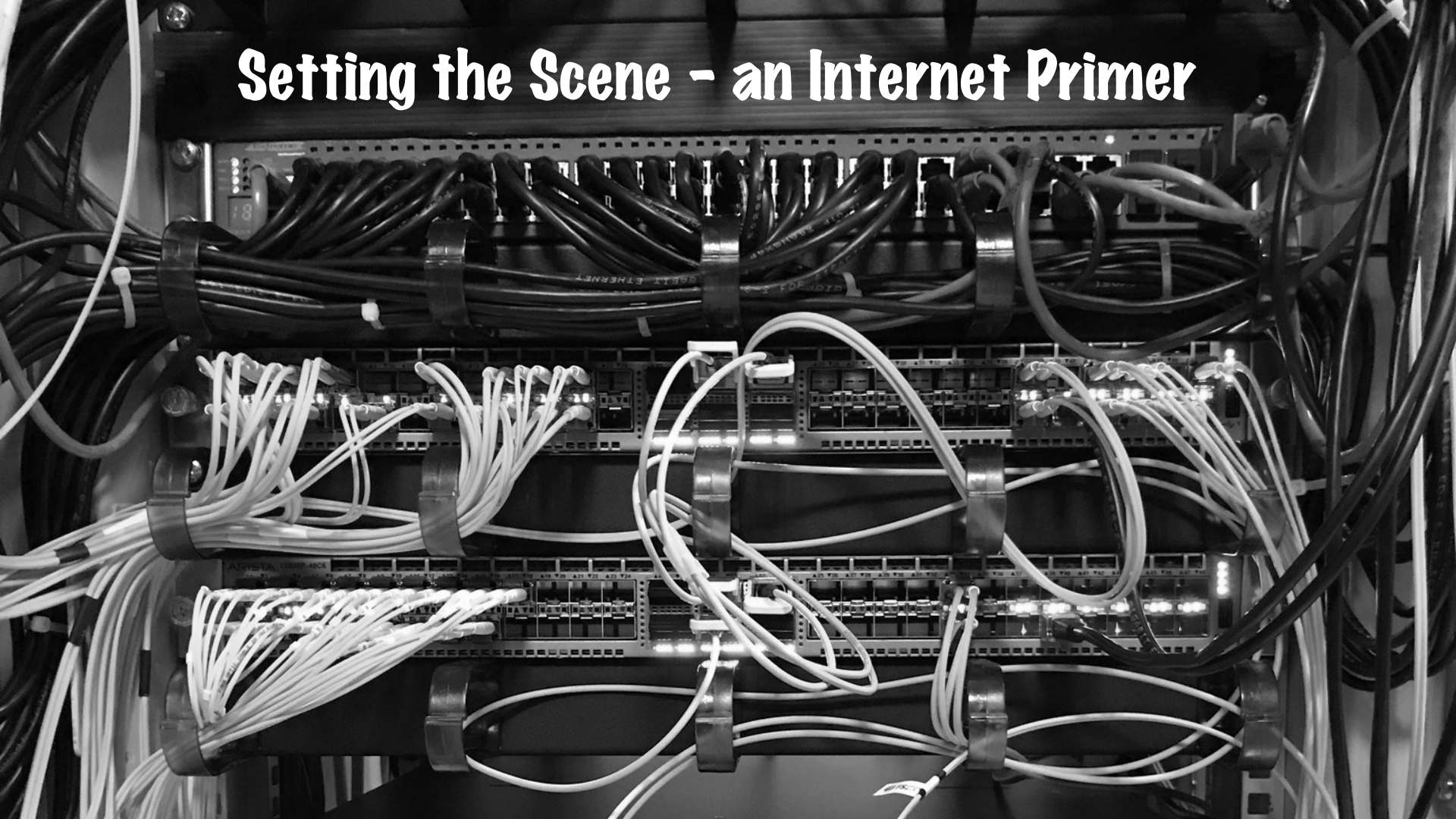
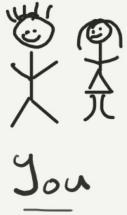
# IXPs, BGP and the Internet

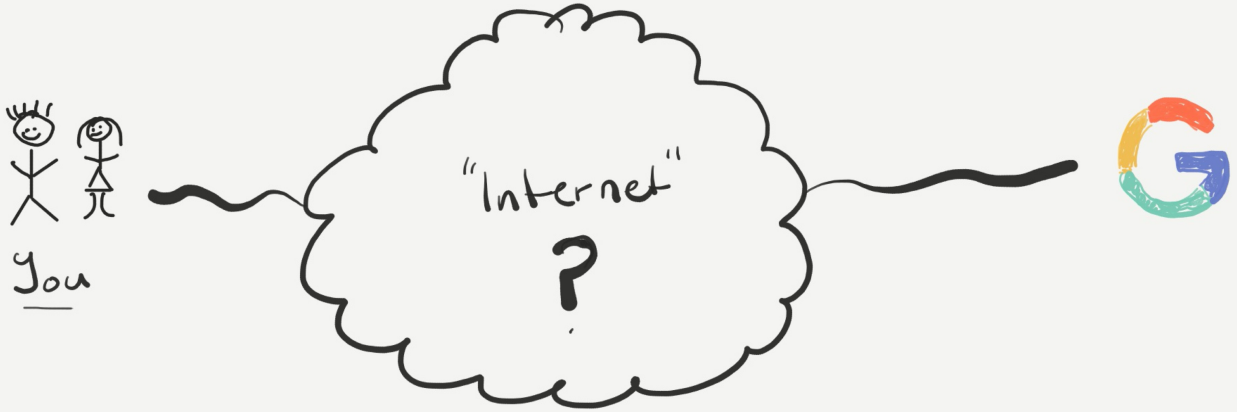
# Agenda

- Introduction to the Internet and IXPs
- IP Allocations, RIPE, BGP and the Global Routing Table
- Introduction to IXPs and INEX
- A Technical Overview of INEX
- Route Servers
- Questions / Discussion

# Setting the Scene - an Internet Primer



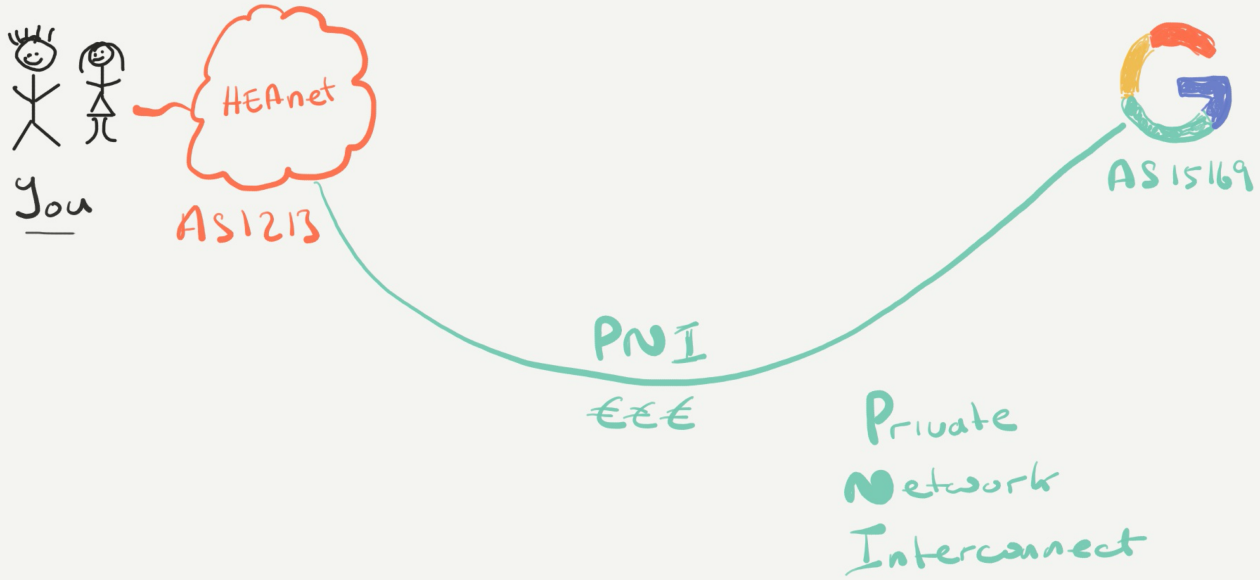

You

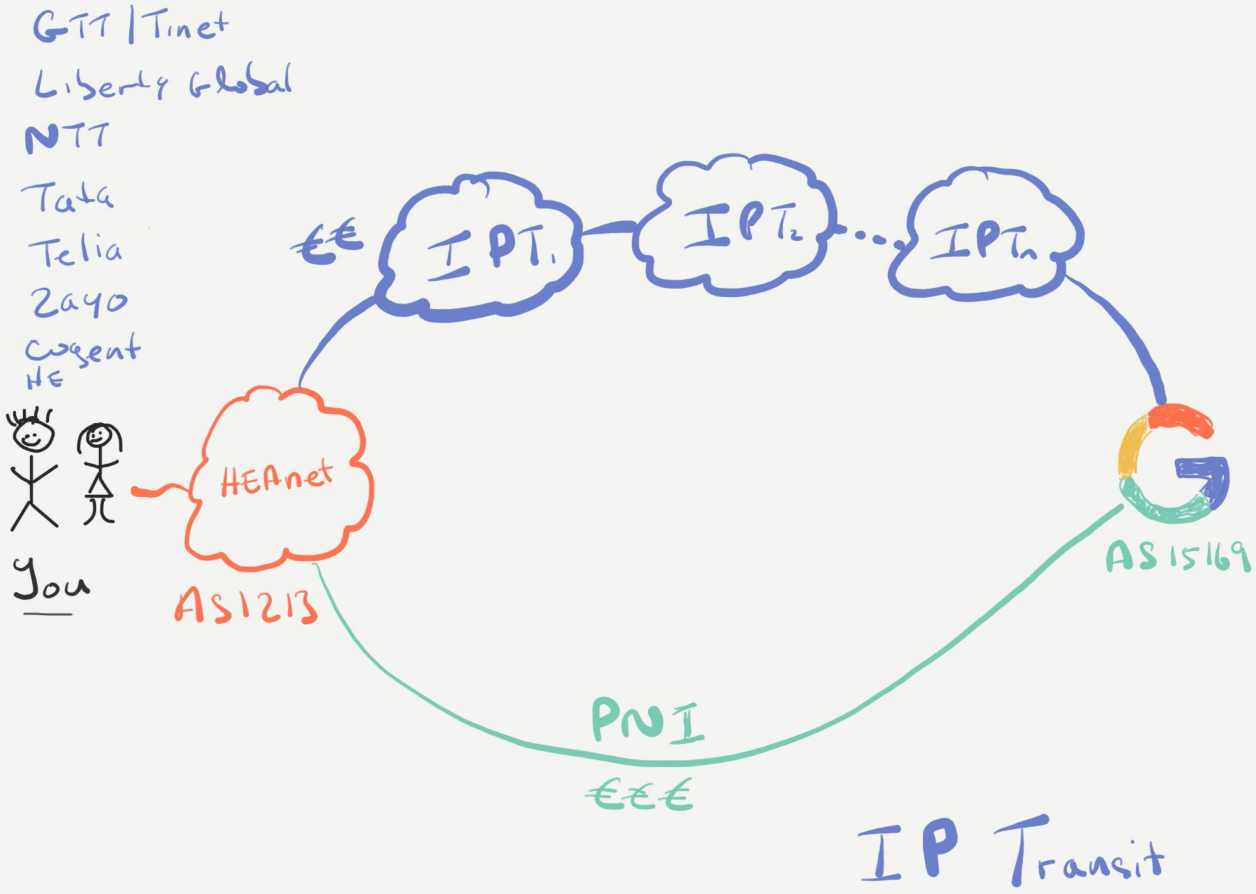

You
"Internet"
?

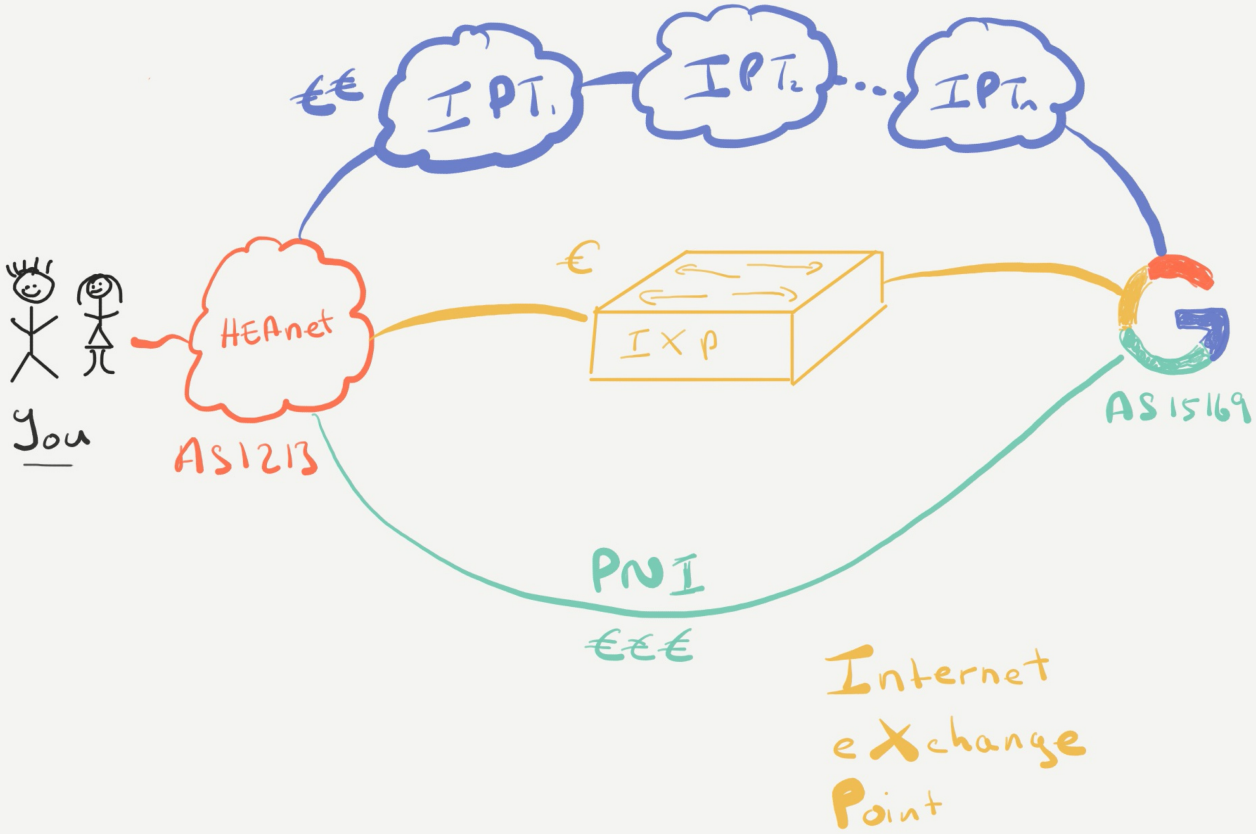
```
barryo@dub-ixn-lab01:~$ mtr -nrc 5 8.8.8.8
Start: 2018-10-25T16:24:09+0100
HOST: dub-ixn-lab01                Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 185.101.240.0               0.0%     5    0.5   0.4   0.3   0.5   0.1
  2.|-- 185.6.36.57                 0.0%     5    1.0   1.4   0.6   2.5   0.7
  3.|-- 74.125.244.1                0.0%     5    2.3   1.9   1.6   2.3   0.3
  4.|-- 72.14.239.219               0.0%     5    1.5   1.5   1.4   1.9   0.2
  5.|-- 8.8.8.8                     0.0%     5    0.7   0.7   0.7   0.7   0.0
barryo@dub-ixn-lab01:~$
```

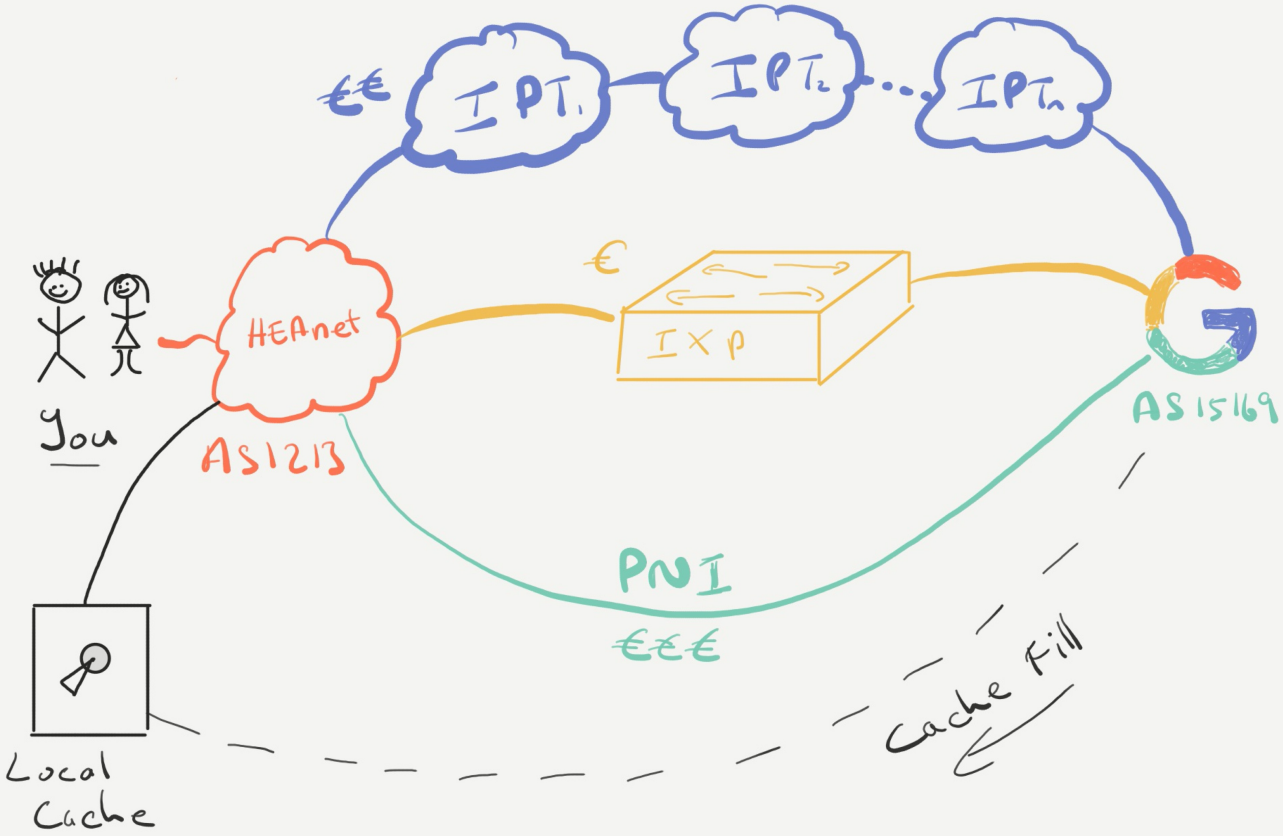

Your Internet Service Provider
HEAnet
You
AS1213

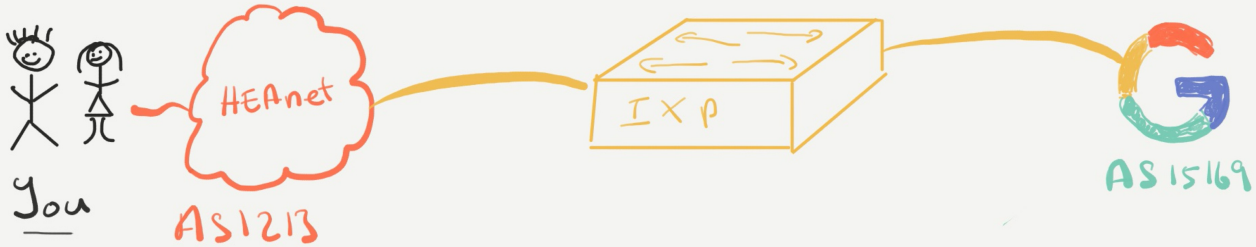

AS15169

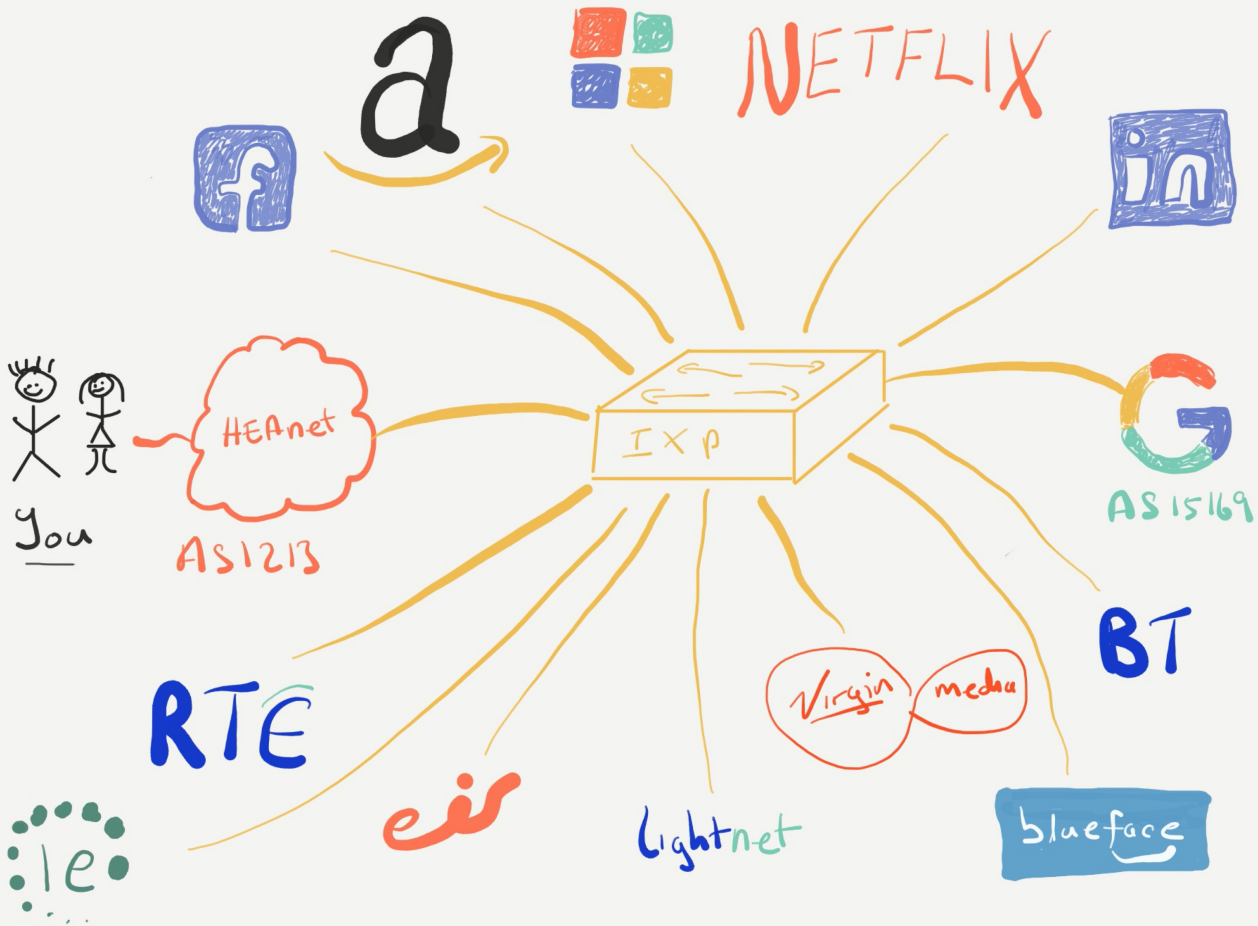

You
HEAnet
AS1213
PNI
€€€
AS15169
Private
Network
Interconnect

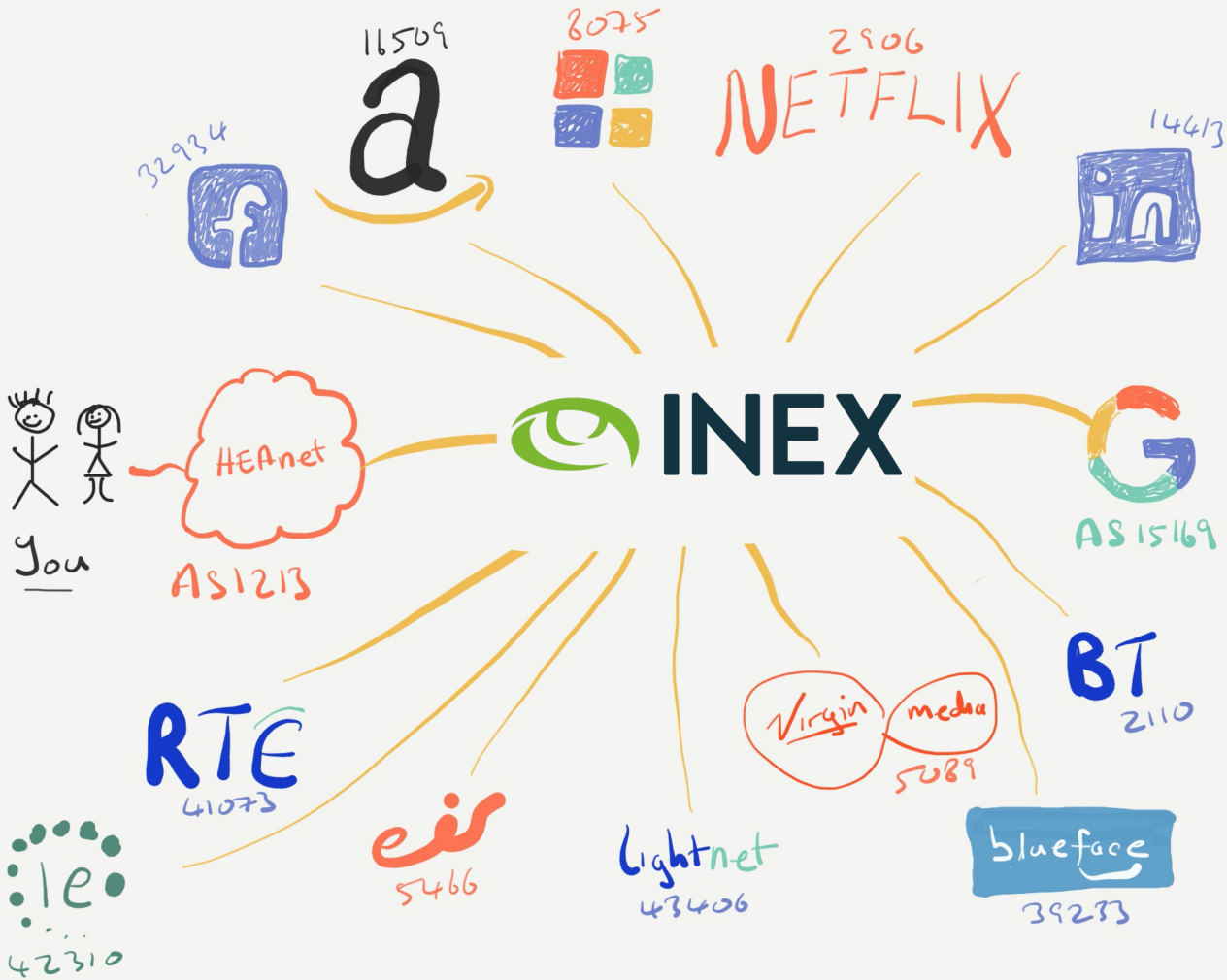

GTT | Tinet
Liberty Global
NTT
Tata
Telia
Zayo
Cogent
HE
You
HEAnet
AS1213
€€
IPT1
IPT2
...
IPTn
AS15169
PNI
€€€
IP Transit


€€
IPT1
IPT2
...
IPTn
€
HEAnet
IXP
You
AS1213
AS15169
PNI
€€€
Internet
eXchange
Point

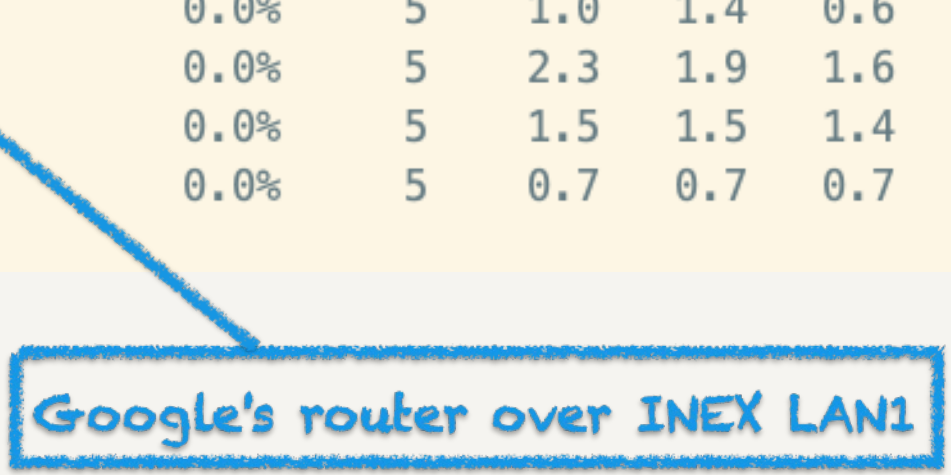

€€
IPT1
IPT2
IPTn
€
IXP
HEAnet
You
AS1213
AS15169
PNI
€€€
Local
Cache
Cache Fill


You
HEAnet
AS1213
IXP
AS15169


NETFLIX
HEAnet
AS1213
You
IXP
AS15169
RTÉ
BT
Virgin
media
eir
lightnet
blueface
ie


16509
8075
2906
NETFLIX
32934
14413
You
HEAnet
AS1213
INEX
AS15169
RTE
41073
Virgin
media
5089
BT
2110
eir
5466
lightnet
43406
blueface
39273
ie
42310

```
barryo@dub-ixn-lab01:~$ mtr -nrc 5 8.8.8.8
Start: 2018-10-25T16:24:09+0100
HOST: dub-ixn-lab01                Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 185.101.240.0               0.0%     5    0.5   0.4   0.3   0.5   0.1
  2.|-- 185.6.36.57                 0.0%     5    1.0   1.4   0.6   2.5   0.7
  3.|-- 74.125.244.1                0.0%     5    2.3   1.9   1.6   2.3   0.3
  4.|-- 72.14.239.219               0.0%     5    1.5   1.5   1.4   1.9   0.2
  5.|-- 8.8.8.8                     0.0%     5    0.7   0.7   0.7   0.7   0.0
barryo@dub-ixn-lab01:~$
```

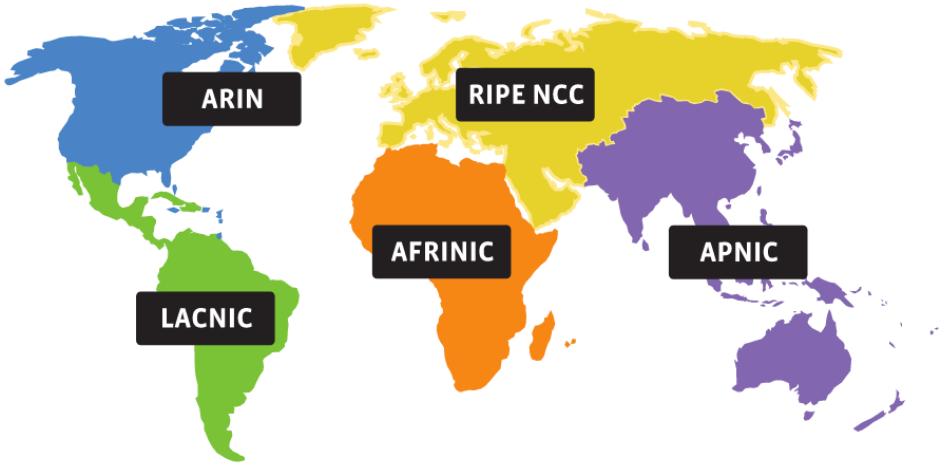


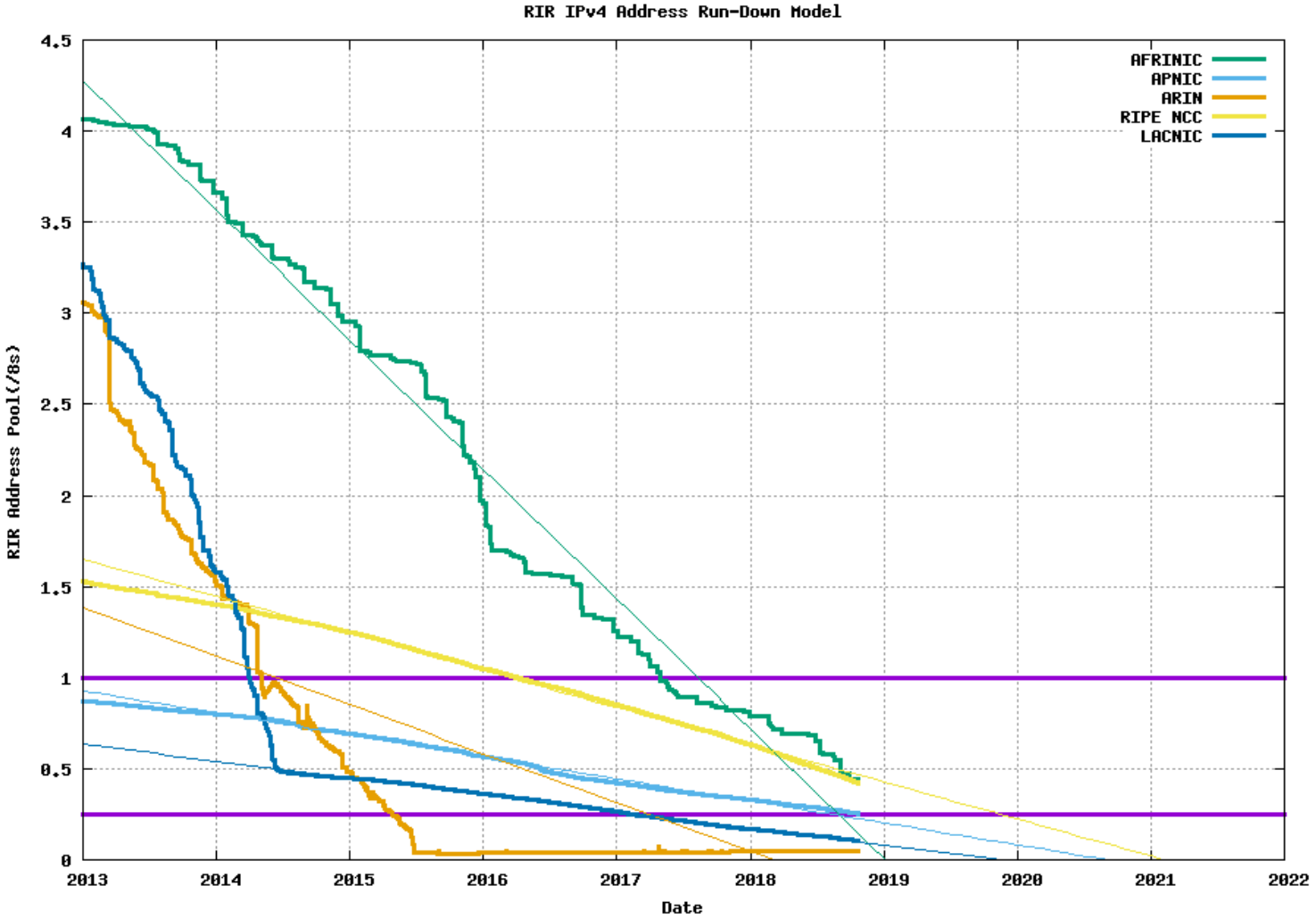

# IP Allocations and Networks

IP ALLOCATIONS

# Who allocates IP addresses?

"coordinate the Internet's globally unique identifiers"




RIR IPv4 Address Run-Down Model
RIR Address Pool(/8s)
Date
AFRINIC
APNIC
ARIN
RIPE NCC
LACNIC
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022

IP ALLOCATIONS

# RIPE - Last /8 Policy

- A /8 IPv4 block is 16.8 million addresses
  - Or 16,384 blocks of 1,024 addresses (/22)
- Each LIR (20k) is entitled to one allocation of /22
- Purpose is to allow for new services, innovation and to allow new entrants into the ISP market while the transition to IPv6 progresses

IP ALLOCATIONS

# HEAnet's Blocks

```
~ dig +short a www.it.nuigalway.ie
140.203.202.90
```

```
~ bgpq3 -b as1213
NN = [
    44.155.0.0/16,
    87.32.0.0/12,
    91.123.224.0/20,
    134.226.0.0/16,
    136.201.0.0/16,
    136.206.0.0/16,
    140.203.0.0/16,
    143.239.0.0/16,
    147.252.0.0/16,
    149.153.0.0/16,
    149.157.0.0/16,
    157.190.0.0/16,
    160.6.0.0/16,
    193.1.0.0/16,
    194.26.0.0/24
];
```

IP ALLOCATIONS

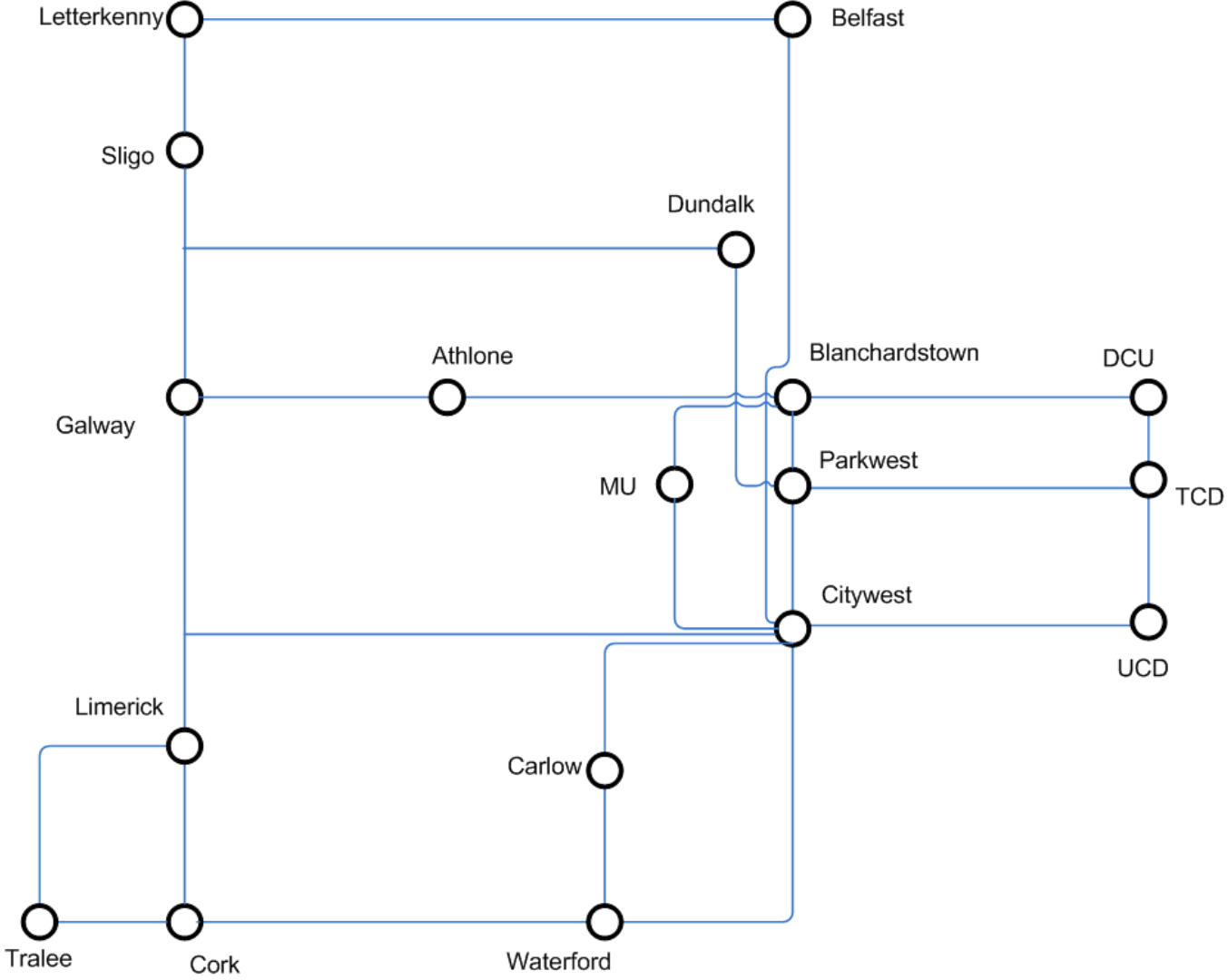
# HEAnet's Blocks - Traceroute to 140.203.202.90

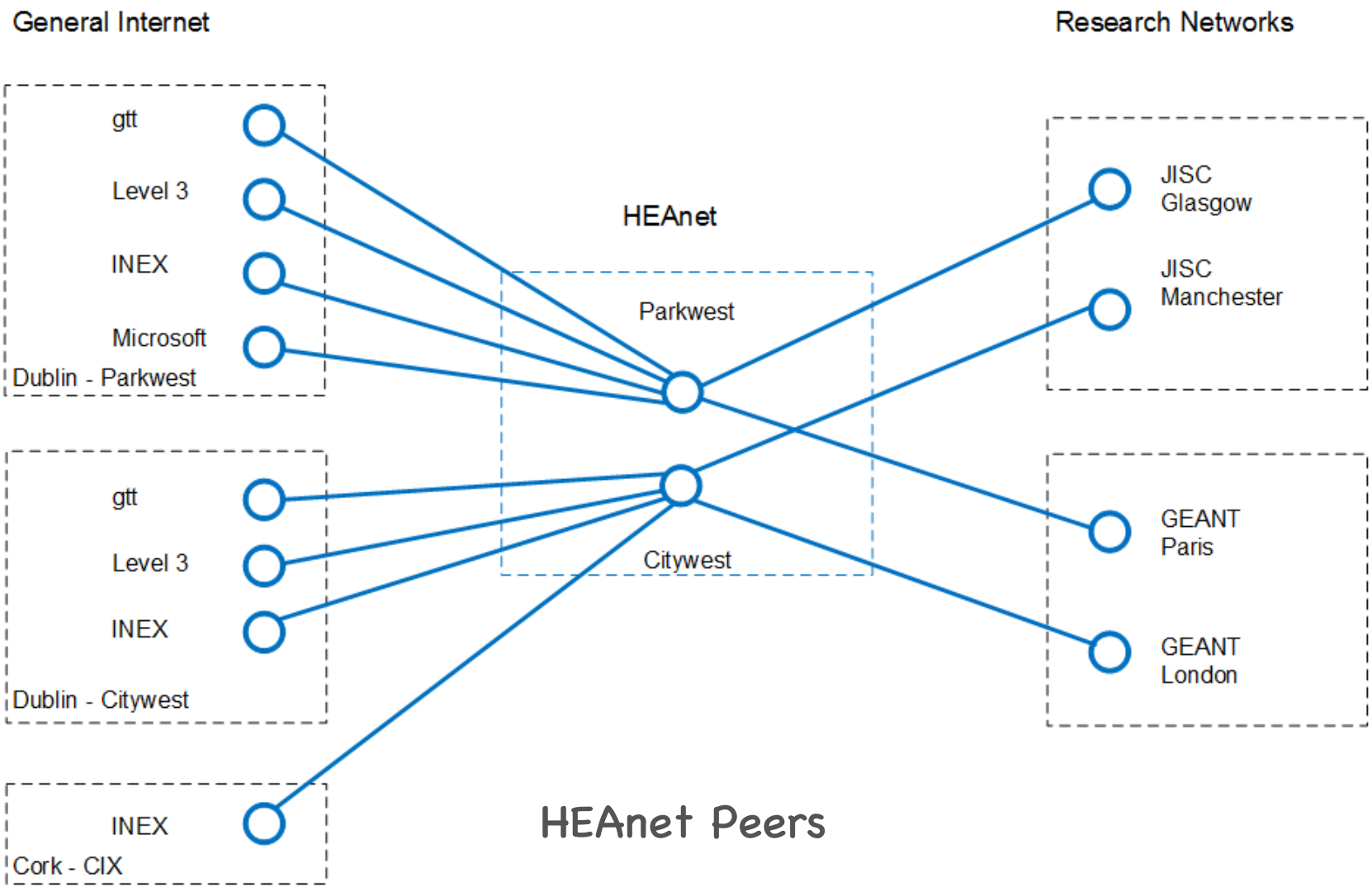
IP ALLOCATIONS

# HEAnet's Blocks - Traceroute to 140.203.202.90

```
Start: 2018-10-25T20:01:43+0100
HOST: barryo-mac1.local          Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 192.168.140.2            30.0%    10    0.3   0.3   0.2   0.3   0.0
  2.|-- 31.187.45.1              30.0%    10   11.3   9.7   7.6  12.0   1.7
  3.|-- 109.255.254.29            0.0%    10   12.6   9.4   6.3  12.6   1.9
  4.|-- 84.116.238.70             0.0%    10    8.9  14.3   6.5  37.0   9.8
  5.|-- 84.116.134.174            0.0%    10    7.9   8.7   7.4  11.3   1.2
  6.|-- 194.88.240.15             0.0%    10    8.1  11.0   7.6  23.0   4.8
  7.|-- 87.44.50.5                0.0%    10   13.7  12.3   8.7  21.8   4.6
  8.|-- 87.44.50.72               0.0%    10   13.2  12.6   9.8  16.0   2.1
  9.|-- 87.44.50.79               0.0%    10   12.6  24.5  10.9 120.7  34.0
 10.|-- 87.44.56.38               0.0%    10   11.3  25.6  11.3 123.7  34.5
 11.|-- 140.203.202.90            0.0%    10   12.6  26.3  10.5 125.2  35.0
```

# HEAnet National Backbone




General Internet
Research Networks
gtt
Level 3
INEX
Microsoft
Dublin - Parkwest
HEAnet
Parkwest
JISC
Glasgow
JISC
Manchester
gtt
Level 3
INEX
Dublin - Citywest
Citywest
GEANT
Paris
GEANT
London
INEX
Cork - CIX
HEAnet Peers

IP ALLOCATIONS

# The "Internet"

140.203.0.0/16,

- Each IP block / prefix is routed by a single network (AS)
  - A network may have many IP prefixes
  - Each network is described by an AS number

IP ALLOCATIONS

# The "Internet"

- Each IP block is routed by a single network
  - A network may have many IP blocks (but not vice versa)
  - Each network is described by an AS number

Within the Internet, an autonomous system (AS) is a collection of Internet Protocol (IP) routing prefixes (address blocks) under the control of an Internet Service Provider (ISP) [that presents a common, clearly defined routing policy to the Internet].

IP ALLOCATIONS

# The "Internet"

- Each IP block is routed by a single network
  - A network may have many IP blocks (but not vice versa)
  - Each network is described by an AS number

Within the Internet, an autonomous system (AS) is a collection of Internet Protocol (IP) routing prefixes (address blocks) under the control of an Internet Service Provider (ISP) [that presents a common, clearly defined routing policy to the Internet].

- So, the internet is ***a network of interconnected networks***


# Border Gateway Protocol

BGP

# BGP - Border Gateway Protocol

- BGP – 1989 (RFC1105)
  - BGP-2 – 1990 (RFC1163)
  - BGP-3 – 1991 (RFC1267)
  - BGP-4 – 1995 (RFC1654, 1771, 4271, …)

- AS – Autonomous System: a network managed by a single entity; uniquely identified by an AS number (ASN)

- BGP is an EGP – Exterior Gateway Protocol
  - Sets up inter-AS routing - routing between different AS'
  - IGPs are used for intra-AS routing - routing within an AS

BGP

# BGP - Border Gateway Protocol

- BGP is a routing protocol that allows one network (AS) to signal to other networks the IP prefixes that can be reached *through* it

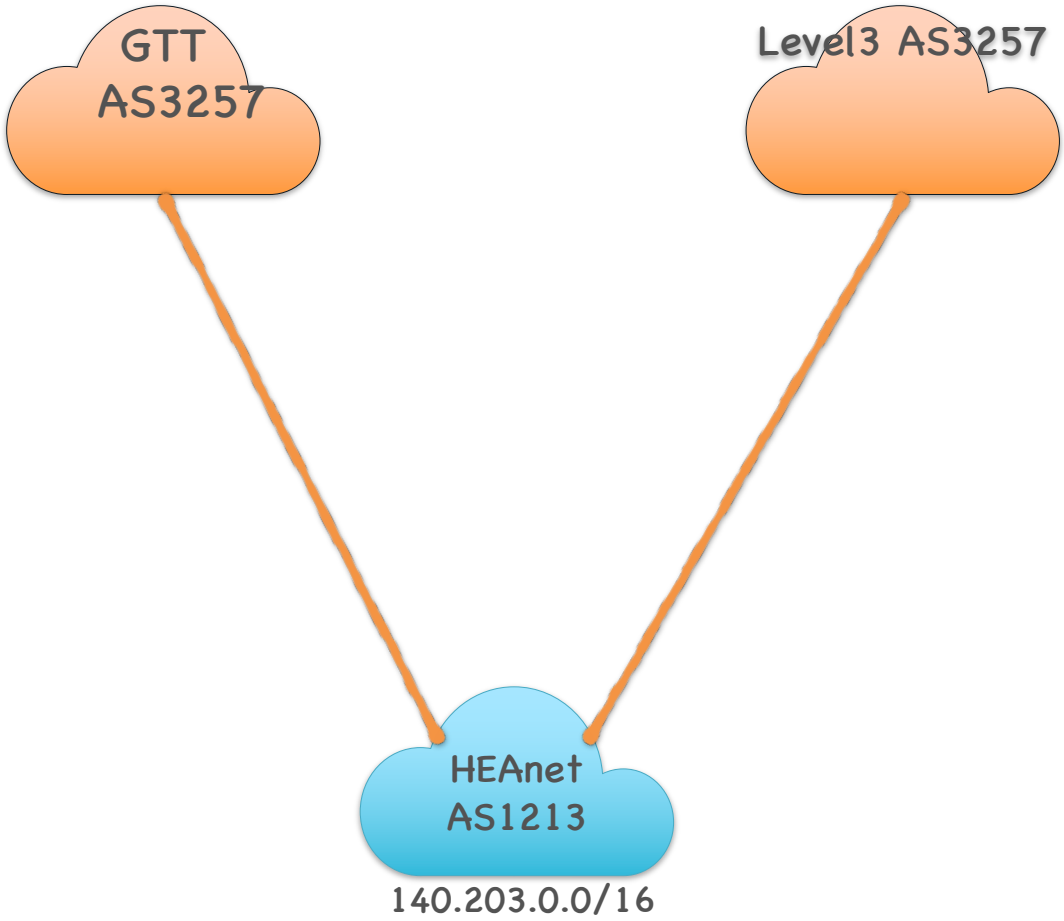
- These relationships are called peers / neighbors. Three main types:
  - Transit - you are the customer to an upstream ISP
  - Peerings - normally settlement free; IXPs or PNIs
  - Customers - you are the ISP providing downstream connectivity

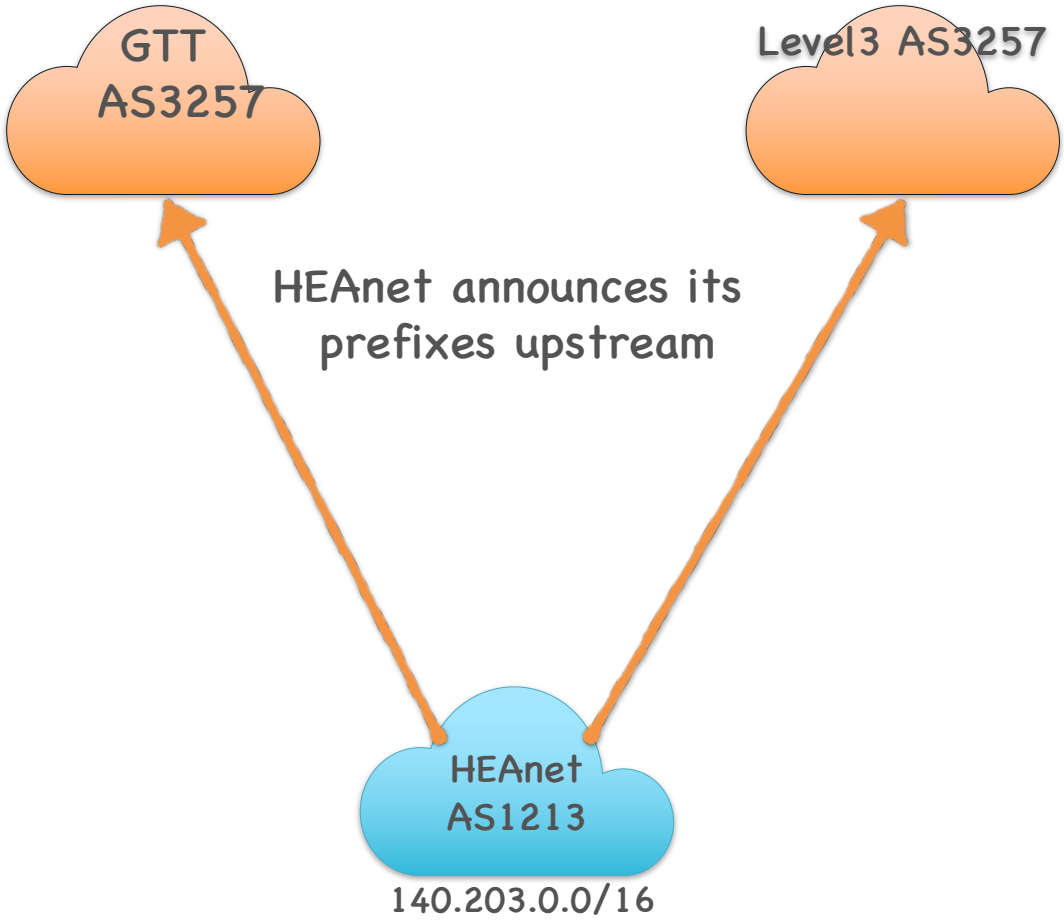
- Default route - gateway of last resort
- Default Free Zone (DFZ) - the 'full' routing table, no default route

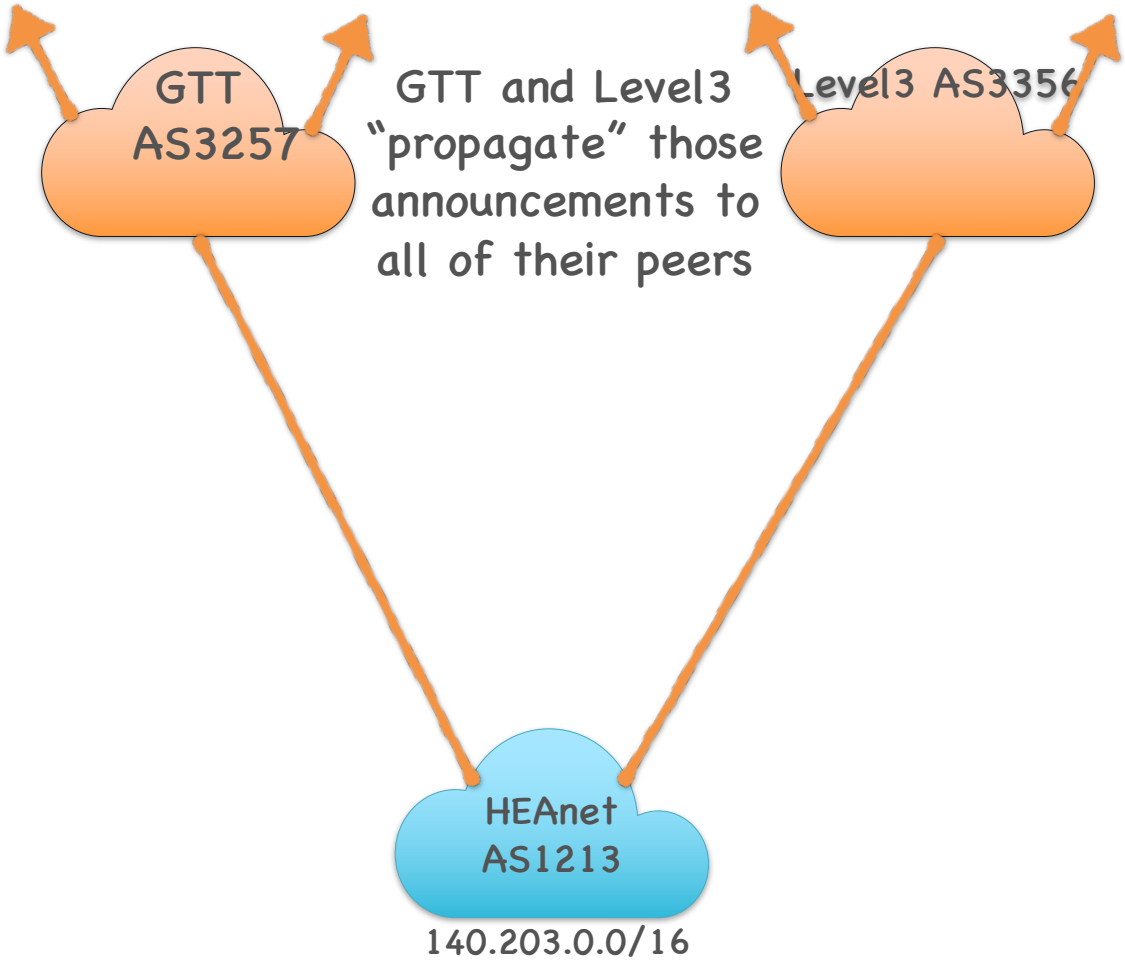
BGP

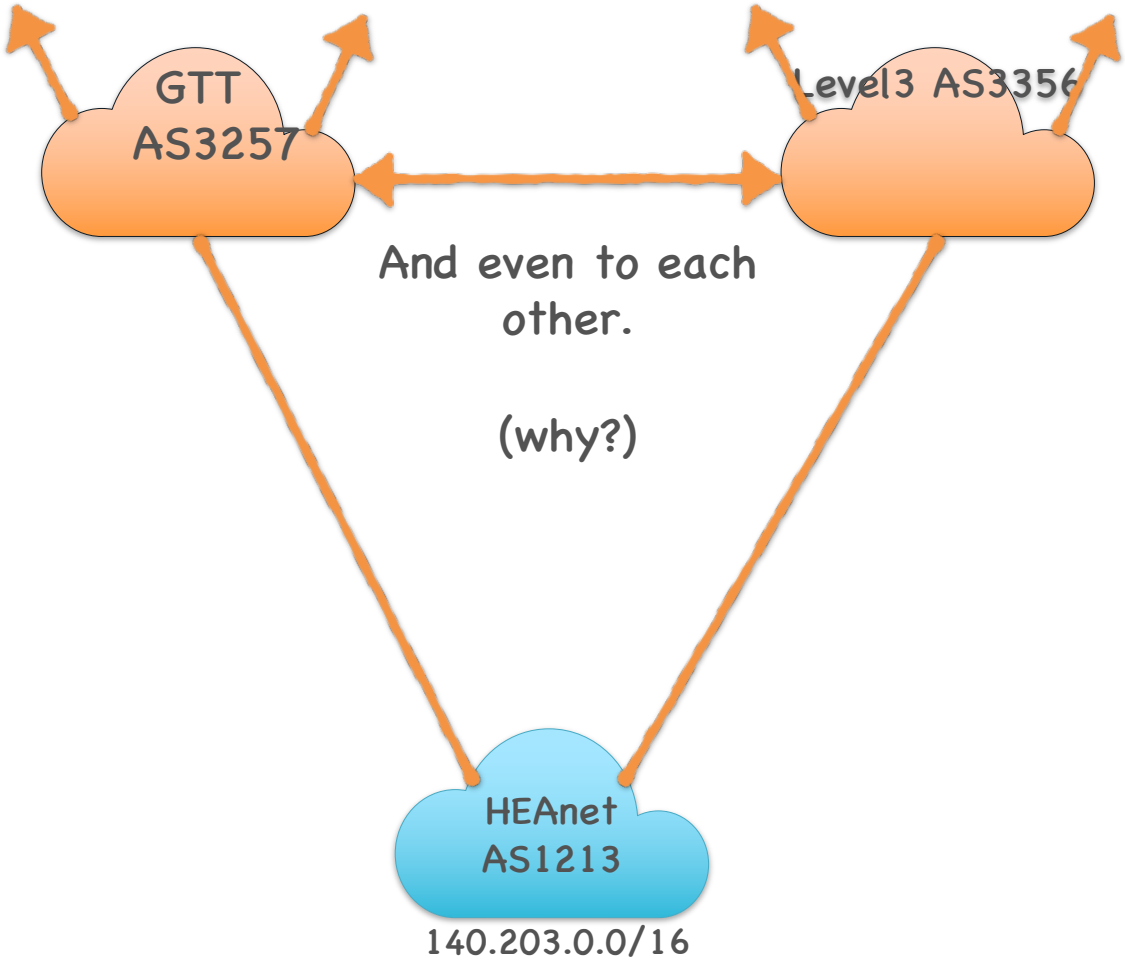
# Prefix Propagation

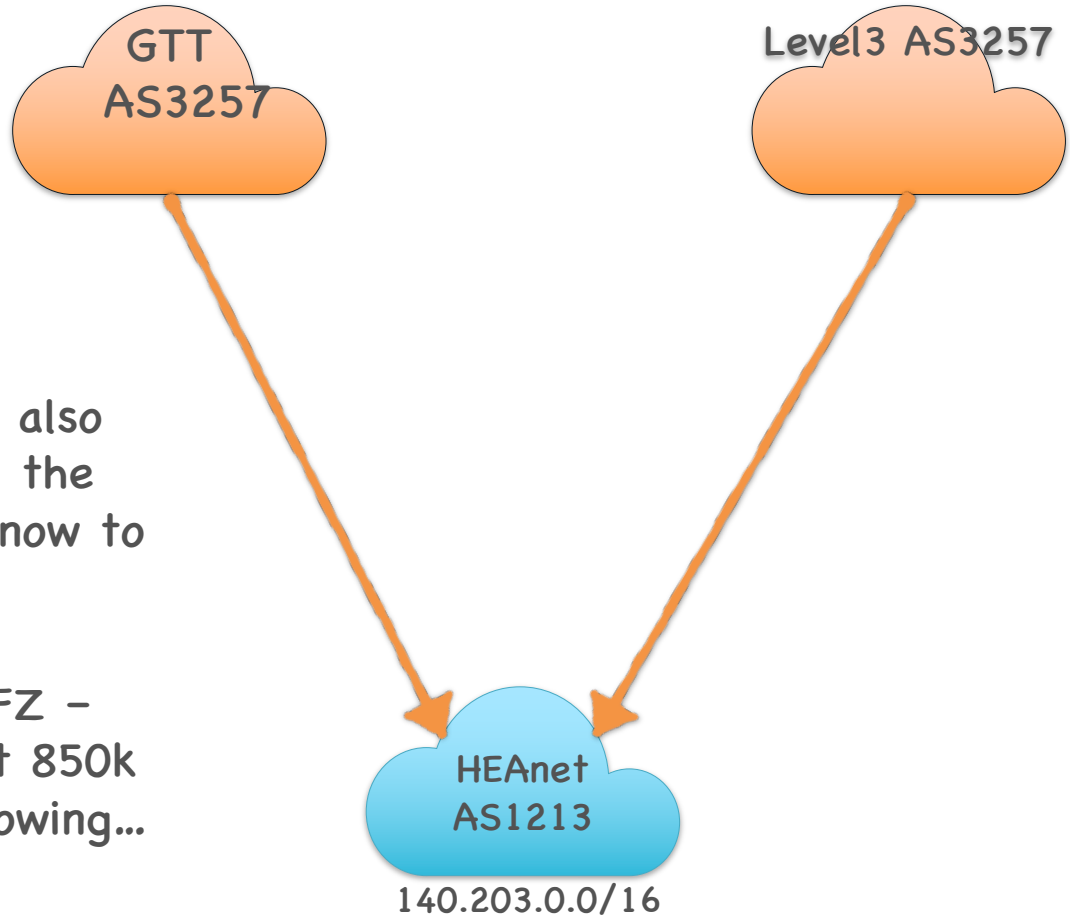
How does every other network on the internet learn what prefixes my network (AS) has and (b) how to get to my network?


GTT
AS3257
Level3 AS3257
HEAnet
AS1213
140.203.0.0/16


GTT
AS3257
Level3 AS3257
HEAnet announces its
prefixes upstream
HEAnet
AS1213
140.203.0.0/16

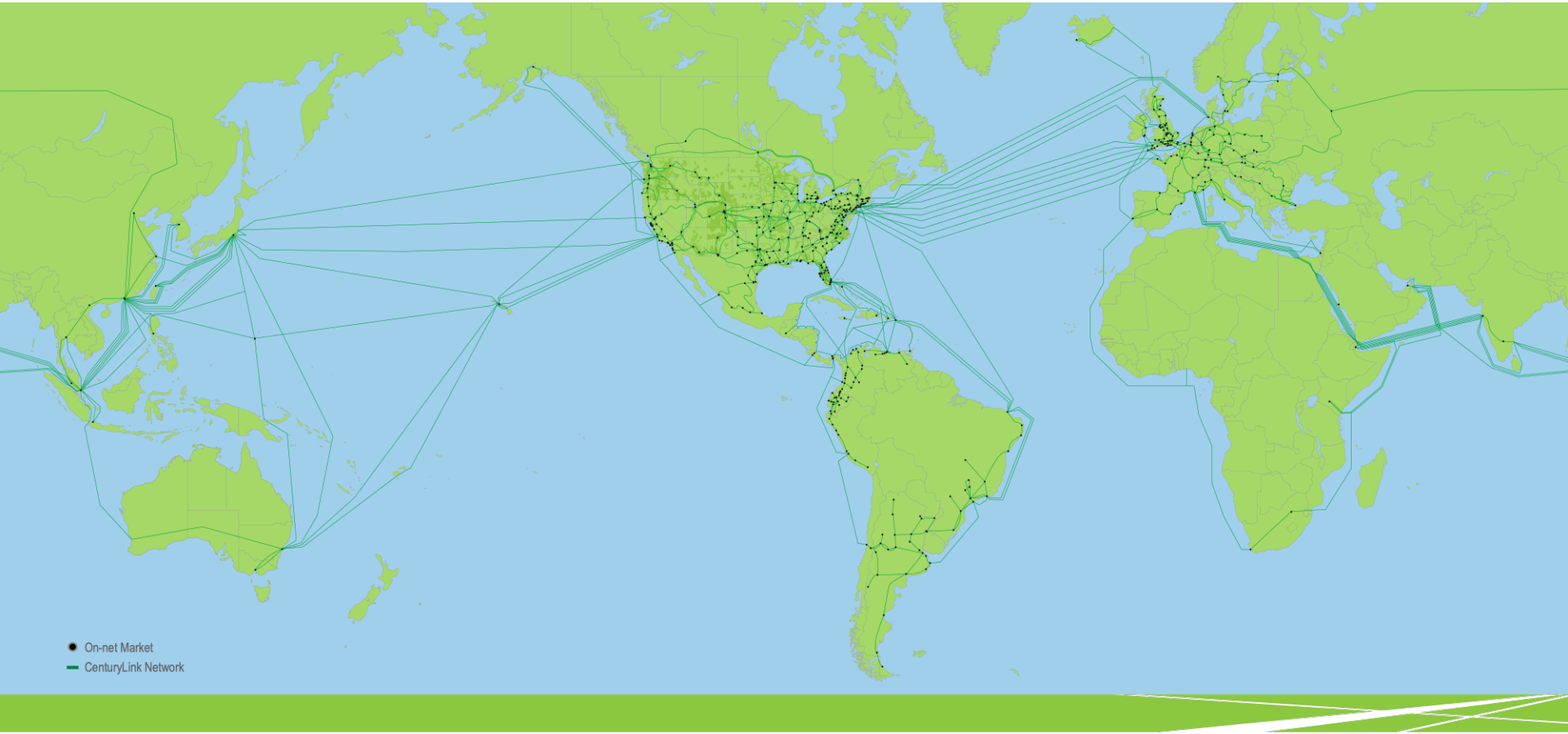

GTT
AS3257
GTT and Level3 "propagate" those announcements to all of their peers
Level3 AS3356
HEAnet
AS1213
140.203.0.0/16

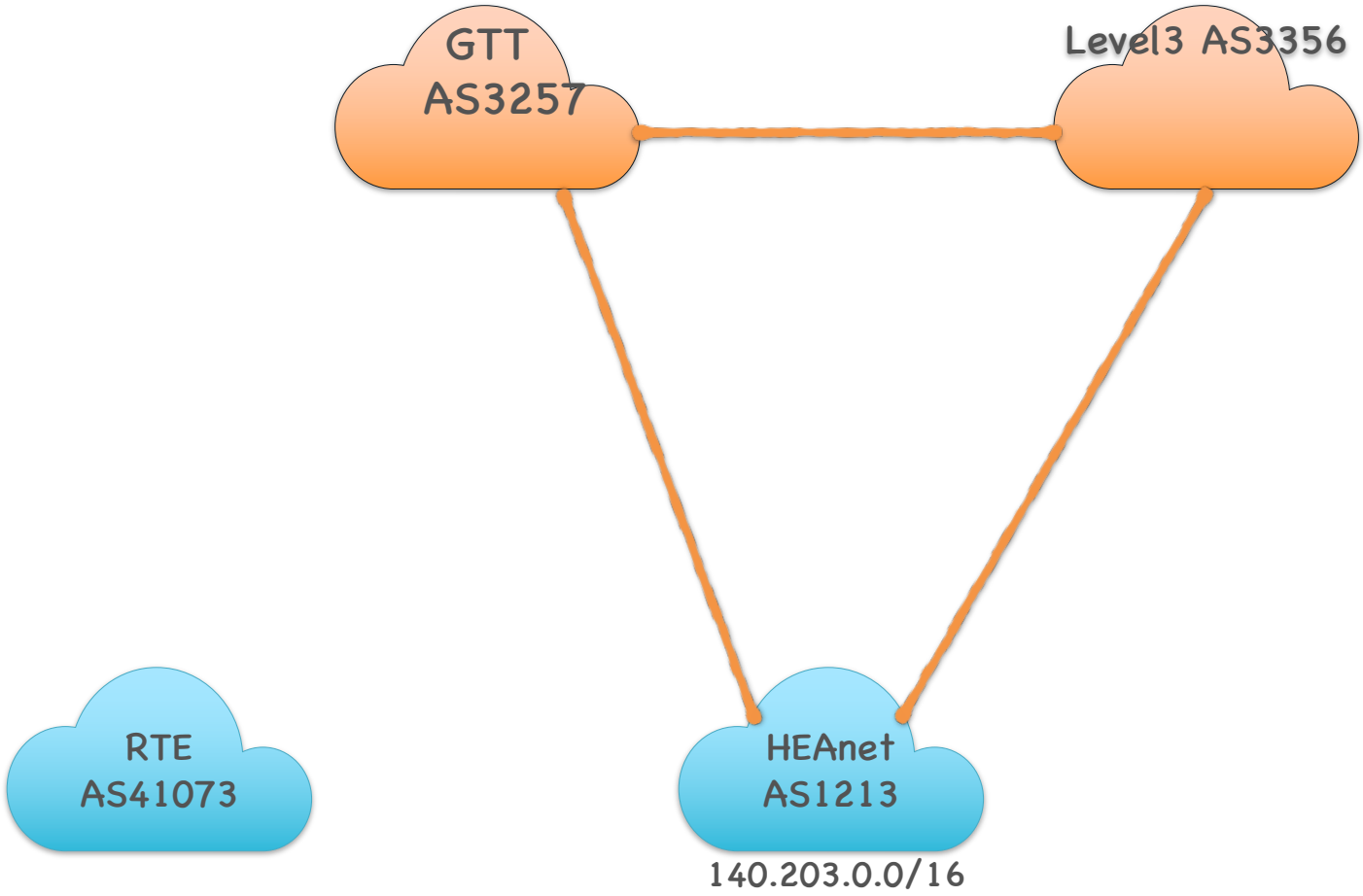

GTT
AS3257
Level3 AS3356
And even to each other.
(why?)
HEAnet
AS1213
140.203.0.0/16

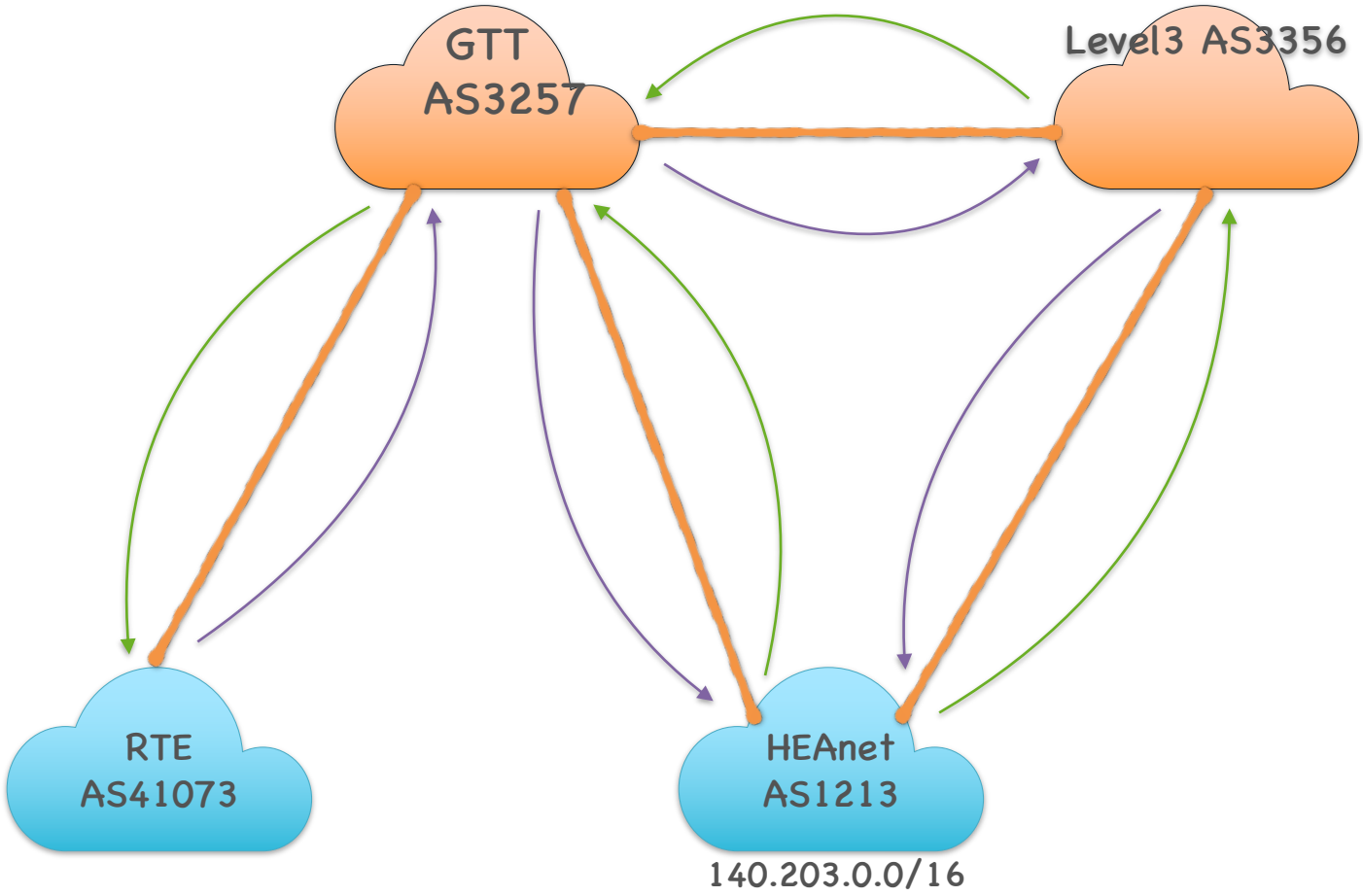


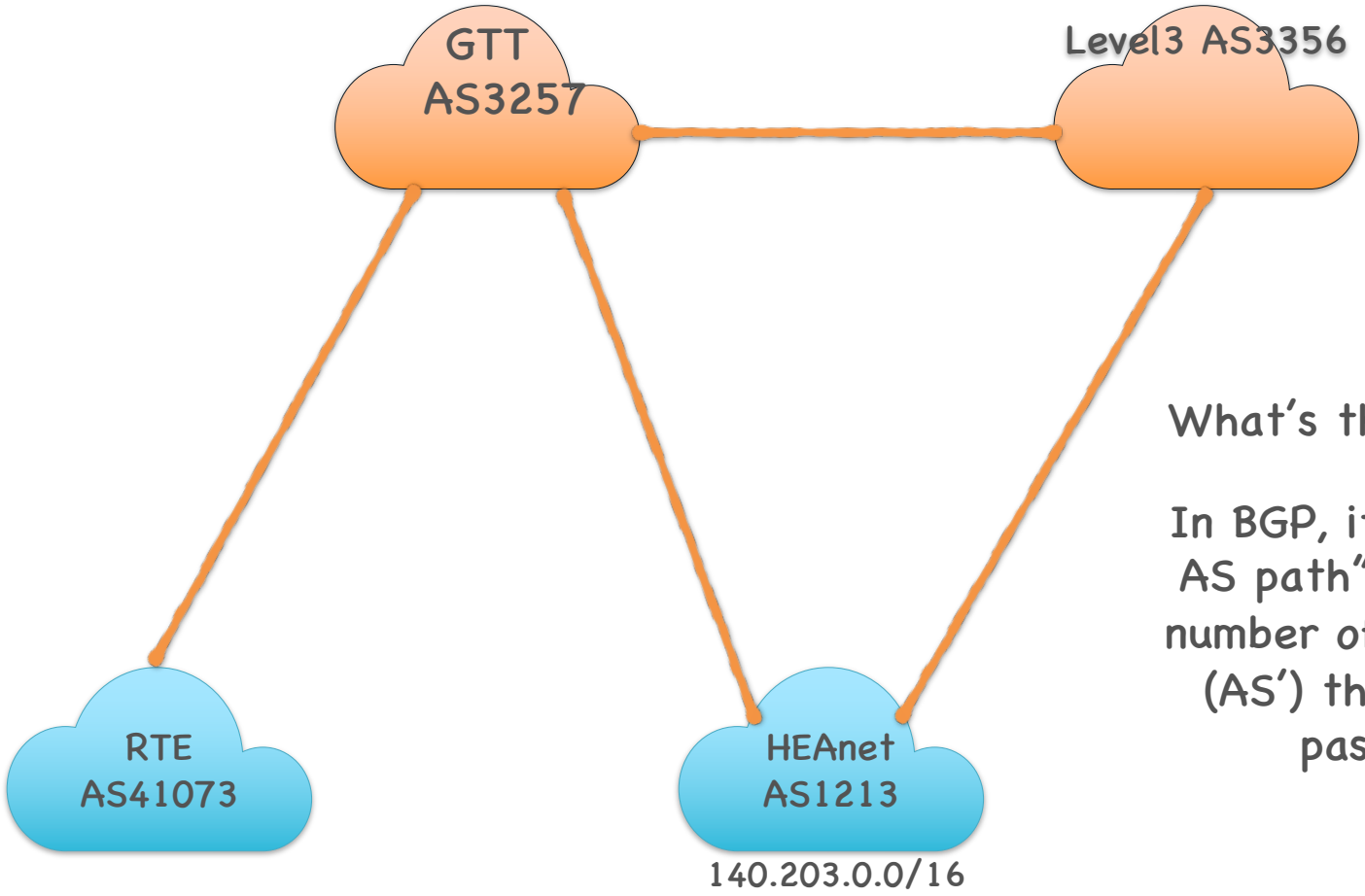

GTT & Level3 also propagate all the prefixes they know to HEAnet.

This is the DFZ - currently about 850k prefixes and growing...


On-net Market
CenturyLink Network

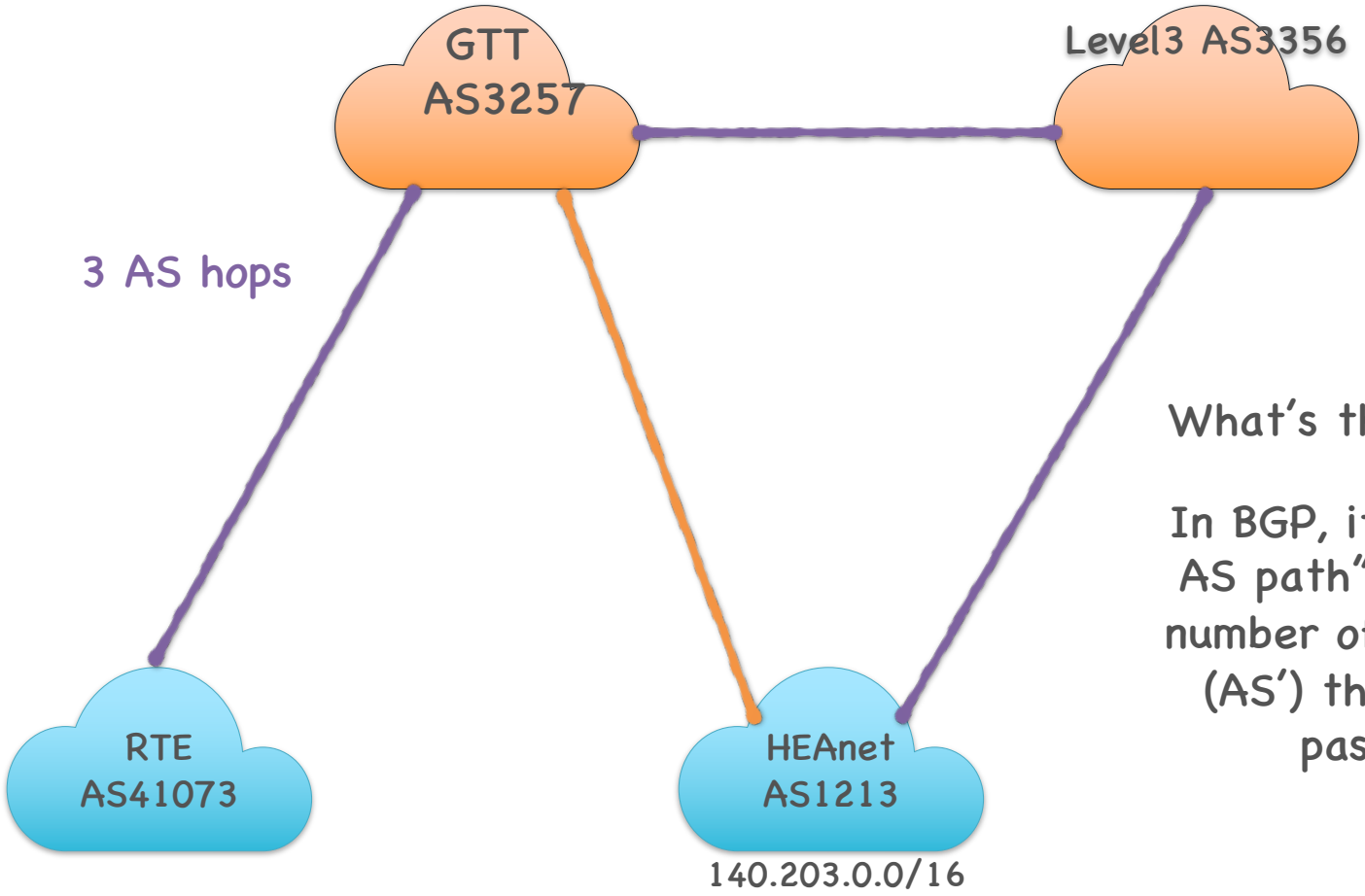

GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
140.203.0.0/16


GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
140.203.0.0/16



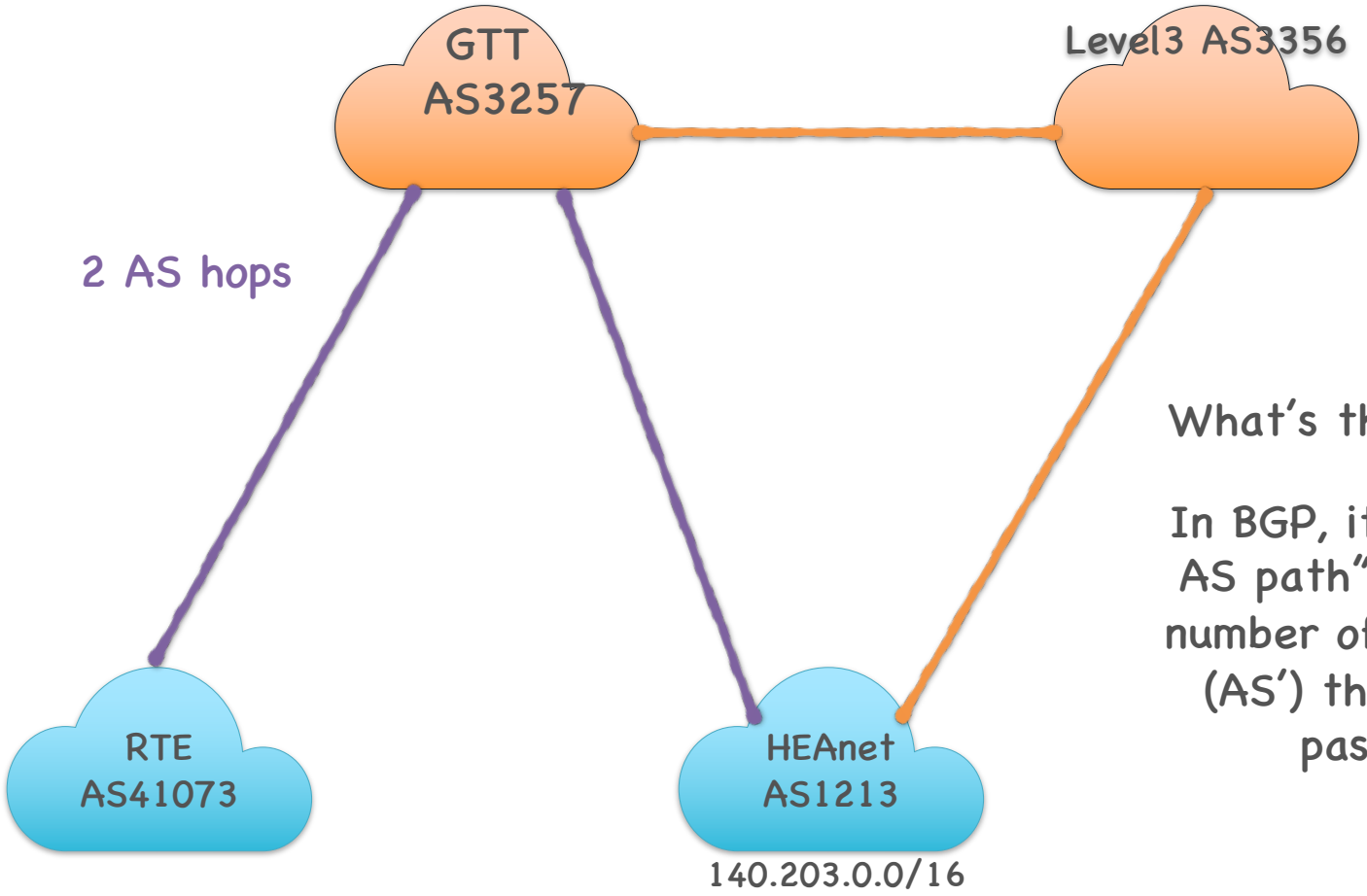

What's the quickest path?

In BGP, it's the "shortest AS path" - i.e. the least number of other networks (AS') that you have to pass through.



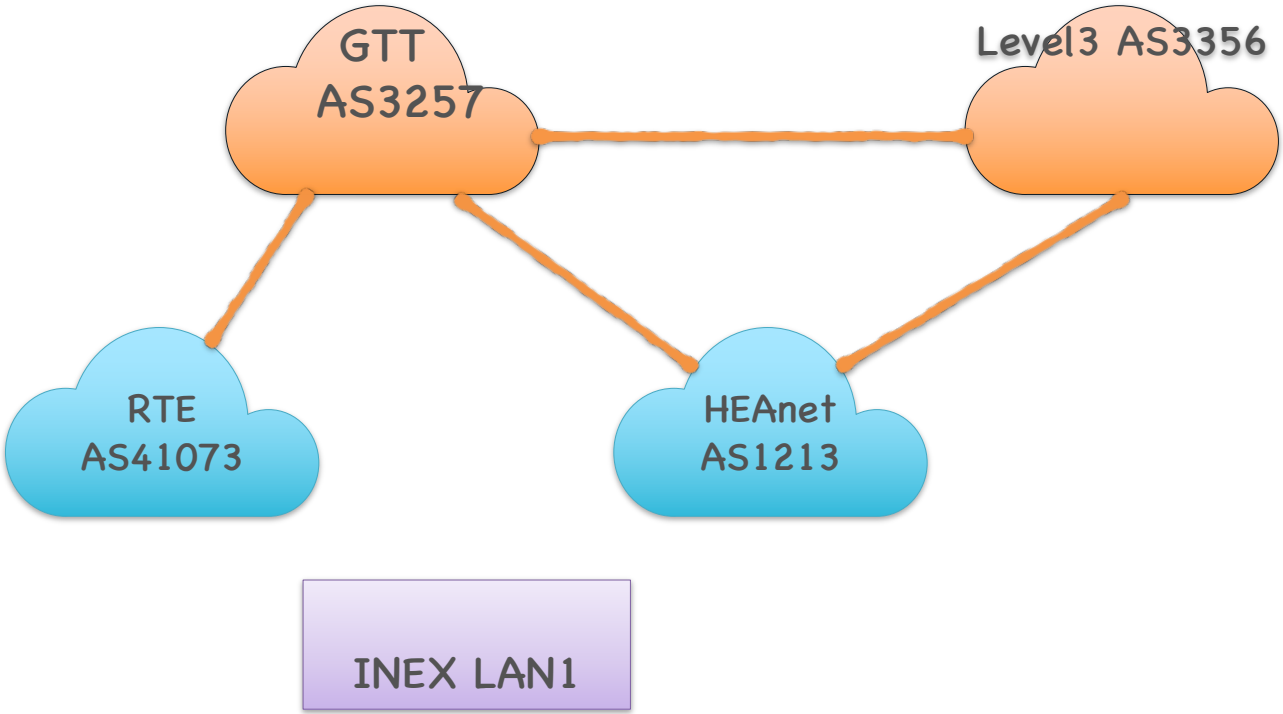

What's the quickest path?

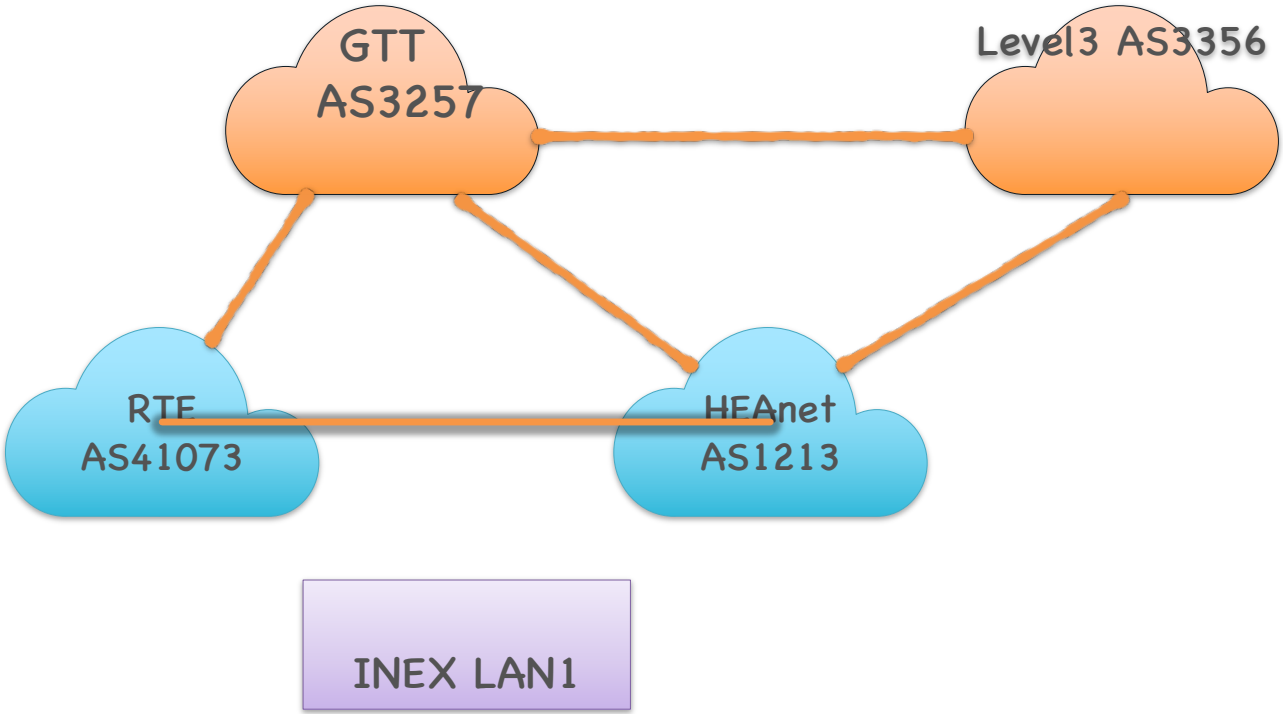
In BGP, it's the "shortest AS path" - i.e. the least number of other networks (AS') that you have to pass through.

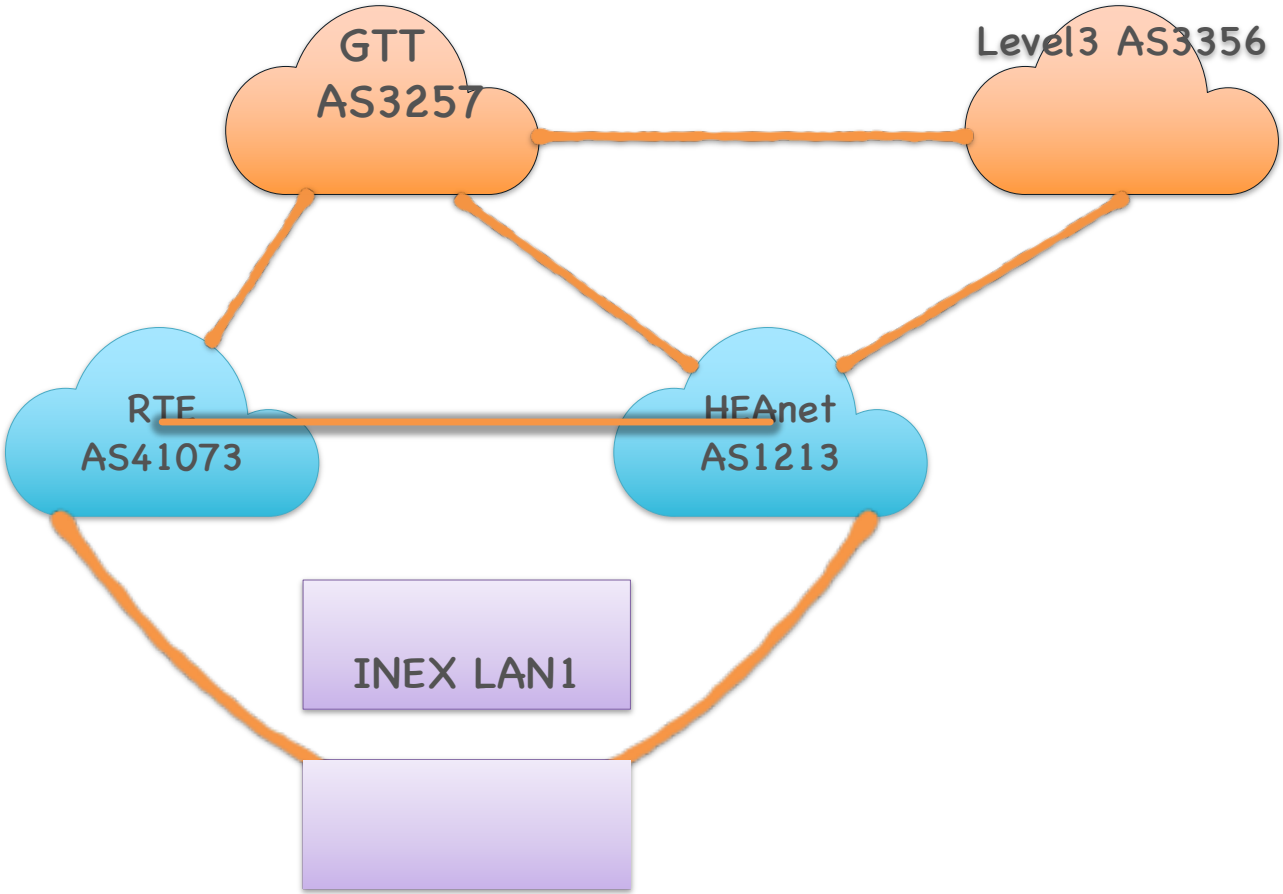


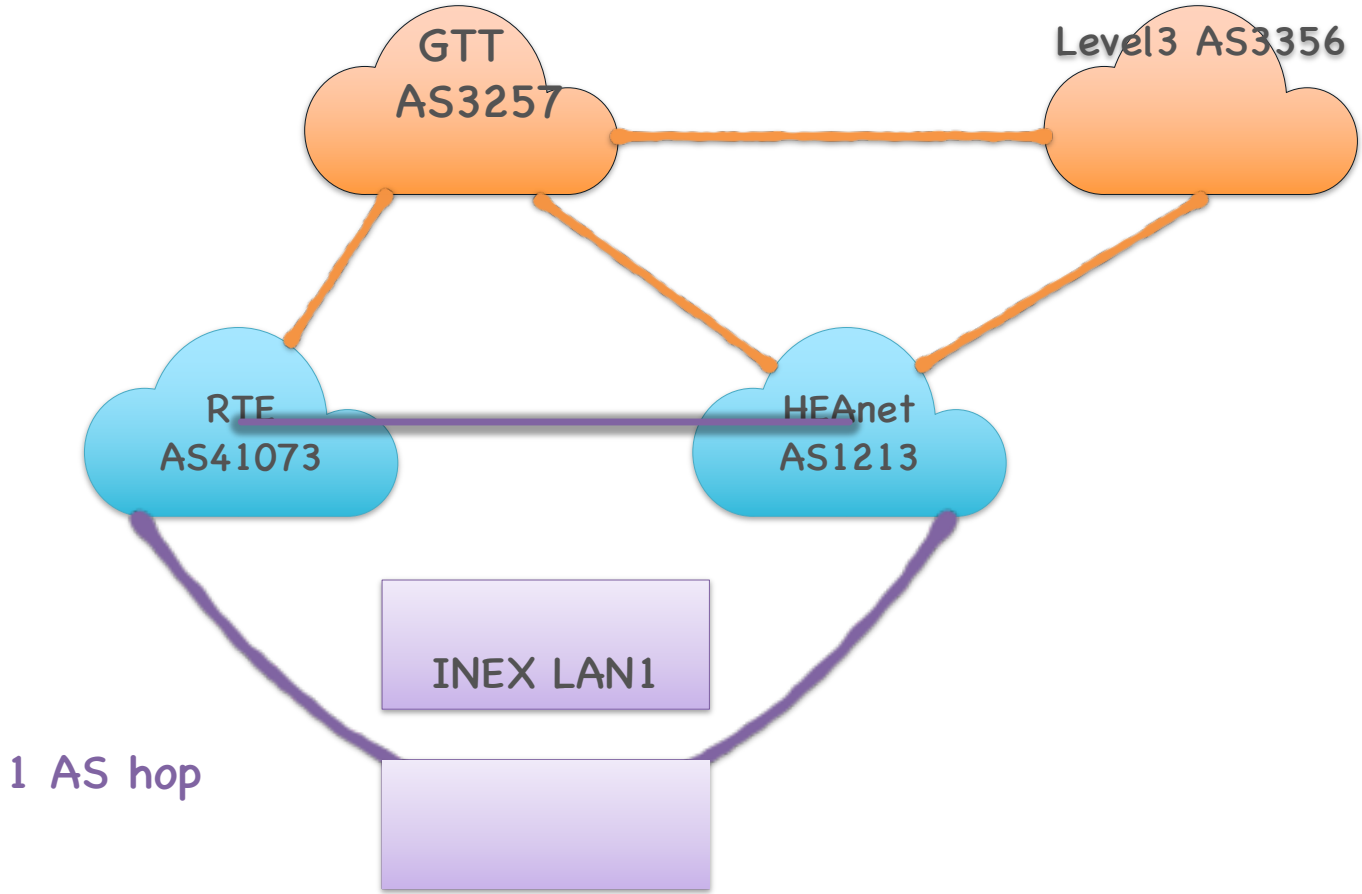

What's the quickest path?

In BGP, it's the "shortest AS path" - i.e. the least number of other networks (AS') that you have to pass through.


GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
INEX LAN1


GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
INEX LAN1


GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
INEX LAN1


GTT
AS3257
Level3 AS3356
RTE
AS41073
HEAnet
AS1213
INEX LAN1

1 AS hop

BGP

# What We Will Look At

- Configuring a BGP session - step by step
- Securing a BGP session
- Route (Best Path) Selection Algorithm
- Routing Examples
- Traffic Engineering
  - Local preferences
  - MEDs
  - AS path prepending

BGP

# What We Will Not Look At

- iBGP
- Multihop eBGP
- IGPs and Redistribution
- Protocol Internals
- Route Reflectors
- Communities
- Examples Will be IPv4 Only
- Examples Will be Mikrotik

BGP

# Ingredients for a BGP Session

- Layer 2 connectivity between routers (physical link, IXP, …)
- Layer 3 subnet for IP communication
  - E.g. 194.88.240.0/25 (INEX LAN2)
  - Typically a /30 for single router IPT
- Prefixes (routes) to advertise over BGP
- AS number

BGP

# Ingredients for a BGP Session

- Security
  - Inbound prefix filters
  - Outbound prefix filters
  - *AS path filters*
  - MD5 shared secret
  - Maximum prefixes
  - *Next hop verification*

BGP

# Ingredients for a BGP Session

- Our ASN: 65550
- We advertise:
  - 192.0.2.0/24
- Peering Network:
  - 203.0.113.0/30
  - Our end: 203.0.113.2

- Peer ASN: 65551
- They advertise:
  - DFZ
- Peering Network:
  - 203.0.113.0/30
  - Their end: 203.0.113.1

BGP

# Configure Our Router

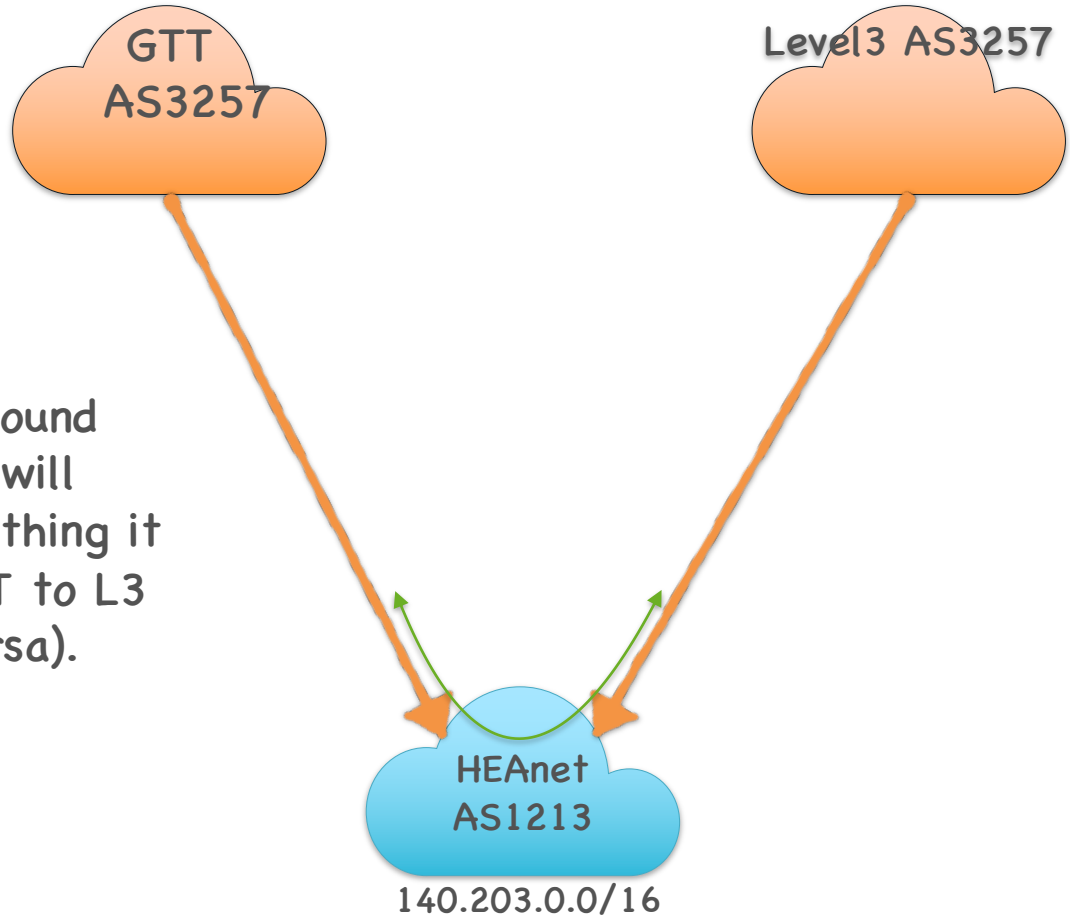
```
/ip address add address=203.0.113.2/30 interface=ether1

; ping the other end:

/ping 203.0.113.1 count=1
  SEQ HOST                                     SIZE TTL TIME  STATUS
    0 203.0.113.1                                56  64 0ms
    sent=1 received=1 packet-loss=0% min-rtt=0ms avg-rtt=0ms max-rtt=0ms

; null route for our prefix and the default route

/ip route add dst-address=192.0.2.0/24 type=blackhole
/ip route add dst-address=0.0.0.0/0    type=blackhole distance=250
```

BGP

# Configure Our Router

```
/routing bgp instance
set default as=65550 redistribute-static=yes router-id=192.0.2.0

/routing bgp peer
add name=mypeer1 nexthop-choice=force-self remote-address=203.0.113.1 \
    remote-as=65551 update-source=ether1
```

Without outbound filters, BGP will propagate everything it learns from GTT to L3 (and vice versa).



BGP

# Configure Our Router - Security

```
/routing filter
add action=accept  chain=ebgp-out prefix=192.0.2.0/24
add action=discard chain=ebgp-out


/routing bgp peer
set [ find name=mypeer1 ] out-filter=ebgp-out
```

BGP

# Configure Our Router - Security

```
/routing filter
add action=discard chain=sane-in prefix=192.168.0.0/16 prefix-length=16-32
…
add action=discard chain=sane-in prefix=192.0.2.0/24 prefix-length=24-32
…
add action=discard chain=sane-in prefix=0.0.0.0/0
…
add action=discard chain=sane-in prefix=0.0.0.0/0 prefix-length=25-32
add action=accept chain=sane-in


/routing bgp peer
set [ find name=mypeer1 ] in-filter=sane-in
```

BGP

# Configure Our Router - Security

```
/routing bgp peer
set [ find name=smallpeer1 ] max-prefix-limit=50 max-prefix-restart-time=10m
```

BGP

# Best Path Selection Algorithm (on prefix length)

- Prefer the path with the highest LOCAL_PREF (def: 100)
- Prefer the path with the shortest AS_PATH
- Prefer the path with the lowest MED
- Prefer the oldest path

Tie-breakers:

- Prefer the path from the router with the lower router-id
- Prefer the path that comes from the lowest neighbor address

(vendor specific and more esoteric decisions omitted)

BGP

# Best Path Selection Algorithm (on prefix length)

- **Prefer the path with the highest LOCAL_PREF (def: 100)**
- **Prefer the path with the shortest AS_PATH**
- **Prefer the path with the lowest MED**
- Prefer the oldest path

Tie-breakers:

- Prefer the path from the router with the lower router-id
- Prefer the path that comes from the lowest neighbor address

Typical default decision. What you can effect.

BGP

# Best Path Selection Algorithm - Example

```
rtr01#show bgp ipv4 unicast 140.203.0.0/16

BGP routing table entry for 140.203.0.0/16

Paths: (3 available, best #x)
  174 3356 1213 1213
    154.50.192.49 from 154.50.192.49 (154.26.32.227)
      Localpref 100, valid, external
  1213
    194.88.240.15 from 194.88.240.8 (194.88.240.8)
      Localpref 400, valid, external
  1213
    83.220.203.172 from 83.220.203.172 (83.220.203.170)
      Localpref 300, valid, internal
```

BGP

# Best Path Selection Algorithm - Example

```
rtr01#show bgp ipv4 unicast 140.203.0.0/16

BGP routing table entry for 140.203.0.0/16

Paths: (3 available, best #x)
 174 3356 1213 1213
    154.50.192.49 from 154.50.192.49 (154.26.32.227)
      Localpref 100, valid, external
 1213
    194.88.240.15 from 194.88.240.8 (194.88.240.8)
      Localpref 400, valid, external
 1213
    83.220.203.172 from 83.220.203.172 (83.220.203.170)
      Localpref 300, valid, internal
```

BGP

# Best Path Selection Algorithm - Example

```
rtr01#show bgp ipv4 unicast 140.203.0.0/16

BGP routing table entry for 140.203.0.0/16

Paths: (3 available, best #x)
 174 3356 1213 1213
    154.50.192.49 from 154.50.192.49 (154.26.32.227)
      Localpref 100, valid, external
 1213
    194.88.240.15 from 194.88.240.8 (194.88.240.8)
      Localpref 400, valid, external
 1213
    83.220.203.172 from 83.220.203.172 (83.220.203.170)
      Localpref 300, valid, internal
```

BGP

# Best Path Selection Algorithm - Example

```
rtr01#show bgp ipv4 unicast 140.203.0.0/16

BGP routing table entry for 140.203.0.0/16

Paths: (3 available, best #x)
 174 3356 1213 1213
    154.50.192.49 from 154.50.192.49 (154.26.32.227)
      Localpref 100, valid, external
 1213
    194.88.240.15 from 194.88.240.8 (194.88.240.8)
      Localpref 400, valid, external
 1213
    83.220.203.172 from 83.220.203.172 (83.220.203.170)
      Localpref 300, valid, internal
```

BGP

# Useful Tools

RIPE Stat: https://stat.ripe.net/

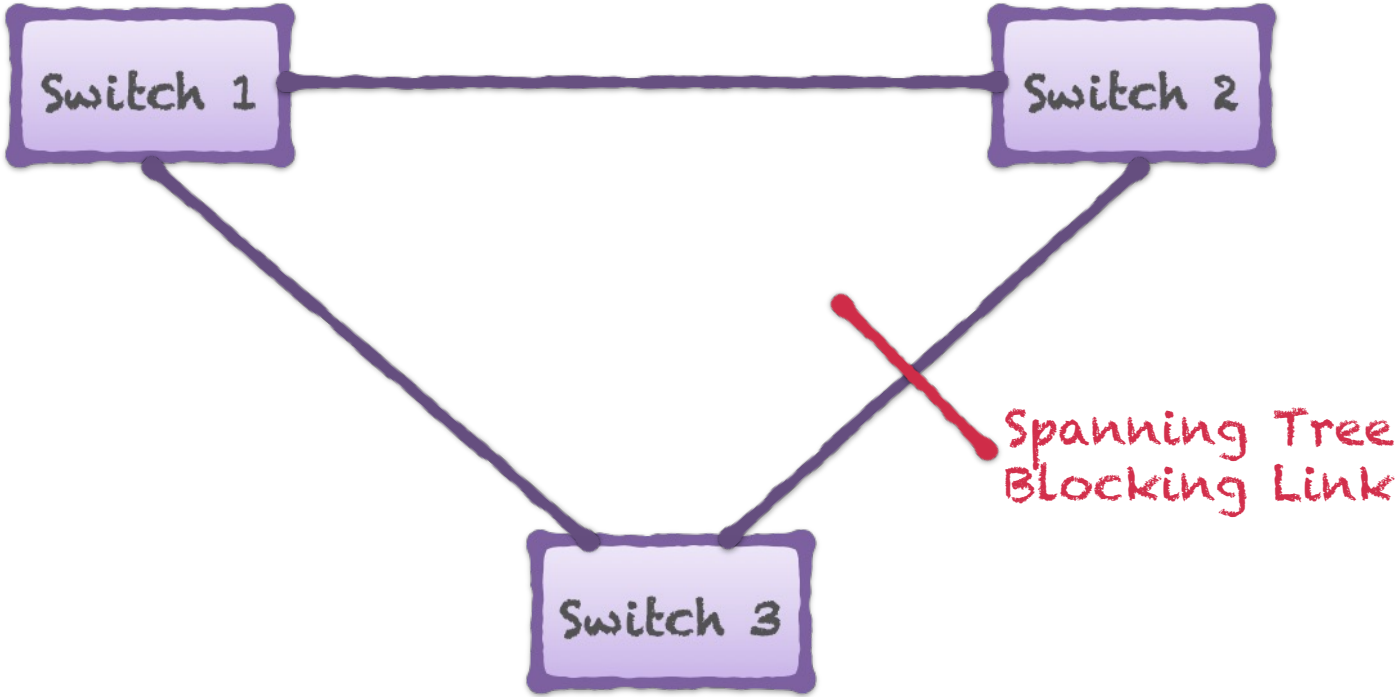

# INEX and IXPs

- INEX and IXPs

# What is an IXP?

**Internet eXchange Point**


"Just" a switching platform...


Switch 1
Switch 2
Switch 3
Spanning Tree
Blocking Link

# IXP Definition

- An Internet eXchange Point (IXP) is a network facility that enables the interconnection and exchange of internet traffic between more than two independent Autonomous Systems.

- An IXP provides interconnection only for Autonomous Systems.

- An IXP does not require the internet traffic passing between any pair of participating Autonomous Systems to pass through any third Autonomous System, nor does it alter or otherwise interfere with such traffic.

INEX AND IXPS

## IXPs

- Limited prefixes / routes
- Member & community
- Greater control
- Greater QoS
  - Traffic remains local
  - Lower latency
  - No congestion
- Greater reliability
- Enhanced security

## IP Transit

- DFZ (all prefixes)
- You're *just* a customer
- Limited or no control
- Main problems:
  - Latency
  - Routing
  - Congestion

# Examples of IXPs

- INEX in Dublin
- LINX in London, UK
- AMS-IX in Amsterdam, NL
- SIX in Seattle, US
- JPNAP in Tokyo, Japan


United Kingdom
Isle of Man
Ireland
Liverpool
Poland
Germany
Belgium
Luxembourg
Czech Republic
Slovakia
Austria
Hungary
Switzerland
Slovenia
Zagreb
Croatia
Bosnia and Herzegovina
Serbia
Sarajevo
Montenegro
Kosovo
Podgorica
Macedonia (FYROM)
Tirana
Albania
Greece
Monaco
Italy
Rome
Andorra
Spain
Portugal
Porto
Bay of Biscay
Tyrrhenian Sea
Vilnius

INEX AND IXPS

# INEX - An Introduction

- Why did Ireland need an exchange?
  - Internet connectivity was extremely expensive
    - Not unusual to pay £000's / Mbps / month
- Speed of access was cripplingly low
  - Local IXP would relieve international links
- Local traffic routed via London / Amsterdam
  - Greater security / resilience for local traffic

INEX AND IXPS

# INEX - An Introduction

- INEX - Internet Neutral Exchange Association (CLG)
  - Emphasis on **neutral**
    - No member is more important than any other
    - Not for profit, limited by guarantee
    - Open to anyone agreeing to and meeting the MoU
    - Owned by the members (currently ~100)

INEX AND IXPS

# INEX - An Introduction

- Mandate Includes:
  - Provide high-speed, reliable and resilient IP traffic exchange facilities
  - Allow our national and international members route traffic more efficiently
  - *Keep Irish IP traffic in Ireland*
  - No *feature creep* into other business areas

INEX AND IXPS

# INEX - Three IXPs

- INEX LAN1 and INEX LAN2 are both in Dublin offering the same services in the same PoPs
  - Members requested resiliency
  - Everyone joins INEX LAN1
  - Most join INEX LAN2
    - Incentives: free 1Gb port on LAN2 when you join LAN1

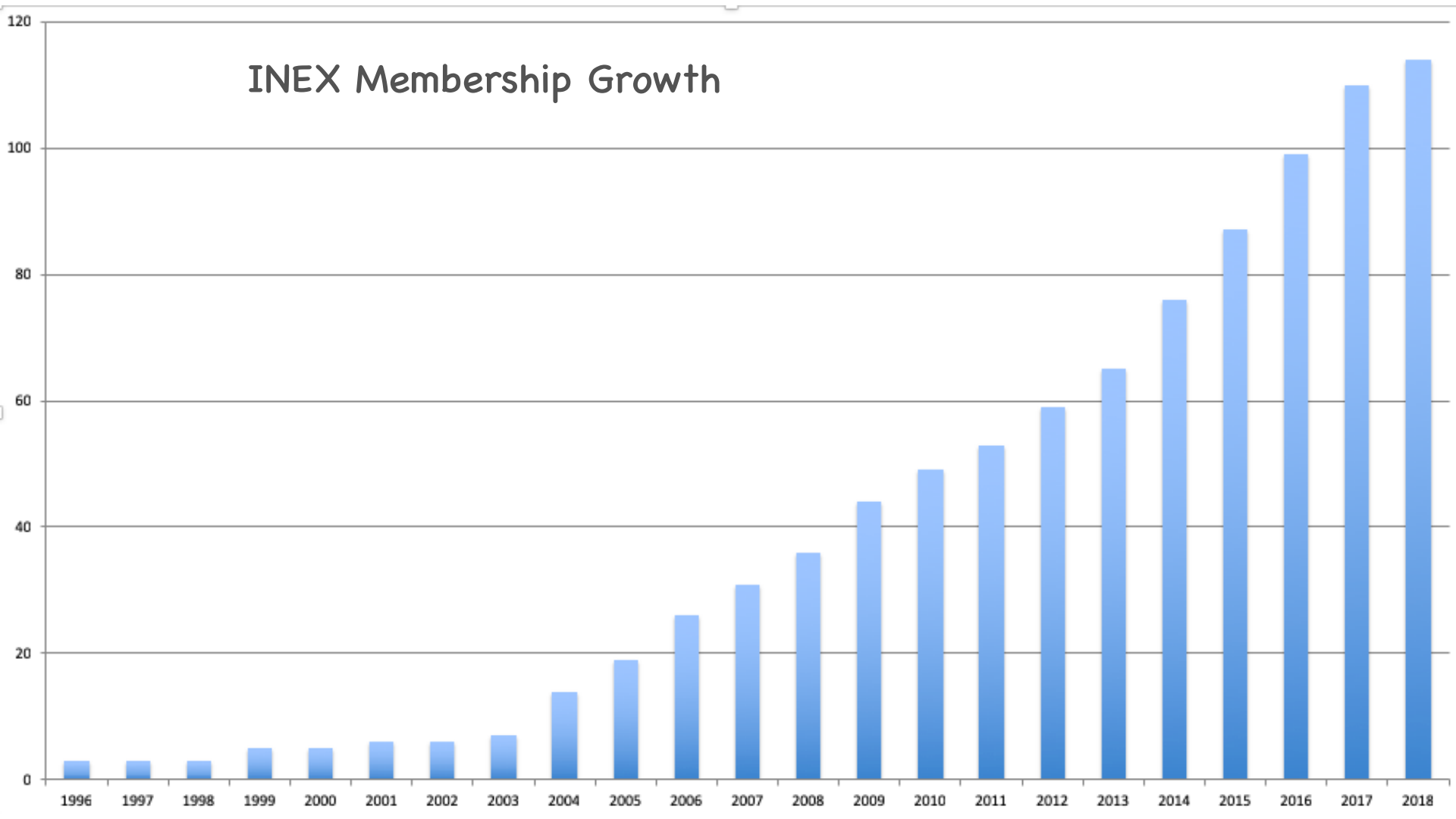
- INEX Cork
  - Opened in 2016
  - Supports 1/10Gb ports
  - All ports are currently free to end of 2019

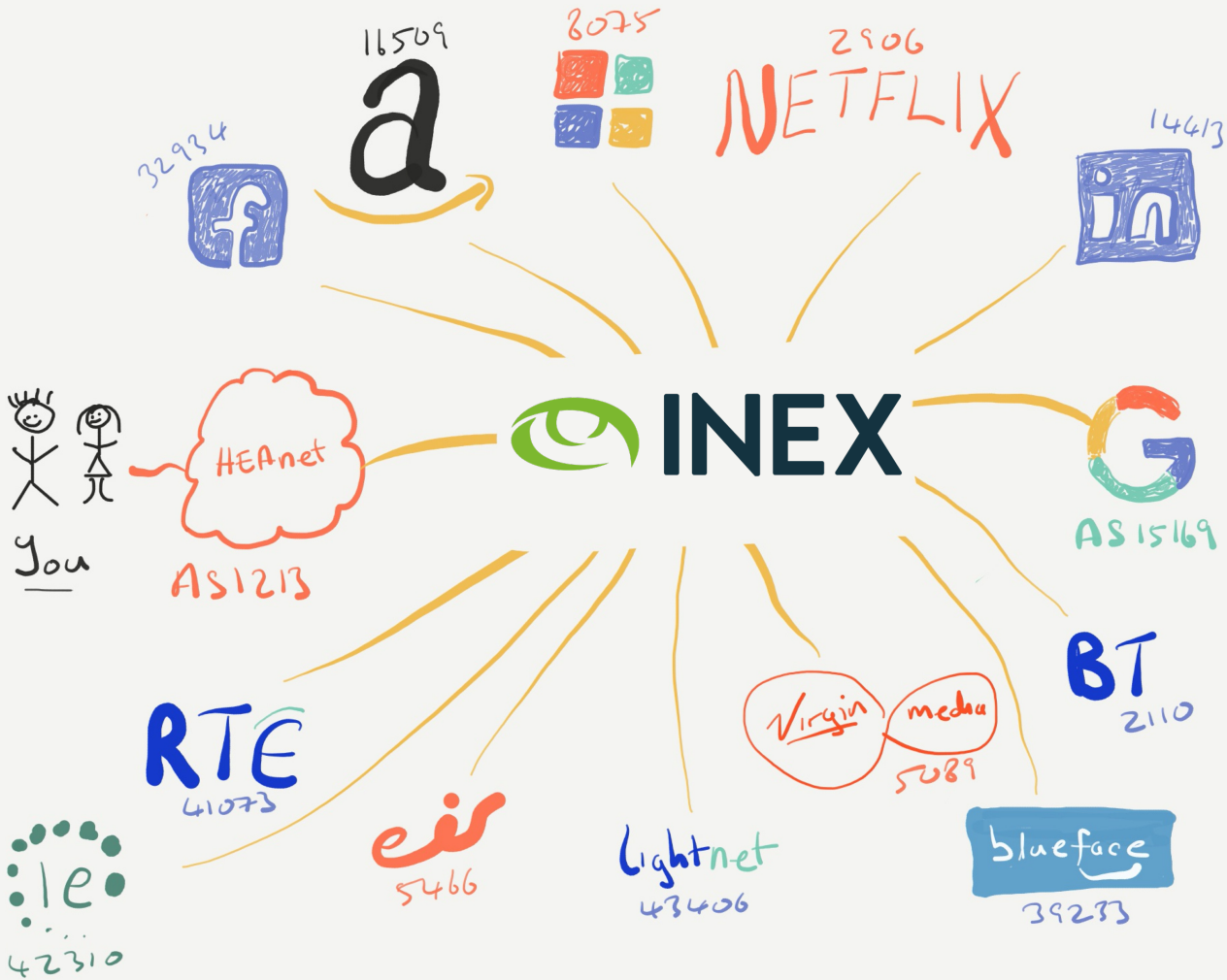
INEX AND IXPS

# INEX - Three IXPs

## Customer Ports by Infrastructure

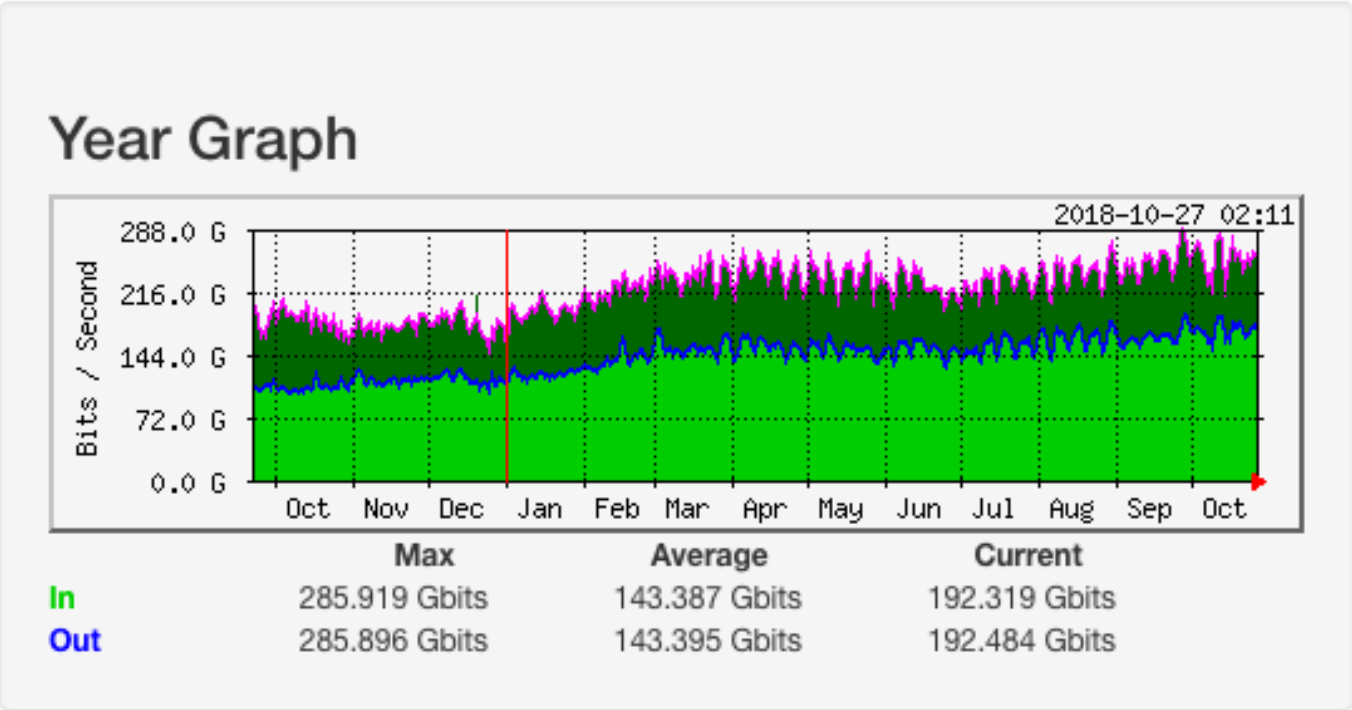
| Infrastructure | 1 Gbits | 10 Gbits | 100 Gbits | Total | Connected Capacity |
|---|---|---|---|---|---|
| INEX LAN1 | 52 | 80 | 6 | 138 | 1.45 Tbits |
| INEX LAN2 | 35 | 39 | 1 | 75 | 525.00 Gbits |
| INEX Cork | 10 | 9 | 0 | 19 | 100.00 Gbits |
| **Totals** | **97** | **128** | **7** | **232** | **2.077 Tbits** |


INEX Membership Growth
120
100
80
60
40
20
0
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018


16509
8075
2906
NETFLIX
32934
14413
You
HEAnet
AS1213
INEX
AS15169
RTE
41073
Virgin
media
5089
BT
2110
eir
5466
lightnet
43406
blueface
39273
ie
42310

INEX AND IXPS

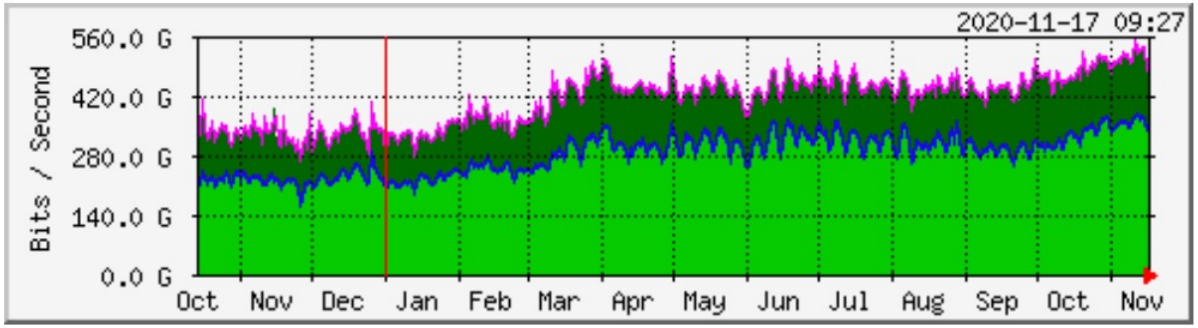
# Traffic At INEX (2018)



INEX AND IXPS

# Traffic At INEX (2020)

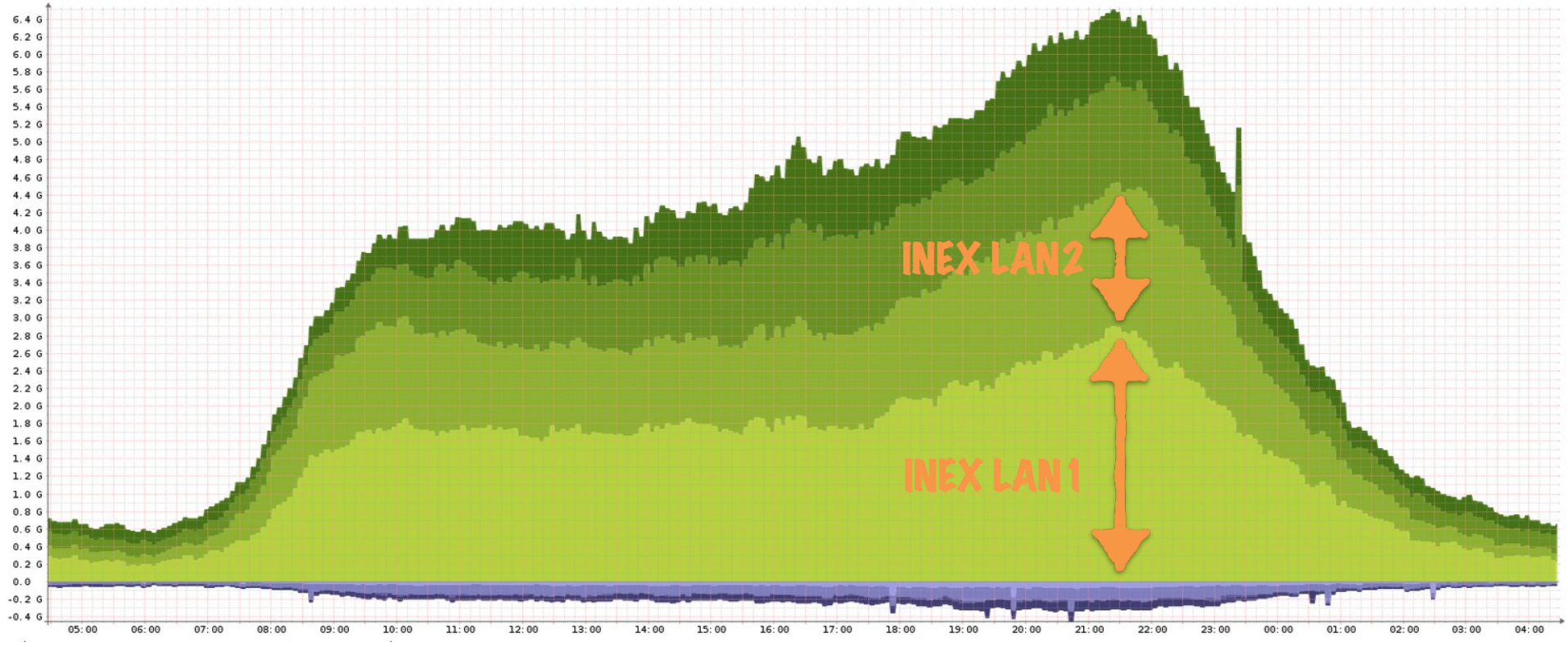
## Year Graph




| | Max | Average | Current |
|---|---|---|---|
| In | 552.811 Gbits | 283.988 Gbits | 426.073 Gbits |
| Out | 552.758 Gbits | 284.461 Gbits | 426.166 Gbits |

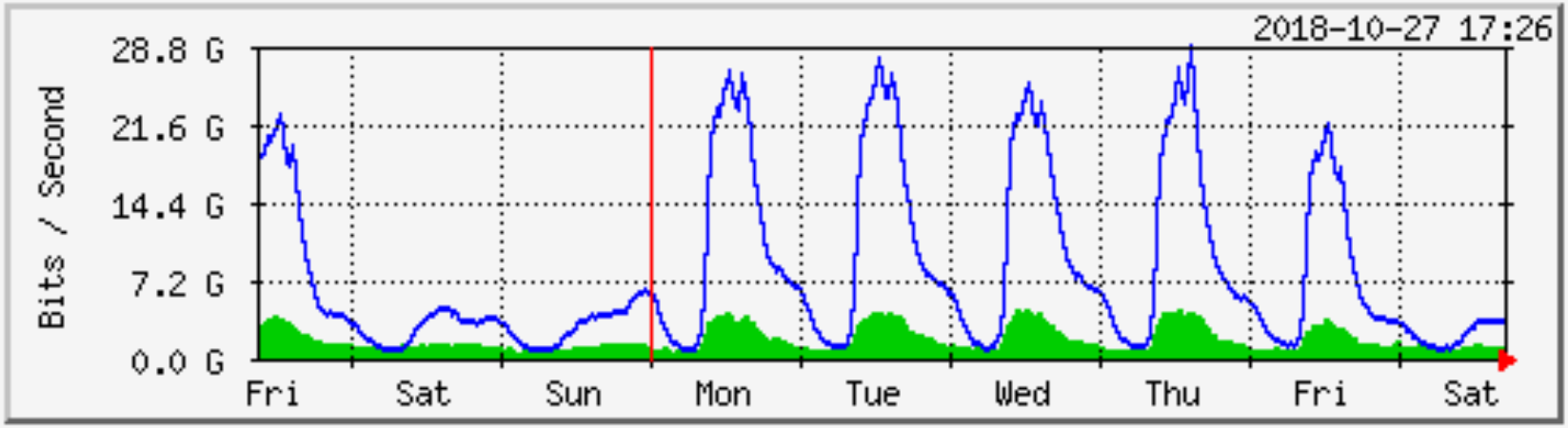
INEX AND IXPS

# Sample INEX Member - WISP



INEX AND IXPS

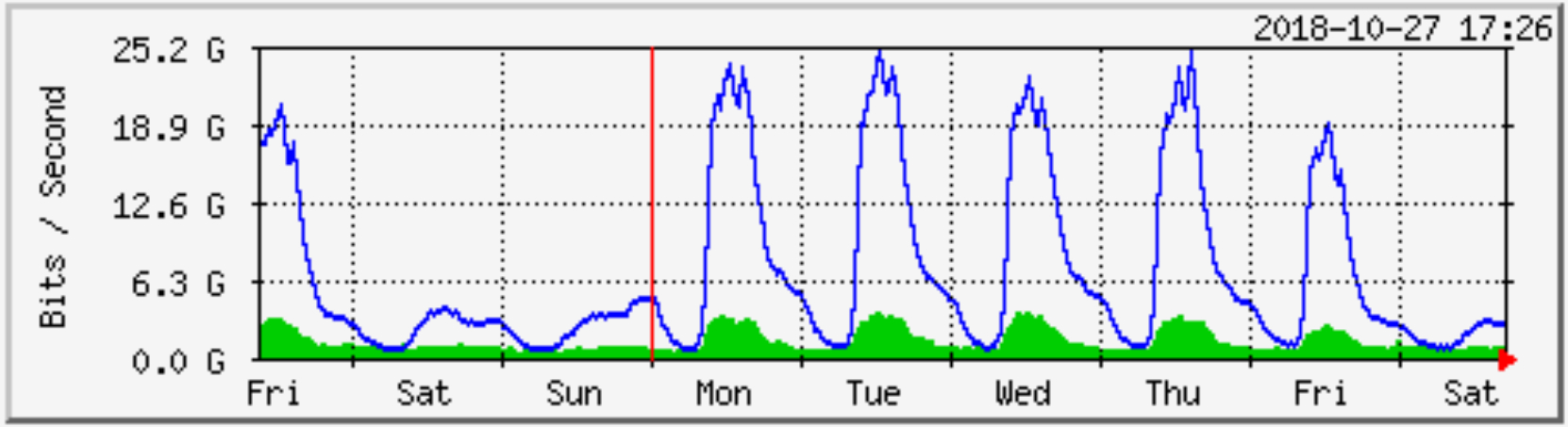
# HEAnet Traffic Over INEX (All LANs)




| | Max | Average | Current |
|---|---|---|---|
| In | 4.670 Gbits | 1.634 Gbits | 798.671 Mbits |
| Out | 29.831 Gbits | 7.629 Gbits | 2.830 Gbits |

INEX AND IXPS

# HEAnet Traffic Over INEX LAN1




| | Max | Average | Current |
|---|---|---|---|
| In | 3.904 Gbits | 1.274 Gbits | 588.441 Mbits |
| Out | 26.168 Gbits | 6.639 Gbits | 2.343 Gbits |

INEX

# What Happens When an AS Joins INEX?

- INEX Operations assigns you an IP address
  - INEX LAN1: 185.6.36.0/23 / 2001:7f8:18::/64
  - INEX LAN2: 194.88.240.0/25 / 2001:7f8:18:12::/64
  - INEX Cork: 185.1.69.0/24 / 2001:7f8:18:210::/64
- They do a quarantine procedure
- Your first BGP peering session is with the route collector
  - It's purely for diagnostic and monitoring purposes - but mandatory
  - It's BGP filters: accept everything, advertise nothing
  - Looking glass: https://www.inex.ie/ixp/lg/rc1-lan1-ipv4

INEX LAN1 - Route Collector - IPv4

# Looking Glass Routes for Protocol

## pb_as1213_vli4_ipv4

*This is the public looking glass. Uncached results and additional routers available when logged in.*

Bird 1.6.3 | API: 1.1.4 | Router ID: 185.6.36.126 | Uptime: 346 days. | Last Reconfigure: 2018-10-28 09:00:12 JSON: [status] [bgp]

Search:

| Network | Next Hop | | Metric | Communities? | AS Path | |
|---|---|---|---|---|---|---|
| 134.226.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 | Details |
| 136.201.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 | Details |
| 136.206.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 | Details |
| 137.43.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 2850 | Details |
| 140.203.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 | Details |
| 143.239.0.0/16 | 185.6.36.16 | P | 100 | 1 | 1213 | Details |

INEX

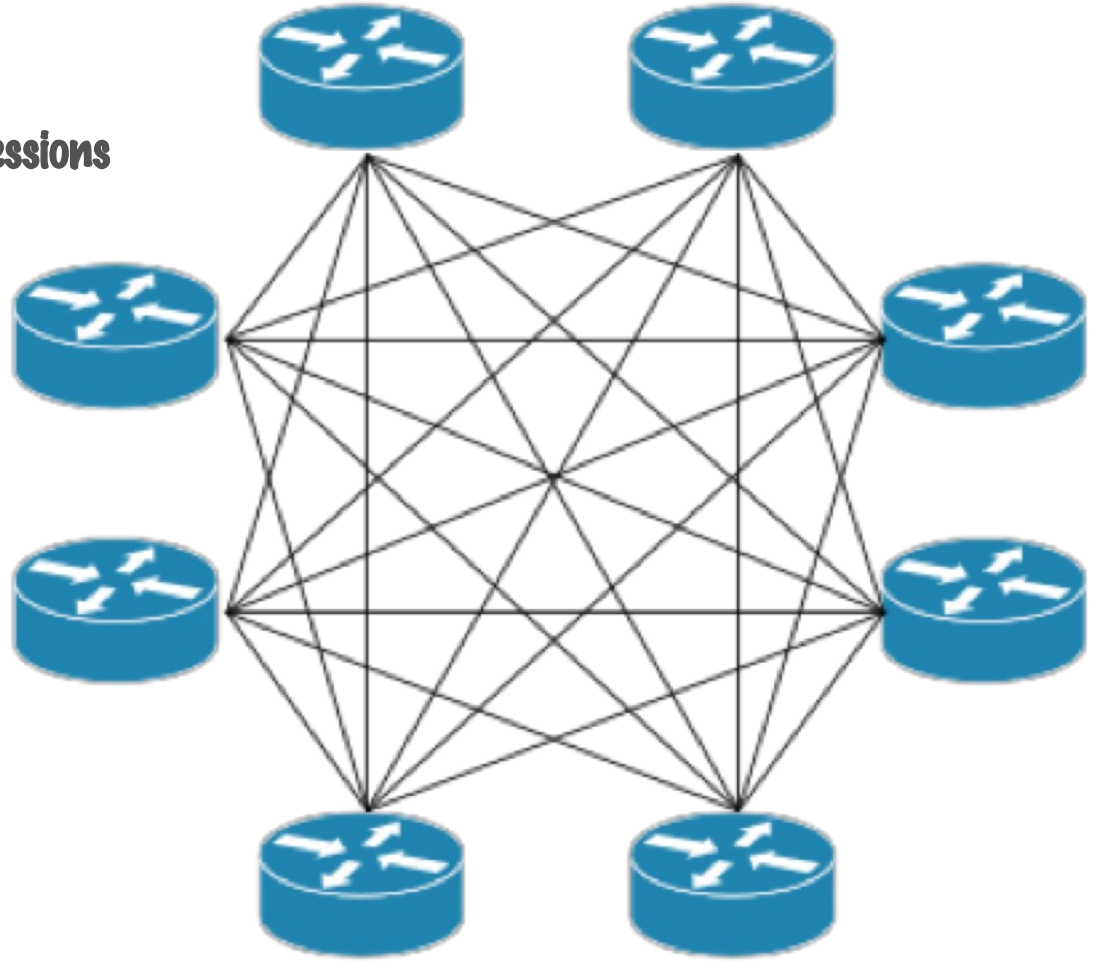
# What Happens When an AS Joins INEX?

- But how do I start peering with Google and Facebook and …
- Usually you email them: peering@example.com
- But with 100 other members, that's a lot of emails and hours (on both ends)
- In fact it's:

$$\frac{n(n+1)}{2} \Rightarrow \frac{100(101)}{2} \Rightarrow 5050$$

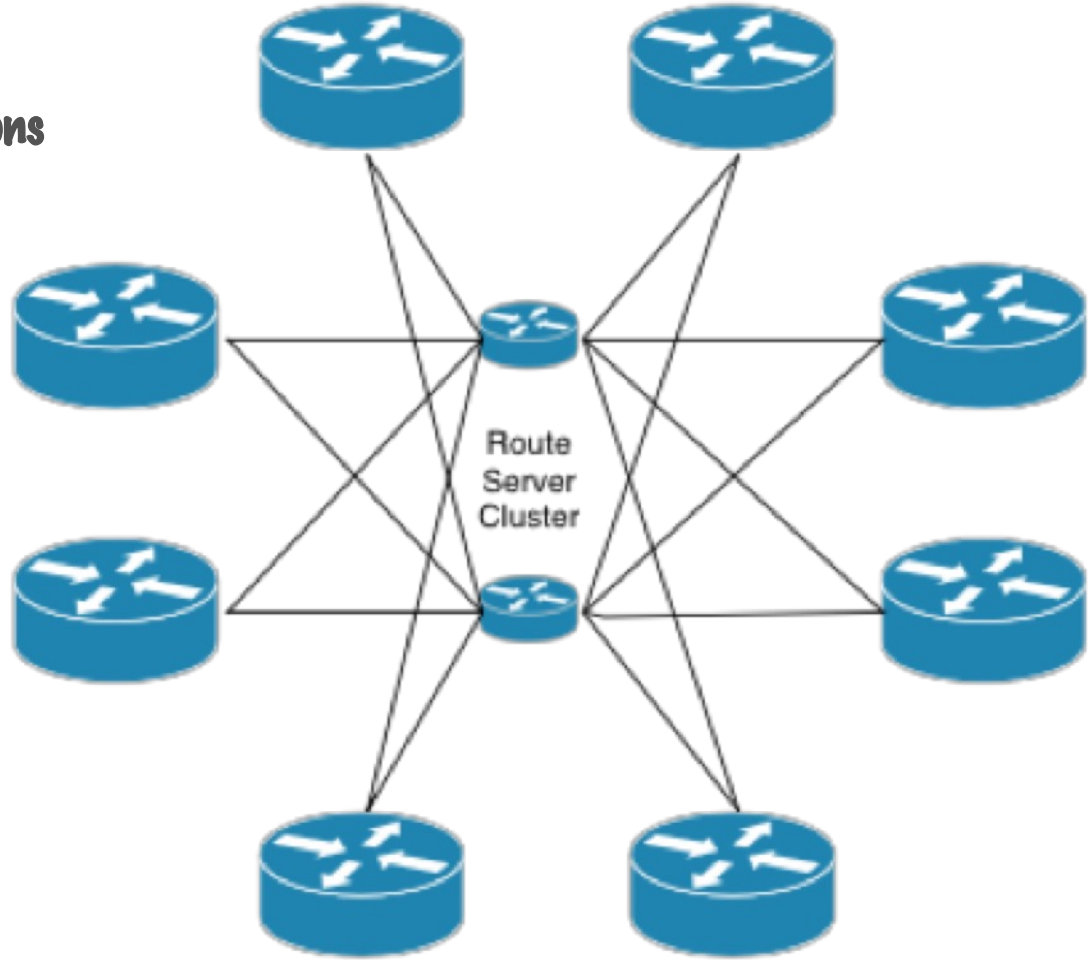
$$\frac{n(n+1)}{2}$$

individual BGP sessions



2n

individual BGP sessions



INEX

# Route Servers

- RFC7947 - Internet Exchange BGP Route Server
- RFC7948 - Internet Exchange BGP Route Server Operations
- "Third party brokering system" run by the IX
  - Does not change advertised routes in anyway
  - Does not put its own ASN in the AS path
  - The next-hop address is the other member's router - route servers do not route traffic
  - Multilateral vs bilateral peering sessions
  - Must be reliable, secure and trust-worthy
- INEX's are IP addresses ending .8/.9/::8/::9 in each LAN, ASN is 43760

BGP

# Best Path Selection Algorithm - Example

```
RP/0/RSP0/CPU0:rtr01#show bgp ipv4 unicast summary

Neighbor        Spk    AS MsgRcvd MsgSent   TblVer  InQ OutQ  Up/Down  St/PfxRcd
83.220.203.172    0 12388 1804398 2846671 114184776    0    0    35w1d     123530
154.50.192.49     0   174 4253889  354696 114184776    0    0    35w1d     711433
194.88.240.6      0   112  405168  354596 114184776    0    0    16w3d          2
194.88.240.8      0 43760  465199  354663 114184776    0    0    16w3d       1729
194.88.240.9      0 43760  464071  354671 114184776    0    0    16w3d       1729
194.88.240.13     0  2110  355881  354542 114184776    0    0     4w6d        118
194.88.240.61     0 13335  731197  709289 114184776    0    0    16w3d        255
194.88.240.65     0 10310  729171  709325 114184776    0    0    16w3d        271
194.88.240.66     0   714  778921  709315 114184776    0    0    16w3d        542
194.88.240.67     0   714  798235  709314 114184776    0    0    16w3d        542
194.88.240.126    0  2128  405253  354672 114184776    0    0    16w3d          0
```

BGP

# Looking Back: Best Path Selection Algorithm

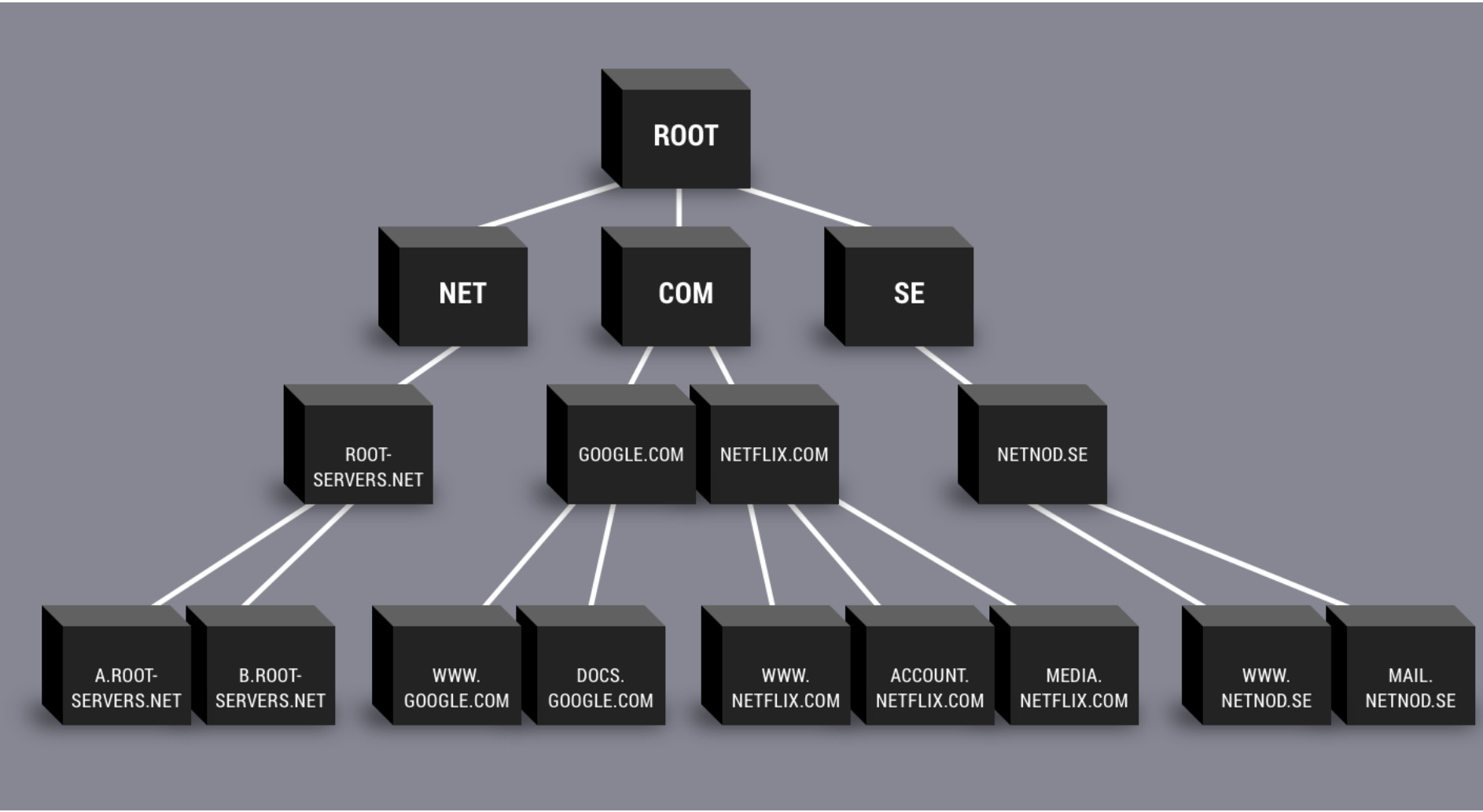
```
rtr01#show bgp ipv4 unicast 140.203.0.0/16

BGP routing table entry for 140.203.0.0/16

Paths: (3 available, best #x)
  174 3356 1213 1213
    154.50.192.49 from 154.50.192.49 (154.26.32.227)
      Localpref 100, valid, external
  1213
    194.88.240.15 from 194.88.240.8 (194.88.240.8)
      Localpref 400, valid, external
  1213
    83.220.203.172 from 83.220.203.172 (83.220.203.170)
      Localpref 300, valid, internal
```

- INEX route server ASN (43760) not in AS path
- Next hop is HEAnet's router, not the route server
- Route was learnt from the route server

INEX

# INEX and Essential Internet Infrastructure

- BGP "Anycast"
  - *Anycast* is a network addressing and routing methodology in which a single destination address has multiple routing paths to two or more endpoint destinations.
    - Usually means advertising the same /24 network from multiple AS'
    - Reliability - if one network or service instance goes offline, your router's BGP best-path algorithm will pick up the next best anycast advertisement
    - Latency - again, BGP best-path will pick the closest anycast advertisement
    - Only appropriate for certain services
      - CDNs - same content
      - DNS - and particularly the DNS root zone and TLDs


ROOT
NET
COM
SE
ROOT-SERVERS.NET
GOOGLE.COM
NETFLIX.COM
NETNOD.SE
A.ROOT-SERVERS.NET
B.ROOT-SERVERS.NET
WWW.GOOGLE.COM
DOCS.GOOGLE.COM
WWW.NETFLIX.COM
ACCOUNT.NETFLIX.COM
MEDIA.NETFLIX.COM
WWW.NETNOD.SE
MAIL.NETNOD.SE

INEX

# INEX and Essential Internet Infrastructure

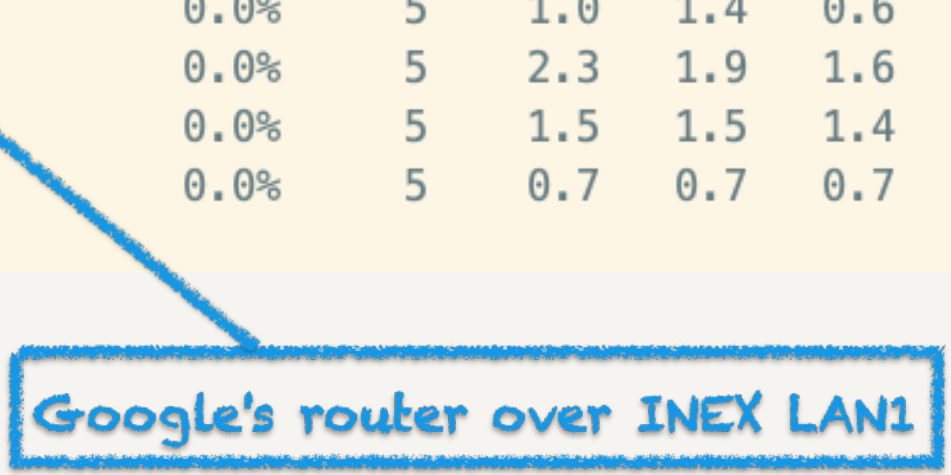
- There are 12 organisations who operate DNS root name servers
  - There are ~750 instances on 13 different IP addresses
  - 
  - These 13 addresses are hard-coded into DNS servers
- INEX offers pro-bono connections for such essential internet infrastructure:
  - Verisign '' DNS root server (and .com/.net authoritative server)
  - PCH 'D' and 'E' DNS root servers and ~400 TLDs (many countries included)
  - RIPE 'K' DNS root (INEX Cork)
- RIPE Atlas anchors, IE domain registry, Cloudflare 'F' DNS root server
- AS112 service hosting and operated by INEX

INEX AND IXPS

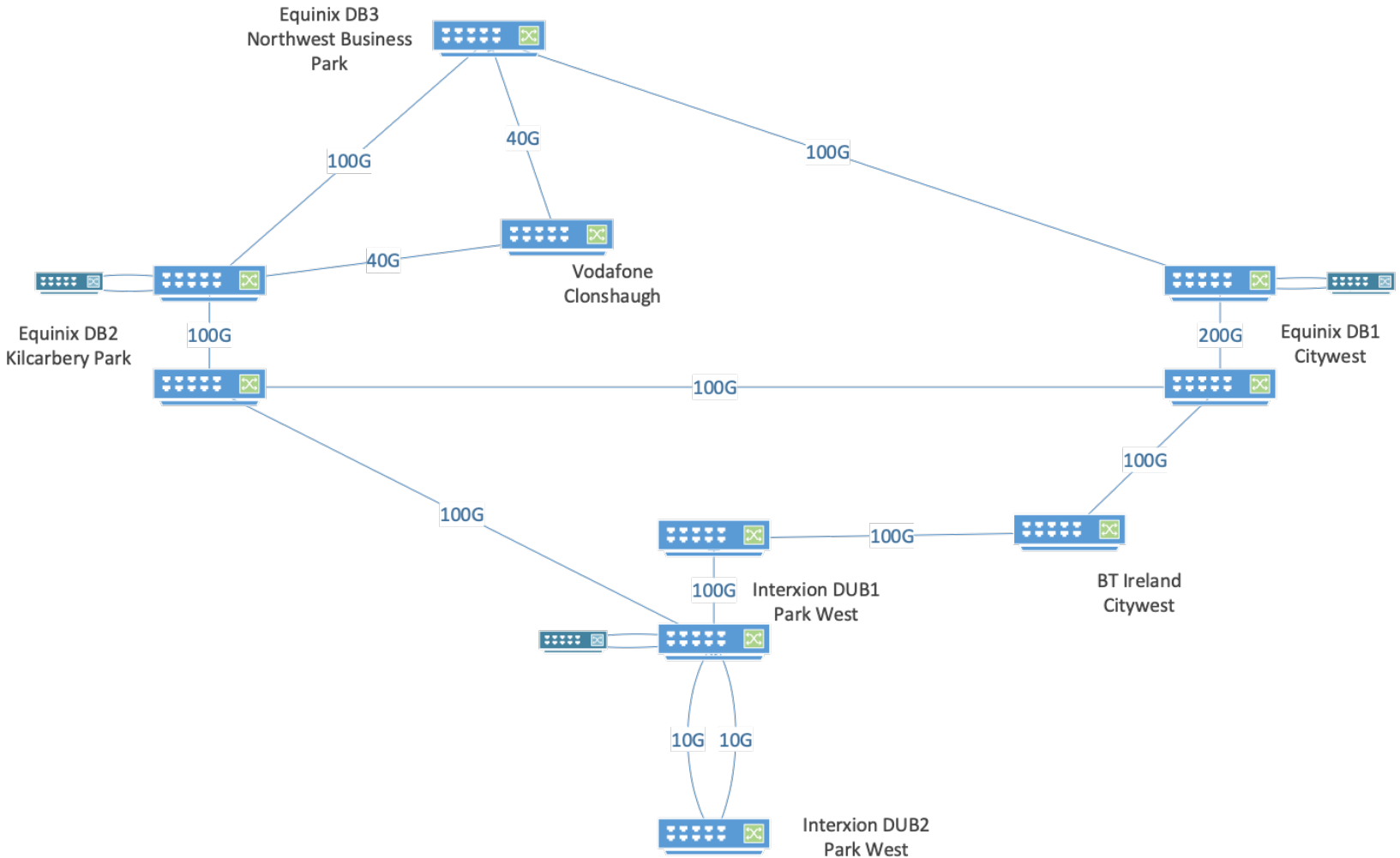
# Is Your Traffic Traversing INEX?

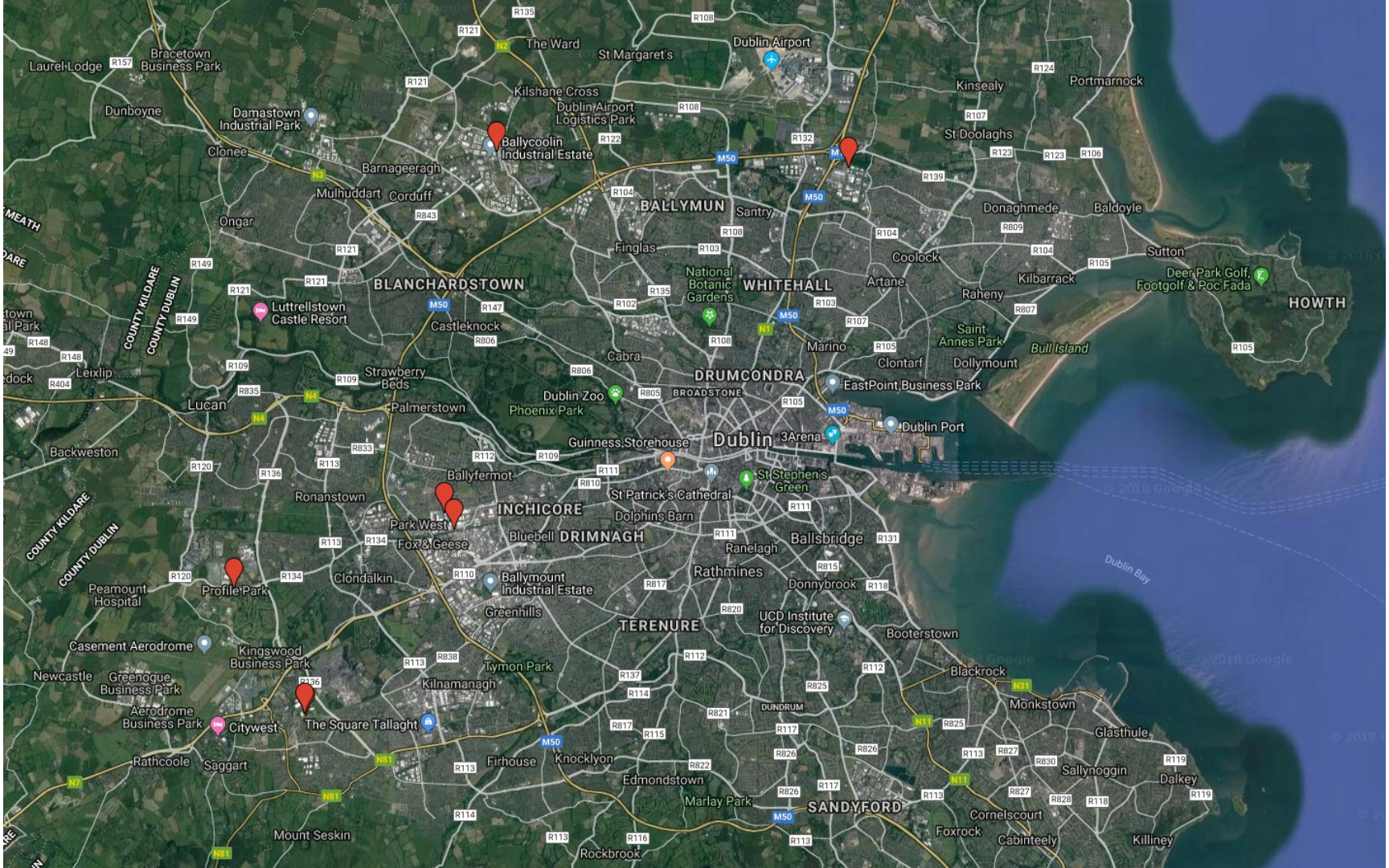
- Do a traceroute and see if an INEX IP shows up:
  - INEX LAN1: 185.6.36.0/23 / 2001:7f8:18::/64
  - INEX LAN2: 194.88.240.0/25 / 2001:7f8:18:12::/64
  - INEX Cork: 185.1.69.0/24 / 2001:7f8:18:210::/64

```
barryo@dub-ixn-lab01:~$ mtr -nrc 5 8.8.8.8
Start: 2018-10-25T16:24:09+0100
HOST: dub-ixn-lab01                Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 185.101.240.0               0.0%     5    0.5   0.4   0.3   0.5   0.1
  2.|-- 185.6.36.57                 0.0%     5    1.0   1.4   0.6   2.5   0.7
  3.|-- 74.125.244.1                0.0%     5    2.3   1.9   1.6   2.3   0.3
  4.|-- 72.14.239.219               0.0%     5    1.5   1.5   1.4   1.9   0.2
  5.|-- 8.8.8.8                     0.0%     5    0.7   0.7   0.7   0.7   0.0
barryo@dub-ixn-lab01:~$
```




# Technical Detail

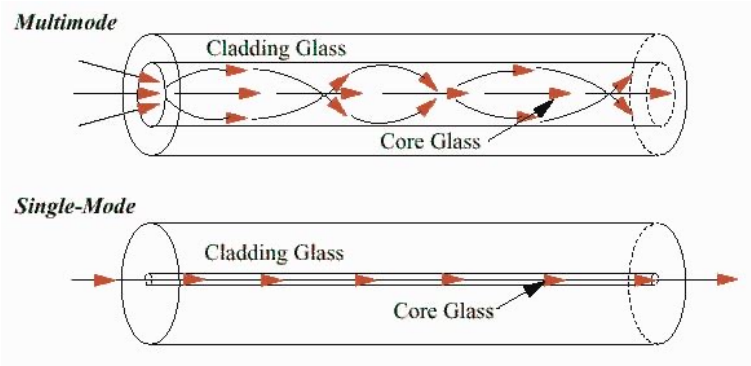

Equinix DB3
Northwest Business
Park
100G
40G
100G
40G
Vodafone
Clonshaugh
Equinix DB2
Kilcarbery Park
100G
200G
Equinix DB1
Citywest
100G
100G
100G
100G
100G
Interxion DUB1
Park West
BT Ireland
Citywest
10G
10G
Interxion DUB2
Park West


Laurel Lodge
Bracetown Business Park
The Ward
St Margaret's
Dublin Airport
Kinsealy
Portmarnock
Dunboyne
Damastown Industrial Park
Kilshane Cross
Dublin Airport Logistics Park
Ballycoolin Industrial Estate
St Doolaghs
Clonee
Barnageeragh
Mulhuddart
Corduff
BALLYMUN
Santry
Donaghmede
Baldoyle
Ongar
Finglas
Coolock
Sutton
Deer Park Golf, Footgolf & Poc Fada
BLANCHARDSTOWN
National Botanic Gardens
WHITEHALL
Artane
Kilbarrack
HOWTH
Luttrellstown Castle Resort
Castleknock
Raheny
Saint Annes Park
Marino
Bull Island
Cabra
Clontarf
Dollymount
Leixlip
Strawberry Beds
DRUMCONDRA
EastPoint Business Park
Lucan
Palmerstown
Dublin Zoo
Phoenix Park
BROADSTONE
Dublin Port
Guinness Storehouse
Dublin
3Arena
Backweston
Ballyfermot
St Stephen's Green
St Patrick's Cathedral
Ronanstown
INCHICORE
Dolphins Barn
Park West
Bluebell
DRIMNAGH
Ballsbridge
Fox & Geese
Ranelagh
Rathmines
Peamount Hospital
Profile Park
Clondalkin
Ballymount Industrial Estate
Donnybrook
Greenhills
TERENURE
UCD Institute for Discovery
Booterstown
Casement Aerodrome
Kingswood Business Park
Tymon Park
Blackrock
Newcastle
Greenogue Business Park
Kilnamanagh
Monkstown
Aerodrome Business Park
Citywest
The Square Tallaght
DUNDRUM
Glasthule
Rathcoole
Saggart
Firhouse
Knocklyon
Edmondstown
Sallynoggin
Dalkey
Marlay Park
SANDYFORD
Cornelscourt
Foxrock
Cabinteely
Killiney
Mount Seskin
Rockbrook
Dublin Bay
COUNTY KILDARE
COUNTY DUBLIN
MEATH

MAIN CONTENT

# ISLs - InterSite Links

- INEX has a single-mode (SM) dark fibre pair between sites
  - Multi-mode Fibre (usually orange (OM2) or blue (OM3)) - 10Gb @ ~300m
  - Single-mode Fibre (yellow) - 10Gb @ ~120Km




MAIN CONTENT

# ISLs - InterSite Links

- INEX has a single-mode (SM) dark fibre pair between sites
  - Multi-mode Fibre (usually orange (OM2) or blue (OM3)) - 10Gb @ ~300m
  - Single-mode Fibre (yellow) - 10Gb @ ~120Km
- Fibre "pair":
  - one strand for Tx (transmit)
  - one strand for Rx (receive)

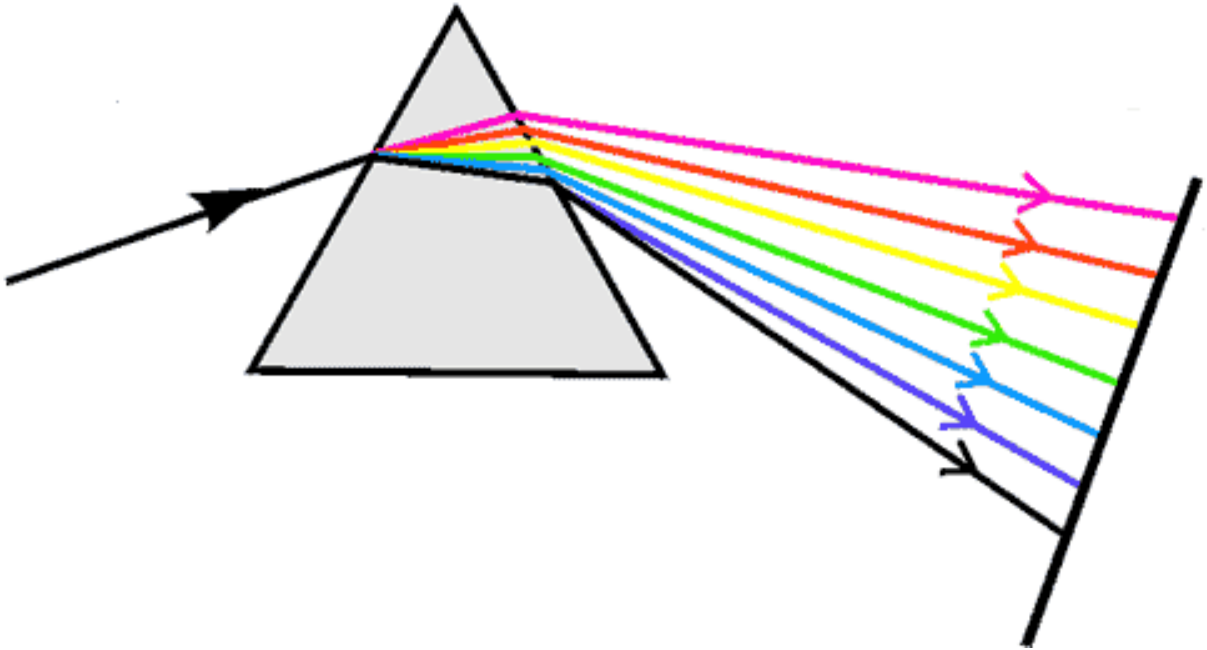
MAIN CONTENT

# ISLs - InterSite Links

- INEX has a single-mode (SM) dark fibre pair between sites
  - Multi-mode Fibre (usually orange (OM2) or blue (OM3)) - 10Gb @ ~300m
  - Single-mode Fibre (yellow) - 10Gb @ ~120Km
- Fibre "pair":
  - one strand for Tx (transmit)
  - one strand for Rx (receive)
- On each fibre pair, INEX needs to run multiple links:
  - n x 100Gb for INEX LAN1
  - n x 100Gb for INEX LAN2
  - 10Gb for management

INEX

# (D)WDM



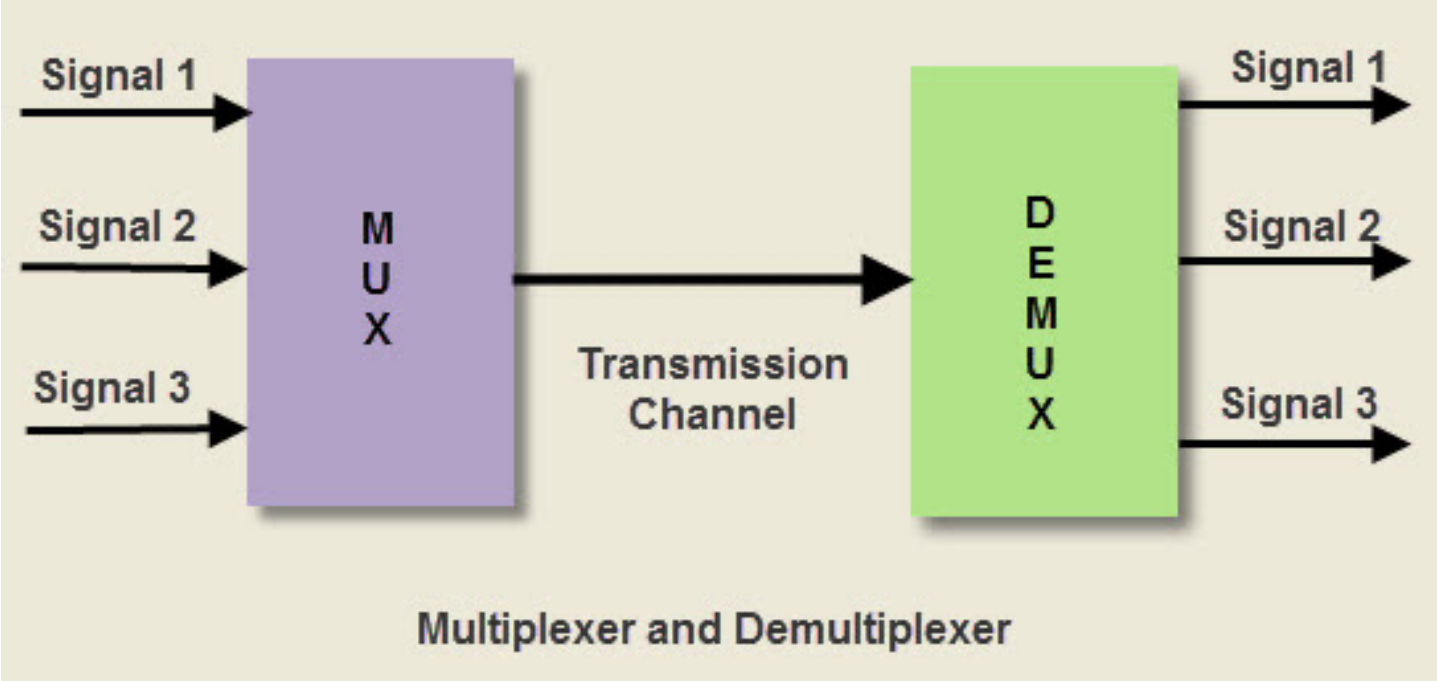
MAIN CONTENT

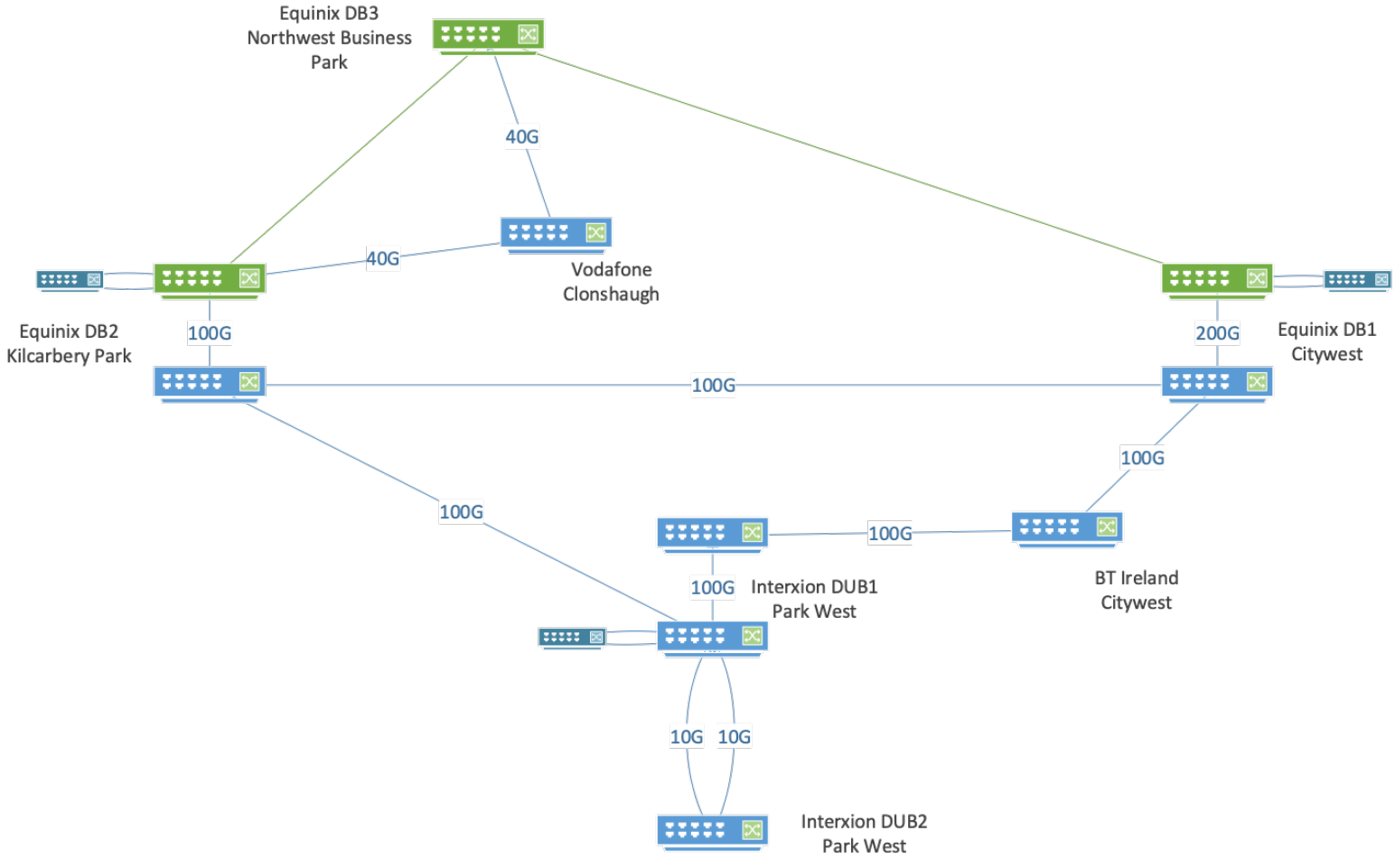
# (D)WDM

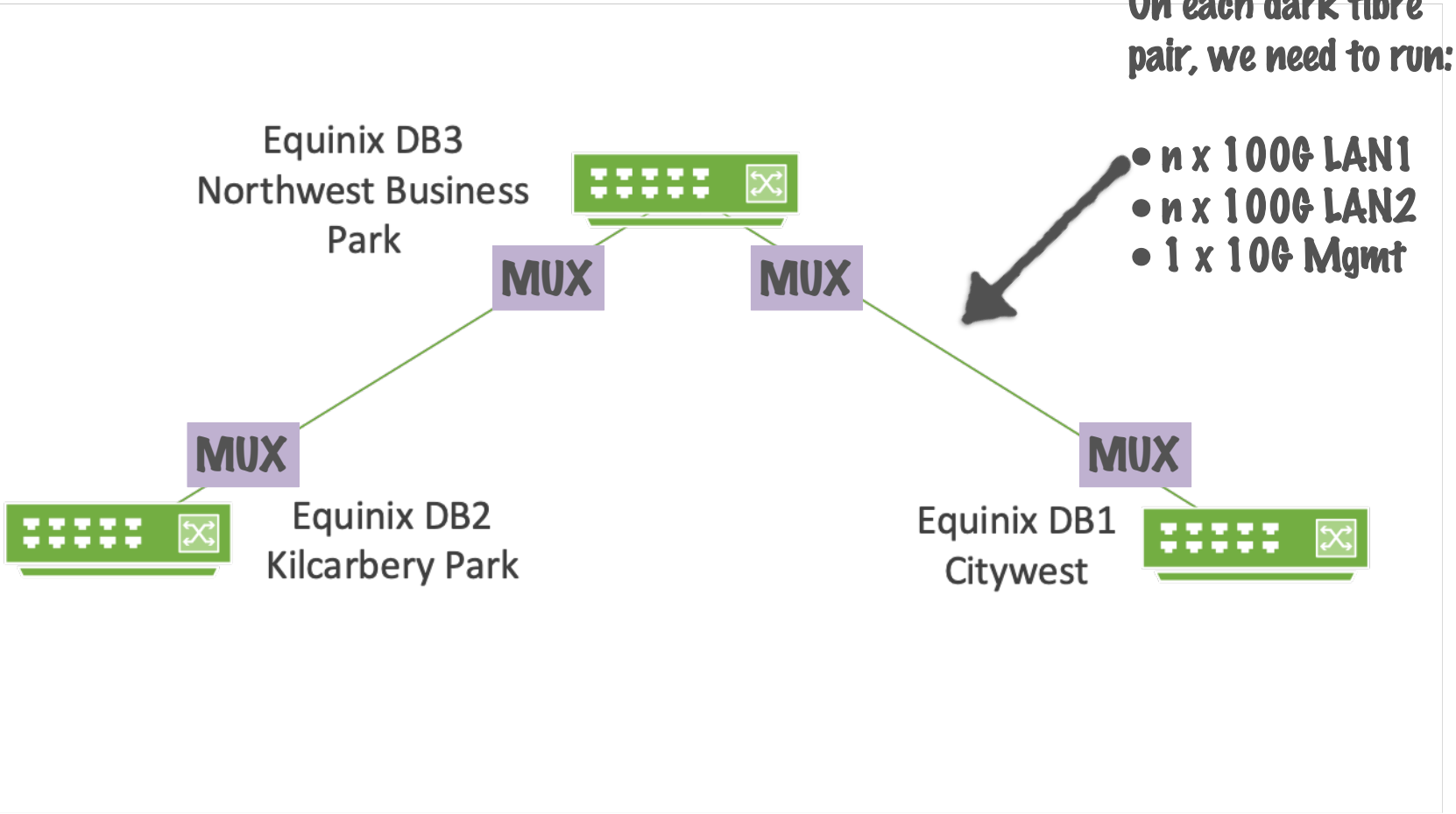
- Wavelength Division Multiplexing
- Using lasers to send light through fibre systems
- When light bends when sent through a prism or a diffraction grating, different wavelengths bend by a different amount
- We can run 1G / 10G / 100G links on different colours (wavelengths)

INEX

# (D)WDM - MUX/DEMUX




Equinix DB3
Northwest Business
Park
40G
40G
Vodafone
Clonshaugh
Equinix DB2
Kilcarbery Park
100G
200G
Equinix DB1
Citywest
100G
100G
100G
100G
100G
Interxion DUB1
Park West
BT Ireland
Citywest
10G
10G
Interxion DUB2
Park West


On each dark fibre pair, we need to run:
• n x 100G LAN1
• n x 100G LAN2
• 1 x 10G Mgmt
Equinix DB3
Northwest Business Park
MUX
MUX
MUX
MUX
Equinix DB2
Kilcarbery Park
Equinix DB1
Citywest