BGP Graceful Shutdown

draft-ietf-grow-bgp-gshut-02

Pierre Francois
Bruno Decraene
Cristel Pelsser
Clarence Filsfils
Keyur Patel

UCL-BE
France Telecom - Orange
IIJ
Cisco Systems
Cisco Systems
What is it about?

- Avoid traffic loss upon planned BGP session shutdown
- Operational procedures
  - Pre-configuration set
  - At maintenance time
- *New:* Open option for a “simplified” procedure
- *New:* Vendor support considerations
  - Less configuration burden
LoC

• Outbound :
  Path hiding in iBGP leads to LoC
  Transient inconsistencies lead to LoC

• Inbound :
  Same problem
  Need to trigger G-Shut at neighboring AS
LoC upon maintenance
Path hiding

$p : lp = 100$

$ASBR_3$

$ASBR_1$

$RR_1$

$RR_2$

$ASBR_2$

$p : lp = 90$
LoC upon maintenance
Transient inconsistencies

\[ p : \text{lp} = 100 \]

\[ p : \text{lp} = 90 \]
Outbound traffic
Outbound LP

Decrease LP on outbound iBGP session

Wait

Shut

p \_lp = 0
Outbound traffic
Outbound LP

No more lack of path
ASBR1 keeps using primary (no forwarding mess)
Outbound traffic
Outbound LP : glitch

Decrease LP on outbound iBGP session

Transient switch to ASBR2

Update on Shut
Outbound traffic
Simple : Inbound LP

Decrease LP on inbound eBGP session

Wait

Shut
Outbound traffic
Inbound LP : glitches

ASBR1 switches to alternates as soon as it knows one

Early primary withdraw

fwd mess if alternate was not installed (Ext. best)
Outbound traffic
Inbound LP glitches: dilemma?

- Solving path hiding problem with Ext Best leads to transient forwarding loops with inbound LP
Outbound traffic

Inbound LP glitches: dilemma?

- Solving path hiding problem with Ext Best leads to transient forwarding loops with inbound LP. Unless you have a proper PIC implementation!
Best External behavior

ASBR1

ASBR3

RR1

RR2

ASBR2

p in:
l_p = 100

p = 90

via ASBR1

via ASBR2

p = 100

p = 90

to

ASBR1

ASBR2

RR1

RR2
Clever Best External behavior

\[ \text{p in: } \quad l_p = 100 \]

\[ \text{p: } l_p = 100 \text{ via ASBR1} \]
\[ \text{p: } l_p = 90 \text{ via ASBR2} \]

\[ \text{p: } l_p = 90 \]

\[ \text{p: } l_p = 90 \text{ via ASBR2}^* \]

\[ * \text{ used for traffic rcvd from the inside} \]

\[ \text{p: } l_p = 100 \text{ via ASBR1} \]
\[ \text{p: } l_p = 90 \text{ via ASBR2} \]
Inbound traffic
Trigger outbound g-shut at the other side of the peering link

z/Z:
tag community: 0xFFFF0000

iBGP out/in-filter:
match 0xFFFF0000 (remove community)
set local pref 0

No need for action from the customer at maintenance time
Inbound traffic
Boring aspect of GSHUT community

- Need to handle its non-transitive essence

- Option to use a *non transitive* extended community
  draft-decraene-idr-reserved-extended-communities-00
Next steps

- Radical draft verbosity decrease
- Simplified procedure when obtained FCFS non transitive extended community
- Needs IDR agreement