

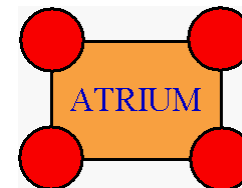
Redistribution Communities for Interdomain Traffic Engineering

B. Quoitin

bqu@infonet.fundp.ac.be

Infonet Group, University of Namur, Belgium

<http://www.infonet.fundp.ac.be>



This work was supported by the European Commission within the IST ATRIUM project

Table of Contents

- Introduction
- BGP Communities attribute
- Community-based Traffic Engineering
 - How does it work ?
 - How often is it used ?
 - Drawbacks
- Solution: Redistribution communities
- Conclusion

Communities

- BGP Communities attribute
 - RFC1997 in 1996
 - list of 32 bits values
 - Transitive
- Used to
 - mark routes which share a common property
 - signal routes which must undergo a given treatment
- Allow more scalable configurations

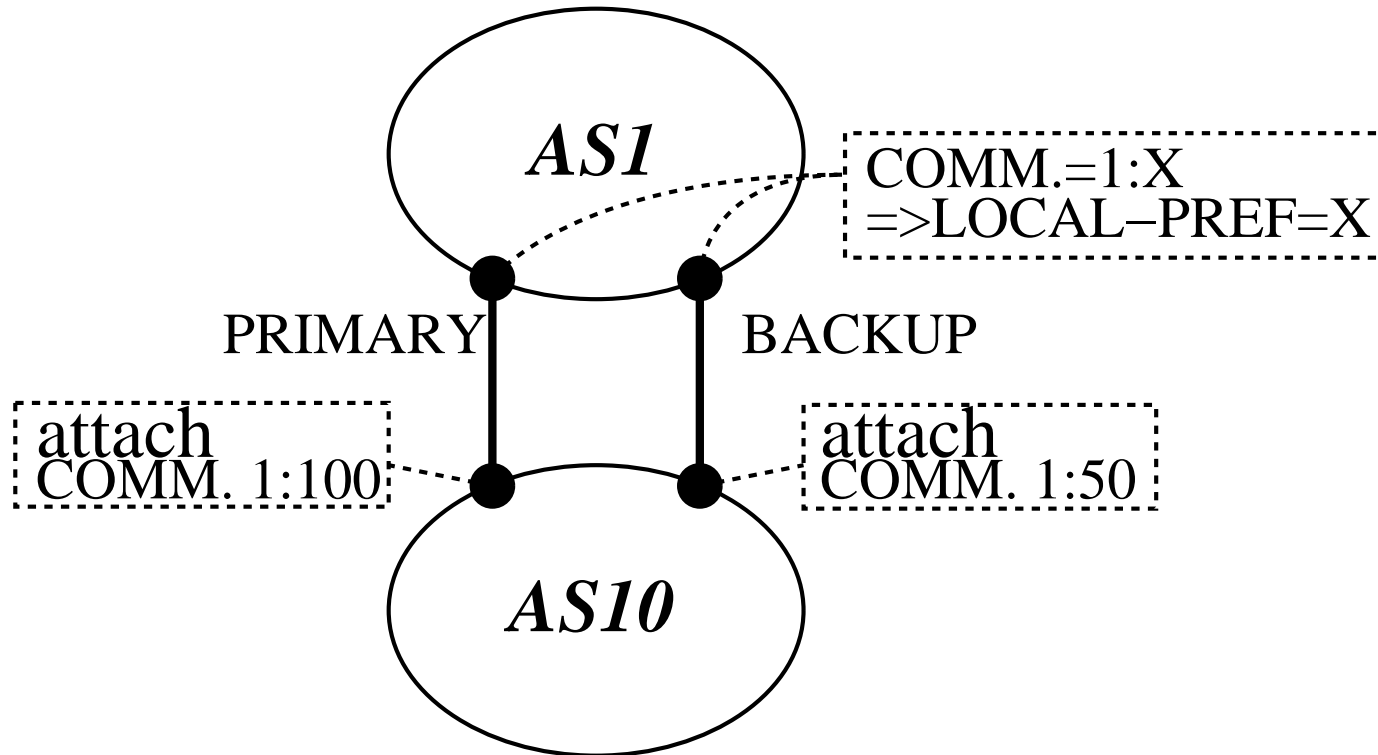
Communities

- Standardized values:
 - **NO_EXPORT, NO_ADVERTISE, NO_EXPORT_SUBCONFED**
 - **0x00000000-0x0000FFFF** and **0xFFFF0000-0xFFFFFFFF** are reserved
- Useable space:
 - Usually structured as

AS-number	0-65535
-----------	---------
 - Unclear for private AS numbers (64512-65534) !

RFC1998

- RFC1998 recommends the use of communities for multi-homed/linked networks
 - set the LOCAL-PREF according to community values



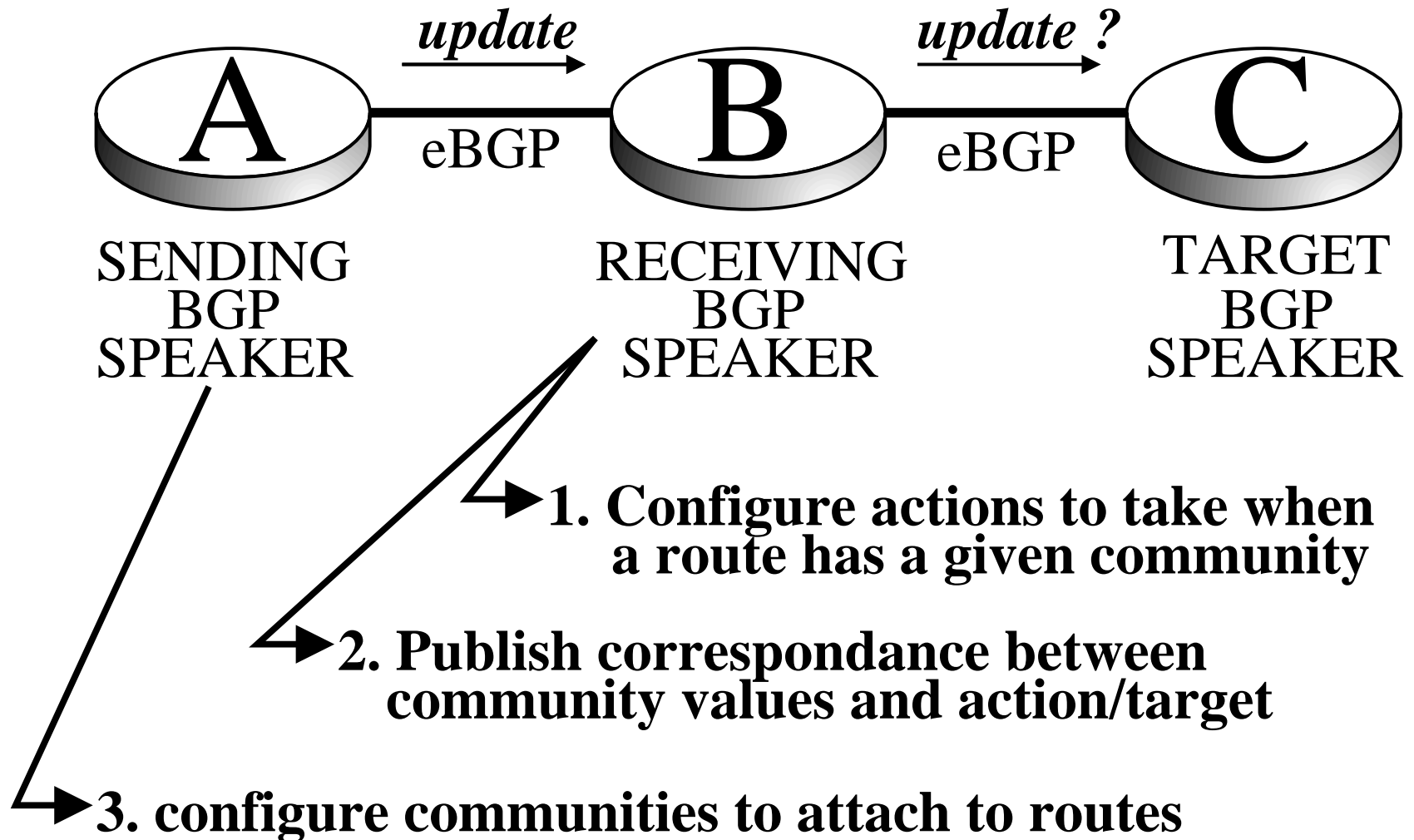
Community-based TE

Communities now appear in the global Internet

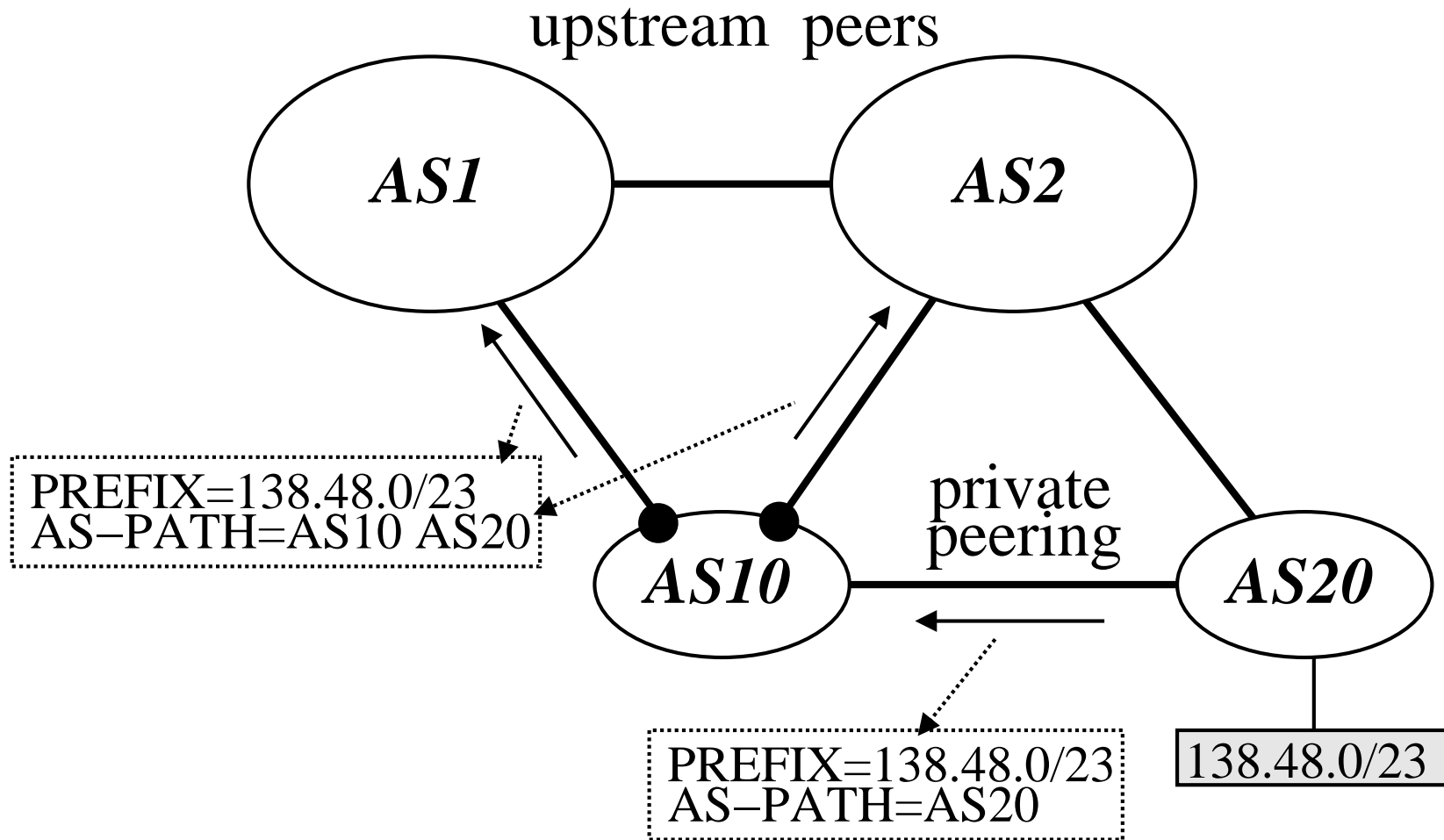
- Used to request a particular treatment of routes announced to a peer:
 - do not announce to specified peers;
 - prepend n times the `as-path` announced to specified peers;
 - set the `local-pref` of the route.

Community-based TE

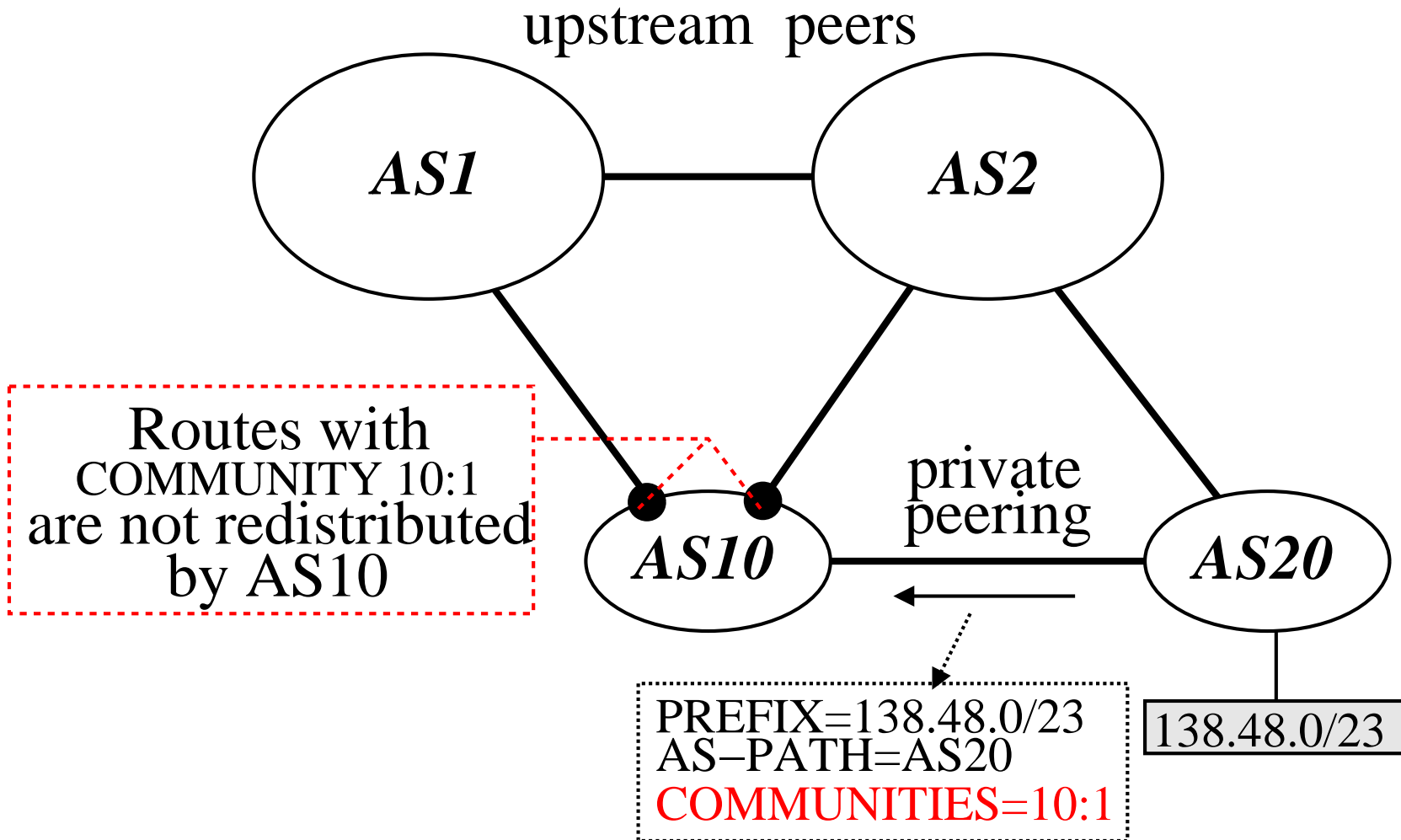
How does it work ?



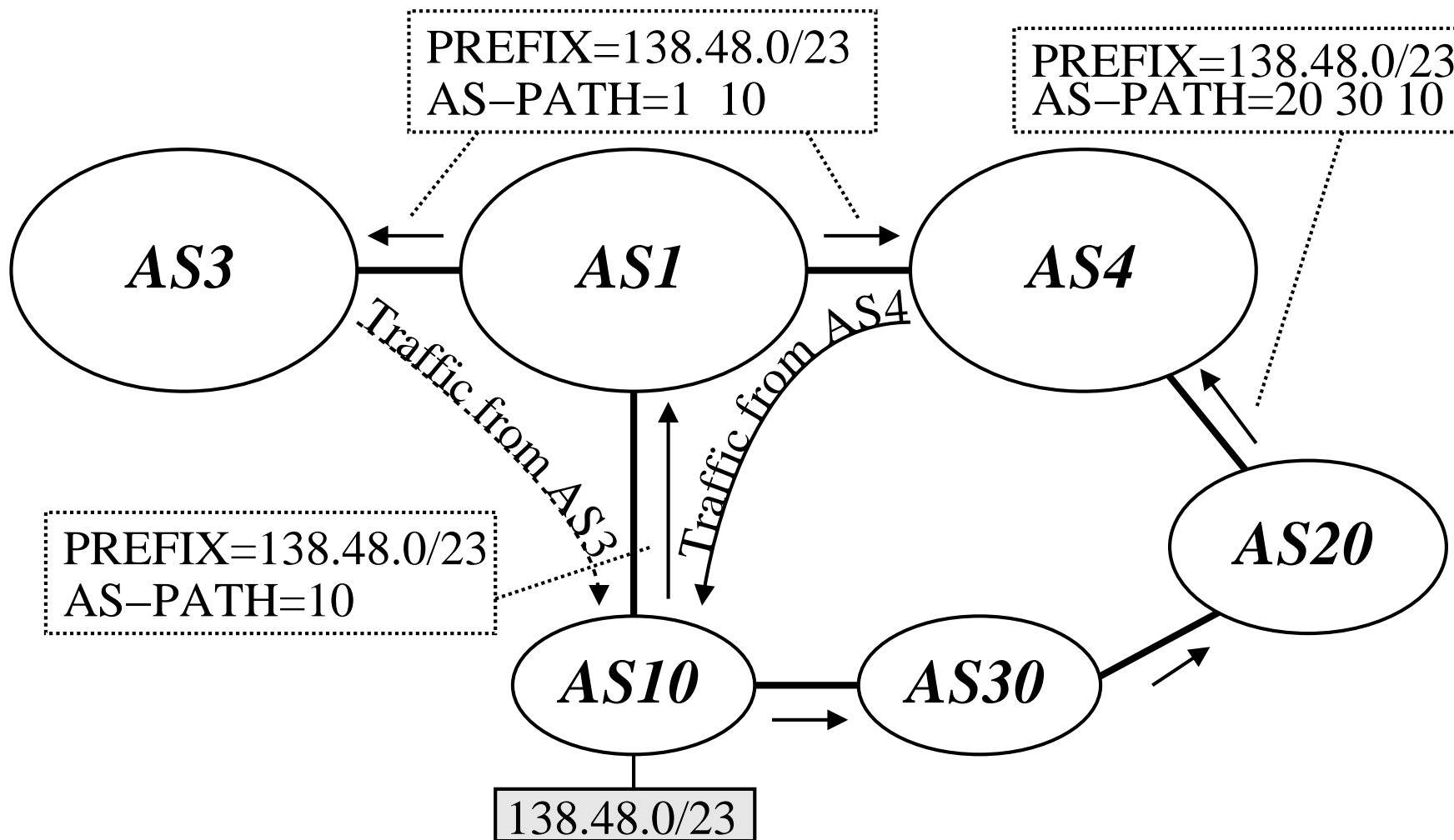
Selective announcements



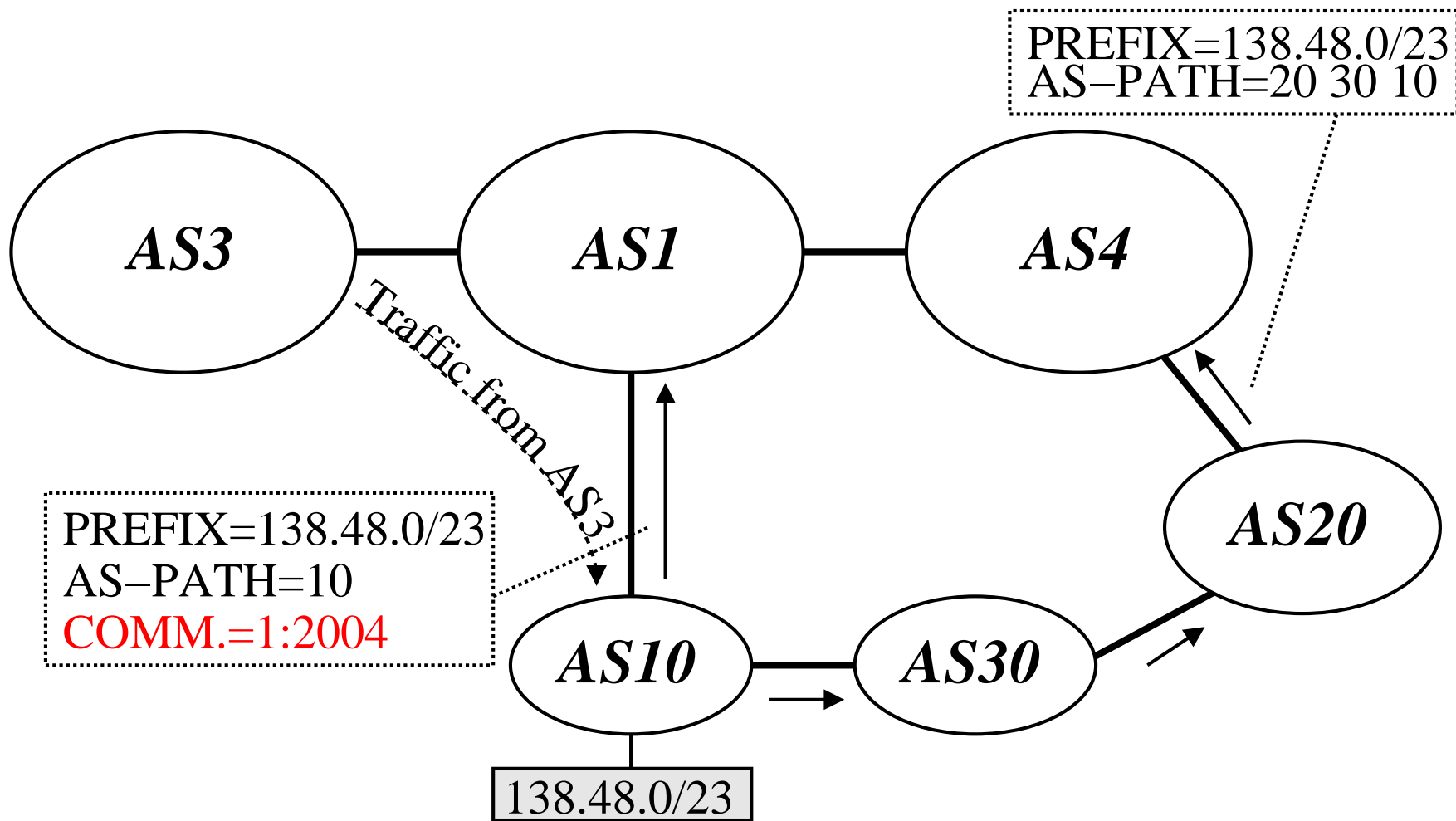
Selective announcements



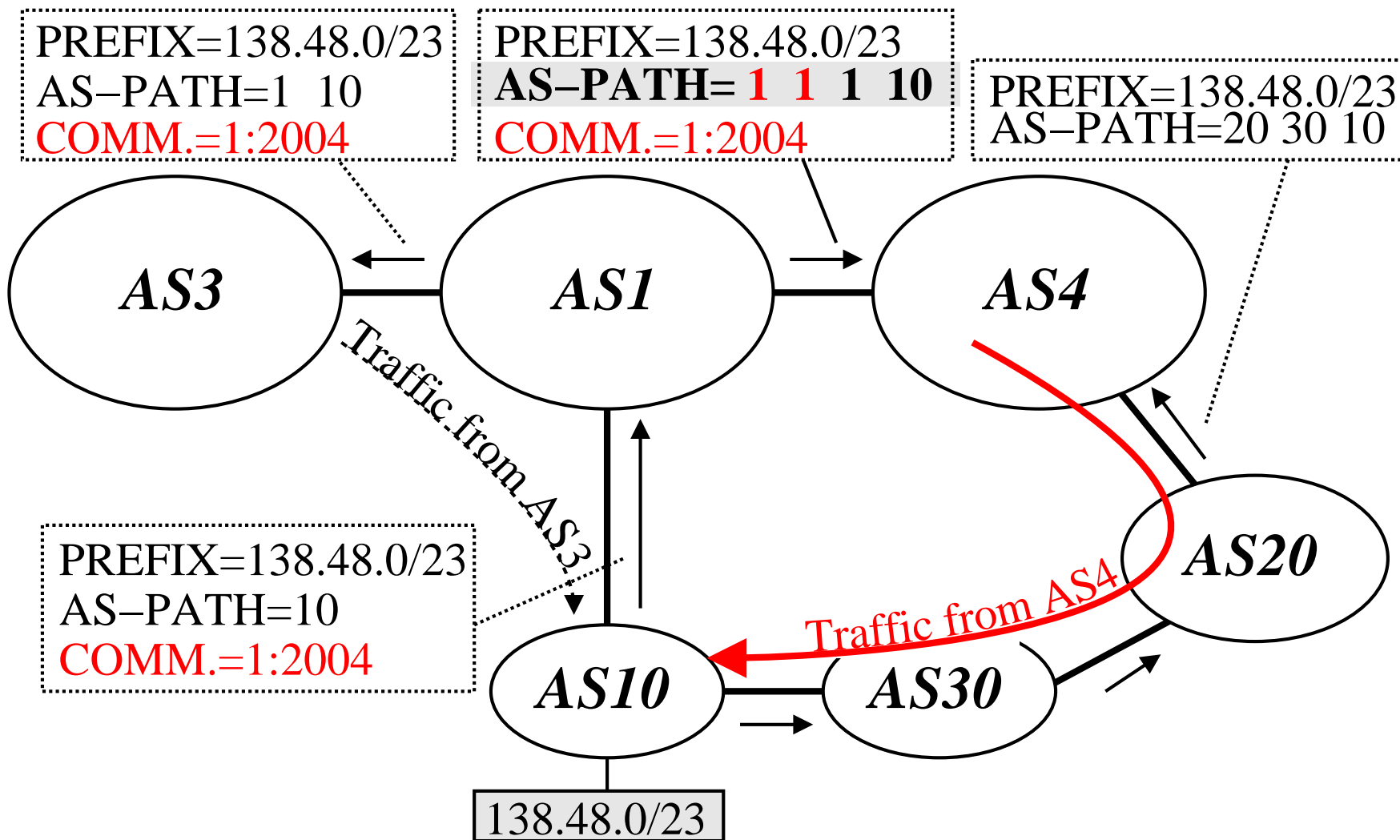
Prepending



Prepending



Prepending

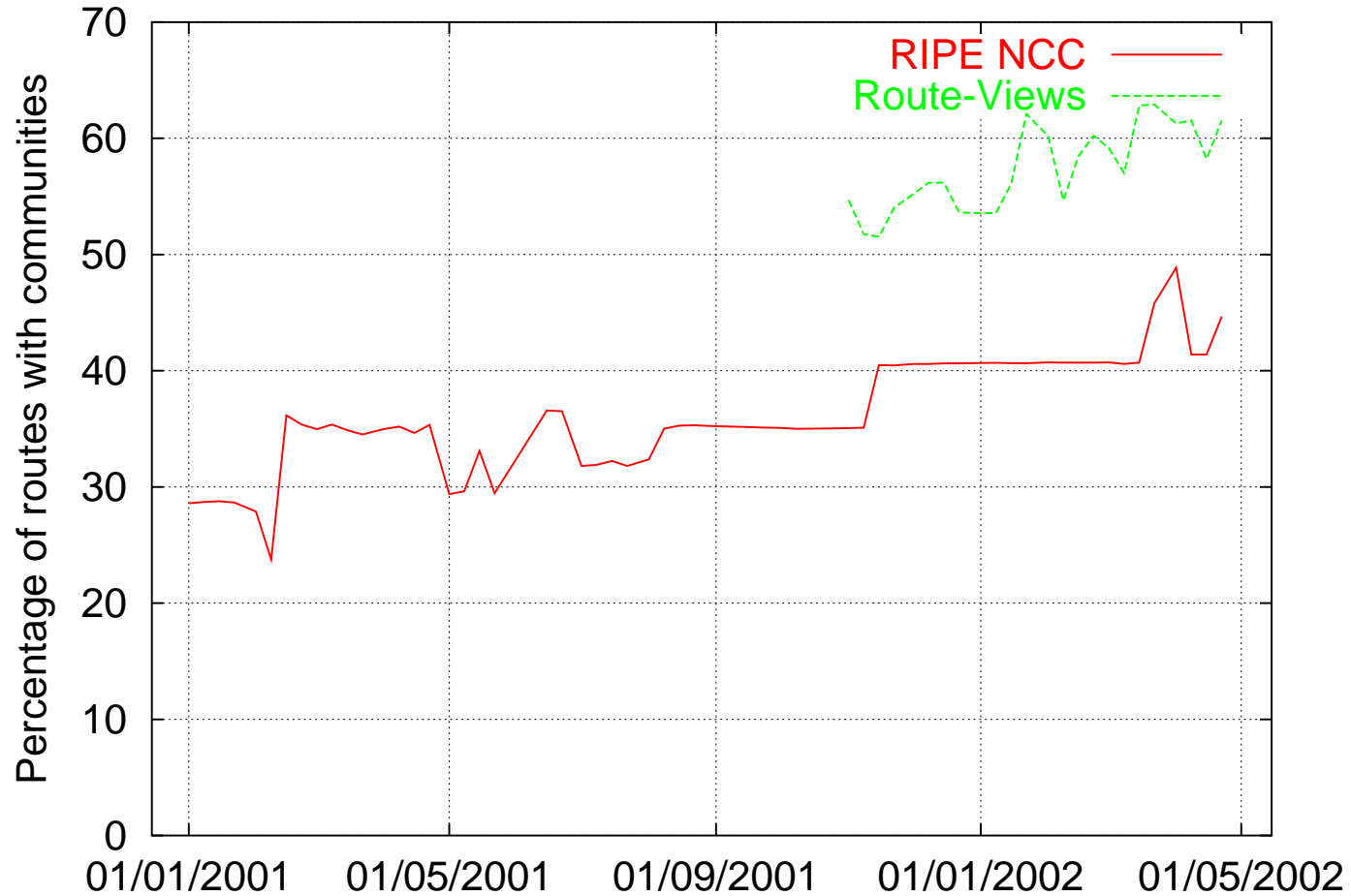


Analysis of BGP tables

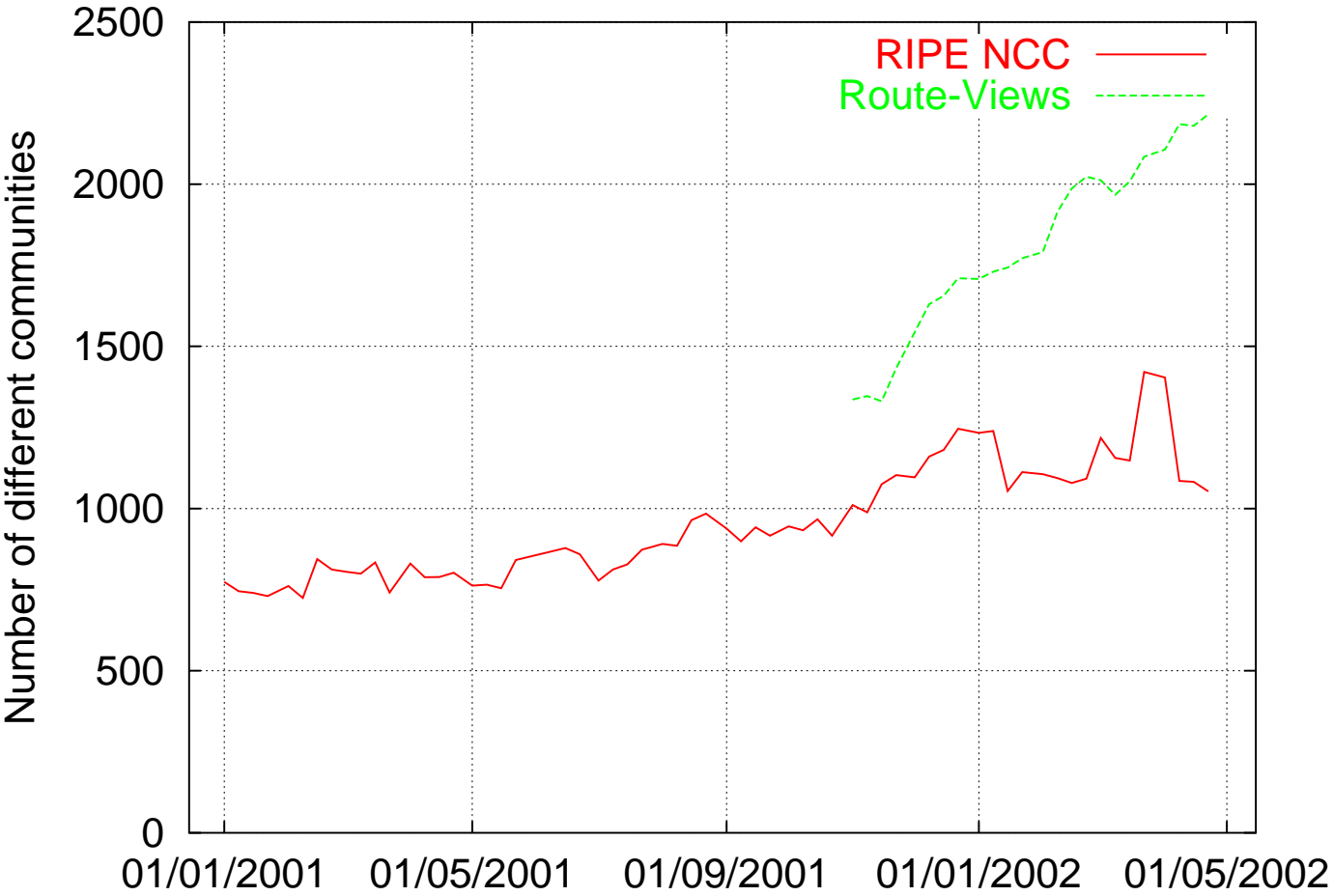
Are these communities often used ?

- Analysis of BGP tables
 - RIPE RIS [RIS02]
 - Route Views [Mey02]
 - during the period January 2001 - April 2002
- in conjunction with
 - whois databases
 - ISPs web sites

Percentage of routes with communities



Number of distinct community values



Community-based pollution

40 communities attached to a single route !

```
TABLE_DUMP | 1019465346 | B | 64.200.199.3 | 7911 | 57.249.147.0/24 | 7911 3561  
5511 3215 | INCOMPLETE | 64.200.199.3 | 0 | 0 | 3215:101 3215:204 3215:500  
3215:589 3215:903 3215:1001 3215:2001 3215:7503 3215:50000 3561:11840  
3561:30010 3561:30020 3561:30030 3561:30040 3561:30050 3561:30060  
3561:30070 3561:30080 3561:30090 3561:30100 3561:30110 3561:30120  
3561:30130 3561:30140 3561:30150 3561:30160 3561:30170 3561:30180  
3561:30190 3561:30200 3561:30410 3561:30420 3561:30430 3561:30440  
3561:30450 3561:30460 5511:500 5511:502 5511:999 7911:999 | NAG | |
```

AS3215 (France Telecom)

AS3561 (Cable & Wireless)

AS5511 (OpenTransit)

Is it useful to keep all these communities in routing tables ?

Published communities

How do ASes publish/define their communities ?

- Information from
 - whois database (remarks in RPSL)
 - ISPs web sites
 - defined in peering contract ?
- short summary (51 ASes observed)

target \ action	do-not-announce	prepending
specified AS	63 %	53 %
specified IX	49 %	37 %
specific peers	38 %	29 %

Selective announcements

Example of communities

Value	Meaning
1755:1000	Do not announce to US upstreams/peers
1755:1101	Do not announce to Sprintlink(US)/AS1239
1755:1102	Do not announce to UUNET(US)/AS701
1755:1103	Do not announce to Abovenet(US)/AS6461
...	
1755:2000	No announcement to european peers
...	

Community values published by Ebone (AS1755).

Structured values

Smart design

3561:30 <i>PPN</i>	<i>PP</i> is the peer code examples: 00=All Peers, 01=Genuity, 02=Sprint, 03=PSINet, 04=Qwest, ... <i>N</i> = 0, do not export to anyone = 1, prepend once to <i>PP</i> = 2, prepend twice to <i>PP</i> = 3, prepend three times to <i>PP</i>
---------------------------	--

Community values published by Cable & Wireless (AS3561).

A possible solution ?

Reuse unallocated community values

Value	Meaning
65000:X	do not announce on peerings to AS X
64970:X	do not announce on Asian/Pacific peerings to AS X
64980:X	do not announce on European peerings to AS X
64990:X	do not announce on North American peerings to AS X

Community values published by Level3 (AS9057).

Can every AS behave as Level3 ?

Community-based TE

Drawbacks

- Semantic of community values must be published,
- Limited size of communities
 - Unstructured values.
- Manual configuration
 - Error-prone configuration
 - Risk of errors
- Transitivity.

Redistribution Communities

How to improve the mechanism ?

Redistribution Communities are the solution !

- Internet draft

`draft-ietf-ptomaine-bgp-redistribution-00.txt`

- new type of extended-communities

- Non-transitive => **no pollution !**

- Larger size: **8 bytes**

Redistribution Communities

- Action

- Prepend
- Attach NO_EXPORT
- Do not announce

- Filter

- AS
- CIDR prefix

(1 byte)	(1 byte)	(6 bytes)
01TBDTBD	Action	Filter

Redistribution Communities

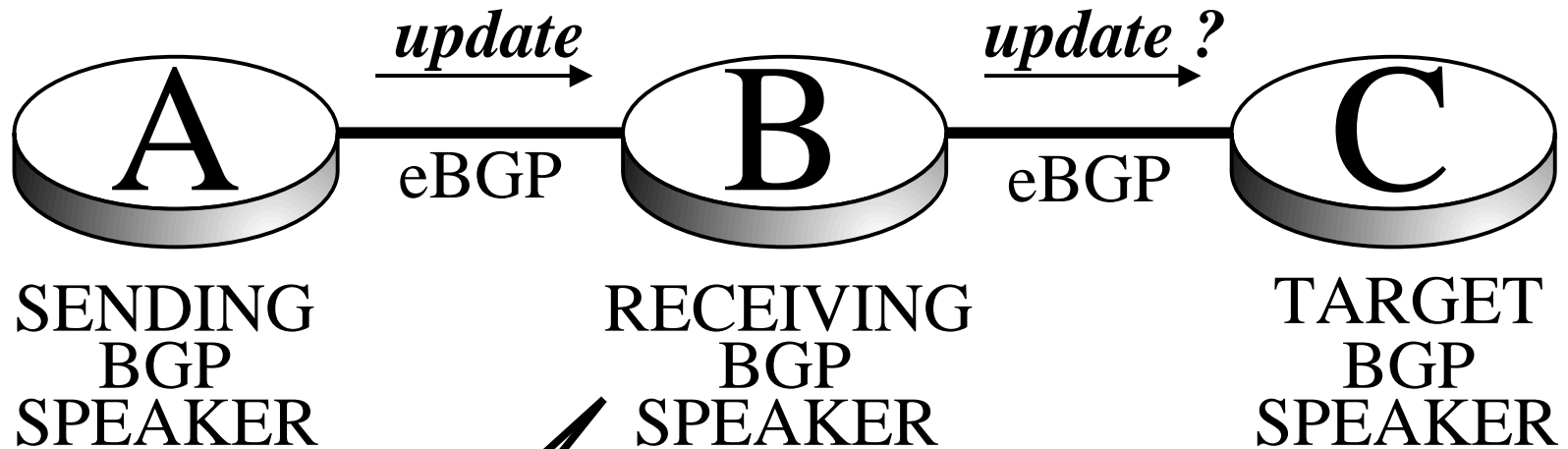
Is it implementable ? **YES**

- Zebra
 - Routing software
- Implementation
 1. Extension of the command-line interface (CLI)
 2. Modification of the redistribution code
 3. < 2000 lines of code

**Your favorite router vendor
could easily do it as well !**

Redistribution Communities

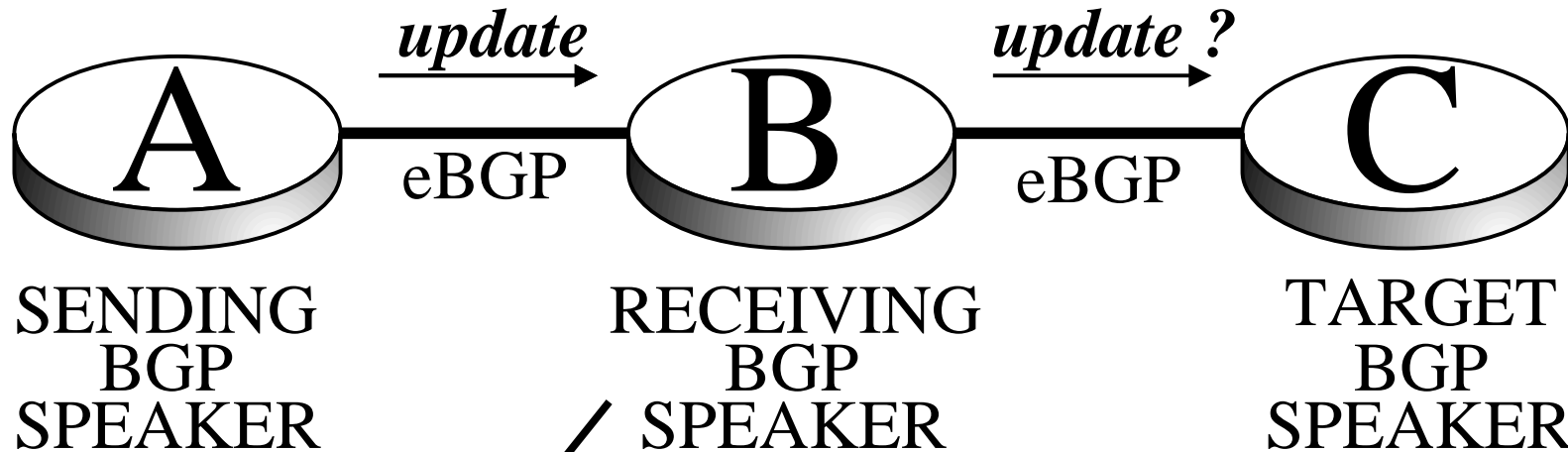
Configuration required



1. Configure actions to take when a route has a given community
2. Publish correspondance between community values and action/target
3. configure communities to attach to routes

Redistribution Communities

Configuration required



1. Enable the support of redistribution communities
2. Configure actions to apply to routes

Redistribution Communities

AS20 requests that AS10 do not re-announce its routes to AS1 and AS2.

```
router bgp 20
neighbor x.x.x.x remote-as 10
neighbor x.x.x.x route-map config-community out
neighbor x.x.x.x send-community
!
route-map config-community permit 10
  match ip address any
  set extcommunity red ignore:as(1) ignore:as(2)
```

Redistribution Communities

AS10 requests that AS1 prepend 2 times when re-announcing to AS4 only.

```
router bgp 10
neighbor x.x.x.x remote-as 1
neighbor x.x.x.x route-map config-community out
neighbor x.x.x.x send-community
!
route-map config-community permit 10
  match ip address any
  set extcommunity red prepend(2):as(4)
```

Conclusion

Advantages of the redistribution communities

- Standardized and structured => reduces risk of misconfiguration,
- Not transitive => reduces table growth and pollution,
- Simple to implement and easy to configure,

Could also be used to

- reduce the impact of denial of service attacks;
- with route optimization tools ...

Thank you for your attention

References

[BCH+02] O. Bonaventure et al. Controlling the redistribution of BGP routes. Internet draft, <http://www.ietf.org/internet-drafts/draft-ietf-ptomaine-bgp-redistribution-00.txt>.

[Mey02] Route-Views project. <http://archive.routeviews.org>.

[QB02] B. Quoitin and O. Bonaventure. A survey of the utilization of the BGP community attribute. Internet draft, <http://www.ietf.org/internet-drafts/draft-quoitin-bgp-comm-survey-00.pdf>.

[RIS02] RIPE RIS project. <http://data.ris.ripe.net>.

[UB02] A study of the macroscopic behavior of Internet traffic. S. Uhlig and O. Bonaventure. Under submission. Available from <http://www.infonet.fundp.ac.be/doc/tr>.