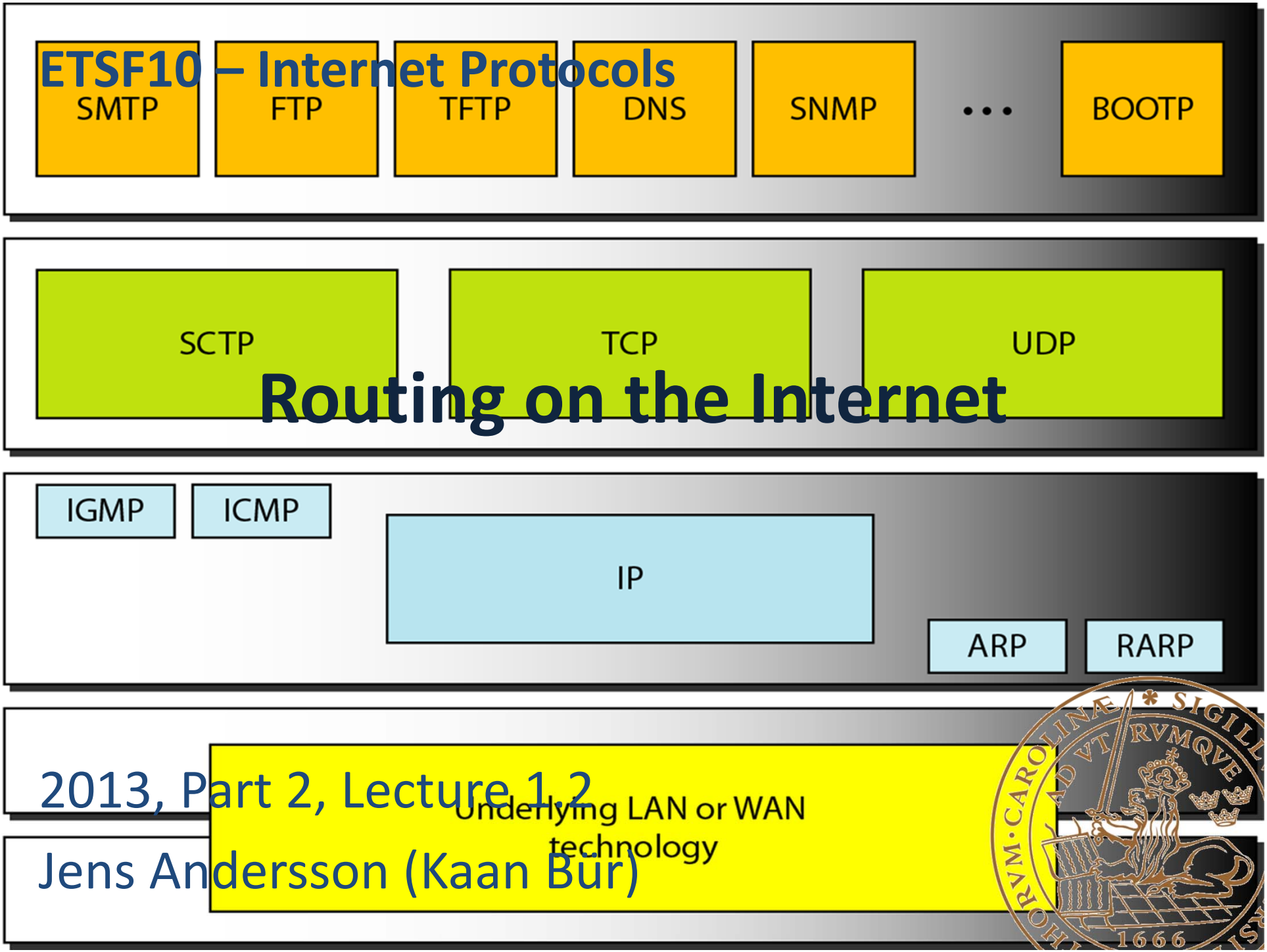
ETSF10 – Internet Protocols



Routing on the Internet

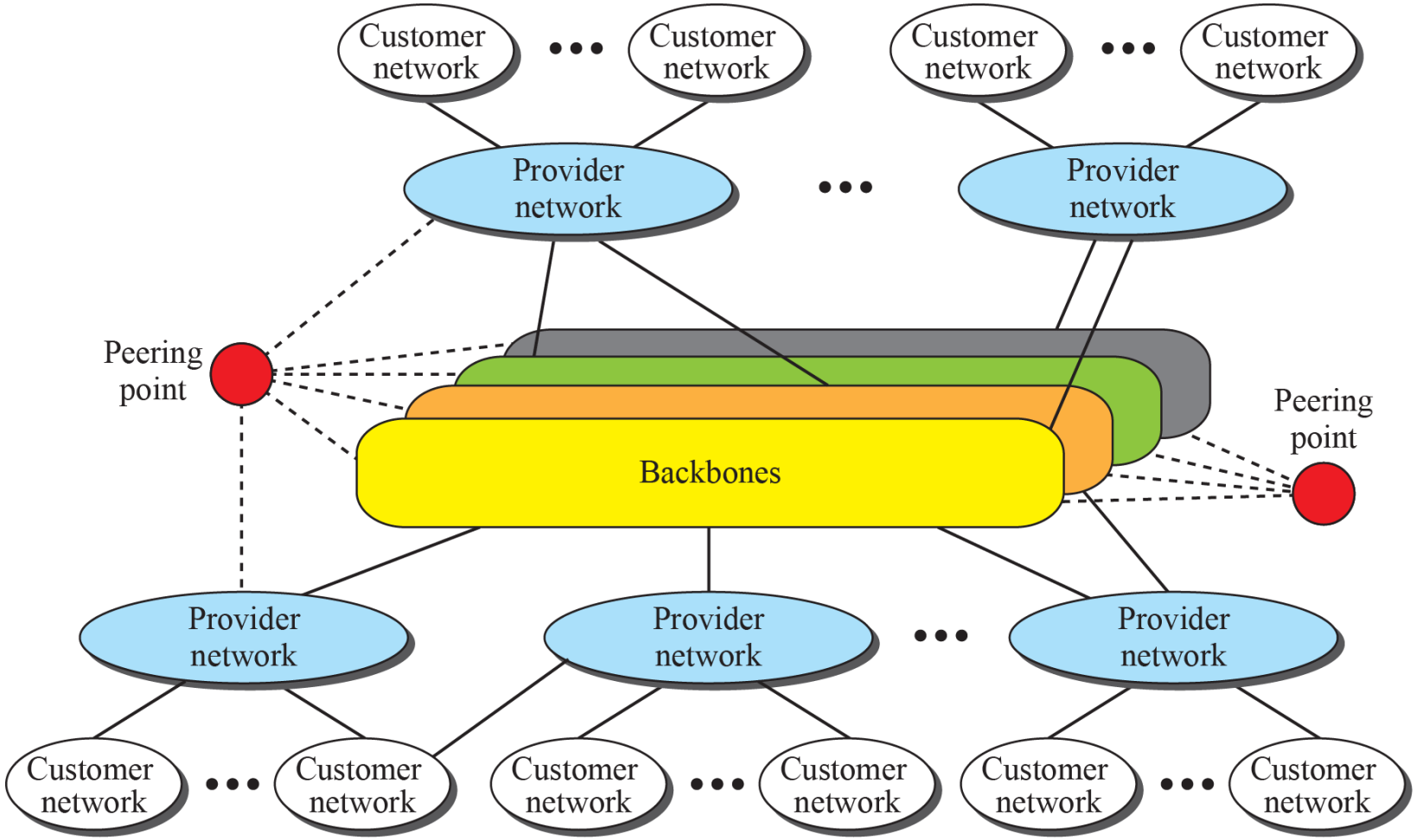
2013, Part 2, Lecture 1.2

Jens Andersson (Kaan Bür)

Routing on the Internet

- Unicast routing protocols (part 2)
  [ed.5 ch.20.3]
- Multicast routing, IGMP
  [ed.5 ch.21.1-5]
- NAT & Firewalls
  [ed.5 ch.18.4.5 & 32.4]

Internet Hierarchy



Hierarchical Routing

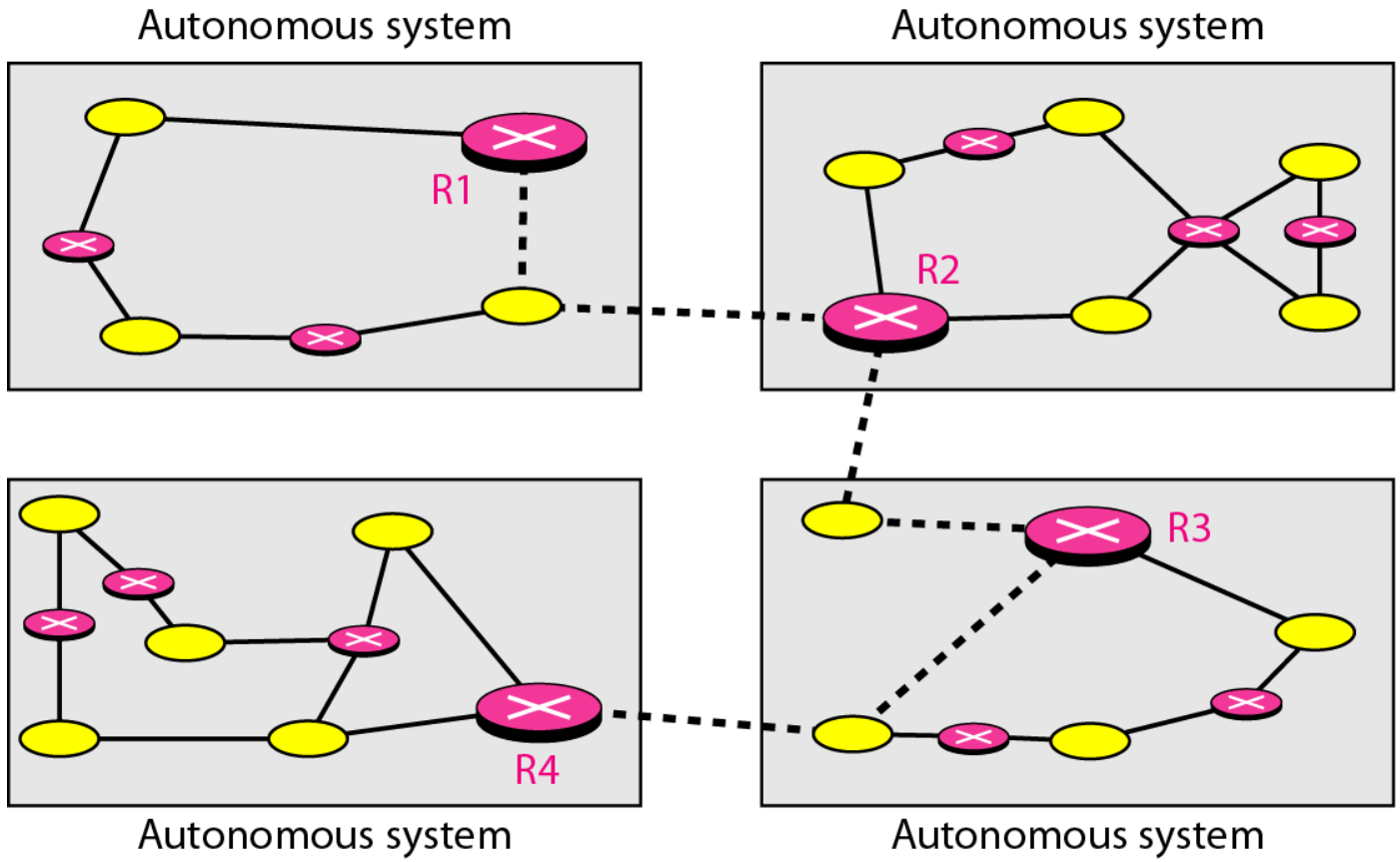
- aggregate routers into “autonomous systems”
- routers in same AS run same routing protocol
  - “intra-AS”
- routers in different AS can run different intra-AS routing protocol

Border Gateway Routers

- special routers in AS
  - run intra-AS routing protocol with all other routers in AS
- also responsible for routing to destinations outside AS
  - run inter-AS routing protocol with other gateway routers

Autonomous Systems

- Inter-AS border (exterior gateway) routers
  - R1
  - R2
  - R3
  - R4



Why different Intra- & Inter-AS routing?

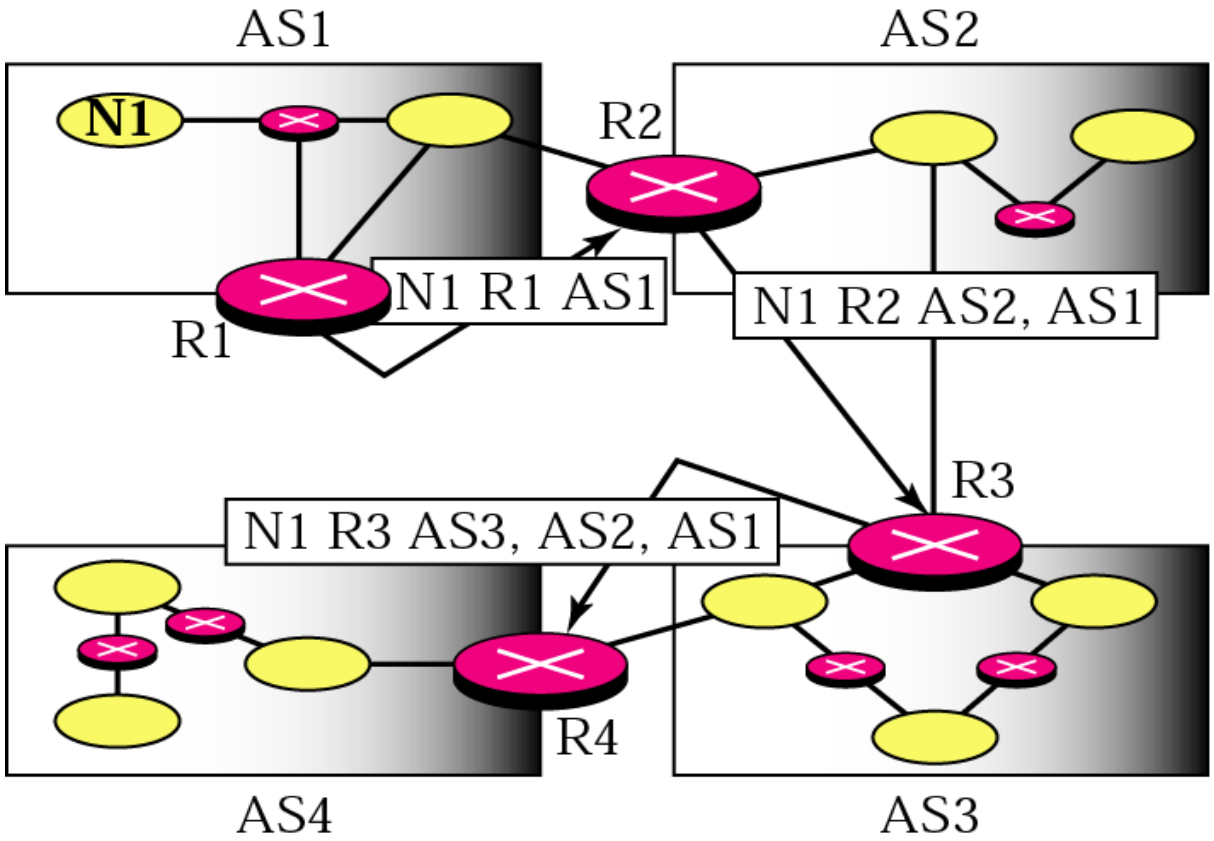
- Policy
  - Inter-AS: admin wants control over how its traffic routed, who routes through its net.
  - Intra-AS: single admin, so no policy decisions needed
- Scale
  - Hierarchical: saves table size, reduced update traffic
- Performance
  - Intra-AS: can focus on performance
  - Inter-AS: policy may dominate over performance

Internet Inter-AS routing: BGP

- Border Gateway Protocol: *de facto* standard
- Path Vector protocol:
  - Similar to *Distance Vector*
  - Border gateways broadcast to peers (not necessarily neighbours) entire path (sequence of AS) to destination
  - BGP routes to networks (AS), not individual hosts

Path Vector Messages

- Same principle as distance vector routing



Path Vector Routing Table

AS = Autonomous System = Organisation

Network	Next Router	Path
N01	R01	AS62, AS23, AS67
N02	R05	AS67, AS22, AS05, AS89
N03	R06	AS67, AS89, AS09, AS34
N03	R12	AS62, AS02, AS34

Network id

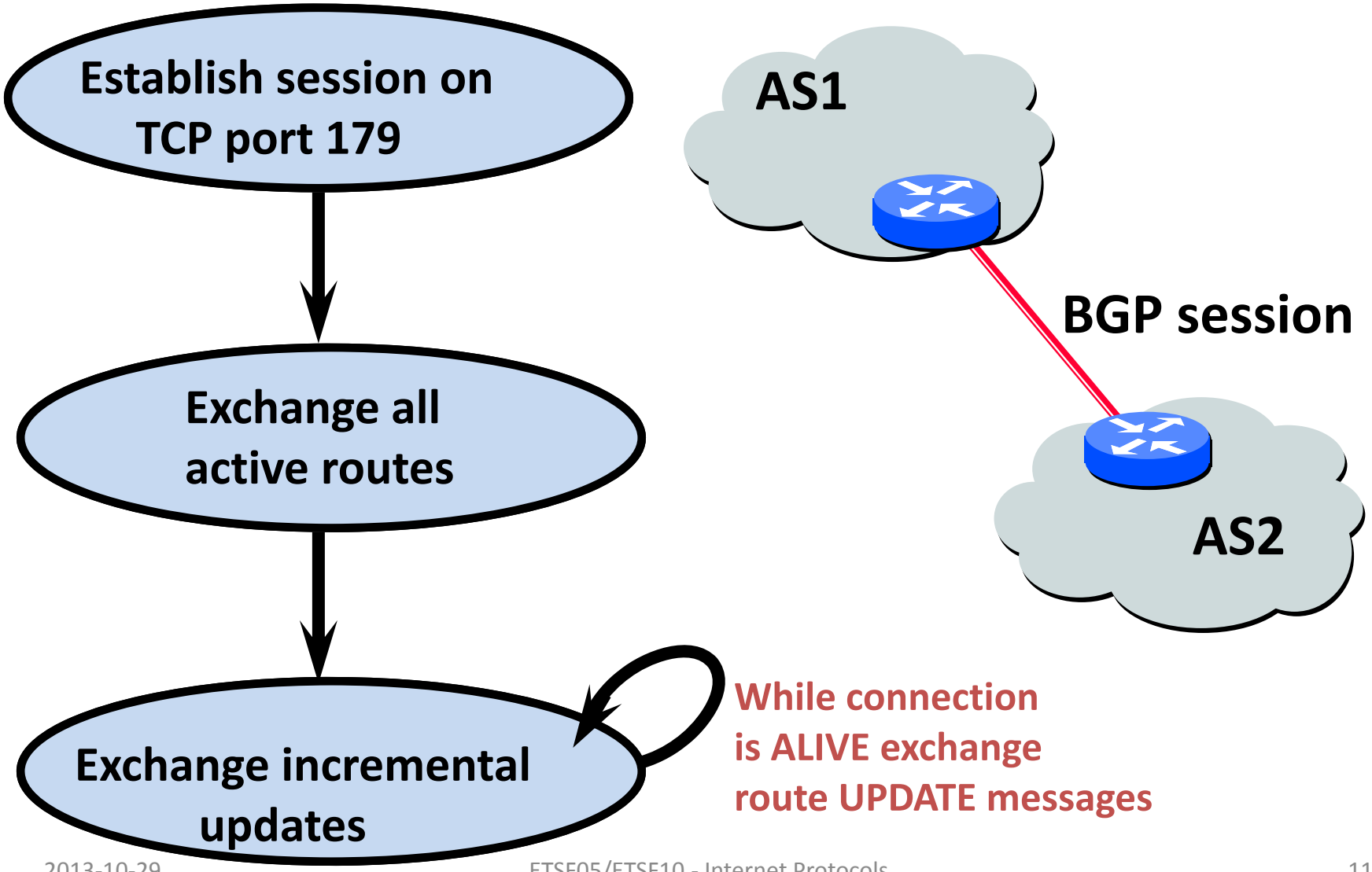
“next hop”

“Metric”
Most valid of many
ATTRIBUTES

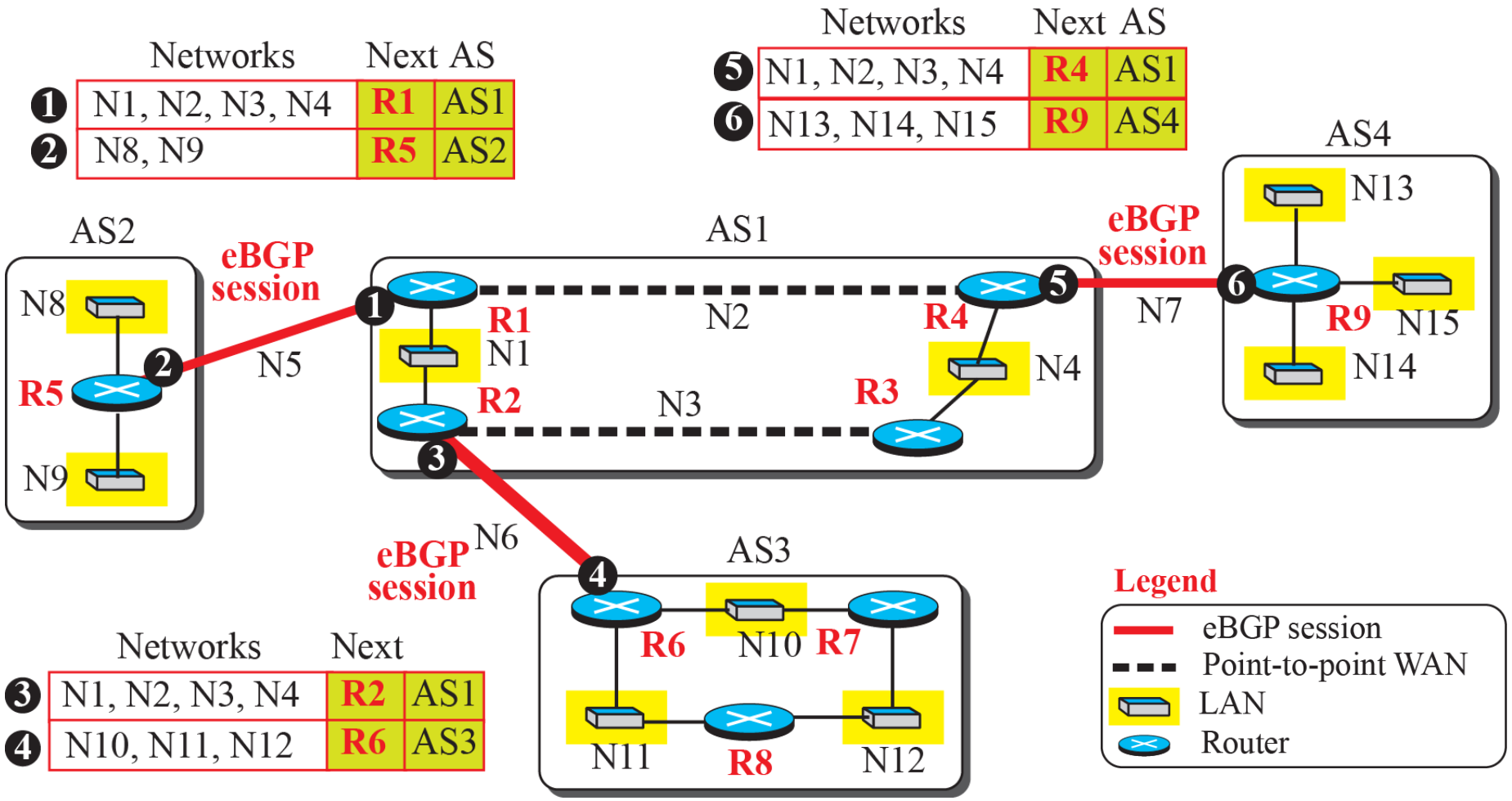
BGP Router Operations

- Receiving and filtering route advertisements from directly attached neighbour(s)
- Route selection
  - To route to destination X, which path (of several advertised) will be taken?
- Sending route advertisements to neighbours

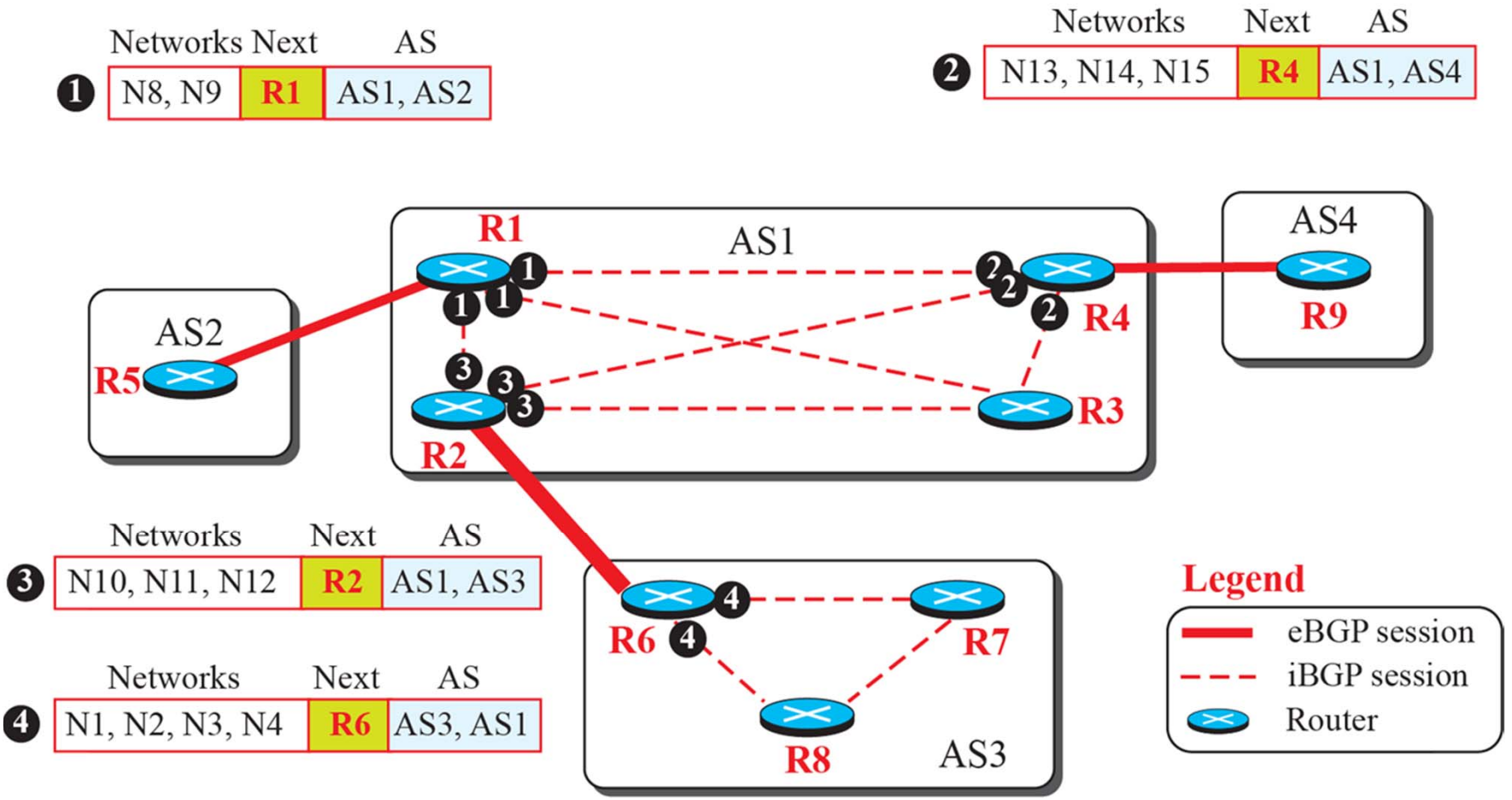
BGP Router Operations



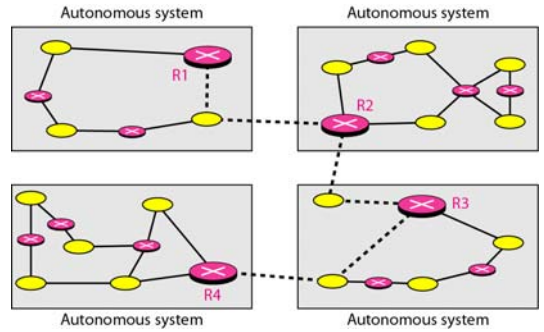
eBGP Operation

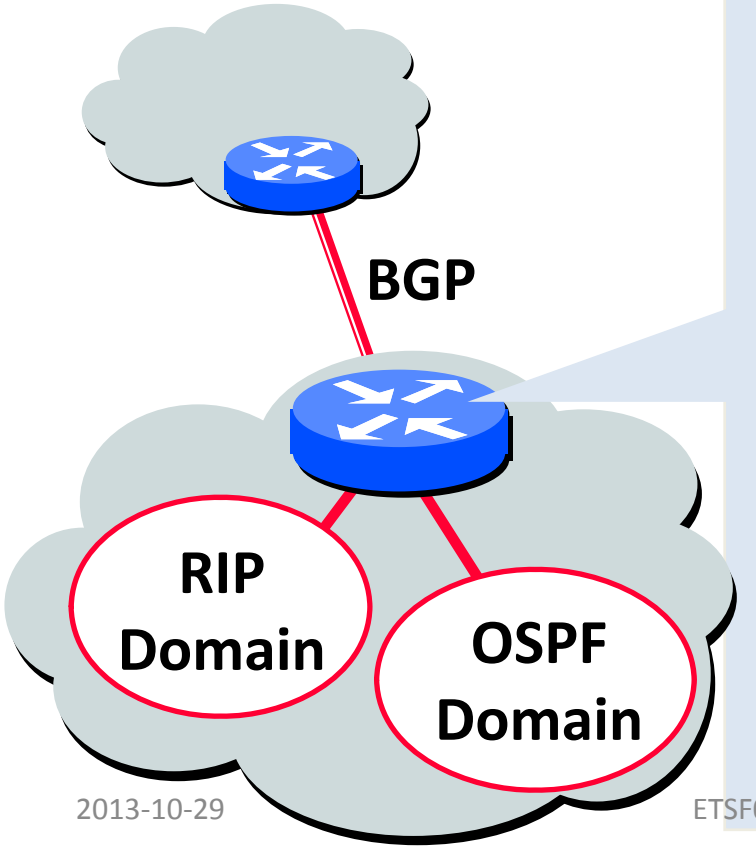


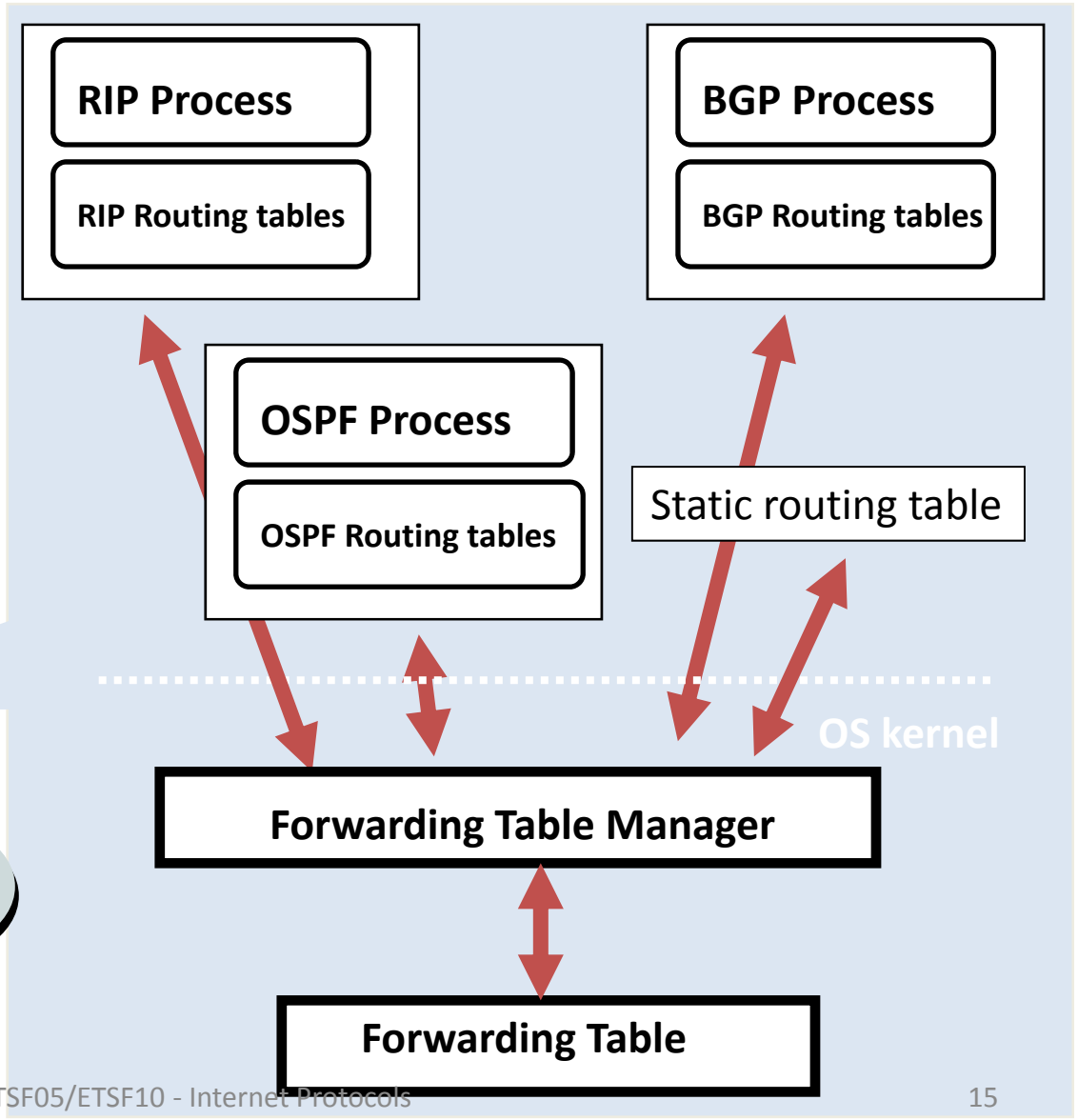
eBGP combined with iBGP



Routing Tables and Forwarding Table



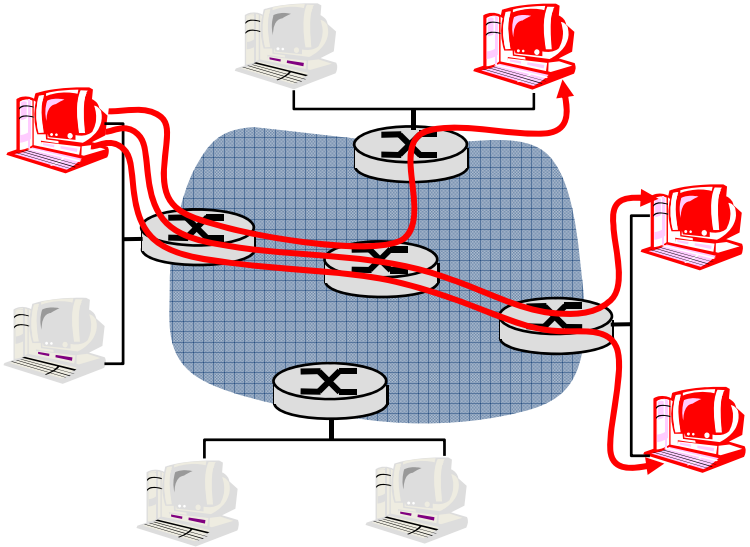




Multicast: One-to-many Routing

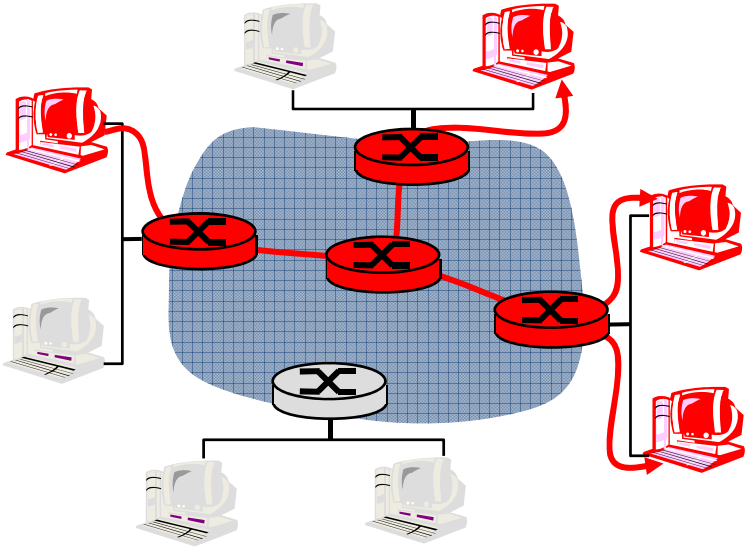
Unicast

- Routers forward multiple unicast datagrams

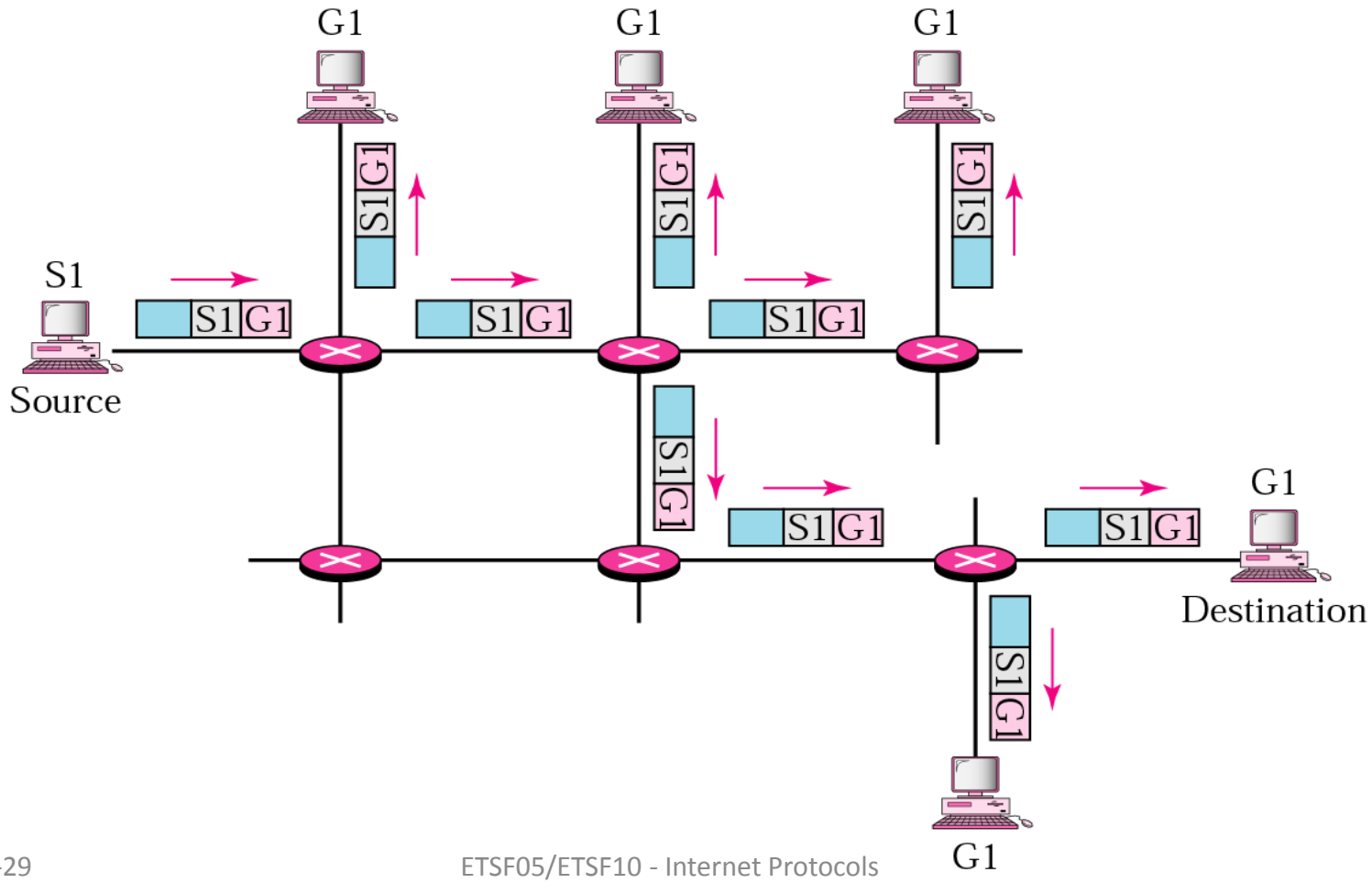


Multicast

- Routers (red) duplicate and forward multicast datagrams

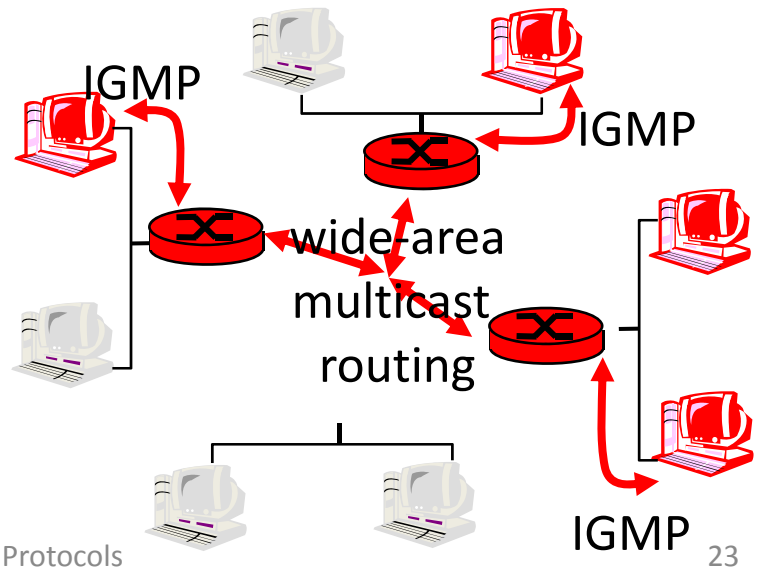


Source and Group Addresses



Joining a Multicast Group

- **Local**: host informs local multicast router
  - IGMP (Internet Group Management Protocol)
- **Wide area**: local router interacts with other routers to build forwarding tree and receive multicast data flow
  - MOSPF, DVMRP, PIM-DM
  - CBT, PIM-SM



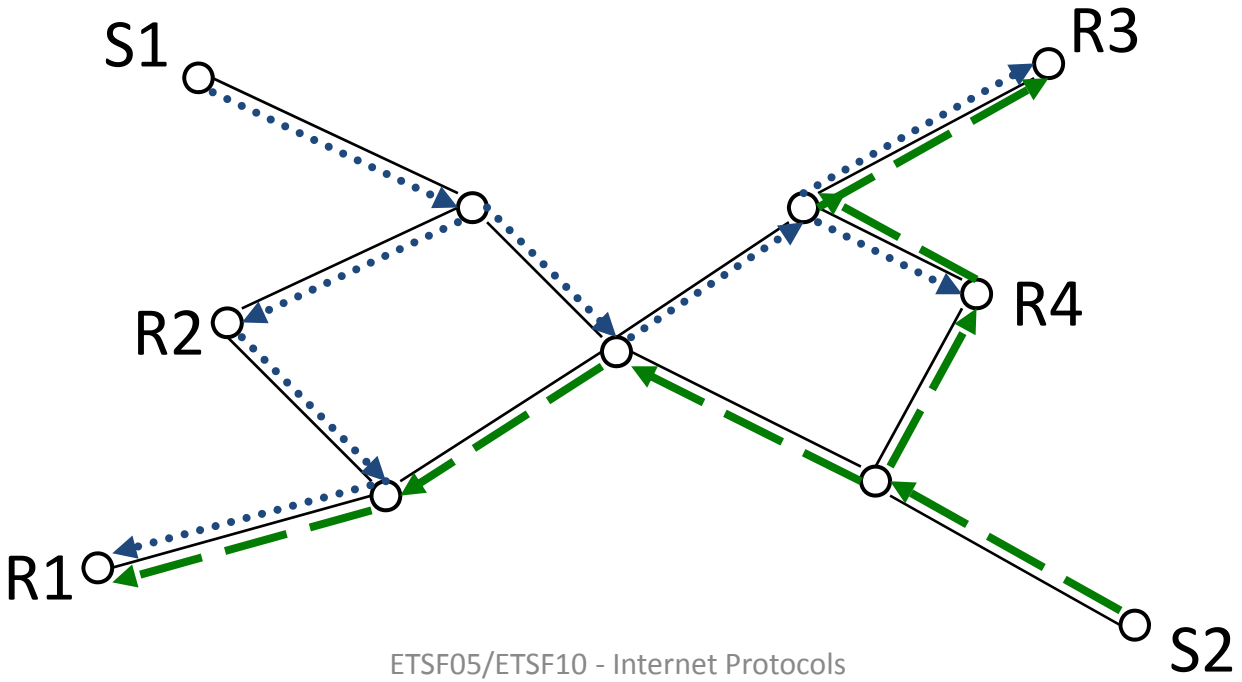
Multicast Routing Protocols

- Shortest path trees, again!

- In unicast routing
  - One path (on tree) used at a time
- In multicast routing
  - Whole tree used each time
  - Each source needs a tree

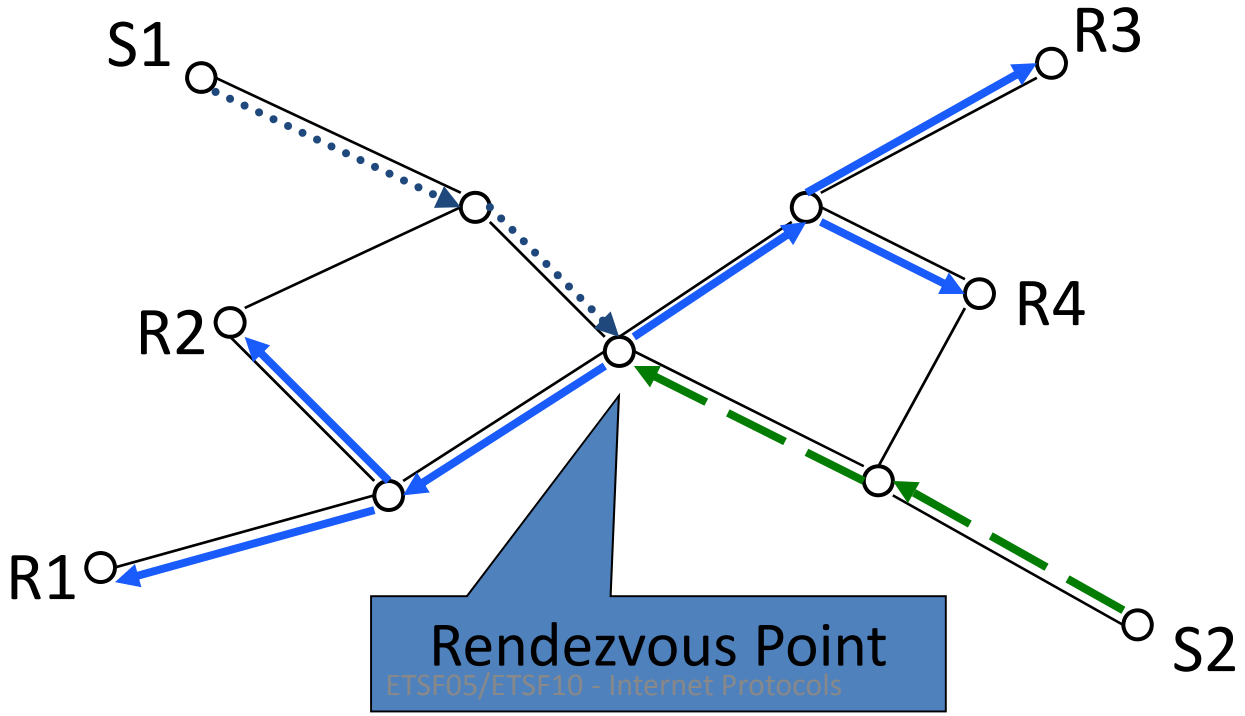
Source-Based Tree

- One tree per source (at each router)
- One source per group
- High complexity, high efficiency

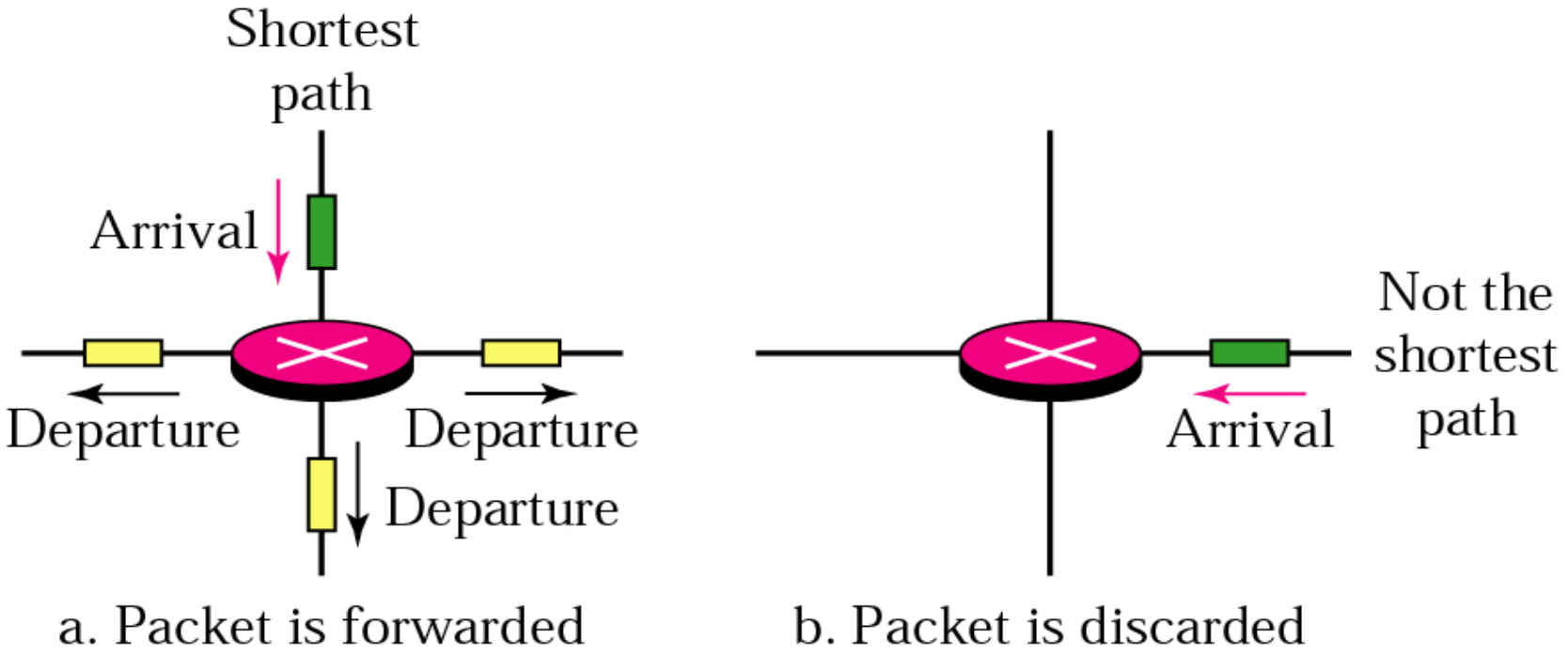


Group-Shared Tree

- One tree per group (at one router)
- Shared by multiple sources in group
- Lower complexity, lower efficiency

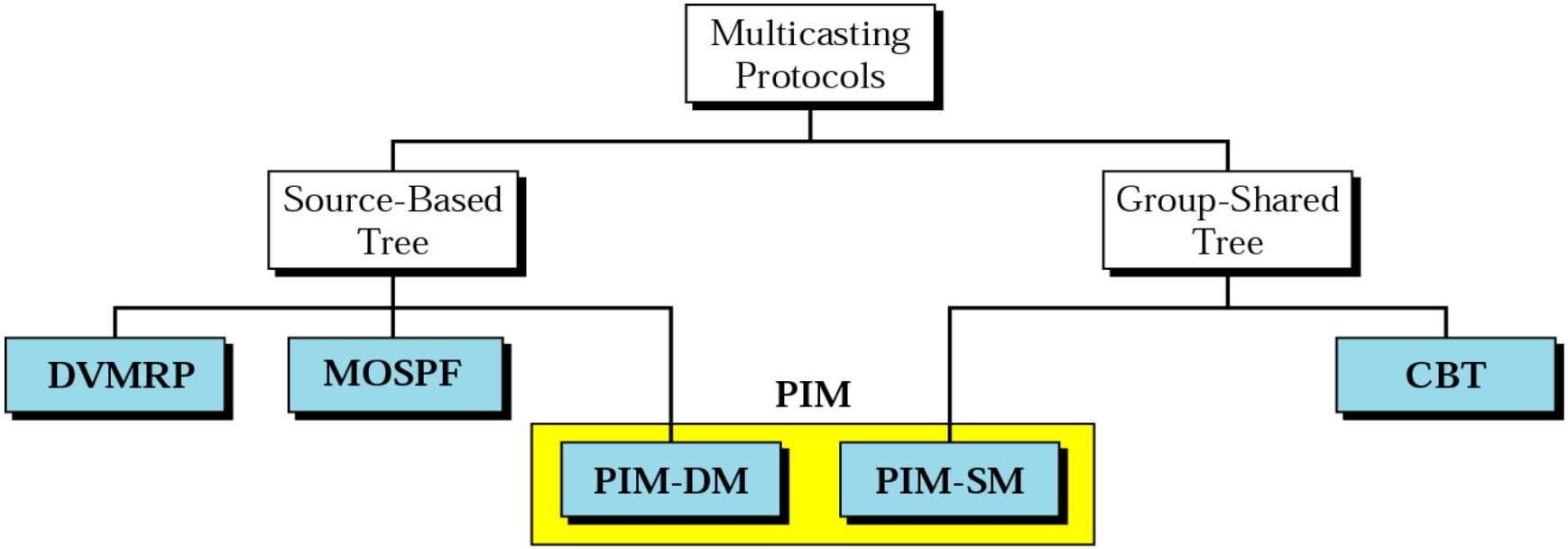


Reverse Path Forwarding



Source address routing!

Classification of Algorithms



PIM

- Independent from unicast protocol
- Uses available routing info for path lookups
- Two modes:
  - Sparse Mode
  - Dense Mode

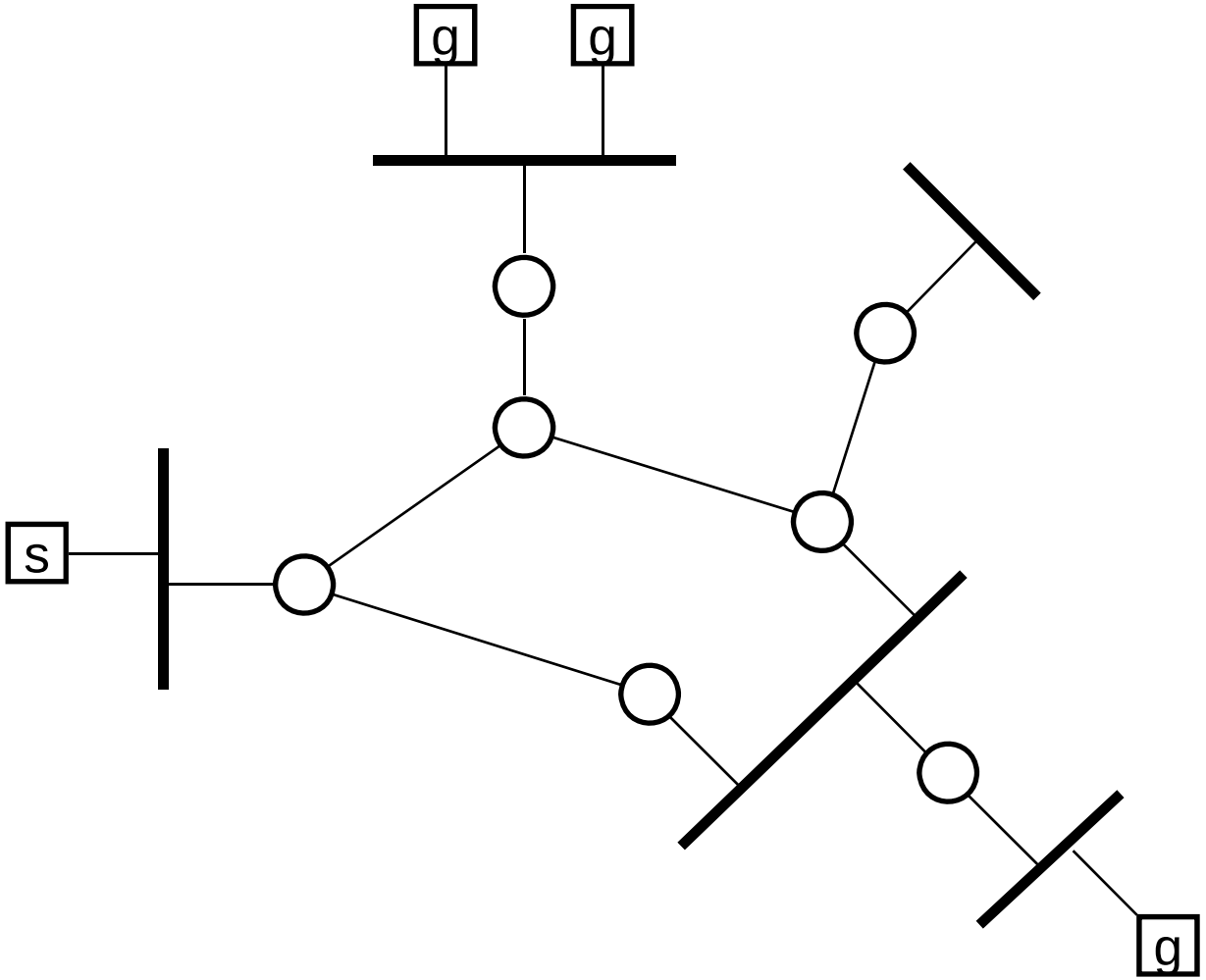
PIM-SM

- Relatively few members assumed
- Trees are built on demand (when needed)
  - Group-shared trees with rendezvous points
- Methods for tree construction
  - Grafting
  - Pruning
- Can switch from group-shared to source-based if more efficient

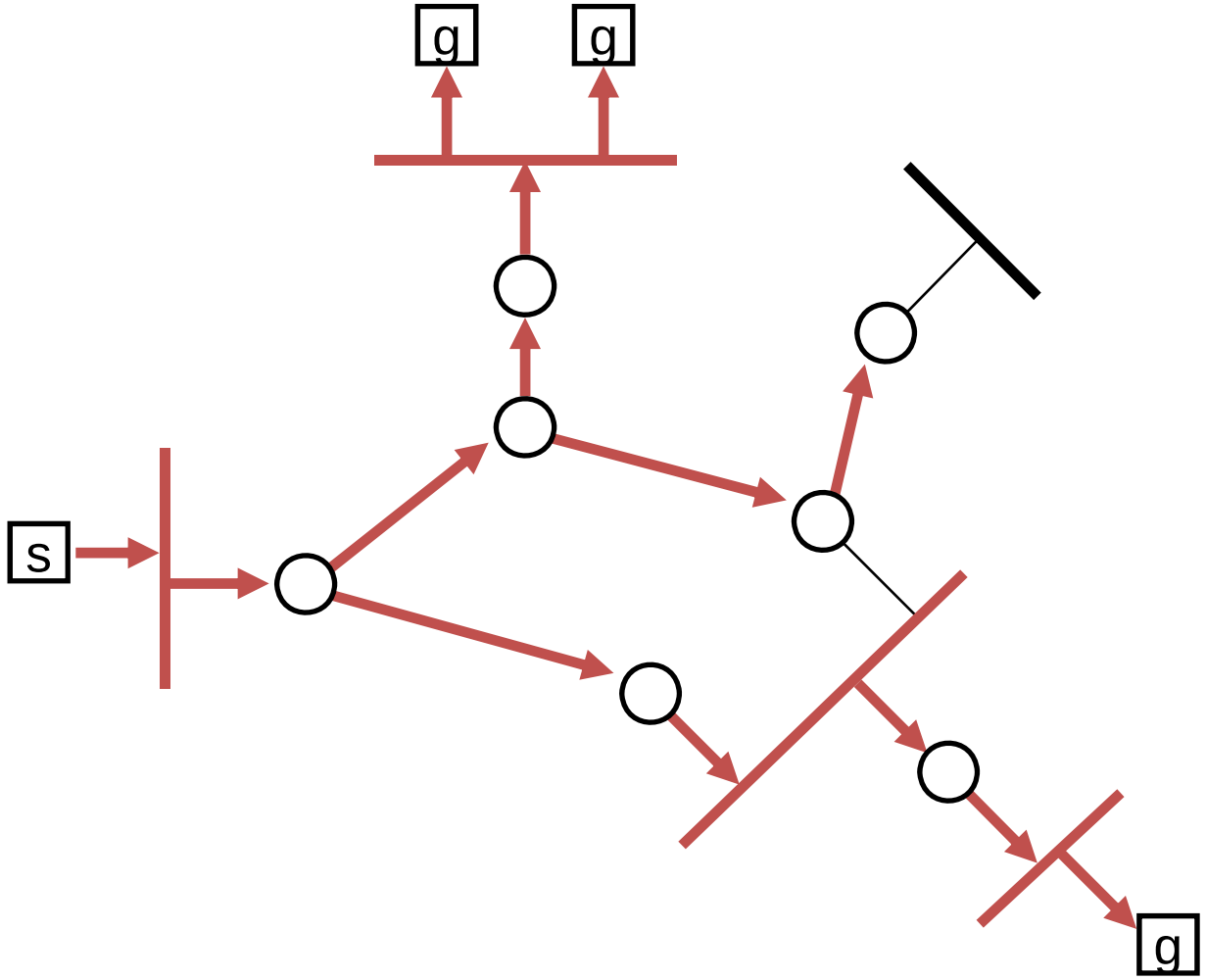
PIM-DM

- All hosts assumed to be members
- Build source-based tree from source
- Routers without members prune tree
- Grafting used to add new members

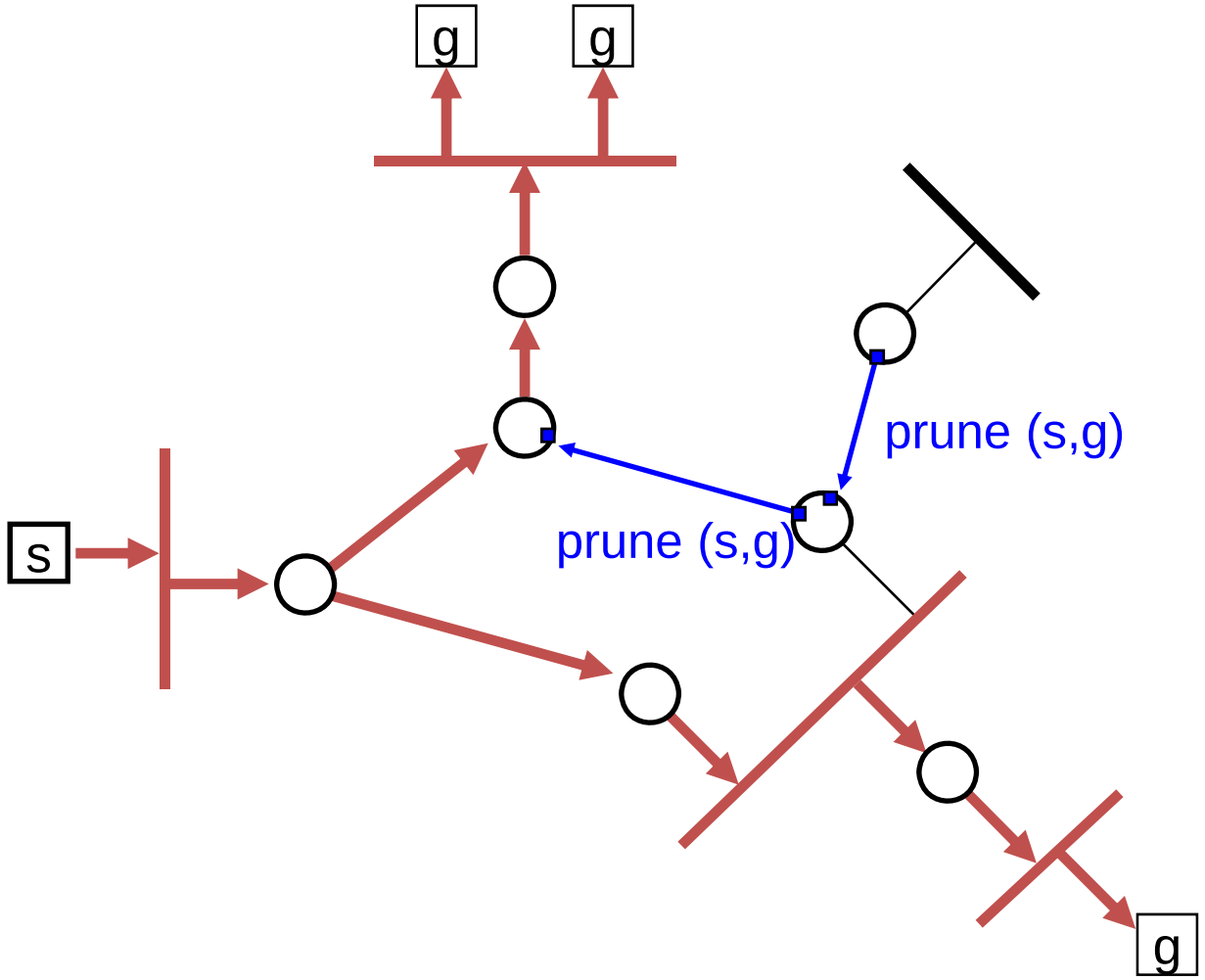
Example Topology



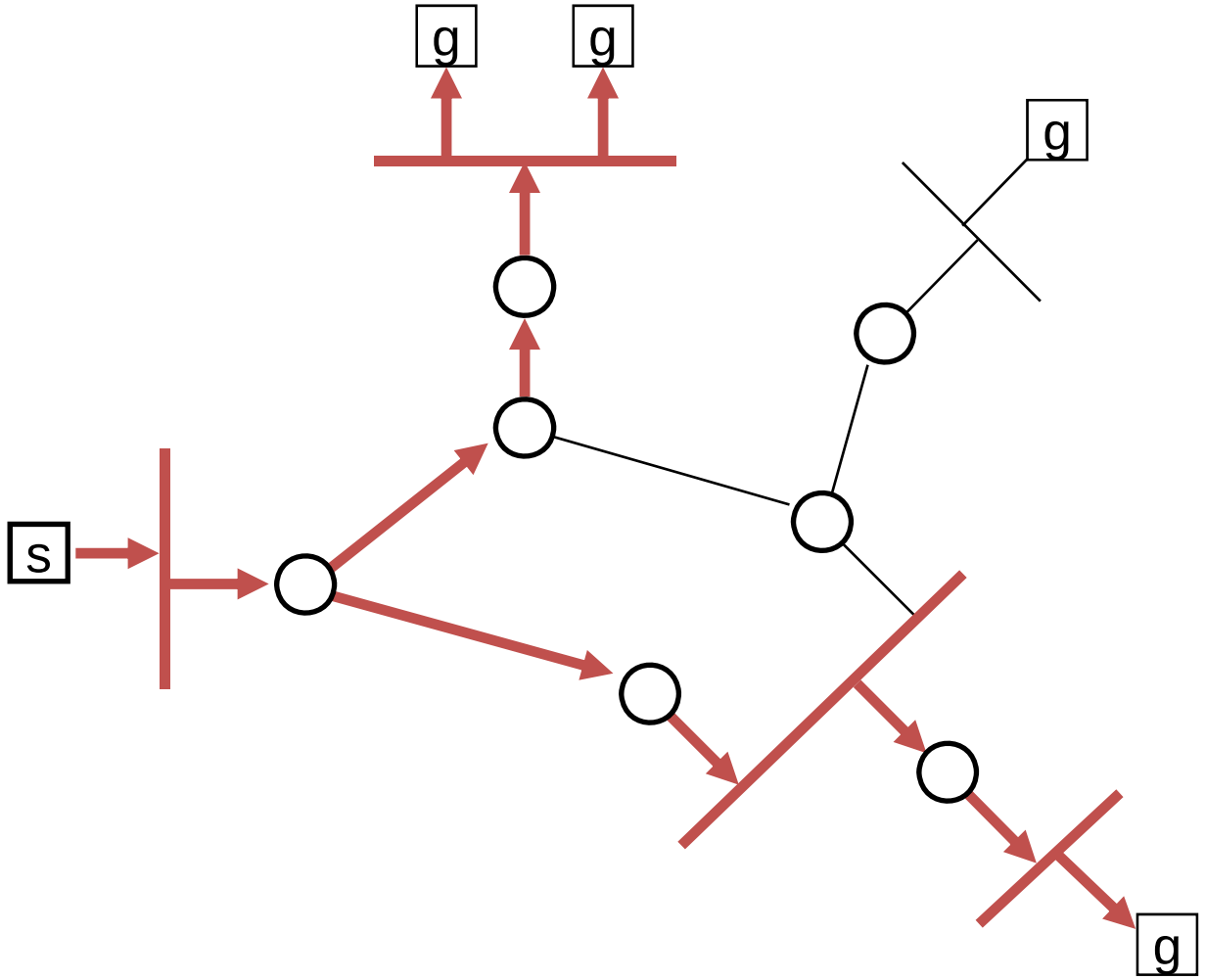
Truncated Broadcast



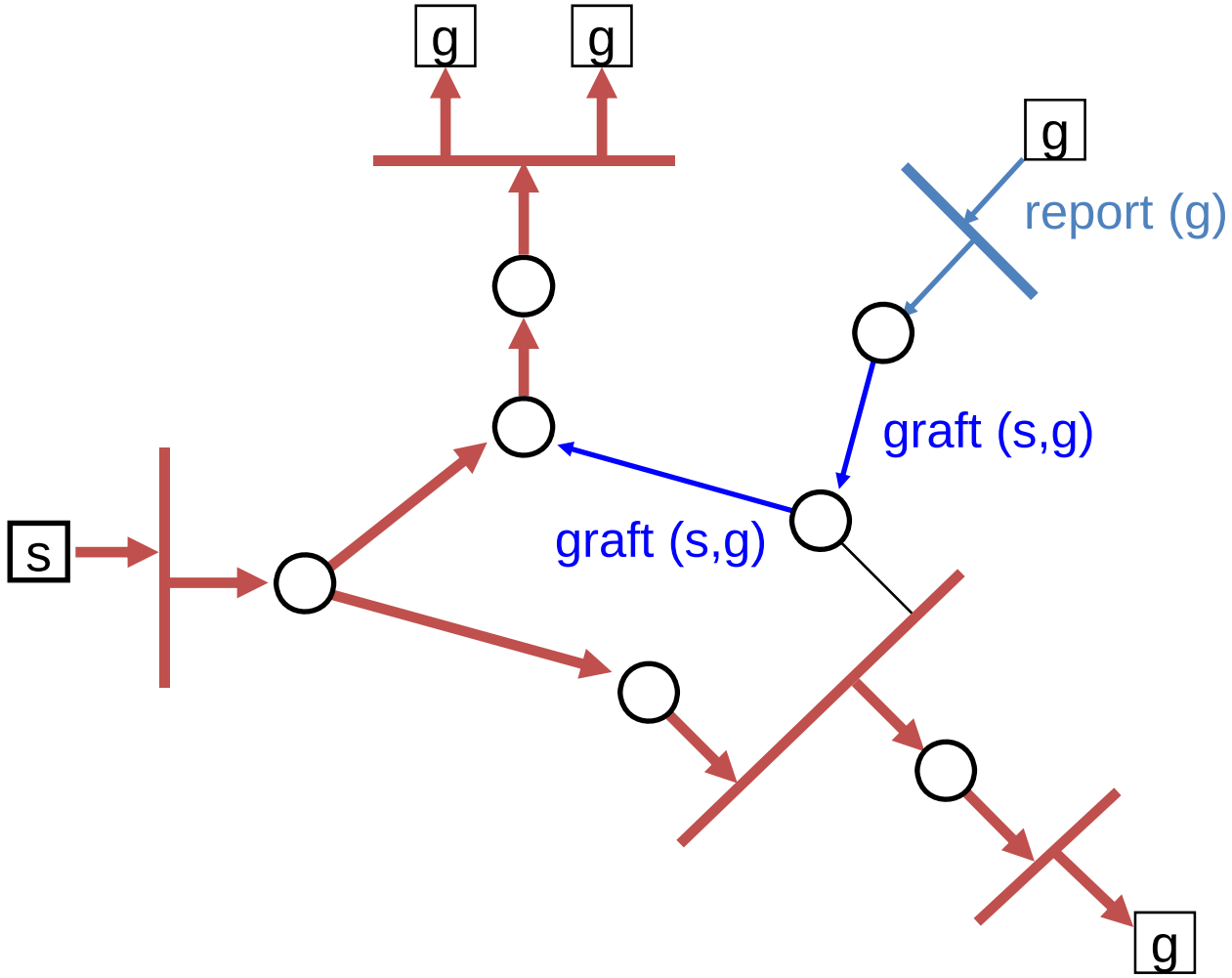
Pruning



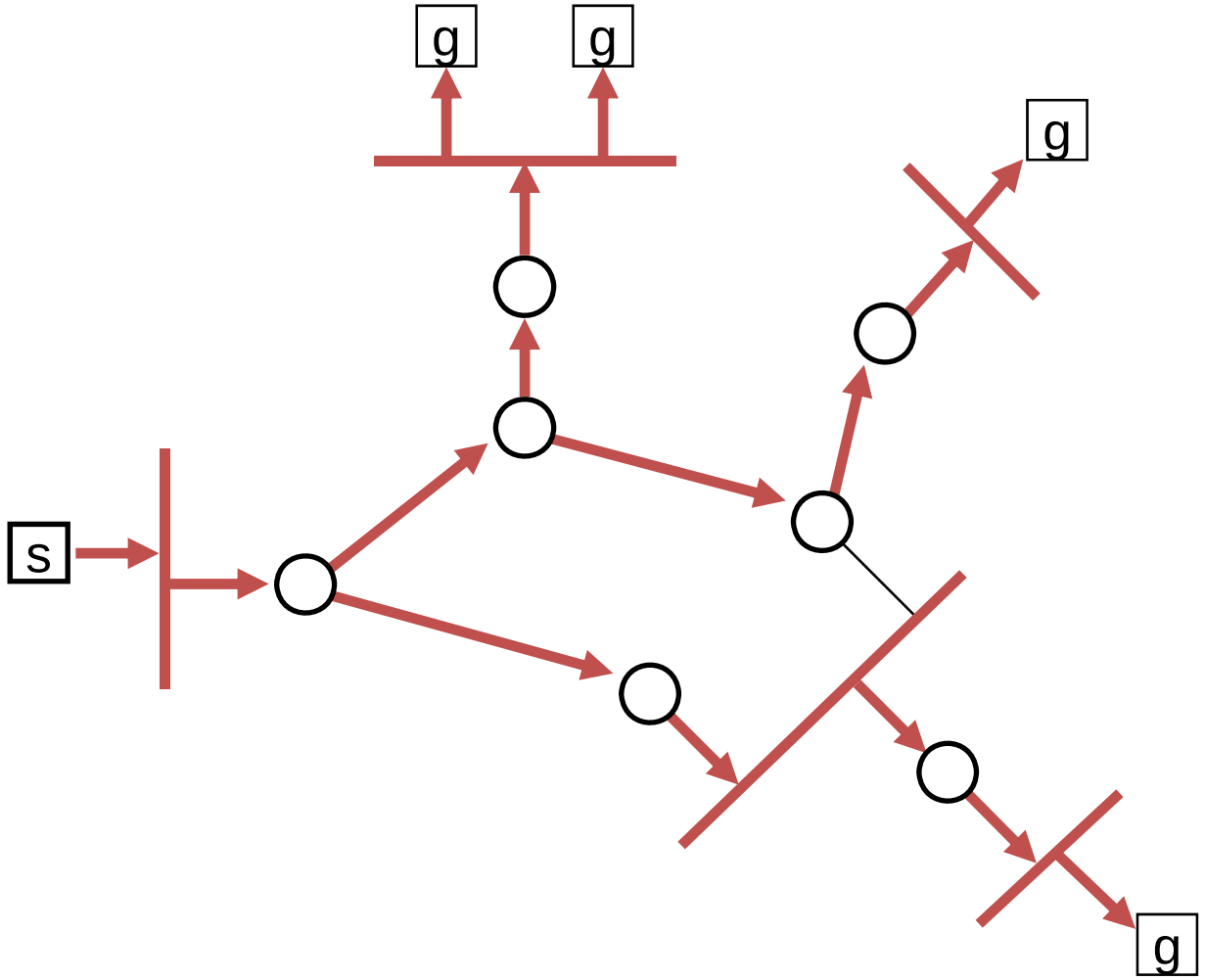
Steady State after Pruning



Grafting on New Receivers

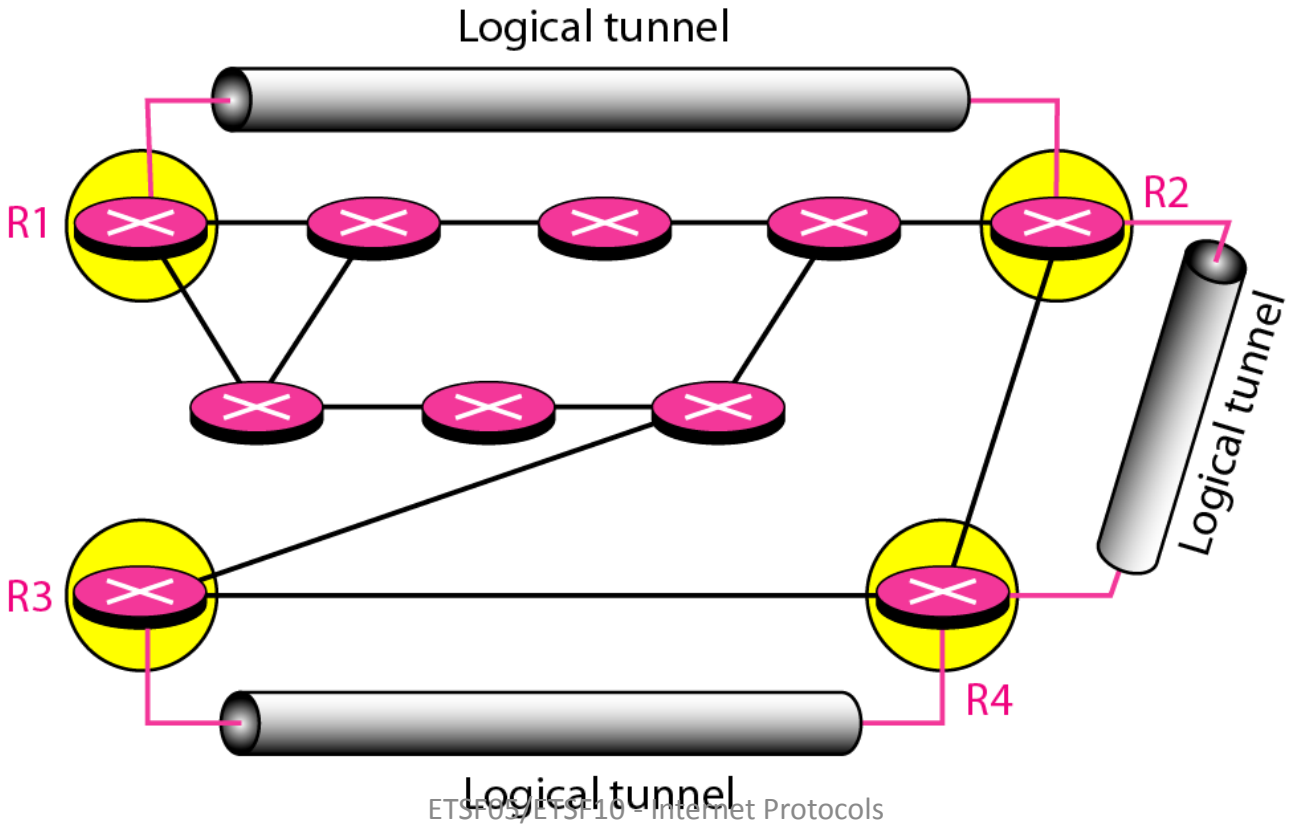


Steady State after Grafting



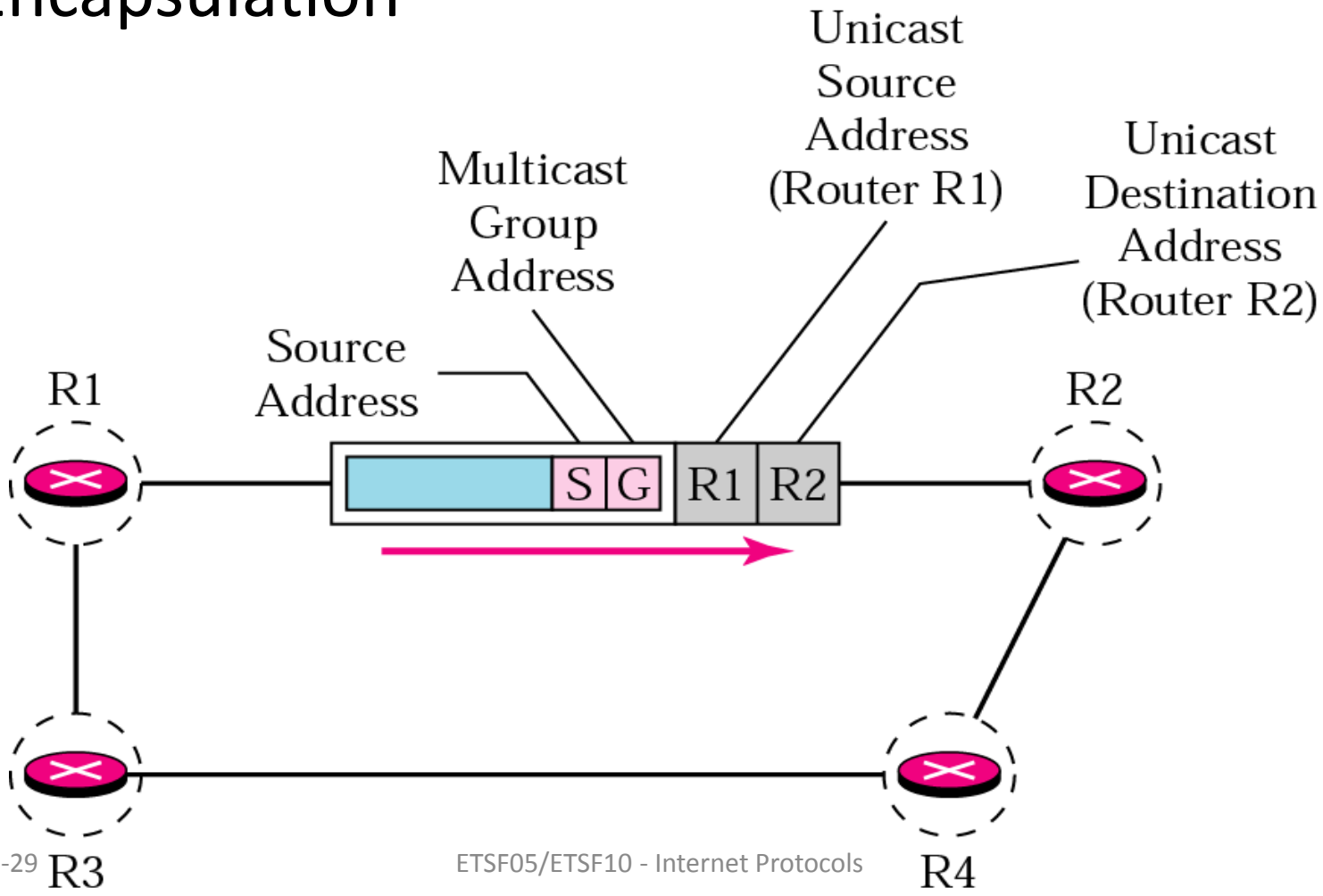
Logical Tunnelling

- Very few Internet routers can multicast
  - How to connect them?



Multicast Backbone (MBONE)

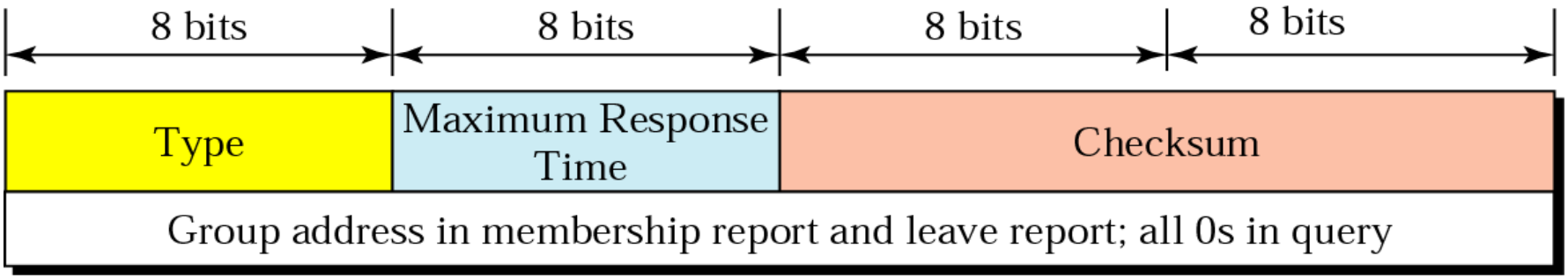
- Encapsulation

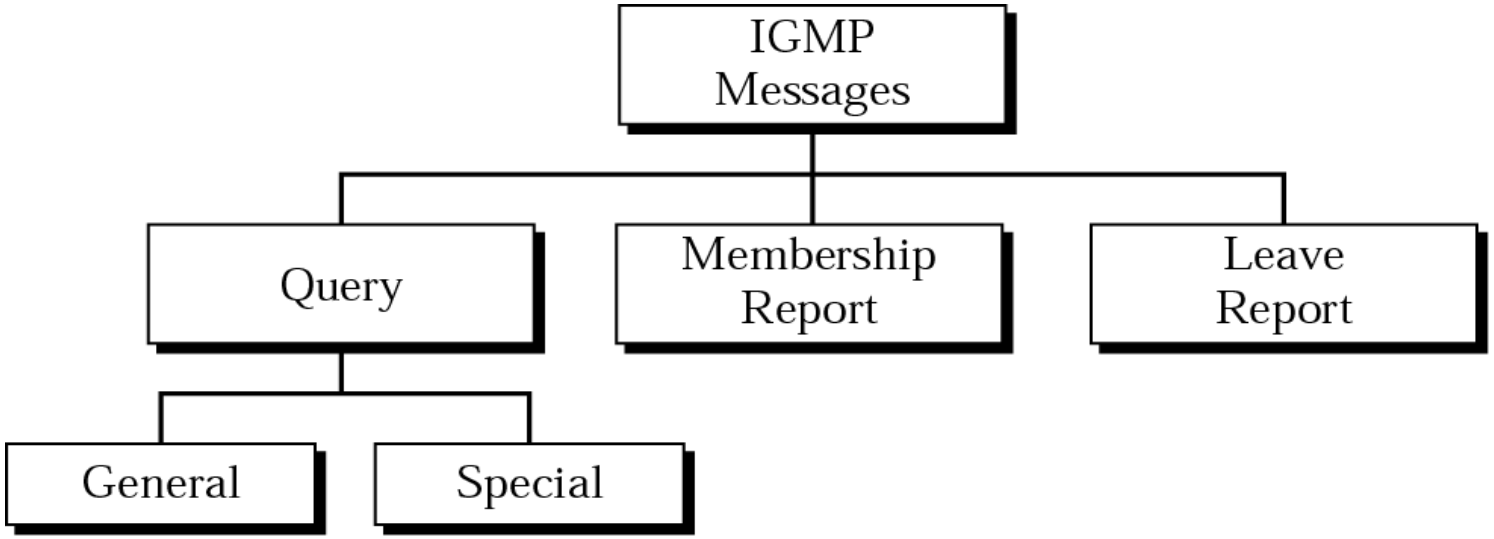


Internet Group Management Protocol

- IGMP, runs on top of IP
- Not a multicast protocol
  - Complementary
  - Runs in the leaves of the network
- Manages group membership
  - Provides multicast router with info

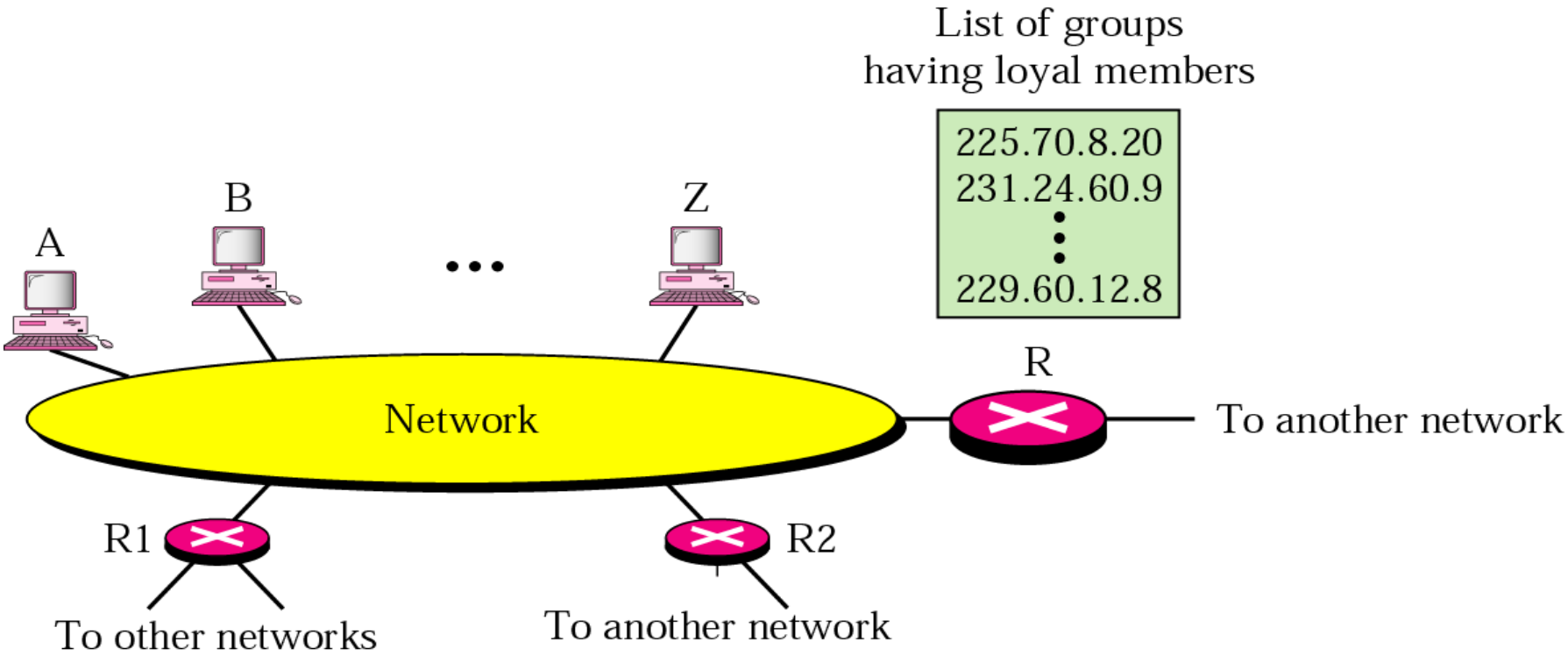
IGMP Message Format





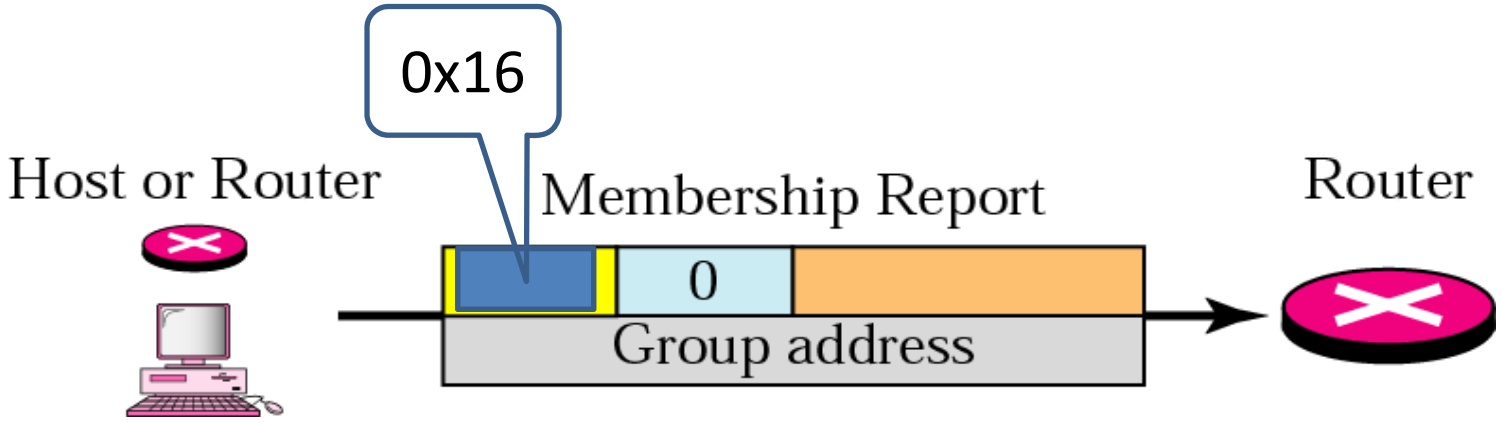
IGMP Operation

- Only one router distributes packets in a group
  - Other routers may be serving their networks

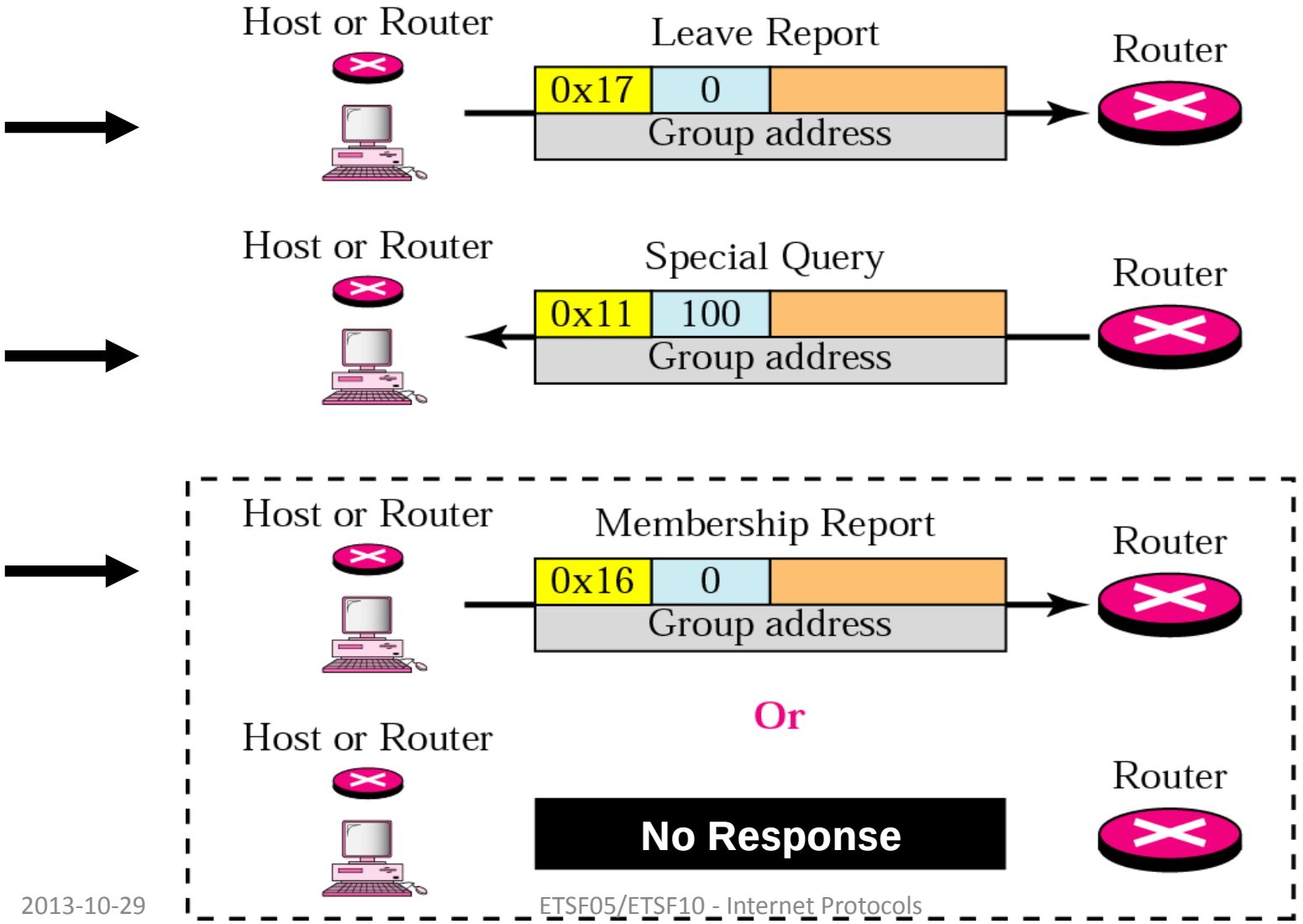


Joining a Group

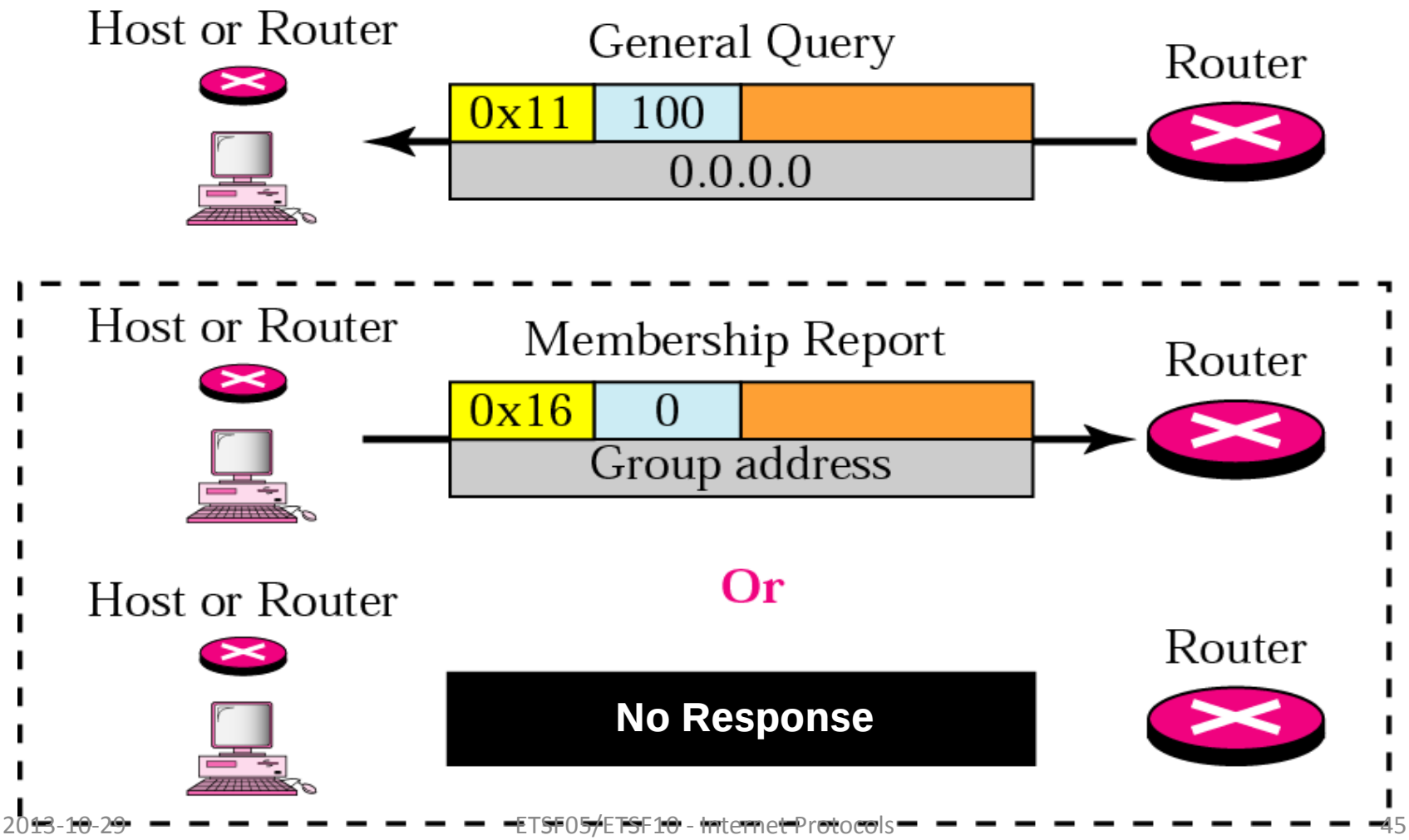
- Request to router
  - Forwarded if first for a group



Leaving a Group

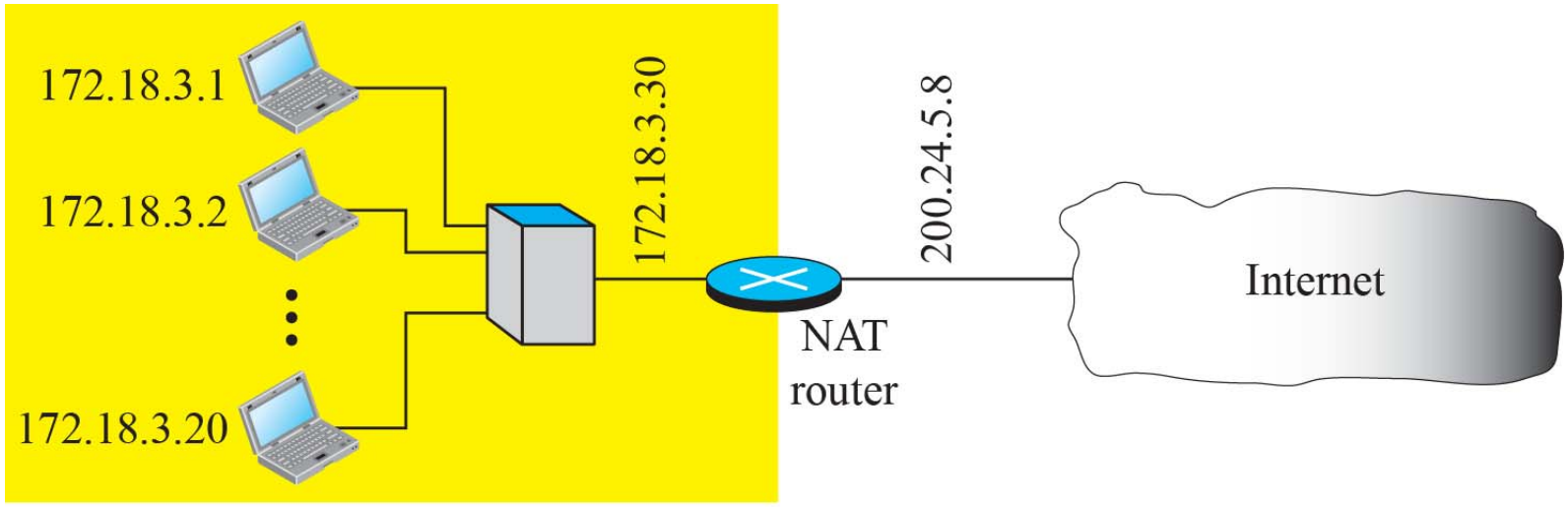


IGMP General Query



NAT - Network Address Translation

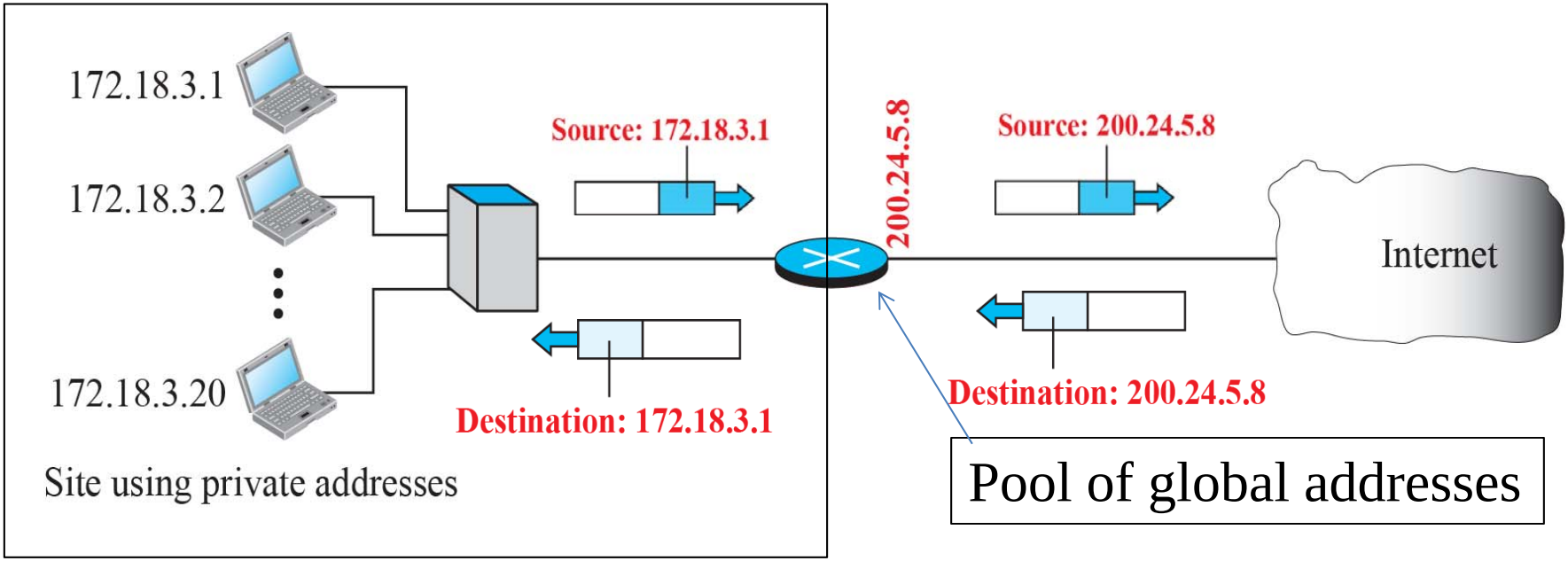
- Sharing of routable addresses (scarse resource)
- Adds some security …



Site using private addresses

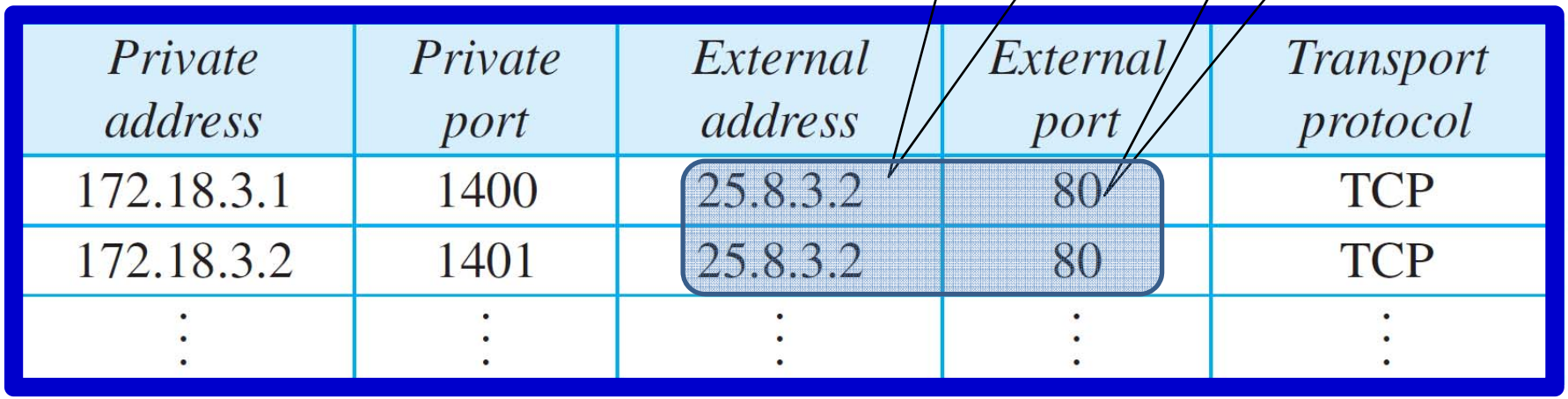
NAT (network address only)

- Change source address on outgoing packets
- Add address pair to active translations table
- Only one internal address per destination



NAT extended

- Add transport layer port

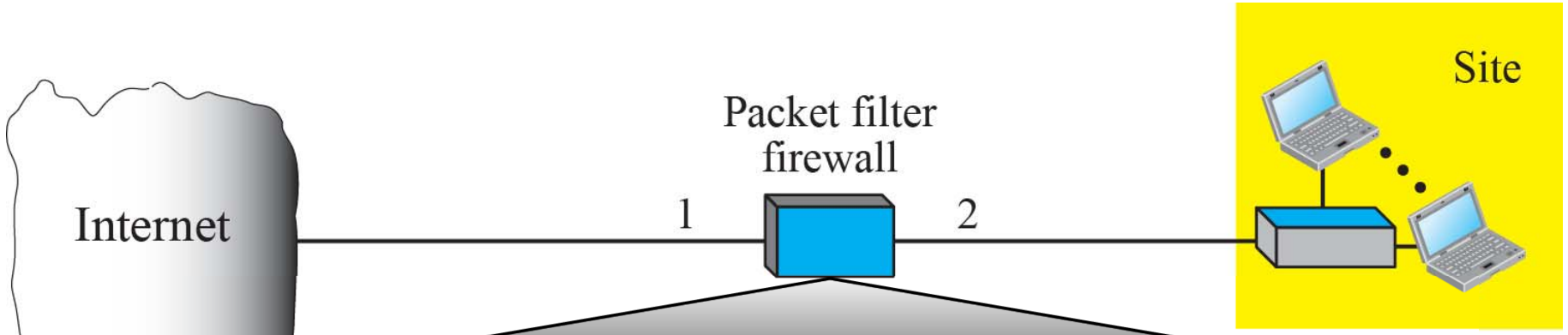


Private address	*Private port*	*External address*	*External port*	*Transport protocol*
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
⋮	⋮	⋮	⋮	⋮

- Normally initiated from inside
- Port forwarding: Setup static entry in table

Firewalls: Filtering

- Accept or reject



Interface	Source IP	Source port	Destination IP	Destination port
1	131.34.0.0	*	*	*
1	*	*	*	23
1	*	*	194.78.20.8	*
2	*	*	*	80

Proxy Firewall

- Filter on message content
- Application gateway acts as proxy for http server