

# Single Border RBridge Nickname for TRILL Multilevel

draft-zhang-trill-multilevel-single-nickname-00.txt

Mingui Zhang, Donald Eastlake, Radia Perlman

[zhangmingui@huawei.com](mailto:zhangmingui@huawei.com)

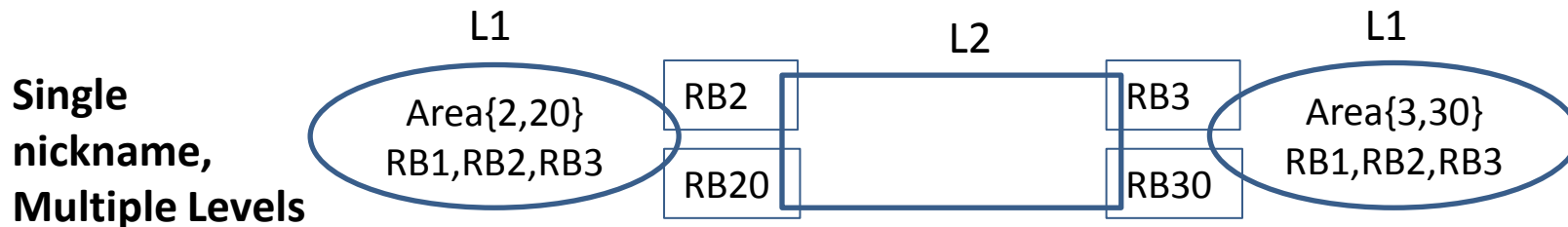
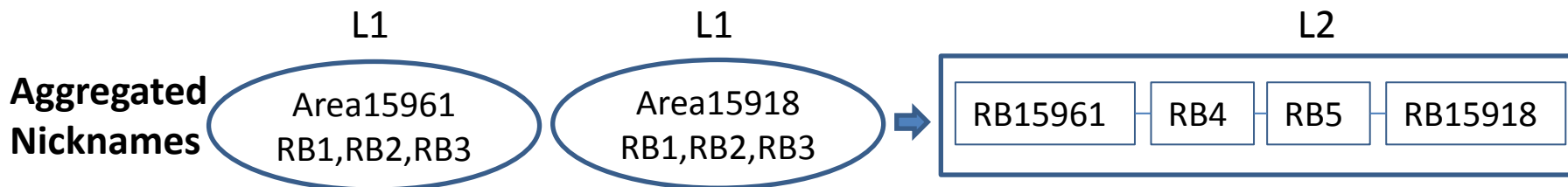
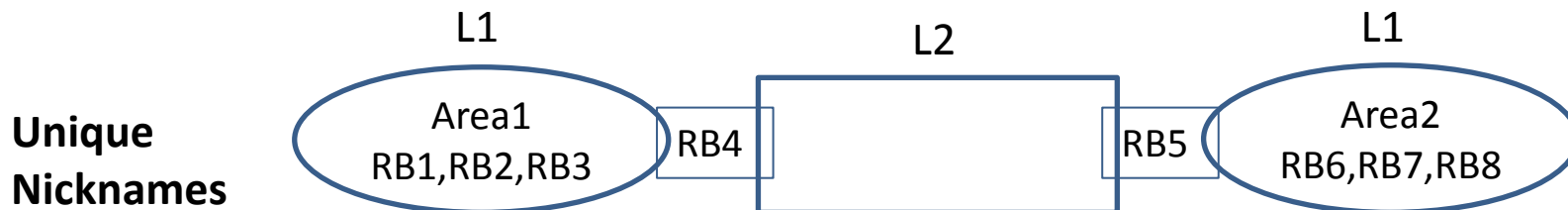
# Why multilevel

- The link state database (LSDB) is split between the Level 1 (L1) areas connected by a Level 2 (L2) area so:
  - Each LSDB part is smaller, has less control traffic, and has fewer topology changes.
  - Optimized routing computation for  $n$  RBridges is reduced from  $O(n \cdot \log(n))$  to  $O(\sqrt{n} \cdot \log(n))$ .
  - Can be specified so as to permit re-use of TRILL nicknames.

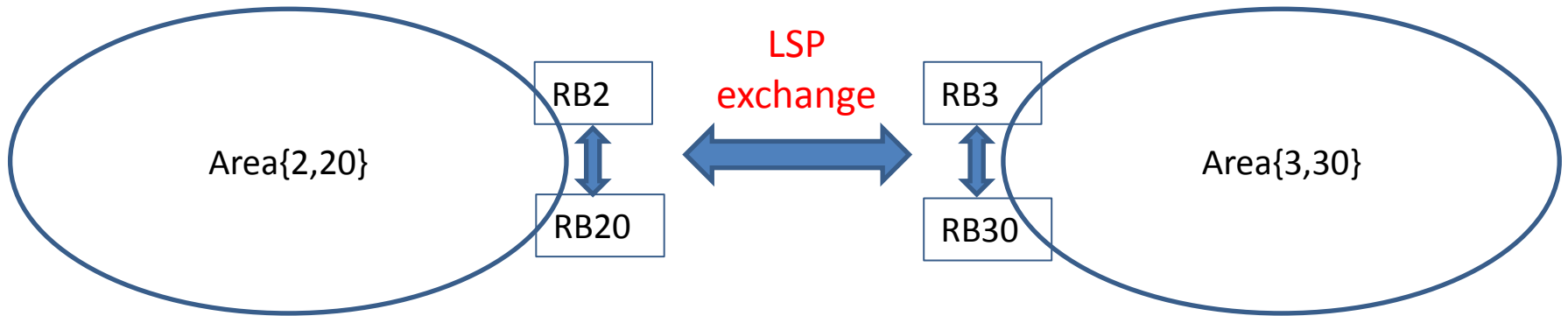
# TRILL multilevel

- A major issue of multilevel is on how to manage RBridge nicknames. Alternatives:
  - Unique nickname
    - Give unique nicknames to all RBridges in all L1/L2 areas.
  - Aggregated nickname
    - Assign nicknames independently in each L1 area.
    - In L2, represent a whole L1 area with one nickname.
  - Single nickname, multiple levels [this document]
    - The border RBridge nickname is used in both L1, L2
    - Other nicknames in each L1 area are assigned independently.
    - In L2, represent an L1 area with the list of border nicknames.

# Examples

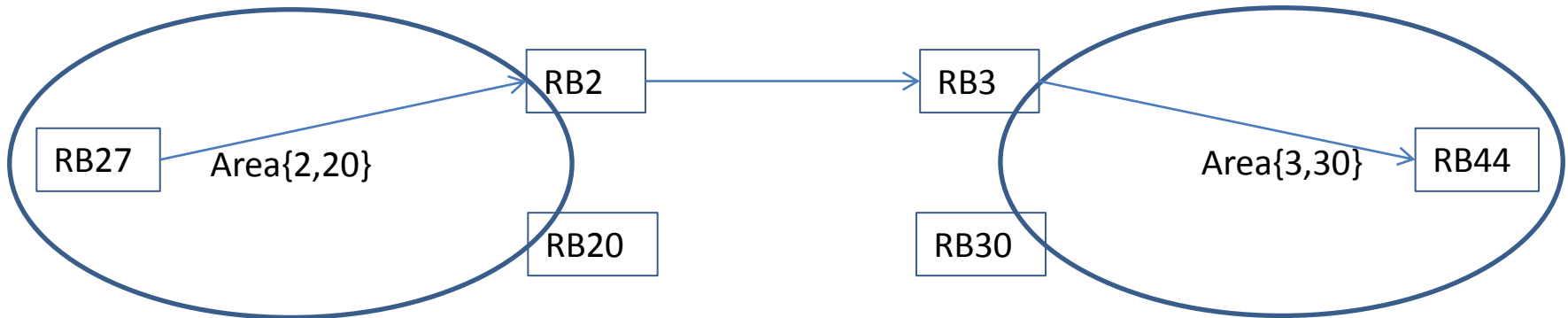


# Area discovery



- In L1, border RBridges RB2, RB20 discover each other via L1 LSP exchanging.
- In L2, an area is represented by the list of border RBridges. The list of border RBridges is announced in L2 LSPs.
- Each border RBridge is aware of which RBridges will be used as border RBridges for an area.

# Actions on unicast forwarding



RB27	egress=RB3	ingress=RB27	MACdst=D	MACsrc=S
RB2	egress= <b>RB3/RB30</b>	ingress= <b>RB2</b>	MACdst=D	MACsrc=S
RB3	egress= <b>RB44</b>	ingress= <b>RB2/RB20</b>	MACdst=D	MACsrc=S

- Border RBridge may replace the egress/ingress nickname of the TRILL Data packets.

# Actions on unicast forwarding

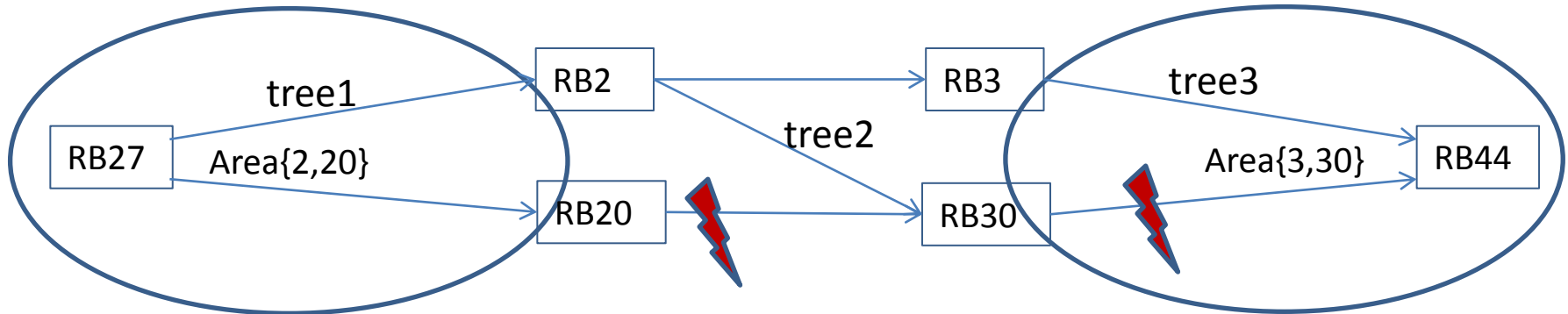
- Note:
  - The address of a TRILL Data packet is the Inner.MacDA and Inner VLAN or FGL. This address never changes from the initial ingress RBridge to the final egress RBridge. The change in ingress / egress nickname at border RBridges is a little like label swapping.

# MAC synchronization

- Since the return traffic may go through any border RBridge, all of them need to learn the MAC.
- The MAC info is synchronized among area border RBridges using L1 ESADI.



# Actions on multicast forwarding



RB27	M-bit=1	egress=RB27(root)	ingress=RB27	MACdst=?	MACsrc=S
------	---------	-------------------	--------------	----------	----------

RB2	M-bit=1	egress= <b>RB2(root)</b>	ingress=RB2	MACdst=?	MACsrc=S
-----	---------	--------------------------	-------------	----------	----------

**RB20** L1 -> L2. Drops the packet to avoid duplication.

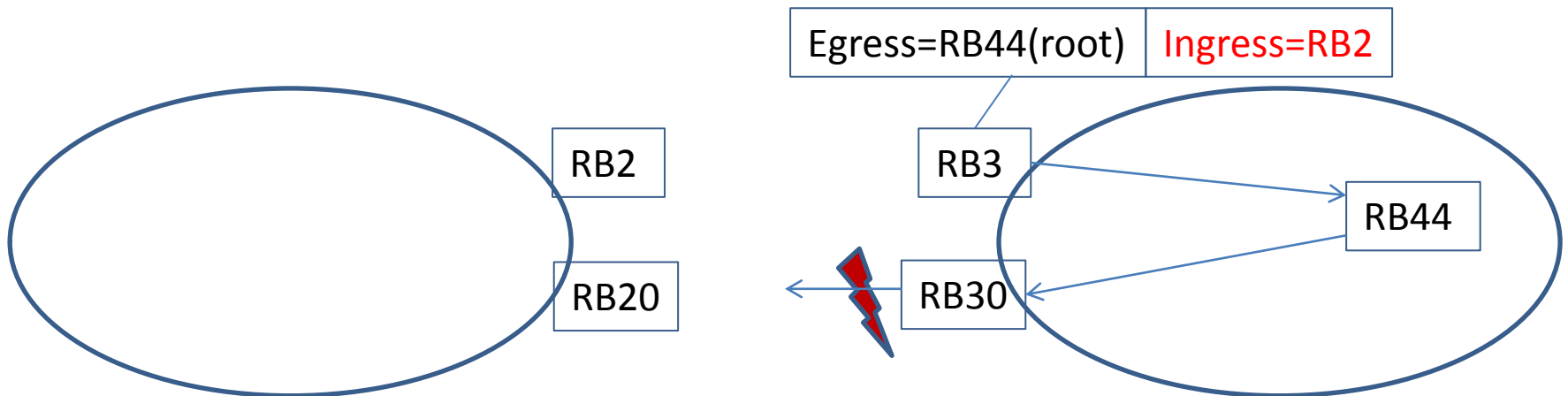
RB3	M-bit=1	egress= <b>RB44(root)</b>	ingress= <b>RB2/RB20</b>	MACdst=?	MACsrc=S
-----	---------	---------------------------	--------------------------	----------	----------

**RB30** L2 -> L1. Drops the packet to avoid duplication.

- Designate an border RBridge for multicast forwarding, according to a pseudorandom algorithm.
- Only the DBRB (Designated Border RBridge) can do the transition between L1 and L2.
- This avoids packet duplication.

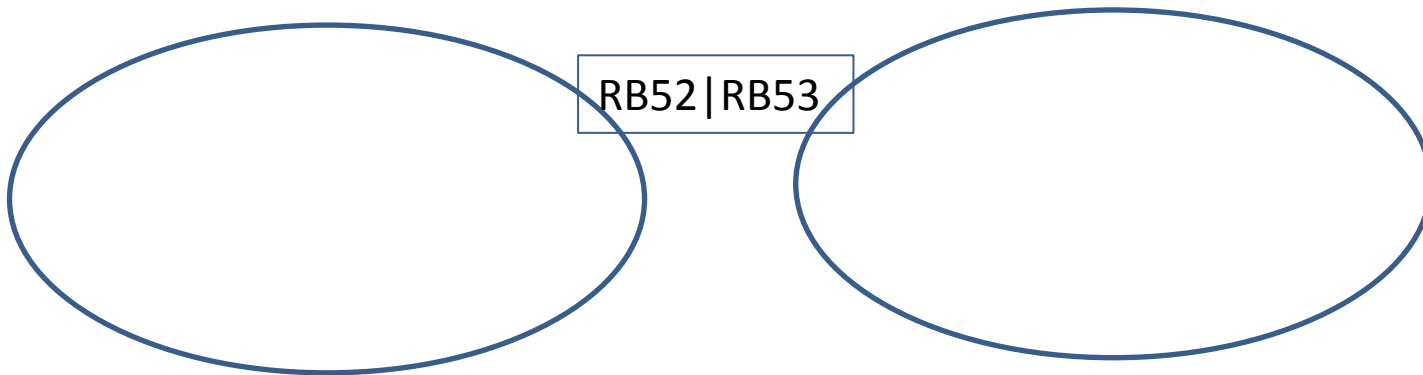
# Multicast scoping

- Based on the ingress nickname, a multi-destination packet with an L2 ingress nickname **MUST NOT** be forwarded back into L2.
- Otherwise, forwarding loops will occur.



# Border nickname per area

- If one border RBridge is connected to multiple areas, this border RBridge needs to obtain nicknames per area.



# Advantages

- Fewer nicknames used
- Less configuration, and less ability to screw up the configuration
- No Reverse Path Forwarding Check (RPFC) issue since the border RBridge uses its own nickname
- Adequate path split on multiple trees

Please read the draft and comment!

Thanks!