Desventuras de un ingeniero de soporte

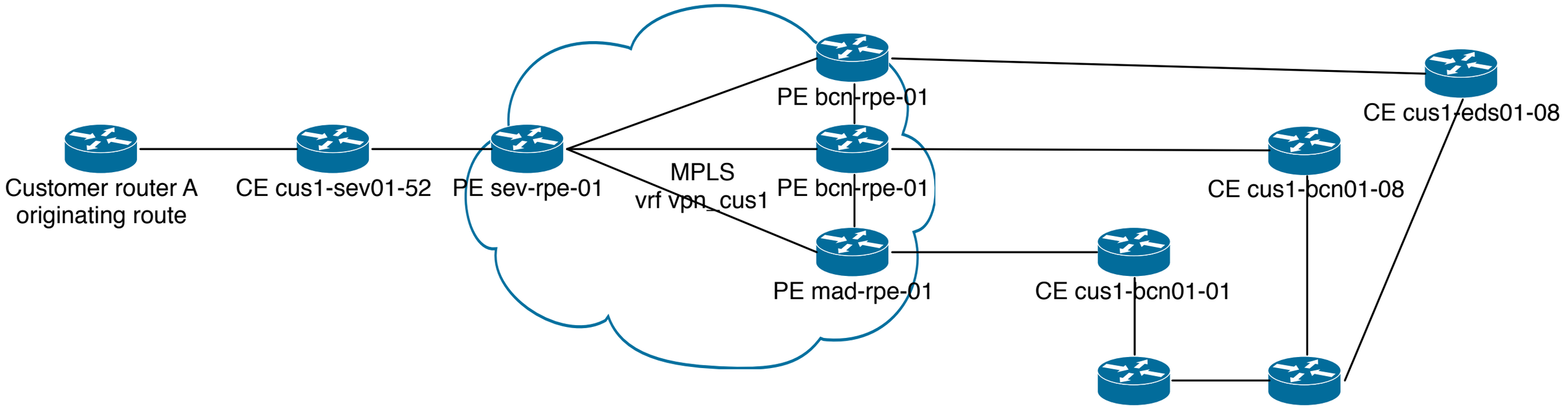
... en la red de un cliente

La conclusión...

Tecnocom

Hoy una de miedo

Escenario inicial



Tecnocom

Escenario inicial (2)

- Core MPLS
- Muestrario de protocolos de routing
  - RIP
  - IS-IS
  - OSPF para el cliente
  - EIGRP
  - MP-BGP por supuesto

Tecnocom

Escenario inicial (3)

- Red pura de Cisco
- Core Cisco 7609

Tecnocom

- De pronto...
- Fusión de empresas

Tecnocom

Diseño político

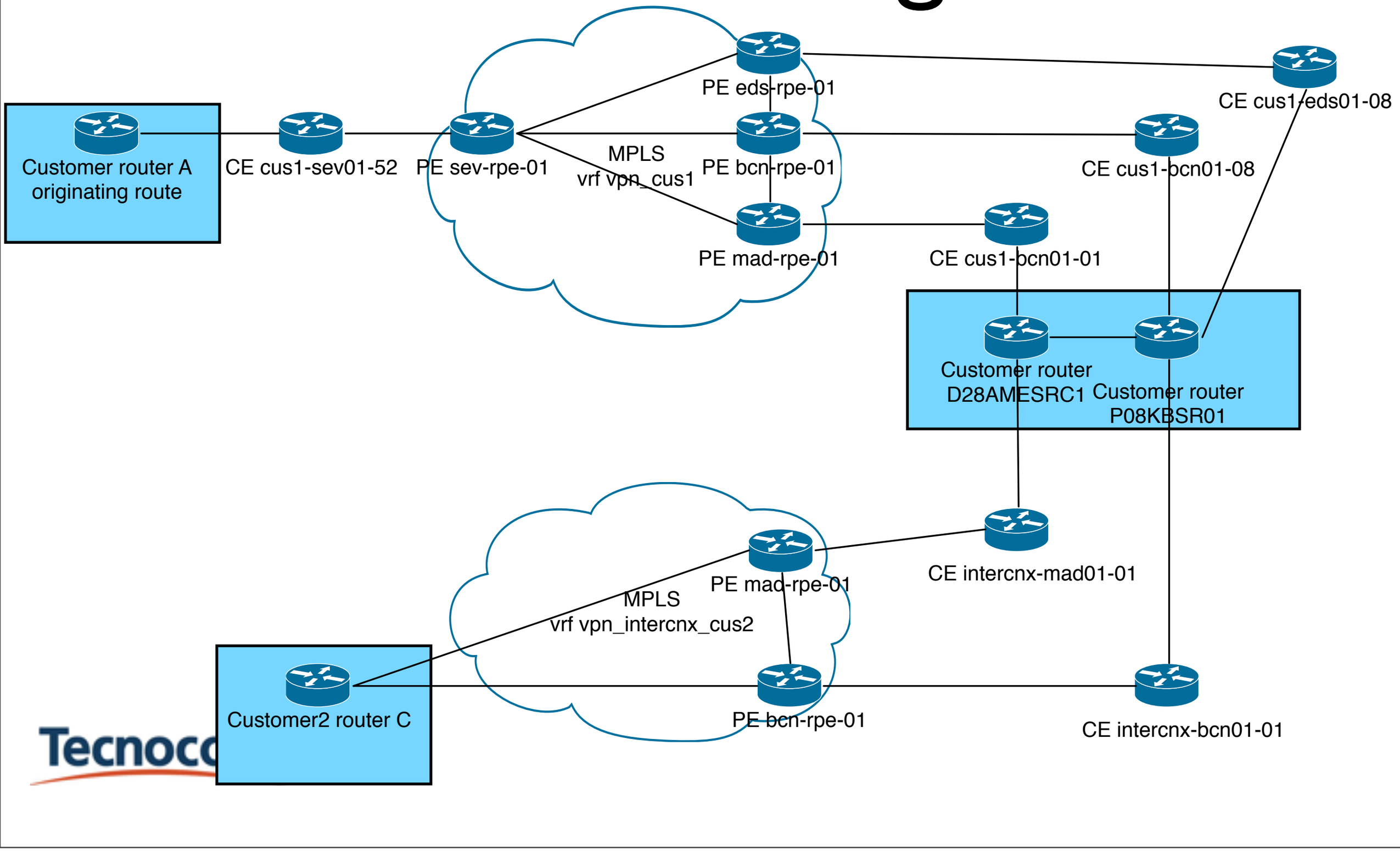
- Se mantienen las redes separadas
- ... pero pensando en fusionarlas

Feliz idea

- Unimos las redes OSPF
- ... sin autoridad administrativa común

Tecnocom

Diseño heterogéneo



Y llegaron las pesadillas

Tecnocom

Llega la incidencia (28/10)

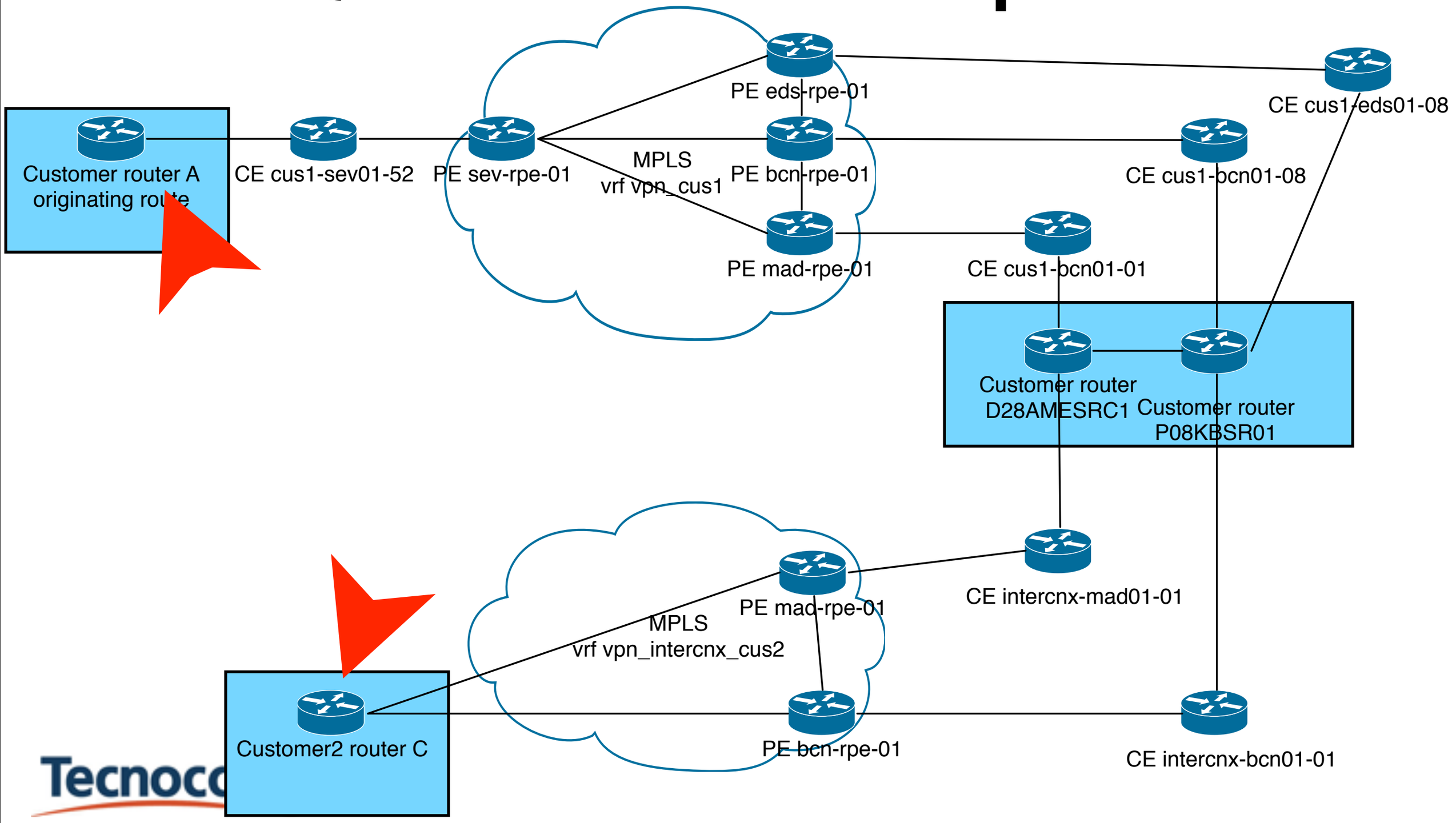
Hola,

Se han detectado cortes intermitentes en las comunicaciones con dos CCC, el CCC de Málaga y el CCC de San Roque. Los cortes son de poca duración y difíciles de localizar. Se solicita que se monitorice de manera continua para detectar el problema cuando se produzca.

Nos es bastante urgente su revisión.

Tecnocom

Quien no ve a quien



Mas información (28/10)

Cuando ocurren los cortes sólo se ven afectadas estas dos centrales y de manera simultánea

Otras centrales que funcionan correctamente con las que se puede comparar son:

- A, B y C

Se ha producido un nuevo corte hoy a las 13:10.

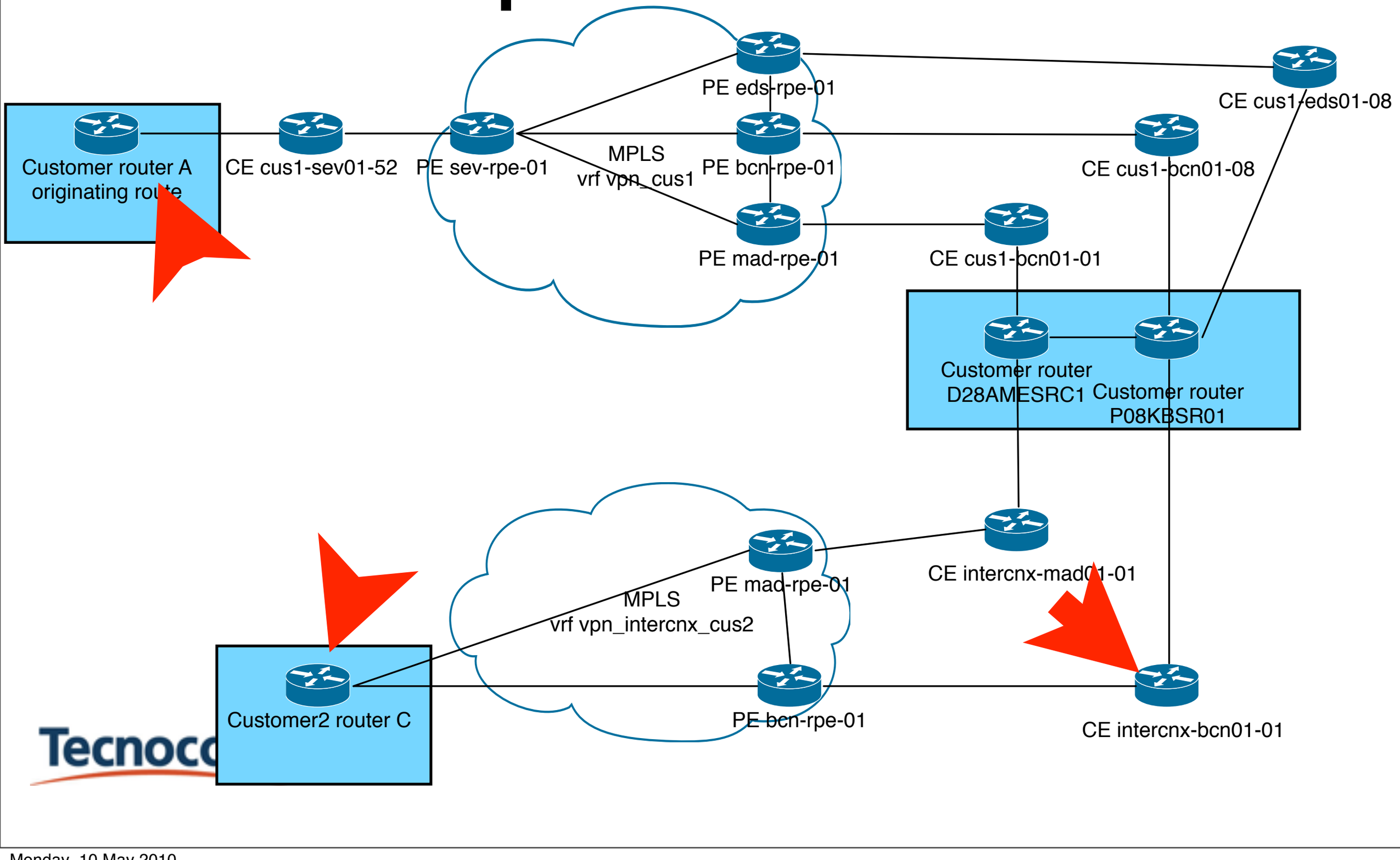
Tecnocom

Traceroute (28/10)

```
1   10.105.17.254 (10.105.17.254)  0.446 ms  0.312 ms  0.379 ms
2   10.105.20.2 (10.105.20.2)  0.593 ms  0.554 ms  0.520 ms
3   10.102.238.252 (10.102.238.252)  1.167 ms  0.969 ms  1.084 ms
4   192.168.253.66 (192.168.253.66)  1.167 ms  1.115 ms  1.085 ms
5   10.55.1.253 (10.55.1.253)  4.936 ms  1.397 ms  2.062 ms
6   10.200.6.13 (10.200.6.13)  1.025 ms  0.981 ms  0.936 ms
7   10.200.6.6 (10.200.6.6)  1.168 ms  1.111 ms  1.082 ms
8   192.168.57.6 (192.168.57.6)  11.853 ms  11.658 ms  11.763 ms
9   10.161.1.125 (10.161.1.125)  12.125 ms  12.077 ms  12.044 ms
10   10.2.254.57 (10.2.254.57)  12.689 ms  12.492 ms  12.046 ms
11   192.168.45.37 (192.168.45.37)  13.394 ms  12.636 ms  12.888 ms
12   192.168.53.6 (192.168.53.6)  32.223 ms  32.175 ms  33.690 ms
13   10.5.34.229 (10.5.34.229)  33.911 ms  34.701 ms  34.148 ms
14   10.5.34.6 (10.5.34.6)  33.434 ms  34.954 ms  33.826 ms
15   10.5.34.70 (10.5.34.70)  34.334 ms  34.280 ms  34.250 ms
```

Tecnocom

Donde perdemos la ruta



Que le pasa a la ruta (28/10)

```
cus1-sev01-138#sh ip rou 10.5.34.70
Routing entry for 10.5.34.64/27
  Known via "ospf 1", distance 110, metric 1020, type extern 1
  Last update from 10.5.34.229 on GigabitEthernet0/1, 00:08:33 ago
  Routing Descriptor Blocks:
  * 10.5.34.229, from 10.5.34.1, 00:08:33 ago, via GigabitEthernet0/1
      Route metric is 1020, traffic share count is 1

cus1-rib01-27#sh ip rou  10.5.30.30
Routing entry for 10.5.30.0/24
  Known via "ospf 1", distance 110, metric 1001, type intra area
  Last update from 10.5.31.233 on GigabitEthernet0/1, 4d22h ago
  Routing Descriptor Blocks:
  * 10.5.31.233, from 10.5.31.1, 4d22h ago, via GigabitEthernet0/1
      Route metric is 1001, traffic share count is 1
```

Tecnocom

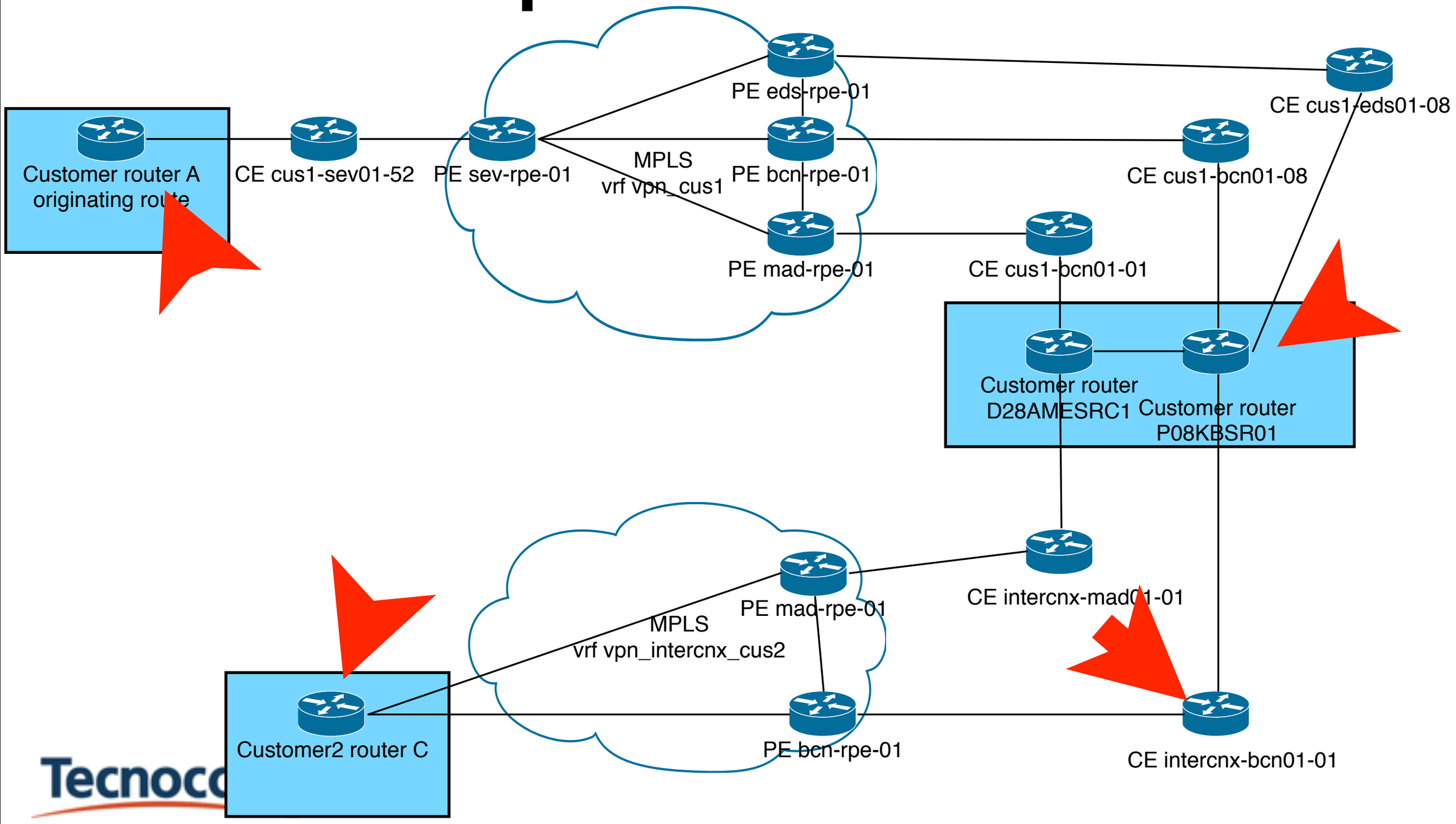
Mas incidencias (29/10)

Acaban de detectar un corte de 4 segundos en la línea (tanto a la IP 10.5.34.1 como a 10.5.34.70)

Nosotros tenemos un ping lanzado desde un servidor en el CPD de Barcelona y no hemos notado ningún corte

Tecnocom

Donde perdemos la ruta



Mas incidencias (30/10-05/11)

Estamos revisando nuestro equipo P08KBWSR01 y estamos viendo que recibimos siempre LSA de equipo se DC. Mañana seguimos revisando.

```
P08KBWSR01#sh ip route 10.133.21.0
Routing entry for 10.133.21.0/27
  Known via "ospf 1", distance 110, metric 3102, type inter area
  Redistributing via ospf 10
  Advertised by ospf 10 metric-type 1 subnets tag 1 match internal
route-map Telefonica-GNI
  Last update from 10.2.254.57 on GigabitEthernet2/9, 00:02:01 ago
  Routing Descriptor Blocks:
  * 10.2.254.57, from 192.168.43.198, 00:02:01 ago, via
GigabitEthernet2/9
      Route metric is 3102, traffic share count is 1
```

Tecnocom

Resumiendo pruebas

- Desconectar intercnx-mad01-01. cus1-mad01-01
  - Sin cambios
  - debug ip opsf spf
  - Tiró abajo intercnx-mad01-01

Tecnocom

Se abre caso a Cisco (05-11)

1. Are all the affected prefixes originated from behind specific PE router?

No. Along the day we made some tests and discover that we lose routes behind several PE routers. And the lost is seen also in several destination PEs. Sometimes simultaneously, but sometimes the routes are lost in one PE and not in another.

2. Are these prefixes lost after the Sham-link flap?

In fact there is NO sham-link flam. We checked along one day and didn't see any flap of the sham-link.

3. Have you observed any activity (like flap ,etc) prior to the issue?

No. The only thing that we could see is that the lost follows more or less a period of 30 minutes. We don't lose the routes every 30 minutes, but many times they follow this pattern.

Please pick up one affected prefix and let me know following information:

1. Where it is originated

One prefix originated in gni-sev01-117 (a CE connected to sev-rpe-01) is 10.133.28.0

2. Which PE device advertisises it to the network.

sev-rpe-01

3. Which devices lost the prefix from the routing table (from which protocol).

bcn-pre-01 lost the prefix announced through the sham-link, but we still see it in the same PE through MP-BGP.

Tecnocom

Después de varias pruebas (12-11)

```
bcn-rpe-01#    sh ip ospf database router 192.168.43.198

            OSPF Router with ID (192.168.43.203) (Process ID 2)

                Router Link States (Area 0)

  Adv Router is not-reachable
  LS age: 1006 (DoNotAge)
  Options: (No TOS-capability, DC)
  LS Type: Router Links
  Link State ID: 192.168.43.198
  Advertising Router: 192.168.43.198
  LS Seq Number: 8000A20F
  Checksum: 0x7D9E
  Length: 132
  Area Border Router
  AS Boundary Router
  Number of Links: 9
```

Tecnocom

Descubrimientos (19-11)

What I have discovered is that in the originating router the LSA Age was correct. It started with 1 and get renewed in about hals an hour (2004 seconds in one of the cases).

Looking at the same LSA in other PEs in the same moment we see that after one timeout the starting Age is 1566.

Tecnocom

Mas descubrimientos (19-11)

I have examined the logs I sent you more in depth and the problem is more strange. You can check this information looking at those logs.
Normally the LSA received in mad-rpe-01 (the destination router) has the Do Not Age bit activated:

```
            OSPF Router with ID (192.168.43.193) (Process ID 1)
                Router Link States (Area 0)

 Routing Bit Set on this LSA
 LS age: 1605 (DoNotAge)
 Options: (No TOS-capability, DC)
 LS Type: Router Links
 Link State ID: 192.168.43.198
 Advertising Router: 192.168.43.198
 LS Seq Number: 8000A36B
```

Tecnocom

Mas descubrimientos (19-11)

```
The same happens in the CE connected to him:
          OSPF Router with ID (10.2.254.17) (Process ID 1)

              Router Link States (Area 0)

  Routing Bit Set on this LSA
  LS age: 1606 (DoNotAge)
  Options: (No TOS-capability, DC)
  LS Type: Router Links
  Link State ID: 192.168.43.198
  Advertising Router: 192.168.43.198
  LS Seq Number: 8000A36B
```

Tecnocom

Mas descubrimientos (19-11)

But the DoNotAge bit doesn't arrive to intercnx-mad01-01 (the CE to the other side of the customer router):

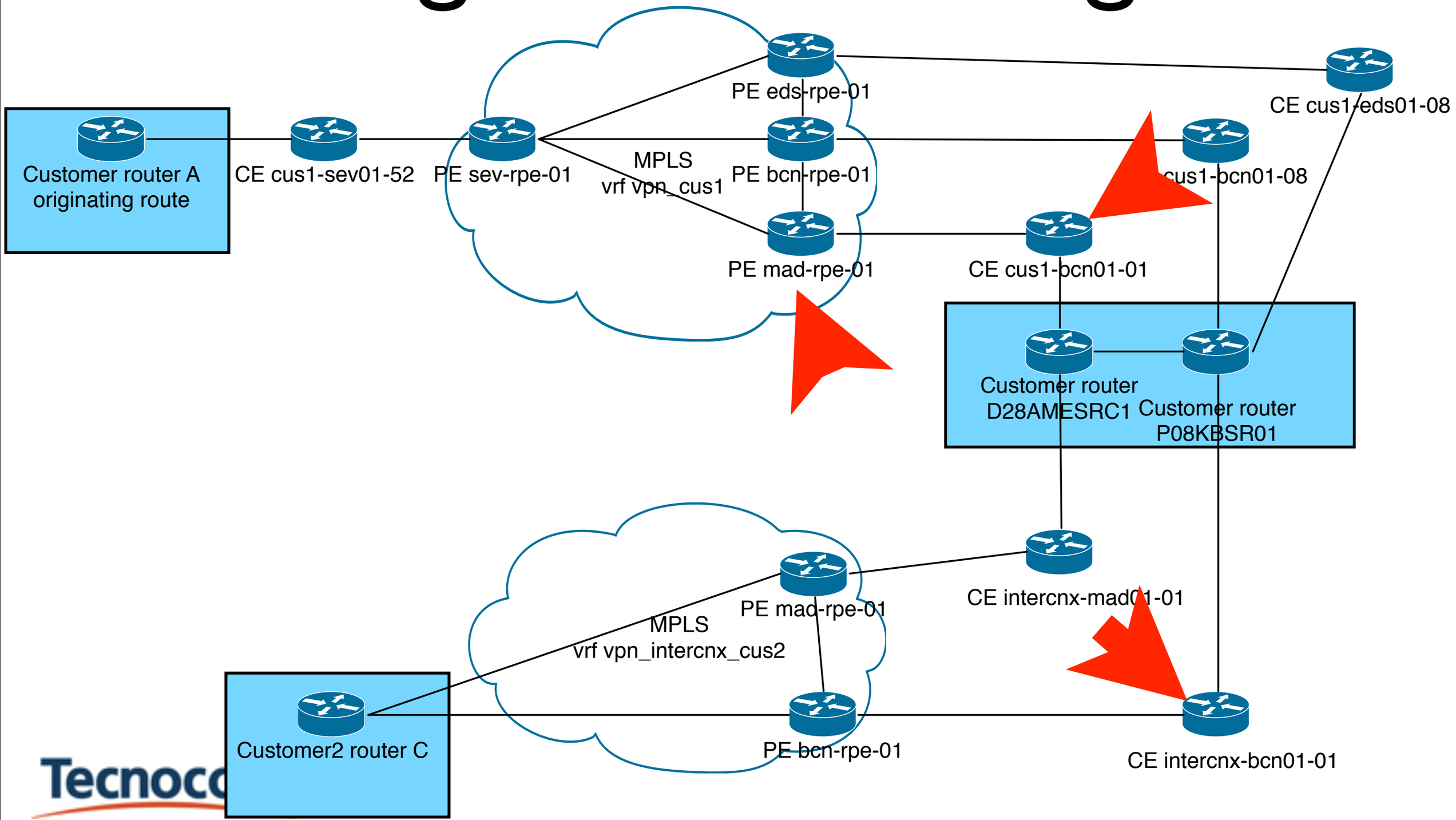
```
intercnx-mad01-01#sh ip ospf database router 192.168.43.198

            OSPF Router with ID (10.161.1.130) (Process ID 1)
                Router Link States (Area 0)

 Routing Bit Set on this LSA
 LS age: 1922
 Options: (No TOS-capability, DC)
 LS Type: Router Links
 Link State ID: 192.168.43.198
 Advertising Router: 192.168.43.198
 LS Seq Number: 8000A36B
```

Tecnocom

To Age or not To Age...



Mas descubrimientos (19-11)

The LS Age in this router nevers goes down of 1600. This is LS Age measured each 10 seconds aprox:

1944
1955
...
....
3573
3584
3595
3606
3617
1620
1631
1642

¡Aja! (20/11)

```
source PE (sev-rpe-01):
OSPF Router with ID
(192.168.43.198) (Process ID 1)
     Router Link States (Area 0)
 LS age: 1625
 Options: (No TOS-capability, DC)
 LS Type: Router Links
 Link State ID: 192.168.43.198
 Advertising Router:
192.168.43.198
 LS Seq Number: 8000A3A0
```

```
destination PE (mad-rpe-01):
OSPF Router with ID
(192.168.43.193) (Process ID 1)
     Router Link States (Area 0)
 Routing Bit Set on this LSA
 LS age: 1655 (DoNotAge)
 Options: (No TOS-capability, DC)
 LS Type: Router Links
 Link State ID: 192.168.43.198
 Advertising Router:
192.168.43.198
 LS Seq Number: 8000A39F
```

Tecnocom

El programador de OSPF (20/11)

I got quick replay from the Developer.
So it is possible that you are hitting CSCej89011. Please Chekc if all the devices in area (including all customer devices) are running the code with fix.
Second thing. Regarding the additional delay of 1600s. Please check also the sequence number in the LSA. Is is changing or it is the same ?
Developer suspects it has been somehow re-floode

Tecnocom

El fin por ahora (23/11)

Today I had also small chat with the developer. The scenario most likely looks like this:
For a network like this:

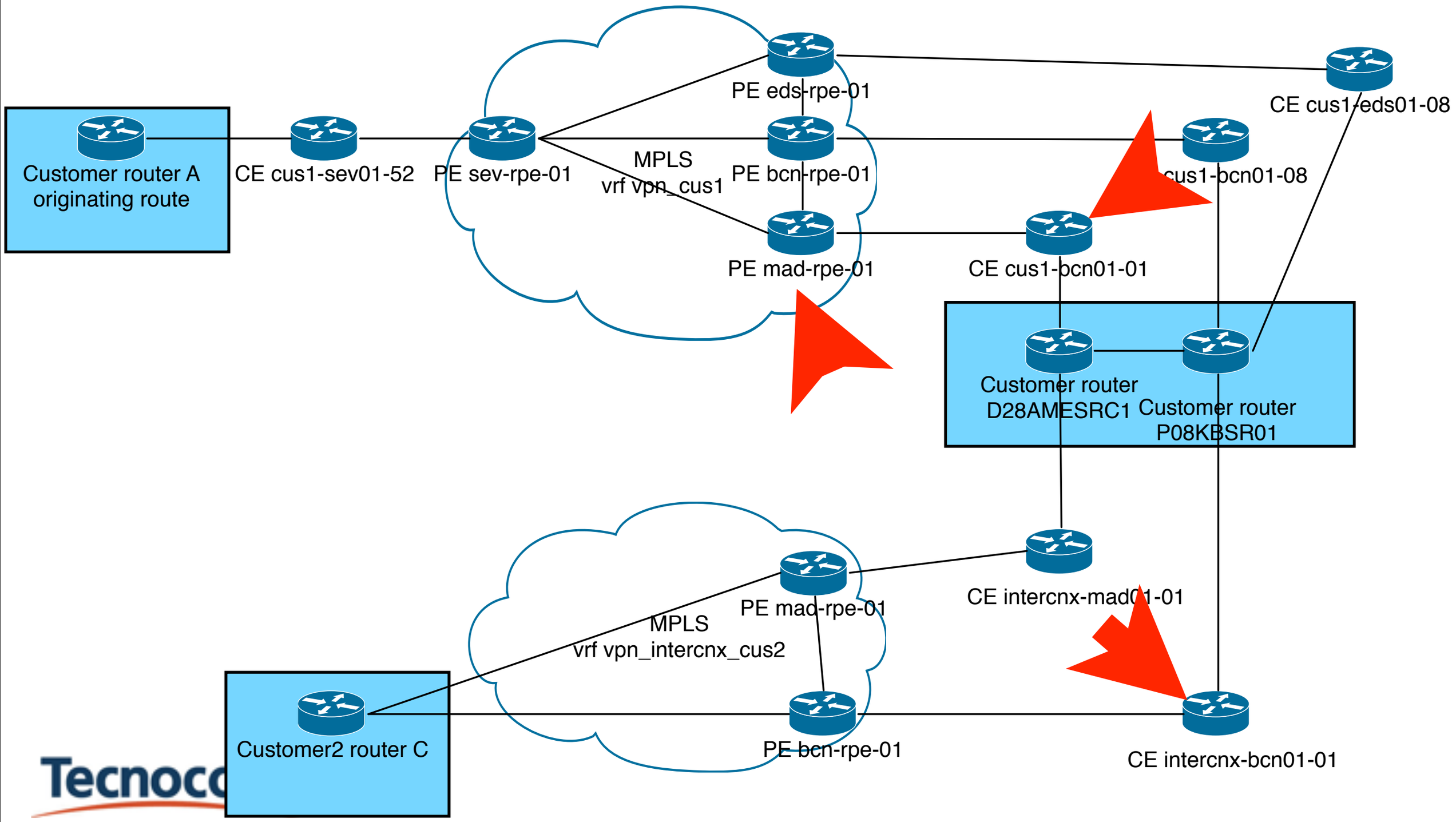
CE1 ---- PE1---{mpls}---PE2---CE2---C1---C2

CE1 is the one originating the LSA for the prefix.

1. C1 and/or C2 have in dbase LSA without DNA bit (probably because of CSCej89011)
2. CE1 periodically refreshes the LSA (on ~1800 sec), but periodic refresh do not go via DC, so destination PE, C1, C2 will not receive it.

Tecnocom

La solución...


Customer router A
originating route
CE cus1-sev01-52
PE sev-rpe-01
MPLS
vrf vpn_cus1
PE eds-rpe-01
PE bcn-rpe-01
PE mad-rpe-01
CE cus1-eds01-08
cus1-bcn01-08
CE cus1-bcn01-01
Customer router
D28AMESRC1
Customer router
P08KBSR01
CE intercnx-mad01-01
PE mad-rpe-01
MPLS
vrf vpn_intercnx_cus2
Customer2 router C
PE bcn-rpe-01
CE intercnx-bcn01-01

El fin por ahora (23/11)

3. LSA reaches the max-age after 3600 sec on C1 or C2, so C1 or C2 try to flush LSA out of the domain.
4. CE1 receives max-age LSA. It has in dbase valid LSA with higher SQN and age about 1600 sec. CE1 sends this LSA in response.
5. Shortly after that (at age 1800+) CE1 re-originate LSA with SQN+1 and age 1. This LSA is flooded to PE1, but it is not flooded over DC so PE2, C1, C2 will not receive it. It's expected.

Therefore the conclusion is to upgrade the IOS of all the devices within both VRFs as you already scheduled.

Tecnocom

Evolución temporal

Hora	sev-rpe	mad-rpe	gni-mad	intercnx
T-10	1586	1610 (DoNotAge)	1611 (DoNotAge)	3589
	8000A36D	8000A36C	8000A36C	8000A36C
T	1597	1610 (DoNotAge)	1611 (DoNotAge)	MAXAGE(3602)
	8000A36D	8000A36C	8000A36C	8000A36C
T+10	1608	1610 (DoNotAge)	1611 (DoNotAge)	MAXAGE(3613)
	8000A36D	8000A36C	8000A36C	8000A36C
T+20	1619	1610 (DoNotAge)	1611 (DoNotAge)	MAXAGE(3624)
	8000A36D	8000A36C	8000A36C	8000A36C
T+30	1630	1610 (DoNotAge)	1611 (DoNotAge)	MAXAGE(3635)
	8000A36D	8000A36C	8000A36C	8000A36C
T+40	1641	1610 (DoNotAge)	1611 (DoNotAge)	MAXAGE(3646)
	8000A36D	8000A36C	8000A36C	8000A36C
T+50	1652	1663 (DoNotAge)	1664 (DoNotAge)	1667
	8000A36D	8000A36D	8000A36D	8000A36D

Tecnocom

Actualizamos nuestra red

Tecnocom

Actualizamos nuestra red

- El problema sigue

Actualizamos nuestra red

- El problema sigue
- Hay equipos de la red del cliente en el area 0 de OSPF

Actualizamos nuestra red

- El problema sigue
- Hay equipos de la red del cliente en el area 0 de OSPF
- Les pedimos que actualicen

Tecnocom

Actualizamos nuestra red

- El problema sigue
- Hay equipos de la red del cliente en el area 0 de OSPF
- Les pedimos que actualicen
- Tardan...

Tecnocom

Para dar mas alegría...

Tecnocom

Para dar mas alegría...

- se ha decidido que se realizarán los upgrades de los routers de Orange y Telefónica que estén afectados (son los únicos equipos routers en nuestra red ya que el resto son switch-routers).

Tecnocom

Para dar mas alegría...

- se ha decidido que se realizarán los upgrades de los routers de Orange y Telefónica que estén afectados (son los únicos equipos routers en nuestra red ya que el resto son switch-routers).
- Orange y Telefonica al revoltillo

Tecnocom

Justificar más

Tecnocom

Justificar más

Telefónica nos está solicitando información sobre el bug de software para proceder a realizar los upgrade en sus equipos

Tecnocom

Justificar más

Telefónica nos está solicitando información sobre el bug de software para proceder a realizar los upgrade en sus equipos

...

Tecnocom

Justificar más

Telefónica nos está solicitando información sobre el bug de software para proceder a realizar los upgrade en sus equipos

...

La información del bug se la hemos proporcionado pero en el propio bug no se especifica la necesidad de realizar el upgrade en todo elemento que pertenece al área 0.

Tecnocom

Justificar más

Telefónica nos está solicitando información sobre el bug de software para proceder a realizar los upgrade en sus equipos

...

La información del bug se la hemos proporcionado pero en el propio bug no se especifica la necesidad de realizar el upgrade en todo elemento que pertenece al área 0.

Podríamos proporcionarles la información dada por el TAC de Cisco?

Tecnocom

Finalmente...

- Se acaba de actualizar la red el 19 de abril
  - Matemáticas elementales:
    - 28/10/09 a 19/04/10 = 173 días
- Todo OK, al fin...

Tecnocom

Sin embargo

YO:

Since last thursday we haven't seen any traffic drop in the network and the routes are stable.

What I have seen is a strange thing. The routes that should have update time of days or weeks never stays up longer than some hours. In this example less than five hours (though this is one of the stables routes,

Tecnocom

Mas curioso

Also we see that all the routes of the same king originated in a remote PE have the same lifetime:

```
mad-rpe-01#sh ip route vrf vpn_gni ospf | i 192.168.255.6
O IA    192.168.45.100/30 [110/25414] via 192.168.255.6, 03:31:54
O IA    192.168.45.24/30 [110/31000] via 192.168.255.6, 03:31:54
O IA    192.168.45.228/30 [110/25414] via 192.168.255.6, 03:31:54
O IA    192.168.45.236/30 [110/11000] via 192.168.255.6, 03:31:54
O IA    192.168.45.232/30 [110/2000] via 192.168.255.6, 03:31:54
O IA    192.168.45.240/30 [110/25414] via 192.168.255.6, 03:31:54
O IA    192.168.45.128/30 [110/11000] via 192.168.255.6, 03:31:54
O IA    192.168.45.156/30 [110/2000] via 192.168.255.6, 03:31:54
O IA    77.72.105.22/32 [110/49829] via 192.168.255.6, 03:31:54
O IA    10.143.132.0/24 [110/3001] via 192.168.255.6, 03:31:54
O IA    10.17.12.0/24 [110/12000] via 192.168.255.6, 03:31:54
O IA    10.133.21.128/27 [110/3001] via 192.168.255.6, 03:31:54
```

Tecnocom

Pasapalabra

If we are not seeing any flapping messages for this vrf , then it should be an issue as it does not affect the routing stability.

Thanks & Regards,

Conclusiones

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP
- KISS

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP
- KISS
- Los shows muestran mucho mas de lo que creemos (el debug no sirvió para nada)

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP
- KISS
- Los shows muestran mucho mas de lo que creemos (el debug no sirvió para nada)
- Pero hay que mirarlos bien

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP
- KISS
- Los shows muestran mucho mas de lo que creemos (el debug no sirvió para nada)
- Pero hay que mirarlos bien
- Huye de soluciones exóticas/nuevas

Tecnocom

Conclusiones

- Interconexión de redes. Usa BGP
- KISS
- Los shows muestran mucho mas de lo que creemos (el debug no sirvió para nada)
  - Pero hay que mirarlos bien
  - Huye de soluciones exóticas/nuevas
- No pongas el CPD en el area 0

Tecnocom

¿Preguntas?

Tecnocom