

IETF'43, 1998/12/7-11, Orlando, Florida, USA

Static Multicast and PIM-SM/CBT/OCBT

Manolo Sola

Waseda University

Masataka Ohta

Tokyo Institute of Technology

`ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt,
draft-sola-pim-static-multicast-00.txt, draft-sola-ocbt-static-multicast-00.txt`

Contents

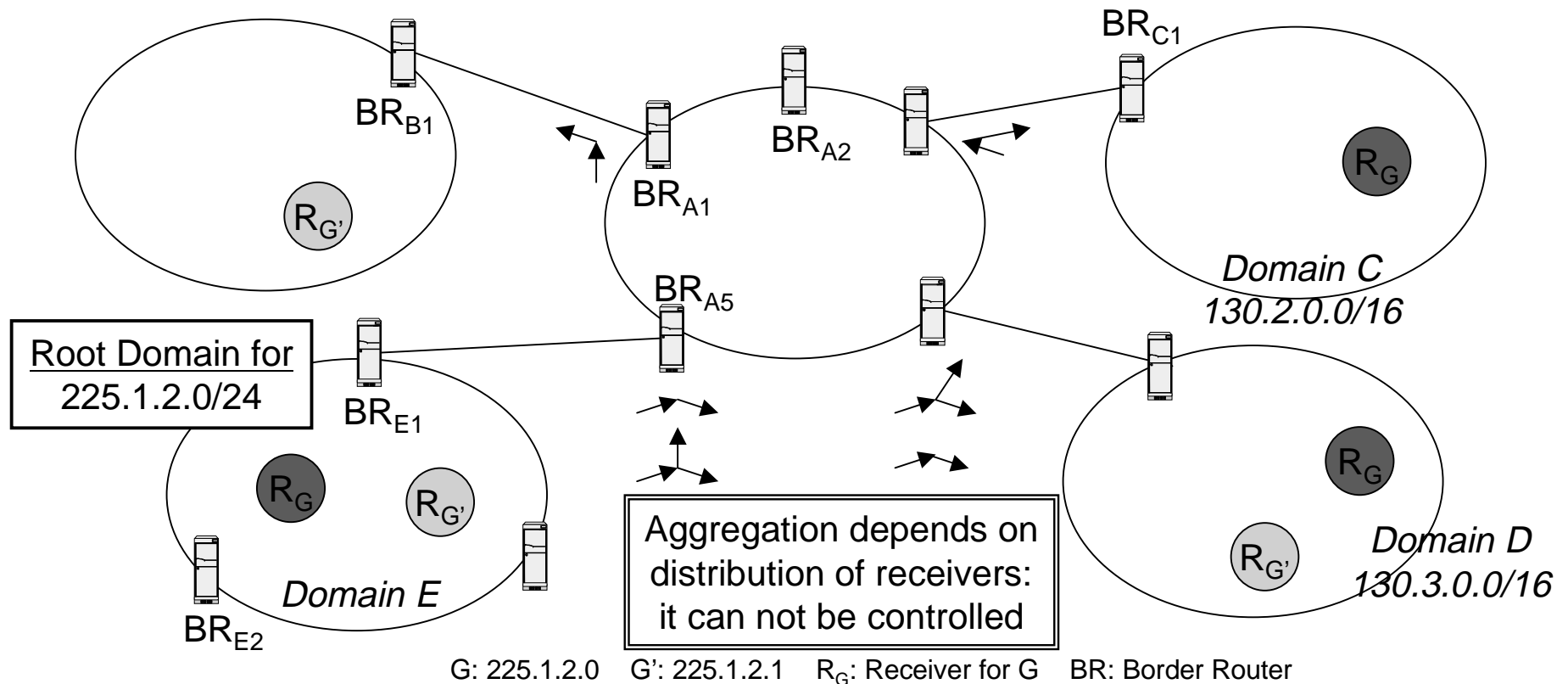
1. Static Multicast
 - 1.1. Multicast Address Allocation
 - 1.2. Multicast Core/RP Location
 - 1.3. Multicast Session Announcement
2. Static Multicast and PIM-SM
3. Static Multicast and CBT/OCBT

1.1. Multicast Address Allocation

- For a multicast group, senders and receivers may be placed anywhere in the internet.
- Consecutive multicast group addresses will not necessarily have identical multicast distribution trees.
- That's why aggregation of multicast addresses at routers can not be controlled.
- This is also true no matter how multicast addresses are allocated.

1.1. Multicast Address Allocation

- Allocation of aggregated multicast addresses does not mean that multicast routing tables can be aggregated

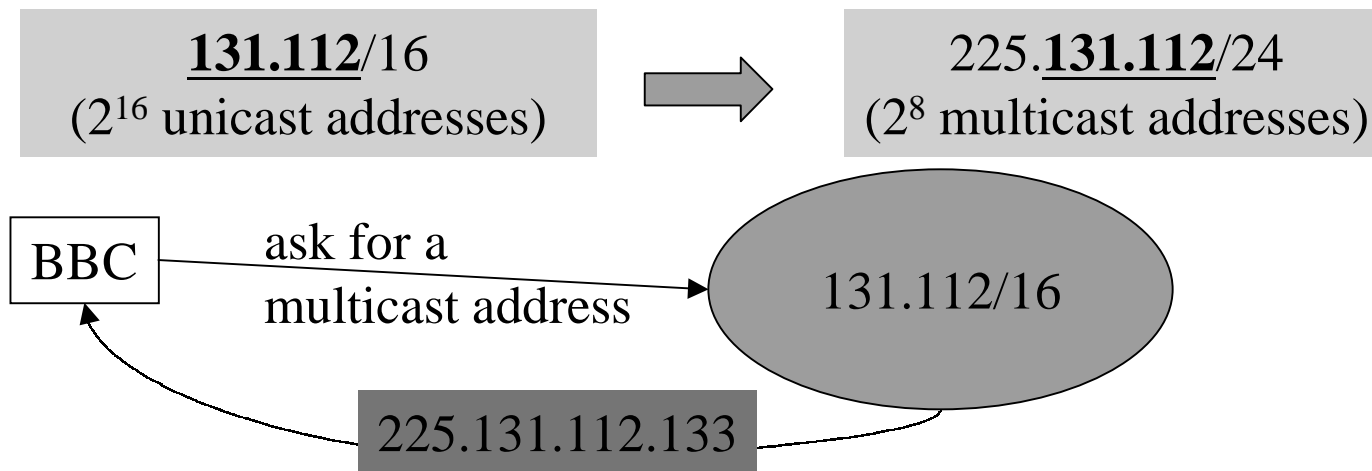


1.1. Multicast Address Allocation

- aggregation of multicast addresses
 - meaningless
- multicast address allocation
 - use the same strategy as in unicast
- multicast also needs a policy for allocation
 - 1 (up to 16) multicast address per 256 unicast addresses
 - multicast addresses can be further delegated

1.1. Multicast Address Allocation

- Example: use 225.0.0.0/8 for static allocation



133.112.131.225.in-addr.arpa.	CNAME	mcast.133.112.131.in-addr.arpa.
mcast.133.112.131.in-addr.arpa.	PTR	ch1.bbc.com.
ch1.bbc.com.	A	225.133.112.131

<ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt>,
[draft-sola-pim-static-multicast-00.txt](ftp://ftp.ietf.org/internet-drafts/draft-sola-pim-static-multicast-00.txt), [draft-sola-ocbt-static-multicast-00.txt](ftp://ftp.ietf.org/internet-drafts/draft-sola-ocbt-static-multicast-00.txt)

1.2. Multicast Core/RP Location

- location of the Core in CBT

```
ch1.bbc.com.      A      255.192.2.3
                  CORE  london-station.bbc.com.
```

- location of the RP in PIM-SM

```
ch1.bbc.com.      A      255.192.2.3
                  RVP   london-station.bbc.com.
```

- multiple RPs

```
ch1.bbc.com.      A      255.192.2.3
                  RVP   london-station.bbc.com.
                  RVP   wales-station.bbc.com.
```

- DNS should continue to scale, is becoming secure and dynamically updated

<ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt>,
[draft-sola-pim-static-multicast-00.txt](ftp://ftp.ietf.org/internet-drafts/draft-sola-pim-static-multicast-00.txt), [draft-sola-ocbt-static-multicast-00.txt](ftp://ftp.ietf.org/internet-drafts/draft-sola-ocbt-static-multicast-00.txt)

1.3. Mult. Session Announcement

- URL containing a session description
 - RFC1738 (Uniform Resource Locator): “information for location and access of resources via the Internet”

<scheme>:<scheme-specific-part>

```
session://mcast.bbc.com/?t=2873397496+2873404696&
m=audio+3456+RTP-AVP+0&m=video+2232+RTP-AVP+31
```

- Example: a user finds that URL inside of a web page, clicks on it, send/receive ? start sending/receiving stream, ...

`ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt,`
`draft-sola-pim-static-multicast-00.txt, draft-sola-ocbt-static-multicast-00.txt`

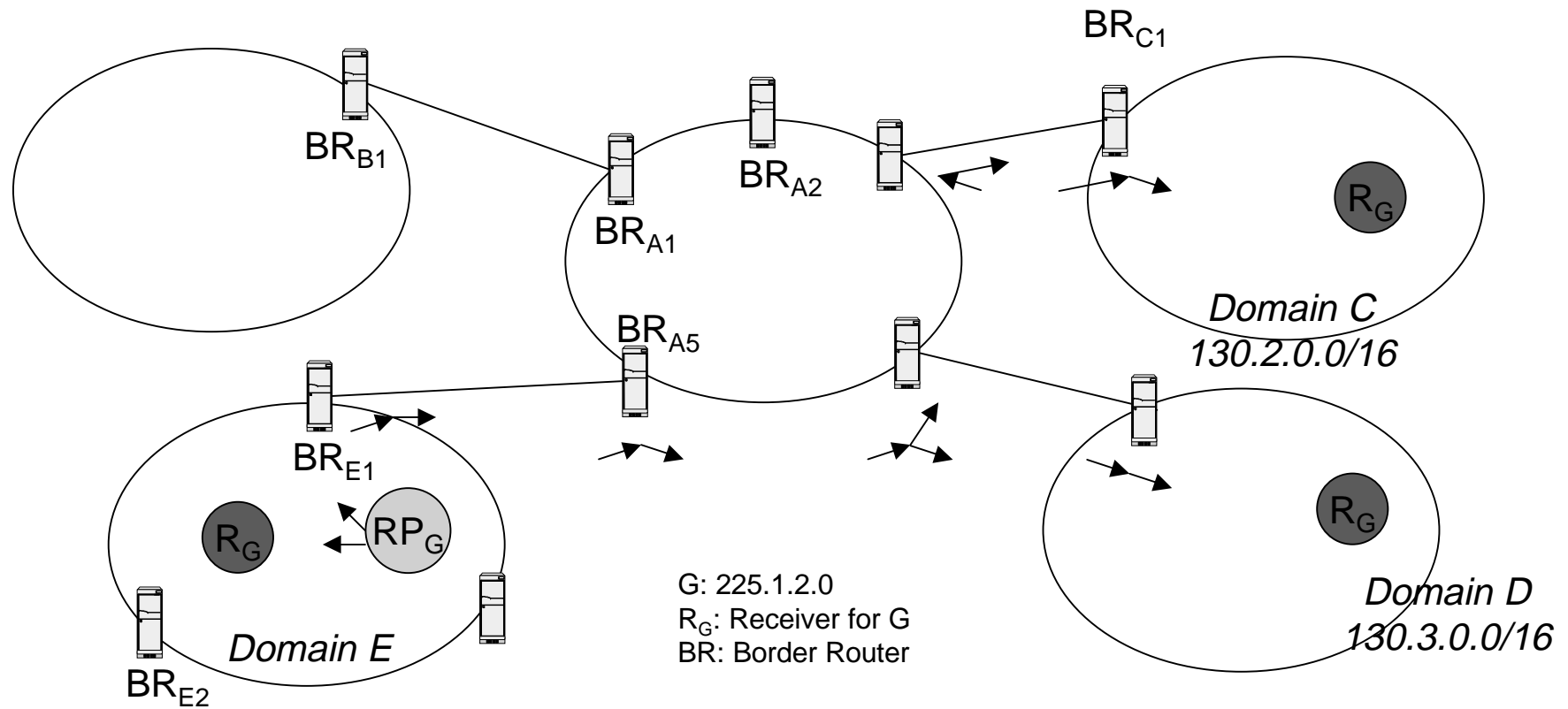
2. Static Multicast and PIM-SM

- Behavior of
 - Receivers
 - Sources
 - DRs
 - RPs
 - Intermediate routers

Receivers/Sources

- PIM-SM: How to discover RPs
 - Replace Bootstrap mechanism by Static Multicast
- Receiver
 - sends IGMP, and if necessary DR queries DNS, and send Join/Prune messages towards the RP
- Source
 - A. at the RP: send directly to the multicast group
 - B. queries DNS and sends Register messages directly to one RP.
Nearest RP ? up/down ?
 - C. sends special IGMP message, DR have time to query and discover nearest RP, source sends data, DR sends Register messages to RP.

Architecture



<ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt>,
<draft-sola-pim-static-multicast-00.txt>, <draft-sola-ocbt-static-multicast-00.txt>

Multicast Traffic Policy

- Policy, as in unicast, can be placed at BRs and may be done over:
 - IP unicast addresses of sources or RPs/Cores (for multicast data traffic going from senders to receivers),
 - or over IP unicast addresses of RPs/Cores (for multicast control traffic going from receivers to the RPs/Cores, Join messages for example).
 - If someone needs a policy for these cases different than the one use for unicast forwarding then a different routing table should be built and consulted by the multicast routing daemon when forwarding multicast traffic.
 - There is no need for different multicast protocols running intra and inter domains.

BSR/RP failure

- How PIM-SM support an BSR/RP failure ?
 - Bootstrap timer expires, C-BSRs flood Bootstrap messages, new BSR is elected, C-RPs send C-RP-Adv to BSR, BSR floods Bootstrap message with RP/group mapping hop-by-hop through the domain, then DRs know RP/group mapping
 - If at the BSR the RP timer expires, the RP is not included in next Bootstrap messages
- This method doesn't scale to the whole internet

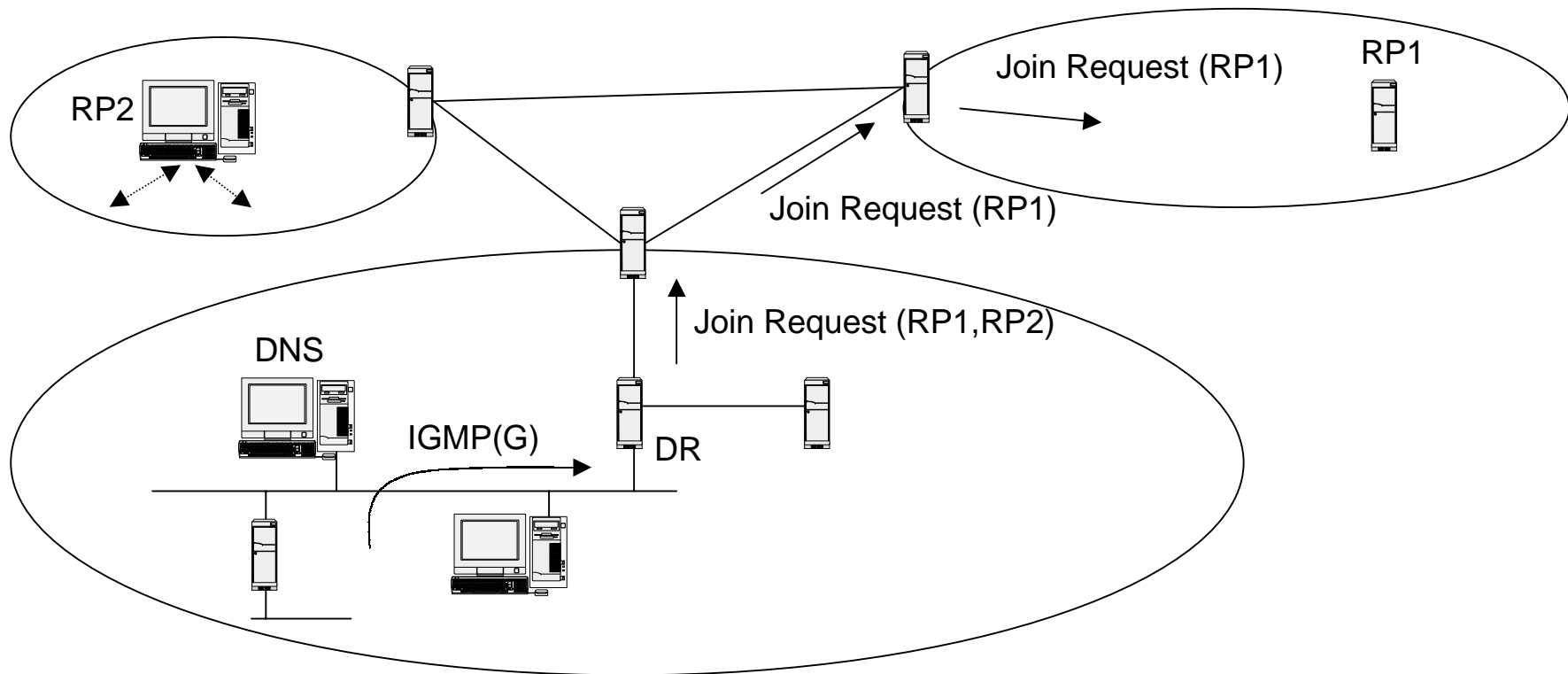
Multiple Active RPs

- Multiple active RPs per multicast group
 - The administrator of a multicast group G must decide how many RPs will exist for G and where they will be located
 - The administrator of the multicast group must contact the administrator of the machine that will behave as an RP so that it can be configured properly.
 - A router must know whether it is an RP by means of a configuration parameter at the router which must be set by the administrator of that router.
 - RPs know about other RPs through DNS
 - DRs know about RPs through DNS

DRs' behavior

- selects the nearest RP
 - If whole set is {RP1, RP2, RP3}
 - {address, oif, cost}={RP1, oif1, 3}, {RP2, oif1, 3}, {RP3, oif2, 4}
 - {RP1, RP2} ... {RP3} ...
 - {RP1} ... {RP2} ... {RP3} ...
- Join Ack containing which RP has been joined

DRs' behavior



<ftp://ftp.ietf.org/internet-drafts/draft-ohta-static-multicast-01.txt>,
<draft-sola-pim-static-multicast-00.txt>, <draft-sola-ocbt-static-multicast-00.txt>

DRs' behavior

- How does a DR with no state for G find an RP to whom forward traffic from a non-member source ?
 - DR joins to receive only Join Ack messages
 - with current proposals
 - intra-domain: all routers must know which is the RP for the multicast group
 - inter-domain: all routers must know which is the root domain for the multicast group

RP behavior

- How to forward traffic arriving to a RP to the rest of RPs for G ?

A.

Unicast tunnel to other RPs

B.

A RP with receivers for G must send a (S,G) Join message towards each of the active RPs placing in S the address of the RP to join. Only RPs can initiate (S,G) joins.

When a Register message from a source arrives to an RP, the RP decapsulates and encapsulates the packet again using as the source address its own address, and forwards the packet as in PIM-SM.

When a receiver receives a packet sent to the multicast group G, it decapsulates the packet and process it as in PIM-SM.

Intermediate routers' behavior

- If a router with Join state for G receives a Join message for multicast group G, and the list of RPs in the message is not equal to the list of RPs at the router,
 - the router must query DNS and build a new list of RPs.
 - If the new list is not equal to the one at the router, the router must prune the current up-branch and send a new Join message based on the new list.
- A new message is defined, the Join Probe message, to detect when an RP marked as down becomes reachable.
 - Routers should send Join Probes to next hops (excluding the one through which the normal Join is sent) towards RPs nearer than the currently joined RP.

3. Static Multicast and CBT/OCBT

- location of the Core in CBT

```
ch1.bbc.com.  A      255.192.2.3
              CORE  london-station.bbc.com.
```

- location of Cores in OCBT

```
ch1.bbc.com.  A      255.192.2.3
              OCBT  0    london-station.bbc.com.
              OCBT  1    wales-station.bbc.com.
```

– the integer number indicates the Core level