iBGP Deceptions: More Sessions, Fewer Routes

Laurent Vanbever
UCLouvain, Louvain-la-Neuve, Belgium
laurent.vanbever@uclouvain.be

Joint work with:
Stefano Vissicchio (Roma Tre University),
Luca Cittadini (Roma Tre University)
Olivier Bonaventure (UCLouvain)
Breaking News

*Adding* a single iBGP session can result in iBGP distributing *fewer* routes.
iBGP Deceptions: More Sessions, Fewer Routes

Introduction and Motivation

Dissemination correctness

Revisiting the state-of-the-art

Conclusion
iBGP Deceptions: More Sessions, Fewer Routes

Introduction and Motivation

Dissemination correctness

Revisiting the state-of-the-art

Conclusion
BGP is *the* inter-domain routing protocol used today
BGP comes in two flavors
external BGP (eBGP) exchanges reachability information between ASes
internal BGP (iBGP) distributes externally learned routes within the AS

In this talk, we take the perspective of a single AS and focus on iBGP
Plain iBGP mandates a full-mesh of iBGP sessions

Fair warning: some sessions are missing

$O(n^2)$ iBGP sessions where $n$ is the number of routers

... quickly becomes totally *unmanageable*
With Route Reflection, iBGP routers are organized in a hierarchy
Route Reflectors relay updates to iBGP neighbors
Route Reflectors relay updates to iBGP neighbors
Several layers of Route Reflection can be built
Several layers of Route Reflection can be built
OVER sessions connect iBGP peers
UP/DOWN sessions connect a Route Reflector to its client(s)
Best routes are allowed to flow on valid signaling paths only

Valid signaling path match the \texttt{UP* OVER? DOWN*} regular expression

\textbf{BGP Propagation rules}

\begin{tabular}{ll}
\textit{From client} & ✓ ✓ \\
\textit{From peer/RR} & ✓ ✗ \\
\end{tabular}
Routes are allowed to flow on valid signaling paths only

Valid signaling path match the `UP* OVER? DOWN*` regular expression

**BGP Propagation rules**

- **To client**
  - From client: ✓ ✓
  - From peer/RR: ✓ ✗

- **To peer/RR**
  - From client: ✓ ✓
  - From peer/RR: ✓ ✗
Routes are allowed to flow on valid signaling paths only

Valid signaling path match the UP* OVER? DOWN* regular expression

BGP Propagation rules

To client  To peer/RR

From client  ✓  ✓
From peer/RR  ✓  ×

UP OVER DOWN
Routes are allowed to flow on valid signaling paths only

Valid signaling path match the UP* OVER? DOWN* regular expression

BGP Propagation rules

To client  To peer/RR

From client  ✓  ✓

From peer/RR  ✓  ✗
Routes are allowed to flow on valid signaling paths only

Valid signaling path match the \text{UP}^* \text{OVER}? \text{DOWN}^* regular expression

```
OVER UP
```

**BGP Propagation rules**

- **To client**
  - From client: ✓ ✓
  - From peer/RR: ✓ ✗
Breaking News

Adding a single iBGP session can result in iBGP distributing fewer routes
Breaking News

Adding *a single spurious OVER* can result in iBGP distributing fewer routes
A spurious OVER is a special type of OVER

Spurious OVER

An OVER session between two routers \( x \) and \( y \) such that \( x \) or \( y \) is not in the RR top layer

Layer 0 (Top)
Layer 1
Layer 2
Let’s consider a simple example

OVER-RIDE GADGET
Signaling path used to learn the best route

OVER-RIDE GADGET
Let’s add a spurious OVER session between R3 and R1
Now, R3 learns P via two signaling paths

OVER-RIDE GADGET
R3 BGP Decision Process is used to select one of them

BGP Decision Process

1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP
BGP Decision Process

1. Higher Local-preference  
2. Shorter AS-Path  
3. Lower Origin  
4. Lower MED  
5. Prefer eBGP over iBGP  
6. Lower IGP metric to NH  
7. Lower Router ID  
8. Shorter cluster-list  
9. Lower neighbor IP
BGP Decision Process

1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP

OVER-RIDE GADGET
BGP Decision Process
1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP
**BGP Decision Process**

1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. **Lower MED**
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP

---

OVER-RIDE GADGET
1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP

BGP Decision Process

(R3 R1) (R3 R2 R1)
OVER-RIDE GADGET

BGP Decision Process
1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP
BGP Decision Process

1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP
(R3 R1) wins since it has no cluster-list

BGP Decision Process
1. Higher Local-preference
2. Shorter AS-Path
3. Lower Origin
4. Lower MED
5. Prefer eBGP over iBGP
6. Lower IGP metric to NH
7. Lower Router ID
8. Shorter cluster-list
9. Lower neighbor IP

OVER-RIDE GADGET
Due to BGP Propagation rules, R3 does not announce the route to R4 anymore.

**BGP Propagation rules**

<table>
<thead>
<tr>
<th>From client</th>
<th>To client</th>
<th>To peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>
R4 does not receive any route for P

OVER-RIDE GADGET
Either R4 does not learn a less-specific route and a forwarding blackhole is created.
Or R4 might use a less specific route which can create forwarding deflections and loops.

Traffic to P

> 0.0.0.0 via R3

OVER-RIDE GADGET
Although uncommon, spurious OVER might appear in real world network

Spurious OVERs

- have been found in real network  
  [Feamster05, Park11]
- act as an easy-visibility fix  
  [Pelsser08, Pelsser10]
- could appear during reconfiguration  
  [Herrero10]
A spurious OVER is an easy and tempting solution to solve route visibility issue

Although preferred, R4 does not receive P1 since R3 prefers P2 (IGP cost), leading to suboptimal routing
A spurious OVER is an easy and tempting solution to solve route visibility issue

Adding a spurious OVER, improves R3’s visibility

[Pelsser08, Pelsser10]
Spurious OVER are likely to appear during iBGP reconfiguration

Best practices: Introduce UPs before tearing OVERs down [Herrero10]
Spurious OVER are likely to appear during iBGP reconfiguration

Best practices: Introduce UP before tearing OVER down [Herrero10]  ➤ potentially spurious OVERs during the process

going from

[Diagram of network connections showing R1, R2, R3, and R4 with arrows indicating connections and potential issues during configuration.]
iBGP Deceptions: More Sessions, Fewer Routes

Introduction and Motivation

Dissemination correctness

Revisiting the state-of-the-art

Conclusion
Route reflection is prone to both routing and forwarding anomalies

An iBGP configuration is correct if it respects the following two properties [Griffin02]:

- **signaling correctness**
  BGP will always converge to a stable, unique routing state

- **forwarding correctness**
  No forwarding deflection arises along any BGP forwarding path
One property is missing: 

*dissemination correctness*

An iBGP configuration is correct if it respects the following two properties [Griffin02]:

- **signaling correctness**
  BGP will always converge to a stable, unique routing state

- **forwarding correctness**
  No forwarding deflection arises along any BGP forwarding path
Dissemination correctness deals with issues in the route propagation process

An iBGP configuration is correct if it respects the following three properties:

- **signaling correctness**
  BGP will always converge to a stable, unique routing state

- **forwarding correctness**
  Absence of deflection along any BGP forwarding path

- **dissemination correctness**
  all BGP routers are guaranteed to receive a route to all prefixes
Signaling, dissemination and forwarding correctness complement each other

- Signaling correct does not imply dissemination correct
Signaling, dissemination and forwarding correctness complement each other

- Signaling correct does not imply dissemination correct

Example of iBGP topology which is *signaling* correct, but not *dissemination* correct
Signaling, dissemination and forwarding correctness complement each other

- Signaling correct does not imply dissemination correct
- Dissemination correct does not imply forwarding correct

Example of iBGP topology which is dissemination correct, but not forwarding correct

[Griffin, SIGCOMM02]
Dealing with dissemination correctness is computationally hard

Dissemination Correctness Problem (DCP):

Given a signaling correct iBGP topology $B$ and the underlying IGP topology $I$,

Decide if $B$ is dissemination correct

DCP is coNP-hard
P-time reduction from 3-SAT complement
Prior knowledge of correctness is useless

One More Session Problem (OMSP):

Given a dissemination correct iBGP topology $B$, and the underlying IGP topology $I$,

Decide if adding a spurious OVER session to $B$ will result in a dissemination correct topology

OMSP is coNP-hard
P-time reduction from 3-SAT complement
There exist sufficient conditions that guarantee *dissemination correctness*

Either of the following conditions guarantees a signaling correct iBGP topology to be dissemination correct

- **prefer-client**
  All iBGP routers strictly prefer client routes

- **no-spurious-OVER**
  The iBGP topology contains no spurious OVERs
iBGP Deceptions: More Sessions, Fewer Routes

Introduction and Motivation

Dissemination correctness

Revisiting the state-of-the-art

Conclusion
Some results already encompass dissemination correctness

Sufficient conditions guaranteeing signaling, forwarding correctness

On the correctness of IBGP configuration

[Griffin, SIGCOMM02]
Some results already encompass dissemination correctness

Sufficient conditions guaranteeing signaling, forwarding correctness

On the correctness of IBGP configuration

1. $B$ has no cycles of UP sessions only
2. Route-reflector prefers paths propagated by clients
3. All-shortest-paths must also be valid signaling paths

implies dissemination correctness
Dissemination is often overlooked

Relaxed sufficient conditions for signaling or forwarding correctness

Preventing persistent oscillations and loops in IBGP configuration with route reflection

Checking for optimal egress points in iBGP routing

Such conditions do not imply dissemination correctness (e.g. OVER-RIDE gadget)

[Rawat, Comput.Netw.06]

[Buob, DRCN07]

[Buob, Networking08]
Dissemination is often overlooked

Guarantee iBGP convergence by modifying the decision process

Stable and flexible iBGP

Modified iBGP does not guarantee dissemination (e.g., OVER-RIDE gadget)
Improving route diversity through the design of iBGP topologies

Providing scalable NH-diverse iBGP route redistribution to achieve sub-second switch-over time

[Pelsser, ICC08]

[Pelsser, Comput. Netw.10]

Dissemination is often overlooked

Improve route diversity by adding spurious OVERs

adding spurious OVERs increase the diversity only \textit{locally}, but may worsen it \textit{globally}
Dissemination is often overlooked

iBGP topology design guidelines

How to Construct a Correct and Scalable iBGP Configuration

[Vutukuru, INFOCOM06]

Lemma 3

“If there exists a signaling chain between routers A and B [...] then A learns of the best route via B [...]”

Not true in presence of spurious OVERs

Having a valid signaling path is necessary, not sufficient
Summary of our contributions

In this work, we

- showed that iBGP Propagation rules play a big role in iBGP
- introduced dissemination correctness
  - studied its complexity
  - provided sufficient conditions and guidelines to enforce it
- showed that dissemination is often overlooked
iBGP Deceptions: More Sessions, Fewer Routes

Introduction and Motivation

Dissemination correctness

Revisiting the state-of-the-art

Conclusion
iBGP semantic is more complex than what is commonly assumed

- Valid signaling path is not a good abstraction to study route propagation

- Spurious OVERs invalidate assumptions that apparently hold in any iBGP topology

- Dissemination correctness provides new motivations for decoupling route propagation from route selection
iBGP Deceptions: More Sessions, Fewer Routes

Laurent Vanbever
UCLouvain, Louvain-la-Neuve, Belgium
laurent.vanbever@uclouvain.be