



## Como otimizar o balanceamento de serviços L2VPN em redes metro utilizando o FAT PW

20 de março de 2020

## **ÍNDICE DA APRESENTAÇÃO**

- **Sobre a PS Network Experts**
- **Quem sou eu**
- **Evolução das redes metro**
- **O problema**
- **ECMP: escolha do caminho**
- **Solução para otimizar o balanceamento**
- **Demonstração**

## SOBRE A PS

---

A PS Network Experts é uma prestadora de serviços profissionais que atua nas áreas de:



REDES DE DADOS E VOZ



PROVEDORES DE SERVIÇO DE INTERNET



DATA CENTER



SEGURANÇA



COLABORAÇÃO



**Diferencial:** desenvolvemos, implantamos e suportamos projetos **em ambiente multivendor, sem qualquer restrição a fabricantes**

# Alberto Noronha, CCIE #17533

- Trabalha há 7 anos na PS Network Experts
- Há mais de 10 anos desenvolve e implanta projetos de backbone IP/MPLS e redes metro em grandes operadoras, operadoras regionais e provedores
- Detentor de certificações Cisco CCIE



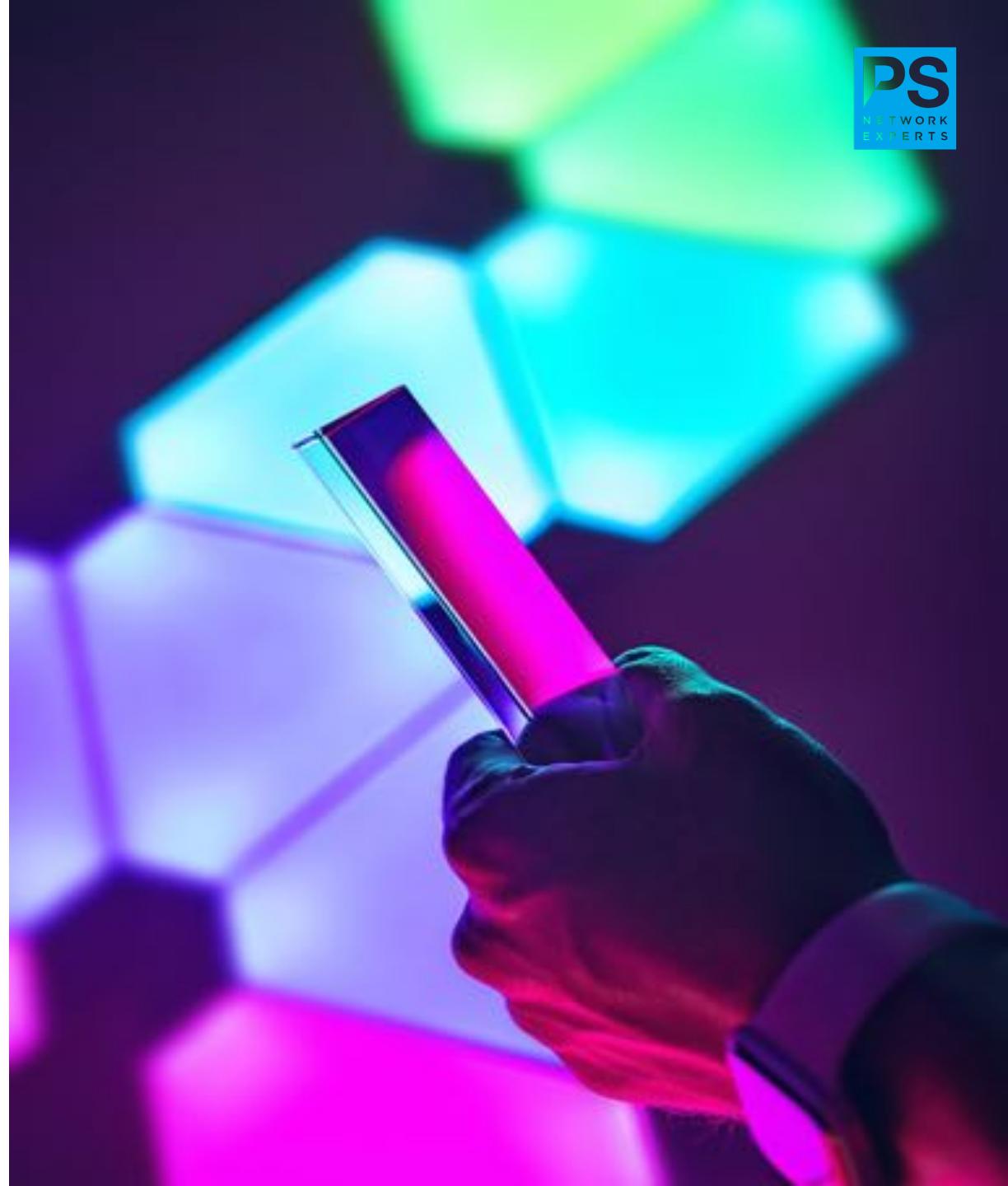
# Evolução das redes metro

- Migração do rádio para fibra obrigou a implementação de redes metro
- A redução do custo de aquisição de equipamentos permitiu a implementação de MPLS nas redes metro
- A concentração de serviços em roteadores de borda e em BRAS foi simplificada pela utilização de VPWS ou VPLS
- Devido ao alcance de interfaces de 10Gbps o custo de iluminar fibras com switches mostrou-se muito eficiente

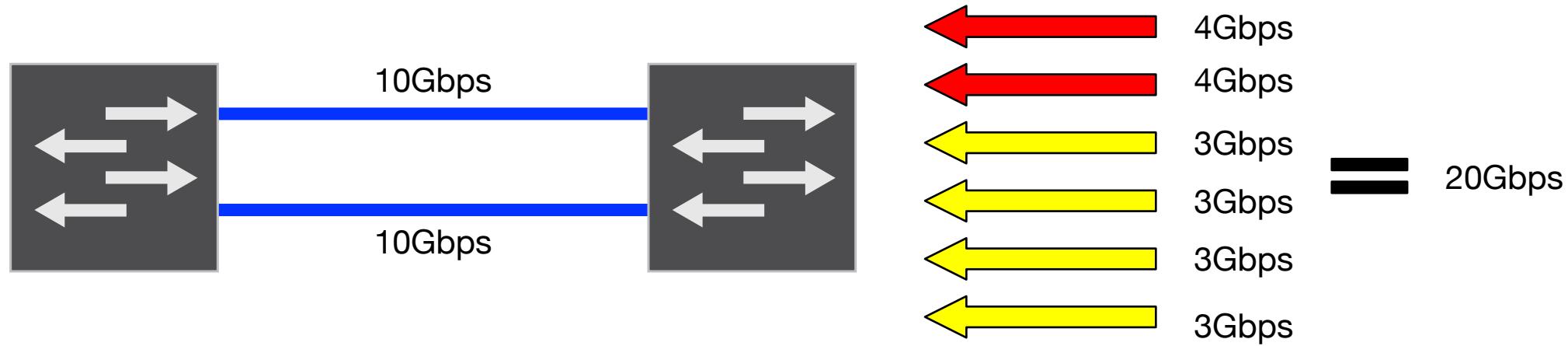


# O problema

- Crescimento exponencial da demanda por mais banda
- PW individuais consumindo cada vez mais banda
- Congestionamento de links, mesmo existindo caminhos paralelos (ineficiência do balanceamento de tráfego)



# Cenário do problema



# Escolha do caminho

---

O caminho, quando há caminhos paralelos (ECMP ou Equal Cost Multi-Path), é determinado pelo cálculo do hash

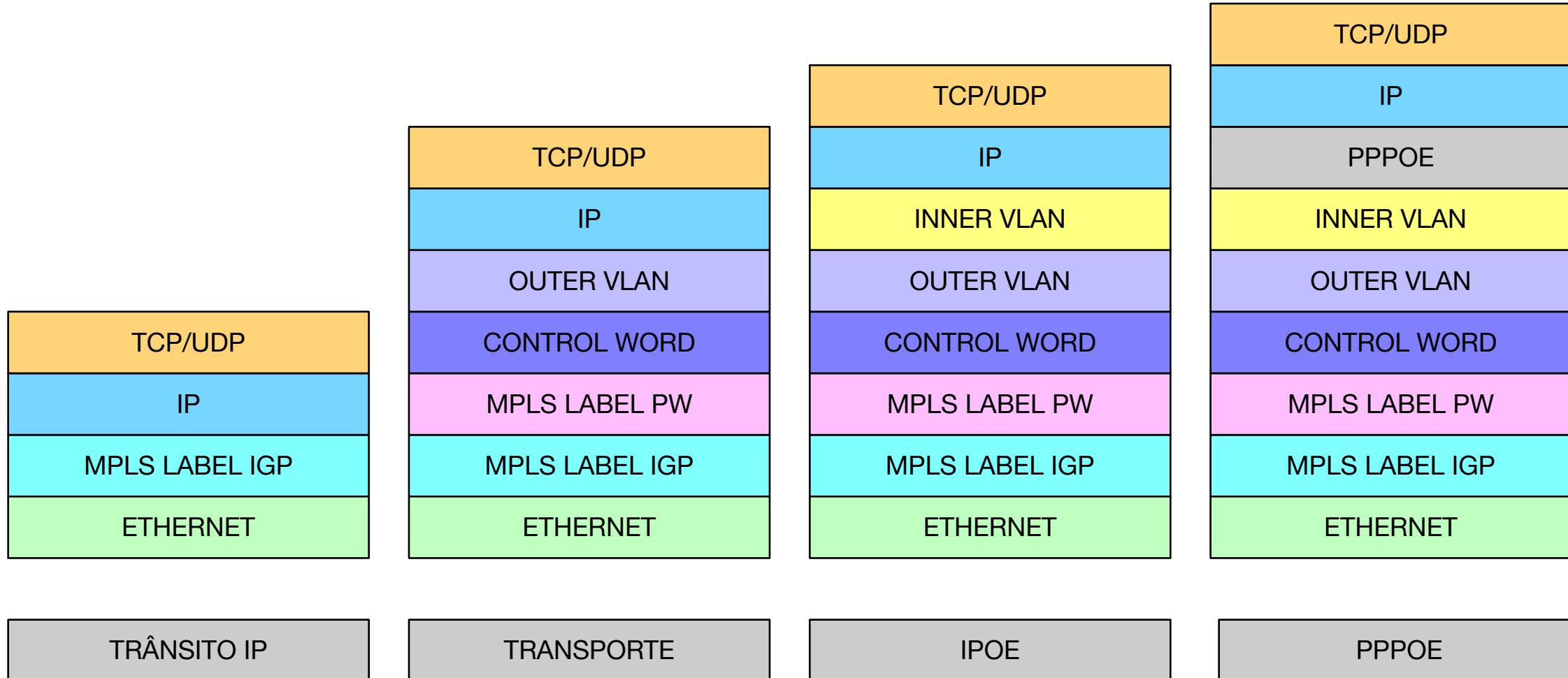
- O hash é computado sobre campos dos pacotes
- Os campos escolhidos dependem do tipo de tráfego sendo transportado
- O tráfego IP oferece muita variação (fluxos), principalmente porque as redes de provedores e operadoras possuem centenas de milhares de fluxos
- Balanceamento baseado em IP (layer 3), normalmente, permite balanceamento otimizado

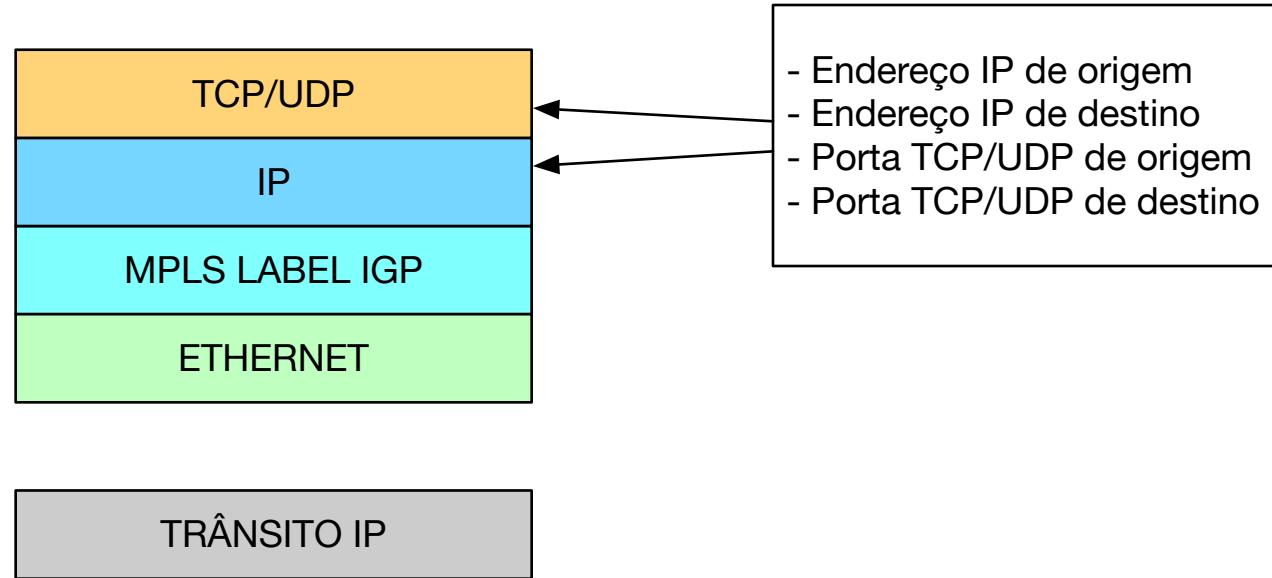
# Escolha dos campos para computação do hash\*

Tipo do tráfego	Campos utilizados para computação do hash
<b>IPv4</b>	Endereço de origem, endereço de destino, porta de origem e porta de destino
<b>IPv6</b>	Endereço de origem, endereço de destino, porta de origem e porta de destino
<b>IP sobre MPLS com menos de 4 labels</b>	Endereço de origem, endereço de destino, porta de origem e porta de destino
<b>IP sobre MPLS com 4 ou mais labels</b>	Último label da pilha (inner most label)
<b>Pacote não IP</b>	Último label da pilha (inner most label)

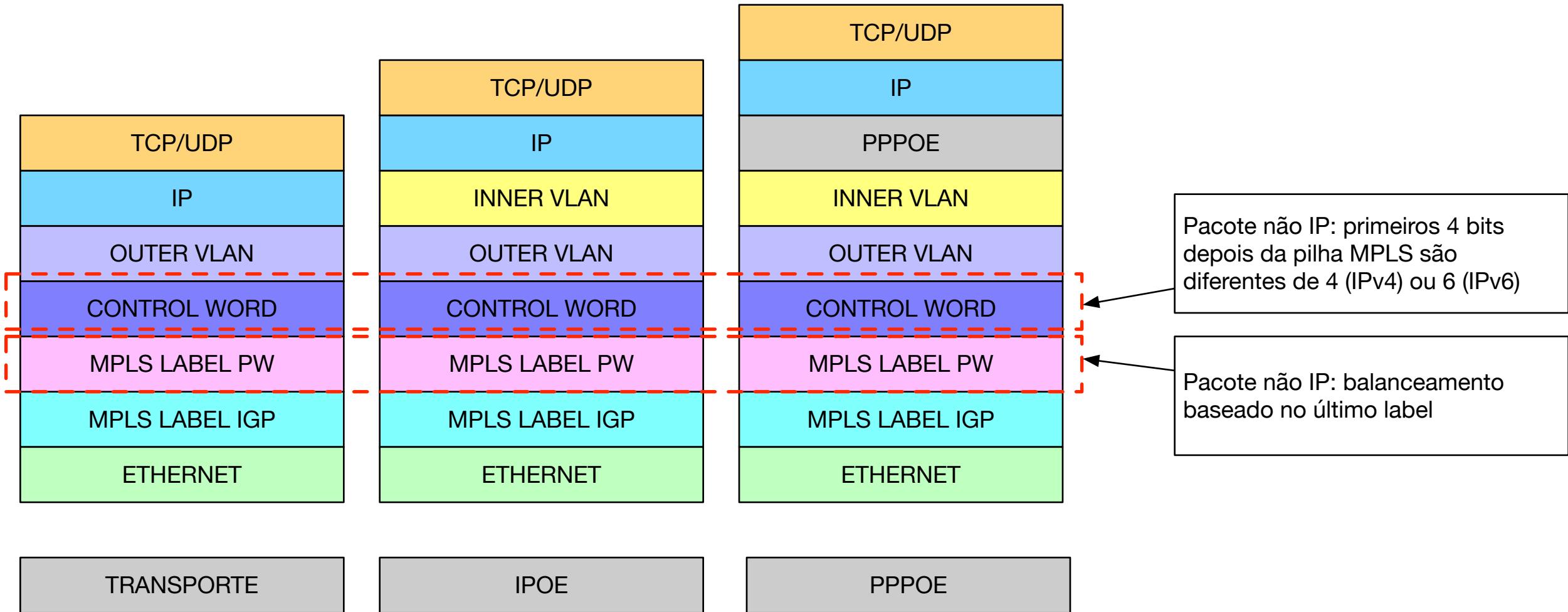
\*<https://community.cisco.com/t5/service-providers-documents/asr9000-xr-load-balancing-architecture-and-characteristics/ta-p/3124809#field>

# Exemplos de pilhas de protocolos





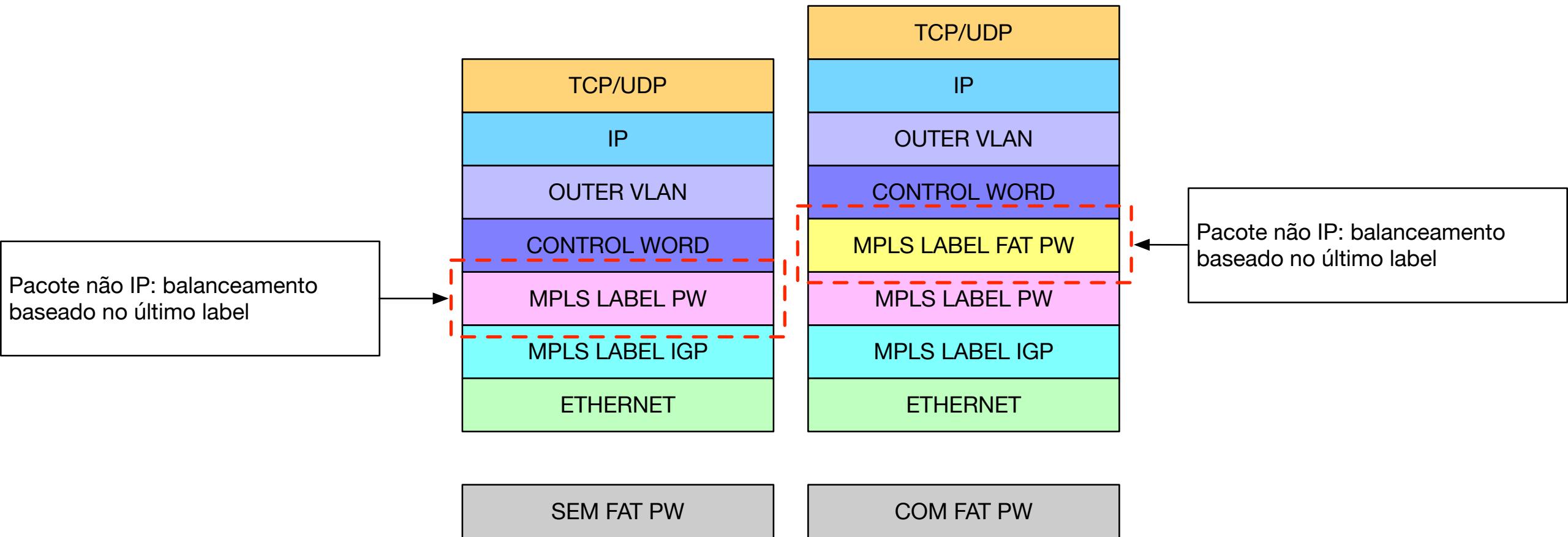
# O que são pacotes não IP



## RFC 6391 FAT PW (Flow-Aware Traffic of Pseudowires)

- Dependente do hardware nos equipamentos de borda
- Transparente para nós intermediários (roteadores P e outros PE)
- Poucos comandos para ativar
- Adiciona um label MPLS extra, cujo valor é baseado nos campos do pacote original (cada fabricante possui seu algoritmo)
- Equipamento que termina o túnel (PW) descarta o label extra

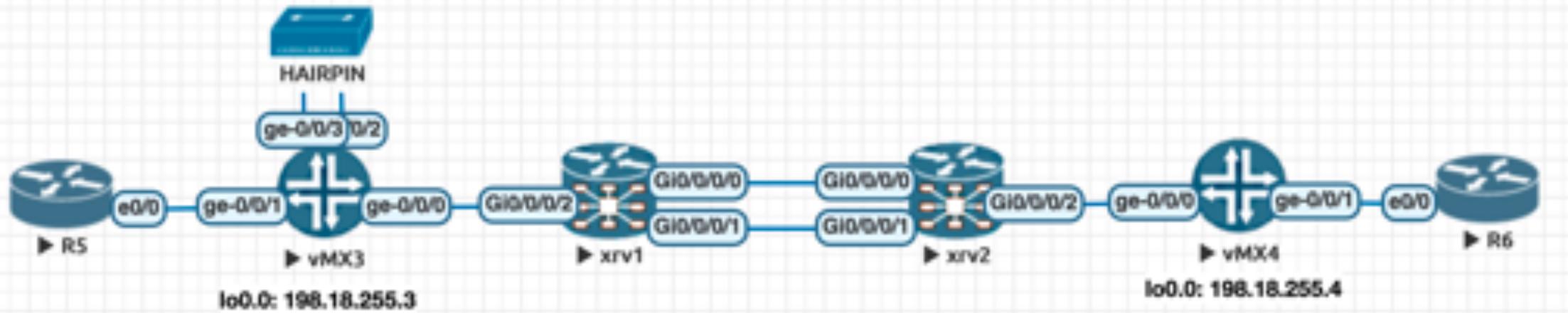
# Solução para otimizar o balanceamento (pacotes com e sem o FAT PW)



## Demonstração

- EVE-NG
- Juniper vMX (roteador PE)
- Cisco XRV (roteador P)
- Cisco IOL (gerador de tráfego)
- Wireshark (captura de pacotes)

# Topologia



# Validação pré FAT PW

```
root@vMX3> show l2circuit connections
```

```
Layer-2 Circuit Connections:
```

```
Neighbor: 198.18.255.4
```

Interface	Type	St	Time last up	# Up trans
ge-0/0/3.0(vc 500)	rmt	Up	Mar 9 19:44:42 2020	1
Remote PE: 198.18.255.4	Negotiated control-word: Yes (Null)			
Incoming label: 299824	Outgoing label: 299824			

```
Negotiated PW status TLV: No
```

```
Local interface: ge-0/0/3.0, Status: Up, Encapsulation: ETHERNET
```

```
Flow Label Transmit: No, Flow Label Receive: No
```

# Validação pré FAT PW

```
RP/0/0/CPU0:xrv1#show mpls forwarding labels 16002
```

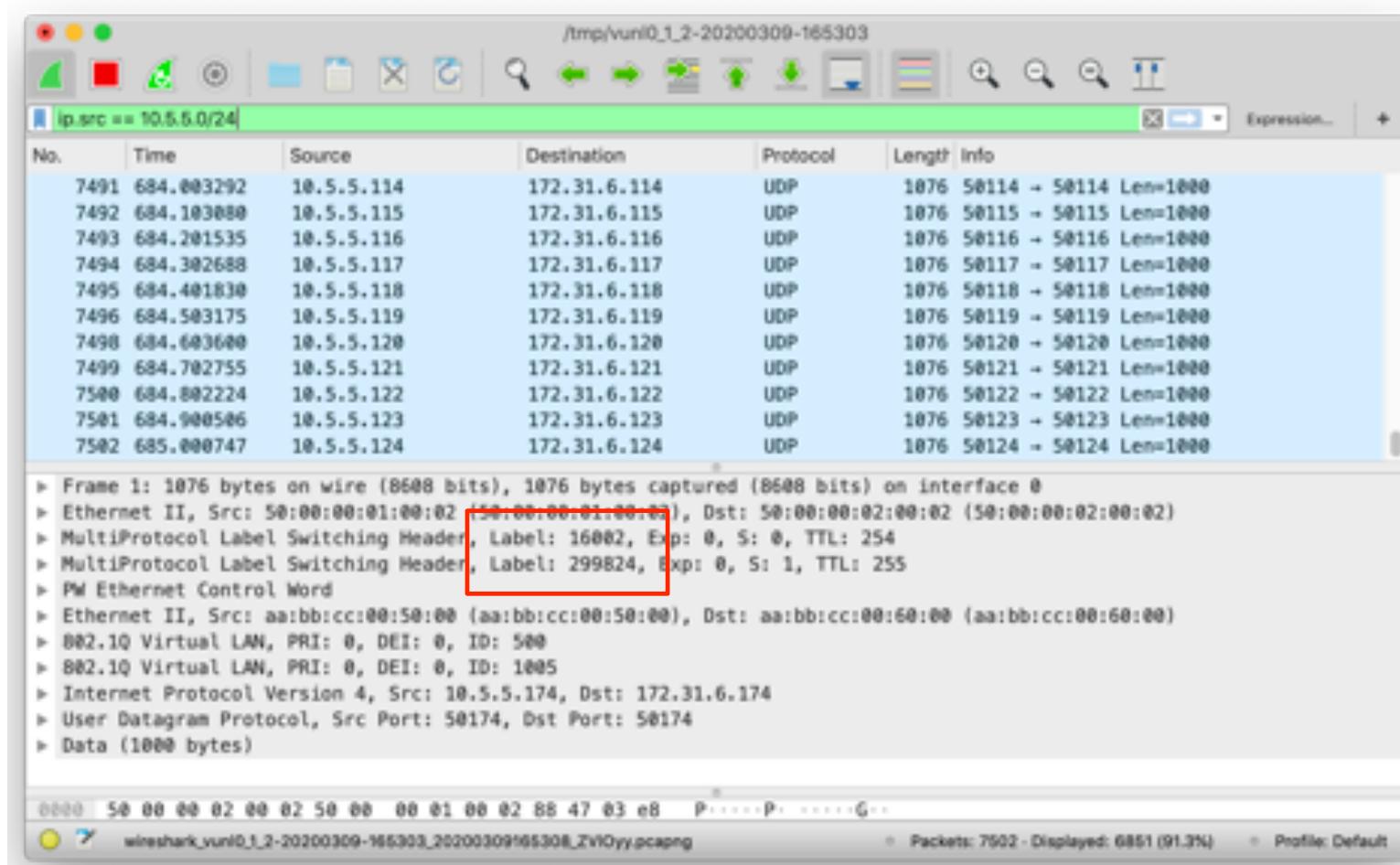
```
Mon Mar 9 20:18:26.046 UTC
```

Local Label	Outgoing Label	Prefix or ID	Outgoing Interface	Next Hop	Bytes Switched
16002	16002	198.18.255.4/32	Gi0/0/0/0	198.18.0.1	36226
	16002	198.18.255.4/32	Gi0/0/0/1	198.18.0.3	21534372

# Validação pré FAT PW

```
RP/0/0/CPU0:xrv1#show mpls forwarding exact-route label 16002 bottom-label 299824
Mon Mar  9 20:15:45.237 UTC
Local  Outgoing      Prefix          Outgoing      Next Hop      Bytes
Label  Label         or ID          Interface      Interface     Switched
-----  -----          -----          -----          -----          -----
16002  16002        198.18.255.4/32  Gi0/0/0/1    198.18.0.3   N/A
Version: 10, Priority: 1
Via: Gi0/0/0/1, Next Hop: 198.18.0.3
MAC/Encaps: 14/18, MTU: 1586
Label Stack (Top -> Bottom): { 16002 }
```

# Validação pré FAT PW



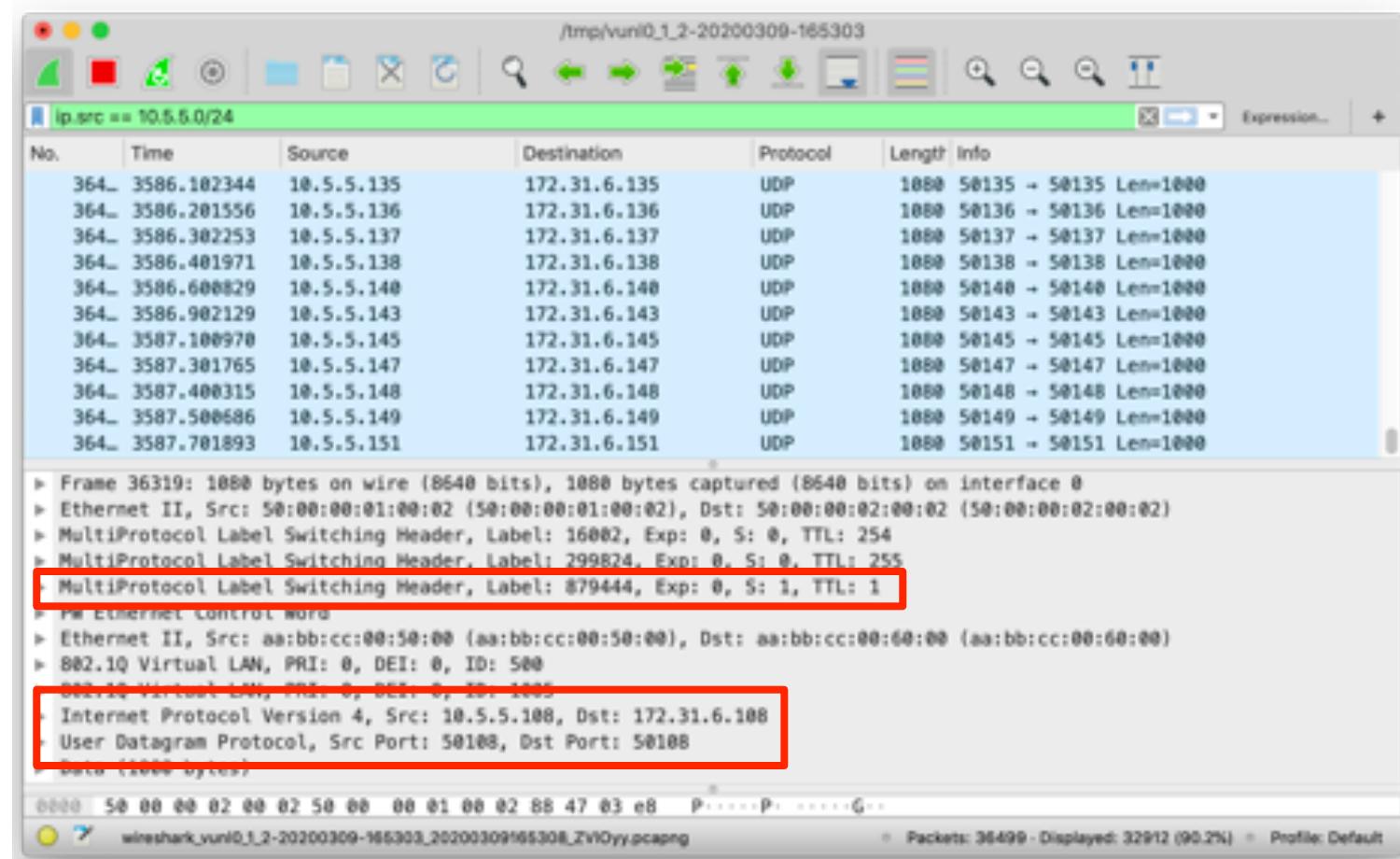
# Ativação do FAT PW

---

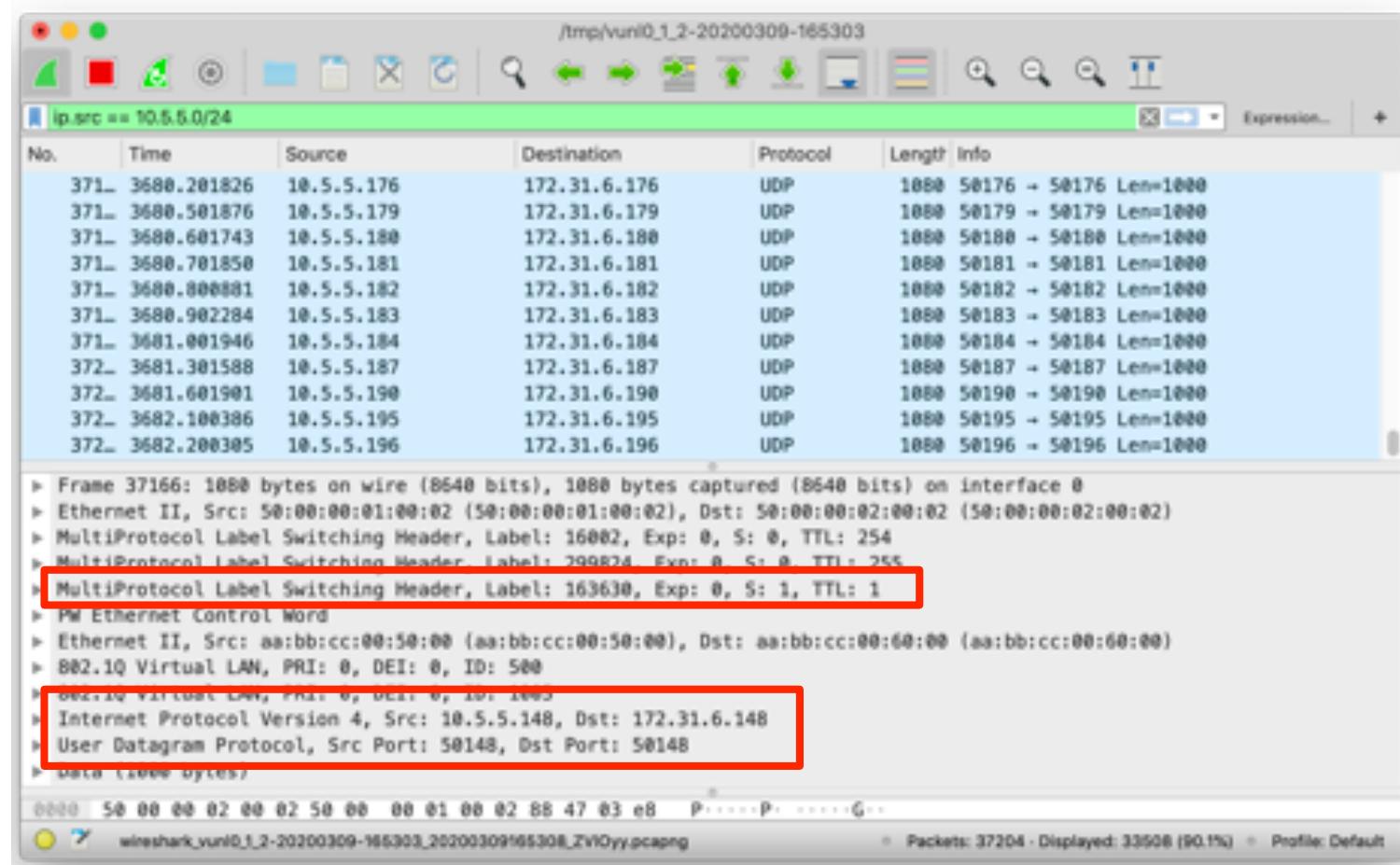
```
root@vMX3# show | compare
[edit protocols l2circuit neighbor 198.18.255.4 interface ge-0/0/3.0]
+      flow-label-transmit;
+      flow-label-receive;
```

```
root@vMX4# show | compare
[edit protocols l2circuit neighbor 198.18.255.3 interface ge-0/0/1.0]
+      flow-label-transmit;
+      flow-label-receive;
```

# Validação – Fluxo 01



# Validação – Fluxo 02



# Dúvidas?

Apresentação e os scripts podem ser encontrados em:  
<https://github.com/avargasn/fatpw>

<https://go.oncehub.com/FaleComExpert>

Schedule now



Online appointment scheduling

Pick a date and time that works for you -  
it's fast and easy.

[go.oncehub.com](https://go.oncehub.com)

# Obrigado

Alberto Noronha

[anoronha@psrv.com.br](mailto:anoronha@psrv.com.br)  
+55 11 97061-6523