MPLS Scale to 100k endpoints with resiliency and simplicity

Clarence Filsfils

Distinguished Engineer

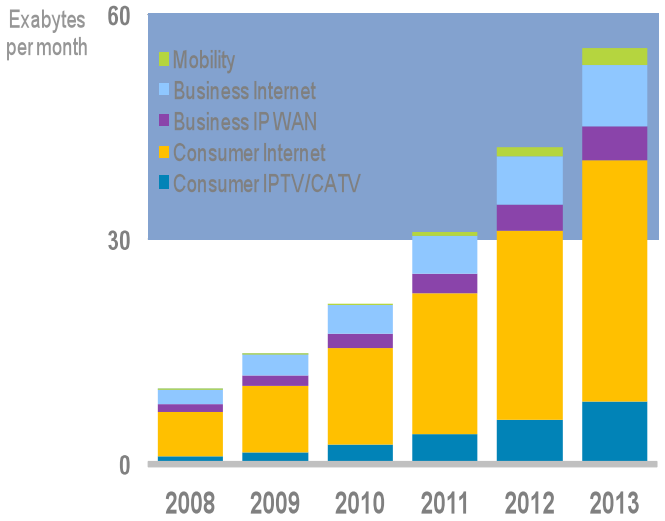
Seamless MPLS Architecture
draft-leymann-mpls-seamless-mpls-02

Outline

- Packet traffic will dominate
- MPLS expansion to Access/Aggregation
- Simplicity
- Scale
- Resilience
- Flexibility

Packet traffic will dominate

- IP services growth rates : 40% globally, nearer 100% for mobile
- All aspects of wireline and mobile solutions moving towards packet
- Packet traffic is the main driver for DWDM upgrades today/future
- Two aspects to packets: L2 transport and L3 routing
- L3 edge and content is extending further out into the network
- Packet switching and transport technology are converging from a cost perspective
- Stat Mux is a key requirement in building cost effective packet networks

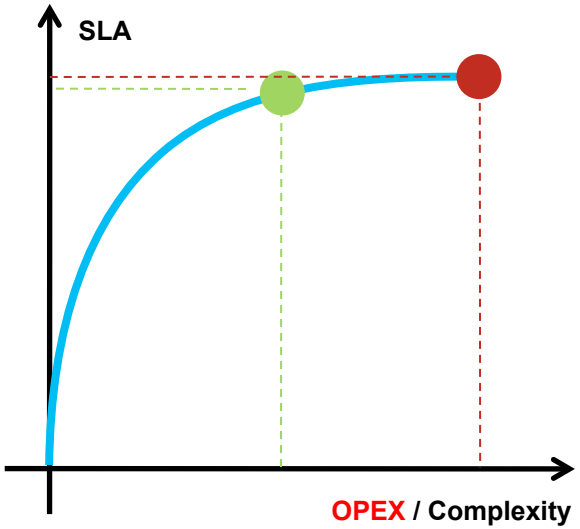


MPLS expansion

- MPLS deployment in the core is a vast success
  - L3VPN, MPLS TE FRR, L2VPN
- MPLS deployment in the access/aggregation
  - Scale
  - Resilience
  - Service Flexibility
  - **Simplicity**

Simplicity

- “Simplicity is prerequisite for reliability”
  Edsger Dijkstra
- "Simplicity is the ultimate sophistication"
  Leonardo da Vinci
- Simplicity to minimize OPEX



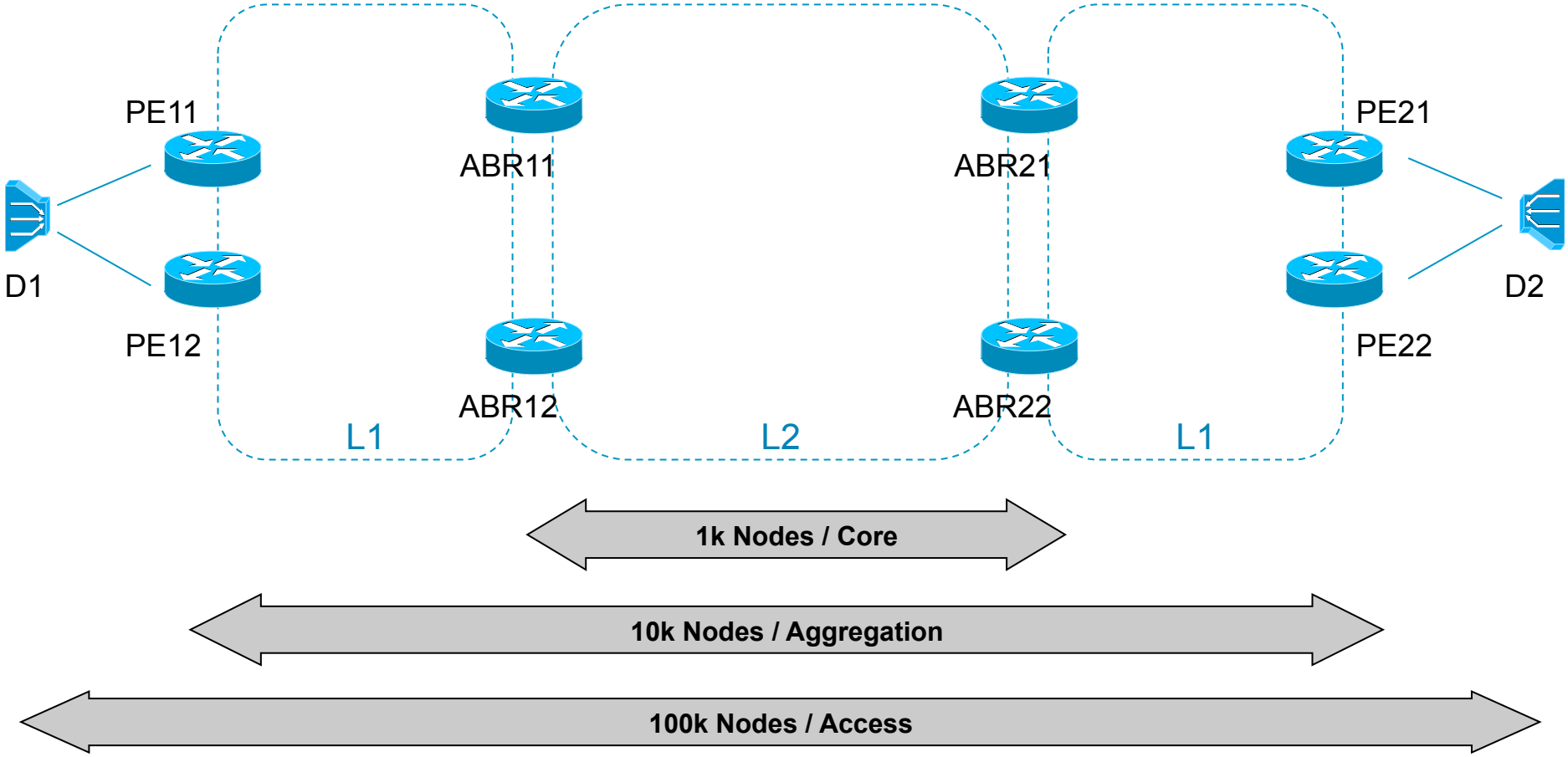
Scale and Resilience with Simplicity

- 100k edge nodes
  - An edge node may have an LSP to any other edge node
  - Simplicity: only requires provisioning on the involved edge node
- 50msec Protection
  - Simplicity: no operator involvement, router optimization which automatically provides 50msec protection

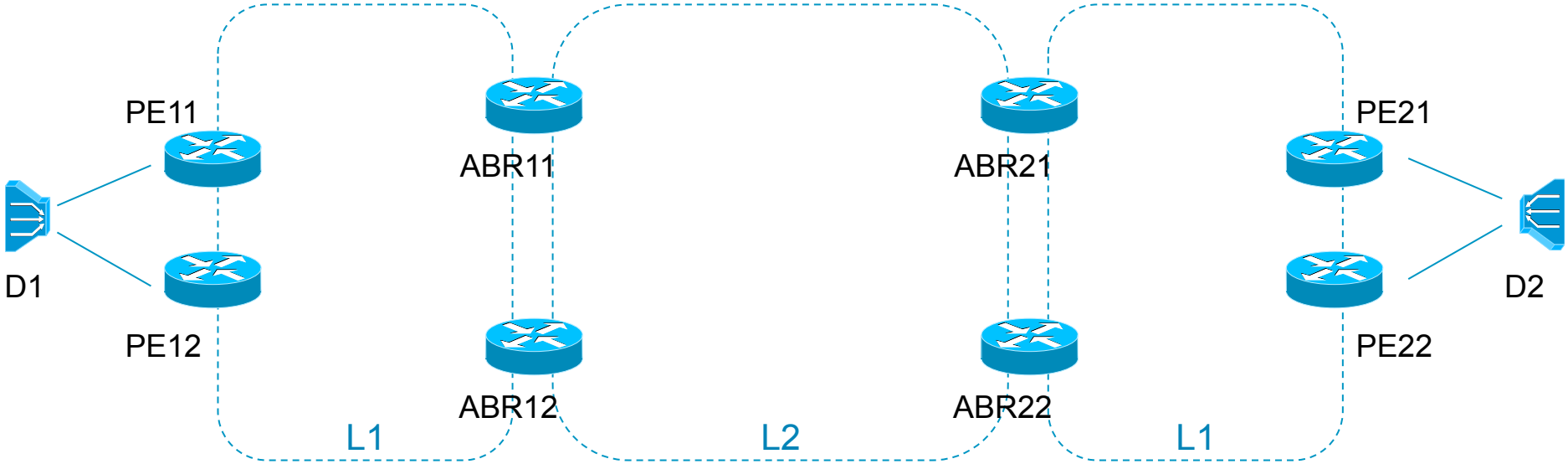
    BGP PIC and LFA FRR

See "Seamless MPLS: Integrating Access and Aggregation into a single MPLS network", N. Leymann, DT

Reference Model

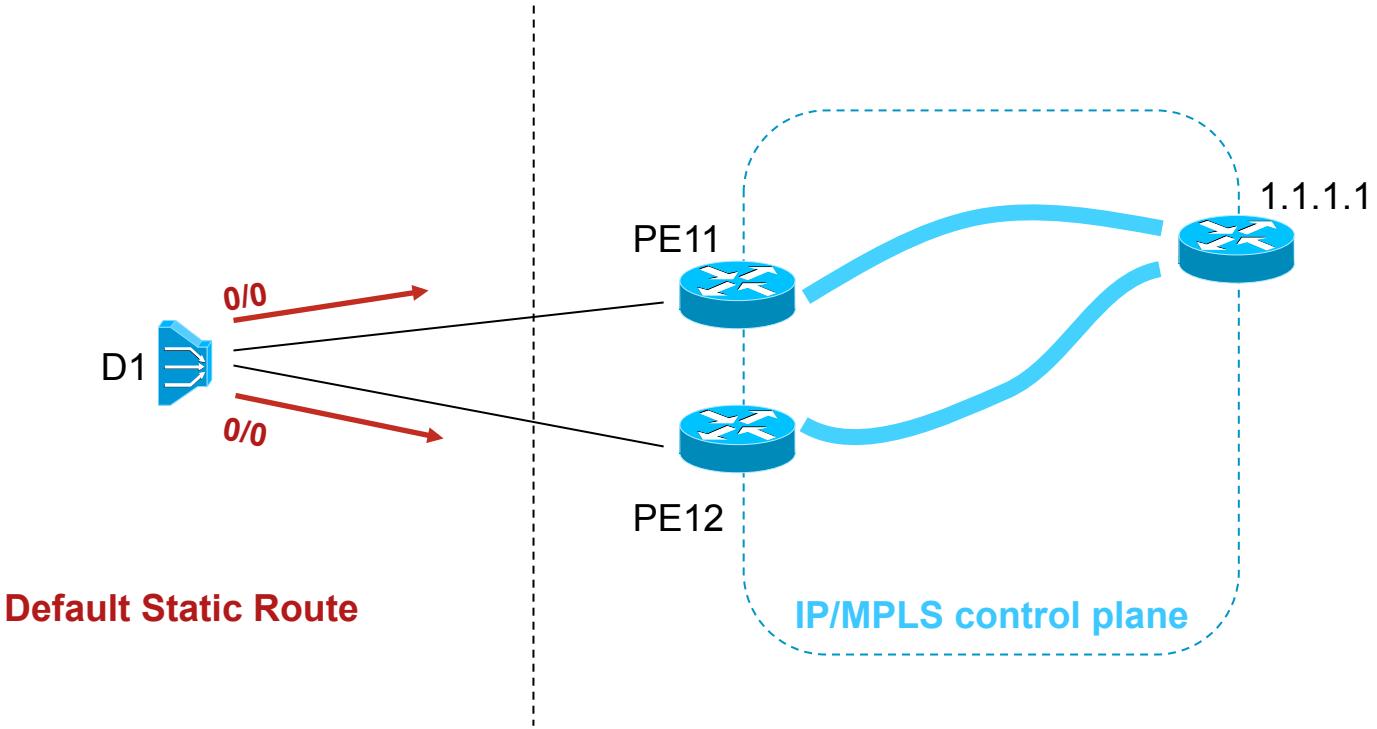


Scaling the Access Nodes



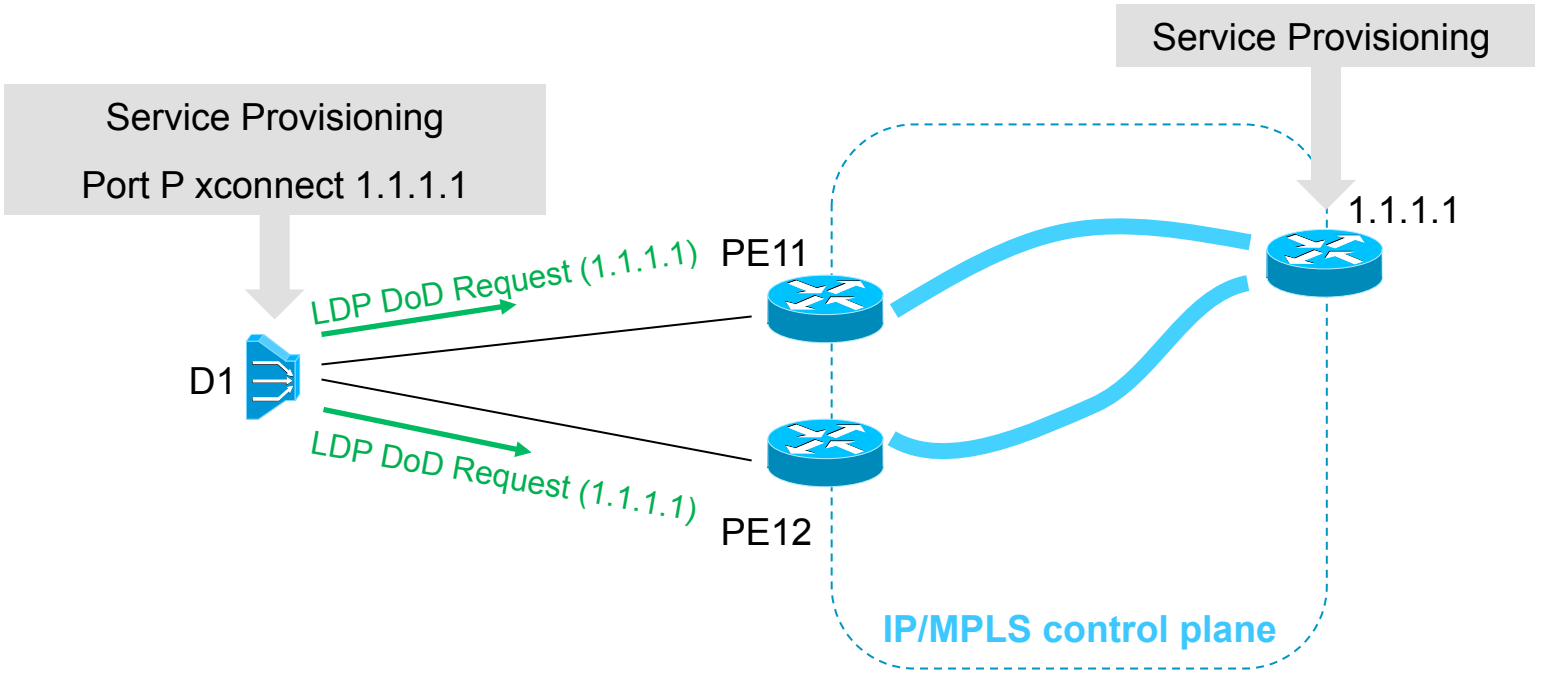
100k Nodes / Access

LDP Downstream on Demand



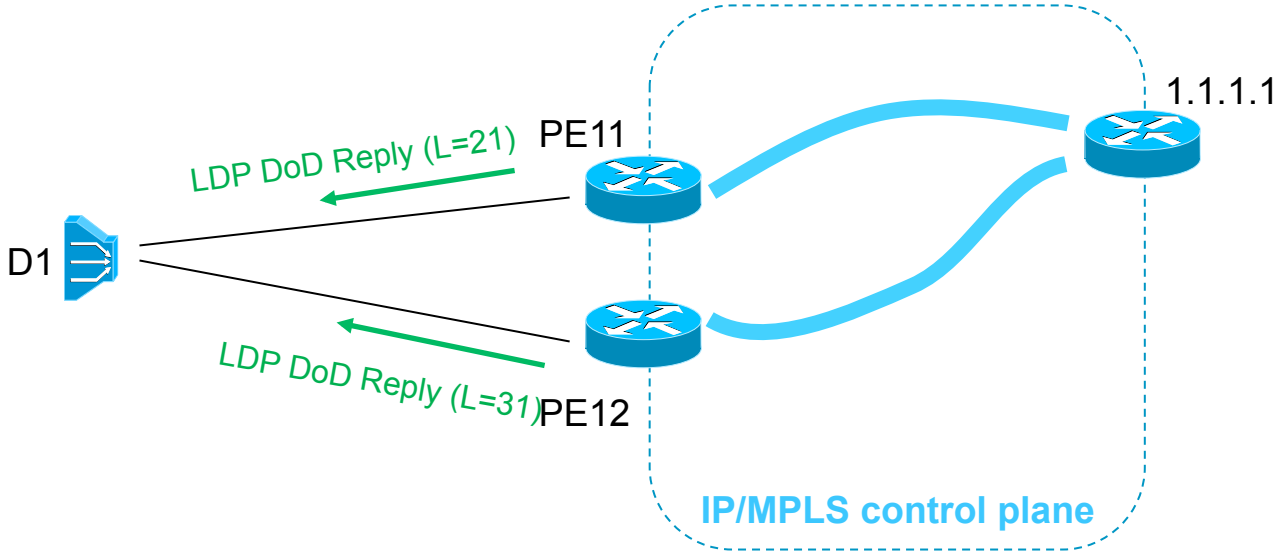
- Access node remains extremely simple
  - no IGP, no BGP

LDP Downstream on Demand



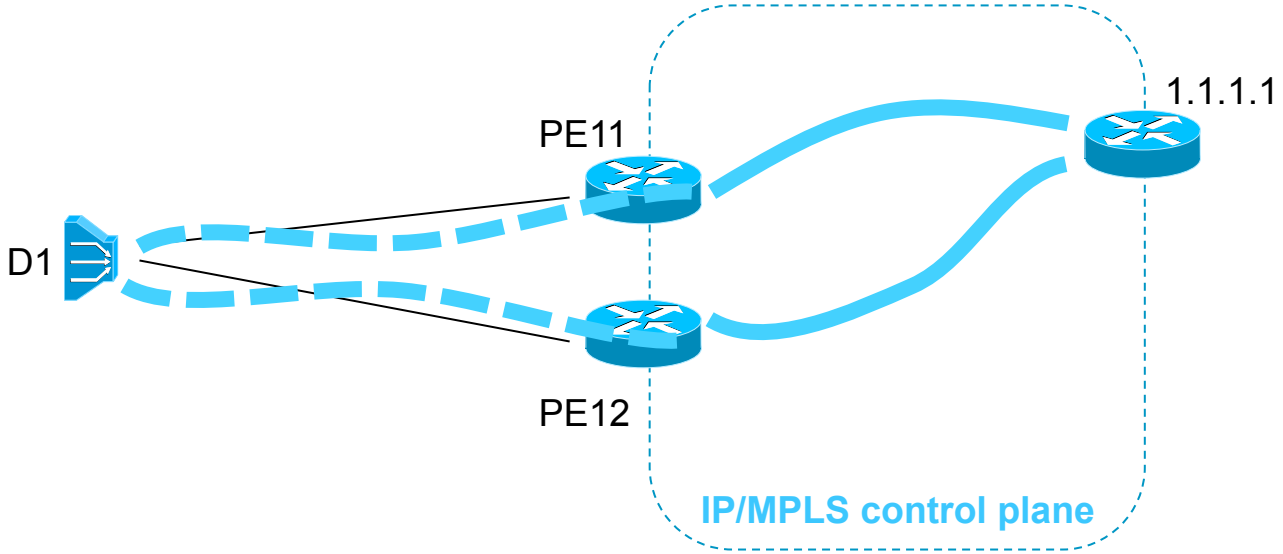
- No service provisioning anywhere else

LDP Downstream on Demand



- No service provisioning anywhere else

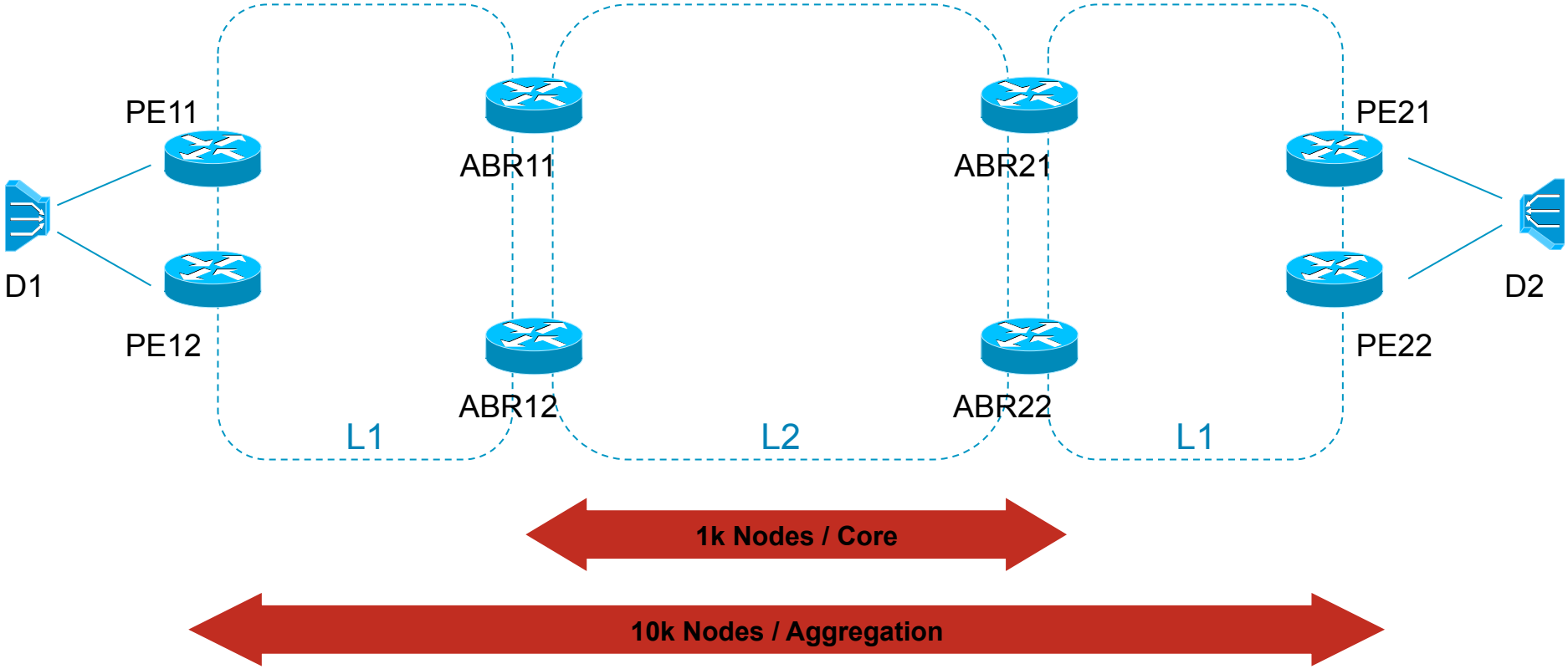
LDP Downstream on Demand



LDP Downstream on Demand

- Access node is extremely simple
  - No IGP, no BGP
- Access node may have an LSP towards any other node
- Access node only knows the labels it needs
- Simple and Scaleable
- Leverage existing technology (simplicity)

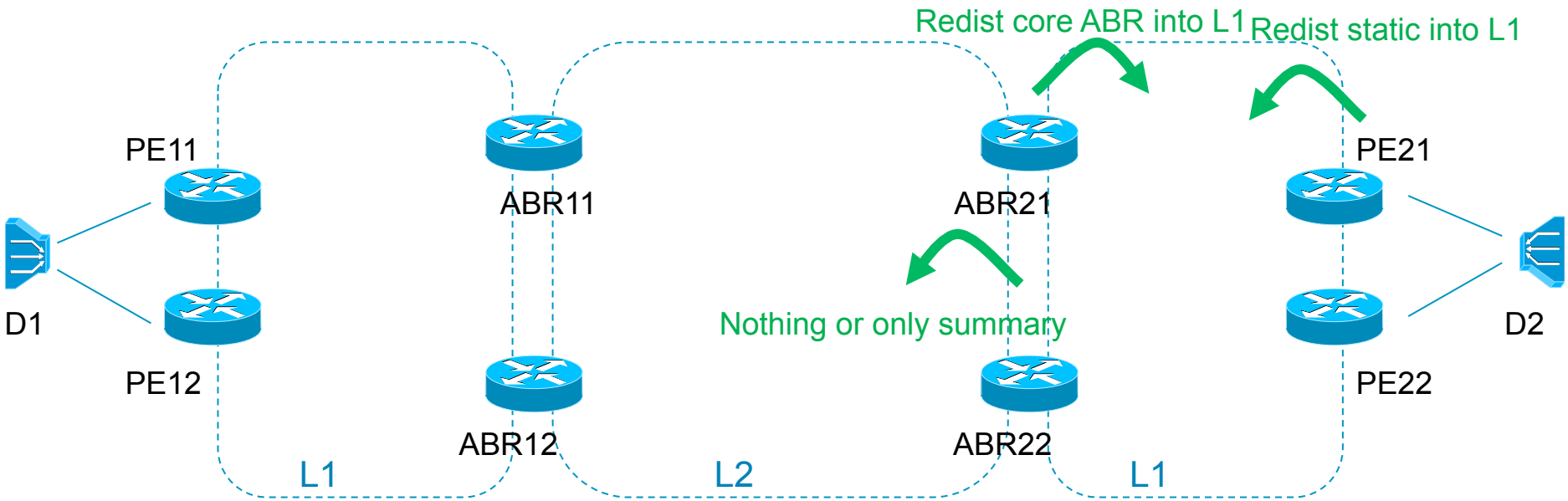
Scaling the IP/MPLS Control Plane



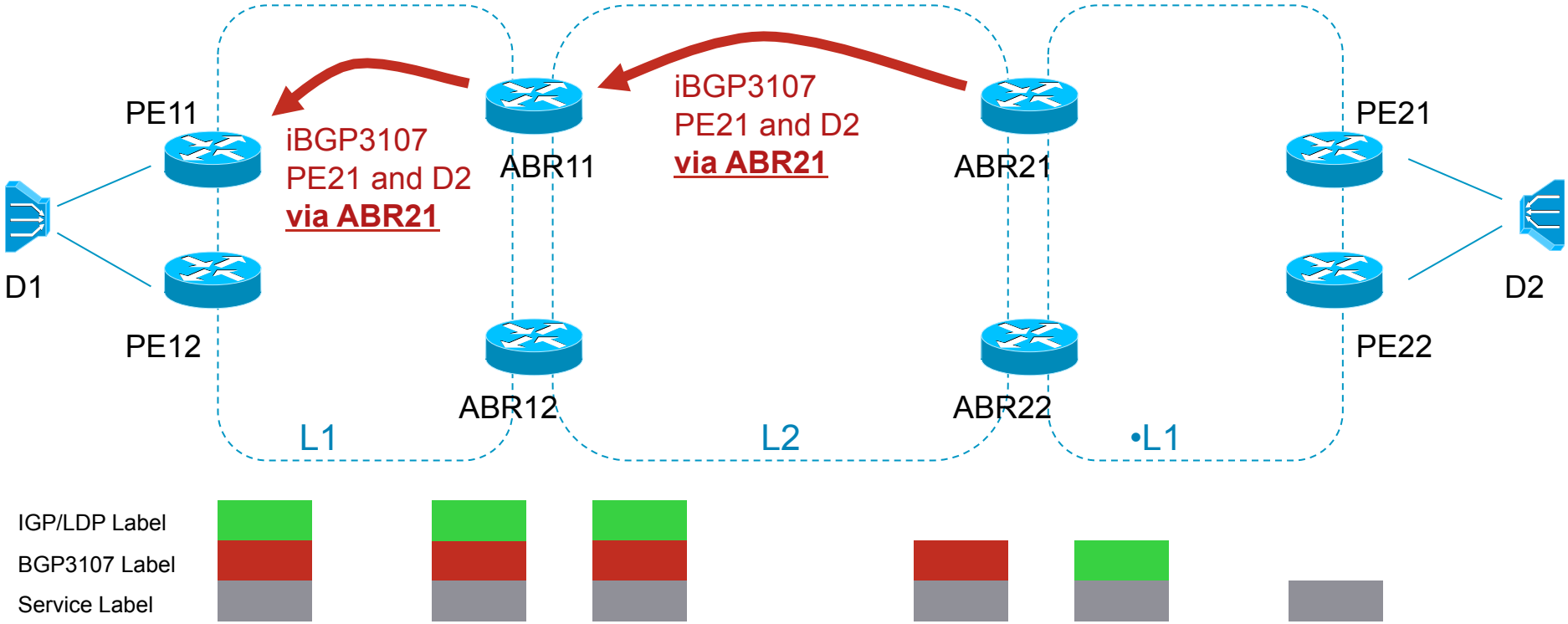
Divide and Conquer

- To scale, introduce a layer of hierarchy
  - BGP
- Possible thanks to key innovation: BGP PIC
  - Scale-Independent BGP FRR
  - **Simple**: default router behavior

IGP ~ K entries

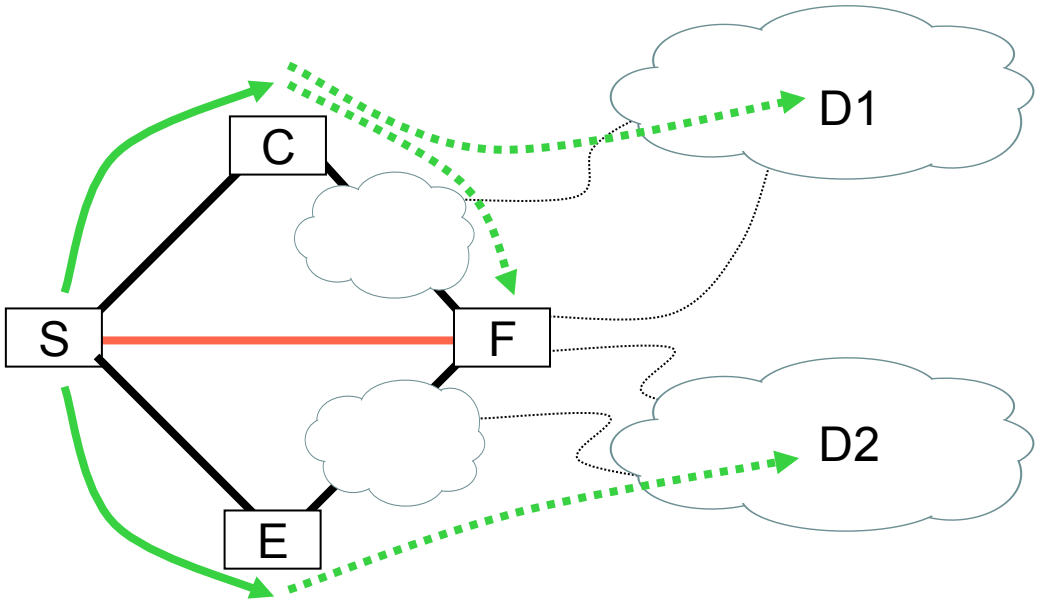


BGP 3107 ~ 100K entries



- Each IGP area has routes for that area only + routes to core ABR's (~1k prefixes)
- LDP labels used to traverse each area and reach core ABR's
- BGP labels used by PEs and ABRs to reach PE's in remote areas
- Service (e.g., PW) labels used by Pes
- Add-Path

IGP FRR: Loop-Free Alternate (LFA)



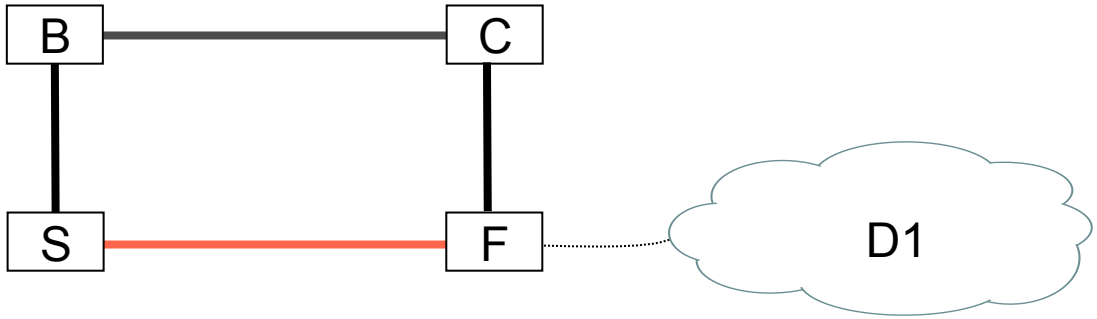
- IGP route D1
  - Primary Path: via F
  - Backup Path: via C because C's primary path is not via S
- IGP route D2
  - Primary Path: via F
  - Backup Path: via E because E's primary path is not via S

LFA Benefits

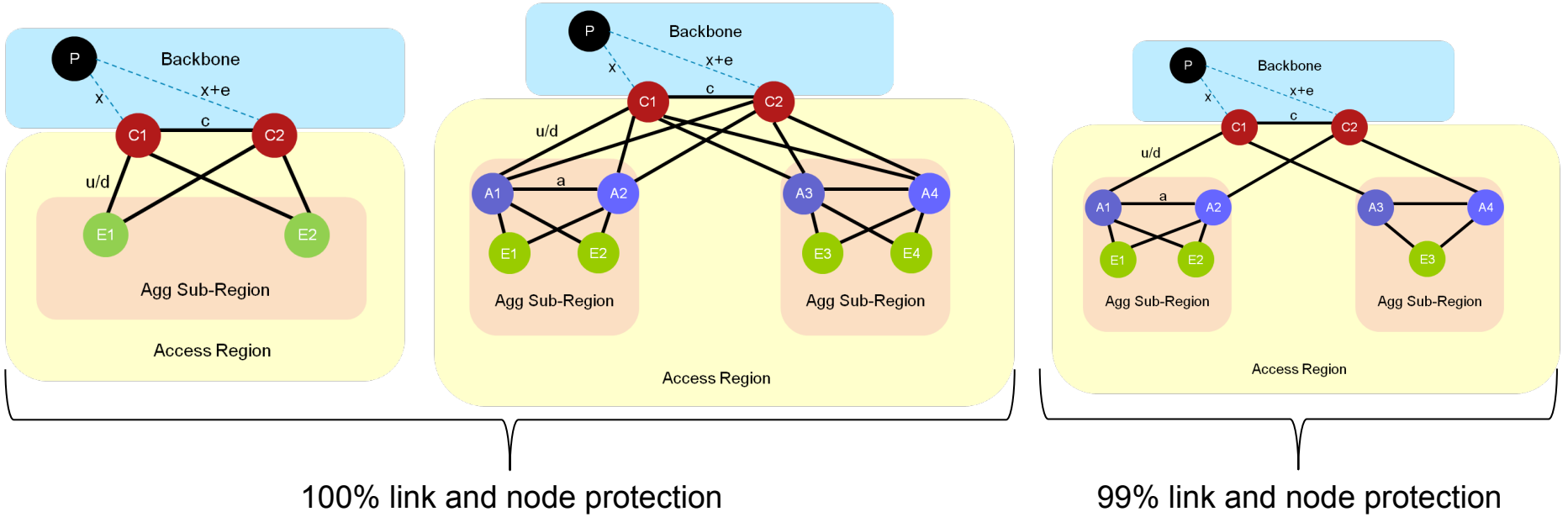
- **Simple**
  - the router computes it automatically
- <50msec
  - pre-computed, pre-installed, enabled on link down in a prefix independent manner
  - Leverage Hierarchical dataplane FIB
- Link and Node Protection
- Deployment friendly
  - no IETF protocol change, no interop testing, incremental deployment

LFA Constraint

- Topology dependent
  - availability of a backup path depends on topology
  - S has no LFA for dest D1



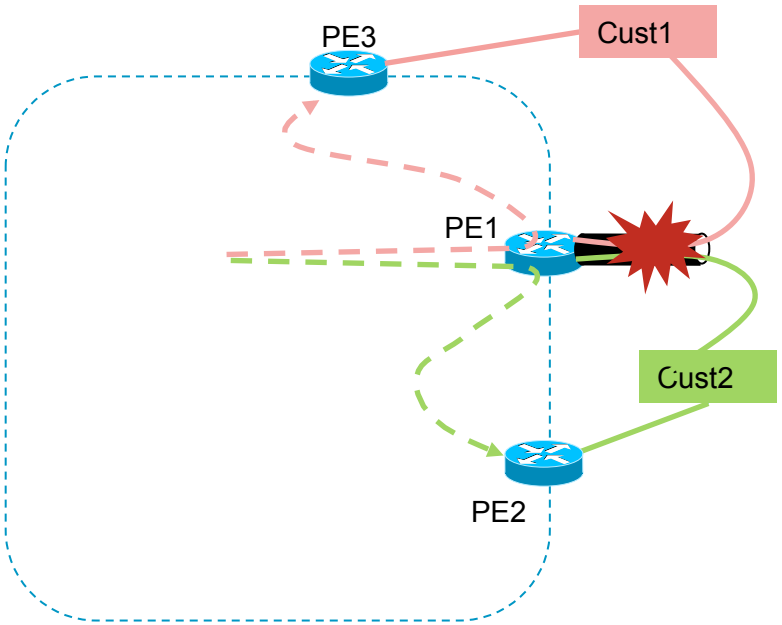
Access/Aggregation Topologies



- See draft-filsfils-lfa-applicability-00

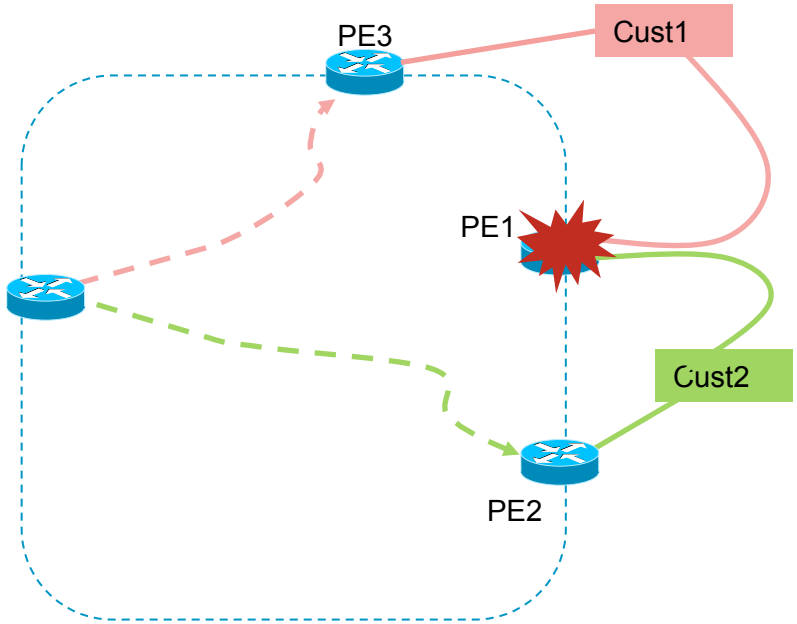
BGP Prefix-Independent Protection (PIC)

- 50msec protection
- Prefix-Independent
- Default behavior, entirely automated computation
- No operator involvement
- **Simple**

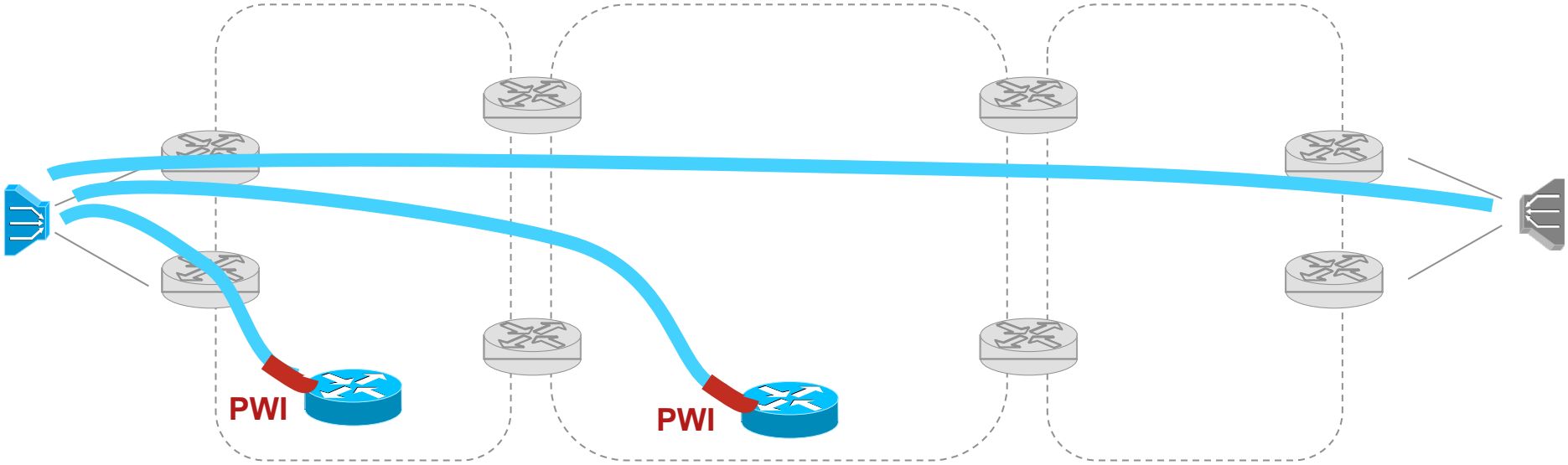


BGP PIC

- x00msec Protection
- Prefix-Independent
- Default behavior, entirely automated computation
- No operator involvement
- **Simple**



Service Flexibility



- Service and Network Architecture are decoupled – No boundary
- **Simplicity** leads to OPEX optimization
  - MPLS as single packet transport technology
  - uniform end-to-end service protection at scale

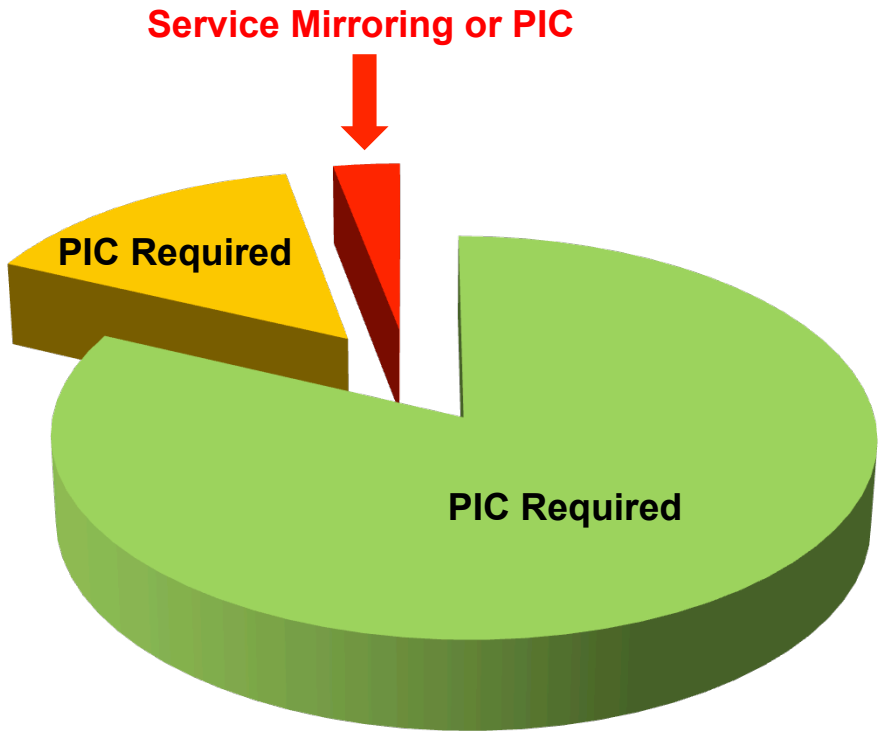
See "Flexible Service Edge Architecture", Le Faucheur

Conclusion

- Packet traffic will dominate
- Innovations support 100k edge nodes in an mpls network with 50msec protection and **simplicity**
- Simplicity to minimize OPEX
  - Operational Convergence
  - Plug&Play 50msec Protection
  - Service Flexibility/Velocity

Backup

Service Mirroring Applicability



- PE-CE Link Failure - BGP PIC required
- PE node failure - strict homing rule does not apply - BGP PIC is required
- PE node failures - strict homing rule applies - Service Mirroring or BGP PIC

Service Mirroring Complexity

- Operator Configuration Intensive
  - Catastrophic service impact if misconfigured
- Restrictive Assumptions
  - small applicability
- Does not replace the need for BGP PIC
  - Additional Technology