Verifying Wide-Area Routing Configuration

Nick Feamster and Hari Balakrishnan
M.I.T. Computer Science and Artificial Intelligence Laboratory
{feamster,hari}@csail.mit.edu

http://nms.lcs.mit.edu/bgp/
BGP Configuration Affects Correctness

- BGP has serious problems
  - Frequently misconfigured [Mahajan2002]
  - Forwarding loops [Dube1999]
  - Persistent route oscillation [Griffin1999, Varadhan2000]
  - Slow convergence/suppressed routes [Labovitz2001, Mao2002]

**BGP’s configuration determines whether the protocol behaves correctly or not.**

*These problems never happen in the "real world", right?*
"A number of...customers went out from 5pm today due to, supposedly, a DDoS (distributed denial of service attack) on a key...data center, which later was described as a route leak (misconfiguration)."

-- dslreports.com
Reported problems:

<table>
<thead>
<tr>
<th>Property</th>
<th>1994-1999</th>
<th>2000-2004</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filtering</td>
<td>42 (64)</td>
<td>56 (109)</td>
<td>98 (173)</td>
</tr>
<tr>
<td>Leaked Routes</td>
<td>23 (25)</td>
<td>41 (42)</td>
<td>64 (67)</td>
</tr>
<tr>
<td>Hijacked Routes</td>
<td>14 (14)</td>
<td>9 (10)</td>
<td>23 (24)</td>
</tr>
<tr>
<td>Global Route Visibility</td>
<td>60 (80)</td>
<td>82 (117)</td>
<td>142 (197)</td>
</tr>
<tr>
<td>Oscillations</td>
<td>0 (0)</td>
<td>0 (4)</td>
<td>0 (4)</td>
</tr>
<tr>
<td>Routing Instability</td>
<td>38 (45)</td>
<td>38 (48)</td>
<td>76 (93)</td>
</tr>
<tr>
<td>Attribute manip.</td>
<td>19 (23)</td>
<td>12 (29)</td>
<td>31 (52)</td>
</tr>
<tr>
<td>iBGP-related</td>
<td>21 (27)</td>
<td>20 (32)</td>
<td>41 (59)</td>
</tr>
<tr>
<td>Routing Loops</td>
<td>11 (11)</td>
<td>13 (17)</td>
<td>24 (28)</td>
</tr>
<tr>
<td>Blackholes</td>
<td>13 (13)</td>
<td>104 (108)</td>
<td>117 (121)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>241 (302)</strong></td>
<td><strong>375 (516)</strong></td>
<td><strong>616 (818)</strong></td>
</tr>
</tbody>
</table>

These problems haven’t gone away.
Some Empirical Evidence: Bogon Route Leaks

BGP route advertisements from July 2003 to May 2004; 8 vantage points.
Possible Remedies

- Protocol is buggy. *Replace.*
  - What to fix?
  - "BGPv5" would have to be as flexible as BGPv4.
  - Will it be any less error-prone?

- Configuration language is too "low-level". *Redesign.*
  - Again, what are the flaws in today’s configuration languages?

*We must understand the problems in BGPv4 before proposing reasonable fixes.*
Approach: Study Today’s Configurations

- Develop a tool that uses static analysis to analyze router configurations.

  - **Operators** can make BGPv4 less error-prone.
    - Find configuration problems before deployment.

  - **Researchers** can learn from the errors we find in today’s configurations.

http://nms.lcs.mit.edu/bgp/rcc/
**rcc Overview**

- **Preprocessor**
- **Parser**
- **Verifier**

*Router Configurations*

*(offline, collected from routers)*

- **Expand macros**
- **Parse configs into intermediate format (mySQL)**
  - Parser reads: Cisco, Juniper, Procket, Zebra/Quagga, Quarry
- **Query intermediate format**

*Extensible design.*
rcc Overview

rcc Error Summary

Network Advertisement
Determinism
iBGP Signaling
Filtering
Parse Errors
Loopback Configuration

network statements without routes
deterministic-med, router ID tiebreak
Possible iBGP partitions
Filtering of bogons and private ASes
Undefined route maps, access-lists, etc.
Duplicate loopbacks, dangling sessions
rcc Overview

rcc Error Summary

Network Advertisement
Determinism
iBGP Signaling

network statements without routes
deterministic-med, router ID tiebreak
Possible iBGP partitions

ERROR: iBGP signaling partition -- rtr-b is missing session to rtr-c
ERROR: iBGP signaling partition -- rtr-c is missing session to rtr-b
rtr-b can't reach (rtr-c)
rtr-c can't reach (rtr-b)
Outline

- Design and implementation of rcc (a.k.a. "RoLex").
  - Correctness definition
  - Description of tests

- Study of configuration errors from 9 ASes.

- Recommended protocol and language changes.

- Appeal for cooperation and feedback.
  - Run rcc on your configurations.
  - Let us know what you find.
  - Suggest new tests and features.
Properties: The Routing Logic

- **Validity**: Does it advertise invalid routes?
  - Bogus route origination, persistent forwarding loops, etc.

- **Visibility**: Does every valid path have a route?
  - Session resets, missing sessions, damped routes, etc.

- **Information-flow control**: Expose information?
  - Accidental route leaks to neighbors, etc.

- **Determinism**: Answer depend on orderings, etc.?
  - Irrelevant route alternatives can affect outcomes.

- **Safety**: Will it converge to a unique, stable answer?
  - Policy-induced oscillation
Applying Correctness Definitions to BGP

1. **Origination:** A router "originates" a route.
2. **Export:** Router advertises route to other routers.
3. **Import:** Other routers learn those routes.
4. **Selection:** Each router selects a single best route.
5. **Modification:** Router modifies attributes.
6. **Propagation:** Propagates route within the AS.
# Putting it together

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Origination</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Export</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>3. Import</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>4. Selection</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>5. Modification</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>6. Intra-AS Prop.</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
</tr>
</tbody>
</table>

- Determine which aspects of correctness apply at each stage of BGP’s operation.
- Express constraints.
- Try to test constraints with static analysis.
rcc Tests: Validity

- Incorrect Origin AS (Origination)
  - Do filters prevent bogon prefixes from being advertised?

- Incorrect AS Path (Export)
  - Mismatches between origin AS and outbound path prependings?
  - Remove private ASes from customers with private sessions?

- Incorrect or Missing Filters (Export/Import)
  - Sessions with no route maps?
  - Route maps with undefined filter-lists, distribute lists, AS path lists, or community lists?

- Incorrect "next-hop" attribute (Modification)
  - Is next-hop-self used when eBGP endpoints are not in the IGR?
rcc Tests: Visibility

- Failure to install valid routes (*Import*)
  - Is synchronization disabled?

- Failure to advertise valid routes (*Export*)
  - Are there "network" statements without routes?
  - [Are filters outdated?]

- iBGP Signaling (*Intra-AS Propagation*)
  - Are there routers with duplicate cluster-ids or loopbacks?
  - Is there an iBGP partition? (*How do we check this?*)
Visibility: iBGP Signaling Overview

- Default: don’t readvertise iBGP-learned routes
  - Complete propagation requires "full-mesh" iBGP.
  - Doesn’t scale.

- "Route reflection" improves scaling (RFC 2796)
  - **Client:** re-advertise as usual
  - **Route reflector:** reflect non-client routes to all clients, client routes to non-clients and other clients.
Visibility: iBGP Signaling
Visibility: iBGP Signaling

Route Reflector (RR)

W

Client

X

Y

Z

route $r_2$ to $d$
Visibility: iBGP Signaling

Route Reflector (RR)

route $r_1$ to $d$

route $r_2$ to $d$

iBGP Signaling Partition!
Visibility: iBGP Signaling

Theorem. (Not Scary)

Suppose the iBGP reflector-client relationship graph contains no cycles.

Then, the AS’s configuration satisfies visibility if, and only if, the set of routers that are not route reflector clients forms a full mesh.

Condition is easy to check with static analysis.
Summary of Errors Discovered in 9 ASes

● Serious Errors (1st Class)
  ▶ Incorrect or missing filters (~ 50 sessions)
  ▶ iBGP signaling partitions (10 instances)
  ▶ Unintentional transit (3 instances)

● Annoyances (2nd Class)
  ▶ Inconsistent export (3 instances)
  ▶ Nondeterministic settings (34 routers)
  ▶ Failure to install valid routes (3 routers)

● Cleanup (3rd Class)
  ▶ Sessions with undefined policies (2 sessions)
  ▶ Policies with undefined distribute lists, etc. (30 policies)
  ▶ Incomplete iBGP sessions (76 sessions)
Summary of Errors
Why are errors happening, and what to do?

- Ad hoc process, intrinsic vulnerabilities
  - *Example:* Filtering is rarely (if ever) done correctly. *(ask me for a copy of recent analysis of bogon advertisements)*
  - *Solution:* Automation; build validity into BGP (e.g., S-BGP).

- Obscure mechanisms
  - *Example:* iBGP signaling partitions
  - *Solution:* Redesign intra-AS route propagation *(ask me for a copy of my proposal)*

- Indirect specification
  - *Example:* Incorrect implementation of information flow policies
  - *Solution:* Better configuration languages
Conclusion

• Our contributions:
  ▶ Correctness constraints for configuration.
  ▶ Design and implementation of rcc.
  ▶ Study of configuration errors in real-world networks.
  ▶ Recommended protocol and language changes.


• rcc is available.
  ▶ More than 30 operators have downloaded the tool.
  ▶ Tested configurations of 9 ASes.

http://nms.lcs.mit.edu/bgp/rcc/
Thanks: Bug fixes, Suggestions, etc.

- Tom Barron
- Rob Beverly
- Randy Bush
- Michael Hallgren
- John Heasley
- Simon Leinen
- Hank Nussbacher
- Michael O’Neill
- Scott Poretsky
- Jennifer Rexford
- Nicolas Strina
The ultimate goal: rcc should be useful to you.

Download rcc.
Report bugs in your configurations.
Report bugs in rcc.
Request new tests.

From this talk alone, what tests are missing from rcc that I should definitely add?

Ideas: IPv6 support, checks against RIR, BGP/MPLS VPNs, etc.

Feel free to help develop, too. :)