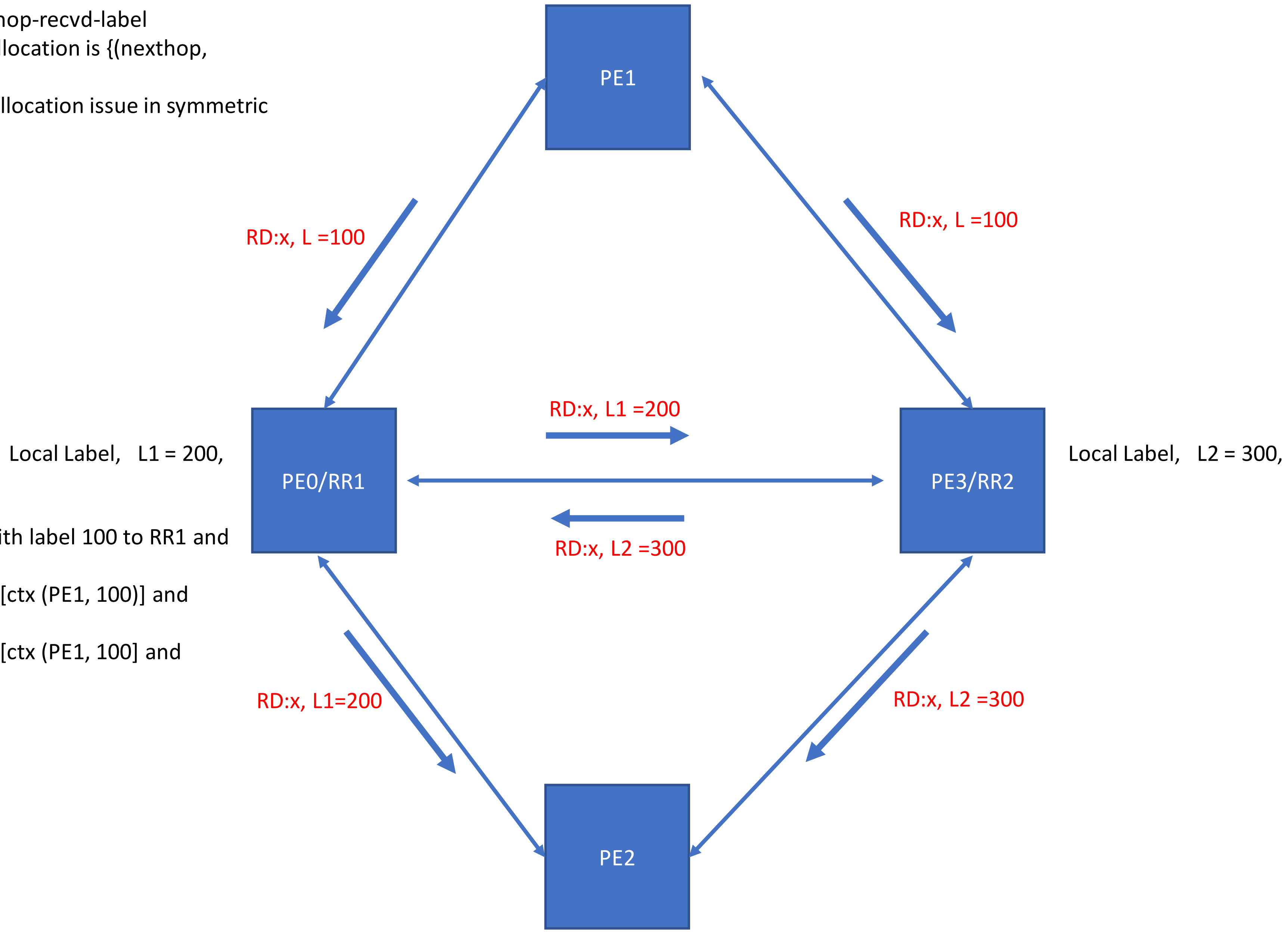
The Secondary Label and its applications
draft-mohanty-idr-secondary-label-00

Satya R Mohanty satyamoh@cisco.com Cisco Systems
Israel Means im8327@att.com AT&T
Praveen Ramadenu pr9637@att.com AT&T
IETF 117, IDR Network Working Group, July 2023

- Various Label Allocation Mode offered by vendors (default VRF labels)
  - **per prefix**
  - **per vrf**
  - **per next-hop-received-label.**
- Per next-hop-received-label
  - Label Allocation **context** for this mode is set of {(nhop, recvd_label)} tuples
  - Useful to conserve labels at the RR with next-hop-self or ASBR

Problem Description #1

1. Label mode used is per-nexthop-recvd-label
2. Context (ctx) for local label allocation is {(nexthop, received label)} tuples
3. Leads to a continuous label allocation issue in symmetric PIC deployments

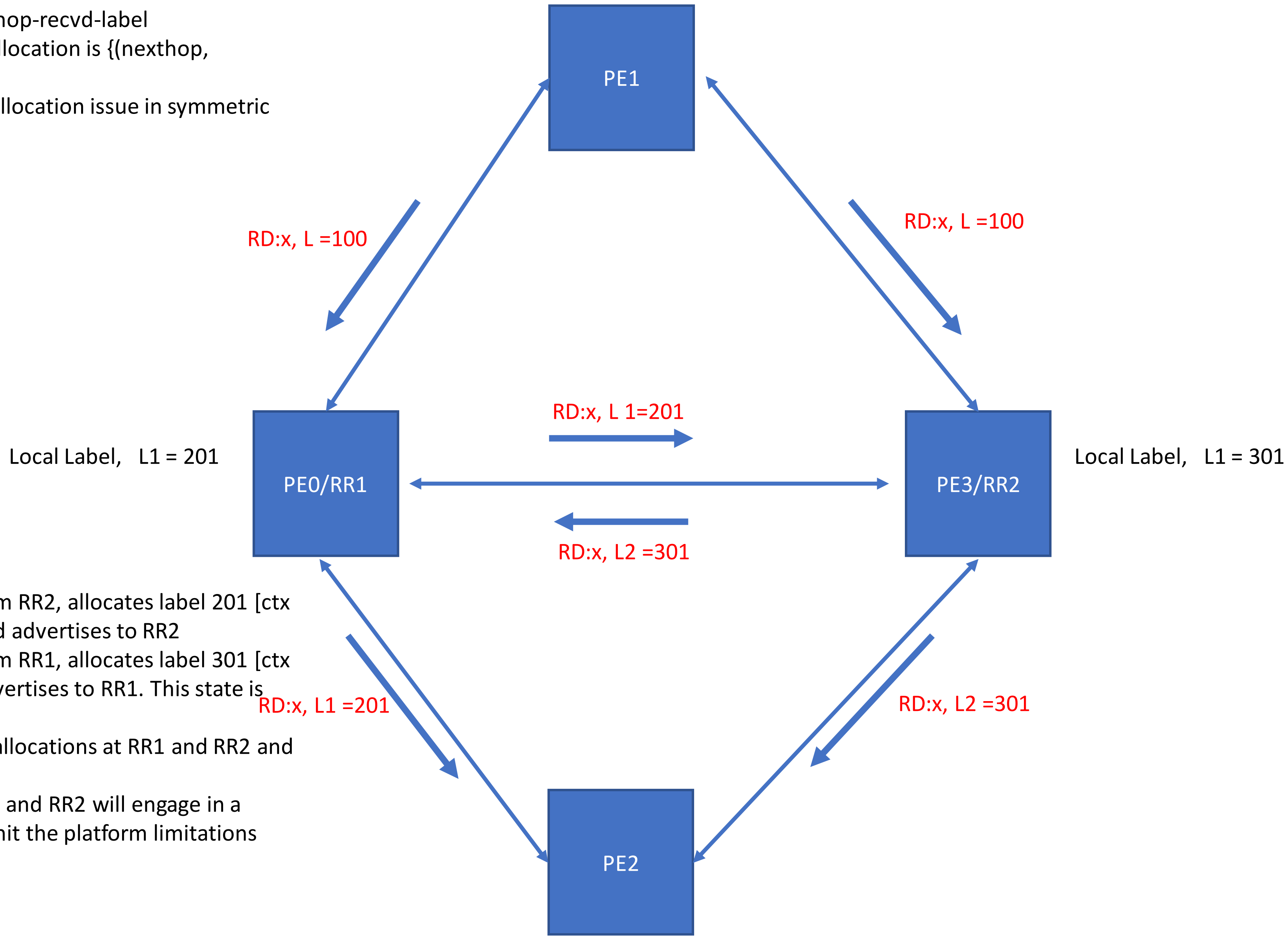


Description

1. PE1 advertises VPN routes with label 100 to RR1 and RR2.
2. RR1 allocates local label 200 [ctx (PE1, 100)] and advertises to RR2.
3. RR2 allocates local label 300 [ctx (PE1, 100] and advertises to RR1.

Problem Description #1 (cont'd)

1. Label mode used is per-nexthop-recvd-label
2. Context (ctx) for local label allocation is {(nexthop, received label)} tuples
3. Leads to a continuous label allocation issue in symmetric PIC deployments



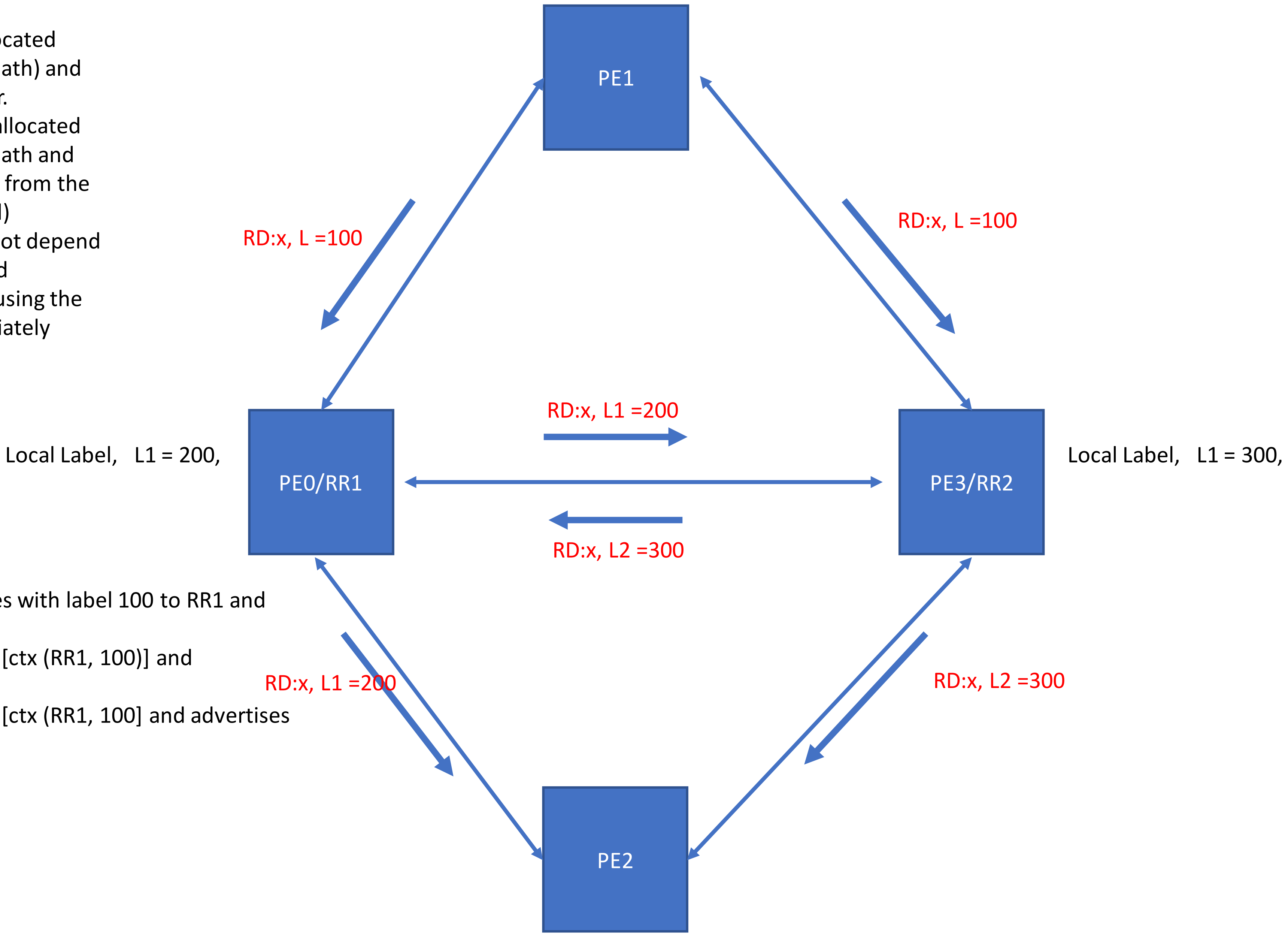
Description

1. RR1 receives the update from RR2, allocates label 201 [ctx {(PE1, 100), (RR2, 300) }] and advertises to RR2
2. RR2 receives the update from RR1, allocates label 301 [ctx {(PE1, 100), (RR1, 200) }] advertises to RR1. This state is shown alongside.
3. This will result in new label allocations at RR1 and RR2 and this process will go on
4. Net result is the devices RR1 and RR2 will engage in a continuous label churn and hit the platform limitations soon

Proposed Solution

High Level Idea

1. A secondary label will be allocated (based only on the primary path) and advertised to the peer router.
2. The usual local label will be allocated (based only on the primary path and the secondary label received from the peer router if one is received)
3. Since secondary label does not depend on the peer router’s local and secondary label, the loop causing the label churn is broken immediately



Solution Walkthrough

1. PE1 is sending the VPN routes with label 100 to RR1 and RR2.
2. RR1 allocates local label 200 [ctx (RR1, 100)] and advertises to RR2
3. RR2 allocates local label 300 [ctx (RR1, 100] and advertises to RR1.

Proposed Solution

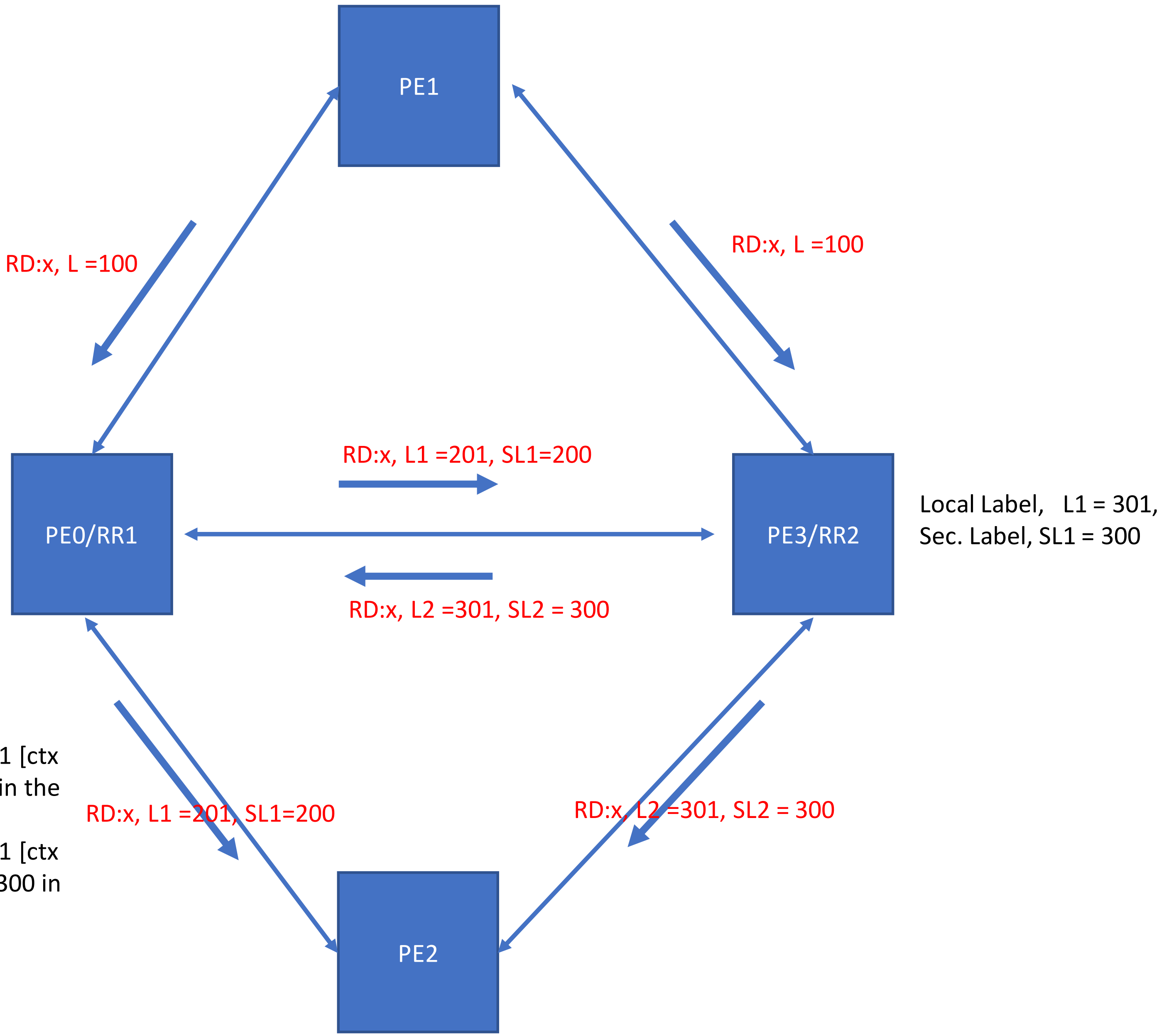
High Level Idea

1. A secondary label will be allocated (based only on the primary path) and advertised to the peer router.
2. The usual local label will be allocated (based only on the primary path and the secondary label received from the peer router if one is received)
3. Since secondary label does not depend on the peer router's local and secondary label, the loop causing the label churn is broken immediately

Solution Walkthrough

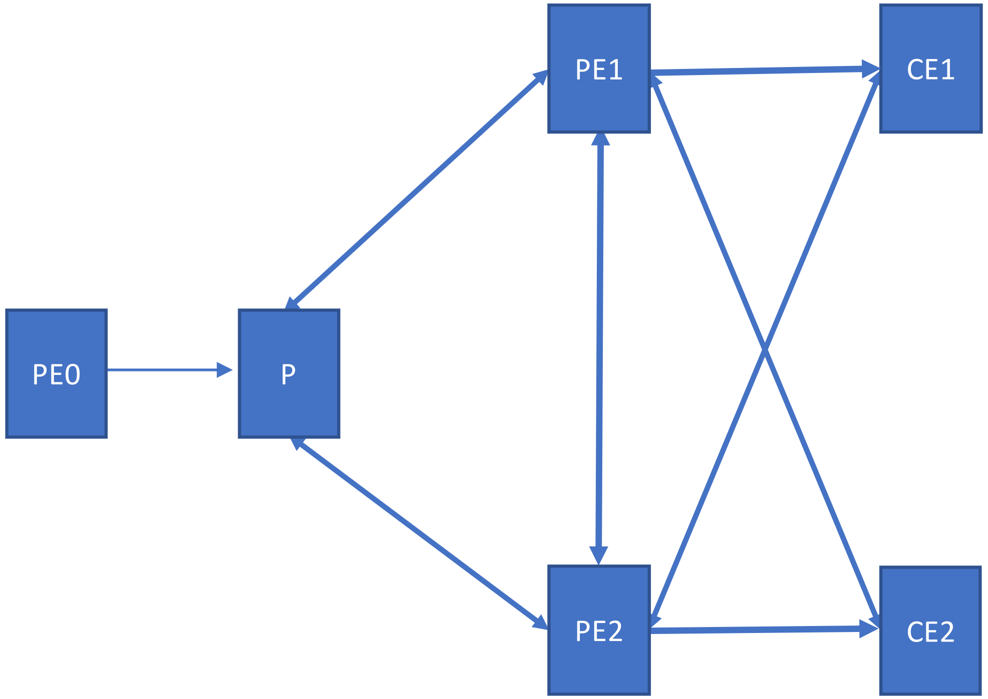
Solution Walkthrough (cont'd)

3. RR1 receives the update from RR2, allocates label 201 [ctx {(PE1, 100), (RR2, 300) }] advertises to RR1 with 200 in the secondary label attribute
4. RR2 receives the update from RR1, allocates label 301 [ctx {(R-leaf-1, 100), (RR1, 200) }] advertises to RR1 with 300 in the secondary label attribute as before.
5. There is no loop from this point on



Problem Description #2

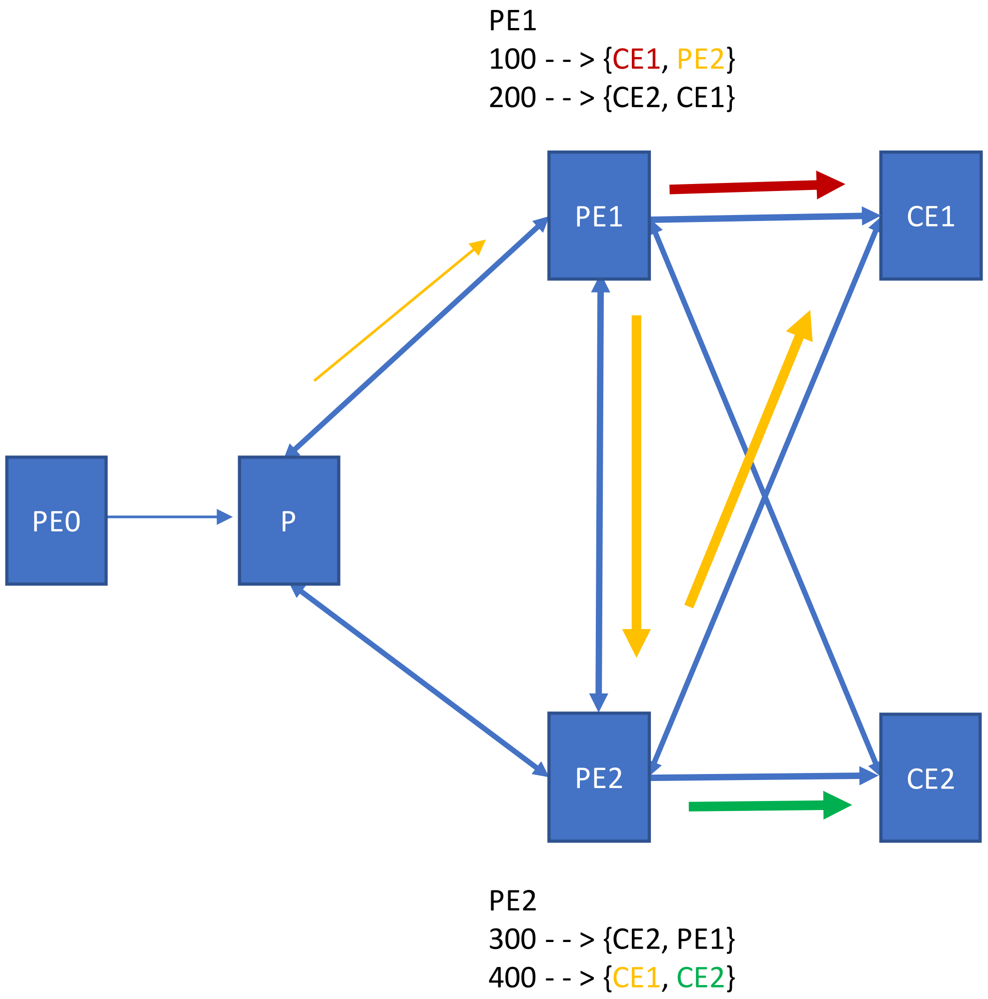
1. CE1 is always preferred to CE2 as long as it has connectivity to at least one of the PEs
2. PE1 will send traffic to CE1 as long as PE1 CE1 link is up
3. If PE1 CE1 link is down, PE1 should send the traffic to PE2
4. PE2 should send the traffic to CE1 if PE2 CE1 link is up.
5. PE2 should send the traffic to CE2 only if PE2 CE1 link is down



Problem Solution #2

High Level Idea

- Allocate a Primary label with the primary path pointing to the directly connected preferred CE (best EBGP path) and the secondary to the less preferred PE (IBGP Path).
- Allocate a 2nd label with primary path to directly connected preferred CE (best EBGP path) and secondary path to the other directly connected CE (less preferred EBGP path). This second label is advertised in the Control Plane along with the primary label



Solution Walkthrough

- PE1allocates a primary label of 100 that points to the primary next-hop CE1 and to the secondary NH, PE2; and, a secondary label of 200 (pointing to primary NH CE1 and secondary NH CE2).
- Similarly, PE2 allocates a primary label of 300 (this primary label points to primary NH CE1 and secondary NH PE1) and a secondary label of 400 (pointing to primary NH CE1 and secondary NH CE2)
- Traffic from the remote PEs always uses the primary label. Traffic sent from one peer PE to another is always sent using the secondary label.
- Therefore, traffic to 10.10.0.0/24 from PE0 is received on PE1 with label 100. In the normal case, this traffic will be sent on the direct PE1 CE1 link.
- If link PE1-CE1 breaks, this traffic is diverted to PE2 with label 400.When this traffic is received by PE2, if the PE2-CE1 link is up, traffic will be forwarded to CE1 on that link. But, now if the PE2-CE1 goes down, the secondary path for the label 400 which points to the NH CE2, is activated immediately and the traffic is directed to CE2 on the PE2 CE2 link

Secondary Label Attribute

- New Optional Transitive Attribute needs to be created for carrying the secondary label
- Type field to define the context of the secondary label

Attr Flags	Attr Code
Length	Flags
Type1	Label L1
Type2	Label L2

Conclusions

- Two use cases are illustrated where the secondary label concepts appears very useful
- Define as an new optional transitive attribute
- Need IANA Assignment
- Questions ?