Influence des bonnes pratiques sur les incidents BGP

François Contat, Sarah Nataf, Guillaume Valadon
francois.contat@ssi.gouv.fr, sarah.nataf@orange.com, guillaume.valadon@ssi.gouv.fr

Agence nationale de la sécurité des systèmes d'information
&
France Télécom Orange

8 juin 2012

Contexte

Border Gateway Protocol (BGP)

- le protocole de routage utilisé par tous les acteurs/opérateurs de l'Internet ;
- interconnecte directement ces acteurs ;
- assure une présence mondiale à ces acteurs ;
- ne protège pas les données échangées.

Les bonnes pratiques limitent les impacts des incidents touchant le protocole BGP.

Comment mettre en œuvre BGP ?

Pour qu'une organisation puisse utiliser BGP, il faut :

- un numéro d'AS : identifiant unique dans Internet ;
- un bloc IP : divisé en préfixes ;
- un routeur BGP : équipement réseau dédié ou simple PC ;
- une interconnexion BGP :
  - transit : l'AS apprend toutes les prefixes de l'Internet ;
  - peering : l'AS apprend seulement les prefixes de l'autre AS.

Fin mai 2012, l'Internet est composé d'environ 41658 AS, 448332 préfixes IPv4 et 9543 préfixes IPv6. Fin mai 2002, il y avait 13088 AS, 110303 préfixes IPv4, et 0 préfixe IPv6.

Sécurité des sessions

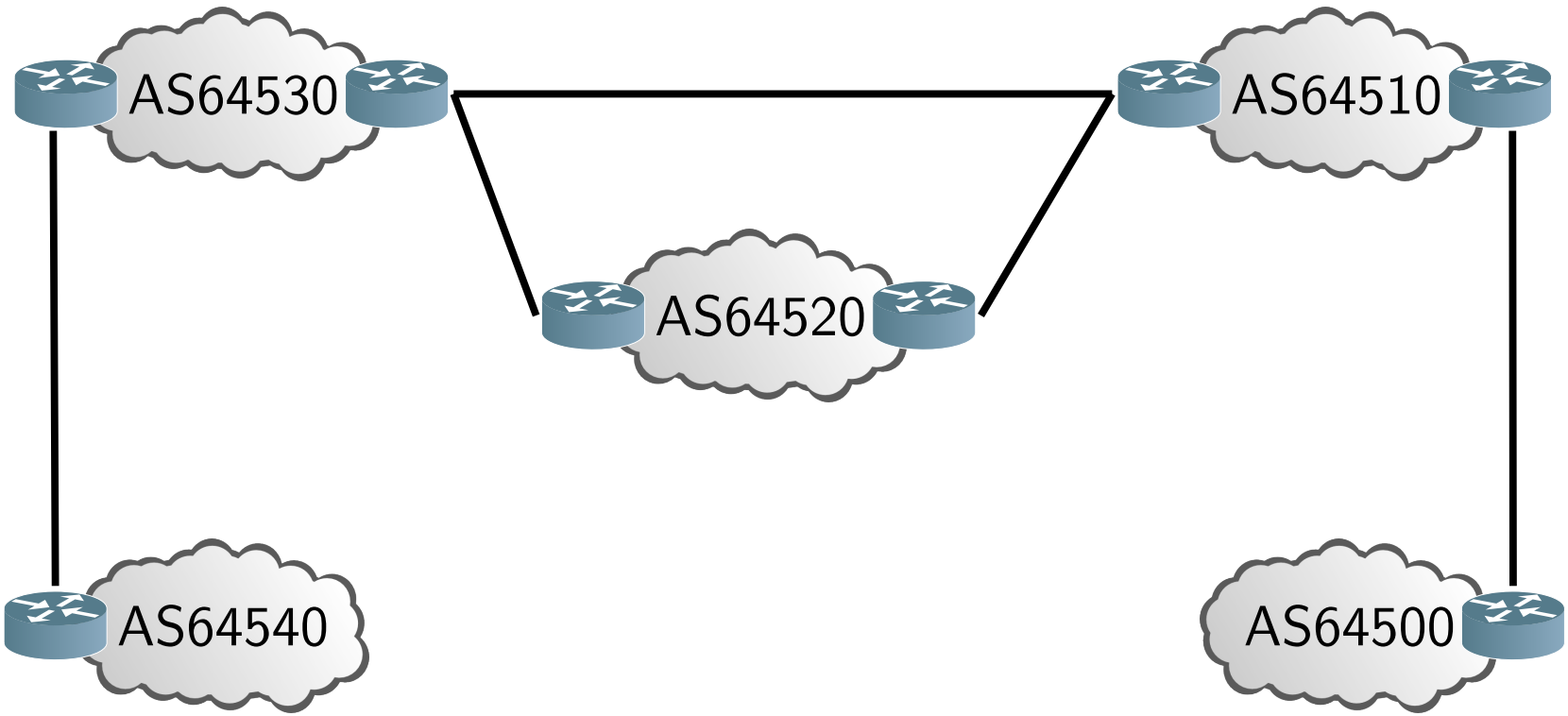
Une session BGP s'établit entre deux routeurs directement connectés.

Mécanismes de protection des sessions :

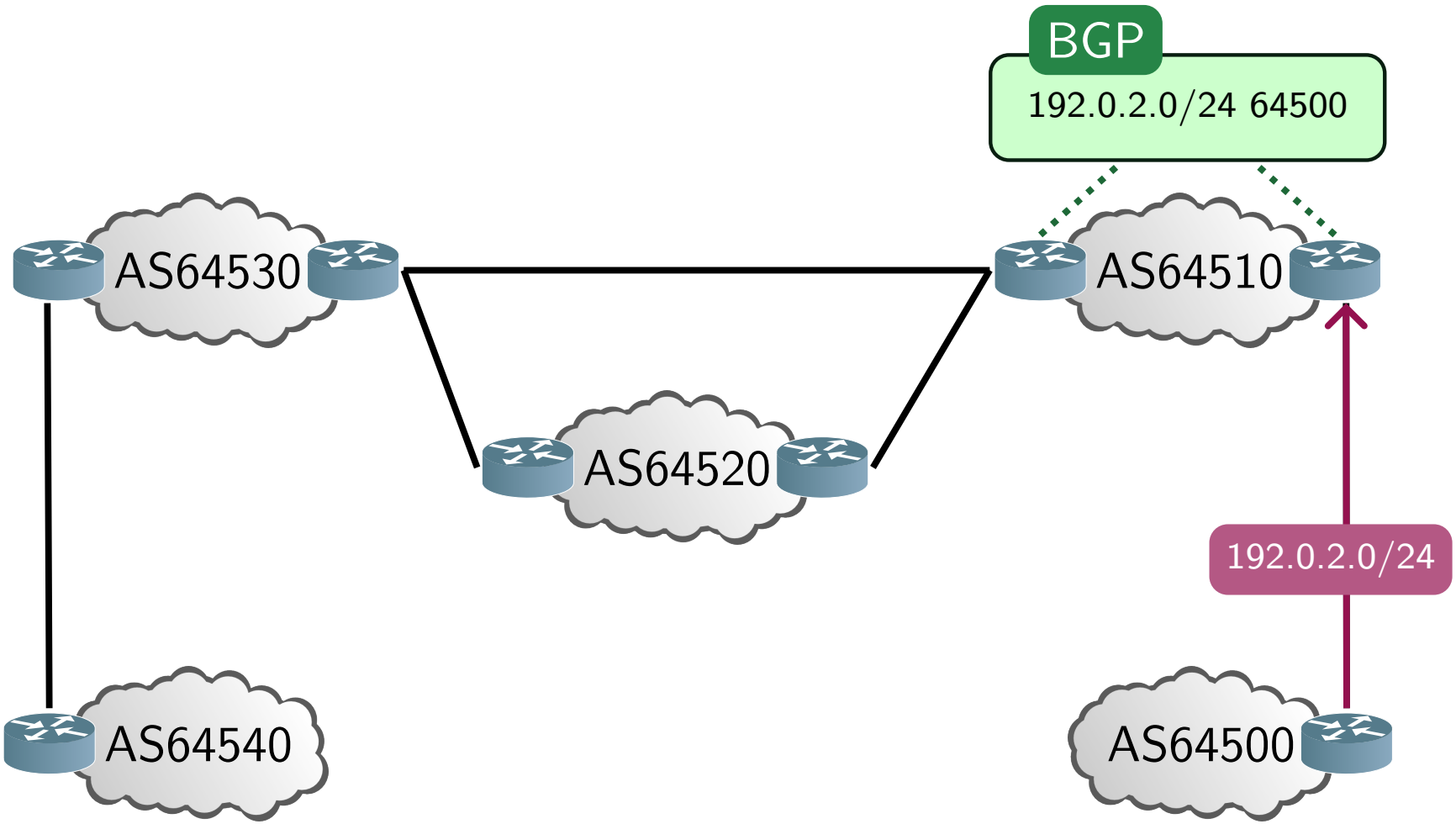
- la vérification de l'adresse IP et du numéro d'AS de son pair ;
- Vérification du TTL : les paquets sont émis avec un TTL à 255. À la réception tous les paquets ayant un TTL différent sont supprimés ;
- TCP MD5 : un HMAC-MD5 est envoyé avec chaque segment TCP émis.

La protection des sessions BGP est particulièrement importante dans les points d'échange où les routeurs sont dans le même LAN.

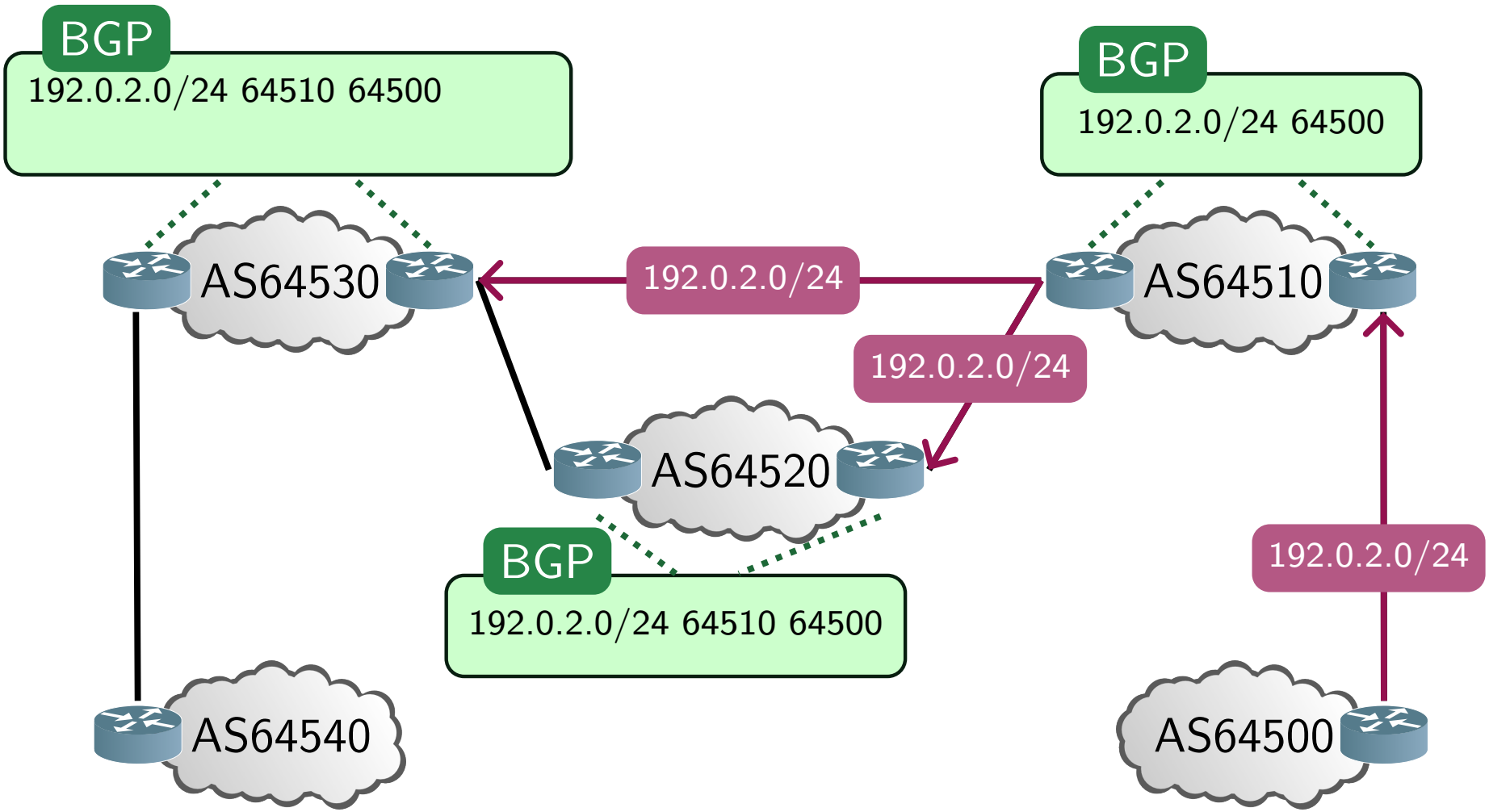
Les annonces de préfixes



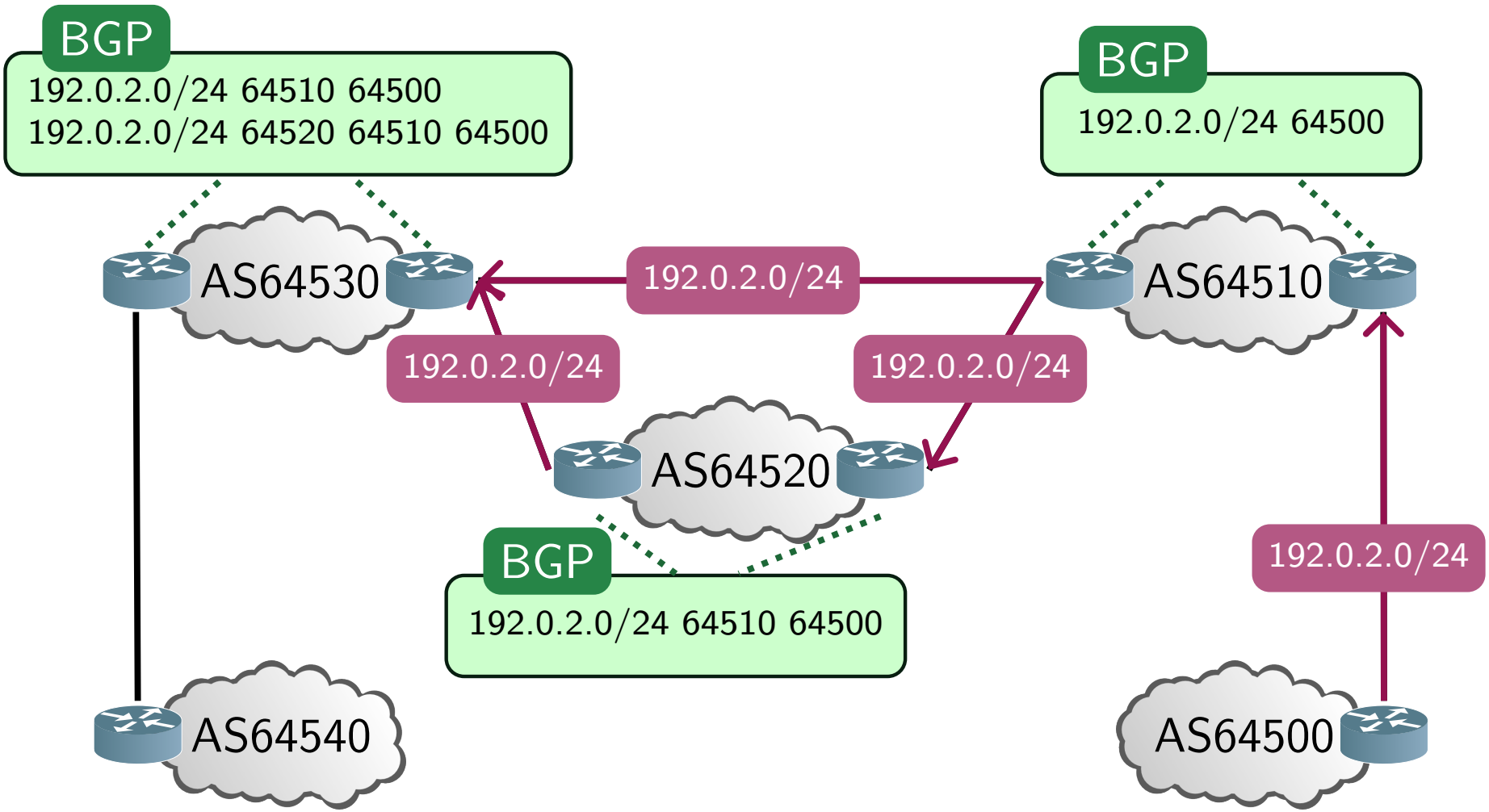
Les annonces de préfixes



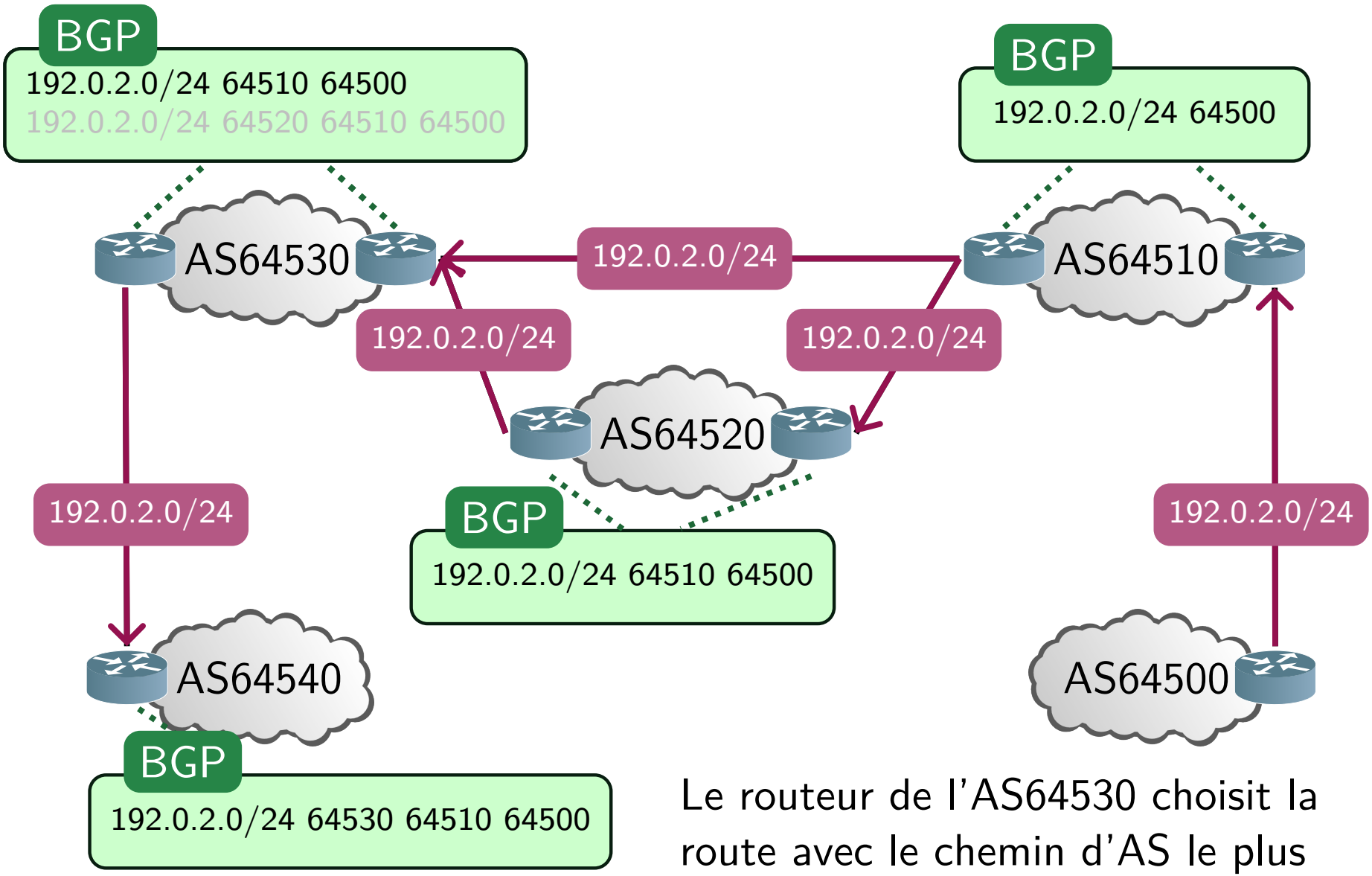
Les annonces de préfixes



Les annonces de préfixes

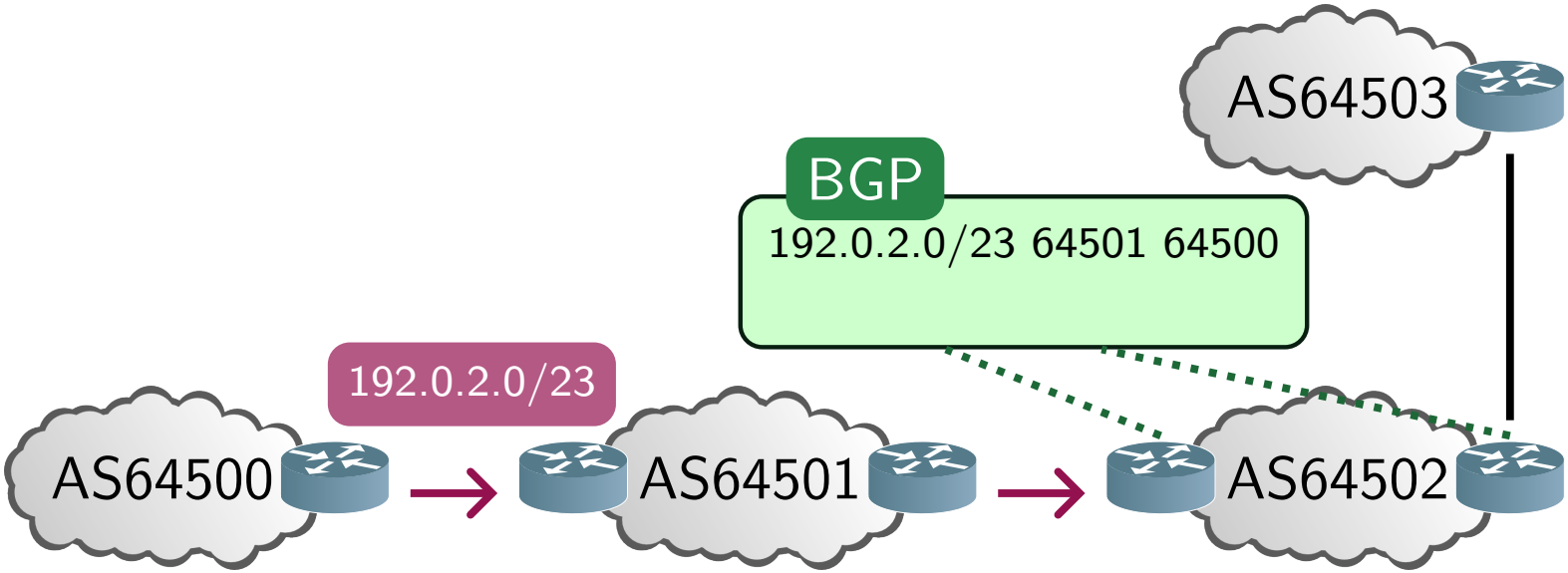


Les annonces de préfixes



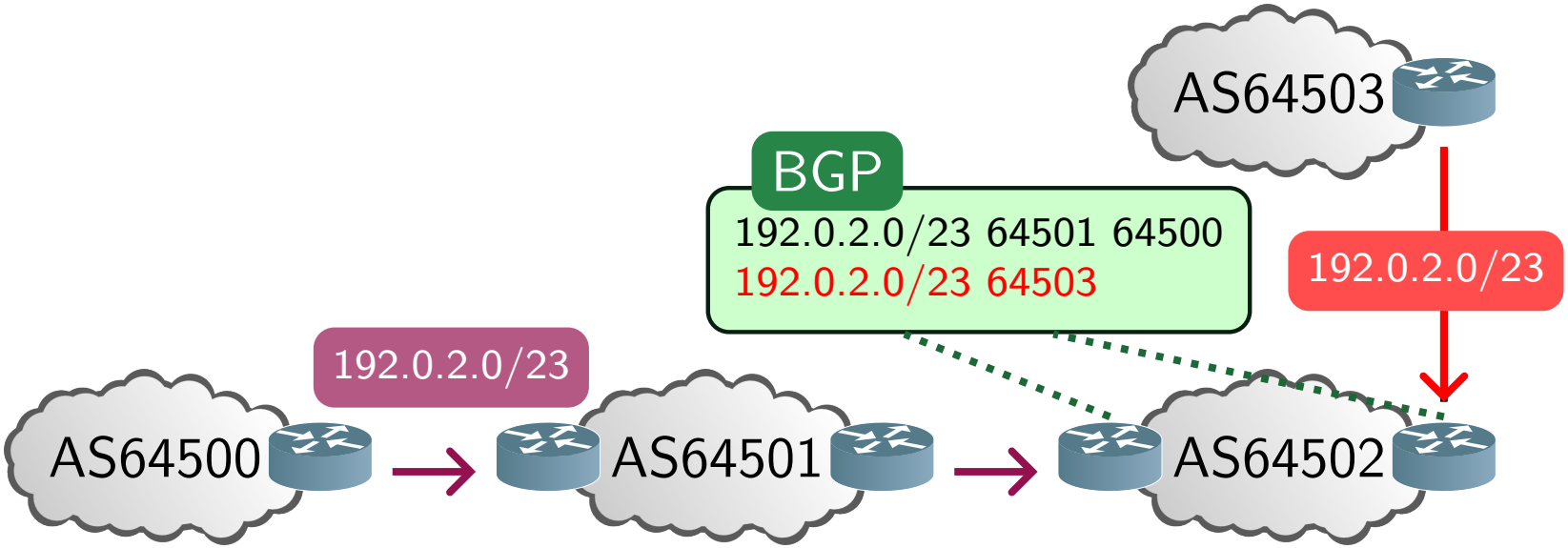
Le routeur de l'AS64530 choisit la route avec le chemin d'AS le plus court.

Usurpation de préfixes



Usurpation : annonce d'un préfixe par un AS ne le gérant pas.
Incidents publics survenus en 2004, 2005, 2006 et 2008.

Usurpation de préfixes



Usurpation : annonce d'un préfixe par un AS ne le gérant pas.
Incidents publics survenus en 2004, 2005, 2006 et 2008.

Par défaut, un routeur choisit les chemins les plus courts.
L'AS64503 reçoit ainsi du trafic pour le préfixe 192.0.2.0/23.

Protection contre les usurpations

1. Mesures préventives

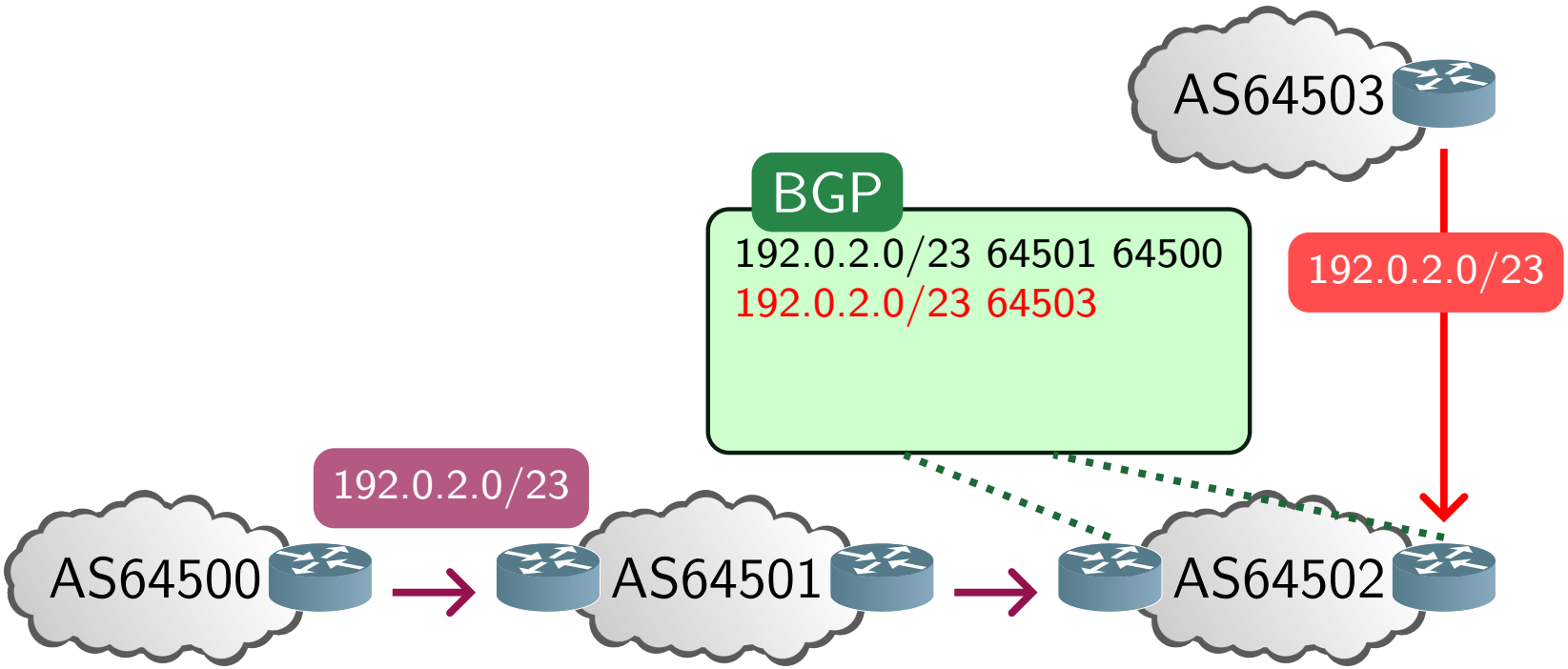
Un AS de transit peut empêcher les usurpations effectuées par ses clients à l'aide de filtres stricts.
Leur création peut être automatisée à l'aide des déclarations effectuées (RIPE) :

```
route:          192.0.2.0/23
origin:         AS-64500
mnt-by:         STIC-MNT
```

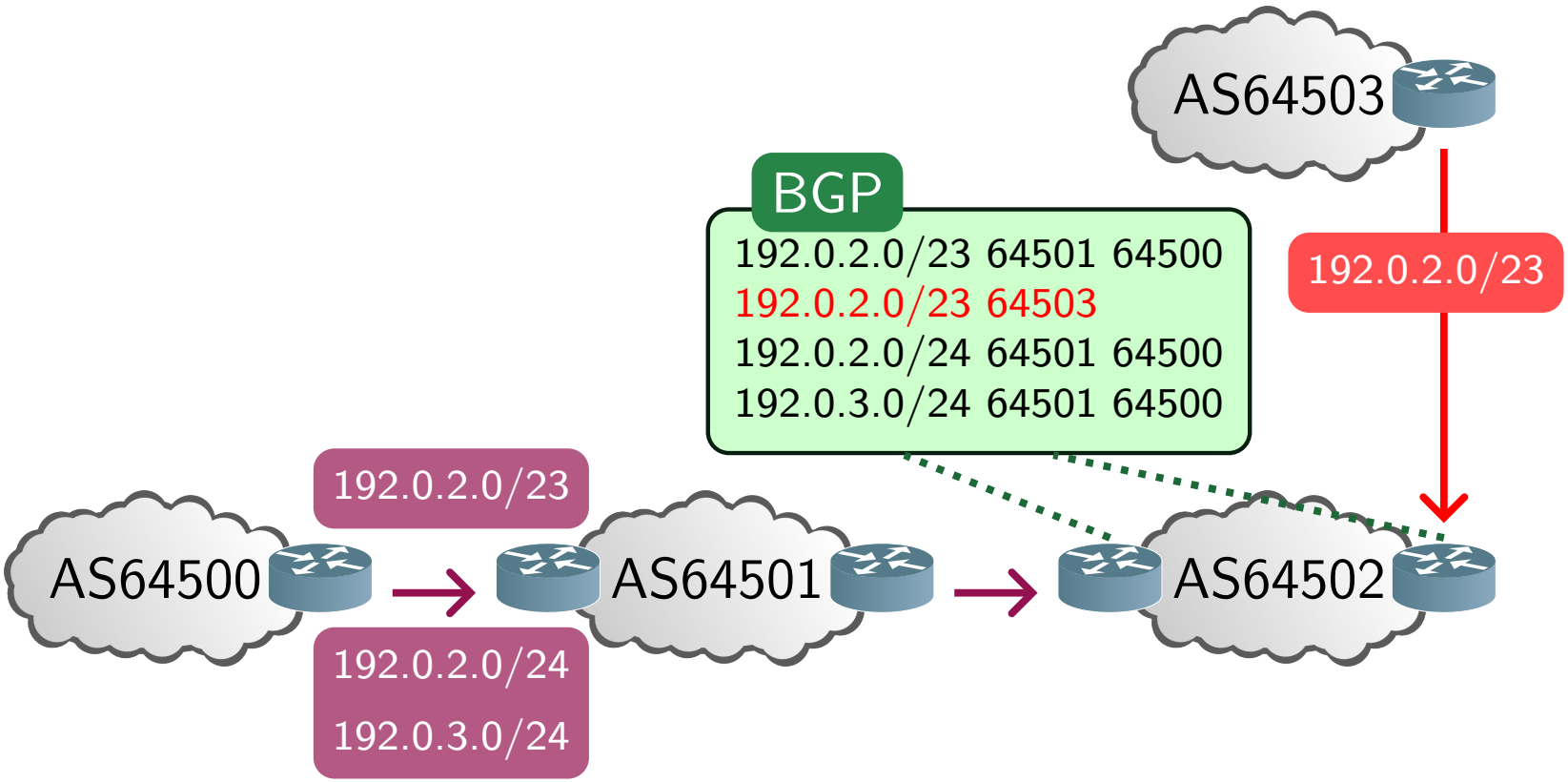
2. Mesures actives

Lors d'une usurpation, l'AS gérant les préfixes peut tenter de limiter les effets d'une usurpation en annonçant des préfixes plus spécifiques.

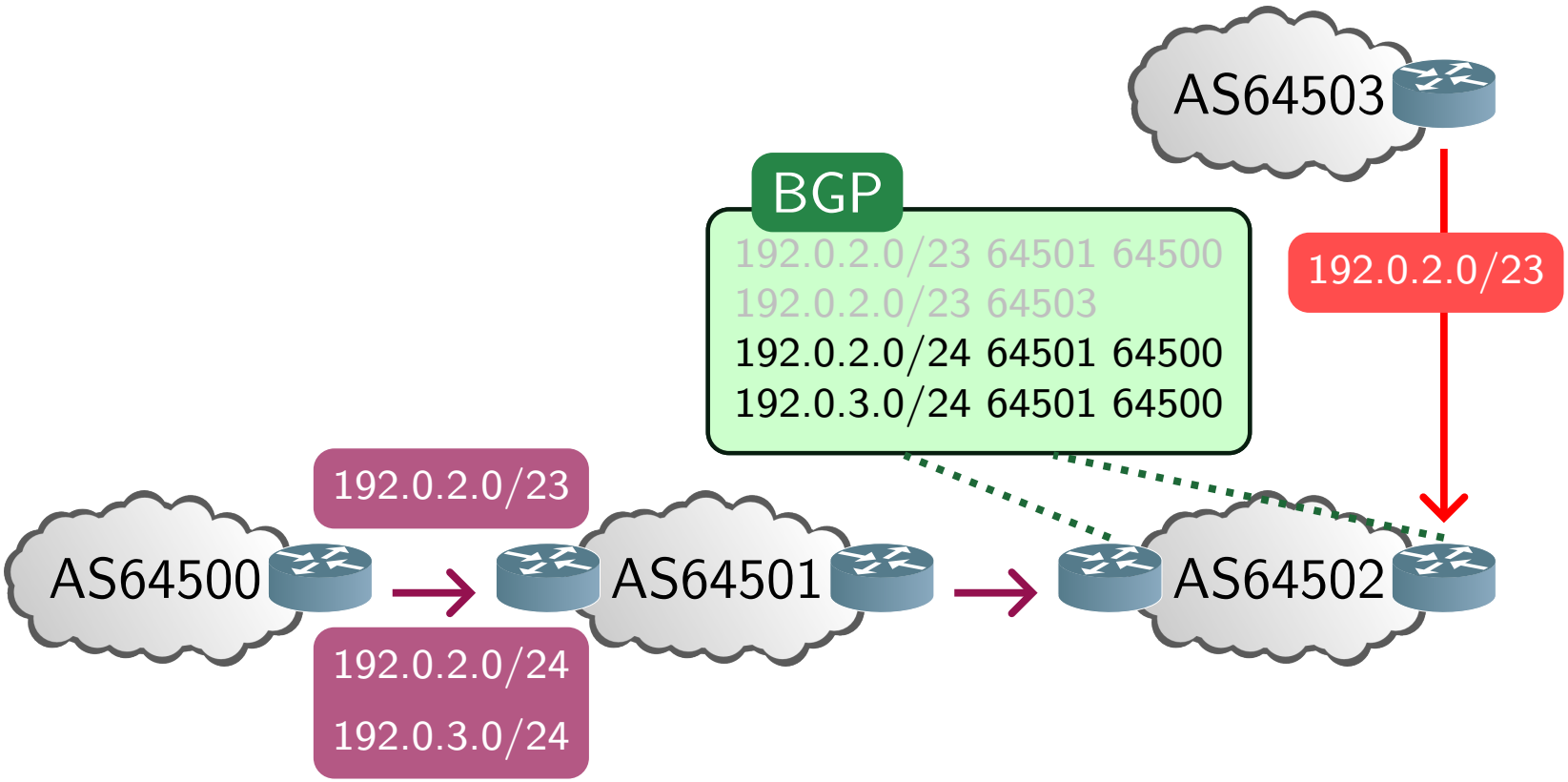
Préfixes plus spécifiques contre l'usurpation



Préfixes plus spécifiques contre l'usurpation



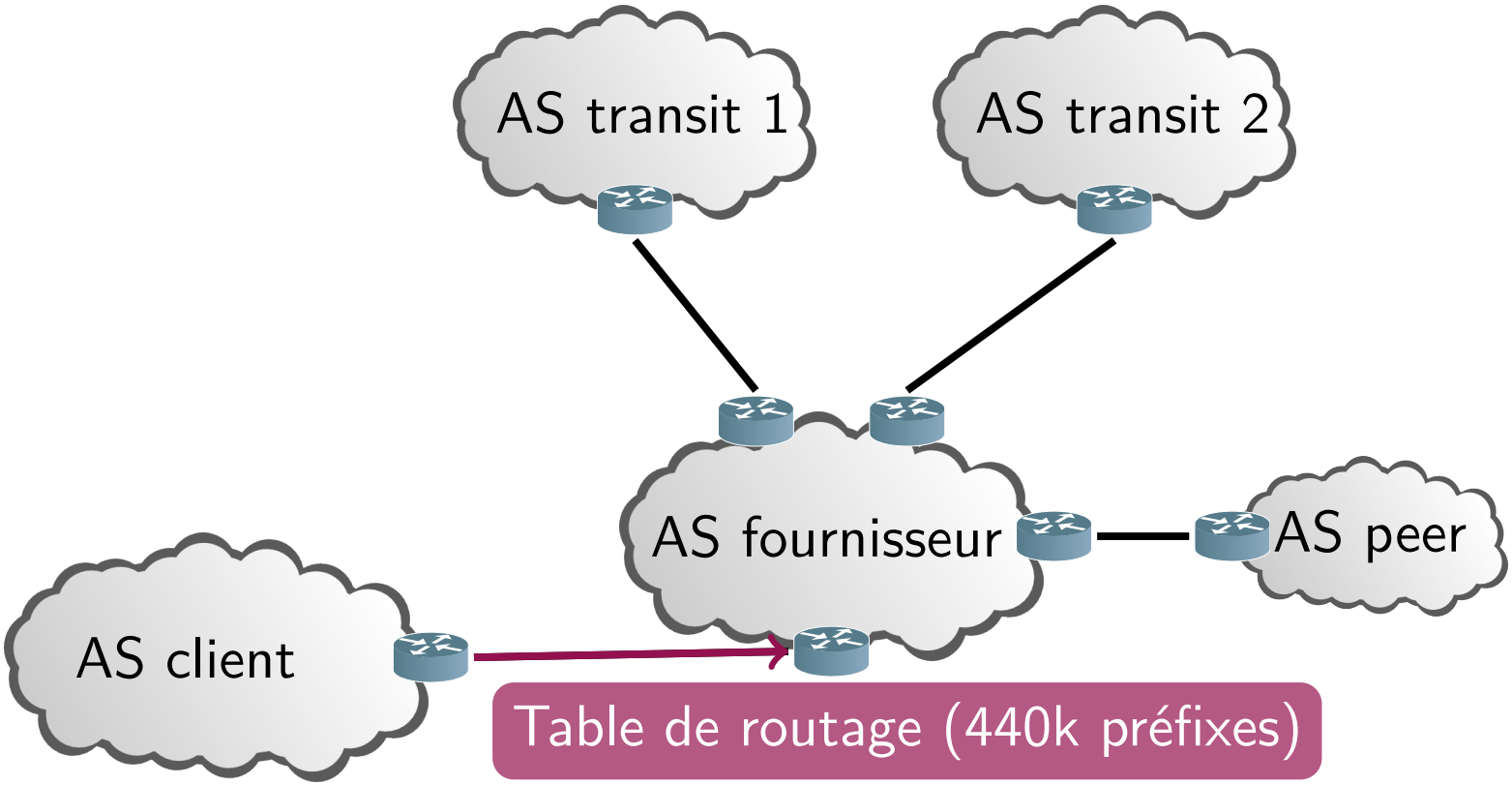
Préfixes plus spécifiques contre l'usurpation



Le routeur de l'AS 64502 choisit les préfixes plus spécifiques.
L'AS 64500 récupère son trafic.

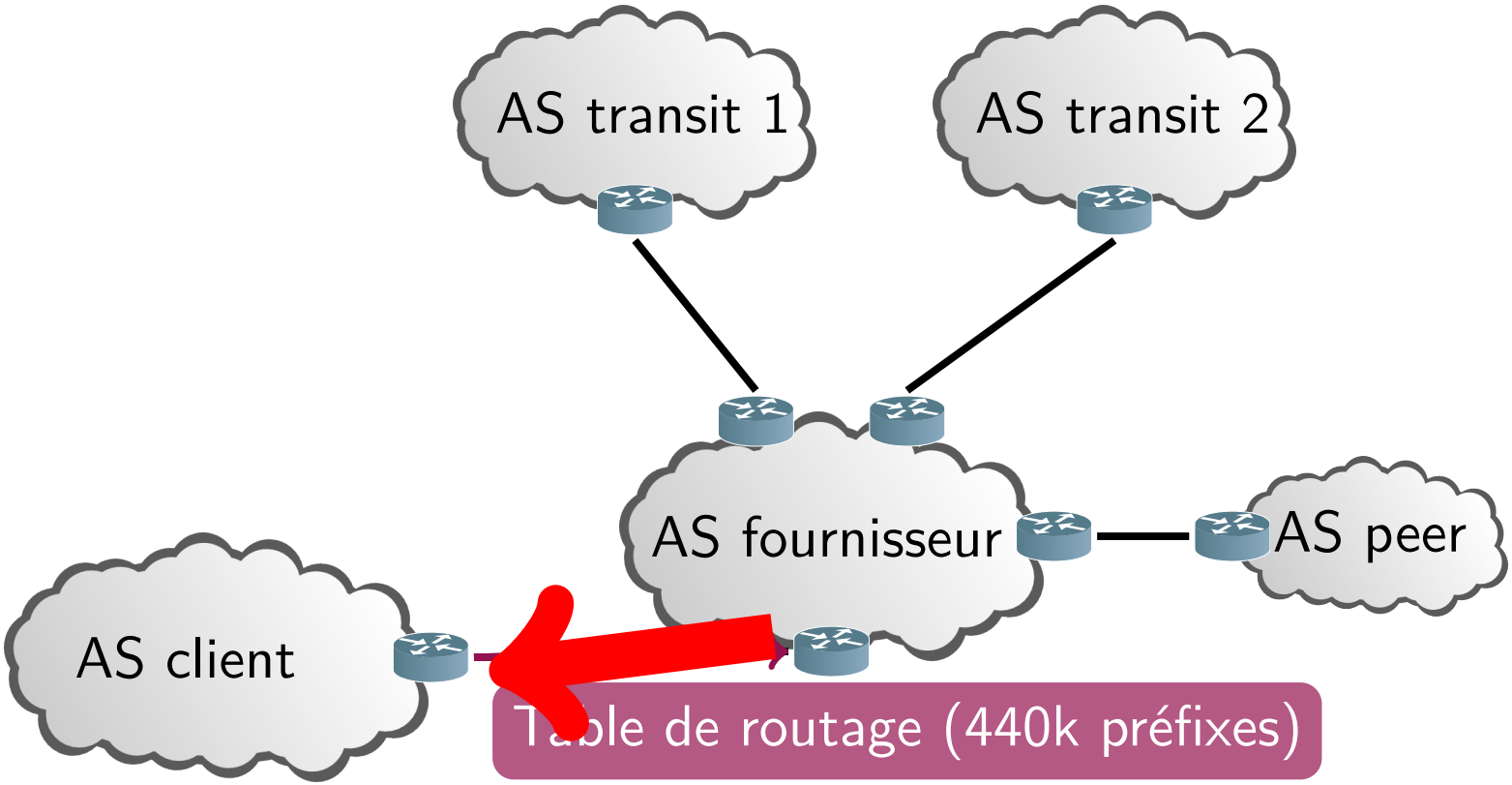
Réannonce d'une table de routage

Un AS réannonce à ses voisins tout ou partie d'une table de routage. Incidents publics survenus en : 2004, 2008, 2010 et 2012.



Réannonce d'une table de routage

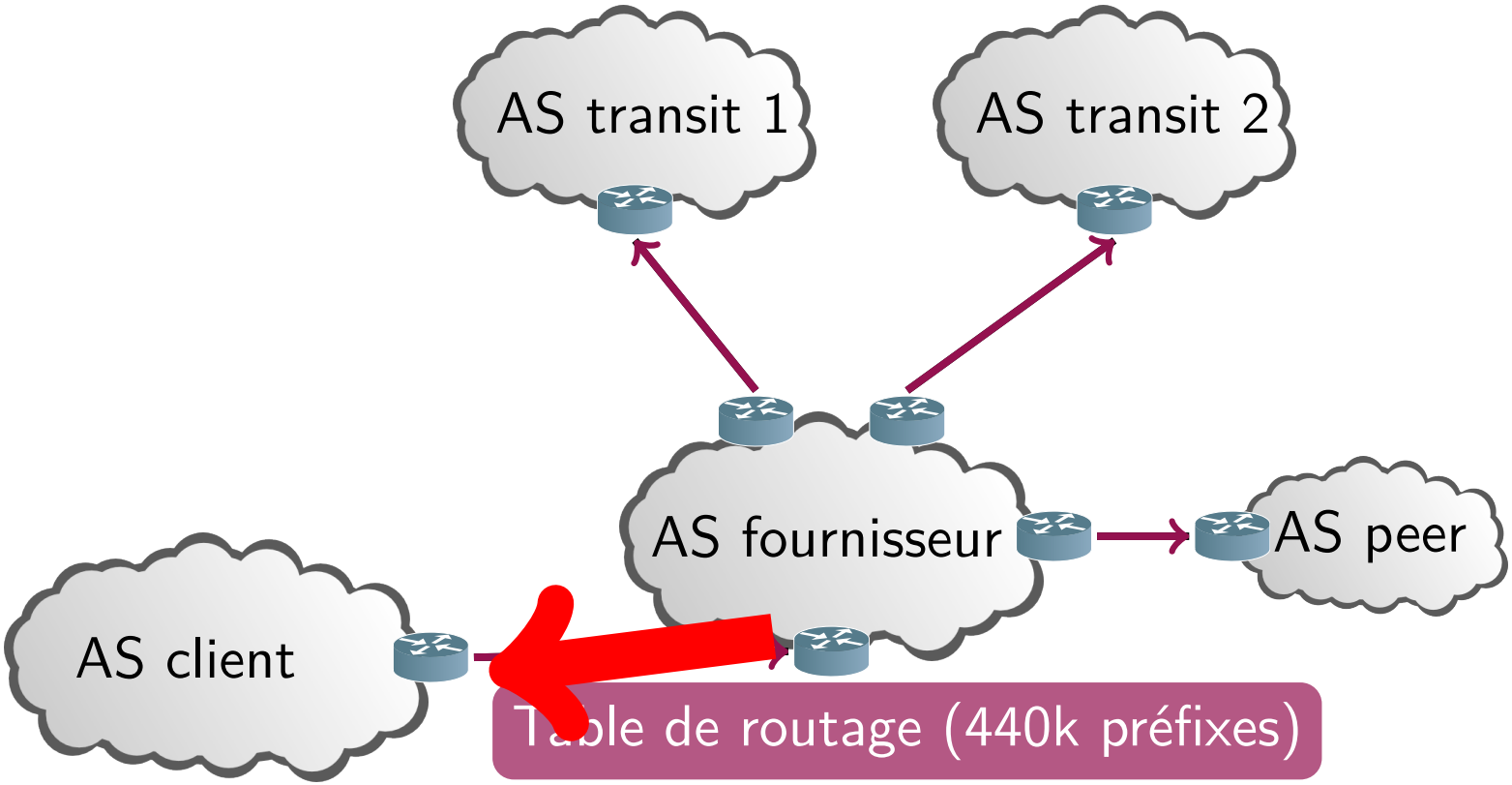
Un AS réannonce à ses voisins tout ou partie d'une table de routage. Incidents publics survenus en : 2004, 2008, 2010 et 2012.



L'AS client reçoit tout le trafic Internet de l'AS fournisseur

Réannonce d'une table de routage

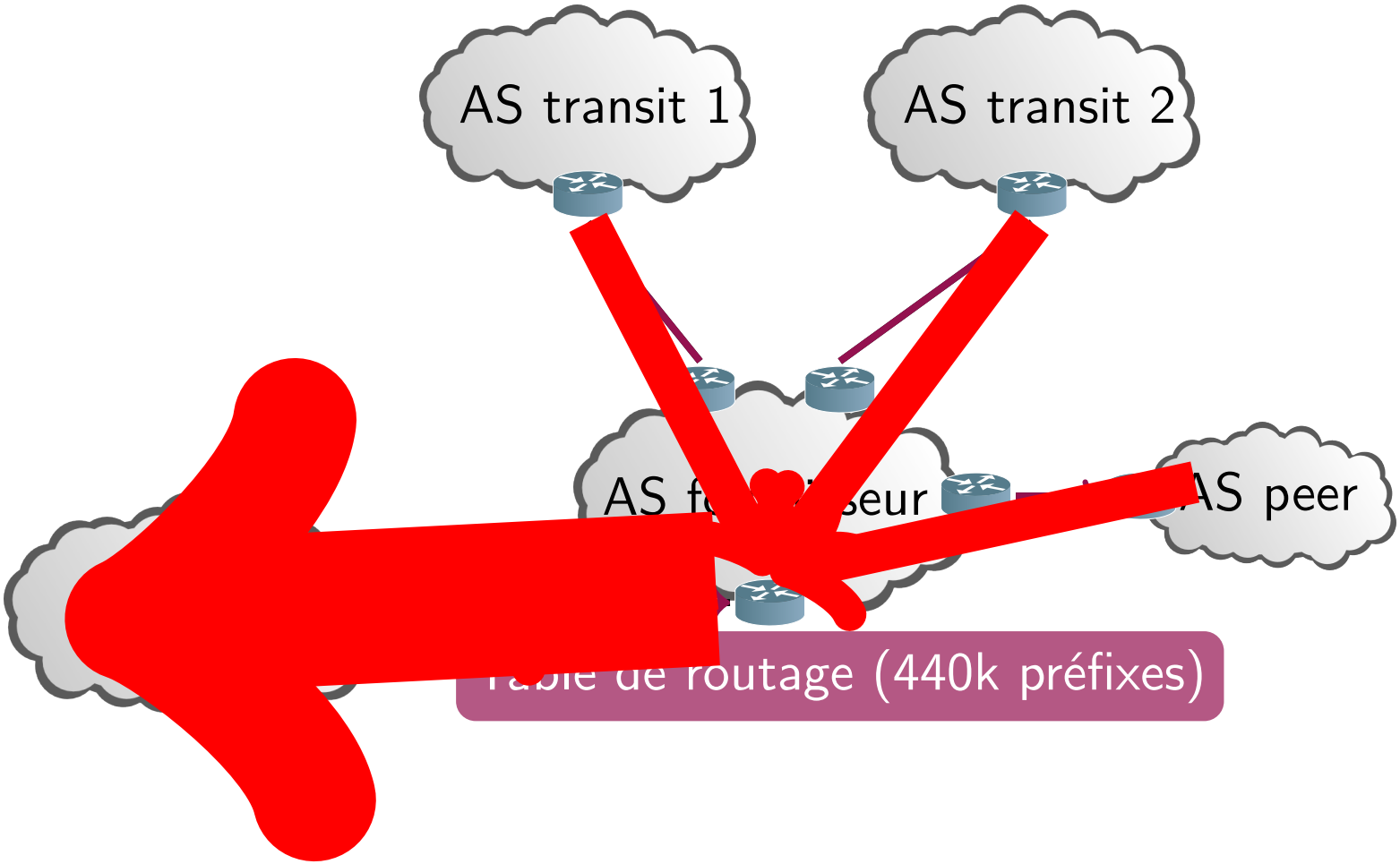
Un AS réannonce à ses voisins tout ou partie d'une table de routage. Incidents publics survenus en : 2004, 2008, 2010 et 2012.



L'AS client reçoit tout le trafic Internet de l'AS fournisseur

Réannonce d'une table de routage

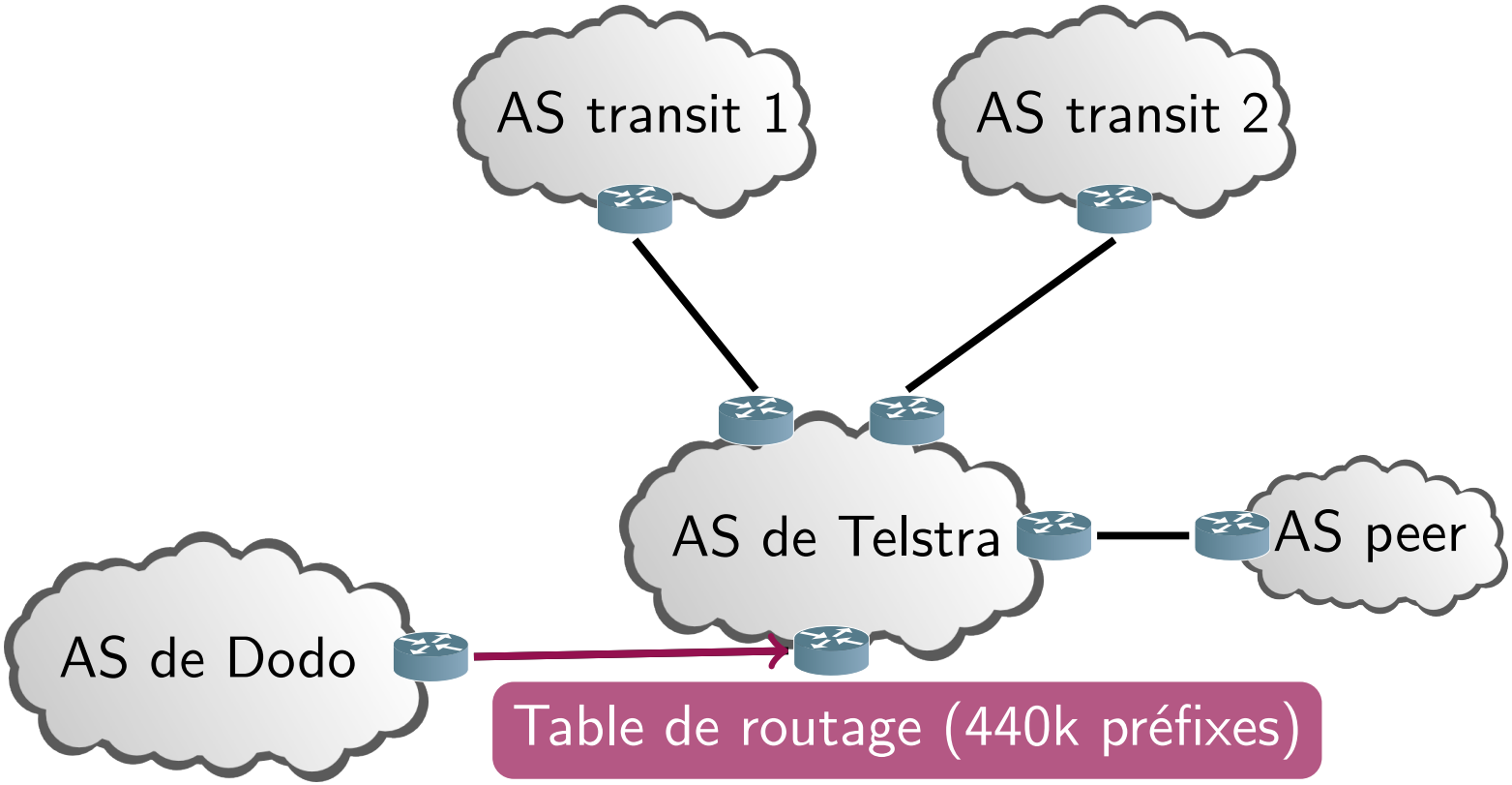
Un AS réannonce à ses voisins tout ou partie d'une table de routage. Incidents publics survenus en : 2004, 2008, 2010 et 2012.



L'AS client reçoit tout le trafic Internet de l'AS fournisseur et de ses voisins.

Protection contre les réannonces : le *max-prefix*

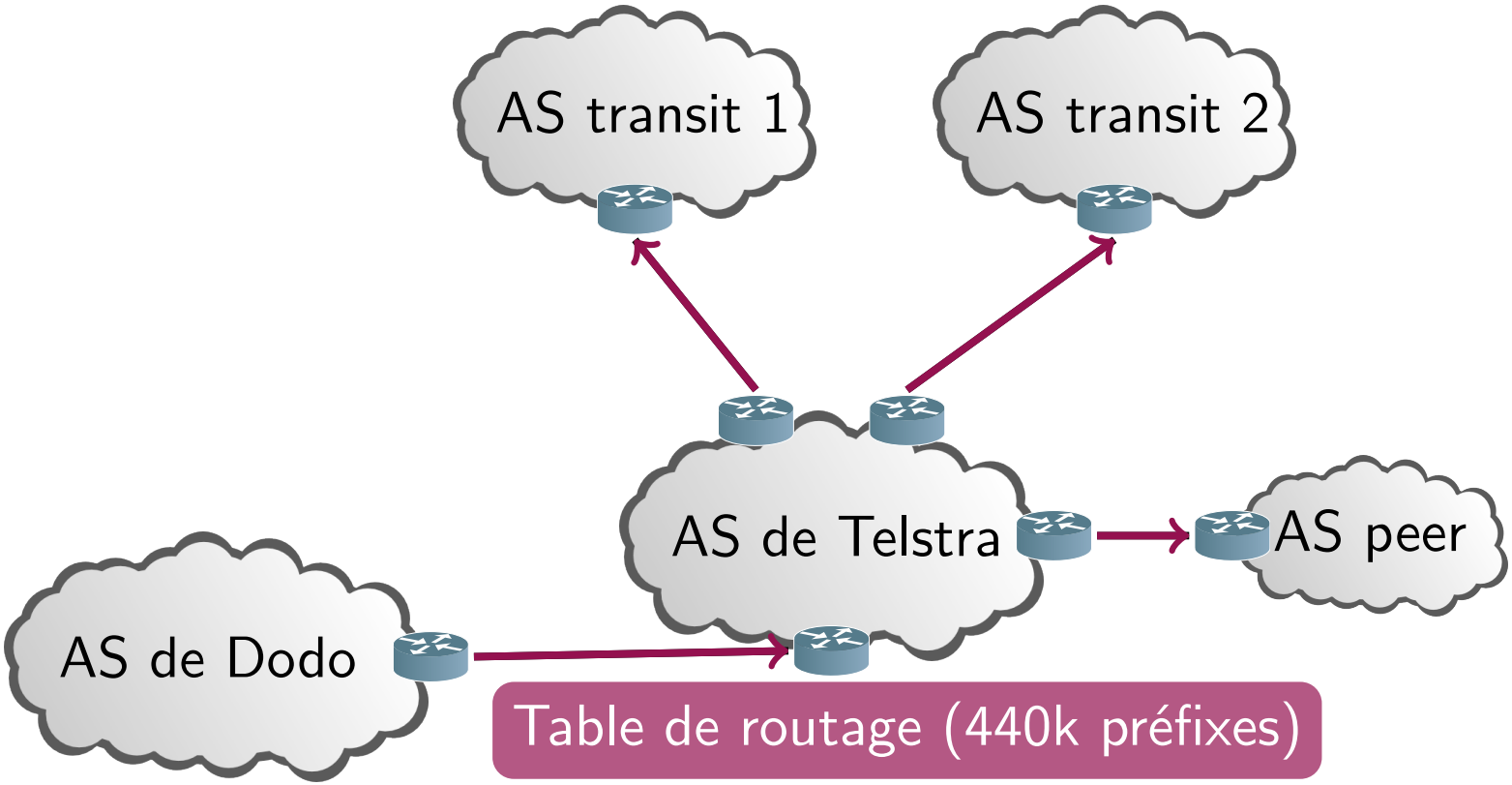
Un filtre *max-prefix* n'accepte qu'un nombre limité de préfixes.



Le 23 février 2012, Dodo a annoncé la table de routage de l'Internet à son transitaire Telstra.

Protection contre les réannonces : le *max-prefix*

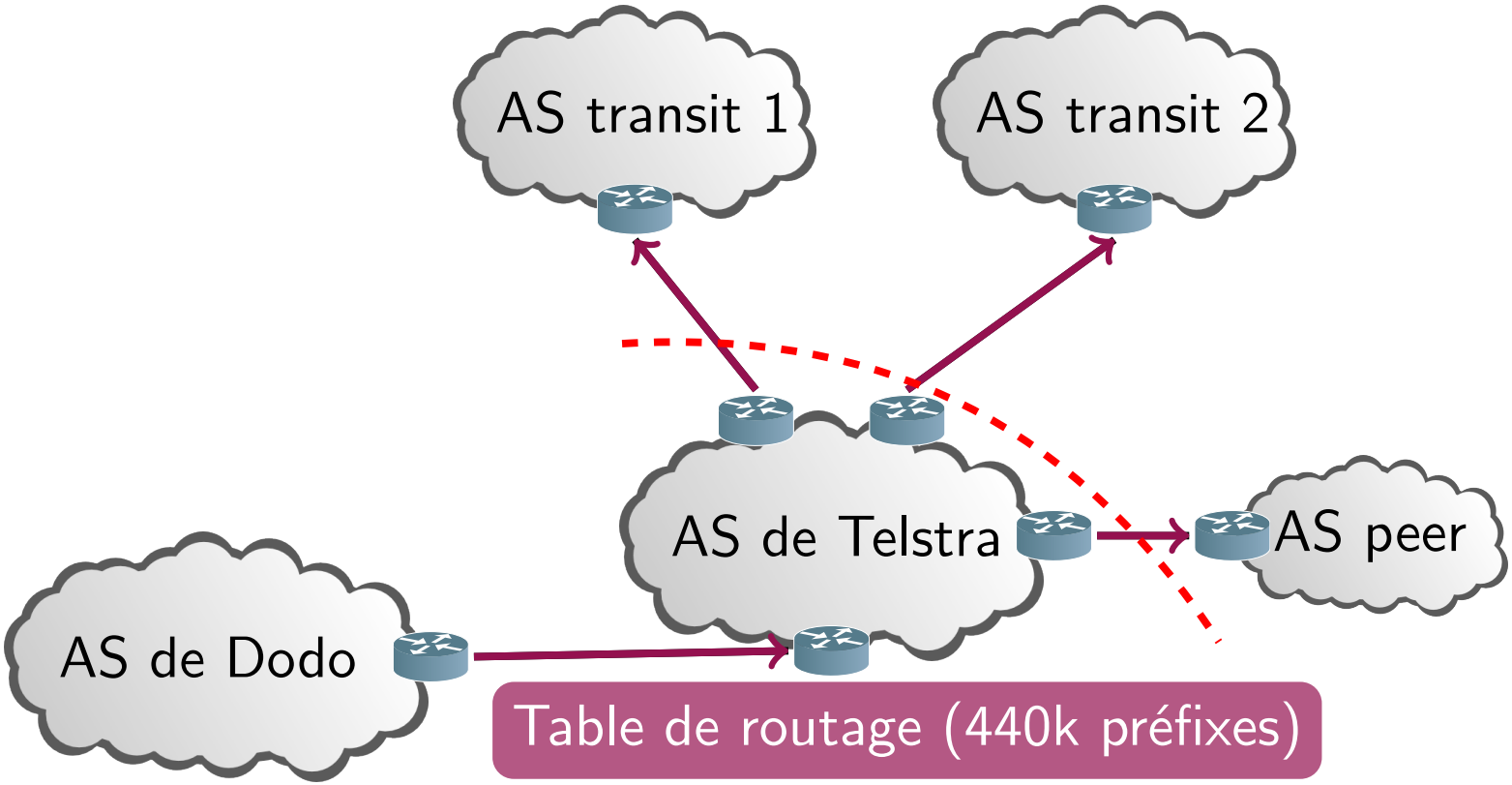
Un filtre *max-prefix* n'accepte qu'un nombre limité de préfixes.



Le 23 février 2012, Dodo a annoncé la table de routage de l'Internet à son transitaire Telstra. Telstra a redistribué ces routes à ses pairs.

Protection contre les réannonces : le *max-prefix*

Un filtre *max-prefix* n'accepte qu'un nombre limité de préfixes.

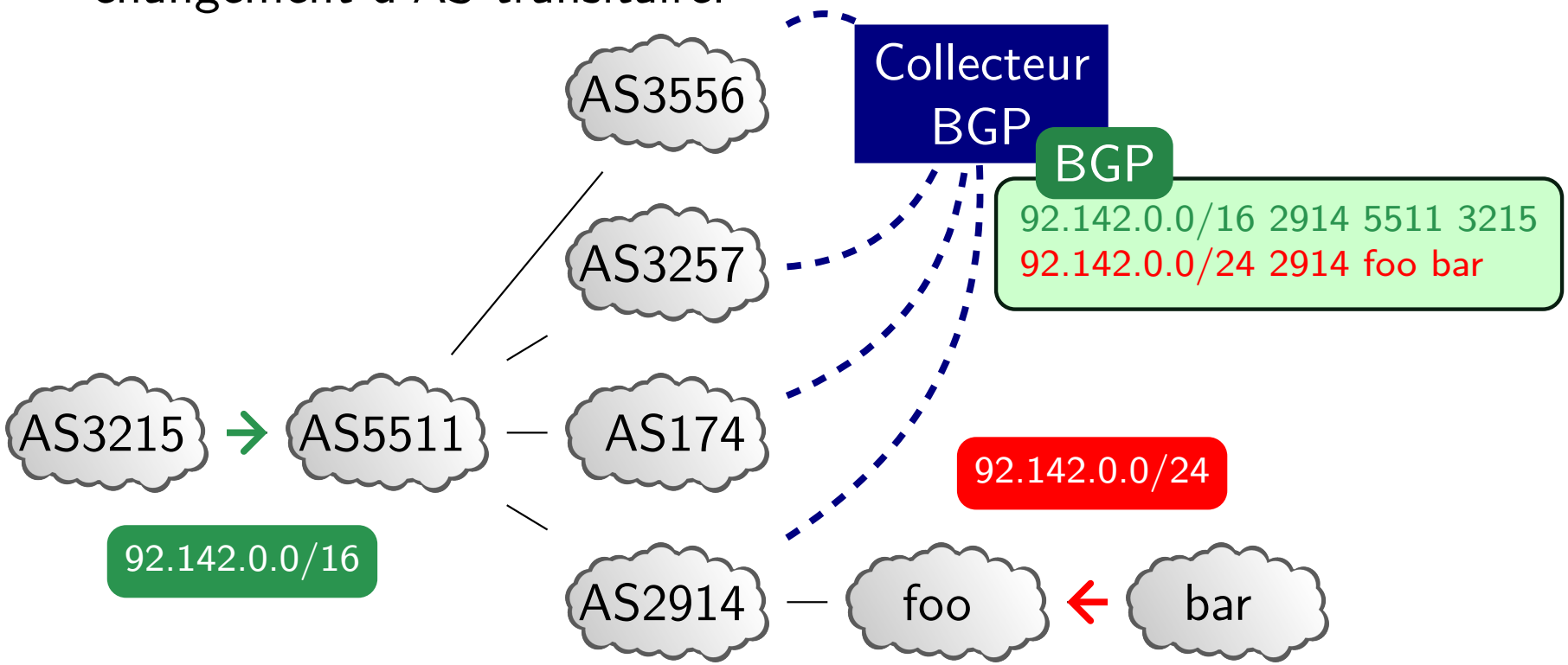


Le 23 février 2012, Dodo a annoncé la table de routage de l'Internet à son transitaire Telstra. Telstra a redistribué ces routes à ses pairs. Ses pairs l'ont isolé de l'Internet grâce au mécanisme *max-prefix*.

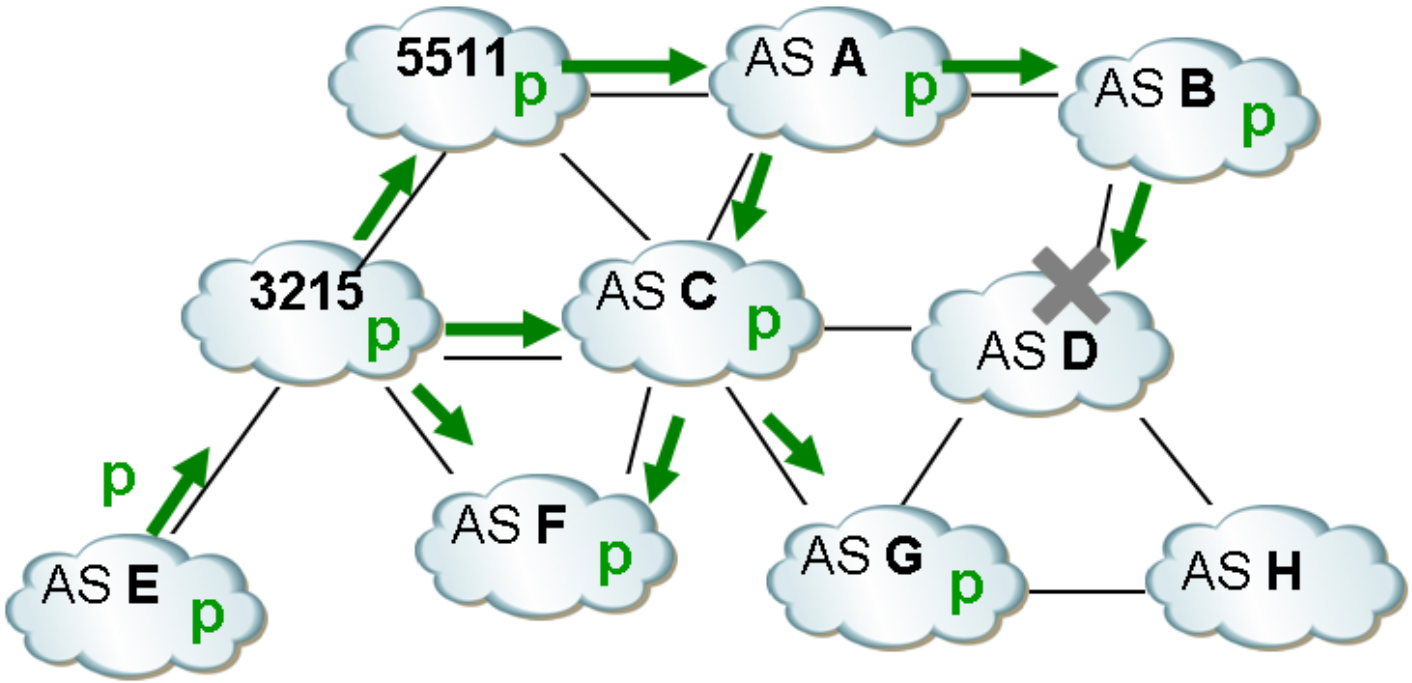
Les outils

Outils construits autour de sondes BGP qui traitent les tables de routage ou les UPDATES BGP et détectent les anomalies :

- disparition d'un préfixe ;
- nouveau préfixe pour un AS donné ;
- préfixe identique annoncé par un autre AS ;
- préfixe plus spécifique (e.g. un /24 appartenant à un /16) ;
- changement d'AS transitaire.



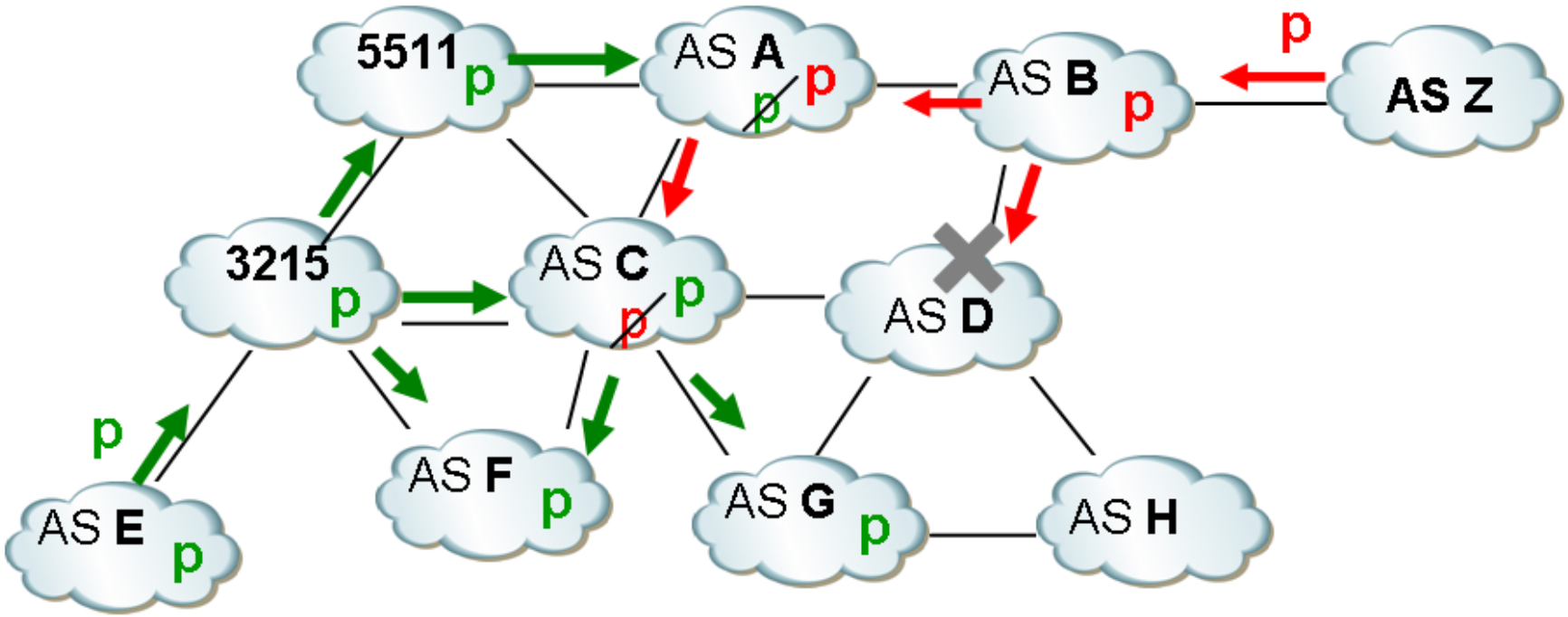
Convergence sur l'annonce d'un préfixe - cas normal



Sur l'Internet, les tables de routage diffèrent du fait des politiques de routage implémentées dans chaque AS.

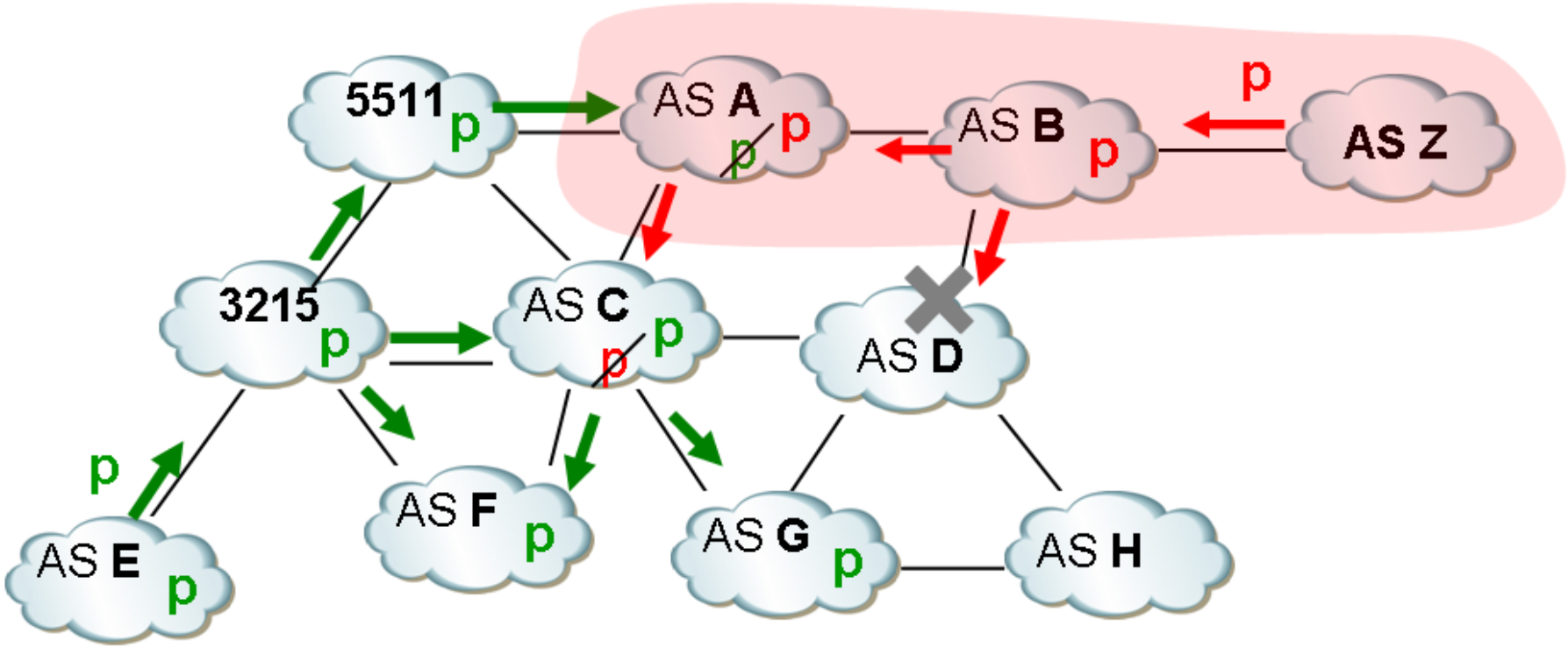
- certains AS ne reçoivent pas certains préfixes ;
- d'où l'intérêt de sondes de monitoring réparties pour améliorer la détection de l'usurpation.

Convergence sur l'annonce d'un préfixe - usurpation



- lorsqu'un préfixe p est usurpé par un AS, il est fréquent que seule une partie de l'Internet converge sur la nouvelle route.

Zone de pollution d'une usurpation



- on appelle zone de pollution l'ensemble des AS pour lequel le trafic à destination du préfixe p diverge vers l'AS origine de l'usurpation ;
- l'impact dépend de la connectivité de l'AS et des politiques de filtrage de chacun.

Usurpations récentes

- lorsque l'usurpateur annonce un préfixe plus spécifique que le préfixe légitime, tous les AS recevant l'UPDATE convergent vers la destination illégitime (en l'absence de filtres) ;
- de rares incidents en 2009 et 2010 ;

```
============================================================
Possible Prefix Hijack (Code: 10)
============================================================
Your prefix:          92.142.0.0/16:
Update time:          2010-06-03 11:15 (UTC)
Detected by #peers:   56
Detected prefix:      92.142.8.0/22
Announced by:         AS1257 (TELE2)
Upstream AS:          AS2119 (TELENOR-NEXTEL T.net)
ASpath:               2119 1257

        ----------------------------------------------

AS 1257 is now announcing 92.142.8.0/22 which is a sub-prefix of
92.142.0.0/16.  92.142.0.0/16 is historically announced by ASes: 3215.
Time: Thu Jun  3 05:15:44 2010 GMT
Observed path: 812 1257
```

- accélération des usurpations depuis début 2011.

Avant annonce du 2.0.0.0/12

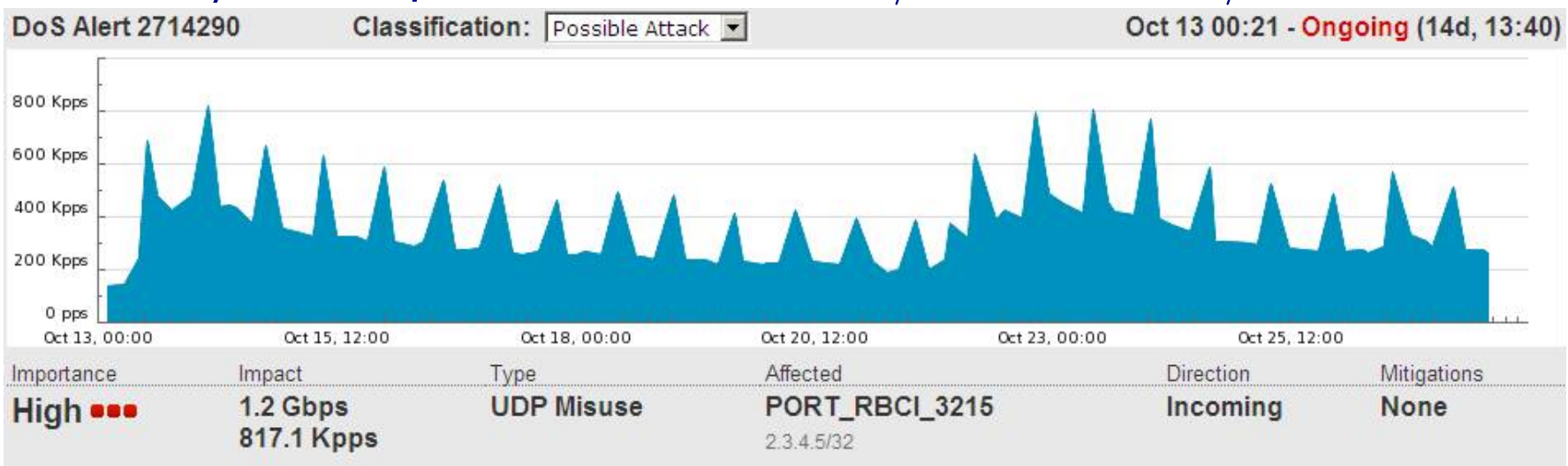
Historique des annonces avant le déploiement

3 entries found for 2.3.0.0/16 related prefixes

This group of prefixes was last seen by RIS on **2011-02-10 10:15:16 UTC**.

Prefix	Origin AS	First seen	Last seen
2.3.0.0/164761		2011-01-14 12:19:09 UTC	2011-01-14 12:20:54 UTC
2.3.0.0/163215		2011-01-13 14:16:37 UTC	2011-02-10 10:15:16 UTC
2.3.4.0/246503		2011-01-20 19:52:09 UTC	2011-01-20 20:05:36 UTC

Trafic reçu au déploiement de 2.2.0.0/16 et 2.3.0.0/16



Après l'annonce du 2.0.0.0/12

Depuis la mise en service :

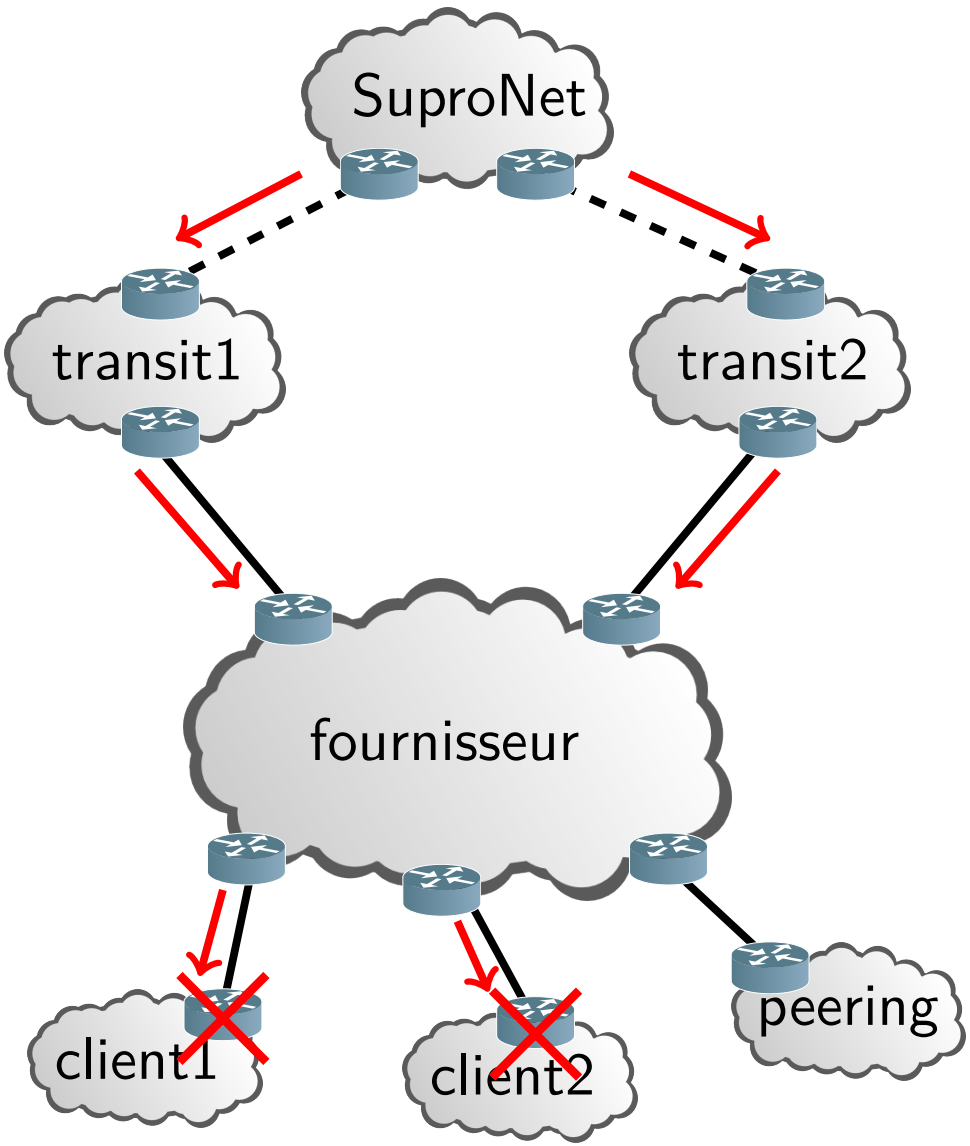
- occurrence des usurpations : environ une annonce par mois ;
- motivations : a priori usurpations involontaires ;
- parfois, il y a un lien entre usurpation et burst : nous avons pu corréler des annonces de /32 très localisées avec des burst.

```
Your prefix:         2.2.0.0/16:
Update time:         2011-08-29 15:10 (UTC)
Detected by #peers: 2
Detected prefix:     2.2.2.2/32
Announced by:        AS12684 (ASTRA-NET SES ASTRA (Astra-Net))
Upstream AS:         AS42652 (DELUNET inexio Informationstechnologie und
                     Telekommunikation KGaA)
ASpath:              51810 25472 1267 42652 12684
```

L'incident MikroTik

En attendant, les mises à jour des constructeurs, les filtres sur certains attributs (comme l'AS PATH) protègent le réseau.

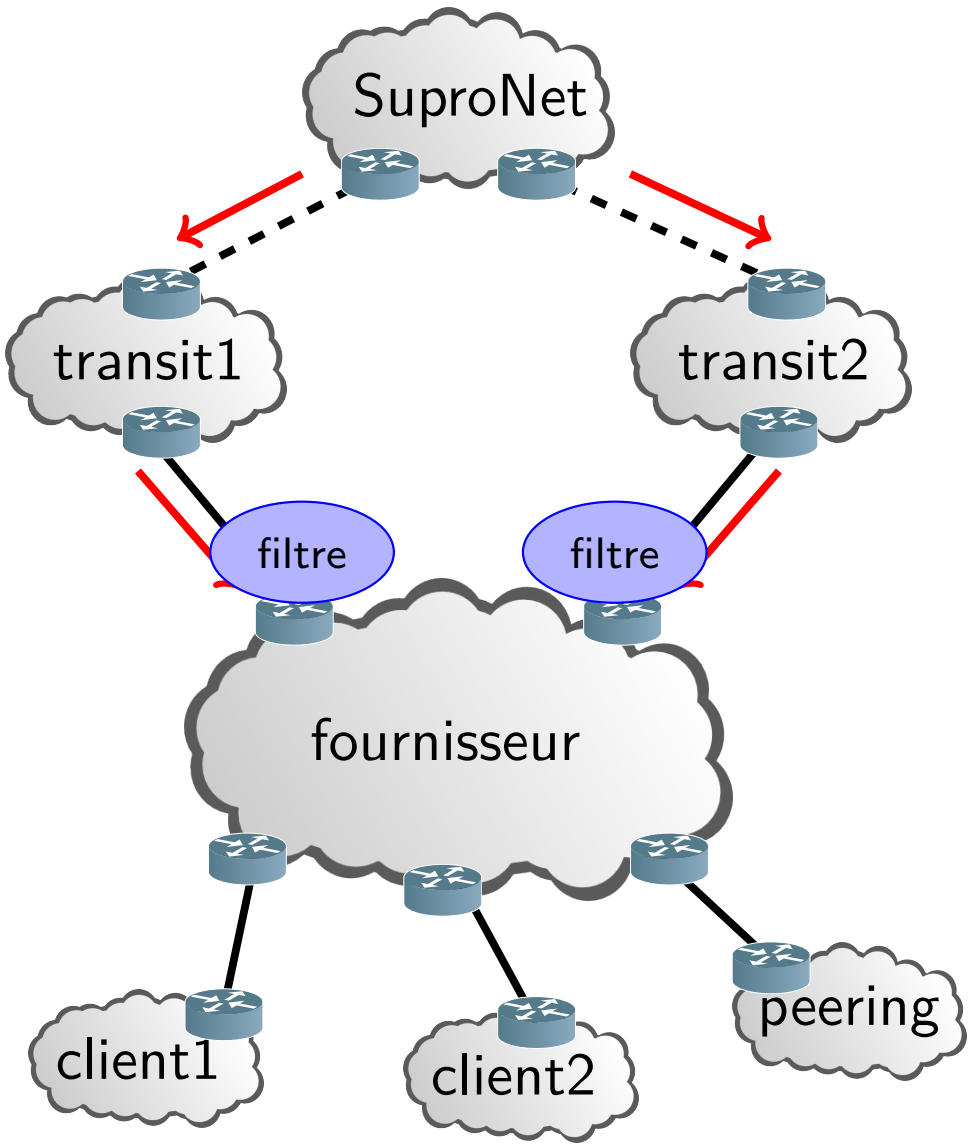
- les tests de validation des piles BGP mettent en évidence les différences de comportement sur réception d'UPDATEs malformés (hors crash de pile) ;
- le temps de convergence sur coupure de session BGP peut être long.



L'incident MikroTik

En attendant, les mises à jour des constructeurs, les filtres sur certains attributs (comme l'AS PATH) protègent le réseau.

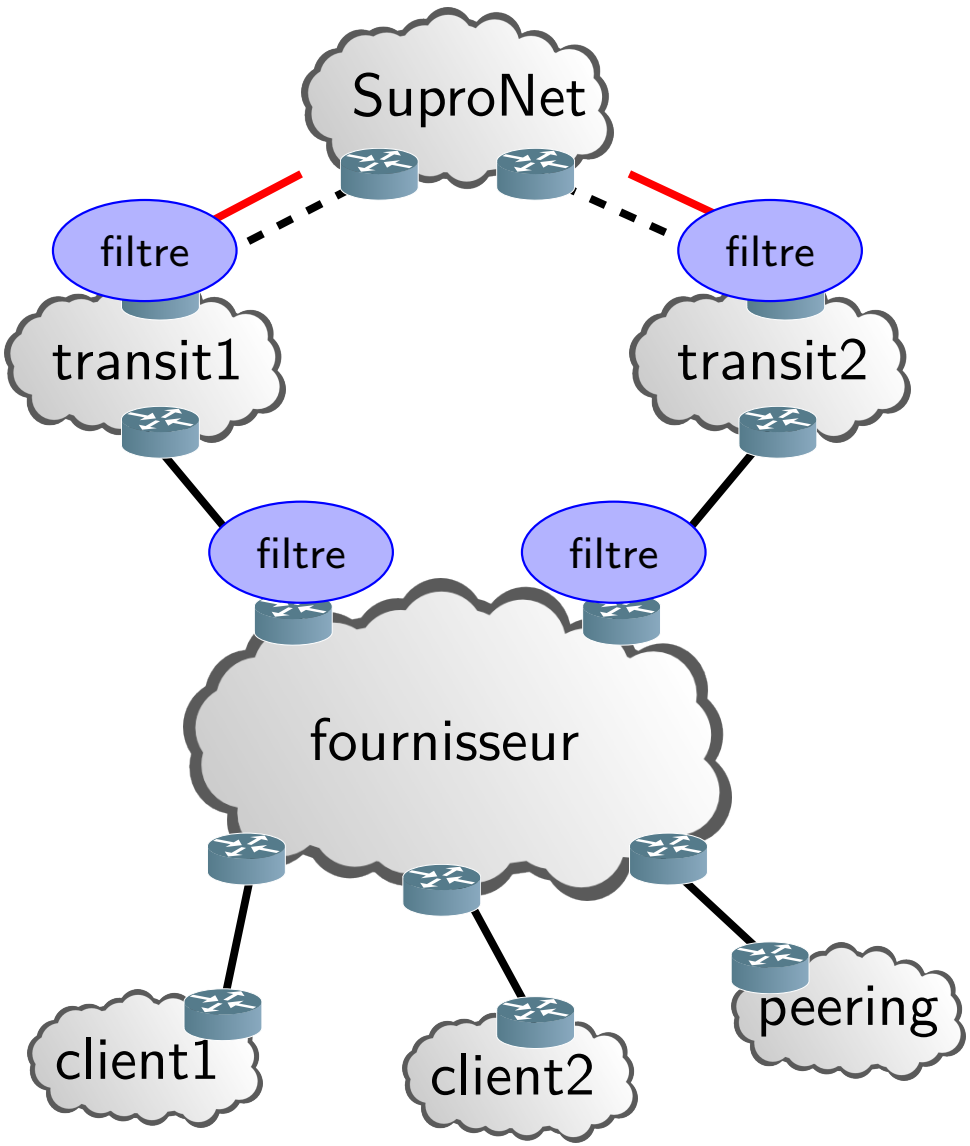
- les tests de validation des piles BGP mettent en évidence les différences de comportement sur réception d'UPDATEs malformés (hors crash de pile) ;
- le temps de convergence sur coupure de session BGP peut être long.



L'incident MikroTik

En attendant, les mises à jour des constructeurs, les filtres sur certains attributs (comme l'AS PATH) protègent le réseau.

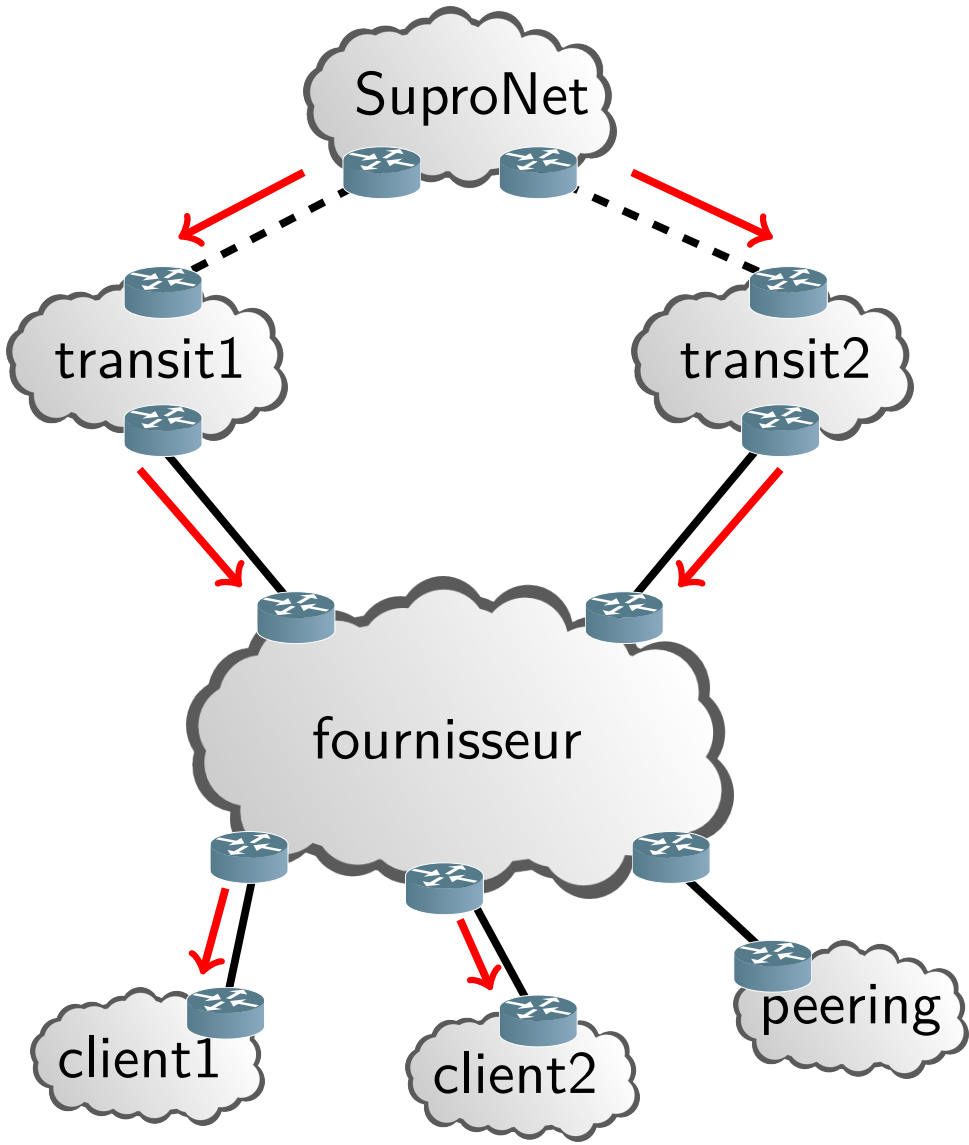
- les tests de validation des piles BGP mettent en évidence les différences de comportement sur réception d'UPDATEs malformés (hors crash de pile) ;
- le temps de convergence sur coupure de session BGP peut être long.



Protection contre les messages malformés

Un travail dans les instances de normalisation est en cours pour homogénéiser les comportements et surtout minimiser l'impact de ces messages :

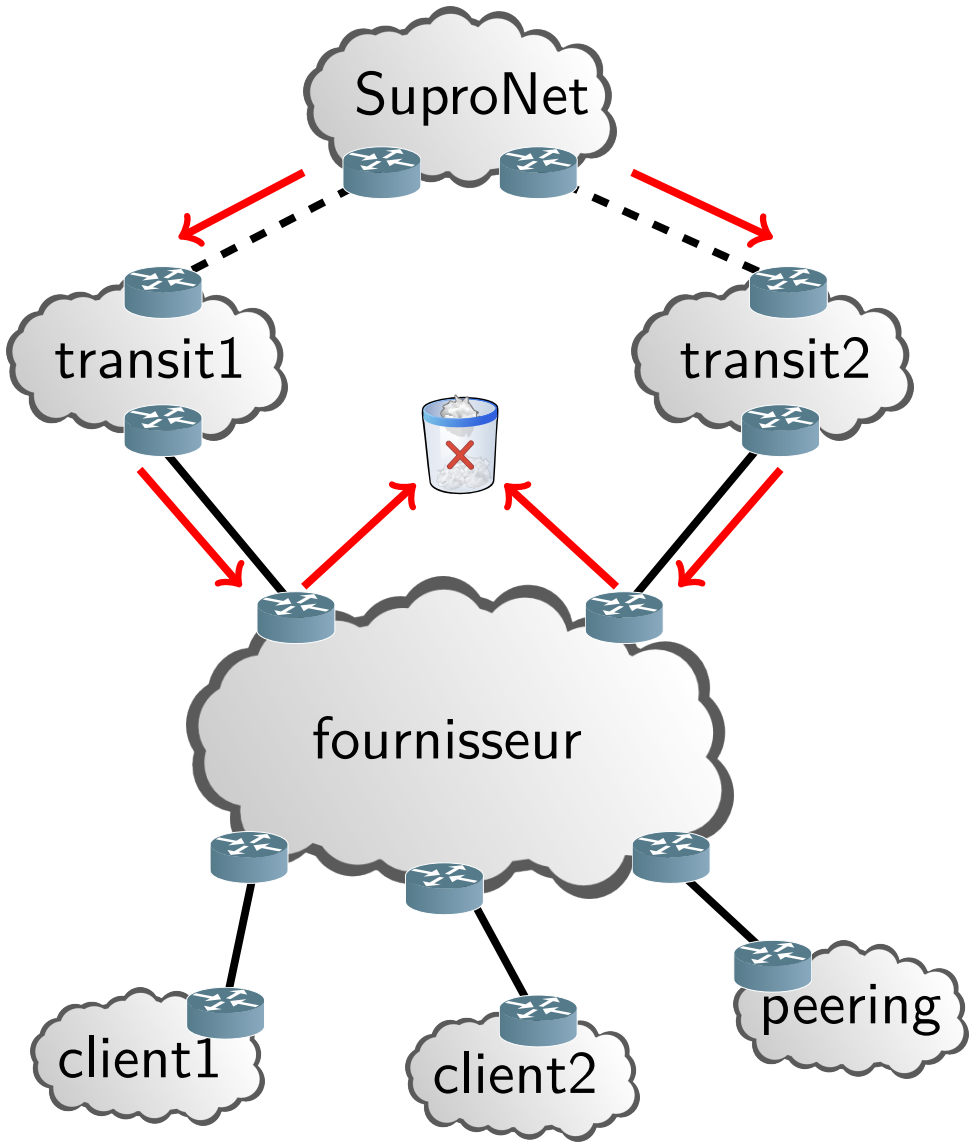
- première bêta en cours de test ;
- lorsque le message est mal interprété, il est rejeté, la session reste active, les préfixes identifiés sont supprimés des tables de routage BGP, l'erreur est journalisée (log + trap SNMP).



Protection contre les messages malformés

Un travail dans les instances de normalisation est en cours pour homogénéiser les comportements et surtout minimiser l'impact de ces messages :

- première bêta en cours de test ;
- lorsque le message est mal interprété, il est rejeté, la session reste active, les préfixes identifiés sont supprimés des tables de routage BGP, l'erreur est journalisée (log + trap SNMP).



Un peu d'architecture

- au sein du réseau cœur IP/MPLS d'un opérateur, BGP supporte de nombreux services : IPv4, IPv6, VPN ;
- le multiplexage des services dans une même session BGP pose problème en cas de mauvaise implémentation ;
  - exemple réel : l'IANA a attribué par erreur la même communauté BGP étendue 0x010a à VPLS et mVPN.

L'architecture est actuellement la seule solution pour protéger l'infrastructure et stabiliser les différents services :

- une seule famille d'adresses sur la même session BGP ;
- aucune mutualisation de services sur les réflecteurs de routes ;
- utiliser MPLS pour soulager les routeurs du cœur.

Conclusion

1. BGP est un protocole éprouvé, mais :
   - il repose sur la confiance entre opérateurs ;
   - les bonnes pratiques demeurent indispensables ;
   - les implémentations évoluent et engendrent de nouveaux bugs.
2. De nouvelles optimisations amélioreront la sécurité de l'infrastructure BGP :
   - spécification des nouveaux comportements attendus lors de messages malformés ;
   - signature cryptographique contre l'usurpation de préfixe :
     - RPKI, ROVER ;
     - BGPsec.
3. La plupart des incidents BGP ont été résolus grâce aux systèmes de surveillance, la coopération entre opérateurs, et la réactivité des équipementiers.