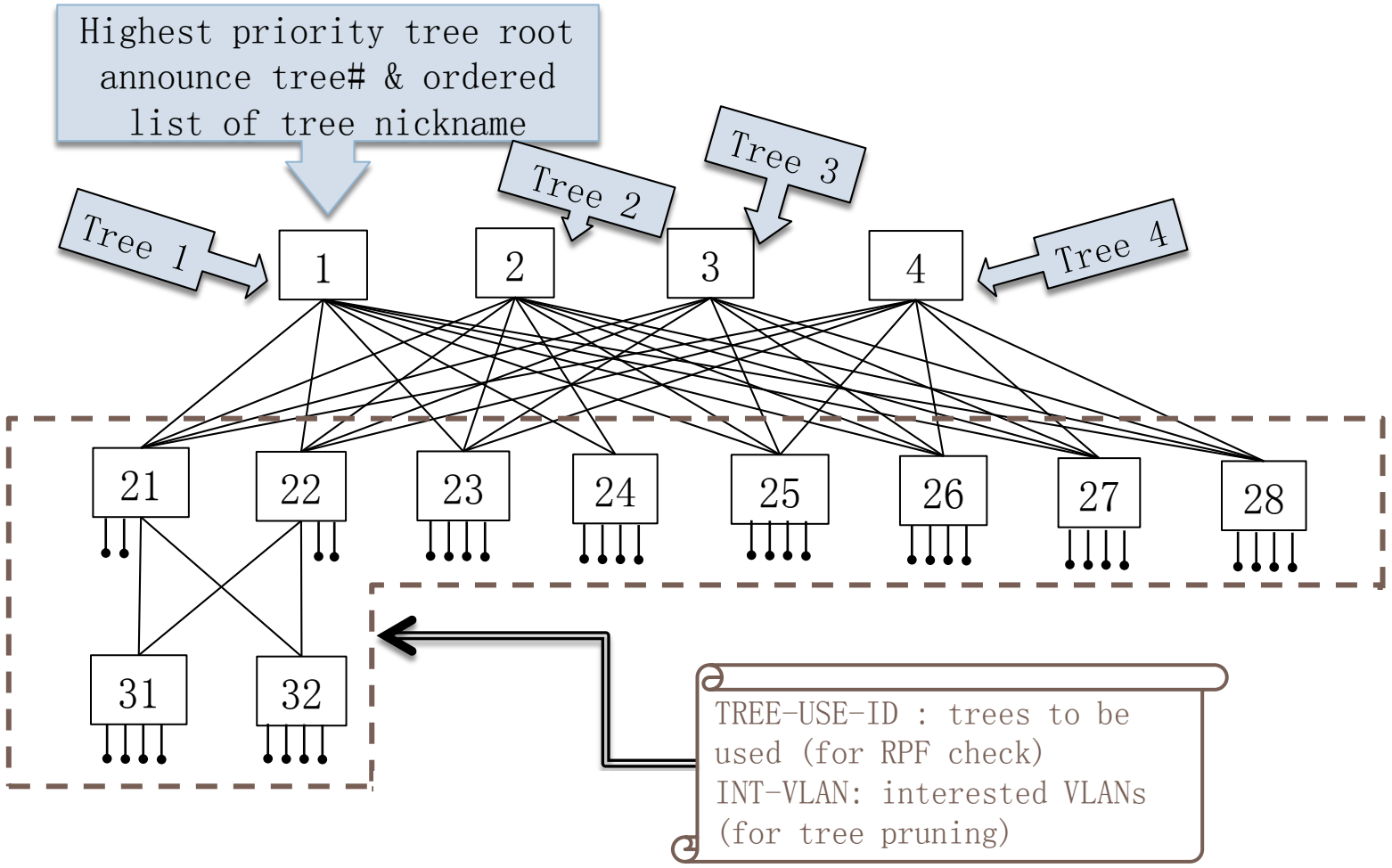
VLAN BASED TREE SELECTION FOR MULTI-DESTINATION

draft-yizhou-trill-tree-selection-00

Yizhou Li (liyizhou@huawei.com)
Weiguo Hao (haoweiguo@huawei.com)
Somnath Chatterjee (somnath.chatterjee01@gmail.com)

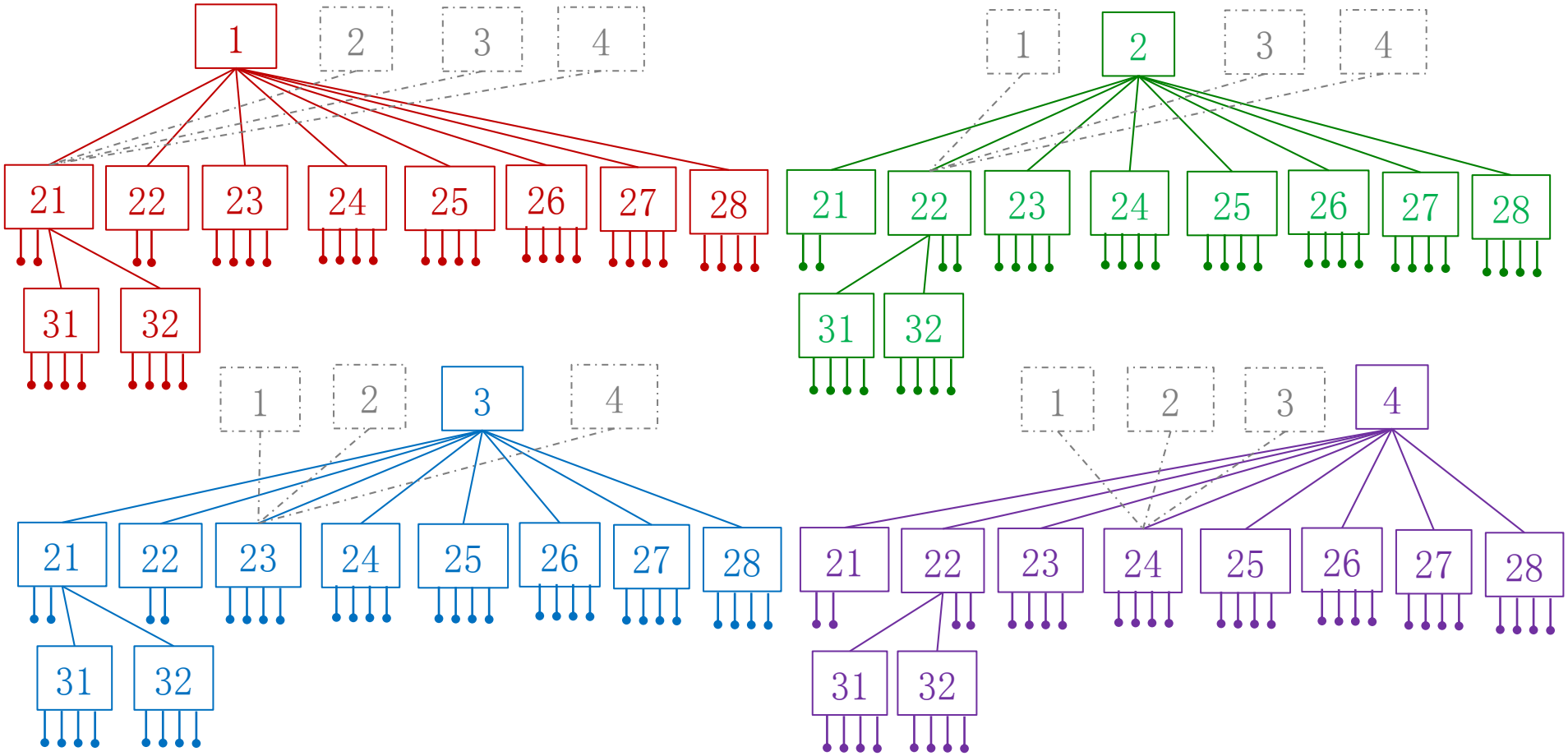
Background (1)

- Each distribution tree SHOULD be pruned per VLAN



Background (2)

- 4 trees are built



Motivations (1)

- Multicast forwarding table on RB21 has 16K entries.

Tree nickname	VLAN	Port list
tree1	1	1, 10, 20, local port 50, local port 60
tree1	…	…
tree1	4095	1, 10, 20, local port 50, local port 60
tree2	1	2, local port 50, local port 60
tree2	…	…
tree2	4095	2, local port 50, local port 60
tree3	1	3, 10, 20, local port 50, local port 60
tree3	…	…
tree3	4095	3, 10, 20, local port 50, local port 60
tree4	1	4, local port 50, local port 60
tree4	…	…
tree4	4095	4, local port 50, local port 60

4K entries for tree1

4K entries for tree2

4K entries for tree3

4K entries for tree4

Motivations (2)

- Table size == n*m entries. (n is #of trees, m is #of VLANs with downstream receivers)
- More entries required if L2/L3 multicast address to be used for finer pruning
  - n*m where m is #of VLANs with downstream receivers * #of multicast groups per VLAN
  - Value of m may exceed 10K in theory
- Linearly increasing with #of trees
- Table size is limited. May share a 8K/16K-entry table with IP multicast/VSI forwarding entries.
- Propose: VLAN based tree selection to reduce the table size
  - still allows the traffic sharing among trees

VLAN based Tree Selection

- Concept:
  - highest priority tree root announces tree-VLAN correspondence which is the value pair of (tree id, VLANs allowed on this tree id)
  - ingress RB selects the tree-VLAN correspondence it is interested in and wishes to use from the list.
  - It should not transmit VLAN x frame on tree y if the highest priority tree root does not say VLAN x is allowed on tree y.
  - Achieved VLAN based load balancing by selecting different trees

VLAN based Tree Selection Example

- If we let the highest priority tree root announces:
  - (tree1, Vlan 1-1000)
  - (tree2, Vlan 1001-2000)
  - (tree3, Vlan 2001-3000)
  - (tree4, Vlan 3001-4095)
- Ingress selects and announces (tree id, interested vlan) from the announced tree-VLAN correspondence
- Multicast table entries are reduced to 4K (maximum).
- Table size shrunk:
  - n * m → m

Tree nickname	VLAN	Port list
tree1	1	
tree1	…	
tree1	1000	
tree2	1001	
tree2	…	
tree2	2000	
tree3	2001	
tree3	…	
tree3	3000	
tree4	3001	
tree4	…	
tree4	4095	

1K entries for tree1

1K entries for tree2

1K entries for tree3

1K entries for tree4

Other issues

- Compatibility:
  - New and old RB announces tree-used-id and interested VLAN per RFC6325
  - New RB additionally announces (tree id, interested-vlan) which must be value combinations allowed by the highest priority tree root
  - Always able to tell new or old RB, thus build the table correspondingly in new or old way
- Failure handling
  - Failure of a tree root
  - Failure of the highest priority tree root
- Extensions
  - extended to (VLAN+L2/L3 multicast group) based tree selection

Next Step

- TLV format for (VLAN + L2/L3 multicast address) based tree selection
- More detailed compatibility consideration