How to operate a VXLAN network

DENOG10 - November 21, 2018
Nicola von Thadden <nicola.vonthadden@pfalzkom-manet.de>

Ein Unternehmen der Pfalzwerke-Gruppe.

PFALZKOM | MANET

Who is PfalzKom|MAnet

- Operate 2 DCs
- Operate a carrier hotel
- 12.000 m² ruled area
- >1.400 km of fibre
- WDM, MPLS, IP, Ethernet...

Planning a VXLAN Setup

- What is your use-case
- Where do you want to install it
- What do you want to connect
- How many ports do you need
- Which bandwidth do you need
- Is your hardware VXLAN aware
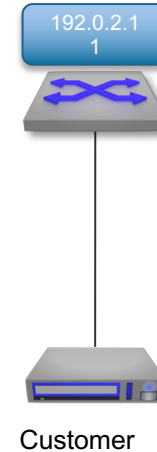
Connecting clients

Do you connect your clients via L2 or L3
How do you handle L2 redundancy

Single port connections are easy

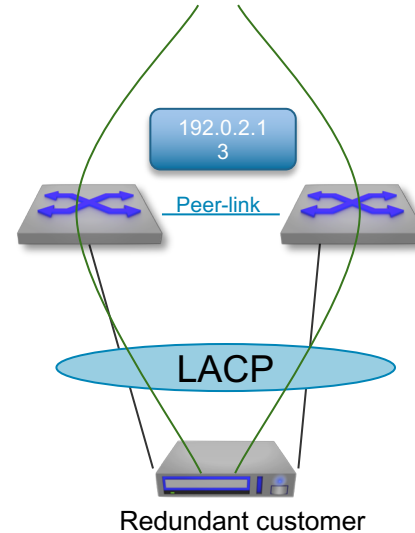Redundant L2 connections without STP?
LACP as MLAG can help

192.0.2.1
1

Customer

MLAG connectivity

One VTEP IP per leaf pair

VXLAN Traffic is spread via ECMP

Peer-link for state sync

Good way to achieve redundancy for L2 clients

192.0.2.1
3

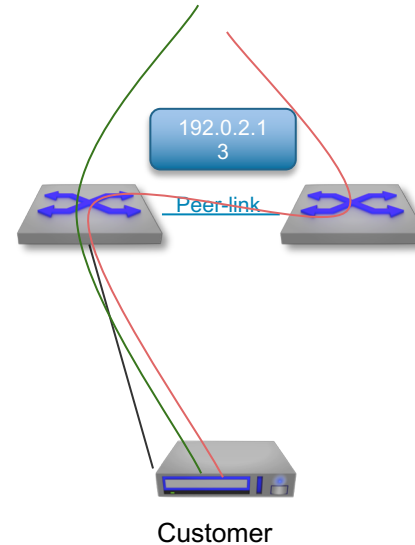Peer-link

LACP

Redundant customer

MLAG connectivity for single homed customer

Still one VTEP IP per leaf pair
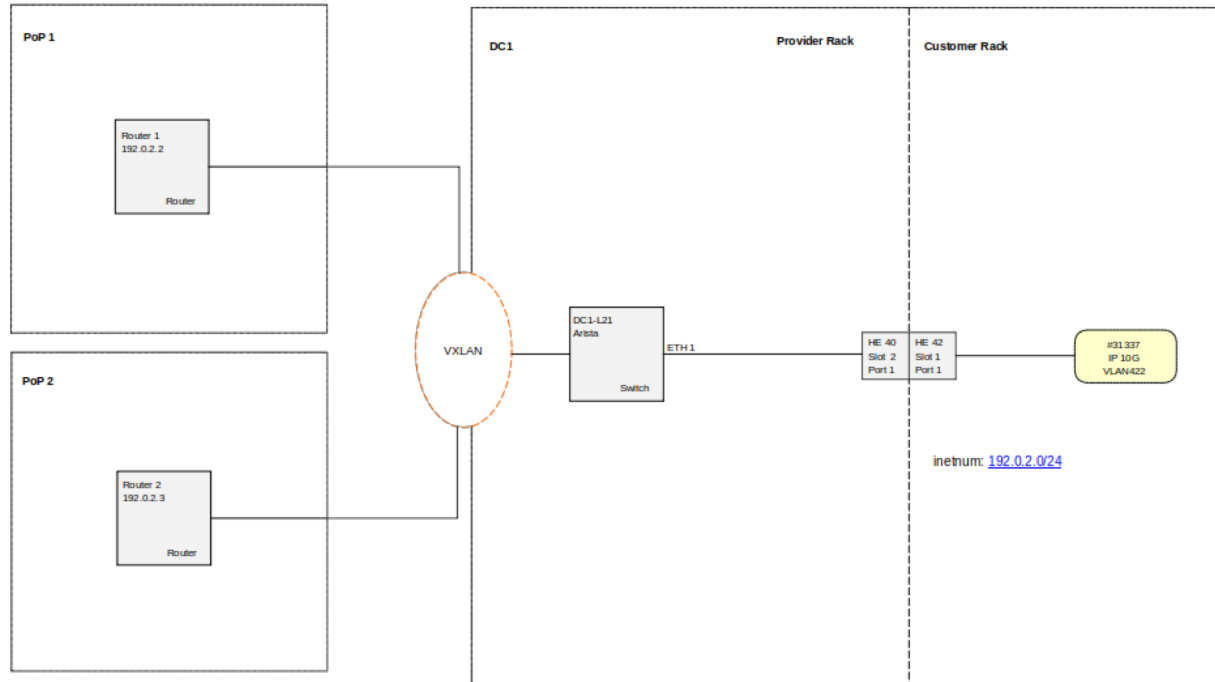VXLAN Traffic is still spread via ECMP

Traffic switched via peer link

Not ideal when connecting single ports
Not ideal when customer hardware is not LACP
capable and prefers different means of redundancy

192.0.2.1
3

Peer-link

Customer

Documentation

Different view for physical topology and services

Our setup

2 spine switches per datacenter

Dedicated leaves for connections leaving the building
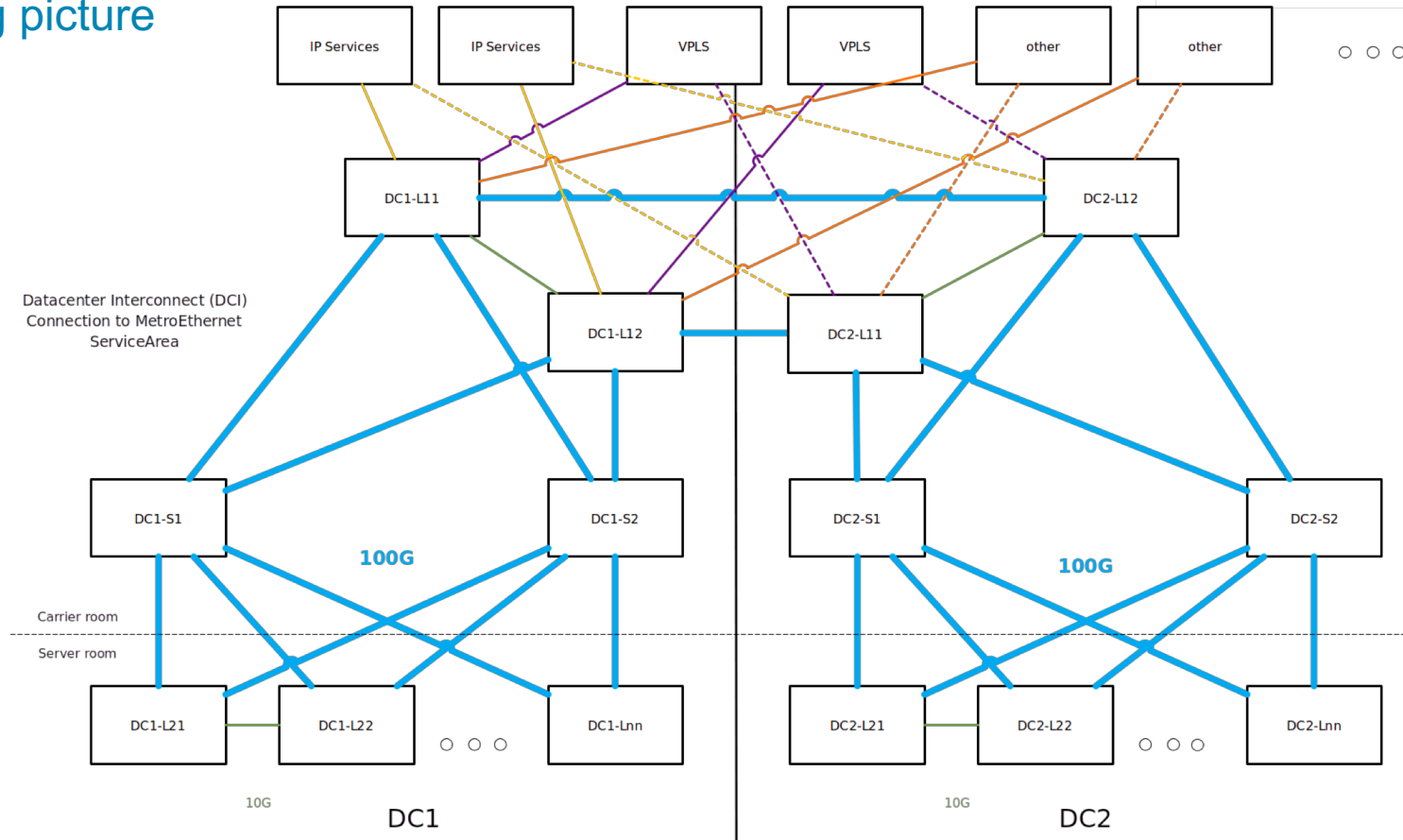
Complete VXLAN based fabric

MLAG to VPLS and routing Hardware

Full MLAG configuration

Controller based setup

Similar setup in our lab

The big picture

Connecting customers

1 or more ports per customer

Per customer & service segmentation

VNI Scheme: VLAN ID + 100000

MLAG preferred when redundancy is ordered

```
!Example config
vlan 539
  name nicoduck_demo
!
interface Ethernet3
  description #1337_nicoduck_demo
  switchport access vlan 539
!
interface Vxlan1
  vxlan vlan 539 vni 100539
```

MLAG connectivity

Port-channel per Port

Port-channel ID derived from physical port

```
vlan 539
  name nicoduck_demo
!
interface Ethernet3
  description #1337_nicoduck_demo
  Channel-group 103 mode active
interface Port-Channel103
  description #1337_nicoduck_demo
  switchport access vlan 1539
  mlag 103
!
interface Vxlan1
  vxlan vlan 539 vni 100539
```

Handling customer-owned VLANs

Assign own VNI range per customer
e.g. 200000 + customer VLAN ID
VLAN translation on shared devices

```
!Example config
vlan 1539
  name nicoduck_internal
!
interface Ethernet4
  description #1338_nicoduck_internal
  switchport access vlan 1539
!
interface Vxlan1
  vxlan vlan 1539 vni 201539
!
```

What really gets connected

01 Non-enterprise switches

02

03

04

- ... but not MLAG between 2 of them
- ...ers don't want ports ($)

...move

...ncy better

Try not to do LACP
Prefer MAC to port binding

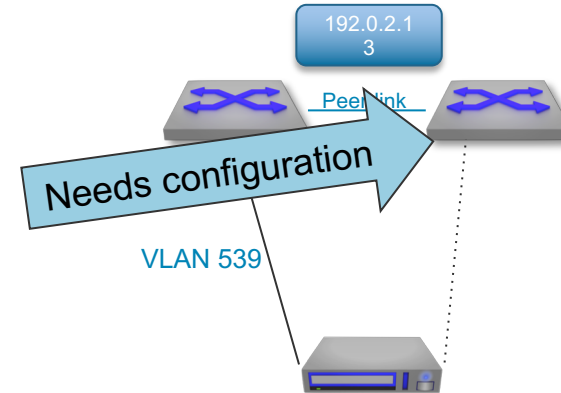Traffic on MLAG peer link

Pitfalls

Configuring VLAN only on one MLAG device

No traffic is flowing until both switches know the VLAN

VLANs must be configured on both MLAG devices
Even for single-port connections

192.0.2.1
3

Peerlink
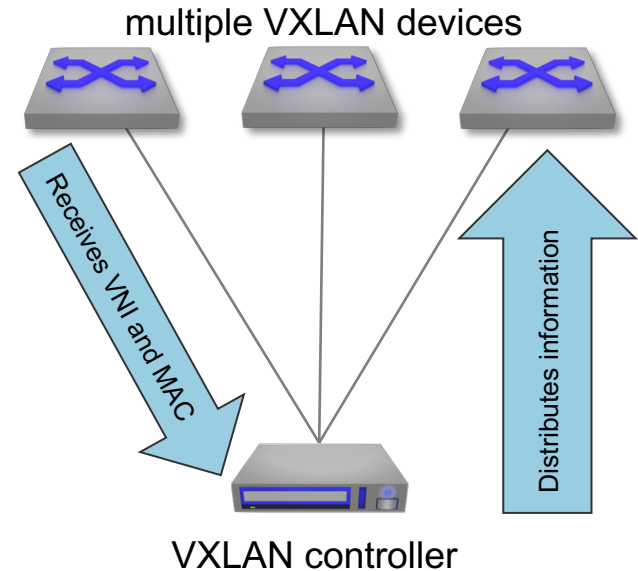
Needs configuration

VLAN 539

VXLAN controller

Managing flood list manually is painful

Control-plane based MAC learning can minimise
BUM traffic

Automatic fallback when controller is unreachable

multiple VXLAN devices

Receives VNI and MAC

Distributes information

VXLAN controller

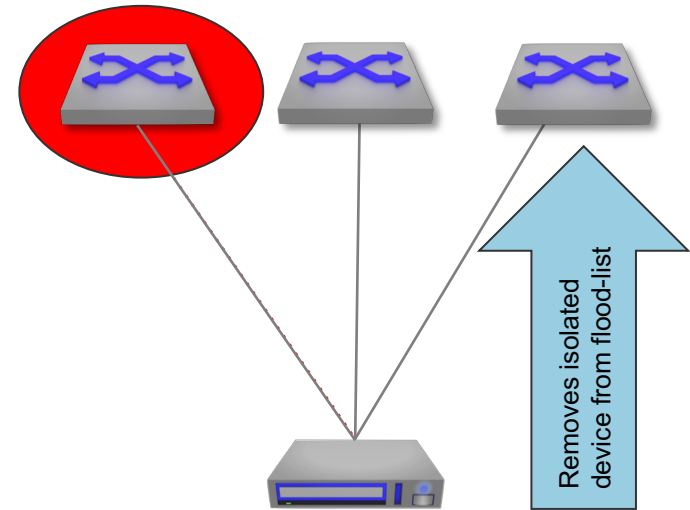# VTEP/VNI Learning via Controller

Controller pushes current flood list to connected devices

Device hold their current list when connection drops

Unicast traffic is not affected

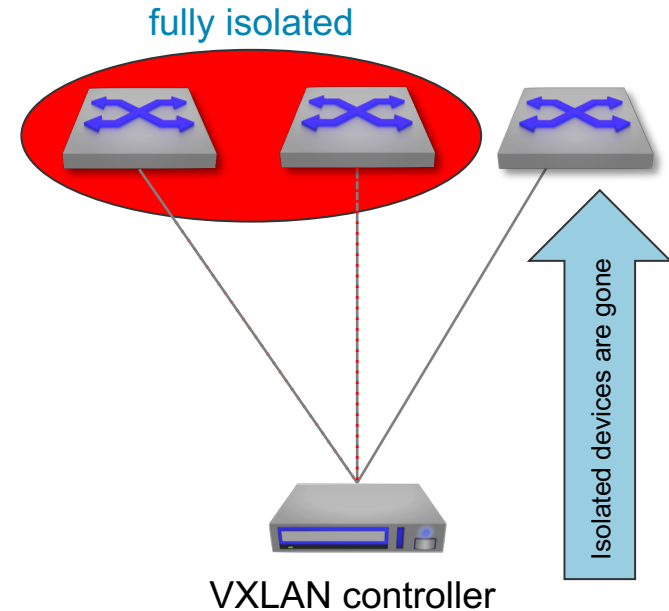BUM traffic dropped since no destination can be found

partially isolated

Removes isolated device from flood-list

VXLAN controller

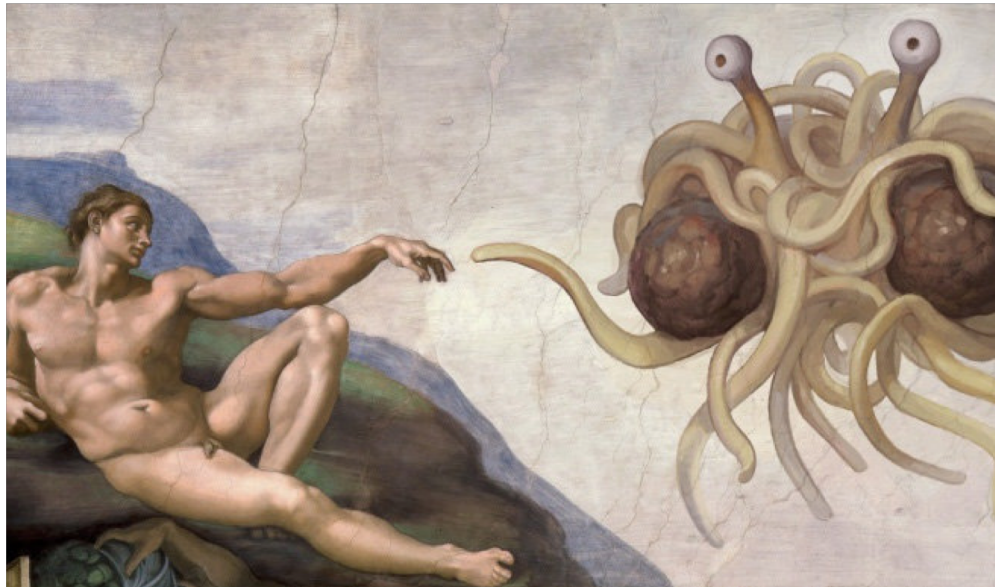# Control-plane based MAC learning

Mutually exclusive to data-plane based learning
Fails over to HER when connection is lost

Controller removes MAC addresses from DB
Isolated device falls back to HER
Isolated devices can reach each other
Non-isolated devices don't do HER
No traffic is flowing between isolated and non-isolated device

fully isolated

Isolated devices are gone

VXLAN controller

How to mitigate

Use fully redundant connections for controller

In-band via BGP in underlay network

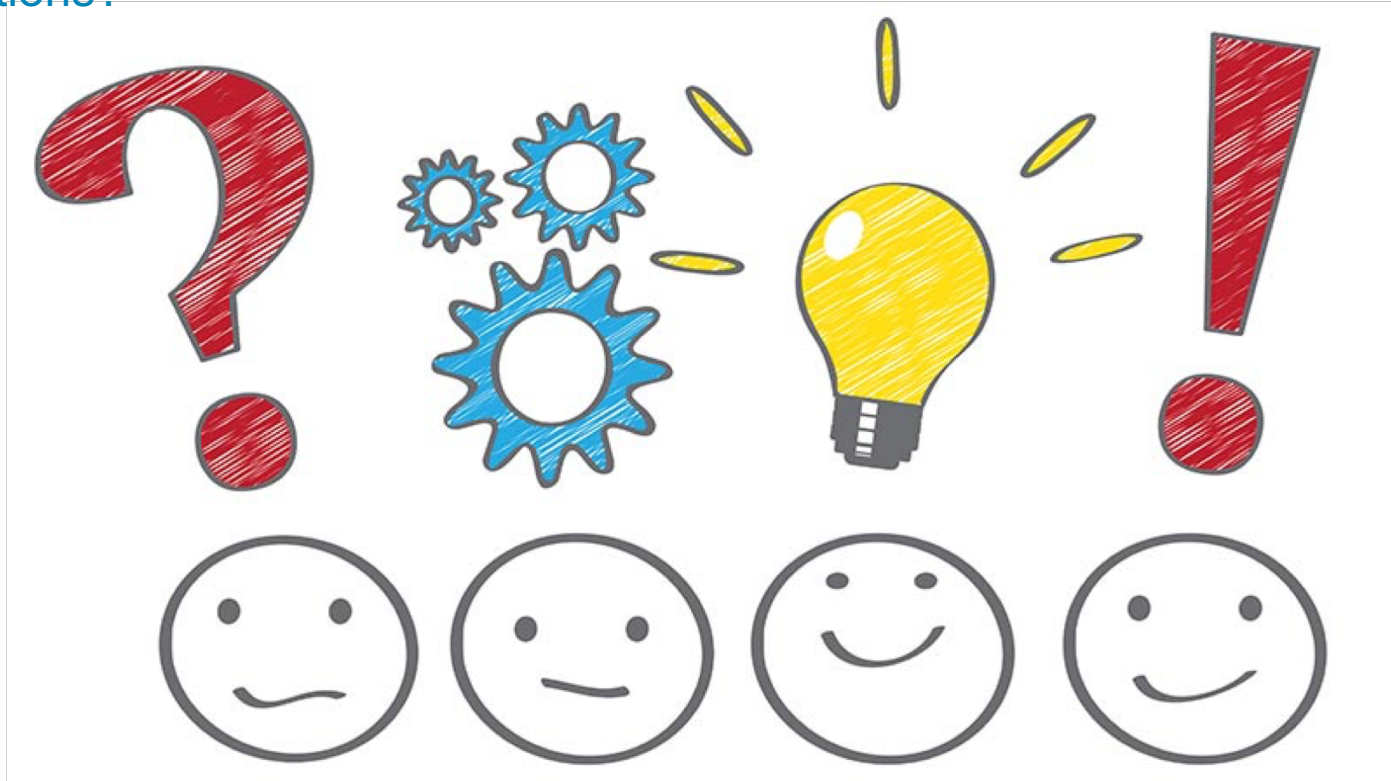**The Flying Spaghetti Monster by Niklas Jansson. Public Domain by the author.**

Conclusion

VXLAN is not that hard

Find out if you want and need MLAG

Make sure your control-plane does not introduce a SPOF

Questions?

THANK YOU FOR YOUR

ATTENTION