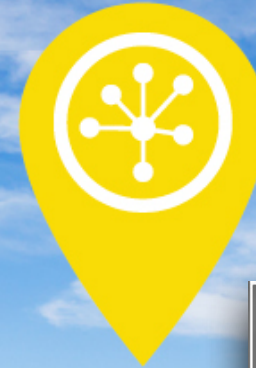


BGP - an introduction

BGP for networks who peer: Part 2

Wolfgang Tremmel

wolfgang.tremmel@de-cix.net

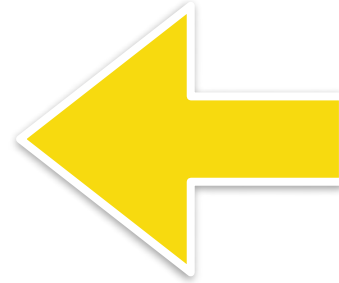


BGP Webinars Overview

- 01 - Prefixes and AS numbers
- 02 - BGP Introduction
- 03a - Setting up iBGP
- 03b - Setting up eBGP
- 04 - Becoming multi-homed
- 05 - BGP Best Path Selection
- ...

BGP Webinars Overview

- 01 - Prefixes and AS numbers
- 02 - BGP Introduction
- 03a - Setting up iBGP
- 03b - Setting up eBGP
- 04 - Becoming multi-homed
- 05 - BGP Best Path Selection
- ...



Today you will learn...



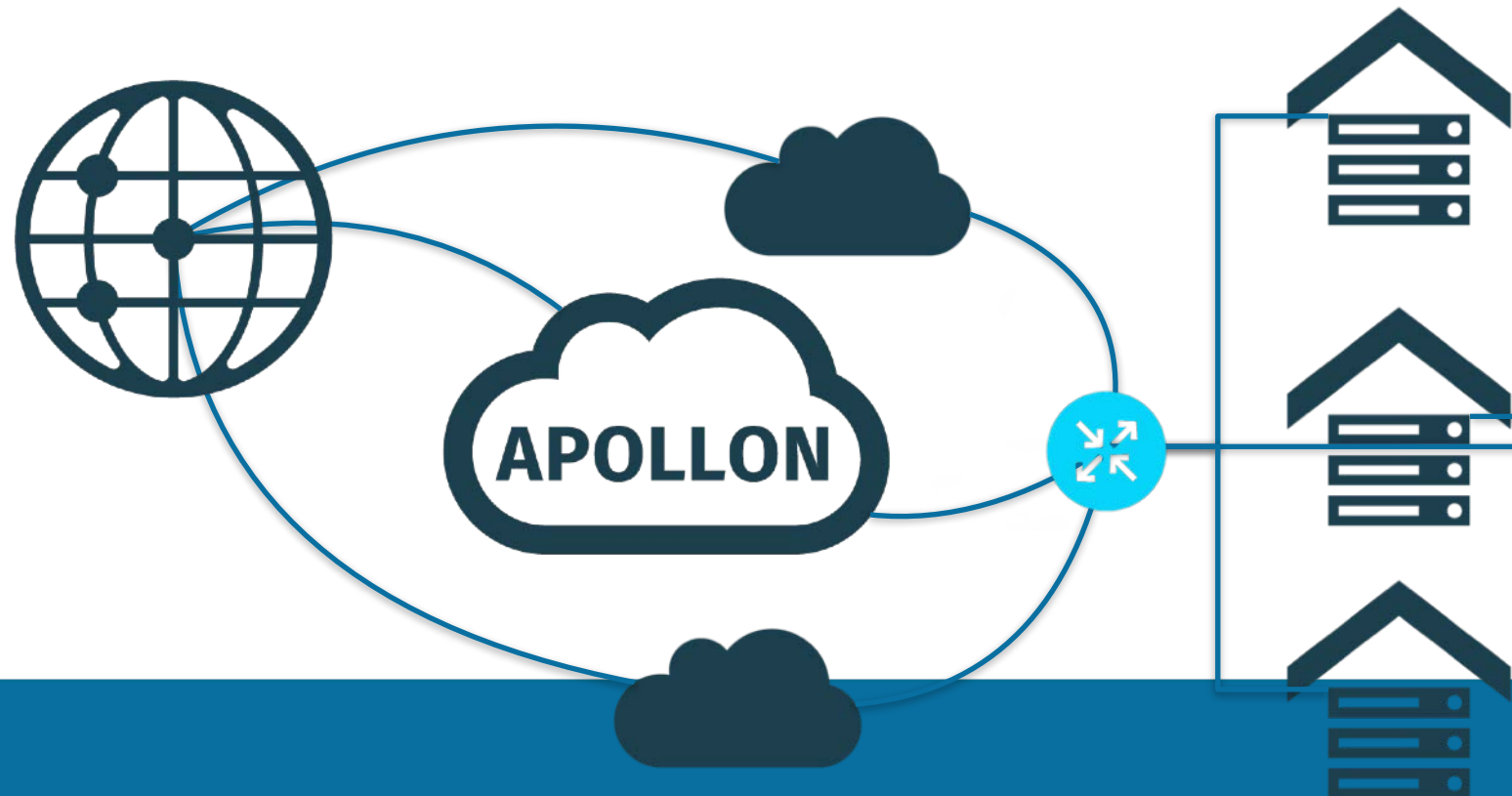
Today you will learn...

→ ... how to build and run a global network



Today you will learn...

- ... how to build and run a global network
- ... how to operate routers with upstreams and peerings



Today you will learn...

- ... how to build and run a global network
- ... how to operate routers with upstreams and peerings
- ... how to reduce cost, increase performance and resilience

\$\$\$ €€€

I am joking!

*But you will learn about **BGP**, the foundation of Internet routing*



BGP

BGP



BGP

Border

Gateway

Protocol

BGP

→ P - a **PROTOCOL**

BGP

→ P - a **PROTOCOL**

- spoken between Internet routers

BGP

- P - a **PROTOCOL**
 - spoken between Internet routers
- B - spoken on the **BORDER** between two providers



BGP

- P - a **PROTOCOL**
 - spoken between Internet routers
- B - spoken on the **BORDER** between two providers
- G - on the **GATEWAYS** - the routers connecting two providers



BGP Key Concepts

- IPv4 and IPv6 prefixes - we already know about them
- Autonomous Systems (AS)

BGP Key Concepts

- IPv4 and IPv6 prefixes - we already know about them
- Autonomous Systems (AS)
- The Autonomous System Path

Characteristics of Prefixes: IPv4

10.3.8.0/22

Prefix-Length: 0-32

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Notation:

- 4 Numbers 0-255
- Separated by "."
- a "/", followed by

Host-part all zero

32 Bits long

Characteristics of Prefixes: IPv4

10.3.8.0/22

Prefix-Length: 0-32

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Notation:

- 4 Numbers 0-255
- Separated by "."
- a "/", followed by

Host-part all zero

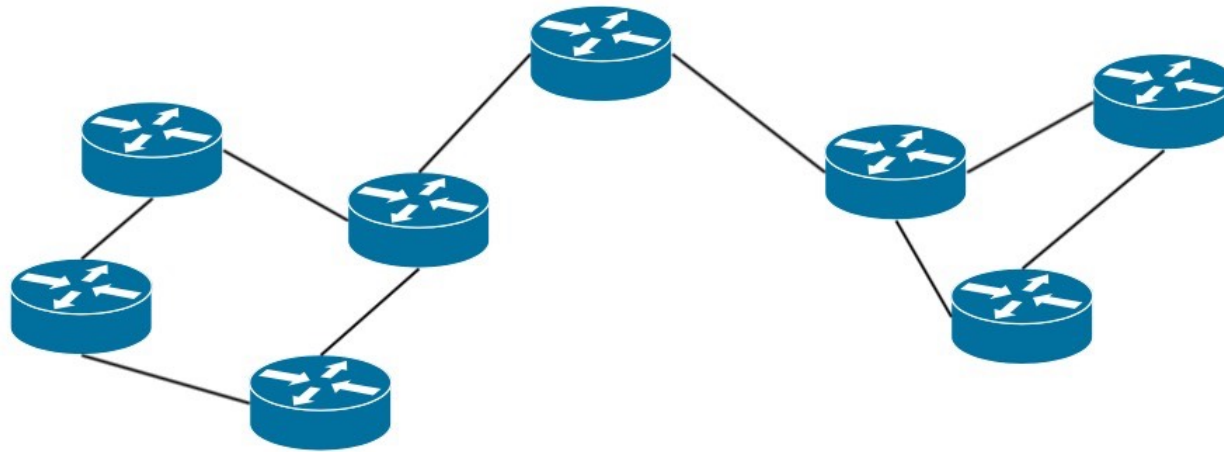
32 Bits long

Autonomous System

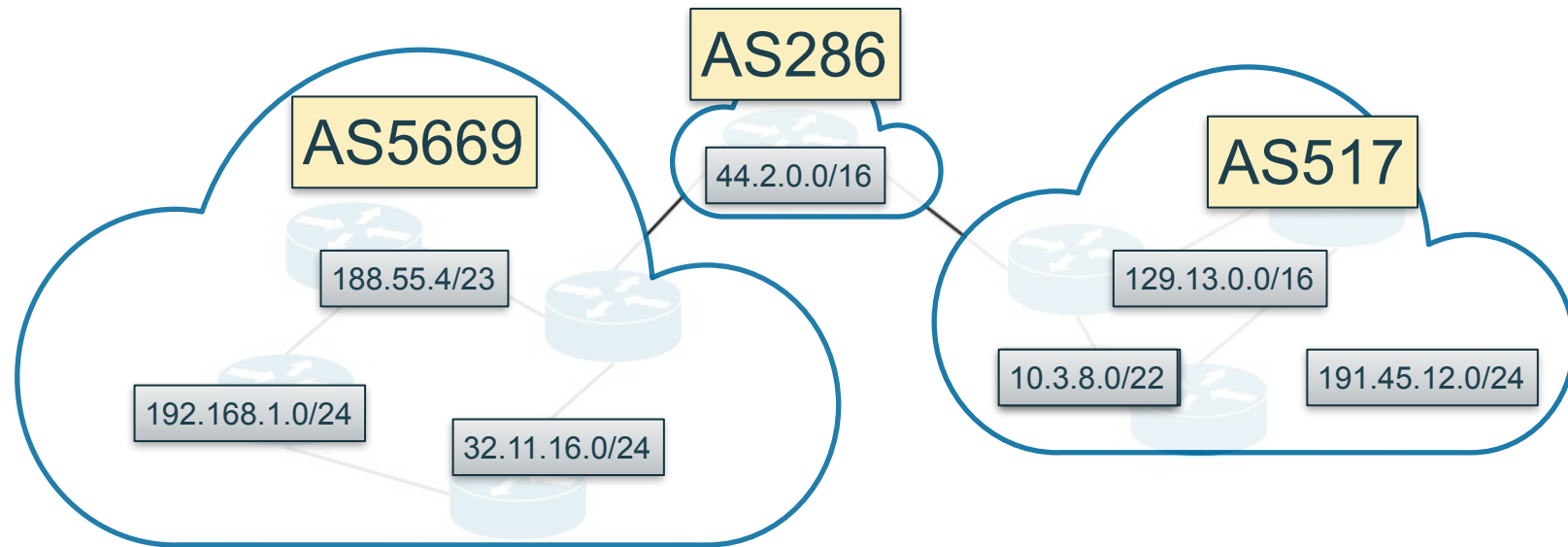
"An AS is a **connected group** of one or more **IP prefixes** run by one or more network operators which has a SINGLE and CLEARLY DEFINED routing policy."

Autonomous System

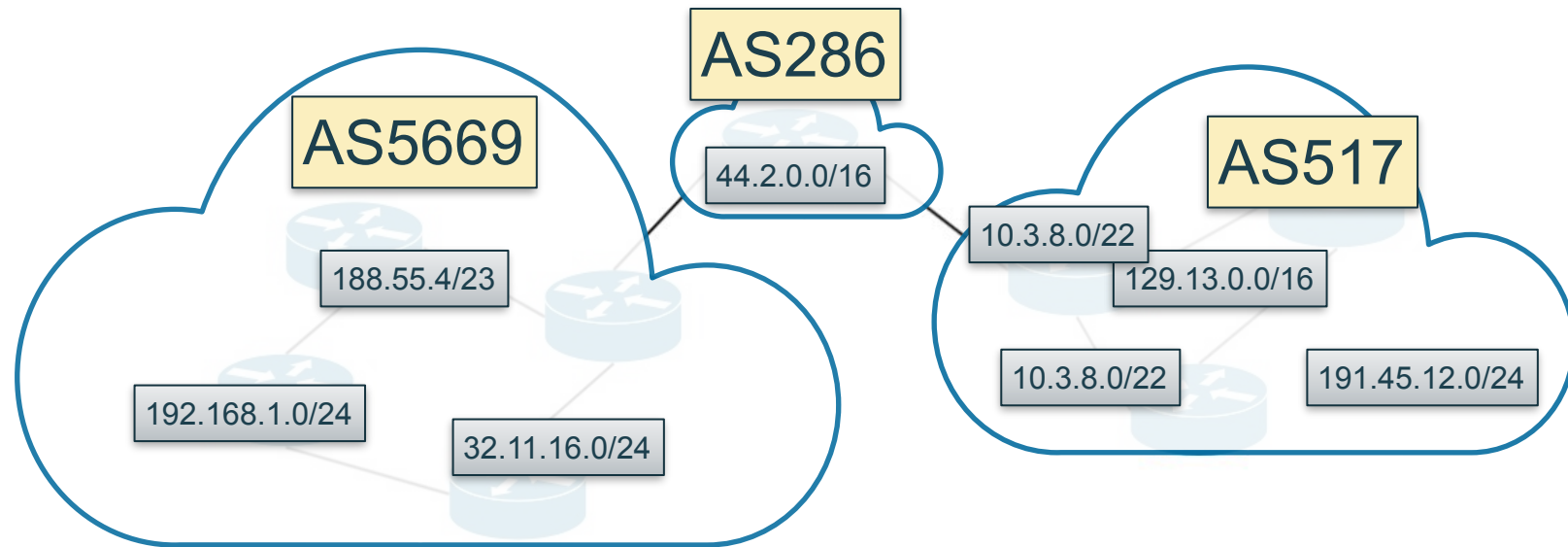
"An AS is a **connected group** of one or more **IP prefixes** run by one or more network operators which has a **SINGLE** and **CLEARLY DEFINED** routing policy."



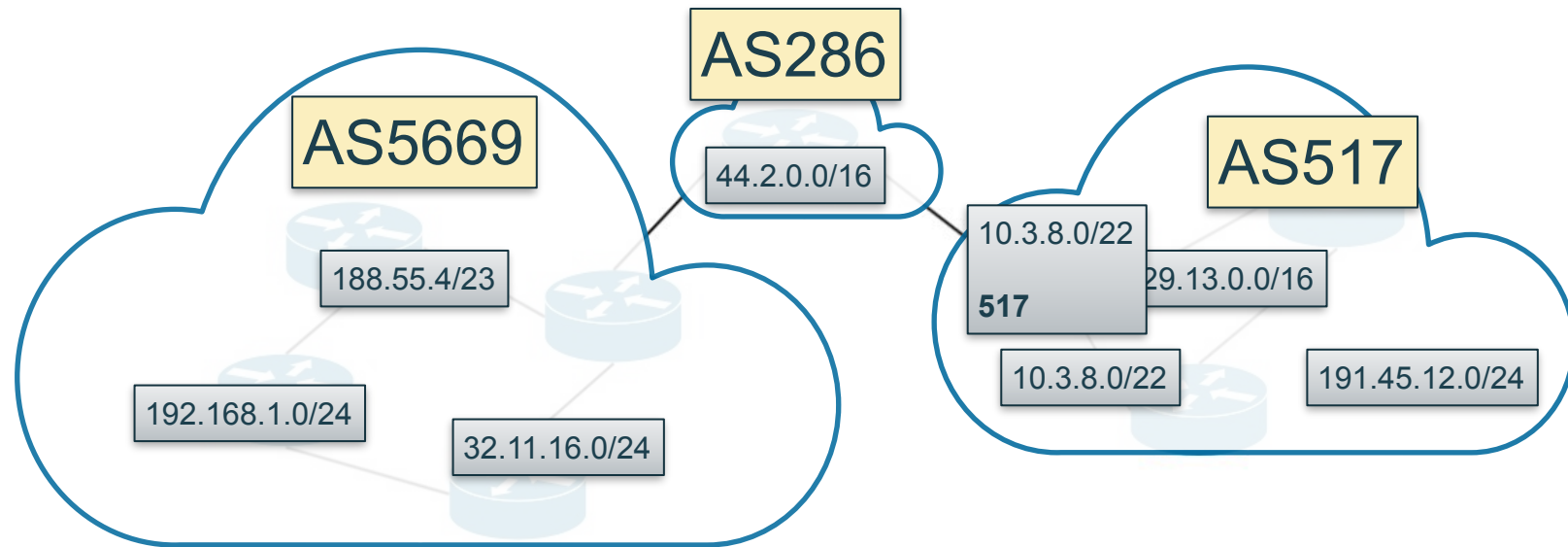
BGP - Key Concepts



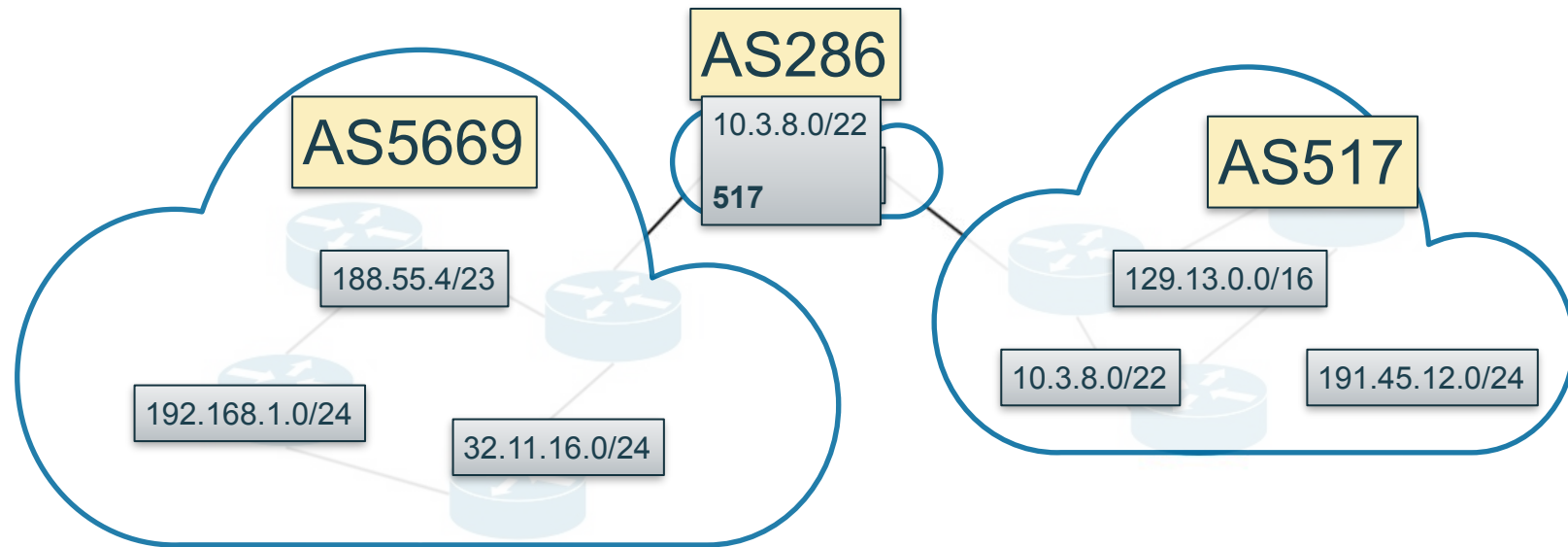
BGP - Key Concepts



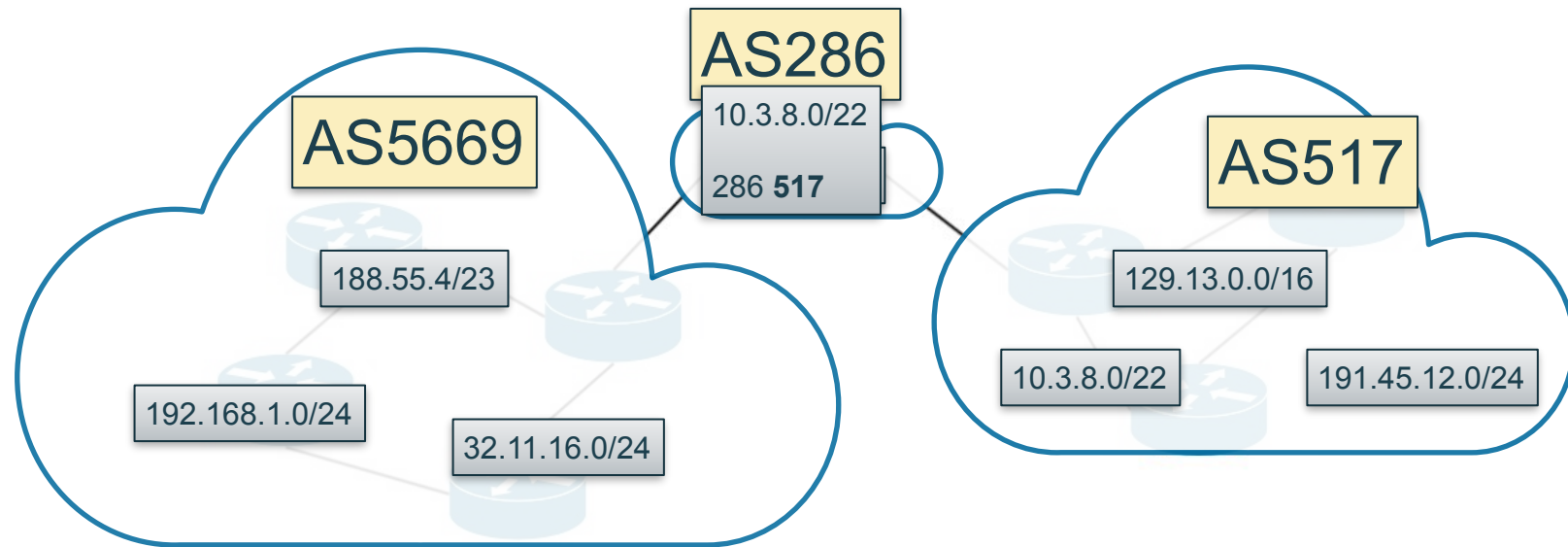
BGP - Key Concepts



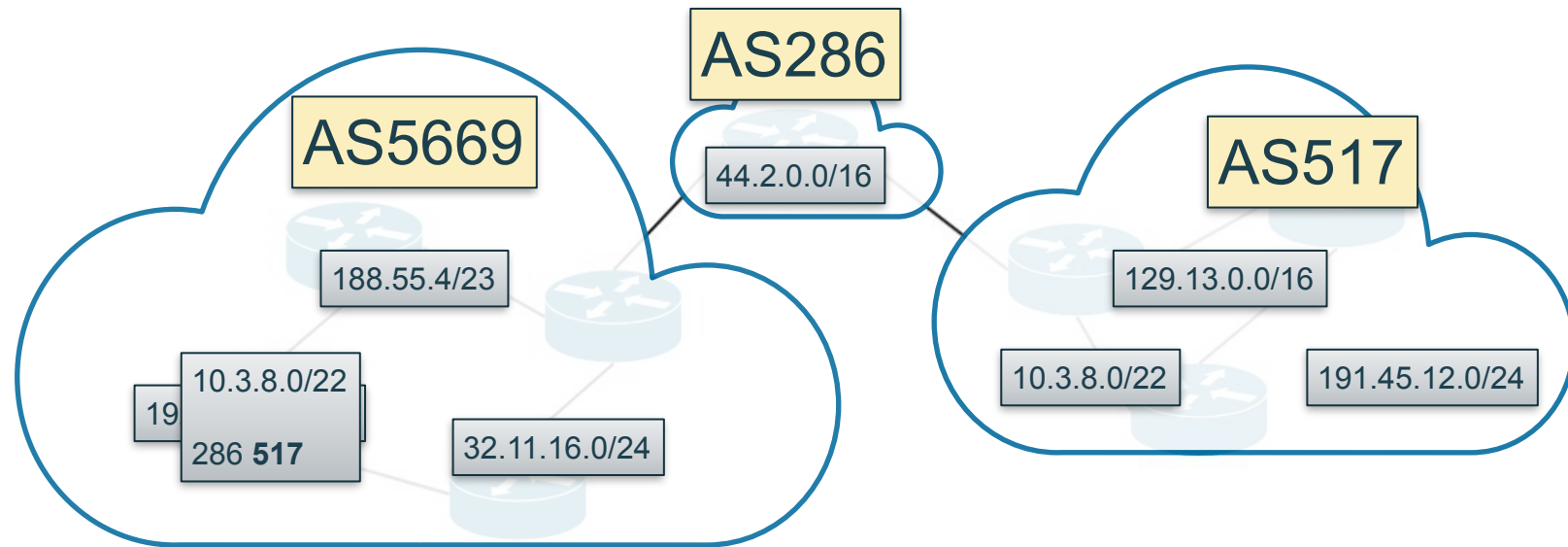
BGP - Key Concepts



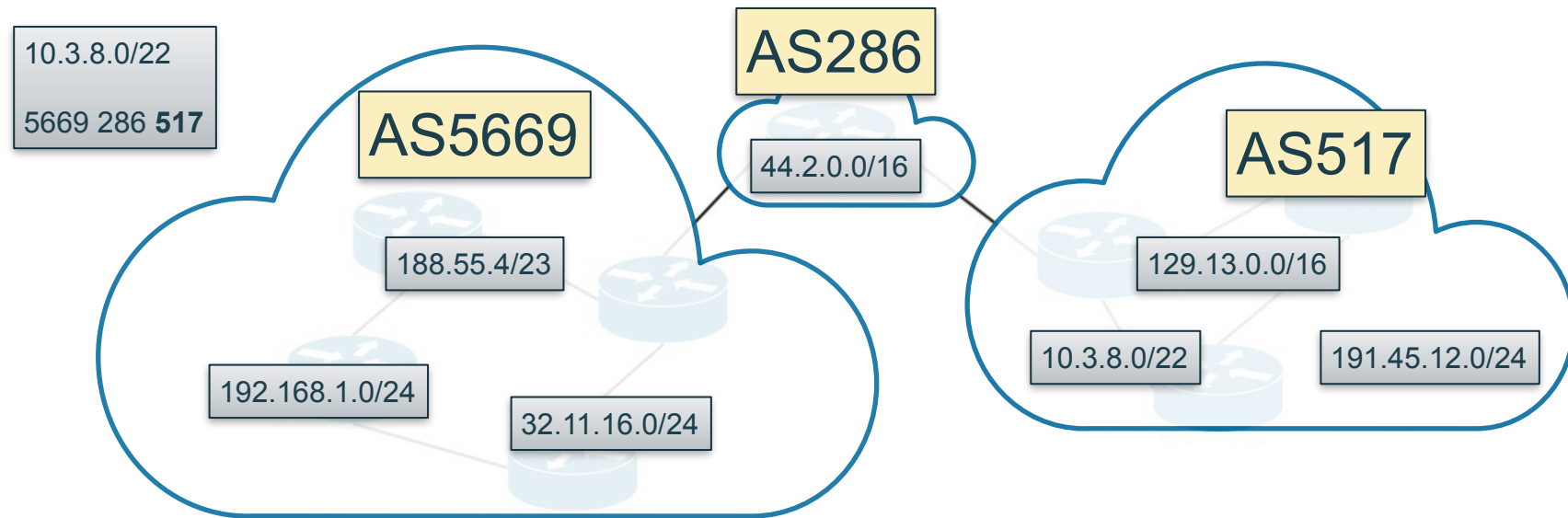
BGP - Key Concepts



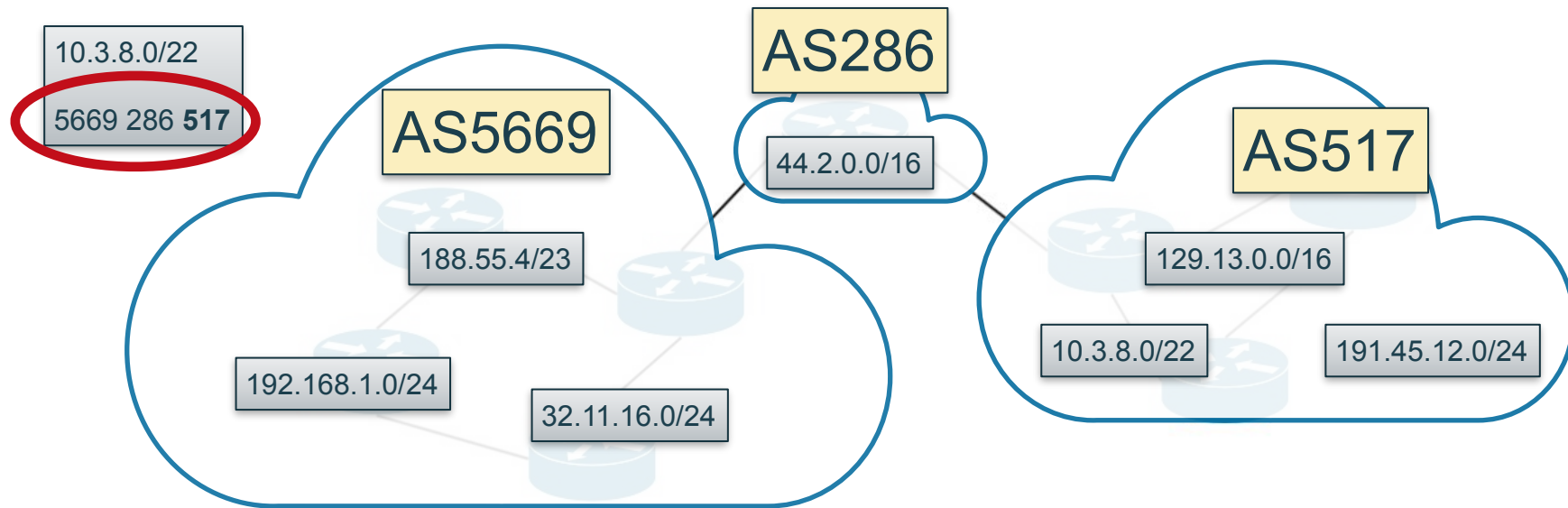
BGP - Key Concepts



BGP - Key Concepts



BGP - Key Concepts: The AS Path



A real live example



A real live example

```
asd2-rs-02>show bgp ipv4 unicast 129.13.0.0
Load for five secs: 1%/0%; one minute: 4%; five minutes: 5%
Time source is NTP, 09:14:07.268 UTC Thu Aug 17 2017
BGP routing table entry for 129.13.0.0/16, version 2944571
Paths: (13 available, best #10, table default)
....
 125 286 517
    134.222.85.126 from 134.222.85.126 (134.222.85.126)
      Origin IGP, metric 0, localpref 80, valid, internal
      Community: 286:18 286:19 286:28 286:29 286:49 286:800 286:888
```

A real live example

Prefix

```
asd2-rs-02>show bgp ipv4 unicast 129.13.0.0
Load for five secs: 1%/0%; one minute: 4%; five minutes: 5%
Time source is NTP, 09:14:07.268 UTC Thu Aug 17 2017
BGP routing table entry for 129.13.0.0/16, version 2944571
Paths: (13 available, best #10, table default)
....
 125 286 517
    134.222.85.126 from 134.222.85.126 (134.222.85.126)
      Origin IGP, metric 0, localpref 80, valid, internal
      Community: 286:18 286:19 286:28 286:29 286:49 286:800 286:888
```

A real live example

Prefix

```
asd2-rs-02>show bgp ipv4 unicast 129.13.0.0
Load for five secs: 1%/0%; one minute: 4%; five minutes: 5%
Time source is NTP, 09:14:07.268 UTC Thu Aug 17 2017
BGP routing table entry for 129.13.0.0/16, version 2944571
Paths: (13 available, best #10, table default)
...
125 286 517
  134.222.85.126 from 134.222.85.126 (134.222.85.126)
    Origin IGP, metric 0, localpref 80, valid, internal
    Community: 286:18 286:19 286:28 286:29 286:49 286:800 286:888
```

AS-Path

A real live example

```
asd2-rs-02>show bgp ipv4 unicast 129.13.0.0
Load for five secs: 1%/0%; one minute: 4%; five minutes: 5%
Time source is NTP, 09:14:07.268 UTC Thu Aug 17 2017
BGP routing table entry for 129.13.0.0/16, version 2944571
Paths: (13 available, best #10, table default)
```

....

125 286 517

134.222.85.126 from 134.222.85.126 (134.222.85.126)

Origin IGP, metric 0, localpref 80, valid, internal

Community: 286:18 286:19 286:28 286:29 286:49 286:800 286:888

Prefix

AS-Path

Originator AS

A real live example

```
asd2-rs-02>show bgp ipv4 unicast 129.13.0.0
Load for five secs: 1%/0%; one minute: 4%; five minutes: 5%
Time source is NTP, 09:14:07.268 UTC Thu Aug 17 2017
BGP routing table entry for 129.13.0.0/16, version 2944571
Paths: (13 available, best #10, table default)
```

....

125 286 517

134.222.85.126 from 134.222.85.126 (134.222.85.126)

Origin IGP, metric 0, localpref 80, valid, internal

Community: 286:18 286:19 286:28 286:29 286:49 286:800 286:888

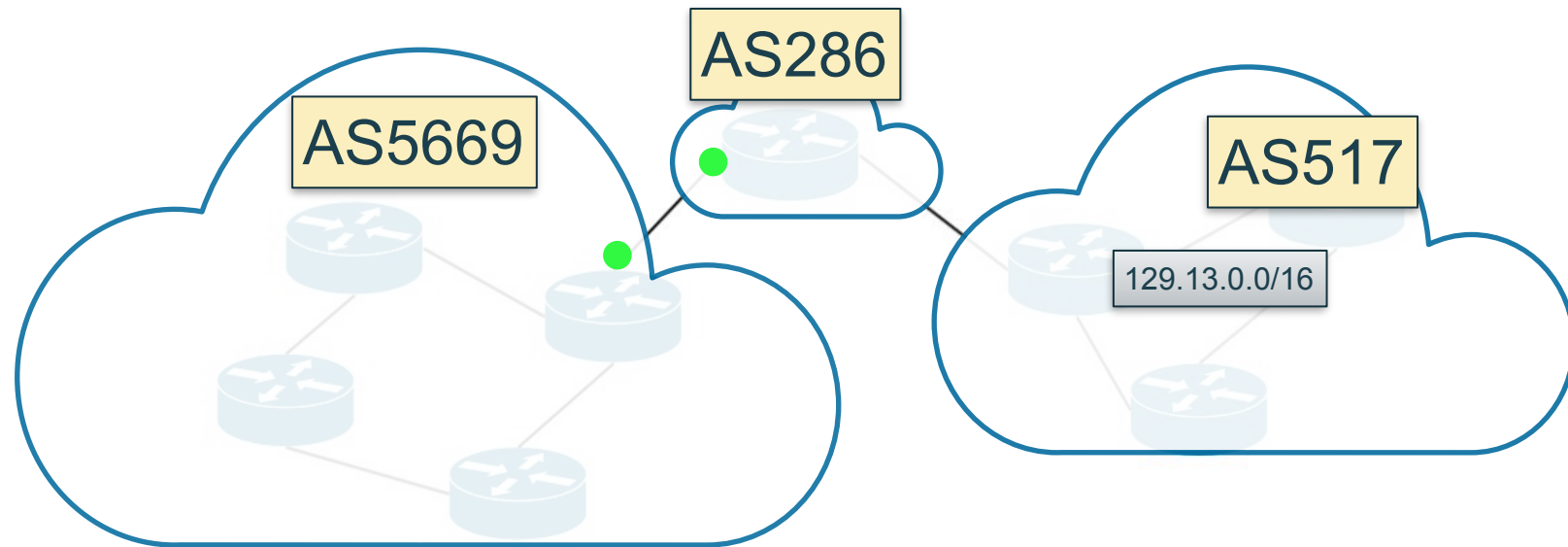
Prefix

AS-Path

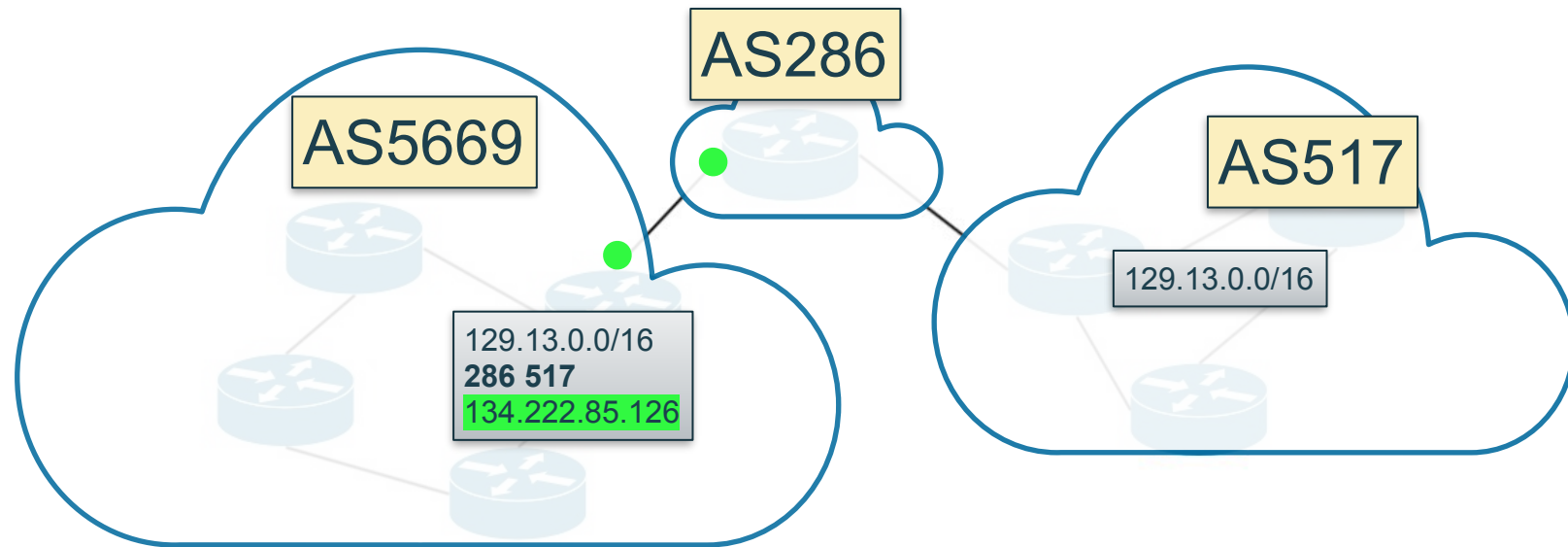
Next Hop IP

Originator AS

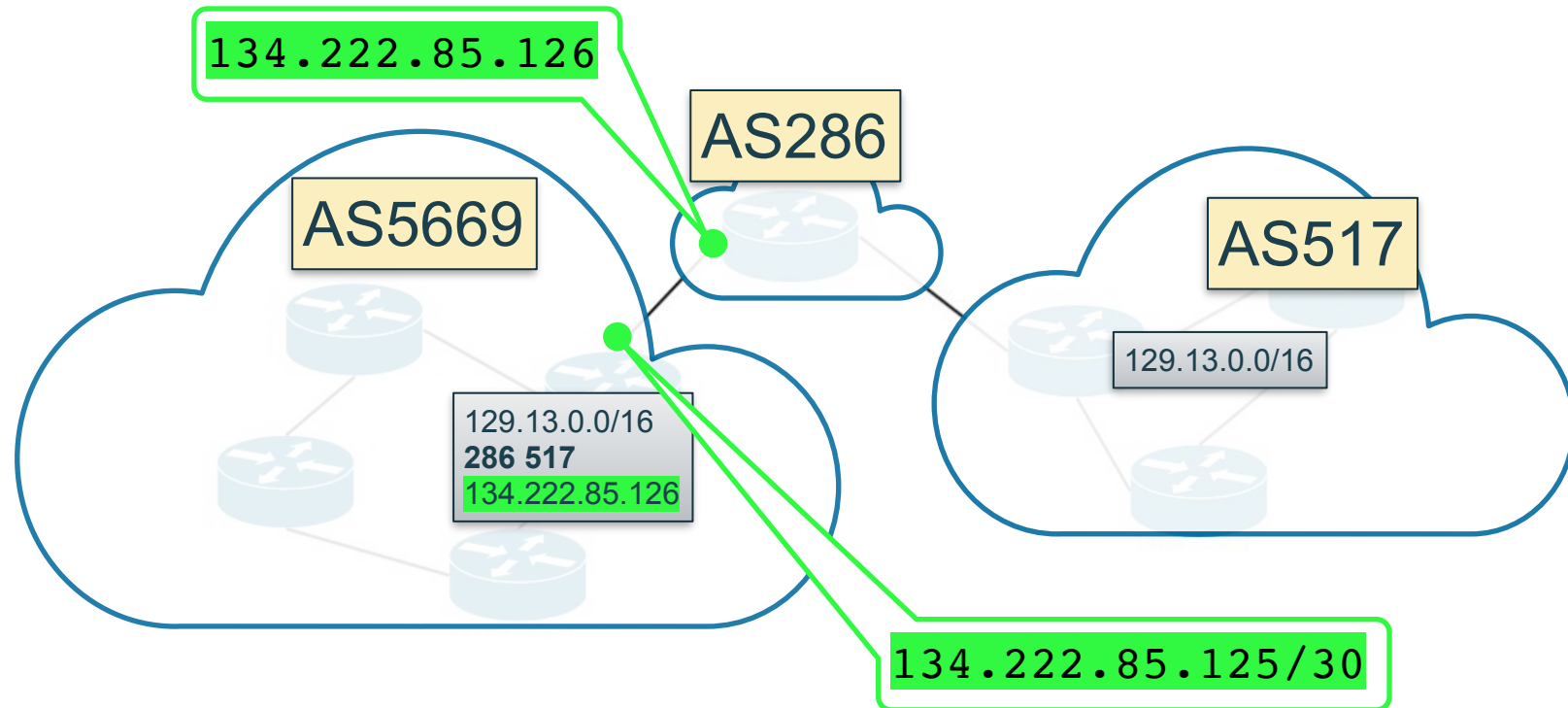
BGP - Key Concepts



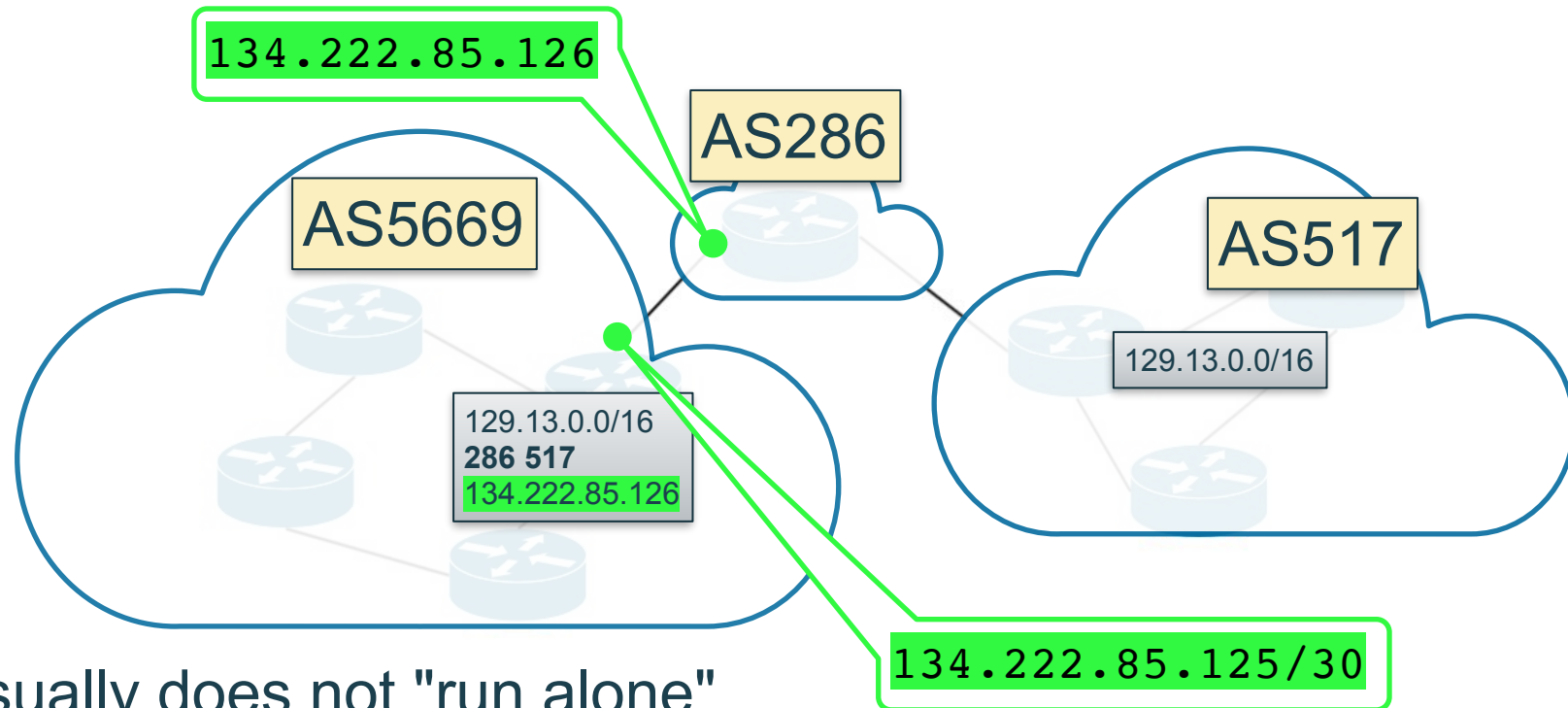
BGP - Key Concepts



BGP - Key Concepts: Next Hop Address

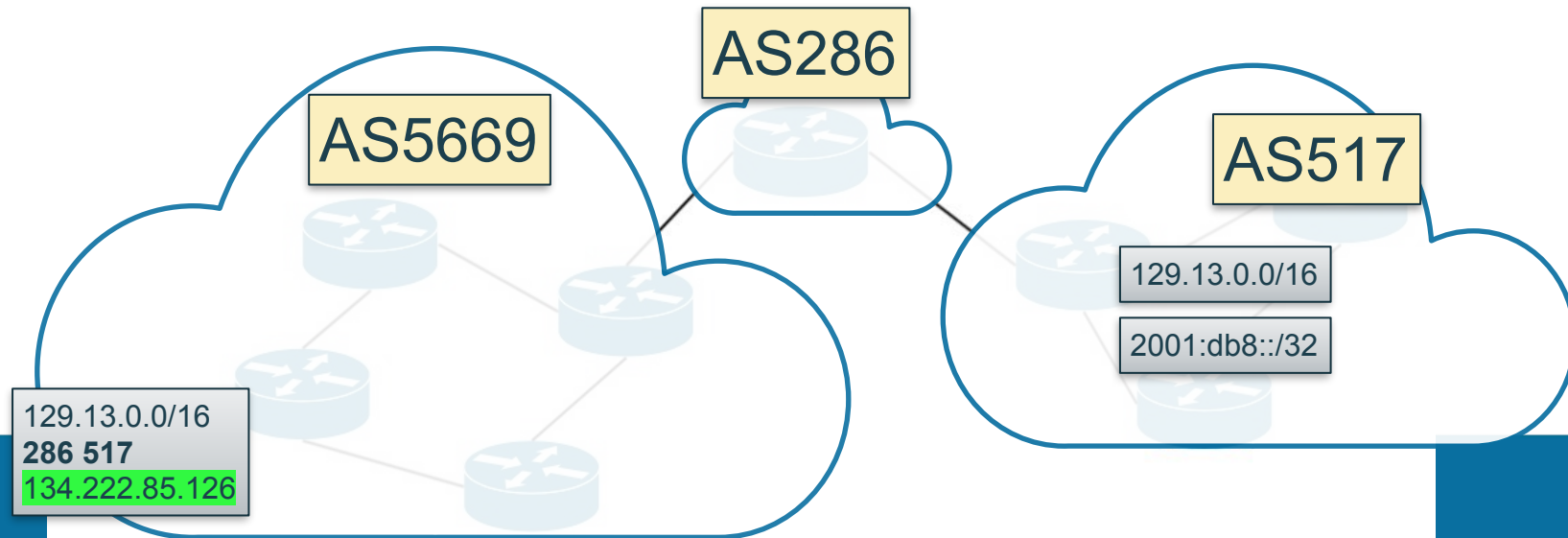


BGP - Key Concepts: Next Hop Address



BGP usually does not "run alone"
Another routing protocol is needed to distribute interface addresses

BGP - Key Concepts: Summary

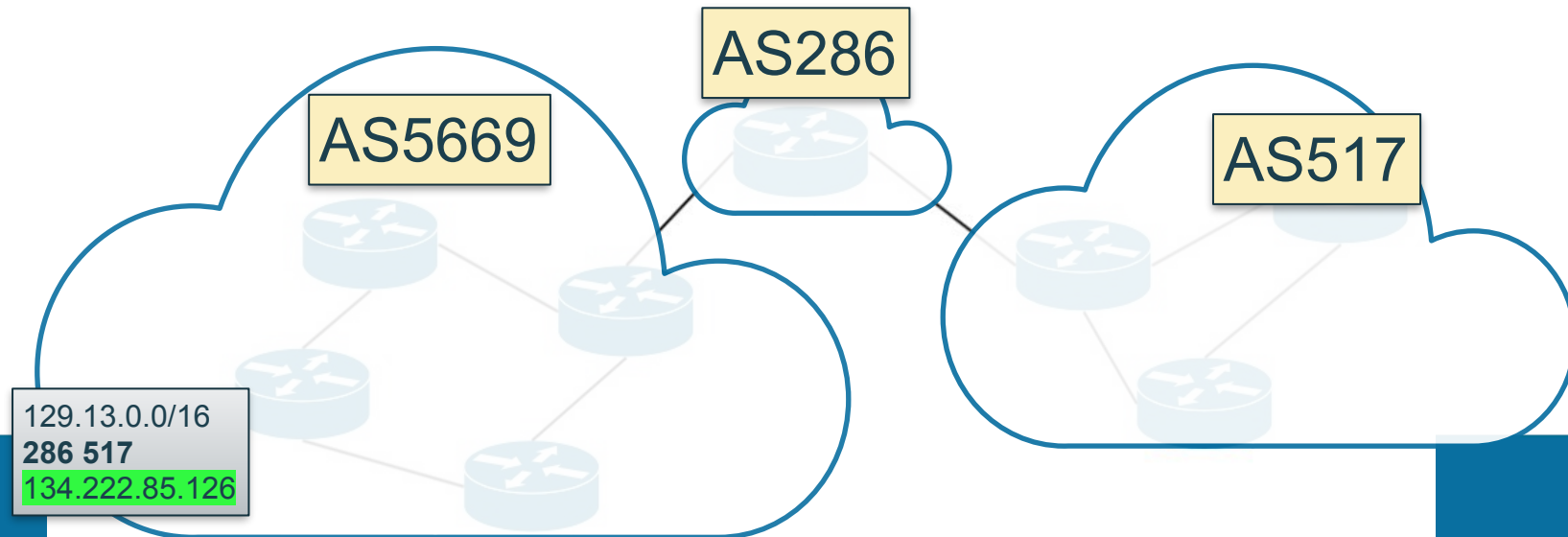


BGP - Key Concepts: Summary

→ Prefixes

129.13.0.0/16

2001:db8::/32



BGP - Key Concepts: Summary

→ Prefixes

129.13.0.0/16

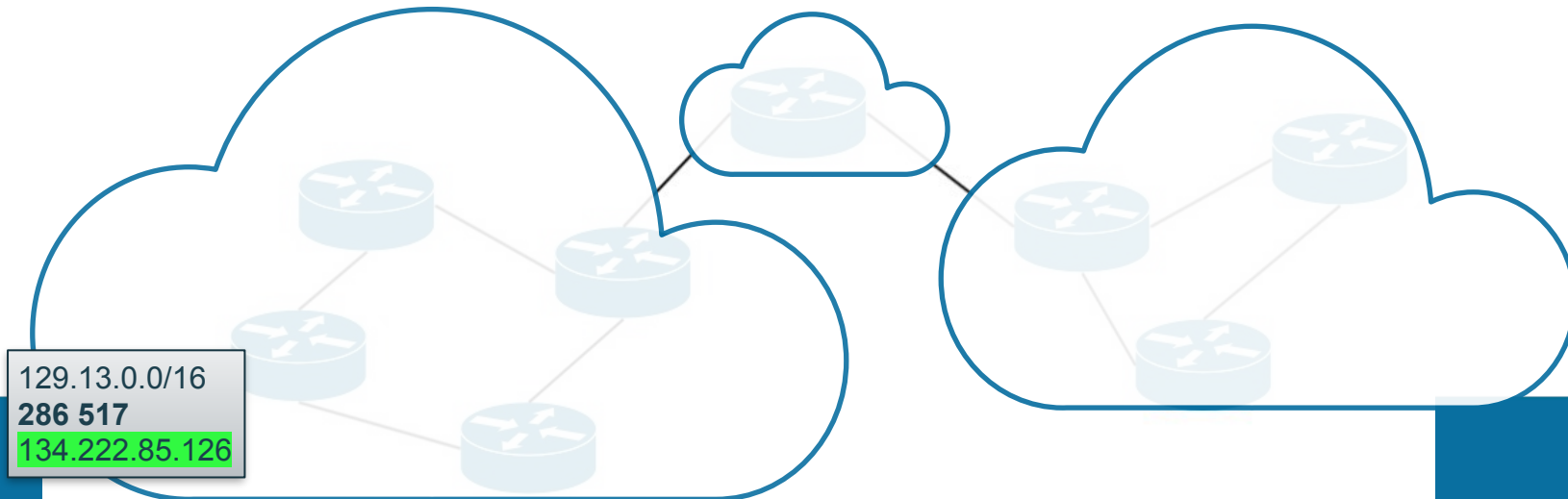
2001:db8::/32

→ AS Numbers

AS5669

AS286

AS517



BGP - Key Concepts: Summary

→ Prefixes

129.13.0.0/16

2001:db8::/32

→ AS Numbers

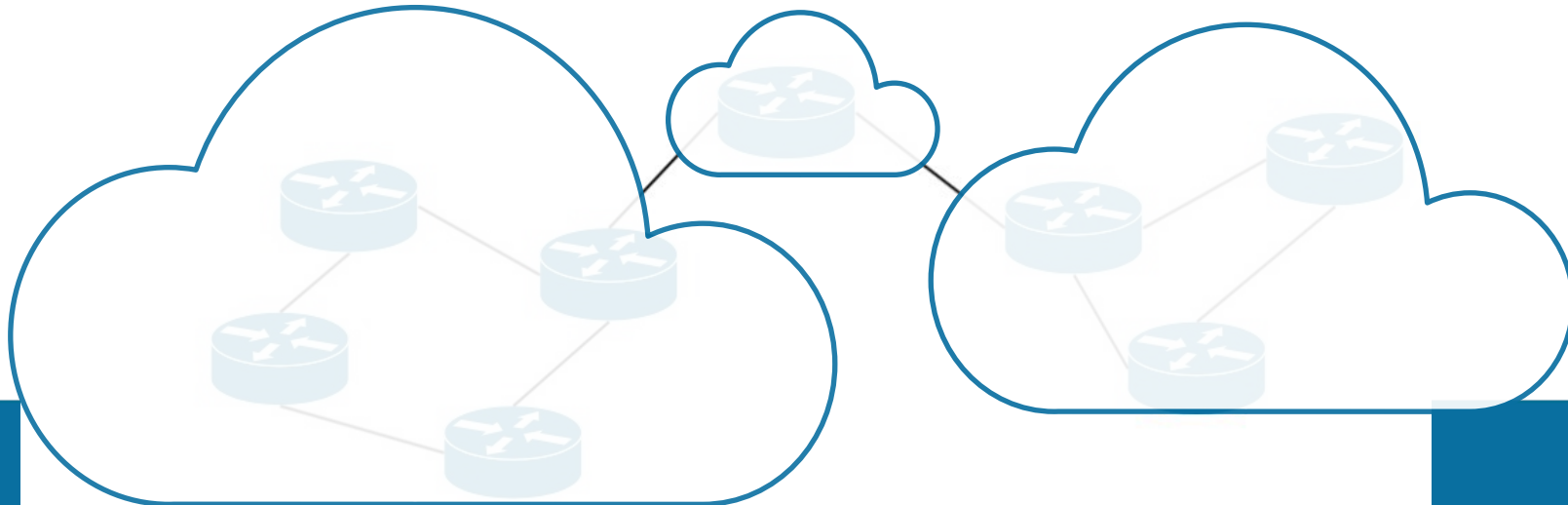
AS5669

AS286

AS517

→ AS Path

129.13.0.0/16
286 517
134 222.85.126



BGP - Key Concepts: Summary

→ Prefixes

129.13.0.0/16

2001:db8::/32

→ AS Numbers

AS5669

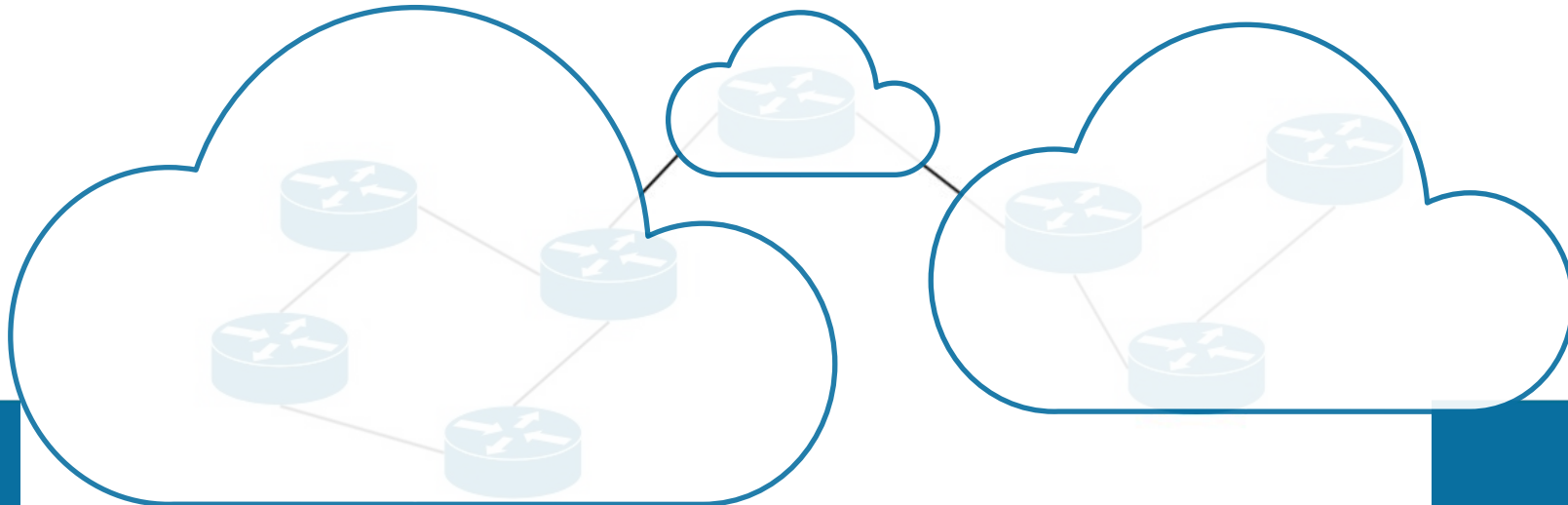
AS286

AS517

→ AS Path

129.13.0.0/16
286 517
134.222.85.126

Originator AS



BGP - Key Concepts: Summary

→ Prefixes

129.13.0.0/16

2001:db8::/32

→ AS Numbers

AS5669

AS286

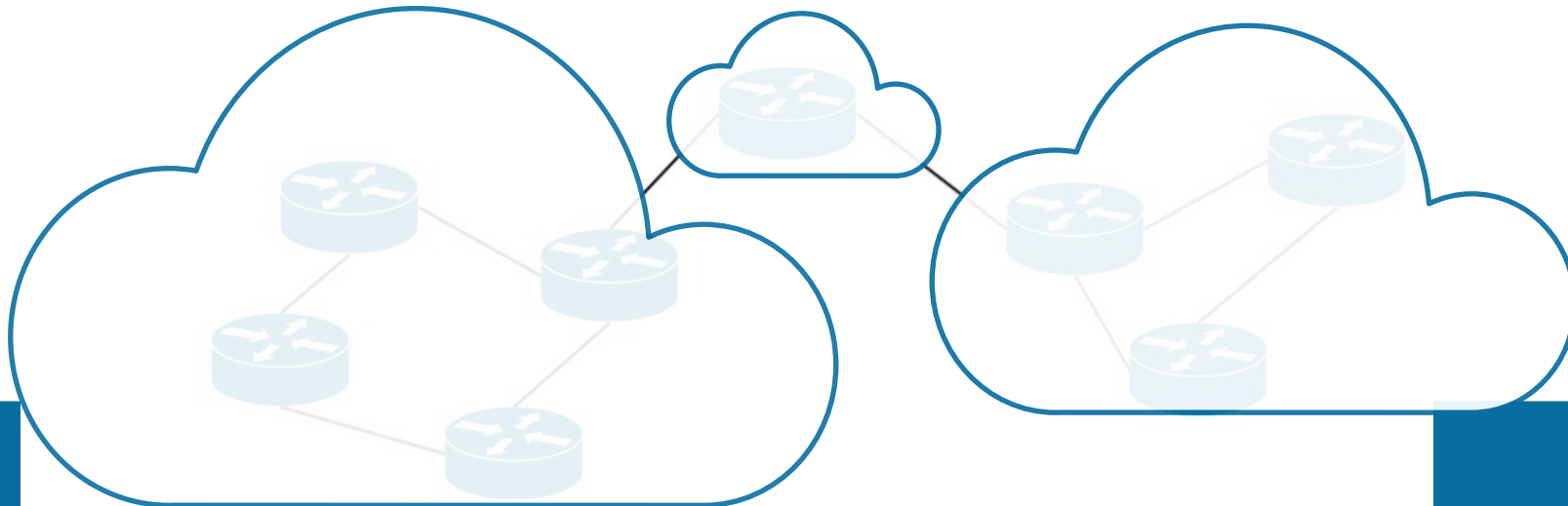
AS517

→ AS Path

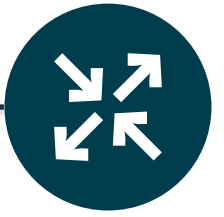
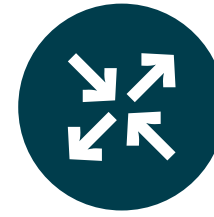
129.13.0.0/16
286 517
134.222.85.126

→ Next Hop

Originator AS



BGP: Example

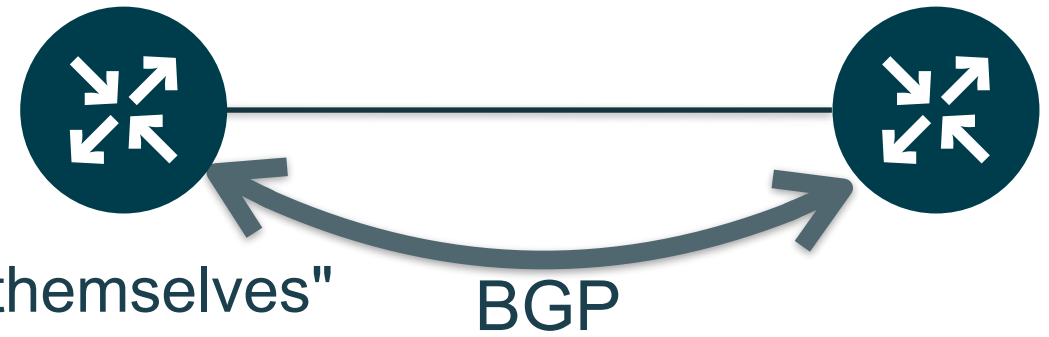


→ BGP speaking routers do not "find themselves"

BGP: Example

→ BGP speaking routers do not "find themselves"

- Everything needs to be configured!

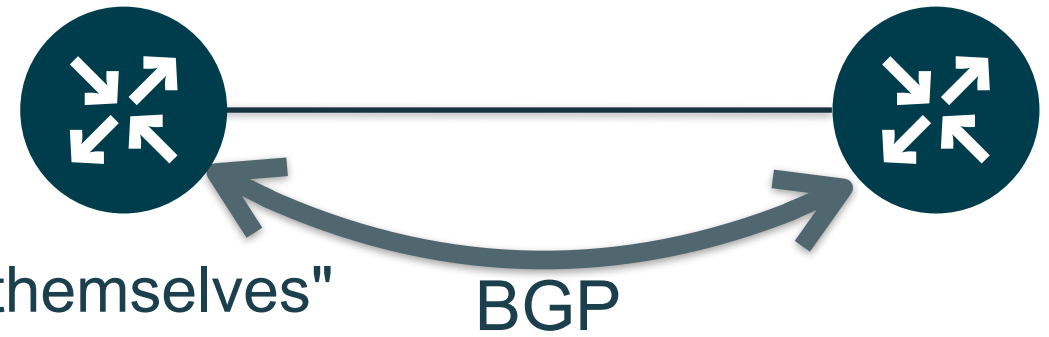


BGP: Example

→ BGP speaking routers do not "find themselves"

- Everything needs to be configured!

→ If you want to try yourself:



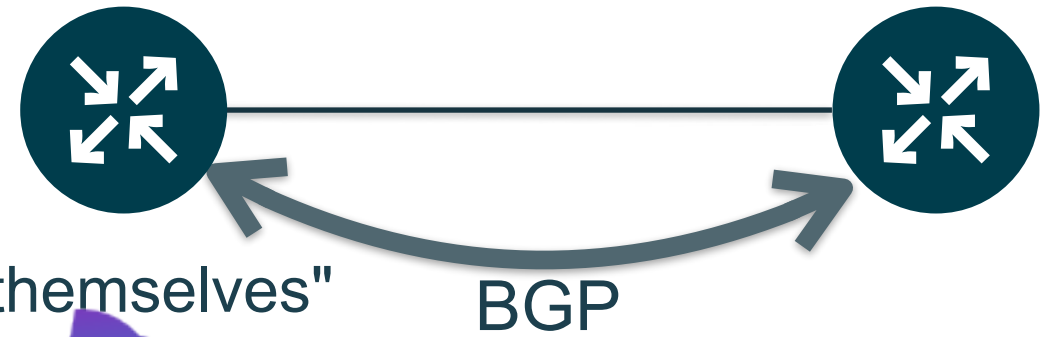
BGP: Example

→ BGP speaking routers do not "find themselves"

- Everything needs to be configured

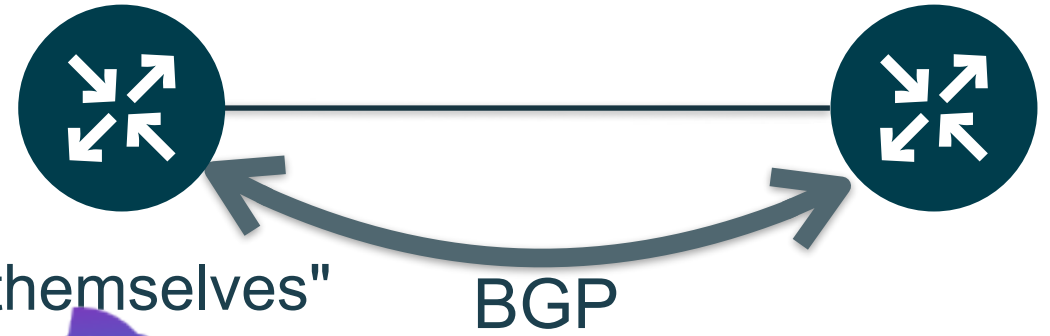
→ If you want to try yourself:

- Install GNS3: <https://gns3.com>



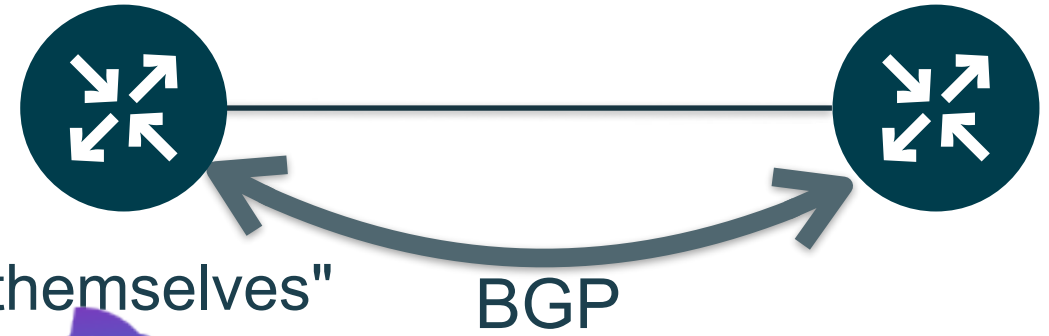
BGP: Example

- BGP speaking routers do not "find themselves"
 - Everything needs to be configured
- If you want to try yourself:
 - Install GNS3: <https://gns3.com>
 - Add a few routers (you need router software for this)



GNS3

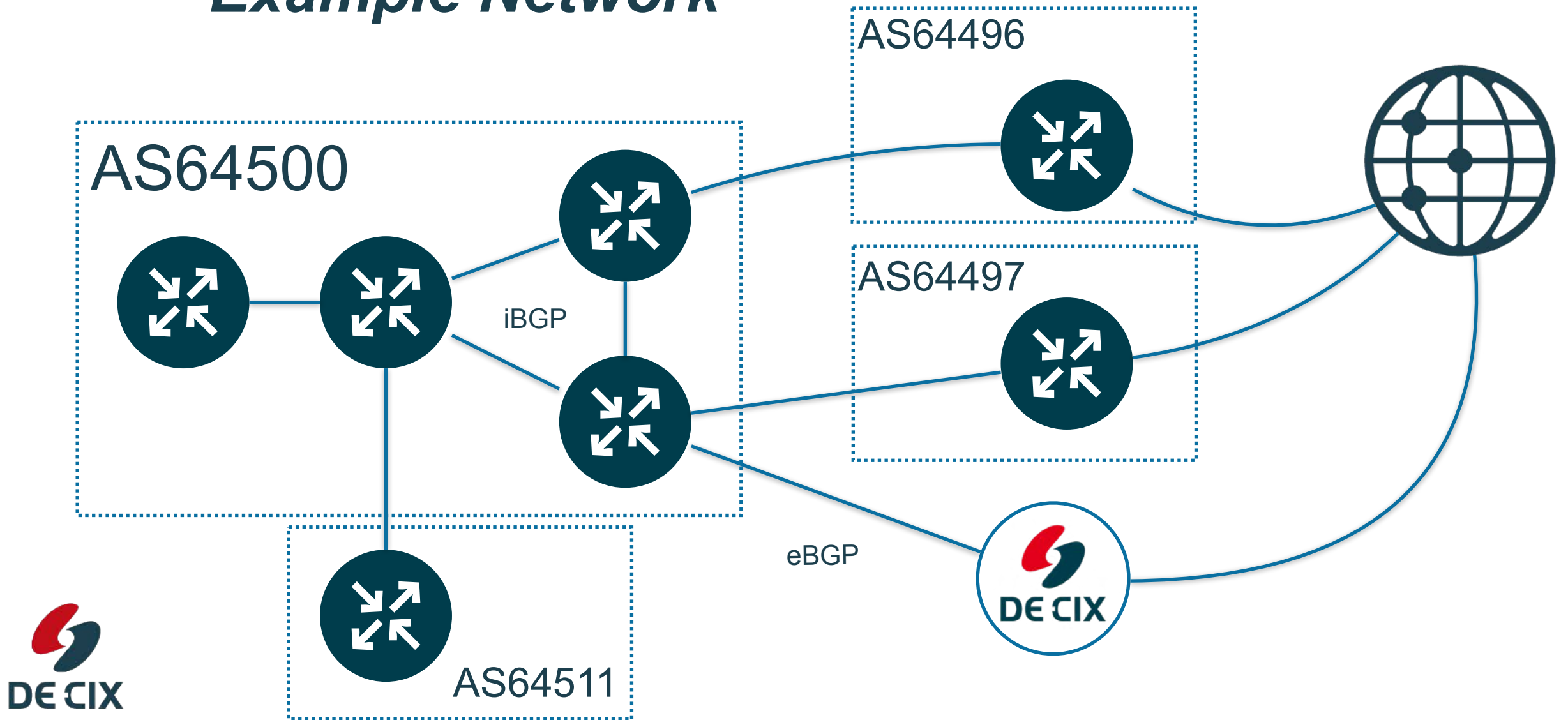
BGP: Example



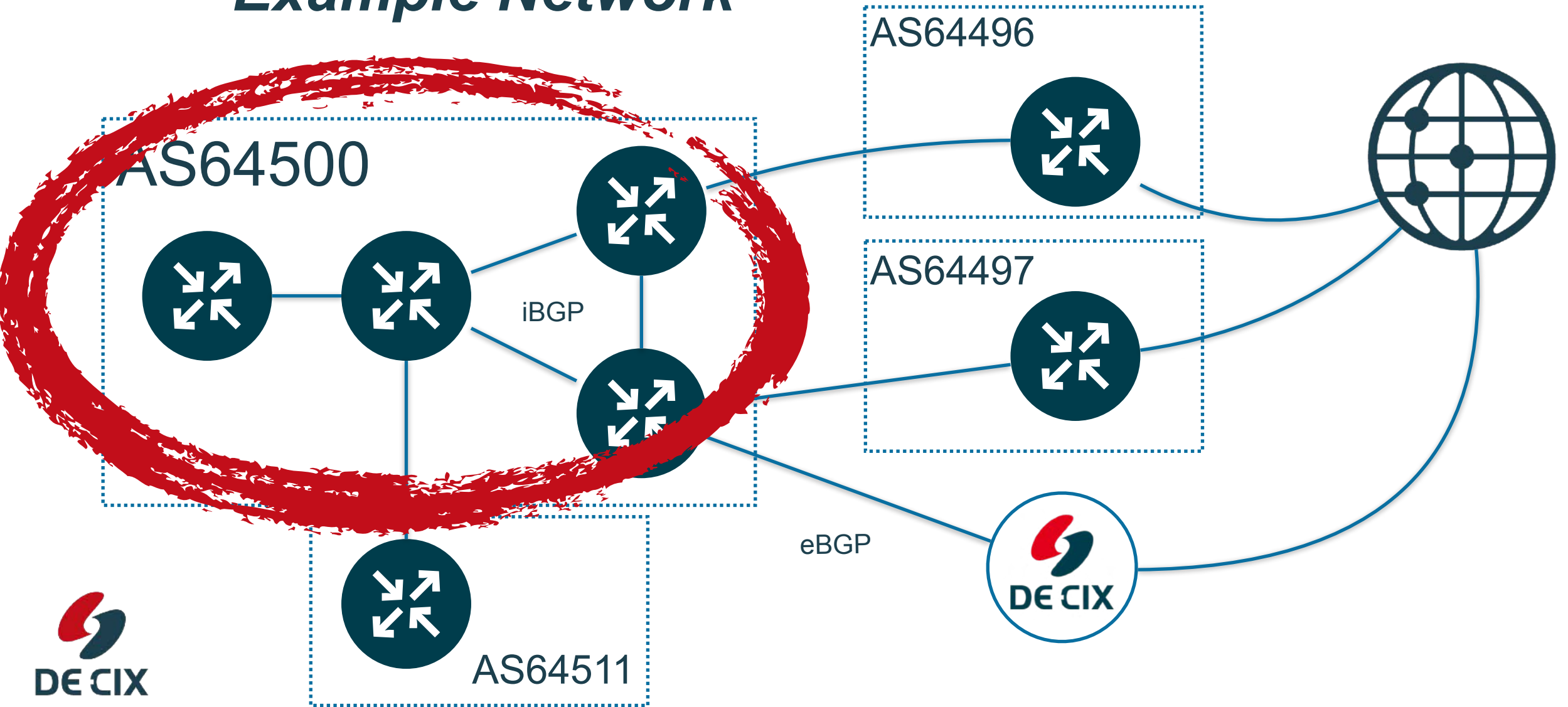
- BGP speaking routers do not "find themselves"
 - Everything needs to be configured
- If you want to try yourself:
 - Install GNS3: <https://gns3.com>
 - Add a few routers (you need router software for this)
 - Start configuring



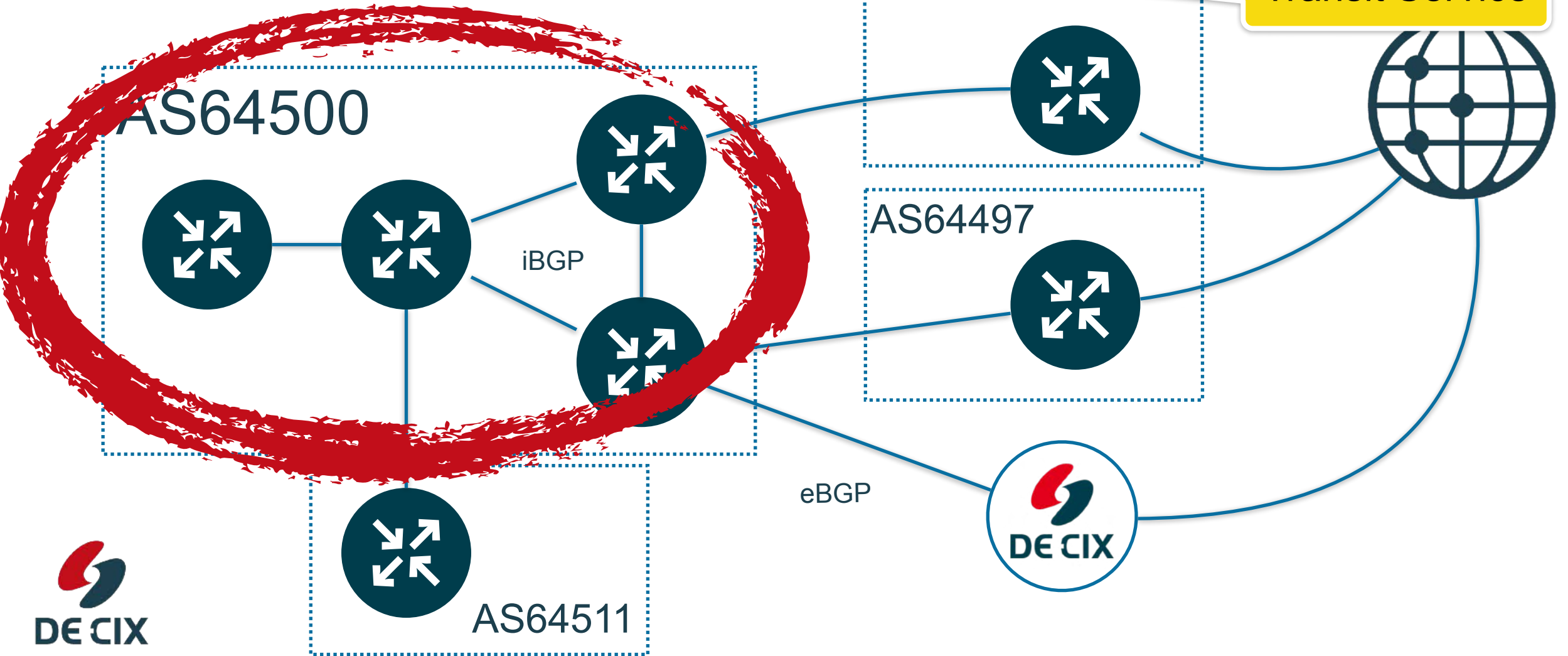
Example Network



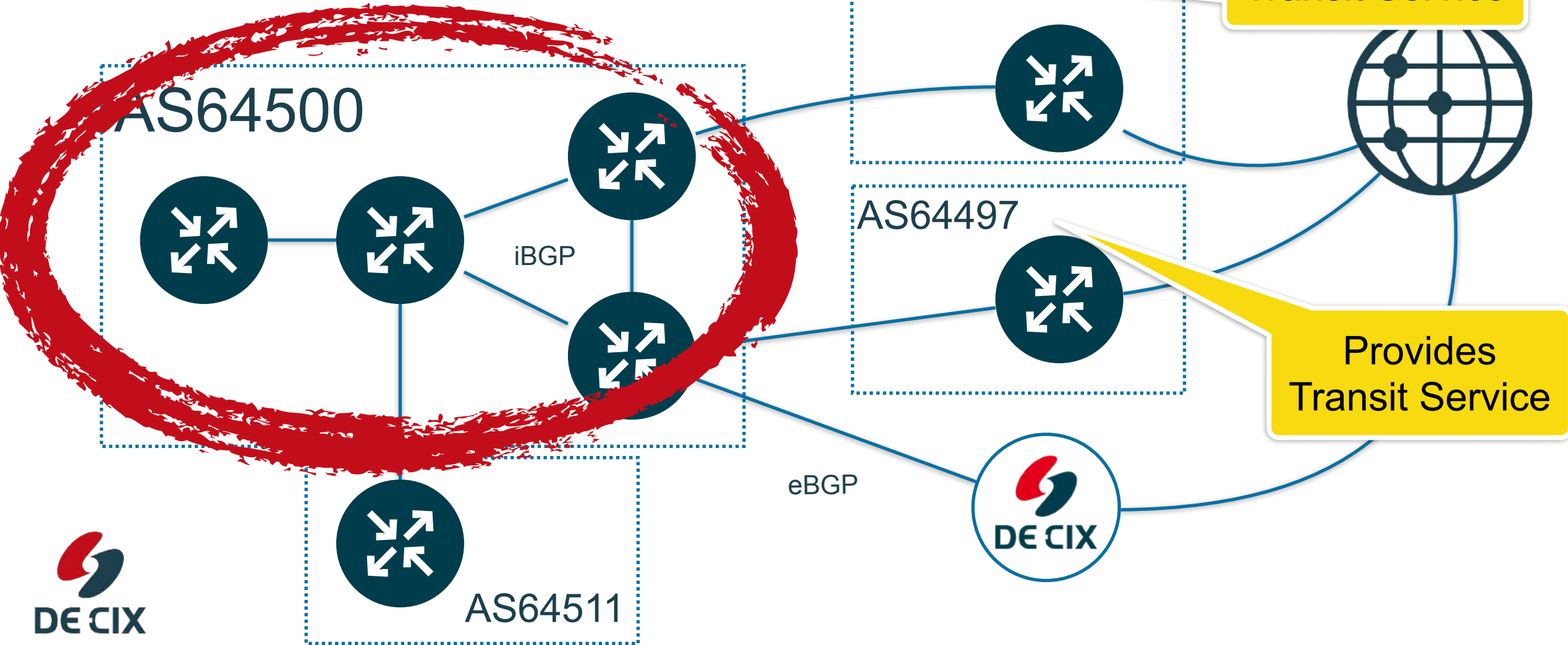
Example Network



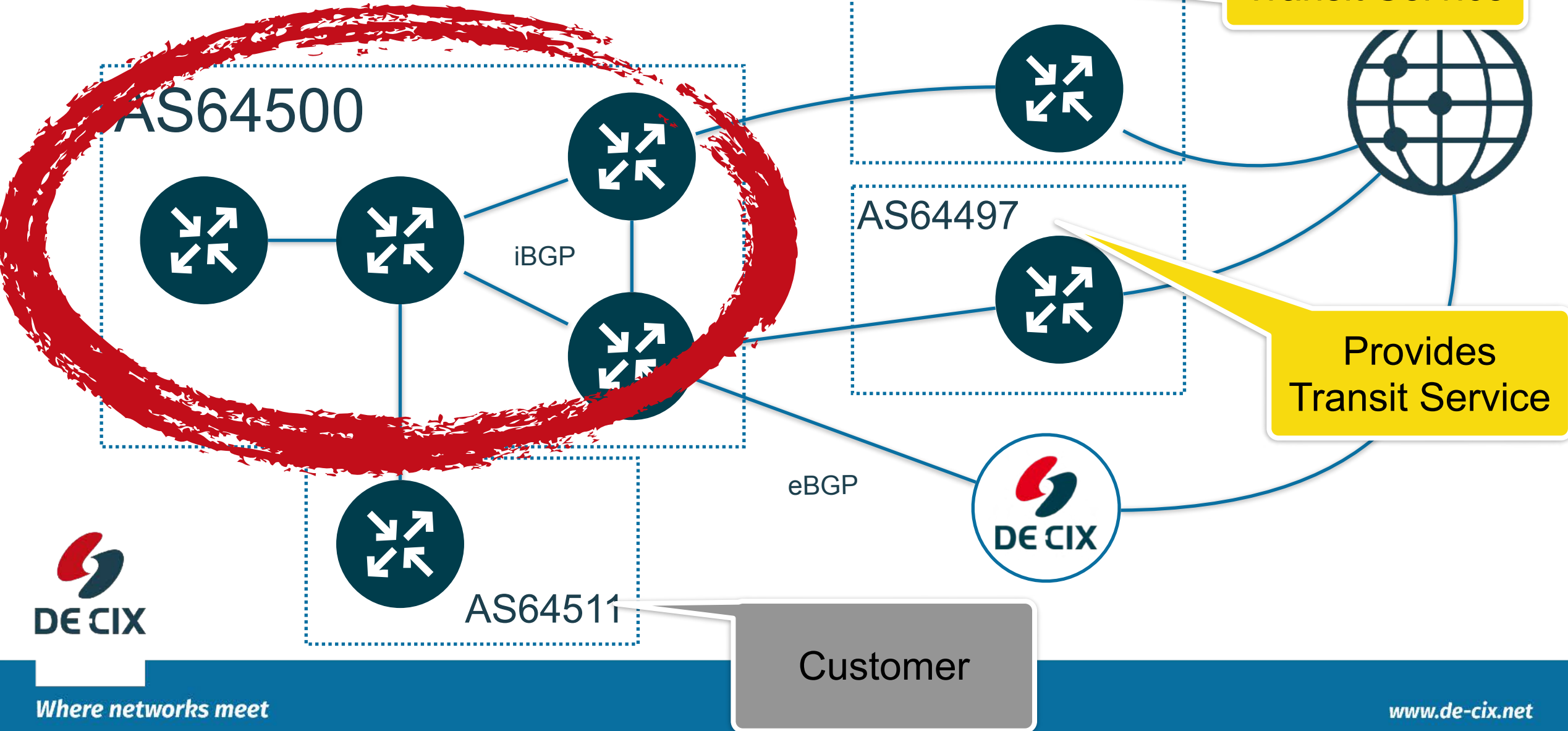
Example Network



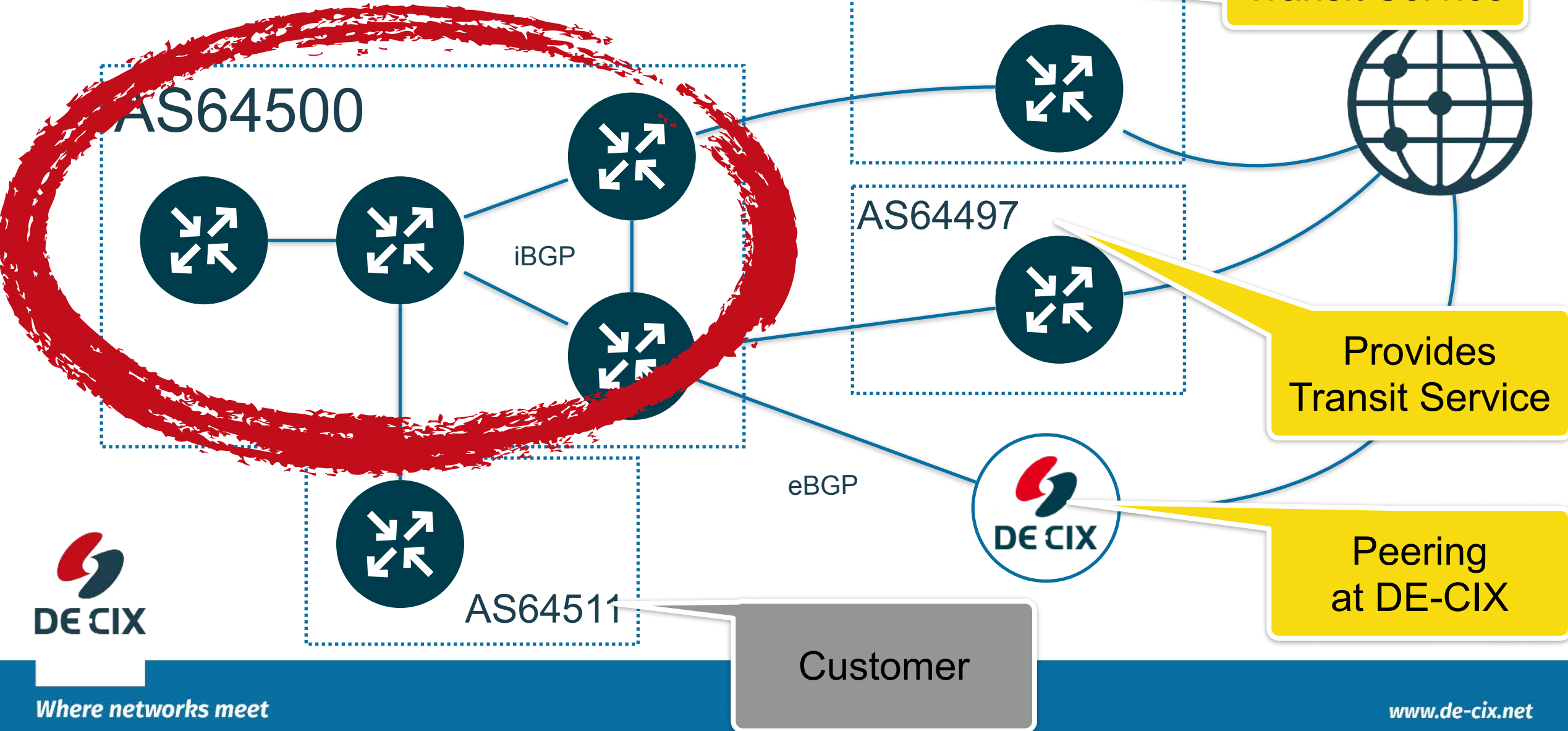
Example Network



Example Network



Example Network



BGP - Characteristics

→ Routers setup **BGP sessions** between each other

BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol

BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol
- BGP works **incremental**

BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol
- BGP works **incremental**
 - **after setup**, each router tells the other **all prefixes** it wants to announce

BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol
- BGP works **incremental**
 - **after setup**, each router tells the other **all prefixes** it wants to announce
 - then **only updates** are sent

BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol
- BGP works **incremental**
 - **after setup**, each router tells the other **all prefixes** it wants to announce
 - then **only updates** are sent
 - **withdraws**, if a prefix goes away

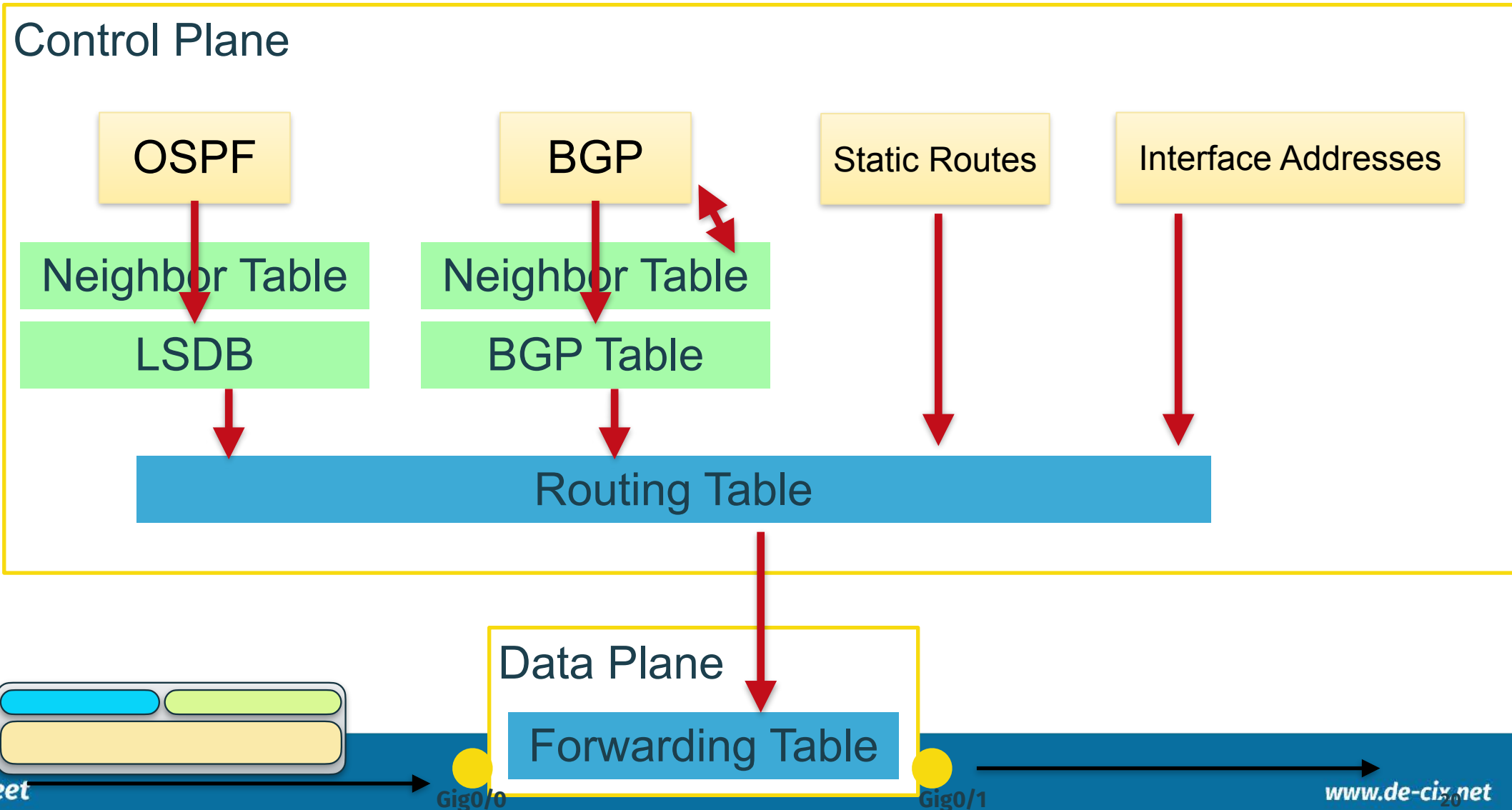
BGP - Characteristics

- Routers setup **BGP sessions** between each other
- BGP uses **TCP** as transport protocol
- BGP works **incremental**
 - **after setup**, each router tells the other **all prefixes** it wants to announce
 - then **only updates** are sent
 - **withdraws**, if a prefix goes away
 - **adds**, if a prefix is added

BGP - not re-inventing the wheel

- BGP uses TCP for transport
- so no need to re-implement features TCP already provides, like
 - reliable transport
 - flow control
 - framing
- as long as the TCP session is up, BGP assumes its neighbors are still there
- and have all the information sent to them

How a router works



DE-CIX Academy BGP Lab

Connecting to the experiments

Wolfgang Tremmel
academy@de-cix.net



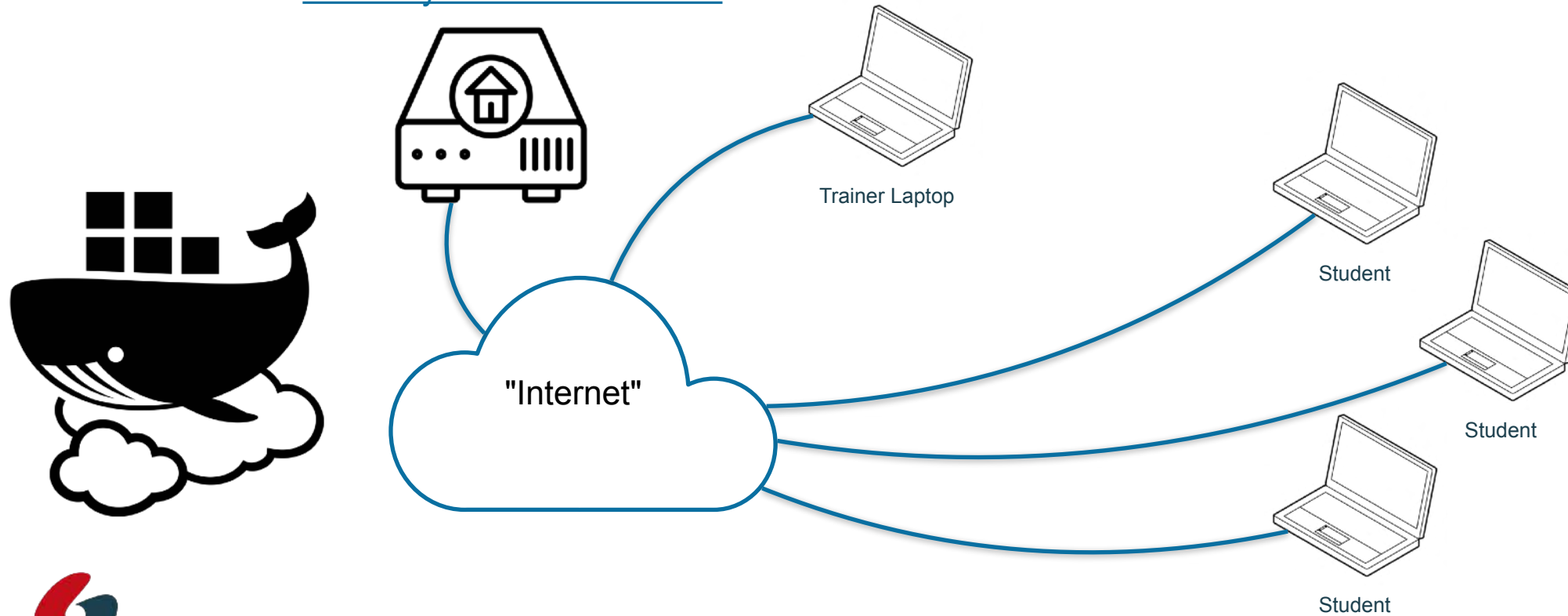
Network setup: Using Docker

academyserver01.de-cix.net



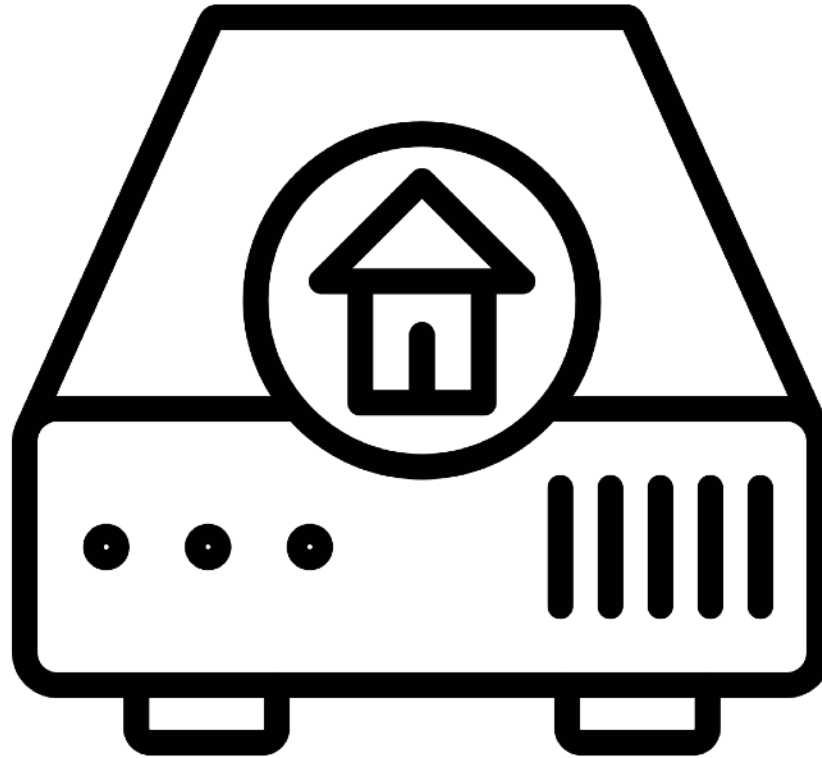
Network setup: Physical Setup

academyserver01.de-cix.net

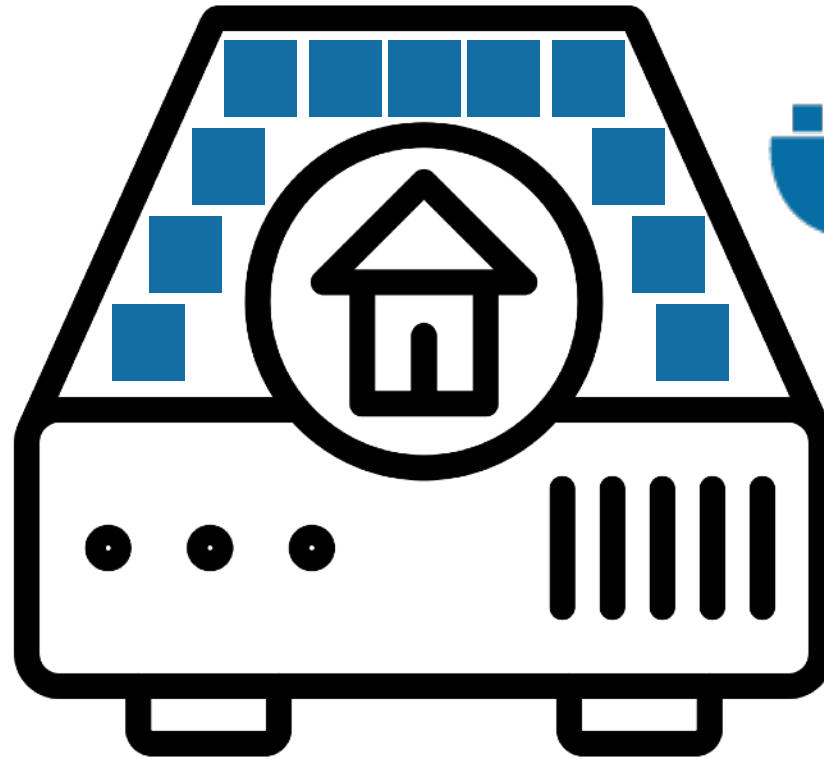


<http://bgplab.de-cix.net:9000/> or <http://go.de-cix.net/webinarlab>

Network setup: Using Docker



Network setup: Using Docker



Network setup: Using Docker

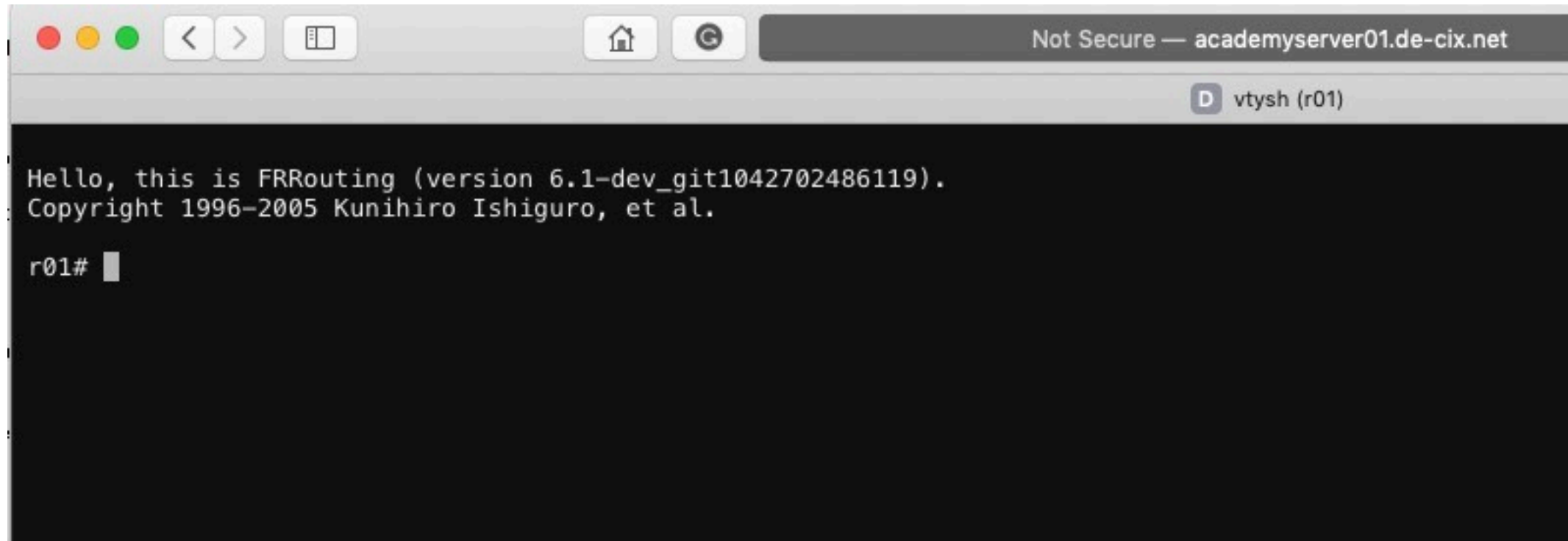
Docker Container

- Alpine Linux
- FRRouting Software
- Supervisord
- TTYd

Network setup: FRRouting



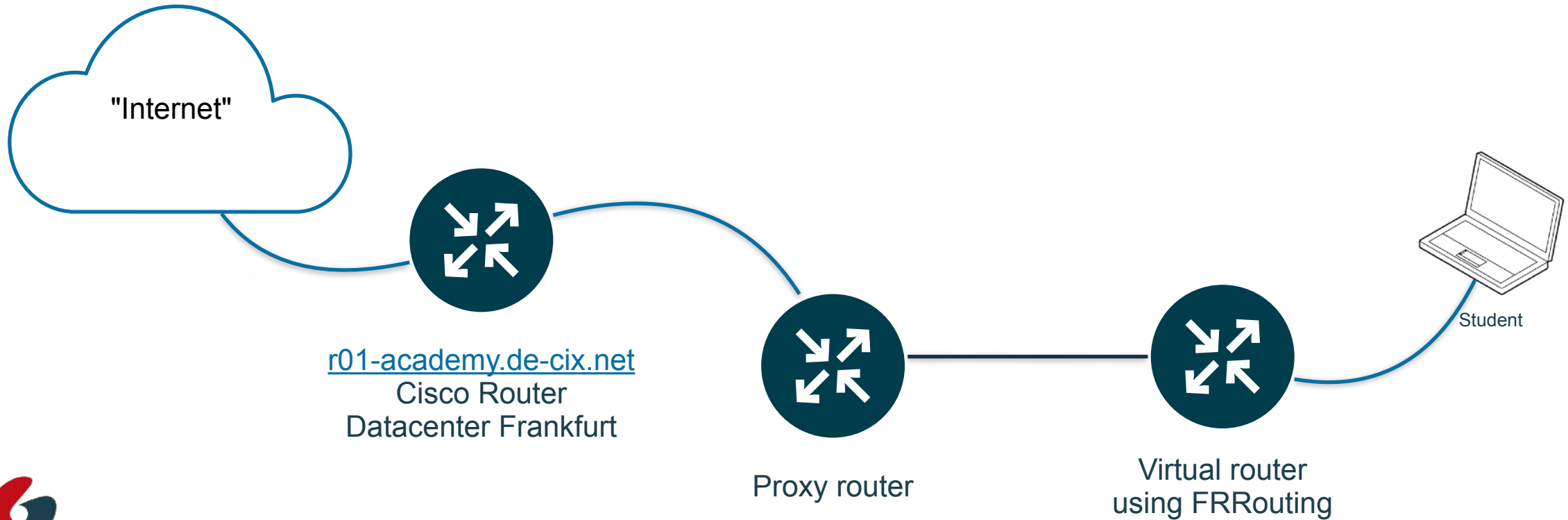
- Open Source routing daemon
 - based on Quagga
- Actively developed
- "Cisco-like" configuration syntax
- Not only BGP, but a lot of other protocols as well
- See frrouting.org



The screenshot shows a web browser window with a macOS-style title bar. The address bar displays "Not Secure — academyserver01.de-cix.net". The page content is a terminal window titled "vtysh (r01)". The terminal text reads: "Hello, this is FRRouting (version 6.1-dev_git1042702486119). Copyright 1996-2005 Kunihiro Ishiguro, et al." followed by a prompt "r01#" and a cursor.

```
Not Secure — academyserver01.de-cix.net
vtysh (r01)
Hello, this is FRRouting (version 6.1-dev_git1042702486119).
Copyright 1996-2005 Kunihiro Ishiguro, et al.
r01#
```

Today's experiment



Connect now

- use <http://go.de-cix.net/webinarlab>
- redirects you to an automatic distributor
 - please open this **once only!**
 - system is still experimental
- automatic distributor redirects you to a virtual router

Experiment: Connecting to your router



experiment 00

Thank you!

academy@de-cix.net



Interested in more webinars? Please subscribe to our mailing list at <https://lists.de-cix.net/www/subscribe/academy>

Links and further reading

Links visited during the webinar

→ RFCs

RFCs are Internet standards issued by the [Internet Engineering Task Force \(IETF\)](#)

- [RFC4632](#): Classless Inter-domain routing (CIDR)
- [RFC4291](#): IPv6 addressing architecture
- [RFC827](#): Exterior Gateway Architecture (EGP) (historical, obsolete)
- [RFC1930](#): Guidelines for creation, selection, and registration of an Autonomous System (AS)
- [RFC6793](#): BGP Support for Four-Octet Autonomous System (AS) Number Space

→ AS Numbers

- Giving AS numbers to the RIRs: [iana.org](#)
- Requesting an AS number, links for:
 - [ARIN](#)
 - [Lacnic](#)
 - [APNIC](#)
 - [RIPE NCC](#)
 - [Afrinic](#)



Interested in more webinars? Please subscribe to our mailing list at <https://lists.de-cix.net/www/subscribe/academy>

Links visited during the webinar

→ RFCs

RFCs are Internet standards issued by the [Internet Engineering Task Force \(IETF\)](#)

- [RFC827](#): Exterior Gateway Architecture (EGP) (historical, obsolete)
- [RFC1930](#): Guidelines for creation, selection, and registration of an Autonomous System (AS)
- [RFC4271](#) - A Border Gateway Protocol 4 (BGP-4)
- [RFC4632](#): Classless Inter-domain routing (CIDR)
- [RFC6793](#): BGP Support for Four-Octet Autonomous System (AS) Number Space

→ AS Numbers

- Giving AS numbers to the RIRs: [iana.org](#)
- Requesting an AS number, links for:
 - [ARIN](#)
 - [Lacnic](#)
 - [APNIC](#)
 - [RIPE NCC](#)
 - [Afrinic](#)



Interested in more webinars? Please subscribe to our mailing list at <https://lists.de-cix.net/www/subscribe/academy>



<https://de-cix.net/academy>



DE CIX

DE-CIX Management GmbH | Lindleyst. 12 | 60314 Frankfurt | Germany
Phone + 49 69 1730 902 0 | sales@de-cix.net | www.de-cix.net

Interested in more webinars? Please subscribe to our mailing list at <https://lists.de-cix.net/www/subscribe/academy>