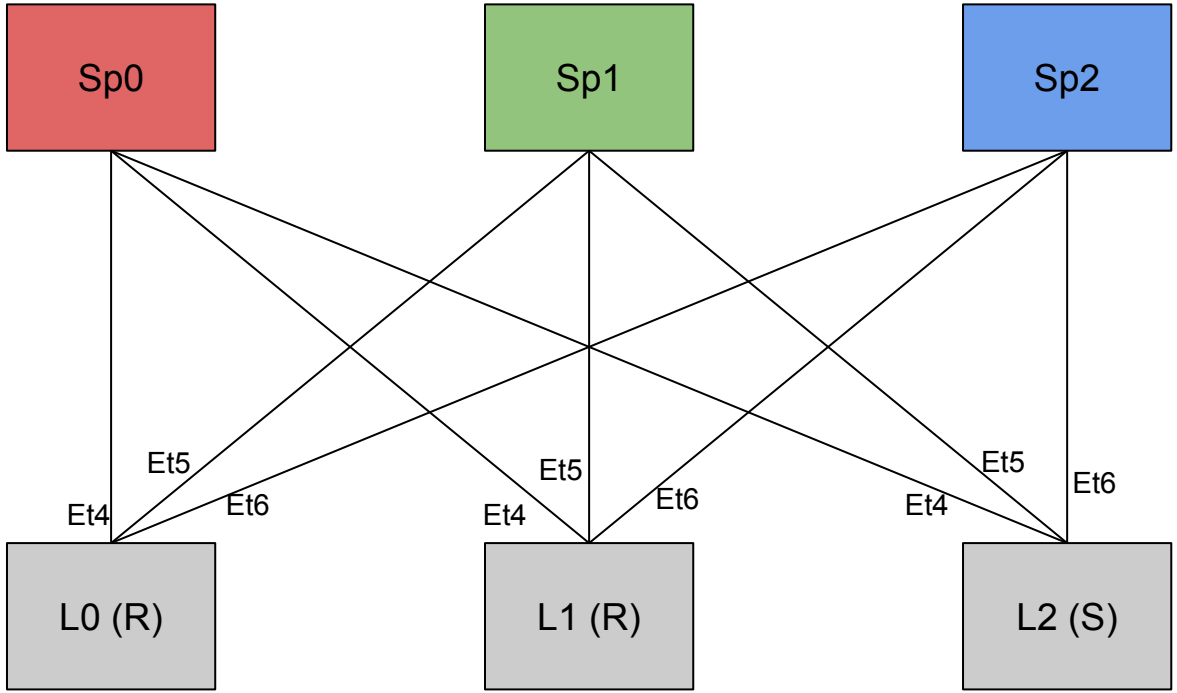
PIM Deterministic ECMP

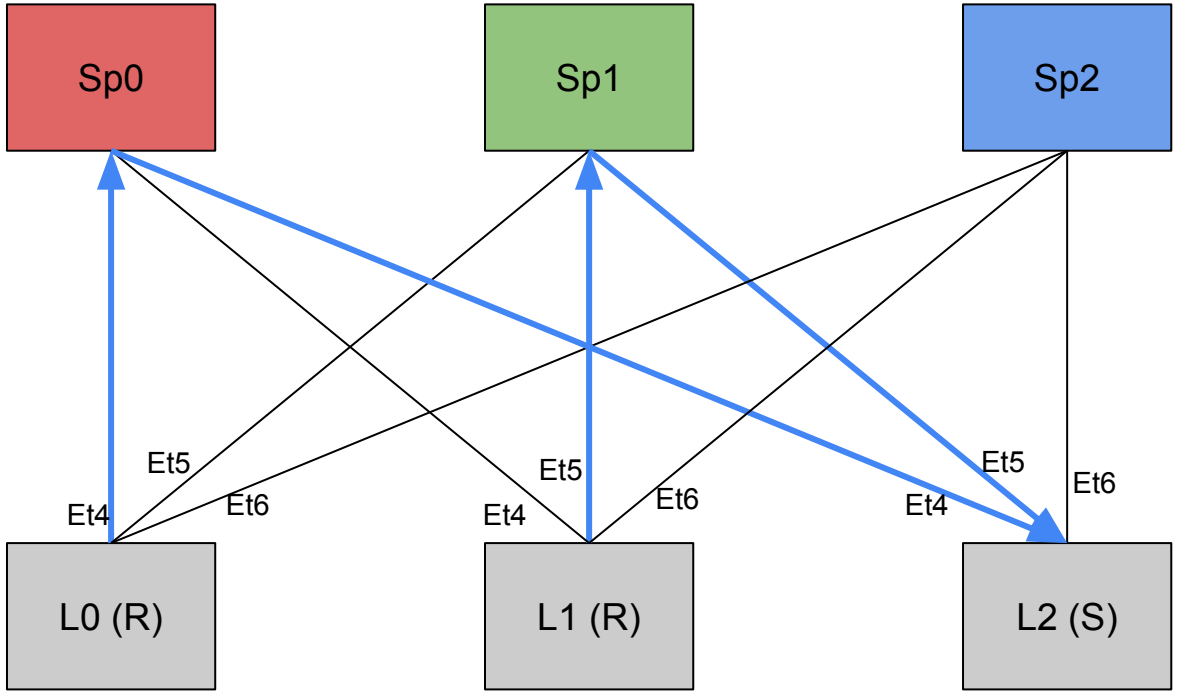
Bill Fenner

PIM WG - IETF 122 Bangkok

Deployment scenario 1: Leaf/Spine



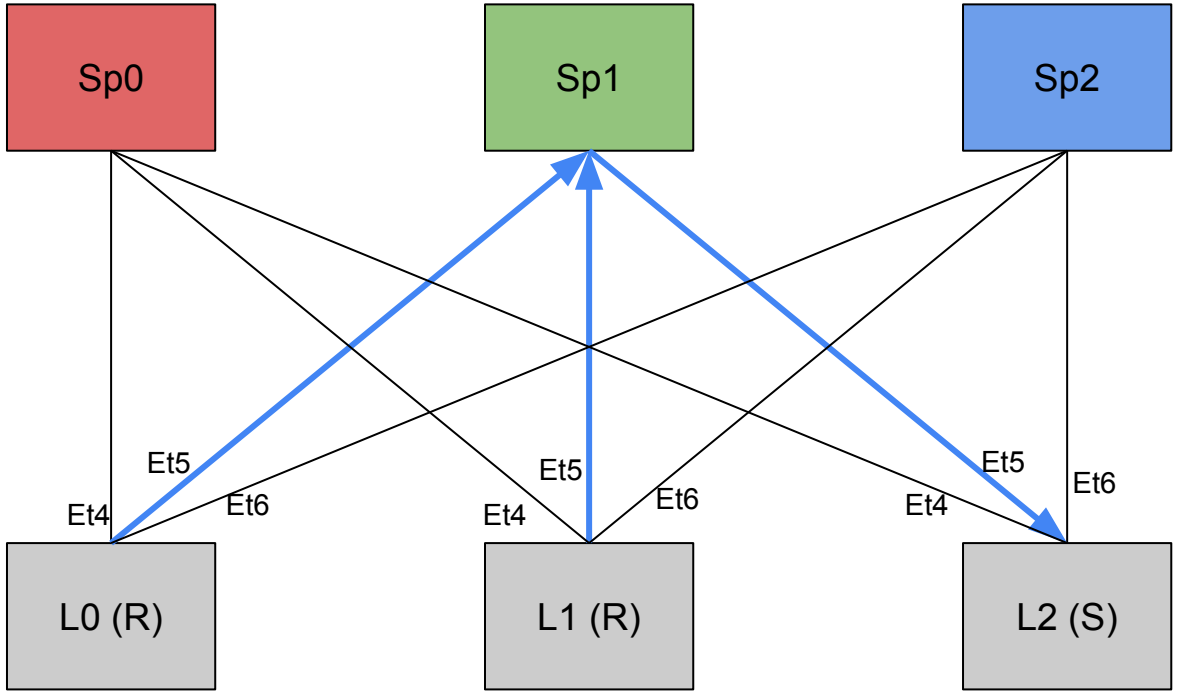
Deployment scenario 1: Traffic over multiple links from L2



L0 and L1 make independent decisions; L2 has to forward through both Sp0 and Sp1

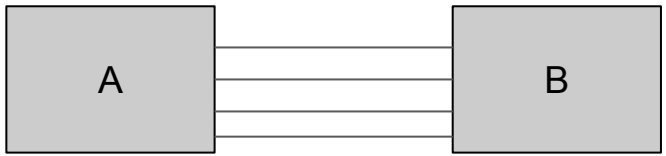
Deployment scenario 1: Leaf/Spine

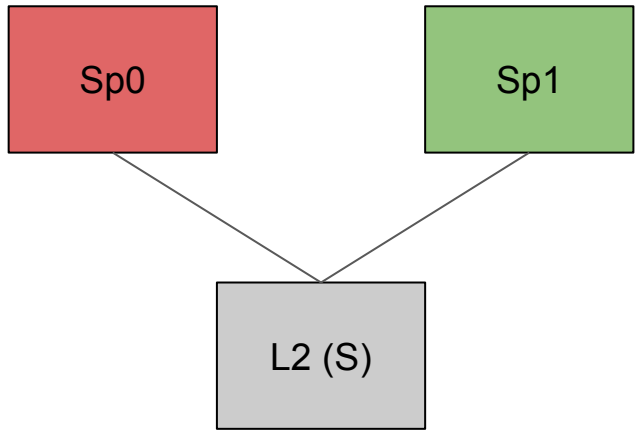
Hash based on (S, G, Router ID), so L0 and L1 make the same decision



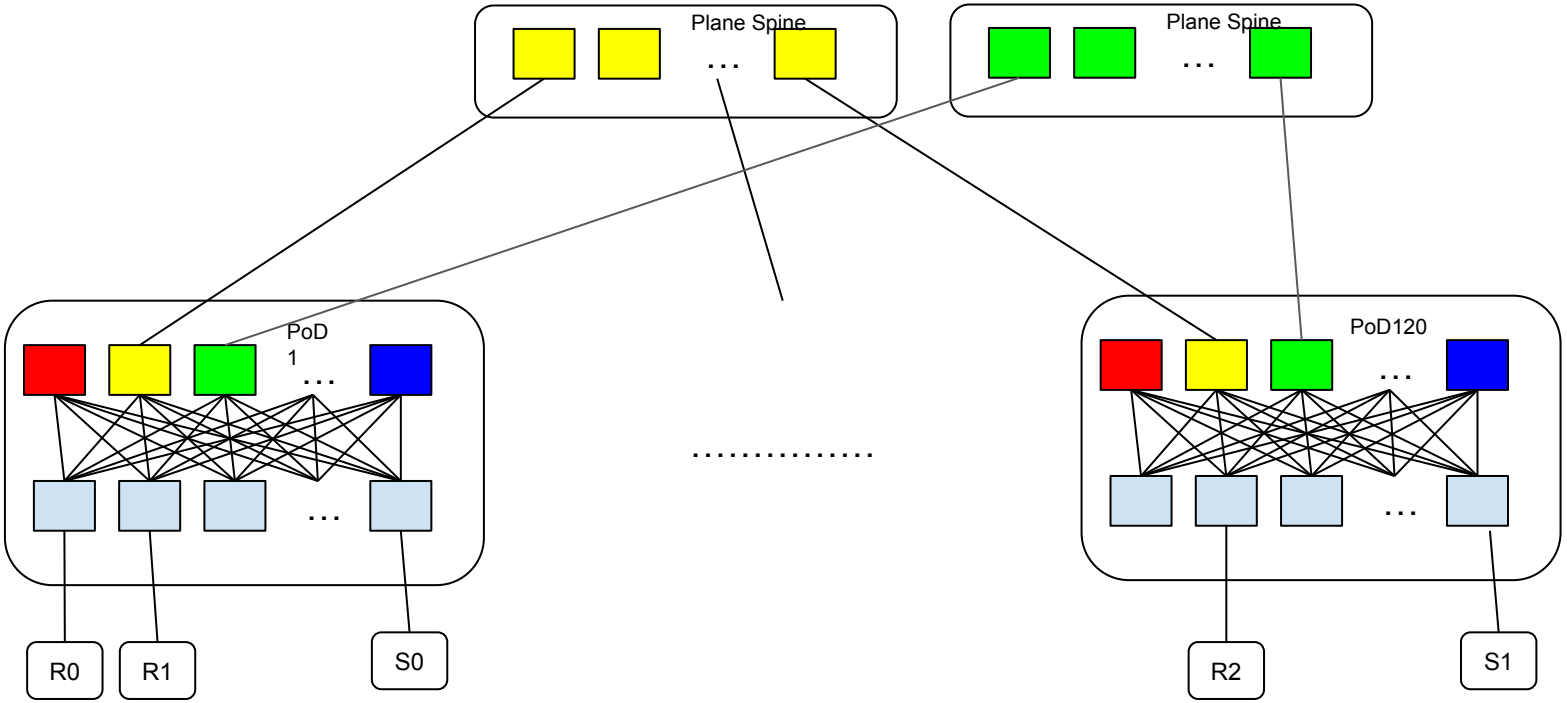
Why not use RFC6754: PIM ECMP Redirect?

PIM ECMP Redirect is designed for an ECMP group between two routers, to move traffic from a loaded path to a free path. It tells a neighbor to switch to a different link. But in this case, each neighbor only has a single link; L2 can not tell Sp1 to switch to using the Sp0 link.



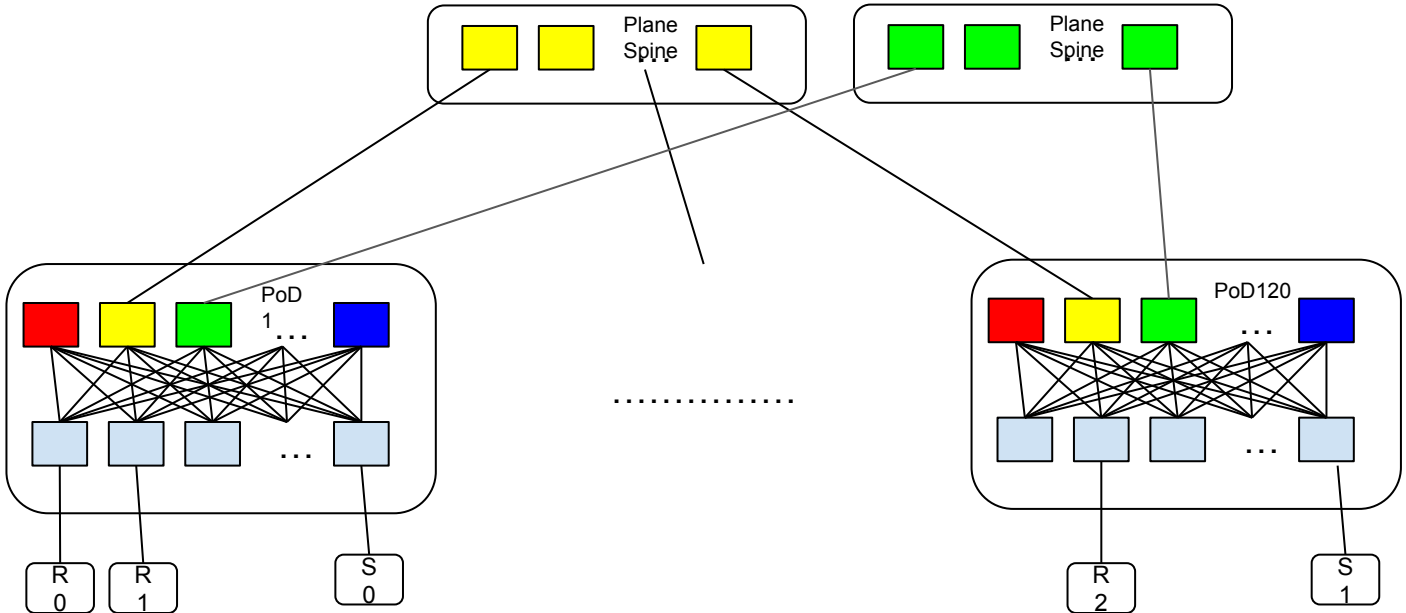


How about a multi-tier fabric?



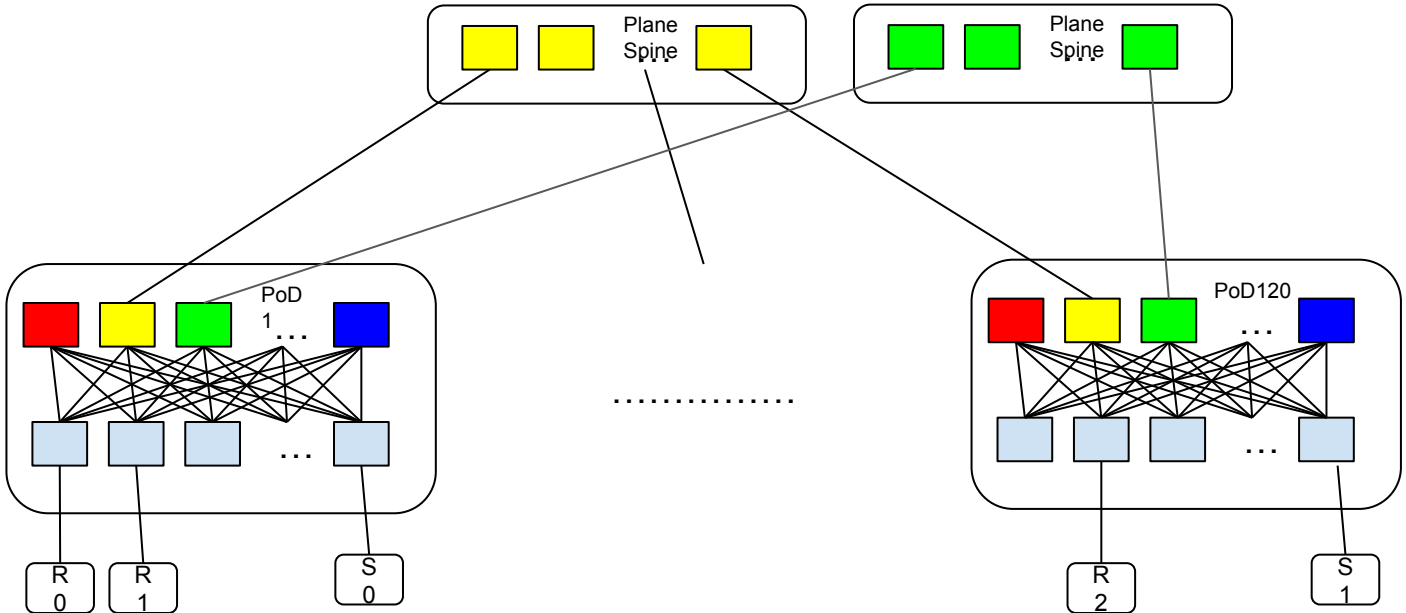
Same problem at the next layer

Router-ID based hashing works within a “Pod”, but when we get to the next layer one pod might use the “Yellow” spine and one pod might use the “Green” spine, replicating the problem at the next layer



Introduce “Color”

Spine routers inside a “Pod” advertise their “Color” to the leaves using a PIM Hello Option; leaves hash using (S, G, Color)



draft-fenner-pim-deterministic-ecmp

Introduces the problem

Describes router-ID hashing and Color-based hashing

Describes what Arista has deployed and how to interoperate

Next Steps

Is there interest in making this a WG work item, and standardizing the Color Hello option?