Technical Manual
MTU, PMTU and PMTUD for VXLAN between two Aruba 6300

Author: Antonio Perez (aperez137@gmail.com)

October 4, 2025

Executive summary

This manual presents a methodological approach to configure MTU/PMTU/PMTUD in a static L2 VXLAN deployment between two Aruba 6300 (the only VTEPs), where the IP underlay traverses a Carrier point-to-point (PTP) link. It explains: (1) why underlay links carrying VXLAN must use jumbo MTU, (2) the root cause that led to this decision (broken PMTUD and effective PMTU ∼1450 bytes due to VXLAN overhead), and (3) how to apply the configuration on AOS-CX with comments that clearly distinguish MTU 9128 vs 1500, including the LAG towards the LAN.

Context, objective and scope

Objective: ensure continuity of L4-L7 applications (HTTP/HTTPS, etc.) when transporting L2 domains over VXLAN between two sites. Scope: only two Aruba 6300 as VTEPs; underlay via a PTP Carrier link. Core criterion: homogeneous MTU across the underlay connecting VTEP↔VTEP.

Nomenclature

Term	Description
MTU	Maximum frame/packet size per interface (Ethernet default: 1500 B).
PMTU	Path MTU: effective end-to-end MTU along the path.
PMTUD	Path MTU Discovery: discovers PMTU via ICMP Type 3 Code 4 (Fragmentation Needed).
MSS	TCP Maximum Segment Size (payload). In IPv4, MSS $\approx$ PMTU $-$ 40.
VXLAN	L2 over IP/UDP (port 4789); effective overhead $\sim$50 B (IP+UDP+VXLAN+outer Ethernet).
VTEP	VXLAN Tunnel End-Point (per-site loopback).
VNI	VXLAN Network Identifier; maps a VLAN in the overlay.
Underlay	IP network transporting encapsulated VXLAN (uplinks, Carrier PTP).
Overlay	L2 domain traveling inside VXLAN (remote VLANs).

Root cause of the incident

Observed: ICMP/DNS worked but HTTP/HTTPS failed. DF ping tests showed: - size 1416 OK, 1424 FAIL, pinning effective PMTU near 1450 B. Interpretation: with physical links at 1500 B and VXLAN enabled, $\sim$50 B extra overhead reduces PMTU to $\sim$1450 B. If the path also blocks ICMP Fragmentation Needed, PMTUD cannot adjust MSS and TCP blackholing occurs (SYN with DF and MSS 1460 does not progress). Hence, web traffic failed while ping/DNS seemed normal.

Why jumbo MTU on ports carrying VXLAN

1) VXLAN adds overhead ($\sim$50 B). To carry a 1500 B inner frame without fragmentation, the underlay should support 1550 B. 2) If any underlay hop is 1500 B, effective PMTU drops to $\sim$1450 B. HTTP/HTTPS breaks when PMTUD is blocked/degraded. 3) Homogeneous jumbo MTU (e.g., 9128 B) across the underlay gives enough headroom to carry 1500 B inner frames without relying on PMTUD/MSS clamping. 4) If jumbo end-to-end is not possible with the Carrier, enable TCP MSS clamping (e.g., 1410 B) or allow ICMP Fragmentation Needed along the path.

Design criteria: where MTU 9128 vs 1500

- MTU 9128 (jumbo): only on underlay links that *carry VXLAN*. In this scenario, the Carrier PTP connecting both 6300. If the underlay is L3 (P2P/SVI), align `ip mtu` as well. - MTU 1500: any LAG towards the LAN/campus and user access ports. These links do *not* encapsulate VXLAN and should remain at 1500 for interoperability.

Example addressing (documentation-safe)

Using documentation blocks (RFC 5737) and an illustrative /30:

- VTEP loopbacks: Site A → 198.51.100.5/32 ; Site B → 198.51.100.6/32
- Underlay PTP (Carrier): 10.250.10.0/30 - Site A: 10.250.10.1/30 - Site B: 10.250.10.2/30
- VLAN/VNI (overlay): VLAN 718 ↔ VNI 10718 ; VLAN 708 ↔ VNI 10708

AOS-CX configuration with in-line highlights of key settings

Legend: highlighted entries (yellow) are the key configurations.

Aruba 6300 - Site A (VTEP = 198.51.100.5)

```
! ---- LAG towards LAN (does NOT carry VXLAN) -> keep MTU 1500 (default)
interface lag 15
    description LAG15-to-LAN
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 1,708,718   ! L2 user VLANs; no VXLAN here (MTU
       1500)

interface 1/1/21
    description Member LAG15
    no shutdown
    lag 15

interface 1/1/22
    description Member LAG15
```

```
    no shutdown
    lag 15

! ---- UNDERLAY PTP towards Carrier (THIS carries VXLAN) -> jumbo MTU
interface 1/1/48
    description PTP-Carrier A-B (UNDERLAY VXLAN)
    no shutdown
    mtu 9128                ! L2 jumbo required by VXLAN
    routing
    ip address 10.250.10.1/30
    ip mtu 9128             ! L3 jumbo aligned with L2
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point

! ---- VTEP loopback
interface loopback 1
    ip address 198.51.100.5/32

! ---- Static VXLAN (VLAN<->VNI mapping + remote VTEP)
interface vxlan 1
    vxlan source-interface loopback 1
    vxlan udp-port 4789
    no shutdown
    vxlan vlan 718 vni 10718
    vxlan vlan 708 vni 10708
    vxlan vtep 198.51.100.6   ! remote VTEP (Site B)

! ---- Deterministic /32 route to remote VTEP over PTP underlay
ip route 198.51.100.6/32 10.250.10.2

router ospf 1
    router-id 198.51.100.5
    area 0.0.0.0
    passive-interface loopback 1
```

Aruba 6300 - Site B (VTEP = 198.51.100.6)

```
! ---- LAG towards LAN (does NOT carry VXLAN) -> keep MTU 1500 (default)
interface lag 15
    description LAG15-to-LAN
    no shutdown
    no routing
    vlan trunk native 1
```

```
    vlan trunk allowed 1,708,718

interface 1/1/31
    description Member LAG15
    no shutdown
    lag 15

interface 1/1/32
    description Member LAG15
    no shutdown
    lag 15

! ---- UNDERLAY PTP towards Carrier (THIS carries VXLAN) -> jumbo MTU
interface 1/1/48
    description PTP-Carrier B-A (UNDERLAY VXLAN)
    no shutdown
     mtu 9128
    routing
    ip address 10.250.10.2/30
     ip mtu 9128
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point

! ---- VTEP loopback
interface loopback 1
     ip address 198.51.100.6/32

! ---- Static VXLAN
interface vxlan 1
     vxlan source-interface loopback 1
    vxlan udp-port 4789
    no shutdown
     vxlan vlan 718 vni 10718
     vxlan vlan 708 vni 10708
     vxlan vtep 198.51.100.5

! ---- Deterministic /32 route to remote VTEP over PTP underlay
ip route 198.51.100.5/32 10.250.10.1

router ospf 1
    router-id 198.51.100.6
    area 0.0.0.0
    passive-interface loopback 1
```

Validation and tests

On the 6300

```
show interface 1/1/48 | include MTU
ping 198.51.100.6 source 198.51.100.5
ping 198.51.100.5 source 198.51.100.6
show vxlan interface vxlan 1
show vxlan statistics
show mac-address-table vlan 718
show mac-address-table vlan 708
```

From a Linux host

```
ping -M do -s 1472 8.8.8.8
tracepath -n 8.8.8.8
```

Troubleshooting matrix (summary)

Symptom	Probable cause	Action
ICMP/DNS OK; HTTP/HTTPS fails	Broken PMTUD; PMTU ∼1450 B due to VXLAN over 1500 B link	Make underlay jumbo or apply TCP MSS clamping (∼1410 B) or allow ICMP Fragmentation Needed.
Unidirectional traffic	Missing/asymmetric /32 routes to VTEP	Create deterministic /32 via the PTP; verify show ip route <VTEP>.
No remote MAC in VLAN	VLAN missing or mistagged	Check show vlan <id>, mac-address-table; fix tagging.

Conclusions

The root cause of browsing issues was an effective PMTU reduced to ∼1450 B by VXLAN overhead, compounded by blocked PMTUD along the path. To eliminate TCP blackholing, underlay links carrying VXLAN must operate with uniform jumbo MTU (e.g., 9128), while

the LAG-to-LAN and access ports remain at 1500. When end-to-end jumbo is not feasible, TCP MSS clamping or enabling ICMP Fragmentation Needed mitigates the risk.

Manual academico
MTU, PMTU y PMTUD en VXLAN entre dos Aruba 6300

Autor: Antonio Perez (aperez137@gmail.com)

4 de octubre de 2025

Resumen ejecutivo

Este manual presenta un enfoque metodologico para configurar MTU/PMTU/PMTUD en un despliegue VXLAN L2 estatico entre dos Aruba 6300 (unicos VTEP), cuyo underlay IP cruza un enlace PTP de un Operador. Se explica: (1) por que los enlaces del underlay que transportan VXLAN deben operar con MTU jumbo, (2) la causa raiz que llevo a esta decision (PMTUD roto y PMTU efectiva ∼1450 B por overhead de VXLAN), y (3) como aplicar la configuracion en AOS-CX con comentarios que distinguen claramente MTU 9128 vs 1500, incluyendo el LAG hacia la LAN.

Contexto, objetivo y alcance

Objetivo: garantizar continuidad de aplicaciones L4-L7 (HTTP/HTTPS, etc.) al transportar dominios L2 sobre VXLAN entre dos sitios. Alcance: solo dos Aruba 6300 como VTEP; underlay por un enlace PTP. Criterio central: uniformidad de MTU en el underlay que conecta VTEP↔VTEP.

Nomenclatura

Termino	Descripcion
MTU	Tamano maximo de trama/paquete por interfaz (Ethernet tipico: 1500 B).
PMTU	Path MTU: MTU efectiva del camino extremo a extremo.
PMTUD	Path MTU Discovery: descubre la PMTU via ICMP Type 3 Code 4 (Fragmentation Needed).
MSS	Maximum Segment Size de TCP (payload). En IPv4, MSS $\approx$ PMTU $-$ 40.
VXLAN	L2 sobre IP/UDP (puerto 4789); overhead efectivo $\sim$50 B (IP+UDP+VXLAN+Ethernet exterior).
VTEP	VXLAN Tunnel End-Point (loopback por sitio).
VNI	VXLAN Network Identifier; mapea VLAN del overlay.
Underlay	Red IP que transporta el VXLAN encapsulado (uplinks, PTP del Operador).
Overlay	Dominio L2 que viaja dentro de VXLAN (VLAN remotas).

Causa raiz del incidente

En el caso observado, ICMP/DNS funcionaban pero HTTP/HTTPS fallaba. Las pruebas con ping con DF mostraron: - tamano 1416 OK, 1424 FAIL, lo que fija una PMTU efectiva cercana a 1450 B. Interpretacion: con enlaces fisicos en 1500 B y VXLAN activo, el overhead de $\sim$50 B reduce la PMTU a $\sim$1450 B. Si ademas el camino bloquea ICMP Fragmentation Needed, PMTUD no puede ajustar MSS y aparecen blackholes selectivos en TCP (SYN con DF y MSS 1460 no progresa). Por ello, el trafico web fallaba a pesar de que ping y resolucion DNS parecian normales.

Por que MTU jumbo en puertos que transportan VXLAN

1) VXLAN agrega overhead adicional ($\sim$50 B). Para transportar una trama interna de 1500 B sin fragmentar, el underlay deberia soportar al menos 1550 B. 2) Si algun salto underlay queda en 1500 B, la PMTU se reduce a $\sim$1450 B. HTTP/HTTPS rompe si PMTUD esta bloqueado o degradado. 3) Configurar MTU jumbo homogenea en el underlay (por ejemplo 9128 B) crea margen suficiente para encapsular tramas internas de 1500 B sin depender de

PMTUD ni de MSS clamping. 4) Cuando no sea posible jumbo end-to-end en el Operador, se debe habilitar TCP MSS clamping coherente (por ejemplo 1410 B) o permitir ICMP Fragmentation Needed a lo largo del camino.

Criterios de diseno: donde va MTU 9128 y donde 1500

- MTU 9128 (jumbo): solo en enlaces del underlay que *transportan VXLAN*. En este escenario, el enlace PTP que conecta ambos 6300. Si el underlay es L3 (P2P o SVI), alinear tambien `ip mtu`. - MTU 1500: cualquier LAG hacia la LAN/campus y puertos de acceso a usuarios. Estos enlaces no encapsulan VXLAN y deben permanecer en 1500 por interoperabilidad.

Direccionamiento de ejemplo (documentacion, seguro)

Se usan bloques de documentacion (RFC 5737) y un /30 ilustrativo.

- VTEP loopbacks: Sitio A → 198.51.100.5/32 ; Sitio B → 198.51.100.6/32
- Underlay PTP Operador: 10.250.10.0/30 - Sitio A: 10.250.10.1/30 - Sitio B: 10.250.10.2/30
- VLAN/VNI (overlay): VLAN 718 ↔ VNI 10718 ; VLAN 708 ↔ VNI 10708

Configuracion AOS-CX con resaltado en linea de claves

Leyenda: las entradas resaltadas (amarillo) son las configuraciones clave.

Aruba 6300 - Sitio A (VTEP = 198.51.100.5)

```
! ---- LAG hacia LAN (NO transporta VXLAN) -> mantener MTU 1500 (default)
interface lag 15
    description LAG15-to-LAN
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 1,708,718   ! VLANs de usuario; sin VXLAN (MTU
       1500)
```

```
interface 1/1/21
    description Member LAG15
    no shutdown
    lag 15

interface 1/1/22
    description Member LAG15
    no shutdown
    lag 15

! ---- UNDERLAY PTP hacia Operador (SI transporta VXLAN) -> jumbo
interface 1/1/48
    description PTP-Operador A-B (UNDERLAY VXLAN)
    no shutdown
     mtu 9128                 ! L2 jumbo requerido por VXLAN
    routing
    ip address 10.250.10.1/30
     ip mtu 9128              ! L3 jumbo alineado con L2
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point

! ---- Loopback del VTEP
interface loopback 1
     ip address 198.51.100.5/32

! ---- VXLAN estatico (mapeo VLAN<->VNI + VTEP remoto)
interface vxlan 1
     vxlan source-interface loopback 1
    vxlan udp-port 4789
    no shutdown
     vxlan vlan 718 vni 10718
     vxlan vlan 708 vni 10708
     vxlan vtep 198.51.100.6   ! VTEP remoto (Sitio B)

! ---- Ruta /32 determinista al VTEP remoto por el PTP
ip route 198.51.100.6/32 10.250.10.2

router ospf 1
    router-id 198.51.100.5
    area 0.0.0.0
    passive-interface loopback 1
```

Aruba 6300 - Sitio B (VTEP = 198.51.100.6)

```
! ---- LAG hacia LAN (NO es underlay VXLAN) -> MTU 1500 (default)
interface lag 15
    description LAG15-to-LAN
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 1,708,718

interface 1/1/31
    description Member LAG15
    no shutdown
    lag 15

interface 1/1/32
    description Member LAG15
    no shutdown
    lag 15

! ---- UNDERLAY PTP hacia Operador (transporta VXLAN) -> jumbo
interface 1/1/48
    description PTP-Operador B-A (UNDERLAY VXLAN)
    no shutdown
     mtu 9128
    routing
    ip address 10.250.10.2/30
     ip mtu 9128
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point

! ---- Loopback del VTEP
interface loopback 1
     ip address 198.51.100.6/32

! ---- VXLAN estatico
interface vxlan 1
     vxlan source-interface loopback 1
    vxlan udp-port 4789
    no shutdown
     vxlan vlan 718 vni 10718
     vxlan vlan 708 vni 10708
     vxlan vtep 198.51.100.5
```

```
! ---- Ruta /32 determinista hacia el VTEP remoto por el PTP
ip route 198.51.100.5/32 10.250.10.1

router ospf 1
    router-id 198.51.100.6
    area 0.0.0.0
    passive-interface loopback 1
```

Validacion y pruebas

En los 6300

```
show interface 1/1/48 | include MTU
ping 198.51.100.6 source 198.51.100.5
ping 198.51.100.5 source 198.51.100.6
show vxlan interface vxlan 1
show vxlan statistics
show mac-address-table vlan 718
show mac-address-table vlan 708
```

Desde un host Linux

```
ping -M do -s 1472 8.8.8.8
tracepath -n 8.8.8.8
```

Matriz de troubleshooting (resumen)

Sintoma	**Causa probable**	**Accion**
ICMP/DNS OK; HTTP/HTTPS falla	PMTUD roto; PMTU ∼1450 B por VXLAN y enlace en 1500 B	Hacer jumbo en underlay o aplicar TCP MSS clamping (∼1410 B) o permitir ICMP Fragmentation Needed.
Trafico unidireccional	Rutas /32 a VTEP ausentes o asimetricas	Crear /32 deterministas por el PTP; verificar show ip route <VTEP>.
No hay MAC remota en VLAN	VLAN no presente o tagging incorrecto	Revisar show vlan <id>, mac-address-table; corregir tagging.

Conclusiones

La causa raiz del problema de navegacion fue una PMTU efectiva reducida (∼1450 B) por el overhead VXLAN, agravada por PMTUD bloqueado en el camino. Para eliminar el blackholing TCP, los enlaces del underlay que transportan VXLAN deben operar con MTU jumbo uniforme (p. ej., 9128), mientras que el LAG hacia la LAN y los accesos se mantienen en 1500. Cuando no sea viable jumbo end-to-end, TCP MSS clamping o la habilitacion de ICMP Fragmentation Needed mitigan el riesgo.