nic.br cgi.br | ptt.br

Lima, Perú

May 22th, 2015

PTT.br - Solution for multiple connection of participants in one link

Regional Interconnection Forum 2015

ptt.br nic.br cgi.br

Julimar Lunguinho Mendes <julimar@nic.br>
PTT.br Engineering Team <eng@ptt.br>

Summary

The presentation intend to display the solution found by PTT.br team for multiple remote participants in connection peering structure, sharing the same link, mainly in the São Paulo city.
The main idea is to spend a bit of PTT.br team experience with this type of solution showing the advantages and disadvantages of each.

Reference

IXP - Internet eXchange Point

PTT – Ponto de Troca de Tráfego

Definitions

Participant
Autonomous System (AS) connected at least to one of the different PTT.br Locations (current 25)

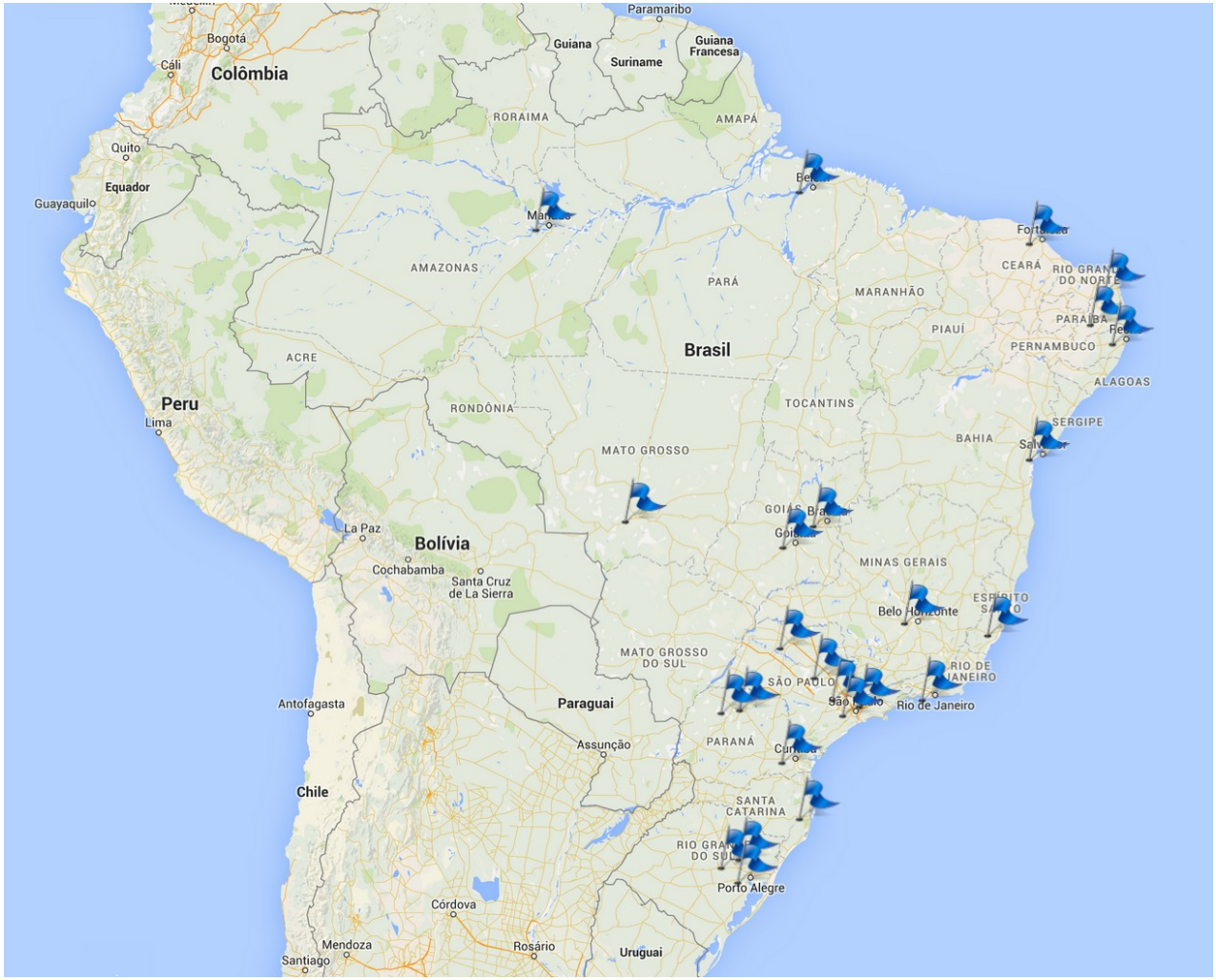
Transport
Physical link connecting the participant router to the selected PIX

PIX – Interconnection Point
PTT.br peering fabric edge that receive participants connections

MLPA - Multi-Lateral Peering Agreements
IP (v4/v6) traffic exchange between all participants.
BGP sessions established between participants and IXP Route Servers.
Different MLPA VLANs for IPv4 and IPv6.

PTT.br – Current locations

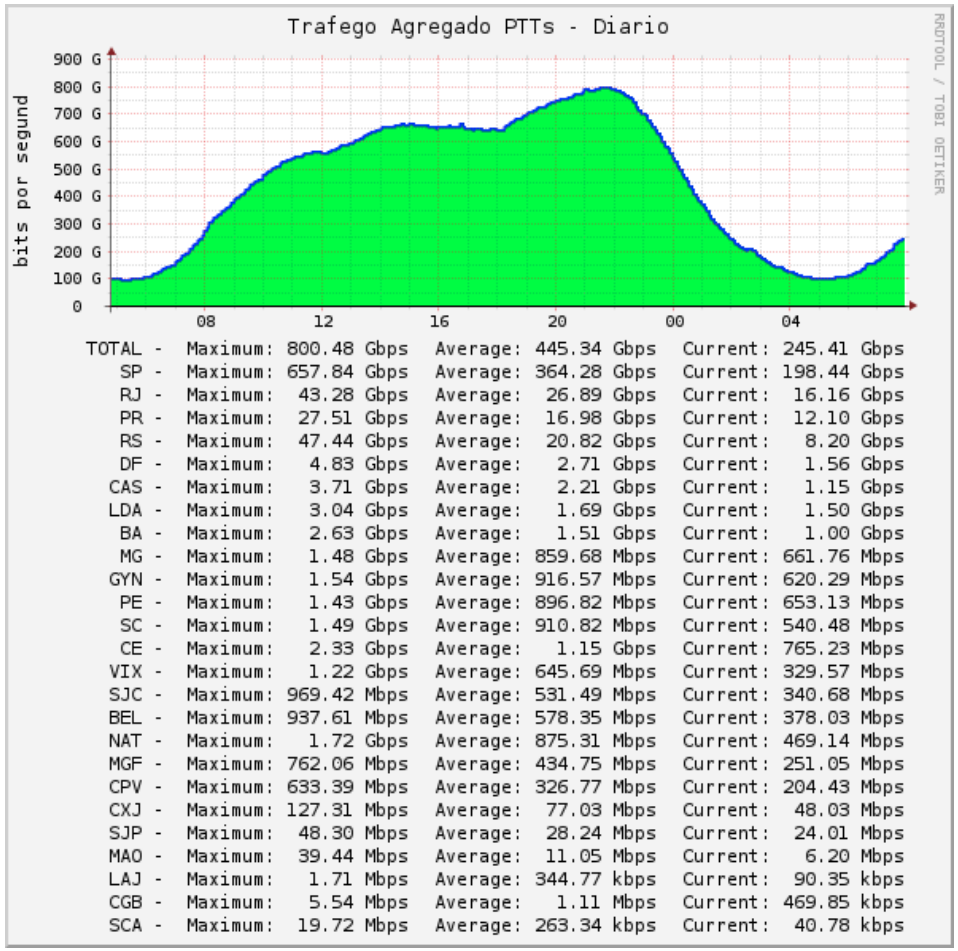
Each location is an independent IX



1	Belém
2	Belo Horizonte
3	Brasília
4	Campina Grande
5	Campinas
6	Caxias do Sul
7	Cuiabá
8	Curitiba
9	Florianópolis
10	Fortaleza
11	Goiânia
12	Lajeado
13	Londrina
14	Manaus
15	Maringá
16	Natal
17	Porto Alegre
18	Recife
19	Rio de Janeiro
20	Salvador
21	São Carlos
22	São José do Rio Preto
23	São José dos Campos
24	São Paulo
25	Vitória

The PTT.br currently has 25 locations not interconnected

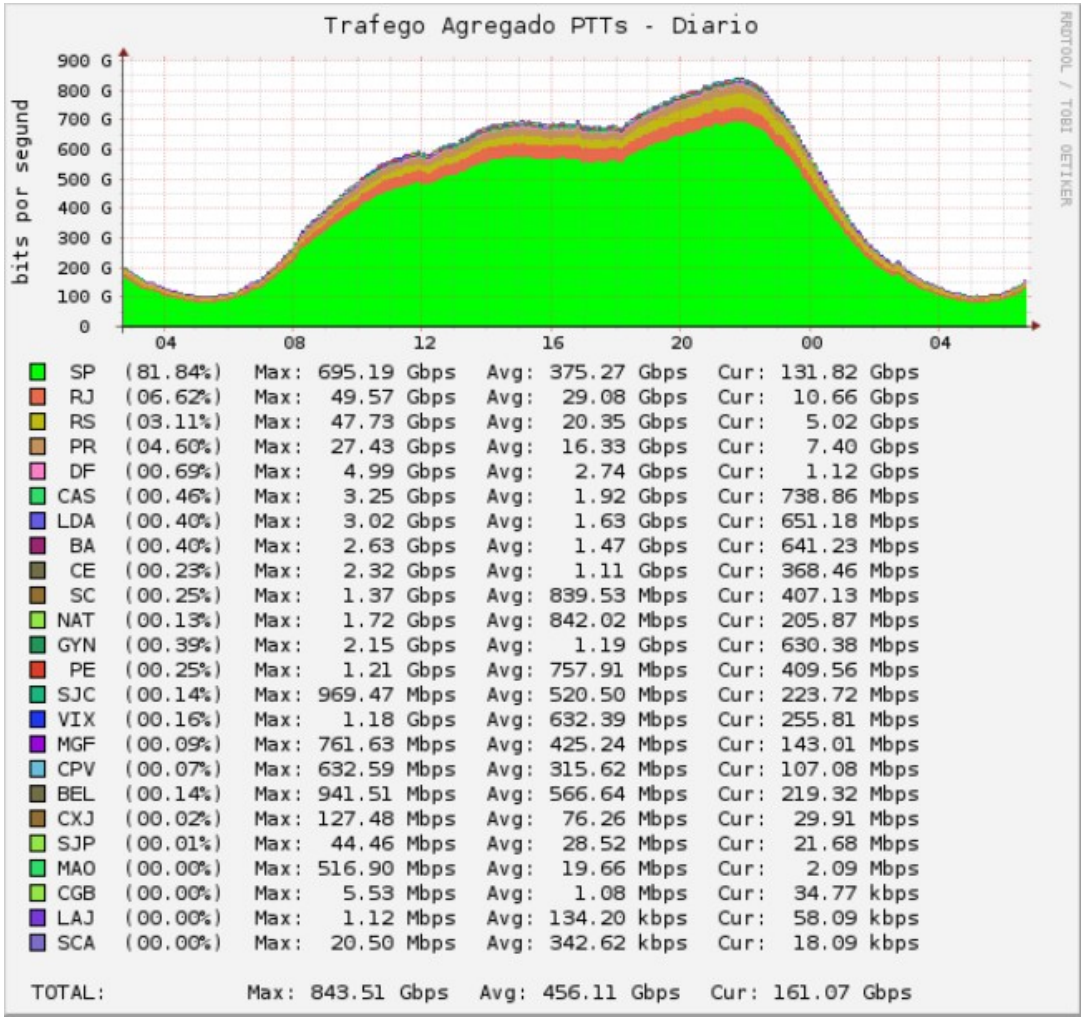
PTT.br – Aggregate all traffic locations



TOTAL -	Maximum:	800.48 Gbps	Average:	445.34 Gbps	Current:	245.41 Gbps
SP -	Maximum:	657.84 Gbps	Average:	364.28 Gbps	Current:	198.44 Gbps
RJ -	Maximum:	43.28 Gbps	Average:	26.89 Gbps	Current:	16.16 Gbps
PR -	Maximum:	27.51 Gbps	Average:	16.98 Gbps	Current:	12.10 Gbps
RS -	Maximum:	47.44 Gbps	Average:	20.82 Gbps	Current:	8.20 Gbps
DF -	Maximum:	4.83 Gbps	Average:	2.71 Gbps	Current:	1.56 Gbps
CAS -	Maximum:	3.71 Gbps	Average:	2.21 Gbps	Current:	1.15 Gbps
LDA -	Maximum:	3.04 Gbps	Average:	1.69 Gbps	Current:	1.50 Gbps
BA -	Maximum:	2.63 Gbps	Average:	1.51 Gbps	Current:	1.00 Gbps
MG -	Maximum:	1.48 Gbps	Average:	859.68 Mbps	Current:	661.76 Mbps
GYN -	Maximum:	1.54 Gbps	Average:	916.57 Mbps	Current:	620.29 Mbps
PE -	Maximum:	1.43 Gbps	Average:	896.82 Mbps	Current:	653.13 Mbps
SC -	Maximum:	1.49 Gbps	Average:	910.82 Mbps	Current:	540.48 Mbps
CE -	Maximum:	2.33 Gbps	Average:	1.15 Gbps	Current:	765.23 Mbps
VIX -	Maximum:	1.22 Gbps	Average:	645.69 Mbps	Current:	329.57 Mbps
SJC -	Maximum:	969.42 Mbps	Average:	531.49 Mbps	Current:	340.68 Mbps
BEL -	Maximum:	937.61 Mbps	Average:	578.35 Mbps	Current:	378.03 Mbps
NAT -	Maximum:	1.72 Gbps	Average:	875.31 Mbps	Current:	469.14 Mbps
MGF -	Maximum:	762.06 Mbps	Average:	434.75 Mbps	Current:	251.05 Mbps
CPV -	Maximum:	633.39 Mbps	Average:	326.77 Mbps	Current:	204.43 Mbps
CXJ -	Maximum:	127.31 Mbps	Average:	77.03 Mbps	Current:	48.03 Mbps
SJP -	Maximum:	48.30 Mbps	Average:	28.24 Mbps	Current:	24.01 Mbps
MAO -	Maximum:	39.44 Mbps	Average:	11.05 Mbps	Current:	6.20 Mbps
LAJ -	Maximum:	1.71 Mbps	Average:	344.77 kbps	Current:	90.35 kbps
CGB -	Maximum:	5.54 Mbps	Average:	1.11 Mbps	Current:	469.85 kbps
SCA -	Maximum:	19.72 Mbps	Average:	263.34 kbps	Current:	40.78 kbps

Daily peaks traffic of 800 Gbps
101 PIXes
861 participants

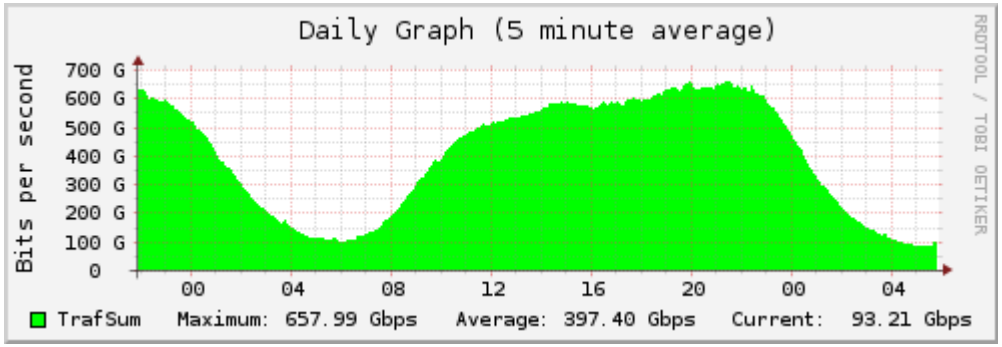
PTT.br – Stacked Locations Traffic



SP	(81.84%)	Max:	695.19 Gbps	Avg:	375.27 Gbps	Cur:	131.82 Gbps
RJ	(06.62%)	Max:	49.57 Gbps	Avg:	29.08 Gbps	Cur:	10.66 Gbps
RS	(03.11%)	Max:	47.73 Gbps	Avg:	20.35 Gbps	Cur:	5.02 Gbps
PR	(04.60%)	Max:	27.43 Gbps	Avg:	16.33 Gbps	Cur:	7.40 Gbps
DF	(00.69%)	Max:	4.99 Gbps	Avg:	2.74 Gbps	Cur:	1.12 Gbps
CAS	(00.46%)	Max:	3.25 Gbps	Avg:	1.92 Gbps	Cur:	738.86 Mbps
LDA	(00.40%)	Max:	3.02 Gbps	Avg:	1.63 Gbps	Cur:	651.18 Mbps
BA	(00.40%)	Max:	2.63 Gbps	Avg:	1.47 Gbps	Cur:	641.23 Mbps
CE	(00.23%)	Max:	2.32 Gbps	Avg:	1.11 Gbps	Cur:	368.46 Mbps
SC	(00.25%)	Max:	1.37 Gbps	Avg:	839.53 Mbps	Cur:	407.13 Mbps
NAT	(00.13%)	Max:	1.72 Gbps	Avg:	842.02 Mbps	Cur:	205.87 Mbps
GYN	(00.39%)	Max:	2.15 Gbps	Avg:	1.19 Gbps	Cur:	630.38 Mbps
PE	(00.25%)	Max:	1.21 Gbps	Avg:	757.91 Mbps	Cur:	409.56 Mbps
SJC	(00.14%)	Max:	969.47 Mbps	Avg:	520.50 Mbps	Cur:	223.72 Mbps
VIX	(00.16%)	Max:	1.18 Gbps	Avg:	632.39 Mbps	Cur:	255.81 Mbps
MGF	(00.09%)	Max:	761.63 Mbps	Avg:	425.24 Mbps	Cur:	143.01 Mbps
CPV	(00.07%)	Max:	632.59 Mbps	Avg:	315.62 Mbps	Cur:	107.08 Mbps
BEL	(00.14%)	Max:	941.51 Mbps	Avg:	566.64 Mbps	Cur:	219.32 Mbps
CXJ	(00.02%)	Max:	127.48 Mbps	Avg:	76.26 Mbps	Cur:	29.91 Mbps
SJP	(00.01%)	Max:	44.46 Mbps	Avg:	28.52 Mbps	Cur:	21.68 Mbps
MAO	(00.00%)	Max:	516.90 Mbps	Avg:	19.66 Mbps	Cur:	2.09 Mbps
CGB	(00.00%)	Max:	5.53 Mbps	Avg:	1.08 Mbps	Cur:	34.77 kbps
LAJ	(00.00%)	Max:	1.12 Mbps	Avg:	134.20 kbps	Cur:	58.09 kbps
SCA	(00.00%)	Max:	20.50 Mbps	Avg:	342.62 kbps	Cur:	18.09 kbps
TOTAL:		Max:	843.51 Gbps	Avg:	456.11 Gbps	Cur:	161.07 Gbps

PTT.br São Paulo represents approximately 81% of the all locations aggregated exchanged traffic.

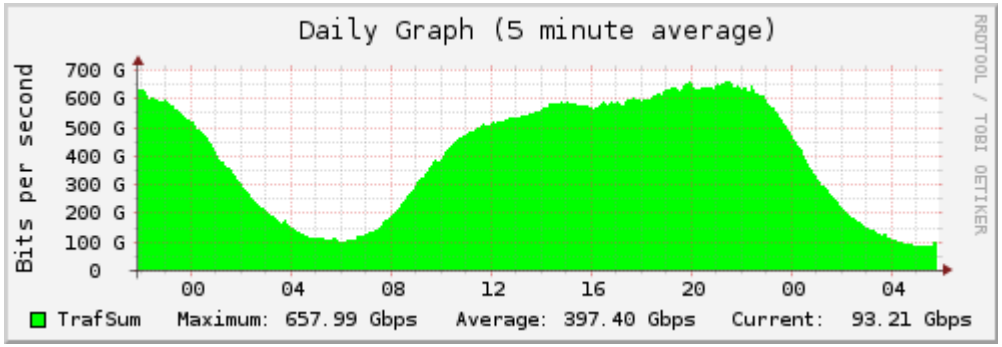
PTT.br – São Paulo is the biggest location



Daily peaks traffic of 657Gbps
28 PIXes
632 participants
4 Route Servers
572 active BGP sessions

Almost 27% of Brazilian IPv4 routed address space is present at PTT.br São Paulo

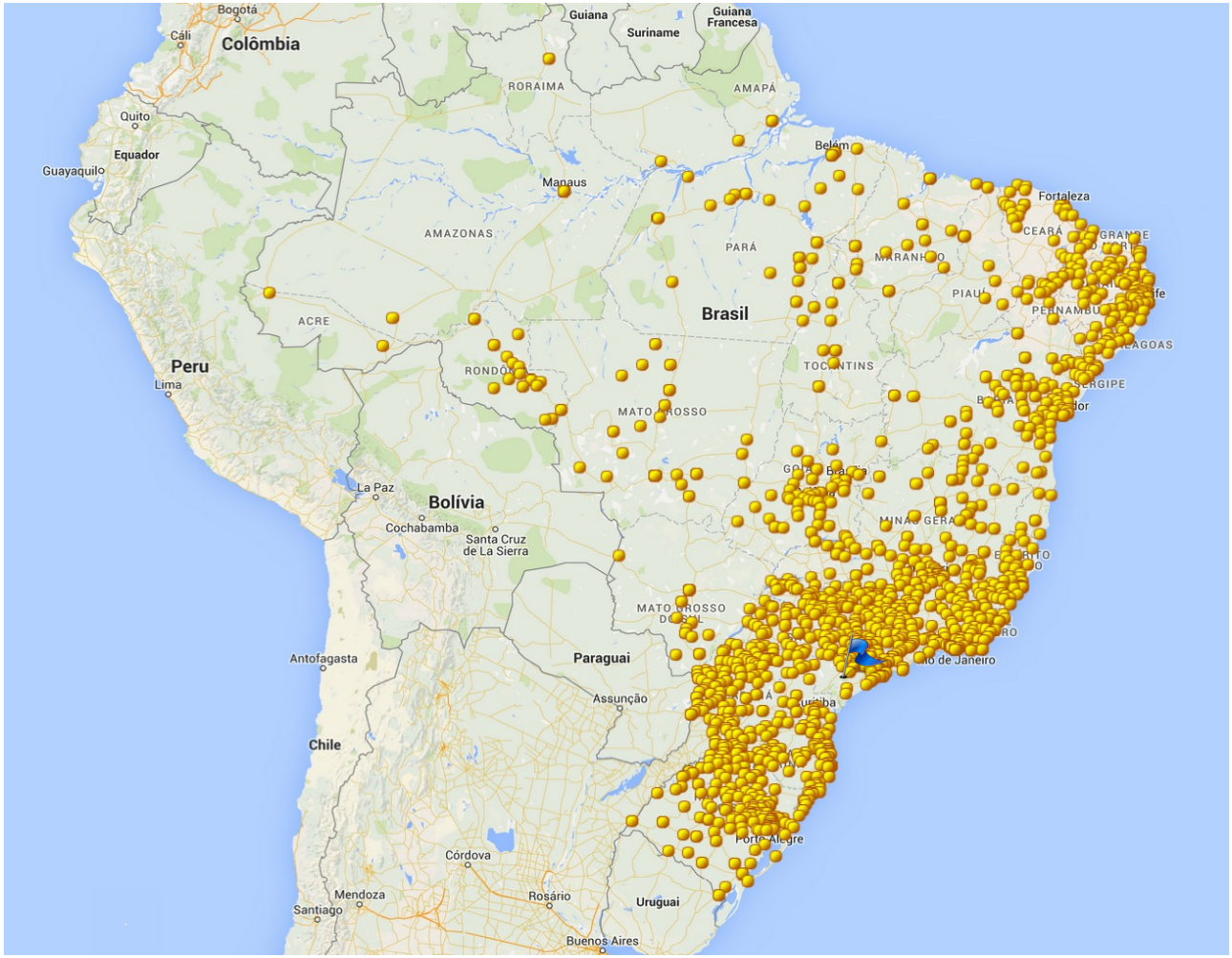
PTT.br – São Paulo is the biggest location



The PTT.br São Paulo attracts the connection interest of several autonomous systems.

PTT.br – Brazilian AS Distribution

Total allocated ASN: 3278



Some Autonomous Systems are

a few Km distance from PTT.br São Paulo

PTT.br – Peering interest

Normally an AS connects to a PTT local and PTT.br São Paulo that is currently considered a National PTT.

Autonomous Systems geographically distant is interested in connecting to PTT.br São Paulo mainly for 3 reasons:

- Available streaming content providers
- Amount Exchanged Traffic
- Possibility of buy and sell Internet Transit of Providers connected

PTT.br – Geographically distant locations

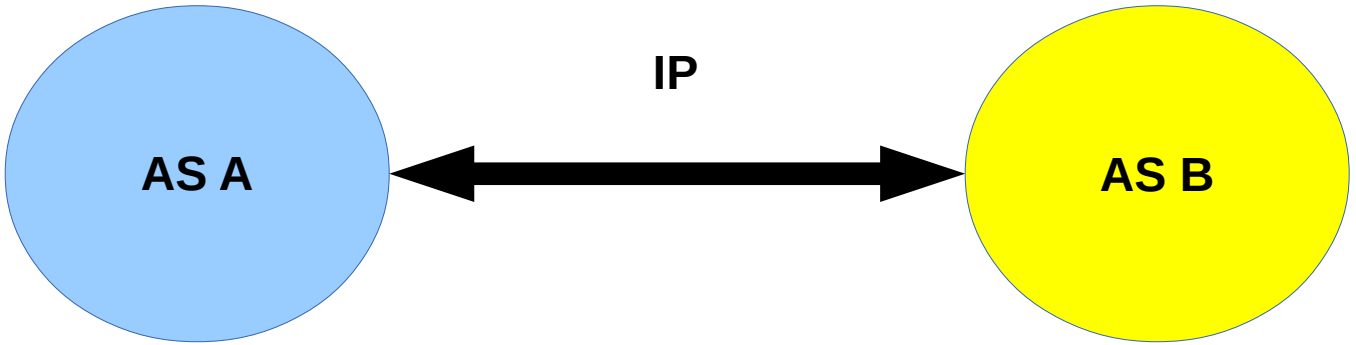
A transport is necessary (Lan-to-Lan) for an Autonomous System can participate in a remote location.

For example for an AS that is in Manaus can connect to PTT.br São Paulo it will have to pay a transport with a distance of about 3875 km.

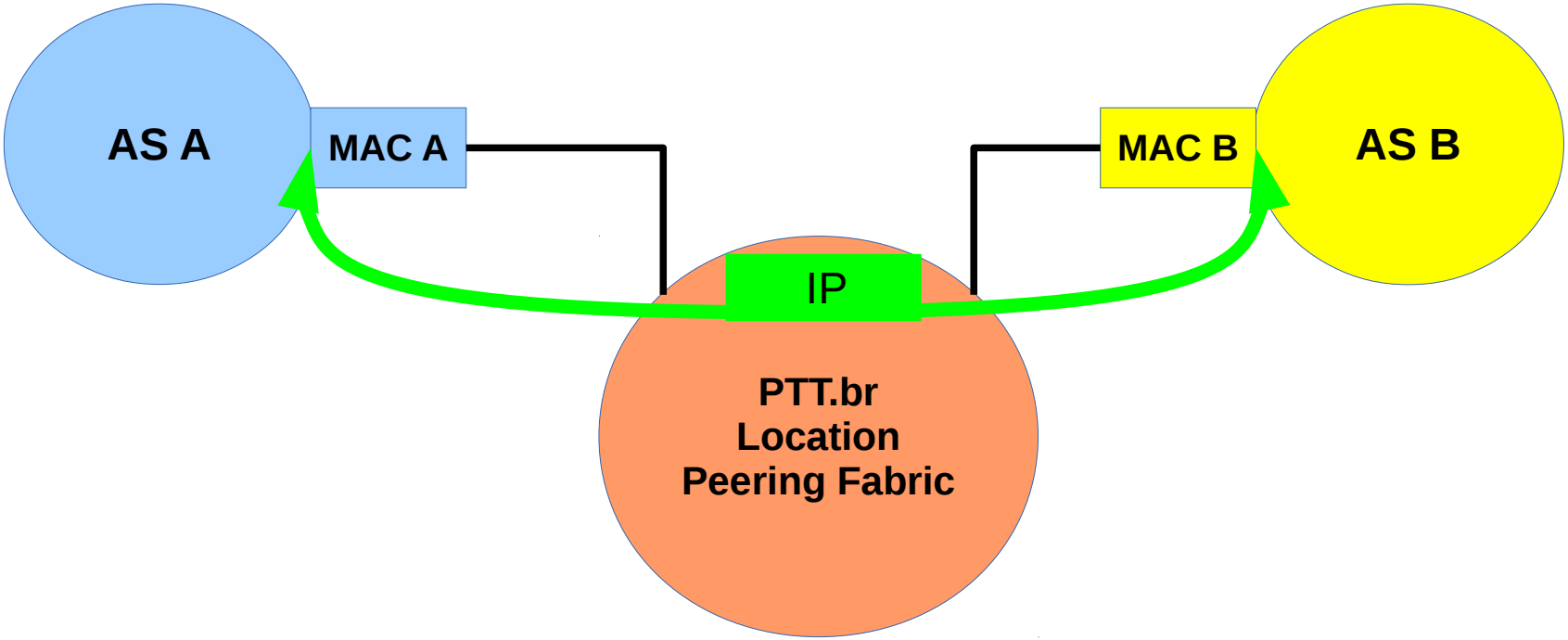
Transport between locations in Brazil usually have high costs, but the cost is reduced when shared by more than one company or Providers Association.

PTT.br – Reference Model

Fundamental Definition



PTT.br – Basic Operational Model – Allowed Traffic

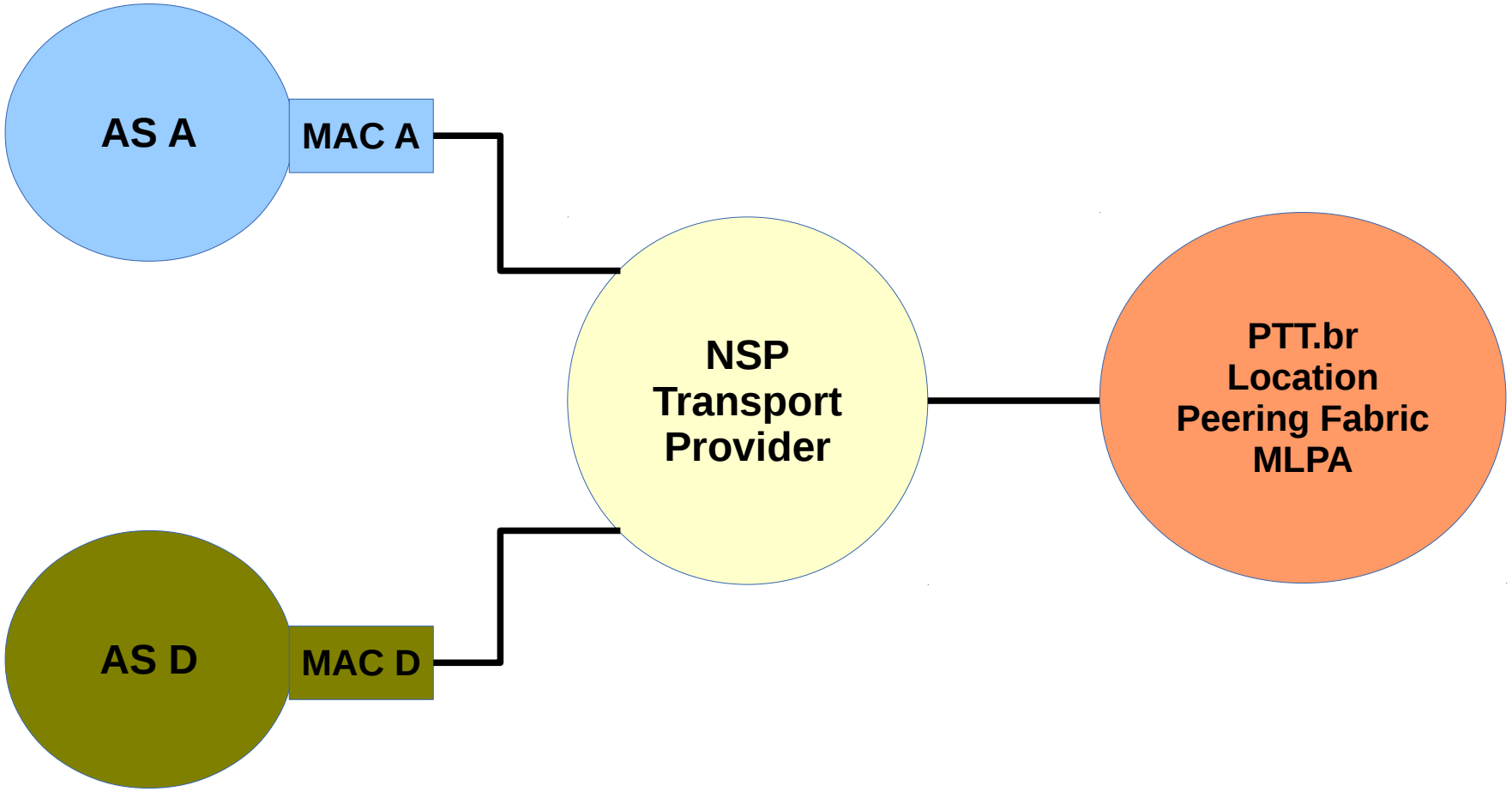


Allowed Frames with Ethertypes:

- 0x0800 - IPv4
- 0x0806 - ARP
- 0x86dd - IPv6

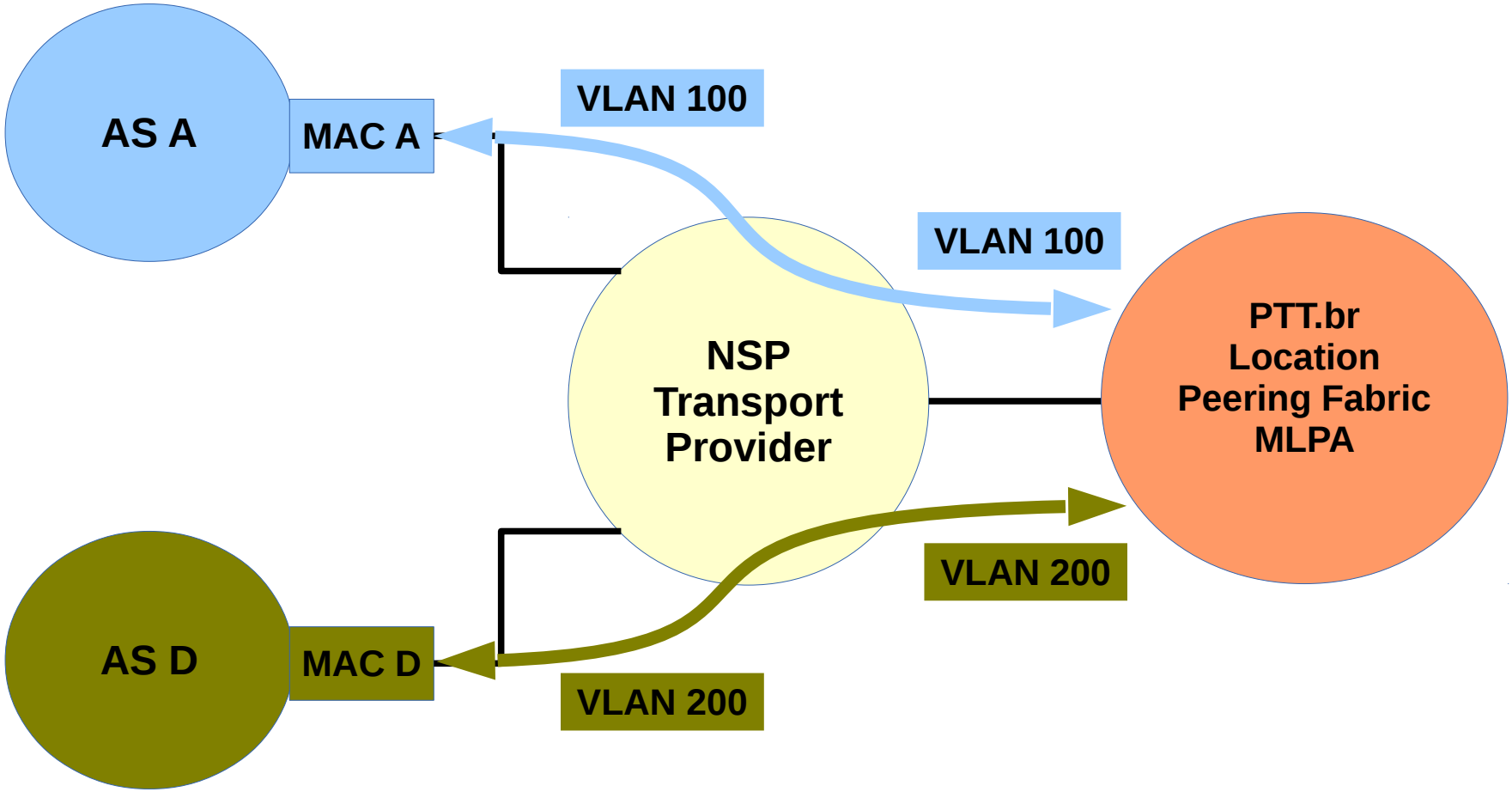
PTT.br – Extended Operational Model – Multiple AS per Single Port

Shared L2 structure outside PTT.br operation domain.



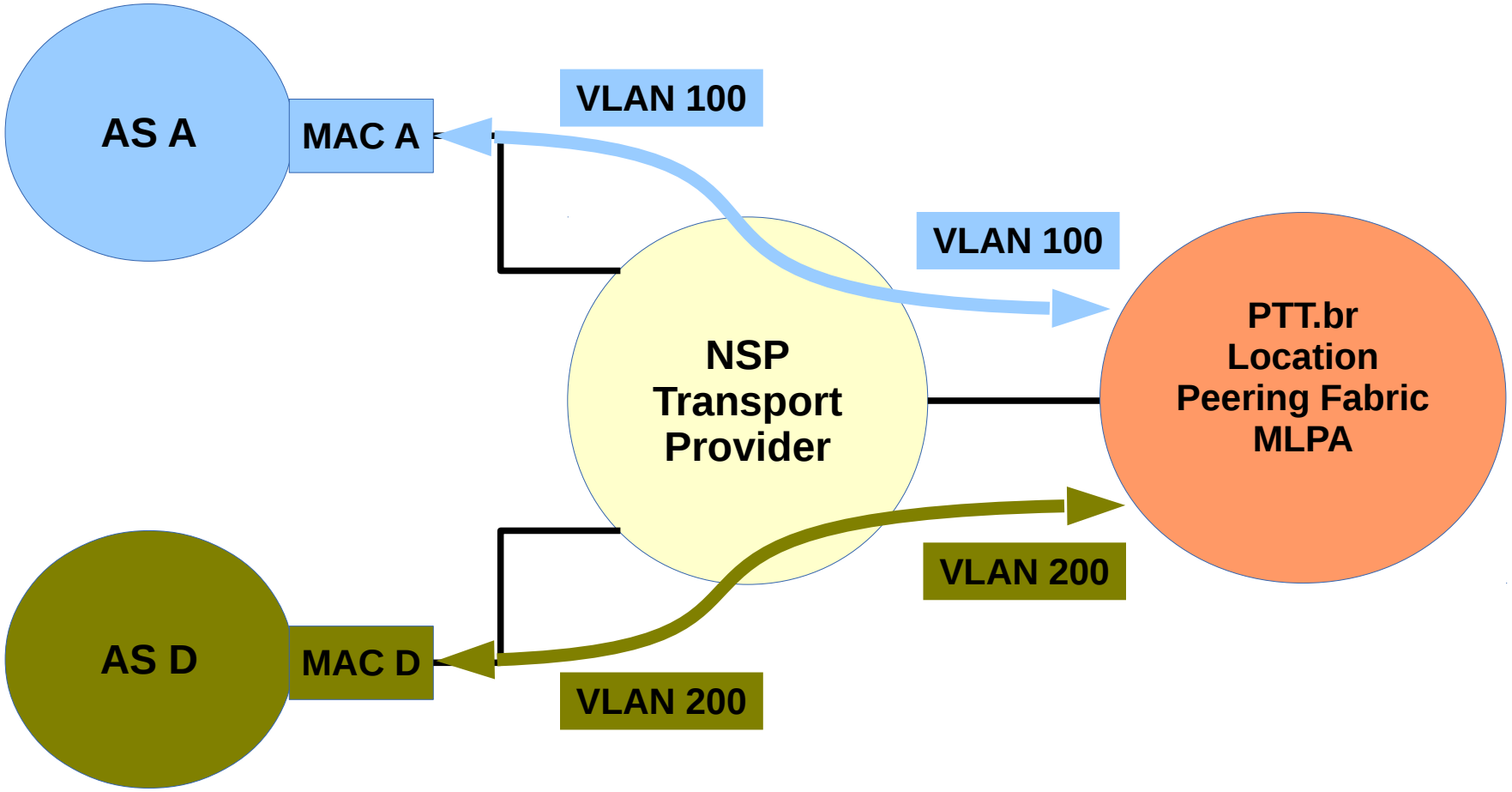
PTT.br – Extended Operational Model – Multiple AS per Single Port

For multiple participants connection in one Lan-to-Lan can occur is necessary to provide logical isolation between connections, in this case using VLANs with different tags (IEEE 802.1Q).



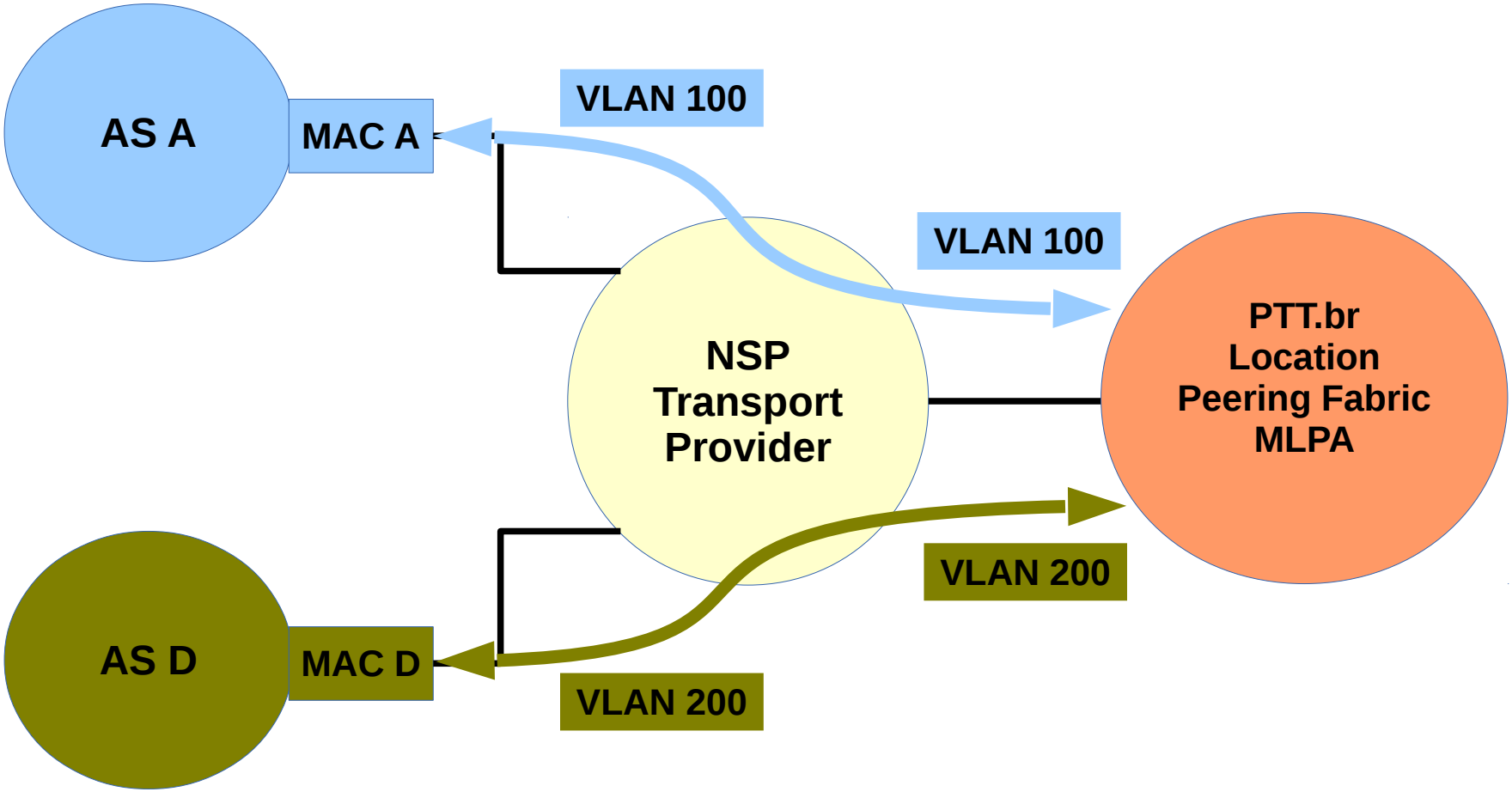
PTT.br – Extended Operational Model – Multiple AS per Single Port

PTT.br is a single domain broadcast with a single vlan tag and for communication occur, it is necessary to adopt VLANs translation mechanisms.



PTT.br – Extended Operational Model – Multiple AS per Single Port

Some telecom operators can mix traffic from multiple connections causing problems for the other participants, so currently all new participants PTT.br São Paulo are activated with different vlan tags each other.



PTT.br – The first request to share L2 structure

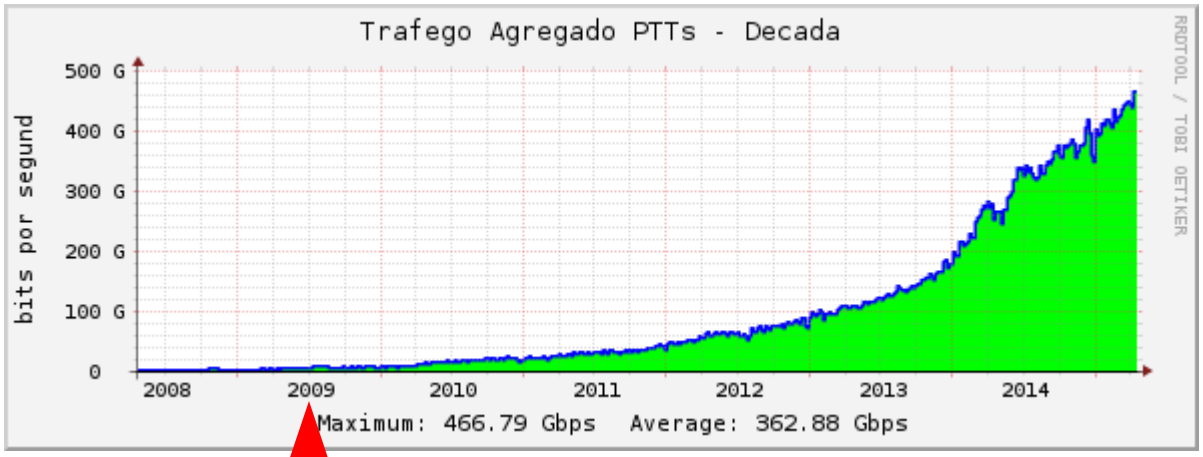
The first requests for share L2 structure to connect multiple remote participants occurred in 2009:

- Fasternet carrying Autonomous Systems of Campinas and region to São Paulo
- ANID carrying Autonomous Systems of Campina Grande and region to São Paulo
  Reference: PTT Fórum 3 - Panel - Transport between cities in the PTTMetro

The idea was to connect remote participants from other location to get content not available in their cities in PTT.br São Paulo.

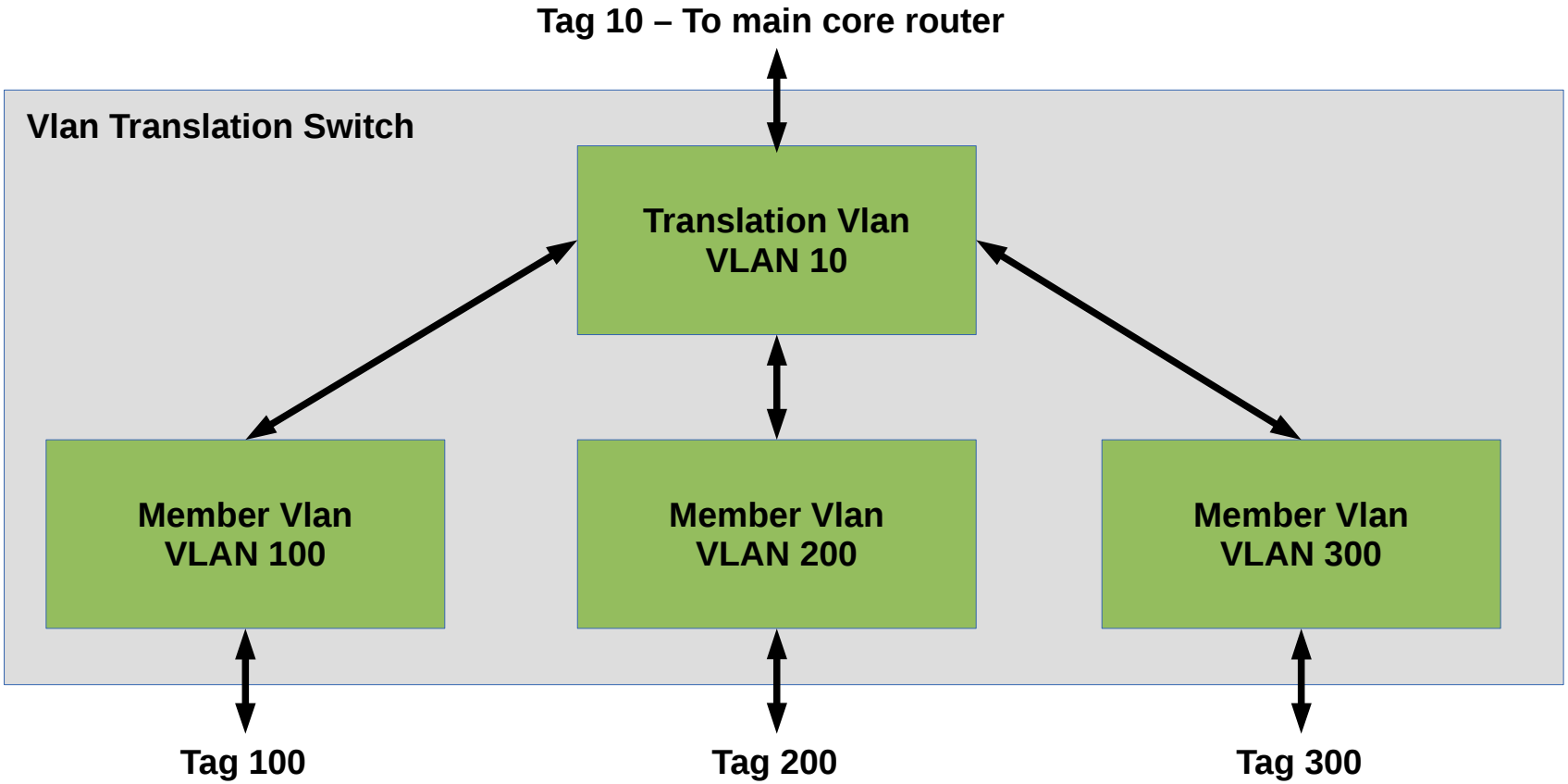
PTT.br – Traffic Growth

The possibility of connecting many participants sharing the same physical connection was one of the factors permited the PTT.br reach 80% of average growth in last years.

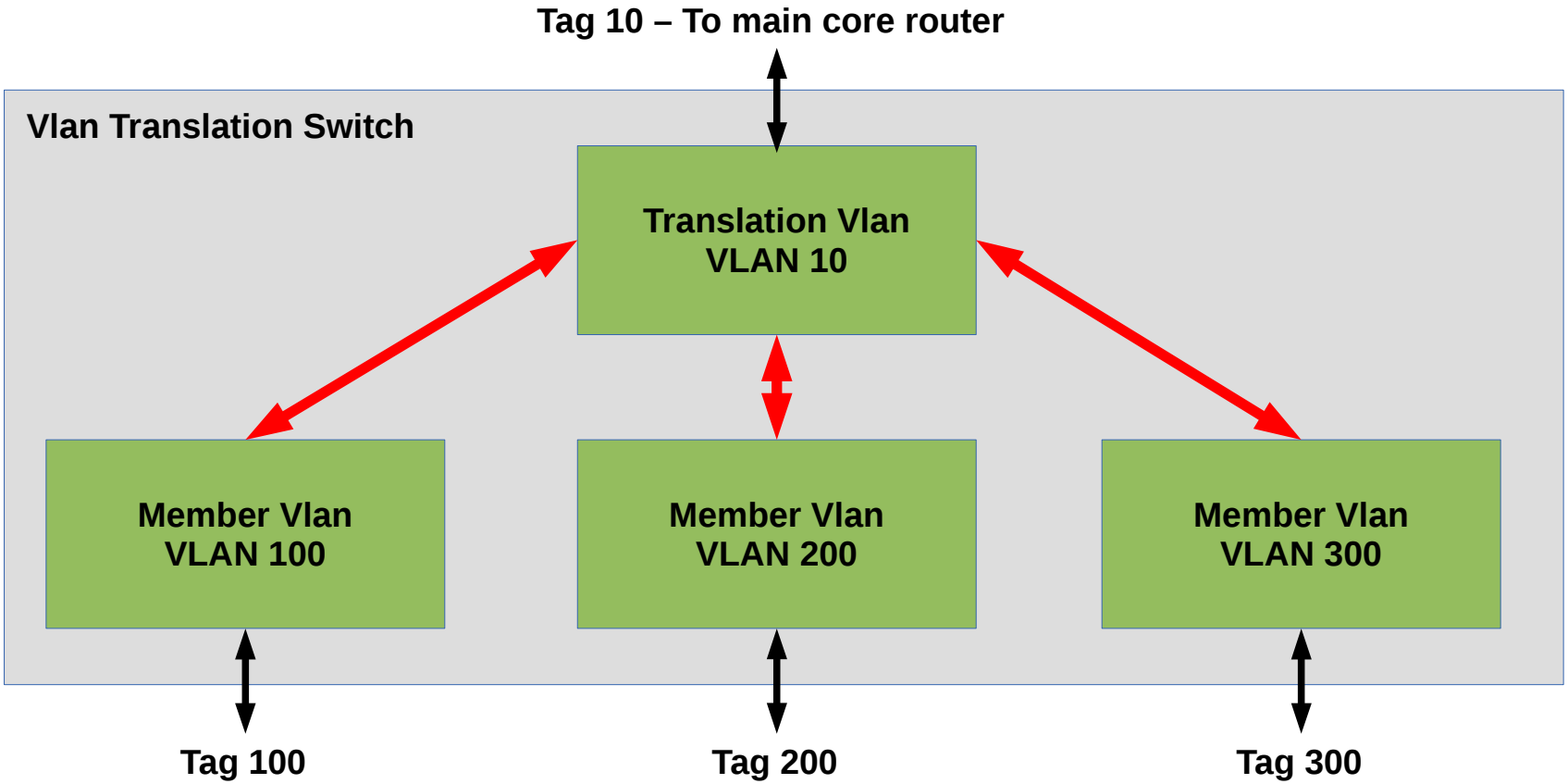


Start multiple participants connection in a same Lan-to-Lan

PTT.br – Vlan translation feature

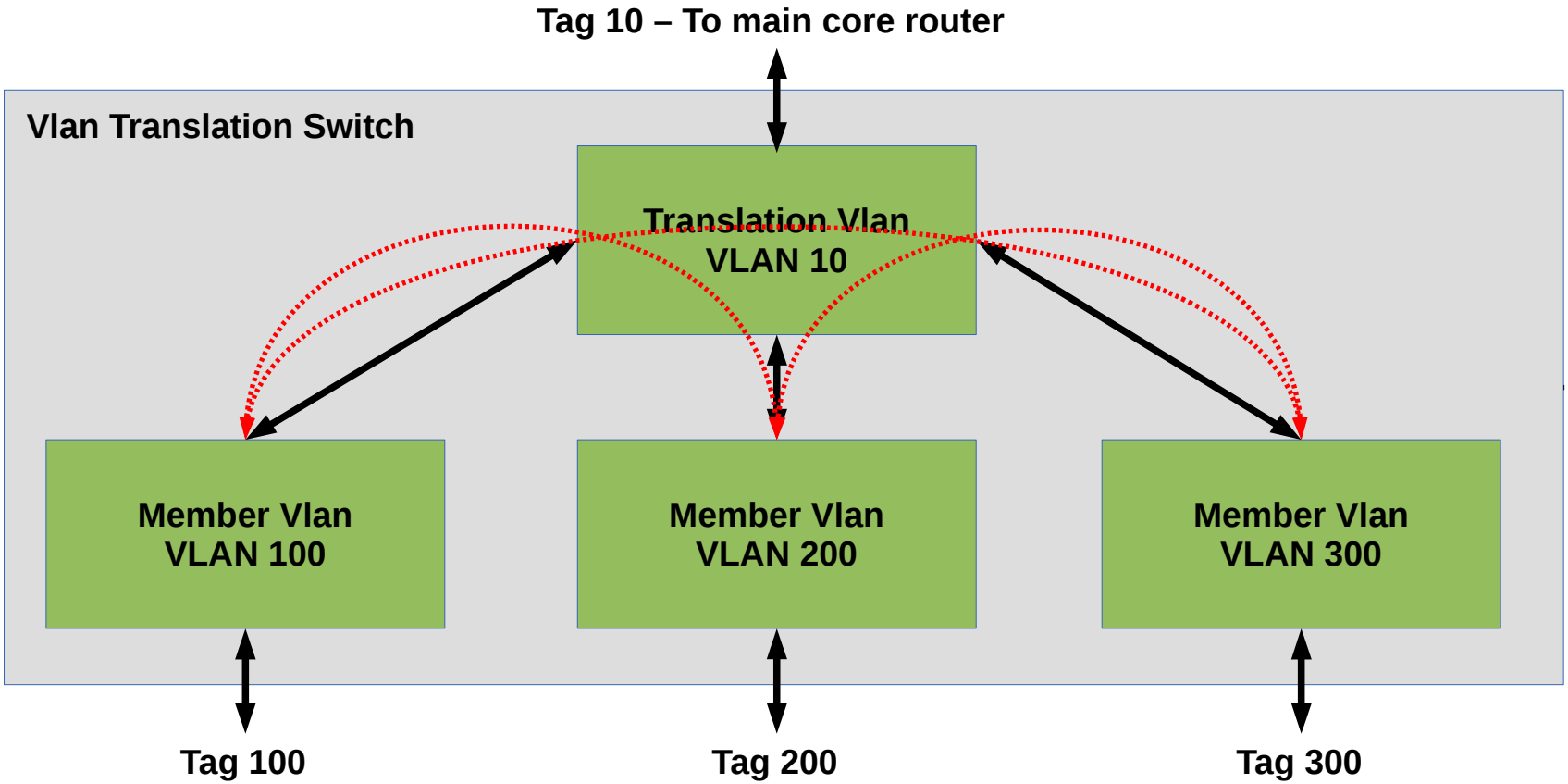


PTT.br – Vlan translation feature



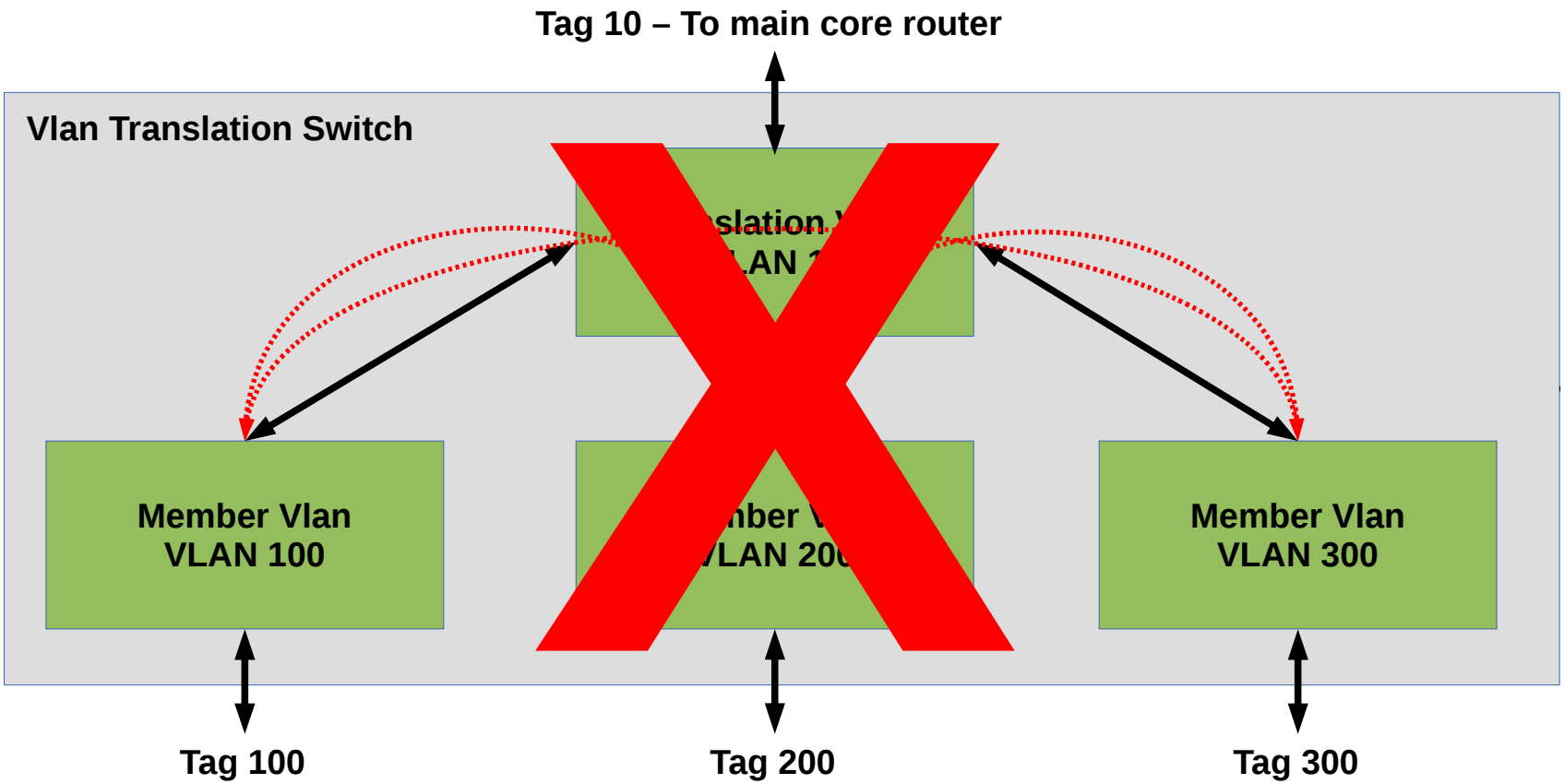
There is communication between Vlans 100,200 and 300 with vlan 10

PTT.br – Vlan translation feature



There is no communication between Vlans 100 with 200 and 300

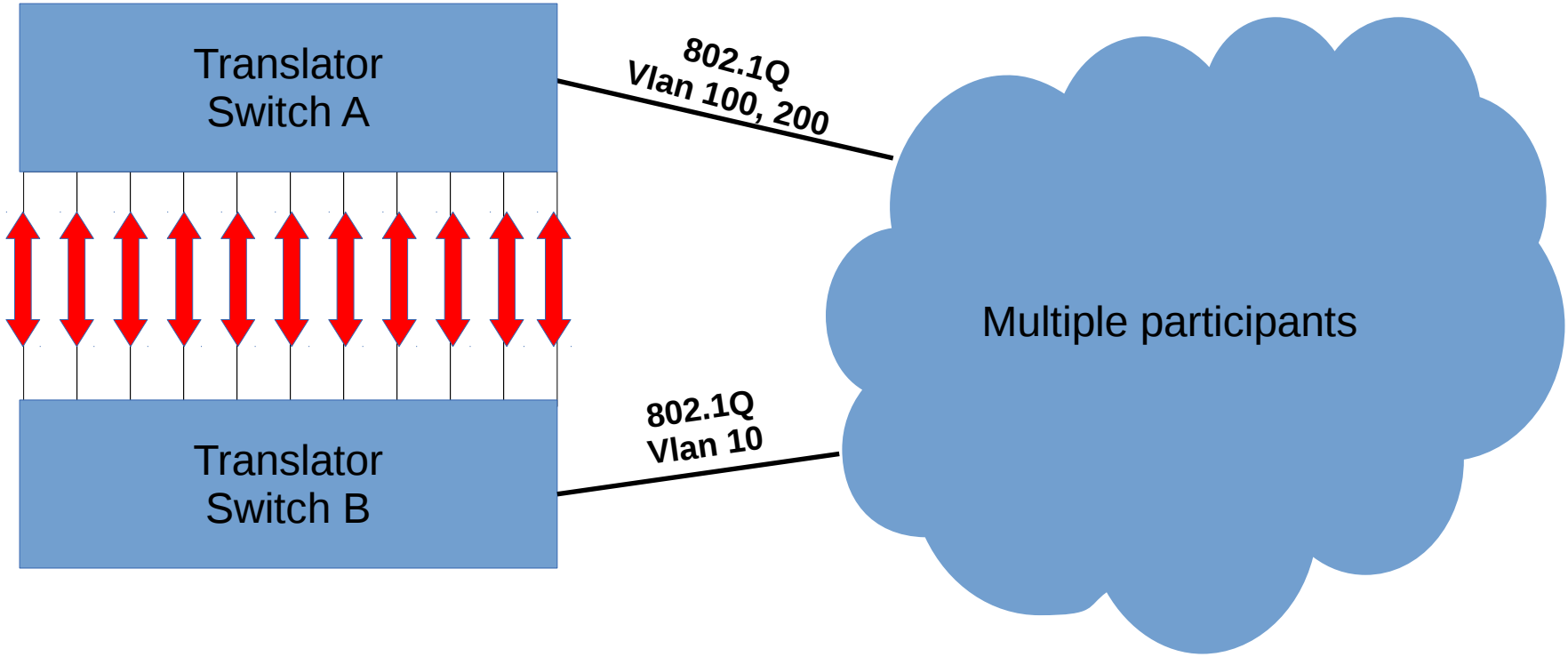
PTT.br – Vlan translation feature



Does not allow communication between all Vlans.
Therefore can not be used for PTTs

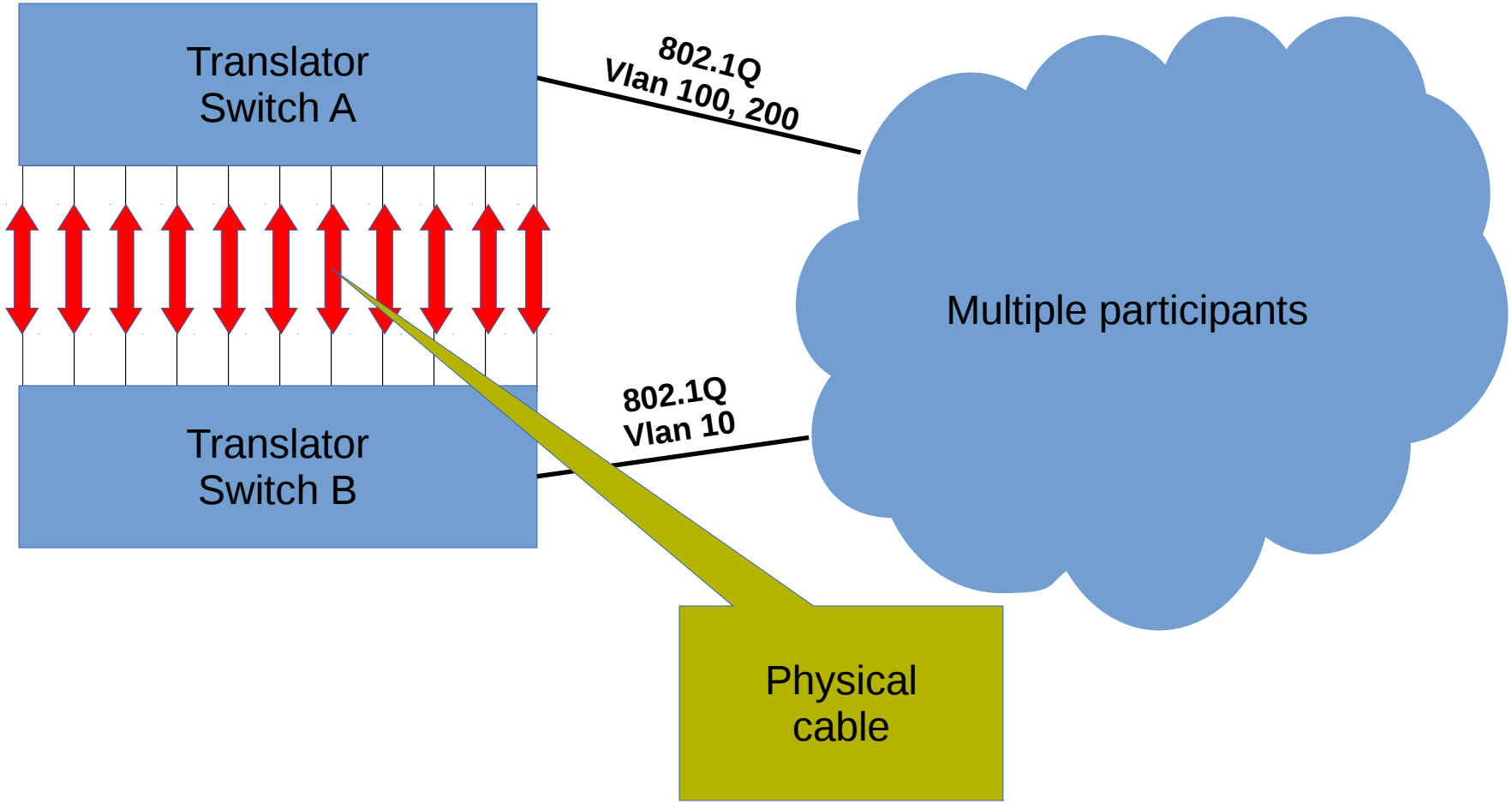
PTT.br – Solution A

PTT.br – Solution A

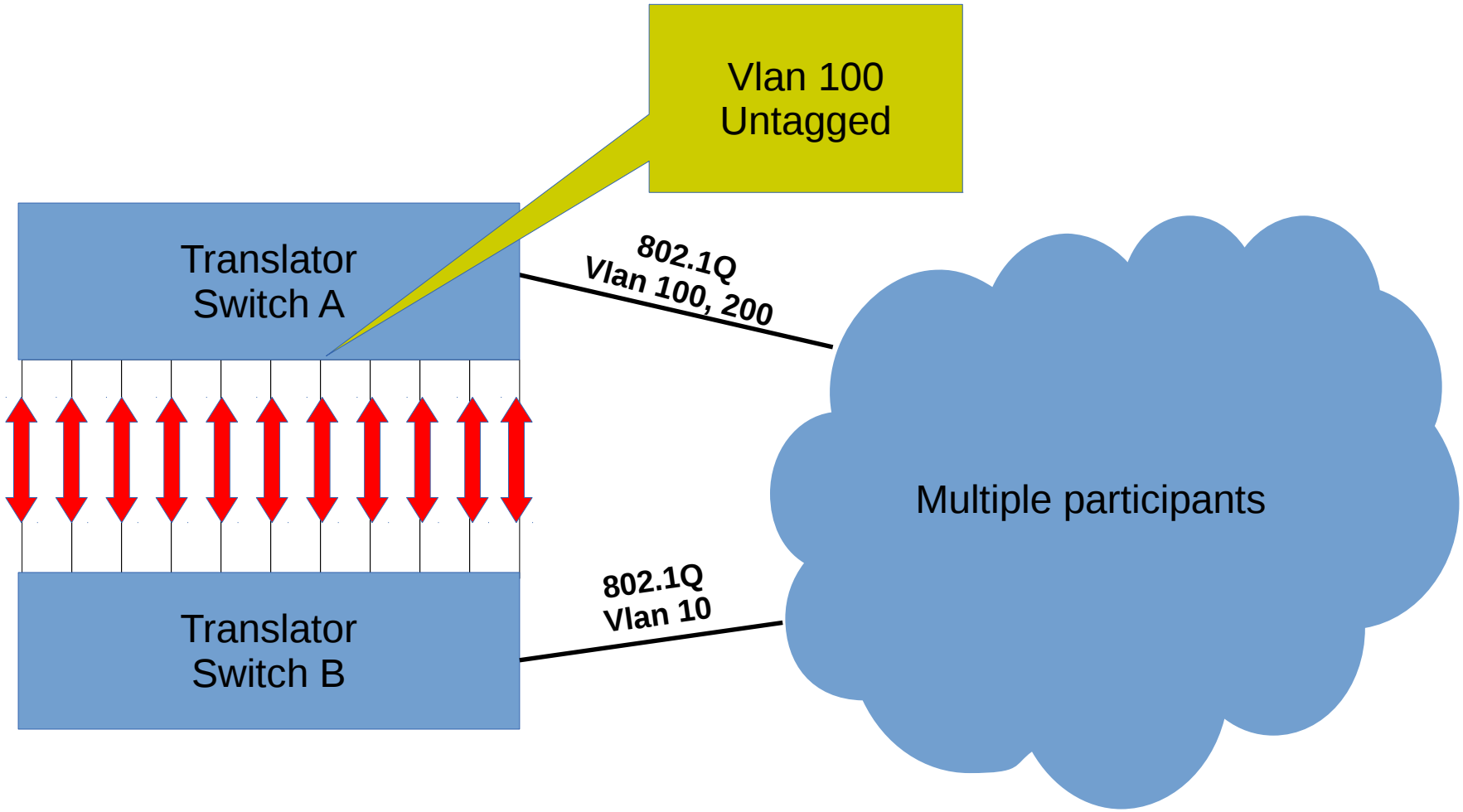


Workaround developed by the PTT.br team

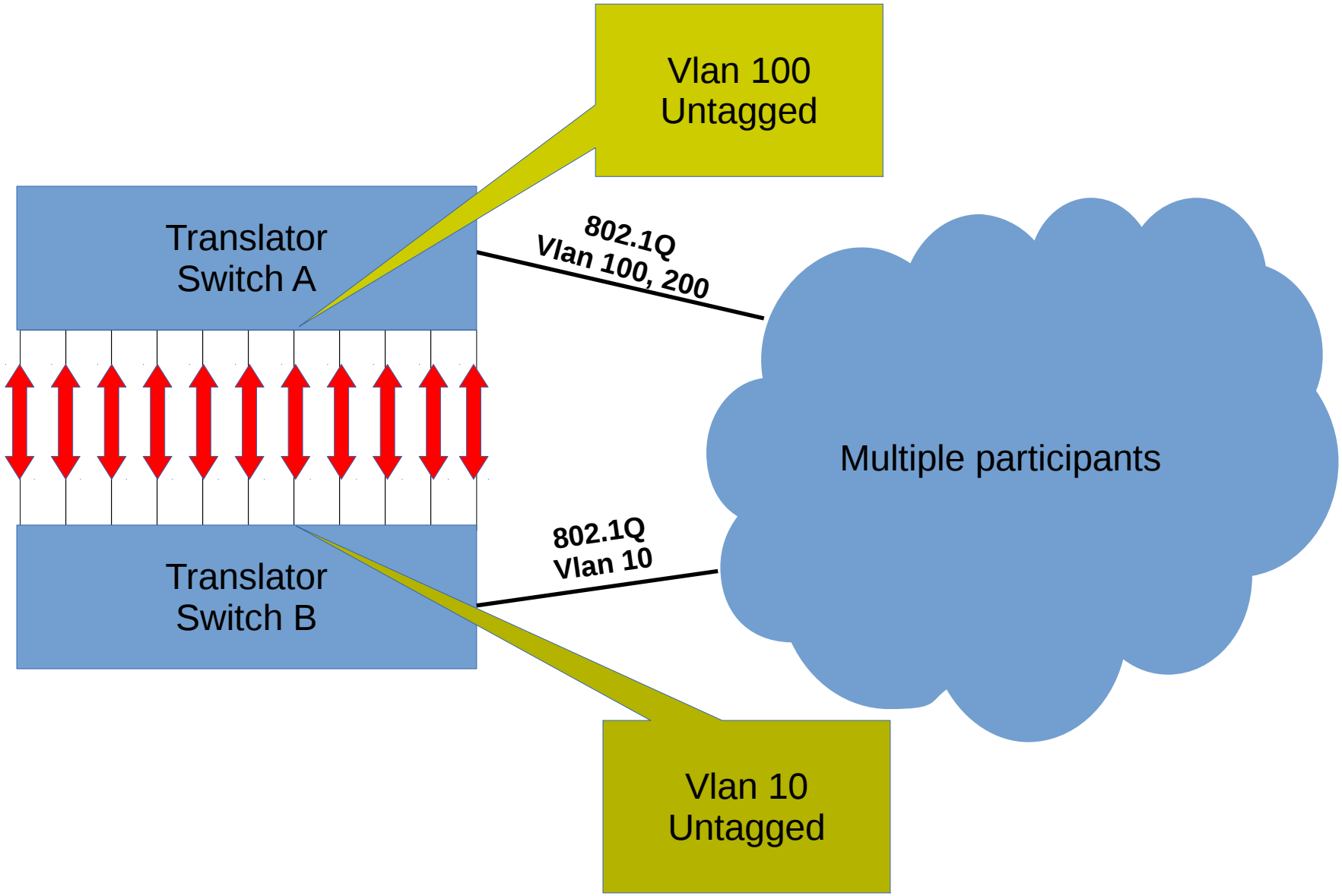
PTT.br – Solution A



PTT.br – Solution A



PTT.br – Solution A

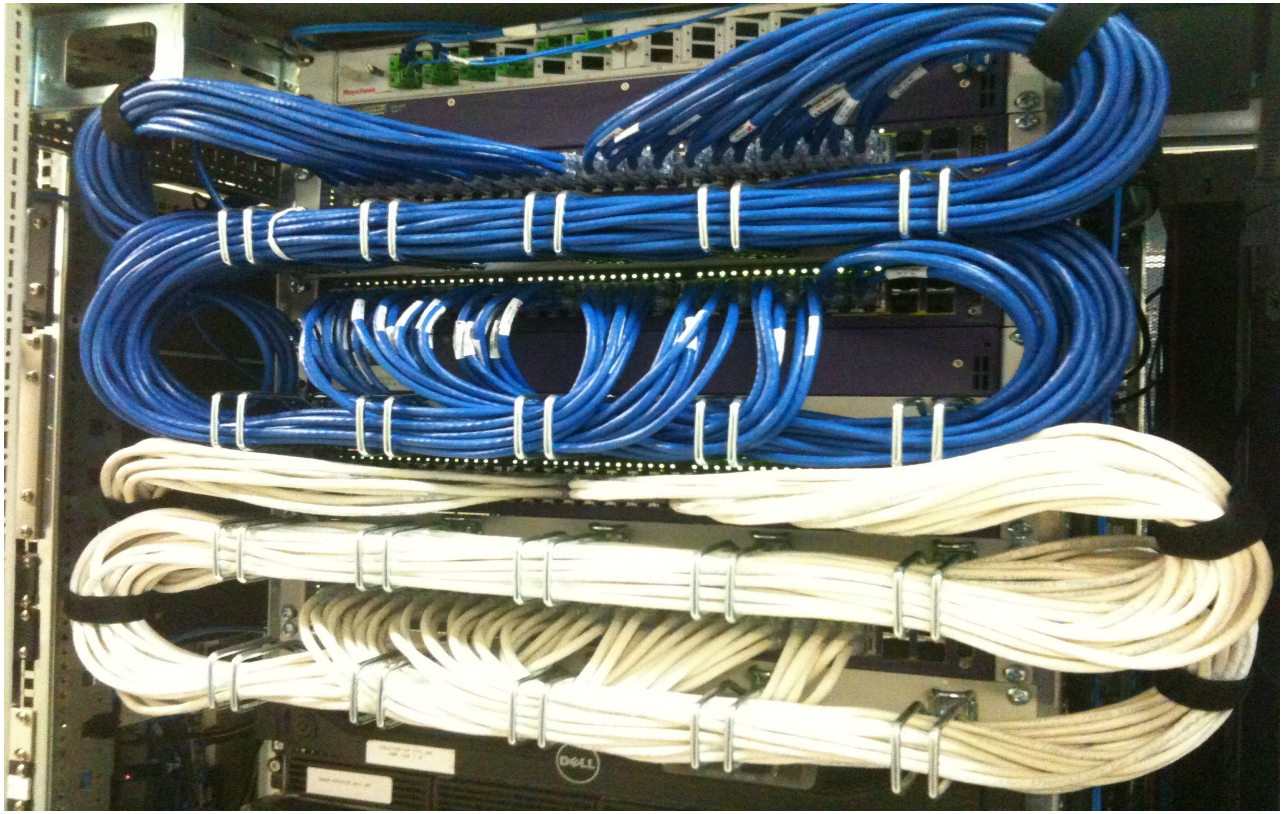


PTT.br – Solution A - Configuration

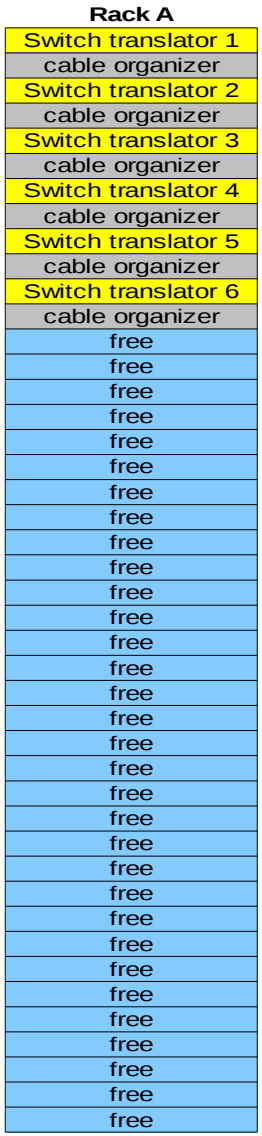
```
Translator-Switch-A#create vlan V100
Translator-Switch-A#configure V100 tag 100
Translator-Switch-A#configure V100 add ports x tagged
Translator-Switch-A#configure V100 add ports y untagged

Translator-Switch-B#create vlan V10
Translator-Switch-B#configure V10 tag 10
Translator-Switch-B#configure V10 add ports x tagged
Translator-Switch-B#configure V10 add ports y untagged
```

PTT.br – Solution A

Two translation systems (Blue and White cables - each group connecting a par of switches).



PTT.br – Solution A



At the time of deactivation there were 3 pairs of translators switches at PTT.br São Paulo, 144 1G interfaces used to attend around 50 participants.

Each switch consumes 182W.
6 switches consumed 1092W.

The solution A was discontinued in Sao Paulo because it didn't scaled.

PTT.br – Solution A

Rack A
Switch translator 1
cable organizer
Switch translator 2
cable organizer
Switch translator 3
cable organizer
Switch translator 4
cable organizer
Switch translator 5
cable organizer
Switch translator 6
cable organizer
Switch translator 7
cable organizer
Switch translator 8
cable organizer
Switch translator 9
cable organizer
Switch translator 10
cable organizer
Switch translator 11
cable organizer
Switch translator 12
cable organizer
Switch translator 13
cable organizer
Switch translator 14
cable organizer
Switch translator 15
cable organizer
Switch translator 16
cable organizer
Switch translator 17
cable organizer
Switch translator 18
cable organizer
Switch translator 19
cable organizer
Switch translator 20
cable organizer
Switch translator 21
cable organizer
Switch translator 22
cable organizer

Rack B
Switch translator 23
cable organizer
Switch translator 24
cable organizer
Switch translator 25
cable organizer
Switch translator 26
cable organizer
Switch translator 27
cable organizer
Switch translator 28
cable organizer
Switch translator 29
cable organizer
Switch translator 30
cable organizer
Switch translator 31
cable organizer
Switch translator 32
cable organizer
Switch translator 33
cable organizer
Switch translator 34
cable organizer
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free
free

Extrapolation for São Paulo nowadays

If PTT.br were using today the solution A 17 translation systems would be needed

Resources Needed:

Equipments: 34 switches
Cabling: 814 UTP CAT6
Space: 68 U (Rack Units) in Almost 2 racks
Power Draw: > 6 KW

PTT.br – Solution A

PTT.br – Solution A

Disadvantages:

- Number of interfaces (2 per translation);
- Equipments that support high number of MAC address (more expensive):

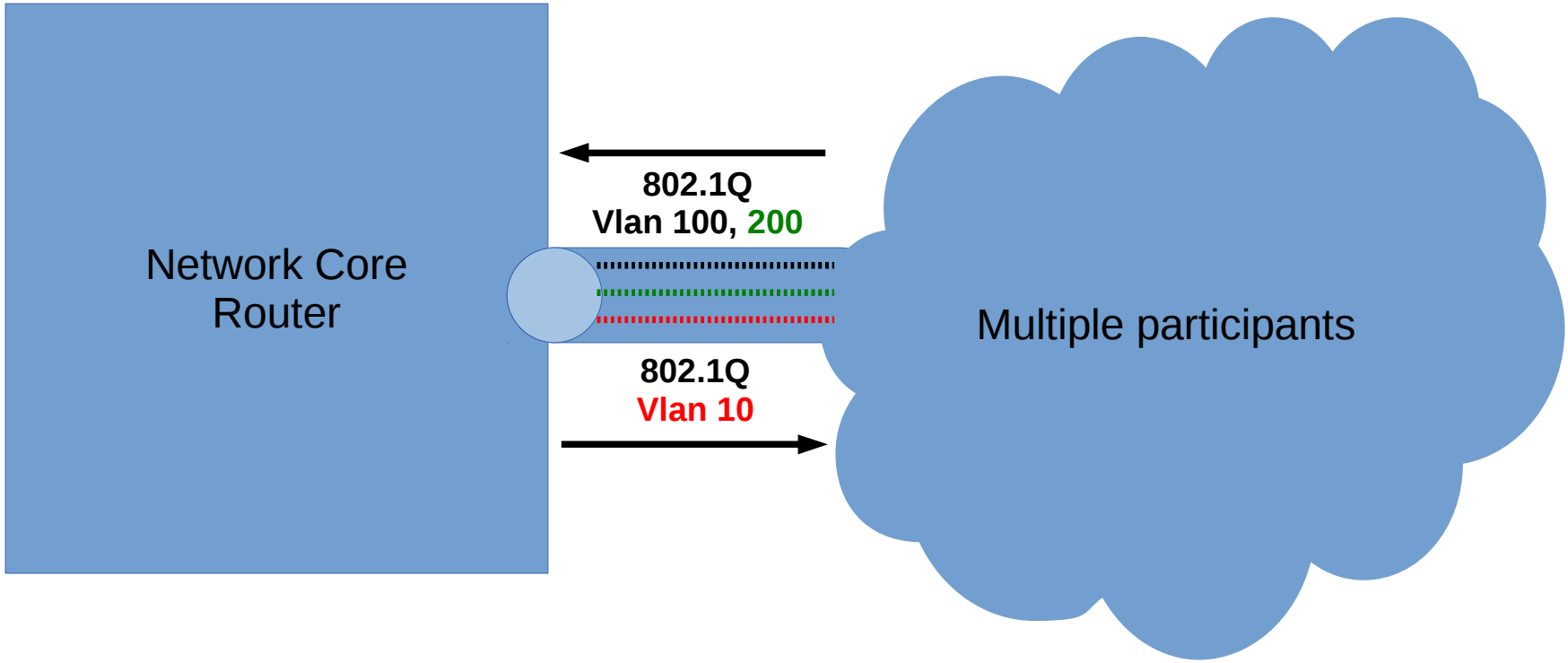
  Translation switch: 48*2000 = 96 K
  Core switch: 2000*2000 = 4 M
  * considering 2000 participants and 2000 routers on each MLPA bridge domain.

- High cost of operating / infrastructure;
- Low bandwidth;
- Physical space.

PTT.br – Solution B

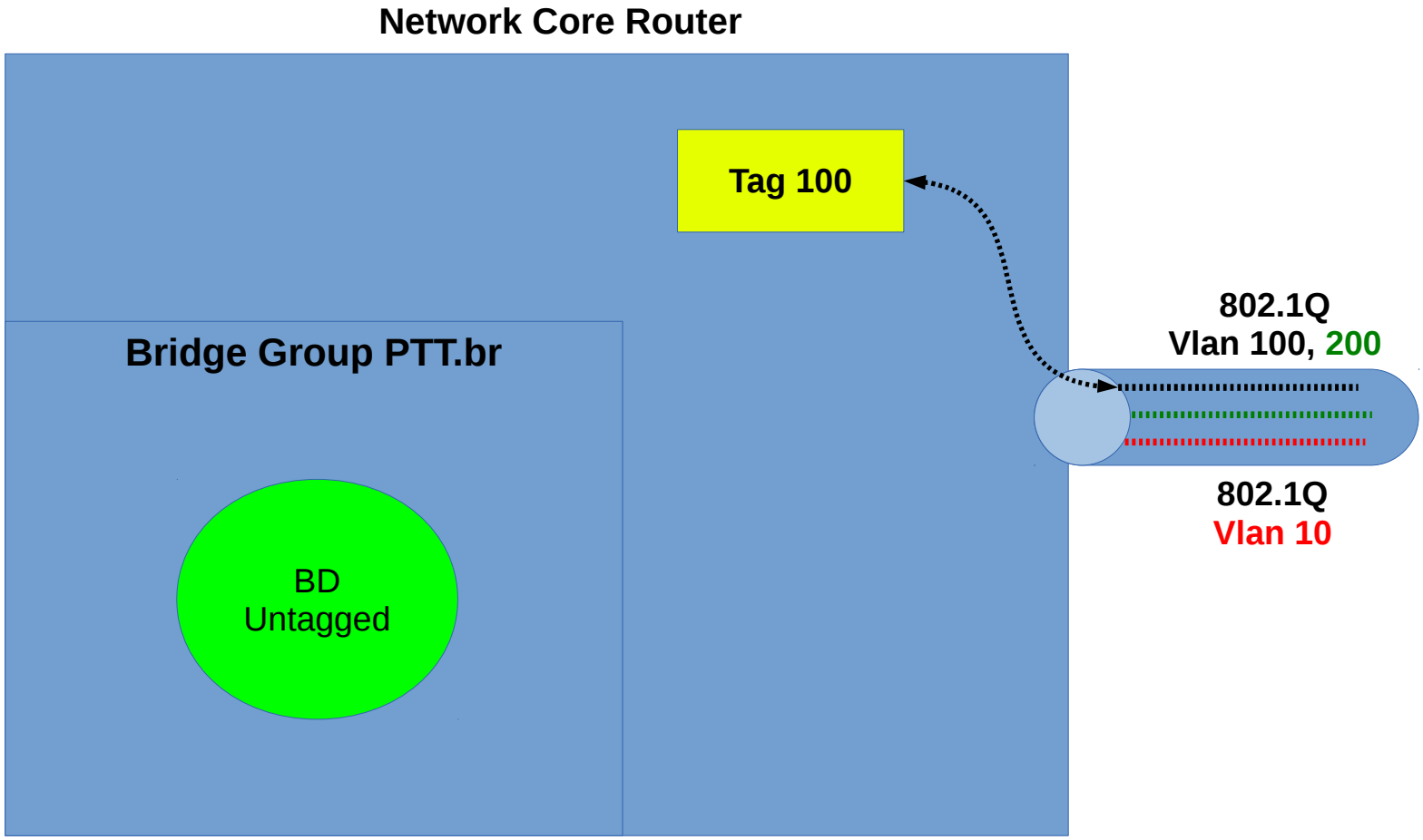
Logical Router Untagged Bridge Domain (L2VPN)

PTT.br – Solution B



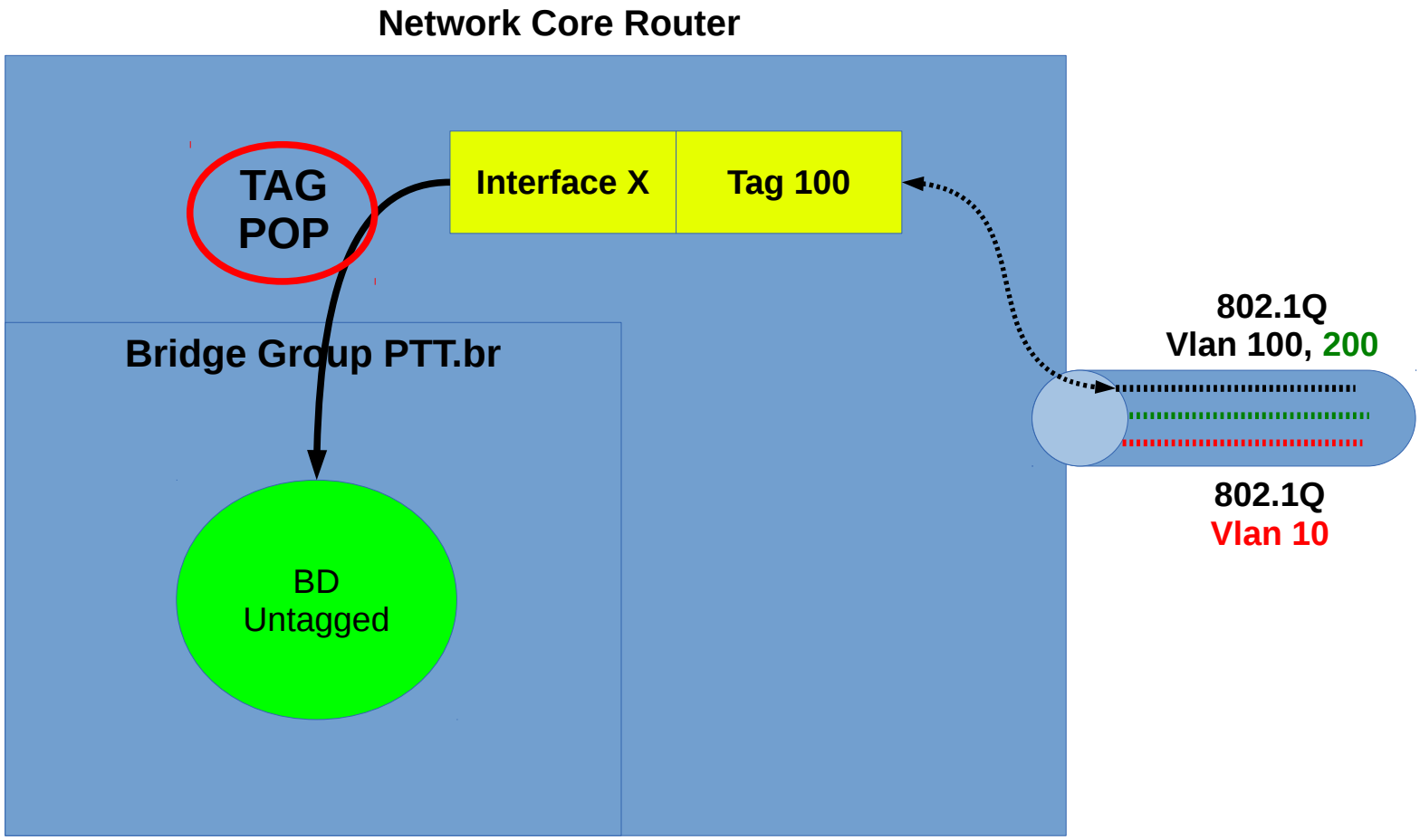
Implementation of router based peering fabric with untagged bridge domains

PTT.br – Solution B



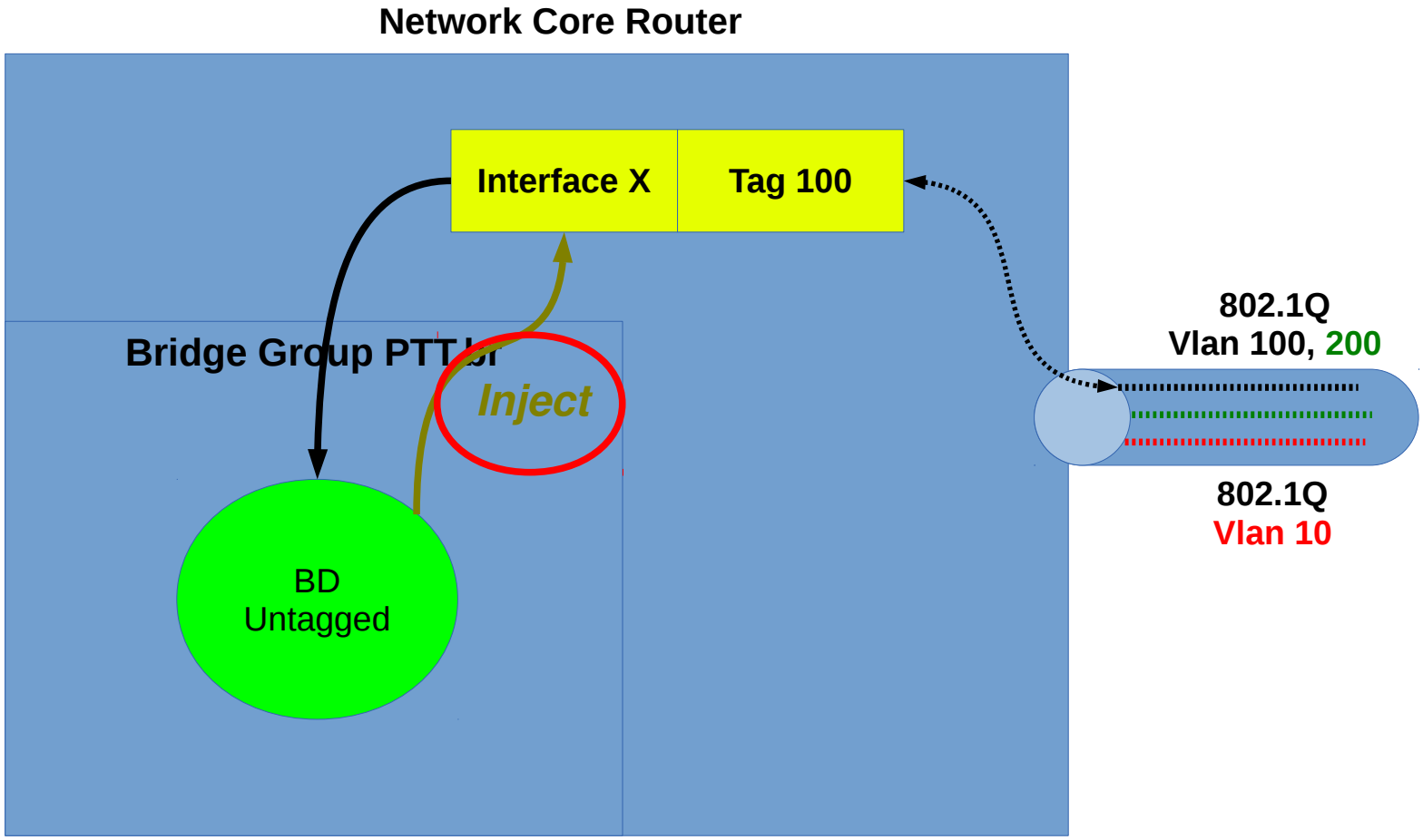
BD: Bridge Domain *(L2 switch)*

PTT.br – Solution B



Removing the Vlan tag

PTT.br – Solution B

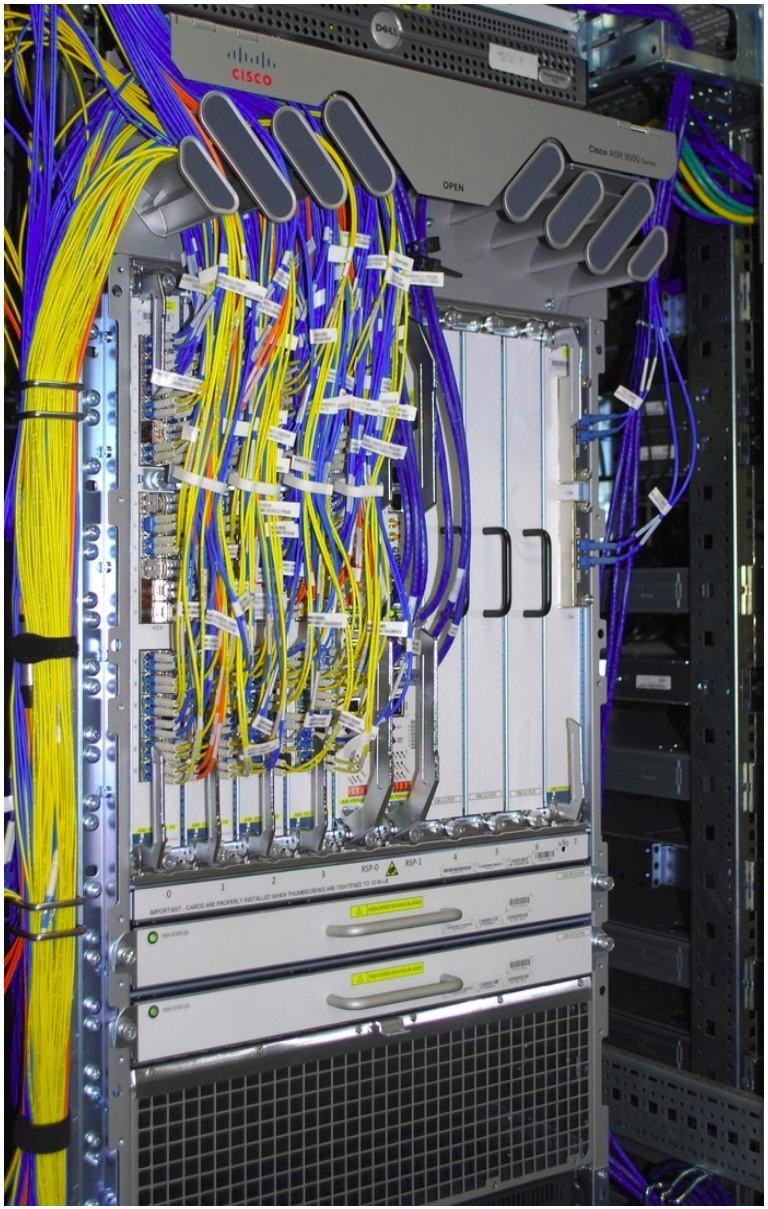


Injecting Vlan tag

PTT.br – Solution B - Configuration

```
Network-Core-Router#

interface X.100 l2transport
 encapsulation dot1q 100
 rewrite ingress tag pop 1 symmetric
!
l2vpn
 !
 bridge group PTT.br
  bridge-domain Untagged
   interface X.100
   !
  !
!
```

PTT.br – Solution B



PTT.br – Solution B

This solution is operating in PTT.br São Paulo since 2013 with use L2VPN in Cisco router model ASR9010.

The physical structure has 814 VLANs translated this model in PTT.br São Paulo.

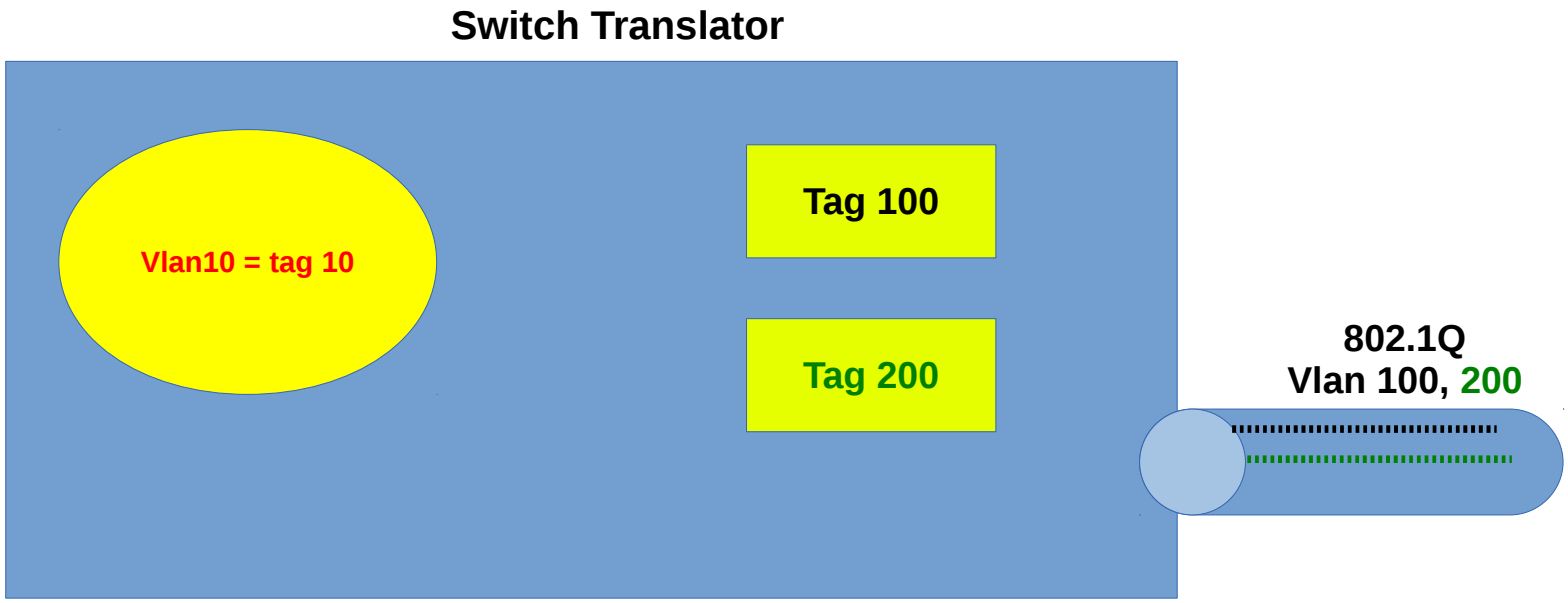
The model with Bridge Domain allows growth because the vlan tag has value for interface and not by equipment.

This solution can be used in another PTT.br locations if there is demand.

PTT.br – Solution C

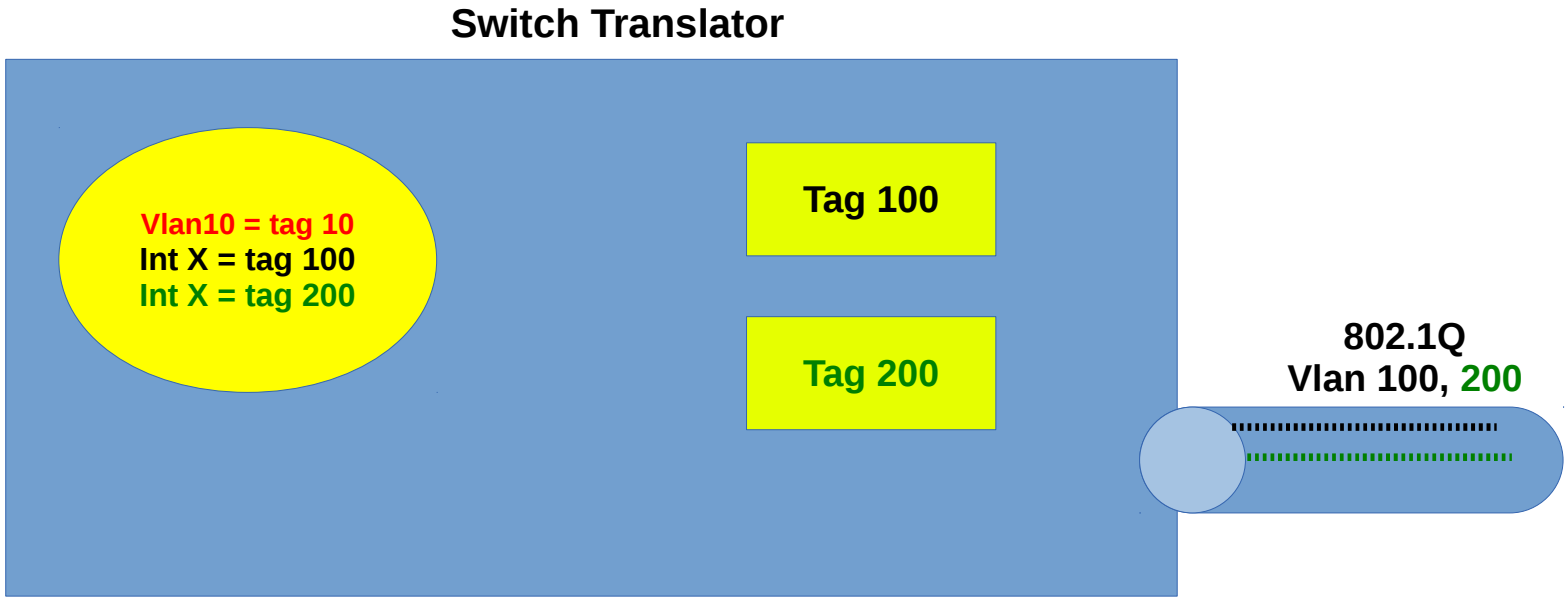
New switch Feature

PTT.br – Solution C



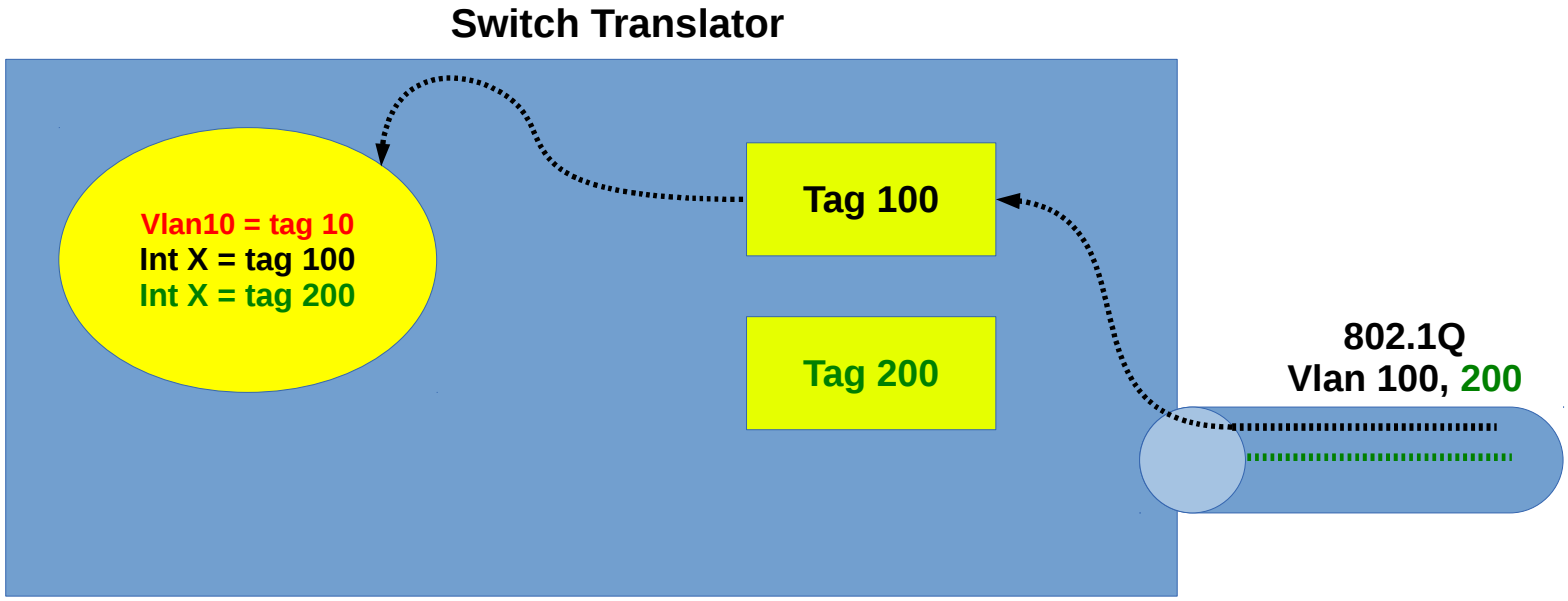
The solution was developed by Extreme Networks to attend our need. This feature is known as Port-specific VLAN tag and it is being used for other IXPs around the world.

PTT.br – Solution C



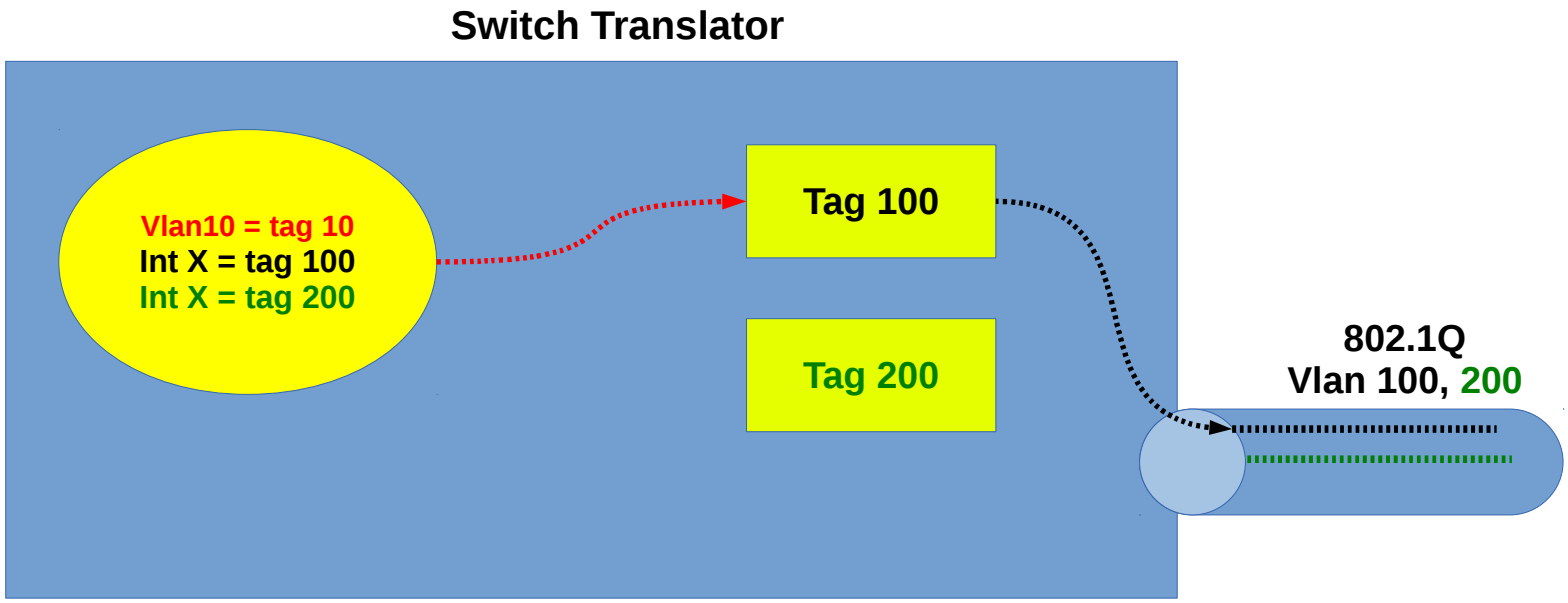
One Vlan can have 4094 tags per physical interface

PTT.br – Solution C



If the switch receives a frame with tag 100 by the interface X the tag will be changed to 10

PTT.br – Solution C



If the switch receives a frame with tag 10 and a destination MAC learned by interface X tag 100, it will change the tag to 100.

PTT.br – Solution C - Configuration

```
Switch-Translator#create vlan V10
Switch-Translator#configure V10 tag 10
Switch-Translator#configure V10 add ports x tagged 100
Switch-Translator#configure V10 add ports x tagged 200
```

PTT.br – Solution C

There are some traffic accounting limitations in discussion with the switch vendor.

This solution is very important for other locations that no have router to change tags.

The Port-specific functionality is only supported by newer equipments:

- Summit (460 / 480 / 670)
- Black Diamond

PTT.br – Conclusions

Solution	Cost	Size IXP	Scalability
A	low	small	low
B	high	big	high
C	medium	medium	medium

PTT.br – Conclusions

Solution	Cost	Size IXP	Scalability
A	low	small	low
B	high	big	high
C	medium	medium	medium

What is the best solution?

A: The PTT.br uses three solutions depending on the size of the location.

PTT.br – Conclusions

Transport between locations in Brazil usually have high costs, but the cost is reduced when shared by more than one company or Providers Association.

The membership growth PTT.br São Paulo is directly related to shared transport.

PTTs with less traffic and number of participants, simple solutions can be used for connection sharing between multiple participants in a same connection.

When the structure is greater the solution adopted needs to be scalable and provide security structure and other participants connected.

GTER 39 | GTS 25

http://gtergts.nic.br/

The GTER 39 | GTS 25 event will take place in Rio de Janeiro, Brazil on the 28th and 30th of May 2015.

The PTT.br team will do two presentations in GTER 39 on the 29th of may 2015:

- Infrastructure challenges and solutions for growth.
- CIX - An Intermediate Category.

The event will be transmitted over internet.

PTT Forum 9

http://ptt.br/pttforum/

The PTT Fórum 9 event will take place in São Paulo, Brazil on the 7th and 8th of December 2015.

PTT Fórum normal audience is composed by Brazilian Autonomous Systems (AS) and PTTMetro/PTT.br IX participants (some of them are international companies).

The event will be translated between Portuguese and English.

The event will be part of the **Brazilian Internet Infrastructure week** together with an IPv6 Forum and GTER/GTS (Network Engineering, Operation and Security Working Groups).

Thank you

ptt.br

eng@ptt.br

May 22th, 2015

nic.br cgi.br

www.nic.br | www.cgi.br