

# **Solution for Route Leaks Using BGP Communities**

**ietf-idr-route-leak-detection-mitigation-10**

**K. Sriram (Ed.), A. Azimov (Ed.), D. Montgomery, B. Dickson, K. Patel,  
A. Robachevsky, E. Bogomazov, and R. Bush**

**Authors' Team Discussion Slides  
October 2018**

Acknowledgements: The authors are grateful to many folks in various IETF WGs for commenting, critiquing, and offering very helpful suggestions (see acknowledgements section in the draft.)

# Design C: Solution for Route Leaks Using BGP Communities

Background: In the Montreal face-to-face meeting of authors, John and Sue advised the team to explore a BGP Community based solution. They envision the possibility of faster adoption if there are no changes required in commercially shipped BGP code.

- This set of slides are based in part on conversations many of us had in Montreal (face-to-face and emails) and my one-to-one discussions with Alex. Doug and I reviewed the content in the slides several times at NIST.
- Attempt is made to narrow the design down to one set of semantics and one way of encoding using Community
- Many scenarios are analyzed to examine if the semantics work
- Design choices for encoding using Large Community and Extended Community are presented
- Basic policy is described
- Sender and receive actions are specified
- Pseudo code is provided
- The idea is put down some details on paper and invite comments / discussion

# General Principles of Design C: Solution Using BGP Communities

- Considering **Community** based encoding of RLP info for **faster adoption**
- Wish to **limit the number of RLP** entries so that they can be accommodated in 1 or 2 Community attributes per update.
  - Reason: Avoid having a long string of Community attributes per BGP update because the more they are, the lesser the chance that they will all make it through. If some get dropped, then the rest become useless. Also, save memory, simplify processing, and improve robustness.
- Based on the analysis and knowledge we have so far about RLP/eOTC, independent of encoding (Attribute or Community), at the minimum the RLP info must include:
  - ASN of the RLP-aware AS that **most recently** asserted that it sent update to a customer or lateral peer; let us call this **DO = Down Only indication**
  - Leak warning: **L = Leak indication**
    - **L = ASN of the first RLP-aware AS in the path that is forwarding route from customer or lateral peer in spite of detecting a leak**
      - AS in question is avoiding unreachability (absence of alternative route)

Note: RLP = Route Leak Protection; DO alone or DO and L together constitute RLP

## Limitations:

In the following circumstances, a leaked route may not be detected:

- A leak between two or more **consecutive ASes** that are **not participating**
- **AS dropping** a transitive BGP **Community used for RLP**
- **Implementation errors** (ideally there should be none)

## Design assumptions:

- In the absence of an alternative route, an AS **may forward a route** that is detected to be a leak.

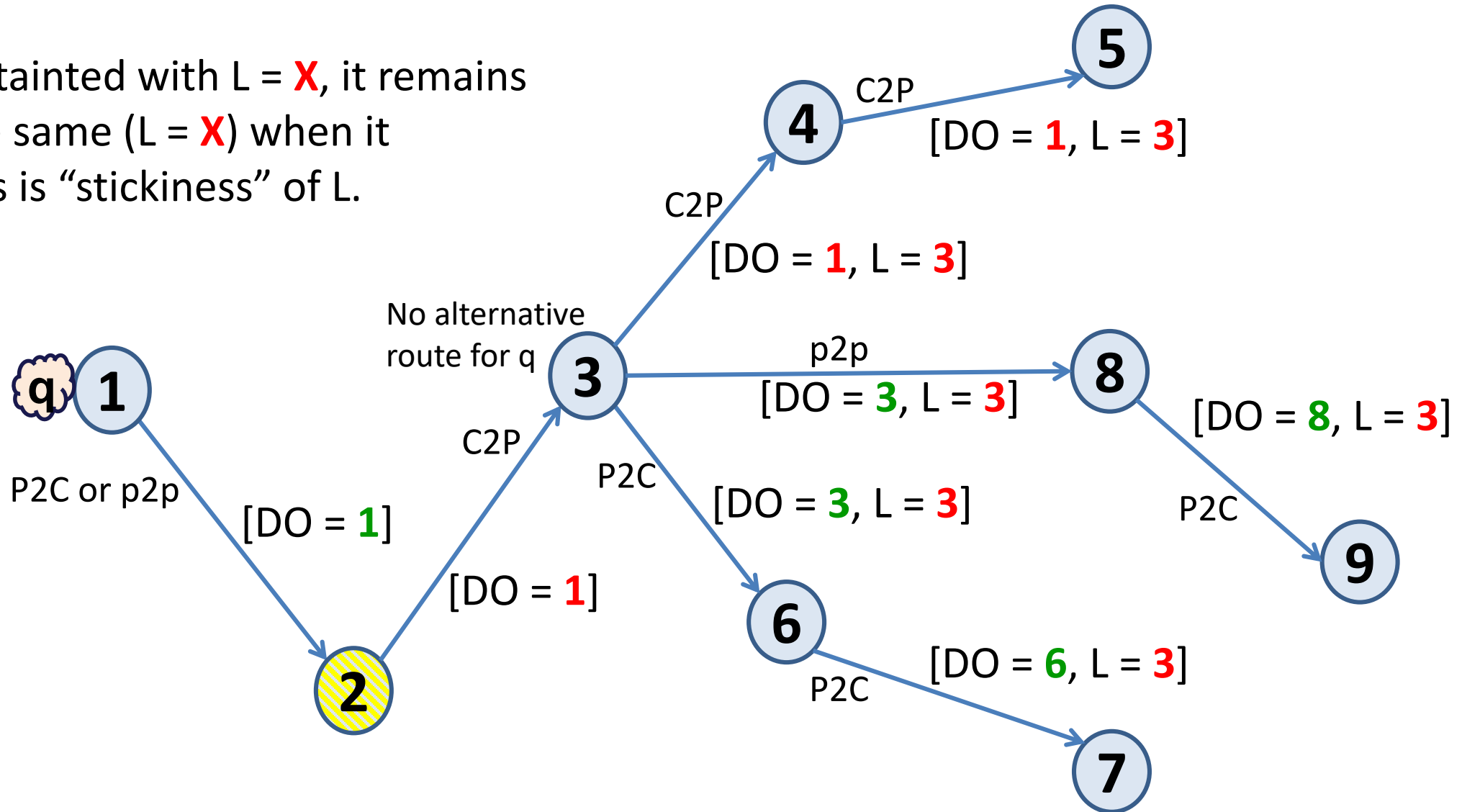
# Illustration of Down Only (DO) and Leak (L) indications – 1 of 2

Once a route is tainted with L = **X**, it remains tainted with the same (L = **X**) when it propagates. This is “stickiness” of L.

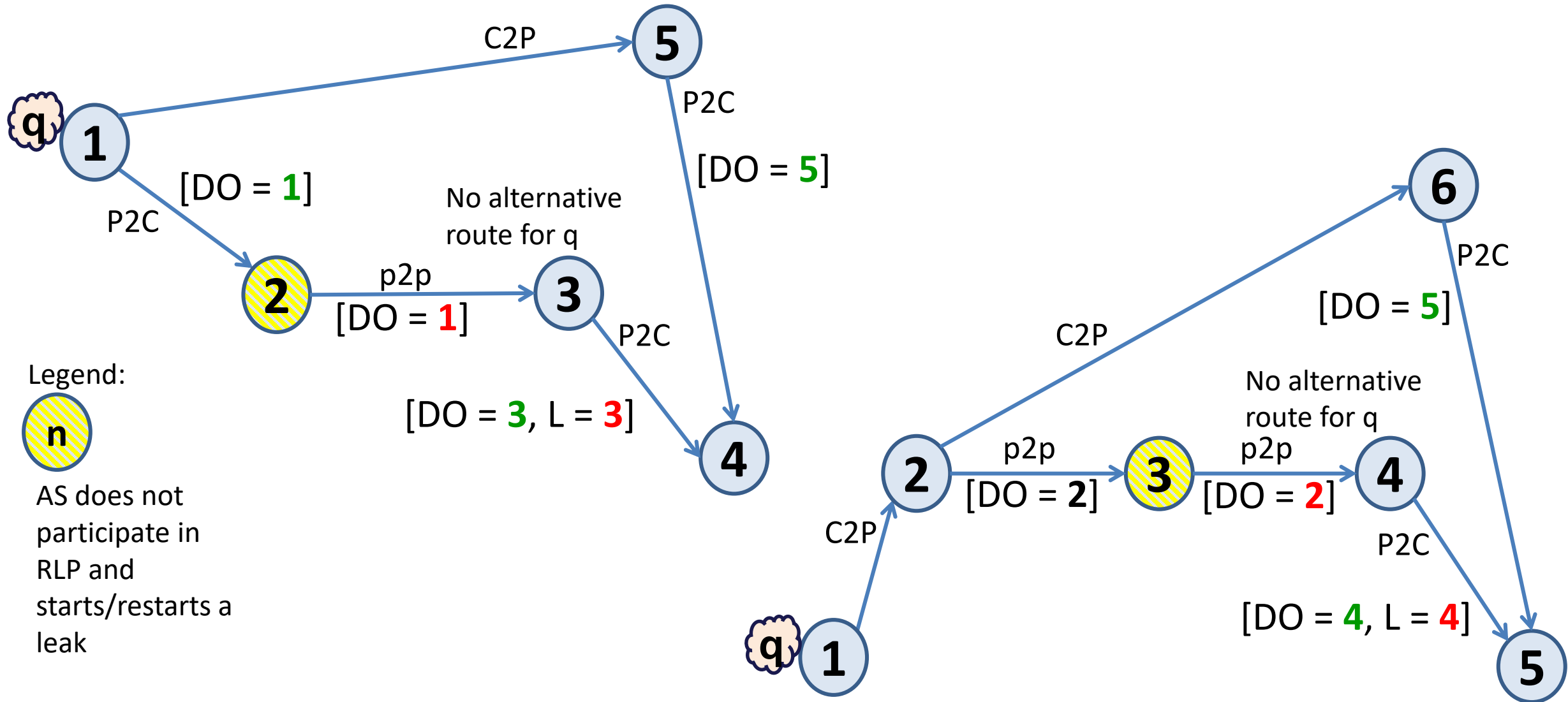
Legend:



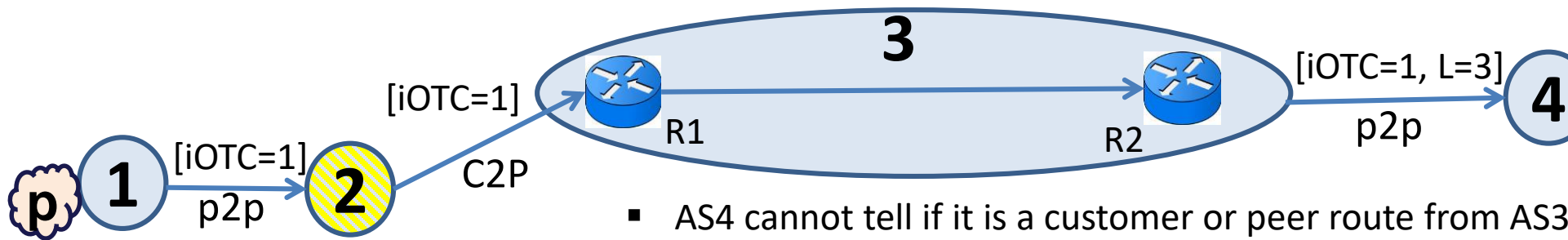
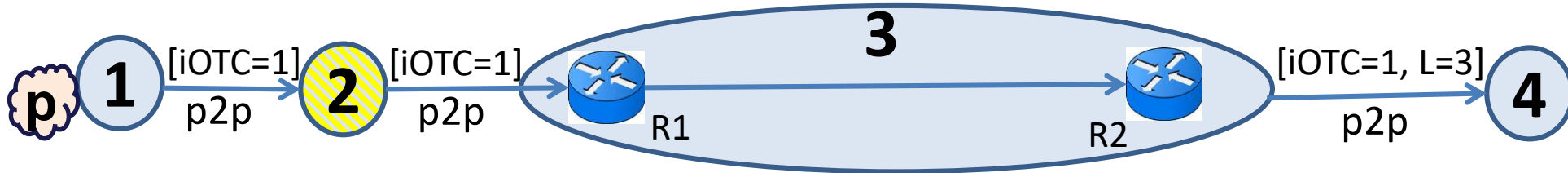
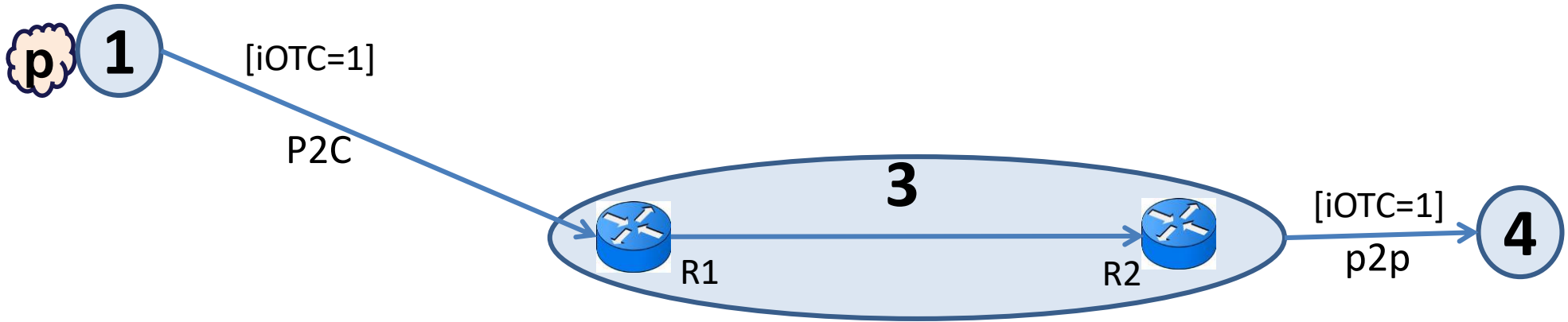
AS does not participate in RLP and starts/restarts a leak




# Illustration of Down Only (DO) and Leak (L) indications – 2 of 2



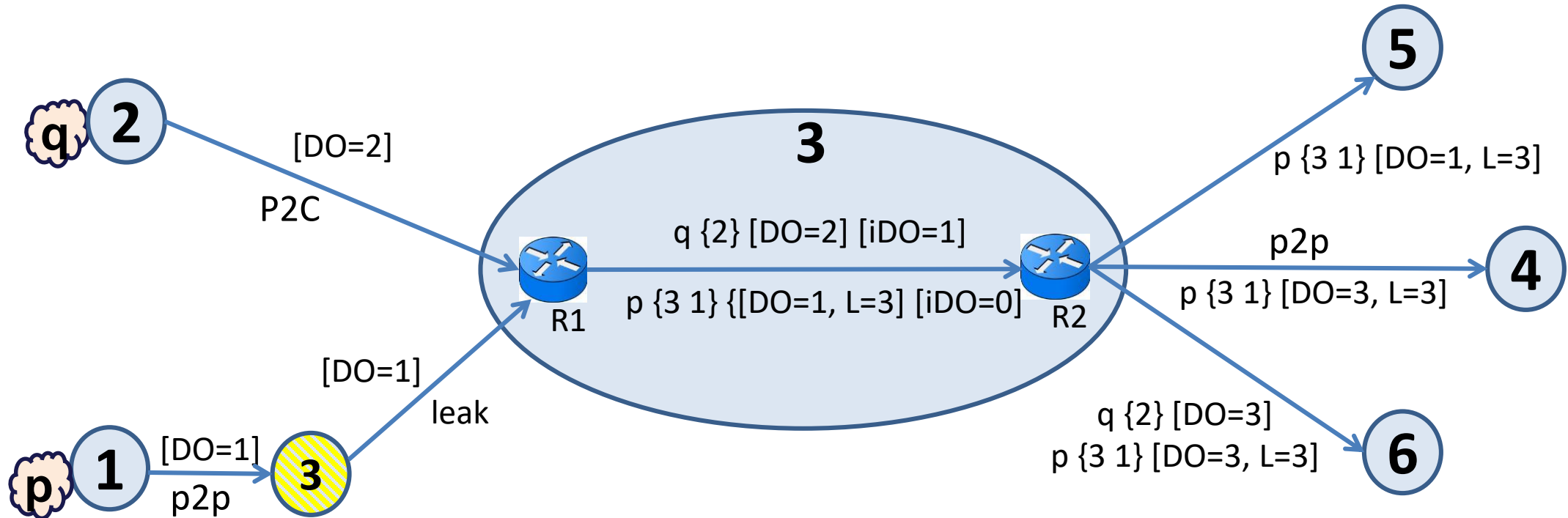
# This is not part of the design; this is just for illustration of a point about the original iOTC



Legend:  
 n  
 AS does not participate in RLP and starts/restarts a leak

- AS4 cannot tell if it is a customer or peer route from AS3
- Hence, it is mandatory for iOTC/RLP-aware AS (AS3 here) to implement both inter-AS and intra-AS solutions. Then, AS3 will simply never forward any p2p or P2C routes (received at R1) to AS4.

# Design: An RLP-aware AS must perform both Inter- and Intra-AS RLP



Legend:

AS does not participate in RLP and starts/restarts a leak

iDO = internal (local) Down Only

iDO=0 means intra-AS (local) DO does not apply


iDO=1 means intra-AS (local) DO applies

- R2 does not send non-customer routes to AS4 and AS5.



# Choices regarding Leak (L) indication

- DO must reflect the most recent AS in the path that sets DO – this is understood to be better based on previous analysis.

	<b>Down Only (DO)</b>	<b>Leak (L)</b>	<b>Choice</b>
Choice 1	ASN value updated to show the most recent AS in the path that sets DO.	ASN of the first AS that set L (sticky)	
Choice 2	- same as above -	Replaceable	Benefit?

With Choice 1, there is the benefit that L provides information about how far back in the path the initial leak occurred. Thus, L complements DO. Also, Choice 1 has less processing cost.

# Detection Rules:

- Semantics: Route is a leak = RLP is violated
- A received route violates RLP
  - if L is present in the received route\*
  - else (L is absent), the route is received from a customer and DO is present
  - else (L is absent), the route is received from a lateral peer and DO is present that is not the lateral peer's ASN
- Note: Here by "L is present" we mean that its value is not the default value (all zeros) but is a proper ASN. Effectively "L is absent" if its value is the default value.
- Note: In a correct implementation, L cannot be present without a DO.

## Minimum Default Policy:

- Whenever there is choice between a customer route and a provider route, and both are detected to be in RLP violation, then lower the LocalPref to X (TBD) for each of them. Then shortest path criterion would typically make the customer route preferred\*.

\* This mitigates persistent oscillation possibility

- Caveat 1: This has an unfortunate downside that in some cases this may result in choosing route from provider over customer even when the provider route is a detour of the customer route. This may be due to prepends by the customer (customer P0 in Scenario 8, slide 15). (Note: Applying the Route Leak Theorem can help avoid this. But we let go of that for simplicity of implementation.)
- Caveat 2: Also, in some cases this would cause customer route to be preferred over the provider route even when evidently the customer route has two valley-free violations while the provider route has only one such violation. Both routes have L (leak indication) in them. See Scenario 3, slide 11.
- We can possibly live with these caveats although we can avoid them if the Route Leak Detection Theorem (Slide 32) is put to use.

# Generalized Minimum Default Policy

- Whenever there is choice between multiple routes (customer/peer/provider), and each is detected to be in RLP violation, then lower the LocalPref to X (TBD) for each of them. Then apply shortest path criterion\*.

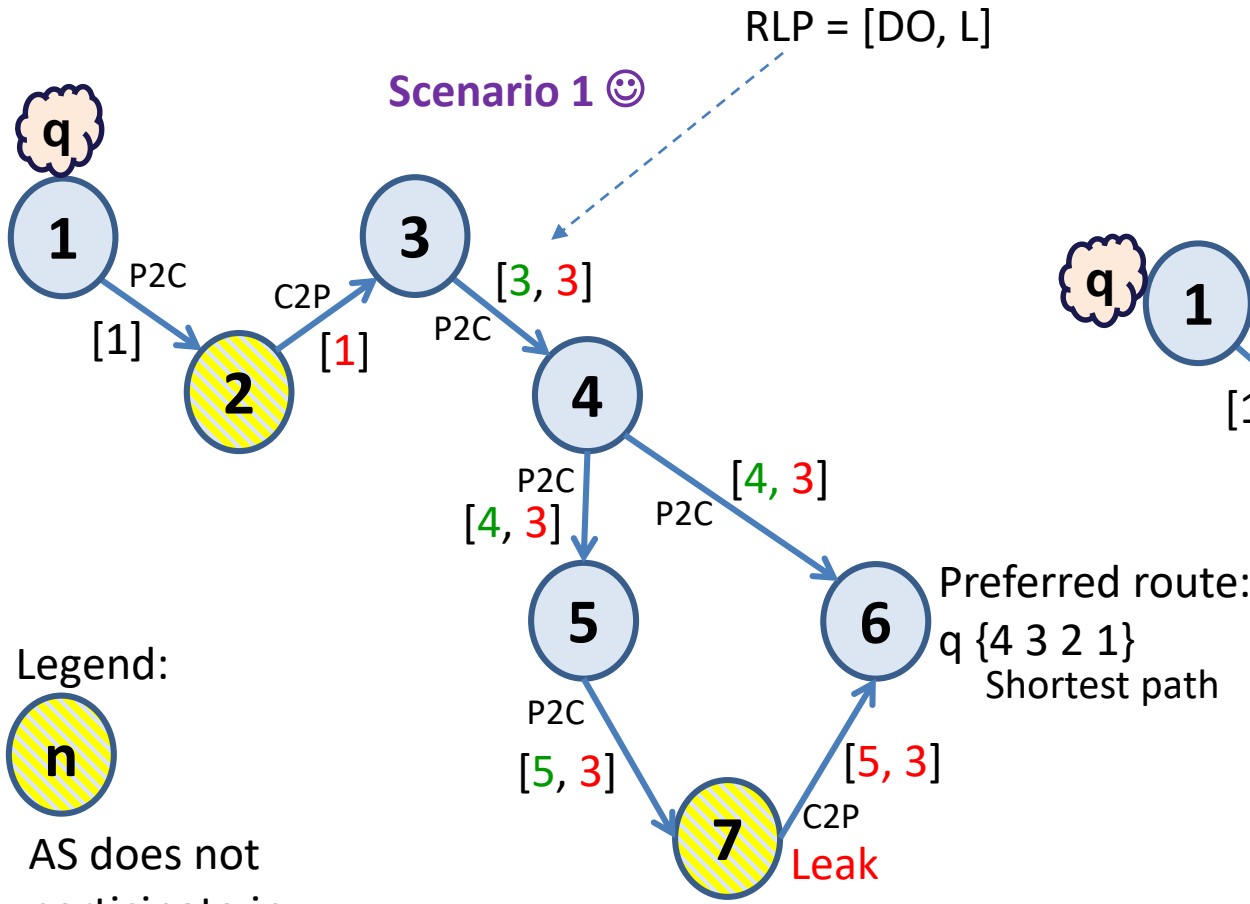
- \* Some network operators may find this inadequate (see the analyzed scenarios)
- \* But they can locally modify their policy while respecting the basic principle

**Scenario analyses:**

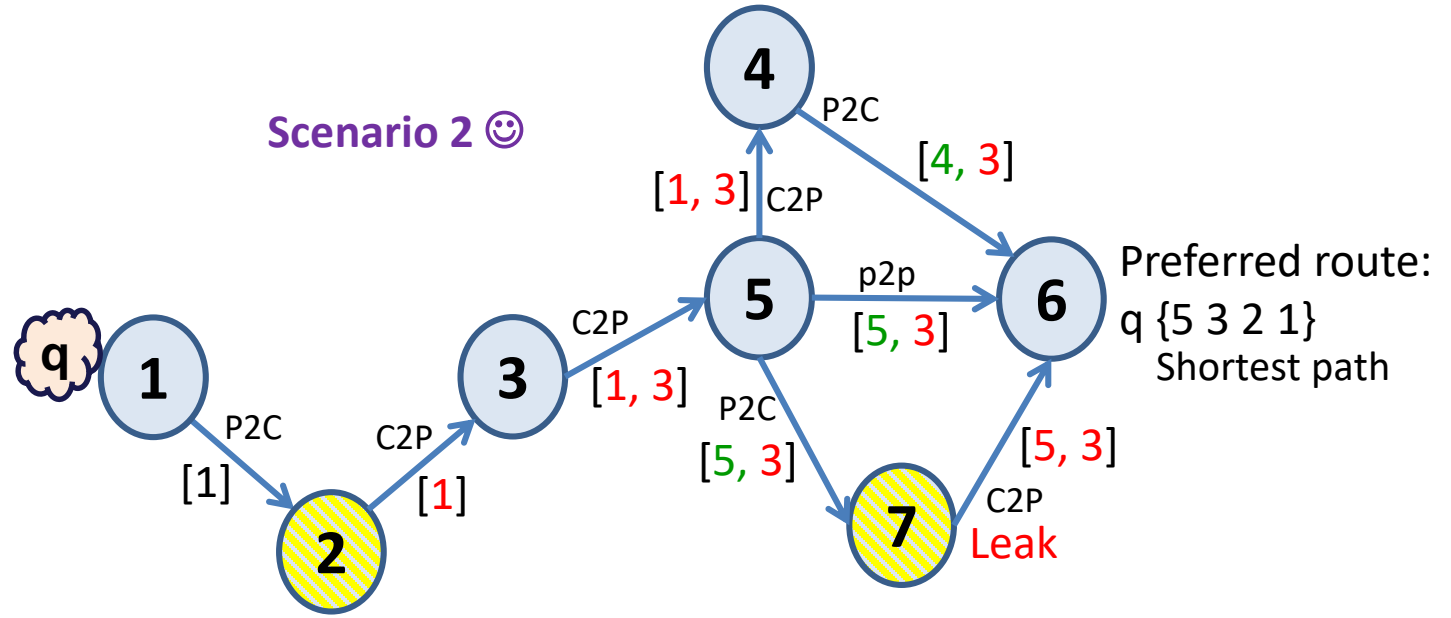
**Does this scheme with RLP = [DO, L] along with the policy work?**

# Scenarios:

Scenario 1 😊



Scenario 2 😊



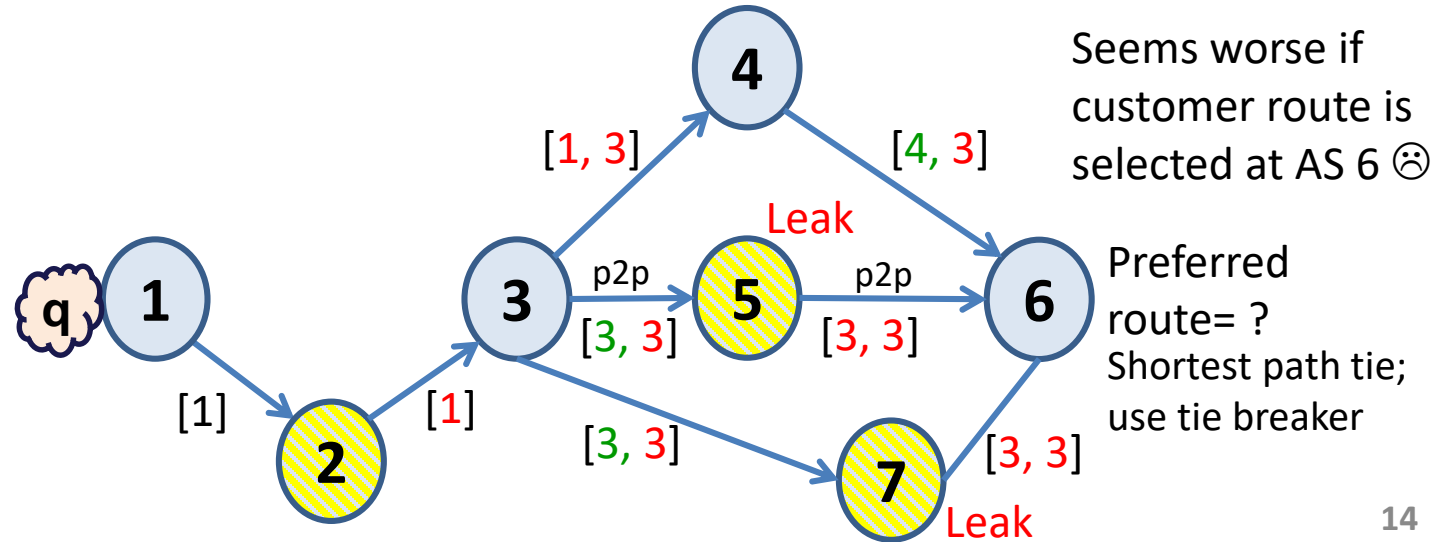
## Legend:



AS does not participate in RLP and starts/restarts a leak

**Green** – not violation  
**Red** – violation

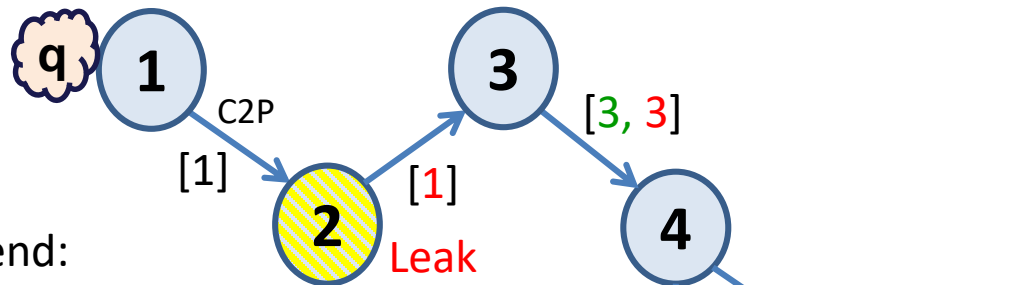
Scenario 3 😞



# More Scenarios:

Green – not violation  
 Red – violation

Scenario 4 ☺



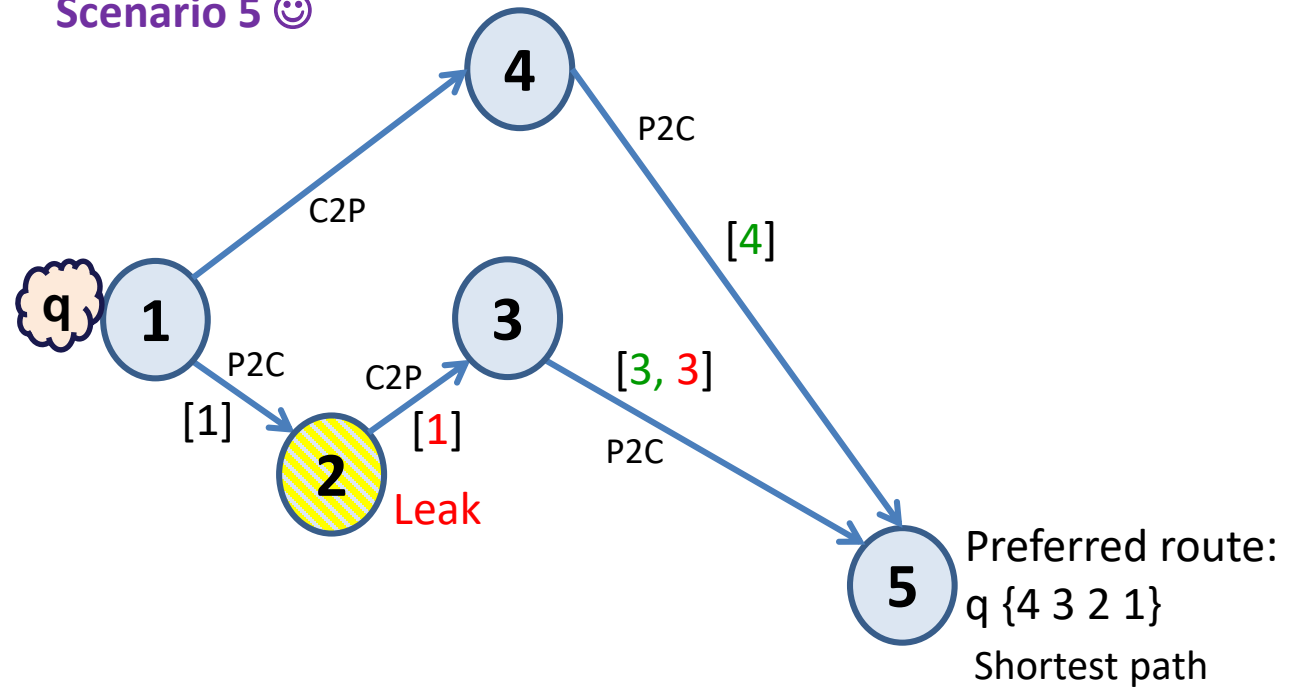
Legend:



AS does not participate in RLP and starts/restarts a leak

RLP = [DO, L]

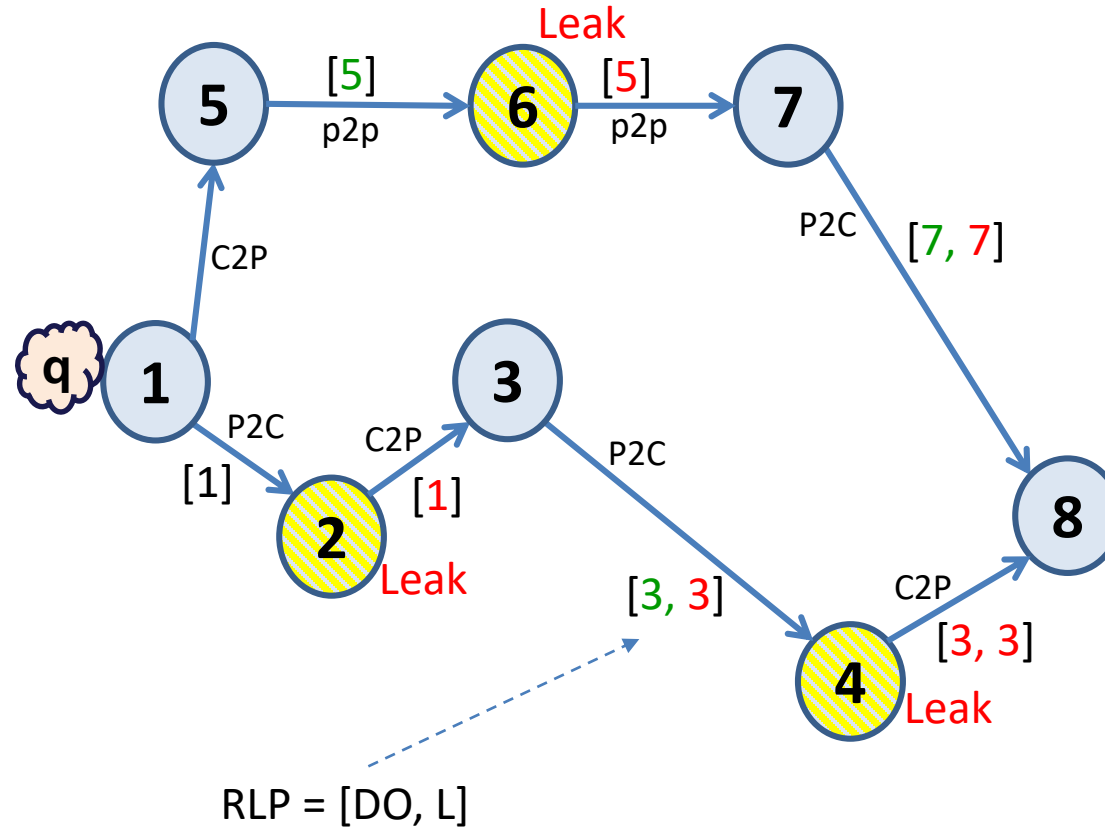
Scenario 5 ☺



# More Scenarios:

**Green** – not violation  
**Red** – violation

## Scenario 6 ☹️



Seems worse if customer route is selected at AS 8 ☹️

Preferred route = ?  
 Shortest path tie;  
 use tie breaker

Legend:



AS does not participate in RLP and starts/restarts a leak



# Leak not detectable if consecutive ASes not participating

Green – not violation

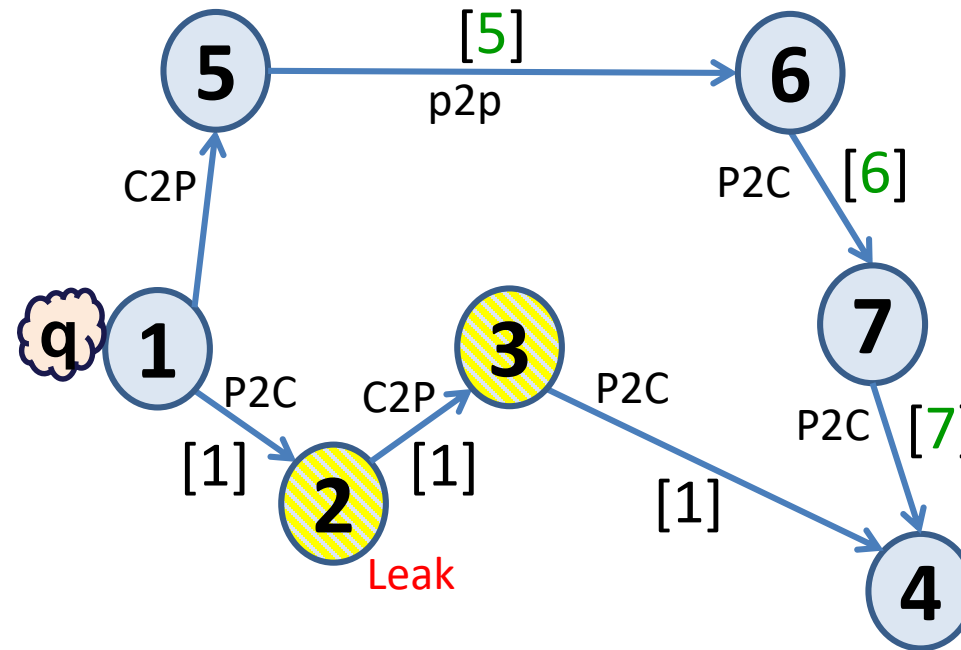
Red – violation

Legend:



AS does not participate in RLP and starts/restarts a leak

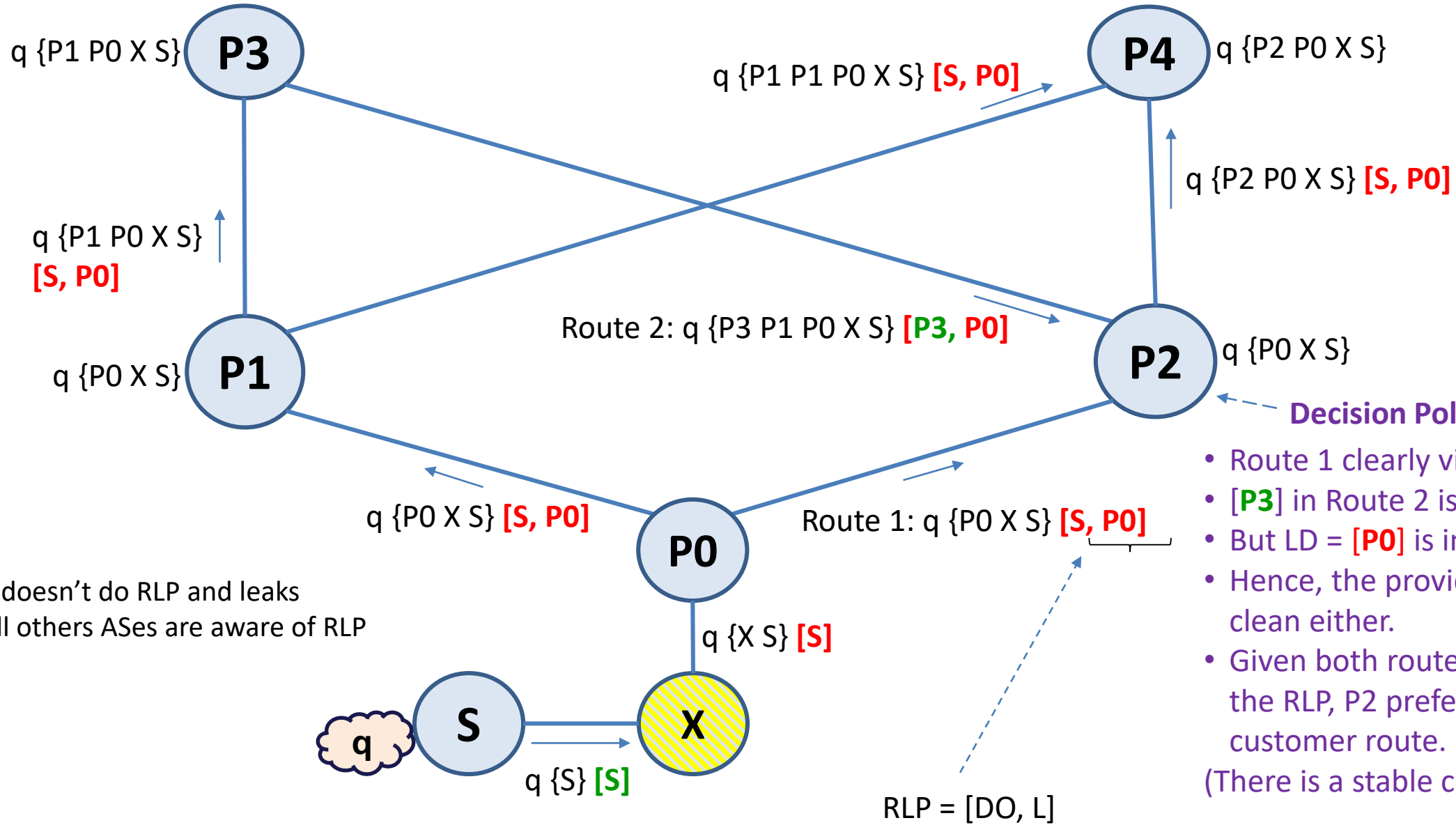
## Scenario 7 ☺



AS4 selects the bad path. It cannot detect that the route from AS 3 is a leak.

# Alexander's scenario

## Scenario 8 😊



X doesn't do RLP and leaks  
All others ASes are aware of RLP

# Encoding RLP in BGP Communities

Relevant RFCs:

[RFC 4360](#): BGP Extended Communities Attribute

[RFC 7153](#): IANA Registries for BGP Extended Communities

[RFC 8092](#): BGP Large Communities Attribute

# Encoding RLP in BGP Communities – 3 Choices

Three choices:

Choice X: One Transitive Large Community: Global Administrator, DO (ASN value), L (ASN value)

Choice Y: Two Transitive Large Communities:

1<sup>st</sup> one: Global Administrator, 16-bit Type (value assigned for DO), DO (ASN value)

2<sup>nd</sup> one: Global Administrator, 16-bit Type (value assigned for L), L (ASN value)

(Choice Y is similar to what John suggested)

Choice Z: Two Transitive Extended Communities (Opaque):

1<sup>st</sup> one: 0x03, 8-bit Sub-Type (value assigned for DO), DO (ASN value)

2<sup>nd</sup> one: 0x03, 8-bit Sub-Type (value assigned for L), L (ASN value)

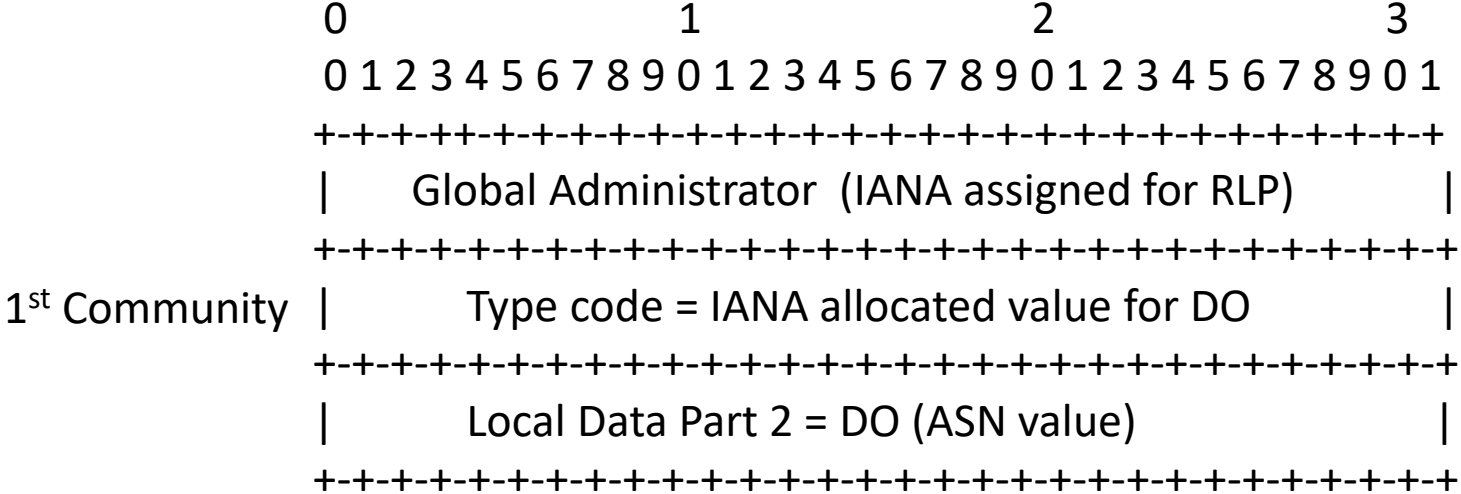
DO = Down Only indication

L = Leak indication



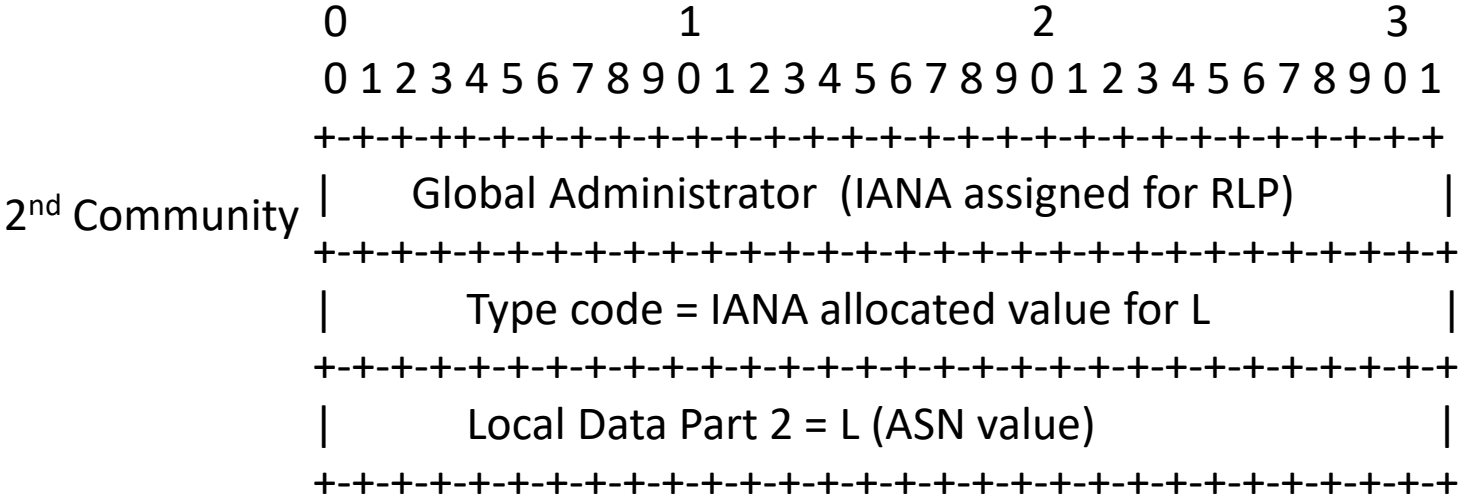
# Encoding Choice Y: Two Transitive Large Communities

(Choice Y is similar to what John suggested)



Global Admin. AS number is shared across RLP and other similar applications.

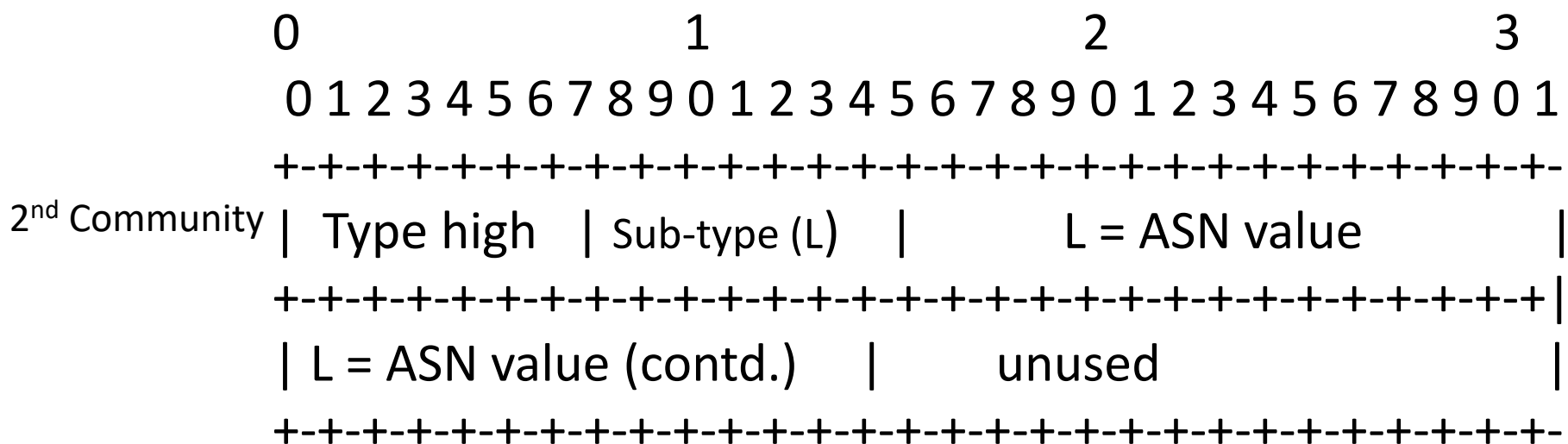
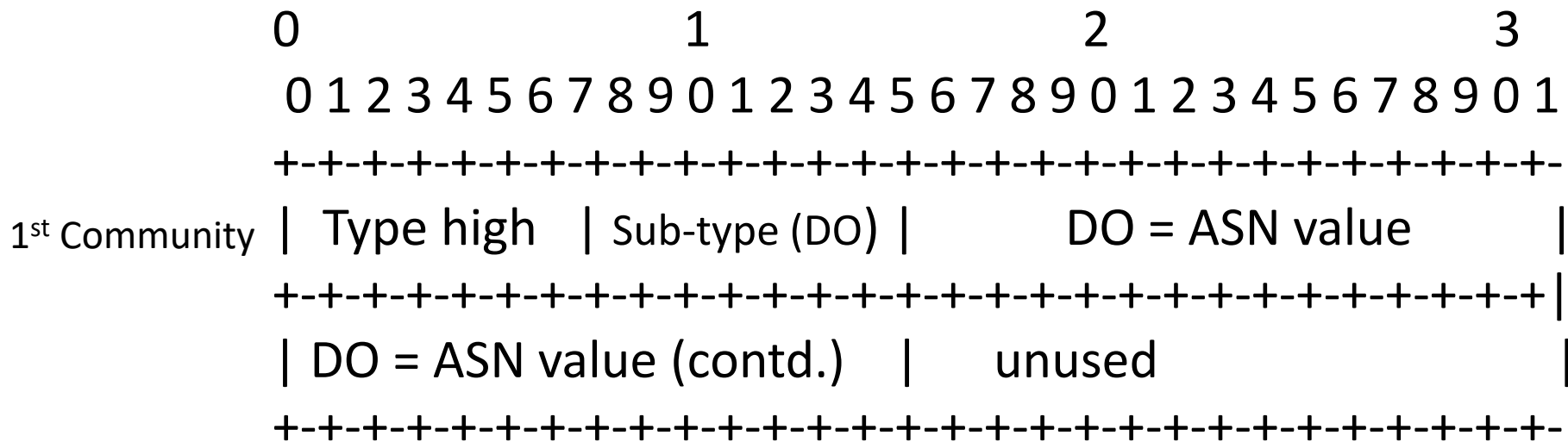
DO = Down Only indication  
L = Leak indication



If no leak was detected by RLP-aware ASes up to the current AS, then L (i.e., the 2<sup>nd</sup> Community) is absent in the received update .

# Encoding Choice Z: Two Transitive Extended Communities

(Opaque: provides 48 bits for data)



[RFC 4360](#): BGP Extended Communities Attribute  
[RFC 7153](#): IANA Registries for BGP Extended Communities

IANA allocated Type high value for RLP

DO = Down Only indication  
 L = Leak indication

If no leak was detected by RLP-aware ASes up to the current AS, then L (i.e., the 2<sup>nd</sup> Community) is absent in the received update .

# Choosing Between Encoding Choices X, Y, and Z

- In Choice X, both DO and X are accommodated in only one Community attribute. Hence, it is more economical than Choices Y and Z in terms of memory and possibly processing.
- Also, may be there is better chance that the single RLP Community attribute in Choice X survives farther (i.e., over greater number of hops) in the update propagation (as compared to two Community attributes in Choices Y and Z).
- Choices Y and Z have more bits to play with in case they're necessary for richer semantics (though the need for that is not evident at this point).



# Pseudo Code – simple default code

```
<receiver action for leak detection>  
<!-- this precedes route selection policy -->  
if received route includes L, then save the route in RIB-in as is;  
else (L is absent), if route is received from a customer and DO is preset, then add L = local ASN;  
else (L is absent), if route is received from a lateral peer and DO is present that is not the lateral peer's ASN, then add L  
= local ASN  
</receiver action for leak detection>
```

Comment: “Route does not include L” or “L is absent” if L is either literally absent or has the default (all zeros) value.

```
<route selection policy>  
for each route that includes L, lower the LocalPref to X (TBD);  
apply best path selection policy*;  
</route selection policy>
```

\* E.g., best path selection based on LocalPref first and then shortest path.

```
<sender action>  
<!-- note: RLP (includes DO and L or just DO) is a *transitive* BGP Community -->  
when propagating a route originated by local AS to a customer or lateral peer, add DO = local ASN;  
when propagating a route that includes a DO (i.e., was received with a DO) to a customer or lateral peer, replace  
the DO value with the local ASN;  
</sender action>
```

# Pseudo Code – operator preferences (if any)

```
<receiver action for leak detection>  
<!-- this precedes route selection policy -->  
if received route includes L, then save the route in RIB-in as is;  
else (L is absent), if route is received from a customer and DO is preset, then add L = local ASN;  
else (L is absent), if route is received from a lateral peer and DO is present that is not the lateral peer's ASN, then add L  
= local ASN  
</receiver action for leak detection>
```

Comment: “Route does not include L” or “L is absent” if L is either literally absent or has the default (all zeros) value.

```
<route selection policy>  
[insert code according to operator preferences here]*  
</route selection policy>
```

\* E.g., Examples: (1) Operator may prefer route from transit provider over customer if both have L present; (2) Operator may prefer route from customer over transit provider if both have L present, and the latter is a detour of the former (i.e., the customer AS is common to both paths).

```
<sender action>  
<!-- note: RLP (includes DO and L or just DO) is a *transitive* BGP Community -->  
when propagating a route originated by local AS to a customer or lateral peer, add DO = local ASN;  
when propagating a route that includes a DO (i.e., was received with a DO) to a customer or lateral peer, replace  
the DO value with the local ASN;  
</sender action>
```

**Thank you.**  
**Comments /discussion?**

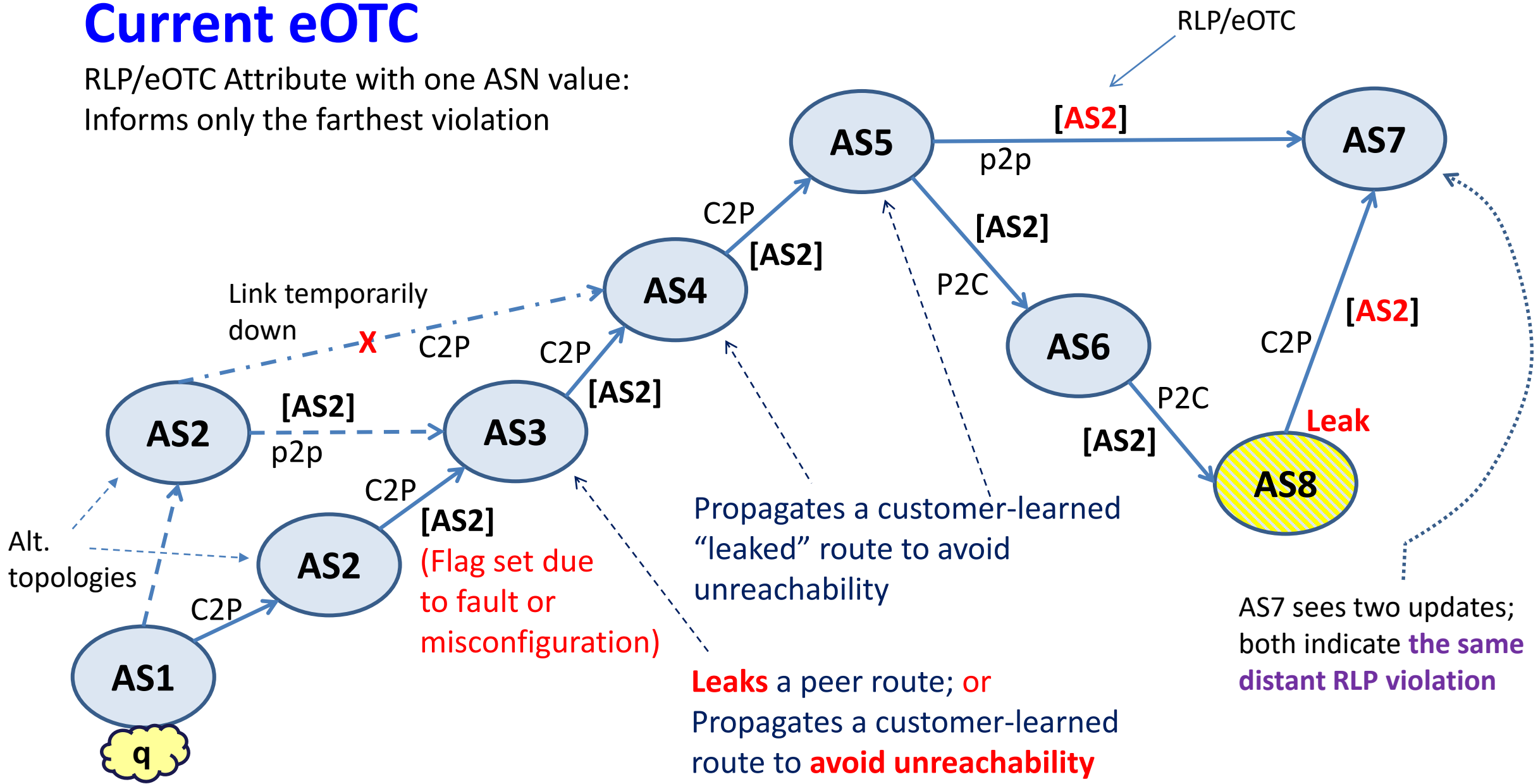
## Stop here

The rest are earlier design slides – prior to IETF 102

Just in any case anyone needs to refer back.

# Current eOTC

RLP/eOTC Attribute with one ASN value:  
Informs only the farthest violation



**Design A:** Same as what is in the RLP draft currently.

Variable size RLP Attribute (an RLP vector) – each AS inserts its ASN value and sets RLP = 00 (default) or RLP = 1 (route sent to customer or peer) in the RLP Attribute.

## **Design A1: Description**

Small and fixed size RLP Attribute (still an RLP vector)

- The RLP Attribute is 32-bits (4 octets) long – same length as for the current eOTC. In this design, each AS in the path computes the path length (after compressing prepends and including itself) of the route it is advertising. If this path length is  $k$  and the AS wants to set an RLP flag, then the AS sets the  $k$ -th bit in the 32-bit RLP Attribute to 1. If the current AS wants to set an RLP flag, but the route does not yet have the RLP attribute, then the current AS adds the 32-bit RLP Attribute. At that point, all bits in the Attribute are zero except the one bit corresponding to the current AS. Each subsequent AS in the path leaves the bit corresponding to itself unchanged at 0 if it is not setting an RLP flag.
- This design eliminates the memory cost argument against the original per-hop RLP proposal.
- RLP Attribute length of 32 bits seems more than sufficient based on Geoff Huston's measurements which show the max AS path length in the teens (for years). If we want to be very conservative, we can make the RLP Attribute 64 bits (still small and fixed size).

## Design B1: Description

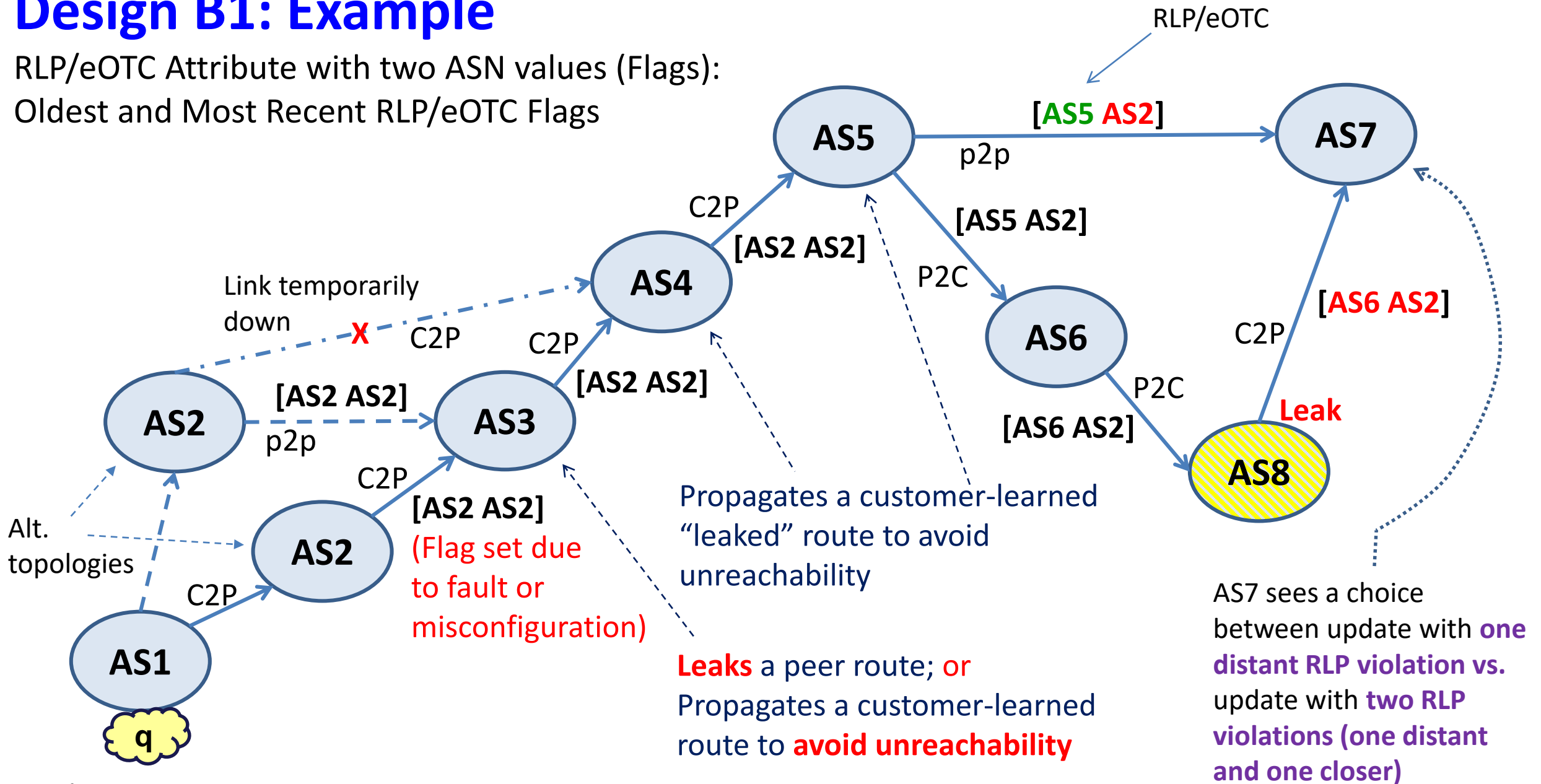
RLP/eOTC Attribute with two ASN values (Flags):  
Oldest and Most Recent RLP/eOTC Flags

- The RLP (or eOTC) Attribute accommodates two ASN values. The left ASN is the closest AS that set its RLP flag. The right ASN is the farthest AS that set its RLP flag. The Attribute informs the closest and the farthest RLP violations.
- The first AS in the path to set the RLP/eOTC flag inserts its ASN value in both places (left and right). A subsequent AS that wants to set its RLP/eOTC flag replaces the left ASN with its own ASN, and leaves the right ASN untouched.

This design is illustrated in the next slide.

# Design B1: Example

RLP/eOTC Attribute with two ASN values (Flags):  
 Oldest and Most Recent RLP/eOTC Flags





## Design B2: Description

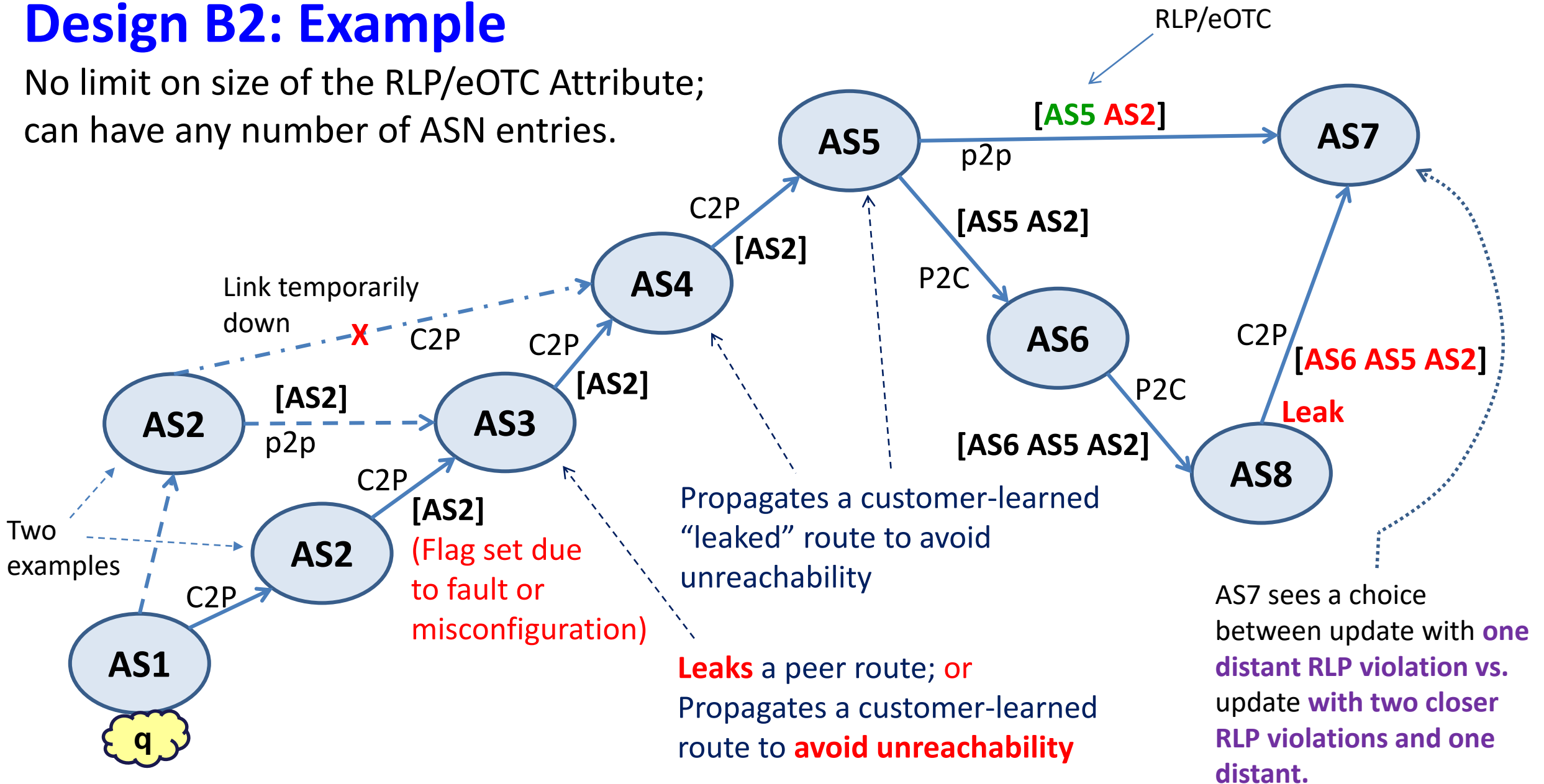
RLP/eOTC Attribute with variable number of ASN values (Flags): All ASes that participate can set their RLP/eOTC Flag (no limit on the size of the attribute)

- Here the RLP (or eOTC) Attribute accommodates any number of ASN values. The left ASN is the nearest AS that set RLP flag. The right ASN is the farthest AS that set RLP flag.

This design is illustrated in the next slide.

# Design B2: Example

No limit on size of the RLP/eOTC Attribute; can have any number of ASN entries.



# Route-Leak Detection Theorem

The “Gao-Rexford” Stability Conditions

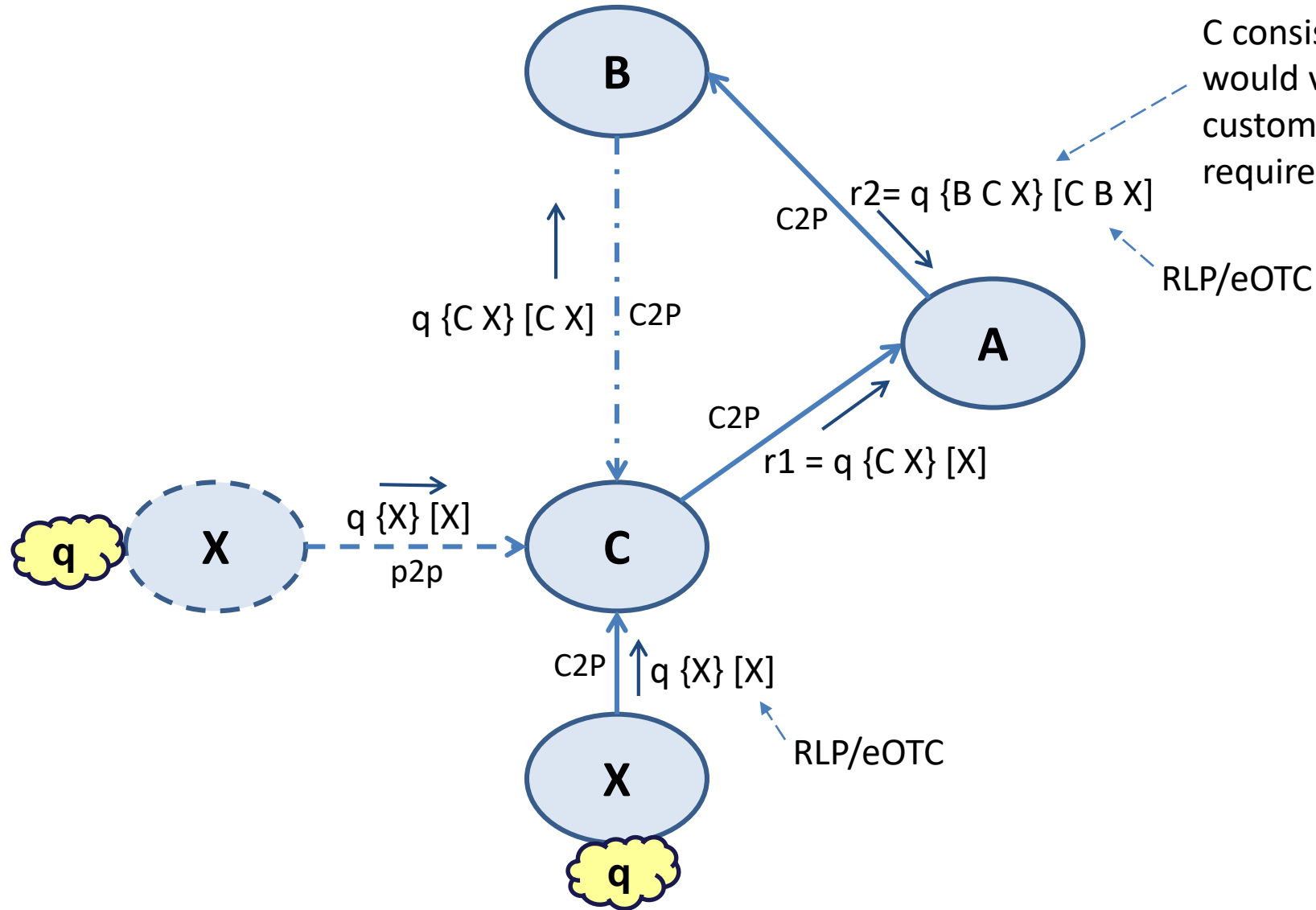
[Gao-Rexford] <http://www.cs.princeton.edu/courses/archive/spr11/cos461/docs/lec17-bgp-policy.ppt>

- **Topology** condition (acyclic) (slide 27)
  - No cycle of customer-provider relationships

**Route-Leak Detection Theorem:** Let it be given that ISP A receives a route  $r_1$  from customer AS C and another route  $r_2$  from provider AS B (for the same prefix), and both routes  $r_1$  and  $r_2$  contain AS C and AS X in the path and also contain [X] in their RLP/eOTC. Then, clearly  $r_1$  is in violation of [X]. It follows that  $r_2$  is also necessarily in violation of [X].

**Proof:** Let us suppose that  $r_2$  is not in violation of [X]. That implies that  $r_2$ 's path from C to B to A included only P2C links. That would mean that there is a cycle of customer-provider relationships involving the ASes in the AS path in  $r_2$ . However, any such cycle is ruled out in practice as a necessary stability condition [Gao-Rexford]. QED.

# Route-Leak Detection Theorem: Illustration



The only possible way that [X] is not violated in r2 is if the path from B to C consists of C2P links only. But that would violate the “No cycle of customer-provider relationships” requirement [Gao-Rexford].

RLP/eOTC

RLP/eOTC



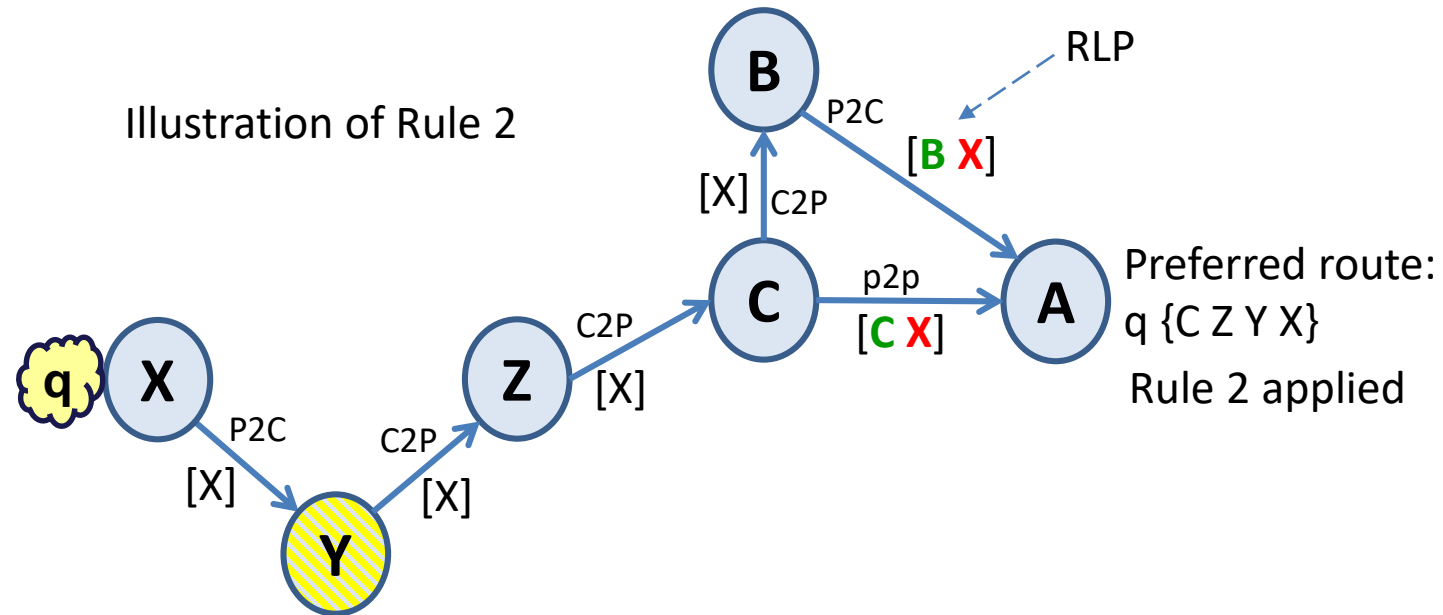
# Route-Leak Mitigation Rules

**Rule 1:** If ISP A receives a route r1 from customer AS C and another route r2 from provider (or peer) AS B (for the same prefix), and both routes r1 and r2 contain AS C and AS X (any X not equal to C) in the path and also contain [X] in their RLP, then prioritize the customer (AS C) route over the provider (or peer) route.

(Rationale: This rule is based on the theorem (slide 7) and its extension to customer vs. peer case. Example of application of Rule 1 is on slide 9.)

**Rule 2:** If ISP A receives a route r1 from peer AS C and another route r2 from provider AS B (for the same prefix), and both routes r1 and r2 contain AS C and AS X (any X not equal to C) in the path and also contain [X] in their RLP, then prioritize the peer (AS C) route over the provider (AS B) route.

(Rationale: The peer (AS C) route is in violation of [X], and so is the provider (AS B) route. Also, realize that the provider route is necessarily a detour of the peer route (with longer AS path). See illustration below.)

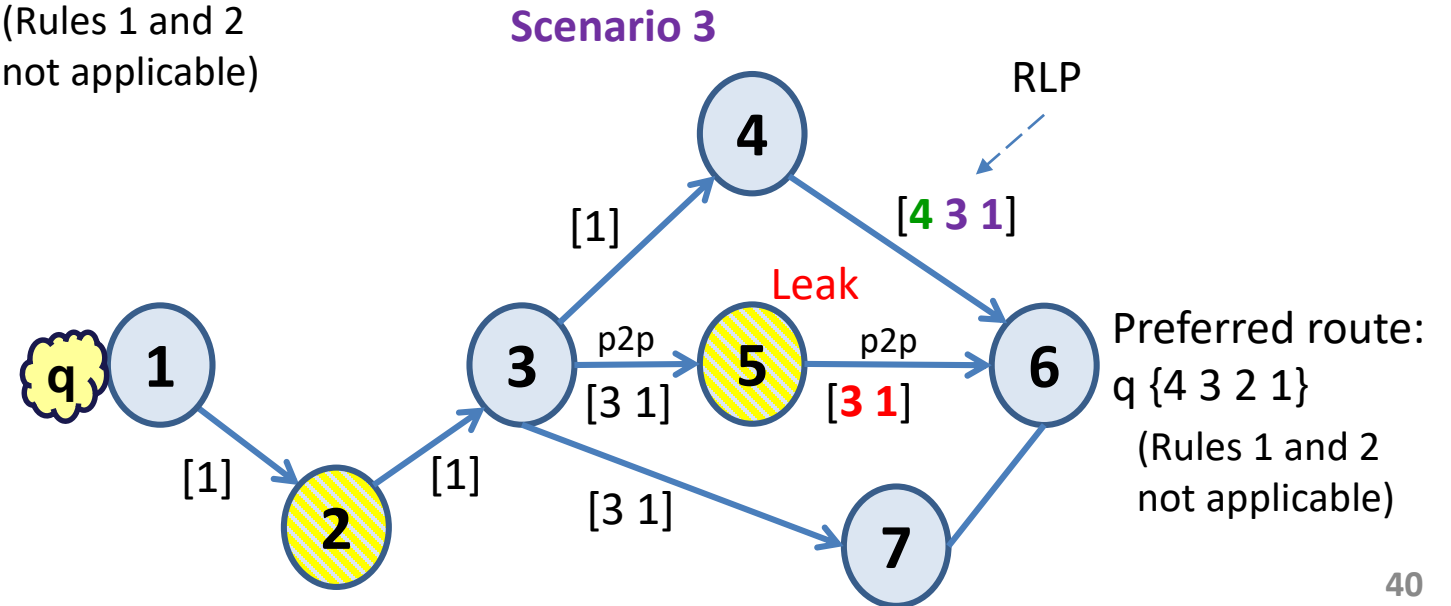
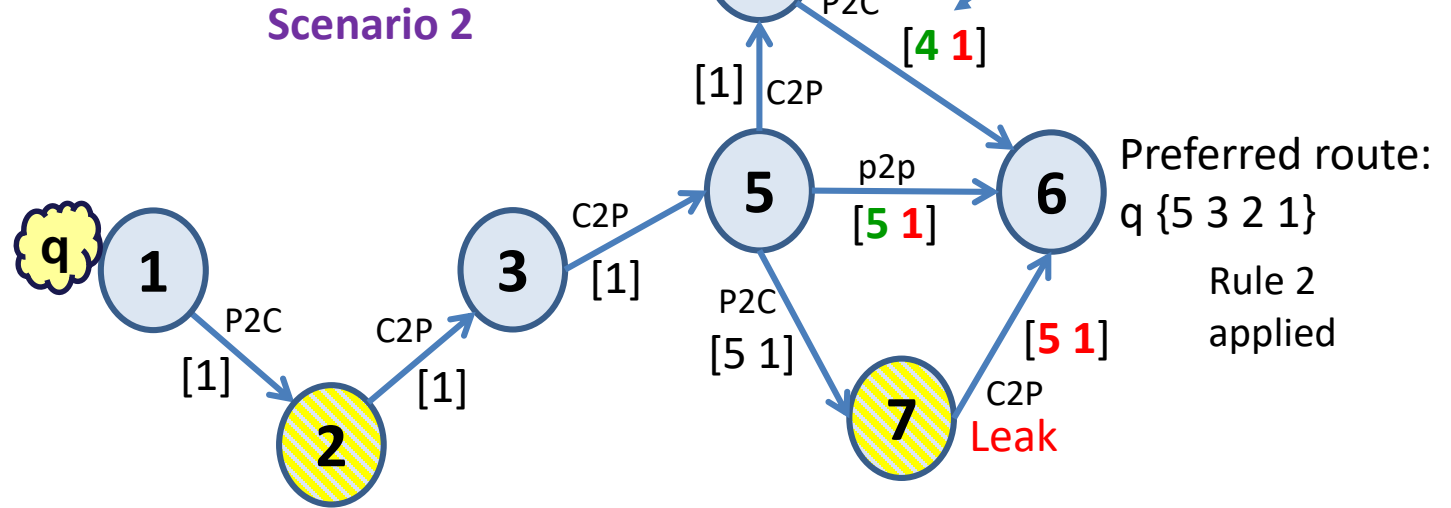
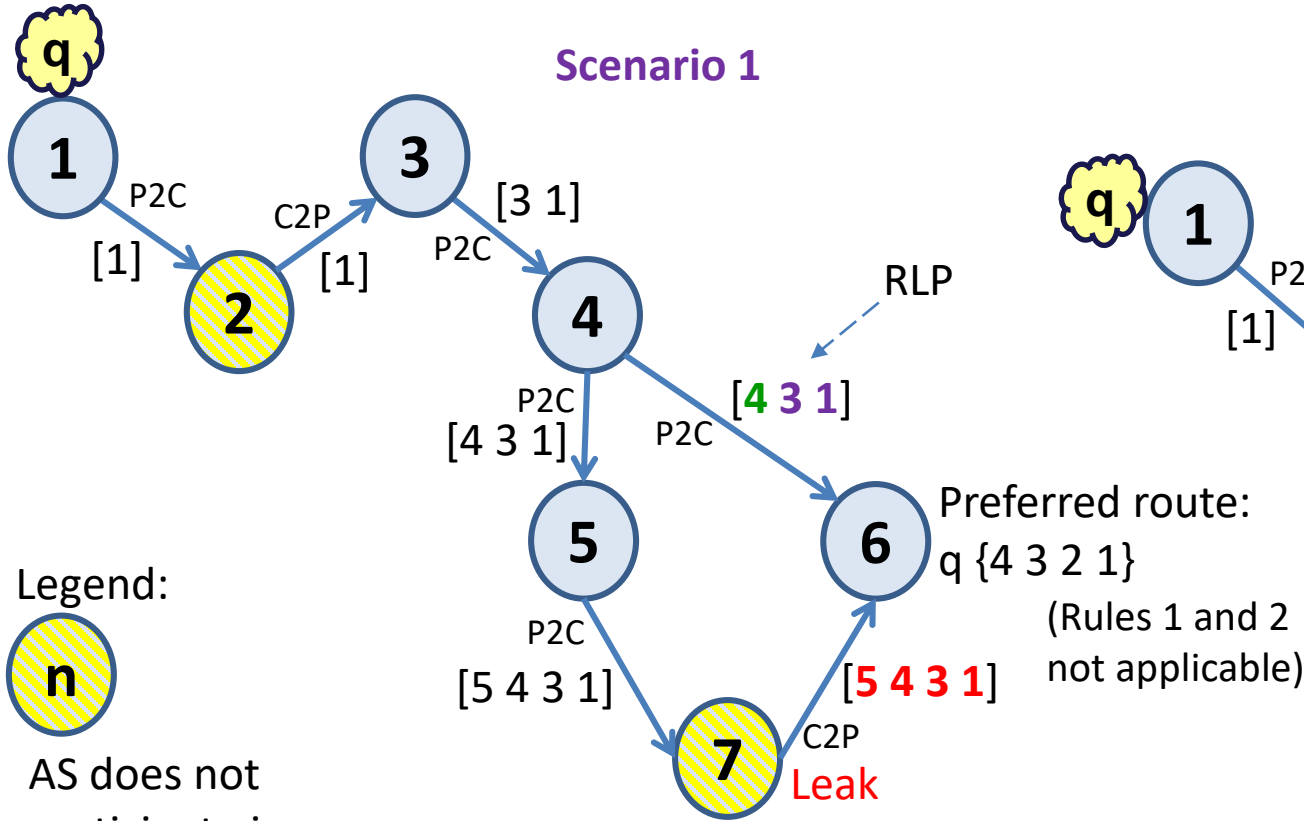


# Default Route-Leak Mitigation Policy

Given a choice between a customer route versus a provider (or peer) route,  
if no route leak is detected in the customer route, then prioritize the customer over the provider (or peer);  
else (i.e., when route leak is detected in the customer route) and the conditions of Rule 1 apply, then too  
prioritize the customer over the provider (or peer);  
else (i.e., when route leak is detected in the customer route and the conditions of Rule 1 DO NOT apply),  
then prioritize the provider (or peer) over the customer.

Given a choice between a peer route versus a provider route,  
if no route leak is detected in the peer route, then prioritize the peer over the provider;  
else (i.e., when route leak is detected in the peer route) and the conditions of Rule 2 apply, then too  
prioritize the peer over the provider;  
else (i.e., when route leak is detected in the peer route and the conditions of Rule 2 DO NOT apply),  
then prioritize the provider over the peer.

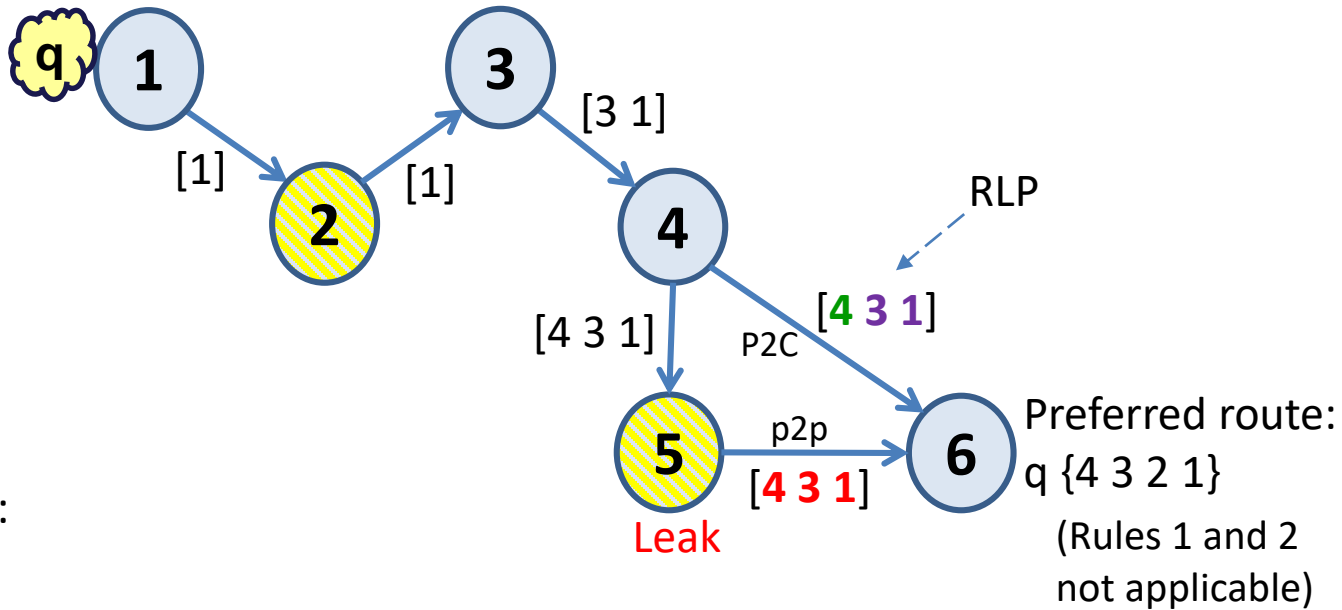
# Examples Showing Policy in Action (1 of 2)





# Examples Showing Policy in Action (2 of 2)

Scenario 4



Legend:



AS does not participate in RLP and starts/restarts a leak

Green – not violation  
 Red – violation  
 Purple – can't tell