

Solution for Route Leaks Using BGP Communities

<https://tools.ietf.org/html/draft-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

**IETF IDR Interim Meeting
October 26, 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.)

General Principles of the Design

- Why BGP Community: Faster deployment without dependence on vendor implementation changes
- 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 a route 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

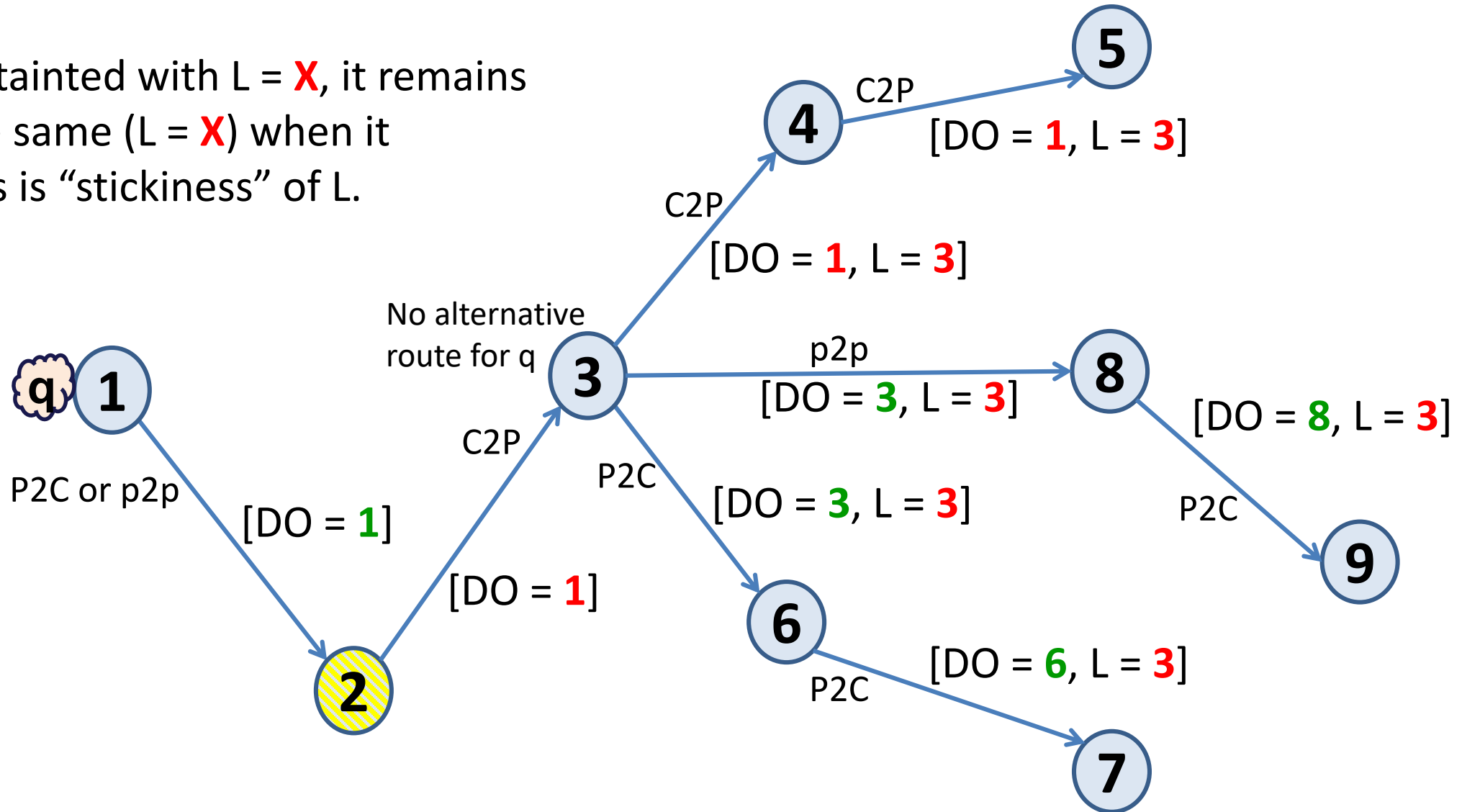
Illustration of Down Only (DO) and Leak (L) indications – 1 of 2

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:

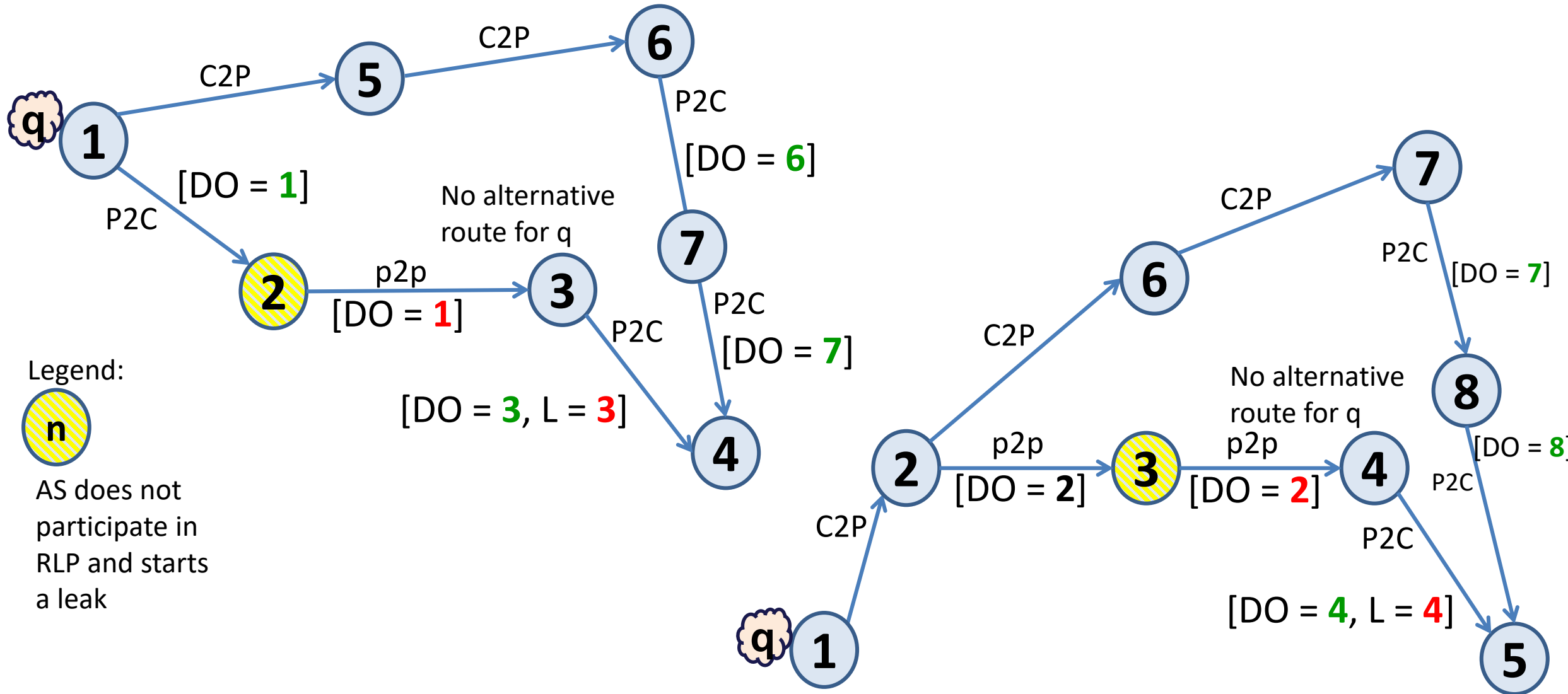


AS does not participate in RLP and starts/restarts a leak



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

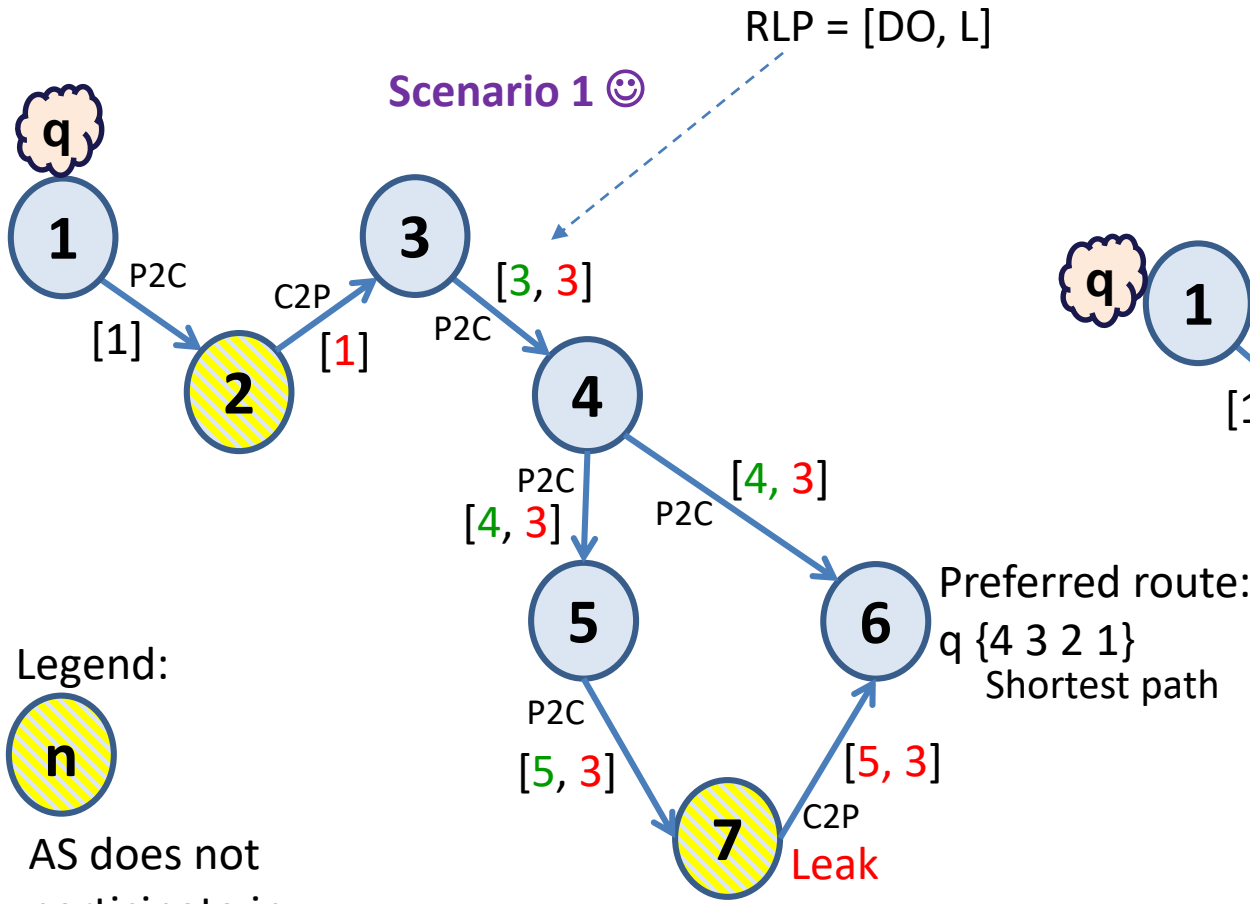
Illustration of Down Only (DO) and Leak (L) indications – 2 of 2



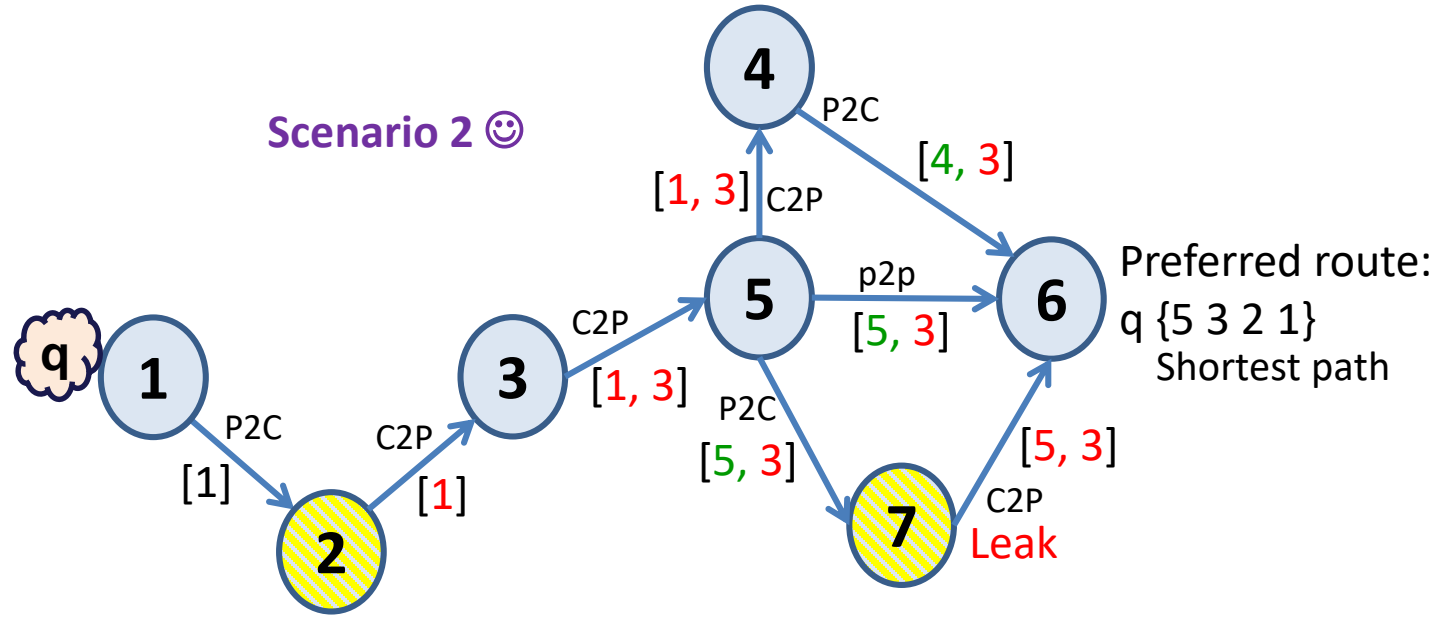
Note: RLP = Route Leak Protection; DO alone or DO and L together constitute RLP

Scenarios:

Scenario 1 😊



Scenario 2 😊

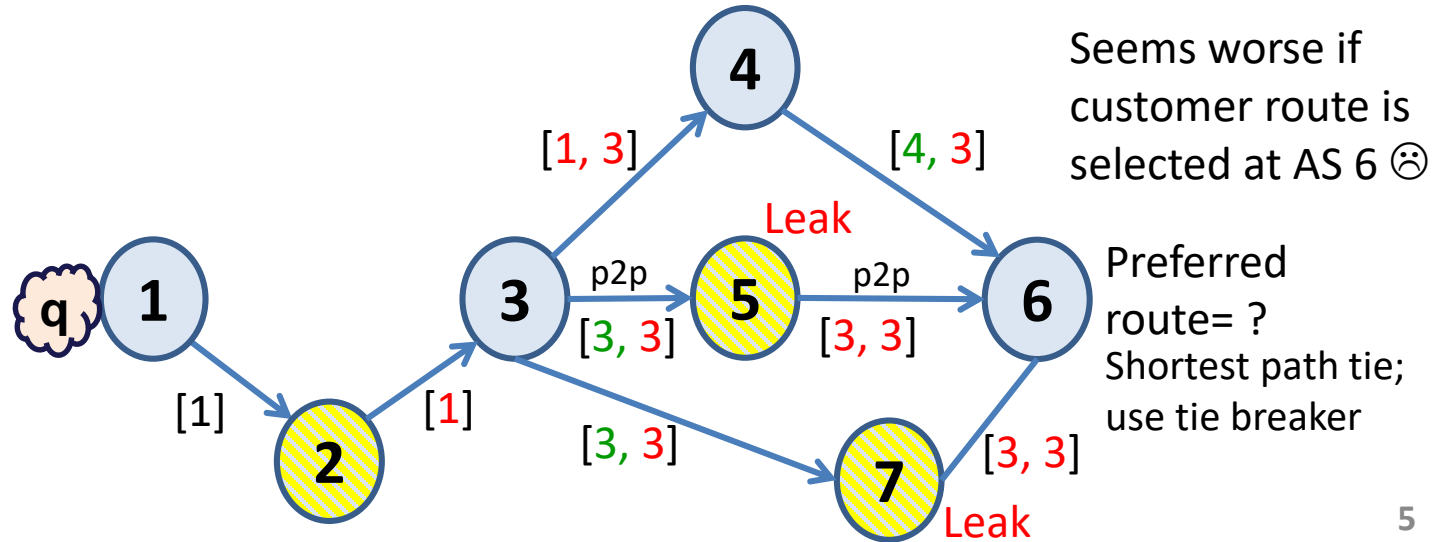


Legend:

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

Green – not violation
Red – violation

Scenario 3 😞

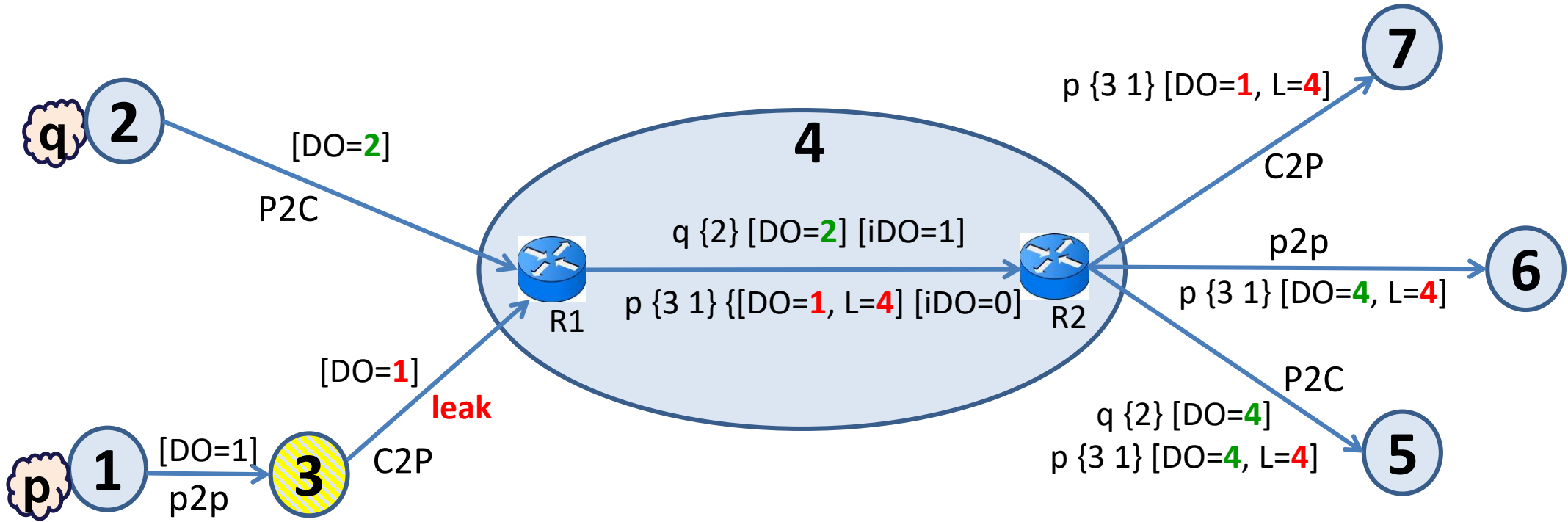


Prefixes with Single Path to Tier1


- 1 hop length 91152
- 2 hops length 56558
- 3 hops length 16348
- 4 hops length 2755
- 5 hops length 274s
- 6 hops length 54
- 7 hops length 5

Measurements by Qrator Labs

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 (AS4) MUST not send non-customer routes to lateral peer AS6 or transit AS7.

iDO here is similar to iOTC

<https://tools.ietf.org/html/draft-ietf-idr-bgp-open-policy-03>

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

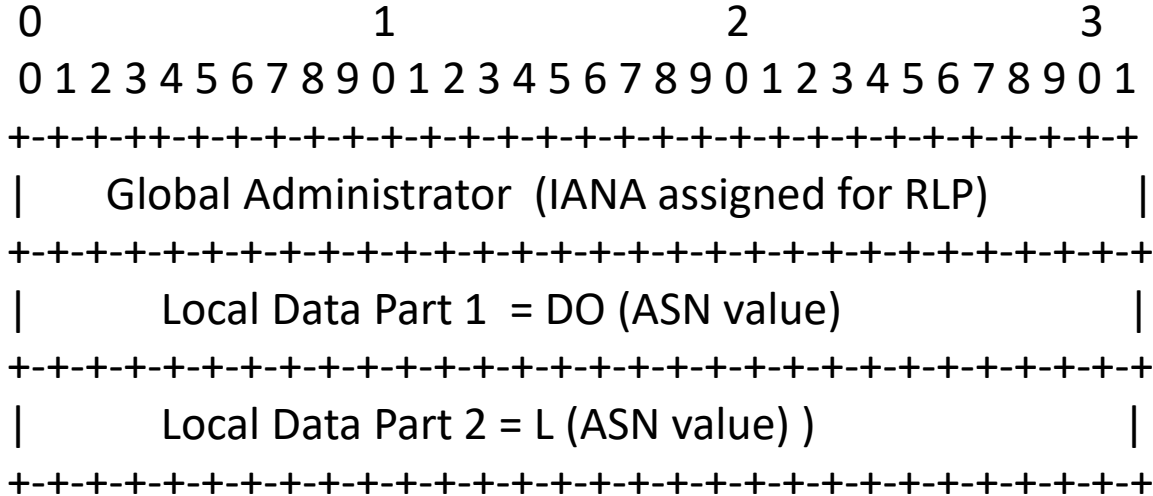
- Whenever there is choice between multiple routes (customer/peer/provider), and each is detected to be in RLP violation, then lower the LocalPref to X (TBD) for each of them. Then apply shortest path criterion*.

* Some network operators may find this inadequate (see the analyzed scenarios)

* But they can locally modify their policy while respecting the basic principle

Encoding Choice X: Single Transitive Large Community

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



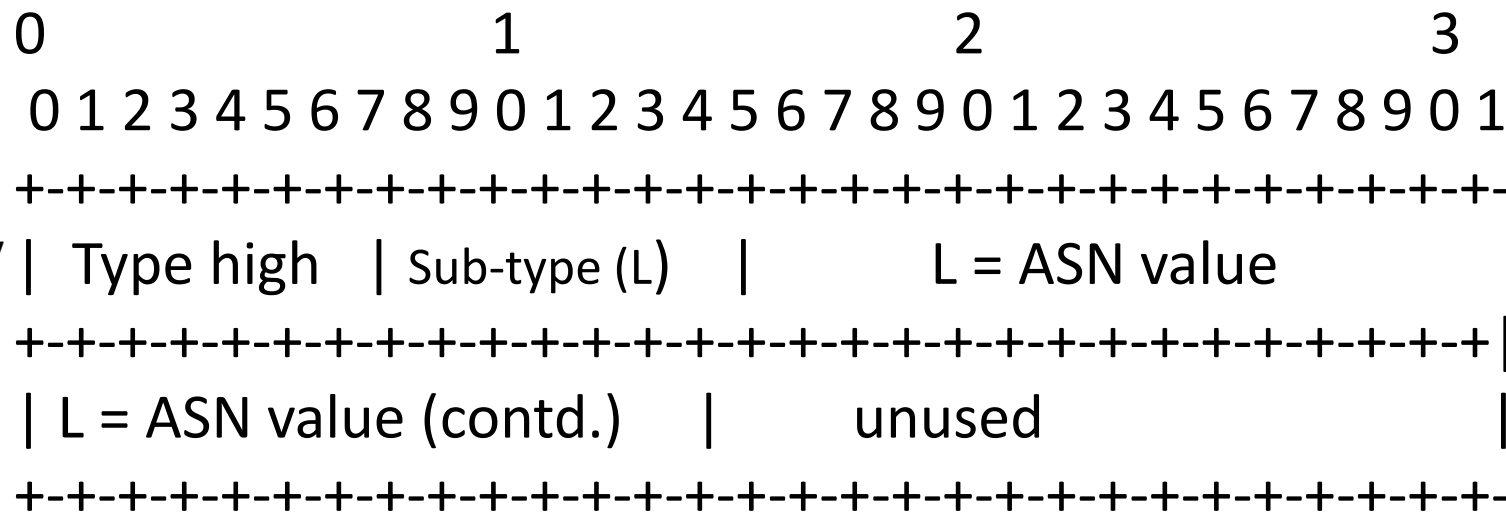
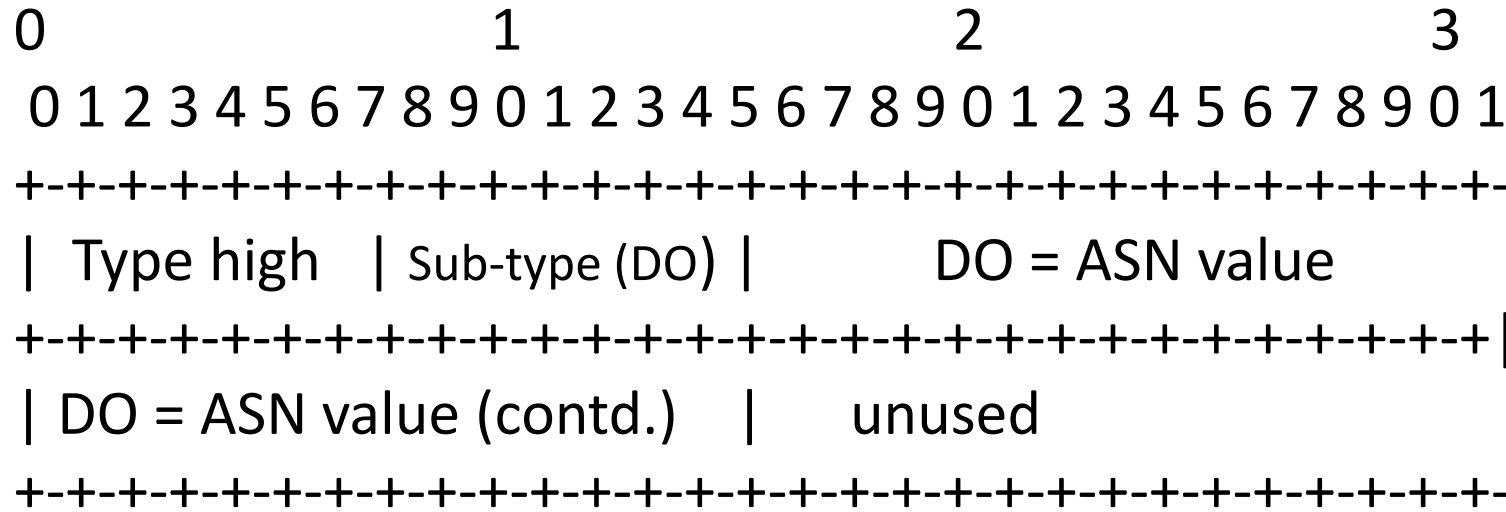
DO = Down Only indication
L = Leak indication

When L is not present, then the Local Data Part 2 is set to some default value such as all zeros (TBD).

For leak indication (L) value, it is better to inform which AS detected the leak rather than simply say that a leak was detected.

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 2nd Community) is absent in the received update .

Request WG Inputs

Questions:

- Which type of Community (transitive) is best?
 - Regular Community vs. Large Vs. Extended
- Which has the best chance to propagate farthest?
- How do they compare in terms of deployment speed?

Pseudo Code: Receiver/Sender Actions and Policy

```
<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>
```

Thank you.
Comments / questions?

Backup slides

Solution for Route Leaks Using BGP Communities

Background: In Montreal face-to-face meeting of authors, John and Sue advised the team to explore a BGP Community based solution

- Motivation: Quicker deployment without dependence on vendor implementation changes
- Detection and mitigation semantics are defined
- Many scenarios are analyzed to examine if the semantics work
- Design choices for encoding using Large Community and Extended Community are presented
- Basic operator 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

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:

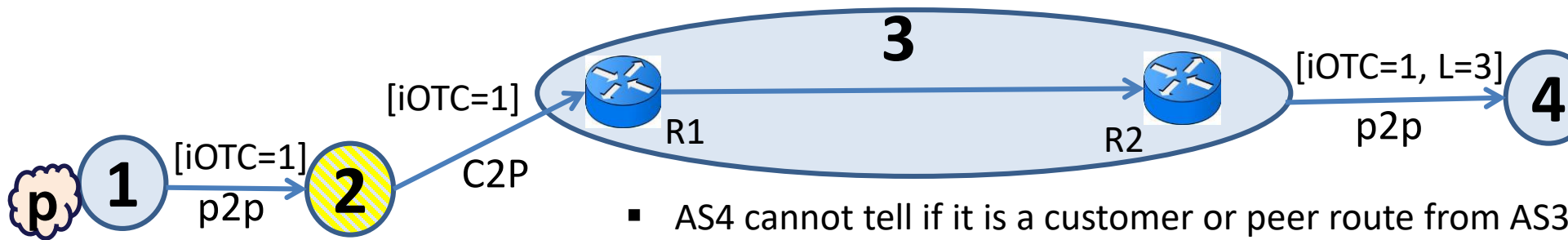
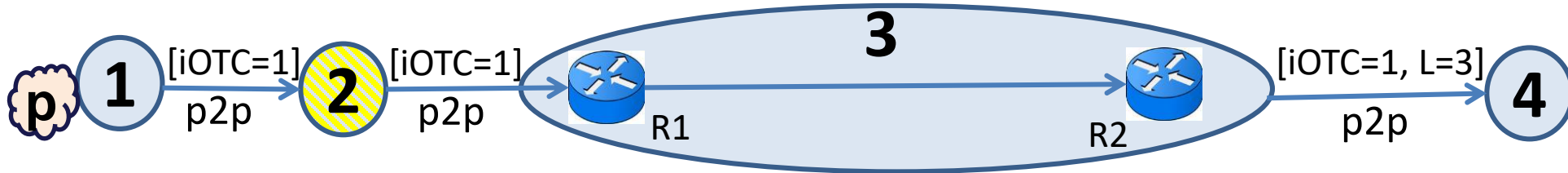
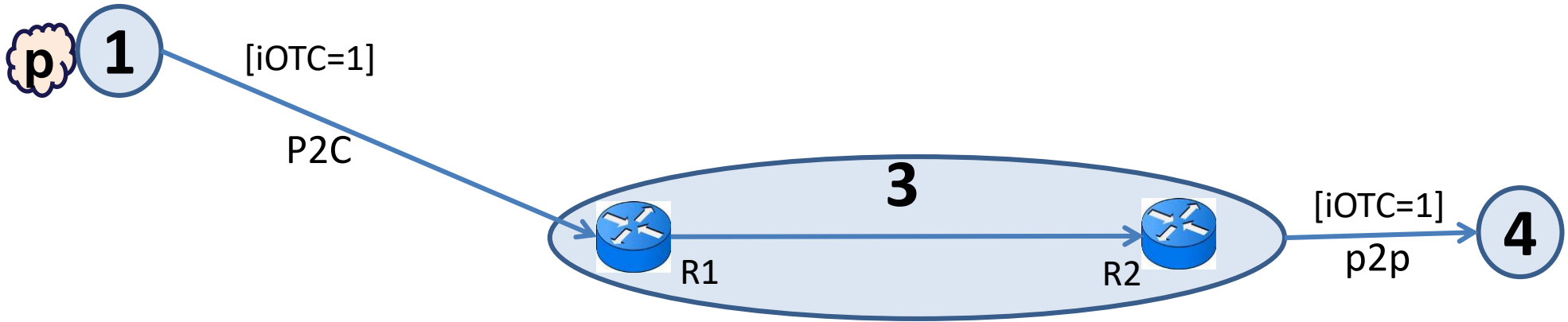
In the following circumstances, a leaked route may not be detected:


- 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.

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.

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.

	Down Only (DO)	Leak (L)	Choice
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.

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:

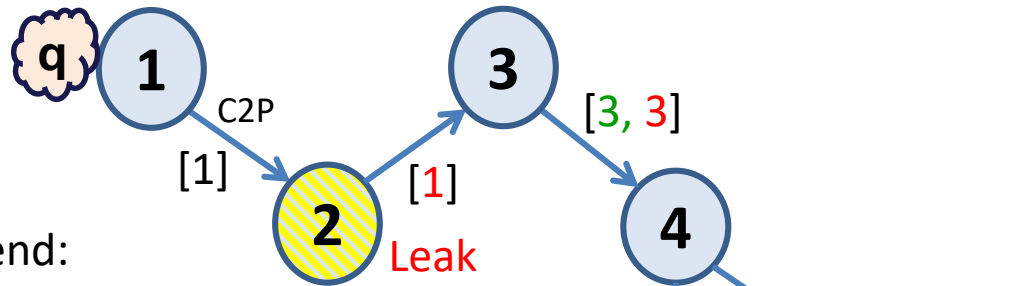
Does this scheme with RLP = [DO, L] along with the policy work?

More Scenarios:

Green – not violation

Red – violation

Scenario 4 ☺



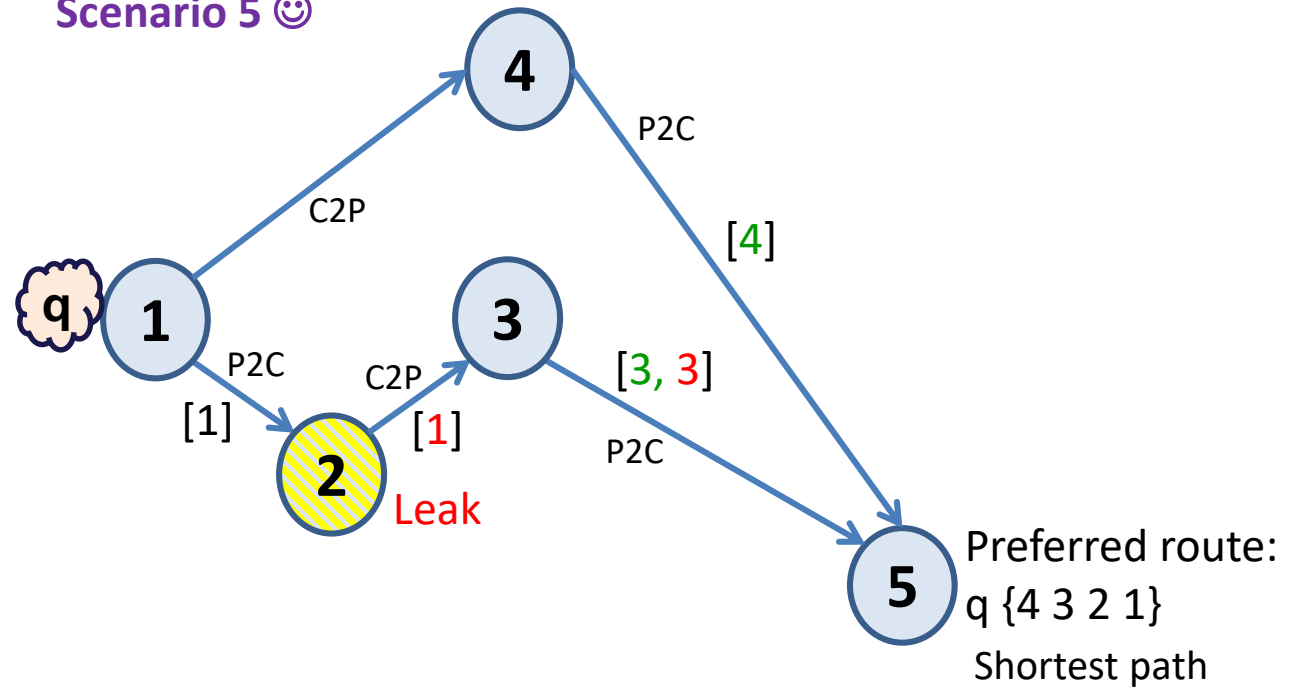
Legend:



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

RLP = [DO, L]

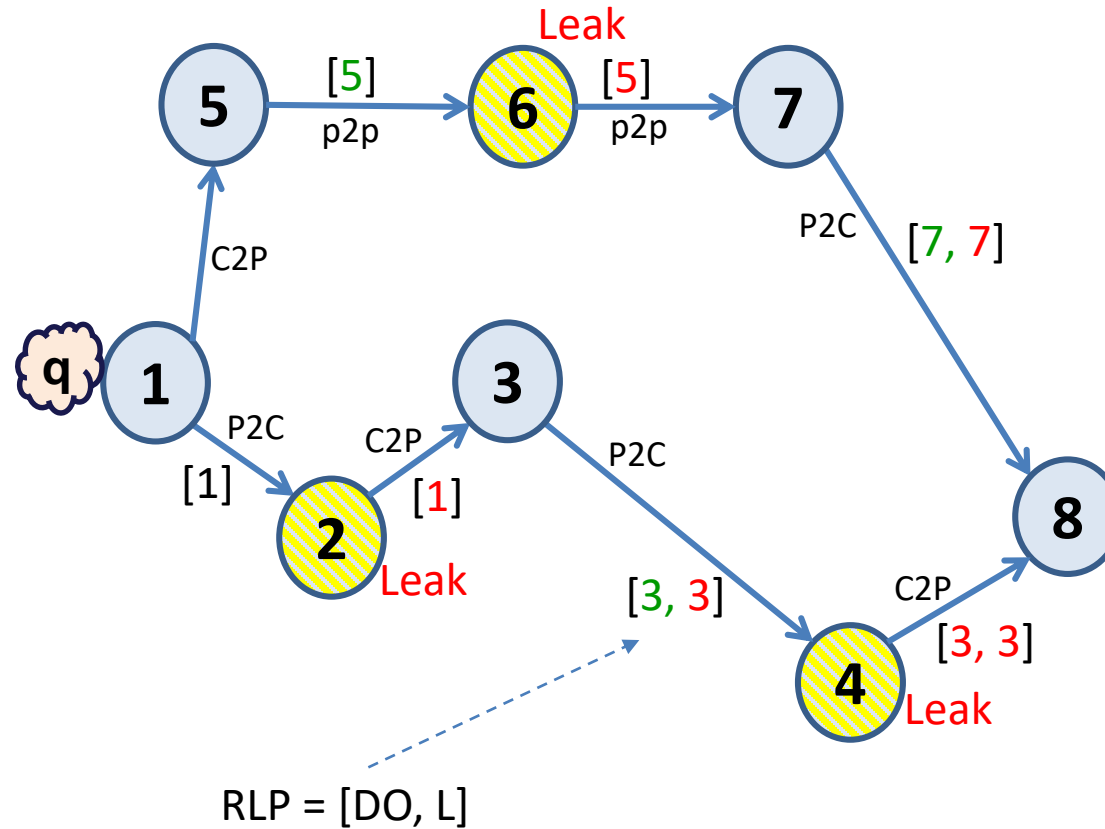
Scenario 5 ☺



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

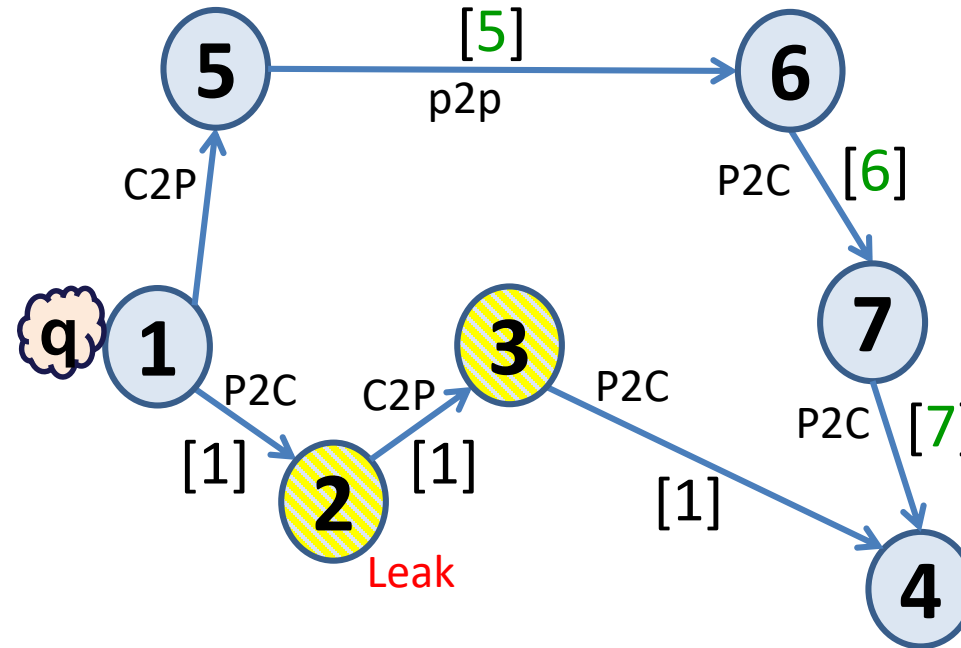
Red – violation

Legend:



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

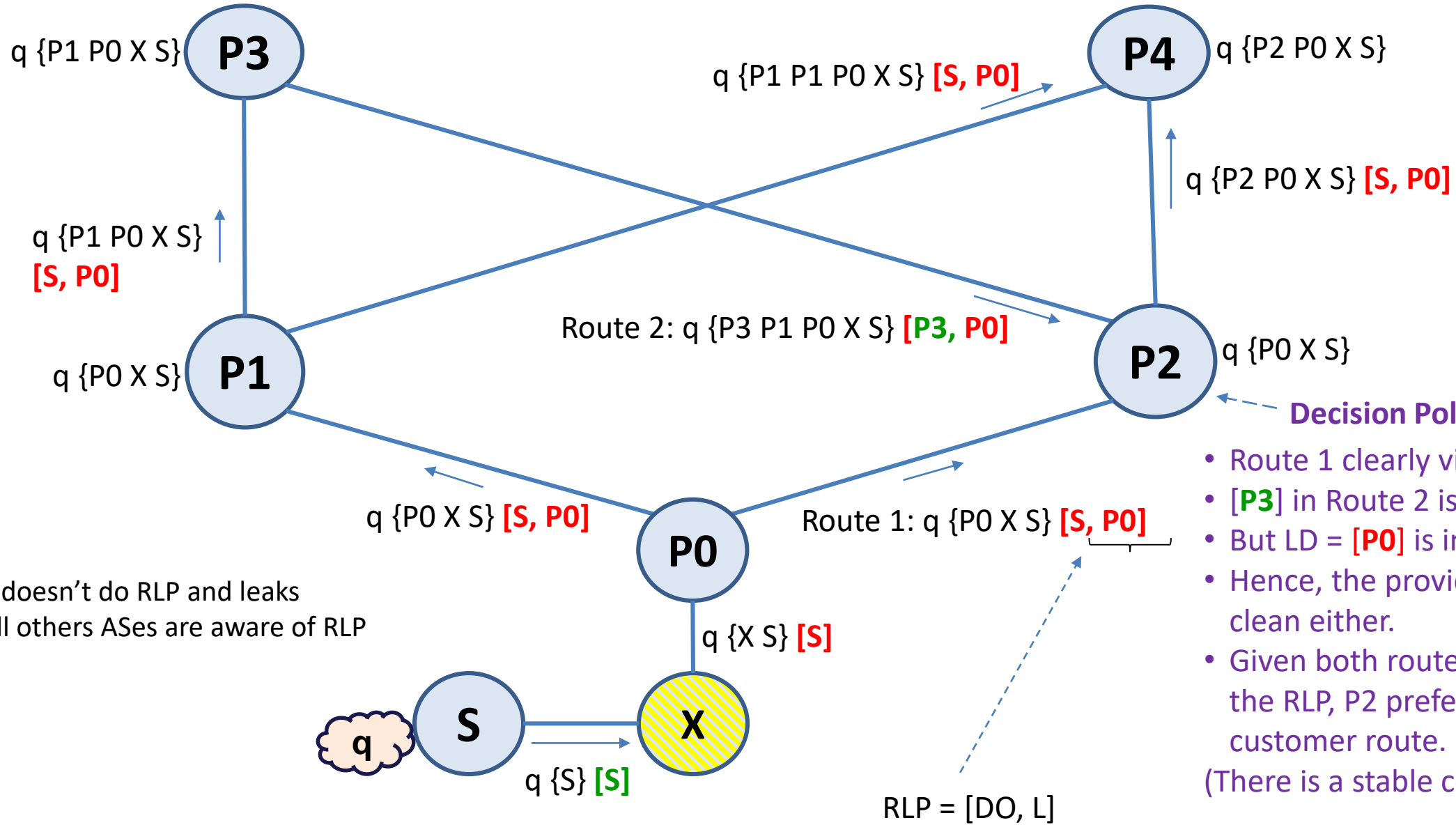
Scenario 7 ☺



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

Alexander's scenario

Scenario 8 ☺



Decision Policy (Algorithm):

- Route 1 clearly violated [S].
 - [P3] in Route 2 is expected (good).
 - But LD = [P0] is in both routes.
 - Hence, the provider route is not clean either.
 - Given both routes are in violation of the RLP, P2 prefers the shorter customer route.
- (There is a stable convergence.)

RLP = [DO, L]

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:

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

2nd 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):

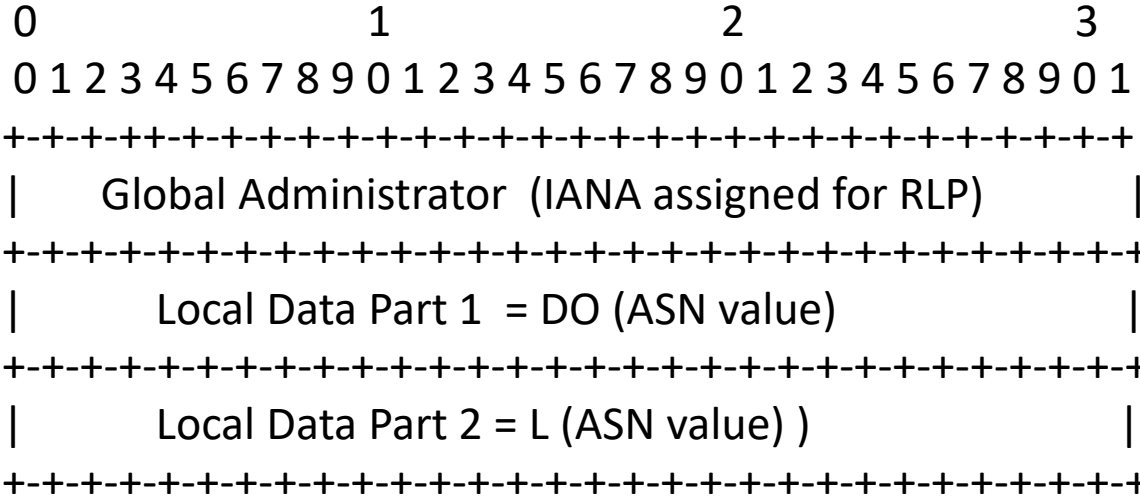
1st one: 0x03, 8-bit Sub-Type (value assigned for DO), DO (ASN value)

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

DO = Down Only indication

L = Leak indication

Encoding Choice X: Single Transitive Large Community



DO = Down Only indication
L = Leak indication

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

When L is not present, then the Local Data Part 2 is set to some default value such as all zeros (TBD).

For leak indication (L) value, it is better to inform which AS detected the leak rather than simply say that a leak was detected.

Encoding Choice Y: Two Transitive Large Communities

(Choice Y is similar to what John suggested)

```

0             1             2             3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|   Global Administrator (IANA assigned for RLP)   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
1st Community |   Type code = IANA allocated value for DO   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|   Local Data Part 2 = DO (ASN value)   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

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

DO = Down Only indication
L = Leak indication

```

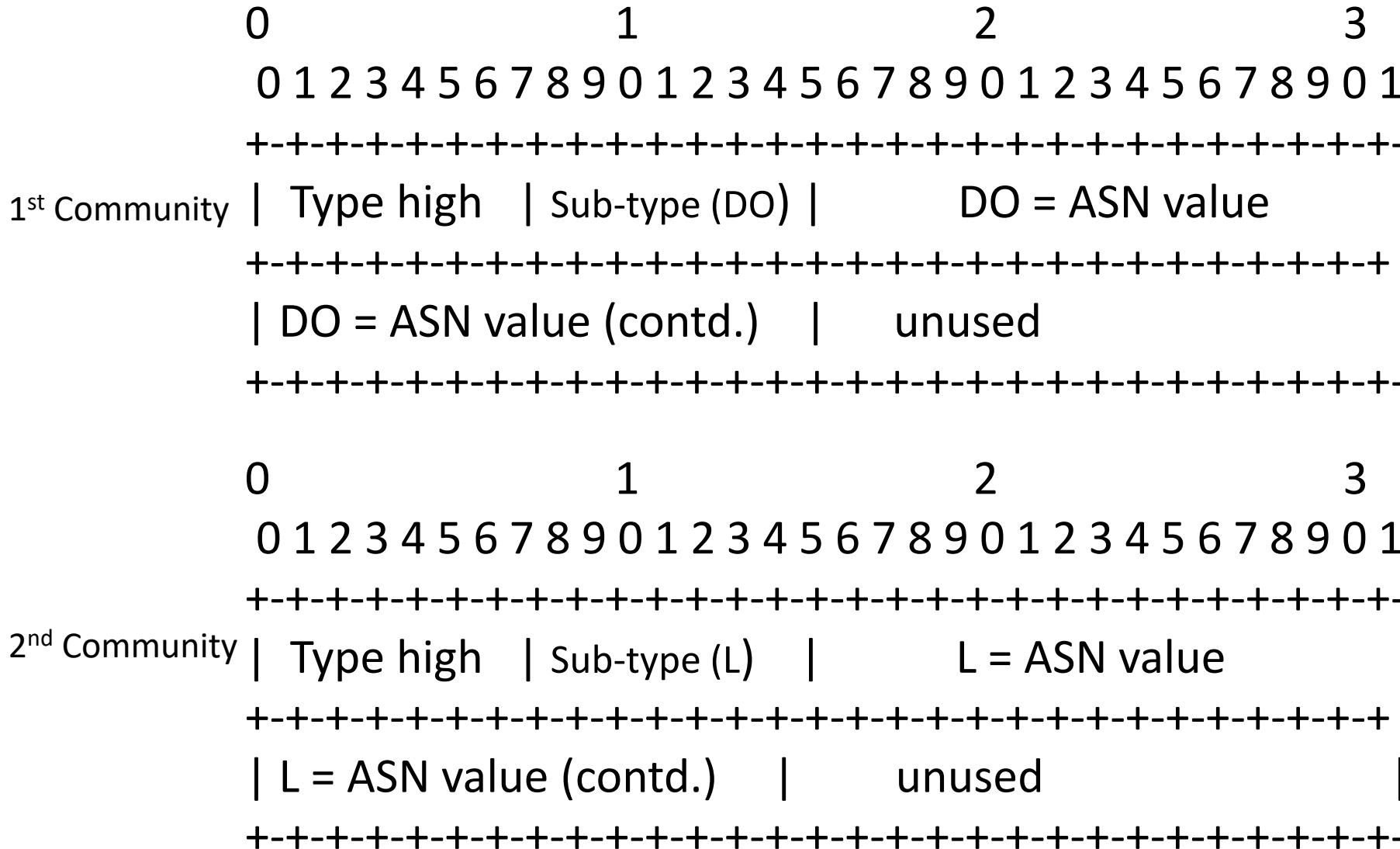
0             1             2             3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|   Global Administrator (IANA assigned for RLP)   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|   Type code = IANA allocated value for L   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|   Local Data Part 2 = L (ASN value)   |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

If no leak was detected by RLP-aware ASes up to the current AS, then L (i.e., the 2nd 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 2nd 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 – 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>
```