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]
  - Useless routing messages [Labovitz1999, Wang2002]
  - Security weaknesses [Beard2002, Kent2000]

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

These problems never happen in the "real world", right?

Monday, February 23, 2004

> **"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**

10 Years of NANOG...

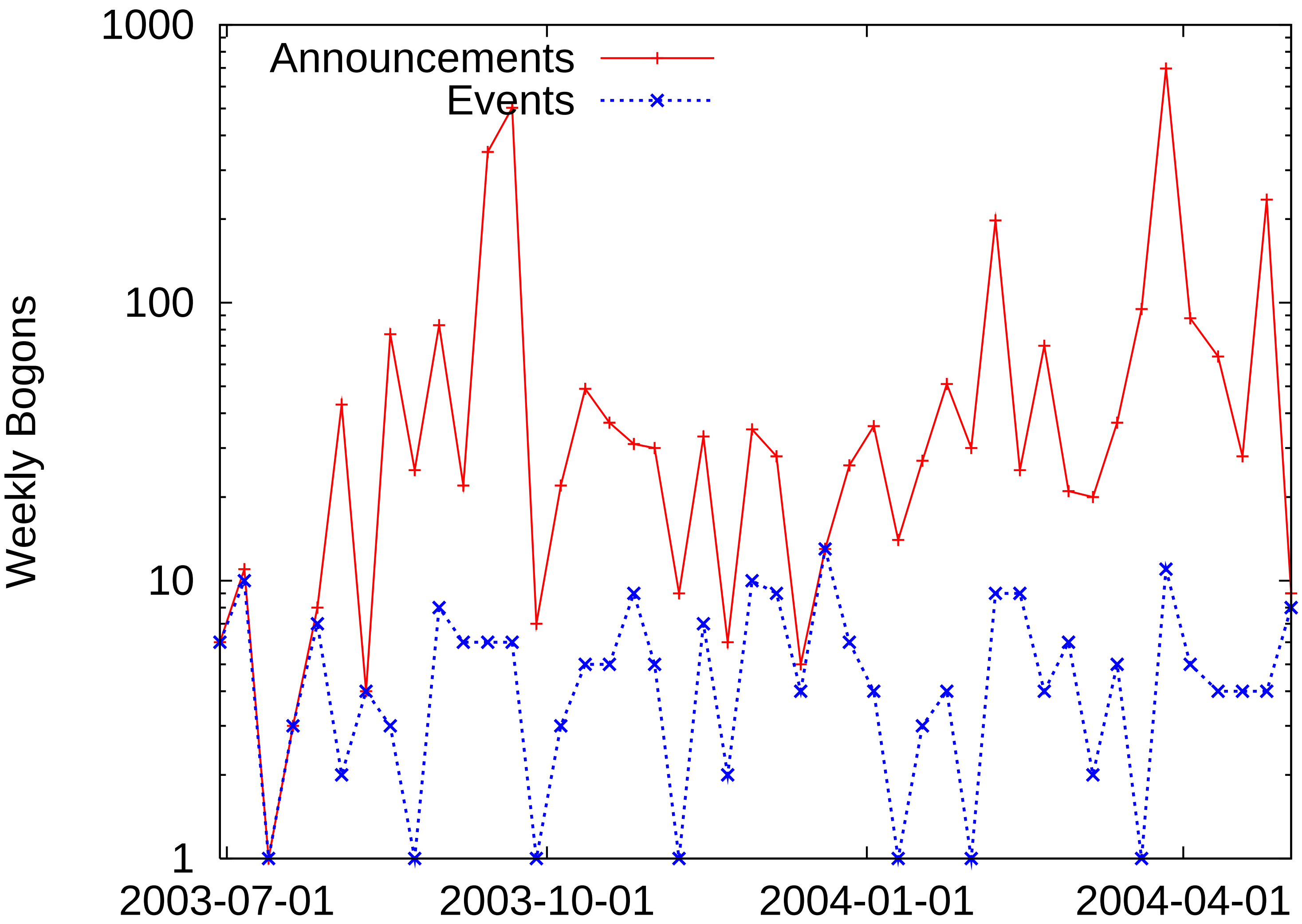
Reported problems:

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

These problems haven't gone away.

Some Empirical Evidence: Bogon Route Leaks

BGP route advertisements from July 2003 to May 2004; 8 vantage points.
(http://bgp.lcs.mit.edu/bogons.cgi)



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

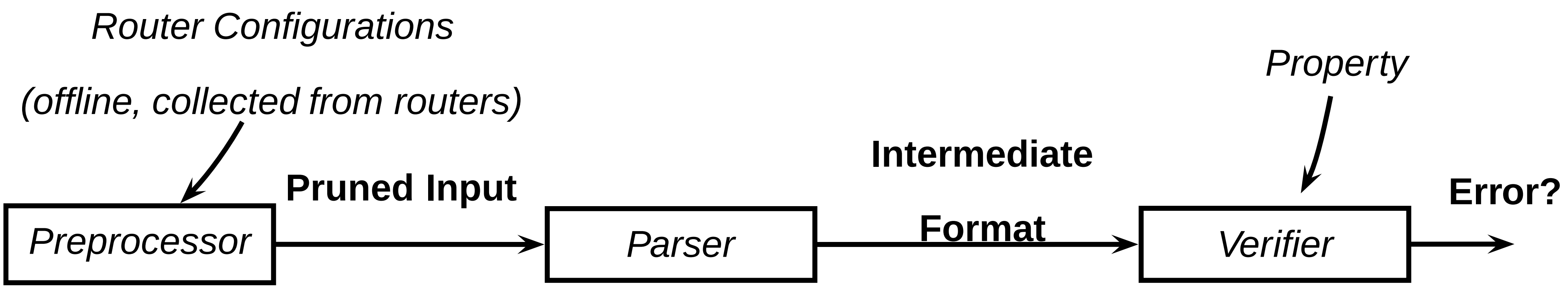
- Develop a tool that uses static analysis to analyze router configurations.

- 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

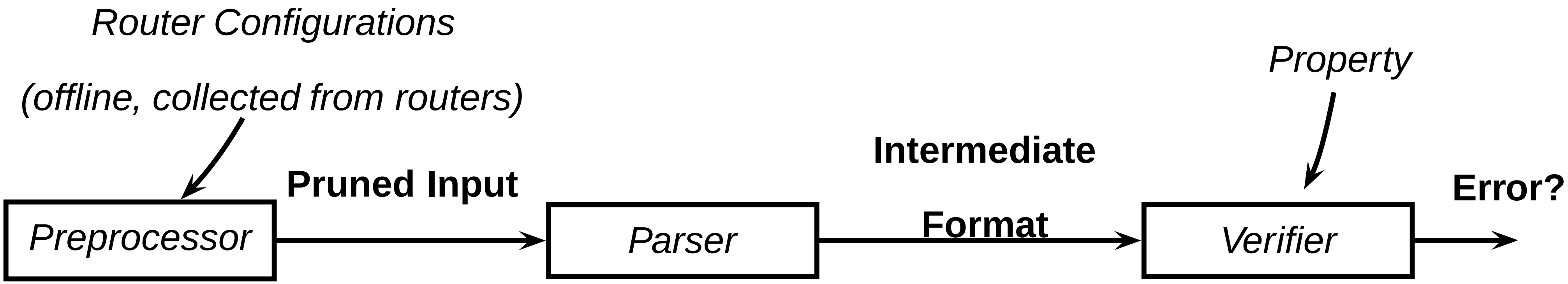


- Expand macros
- Parse configs into intermediate format (mySQL)
  - Parser reads: Cisco, Juniper, Procket, Zebra/Quagga, Quarry
- Query intermediate format

Extensible design.

All configurations depicted, ASes used, and incidents portrayed in this demo are fictitious. No identification with actual IP addresses or routing policies is intended nor should be inferred. Any resemblance of the configuration portrayed to actual ASes, living or dead, is purely coincidental.

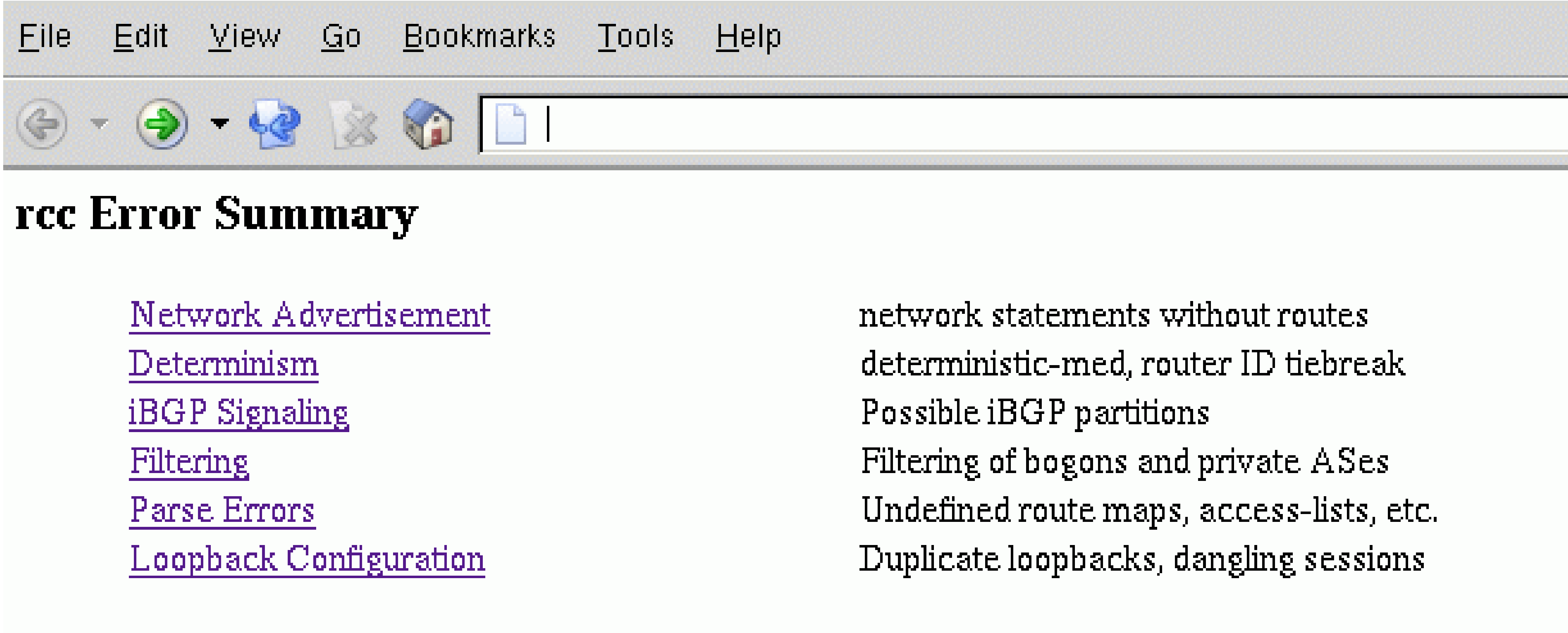
rcc Overview



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

Extensible design.

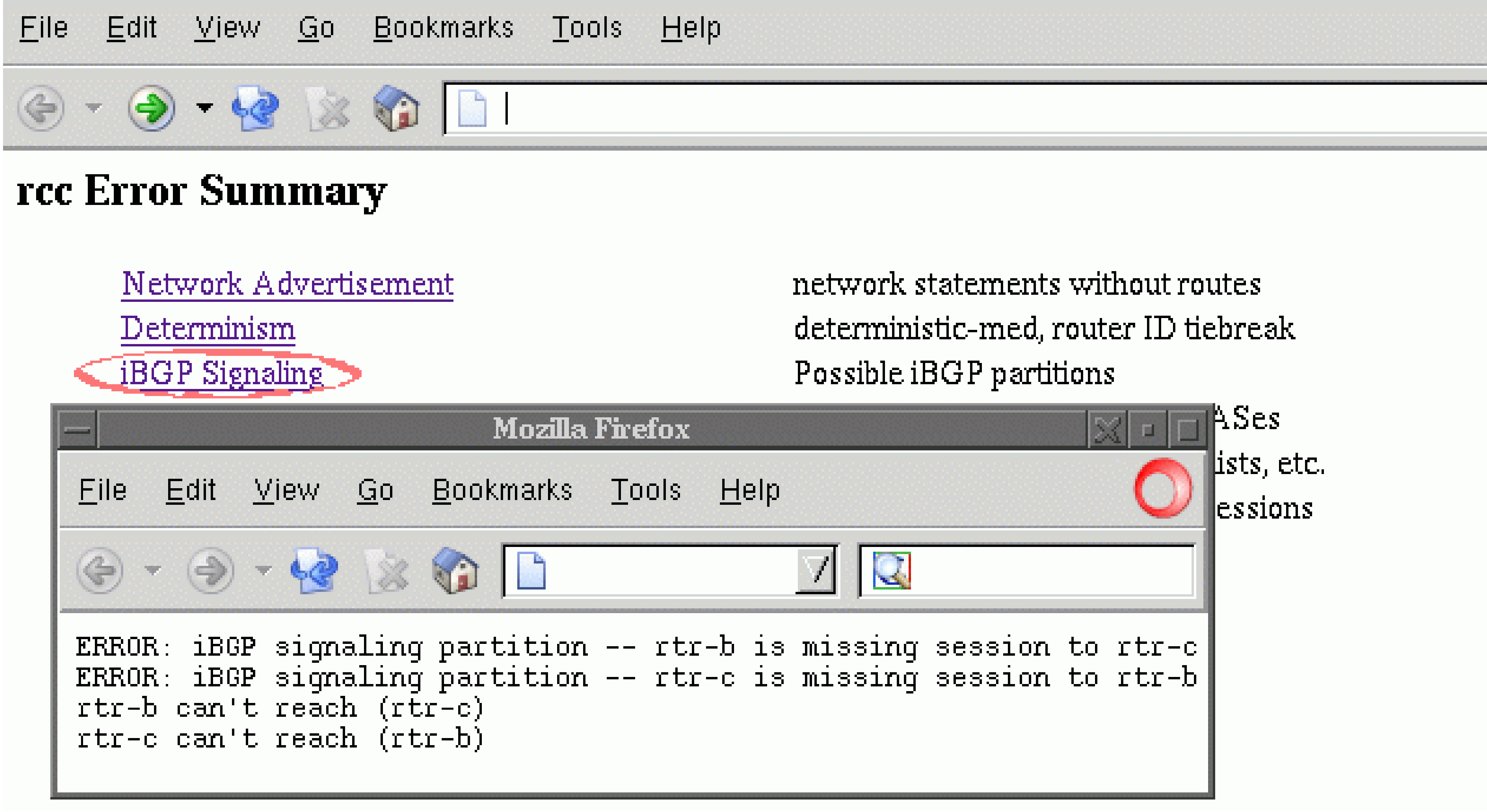
rcc Overview



rcc Error Summary

Network Advertisement	network statements without routes
Determinism	deterministic-med, router ID tiebreak
iBGP Signaling	Possible iBGP partitions
Filtering	Filtering of bogons and private ASes
Parse Errors	Undefined route maps, access-lists, etc.
Loopback Configuration	Duplicate loopbacks, dangling sessions

rcc Overview


File Edit View Go Bookmarks Tools Help
Mozilla Firefox
File Edit View Go Bookmarks Tools Help

rcc Error Summary

Network Advertisement	network statements without routes
Determinism	deterministic-med, router ID tiebreak
iBGP Signaling	Possible iBGP partitions
	ASes
	ists, etc.
	essions

```
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

Step	Valid.	Visib.	Info Flow.	Det.	Safety
1. Origination	●				
2. Export	●		●		
3. Import	●	●	●		
4. Selection				●	●
5. Modification	●		●		
6. Intra-AS Prop.	●	●		●	

- 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 prepending?*
  - *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?*

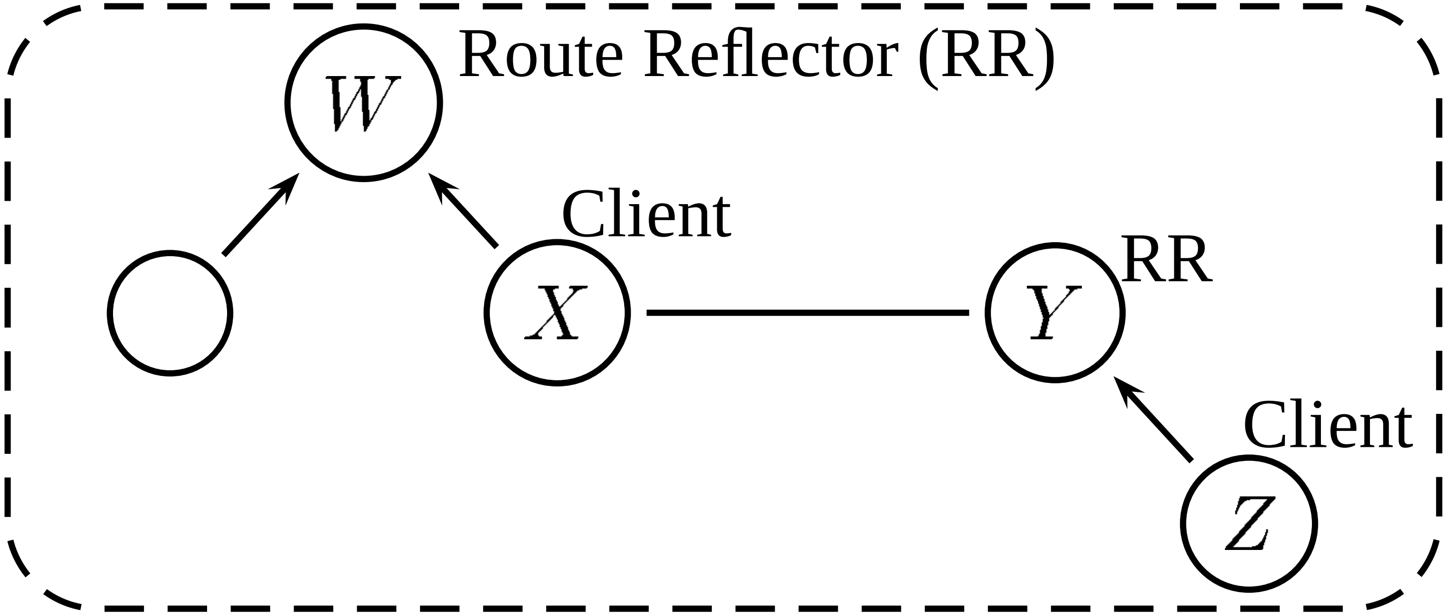
rcc Tests: Visibility

- 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?)**

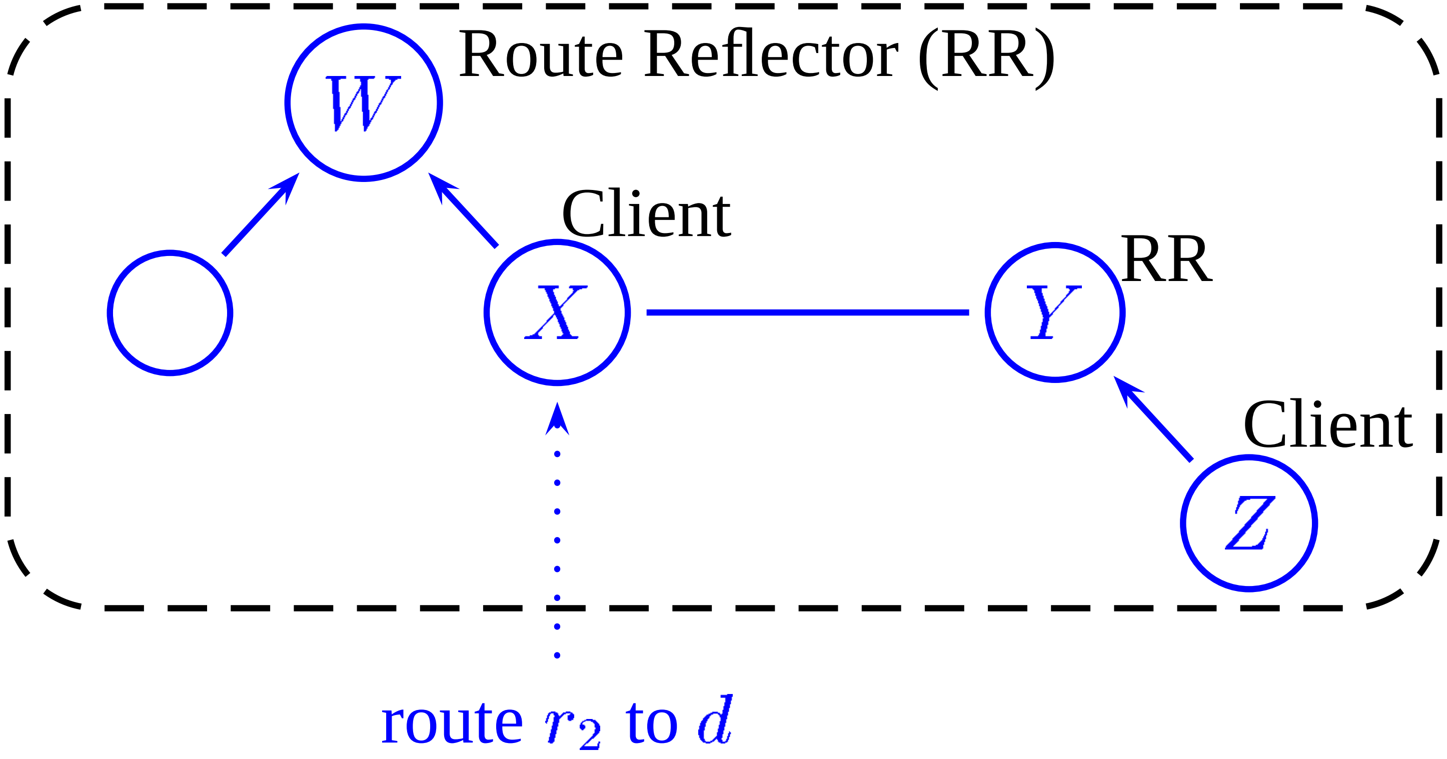
Visibility: iBGP Signaling Overview

- 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.

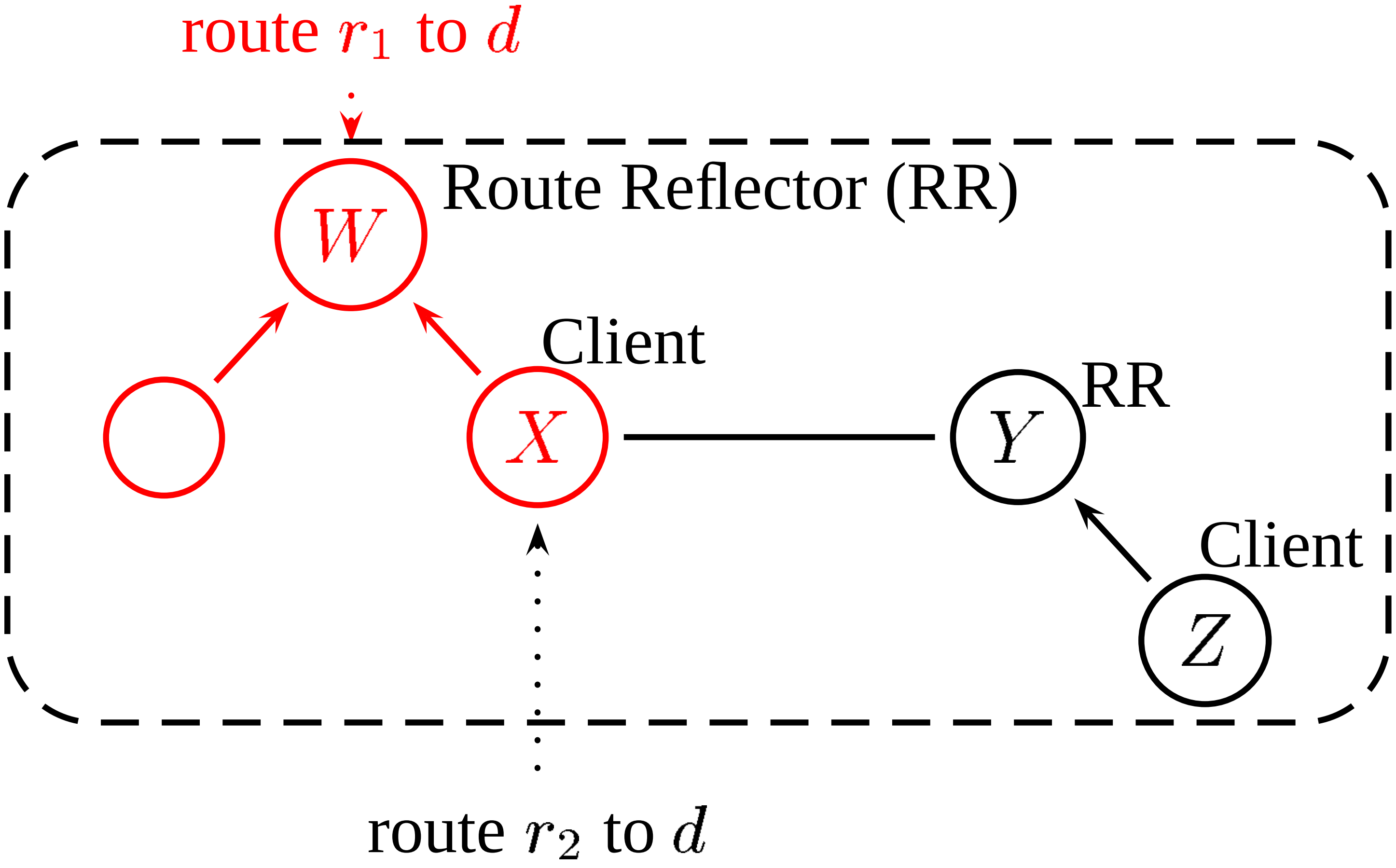
Visibility: iBGP Signaling



Visibility: iBGP Signaling



Visibility: iBGP Signaling



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.

rcc Tests: Information-flow Control

Verification requires a specification of intended policy.
(We don’t have this today, but we can make reasonable assumptions.)

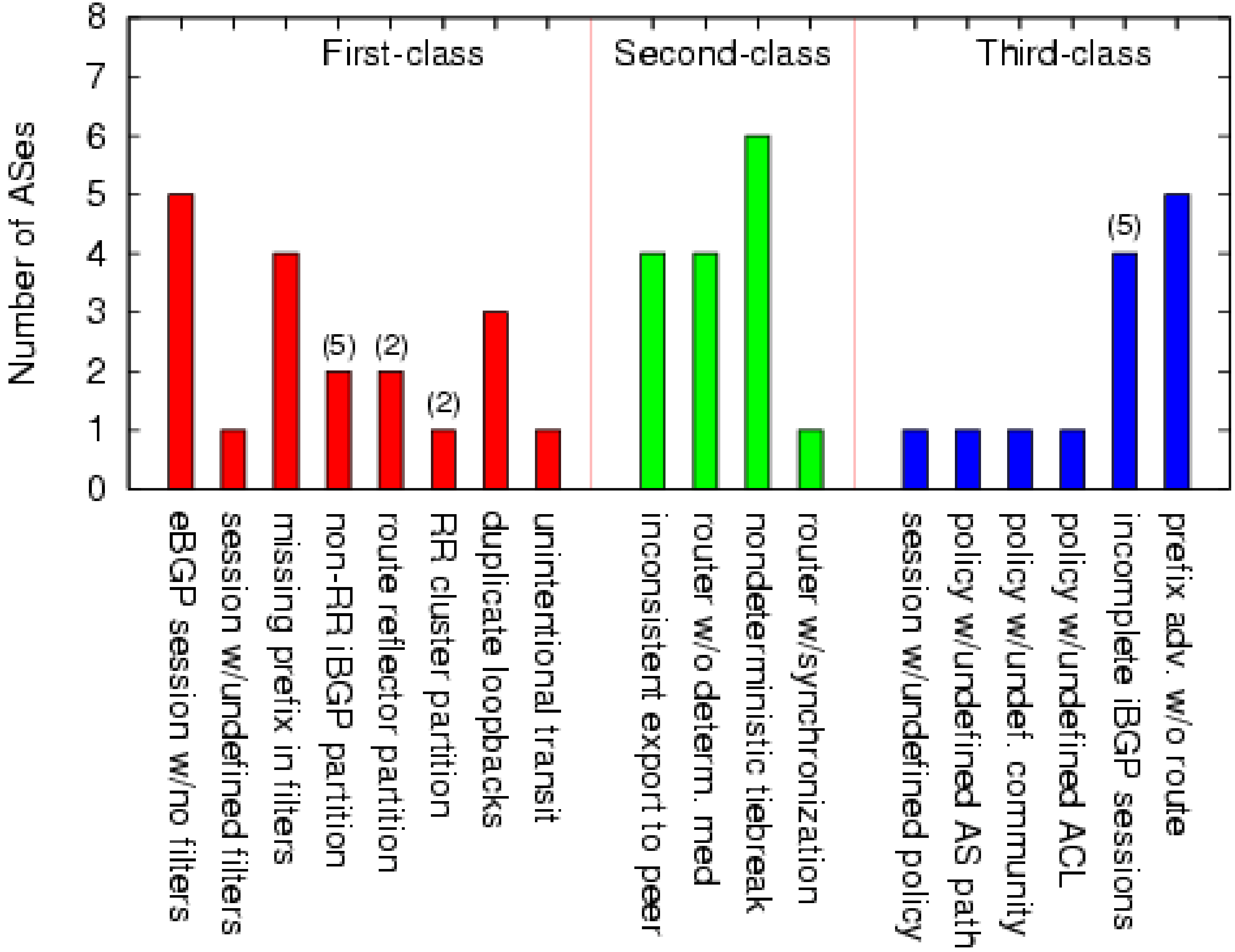
- Controlled export *(Export)*
  - *Unintentionally advertising routes between peers?*
- Consistent export *(Export)*
  - *Unintentionally forcing a peer to "cold potato"?*
- Consistent import *(Import)*
  - *Unintentionally forcing "cold potato" on your own network?*

These conditions are difficult to "eyeball" in practice,
but 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.

http://nms.lcs.mit.edu/papers/rcc-tr.pdf

- **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

Request for feedback

- 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.
    - http://nms.lcs.mit.edu/bgp/rcc/feedback.cgi

- Feel free to help develop, too. :)

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