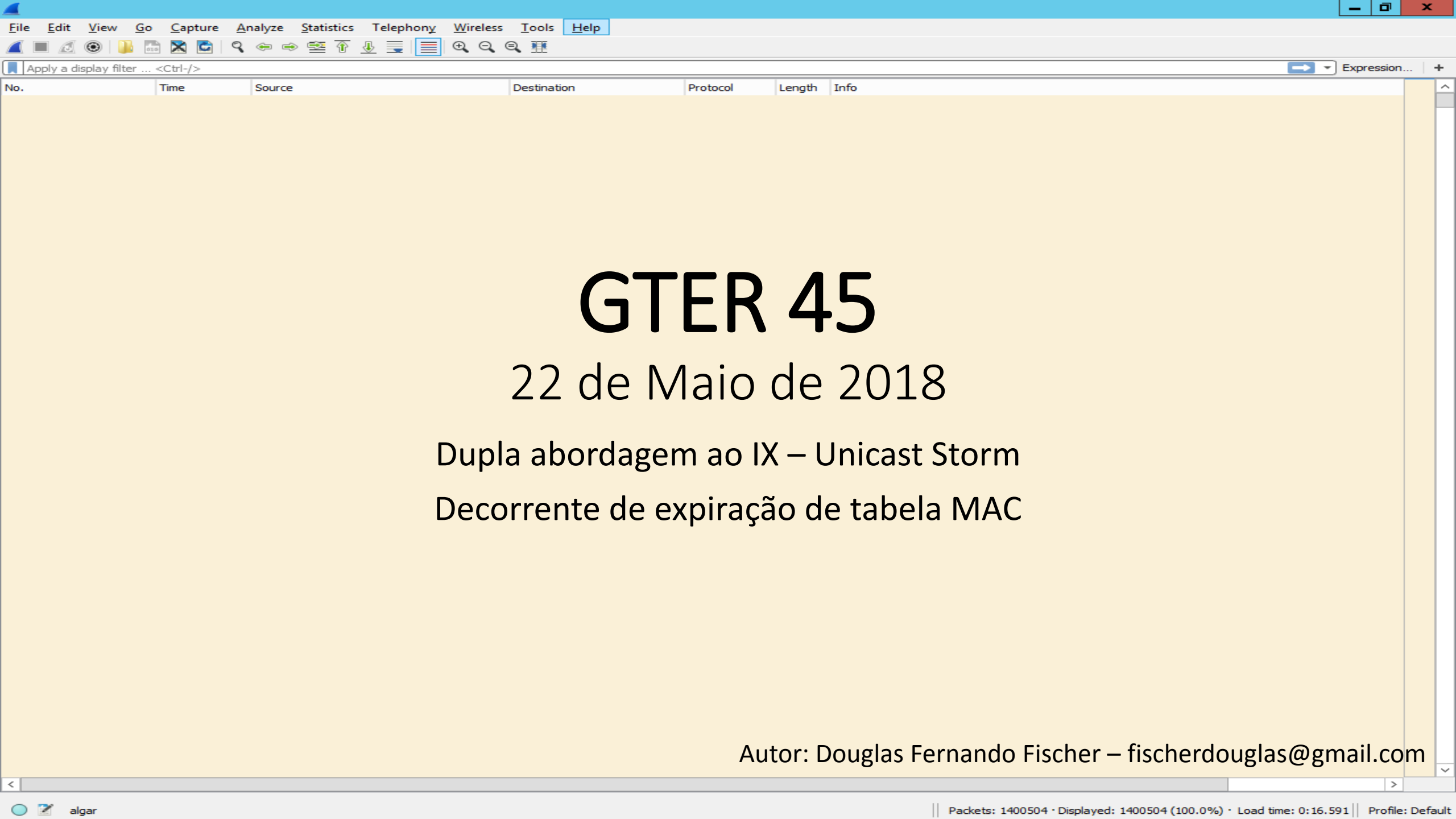


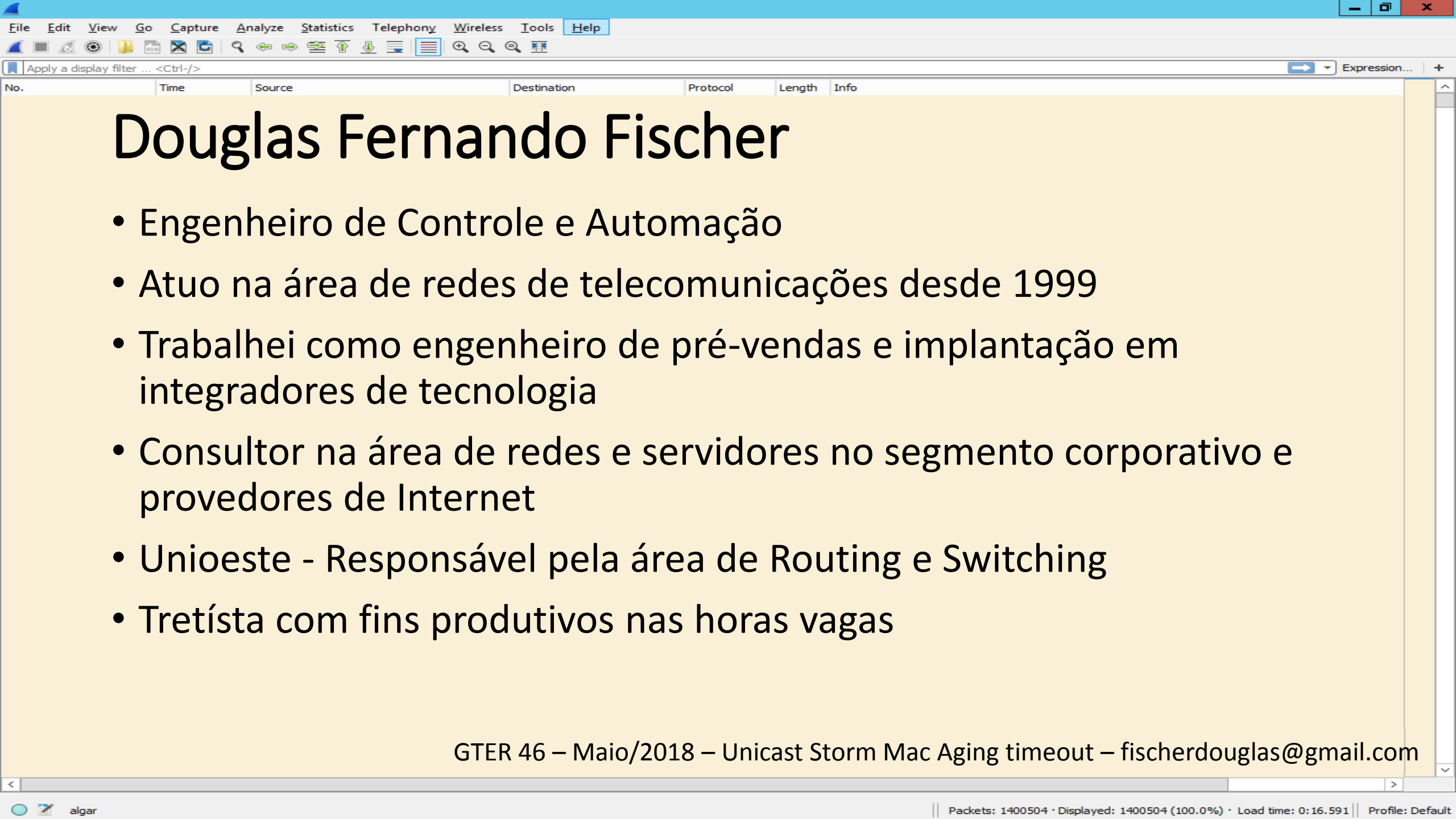
GTER 45

22 de Maio de 2018

Dupla abordagem ao IX – Unicast Storm

Decorrente de expiração de tabela MAC

Autor: Douglas Fernando Fischer – fischerdouglas@gmail.com



Douglas Fernando Fischer

- Engenheiro de Controle e Automação
- Atuo na área de redes de telecomunicações desde 1999
- Trabalhei como engenheiro de pré-vendas e implantação em integradores de tecnologia
- Consultor na área de redes e servidores no segmento corporativo e provedores de Internet
- Unioeste - Responsável pela área de Routing e Switching
- Tretísta com fins produtivos nas horas vagas

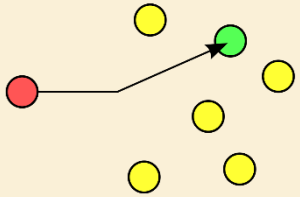
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

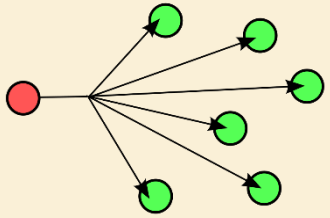
No.	Time	Source	Destination	Protocol	Length	Info

Unicast Flood

Quando um switch sabe em que porta o Mac-Address de destino está, ele encaminha aquele pacote **somente para aquela porta**.



Quando um switch não sabe em qual porta está aquele Mac-Address, ele manda aquele pacote **para todas as portas**.



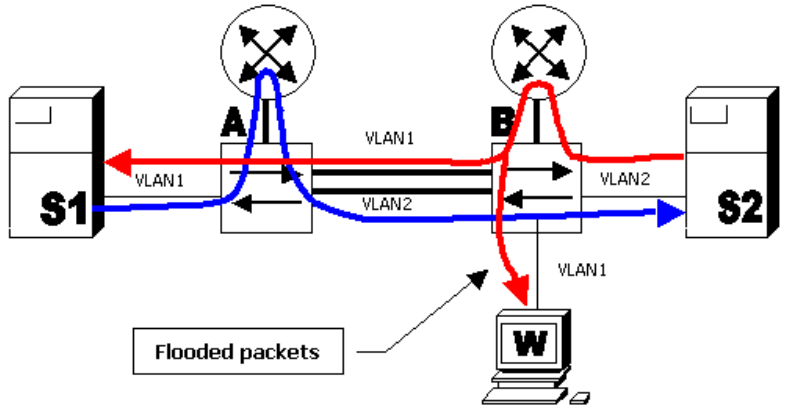
https://en.wikipedia.org/wiki/Unicast_flood

Causes of Flooding

The very cause of flooding is that destination MAC address of the packet is not in the L2 forwarding table of the switch. In this case the packet will be flooded out of all forwarding ports in its VLAN (except the port it was received on). Below case studies display most common reasons for destination MAC address not being known to the switch.

Cause 1: Asymmetric Routing

Large amounts of flooded traffic might saturate low-bandwidth links causing network performance issues or complete connectivity outage to devices connected across such low-bandwidth links. Consider the following diagram:



In the diagram above, server S1 in VLAN 1 is running backup (bulk data transfer) to server S2 in VLAN 2. Server S1 has its default gateway pointing to router A's VLAN 1 interface. Server S2 has its default gateway pointing to router B's VLAN 2 interface. Packets from S1 to S2 will follow this path:

- S1--VLAN 1--switch A--router A--VLAN 2--switch B--VLAN 2--S2 (blue line)

Packets from S2 to S1 go along the following path:

- S2--VLAN 2--switch B--router B--VLAN 1--switch A--flooded to VLAN 1--S1 (red line)

https://www.cisco.com/c/en/us/support/docs/switches/catalyst-6000-series-switches/23563-143.html

algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

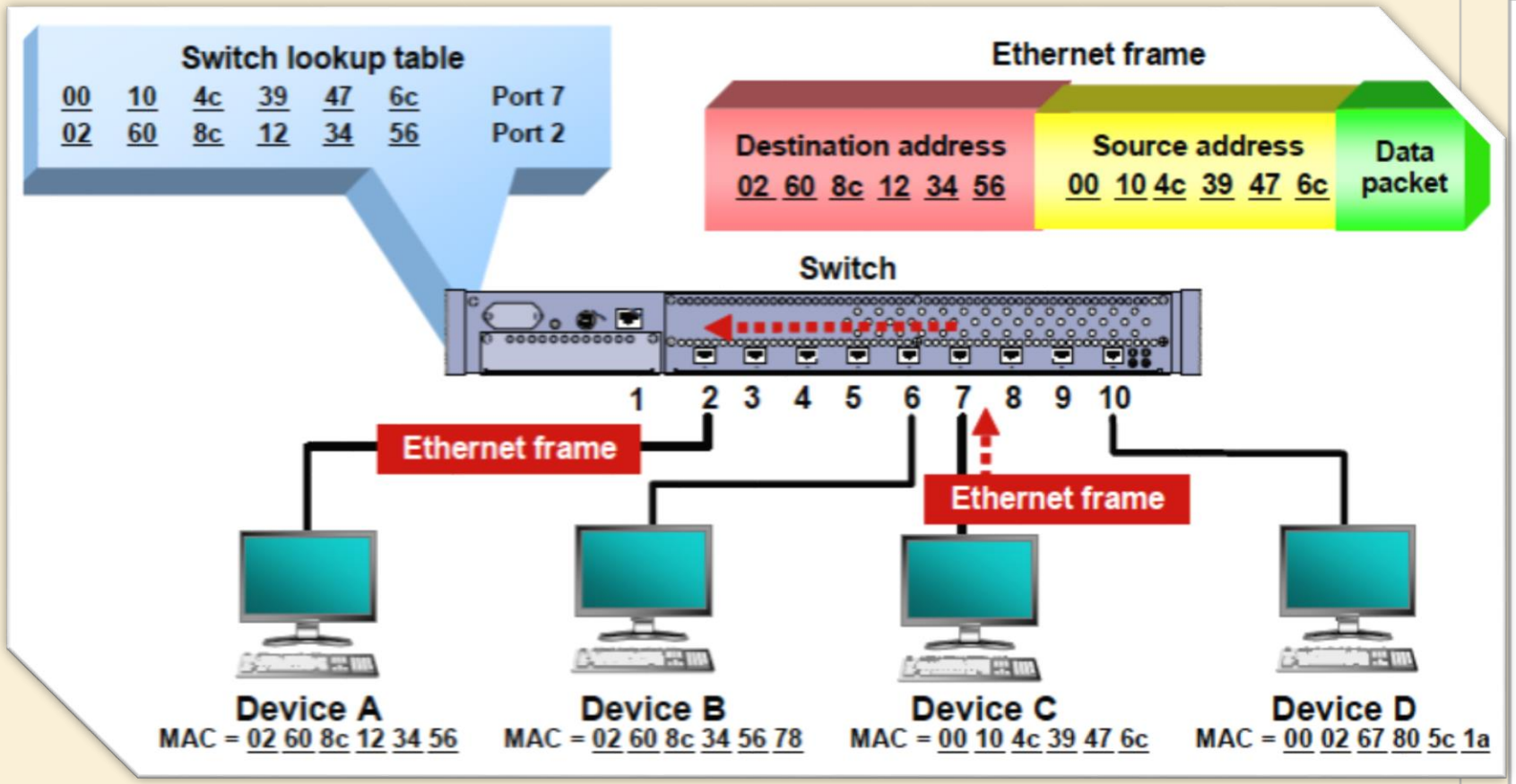
Pausa para relembrar o conceito de Switch.

Switch Monta tabela com base em:

- Endereço de origem do pacote
- Por qual porta entrou aquele pacote

Quando uma entrada é removida dessa tabela?

- Quando um mac que já está na tabela entra por uma porta diferente.
- Quanto o tempo de registro daquela entrada expira.
  (Geralmente 5 minutos)

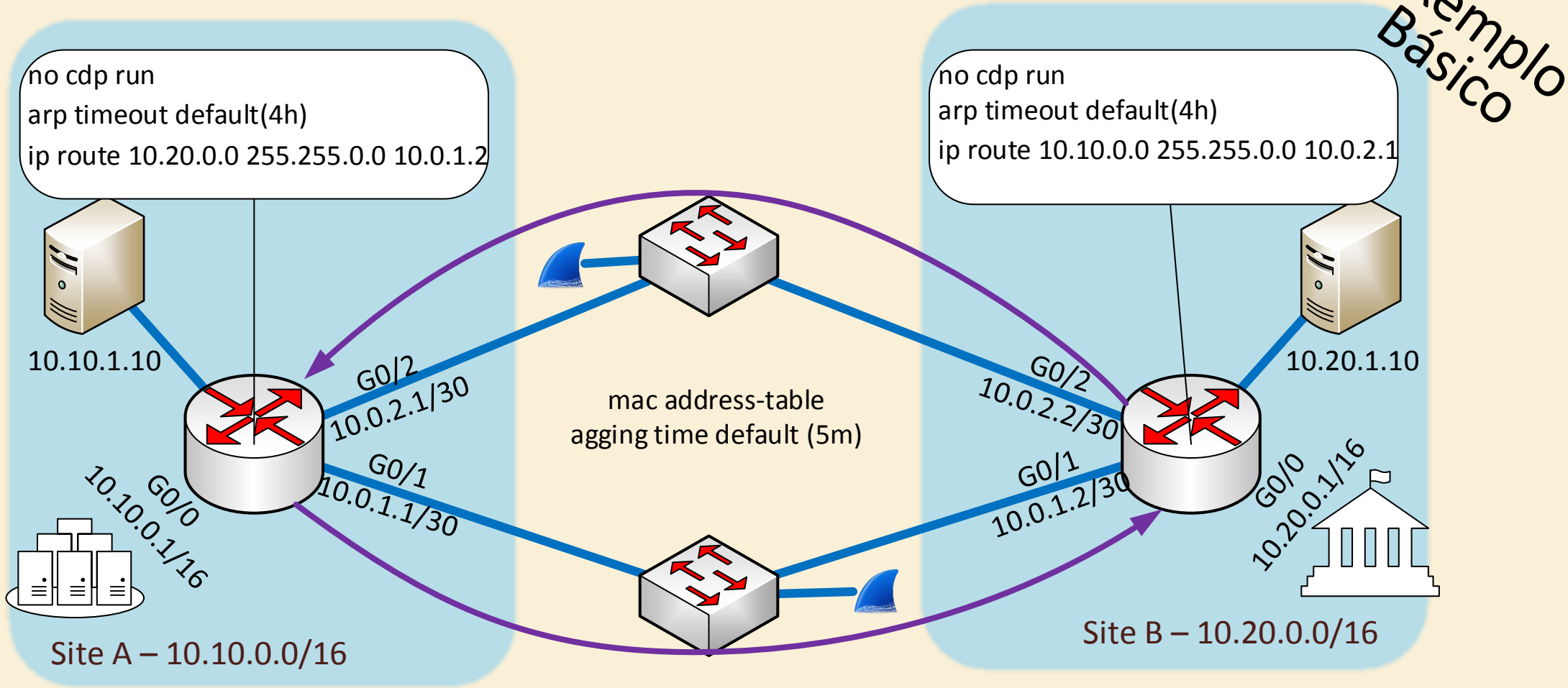


https://www.globalknowledge.com/blog/2012/08/22/how-do-switches-work/

algar Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

Unicast Flood decorrente de roteamento assimétrico

Exemplo Básico



File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Como essa encrenca caiu no meu colo?

https://bitbucket.org/fischerdouglas/arpoador/src/master/

arp@arpoador: ~

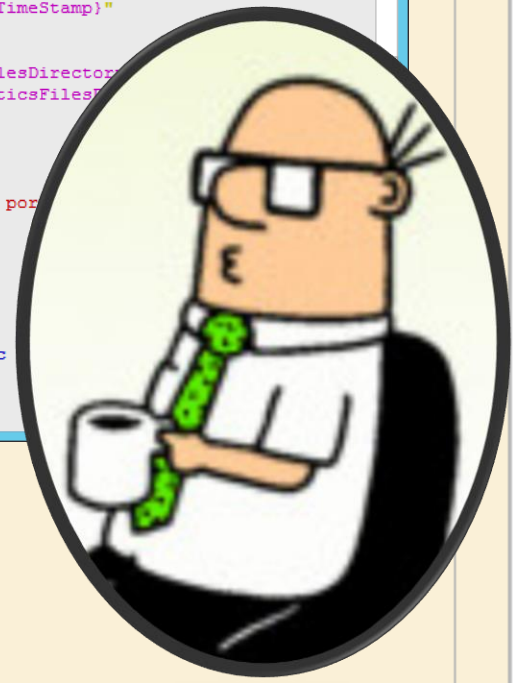
```
#Receive configfile from CLI call, or define default configfile.
if [ -f "$1" ]; then configfile=$1; else   configfile='arpoador.conf'; fi

#Remove basic malicious entry on configfile
configfile_secured='/tmp/arpoador.conf.secure'
egrep '^#|^[^ ]*=[^;&]*'  "$configfile" > "$configfile_secured"

#Get User defined variables from configuration file
source $configfile_secured

#Script Defined Variables
TimeStamp=`date +%Y%m%d%H%M`
BaseCaptureFile="${CaptureFilesDirectory}capture-${IXName}-${TimeStamp}"
StatistcsFile="${StatisticsFilesDirectory}statistics-${IXName}-${TimeStamp}"

#Test existency of Working Directories
if [ ! -d "${CaptureFilesDirectory}" ]; then mkdir -p ${CaptureFilesDirector
if [ ! -d "${StatisticsFilesDirectory}" ]; then mkdir -p ${StatisticsFiles

#Define if capture come from Stream(TZSP) or Directly
if [ ${ReceiveFromStream} = true  ]
    then
        StreamFilterString="host ${StreamSenderIPAddress} and udp por
    else
        StreamFilterString=""
fi


#Does the Capture on the Interface where comes the bridged traffic
tshark -q  -a duration:$CaptureDuration -i $InterfaceOfCapture \
       -w $BaseCaptureFile.pcap -F pcap $StreamFilterString
```

arp@arpoador: ~/arpoador/data/IX-BR-SP/captures

```
arp@arpoador:~/arpoador/data/IX-BR-SP/captures$ ls -hal
total 69M
drwxrwxr-x 2 arp arp 4.0K May 19 18:50 .
drwxrwxr-x 4 arp arp 4.0K May 19 18:12 ..
-rw-rw-r-- 1 arp arp  12M May 19 18:50 capture-IX-BR-SP-201805191849-IPv4.pcap
-rw-rw-r-- 1 arp arp 6.1M May 19 18:50 capture-IX-BR-SP-201805191849-IPv6.pcap
-rw-rw-r-- 1 arp arp 4.0M May 19 18:50 capture-IX-BR-SP-201805191849-NonConform.pcap
-rw------- 1 arp arp  48M May 19 18:50 capture-IX-BR-SP-201805191849.pcap
arp@arpoador:~/arpoador/data/IX-BR-SP/captures$
```

48M - 22,1M = 25,9M

Cadê?



Arpoador

algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info
125447	50.050752	1 .163.162	1 39.214	TCP	107	[TCP Retransmission] 43182 → 38465 [FI
125536	50.090816	1 .84.46	1 136.67	TCP	107	28855 → 60949 [RST, ACK] Seq=1 Ack=1 W
125552	50.093087	1 .184.80	1 45.254	UDP	985	26845 → 20502 Len=892
125555	50.093364	1 .84.169	1 20.230	TCP	125	55672 → 5238 [SYN, ACK] Seq=0 Ack=1 Wi
125648	50.130236	1 5.177	1 4.3	UDP	113	34236 → 12901 Len=20
125650	50.130964	1 .184.108	1 42.118	TCP	107	49818 → 5127 [ACK] Seq=16449 Ack=42215
125675	50.139569	1 5.177	1 4.3	UDP	113	34236 → 12901 Len=20
125700	50.152658	1 157.67	1 236.179	TCP	117	51324 → 30248 [ACK] Seq=875 Ack=530264
125718	50.155257	1 .85.92	4 .3	TCP	117	[TCP Retransmission] 61665 → 59562 [SY
125757	50.167511	1 .184.108	1 42.118	TCP	107	49818 → 5127 [ACK] Seq=16449 Ack=42364
125763	50.168555	1 .84.169	1 20.230	BitTo…	205	Handshake
125767	50.172577	1 6.58	1 144.172	UDP	113	41255 → 57809 Len=20
125782	50.174821	1 157.67	1 236.179	TCP	117	51324 → 30248 [ACK] Seq=875 Ack=531624
125783	50.175002	1 6.58	1 144.172	UDP	113	41255 → 57809 Len=20
125786	50.176119	1 6.58	1 144.172	UDP	113	41255 → 57809 Len=20
125840	50.196876	1 157.67	1 236.179	TCP	117	51324 → 30248 [ACK] Seq=875 Ack=532984
125864	50.209109	4 122.33	1 42.118	TCP	107	50295 → 5127 [ACK] Seq=27769 Ack=40851
125920	50.233301	1 .184.108	1 42.118	TCP	197	49818 → 5127 [PSH, ACK] Seq=16449 Ack=
125926	50.236662	1 .185.19	4 75.75	TCP	210	29026 → 4264 [PSH, ACK] Seq=1 Ack=119
125927	50.236712	1 .185.19	4 75.75	UDP	113	29026 → 4295 Len=20
125931	50.240488	1 .84.169	1 20.230	BitTo…	666	Extended Bitfield, Len:0x4e Have, Pi
125934	50.240995	1 .185.19	4 75.75	UDP	214	29026 → 4295 Len=121
125938	50.242054	1 .84.169	1 20.230	TCP	117	55672 → 5238 [FIN, ACK] Seq=638 Ack=31
125953	50.250146	1 .84.50	1 164.2	ICMP	225	Destination unreachable (Port unreacha
125964	50.255371	1 .186.179	1 91.5	TCP	117	[TCP Window Update] 56009 → 80 [ACK] S
126051	50.287978	1 .184.108	1 42.118	TCP	107	49818 → 5127 [ACK] Seq=16541 Ack=42429
126057	50.288546	1 157.67	1 236.179	TCP	107	51324 → 30248 [ACK] Seq=875 Ack=535704
126104	50.313287	1 .84.169	1 20.230	TCP	117	55672 → 5238 [ACK] Seq=639 Ack=315 Win
126145	50.329543	1 5.177	1 4.3	UDP	113	34236 → 12901 Len=20
126146	50.329545	1 6.58	1 144.172	UDP	113	41255 → 57809 Len=20
126147	50.329546	1 6.58	1 144.172	UDP	113	41255 → 57809 Len=20
126148	50.330144	1 5.177	1 4.3	UDP	113	34236 → 12901 Len=20
126159	50.334256	1 157.67	1 236.179	TCP	107	51324 → 30248 [ACK] Seq=875 Ack=538299
126207	50.348543	1 247.246	1 91.13	TCP	107	50060 → 8291 [SYN] Seq=0 Win=14600 Len
126214	50.349452	1 247.246	1 90.158	TCP	107	6595 → 7547 [SYN] Seq=0 Win=14600 Len=

Interface List

Interface | Interface List | Ethernet | EoIP Tunnel | IP Tunnel | GRE Tunnel | VLAN | VRRP | Bonding | ...

Find

Comment | contains | ix-sp ipv4 | Filter

	Name	Type	Tx	Rx	Tx ...	Rx Pa...	FP Tx	FP Rx
	;;; ix-sp ipv4 - gvt2							
RS	sfpplus1.2	VLAN	0 bps	638.9 kbps	0	1 330	0 bps	638.9 kbps
	;;; ix-sp ipv4 - gvt1							
RS	sfpplus1.1	VLAN	0 bps	7.4 Mbps	0	2 133	0 bps	7.4 Mbps

3 items out of 49 (1 selected)

GTER 45 – Maio/2018 – Unicast Storm em link para o IX – fischerdouglas@gmail.com

algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Pausa para contar causo.

Conclusão:
Os Fabricantes de equipamentos e seus padrões mudam.

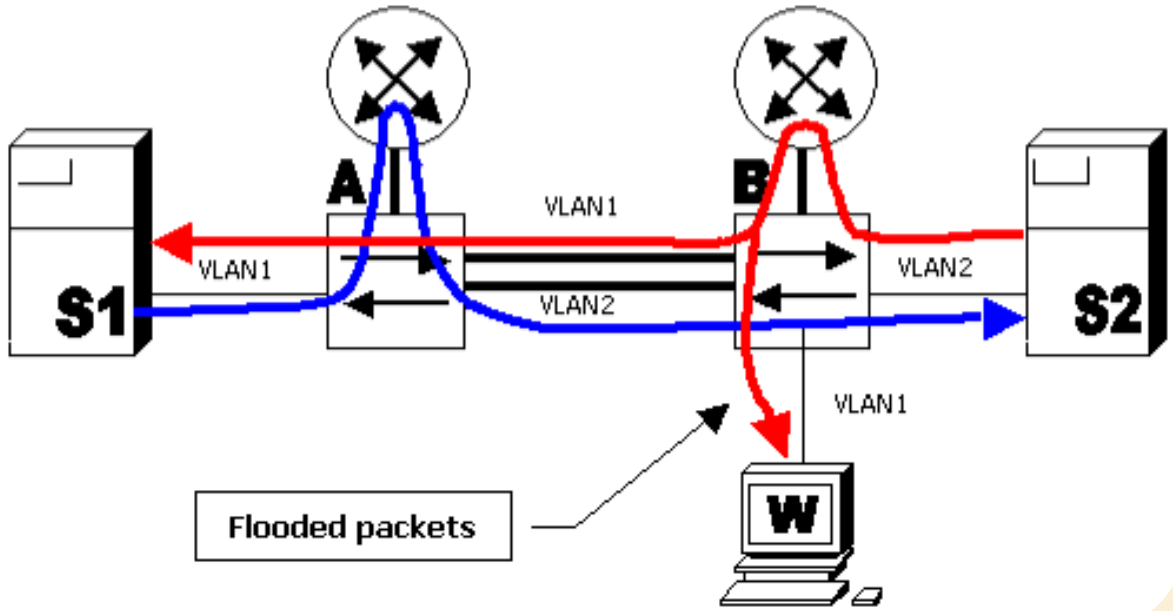
Os protocolos e conceitos de redes são perenes.

Causes of Flooding

The very cause of flooding is that destination MAC address of the packet is no
packet will be flooded out of all forwarding ports in its VLAN (except the port it
reasons for destination MAC address not being known to the switch.

Cause 1: Asymmetric Routing

Large amounts of flooded traffic might saturate low-bandwidth links causing n
to devices connected across such low-bandwidth links. Consider the following



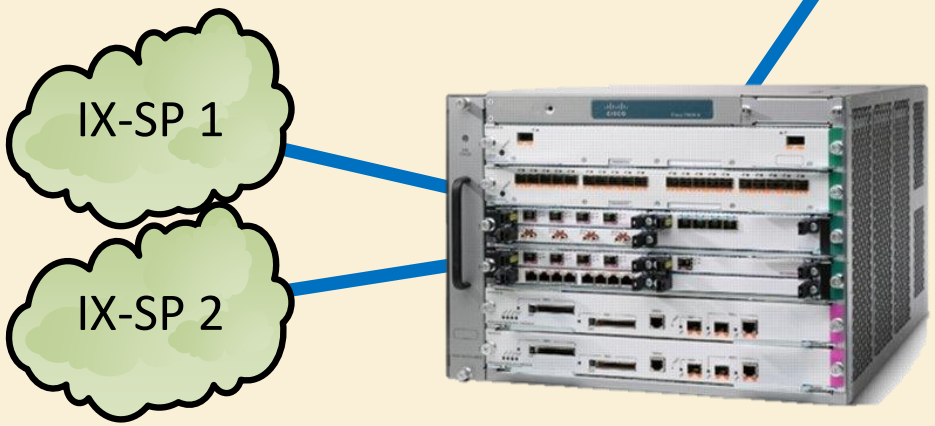
algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

Ajustar MAC Agging Time

Paliativo para redução do Unicast Flood

```
cac-c7600#sh mac-address-table aging-time
Vlan    Aging Time
----    ----------
Global  3600
no vlan age other than global age configured

cac-c7600#show mac address-table vlan 2
Legend: * - primary entry
        age - seconds since last seen
        n/a - not available

  vlan   mac address     type     learn     age
------+---------------+-------+-----+----------+
Module 1:
* 2     e48d.8c15.58ec   dynamic  Yes        240
* 2     4c5e.0cd1.6f52   dynamic  Yes        120
* 2     001e.6706.a301   dynamic  Yes          0
* 2     1005.ca18.ee10   dynamic  Yes       2220
* 2     204e.7150.a54b   dynamic  Yes          0
* 2     6c3b.6bbf.ef30   dynamic  Yes       2700
* 2     e024.7fb6.5abd   dynamic  Yes          0
```



Interface List

Interface	Interface List	Ethernet	EoIP Tunnel	IP Tunnel	GRE Tunnel	VLAN	VRRP	Bonding	LTE

Comment | contains | ix-sp ipv4 | Filter

	Name	Type	Tx	Rx	Tx ...	Rx Pa...	FP Tx	FP Rx
	;;; ix-sp ipv4 - gvt2							
RS	sfpplus1.2	VLAN	0 bps	650.5 kbps	0	1 354	0 bps	650.5 kbps
	;;; ix-sp ipv4 - gvt1							
RS	sfpplus1.1	VLAN	0 bps	693.3 kbps	0	1 444	0 bps	693.3 kbps

3 items out of 49 (1 selected)

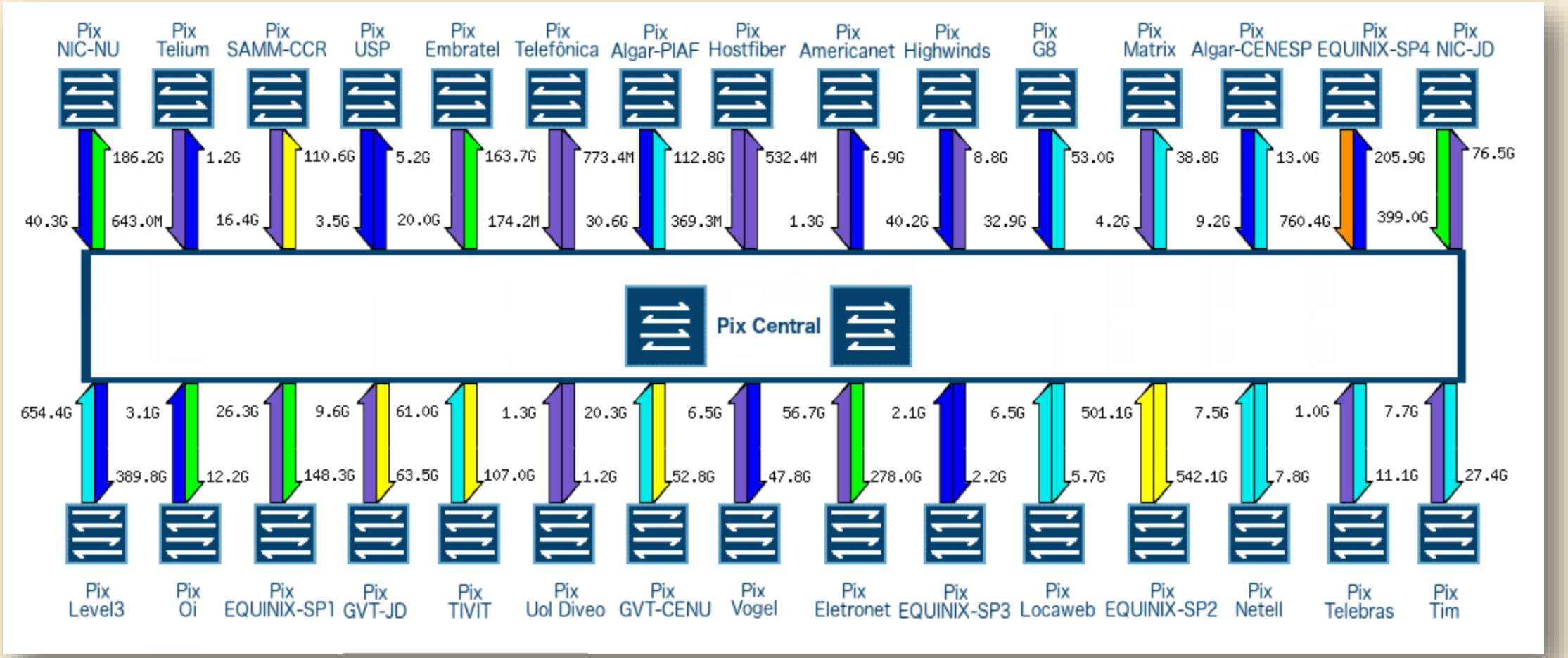
P.S.: Recomendo aplicar especificamente nas vlans envolvidas.
Router(config)#mac address-table agIng-time <tempo> vlan <vlan>

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

OK Douglas.... Mas oque isso tem ver com IX?



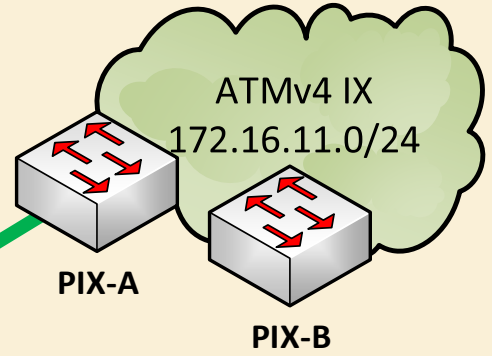
algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

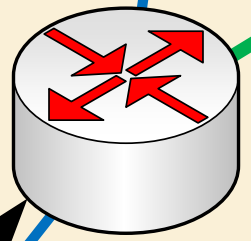
Unicast Flood
Exemplo no IX

Exemplo clássico de um ISP Brasileiro

- Um link de trânsito.
- Um link de transporte até o IX.

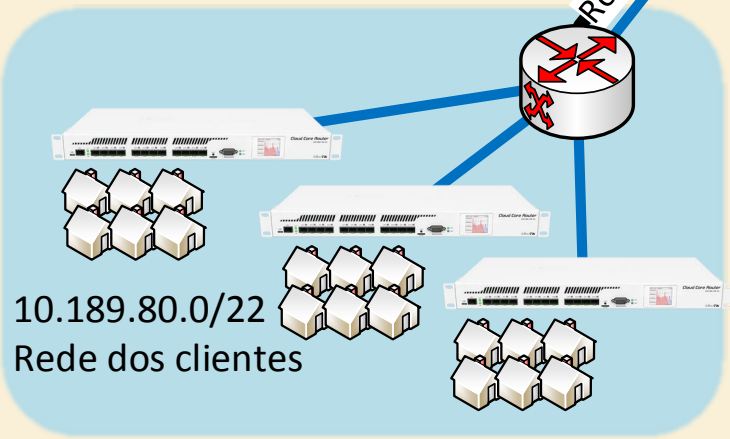


Link A
1Gbps



Roteador de
Borda da Internet

Rota Default



Trânsito
- Anuncia seu prefixos menos específicos
  Ex.: 10.189.80.0/22
- Recebe Full Routing ou default

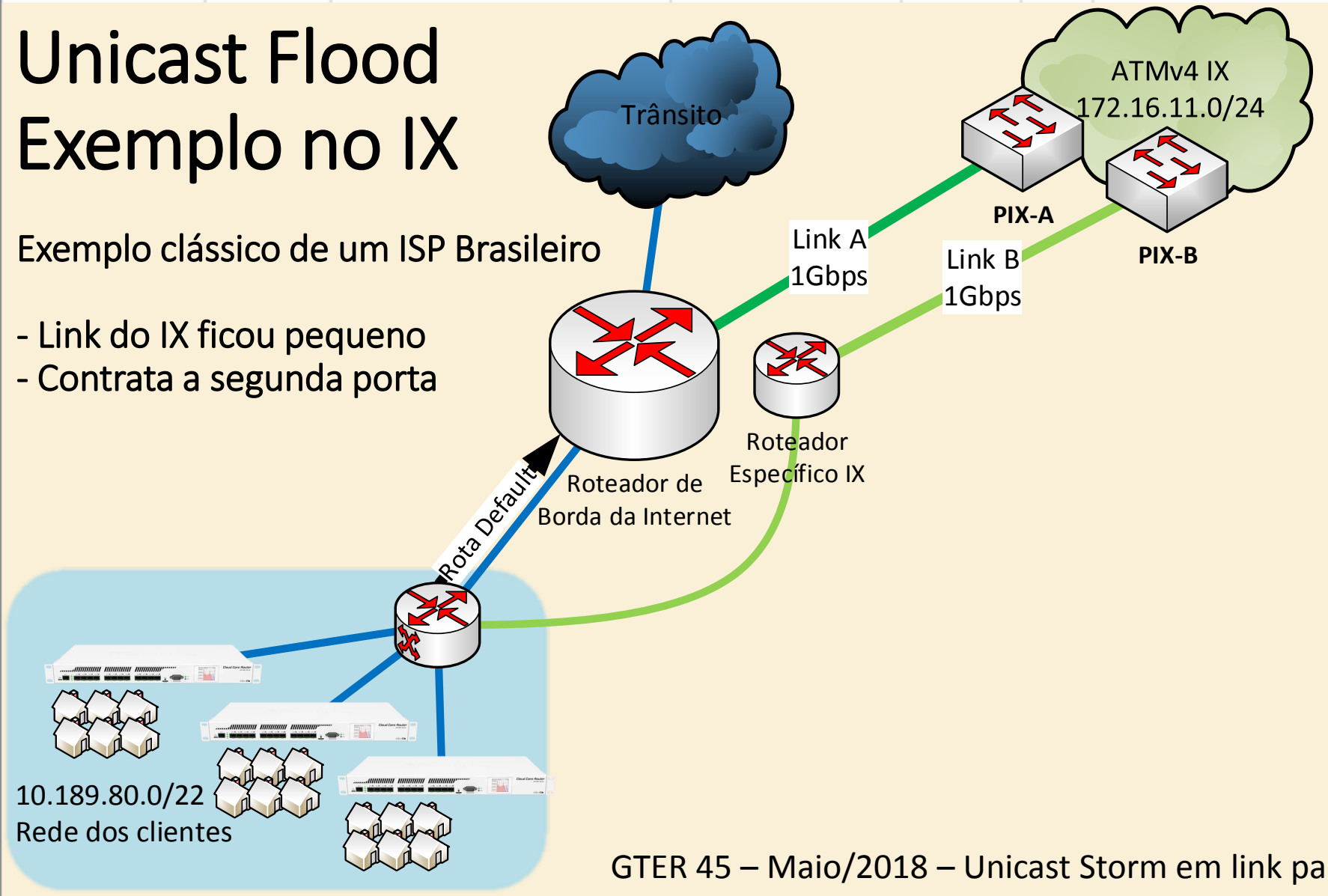
IX
- Anuncia seu prefixos mais específicos
  Ex.: 10.189.80.0/22 e 10.189.82.0/22
- Recebe as rotas dos Route-Servers (set Local-Preference)

Exemplo só com v4 para explanar o raciocínio

Unicast Flood
Exemplo no IX

Exemplo clássico de um ISP Brasileiro

- Link do IX ficou pequeno
- Contrata a segunda porta

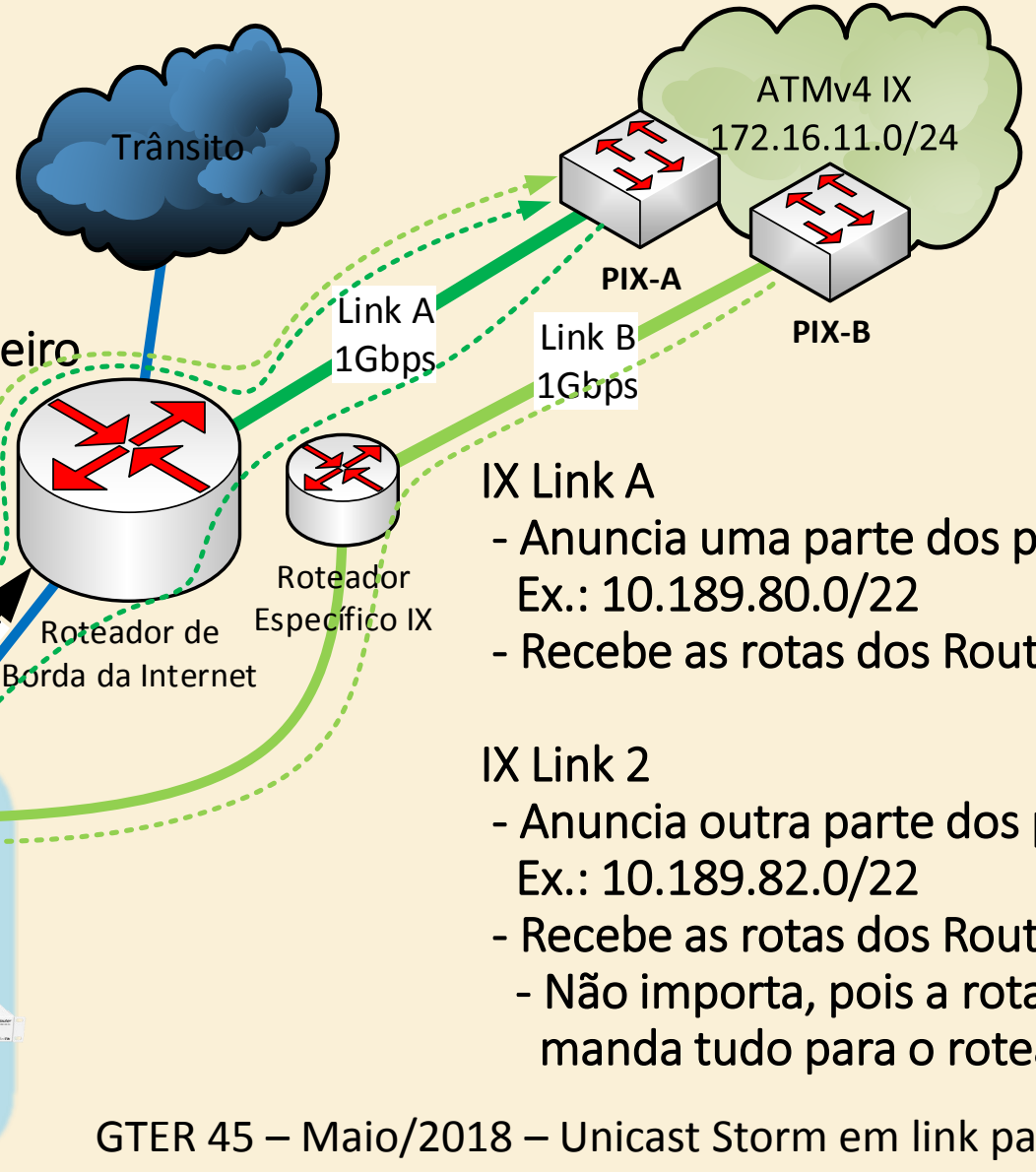


Exemplo só com v4 para explanar o raciocínio

Unicast Flood
Exemplo no IX

Exemplo clássico de um ISP Brasileiro

- Link do IX ficou pequeno
- Contrata a segunda porta



Exemplo só com v4 para explanar o raciocínio

IX Link A
- Anuncia uma parte dos prefixos mais específicos
  Ex.: 10.189.80.0/22
- Recebe as rotas dos Route-Servers (set Local-Preference)

IX Link 2
- Anuncia outra parte dos prefixos mais específicos
  Ex.: 10.189.82.0/22
- Recebe as rotas dos Route-Servers
  - Não importa, pois a rota default da rede interna manda tudo para o roteador principal

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Unicast Flood - Exemplo no IX

Algumas constatações...

algar Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

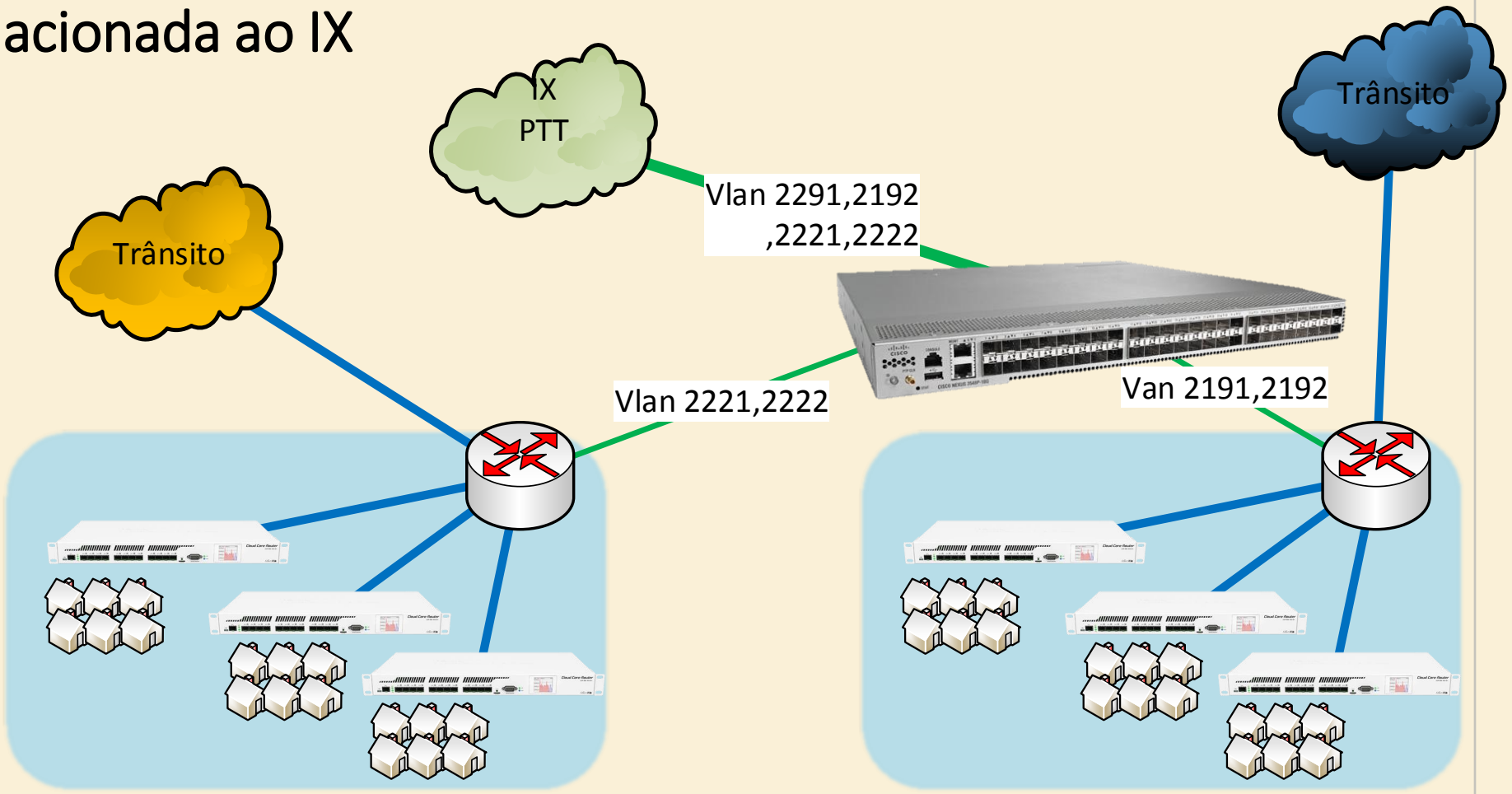
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Unicast Flood – Porta Compartilhada

Outra possibilidade relacionada ao IX



Já houveram relatos na Lista GTER

Seria evitado com Swtchport Alowed Vlan

algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Unicast Flood em ambiente de IX

Opiniões que pude formar até agora.

Os problemas que pude averiguar até agora estão no lado do participante.
Logo a responsabilidade direta é do participante.

A equipe técnica do IX.BR (e outros IX) poderiam colaborar com a orientações técnicas para balizar e evitar que problemas assim aconteçam e afetem a experiência do usuários final.

É um problema pouco representative no momento.
Mas se não forem tomados os devidos cuidados, pode ganhar escala e se tornar um “pequeno monstrinho” como o caso do ARP Excessivo.

Aumetar o Mac Aging-time faz sentido, ameniza bem, mas não resolve completamente.
Casar o Mac Aging-time e o Arp-Timeout faz muito sentido em MINHA OPINIÃO (até o momento).

Alguma comunicação Unicast entre os participantes pode resolver esse problema:
Check-Gateway? Track SLA? “bgp nexthop trigger enable” do Route Dampening?

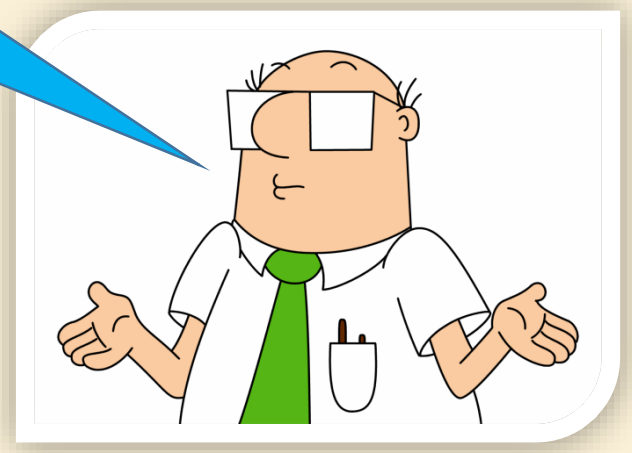
algar Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Length	Info

Perguntas?



“Quando você não sabe o que fazer,

ande rápido e pareça preocupado.”

Wally

algar

Packets: 1400504 · Displayed: 1400504 (100.0%) · Load time: 0:16.591 Profile: Default