MAP-E Residential Deployment and Issues Encountered

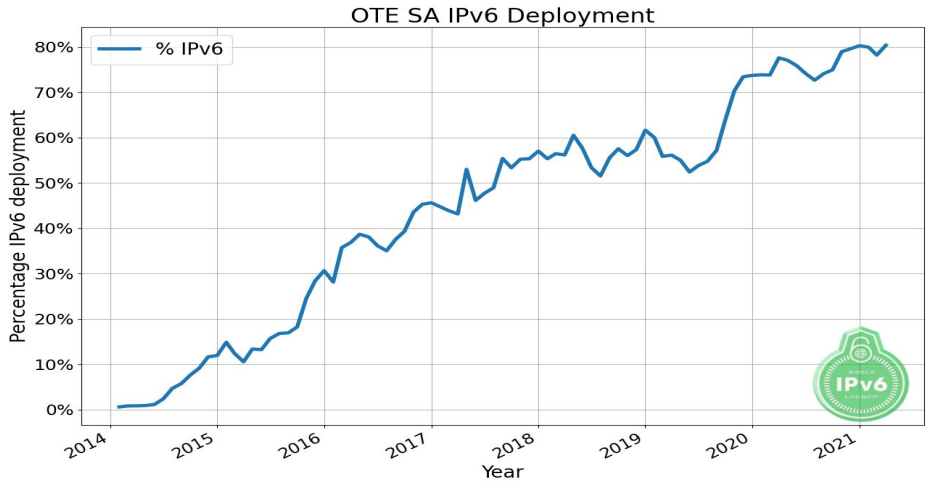
Yannis Nikolopoulos
(yanodd@otenet.gr)
COSMOTE

the rocky road towards IPv4aaS (chapter III: MAP-E)

Yannis Nikolopoulos
(yanodd@otenet.gr)
COSMOTE

to get you up to speed

Rank	Participating Network	ASN(s)	IPv6 deployment
27	OTE SA	6799	80.37%
81	Cosmote Mobile Telecommunications S.A	29247	76.81%



to get you up to speed

- Dual-stack network for several years (IPv6 project running since 2009)
  - core/edge/services
- Been providing IPv6 end-user services since 2013
  - Residential (xDSL)
  - Business (xDSL, fixed line)
  - mobile
- IPv6 in almost 98% of pppoe users (2.1M out of 2.15M)
- Convinced that IPv6-only is the only way
- Skipped dslite/6rd while preparing for next-gen ipv4aaS mechanisms
- Deployed CGN in the meantime, due to IPv4 exhaustion :(
- https://www.ietf.org/proceedings/93/slides/slides-93-v6ops-6.pdf (for more info on our IPv6 deployment)

a timeline for IPv4aaS

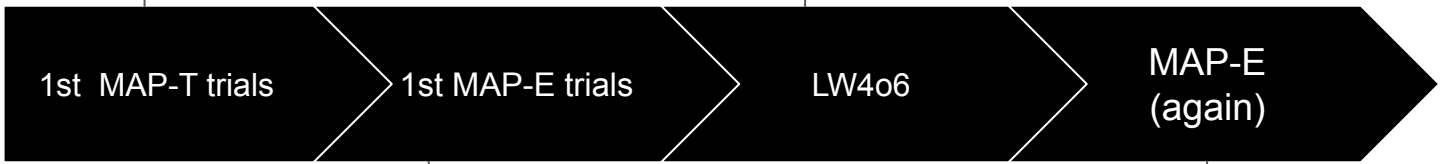
Cisco ASR1k (IOS-XE) w/ OpenWRT

Apr 2015

Juniper lwaftr (SNABB)

https://ripe76.ripe.net/wp-content/uploads/presentations/11-lw4o6-deployment-as6799.pdf

Mar 2017 - Apr 2019 (R.I.P)



Cisco ASR9k (IOS-XR) w/ service module & OpenWRT

May 2015

Cisco ASR99 (IOS-XR) w/o service module

2020

what are you competing against

- At first, NAT444 (single-stack): easy
  - It’s just a temp solution (honestly)
  - It’s expensive to scale
  - Can’t have a future w/o IPv6
- Then, CGN became dual-stack: not so easy
  - Get rid of expensive service modules before their EoL
  - What else?
- Dual-stack everywhere for 15 years, but
  - Reluctance towards change (support, ops, even engineers)
  - 2021 and still “turn-off IPv6” is a thing

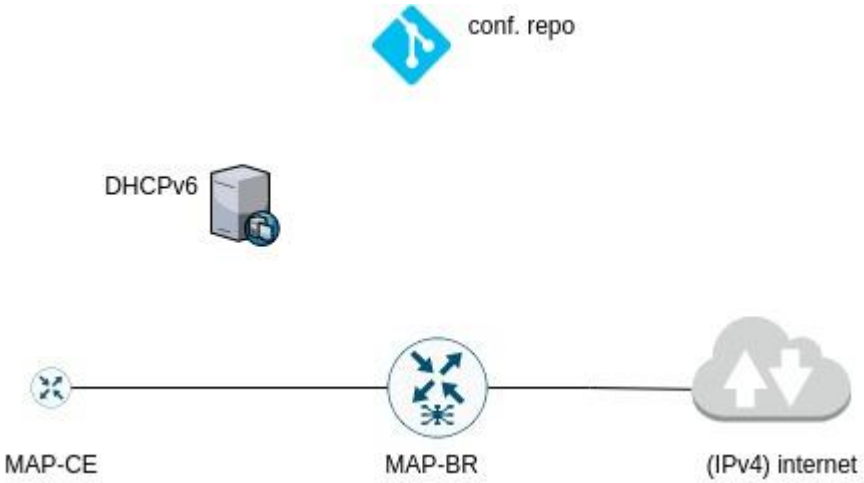
making the correct choice

- The wrong way to go about it
  - MAP-E makes more sense than MAP-T (really? why?)
  - It feels right …
  - Let’s not spend a little bit more (time/money)on the CPE dev/ment
  - Let’s keep it under the radar (MGMT does not care)
- The slightly less wrong way to go about it
  - Research all available techs, in **depth** (a significant amount of time is required)
  - Find the best fit in **your** network
  - Develop a future-proof CPE
  - Engage mgmt, make them part of the journey

MAP-E

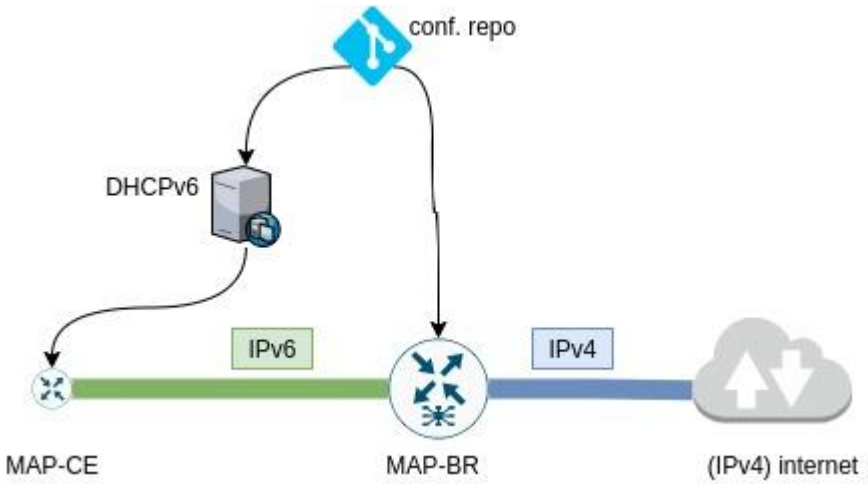
- As per rfc7597
  - mechanism for transporting IPv4 packets across an IPv6 network using IP encapsulation
  - generic mechanism for mapping between IPv6 addresses and IPv4 addresses as well as transport-layer ports
- Border Relay (BR)
  - Receives IPv4 traffic from the internet, encapsulates it into IPv6 and sends it to CE
- CE function
  - nat44+IPv6 encapsulation
- MAP Domain (exist between CE and BR)
- Provisioning mechanism, such as DHCPv6 (rfc7598)

MAP-E



MAP-E

- BR address
- IPv6/IPv4 prefix
- Port Set info

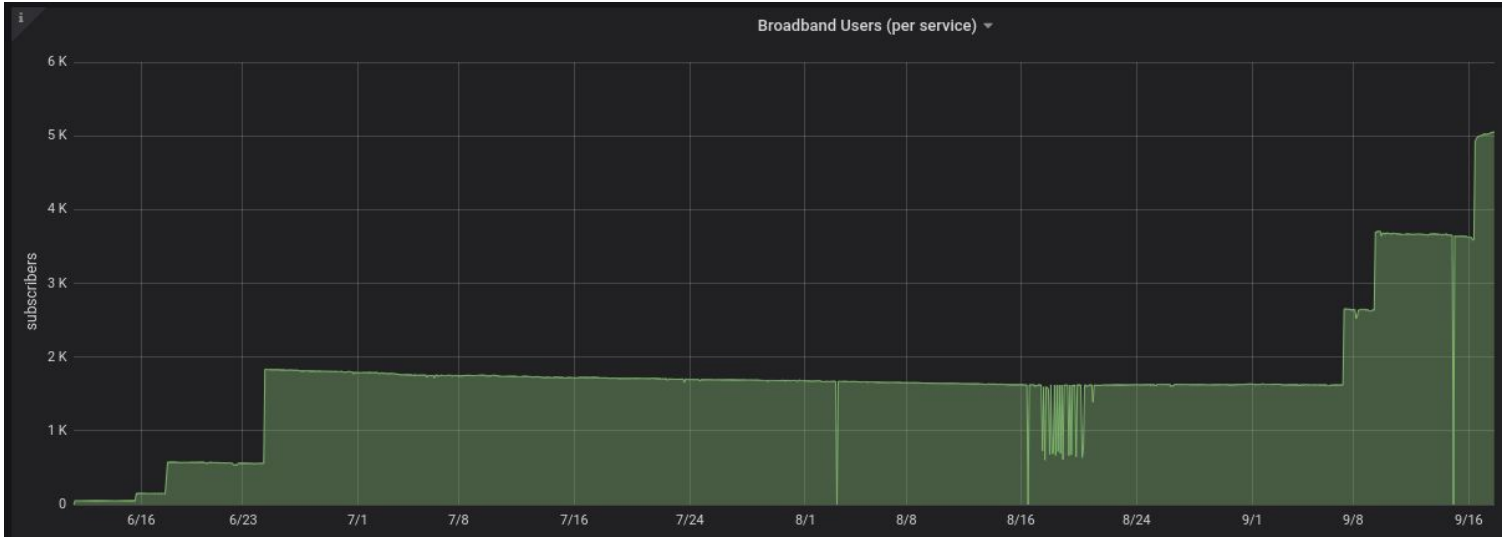


our building blocks

- DHCP based on latest ISC dhdp6 (BNG acts as DHCP proxy)
  - 2 locations, active-standby mode (semi-automatic)
  - option dhcp6.map-option
- 2 MAP-CEs (representing ~90% of total CPEs deployed)
- Automated user selection/provisioning (tr069/LDAP)
- Configuration repo (to feed dhcpd and BR)
- 2 Cisco ASR99, acting as MAP BRs
  - Anycast BR (for now)
  - Dedicated, due to various limitations (MAP function is PBR)
  - MAP domain per BNG (~128), due to nature of addressing

initial Deployment

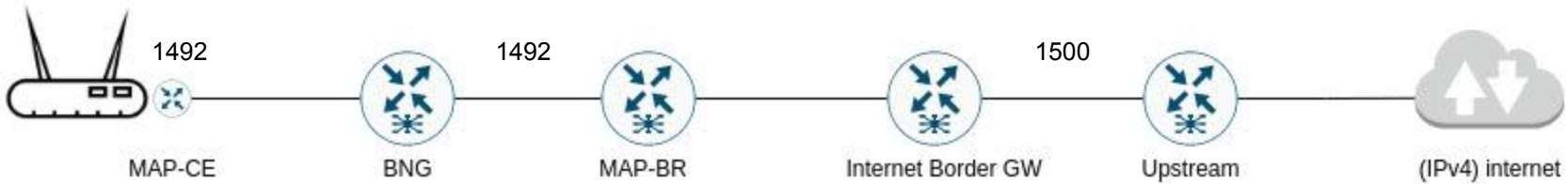
- Successful limited trials (~20 users)
- Rolling out in small batches (~500 users per batch)
- If needed, automated user roll-back to previous profile (CGN most probably)
- Initial % of rolled-back users: ~2% (smaller than CGN's)



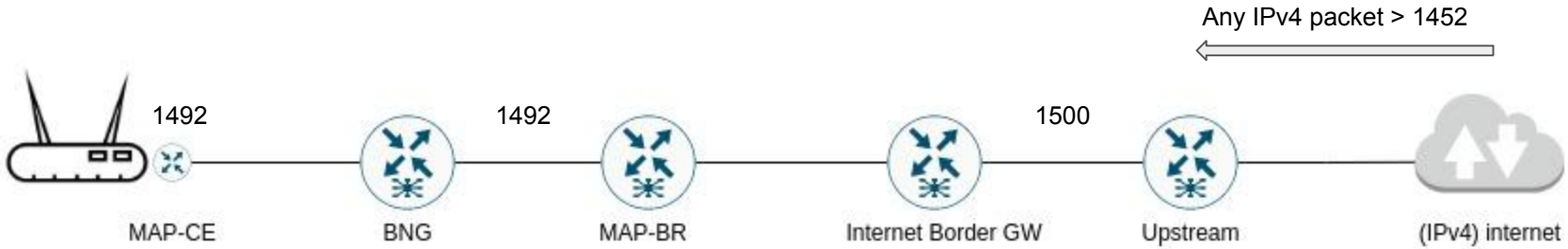
knowing when to stop

- As deployment progressed
  - % of rolled-back users started to increase
  - Customers' complaints became "louder"
- Cases seemed to be unrelated
  - Some were unable to browse specific sites
  - Others were unable to attach files on webmail services
  - Others were unable to use their PoS
  - Oh, and some could not access their VPN services…
  - OK, maybe not so unrelated after all
-

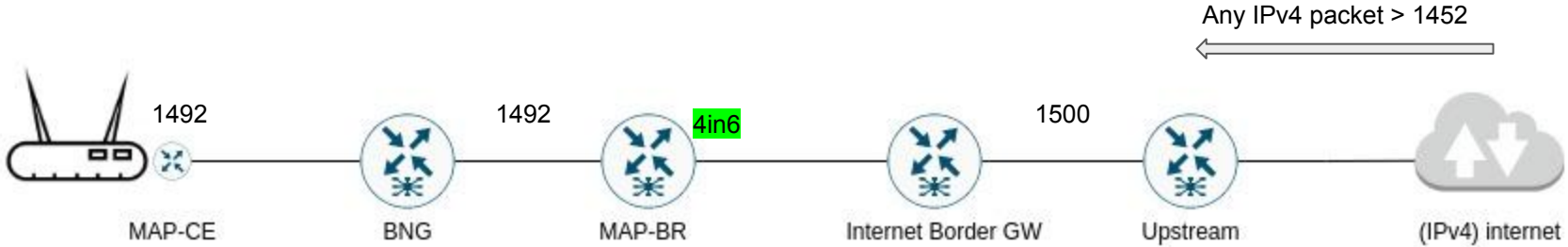
the problem


1492
1492
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

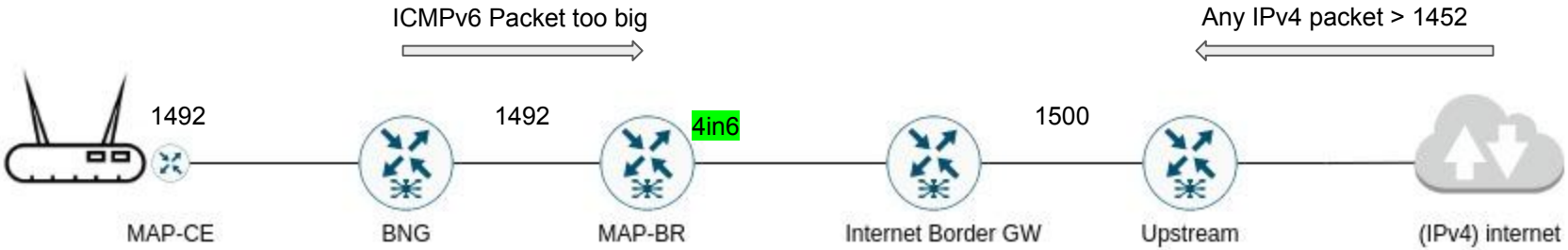
the problem


Any IPv4 packet > 1452
1492
1492
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

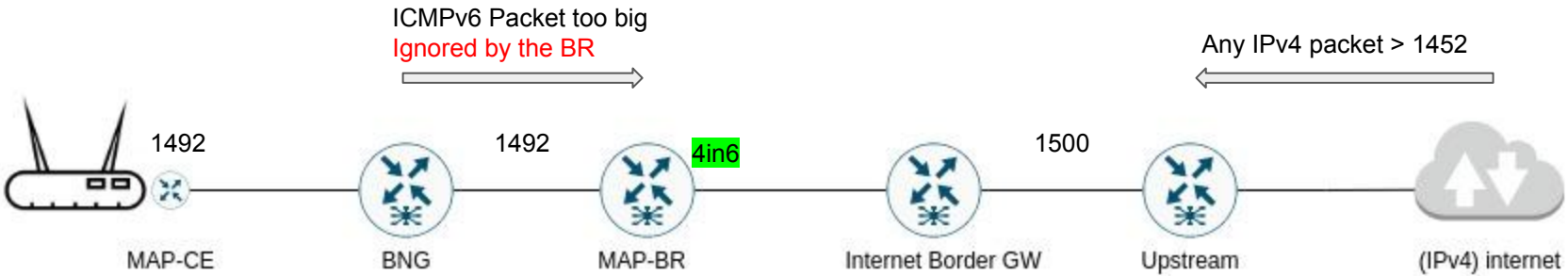
the problem


Any IPv4 packet > 1452
1492
1492
4in6
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

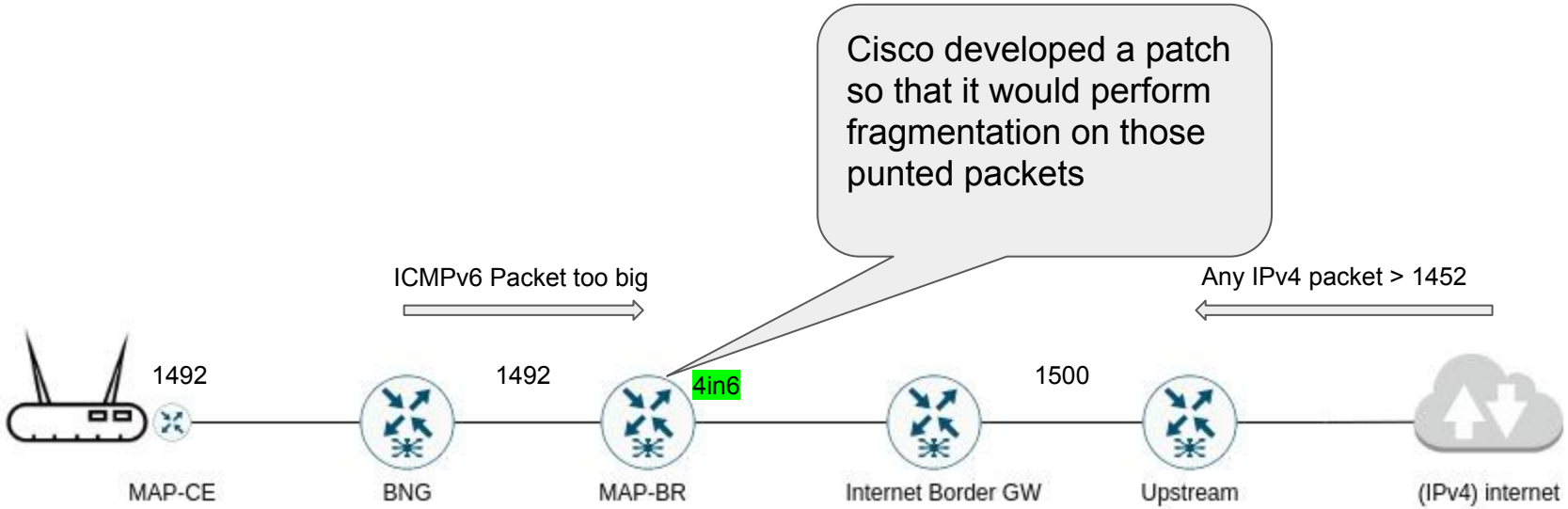
the problem


ICMPv6 Packet too big
Any IPv4 packet > 1452
1492
1492
4in6
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

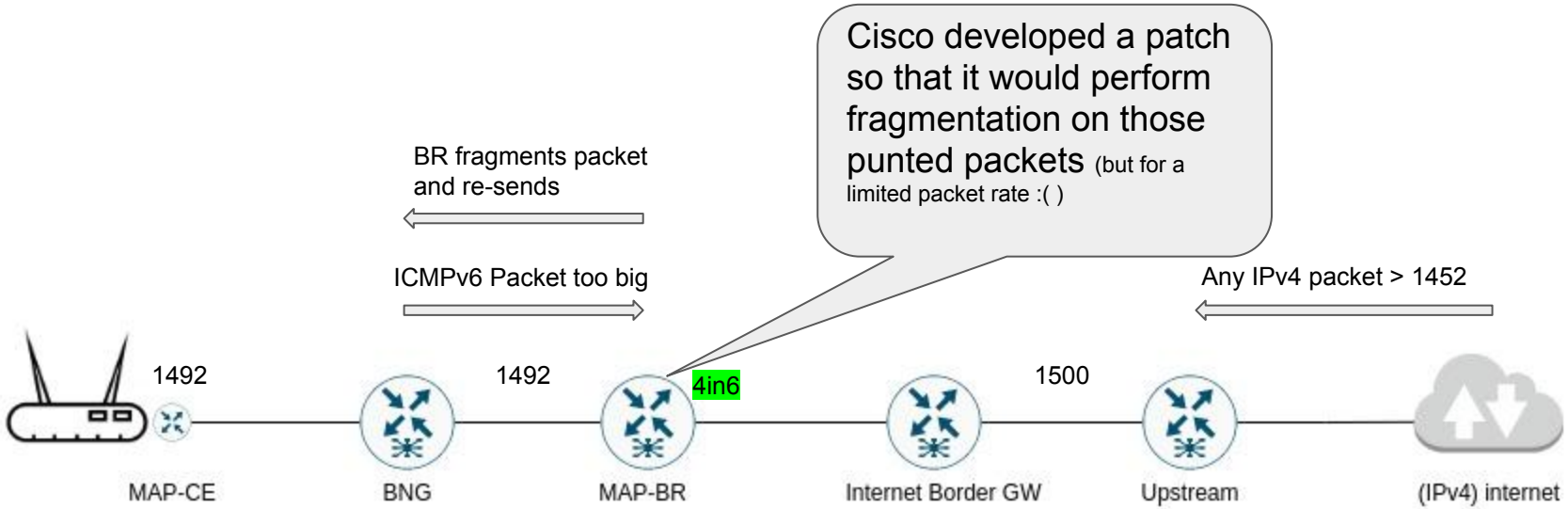
the problem


ICMPv6 Packet too big
Ignored by the BR
Any IPv4 packet > 1452
1492
1492
4in6
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

the problem


Cisco developed a patch so that it would perform fragmentation on those punted packets
ICMPv6 Packet too big
Any IPv4 packet > 1452
1492
1492
4in6
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

the problem

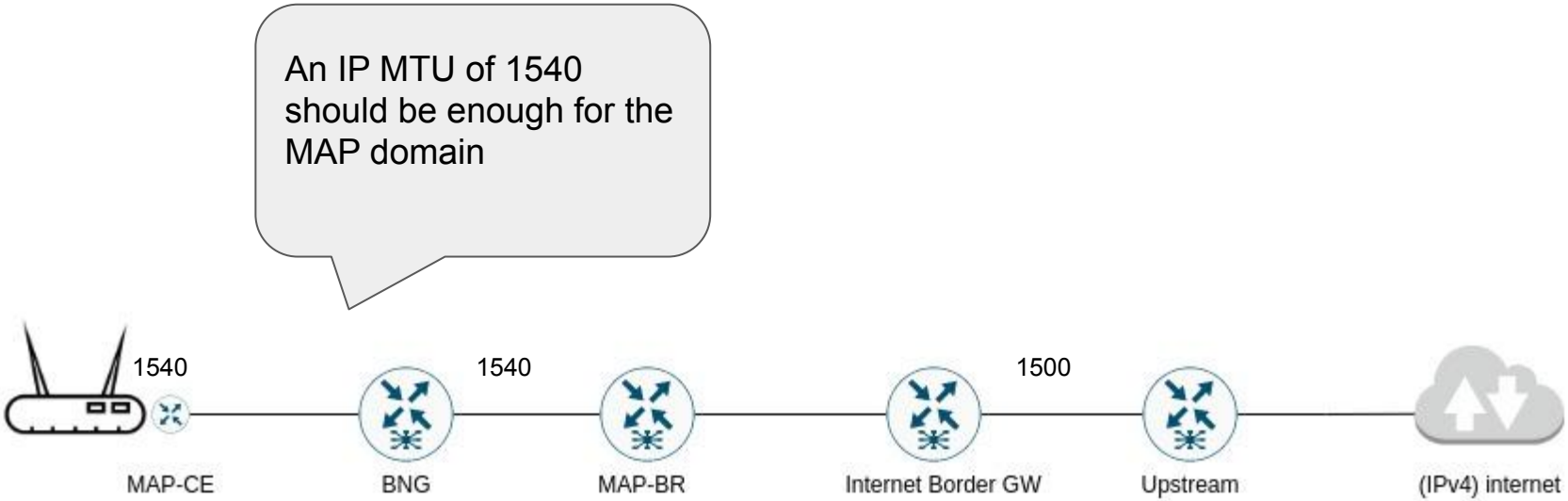

Cisco developed a patch so that it would perform fragmentation on those punted packets (but for a limited packet rate :( )
BR fragments packet and re-sends
ICMPv6 Packet too big
Any IPv4 packet > 1452
1492
1492
4in6
1500
MAP-CE
BNG
MAP-BR
Internet Border GW
Upstream
(IPv4) internet

RTFM

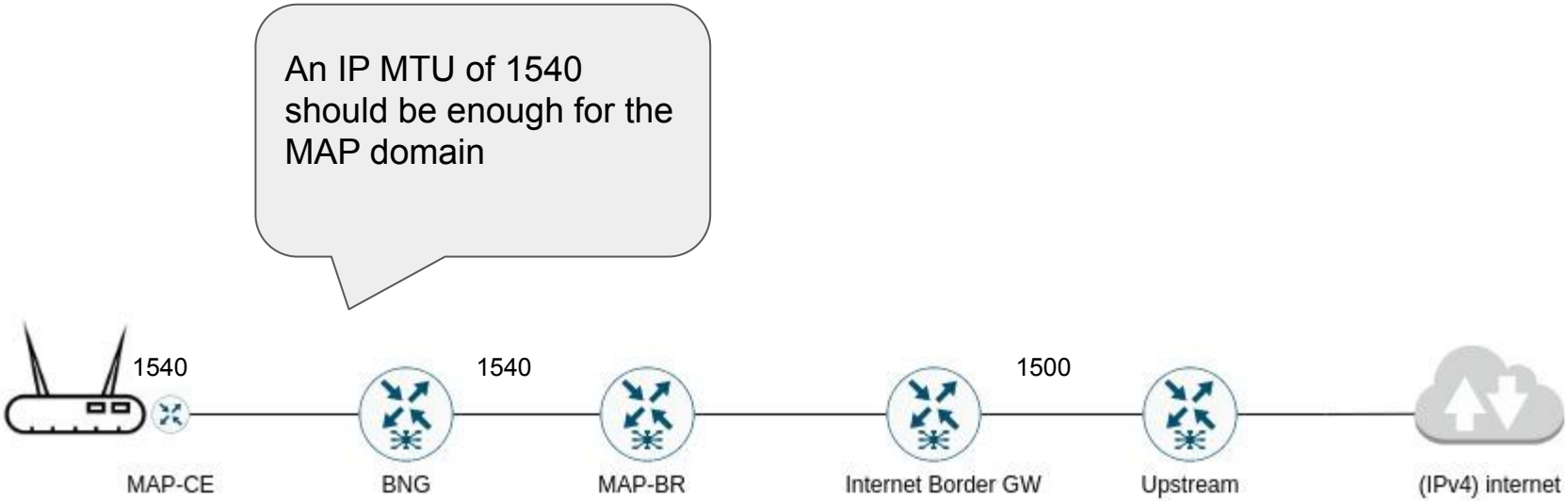
8.3.1. Fragmentation in the MAP Domain

Encapsulating an IPv4 packet to carry it across the MAP domain will increase its size (typically by 40 bytes). It is strongly recommended that the MTU in the MAP domain be well managed and that the IPv6 MTU on the CE WAN-side interface be set so that no fragmentation occurs within the boundary of the MAP domain.

the solution (probably)



the solution (probably)



- All equipment between BNG - CE should be checked (multitude)
- Feature should be developed for the CE
- Investigate possibility of different IP MTUs between services (DS and MAP-E)
- BNG to respect MTU advertised on PPP link
- Each time a user is provisioned to MAP-E, their CPE's MTU needs to be adjusted also (during roll-back as well)

lessons learnt

- When procuring a new CPE, include as many S46 mechanisms as possible (e.g. MAP-E/T and LW4o6)
  - rfc8026 for provisioning multiple mechanisms
  - Slightly bigger cost+time to develop but the flexibility will pay off (e.g.MAP-T with -20 bytes...)
- Assume that the MAP algorithm in the CPE, will be implemented incorrectly
- Extra attention should be paid to the MTU across the MAP domain
- Automate provisioning, automate rollbacks
- When it’s not the DNS, it’s the MTU
-

lessons learnt

- When procuring a new CPE, include as many S46 mechanisms as possible (e.g. MAP-E/T and LW4o6)
  - rfc8026 for provisioning multiple mechanisms
  - Slightly bigger cost+time to develop but the flexibility will pay off (e.g.MAP-T with -20 bytes...)
- Assume that the MAP algorithm in the CPE, will be implement incorrectly
- Extra attention should be paid to the MTU across the MAP domain
- Automate provisioning, automate rollbacks
- ~~When it's not the DNS, it's the MTU~~
- **It's always the MTU**

lessons learnt

- When procuring a new CPE, include as many S46 mechanisms as possible (e.g. MAP-E/T and LW4o6)
  - rfc8026 for provisioning multiple mechanisms
  - Slightly bigger cost+time to develop but the flexibility will pay off (e.g.MAP-T with -20 bytes...)
- Assume that the MAP algorithm in the CPE, will be implement incorrectly
- Extra attention should be paid to the MTU across the MAP domain
- Automate provisioning, automate rollbacks
- ~~when it's not the DNS, it's the MTU~~
- It's always the MTU
- People should be more open with such implementations

a call for collaboration

- A couple of rfcs exist, serving as guidelines or best practices
- No collaboration between operators
- Elusive deployments (MAP-E/T, lw4o6)

a call for collaboration

- A couple of rfcs exist, serving as guidelines or best practices
- No collaboration between operators
- How can we fix this?

~~https://ipv6.ote.gr~~ R.I.P

https://twitter.com/oteipv6

ipv6@otenet.gr