

# Handling BGP Attribute Errors

Rob Shakir (GX Networks)  
[rjs@eng.gxn.net](mailto:rjs@eng.gxn.net) / RJS-RIPE

# Outline / Motivation

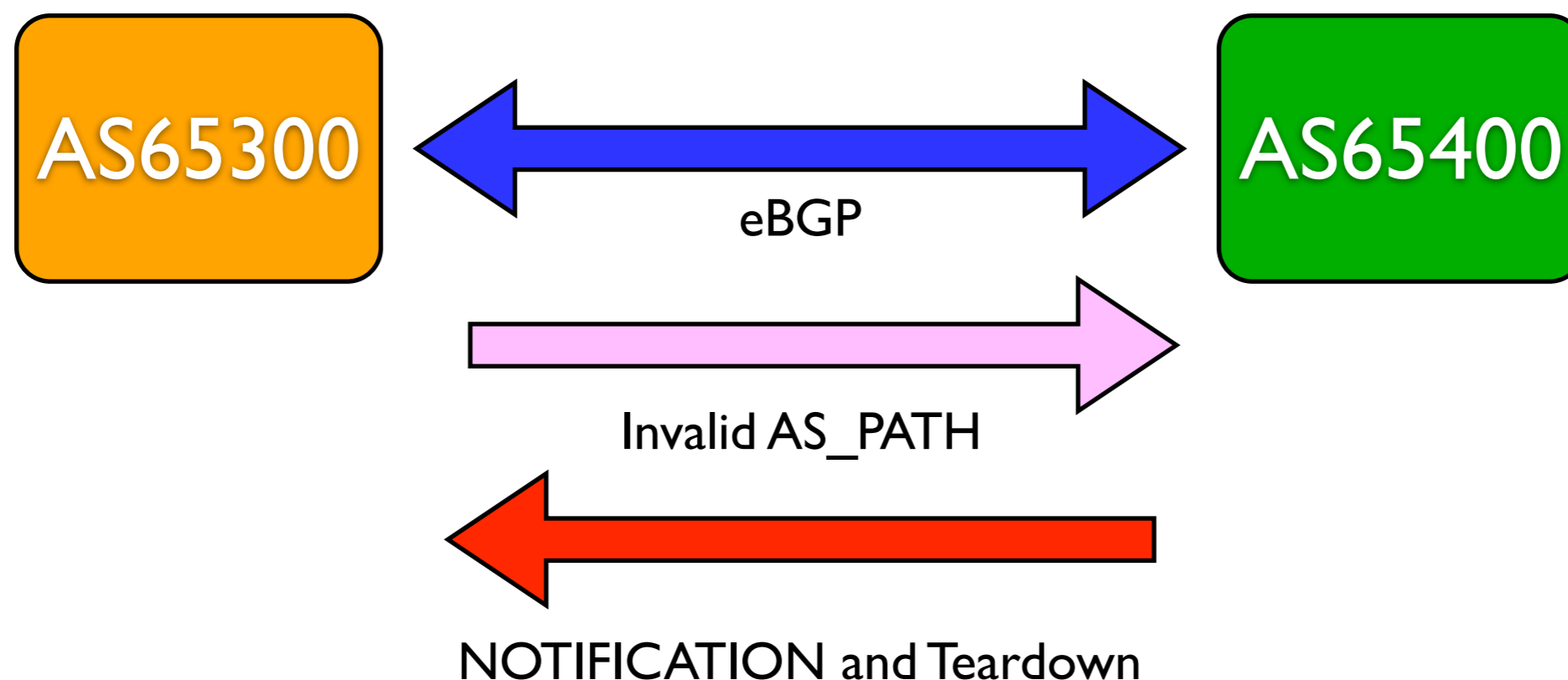
- BGP Errors - Current Handling
- AS4\_PATH Bug and Optional Transitives
  - Update to RFC 4893
- IETF IDR Drafts
- *Why you should care!*

# Attributes and Errors

- Types of BGP Attributes
  - Well-known Mandatory
  - Well-known Discretionary
  - Optional Transitive
  - Optional Non-Transitive
- RFC 4271
  - *“A NOTIFICATION message is sent when an error condition is detected. The BGP connection is closed immediately after it is sent.”*

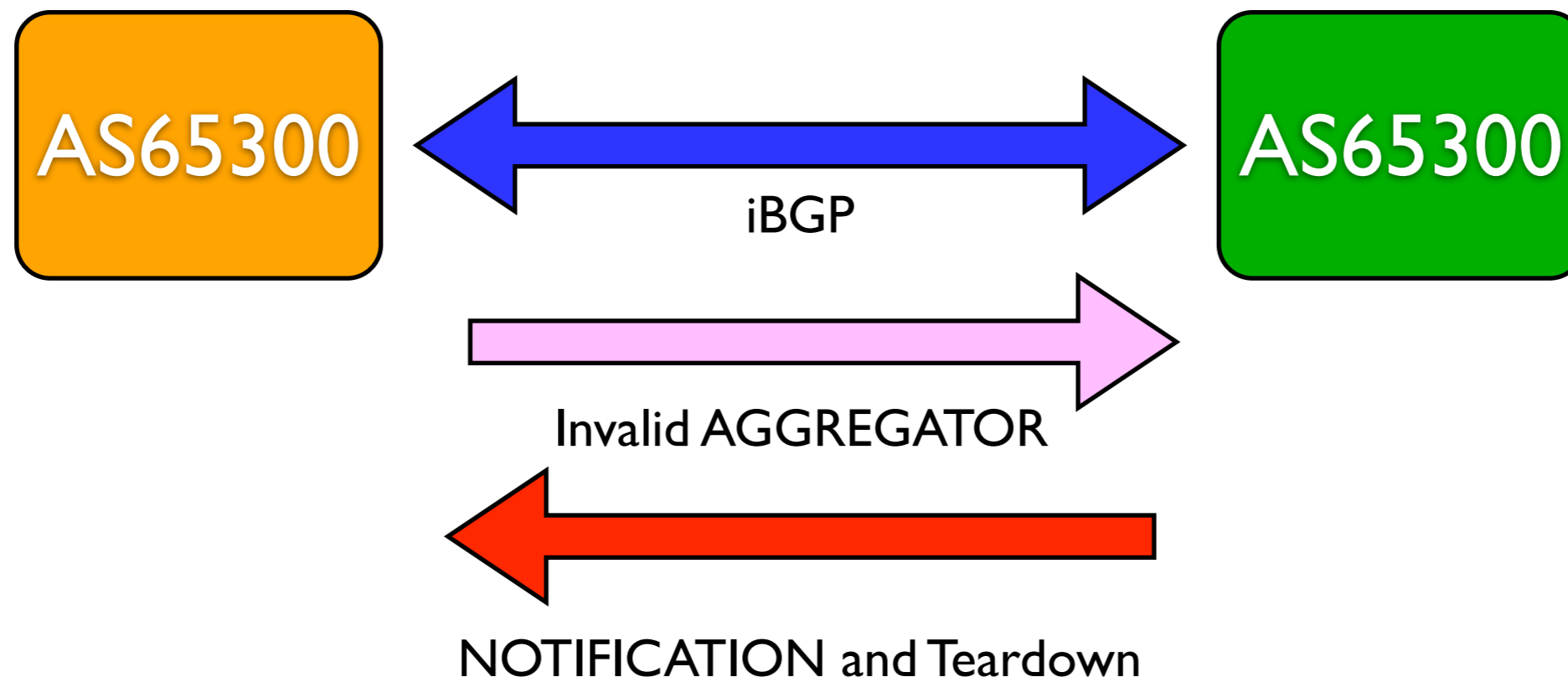
# Current Error Handling (I)

- AS\_PATH Error (Well-known Mandatory)
- Worst case - loops and invalid routing.



# Current Error Handling (2)

- Aggregator Error (Optional Transitive)
  - Worst case? Loss of routing metadata?



# Problem?

- All errors are treated equally.
- Is this the right behaviour?
  - “Good, we’re being cautious!”
  - “Why is my AS suddenly disconnected from the global table?”

# AS4\_PATH

- Defined in RFC 4893 (Optional Transitive)



**AS4\_PATH:**

Not Used

70000

Not Used

**AS\_PATH:**

i

23456 i

65400 70000 i

# Neat! And Errors?

- Shouldn't really see errors!
- Cleaned like AS\_PATH
- Mixed NEW and OLD confederations
- *“To prevent the possible propagation of confederation path segment outside of a confederation, the path segment types AS\_CONFED\_SEQUENCE and AS\_CONFED\_SET [RFC3065] are declared invalid for the AS4\_PATH attribute” (RFC 4893)*



# Whoops!

- December 10th 2008
  - 91.207.218.0/23
  - AS4\_PATH: (65044 65057) 196629 (7 bytes)
  - AS\_PATH: xx xx 35320 23456 (13 bytes)
- Confederation information in AS4\_PATH
  - First RFC-compliant NEW speaker to see the UPDATE tears down the session to where it saw the UPDATE from.

# What went wrong?

- ASN running mixed confeds with mixed OLD/NEW speakers and JunOS.

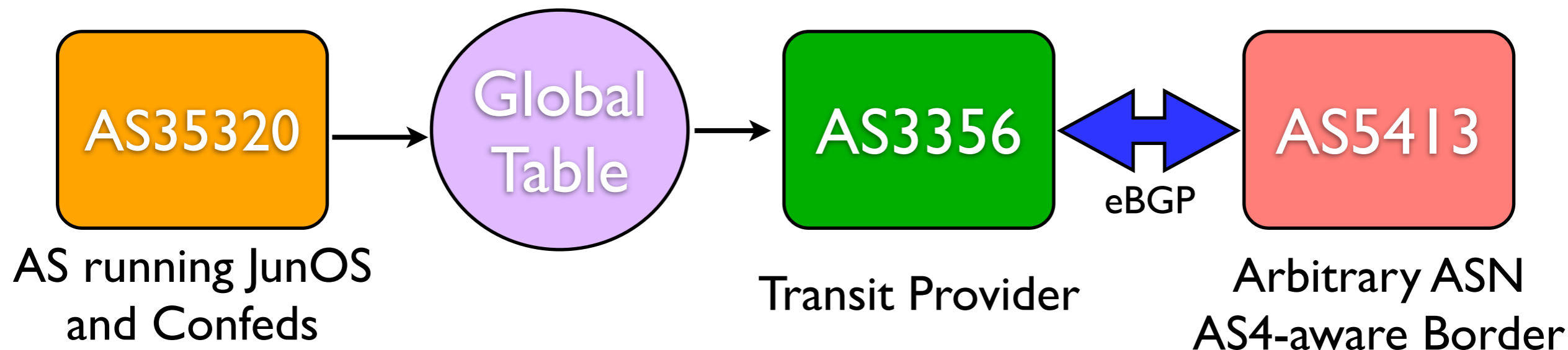


Copies  
AS\_CONFED\_SET  
into AS4\_PATH

Transits AS4\_PATH  
(not checked!)

Invalid AS4\_PATH  
received - sends  
NOTIFICATION  
and teardown

# Why is this concerning?



- First RFC compliant AS4 speaker in the path reacts.
- Teardown can be towards transit (likely, every prefix on these sessions!)
- Can craft an UPDATE to reach via a specific path.

# Our Recommended Fix

- Recommended: Don't send NOTIFICATION, treat UPDATE as withdrawal of prefix via this path.
- “Punish” broken paths without breaking every prefix via a session.
- Prefix might become unreachable.

# Likely RFC Fix

- draft-ietf-rfc4893bis
- Ignore the broken parts of the AS4\_PATH.
- IOS implemented this -I2.0(32)S(Y8|I3)
- Doesn't lose reachability, and recovers from an error "in the wild"
- Some implications in loop detection?

# AS\_PATH and AS4\_PATH

- Last LINX meeting - AS\_PATH length problems.
- Different Case: Well Known Mandatory
- Highlights interesting point relating to AS4\_PATH - loop detection for AS4?
- Bugs will always mean that invalid information is propagated.

# The General Case

- draft-ietf-rfc4893bis fixes this specific - but what about others?
- Errors in *other* optional transitives still cause session teardown.
- Revise this behaviour? Don't require NOTIFICATION be sent.
- Tell our neighbour that someone in their path did something wrong?

# **draft-scudder-idr-optional-transitive**

- Handles the case of Optional Transitives that are not formed or checked by our neighbour
- Partial bit is set to 1 if some BGP speaker passes without checking.
- These are the “tunneled” UPDATES
- Recommended behaviour: Treat as a withdraw of the prefix and log.



# draft-scholl-idr-advisory

- New MP-BGP capability (ADVISORY)- allows a string to be transmitted between two routers.
- NOT a replacement for NOTIFICATION
- Inform our neighbour that we're considering an UPDATE as invalid.
- Not just error handling:
  - “in-band” notification (e.g. maintenance)

# **draft-nalawade-bgp-soft-notify**

- Has been some opposition to ADVISORY
- Humans already have phone and e-mail!
- **SOFT-NOTIFICATION** previous suggestion (2003)
- Intended to allow for graceful recovery from an error.
- Structured payload (no IM via BGP!)

# Implications of these Drafts

- Protocol-wise, this isn't core functionality
  - Vendors and protocol-purists not necessarily interested?
- Operationally, we need to be robust!
  - Do we trust everyone in the global table?
- Easier direct communication of events and settings directly between operators.
- Capability (you can turn it off!)

# Conclusions

- Blanket handling of BGP errors is suboptimal.
- Fix handling optional transitive errors (make the protocol more robust!)
- Add method to communicate these errors without tearing sessions down.
- Operator's voices are *really* needed here!

# Questions, Comments, Corrections?

Many thanks to:

Andy Davidson (NetSumo)

Jonathan Oddy (Hostway)

David Freedman (Claranet)

Will Hargrave (LONAP)

Greg Hankins (Force10)

## **Questions, or comments later?**

[rjs@eng.gxn.net](mailto:rjs@eng.gxn.net)

RJS-RIPE

## **Public Comments?**

IETF IDR - [idr@ietf.org](mailto:idr@ietf.org)

(To Subscribe: [idr-request@ietf.org](mailto:idr-request@ietf.org), In Body: subscribe idr-post)