Happy-Happy L2: Bridges’ Insecurity

Networking

Pierre-Philipp Braun <pbraun@nethence.com>

Table of contents

- Network Segments (recap)
- Linux Bonding (recap)
- Link Aggregation++
- Types of Bridges
- L2 Products
- Switch Security
- L2 Hardening
- VLAN++ & VLAN Hopping

Network Segments

dmz, vlan, stp

What's the difference between perimeter and DMZ?...

==> front-facing vs NAT

network topology

Perimeter (white IP)

- default route –> your ISP’s
- (still protected somehow)
- (this is where you NAT gw lives)
- (–and– your IP6 RA daemon)

DMZ (behind gw / firewall)

- Routed + Firewall
- –or– DNAT & SNAT routed
- –or– DNAT & isolated

What’s a VLAN and how does it work?…

==> a tag that is seen sometimes un-seen

- IEEE 802.1Q – Dot1q / VLAN on Ethernet
- trunk – multiple tags for the uplink
- access – tag is hidden to the hosts

BONUS QUESTION // trunk with only 1 vlan – what happens?

Terminology

Cisco

`trunk vs. access mode`

HPE

`tagged vs. untagged`

Let’s split our switch!

So what would be a physical vlan?…

What does it correspond to?…

==> physical vlan as with

```
cisco -- access mode
hpe -- untagged
```

Did you hear of spanning tree before? Any idea what it is?…

```
+----------+     +-----------+     +-----------+     +----------+
|          |     |           |     |           |     |          |
|   PC-A   +---->+  SW1      +---->+    SW2    +---->+   PC-B   |
|          |     |           |     |           |     |          |
+----------+     +-----------+     +-----------+     +----------+
```

(FR) jusque là tout va bien…

```
+-----------+      +-------------+      +-------------+      +-----------+
|           |      |             |      |             |      |           |
|   PC-A    +----->+     SW1     +----->+     SW2     +----->+   PC-B    |
|           |      |             |      |             |      |           |
+-----------+      +--+----------+      +-----------+-+      +-----------+
                      |                               ^
                      |       +-------------+         |
                      |       |             |         |
                      +------>+     SW3     +---------+
                              |             |
                              +-------------+
```

(FR) plusieurs chemins…

```
+-----------+    +------------+     +------------+    +-----------+
|          |     |            |      |           |    |           |
|   PC-A   +---->+    SW1    +-----^+     SW2    +--->+   PC-B    |
|          |     |           +v-----+            |    |           |
+-----------+    +--+-+-------+     +-----------++    +-----------+
                        | |                        ^
                        | |    +------------+      |
                        | +---^+            |...   |
                        +-----v+    SW3     +------+
                               |            |
                               +------------+
```

(FR) ça tourne en rond…

Spanning Tree Protocol (STP)

- Avoid christmas tree (broadcast storm)
- Plug a wire – delay up to 30 seconds

LAB // PoC & sniff STP on Linux bridge vs OpenvSwitch

LAB // Evaluate the 30 seconds delay caused by STP and try to remediate

Network emulation

- Packet Tracer – Windows only
- GNS3
- EVE-NG Pro
- VirtualBox – Host network manager
- DIY – Linux Bridge
- DIY – OpenvSwitch

// Questions on network segments?

Linux Bonding

```
|     .-.
|    /   \         .-.
|   /     \       /   \       .-.     .-.     _   _
+--/-------\-----/-----\-----/---\---/---\---/-\-/-\/\/---
| /         \   /       \   /     '-'     '-'
|/           '-'         '-'
```

Linux Bonding modes

0	balance-rr	lbs & ha
1	active-backup	active/passive
2	balance-xor	lbs/xmit & ha
3	broadcast	ha
4	802.3ad	lbs & ha
5	Balance-tlb	lbs & ?
6	balance-alb	lbs & ?

LAB // how come round-robin and XOR provide HA here?

Managed vs. un-managed switch

Static port trunk

`balance-rr`

`balance-xor`

Dynamic port trunk

`802.3ad`

Un-managed switch is fine for those

`balance-tlb`

`balance-alb (also RX)`

Linux Bonding - the deprecated way

```
#vi /etc/modprobe.conf
vi /etc/modprobe.d/bonding.conf

alias bond0 bonding
options bond0 miimon=100 mode=X <other option=...>

ifenslave bond0 eth0
ifenslave bond0 eth1
```

check

```
ifenslave –a
```

Linux Bonding - the new way

```
modprobe bonding
echo 100 > /sys/class/net/bond0/bonding/miimon
echo 200 > /sys/class/net/bond0/bonding/downdelay
echo 200 > /sys/class/net/bond0/bonding/updelay
echo X > /sys/class/net/bond0/bonding/mode
echo ... > /sys/class/net/bond0/bonding/other_option
#echo layer3+4 > /sys/class/net/bond0/bonding/xmit_hash_policy
echo +eth0 > /sys/class/net/bond0/bonding/slaves
echo +eth1 > /sys/class/net/bond0/bonding/slaves
```

Status

```
cat /sys/class/net/bonding_masters
cat /proc/net/bonding/bond0
cat /sys/class/net/bond0/bonding/miimon
cat /sys/class/net/bond0/bonding/downdelay
cat /sys/class/net/bond0/bonding/updelay
cat /sys/class/net/bond0/bonding/mode
cat /sys/class/net/bond0/bonding/other_option
cat /sys/class/net/bond0/bonding/xmit_hash_policy
```

Acceptance testing

How to validate

- unplug / replug…
- iPerf3 (does upload/download)
- UDP vs TCP

What about max bandwidth

- multiple iPerf3 instances…

Linux Teaming

- != VMware NIC Teaming
- alternative to Bonding
- user-space daemon

LAB // try-out and validate Linux Teaming

LAB // benchmark Linux Teaming vs. Bonding

FDX vs. HDX

- Full-duplex – dedicated cable for TX/RX
- Half-duplex – only one cable

LAB // search and dig into Half-duplex driver modes

FDX validation

Through a 100Mbit/s poor switch

```
94.1 Mbit/s if only one direction
91.5 Mbit/s with both direction at the same time
```

With direct 1Gbit/s link

```
940 Mbit/s if only one direction
930 Mbit/s with both directions at the same time
```

// Questions on linux bonding?

Link Aggregation++

4x 2.5Gbe cheaper than 10Gbit?…

2x 5Gbe cheaper than 10Gbit?…

==> YES multi-gigabit port trunks are cheaper than 10GbE

- switches are cheaper
- cables are cheaper (CAT5E vs. CAT6)
- no GBIC required

Let’s just load-balance the load!

- disadvantage: cap per connection
- ideal for multiple connections’ load distribution

How does link aggregation’s load distribution work?…

==> load-balance algorithm against **OUTBOUND** traffic

Terminology

Cisco

`EtherChannel`

non-Cisco

`Port trunk`

Port trunking algorithms

- Static trunk round-robin
- Static trunk XOR & xmit
- Dynamic trunk (LACP) & xmit

LAB // switch manufacturers do round-robin or XOR?

xmit

this goes for both XOR and LACP modes

We've got a choice here on ways to outbound balance

```
dst-mac
src-mac
src-dst-mac

dst-ip
src-ip
src-dst-ip
```

Huawei switches default to `src-dst-mac` or `src-dst-ip` depending on model

HPE switches we can also xmit L4/port (non-compliant with LACP)

RR issues

- TCP packets out of order
- *logically the same as for UDP* // LAB
- XOR & LACP's xmit solves the problem

LAB // benchmark RR vs. XOR vs. LACP performance

LACP advantages

1. HA / fail-over
2. negociated between two switches
3. multi-vendor

LACP requirements

- links with same negociated speed
- only FDX (no HDX)
- max 8 ports

LAB // what about HDX for static trunks?

LACP restrictions

- conflicts with 802.1X port-access
- conflicts with port-security

The LACP pain

PXE doesn’t work anymore

Any idea why?…

==> switch delivers LACP-encapsulated frames and the NIC firmware doesn’t know about it

LAB // PoC that PXE dies vs. survives through an auto or active resp. passive LACP

LACP static vs. dynamic

Between switches

- `Auto` is recommended on both sides (default setting)

Between a switch and a host system

Pierre's trick (draft)

- LACP passive on the switch
- LACP active on the host

// Questions on link aggregation++?

Types of Bridges

How fast can an RJ45 copper be?…

==> Port bandwidth

CAT5E

- FastEthernet – `fe0/X`
- GigabitEthernet – `gi0/X`

CAT5E / CAT6

- Multi-gigabit – 2.5 & 5 GbE and more (up to 10 GbE says Cisco)

CAT6A

- 10 GbE (shows up as TE for Ten Gigabyte Ethernet)

CAT7

- 30-35-40 GbE
- 100 GbE up to 15m

What about long-distance media?...

==> CAT6A price ~ SFP+

```
 1G (SFP)
10G (SFP+)
25G (SFP28)
40G (QSFP+)
```

By the way, what is a switch?...

==> simple, stupid **repeater**

==> with many wires inside (**fabric** design)

Features we *need* in a switch

- Multi-gigabit
- VLAN
- Port trunking
- CoS / QoS

Features we possibly *want*…

- DHCP snooping
- Port security / MAC filtering
- (ACL)
- stackable – operate multiples RUs as a single switch

Note Cisco's *Stackwise* also does redundancy

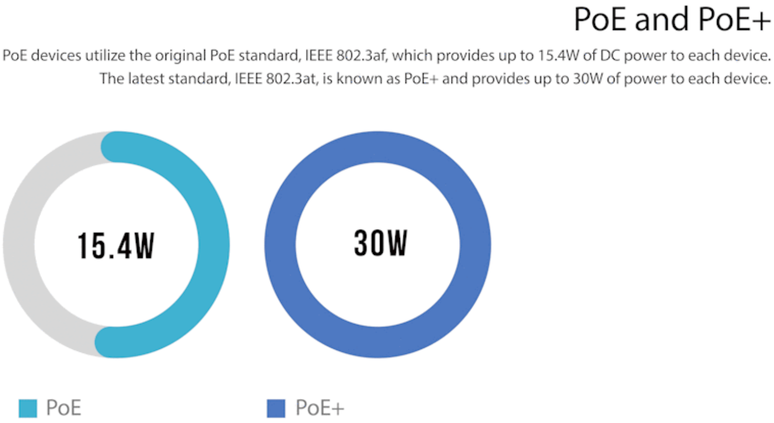
// Questions on types of bridges?

L2 Products

What major kinds of switches there are?…

==> Switch types

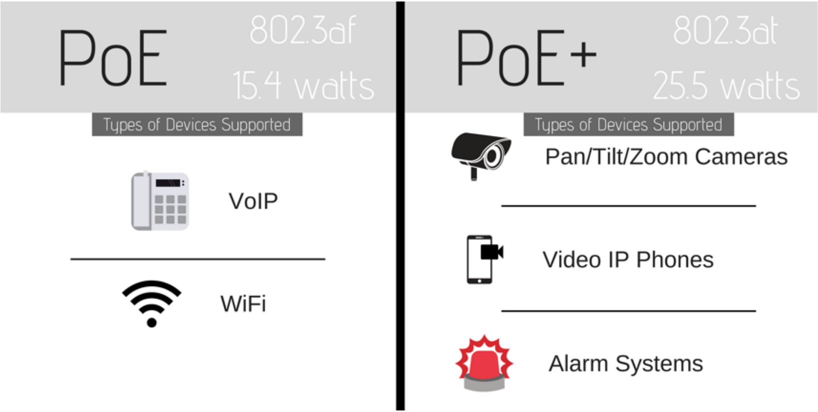
- Unmanaged vs. “Smart” vs. Managed
- Modular vs. fixed-configuration
- Stackable vs. standalone switches
- Fabric architecture & max bandwidth
- PoE, PoE+ and possibly more (Cisco)



PoE and PoE+

PoE devices utilize the original PoE standard, IEEE 802.3af, which provides up to 15.4W of DC power to each device. The latest standard, IEEE 802.3at, is known as PoE+ and provides up to 30W of power to each device.

// twinstate.com



PoE

802.3af

15.4 watts

Types of Devices Supported

VoIP

WiFi

PoE+

802.3at

25.5 watts

Types of Devices Supported

Pan/Tilt/Zoom Cameras

Video IP Phones

Alarm Systems

// twinstate.com

PoE flavors

- PoE/PoE+ – IEEE 802.3at/af
- Cisco Universal Power over Ethernet (UPOE) – 60W
- 24V Passive PoE – *long distance & 5V convert*
- 48V Passive PoE – *idem*

LAB // can an rpi be powered by PoE?

Switch product brands

- Cisco Catalyst
- HPE Procurve
- Brocade / Ethernet switch…

Some new comers (+Wifi)

- Ubiquiti (multi-gigabit!)
- Cisco Mekari
- HPE Aruba

Just cheaper

- FS – CN

LAB // is FS's default fw CLI nice enough?

Cisco switch categories

- Small business
- LAN access
- (LAN compact)
- LAN core and distribution
- Data center
- (Blade)
- (Industrial)

Core switches (SPF+ plugs and more)

fixed, stackable only

all catalyst

- 4500-X series
- 6880-X series
- 9500 series

Catalyst legacy

- 1700, 1900, 2800 series
- 3000 series
- 5000, 6000 series

Back from the future…

- Cisco Catalyst 1000 Series Switches
- Cisco CSR 1000V (virtual & possibly nested KVM)

Today's Catalyst family

- 9300 - branch & campus access
- 9400 - campus access & aggregation
- 9500 - campus core & aggregation
- 9600 - campus core & aggregation

Wireless

- 9100 access points (incl. BLE/IoT)
- 9100 + EWC-AP - embedded controller
- 9800 - wireless

Catalyst 9000 switches

- x86-base + ASIC (UADP)
- 9200 OK got PoE+
- 9300 NOK got docker and cisco umbrella…
- 9400 got some multi-switch HA features (NSF & SSO)
- 9500 campus-ready, VPN, MPLS, NAT
- 9600 supports everything they’ve got

Example pricing

```
2 x Catalyst 6800 Sup6T (440G/slot) with 8x10GE, 2x40GE
2 x Cisco Catalyst 6824-X-Chassis and 2 x 40G (Standard Tables)
2 x Cisco Catalyst 9400 Series 24-Port 10 Gigabit Ethernet(SFP+)
2 x Cisco Catalyst 9400 Series 240GB M2 SATA memory (Supervisor)
2 x Cisco Catalyst 9400 DNA Advantage 3 Year License
+ des accessoires du chassis
+ env. 60 modules optiques (très majoritairement des 1G)
+ maintenance
sur 5 ans => 280ke
```

Credits: Dr|B00BiX on EvoluNET

IOS versions

. . .

Now with all those criterias and categories…

How to choose one?…

==> pick the rarest feature you want e.g. as of 2021, multi-gigabit

Multi-gigabit capable switches

catalyst only

- 2960-CX –> IOS LAN base
- 3560-CX –> IOS IP Base
- 9200
- 9300
- 9400
- 9600

Why the Open Source doesn’t own this market yet?…

==> this is not software

- ASICs
- Switch Fabric

By the way, Open Source is there already...

Open Networking

- FS + Cumulus Linux
- SONiC-compatible models (need 100Gbit/s?…)
- ONIE-compatible models (DELL EMC as of Mar 2021)
- VyOS
- DIY e.g. unmanaged-to-managed conversion

LAB // What FS models are Cumulus Linux powered (or can be firmware upgraded?)

// Questions on layer 2 products?

Switch Security

l2 threats & mitigations

Lateral movement attack vectors

- Rogue DHCP
- MAC flooding
- MAC spoofing
- ARP cache poisoning
- L1 DDoS
- VLAN hopping

LAB // how does a switch react to mac spoofing: blocks or sends to both?

LAB // does MAC flood still work on modern switches?

Mitigate Rogue DHCP

quoting Cisco Data Sheets (SX350X)

DHCP snooping

> *Filters out DHCP messages with unregistered IP addresses and/or from unexpected or untrusted interfaces. This prevents rogue devices from behaving as DHCP servers*

Mitigate MAC flooding

- disable hub failover mode?…

Mitigate MAC spoofing

- port security / MAC filtering
  - mac1,2,3
  - amount of MAC addresses
- EAPOL authentication

Mitigate ARP cache poisoning

Quoting Cisco Data Sheets (SX350X)

Dynamic ARP Inspection (DAI)

> *The switch discards ARP packets from a port if there are no static or dynamic IP/MAC bindings or if there is a discrepancy between the source or destination address in the ARP packet. This prevents man-in-the-middle attacks*

Static ARP entries

OpenBSD example (credits: cryptsus.com)

```
arp -s 85.85.85.1 DE:AD:BE:EF:01:00 permanent
arp -s 192.168.144.1 DE:AD:BE:EF:01:01 permanent
arp -s 192.168.244.1 DE:AD:BE:EF:01:02 permanent
arp -s 192.168.200.1 DE:AD:BE:EF:01:03 permanent
```

Maybe a GOOD PRACTICE

- against gateways and critical servers on that segment
- dynamic MAC-IP pairs are still allowed

LAB // otherwise simply prevent gratuitous ARP? (probably won't be enough anyhow)

Mitigate L1 DDoS

Secure Core Technology (SCT)

Makes sure that the switch will receive and process management and protocol traffic no matter how much traffic is received

// Questions on switch security?

L2 Hardening

- authorize access only to corporate users…

How can we do that?… (hint: WPA/WPA2)

==> EAPOL authentication

encapsulating EAP over LAN

- Better than MAC white list
- Better than PAP - stores and transits passwords in clear…
- Better than CHAP - stores e.g. MD5

EAP methods

Lightweight Extensible Authentication Protocol (LEAP)

EAP Transport Layer Security (EAP-TLS)

EAP-MD5

EAP Protected One-Time Password (EAP-POTP)

EAP Pre-Shared Key (EAP-PSK)

EAP Password (EAP-PWD)

EAP Tunneled Transport Layer Security (EAP-TTLS)

EAP Internet Key Exchange v. 2 (EAP-IKEv2)

EAP Flexible Authentication via Secure Tunneling (EAP-FAST)

Tunnel Extensible Authentication Protocol (TEAP)

EAP Subscriber Identity Module (EAP-SIM)

EAP Authentication and Key Agreement (EAP-AKA)

EAP Authentication and Key Agreement prime (EAP-AKA')

EAP Generic Token Card (EAP-GTC)

EAP Encrypted Key Exchange (EAP-EKE)

Nimble out-of-band authentication for EAP (EAP-NOOB)

What’s used for GSM family networks?…

from weaker to stronger

- EAP-SIM (2G)
- EAP-AKA
- EAP-AKA’

FOSS implementations

- FreeRADIUS
- Osmocom…

LAB // HostAPD + RADIUS or Diameter AAA

EAP-TLS vs. EAP-TTLS

- EAP-TLS – bi-directional auth at once
- EAP-TTLS – server auth then possibly client auth within the ssl tunnel

EAP encapsulations

- IEEE 802.1X / Dot1X – EAP over LAN (EAPOL)
- Protected Extensible Authentication Protocol (PEAP) – EAP over SSL

EAPOL use-cases

- LAN w/ or w/o AAA facilitator
- **WLAN** (Wifi) w/ or w/o AAA facilitator
- FDDI
- MACsec
- IDevID

LAB // wpa_supplicant for ethernet?

Methods for EAPOL

- **EAP-TLS** (mandatory for Wifi compliance)
- many others…

Methods for PEAP

Most commonly

- PEAPv0/EAP-MSCHAPv2 (over SSL)
- PEAPv1/EAP-GTC (over SSL)

AAA facilitators

- RADIUS
- (SS7)
- Diameter

Alternatives

- AEGIS SecureConnect
- Protocol for Carrying Authentication for Network Access (PANA)

Even Yamaha

Ethernet switches optimized for sound infrastructures

> *Mac authentication, Web authentication, and IEEE802.1X authentication can be used with the RADIUS server function. They can be used together by setting them to each port.*

// Questions on l2 hardening?

VLAN++ & VLAN Hopping

- IEEE 802.1Q – Dot1q / VLAN on Ethernet
- IEEE 802.1ad – QinQ
- Virtual Extensible LAN (VXLAN)

QinQ

- basic QinQ – kind of CoS (port based)
- selective QinQ – inner VLAN based on mac/ip/src-ip/vlan-tag

VXLAN

- layer 2 overlay on top of layer 3
- MAC address-in-user datagram protocol (MAC-in-UDP)

quick overview on how ***managed switches and routers*** *communicate with each other*

Dynamic Trunking Protocol (DTP)

Cisco only

Negotiate switch interconnection as access or trunk

```
Access

Trunk

Dynamic Auto (mostly the default)

Dynamic Desirable

No-negotiate
```

VLAN Trunk Protocol (VTP)

Cisco only

- VTP server mode –> distributes VLANs
- VTP client mode –> receives VLANs
- VTP transparent mode –> don’t talk VTP

VLAN hopping — switch spoofing

Dynamic Trunking Protocol (DTP) negociates trunking modes

While we're not even a switch

- we go for `Trunk` or `Desirable` and you will most probably get a trunk
- we say we want VLAN x

VLAN hopping — double tagging

- Not necessarily evil — by design for ISPs (QinQ)
- First tag is the normal one
- Second tag to send the frame to the target VLAN

(Relative) Success

- Works against a native VLAN (VLAN 1)
- –and– works against a port trunk

But the other side cannot answer

- doesn't know about your originating VLAN
- making it an unidirectional flow

LAB // PoC VLAN hopping somehow

LAB // double tag works only against a trunk port, but let’s try with access port anyhow

LAB // other methods for VLAN hopping? CDP? VTP?

// Questions on vlan++ & vlan hopping?