TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (1/31)

S-38.3192 Verkkopalvelujen tuotanto

S-38.3192 Network Service Provisioning

Lecture 9: Multicast

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (2/31)

Communication modes

- Unicast
  - One sender
  - One receiver
- Broadcast
  - One sender
  - All receive
- Multicast
  - One sender
  - N receivers
- Anycast
  - One sender
  - One receiver on a group

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (3/31)

Forwarding modes

- Point-to-point
  - From single ingress to single egress
  - ’link’
- Point-to-multipoint
  - From single ingress to multiple egress
  - ’tree’
- Multipoint-to-multipoint
  - From multiple ingress to multiple egress
  - ’mesh’

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (4/31)

Multicast

- It is all about determining where
  - are the receivers
    - Is the knowleddge of receivers maintained
      - Broadcast vs multicast
  - the replication of packets should be done
    - Is there knowledge about service within the network
      - Source replication vs network replication

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (5/31)

Multicast forwarding

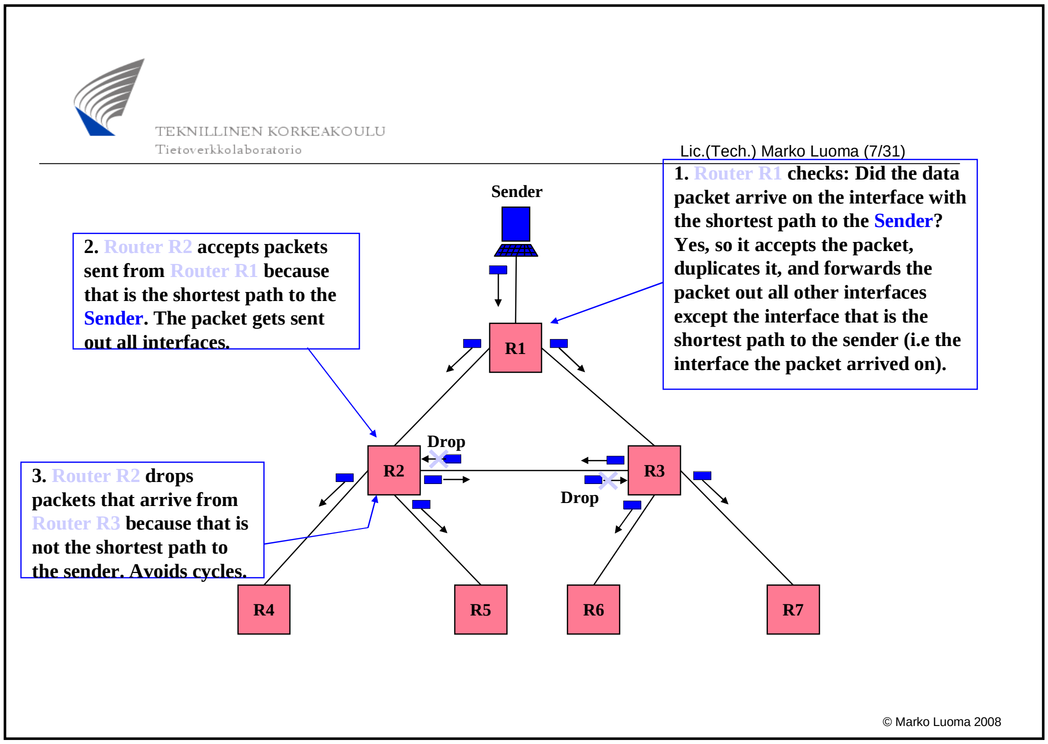
- Problem
  - How to form a point-to-multipoint structure that is efficient
    - e.g. Shortest path tree
- Reverse path forwarding
  - Logic:
    - Send out replicate of the packet to all interfaces except to the one with the shortest path to the sender
      - That is where the packet should have come in the first place
    - Dropped duplicates along the way on the network
      - Wasted resources
    - From the sender to all other stations on the network
      - Wasted resources

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (6/31)

Forwarding

- Unicast forwarding
  - Forwarding of packets towards destination via the shortest path
    - SPF is based on weighting of IGP
- Reverse path forwarding
  - Forwarding of packets away from the shortest path to the sender
    - No knowledge of the destination only about the source
- Reverse path forwarding check
  - Accept packets only from the link which points to SPF towards source



TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (8/31)

Building an intelligent tree

- Requirements
  - We know where the receivers are
    - Group management protocol (IGMP)
      - Receivers express their existence to the network
        - » join, leave
  - We are able to communicate that knowledge within the network
    - Multicast routing protocol (PIM)
      - Activation of RPF on selected downstream links

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

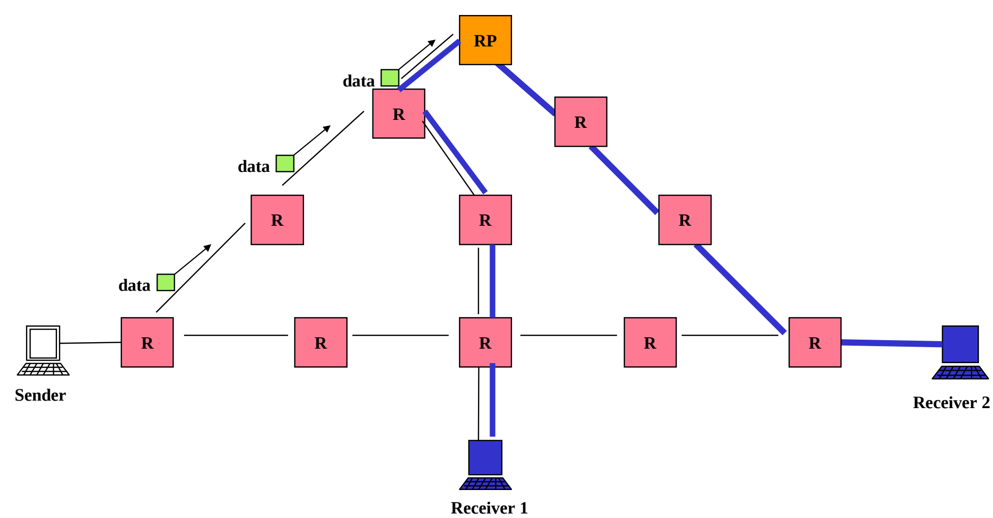
Lic.(Tech.) Marko Luoma (9/31)

Building an intelligent tree

- If we want to minimize the state information within the core we can build a **shared tree** for each sender
  - A single point within the network acts as a relay agent
    - Traffic in unicast sent to relay point which then multicasts it forward in RPF-tree
  - Efficiency depends on the distribution of sources and receivers in relation to location of relay-point
    - Usually very uneffiecient due to sparse nature of receivers and variable locations of sources

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (10/31)



© Marko Luoma 2008

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

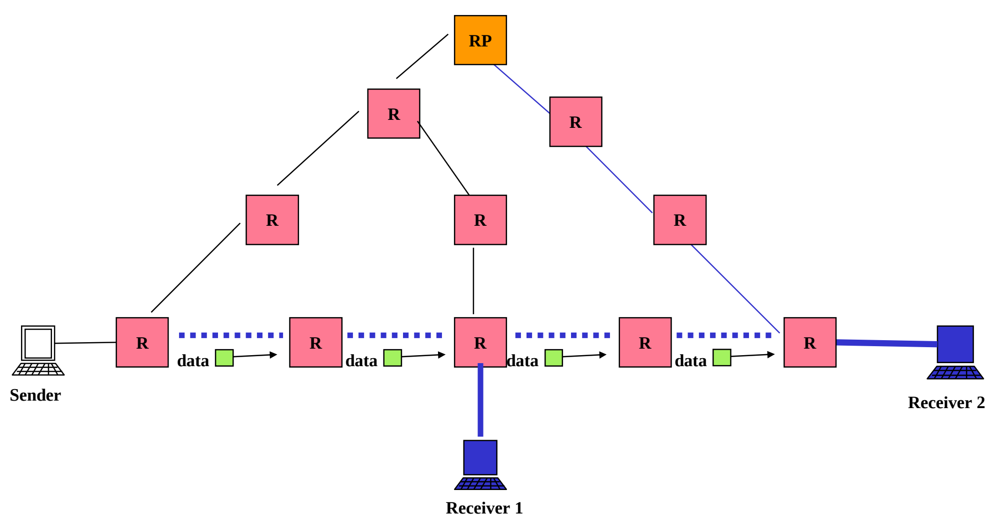
Lic.(Tech.) Marko Luoma (11/31)

Building an intelligent tree

- If we want to achieve maximal efficiency, we must use **shortest paths toward source** of the information not towards common replication point
  - Use explicit source rooted RPF-tree
    - Joining/grafting to multicast tree at the lowest level possible
    - Pruning from the tree if there are no receivers below

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (12/31)



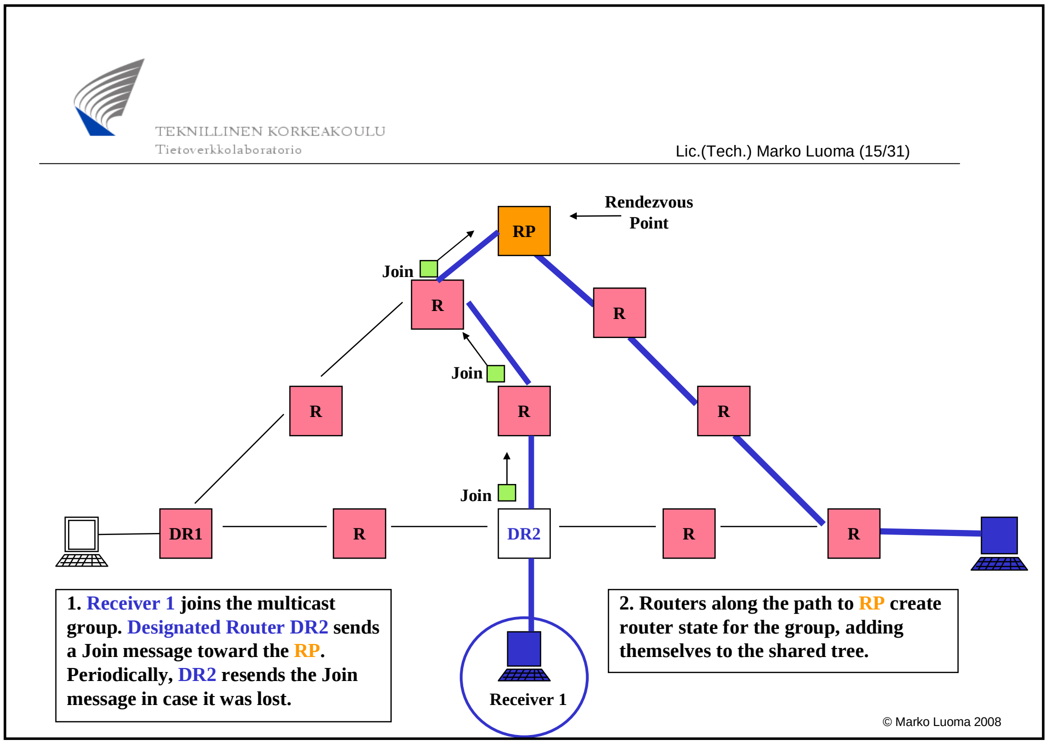
PIM

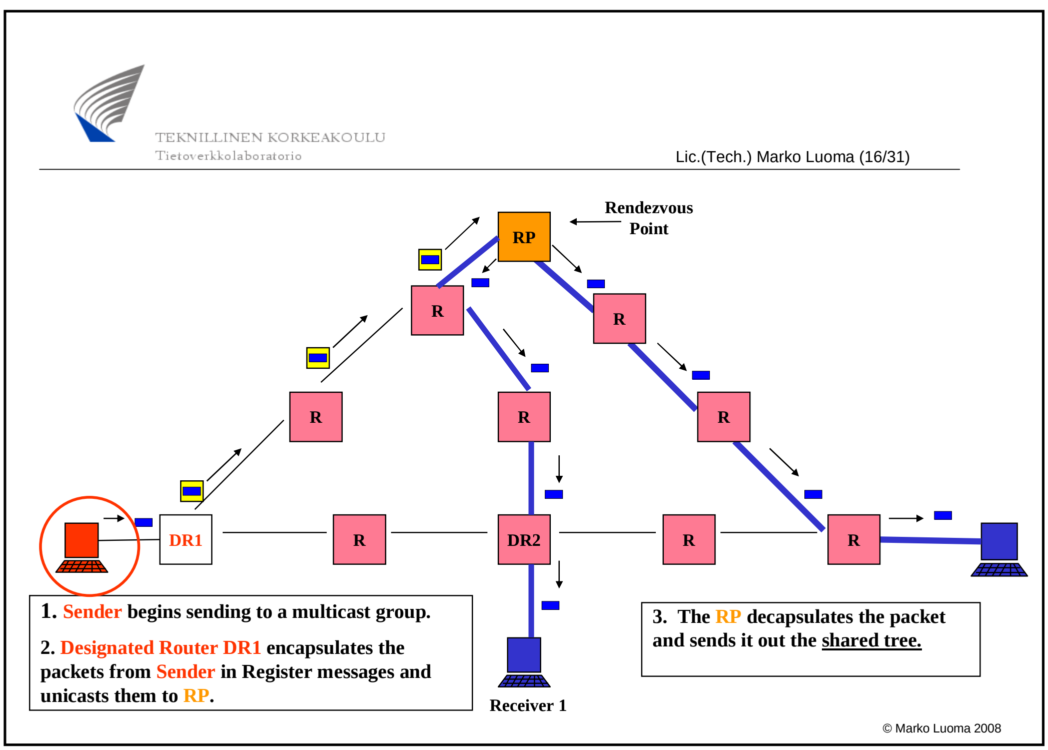
- Is a combination of shared tree and source specific tree operation
  - Infrequent packets are replicated in a single point (shared tree)
    - Randevouz point
  - High data volumes are transferred in more optimal way (source specific tree)
    - Decission based on predefined policy
      - Data rate greater than x
      - Amount of bytes greater than x

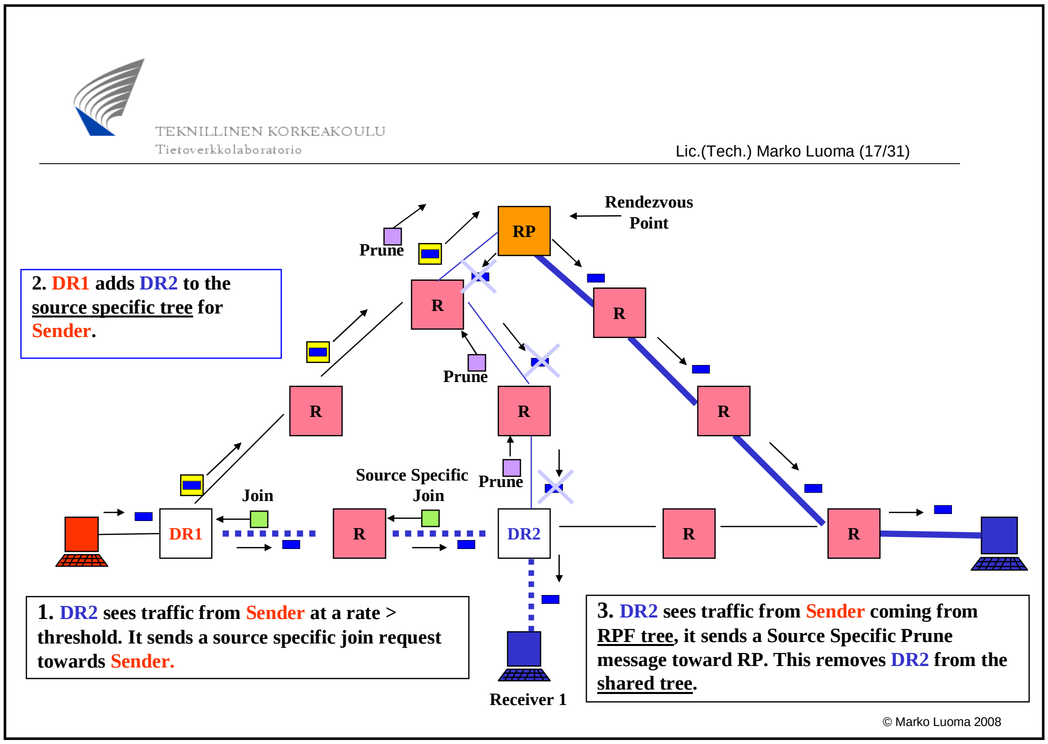
Multicast Distribution Trees

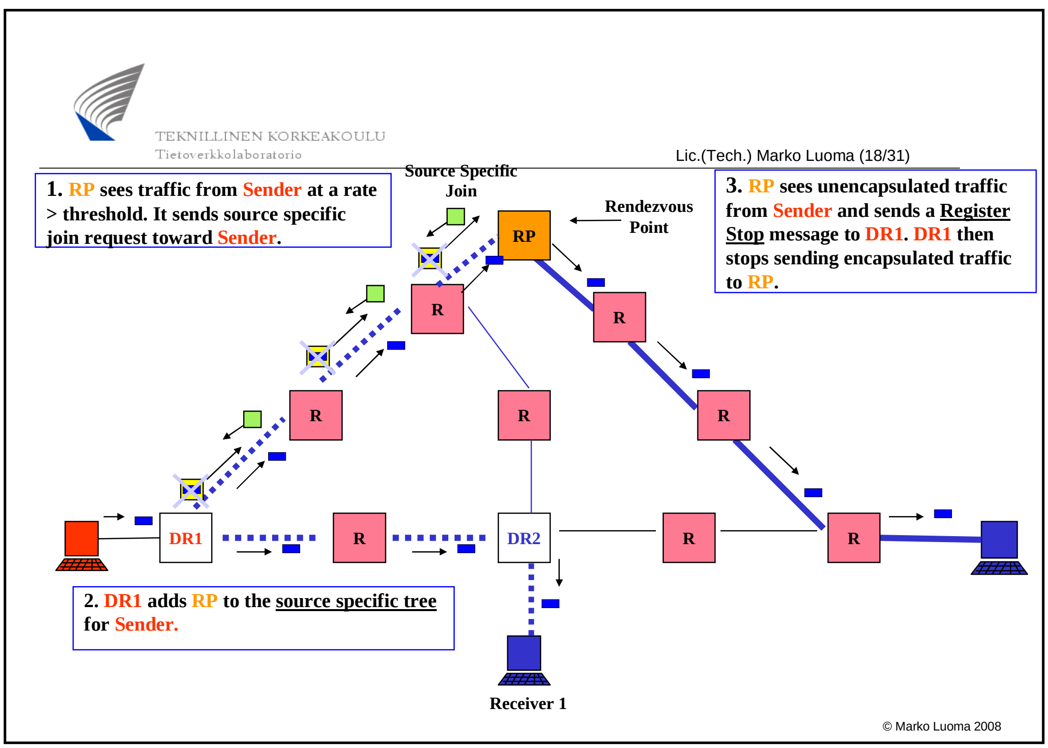
- Source specific tree
  - Takes more resources from the network
    - Amount of state information is dependent of N(S,G)
      - Every source (S) has its own distribution tree per group (G)
    - Optimal path from the source to all receivers within the group
      - Minimizes the IGP cost towards source from each receiver
  - Optimal for one-to-many distribution (point-to-multipoint)

- Shared trees
  - Uses less resources from the network
    - Amount of state information is depent of N(*,G)
      - All sources (*) are mapped into single distribution tree per group (G)
      - Usually suboptimal path from the source to all receivers within the group
        - » Unless RP is rooted at the source


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (15/31)
Rendezvous Point
RP
Join
R
R
Join
R
R
R
Join
DR1
R
DR2
R
R
1. Receiver 1 joins the multicast group. Designated Router DR2 sends a Join message toward the RP. Periodically, DR2 resends the Join message in case it was lost.
Receiver 1
2. Routers along the path to RP create router state for the group, adding themselves to the shared tree.
© Marko Luoma 2008


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (16/31)
Rendezvous Point
RP
R
R
R
R
R
DR1
R
DR2
R
R
1. Sender begins sending to a multicast group.
2. Designated Router DR1 encapsulates the packets from Sender in Register messages and unicasts them to RP.
Receiver 1
3. The RP decapsulates the packet and sends it out the shared tree.
© Marko Luoma 2008


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (17/31)
Rendezvous Point
RP
Prune
Prune
Prune
R
R
R
R
R
R
R
Join
Source Specific Join
DR1
DR2
Receiver 1
2. DR1 adds DR2 to the source specific tree for Sender.
1. DR2 sees traffic from Sender at a rate > threshold. It sends a source specific join request towards Sender.
3. DR2 sees traffic from Sender coming from RPF tree, it sends a Source Specific Prune message toward RP. This removes DR2 from the shared tree.
© Marko Luoma 2008

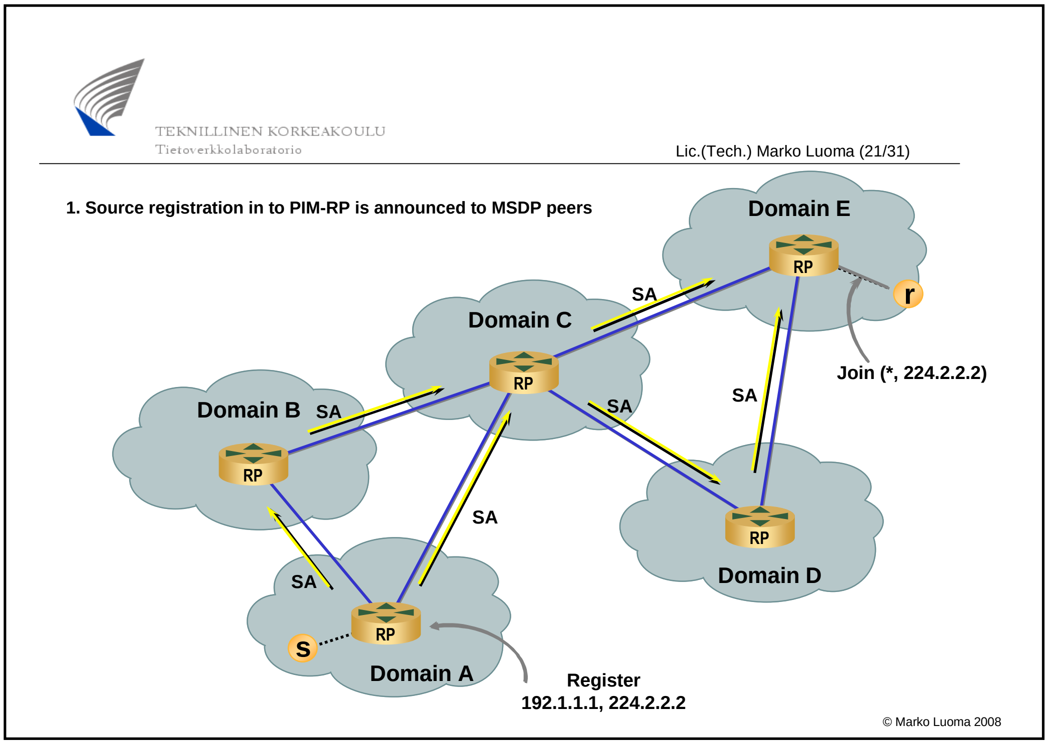

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (18/31)
1. RP sees traffic from Sender at a rate > threshold. It sends source specific join request toward Sender.
Source Specific Join
Rendezvous Point
RP
3. RP sees unencapsulated traffic from Sender and sends a Register Stop message to DR1. DR1 then stops sending encapsulated traffic to RP.
R
R
R
R
R
R
R
DR1
DR2
2. DR1 adds RP to the source specific tree for Sender.
Receiver 1
© Marko Luoma 2008

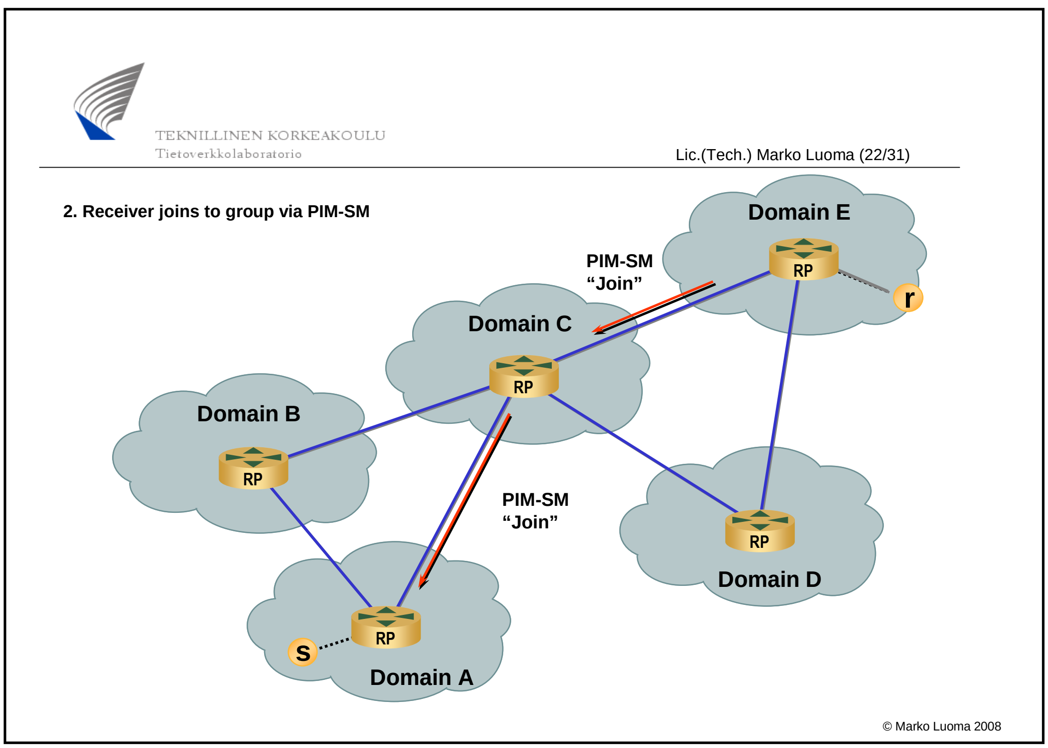
Rendezvous Point (RP)

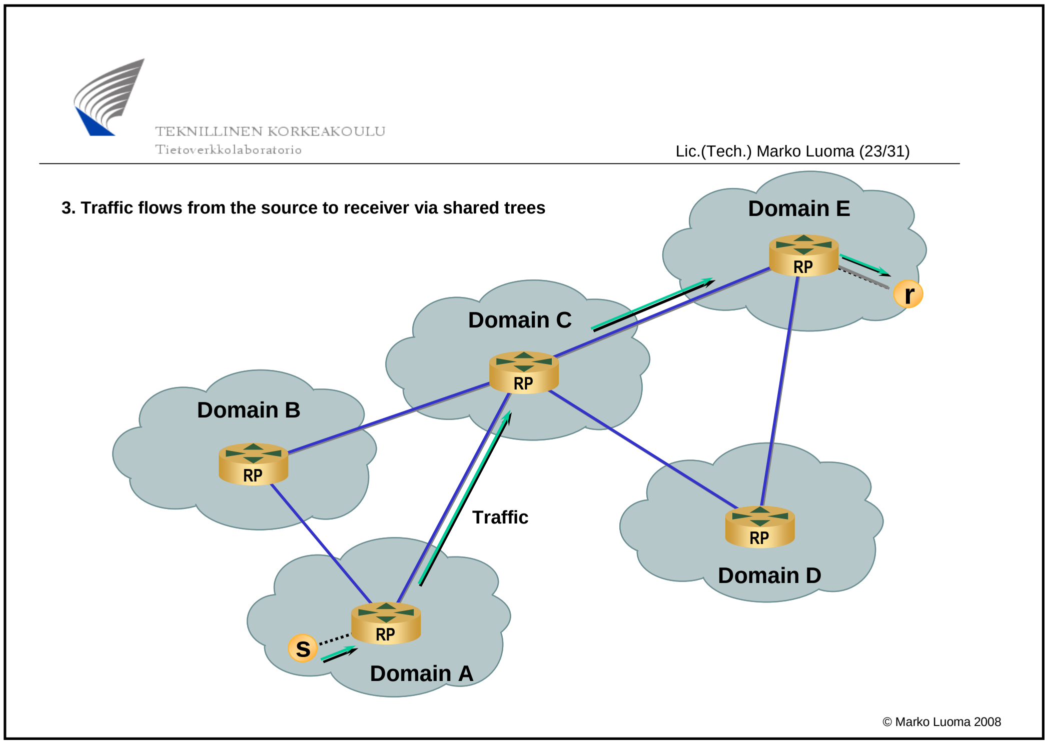
- Location of RP can be based on
  - Static configuration of served groups
  - Bootstrap process with priorities
  - Anycast operation
- Bootstrap Router (BSR)
  - Dynamically elected (like OSPF DR election process)
  - Constructs a set of RP IP addresses based on received Candidate-RP messages

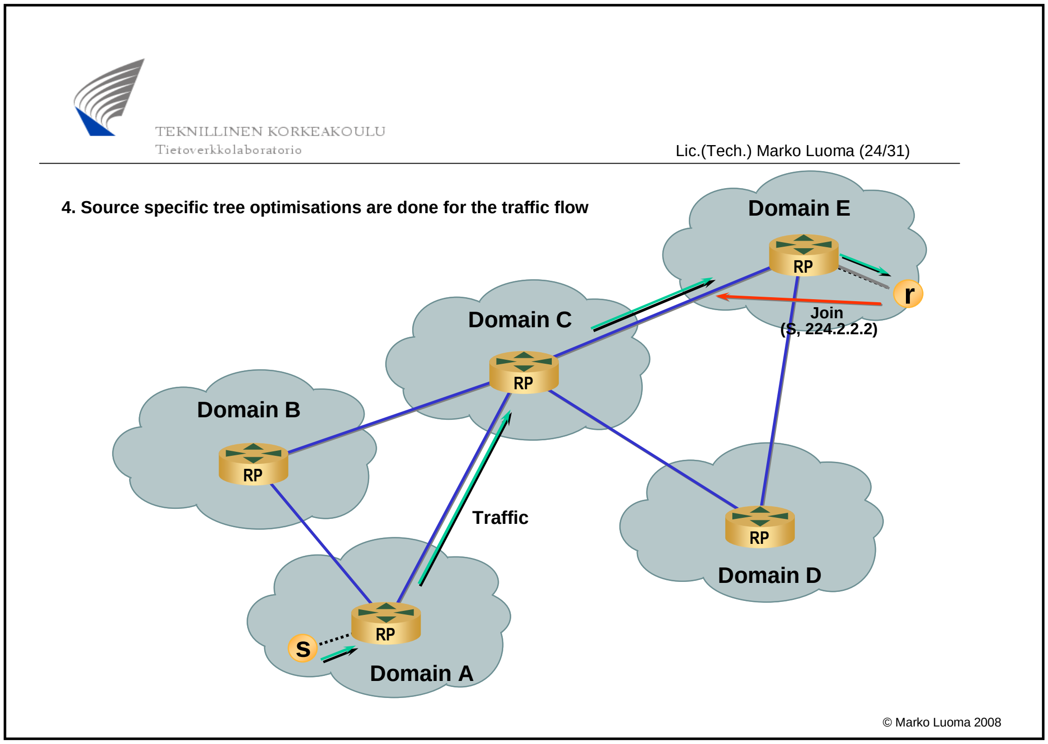
Interdomain multicast

- Need to transfer multicast routing information across domain borders
  - Active groups (senders and receivers)
    - Multicast Source Discovery Protocol (MSDP)
  - Unicast routes for RPF
    - Multicast BGP (MBGP)


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (21/31)
1. Source registration in to PIM-RP is announced to MSDP peers
Domain E
Domain C
Domain B
Domain D
Domain A
RP
r
S
SA
Join (*, 224.2.2.2)
Register
192.1.1.1, 224.2.2.2
© Marko Luoma 2008

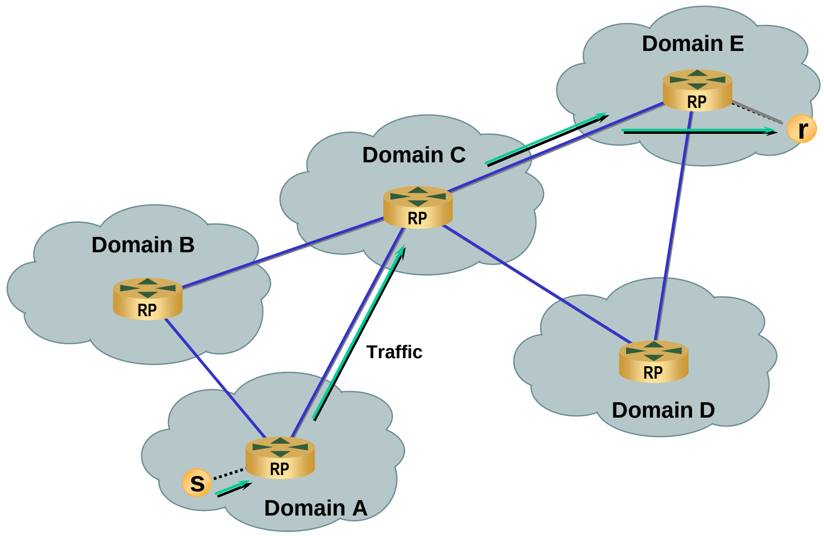

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (22/31)
2. Receiver joins to group via PIM-SM
Domain E
Domain C
Domain B
Domain D
Domain A
RP
r
S
PIM-SM
“Join”
© Marko Luoma 2008


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (23/31)
3. Traffic flows from the source to receiver via shared trees
Domain E
RP
r
Domain C
RP
Domain B
RP
Traffic
RP
Domain D
S
RP
Domain A
© Marko Luoma 2008


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (24/31)
4. Source specific tree optimisations are done for the traffic flow
Domain E
RP
r
Join
(S, 224.2.2.2)
Domain C
RP
Domain B
RP
Traffic
RP
Domain D
S
RP
Domain A
© Marko Luoma 2008

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (25/31)



TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (26/31)

MBGP

- Multiprotocol Extensions to BGP (RFC 2283).
  - Tag unicast prefixes as multicast source prefixes for intra-domain multicast routing protocols to do RPF checks.
  - Why same routes two times
    - Allows for inter-domain RPF checking where unicast and multicast paths are non-congruent
      - Inter-provider relationships
        - » Policies ;-)

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (27/31)

MBGP

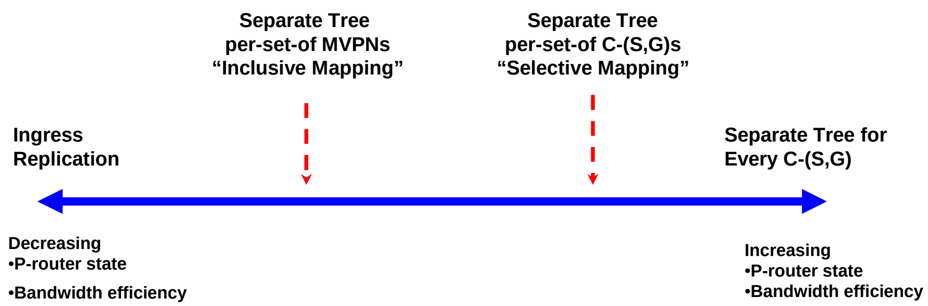
- Different NLRI for unicast and multicast routes
  - Address Family Information (AFI) = 1 (IPv4)
    - SAFI = 1 (NLRI is used for unicast)
    - SAFI = 2 (NLRI is used for multicast RPF check)
    - SAFI = 3 (NLRI is used for both unicast and multicast RPF check)
- Allows for different policies between multicast and unicast
  - For example different ingress/egress point for unicast and multicast traffic

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (28/31)

MPLS VPN multicast

- Multicasting in MPLS VPN is not straight forward
  - Optimal replication requires state in core routers (P)



TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (29/31)

Scalability

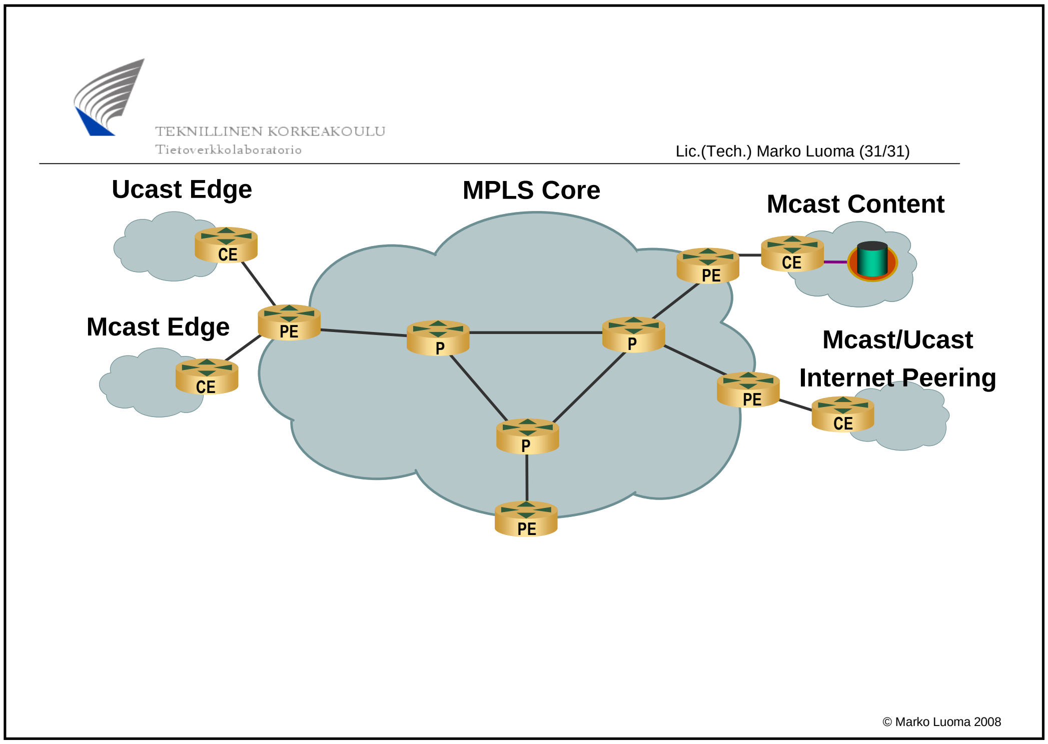
- **Optimizing performance**
  - A given customer (multicast) packet should traverse a given service provider link at most once
  - Deliver customer multicast traffic to only PEs that have (customer) receivers for that traffic
  - Deliver customer multicast traffic along the "optimal" paths within the service provider (from the ingress PE to the egress PEs)
- **Optimizing state-space**
  - The amount of state within the service provider network required to support Multicast in 2547 VPN service should be no greater than what is required to support unicast in 2547 VPN service
  - The overhead of maintaining the state to support Multicast in 2547 VPN service should be no greater than what is required to support unicast in 2547 VPN service

TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio

Lic.(Tech.) Marko Luoma (30/31)

Data plane

- How to tunnel multicast traffic in Service Provider network
  - By using p2mp LSP's
    - Static distribution trees
  - GRE encapsulation and IP multicast forwarding
    - Multicast distribution trees
      - Aka multicast service
  - Unicast tunneling with ingress replication
- How to exchange multicast routing information with service provider network
  - MBGP
  - PIM
  - MSDP
  - Tunneling through


TEKNILLINEN KORKEAKOULU
Tietoverkkolaboratorio
Lic.(Tech.) Marko Luoma (31/31)
Ucast Edge
MPLS Core
Mcast Content
Mcast Edge
Mcast/Ucast
Internet Peering
CE
PE
P
P
PE
CE
CE
PE
P
CE
PE
© Marko Luoma 2008