The Bits Must Flow

netWorking through the abstractions

@crayzeigh@hachyderm.io

THAT®

CONFERENCE

SPECIAL THANKS TO ALL OUR AWESOME CAMP SPONSORS!

TruStage™

HEARTLAND BUSINESS SYSTEMS

Intellibus

aws

Unspecified SOFTWARE CO

Architect.io

ionic

CAMUNDA

alexa

Google Cloud

CORE BTS

GrapeCity

Join Us Next Year!

WI 24

THAT®

CONFERENCE

JULY 29TH - AUG. 1ST

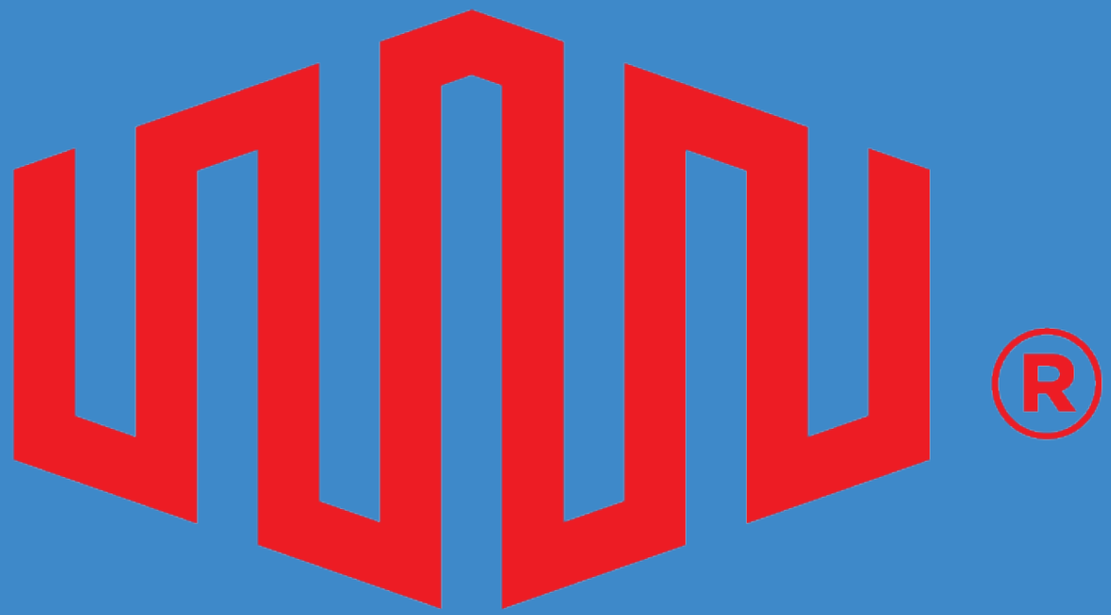
Hi 👋

I'm fen (they/them)

social: @crayzeigh@hachyderm.io

slides: speaking.crayzeigh.com

Developer Advocate:



EQUINIX

Internet Traffic Per Second

167,069 GB

bits per second

1,336,544,000,000,000

What happens when you visit a website?

It's not DNS

There's no way it's DNS

It was DNS

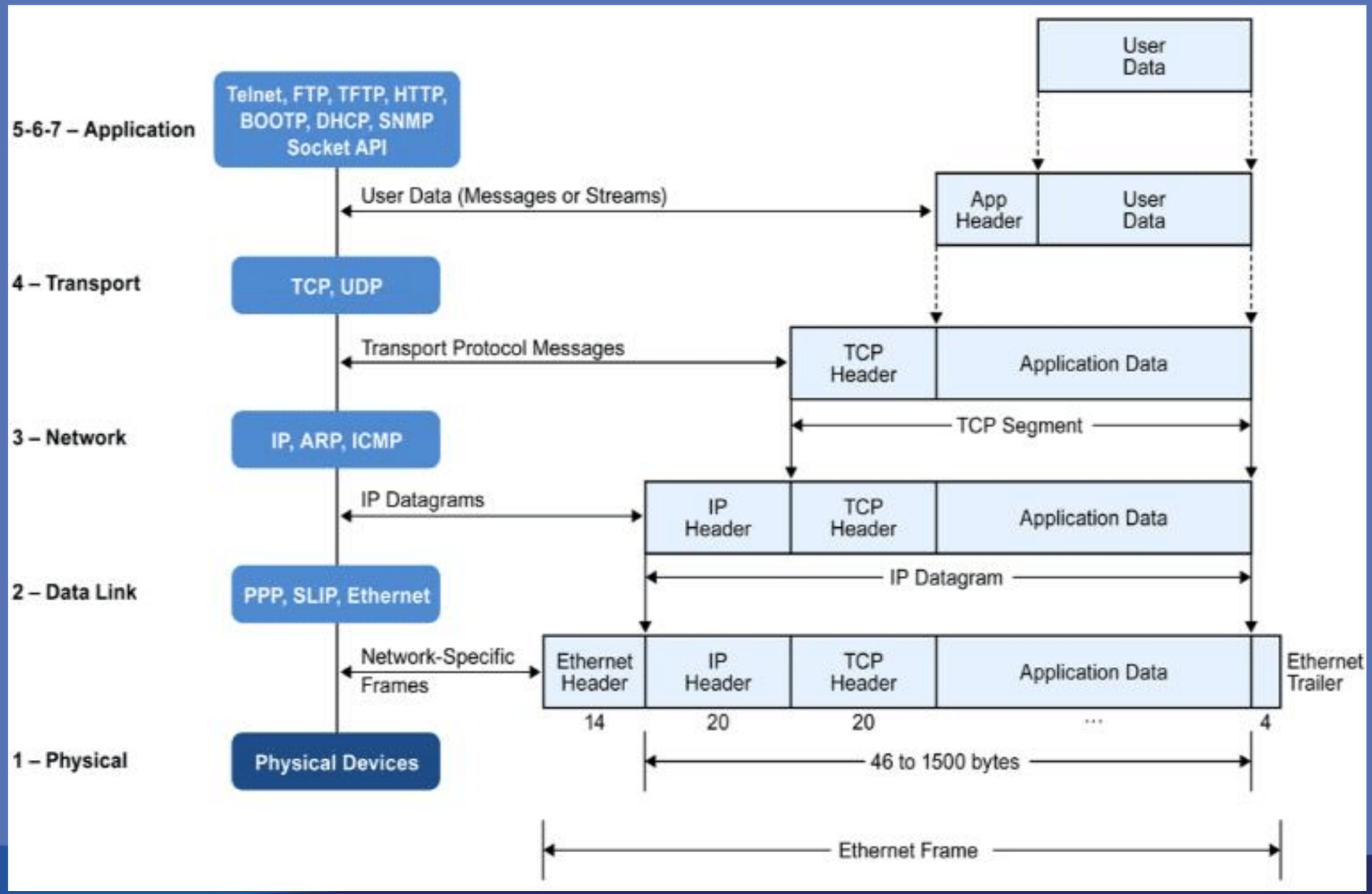
-SSBroski

OSI 7-Layer model

Application
Presentation
Session
Transport
Network
Data Link
Physical

OSI	TCP/IP
Application	Application
Presentation	-
Session	-
Transport	Transport
Network	Internet
Data Link	Network Access
Physical	-

Why all these layers anyway?



Layer 2: Network Access

Data Frames link digital to physical

Layer 2: Switching

Sending data to local devices

Frame Header

Preamble	SFD	Dest. MAC	Src. MAC	Type	Data & Pad	FCS
7 Bytes	1 Byte	6 Bytes	6 Bytes	2 Bytes	46–1500 Bytes	4 Bytes

MAC[1] Address

01:23:45:67:89:ab

1. Media Access Control

MAC Address

- Identifies the (network) device
- For devices on the *local* network

Address Resolution Protocol

ARP

- Mapping IPs and MAC addresses

ARP

- Mapping IPs and MAC addresses
- Necessary for your device to talk to your router

ARP

- Mapping IPs and MAC addresses
- Necessary for your device to talk to your router
- Finds neighbors through broadcast

ARP

- Mapping IPs and MAC addresses
- Necessary for your device to talk to your router
- Finds neighbors through broadcast
- `arp -a` for your local table

Virtual Local Area Networks

VLANs

- Used for limiting broadcast domains

VLANs

- Used for limiting broadcast domains
- up to 4096 VLANs[1]

1. VXLAN addresses this limitation but that's A Whole Other Thing

VLANs

- Used for limiting broadcast domains
- up to 4096 VLANs[1]
- Can be native (default traffic on a switch port) or tagged (logically divided in packet header)

1. VXLAN addresses this limitation but that's A Whole Other Thing

Learn about Layer 2

Manage VLAN Search table + Add VLAN

	VLAN	Devices	Location	Deployed Date	Description
☐	1049	0	DC	Sep 4th, 2020 5:53 AM	provisioning_vlan
☐	1173	0	NY	2 years ago	layer2-testing
☐	1174	0	NY	2 years ago	layer2-testing-2
☐	1254	0	DA	Oct 14th, 2020 3:33 PM	"vSAN"
☐	1336	1	CH	23 days ago	default
☐	1337	1	CH	23 days ago	elite
☐	1338	0	DA	Oct 14th, 2020 3:33 PM	"Management"
☐	1339	0	DA	Oct 14th, 2020 3:33 PM	"VM Private Net"
☐	1340	0	DA	Oct 14th, 2020 3:33 PM	"vMotion"
☐	1440	0	DA	Oct 14th, 2020 3:33 PM	"VM Public Net"
☐	2020	0	CH	22 days ago	Metal-AWS

Layer 3: Internet Protocol

Packets wrap your digital data and route it remotely

Layer 3: Routing

Directing data to remote destinations

IPv4 Packet Header

Internet Protocol Version 4

0			15	16			31
Version		Header L...	Differentiated Services ...	Total Length			
Identification				Flags	Fragment Offset		
Time to Live		Protocol		Header Checksum			
Source Address							
Destination Address							
Options: (4 bytes), Router Alert							

IP Address Classes and CIDR

IP Classes & CIDR

– **C**lassless **I**nter-**D**omain **R**outing

IP Classes & CIDR

- **C**lassless **I**nter-**D**omain **R**outing
- Helps determine destination locality, i.e. routing

IP Classes & CIDR

- **C**lassless **I**nter-**D**omain **R**outing
- Helps determine destination locality, i.e. routing
- CIDR replaced "class a/b/c" IP addressing to help address IP address availability

CIDR Notation

10.10.10.10/24

Host/Network bits

10.10.10/24

IP Address: 10.10.10.10

Subnet Mask: 255.255.255.0

Converts to Binary

```
IP: 00001010.00001010.00001010.00001010
SM: 11111111.11111111.11111111.00000000
```

In the Subnet Mask:

`1` = Network

`0` = Host

Special IPs

- Broadcast (ex. 10.10.10.255)
  - host bits are all 1's
  - For sending data to all hosts in a network
- Network (ex. 10.10.10.0)
  - host bits are all 0's
  - only used for forwarding data between routers

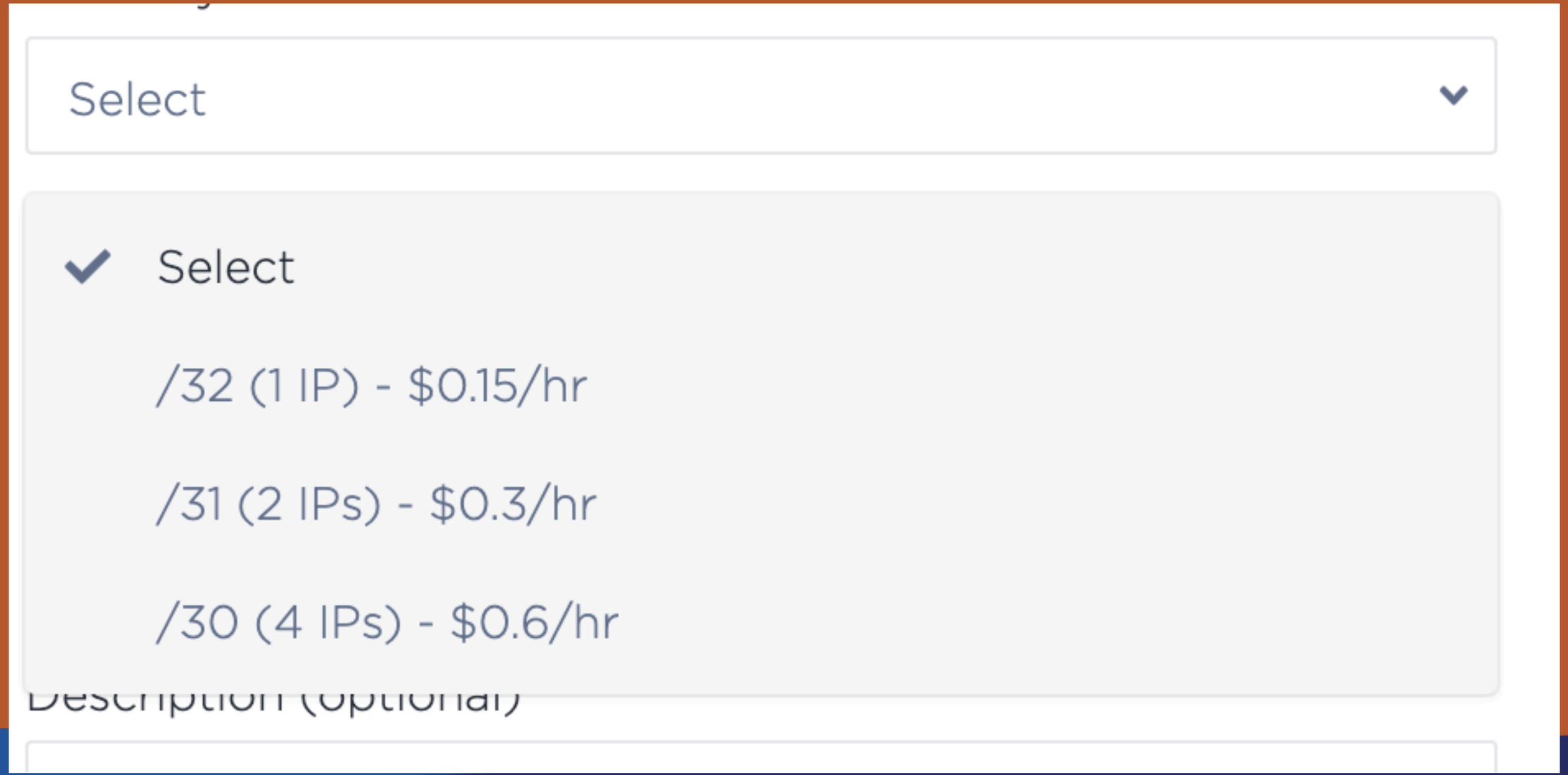
Putting it all together

CIDR	10.10.10.10/24
Network	10.10.10.0/24
Broadcast IP	10.10.10.255
Available Host IPs	10.10.10.1 - 254

Bigger Networks

CIDR	192.168.1.100/22
Network	192.168.0.0/22
Broadcast IP	192.168.3.255
Available Host IPs	192.168.0.1 - 192.168.3.254

Weird Ones



Weird Ones Explained

/30

- "Costs" 4 IPs, but only gives 2 host addresses
- Broadcast & Network IPs still required
- Used for legacy compatibility or you just really like holding IP addresses

Weird Ones Explained

/31

- Creates 2 adjacent host IPs
- Only "costs" 2 IPs
- Proposed in RFC3021 (in 2000) to combat dwindling IP availability

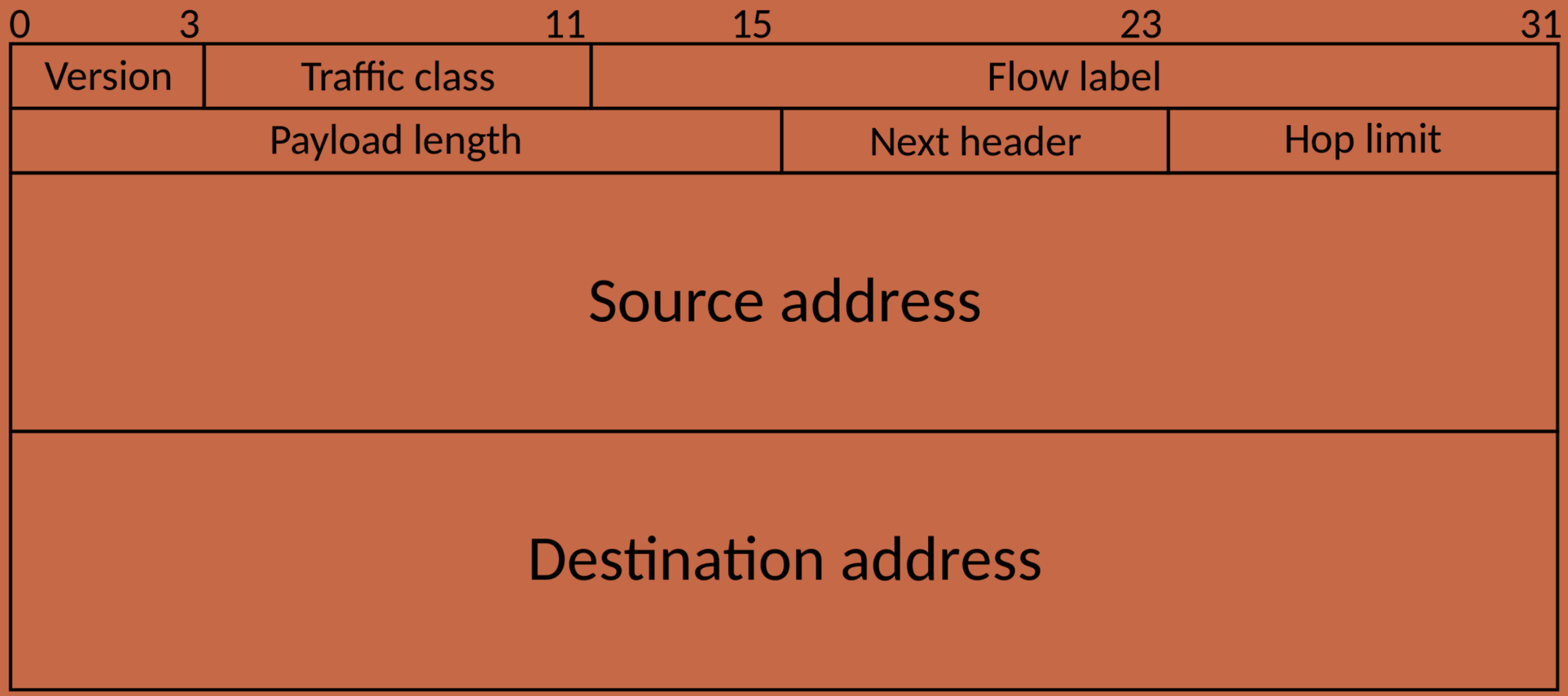
Weird Ones Explained

/32

- Single IP Address
- useful generally for isolating public internet traffic

IPv6

IPv6 Packet Header



IPv6

- 128 bits long (vs 32 bits for v4)

IPv6

- 128 bits long (vs 32 bits for v4)

  $2^{32} \approx 4.3x10^{9}$

IPv6

- 128 bits long (vs 32 bits for v4)

  $2^{32} \approx 4.3x10^{9} \approx 4.3$ billion

IPv6

- 128 bits long (vs 32 bits for v4)

$2^{32} \approx 4.3x10^{9} \approx 4.3$ billion

$2^{128} \approx 3.4x10^{28}$

IPv6

- 128 bits long (vs 32 bits for v4)

  $2^{32} \approx 4.3x10^{9} \approx 4.3$ billion

  $2^{128} \approx 3.4x10^{28} \approx 27$ orders of magnitude larger

IPv6

- 128 bits long (vs 32 bits for v4)
- introduced to handle IP shortage
- written as 8 groups of 4 bytes in hex

Abbreviations

fddd:f00d:b33f:0000:0000:0000:0000:0001

Abbreviations

fddd:f00d:b33f:0:0:0:0:1

Abbreviations

fddd:f00d:b33f::1

Abbreviations

00ff:0000:0000:0000:0001:0000:0000:0001

Which one is correct?

ff:0:0:0:1:0:0:1

1. ff::1:0:0:1
2. ff::1::1
3. ff:0:0:0:1::1

Which one is correct?

ff:0:0:0:1:0:0:1

1. ff::1:0:0:1
2. ~~ff::1::1~~ (ambiguous)
3. ~~ff:0:0:0:1::1~~ (longest 0's must be shortened[1])

1. RFC 5952

Subnetting

Network Bits				Host Bits			
Routing Prefix			Subnet ID	Interface Identifier			
XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX

Subnetting

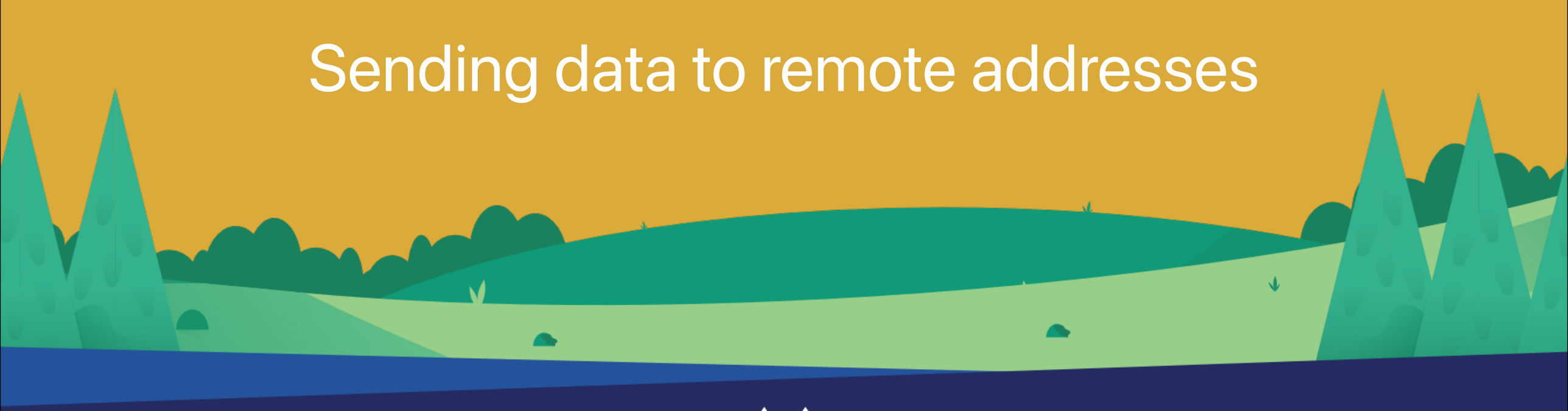
```
ifconfig | grep inet6
    inet6 ::1 prefixlen 128
    inet6 fe80::1%lo0 prefixlen 64 scopeid 0x1
    inet6 fe80::1470:43d6:1243:7a20%en0 prefixlen 64 secured
scopeid 0x4
    inet6 2601:19e:8380:20dc:4ea:6990:6d8e:68a3 prefixlen 64
autoconf secured
    inet6 2601:19e:8380:20dc:5ce:b0e:70ad:988 prefixlen 64
deprecated autoconf temporary
    inet6 2601:19e:8380:20dc::f42a prefixlen 64 dynamic
    inet6 2601:19e:8380:20dc:141b:7764:190d:9240 prefixlen 64
deprecated autoconf temporary
    inet6 2601:19e:8380:20dc:d477:fa12:2a44:7983 prefixlen 64
autoconf temporary
...
```

Subnetting

address	obtained
::1	loopback (special)
fe80::1470:43d6:1243:7a20	self-assigned (local)
2601:19e:8380:20dc:4ea:6990:6d8e:68a3	self-assigned (global)
2601:19e:8380:20dc::f42a	DHCPv6 (global)

Routing

Sending data to remote addresses



Source: 10.10.10.10/24

Destination: 10.10.10.100

Source IP: 10.10.10.10

Network: 255.255.255.0

Destination: 10.10.10.100

Source: 10.10.10.10/24

Destination: 10.10.10.100

1. Checks network space to see address is local
2. Sends local ARP broadcast to find MAC of destination
3. wraps packet in frame with newly discovered MAC
4. sends data frame through switch to destination "directly"

Source: 10.10.10.10/24

Destination: 1.1.1.1

Source IP: 10.10.10.10

Network: 255.255.255.0

Destination: 1.1.1.1

Source: 10.10.10.10/24

Destination: 1.1.1.1

1. Checks network space and see's address is remote
2. Wraps packet in frame with router's mac address and desired destination IP, forwards to router
3. Forwards frame to the router, router re-wraps packet with a frame pointing to the next router in line
4. And so on until the local router for 1.1.1.1 receives the packet and wraps in the final destination frame

How does the router know where the next hop in line is?

Routing Tables

3 Types of Routes:

- **Connected**: The networks connected to your interfaces
- **Static**: Manually set, e.g. Default Route/Gateway
- **Learned**: Learned from advertising peers, e.g. via BGP

BGP

fen

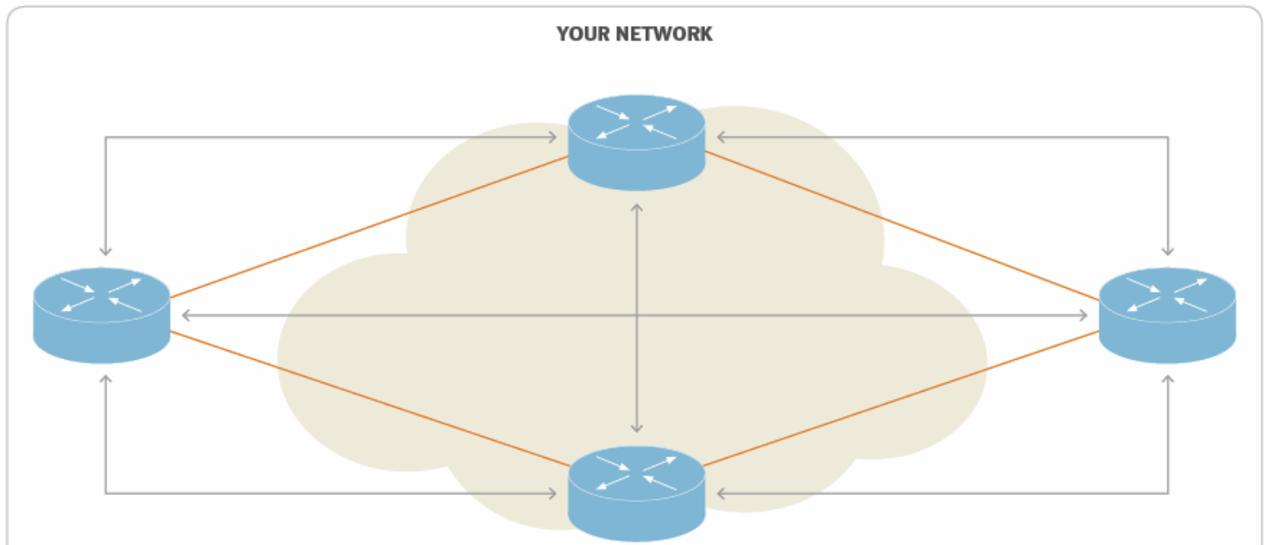
@crayzeigh@hachyderm.io

@malanalysis to err is dns, but to globally f
things up takes bgp

Dec 15, 2022 at 11:05 · · Toot! · 10 · 19

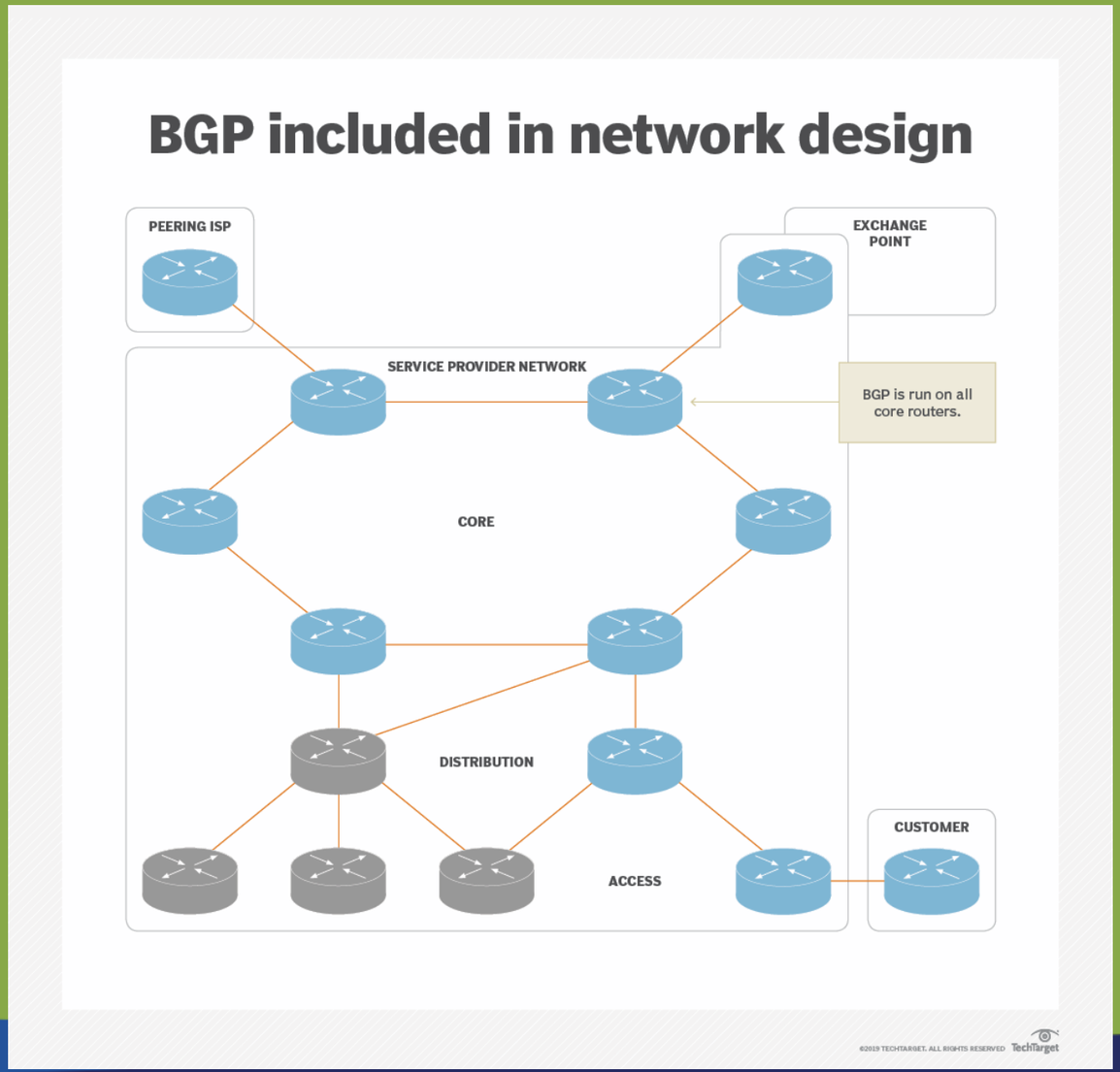
BGP

Full-mesh internal BGP

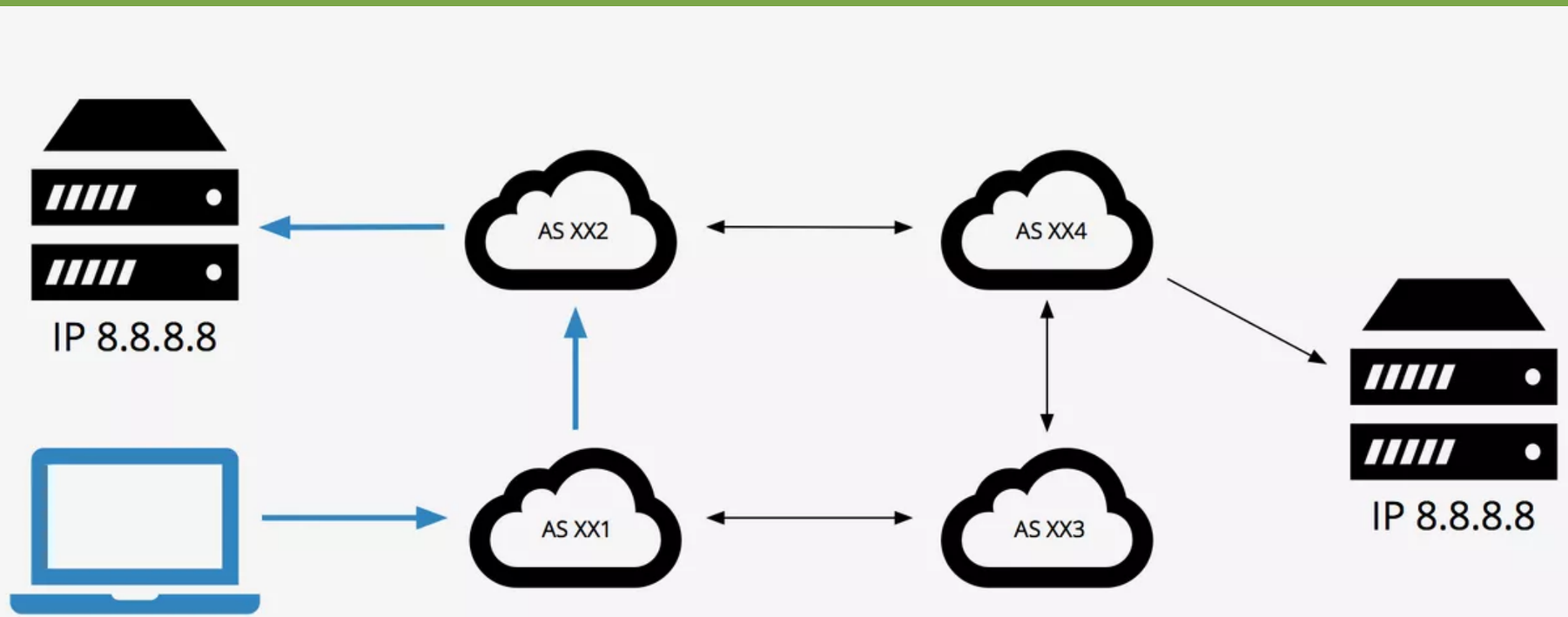


©2019 TECHTARGET. ALL RIGHTS RESERVED TechTarget

BGP


BGP included in network design
PEERING ISP
EXCHANGE POINT
SERVICE PROVIDER NETWORK
BGP is run on all core routers.
CORE
DISTRIBUTION
ACCESS
CUSTOMER
©2019 TECHTARGET. ALL RIGHTS RESERVED
TechTarget

Anycast



Anycast

Shortest Path: 8.8.8.8 AS path xx1 xx2

8.8.8.8 AS path xx1 xx3 xx4
8.8.8.8 AS path xx1 xx2 xx4
8.8.8.8 AS path xx1 xx3 xx4 xx2

Anycast Benefits

IP address or hostname

147.75.40.38

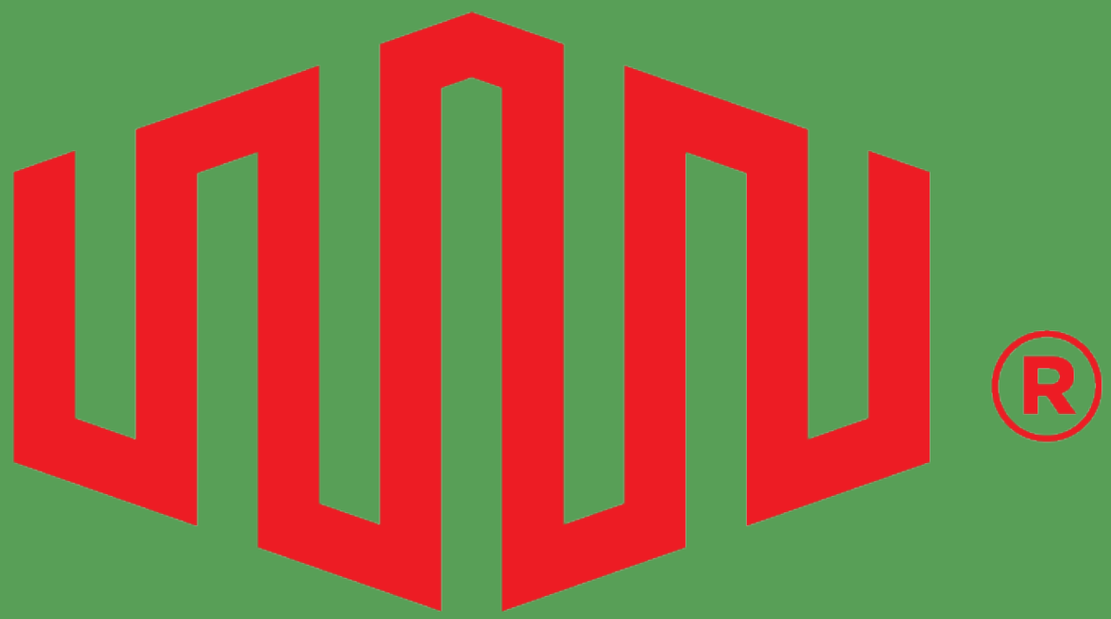
Test

LOCATION	REQ	MIN	MAX	AVG	STD DEV	LOSS
Frankfurt 147.75.40.38	3	8.42 ms	8.52 ms	8.47 ms	0.04 ms	0%
Amsterdam 147.75.40.38	3	0.73 ms	1.19 ms	0.97 ms	0.19 ms	0%
London 147.75.40.38	3	7.9 ms	8.88 ms	8.27 ms	0.43 ms	0%
New York 147.75.40.38	3	1.71 ms	2.53 ms	2.06 ms	0.34 ms	0%
Dallas 147.75.40.38	3	39.06 ms	39.95 ms	39.42 ms	0.38 ms	0%
San Francisco 147.75.40.38	3	2.2 ms	2.65 ms	2.38 ms	0.2 ms	0%
Singapore 147.75.40.38	3	1.23 ms	1.83 ms	1.46 ms	0.27 ms	0%
Sydney 147.75.40.38	3	92.49 ms	92.57 ms	92.54 ms	0.04 ms	0%
Tokyo 147.75.40.38	3	74.49 ms	74.6 ms	74.57 ms	0.05 ms	0%
Bangalore 147.75.40.38	3	147.75 ms	148.62 ms	148.05 ms	0.41 ms	0%

Anycast Benefits

Amsterdam 147.75.40.38	3	0.73 ms	1.19 ms	0.97 ms	0.19 ms	0%
London 147.75.40.38	3	7.9 ms	8.88 ms	8.27 ms	0.43 ms	0%
New York 147.75.40.38	3	1.71 ms	2.53 ms	2.06 ms	0.34 ms	0%
Dallas 147.75.40.38	3	39.06 ms	39.95 ms	39.42 ms	0.38 ms	0%
San Francisco 147.75.40.38	3	2.2 ms	2.65 ms	2.38 ms	0.2 ms	0%
Singapore 147.75.40.38	3	1.23 ms	1.83 ms	1.46 ms	0.27 ms	0%

What Happens when you visit a website?



EQUINIX

Try it!

use discount code **THATCONF23**
for $800 credit

http://eqix.co/metal

/assets/Simple_Footer_center.png

Getting Started

check out our youtube channel for tips and ideas

http://eqix.co/start-metal

Session Feedback

thanks!

👋

social:

@crayzeigh@hachyderm.io

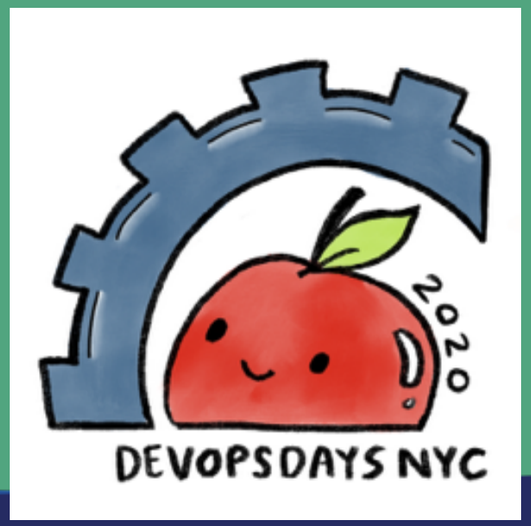
slides:

speaking.crayzeigh.com

organizer:


DEVOPSDAYS
Hartford


DEV
OPS
DAYS
BOS
22


2020
DEVOPSDAYS NYC

host:


I am
>Code
My superpower is:
greaterthancode.com


TRACE
ROUTE
A PODCAST ABOUT THE INNER
WORKINGS OF OUR DIGITAL WORLD

THAT®

CONFERENCE