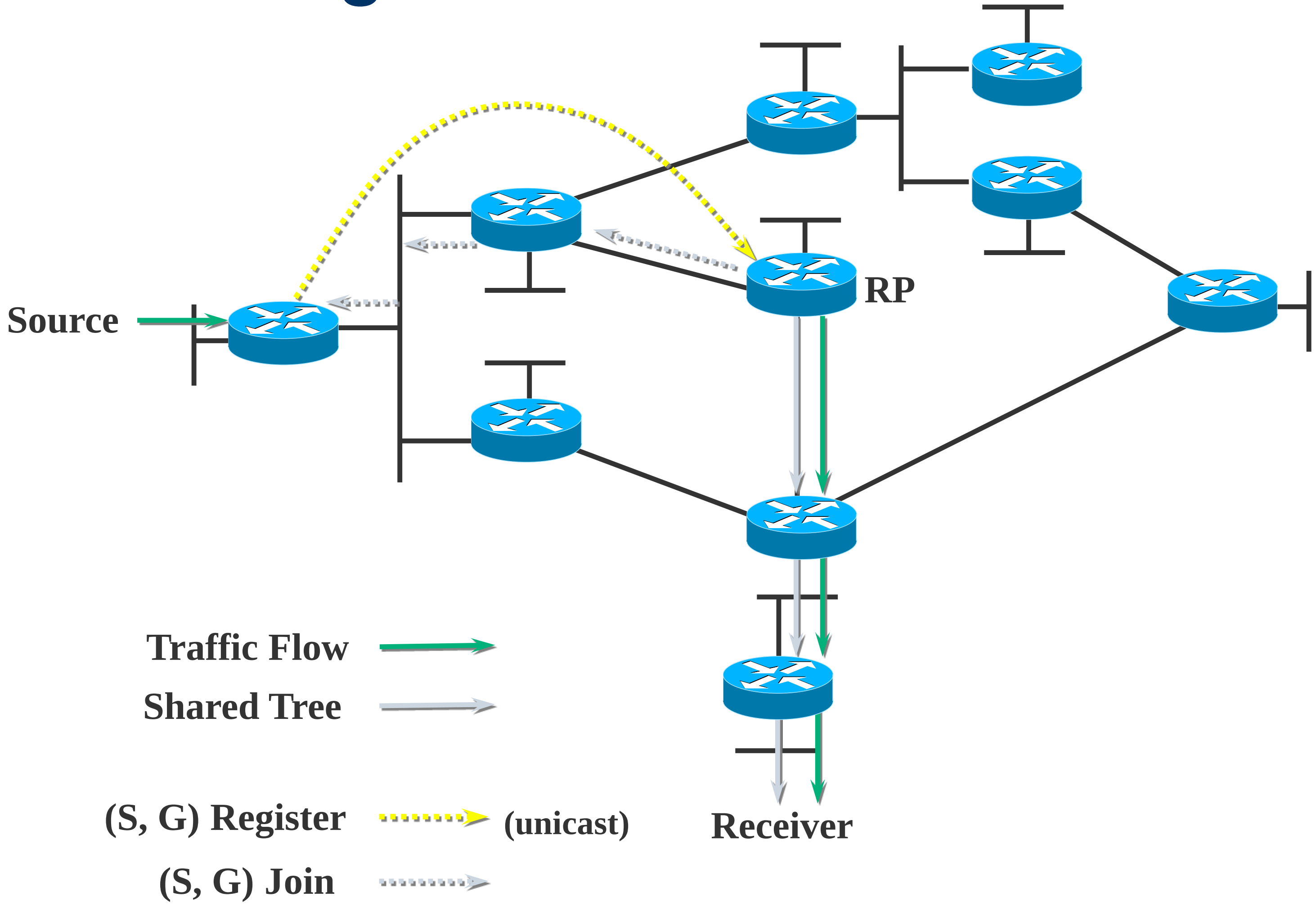
PIM SM Sender Registration



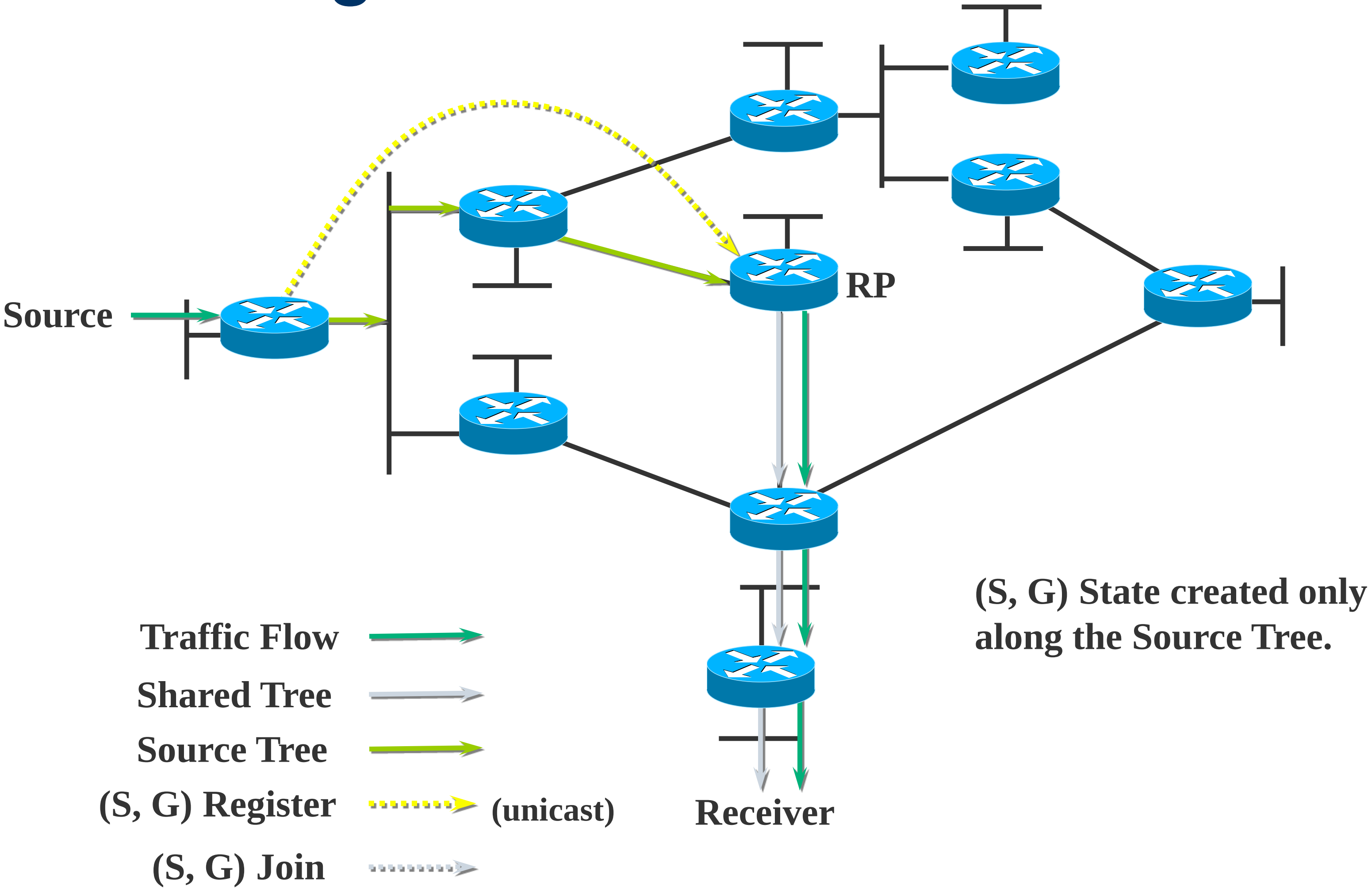
PIM SM Sender Registration



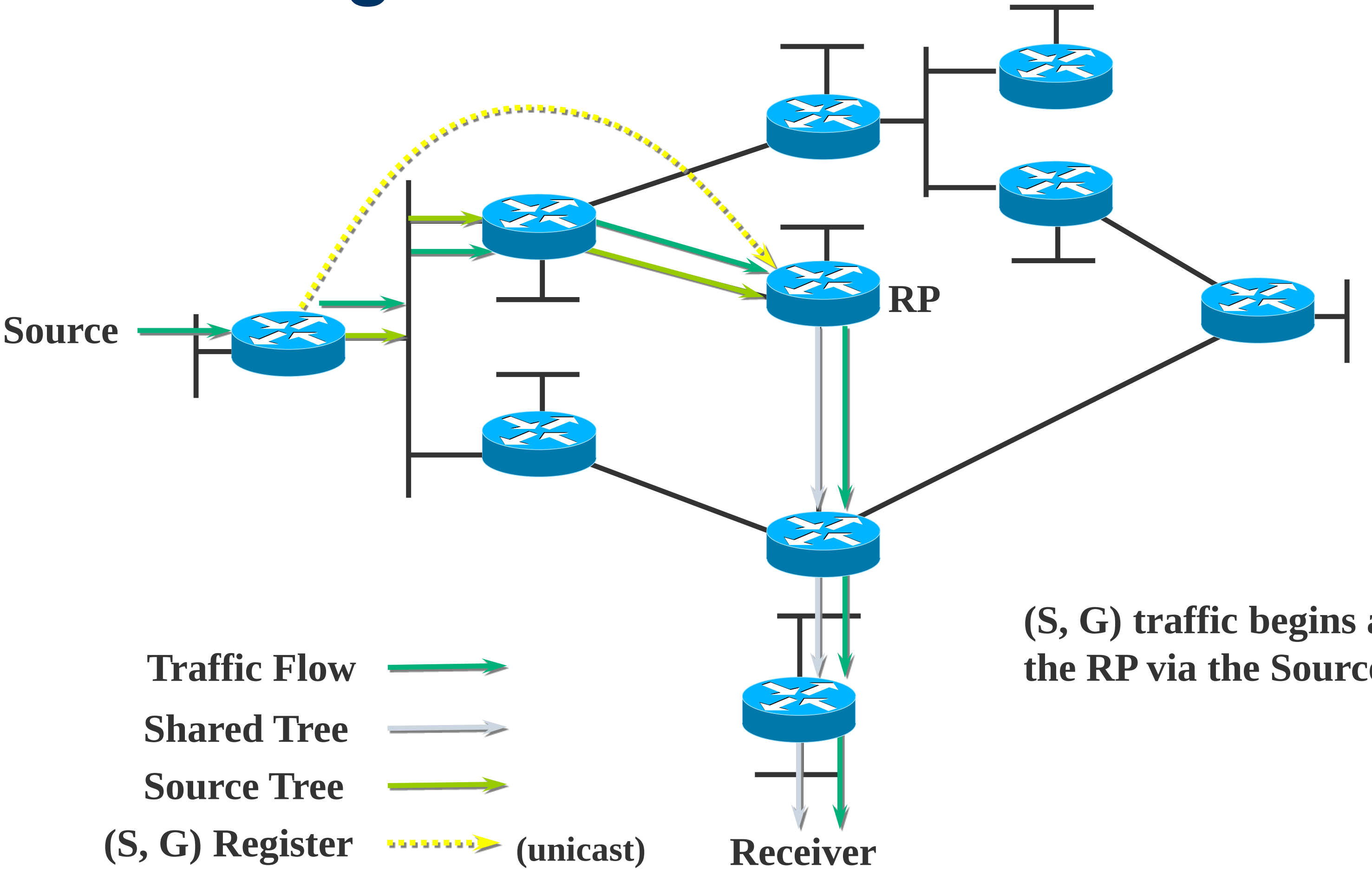
PIM SM Sender Registration



PIM SM Sender Registration

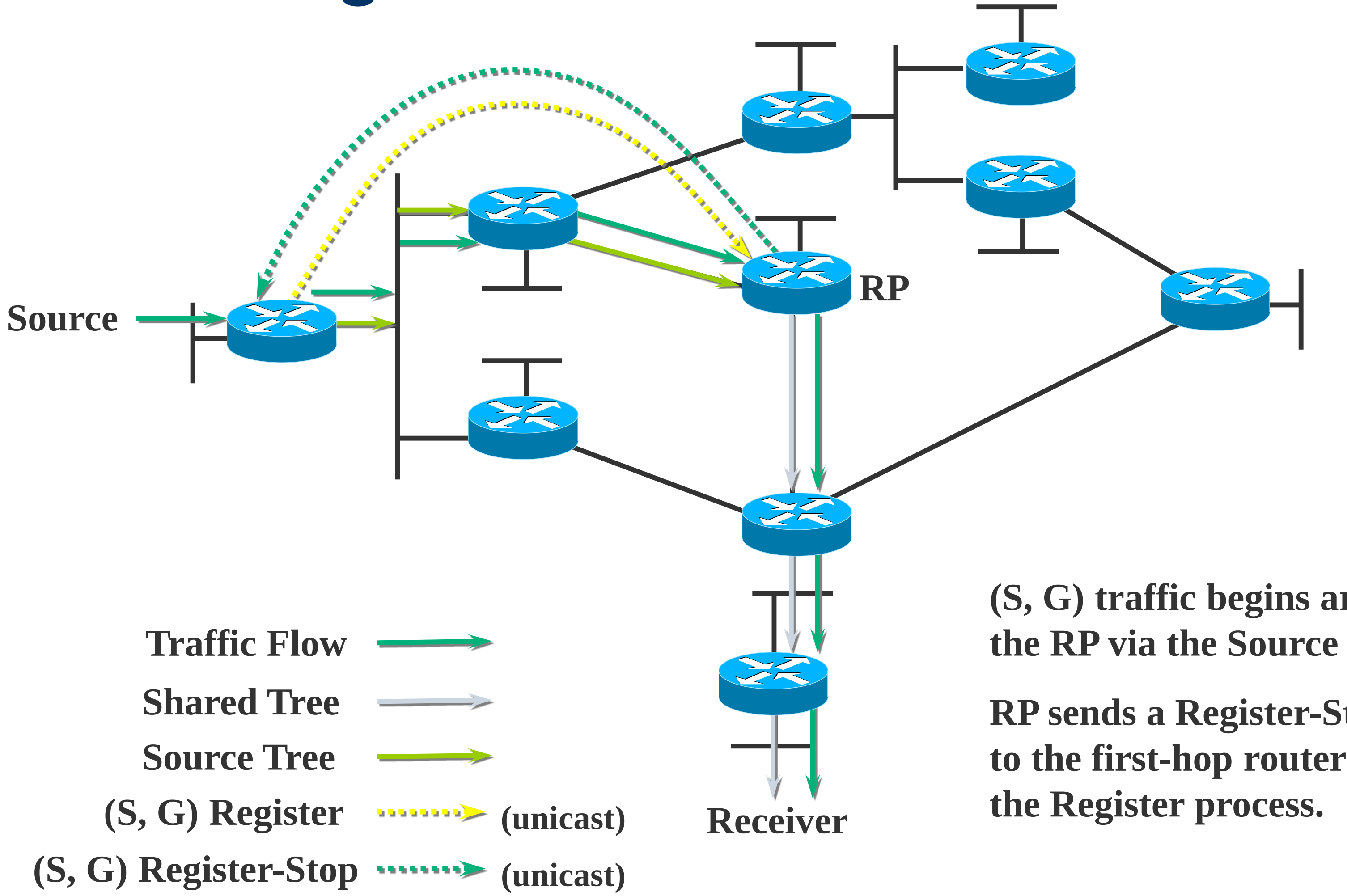


PIM SM Sender Registration



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

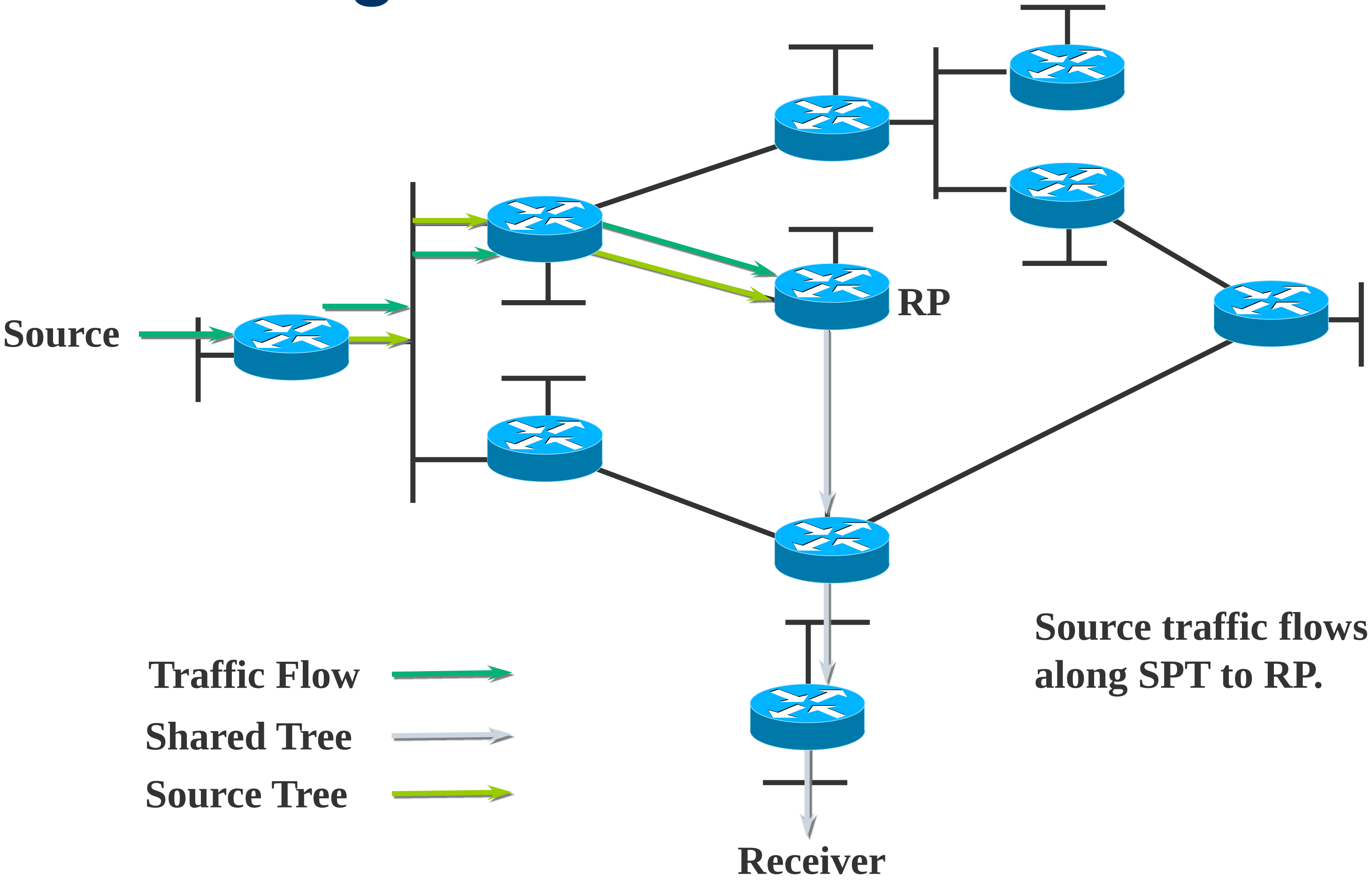
PIM SM Sender Registration



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

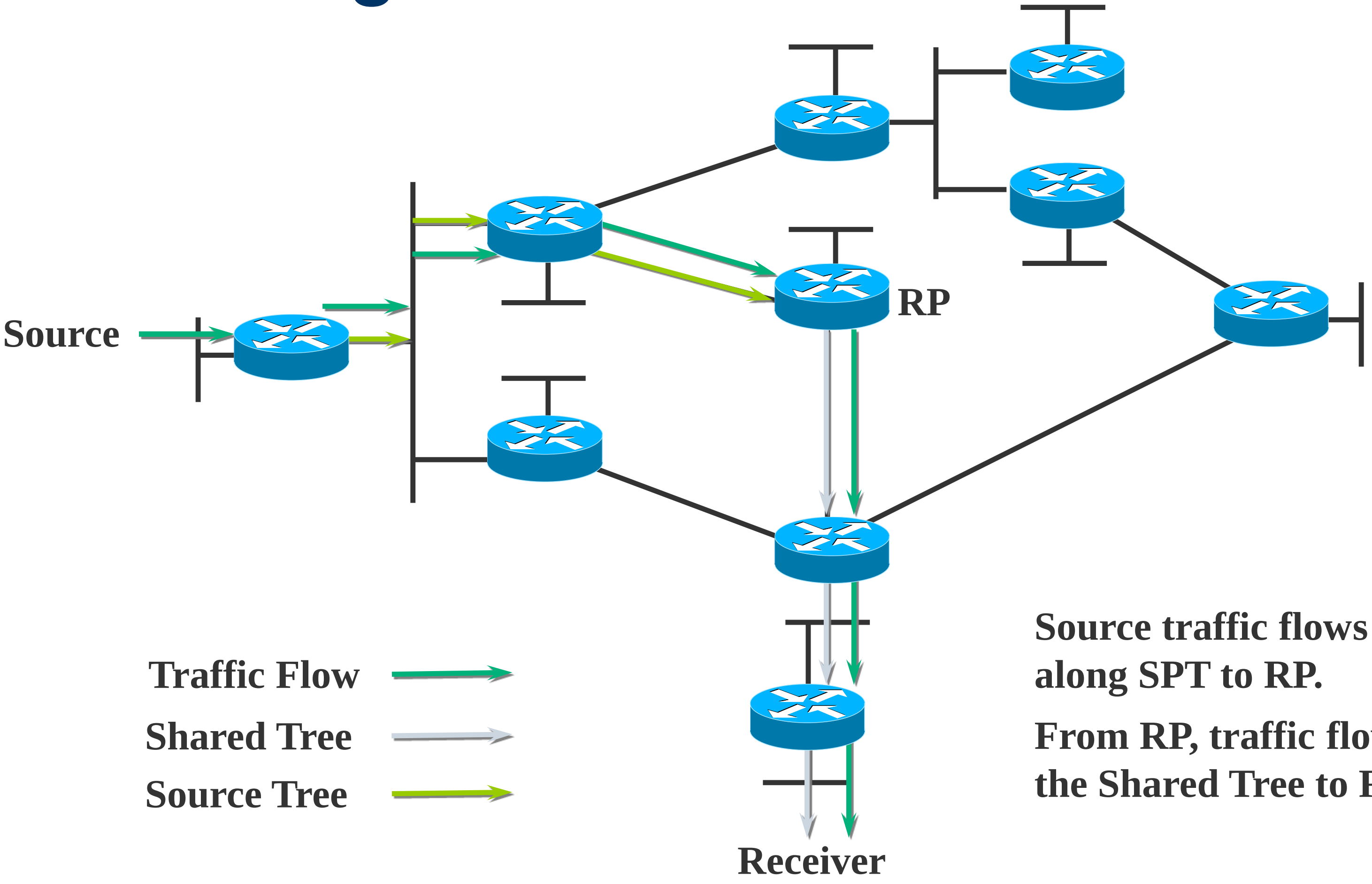
RP sends a Register-Stop back to the first-hop router to stop the Register process.

PIM SM Sender Registration



Source traffic flows natively along SPT to RP.

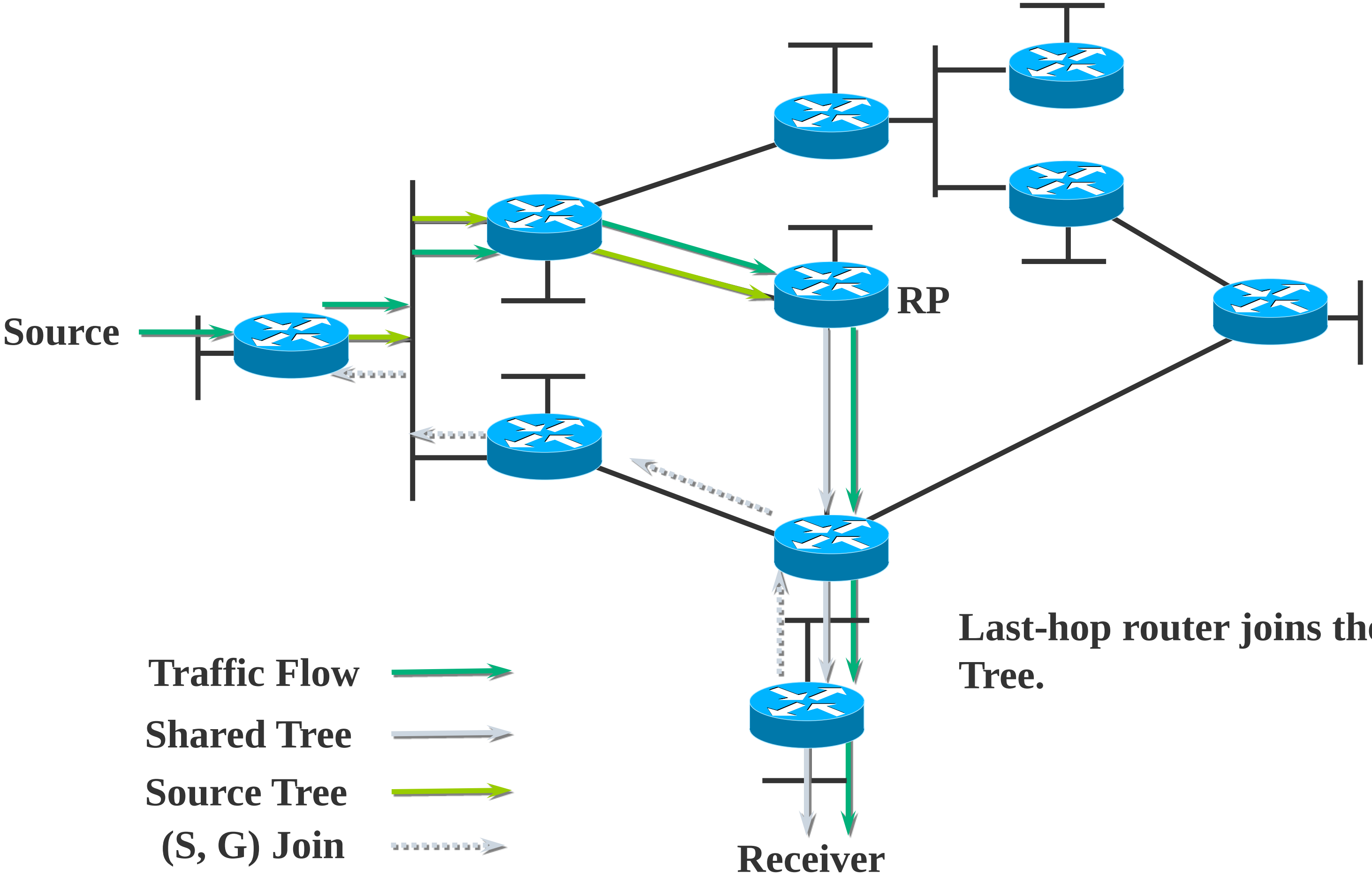
PIM SM Sender Registration



Source traffic flows natively along SPT to RP.

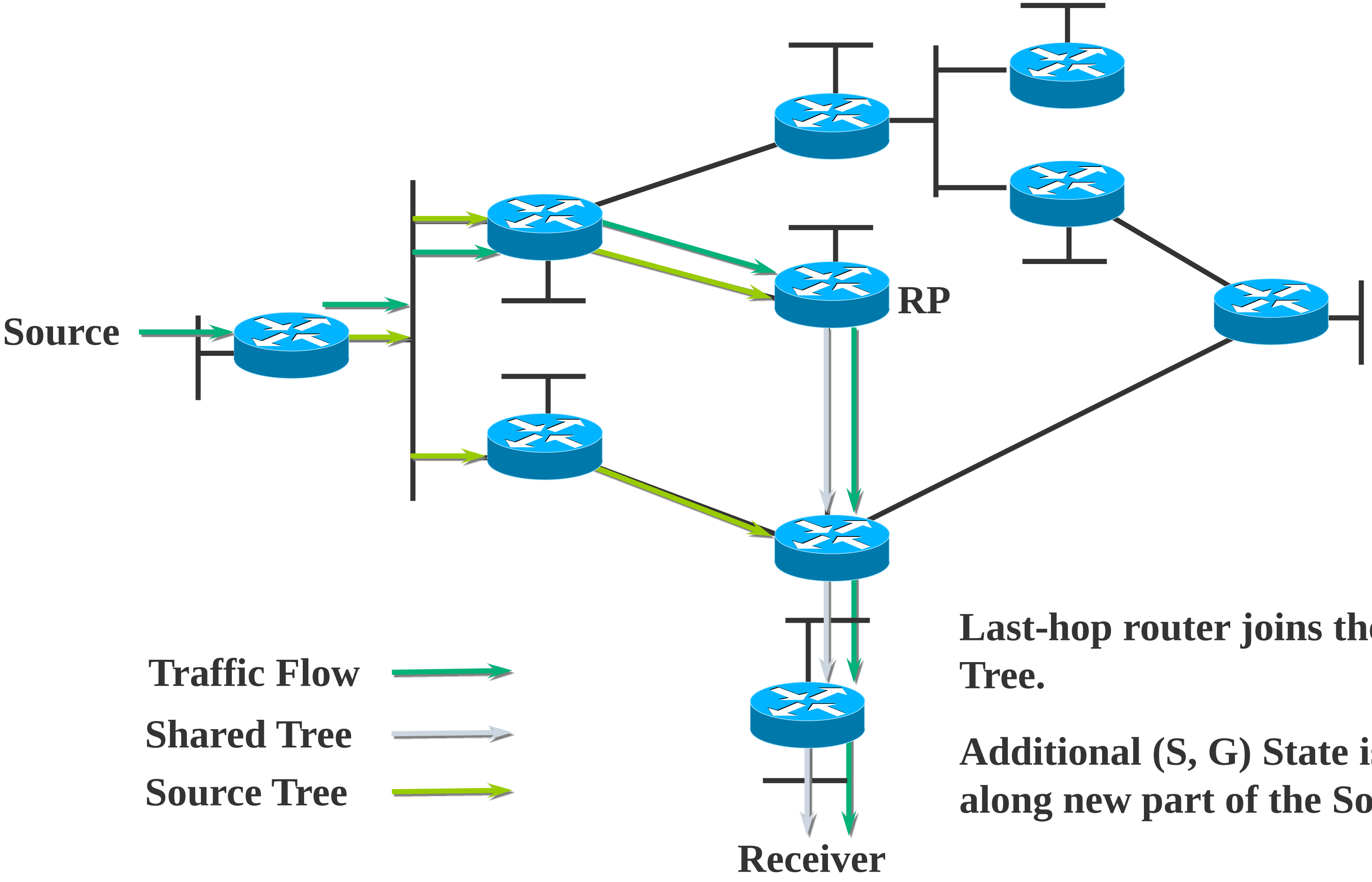
From RP, traffic flows down the Shared Tree to Receivers.

PIM SM Short Cut



Last-hop router joins the Source Tree.

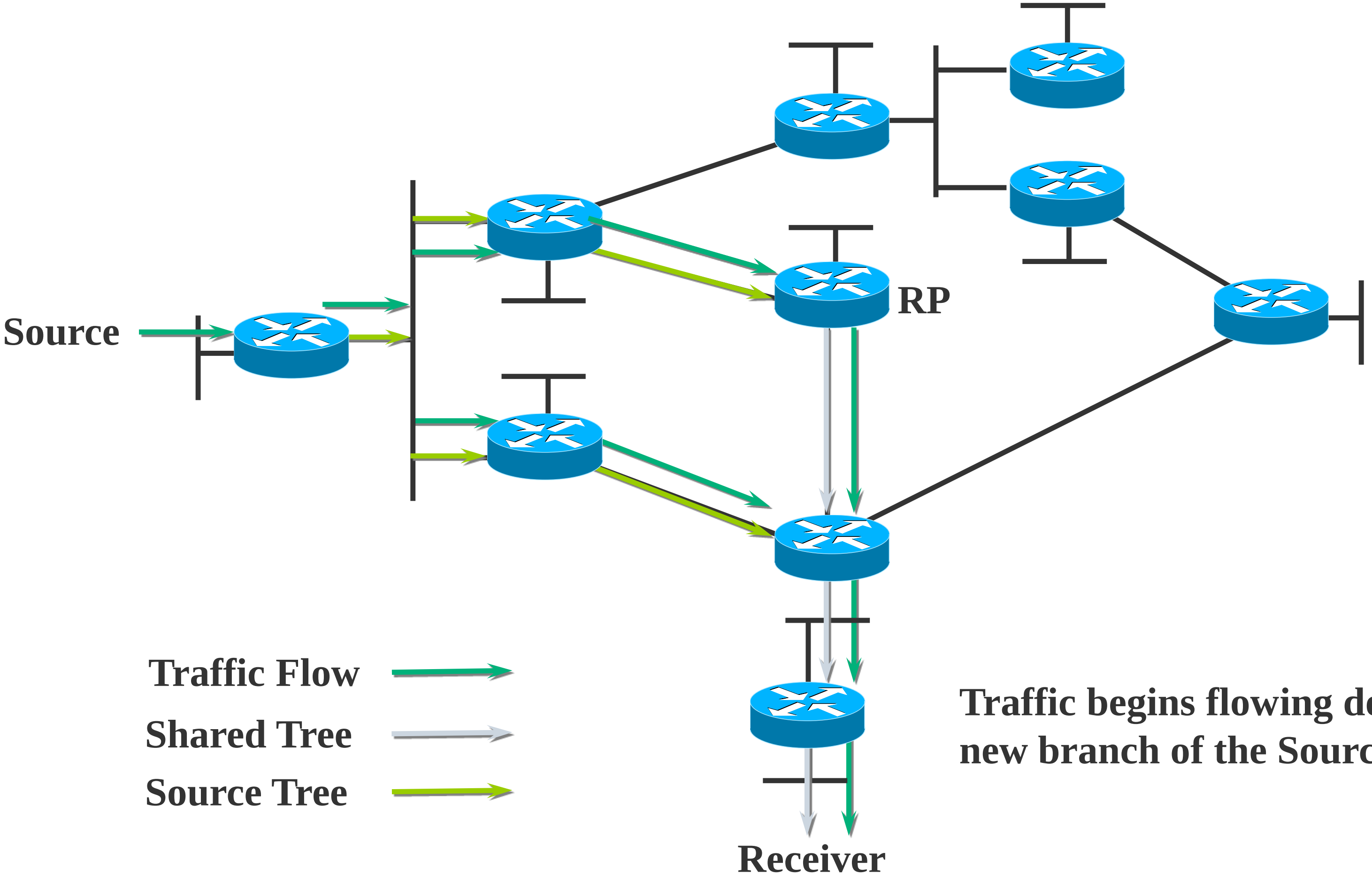
PIM SM Short Cut



Last-hop router joins the Source Tree.

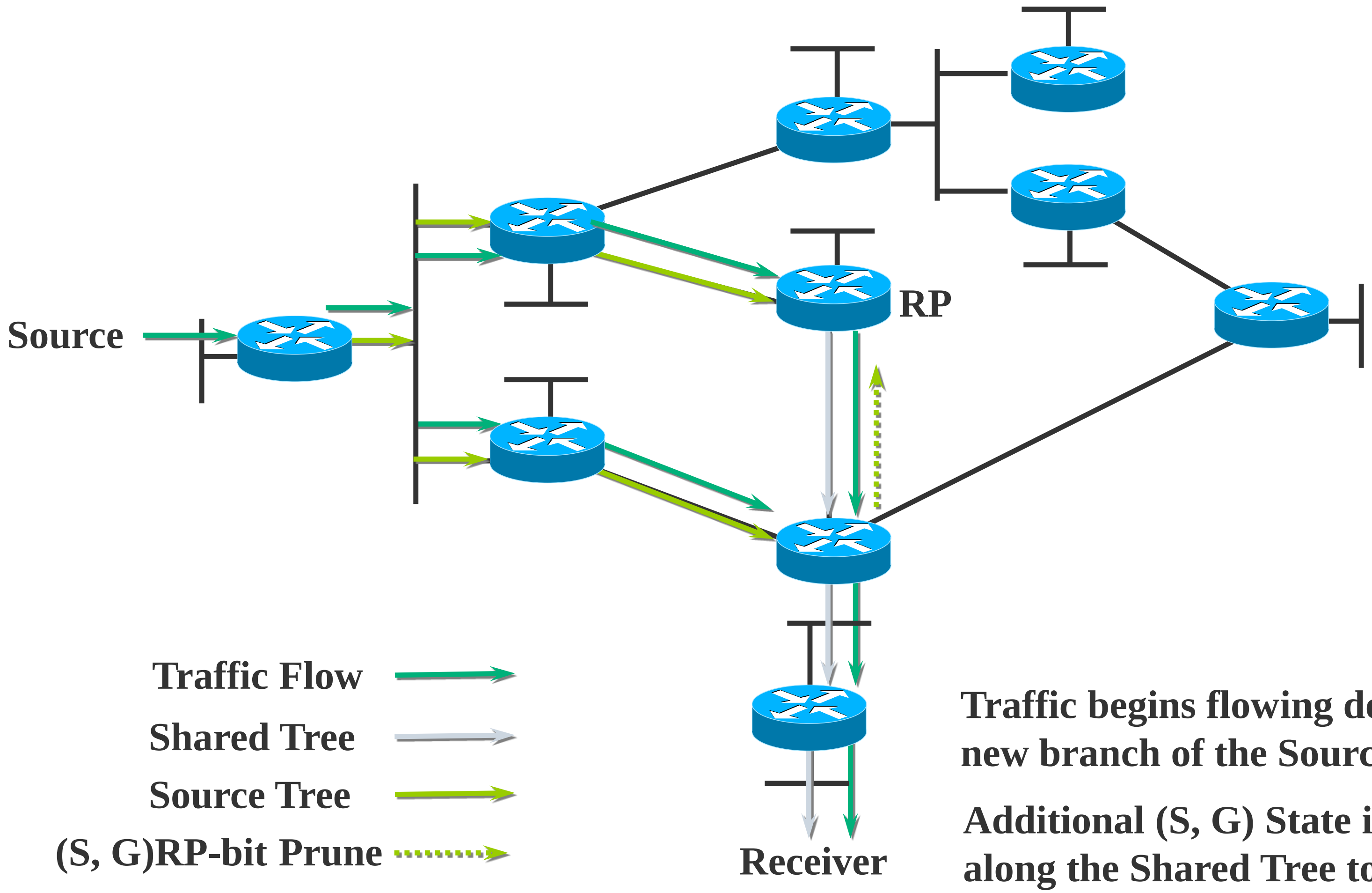
Additional (S, G) State is created along new part of the Source Tree.

PIM SM Short Cut



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

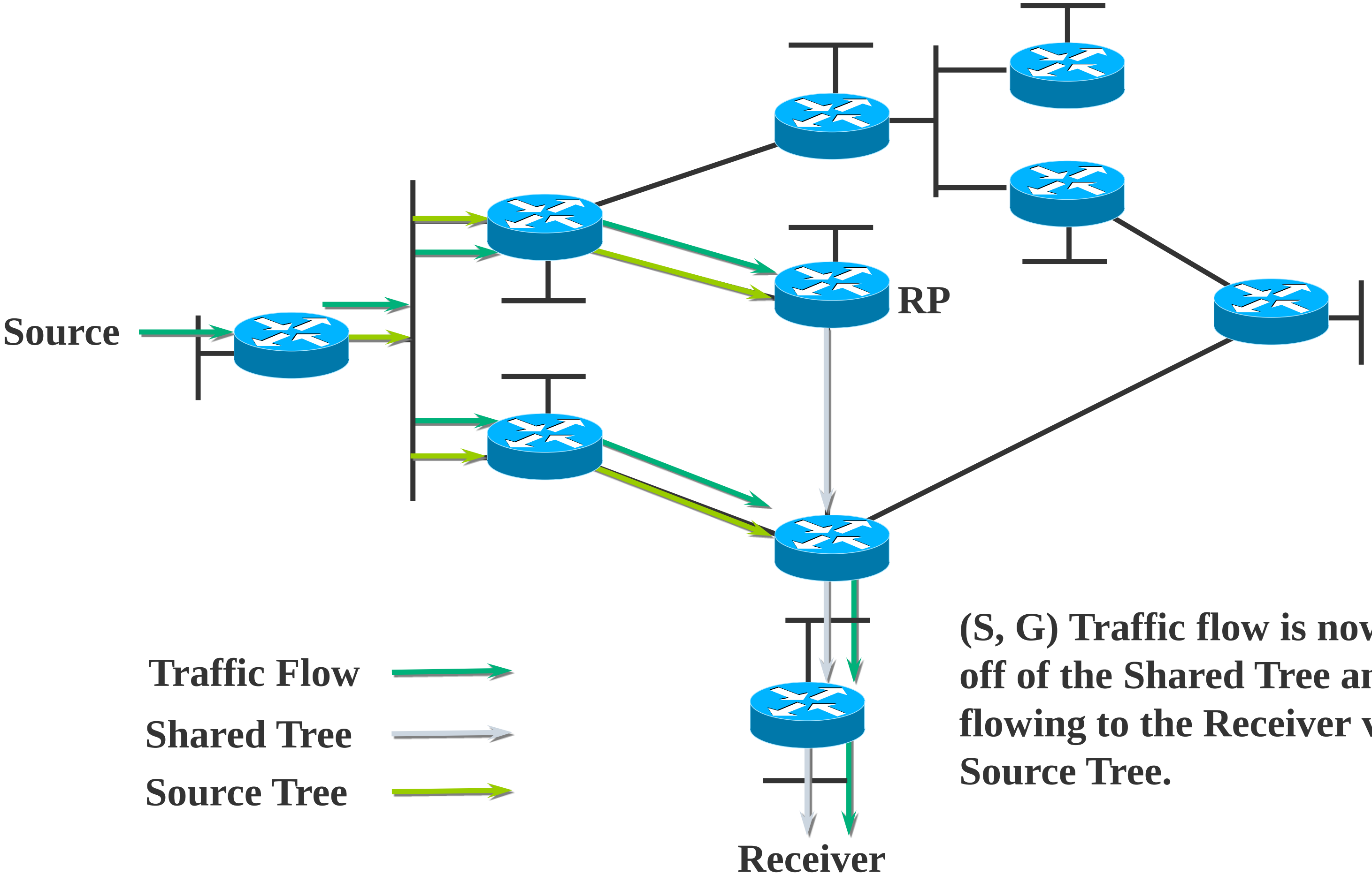
PIM SM Short Cut



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

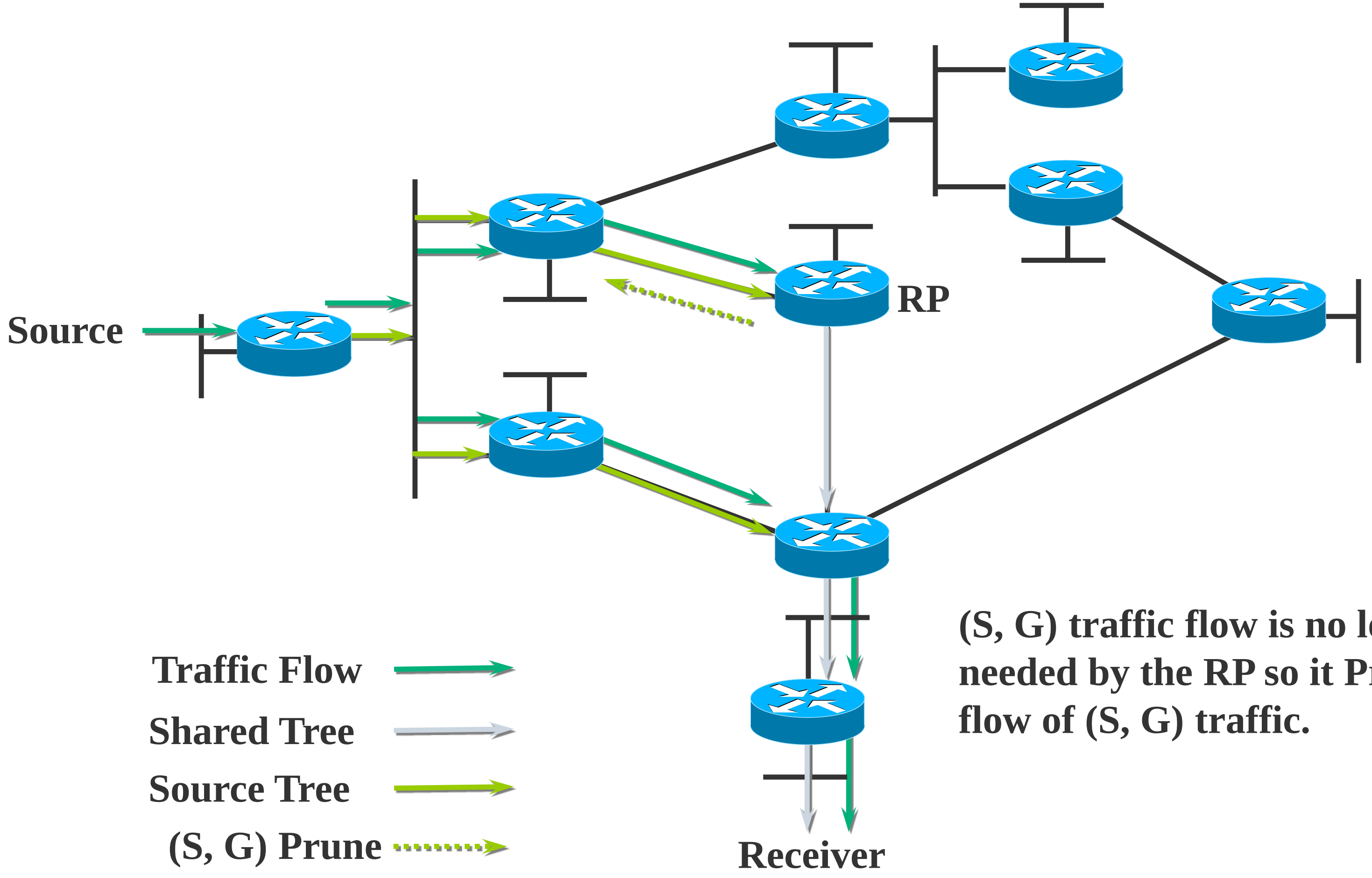
Additional (S, G) State is created along the Shared Tree to prune off (S, G) traffic.

PIM SM Short Cut



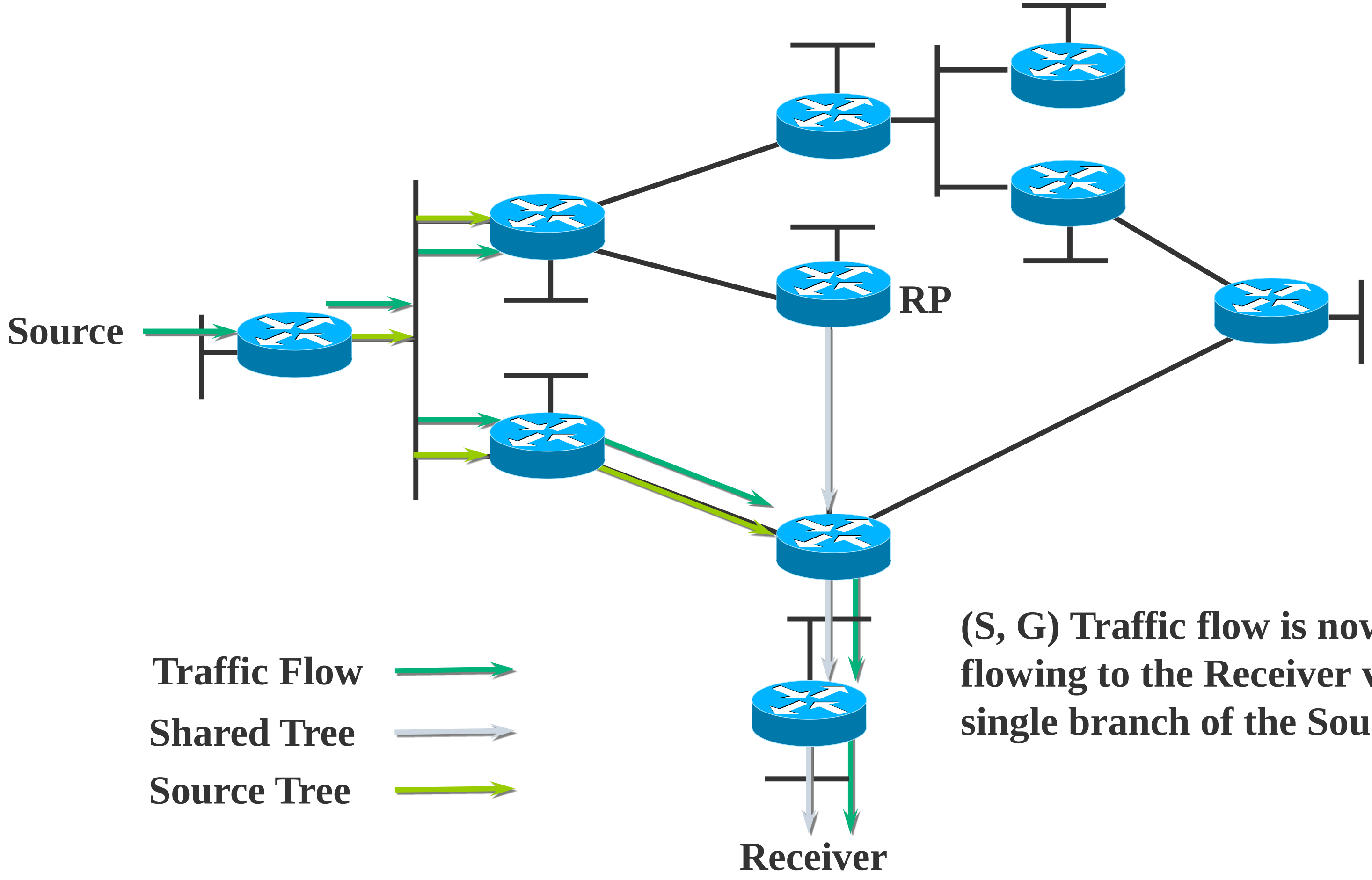
(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver via the Source Tree.

PIM SM Short Cut



(S, G) traffic flow is no longer needed by the RP so it Prunes the flow of (S, G) traffic.

PIM SM Short Cut



(S, G) Traffic flow is now only flowing to the Receiver via a single branch of the Source Tree.

Bidirectional PIM - 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)
- RPA roots a shared tree of designated forwarders (DFs):
  - One router per link with best route to RPA
  - Forwarding on this shared tree is bidirectional
- Explores a domain by per group shared forwarding states:
  - "NoInfo“ or "Include"
  - Decouples state management from data plane

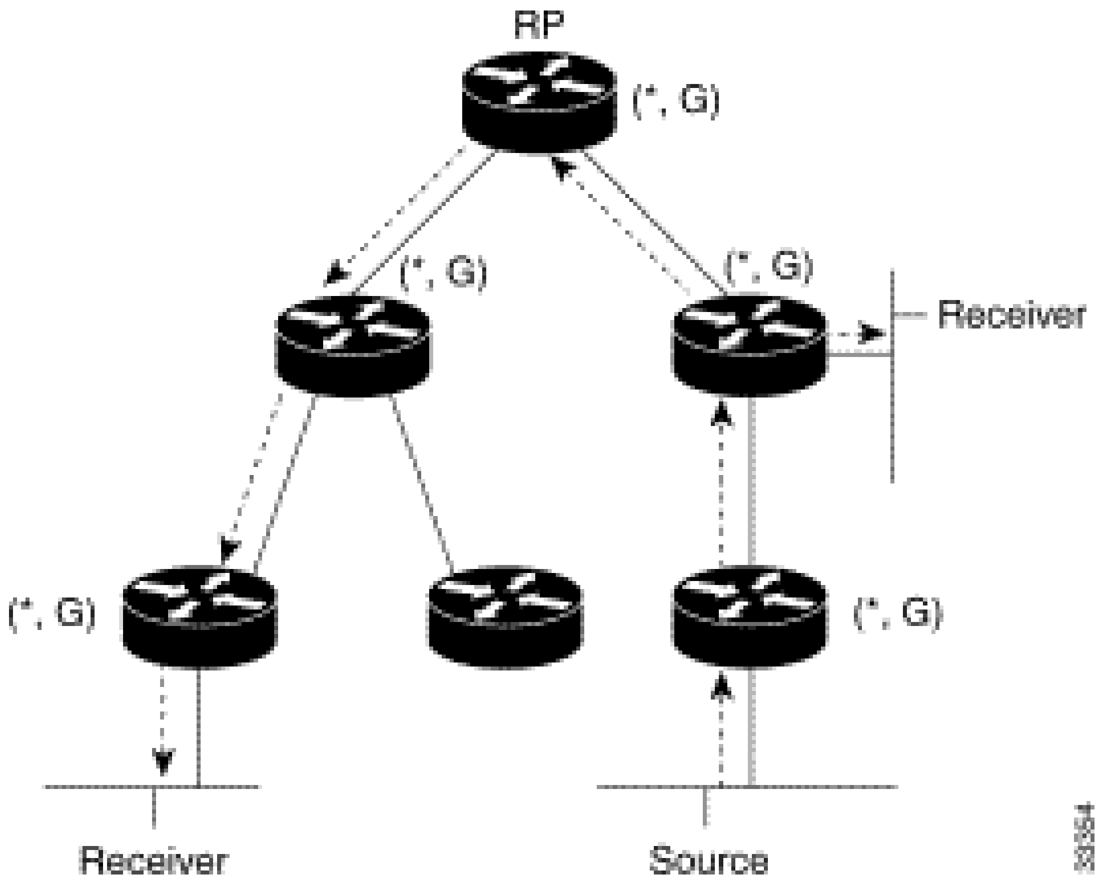
Bidirectional PIM (2)

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



Source Specific Multicast - SSM

Standardised with PIM (RFC 3569, 4607, 7761)

Assumes source address known at receiver

- Allows for source selection
- Source discovery offline or via MSDP

Receiver subscribes to (S,G) using IGMPv3/MLDv2

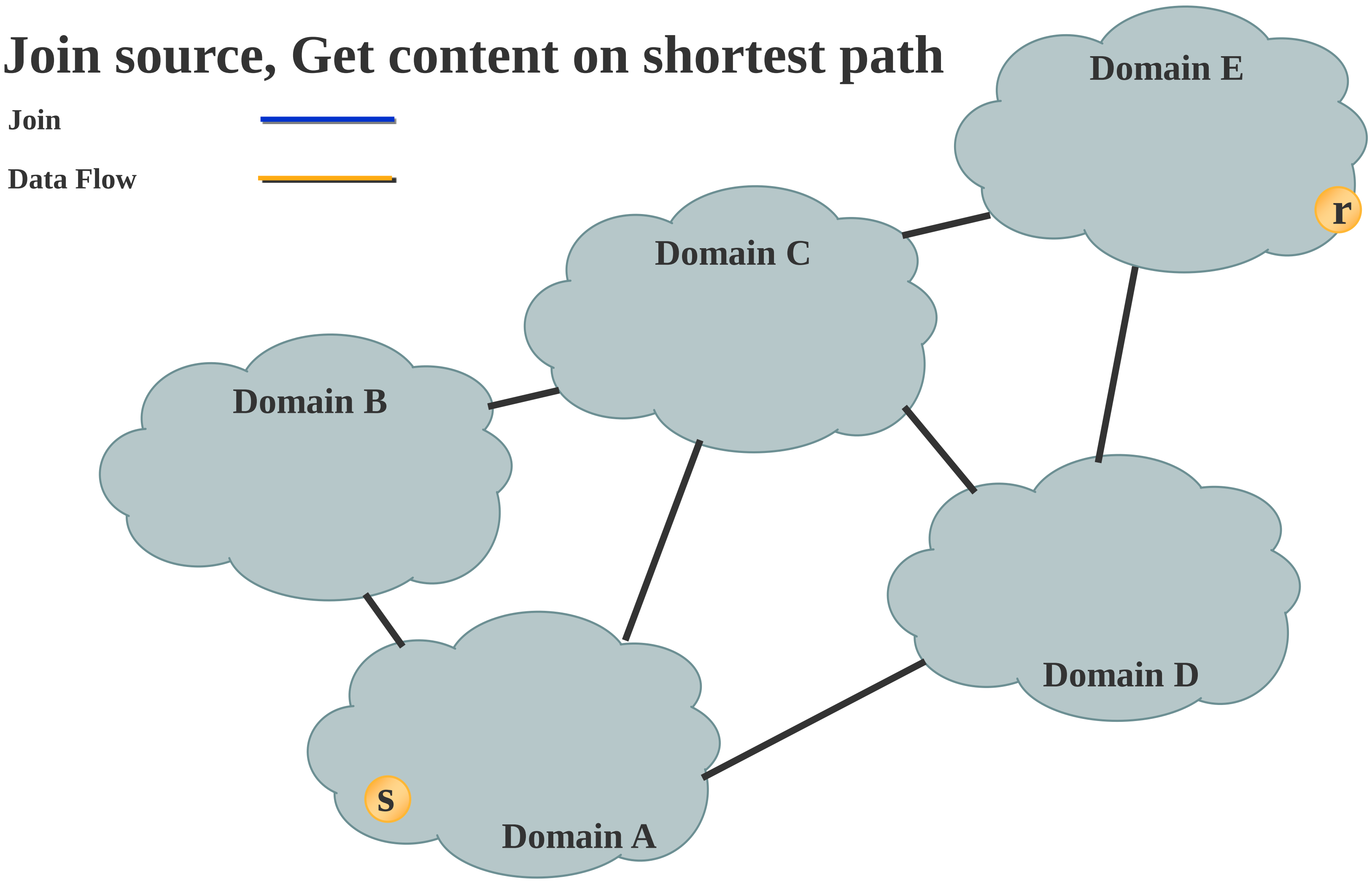
- No state aggregation on shared trees

Routing: PIM-SSM, a subset of PIM-SM

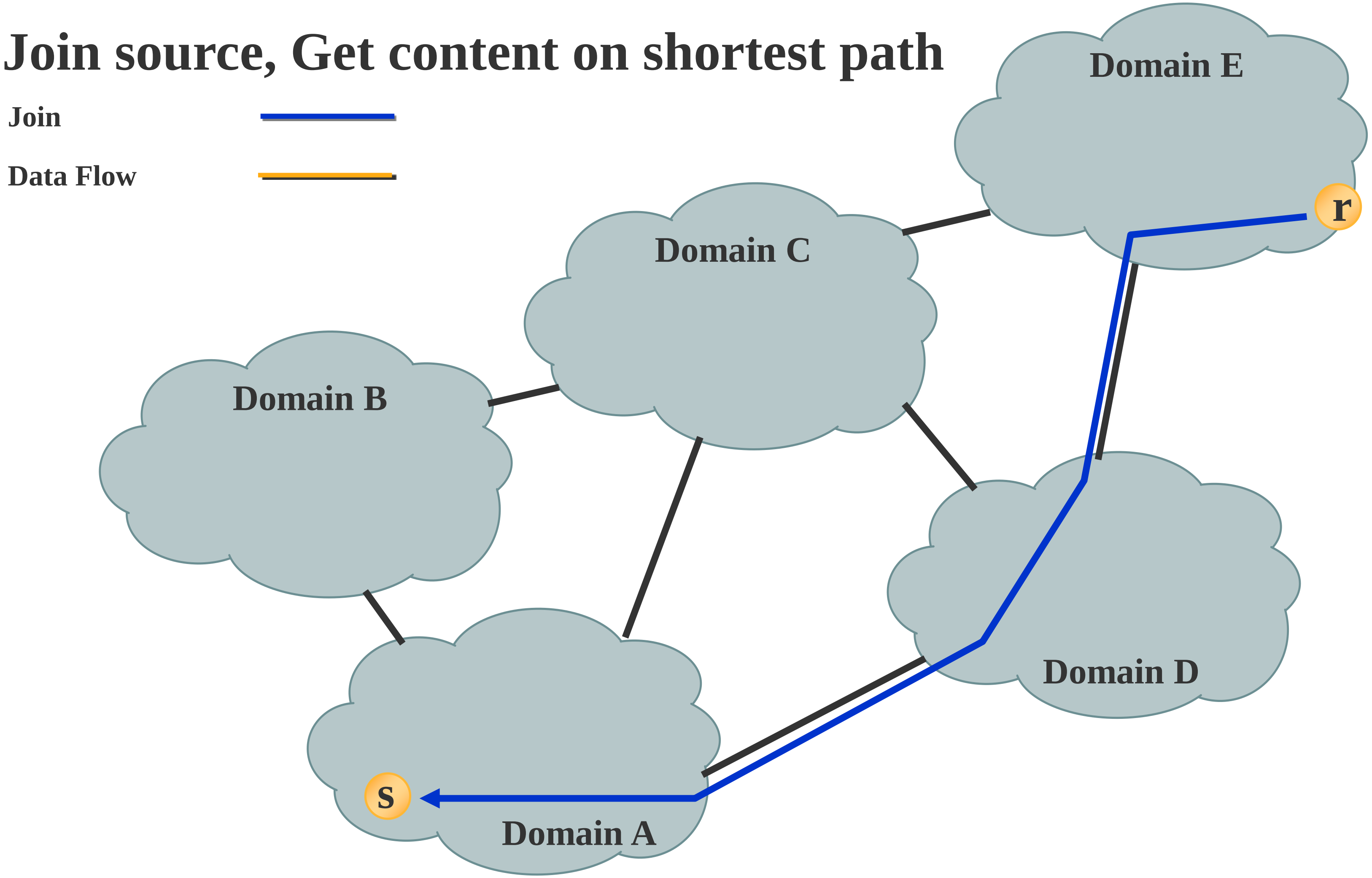
- Obsoletes rendezvous points & flooding

Simpler, well suited for single source media broadcast or inter-domain apps

SSM Routing

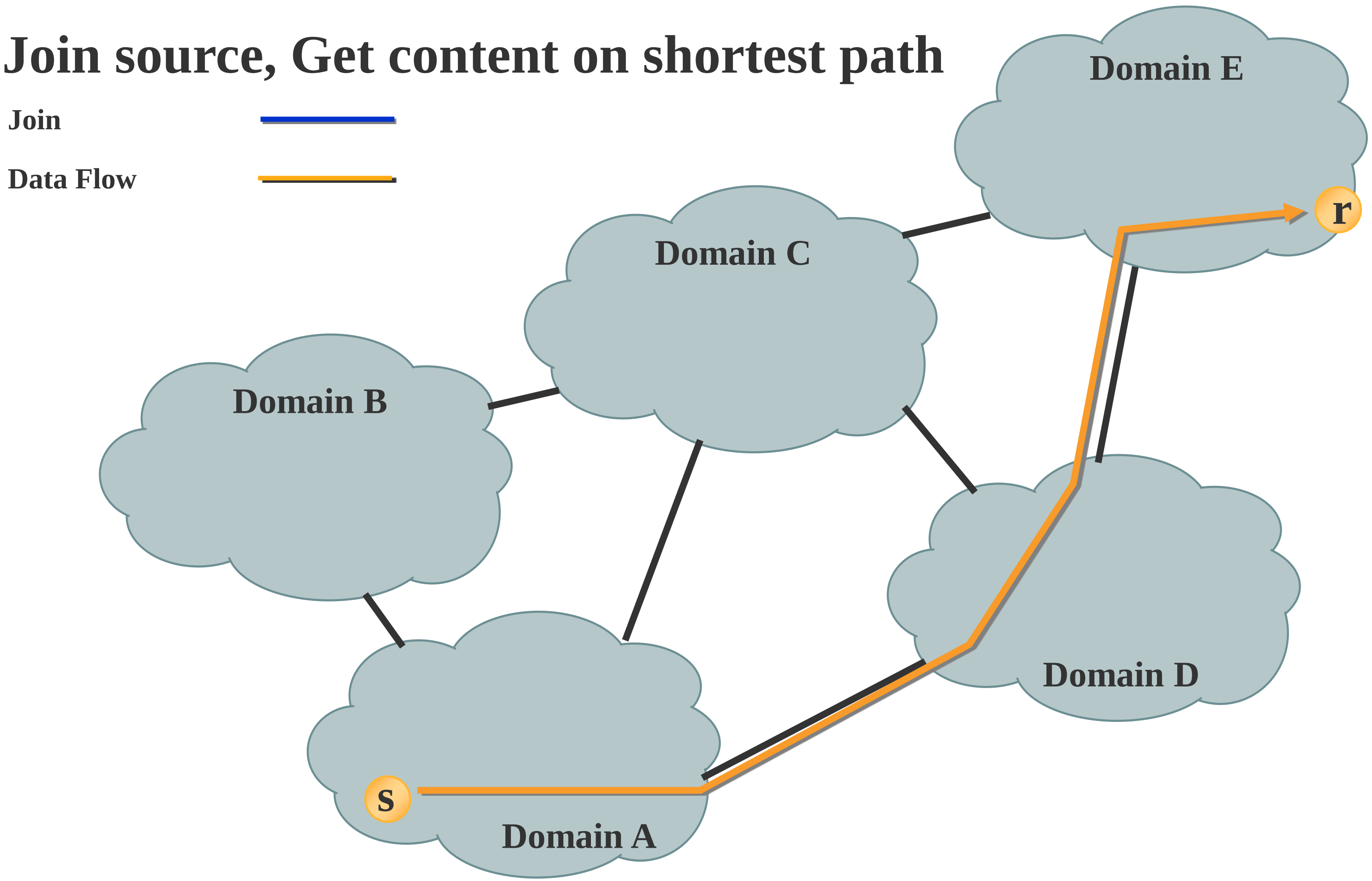

Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
Domain D
s
Domain A

SSM Routing

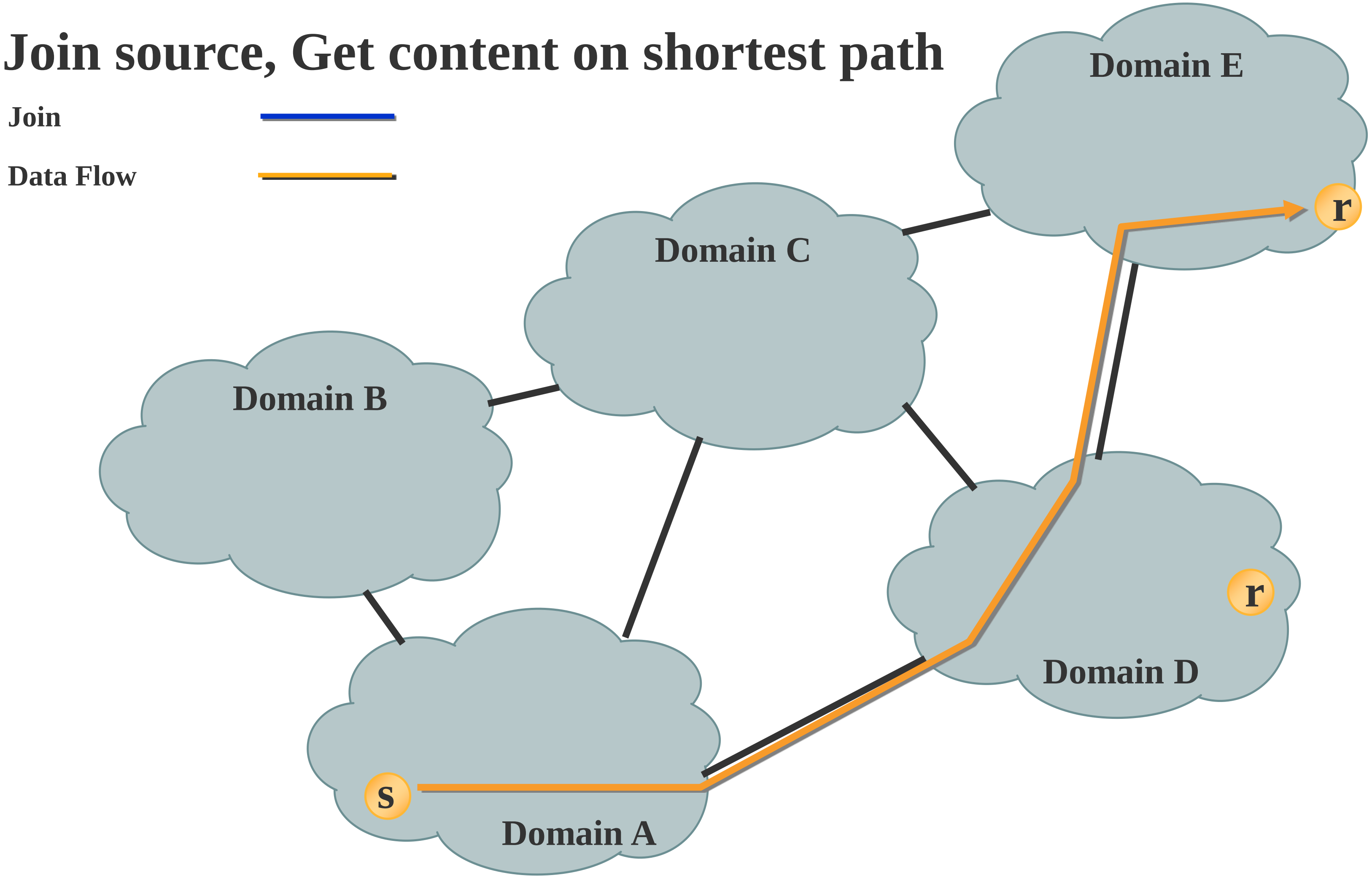


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
Domain D
S
Domain A

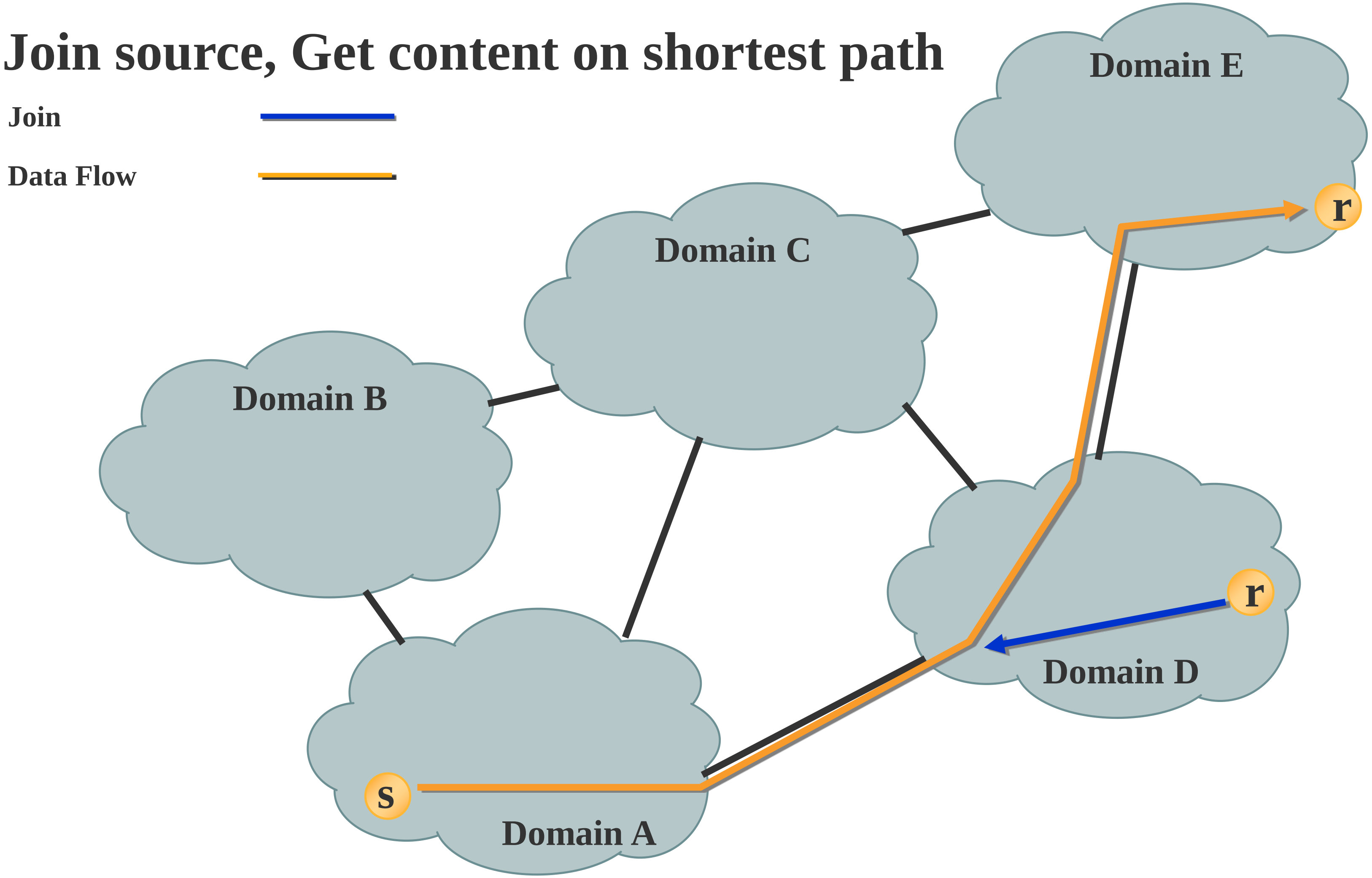
SSM Routing


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
Domain D
s
Domain A

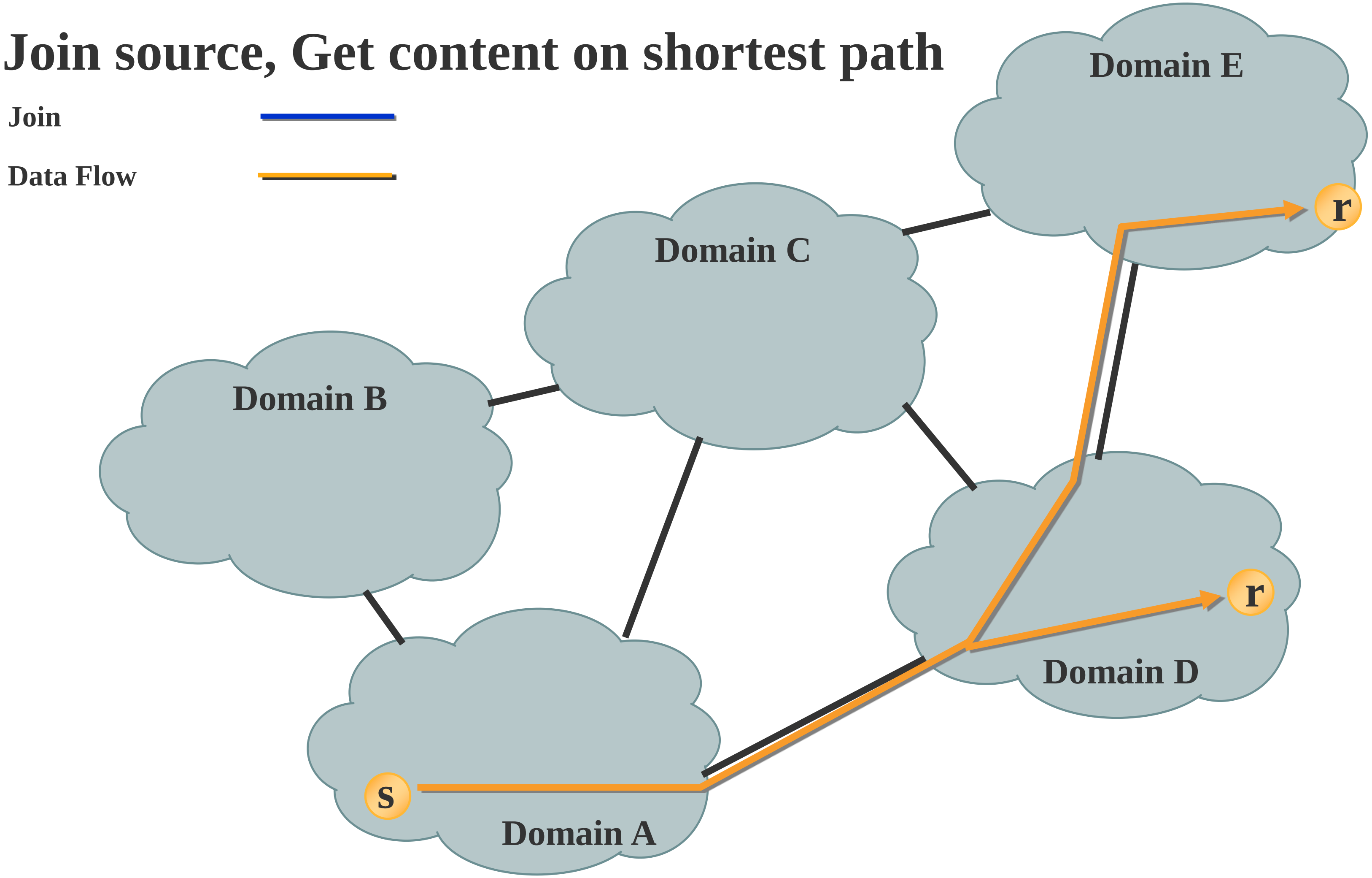
SSM Routing


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
r
Domain D
s
Domain A

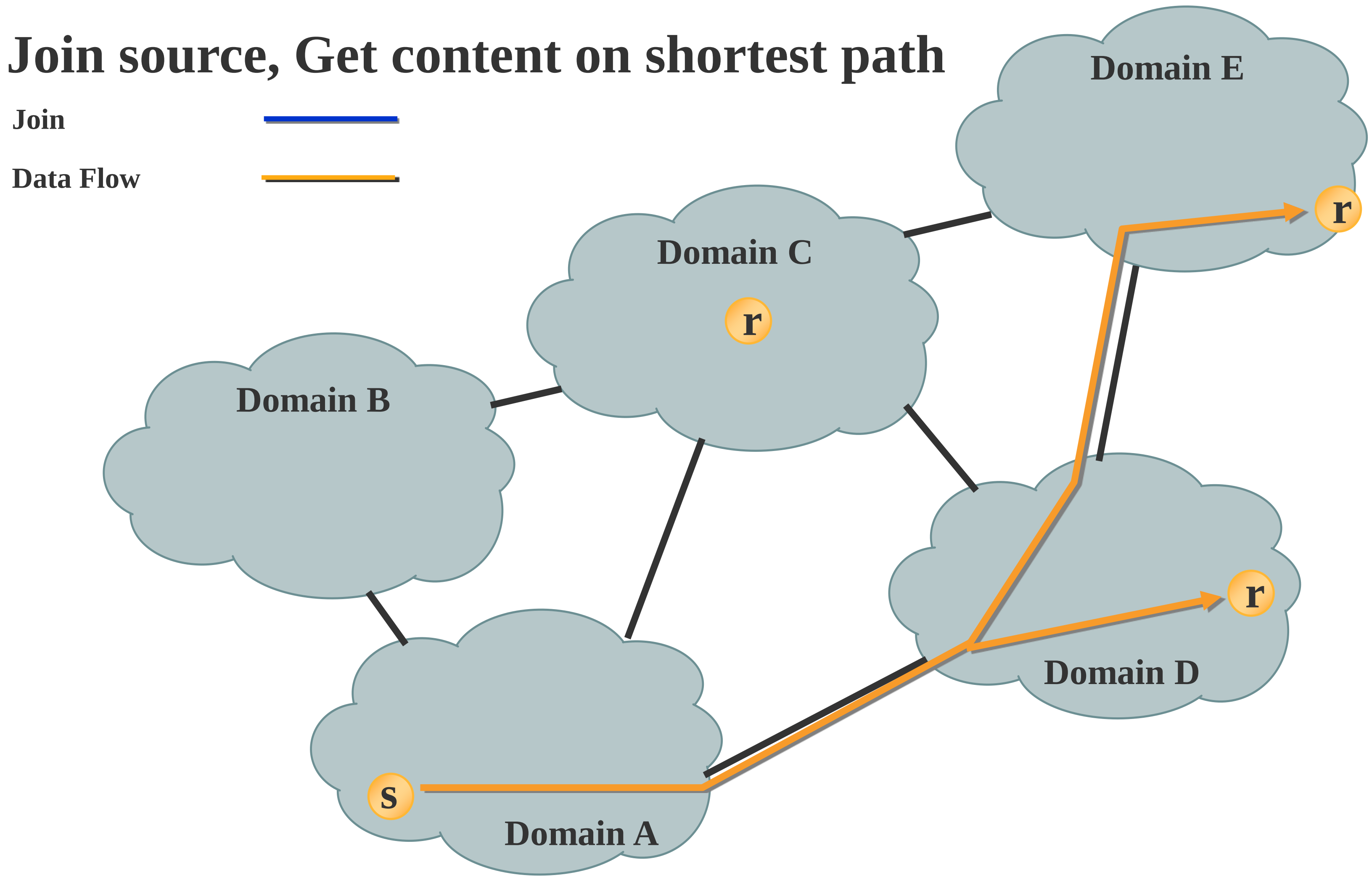
SSM Routing


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
r
Domain D
s
Domain A

SSM Routing

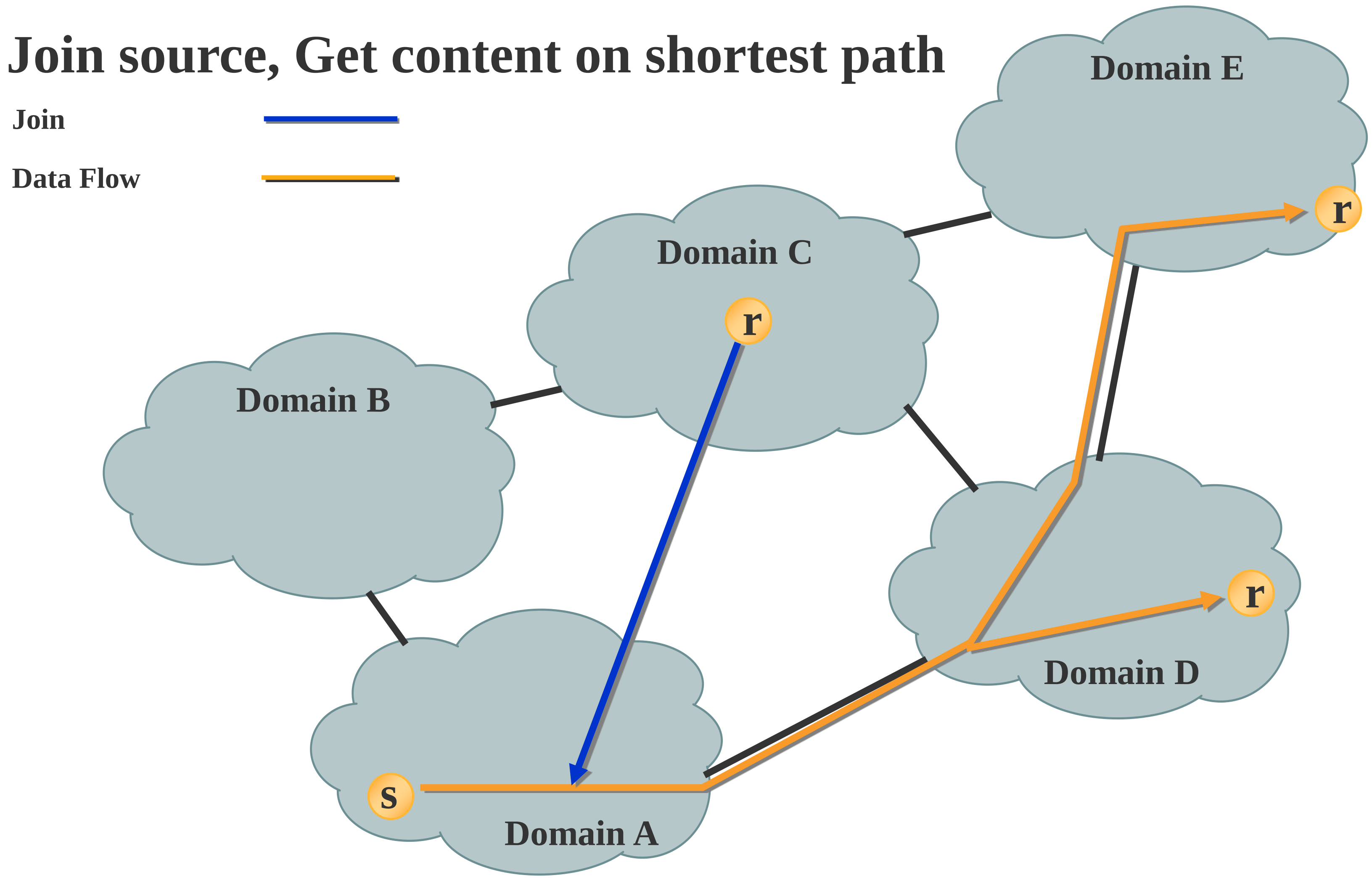

Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
Domain B
Domain D
r
s
Domain A

SSM Routing



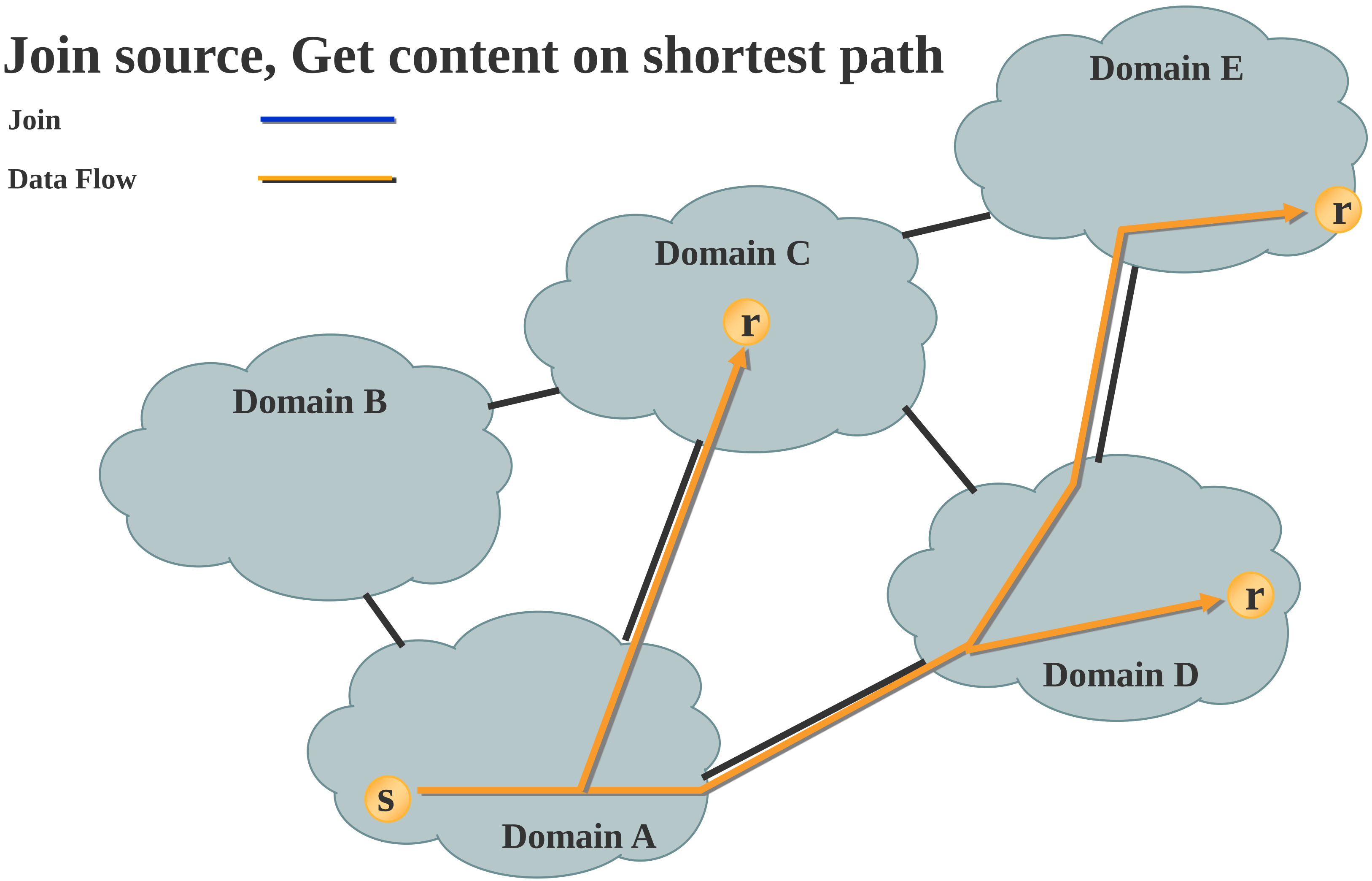
Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
r
Domain B
Domain D
r
s
Domain A

SSM Routing

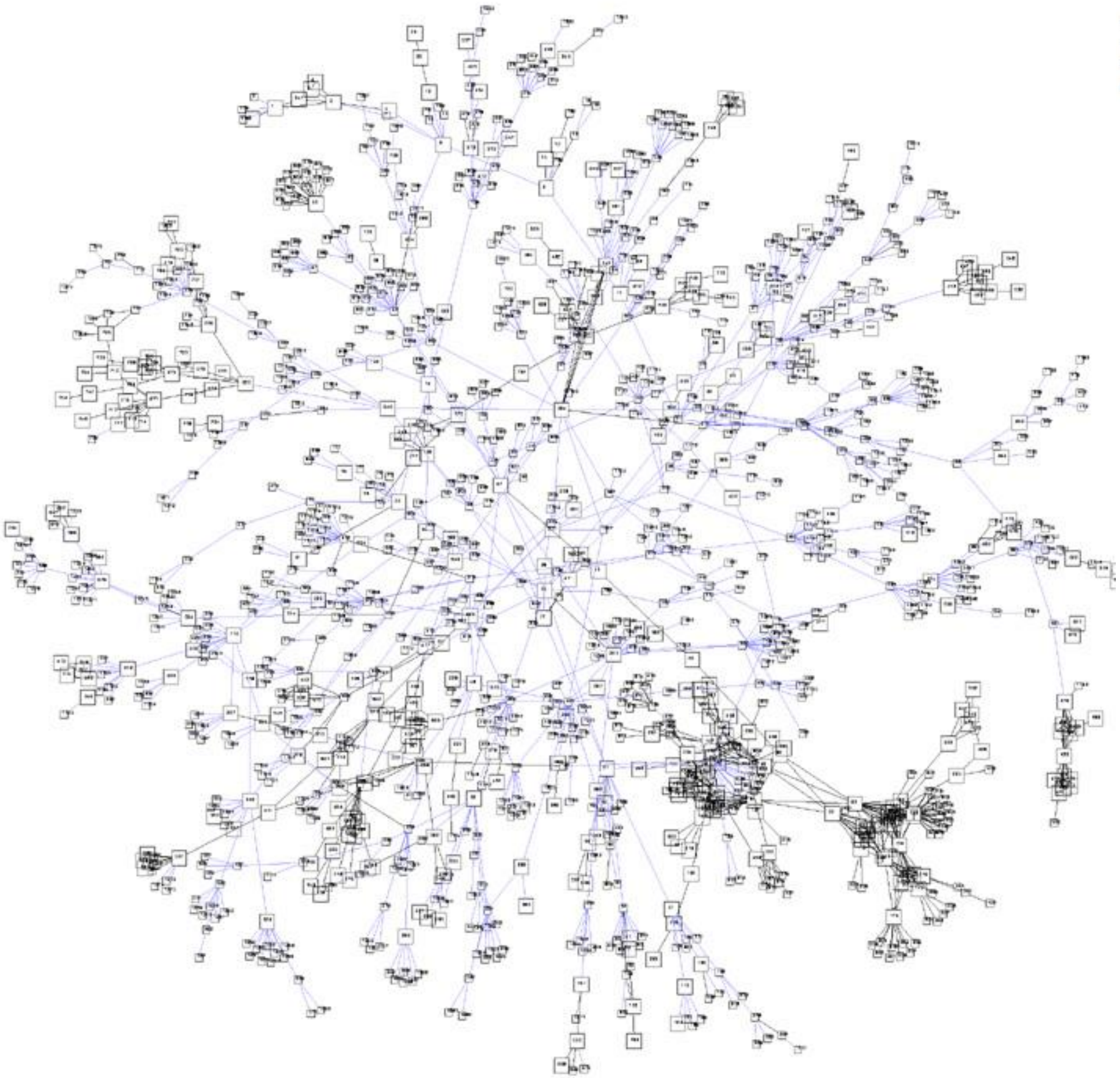


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
r
Domain B
r
Domain D
s
Domain A

SSM Routing


Join source, Get content on shortest path
Join
Data Flow
Domain E
r
Domain C
r
Domain B
Domain D
r
s
Domain A

MBone



Visualisation of Multicast Group

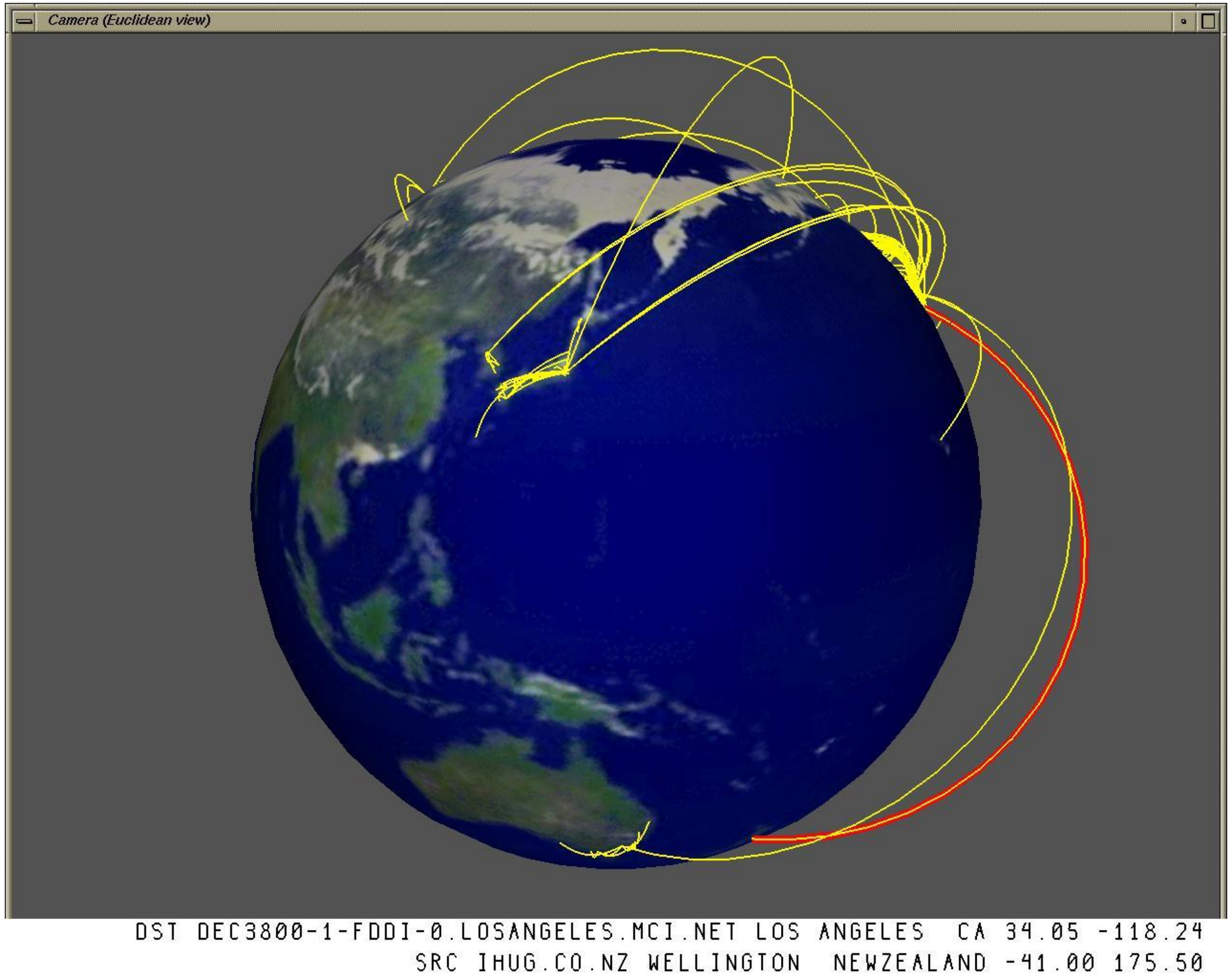


Image & Video by Tamara Munzer, Univ. of British Columbia

HAW HAMBURG

IP Mcast Deployment Issues

- Complexity versus performance efficiency
  - IP Multicast most efficient, but burdens infrastructure
- Provider costs
  - Provisioning of knowledge, infrastructure & maintenance
- Provider revenue
  - Providers sell bandwidth : multicast saves bandwidth
  - Exception: provider offers s.a. IPTV
- Security
  - ASM simplifies DDoS-attacks
- Multicast distributes synchronously
  - VoD supersedes IPTV

HAW HAMBURG

Agenda

- Motivation
  - Content Distribution to Groups of Receivers
- IP Multicast
  - Host Group Model
  - Multicast Addressing
  - Group Membership Management
- Multicast Routing
  - Routing Algorithms
  - ASM Routing Protocols
  - SSM Routing
- Information Centric Networking
  - Motivation
  - ICN Approaches
  - Routing & Forwarding
  - ICN in the IoT

How Can We Improve?

- Content distribution is major Internet goal
  - Still highly redundant, but not synchronous
- Infrastructures have changed
  - Most popular content hosted on CDNs
- Security: DDoS threatens the Internet
  - Desire to prevent unwanted packets
- New use cases: Low-power wireless & IoT
  - Requires some delay / disruption tolerance

Information-centric Networking: Idea

Access content instead of nodes following a request/response paradigm

- Address content directly by name
- Augment content with (self-)authenticating credentials
- Provide ubiquitous in-network storage (caching)

Various approaches

- Seminal: TRIAD (Gritter & Cheriton 2001)
- Most popular: NDN (Van Jacobson et al. 2009)

Approaches to ICN

TRIAD
DONA
CCNx/NDN

} Routing on names

PSIRP/PURSUIT

Name resolution system publishes source routing identifiers (Bloom filters)

NetINF

Name resolution system refers to publisher IDs, routes to pub. locators

TRIAD

Gritter & Cheriton, 2001

Stanford started Future Internet Initiative
with a Multicast rework

Starting point: Make content replication better than CDNs – and open:

- Routing on names by augmenting IP routing
- Content delivery by HTTP/TCP/IP
- Architecture of Content Routers and Content Servers

Early concept of name aggregates

Community was not ready then

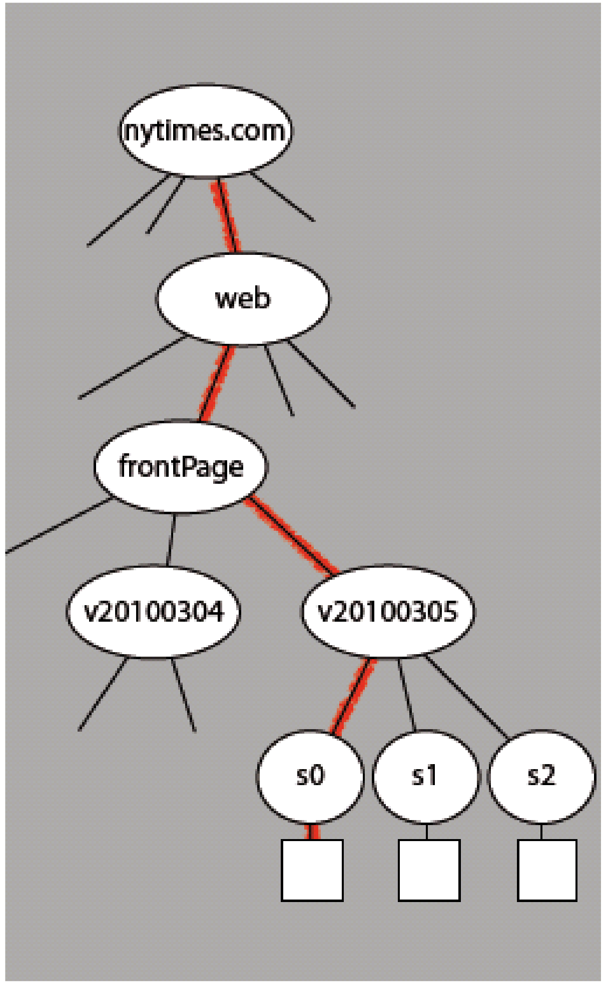
Named Data Networking

Van Jacobson et.al., 2009

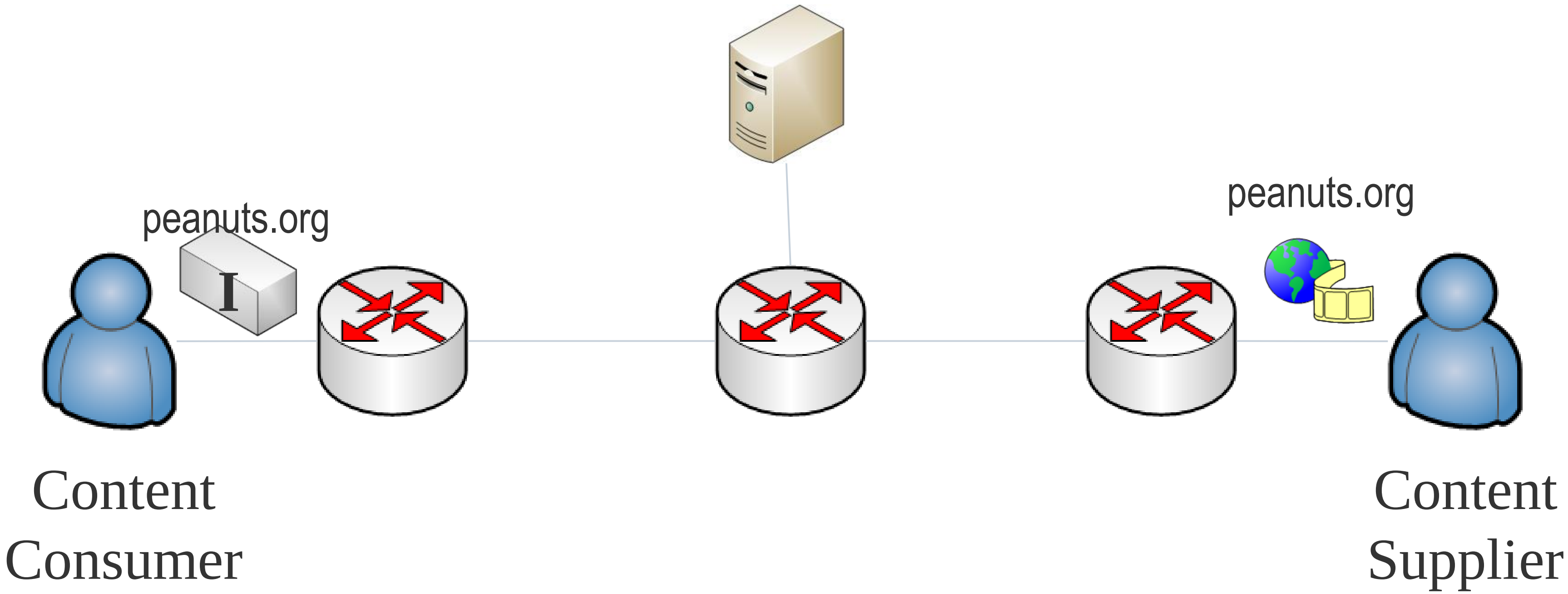
Routes on Names

- Source publishes Content ‚to a network‘ that caches and replicates
- Network distributes names in its routing protocol
- Subscriber requests content from network by name
- Request places ‘trail of breadcrumbs’ in the network

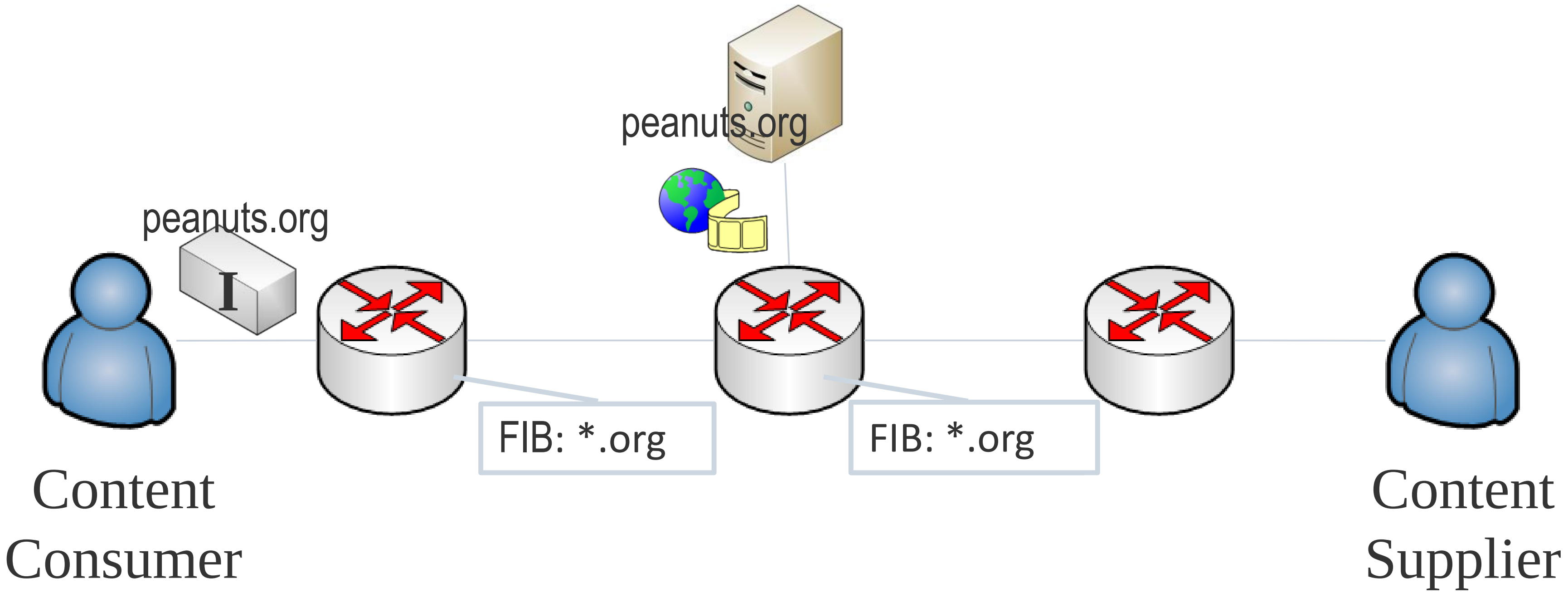
- Forwarding on reverse path
  - No IP layer, no source addresses
- Universal On-Path Caching



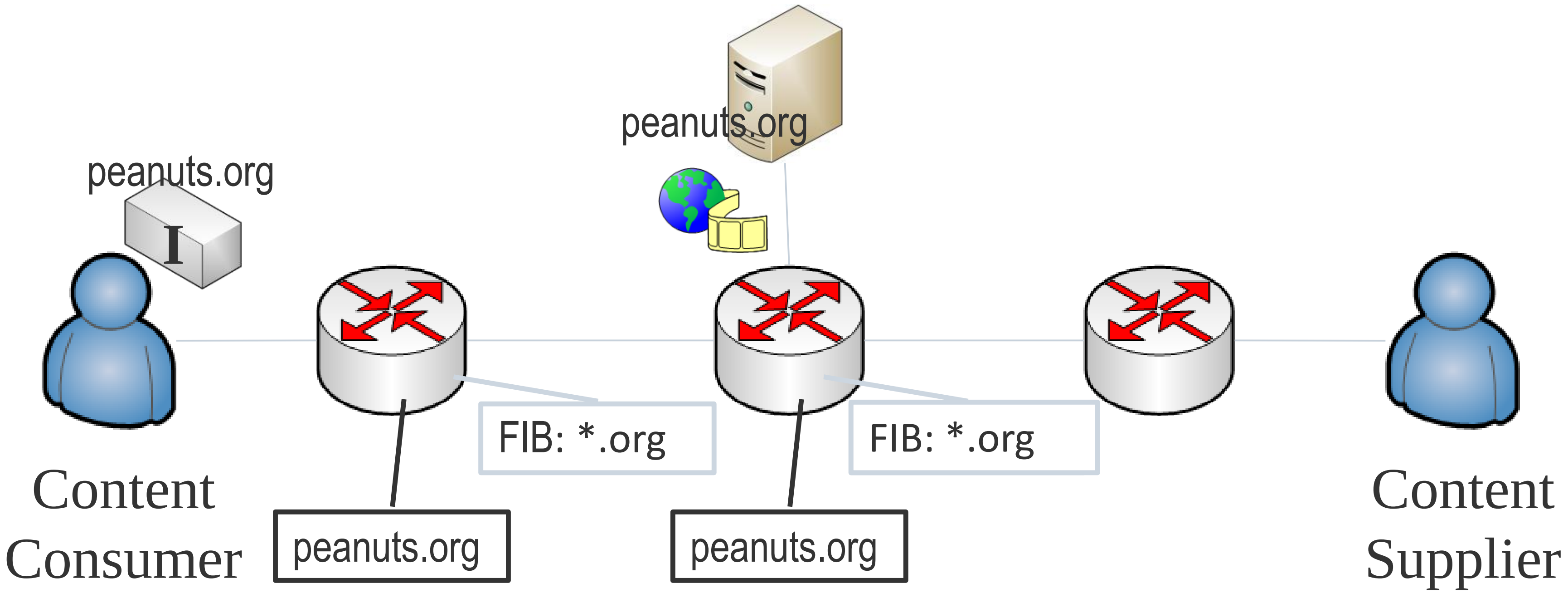
Basic Content Centric Routing and Forwarding



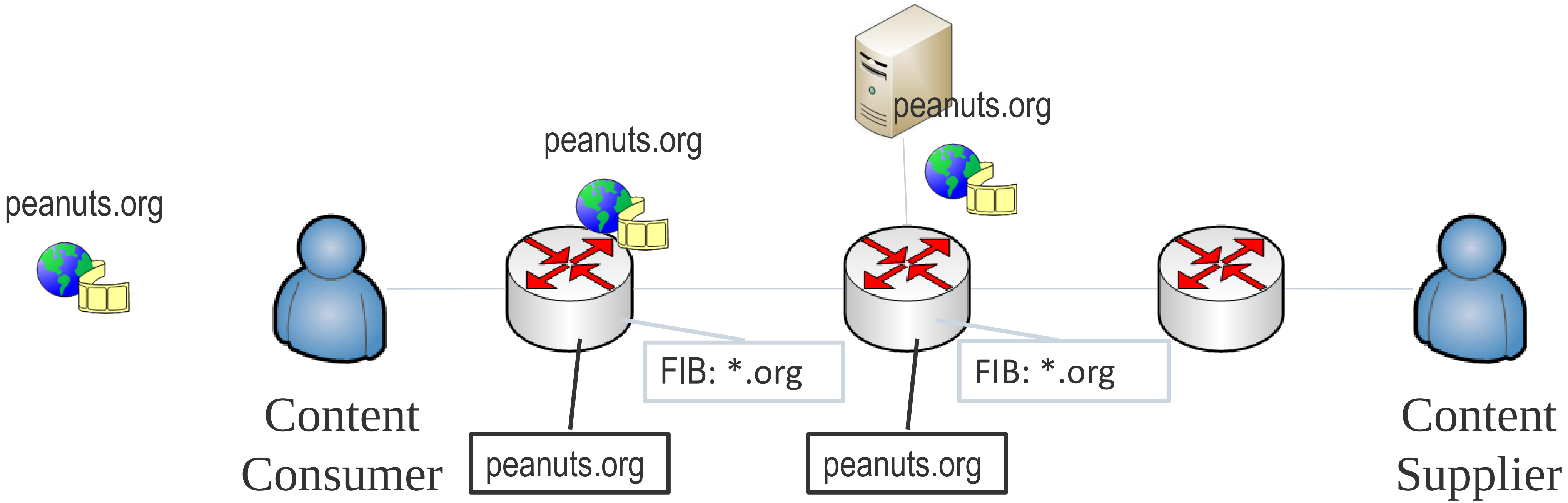
Basic Content Centric Routing and Forwarding



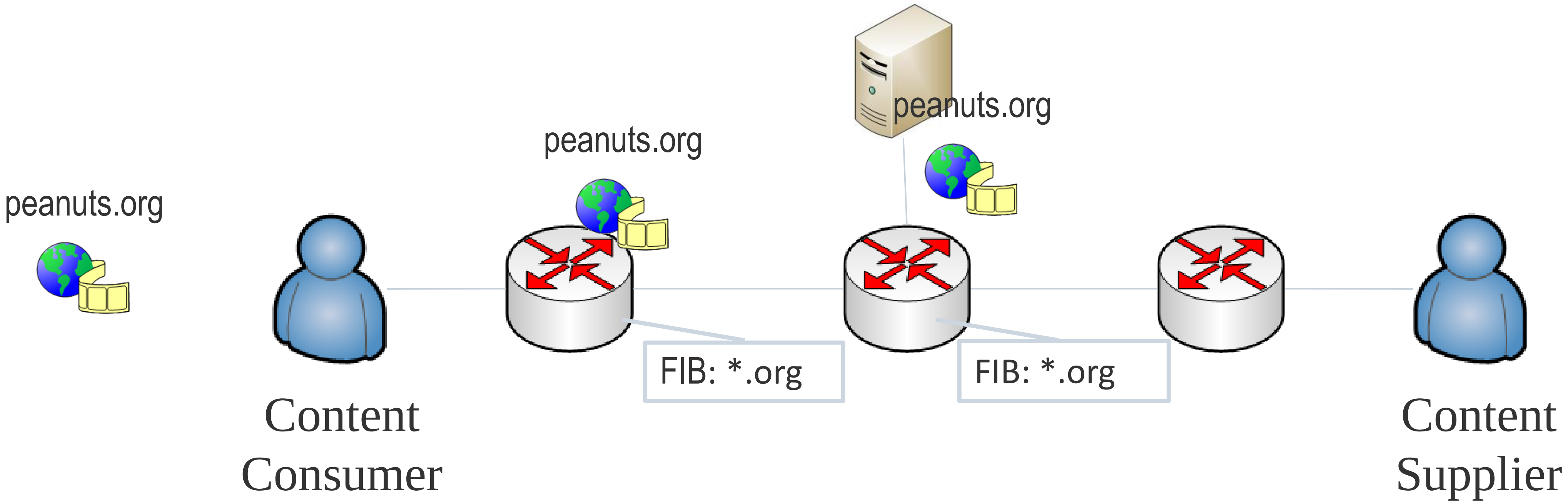
Basic Content Centric Routing and Forwarding



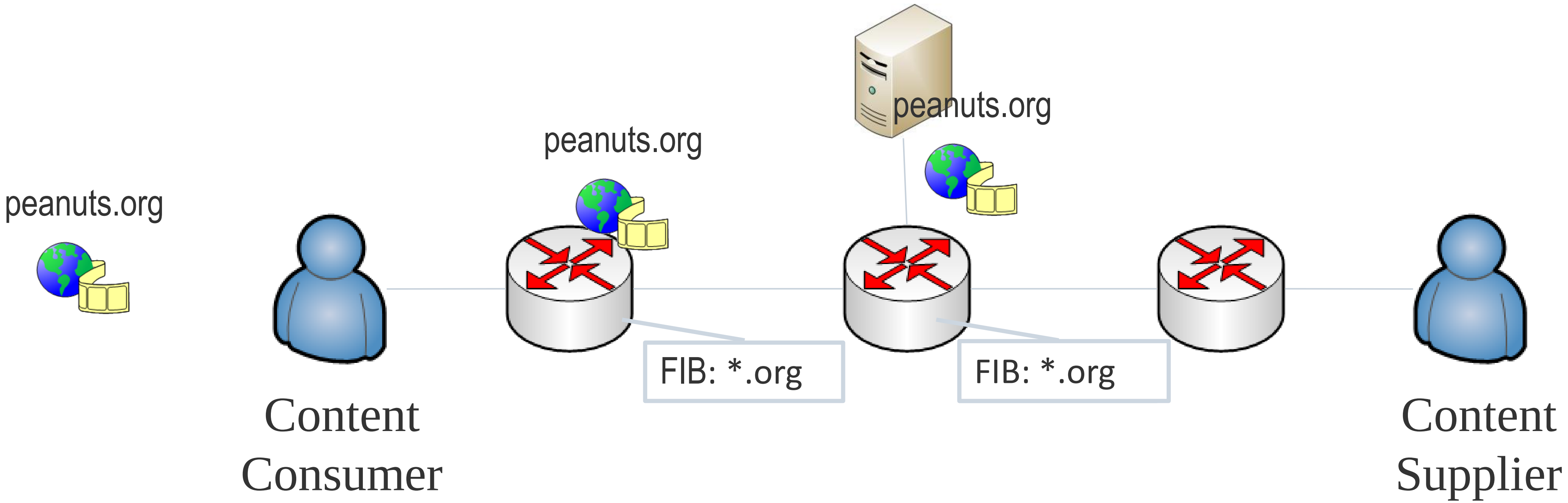
Basic Content Centric Routing and Forwarding



Basic Content Centric Routing and Forwarding



Basic Content Centric Routing and Forwarding



Observation 1: In-network states driven by data

Observation 2: End-users affect backbone states

NDN Stateful Routing and Forwarding

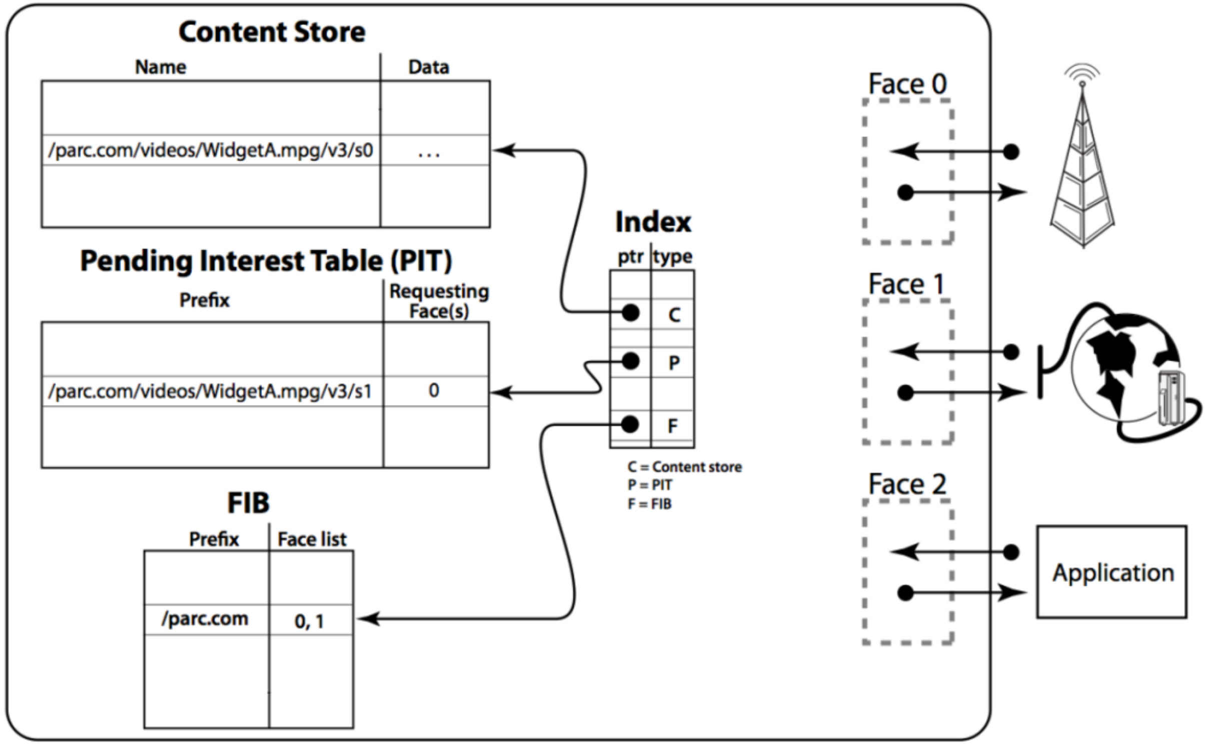
Details on state management:

Each router holds

- Forwarding states (FIB)
- Pending Interest Table (PIT)
- In-network storage

States describe data chunks

- Updates at wire-speed



The Problem of State

Two kinds of states:

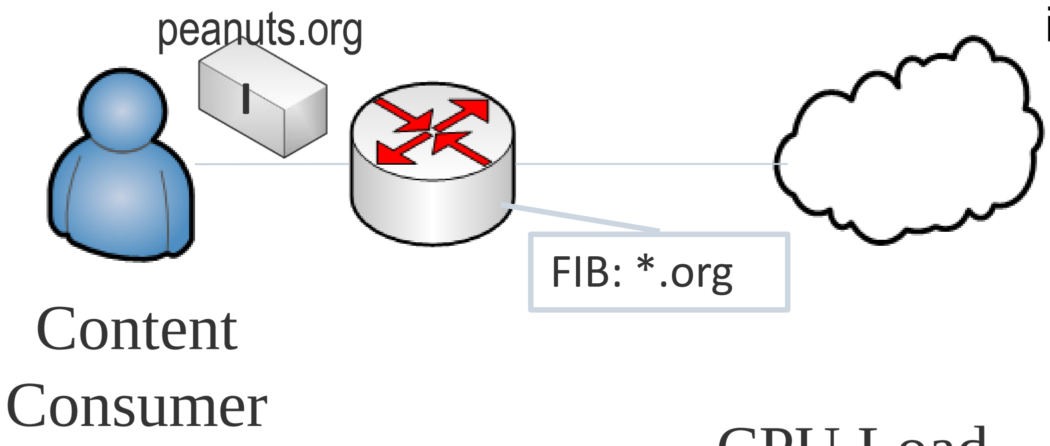
1. Content publication (and caching)
   → routing
2. Content request trails (breadcrumbs)
   → forwarding

Both kinds are ‘content-aware’:

Control states are open to user activities

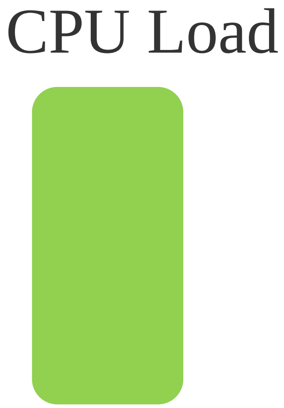
State management relies on data-driven events

Threat: Resource Exhaustion

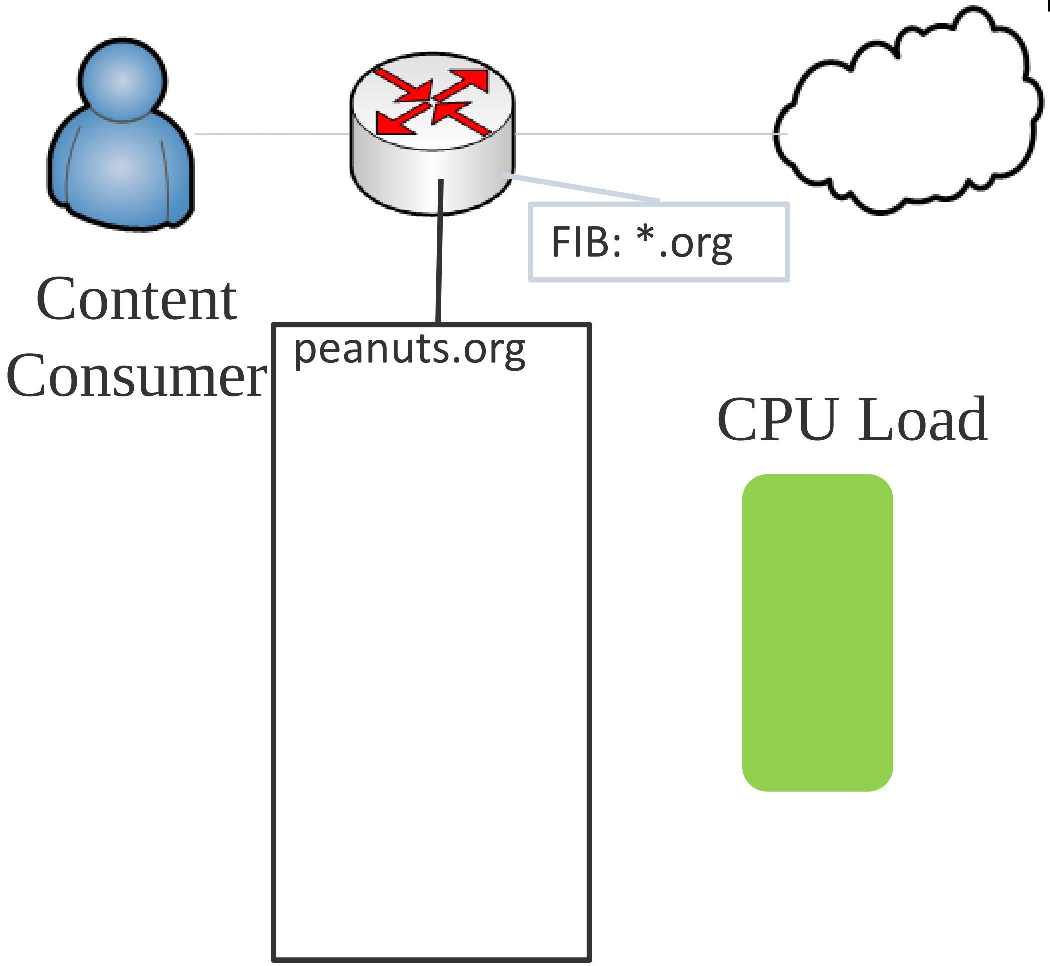


Consumer initiates many interests

- Content need not exist



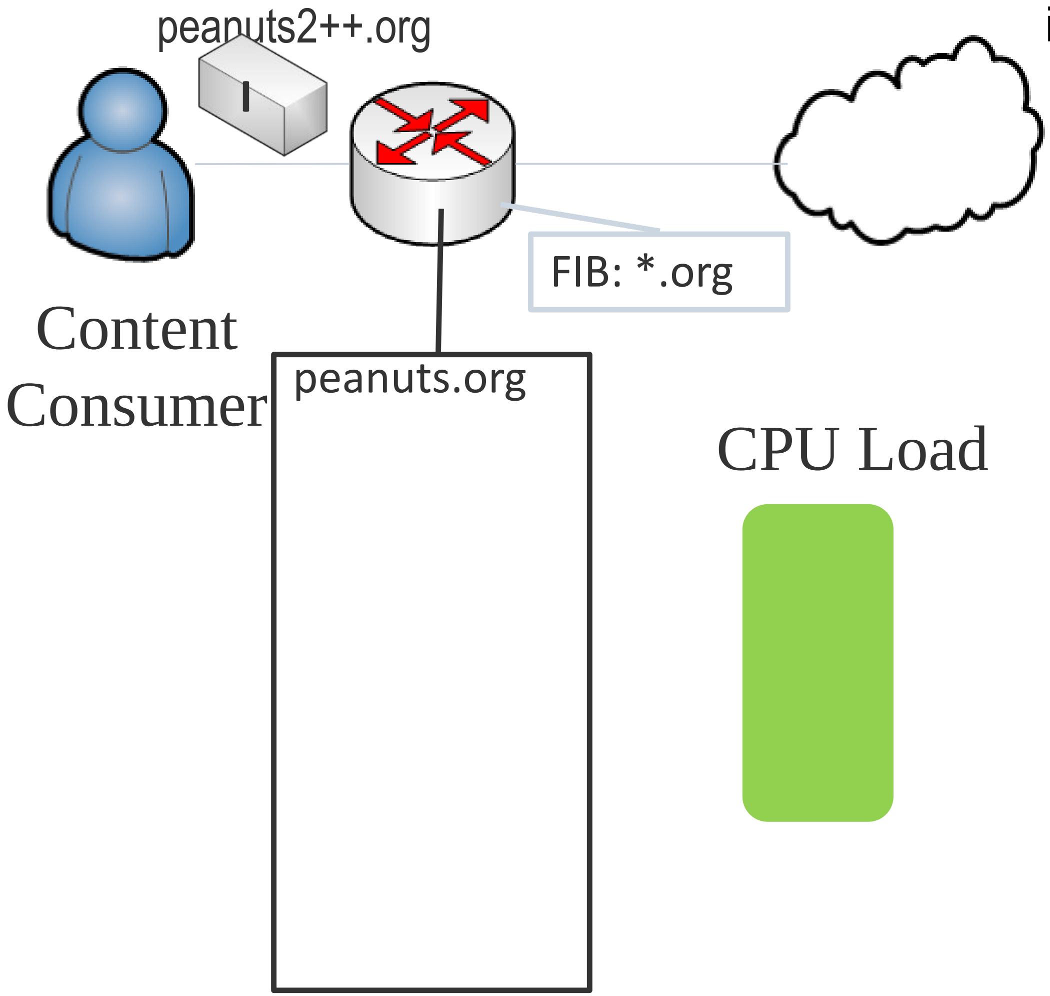
Threat: Resource Exhaustion



Consumer initiates many interests

- Content need not exist

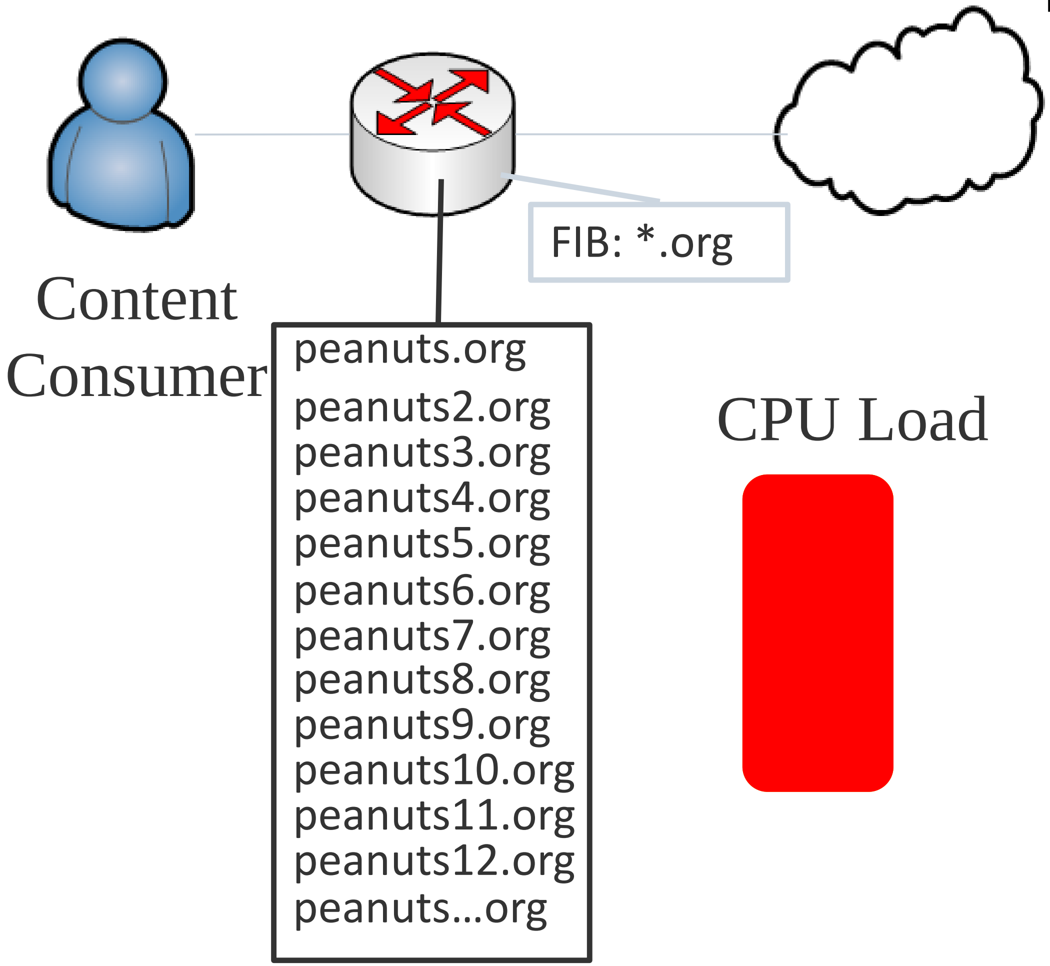
Threat: Resource Exhaustion



Consumer initiates many interests

- Content need not exist

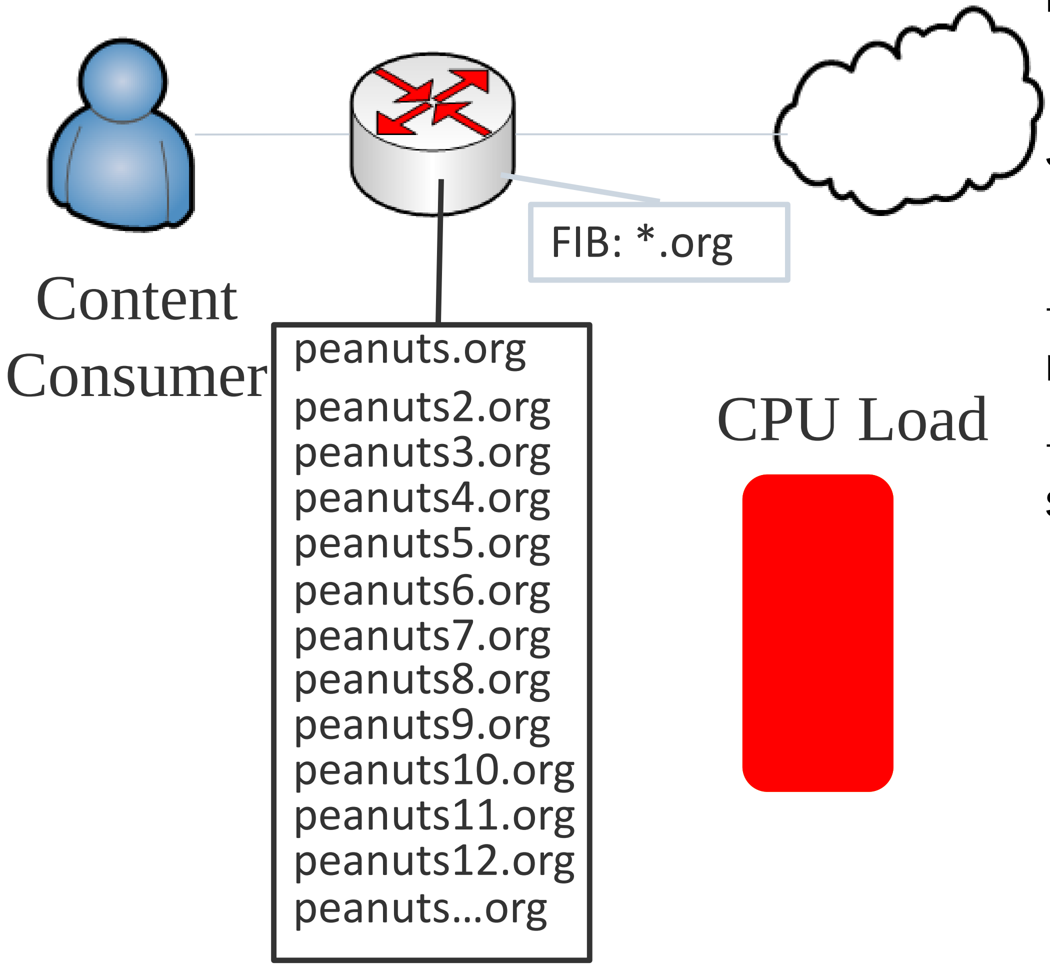
Threat: Resource Exhaustion



Consumer initiates many interests

- Content need not exist

Threat: Resource Exhaustion



Consumer initiates many interests

- Content need not exist

Supplier uploads content

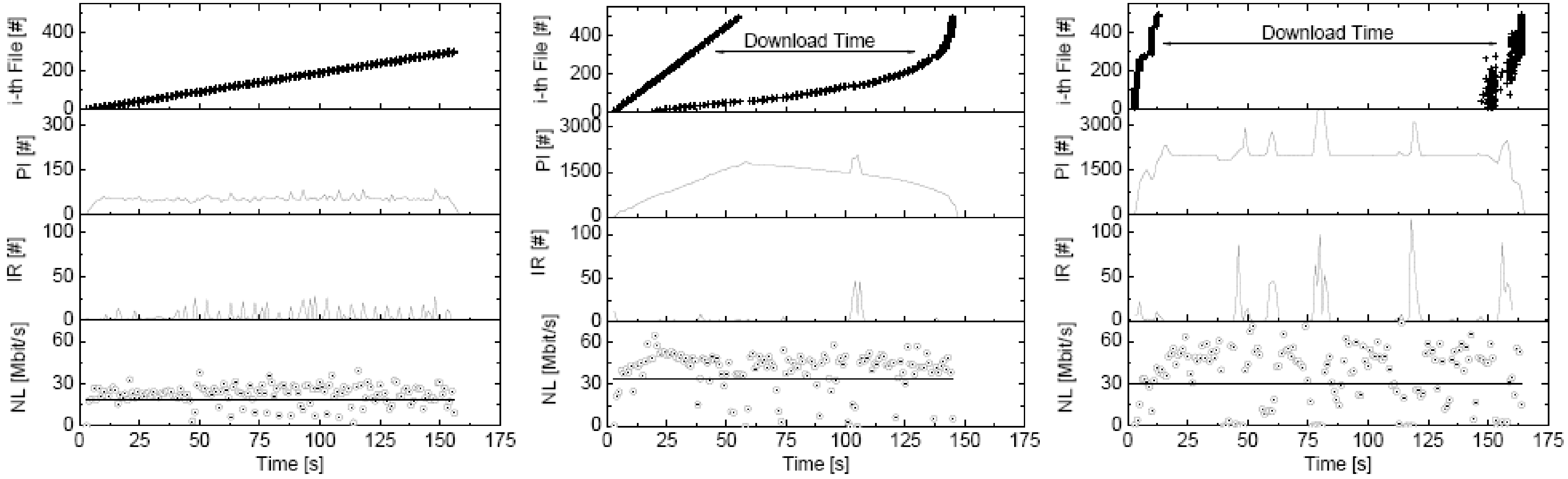
- Might be micro-content

→ Do this for an ‘unlimited’ number of items

→ Affects routing or content states

Data-driven States in Praxis

Example: Experimental Analysis for CCNx



(a) 2 files per second (b) 10 files per second (c) 100 files per second

Bulk of Interest: Performance Measurement of Content-Centric Routing,
In: *Proc. of ACM SIGCOMM Poster, 2012*

Reverse Path Forwarding States (PIT)

Content request rate

$$S_i(t) \approx \langle \alpha_i \rangle \cdot (\langle RTT \rangle + \kappa\, \sigma(RTT))$$

Reverse Path Forwarding States (PIT)

Content request rate → $\langle \alpha_i \rangle$

$$
\begin{aligned}
S_i(t) &\approx \langle \alpha_i \rangle \cdot (\langle RTT \rangle + \kappa\, \sigma(RTT)) \\
&\approx U_i(t)/\langle l \rangle\, (\langle RTT \rangle + \kappa\, \sigma(RTT))
\end{aligned}
$$

Link utilization → $U_i(t)$

Reverse Path Forwarding States (PIT)

Content request rate

$$S_i(t) \approx \langle \alpha_i \rangle \cdot (\langle RTT \rangle + \kappa\, \sigma(RTT))$$

$$\approx U_i(t)/\langle l \rangle\, (\langle RTT \rangle + \kappa\, \sigma(RTT))$$

Link utilization

This can be very bad

Reverse Path Forwarding States (PIT)

Content request rate

$$S_i(t) \approx \langle \alpha_i \rangle \cdot (\langle RTT \rangle + \kappa\, \sigma(RTT))$$

$$\approx U_i(t) / \langle l \rangle \, (\langle RTT \rangle + \kappa\, \sigma(RTT))$$

Link utilization

This can be very bad

⇒ State requirements are proportional network utilization +

⇒ Enhanced by a factor of a global retransmission timeout

Implications

1. The RTT distribution covers Internet-wide traffic:
   A long-tailed Gamma law (unlike TCP that deals with dedicated endpoints)
2. Rapidly varying RTTs are characteristic for ICN interfaces and even for prefixes (multimodal delay distribution due to content replication)
3. Limits of PIT sizes, state timeout, and interest rates are hard to define well – and don’t protect routers without degrading network performance
4. Routing resources (memory, CPU) are required orders of magnitude higher than previously predicted
5. Inverts router design: Highest resources required at edge

Problems of Name-based Routing

Names are many more than active (IP-) Adresses

Names do not aggregate w.r.t. location

Name aggregation is not locally decidable

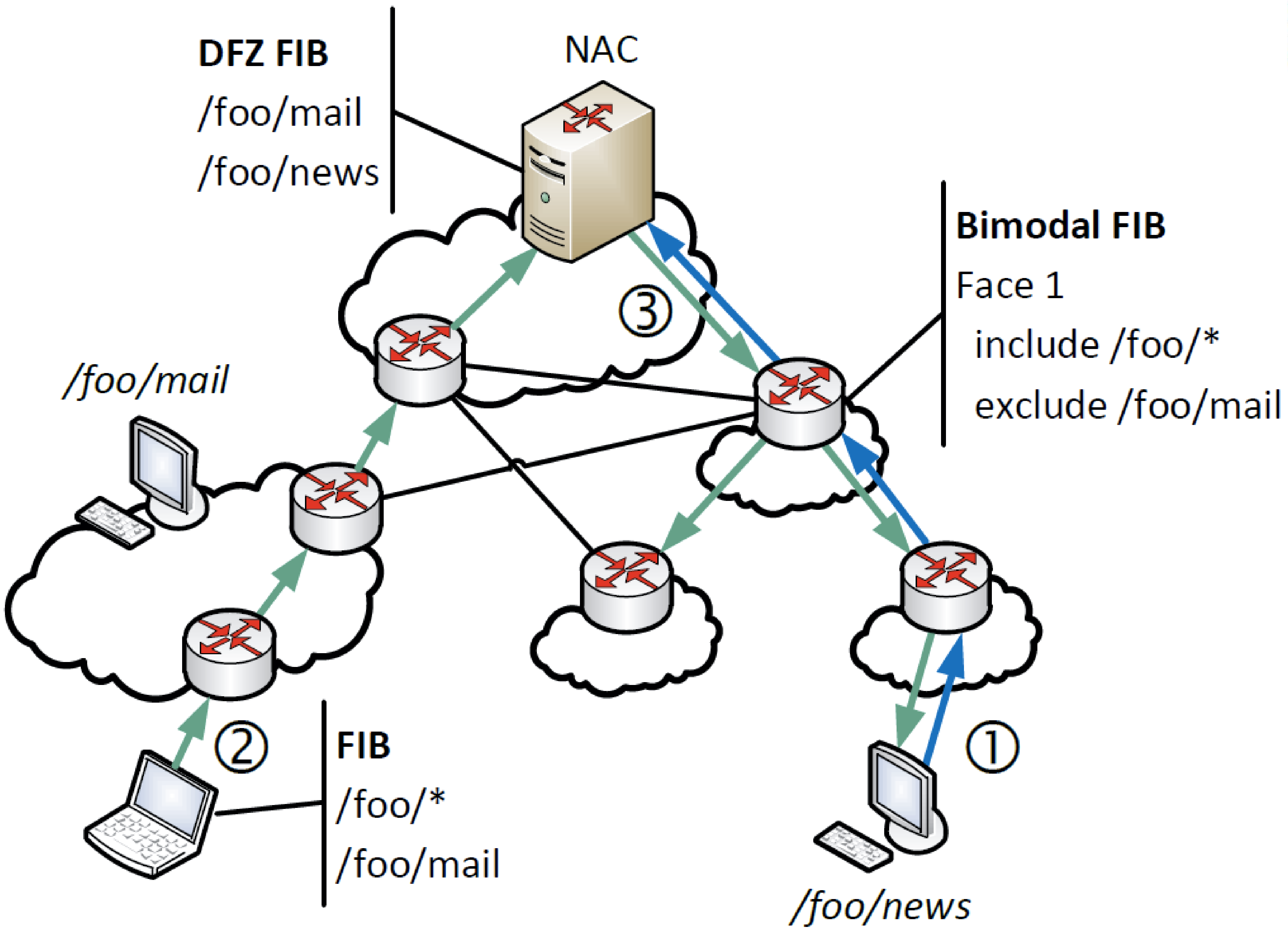
Name update frequency much higher than IP topology

PANINI: Partial Adaptive Name Information in ICN

Intra-domain routing protocol that limits FIBs

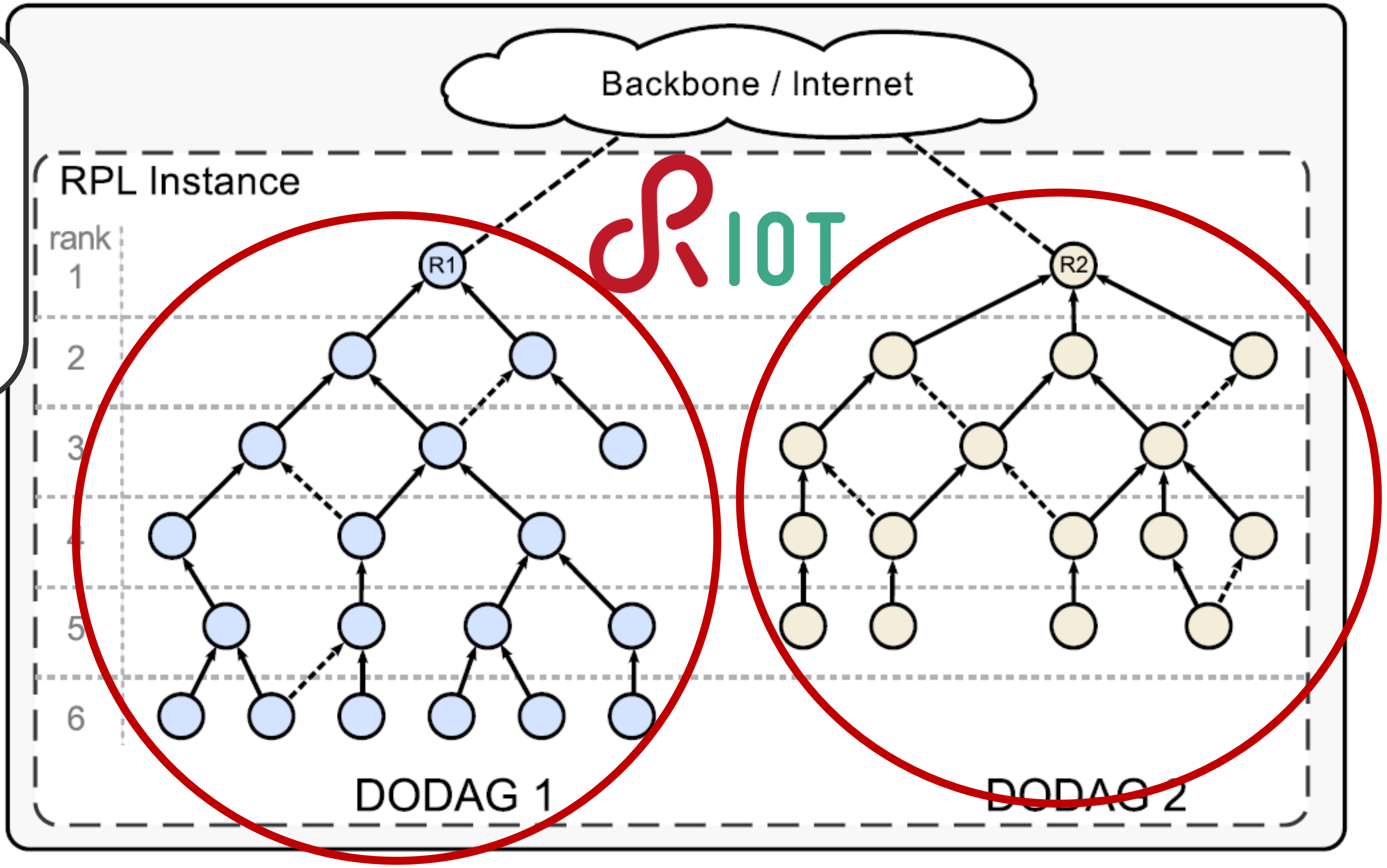
Key ideas

1. Name Collector (NAC): prefix-specific aggregation point
2. Default distribution tree: prefix-specific default routes
3. Adaptive FIB management: adjust to content popularity and local resources
4. Scoped flooding: on FIB miss only, limited to UR-subtrees


DFZ FIB
/foo/mail
/foo/news
NAC
③
Bimodal FIB
Face 1
include /foo/*
exclude /foo/mail
/foo/mail
②
FIB
/foo/*
/foo/mail
①
/foo/news

ICN in the IoT

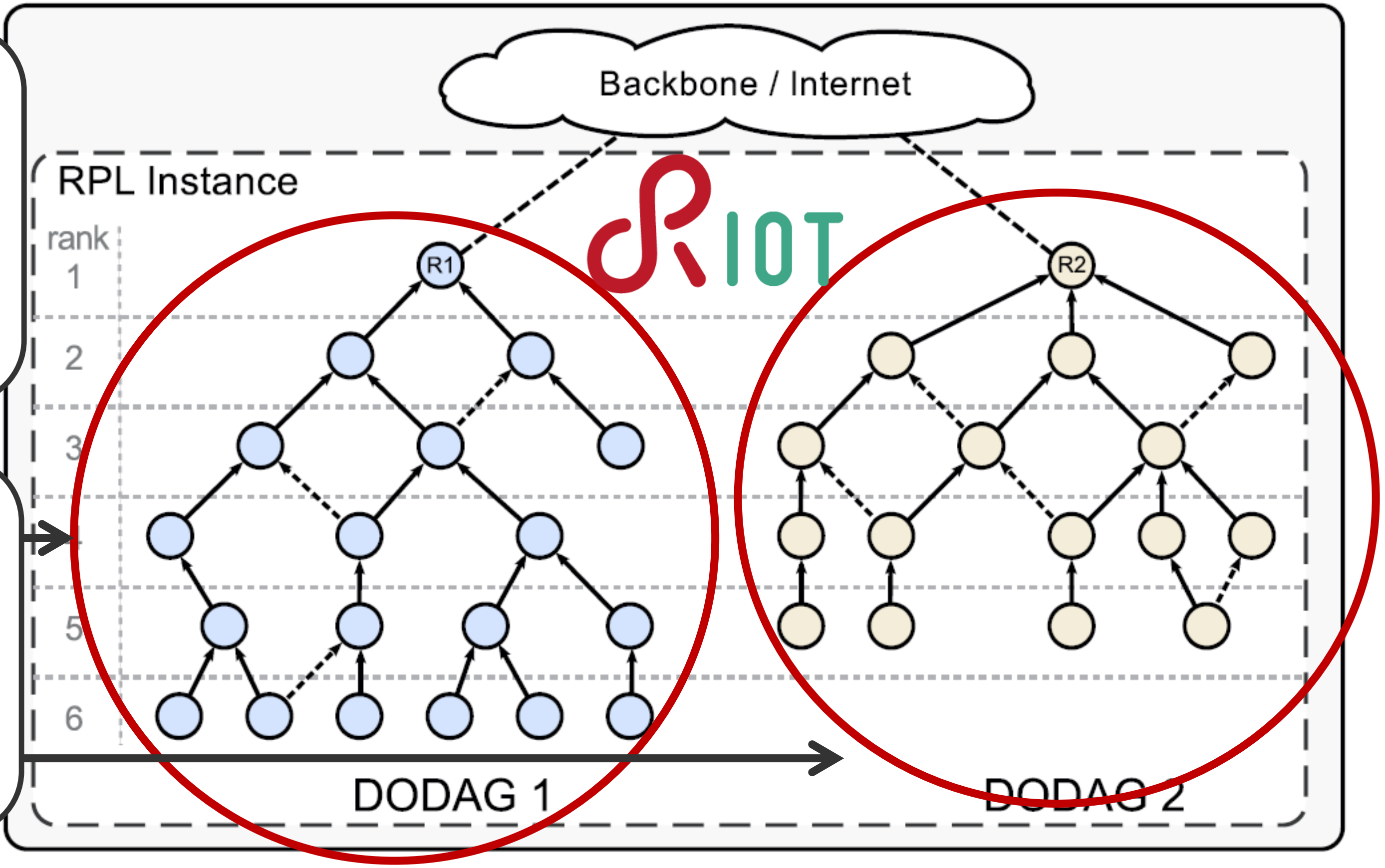
IPv6 Routing Protocol for Low-power and Lossy Networks (RPL)



ICN in the IoT

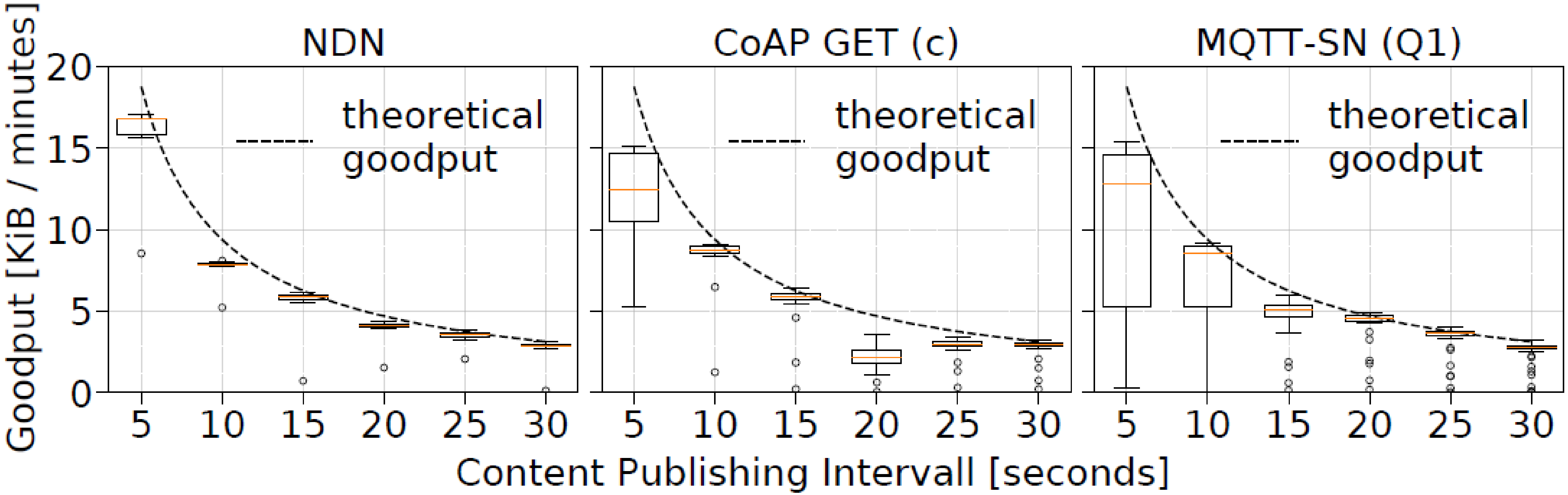
IPv6 Routing Protocol for Low-power and Lossy Networks (RPL)

Hopwise transfer to gateway with caching:
Robust & simple



Performance Comparison: Reliable IoT Protocols

Multihop Network of 50 Nodes:



Further Reading on Multicast

R. Wittmann, M. Zitterbart: *Multicast Communication,* Morgan Kaufmann, 2001

E. Rosenberg: *A Primer of Multicast Routing*, Springer 2012

www.rfc-editor.org

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.

Further Reading on ICN

Mark Gritter and David R. Cheriton: *An Architecture for Content Routing Support in the Internet*. In the *USENIX Symposium on Internet Technologies and Systems*, March 2001

Jacobson, V., Smetters, D., Thornton, J., and M. Plass, "Networking Named Content", 5th Int. Conf. on emerging Networking Experiments and Technologies (ACM CoNEXT), 2009.

B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman: *A Survey of Information-Centric Networking*, IEEE Communications Magazine • July 2012

M. Wählisch, T.C. Schmidt, M. Vahlenkamp: *Backscatter from the Data Plane - Threats to Stability and Security in Information-Centric Networking,* Computer Networks 2013,

T. C. Schmidt, Sebastian Wölke, Nora Berg, Matthias Wählisch: *Let's Collect Names: How PANINI Limits FIB Tables in Name Based Routing*, In: Proc. of 15th IFIP Networking Conference, p. 458–466, IEEE Press, May 2016.

C. Gündogan, P. Kietzmann, M. Lenders, H. Petersen, T. C. Schmidt, M. Wählisch, *NDN, CoAP, and MQTT: A Comparative Measurement Study in the IoT,* In: *Proc. of 5th ACM Conference on Information-Centric Networking (ICN),* pp. 159-171, ACM : September 2018.