WTR
WORKSHOP
DE TECNOLOGIAS DE REDES DO POP-BA

14 A 18 DE SETEMBRO DE 2020

Migrating to MPLS

Not just another MPLS talk – I promise!

Humberto Galiza – Network Engineer

PoP-BA
Ponto de Presença da RNP na Bahia

RNP 30 anos
ORGANIZAÇÃO SOCIAL DO MCTI

Disclaimer

*Opinions expressed are solely my **own** and do not express the views, opinions, products or technologies of my current employer.*

This presentation is intended for educational purposes only and do not replace independent professional judgment.

The information being shared in the presentation haven't hammered any NDAs. All the references mentioned are public information and can be found on the Internet or as form of Academic Papers and/or IETF Request For Comments (RFCs).

What to expect?

What this talk is

- Overview of possible MPLS use cases for metro, campus, datacenter, etc.
- MPLS tools as they apply to those use cases
- Assumes you have some familiarity with IPv4/IPv6 and MPLS fundamentals

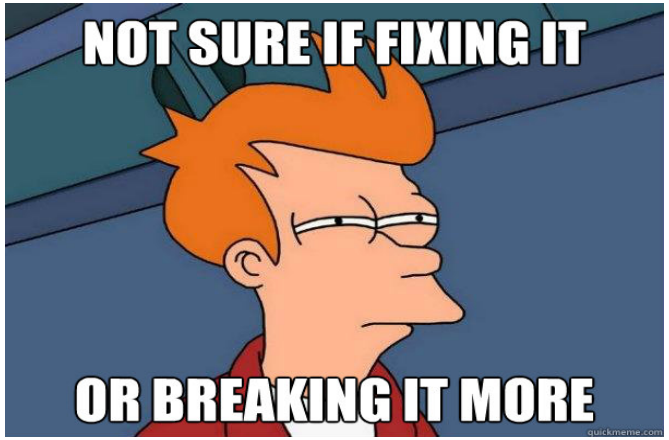
What this talk is NOT:

- MPLS design workshop
- MPLS in depth
- Configs details

Warning: Information overload

- There's a lot to say about MPLS and TE
- But we only have ~25 min – It's going to be fast and furious!
- Just a taster and lots of pointers
- If I leave out your favourite protocol (or vendor pet), please forgive me :p

Who am I?

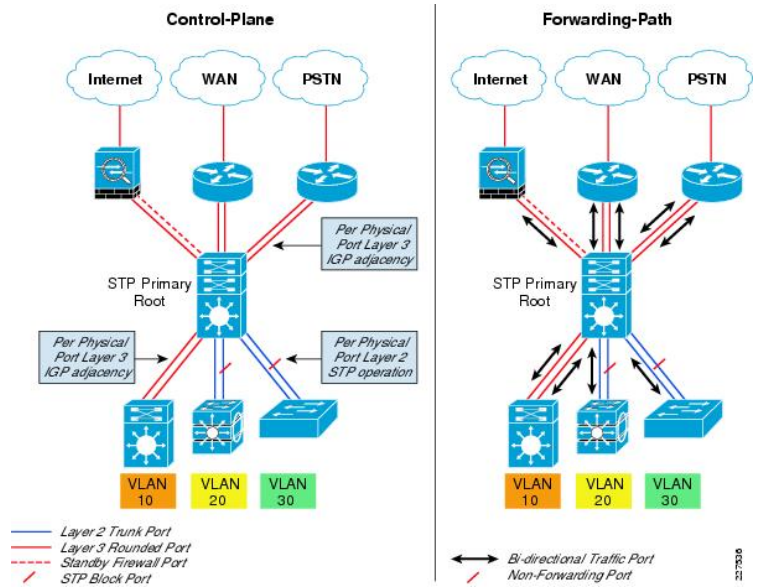


- I’m a nice guy - (*my wife used to say whenever she wants to shop some stuff online at amazon.com*) ☺
  - Ex UFBA alumni (BSc Computer Science, 2009)
  - Ex PoP-BA/RNP/AmLight staff
  - Couple of years dealing with networks of different sizes and complexity levels
  - NDE @ AWS Global BB-Eng Team (Dublin, IE)
    - Breaking & fixing super cool/massive-scale stuff :D

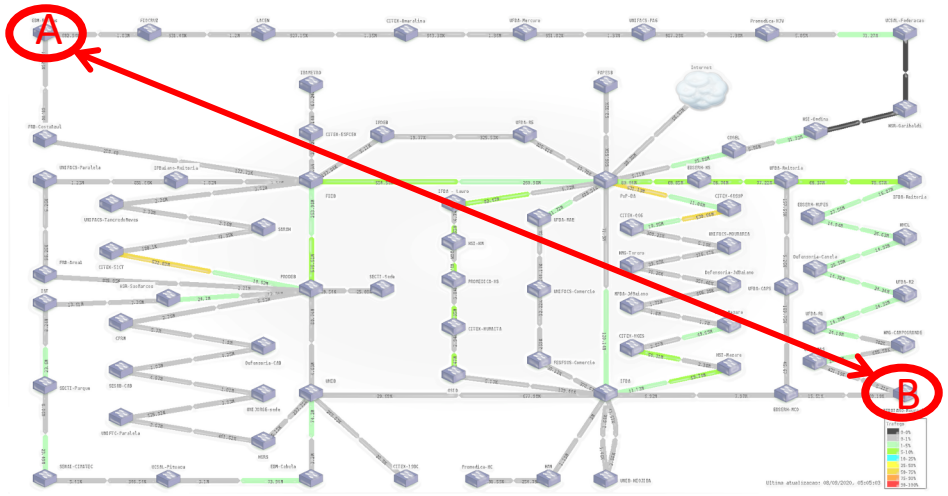
Problem statement

Layer-2 only networks

- Single failure/broadcast domain, slow convergence, network loops, etc.
- Single tasks like **VLAN provisioning** can become a nightmare in bigger/complex topologies
- Adding redundancy to an extended Ethernet network typically means relying on STP to keep the topology loop free
- Troubleshooting can be tricky when dealing with complex topologies (multi-rings, metro networks, campus networks, etc.)



Source: www.cisco.com



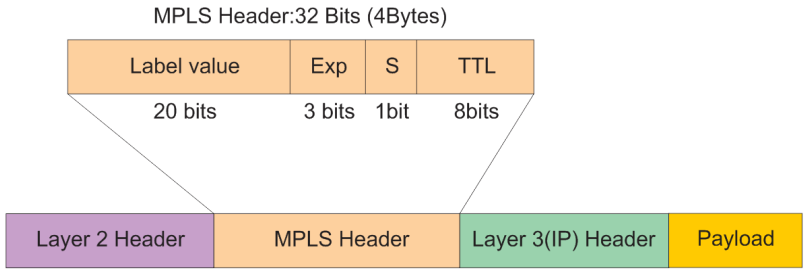
Source: remessa.pop-ba.rnp.br

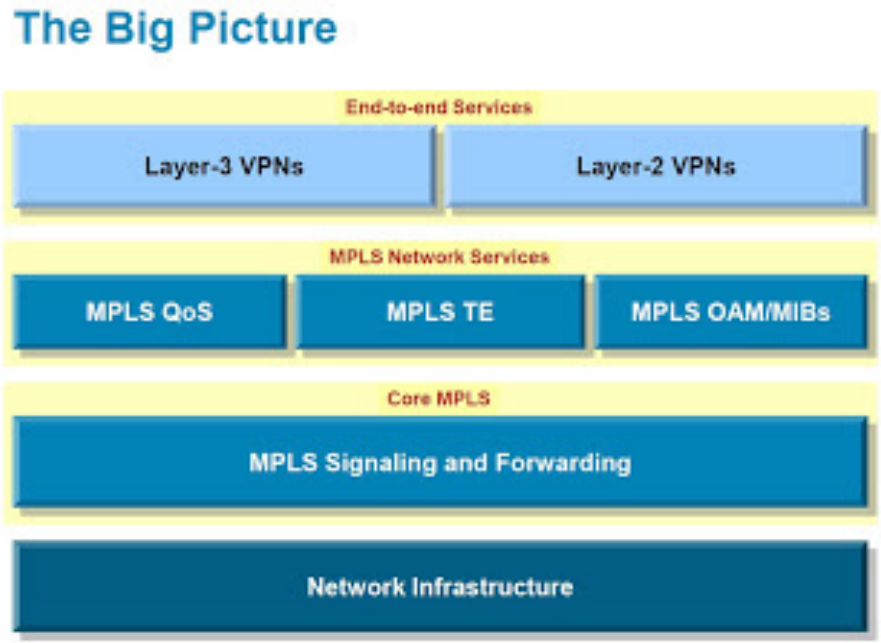
MPLS migration: why we need it?

- **Business perspective**
  - Reduce costs (CAPEX), by consolidating a single network for multiple Layer-2/3 services
  - Support increasingly stringent SLAs
  - Handle increasing scale/complexity of IP/non-IP based services

- **Technical perspective**
  - Need for network segmentation (users, applications, etc.)
  - Need for easier configuration of site-to-site and external WAN connectivity
  - **Need to sleep in peace**, without having to worry about L2-loops, link flaps/failures (and hence, users calling you in the middle of the night ;-))

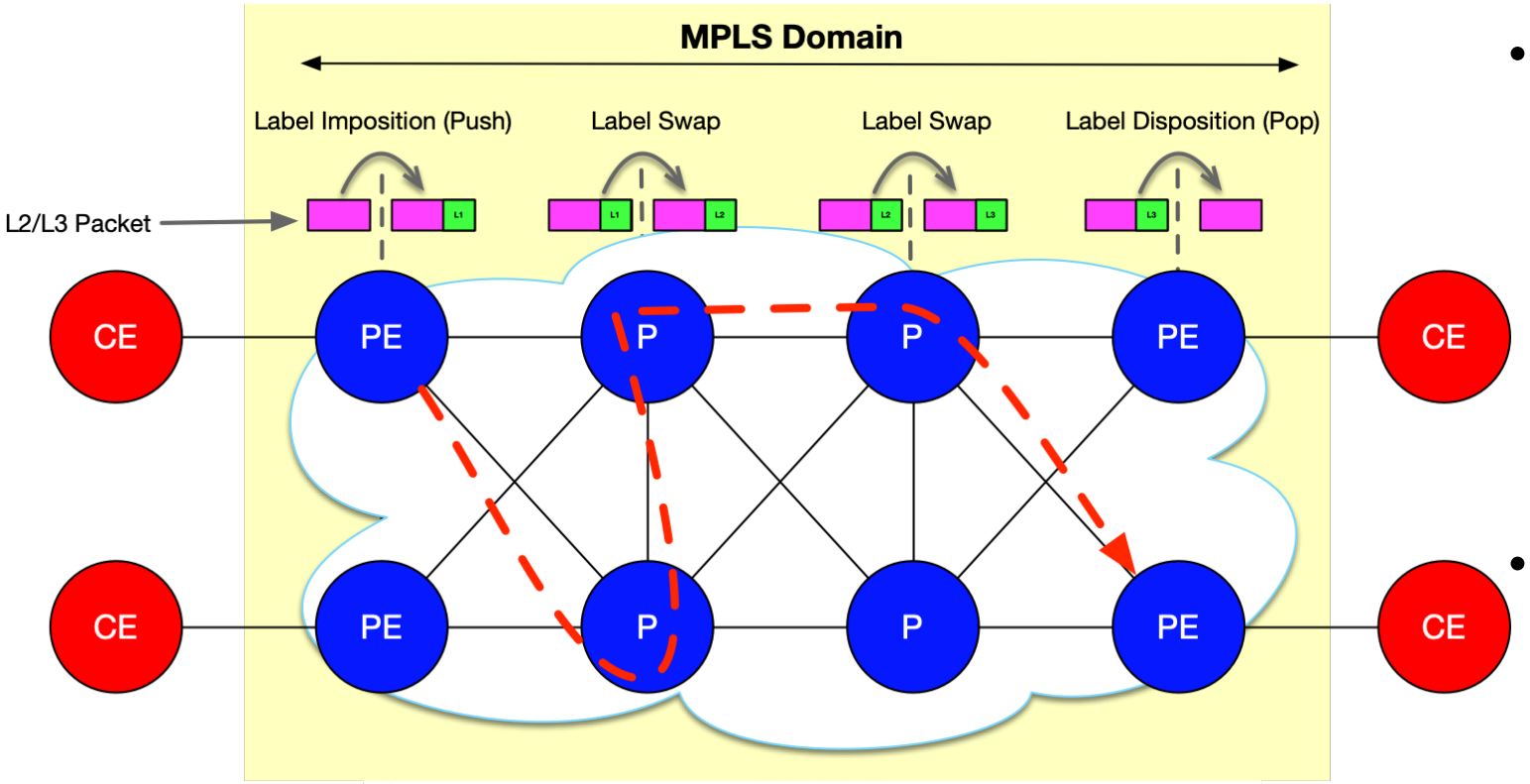
MPLS recap in one slide: basic building blocks



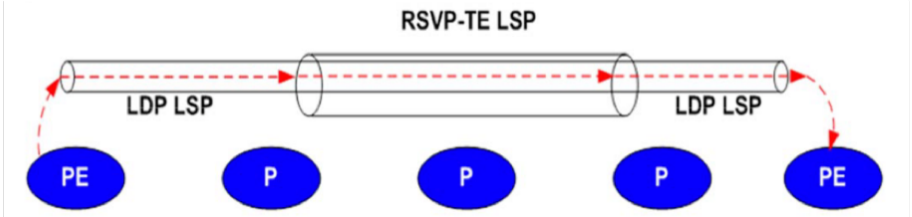


- Network Infrastructure
  - MPLS-enabled network devices: Label Switched Routers (LSR)
- Core MPLS
  - Label Switched Path (LSP)
  - Forwarding Equivalence Class (FEC)
  - Control Plane
    - Label assignment & distribution
      - LDP, RSVP, BGP, etc
  - Forwarding Plane
    - Label operations: push, swap, pop
- Other related protocols and protocols to exchange information
  - Between MPLS-enabled services

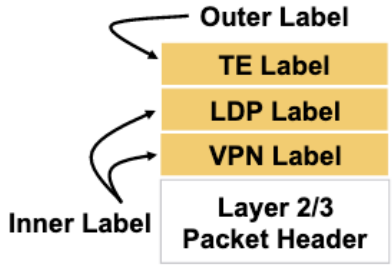
A day in the life of a MPLS packet



CE: Customer Edge
PE: Provider Edge
P: Provider



- It's all about labels
  - More than one label can be used for MPLS packet encapsulation
    - Each label in stack used for different purposes
    - Outer label always user for switching MPLS packets in network
    - Remaining inner labels used to specific services/FECs, etc.
  - **MTU size is important!!!**
- Allows building **services** such as
  - MPLS VPNs: LDP + VPN label
  - Traffic Engineering (FRR): LDP + TE label
  - VPNs over TE core: LDP + TE + VPN label
  - Any transport over MPLS: LDP + PW label



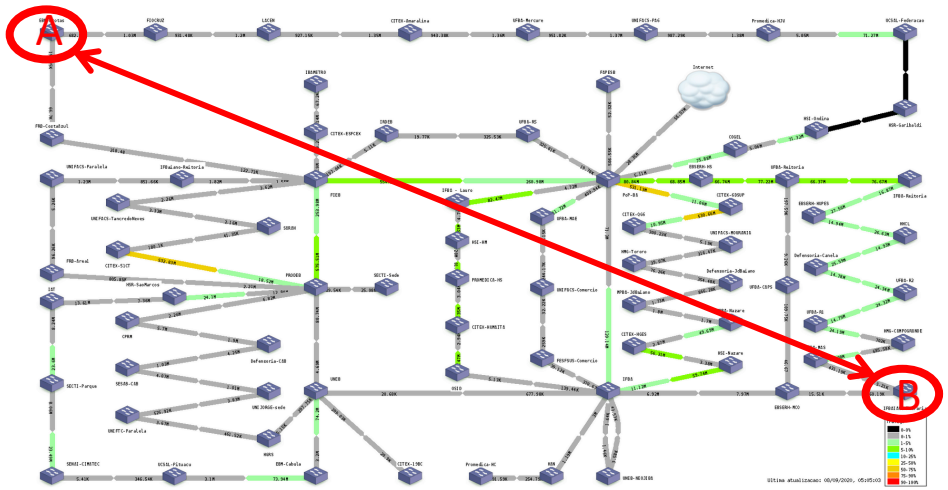
5 MPLS migration challenges

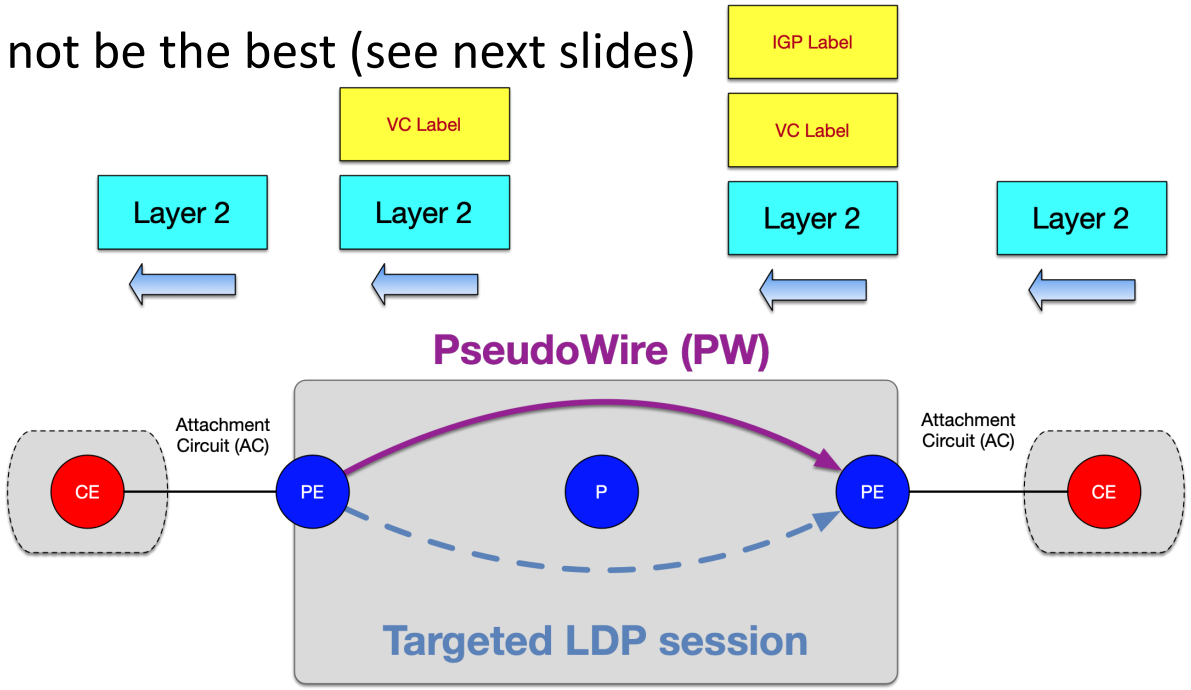
- As in any other migration there are challenges which needs to be addressed. Here I came up with 5 (as the magical number ☺)
  1. Mindset change
  2. Traffic Engineering
  3. Scaling the MPLS network
  4. Maximizing the MPLS network utilization and efficiency
  5. MPLS in the SDN era

MPLS migration: challenge #1

Mindset change

- **Your "VLAN pet" is dead -** I know you will miss STP/EAPS/REP, etc ☺
  - Start thinking now about **services** traversing the network, rather than **VLANs**.
  - L3 p2p links between devices
  - IGP + (labelling protocol) are now your best friends
    - Loopback reachability/label mapping distribution across the MPLS network
    - LDP is the easiest way to start – but may not be the best (see next slides)

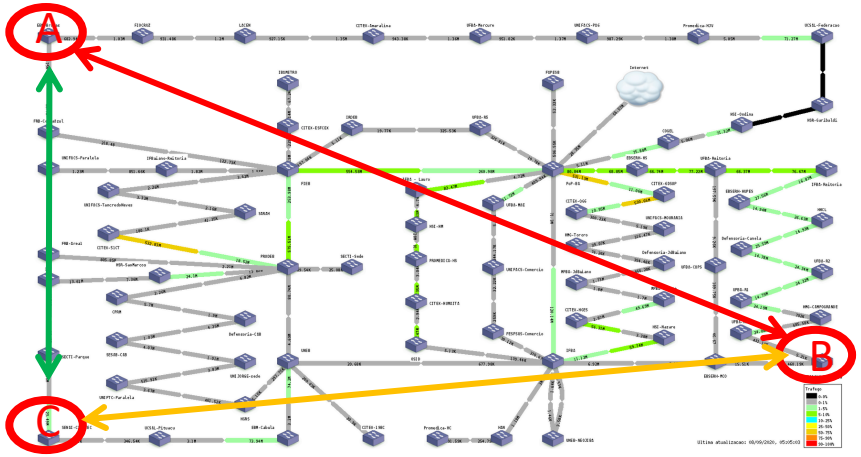


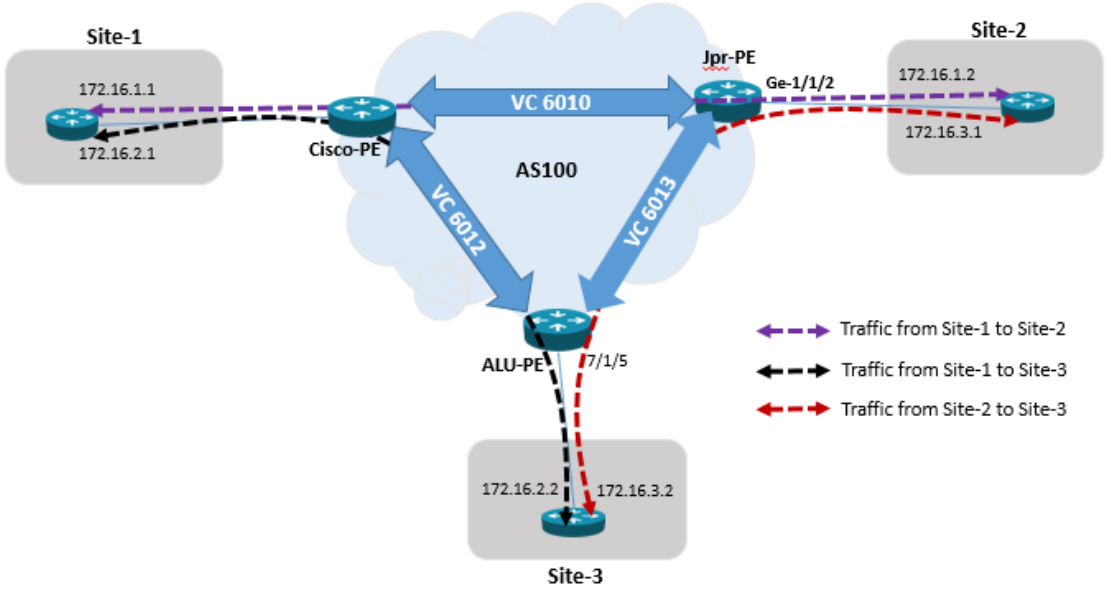


MPLS migration: challenge #1 [cont]

Mindset change

- **Welcome to the (wonderful) VPN world**
  - L2VPN (p2p: VPWS & p2mp: VPLS) – (a.k.a L2 Virtual-Circuit, EoMPLS, etc. depending on your vendor)
    - LDP signalled: easy to configure, but may not scale in large/complex scenarios (high number of pw's)
    - BGP signalled (VPLS only): more complex, has nice features such as PE autodiscovery, etc.
    - Ethernet VPN: next-generation L2VPN (RFC 7432) - use BGP to learn mac-address
  - L3VPN
    - MP-BGP, CsC, etc.

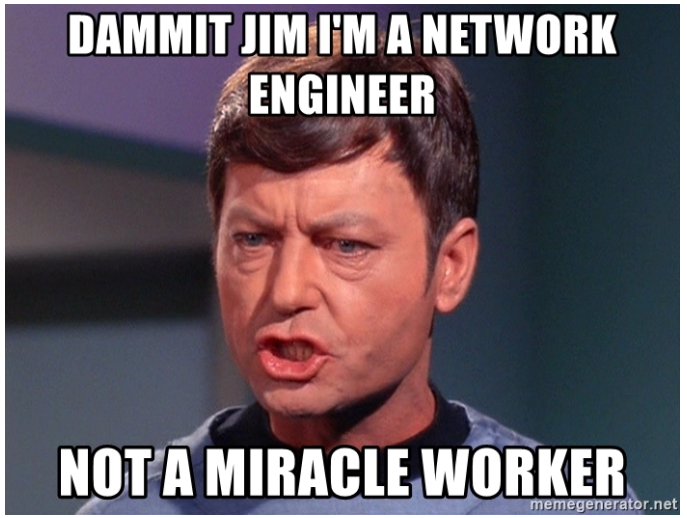




MPLS migration: challenge #2

Network vs Traffic Engineering

- Network Engineering
  - Build your network to carry your predicted traffic
- Traffic Engineering
  - Manipulate your traffic to fit your network



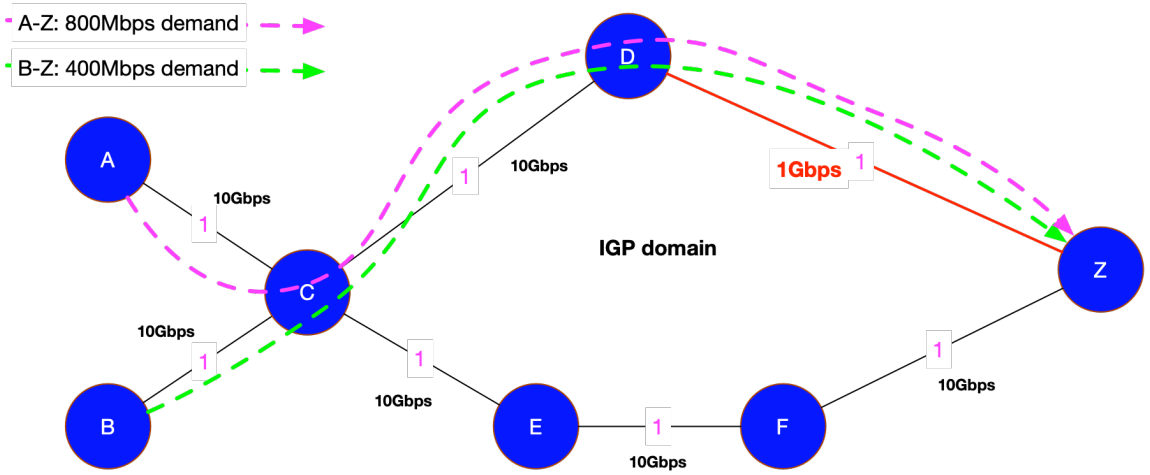
What is the point here?

- Traffic patterns are impossible to accurately predict
- Symmetric bandwidths/topologies, asymmetric load
- IP forwards based on destination IP address
  - This can sometimes not be granular enough, and cause unequal network load
  - This can cause temporary routing loops during network failure

MPLS migration: challenge #2 [cont]

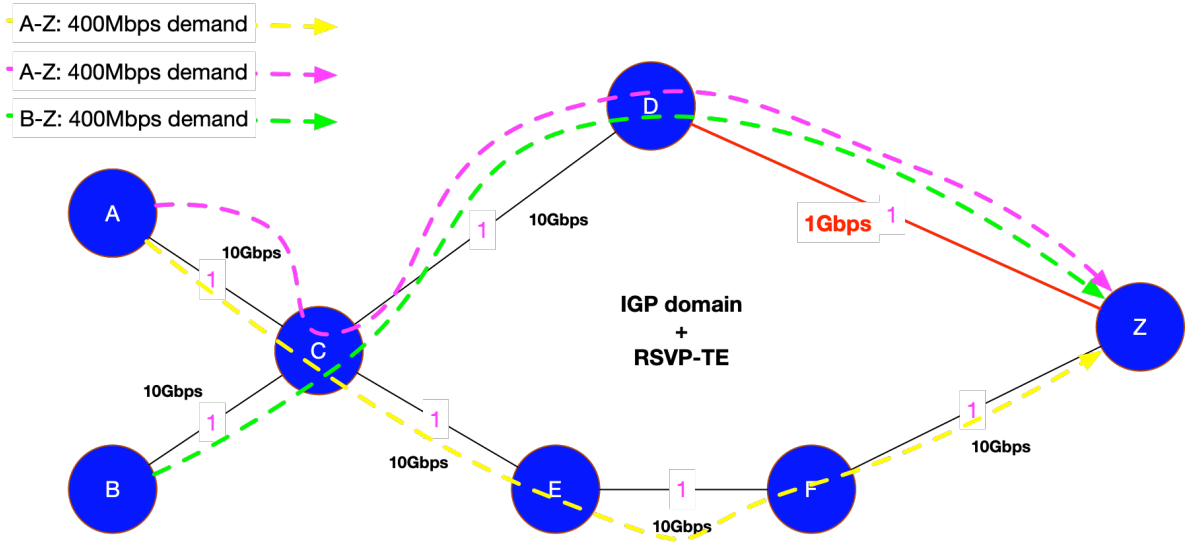
The classical "fish" problem

- IP (mostly) uses Destination-Based Least-Cost Routing
- Alternate path under utilized
- Links Oversubscription / Network congestion ➔ massive packet drops!!
  - Also an issue on LDP-based MPLS networks



How MPLS TE solves the problem?

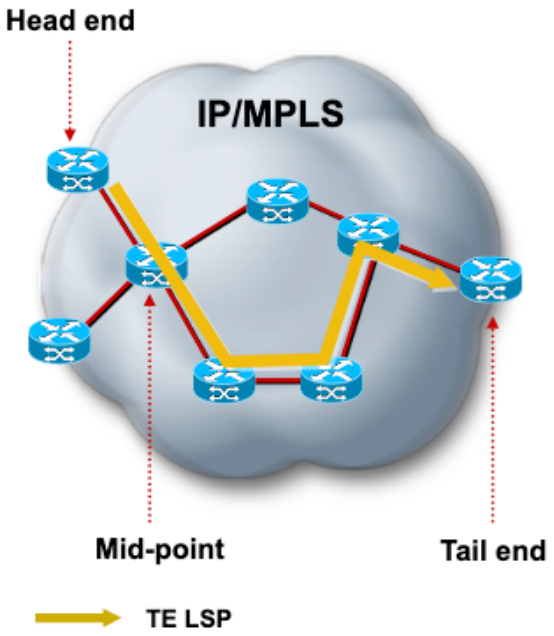
- Router A sees all links
- Router A computes paths on properties other than just shortest cost; and create 2 tunnels
- No link oversubscribed



MPLS migration: challenge #2 [cont]

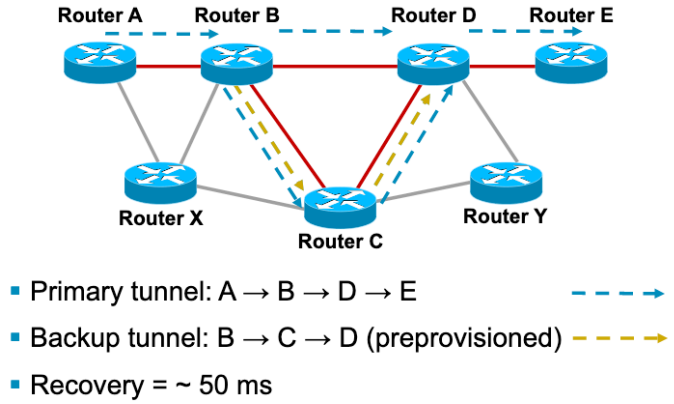
What is MPLS TE?

- Enhanced network availability, utilization, and performance.
- RSVP-TE: a signalling protocol to setup MPLS LSPs
- Widely deployed
- Enables
  - Traffic Engineering
  - Bandwidth Accounting
  - Fast failure protection (<50ms)



MPLS FastReRoute (FRR)

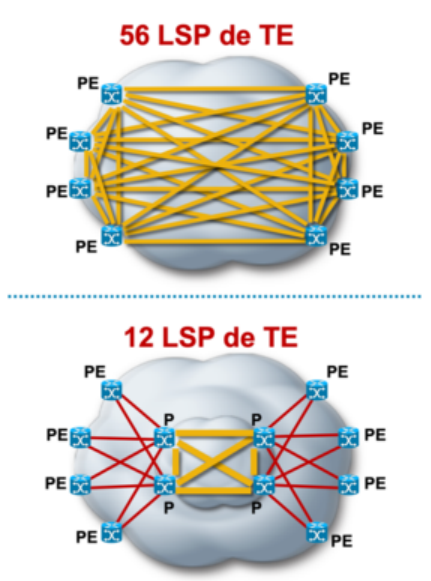
- This is one of the beauties of Traffic Engineering with MPLS
  - Admin-groups, SRLGs, etc.



MPLS migration: challenge #3

Scaling the MPLS network

- Configuration/automation
  - Increase in network-size has led to increased LSP scale
  - Scale magnifies operational issues
    - N^2 LSPs between N routers problem
    - The overhead of maintaining control-plane state – in the entire network, and per LSR
  - Monitoring at scale
    - What events are happening on the LSPs?
    - Which ones needs operator attention?
    - What properties of the LSPs are changing?



- Túneis PE-para-PE
  - Controle mais granular sobre o tráfego de rede
  - Maior quantidade de LSPs para TE
- Túneis P-para-P
  - Requer LDP sobre túneis de TE para transportar tráfego tais como L2 ou L3 VPN, com uso de label stack
  - Menor quantidade de LSPs para TE

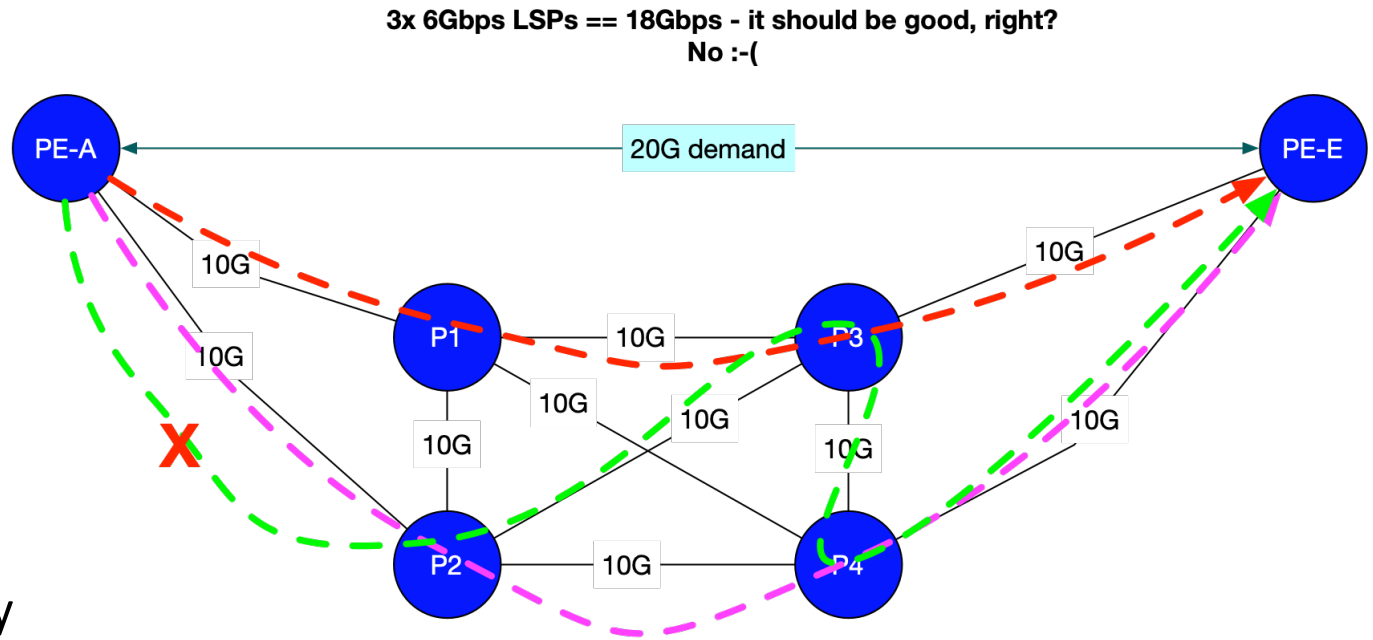
Source: Leonardo Furtado/BPF – wiki.brasilpeeringforum.org

MPLS migration: challenge #3 [cont]

Scaling the MPLS network [cont]

- Large LSPs can't fit small pipes
  - How big of an LSP to set up?
  - A single big LSP or several small ones?
    - If the goal is to maximize link utilization, smaller LSPs are better
  - Fork more parallel LSPs

- Auto-bandwidth issues
  - AutoBW: Automates the process of monitoring and online adjustment of LSP bandwidth
    - No all vendors implement it correctly ☹
  - How to improve "appropriateness" of resizing of the LSPs?

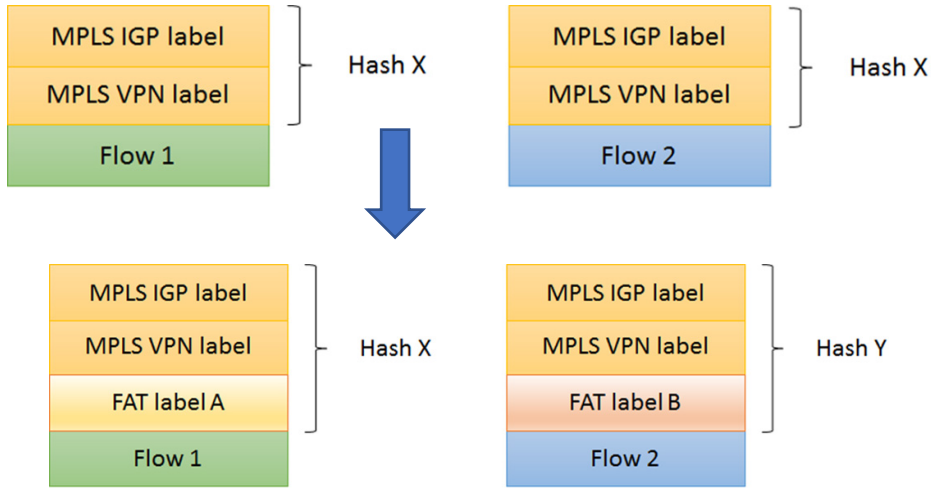


MPLS migration: challenge #4 [cont]

Maximizing the MPLS network utilization

- Per-flow load-balancing model: Traffic polarization
  - LAG/LACP ports
    - if the hashing algorithm is not efficient all the flows will pass through the same member of the LAG
  - ECMP (Equal-Cost Multi-Path) groups can be affected too!

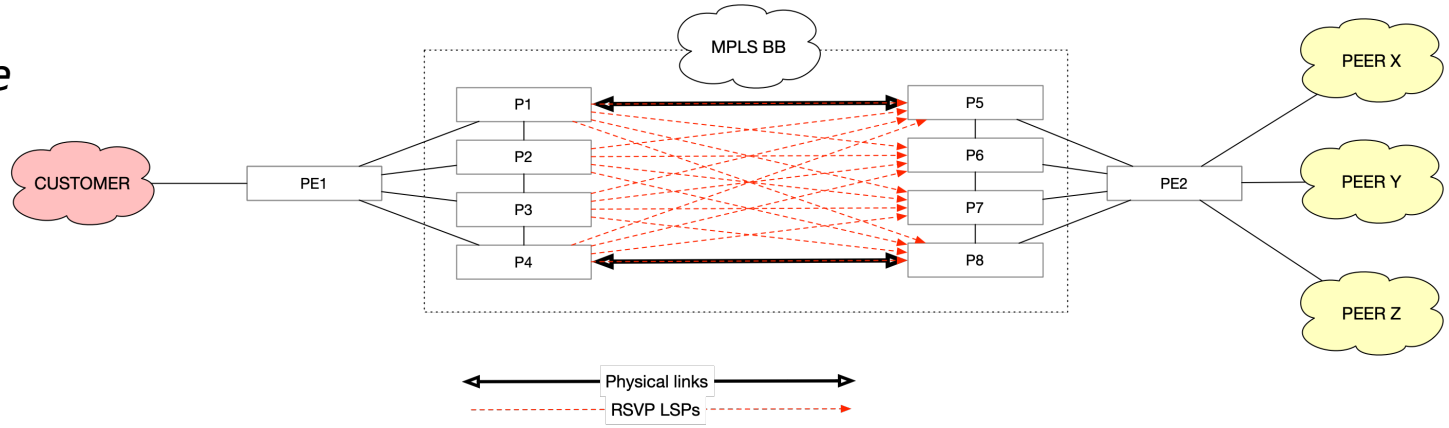
- In the IETF, they worked on solutions in order to fix this issue:
  - **Entropy Label**, RFC 6790
    - L2VPN/L3VPN
    - All LSRs **must** support it ☹
  - **Flow-Aware Transport (FAT) Label (or FAT-PW)**, RFC 6391
    - Easy to deploy (only on PEs), but L2VPN only



MPLS migration: challenge #4 [cont]

Maximizing the MPLS network utilization [cont]

- How to load-share traffic towards remote BGP next-hops/PEs using all LSPs as part of an ECMP group?
  - *"In JunOS if no specific metric is configured, an LSP attempts to track the IGP metric toward the same destination (the "to address" within the LSP configuration)." [1]*
  - In short: only RSVP LSPs signalled over the best/short IGP path will be used to carry traffic (metric inheritance)
    - LSPs in idle state
    - LSP traffic polarization
  - Possible solution: LSP metric manipulation

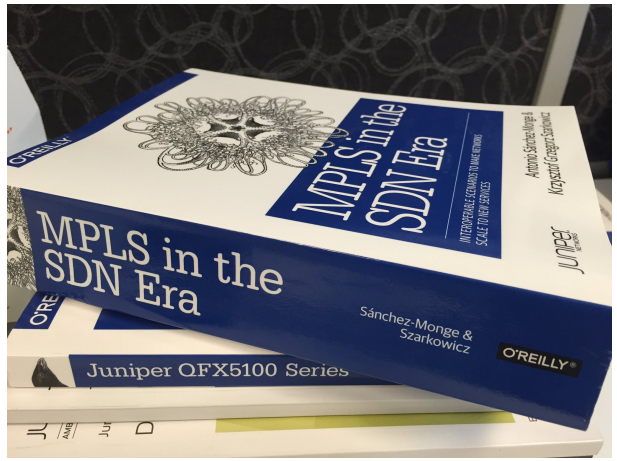


[1] https://www.juniper.net/documentation/en_US/junos/topics/topic-map/basic-lsp-configurtion.html#id-10233

MPLS migration: challenge #5

MPLS in the SDN era :p

- OpenFlow did not replace MPLS
- MPLS and SDN are not competing technologies: MPLS is a key SDN enabler.
- Segment Routing (SR) seems to be a strong candidate to replace RSVP-TE
  - Maximum SID depth hardware/software limitations



Available at: https://www.amazon.com

Key MPLS applications in the SDN era

- Centralized Traffic Engineering
  - Topology discovery
    - BGP, ISIS, OSPF, PCEP, etc
  - Path computation
    - Path computation algorithms
  - Path installation
    - PCEP, SR-TE, Netconf, etc.
  - OpenSource options:
    - ONOS
    - OpenDayLight
  - Some proprietary options:
    - Juniper Northstar controller
    - Cisco Open SDN controller

Summary and Key takeaways

- **MPLS is a mature technology with widespread deployments**
  - Suitable both for SP and enterprise networks
  - Two types of MPLS users
    - Indirect (subscriber): MPLS used as transport for subscribed service
    - Direct (DIY): MPLS implement in (own) SP or Enterprise network

- **It's all about labels…**
  - Label-based forwarding and IP protocol extensions for label exchange
  - Best of both worlds… L2-type forwarding and L3 control plane
  - Mind about MTU size (jumbro frame is your friend)

- **Key applications of MPLS is to implement VPN services**
  - Secure and scalable layer 2 and 3 VPN connectivity

- **MPLS supports advanced traffic engineering capabilities**
  - QoS, bandwidth control, and failure protection

- **MPLS & SDN are your good friends too!**

14 A 18 DE SETEMBRO DE 2020

Send questions, comments, and complaints to:

humbertogaliza [at] gmail [dot] com

Thanks for your patience!

PoP-BA Ponto de Presença da RNP na Bahia

RNP 30 anos ORGANIZAÇÃO SOCIAL DO MCTI

MINISTÉRIO DO TURISMO

MINISTÉRIO DA DEFESA

MINISTÉRIO DA SAÚDE

MINISTÉRIO DA EDUCAÇÃO

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

PÁTRIA AMADA BRASIL GOVERNO FEDERAL