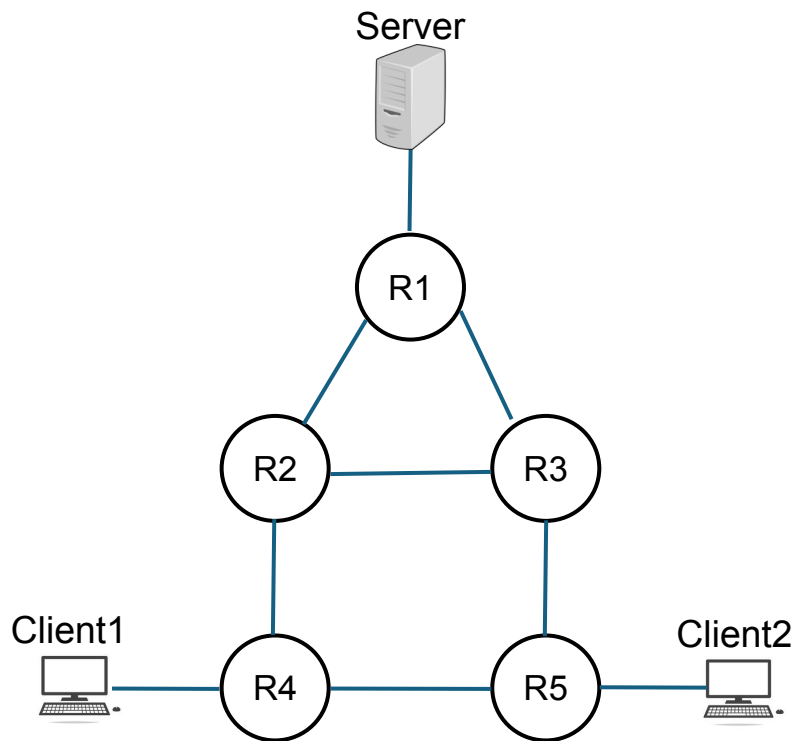


Intradomain Multicast Routing for Modern Routers

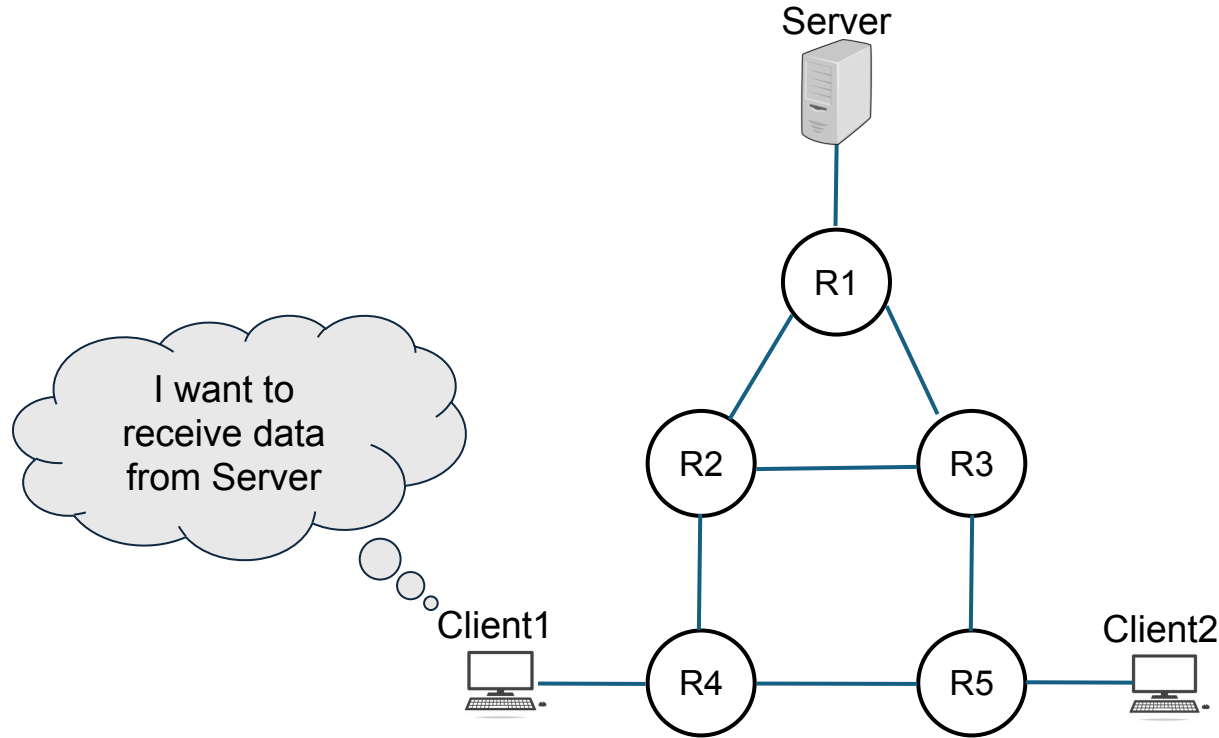
Doeraene Anthony¹ & Olivier Bonaventure²

¹ UCLouvain, ² UCLouvain & WEL-RI

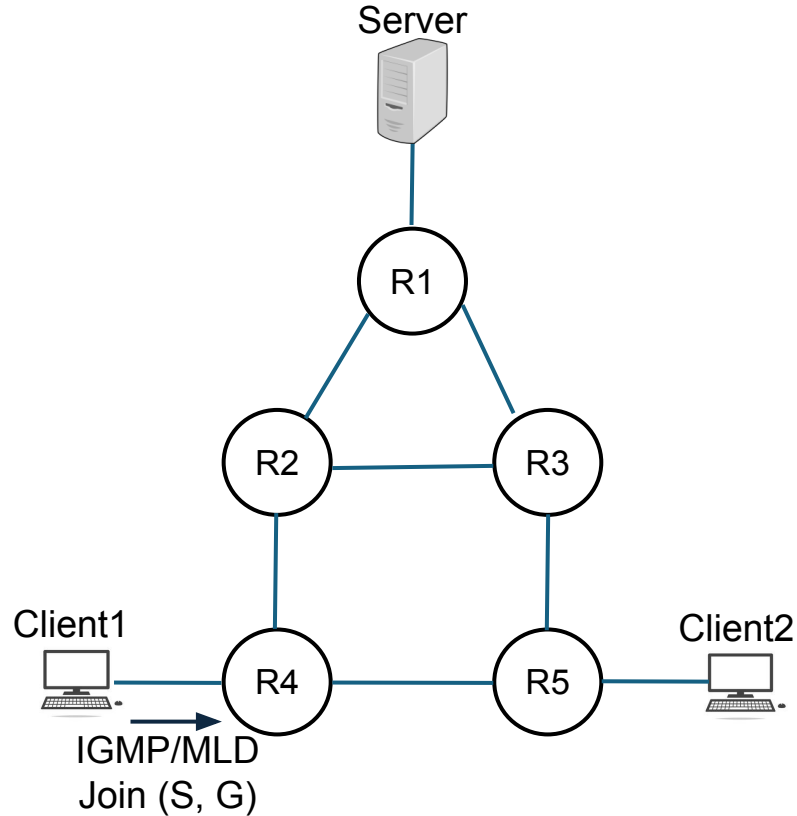
PIM basics



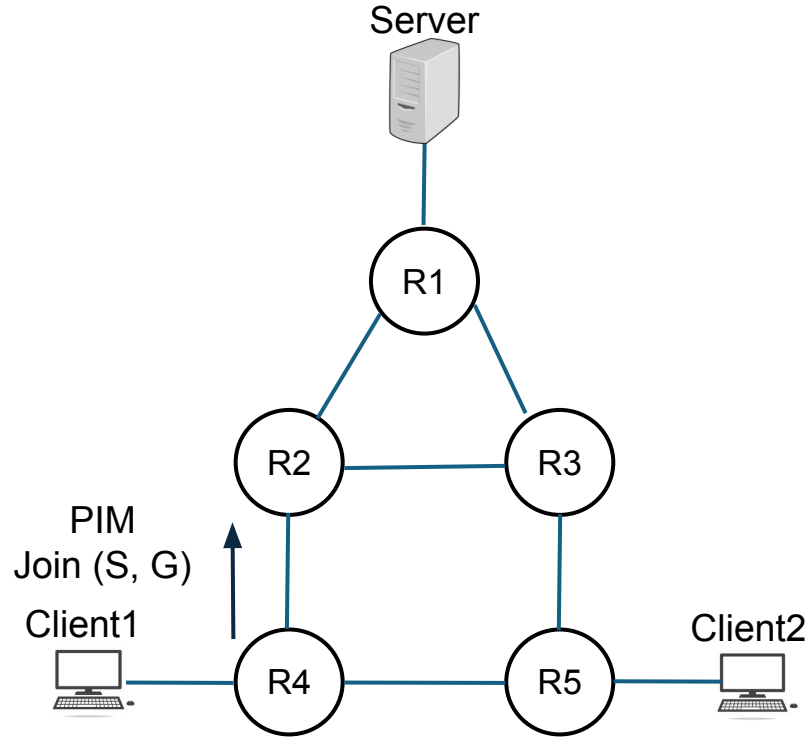
PIM basics: Joins



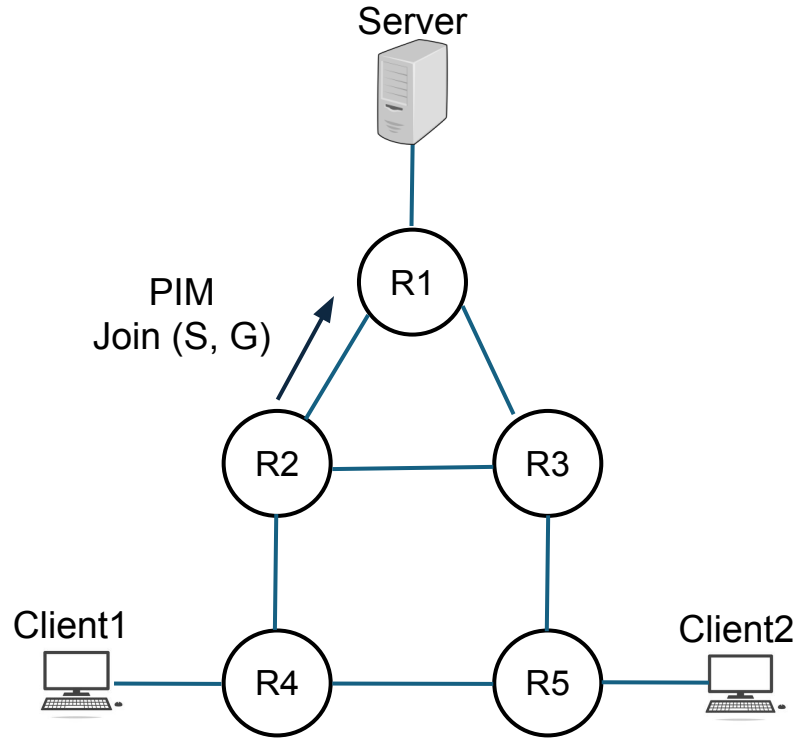
PIM basics: Joins



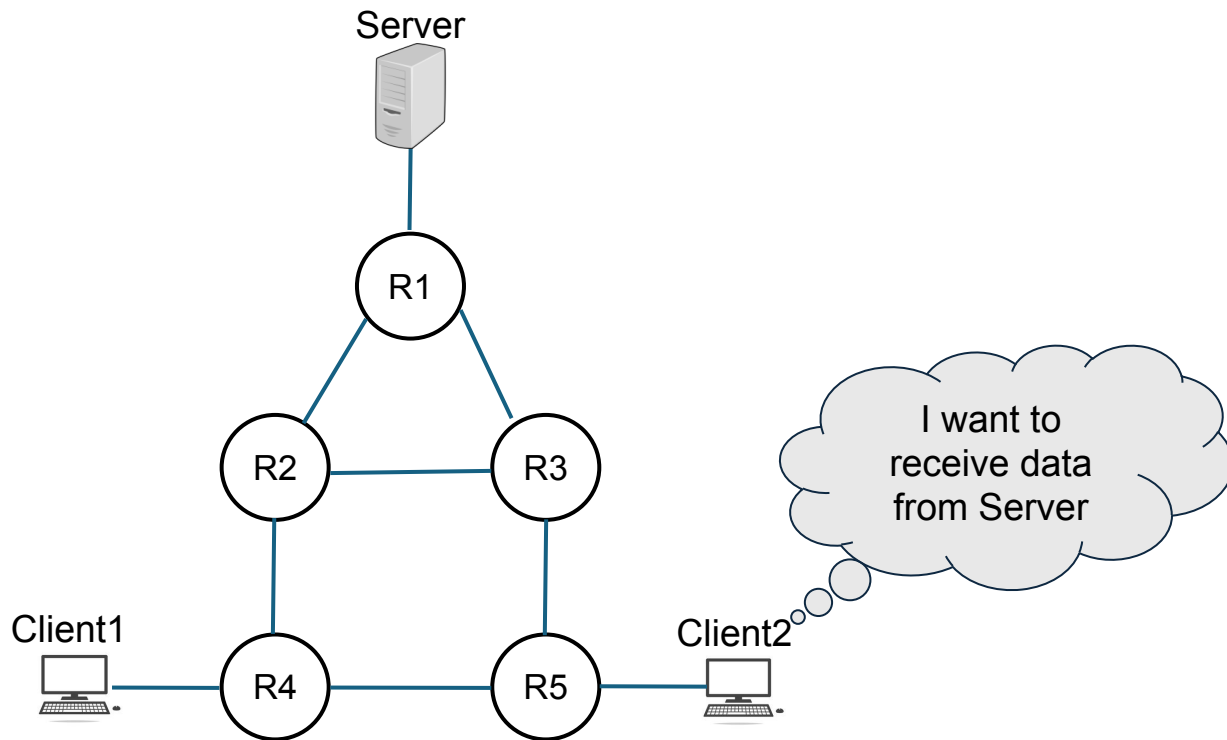
PIM basics: Joins



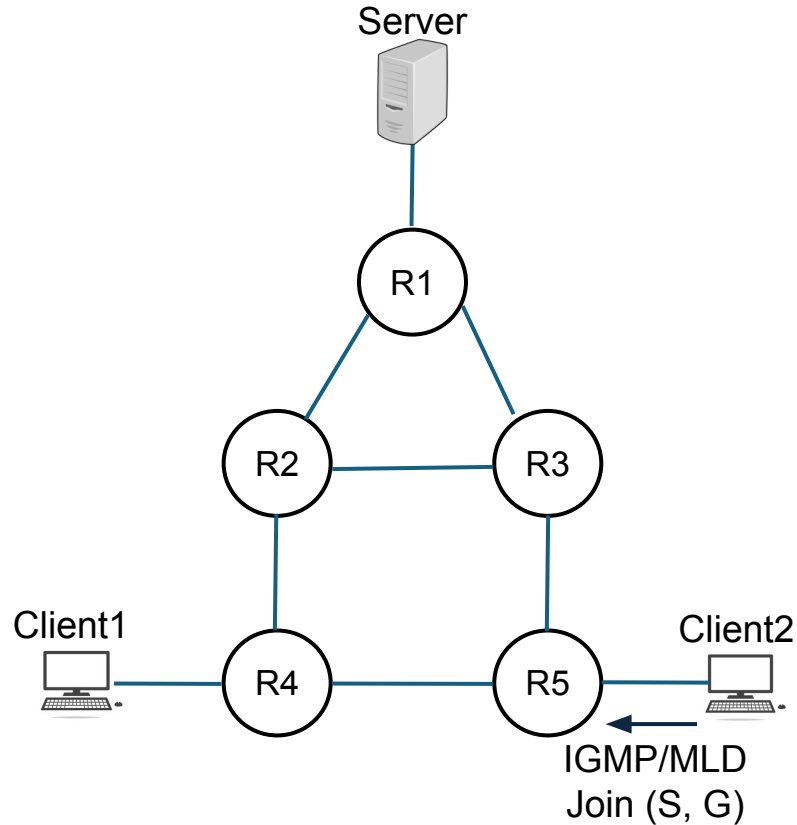
PIM basics: Joins



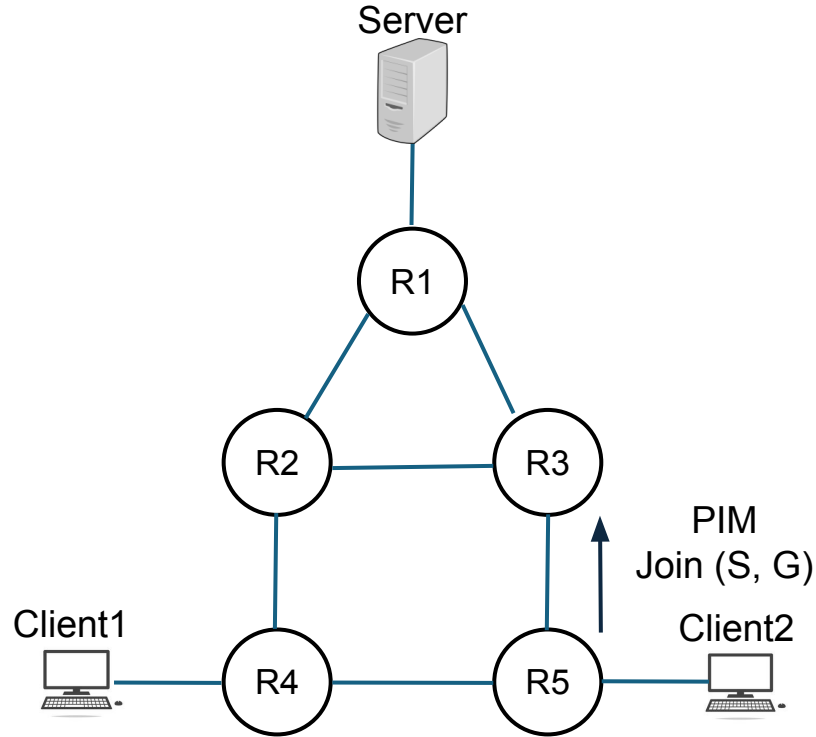
PIM basics: Joins



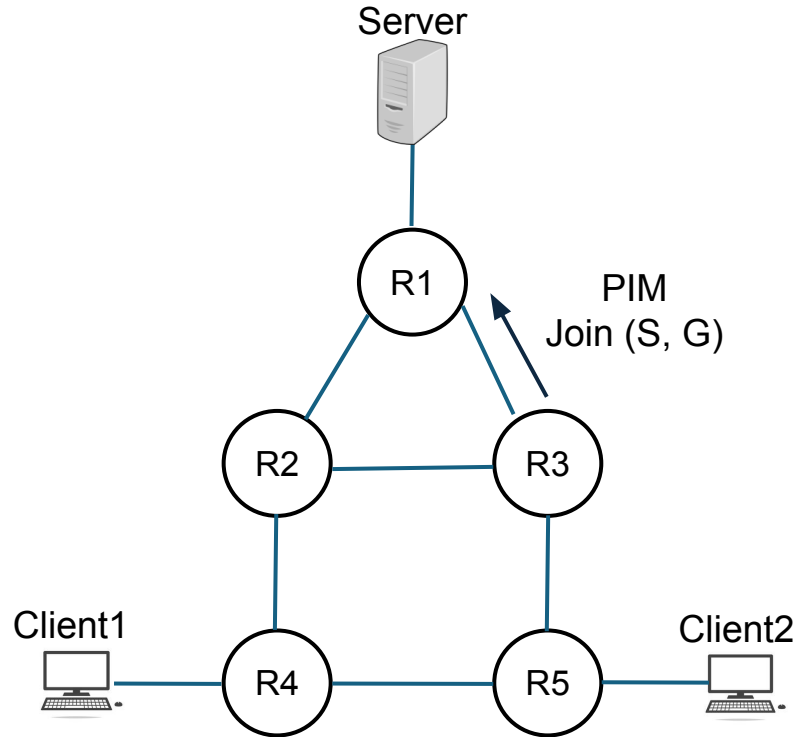
PIM basics: Joins



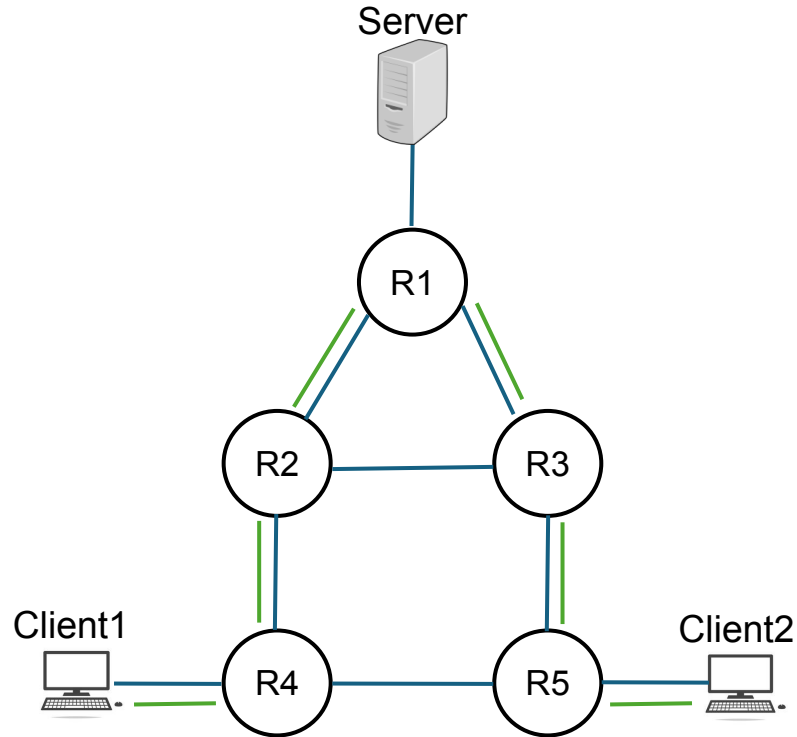
PIM basics: Joins



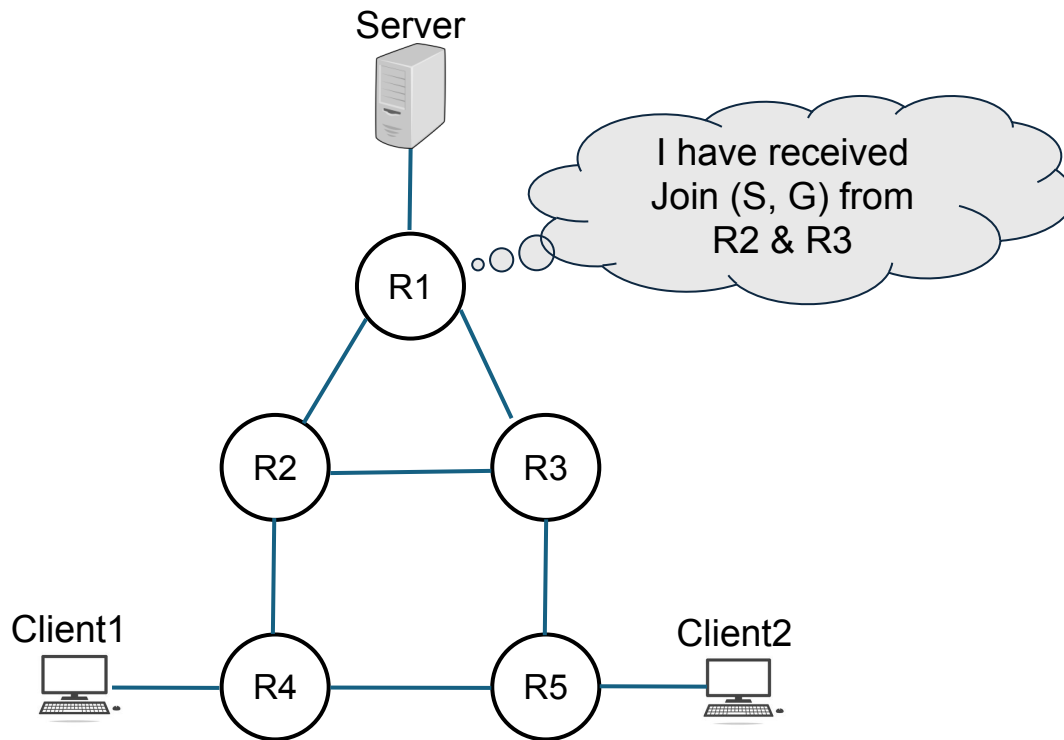
PIM basics: Joins



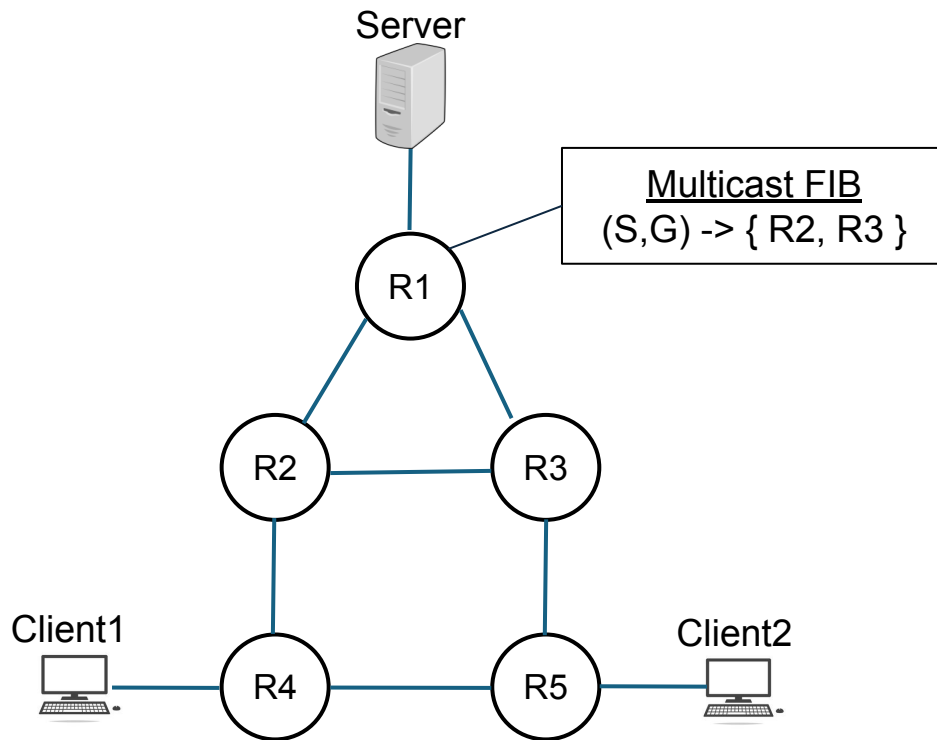
PIM basics: Tree



PIM basics: State



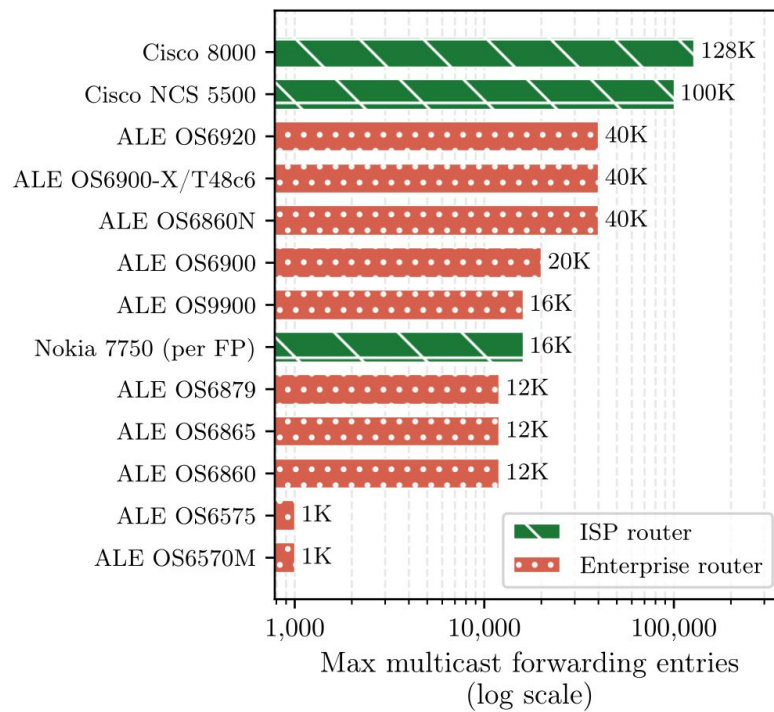
PIM basics: State



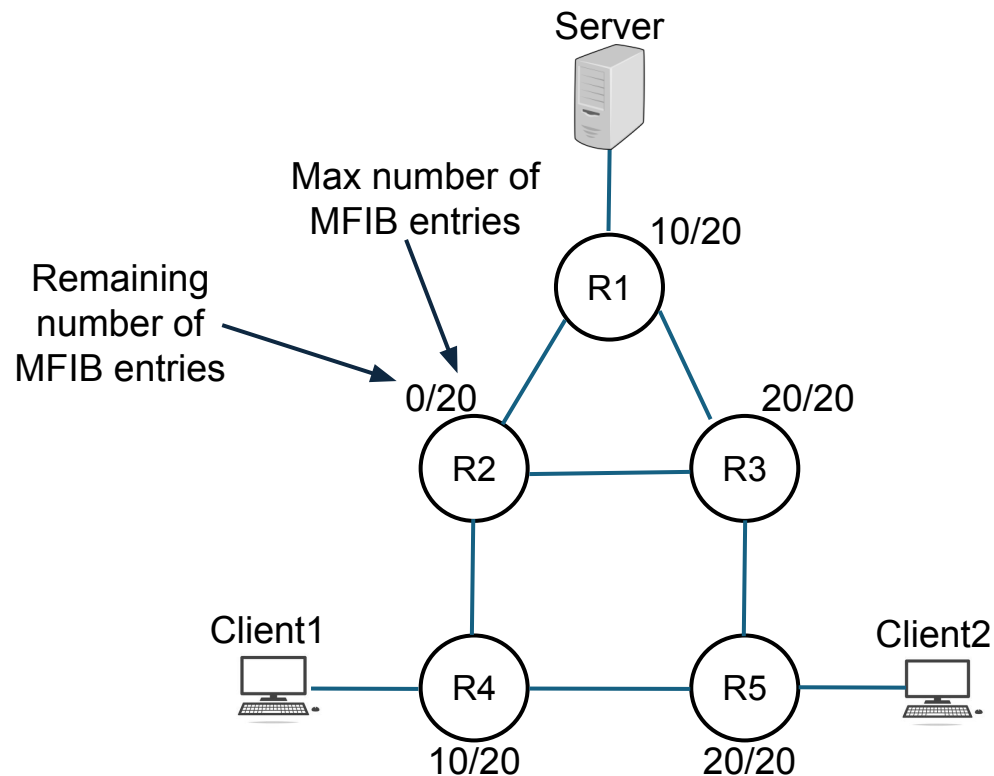
PIM problem: State consumption

Every (S,G) requires an entry in
the multicast FIB !

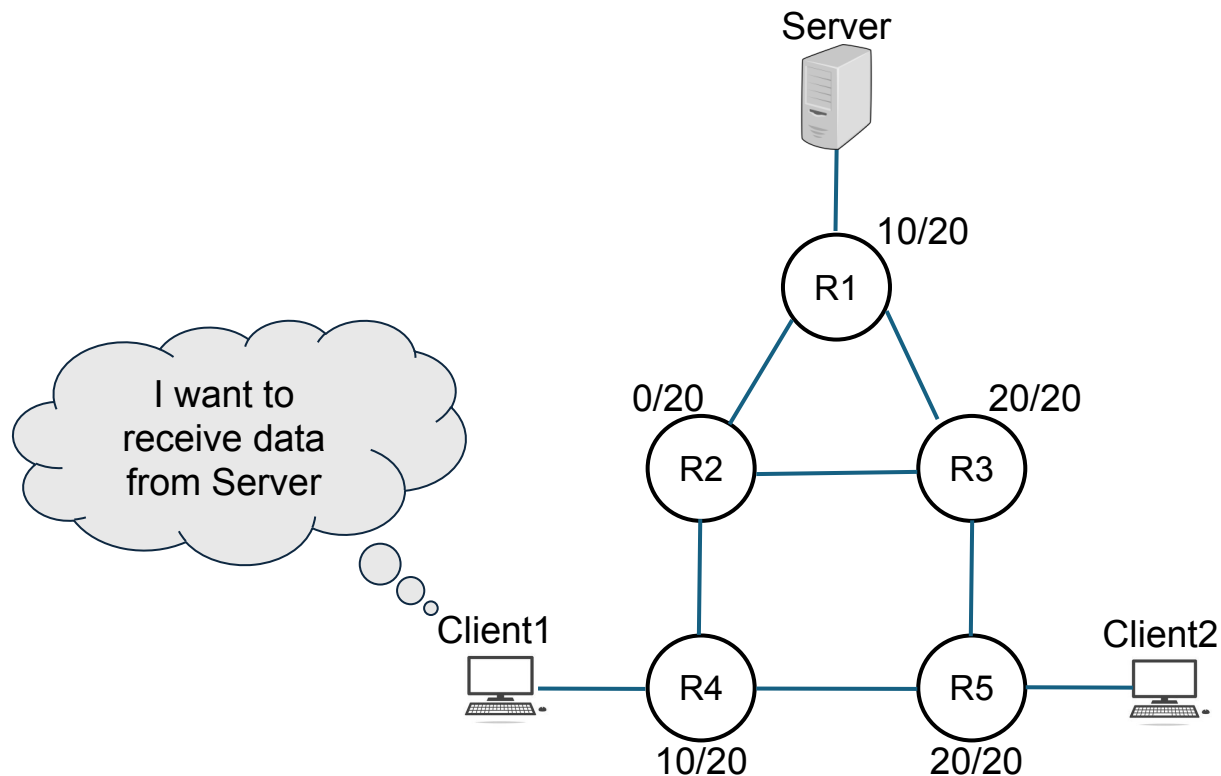
Hardware MFIB capacity limits



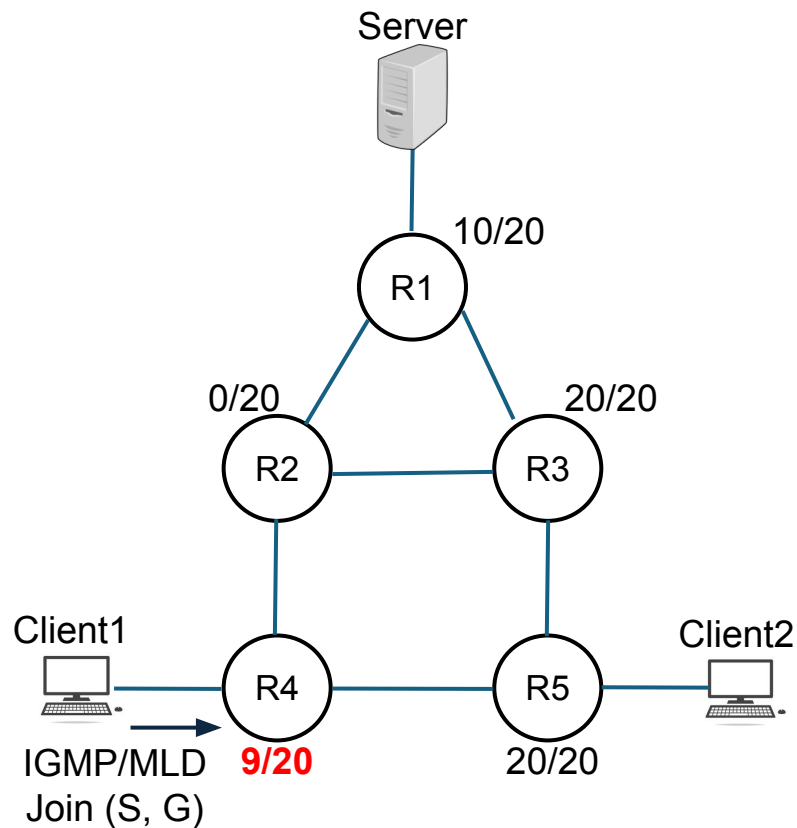
State exhaustion



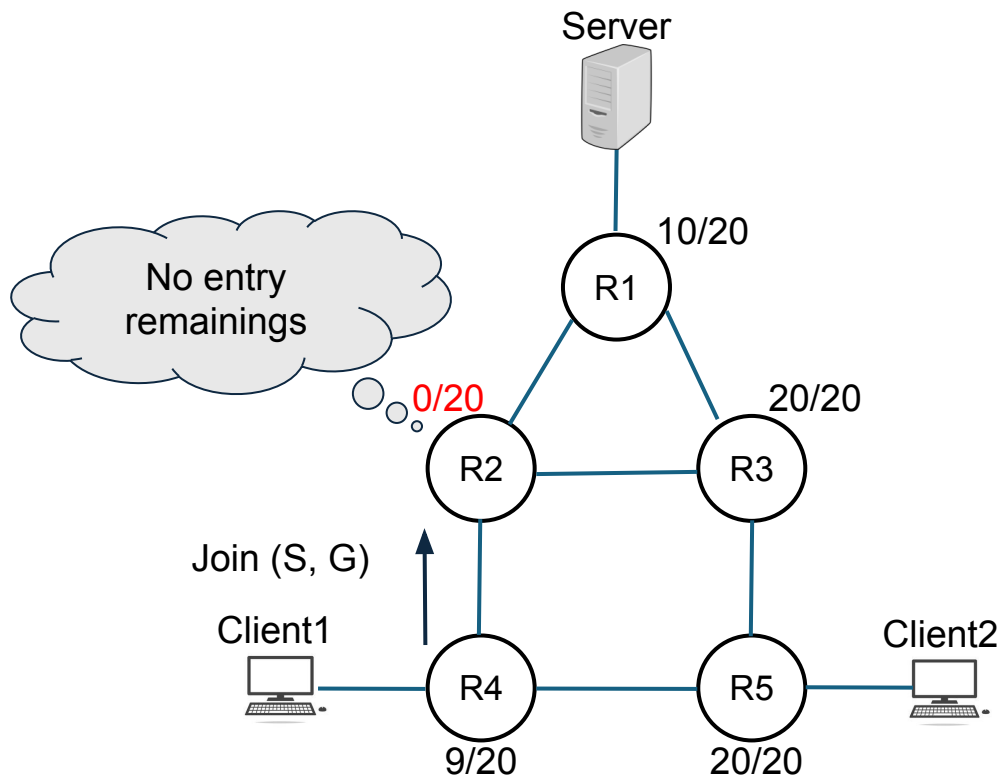
State exhaustion



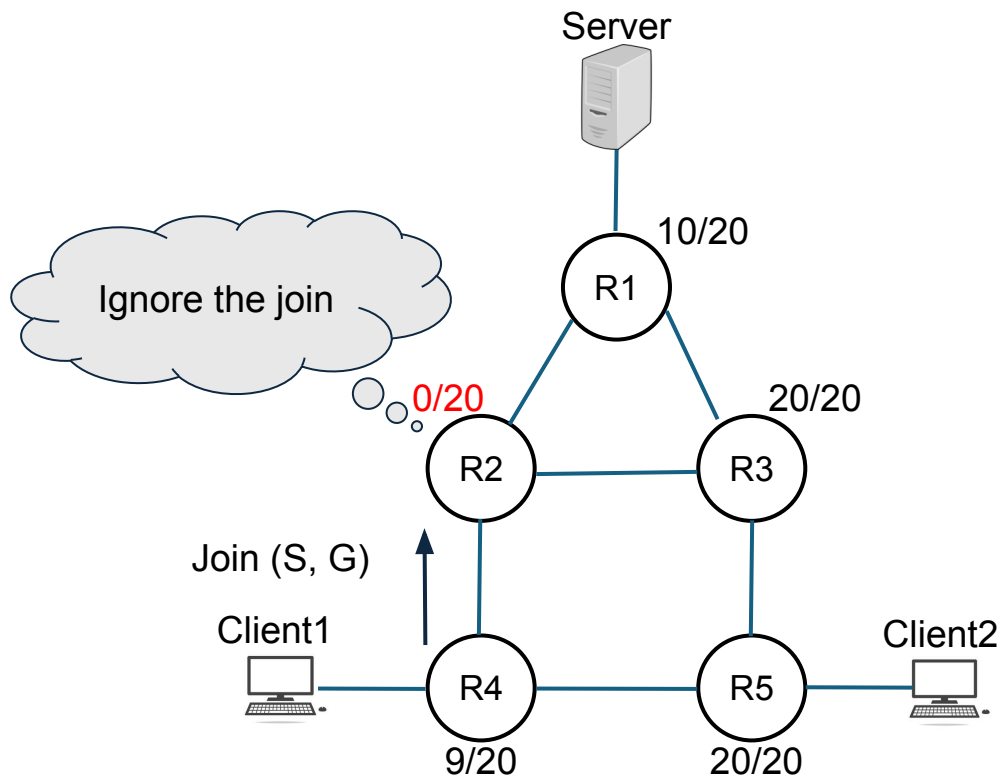
State exhaustion



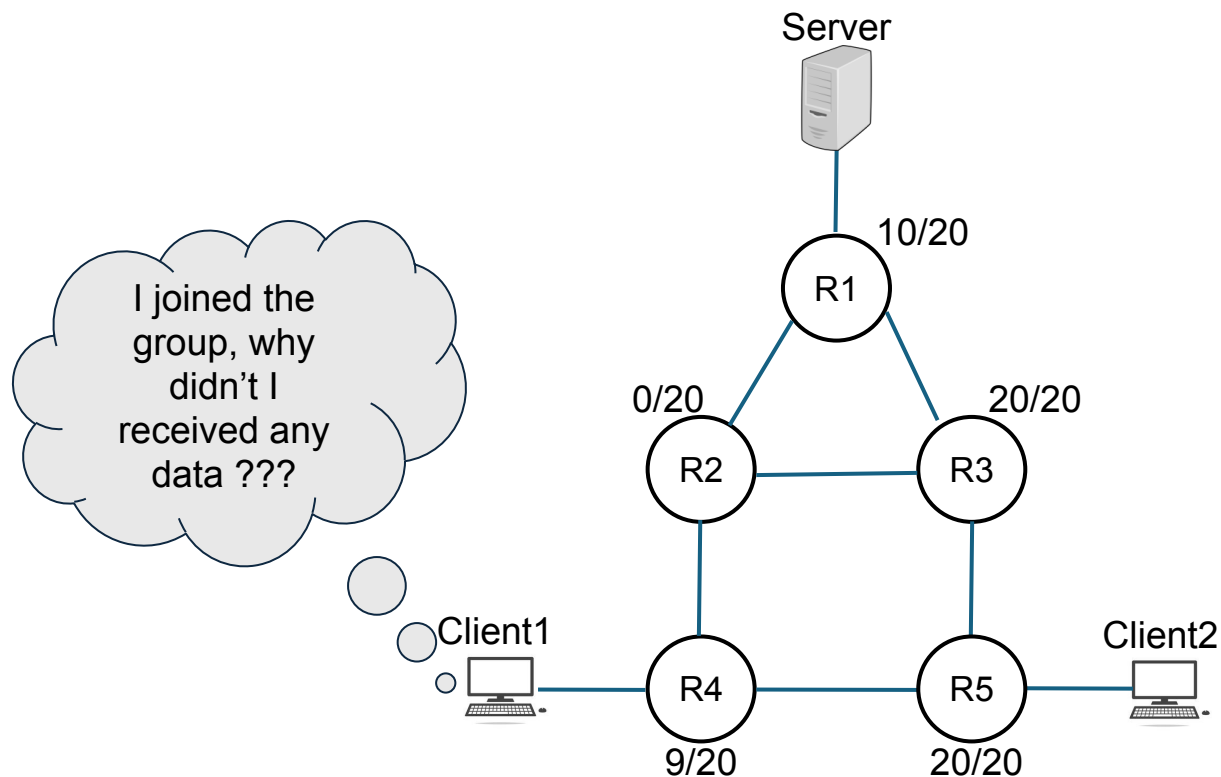
State exhaustion



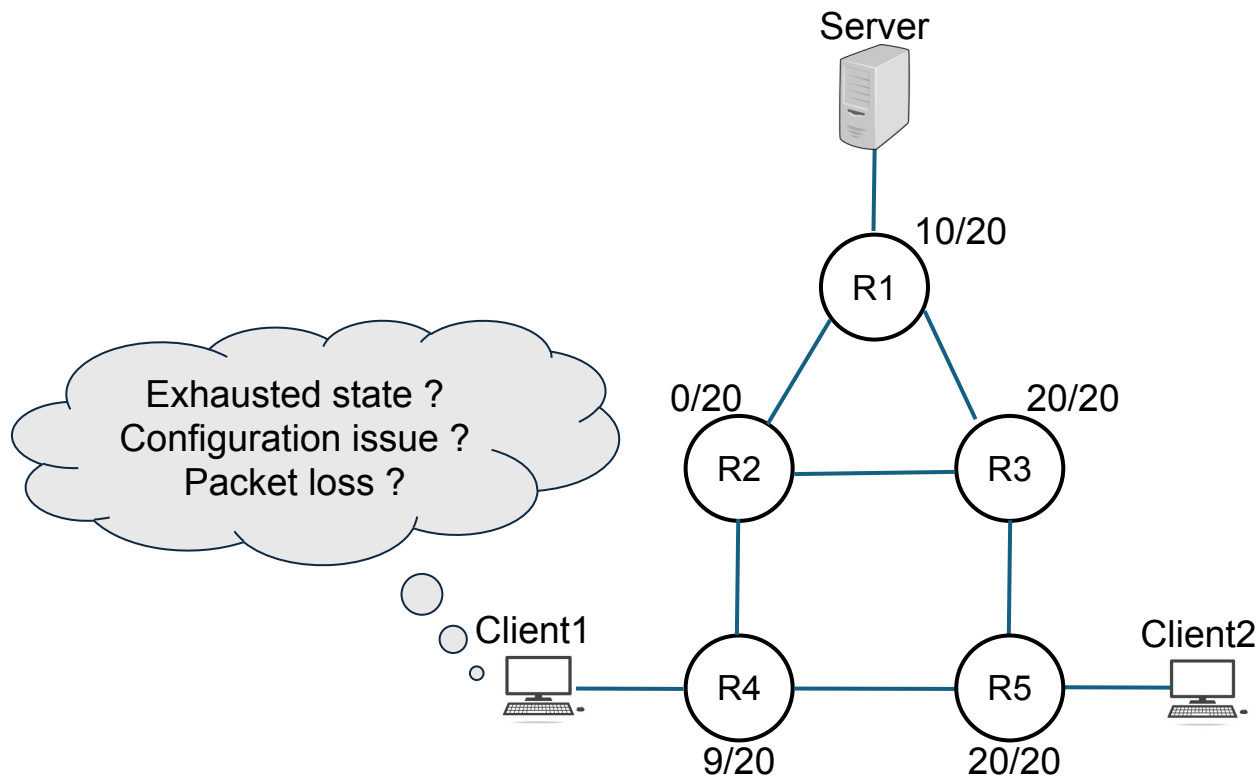
State exhaustion



State exhaustion



State exhaustion



Towards capacity-aware PIM

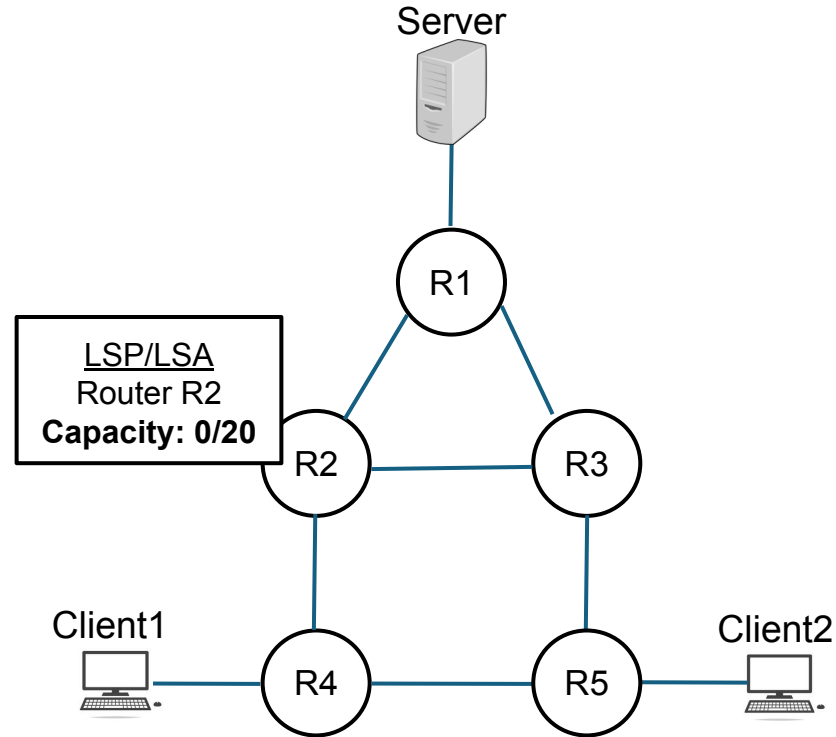
Our proposition:

Expose capacities, route around routers with exhausted entries

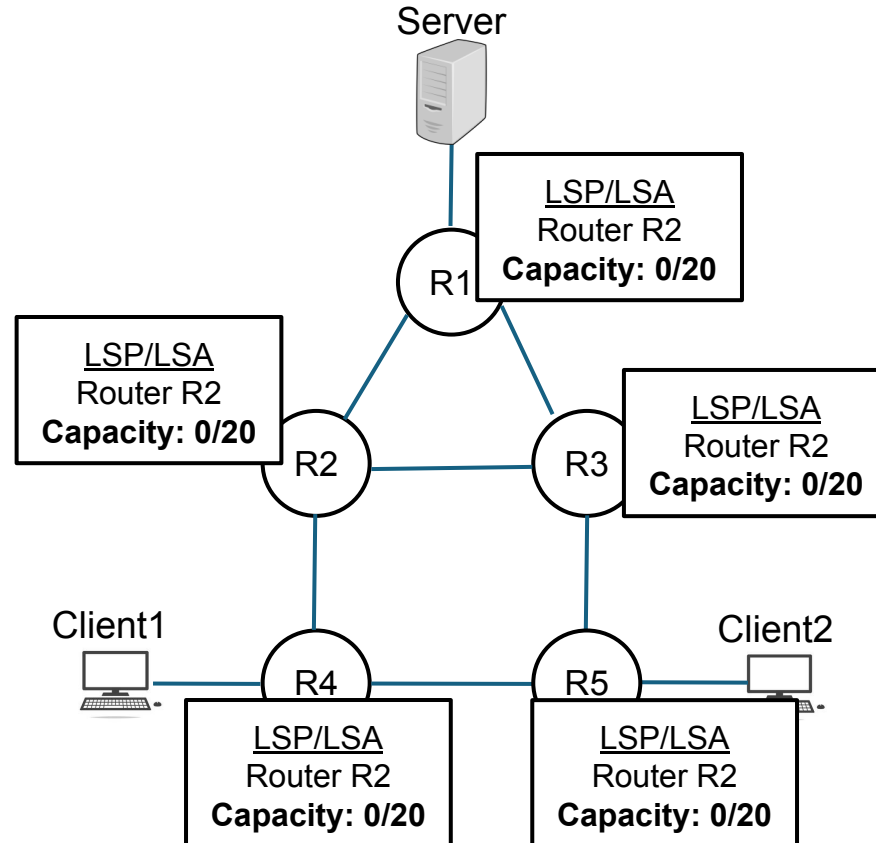
Exposing entries

Use existing link-state protocols

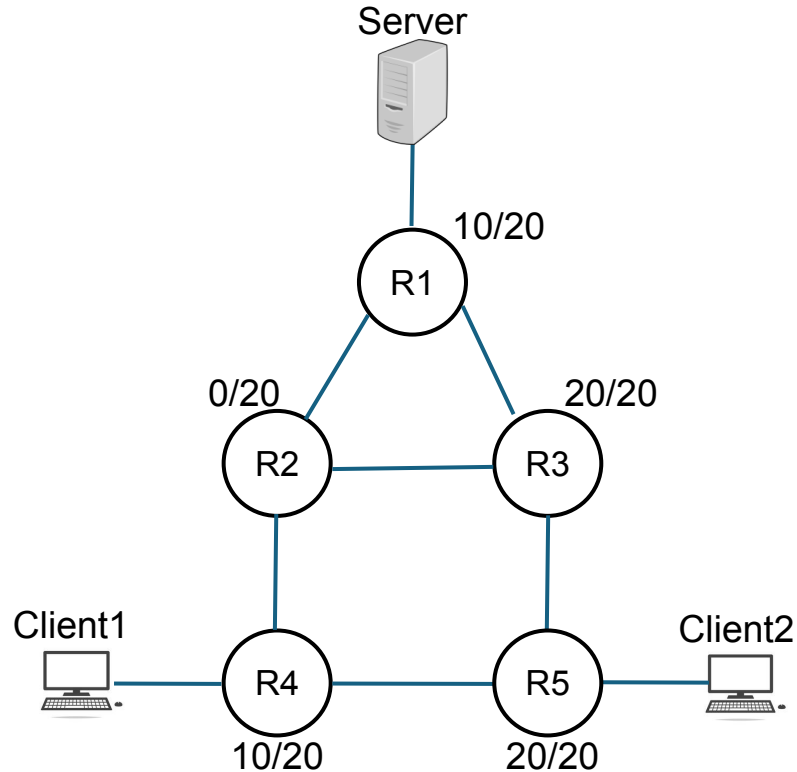
Exposing entries



Flooding



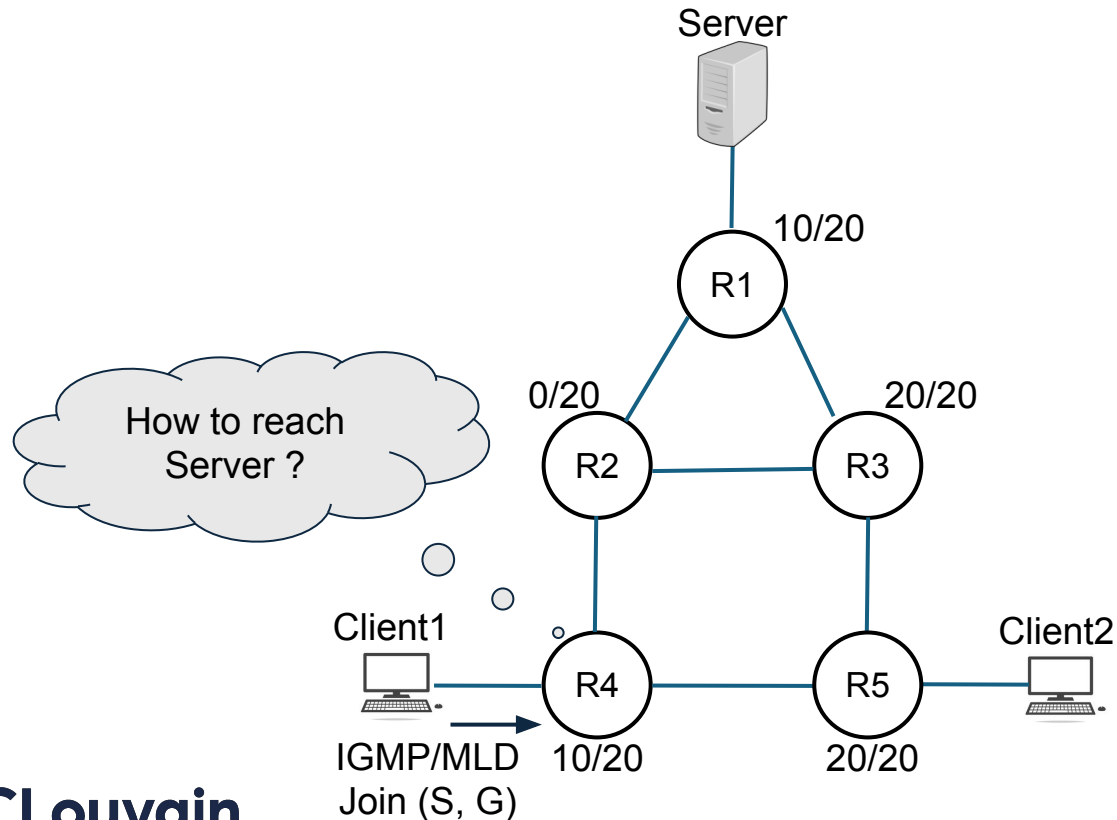
Network view



Exposing entries

Flooding: threshold-based with hysteresis
(similar to TE extensions of OSPF/IS-IS)

Routing around



Path computation

Path computation

Pruned shortest path

Remove routers
without remaining
entries from graph

Path computation

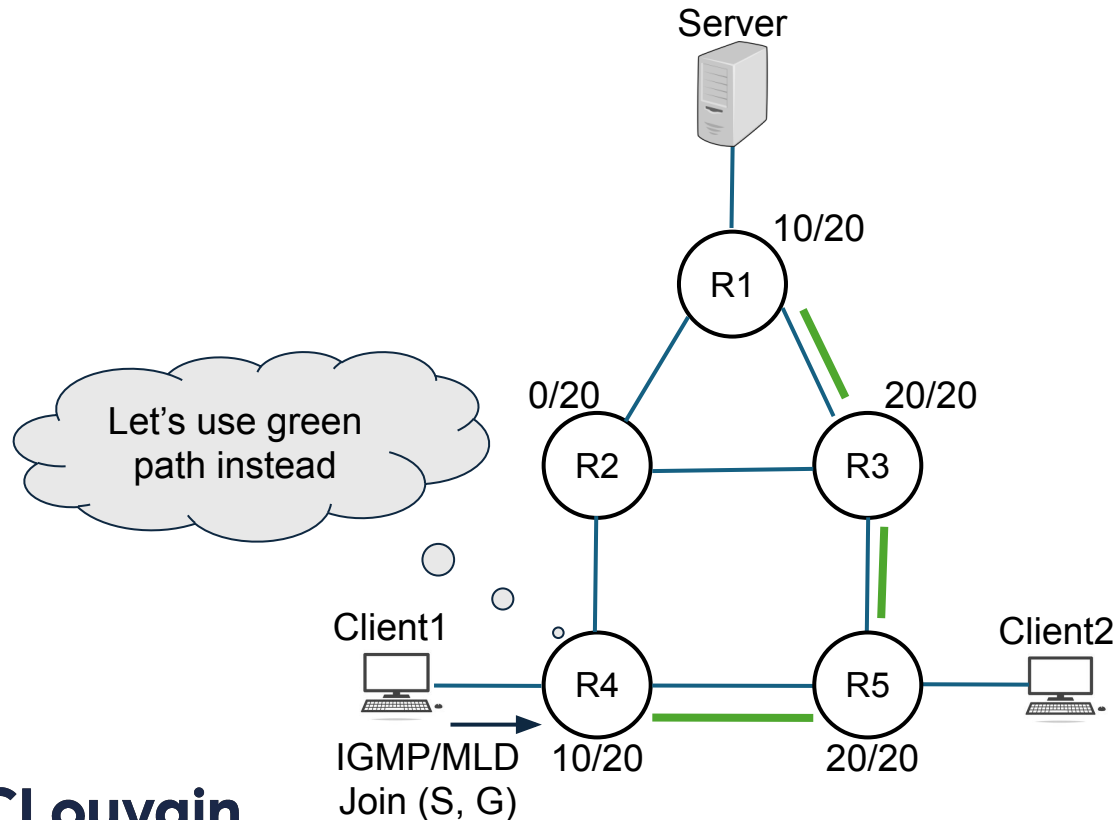
Pruned shortest path

Remove routers
without remaining
entries from graph

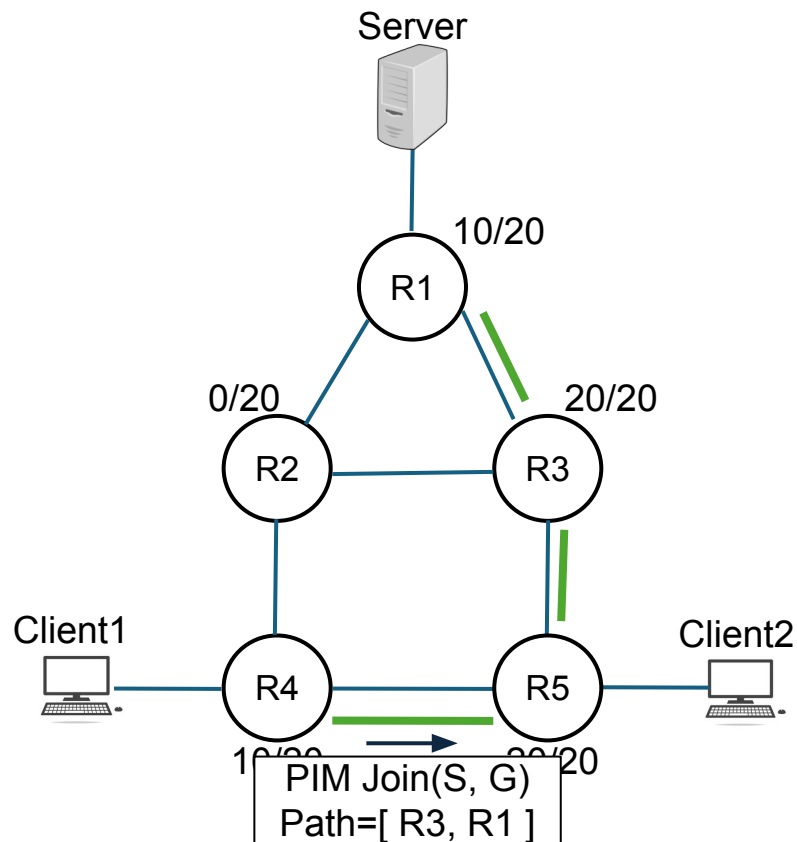
Shortest widest path

Search path with
maximum number of
remaining entries

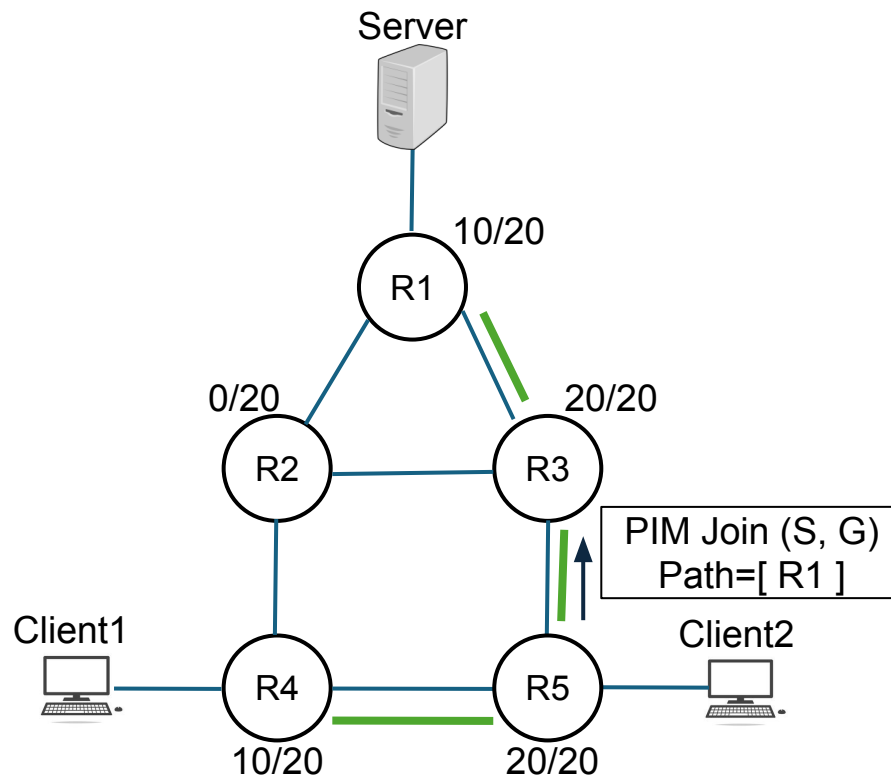
Routing around



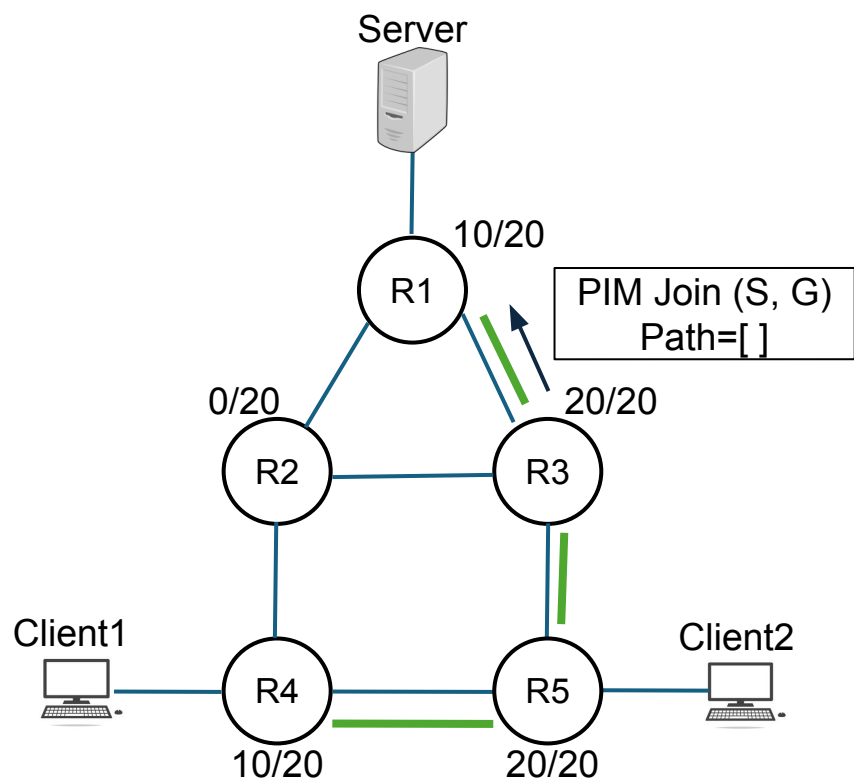
Routing around



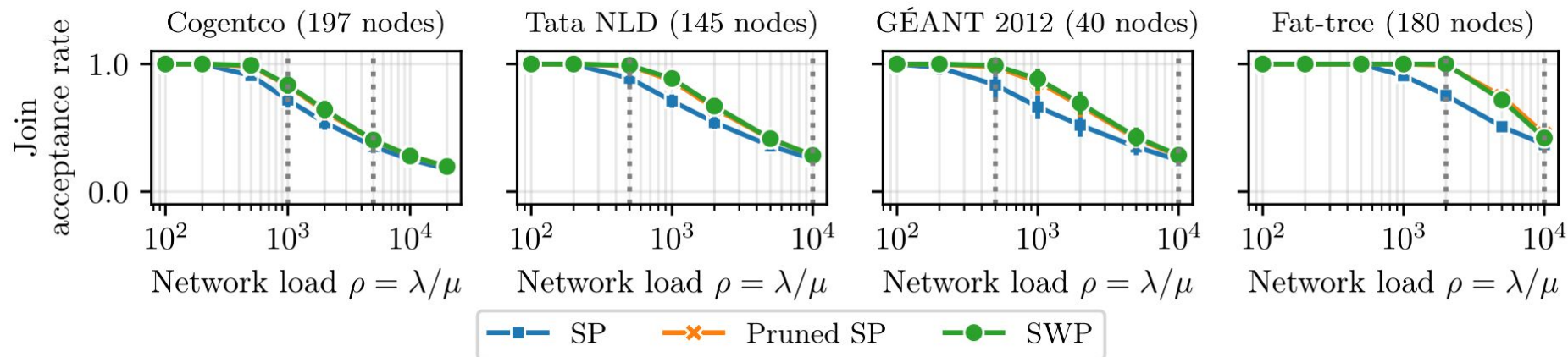
Routing around



Routing around



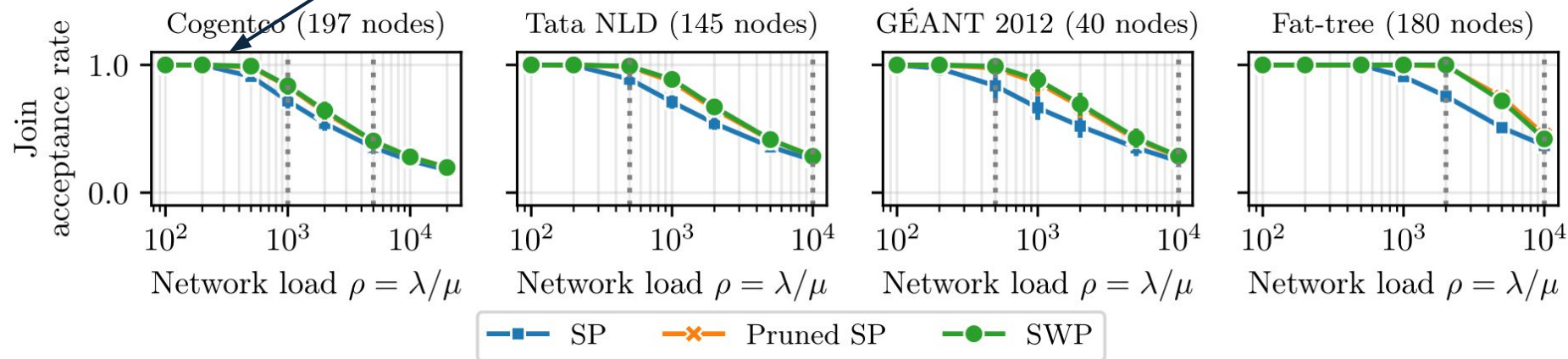
Join acceptance rate



Load = mean number of concurrent multicast trees
Router multicast entries in [50, 5000]

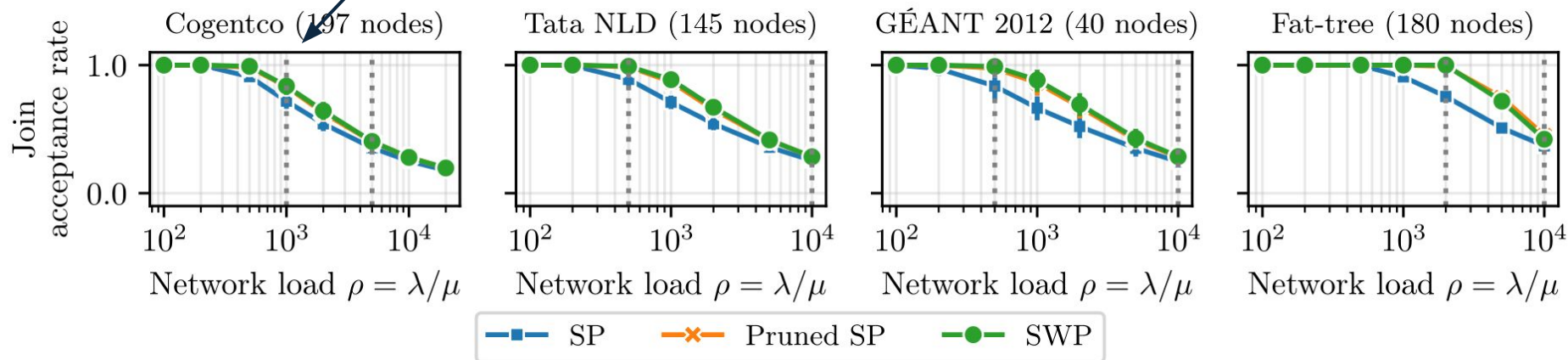
Join acceptance rate

Low loads => no difference



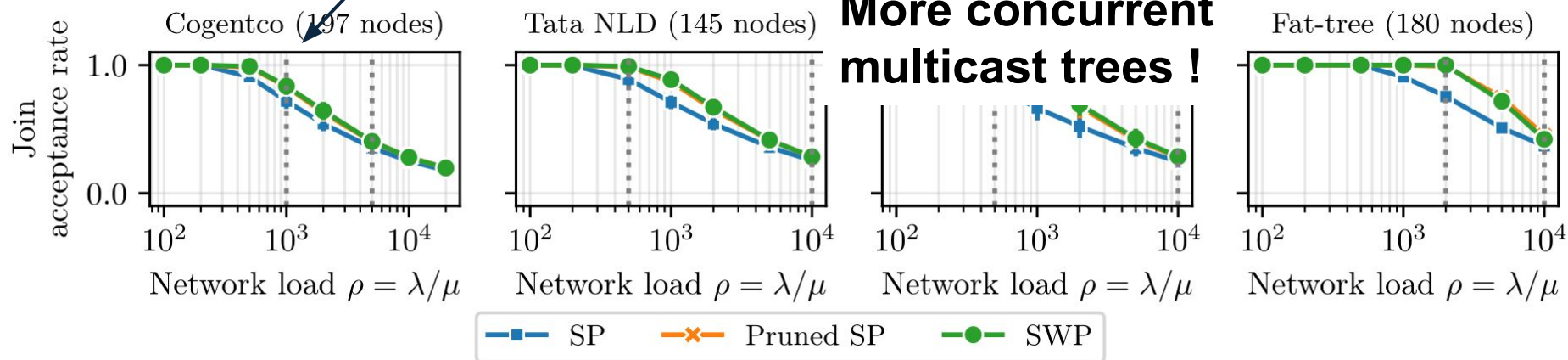
Join acceptance rate

Moderate loads => + ~22% acceptance rate



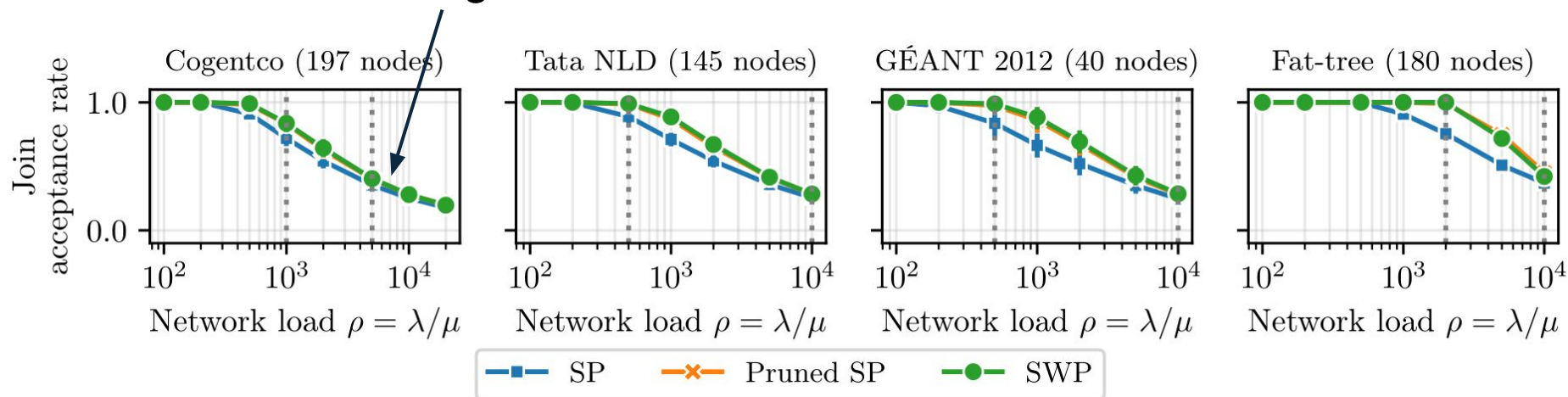
Join acceptance rate

Moderate loads => + ~22% acceptance rate

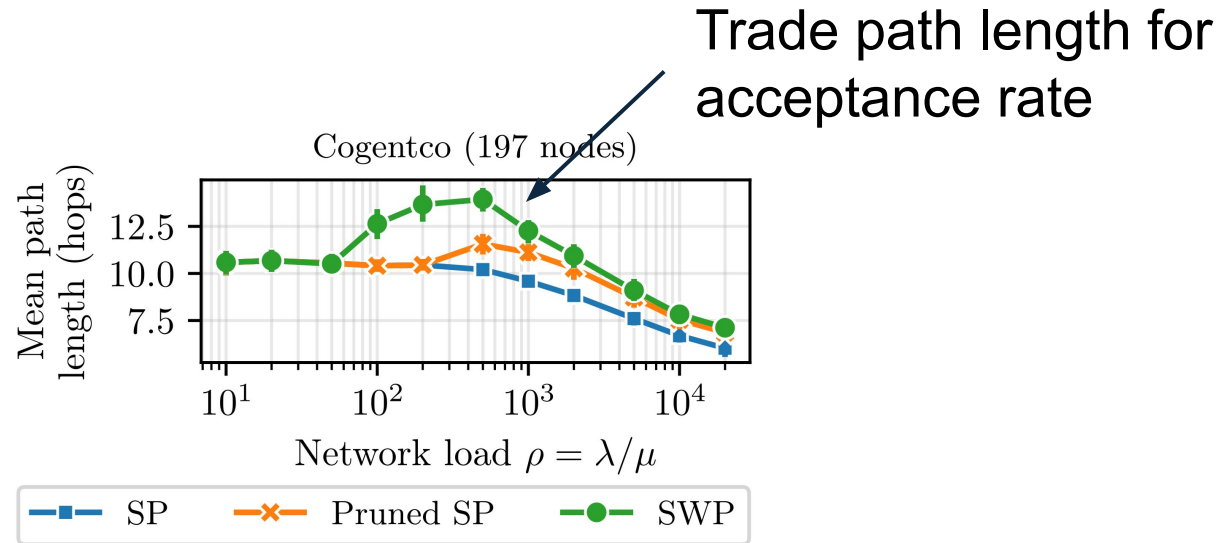


Join acceptance rate

High loads => no difference

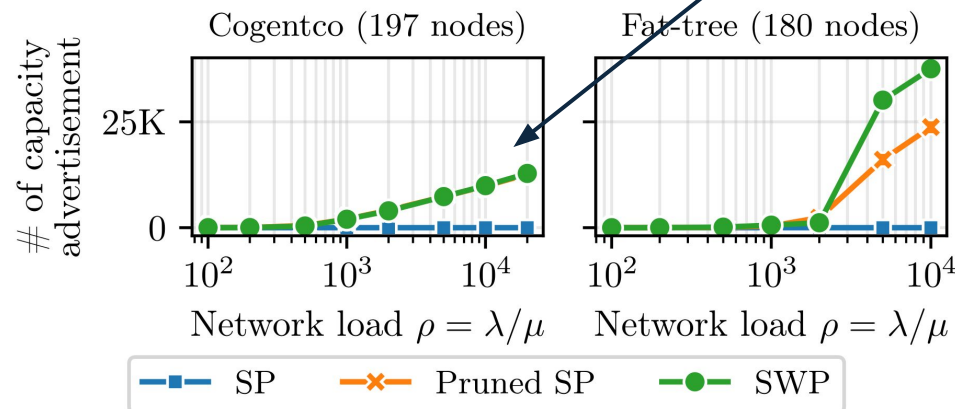


Path lengths



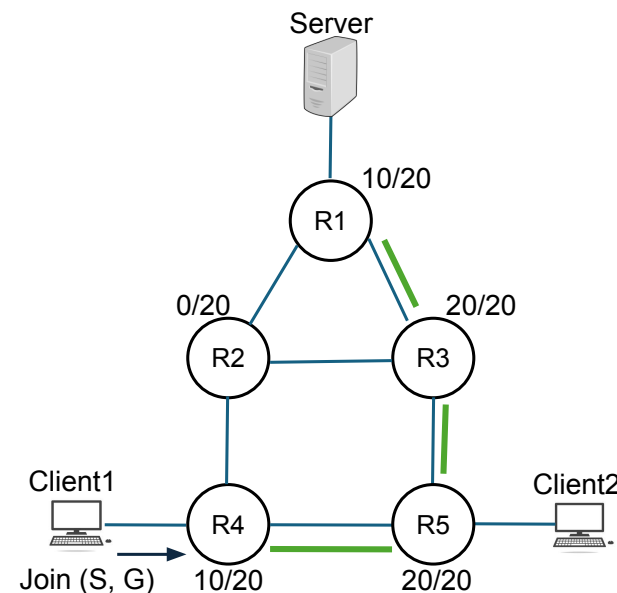
Flooding impact

Flooding to keep fresh capacity information



Conclusion

- PIM does not consider MFIB capacity
=> multicast tree creation failures
- Our proposition
 - Expose capacities with OSPF/IS-IS
 - Route around exhausted routers
- Increase ~22% join acceptance rate
=> More concurrent multicast trees
- Other possibilities: use metrics (BW/Delay) in tree creation



Thank you for your attention !

Thresholds

- Values (%)
[25, 50, 75, 80, 85, 90, 92, 94, 95, 96, 97, 98, 99, 100]



- Narrower bands at high utilisation
=> keep fresh information when near exhaustion
- Hysteresis: avoid oscillations