BGP Graceful Shutdown

draft-ietf-grow-bgp-gshut-01

Pierre Francois 
Bruno Decraene 
Cristel Pelsser 
Clarence Filsfils

UCL-BE
France Telecom - Orange
IIJ
Cisco Systems
Agenda

• Quick reminder
• The G-shut community value
• G-shut behavior concealment
• Local-pref overriding
BGP G-Shut

- Avoid traffic loss upon manual eBGP session shutdown
- Due to transient lack of path during convergence
- Operational procedures
Outbound traffic

p/P in : lp = 100

p/P : lp = 90

nh : ASBR1

RR1

RR2

ASBR1

ASBR2

ASBR3
Outbound traffic

p/P in :
lp = 100

p/P : lp = 90

nh : ASBR1

RR1

RR2

ASBR1

ASBR2

ASBR3

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

nh : ASBR1
Outbound traffic

p/P in:
lp = 100

p/P out:
lp = 0

p/P : lp = 90

nh : ASBR1

ASBR3

ASBR1

RR1

RR2

ASBR2

nh : ASBR1

nh : ASBR1

nh : ASBR1
Outbound traffic

p/P in: lp = 100

p/P out: lp = 0

p/P : lp = 90

ASBR1

ASBR2

RI

RR2

ASBR3

nh : ASBR1

nh : ASBR1

nh : ASBR1

nh : ASBR1
Outbound traffic

ASBR1

p/P in:
l_p = 100

p/P out:
l_p = 0

ASBR2

p/P : l_p = 90

ASBR3

nh : ASBR1

RR1

RR2

nh : ASBR1

nh : ASBR1
Outbound traffic

p/P in: \( l_P = 100 \)

p/P out: \( l_P = 0 \)

p/P: \( l_P = 90 \)
Outbound traffic

- **p/P in**: lp = 100
- **p/P out**: lp = 0
- **p/P**: lp = 90
Outbound traffic

Outbound traffic from ASBR1 to ASBR2.

- **p/P in**: $l_p = 100$
- **p/P out**: $l_p = 0$
- **p/P in**: $l_p = 90$

Routing Relationships:
- ASBR1 to RR1
- ASBR1 to RR2
- RR1 to RR2
- RR2 to ASBR2
- ASBR3 to ASBR1

network labels:
- nh : ASBR1
Outbound traffic

- **p/P in:** \( l_p = 100 \)
- **p/P out:** \( l_p = 0 \)
- **p/P:** \( l_p = 90 \)
Outbound traffic

p/P in : \( l_p = 100 \)

p/P out : \( l_p = 0 \)

p/P : \( l_p = 90 \)
Outbound traffic

- p/P in: lp = 100
- p/P out: lp = 0
- p/P: lp = 90

ASBR1 → ASBR2
ASBR2 → ASBR3
RR1 → RR2
RR2 → RR1
Outbound traffic

p/P in: lp = 100
p/P out: lp = 0
p/P: lp = 90

ASBR1
RR1
RR2
ASBR2
ASBR3

nh: ASBR2
p/P
p/P
p/P
p/P
p/P
p/P
Outbound traffic

......

p/P in :
l_p = 100

p/P out :
l_p = 0

p/P : l_p = 90

nh : ASBR2

nh : ASBR2

hh : ASBR2

nh : ASBR2
Outbound traffic

Shut

p/P in : lp = 100

p/P out : lp = 0

RR1

RR2

ASBR1

ASBR2

ASBR3

Nh : ASBR2

Nh : ASBR2

Nh : ASBR2

Nh : ASBR2

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

p/P

lp = 90

lp = 0

lp = 100

lp = 0
Outbound traffic

Shut

ASBR3

nh : ASBR2

ASBR1

nh : ASBR2

RR1

nh : ASBR2

RR2

nh : ASBR2

ASBR2

p/P : lp = 90

p/P
**Inbound traffic**

Trigger outbound g-shut at the other side of the peering link

---

**ASBR1**

**CE1**

**ASBR2**

**CE2**

**z/Z:**

tag community : GSHUT

iBGP out-filter :
match GSHUT
remove community
set local pref 0

No need for action from the customer at maintenance time
• One g-shut community per neighbor AS is a pain
  • neighbor-specific config
  • help from vendors less practical
• 0xFFFFF0000 value has been reserved for this
  • see http://www.iana.org/assignments/bgp-well-known-communities/
G-shut concealment

• Avoid the g-shut community to be propagated through the whole Internet upon maintenance of a link
• Convergence is kept local to the AS
• Prevents misuses of the community
G-shut concealment

- match FFFF0000
  remove community
  set local pref 0
  on reception from eBGP neighbors that do not support gshut or to whom the service is not provided

- match FFFF0000
  remove community
  set local pref 0
  on iBGP out-filters
Extended community?

- "If the GSHUT community is an extended community, it SHOULD be chosen non-transitive. In that case, some clarification is required about the handling of such communities"

- draft-decraene-idr-rfc4360-clarification-00
What would we get?

- No need to re-implement community non-transitivity of by configuration on ASBRs
- No default behavior of 4360 can lead to misuses of the solution
Local pref overriding

- If some speakers override LP on iBGP sessions
- Use of an AS specific community over the iBGP sessions