BGP Graceful Shutdown

draft-francois-bgp-gshut-01.txt

Pierre Francois   UCL-BE

Bruno Decraene   France Telecom - Orange

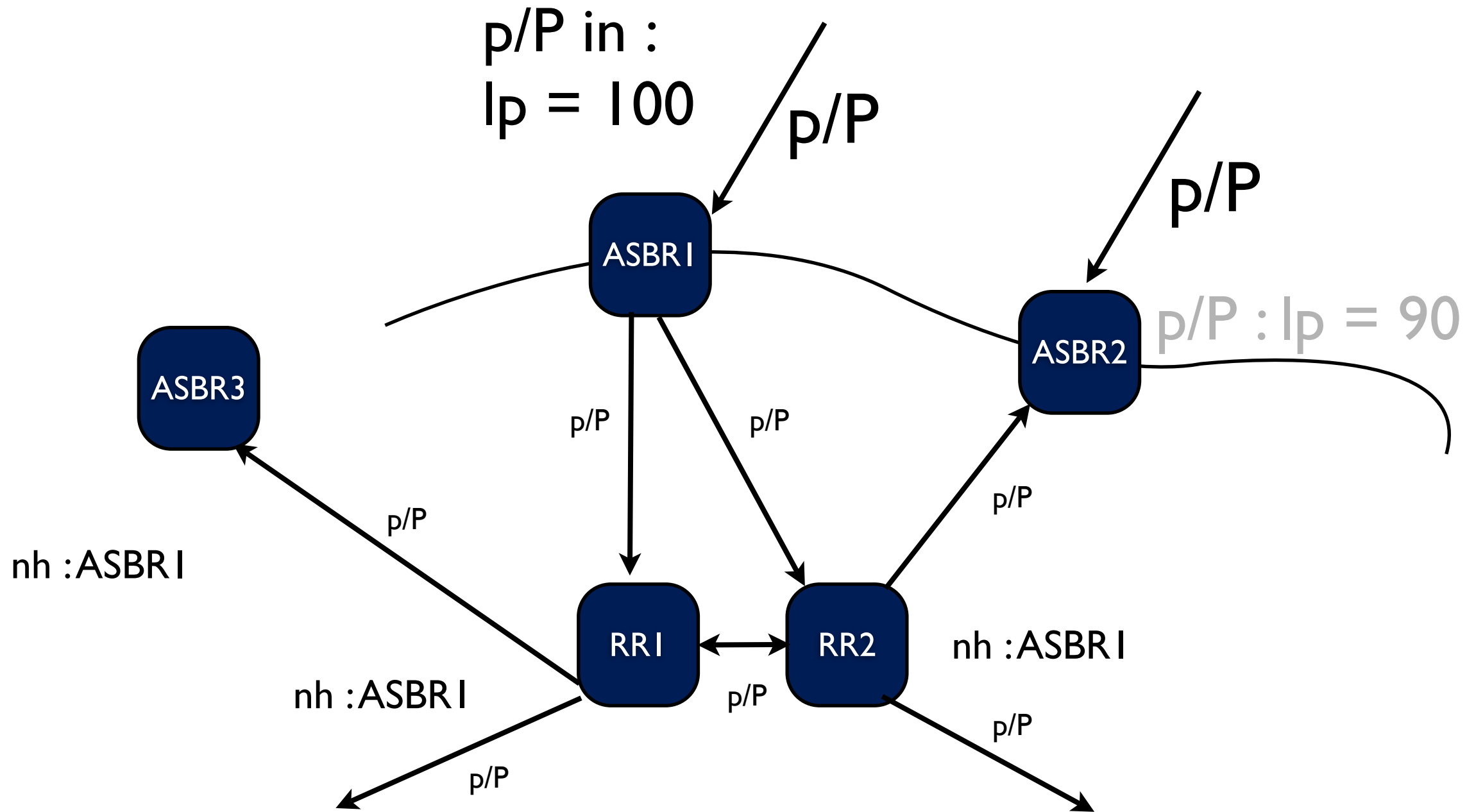Cristel Pelsser   NTT

Clarence Filsfils   Cisco Systems
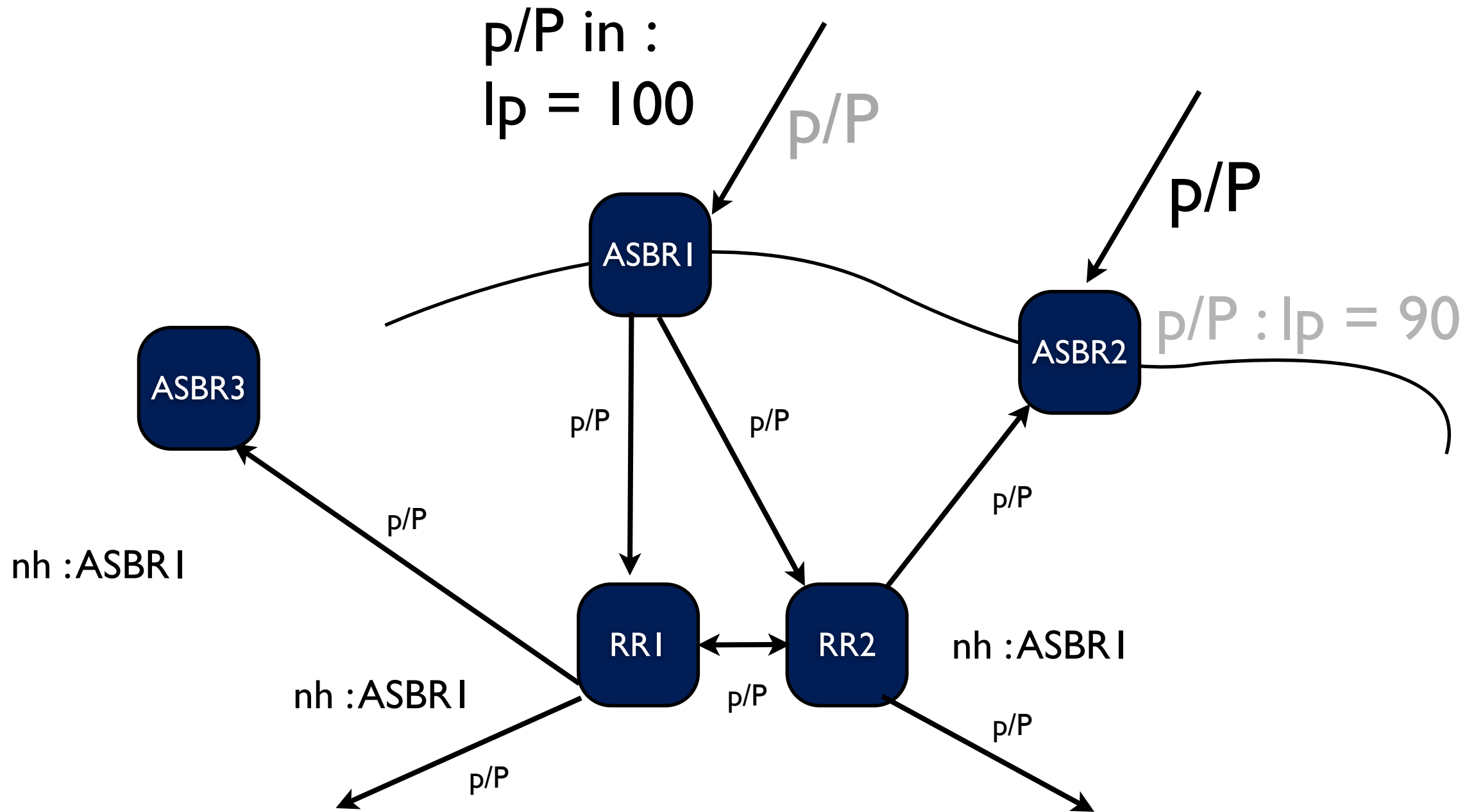
Agenda

- Brief reminder

- Changes since -00

Goal

- Avoid traffic loss upon manual eBGP session shutdown

  - Due to transient lack of path during convergence within the AS
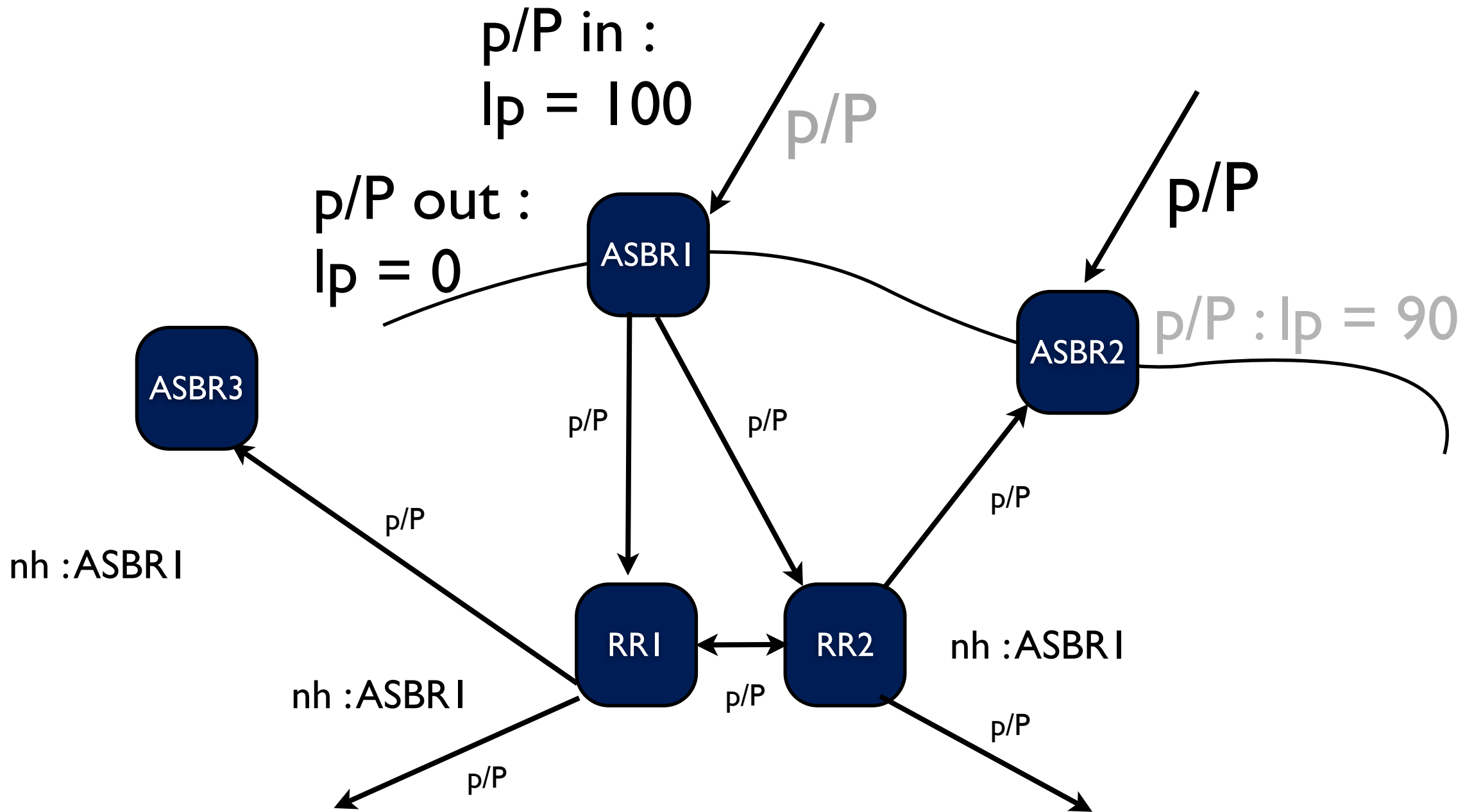
- Operational procedures
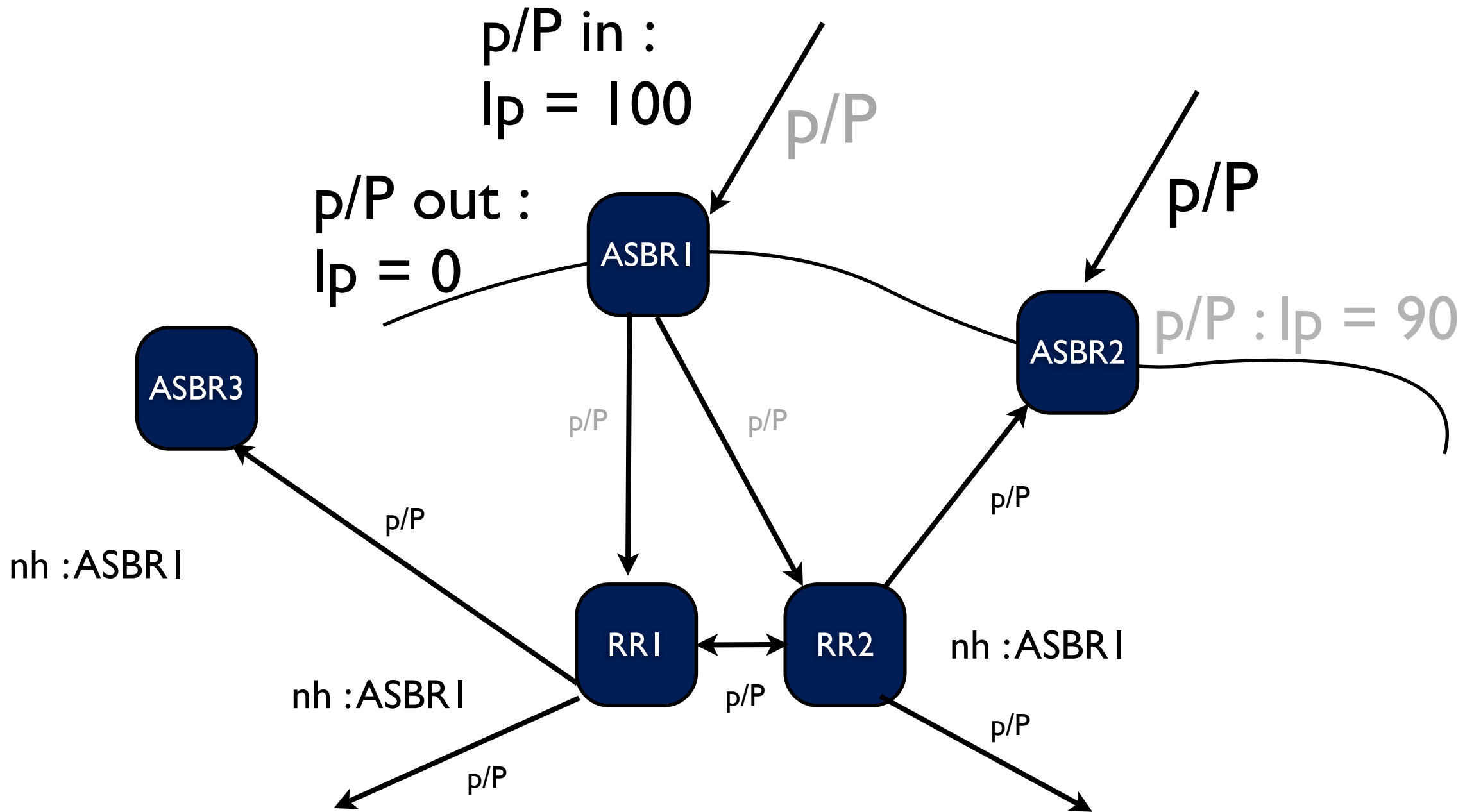
G-Shut : Outbound traffic

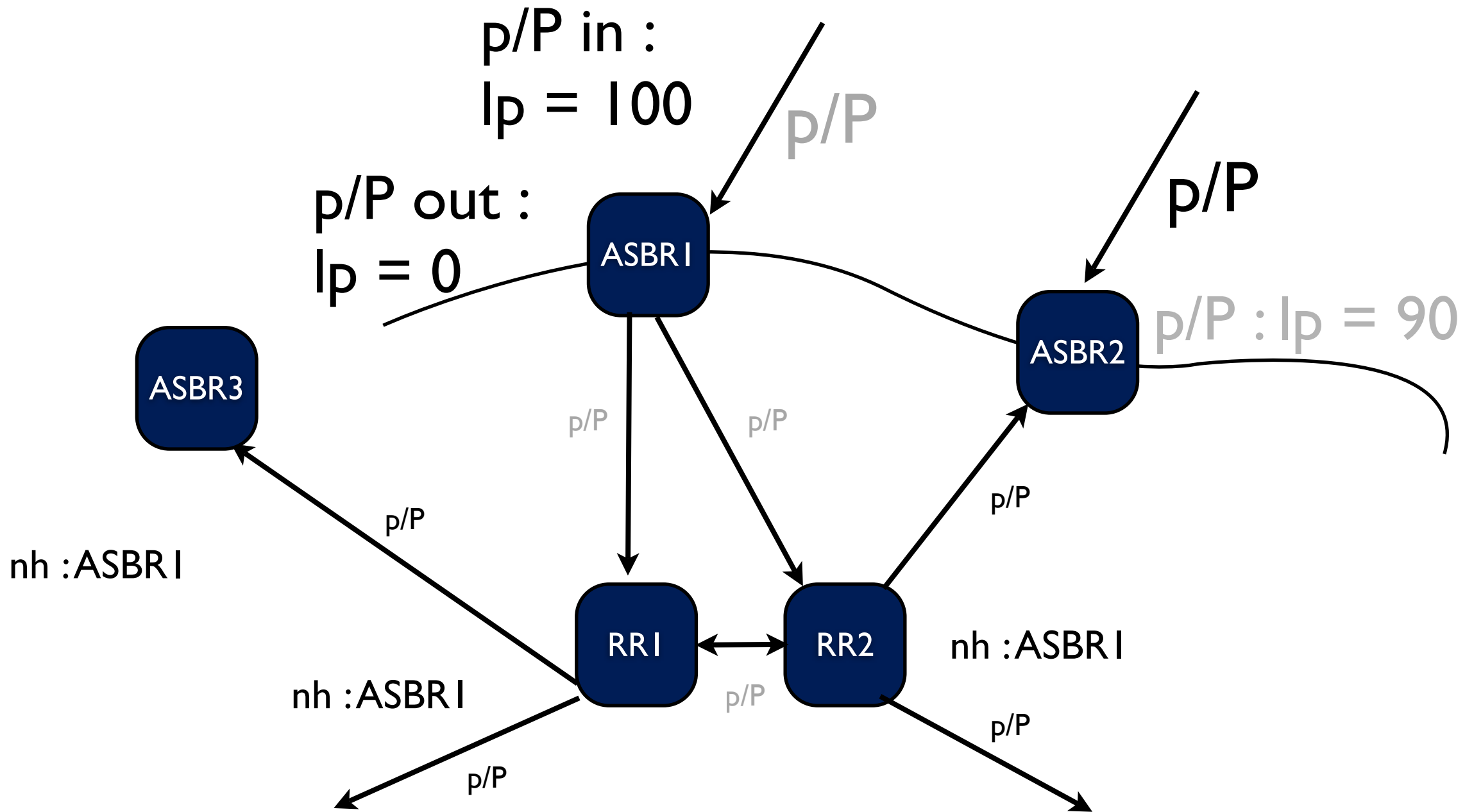# G-Shut : Outbound traffic

G-Shut : Outbound traffic

G-Shut : Outbound traffic

p/P in :
$l_p = 100$

p/P out :
$l_p = 0$

p/P

p/P

p/P : $l_p = 90$

ASBR1

ASBR2

ASBR3

p/P

p/P

p/P

p/P

RR1

RR2

nh :ASBR1

nh :ASBR1

nh :ASBR1

p/P

p/P

p/P

4

G-Shut : Outbound traffic

p/P in :
$I_P = 100$

p/P out :
$I_P = 0$

p/P

p/P

p/P : $I_P = 90$

ASBR1

ASBR2

ASBR3

p/P

p/P

p/P

p/P

RR1

RR2

nh :ASBR1

nh :ASBR1
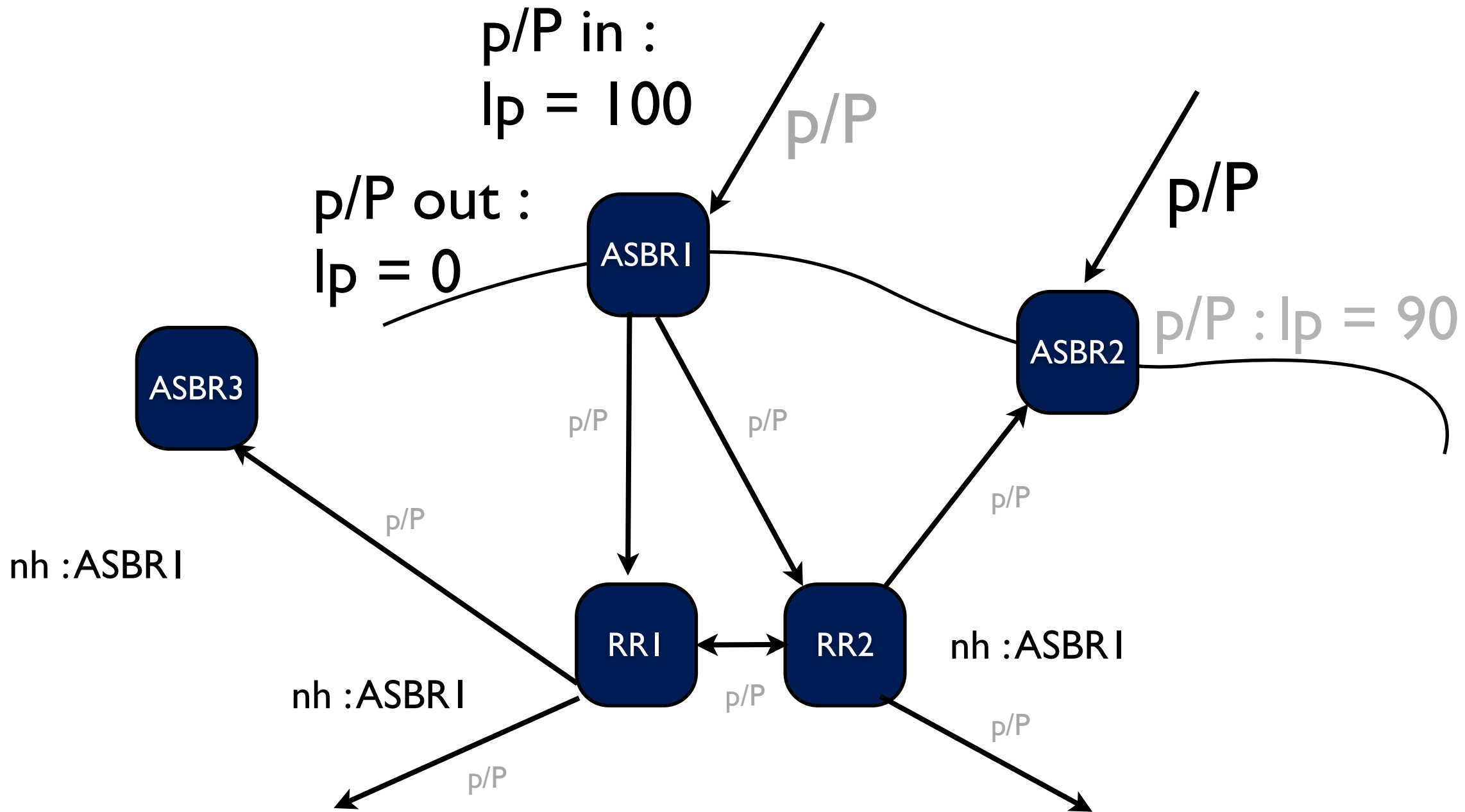
nh :ASBR1

p/P

p/P

p/P

p/P

4

G-Shut : Outbound traffic

G-Shut : Outbound traffic

4

G-Shut : Outbound traffic

p/P in :
Ip = 100

p/P

p/P out :
Ip = 0

p/P

ASBR1

p/P : Ip = 90

ASBR2

ASBR3

nh :ASBR1

p/P

p/P

p/P

p/P

nh :ASBR1

RR1

RR2

nh :ASBR1

p/P

p/P

nh :ASBR1

p/P

4

G-Shut : Outbound traffic

G-Shut : Outbound traffic

p/P in :
$l_P = 100$

p/P out :
$l_P = 0$

p/P

p/P

ASBR1

ASBR2

p/P : $l_P = 90$

ASBR3

p/P

p/P

p/P

p/P

nh :ASBR1

RR1

RR2

nh :ASBR2

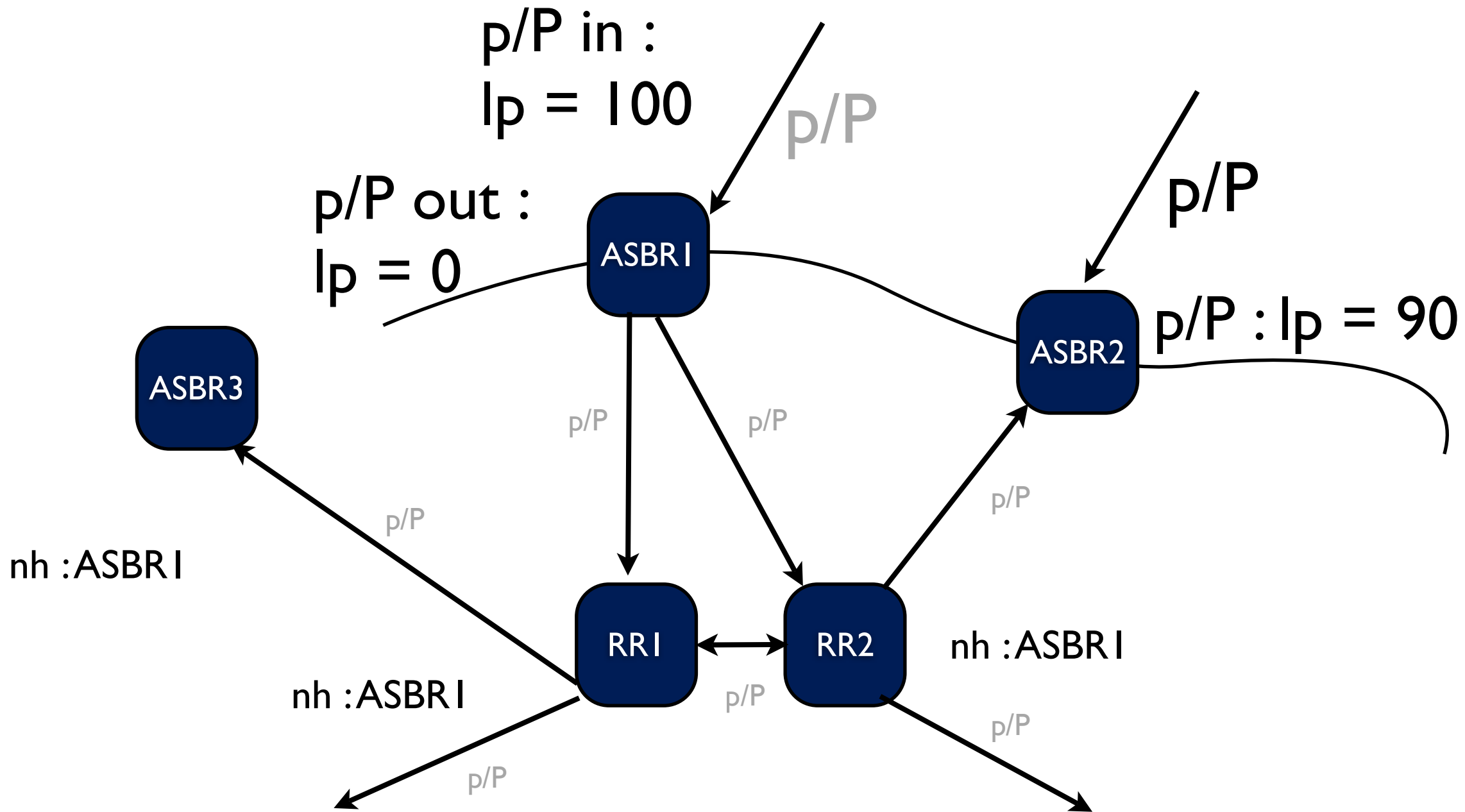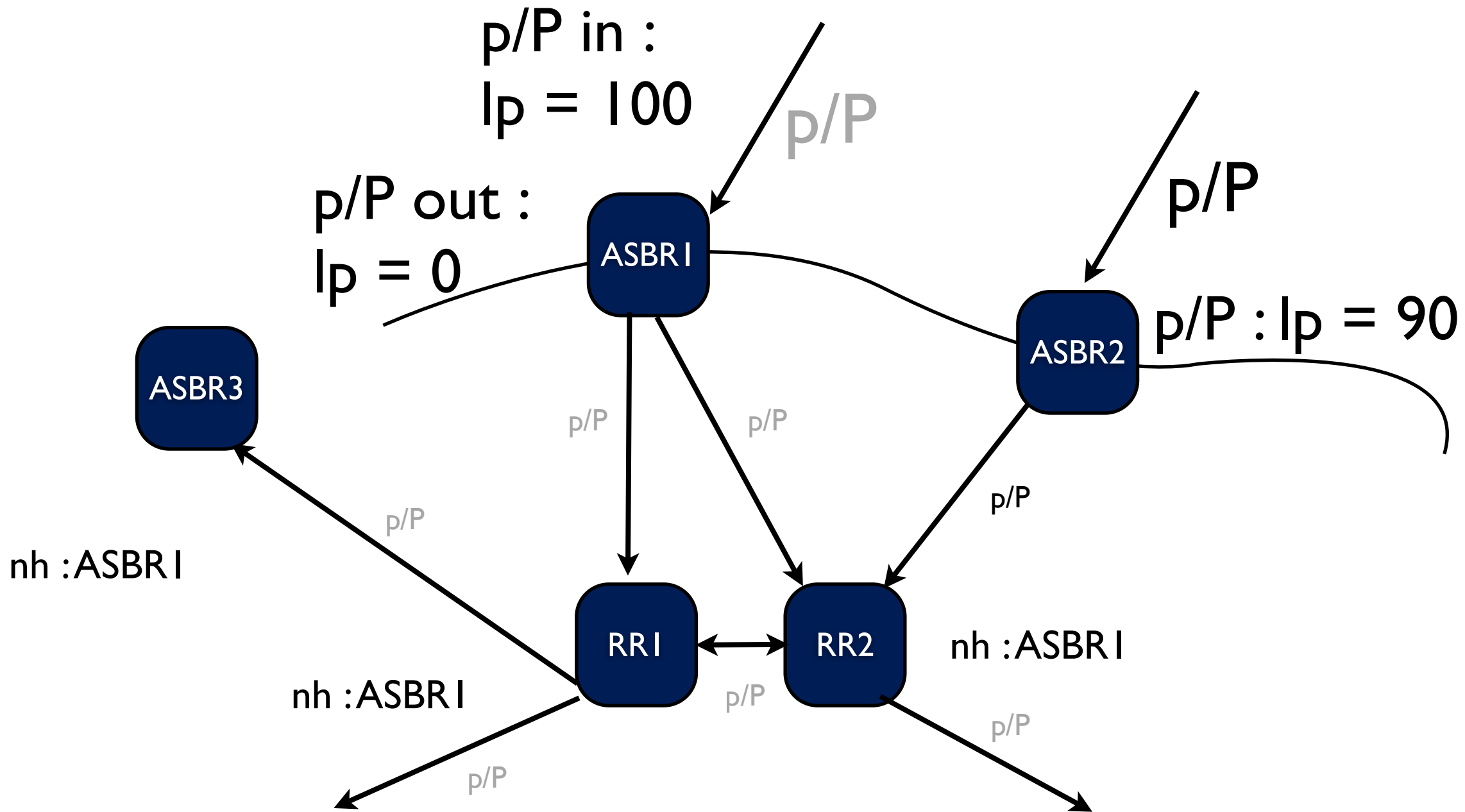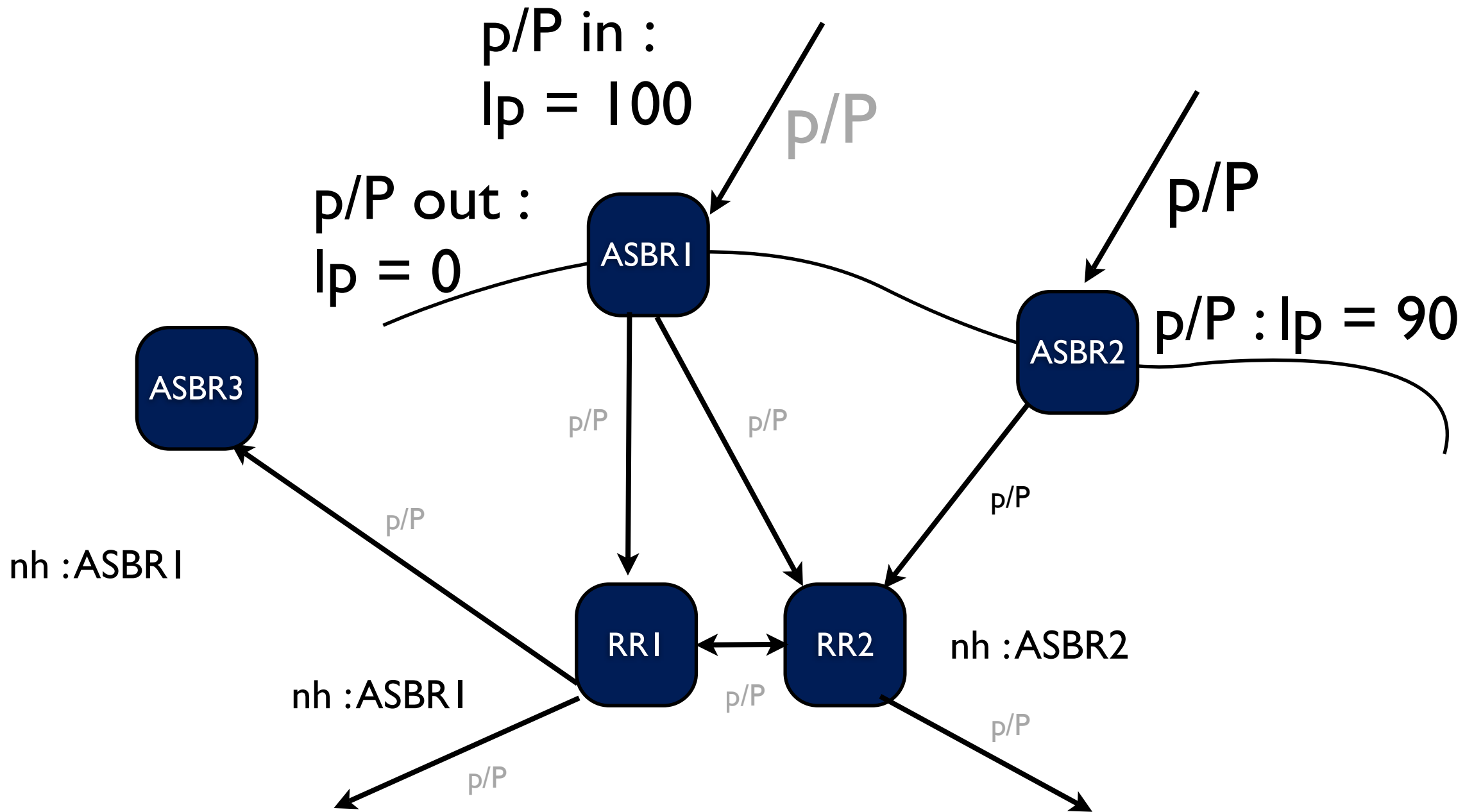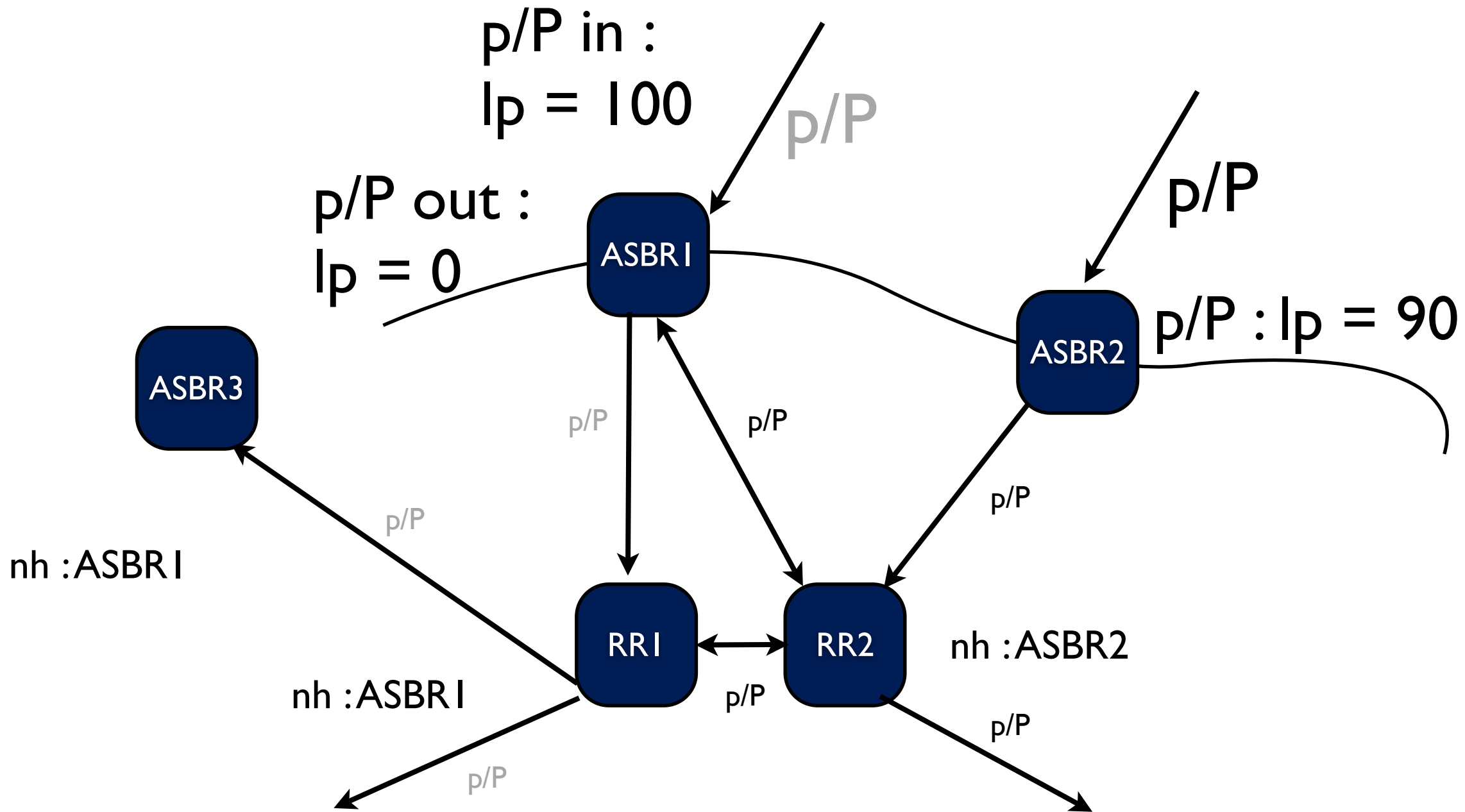nh :ASBR1

p/P

p/P

p/P

4

G-Shut : Outbound traffic

4

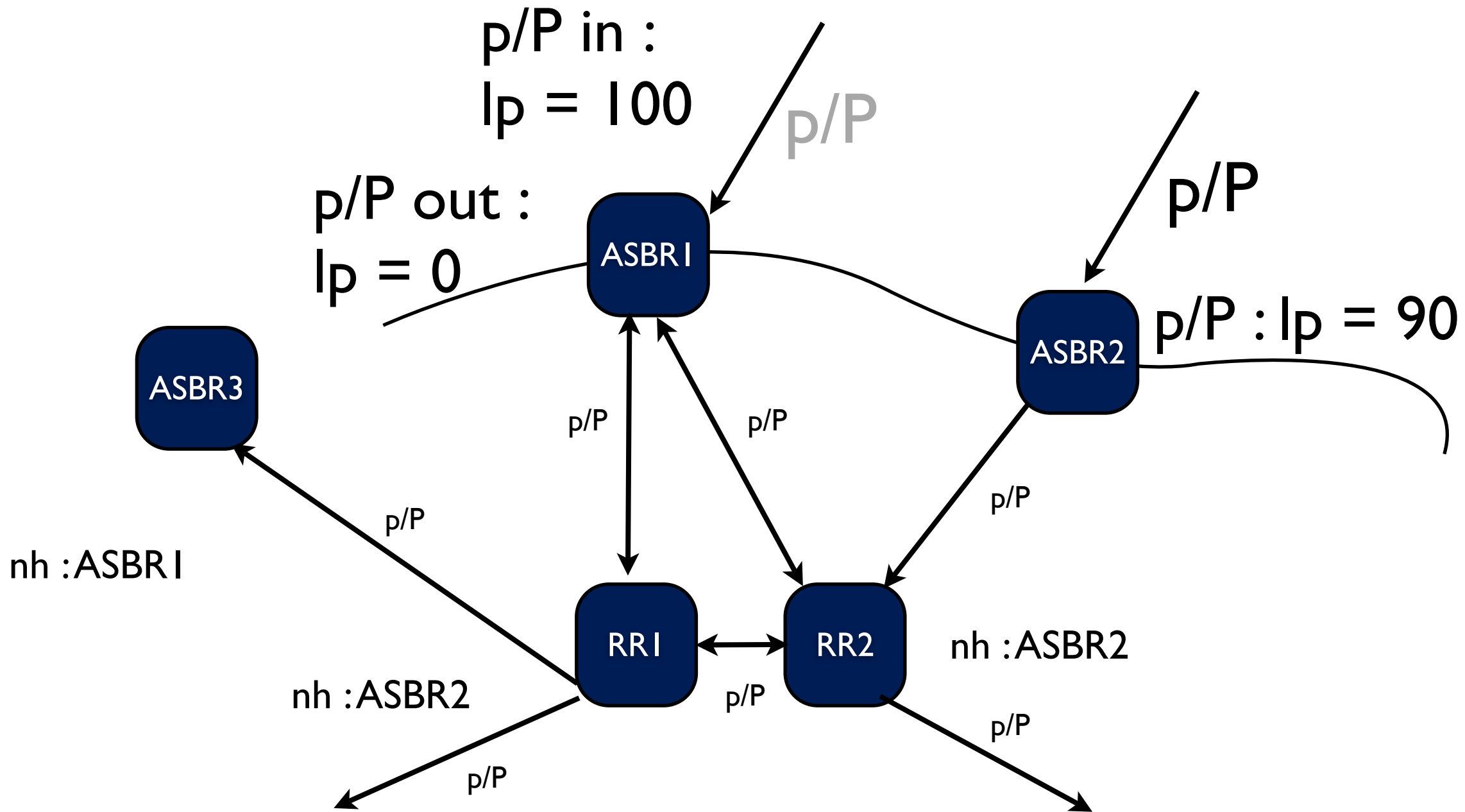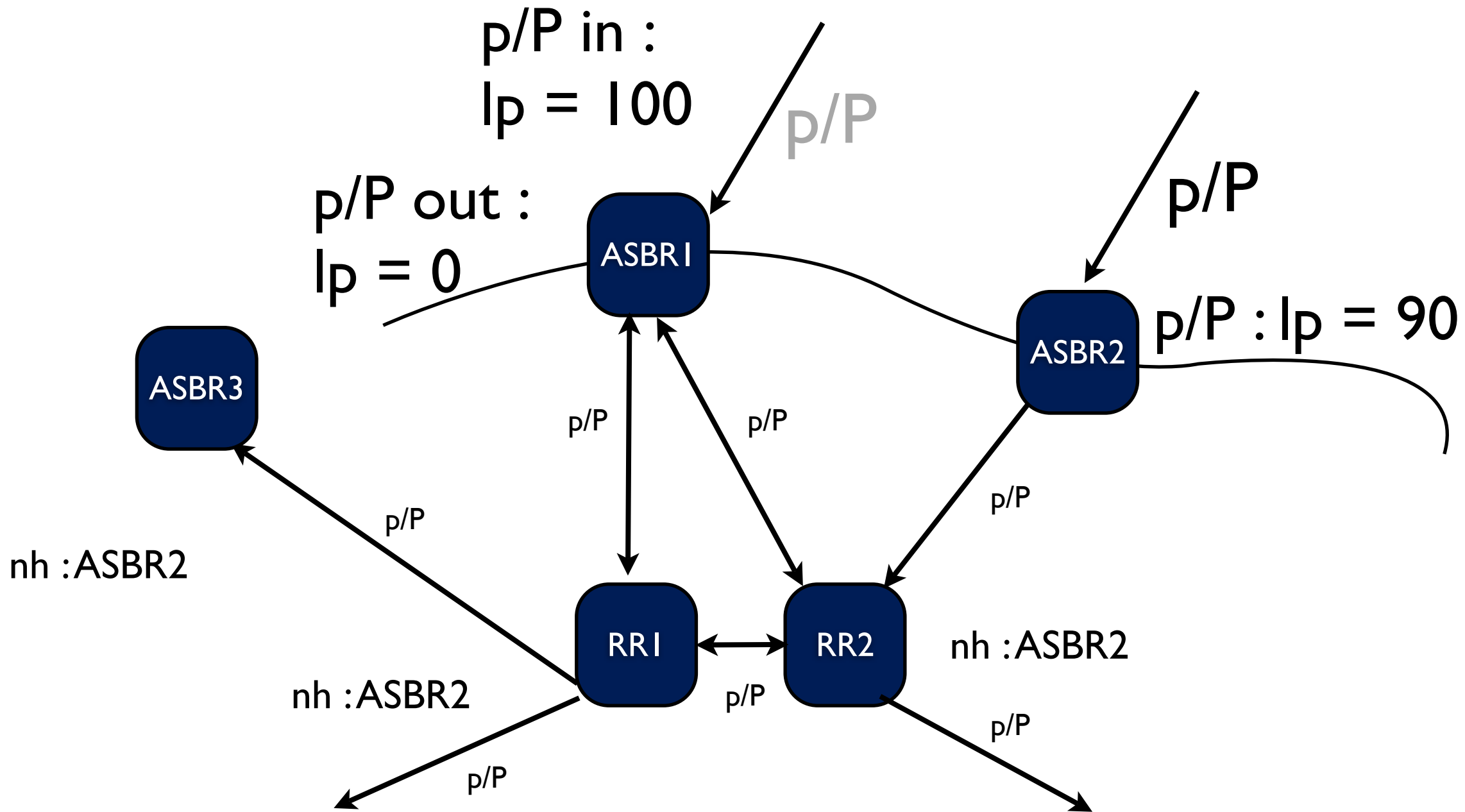# G-Shut : Outbound traffic

G-Shut : Outbound traffic

......

p/P in :
l$_P$ = 100

p/P out :
l$_P$ = 0

p/P

p/P

ASBR1

ASBR2

p/P : l$_P$ = 90

ASBR3

p/P

p/P

nh :ASBR2

RR1

RR2

nh :ASBR2

nh :ASBR2

p/P

p/P

p/P

p/P

4

G-Shut : Outbound traffic

Shut

p/P in :
$l_P = 100$

p/P out :
$l_P = 0$

p/P

p/P

ASBR1

ASBR2

p/P : $l_P = 90$

ASBR3

nh :ASBR2

p/P

p/P

p/P

nh :ASBR2

RR1

RR2

nh :ASBR2

p/P

nh :ASBR2

p/P

p/P

p/P

4
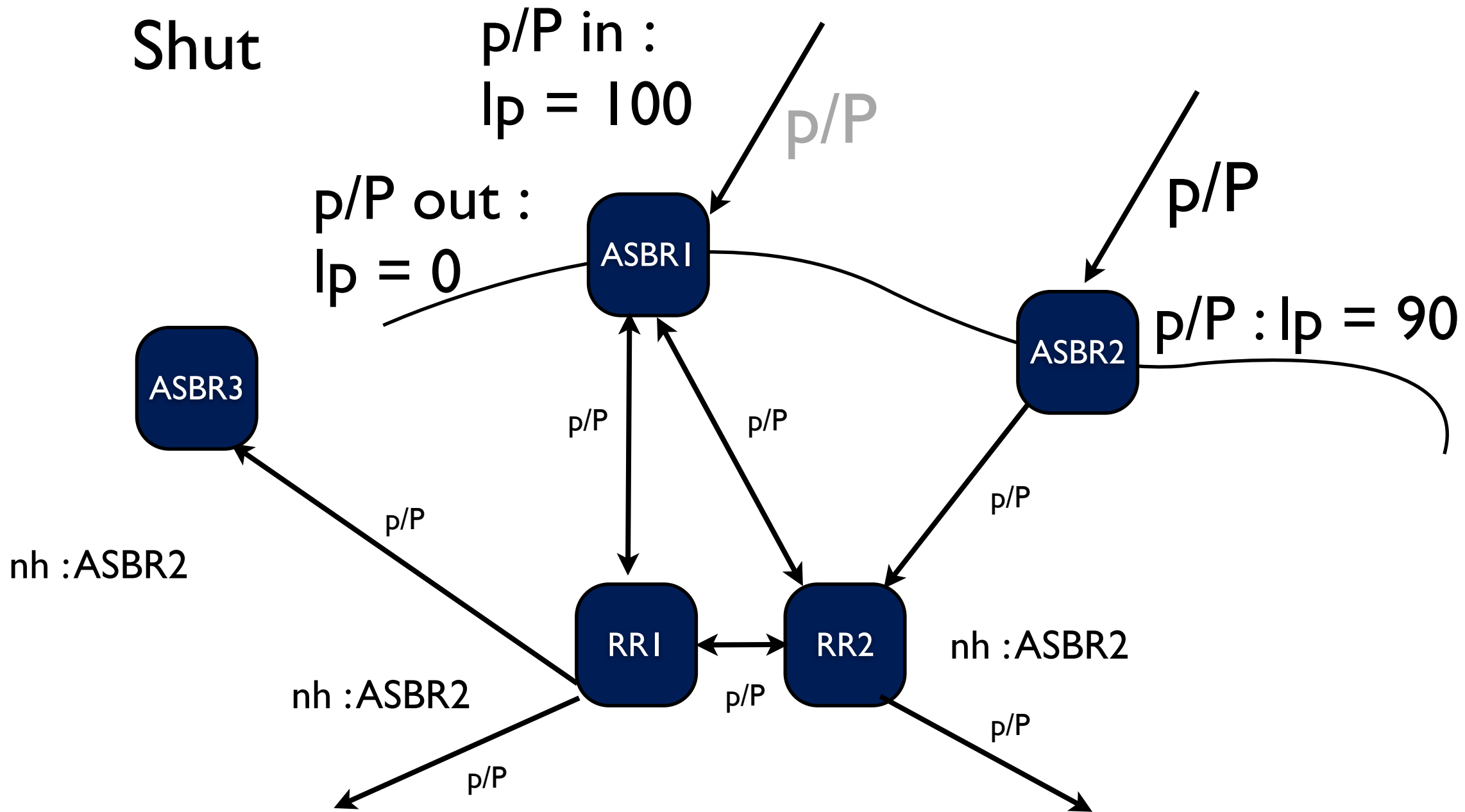
G-Shut : Outbound traffic
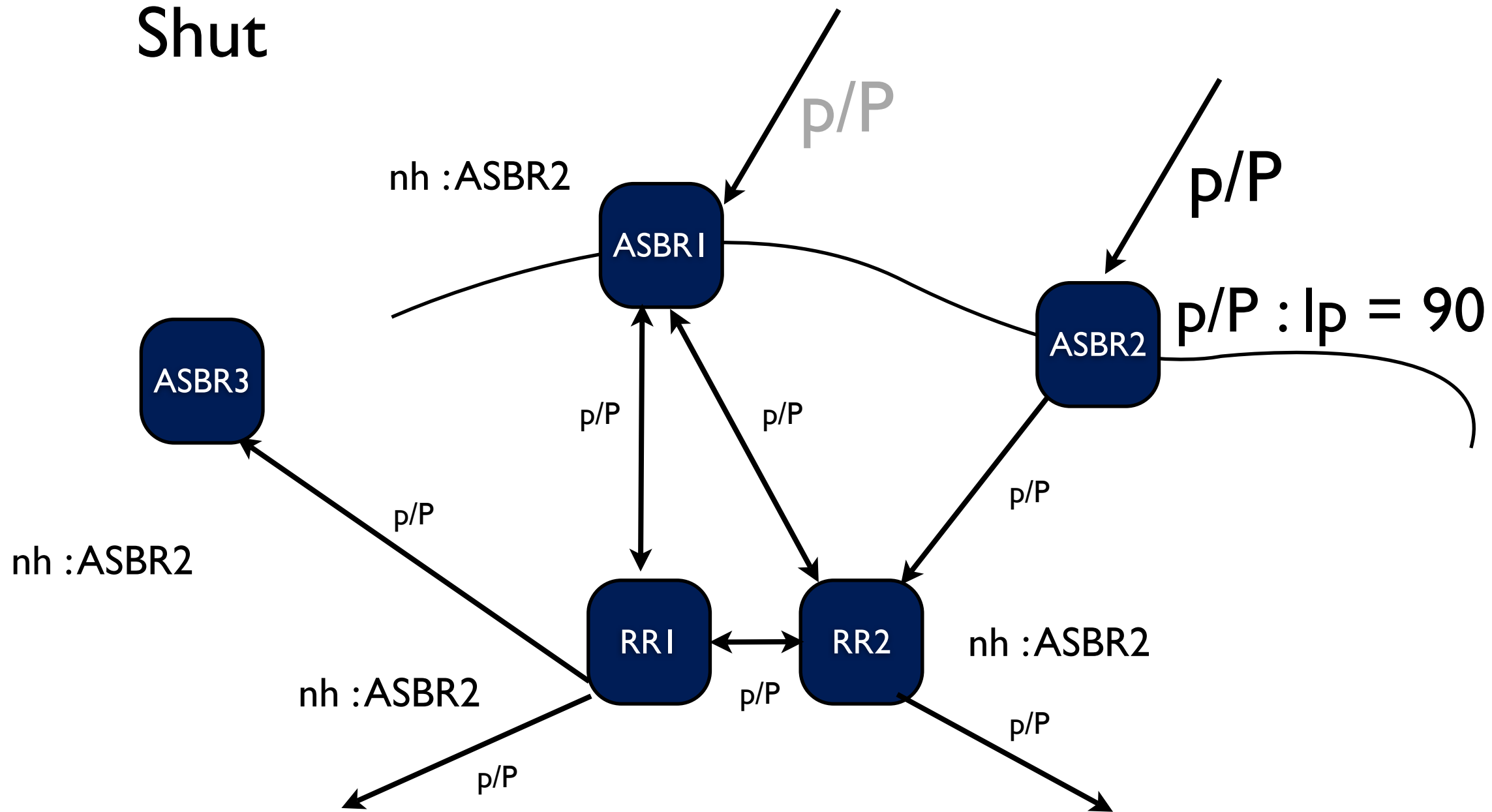
Shut

Inbound traffic

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

ASBR2

ASBR1

z/Z:

tag community : GSHUT

CE2

CE1

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

No need for action from the customer at maintenance time
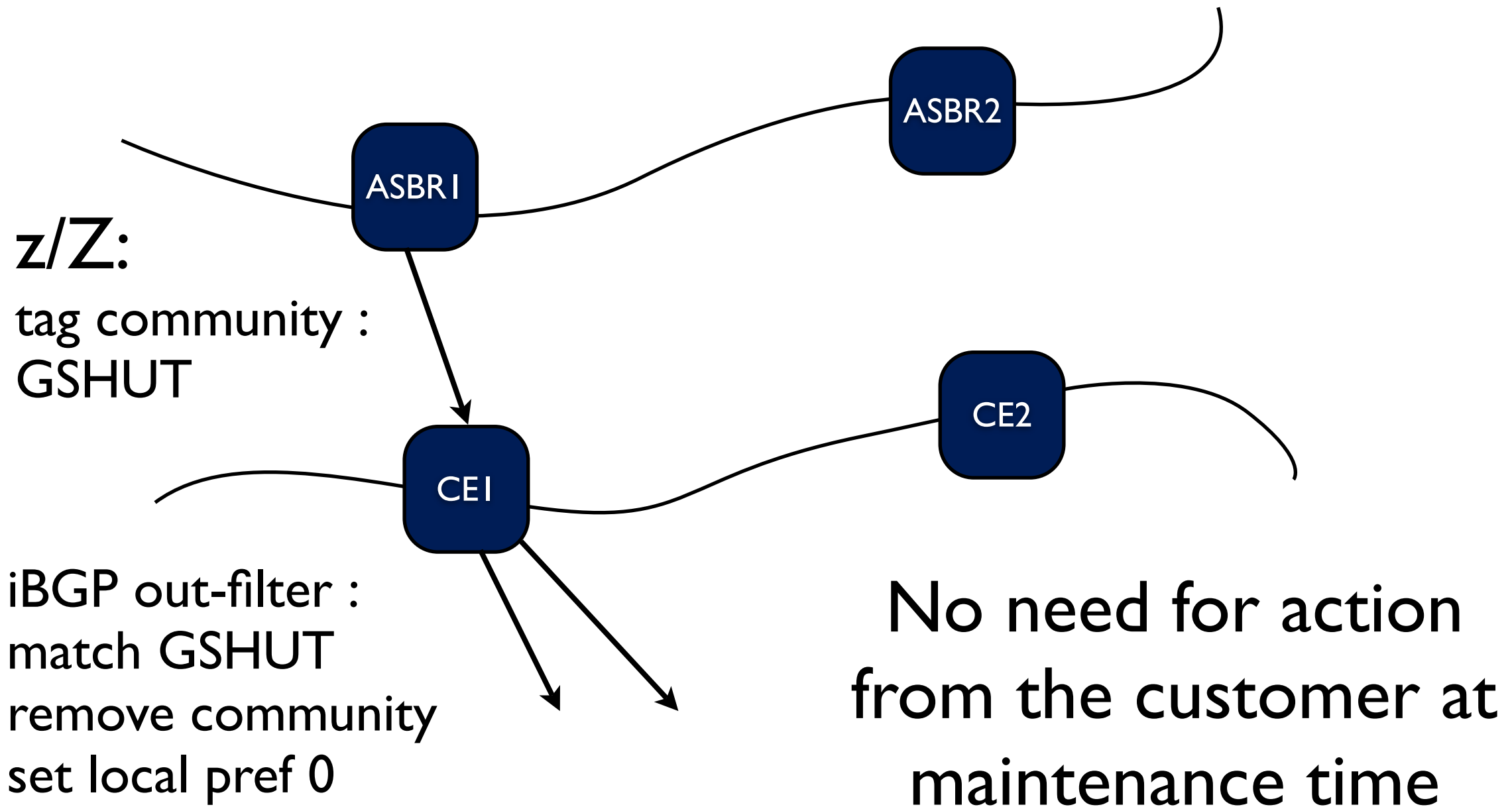
Deployment simplicity

- Incremental deployment

  - per eBGP peering session

- Incremental gain

- Can be implemented now by ISPs through configuration

  - Vendors could help make it simpler

Why documenting this ?

- Gather input and improvements from the community

- Inform about the procedure

  - If you want to implement g-shut

  - If your peers want to implement g-shut

- Let vendors think about how to ease this behavior

- Discuss the value of that community (next slide)

The g-shut community

- One g-shut community value per customer/peer/provider

  - difficult to manage

  - error prone

- We want to **reserve** one for this purpose

Changes

- Status : Informational

- Can an Informational Document describe what is the goal of a reserved FCFS community (0xFFFF0000) ?

Changes(2)

- Take comments from -00 into account

  - Link up

    - LOC

    - What you can do about it

  - Cases where LP of iBGP paths are overridden by ingress ASBR

    - Open the option to use the g-shut community over iBGP

What do we get ?

- 0 packet loss for link down and up cases

  - Some corner case timings require best-external (intra-cluster-best) support

  - still provides very significant loss reduction without support of best-external

Next steps

- Actually reserve the community

- WG Doc ?

?