

BGP Communities

BGP for networks who peer: Part 6

Wolfgang Tremmel
wolfgang.tremmel@de-cix.net



Where networks meet

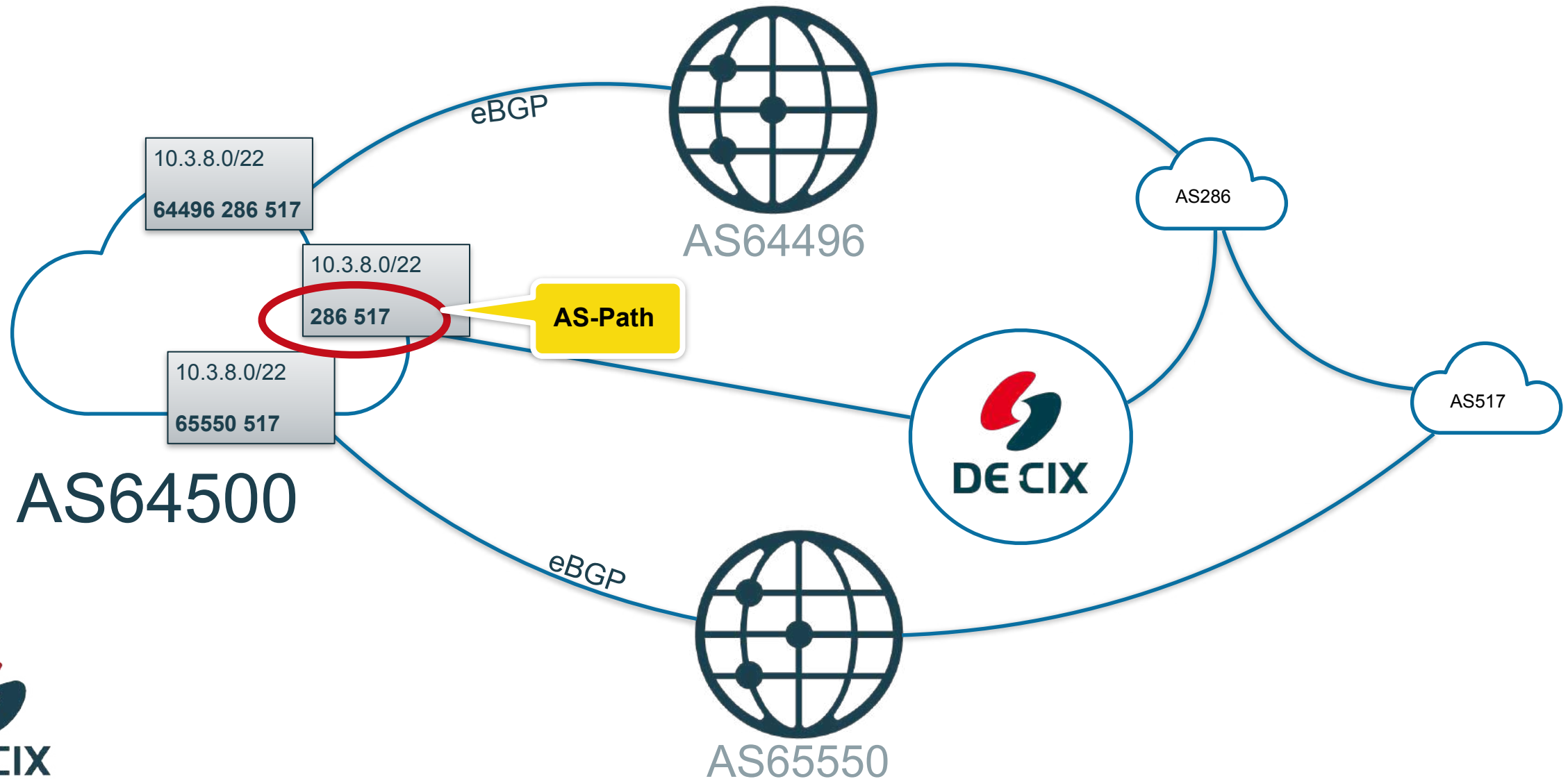
www.de-cix.net

BGP (new) Webinars Overview

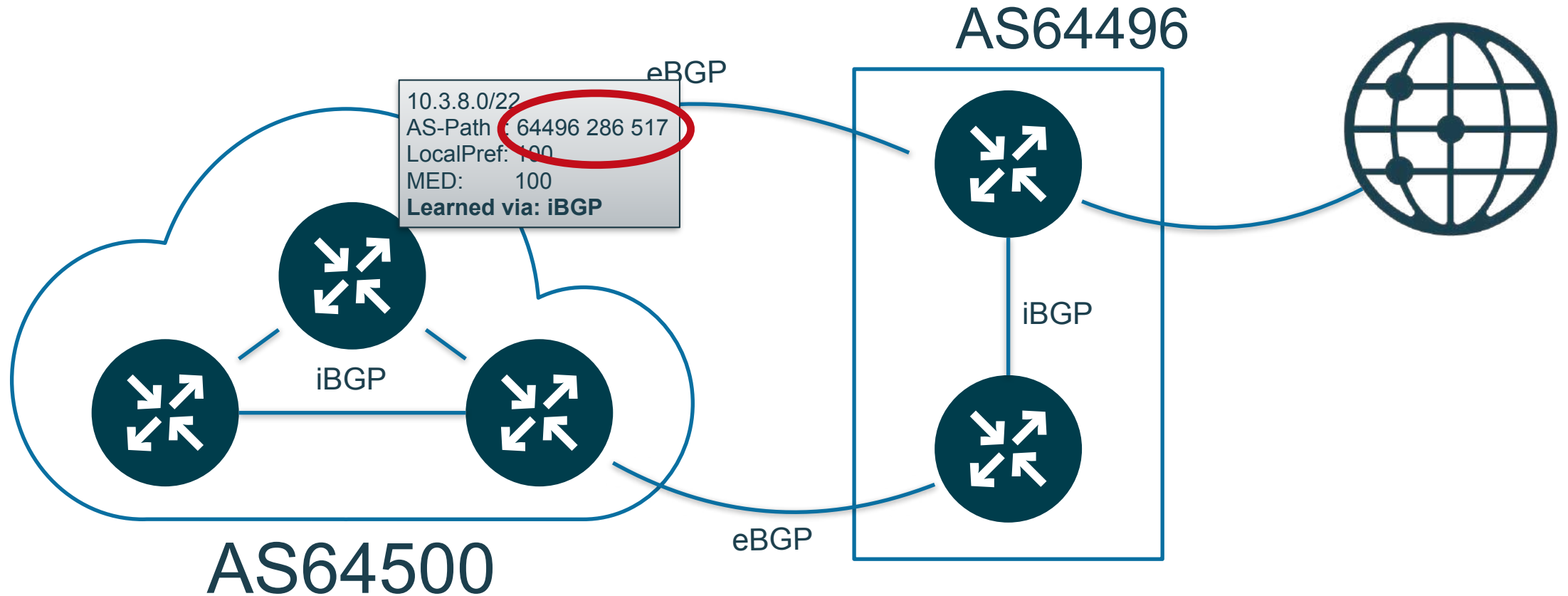
- 01 - Prefixes and AS numbers
- 02 - BGP Introduction
- 03a - Setting up iBGP
- 03b - Setting up eBGP
- 04 - Becoming multi-homed
- 05 - BGP Best Path Selection
- 06 - BGP Communities



We learned about BGP announcing prefixes



...and about some attributes of prefix announcements



Attributes of BGP prefixes

→ **Mandatory** attributes: have to be there

→ Example: AS-Path

→ **Optional** attribute: are, well, optional

→ Example: MED

→ **Transitive** attributes

→ are kept on the prefix and forwarded via BGP

→ **Non-transitive** attributes

→ are added to a prefix and not forwarded by the receiver

About BGP attributes

- **AS path**: mandatory
- **Origin**: mandatory
- **Next Hop**: mandatory
- **MED**: optional, non-transitive
- **Local Preference**: required for iBGP, not sent via eBGP

- IANA keeps track of attribute codes
 - Currently there are more than 40 attributes registered

Introducing:

BGP Communities



BGP Communities

→ A transitive, optional BGP attribute

→ **Transitive**: Once attached, it stays until removed

→ **Optional**: it does not have to be there

→ "BGP Communities are like a sticker on a suitcase"



"Original" BGP Communities

→ Definition:

"A community is a group of destinations which share some common property"

→ Introduced in RFC1997 in year 1996

→ A community is expressed by a 32Bit-Number

→ High 16 bit are the AS defining the low 16 bits

→ Notation: "6695:1000", "5669:32000"

→ You can attach as many communities as you like (within reason)

→ BGP max message size is 4096 Bytes

What are they useful for? Information!

198.51.100.0/24

80.81.192.15

from 80.81.192.15

Path: 1301 286 517

Origin IGP, metric 0, localpref 100, valid, external



Informational Communities

198.51.100.0/24

80.81.192.15

from 80.81.192.15

Path: 1301 286 517

Origin IGP, metric 0, localpref 100, valid, external

Received from: **Upstøngen**

Example: Encode geographical information

65010:1

Example: "1" here means geographical community

You may encode the continent here (if you are global) like:

- 1 = Europe
- 2 = North America
- 3 = Asia ...

ISO-Country-Codes here ...

250 - France

276 - Germany

840 - USA

Just an Example!

Example: Encode logical information

65010:2



Example: "2" here means logical source

Upstream? Peering? Customer?
1 = Upstream
2 = Private Peer
3 = Peer at an IXP
4 = Customer

More details here, like:

- Customer ID
- Upstream location
- up to you!

What are they useful for? Action!

198.51.100.0/24

80.81.192.15

from 80.81.192.15

Path: 65010

Origin IGP, metric 0, localpref 100, valid, external

Encoding up to you!

Announce to  DE CIX



Action Communities: Encoding

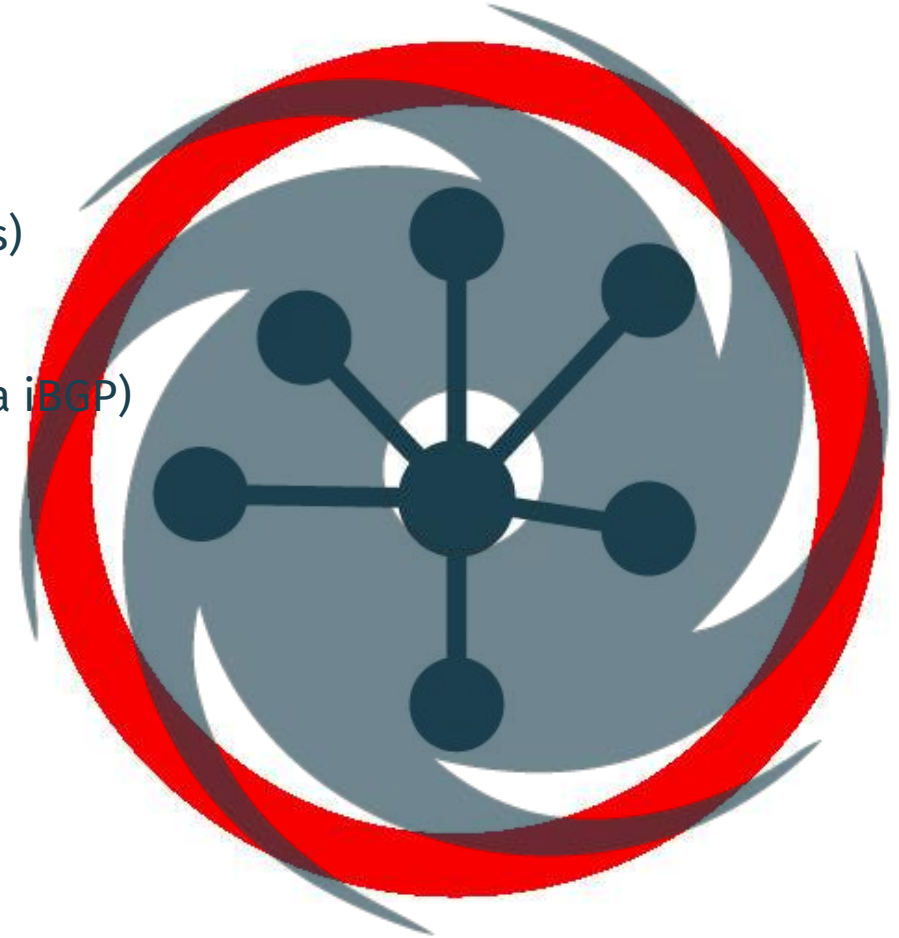
→ Again - you only have two 16bit numbers ... (with original BGP Communities)

→ Some Ideas ...

- If you want your customers to send you "actions"
 - You really should have them put your AS number into the first 16bit number
 - You **must scrub** everything they should not send on incoming
- Possible actions:
 - (not) announce to upstream, peers, customers
 - fine granular announcement control (geographically, by IXP, ...)
 - announce with longer AS path
 - change *local preference*
 - Blackhole

Action Communities: Well-Known

- A couple of communities are pre-defined by RFCs
- NO-EXPORT
 - Do not send the prefix to eBGP neighbours (other ASes)
- NO-ADVERTISE
 - Do not send the prefix to anyone (not even internal via iBGP)
- NO-PEER
 - Do not send to any peers
- BLACKHOLE
 - Sink all traffic to prefixes tagged with this community
 - Most commonly used with host routes
 - Implies NO-EXPORT

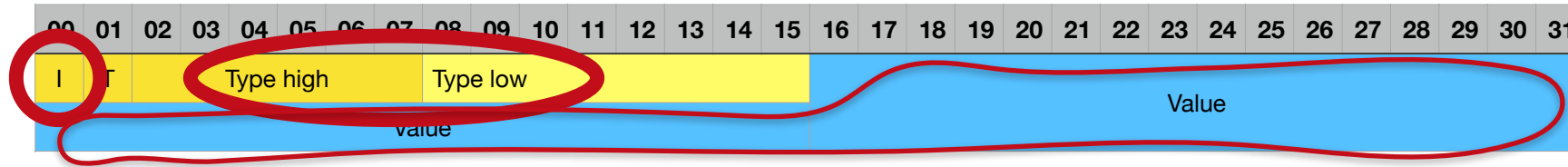


32Bit AS? No luck with original communities

65010:12345

- Two 16-bit numbers
- No way to encode a 32Bit AS number and something else ...
 - [RFC4360](#) - Extended Communities
- Extended Communities - Lots of new features
 - In total 2*32Bits
 - Introducing a "type" field
 - Possible to encode 16Bit Type, 32Bit AS, 16Bit Data

Extended Communities



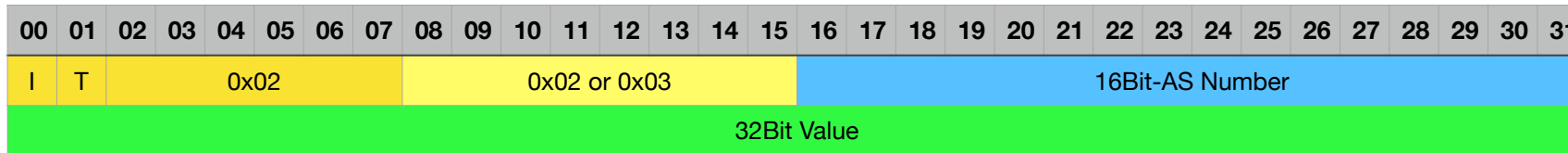
- **I** = Type is IANA assigned (= well known) or private
- **T = 0**: Transitive across AS borders
- **T = 1**: Non-Transitive - should be removed before forwarding to another AS
- **Type**: Types are either IANA-assigned or experimental. For a list of assigned types see the RFC
- **Value**: 48 Bits, meaning is dependent on type
- Standardized in 2006

Extended Communities and 32Bit ASes



- You can encode a 32Bit AS-Number
 - and a 16 Bit value

Extended Communities and 32Bit ASes



→ You can encode a 32Bit AS-Number

→ and a 16 Bit value

→ or a 16Bit AS-Number

→ and a 32 Bit value

→ 32Bit AS and 32Bit Value?

→ **not possible!**



Extended communities use cases

→ Notation:

- Similar to original communities: **RT:6500000:1234** or **RT:1234:6500000**

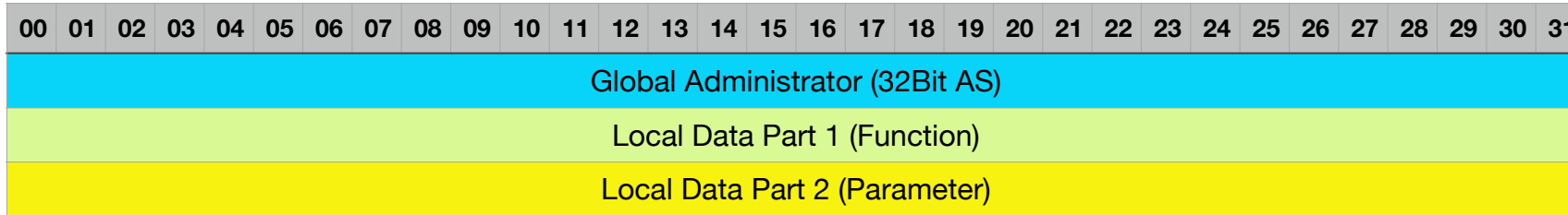
→ Disadvantages

- Only 48bits in total
- Only one 32Bit value is possible (and one 16Bit value)
- RT, RO and other types confusing to many operators

→ Conclusion

- Another community version was needed
- It took the IETF a while to realize that (11 years)

Introducing: Large Communities



→ Very simple - three 32Bit values (finally something useful)

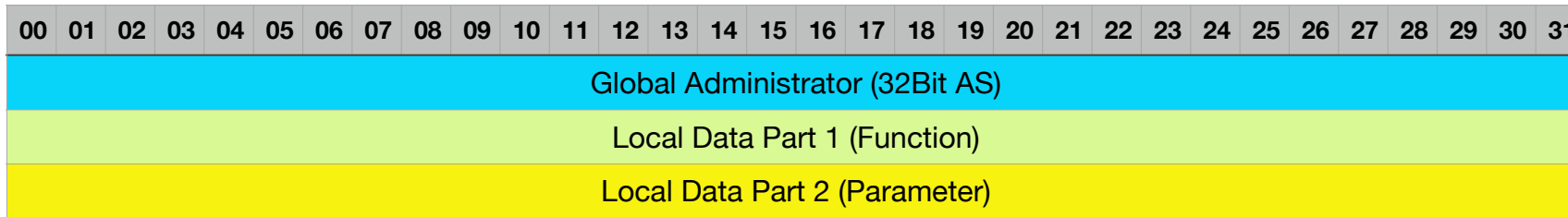
→ Global Administrator:

- An AS number (in 32Bit notation)
- Has defined meaning of two other fields
- May have published that meaning

→ Local Data

- Can be seen as "just two 32Bit numbers"
- Or as "Function" / "Parameter"

Large BGP Communities



→ Notation:

→ Similar to Original Communities: **196610:100:65000010**

→ Defined in two RFCs:

→ [RFC8092](#): BGP Large Communities Attribute

→ [RFC8195](#): Use of BGP Large Communities

→ A dedicated website exists: <http://largebgpcommunities.net>

→ Keeping track of Implementations, News etc.

Experiment: Working with BGP Communities



experiment04

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 and further reading

- BGP attribute types:
 - Registering new types: [RFC2042](#)
 - Published in [BGP Parameters](#) database at IANA
- Well-known communities:
 - Standard: see [IANA website](#)
 - Extended: [see IANA website](#)
- Provider examples:
 - DE-CIX Communities: <https://www.de-cix.net/en/resources/informational-bgp-communities>
 - DE-CIX Routeserver Guides: <https://www.de-cix.net/en/resources/route-server-guides>
 - KPN - AS286 - Community page: <https://as286.net/AS286-communities.html>
 - NTT - AS2914 - Community page: <https://www.us.ntt.net/support/policy/routing.cfm>
 - NTT Looking Glass: <https://www.us.ntt.net/support/looking-glass/>
- Lab: Download the DE-CIX Academy lab here: <https://bitbucket.org/decix-academy/dockerbgp/src/master/>



BGP Communities

	Original Communities	Extended Communities	Large Communities
Defined in	RFC1997	RFC4360	RFC8092
Published	August 1996	February 2006	February 2017
Additional RFCs	RFC1998 , RFC3765 , RFC7999	RFC4384 , RFC5668 , RFC5701 , RFC7153 , RFC8097	RFC8195 also: http:// largebgpcommunities.net
Size	32bit	64bit	96bit
Commonly used	16Bit AS : 16Bit Value	Type : 32Bit AS : 16Bit Value -or- Type : 16Bit Value: 32Bit AS	32Bit AS : 32Bit Value : 32Bit Value
Example	6695:65010	RT:6695:2010223112	6695:65010:2010223112
Intention		Replace original standard	Supplement original standard



DE CIX

BGP Communities for DE-CIX Route Servers

Example is for Frankfurt

For other sites, see [DE-CIX website](#)

	Original Communities	Extended Communities	Large Communities
Announce to all peers	6695:6695	rt:6695:6695	6695:1:0
Do not announce to any peer	0:6695	rt:0:6695	6695:0:0
Redistribute to PEERAS	6695:PEERAS (16Bit only)	rt:6695:PEERAS	6695:1:PEERAS
Do not redistribute to PEERAS	0:PEERAS	rt:0:PEERAS	6695:0:PEERAS
Add NO-EXPORT	6695:65281		6695:901:0
Add NO-ADVERTISE	6695:65282		6695:902:0
Add NO-EXPORT to PEERAS			6695:901:PEERAS
Add NO-ADVERTISE to PEERAS			6695:902:PEERAS
Prepend 1 times to all peers	65001:0		6695:101:0
Prepend 2 times to all peers	65002:0		6695:102:0
Prepend 3 times to all peers	65003:0		6695:103:0
Prepend 1 times to PEERAS	65001:PEERAS	rt:65001:PEERAS	6695:101:PEERAS
Prepend 2 times to PEERAS	65002:PEERAS	rt:65002:PEERAS	6695:102:PEERAS
Prepend 3 times to PEERAS	65003:PEERAS	rt:65003:PEERAS	6695:103:PEERAS
Blackhole (if supported by Peer)	BLACKHOLE		



DE CIX

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

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



DE CIX

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