Detecting inconsistencies in INRDB data

to identify MOAS cases and possible illegitimate Internet resource usage

Peter Ruissen

System and Network Engineering

University of Amsterdam

December 11, 2007

1. Problem: Prefix/ASN Hijacking
2. Research
   - Cryptographic solutions or not?
3. RIPE NCC Data Sources
   - INRDB Prototype
   - Inconsistencies
4. The algorithm: constructing unique trees
5. Results
6. Comments on Results
7. Future work and conclusions
8. Time for discussion (and questions)

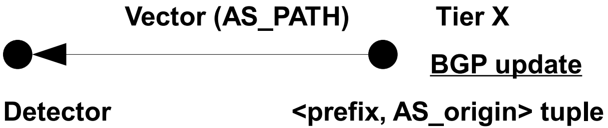
REFERENCES16

Problem: Illegitimate number resource usage (Prefix/ASN)

Properties of illegitimate resource usage

- Prefix hijacking: using someone elses prefix without permission: Possibility: (blackholing, DOS, deception)

  BGP is a distance vector type protocol build on *paths* of trust, without authentication verification by default



- Multiple Origin AS (MOAS) conflicts occur when a prefix originates from more than one AS
- Summary of related work: Hijacked prefixes are mostly dynamic, stealthy and do not correlate with history, have short uptime, are mostly /24 from unaware organizations, can use overlapped sub/supernet address space.

Research question

Properties of illegitimate resource usage

- Short term goal: insight in the current situation,everybody should DEPLOY certification!
- Long term goal: Detect and prevent illegitimate resource usage (detection framework using chosen solution: SBGP, X509 certificates, soBGP).

Research question

How to correlate inconsistencies between INRDB data sources to identify MOAS cases and detect possible illegitimate Internet resource usage?

Next slide: Passive security

Cryptographic solutions or not?

- Ingress access lists, BGP TTL hacks or MD5 hashes are not sufficient
- Secure BGP (S-BGP): X509 (PKI) **centralized** for internet resources. Optional BGP path attribute to carry digital signatures from BGP updates. IPsec to provide data integrity and authenticate BGP routers before exchanging BGP traffic.
- Secure Origin BGP (soBGP) uses **decentralized** Web of Trust model PKI proposed by Cisco.
- Step by step approach: X509 Resource certification is less ambitious without AS-path verification.

This only works if everybody actually DEPLOY this! How do we monitor the current Internet topology? See next slide:

RIPE NCC Data Sources

Routing Information Service (RIS)

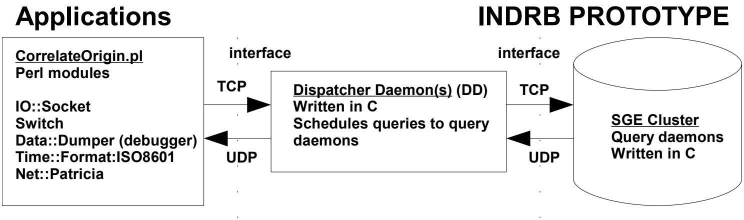
Dynamic databases (similar to Oregon Route Views) provides current view of the Internet by collecting BGP RIB tables. RIPE NCC has 15 remote route collectors (RRC) that peer with 600 collector peers (CP) at various Internet Exchange Points.

RIPE Database (RIPEDB) and RIR Stats

Static databases, like (RADB, RIPEDB) described by (RPSL) contain whois information and policy info. Form the Internet Routing Registry (IRR).

New Internet Number Resources database (INRDB)

INRDB Prototype



- Transparent layer on top of underlying datasources (RIS, RIPEDB, STATS, Reverse DNS lookup)
- Advantages: Fast (in-mem design), Scalability, Historic overview
- Challenges: Different kinds of data, overlap, inconsistencies, quality rating, terabytes of input. More about inconsistencies!: See next slide..

Inconsistencies

- Definition: data that is semantically incorrect, inaccurate or different in comparison with other data.
- IntraDB inconsistencies:overlapping inetnum objects,unreferenced contact info etc, IntraRIS inconsistencies: conflicting origins, overlapping timeframes,overlapping tree, most part: multidimensional
- SCOPE: for now only looking at sample time intervals, number of MOAS, conflicting AS Origins and unregistered prefix usage.

Next slide: The algorithm: constructing unique trees!

The algorithm: constructing unique trees

input prefix list *P*, sampletime list *T*

output percentage of unregistered prefix usage, prefixes with OriginAS not listed in RIPEDB and unique MOAS

difficulties IntraDB/IntraRIS inconsistencies, IntraRIS inconsistencies: conflicting origins, overlapping timeframes,overlapping prefixes, multidimensionality

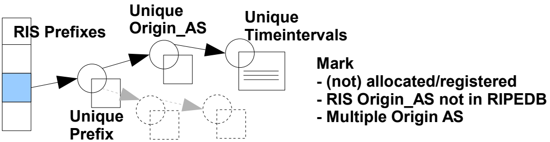


Figure: Sequences of hashes using only hash keys.

Results

Measurements / samples: four year overview of historical RIS data of 62/8

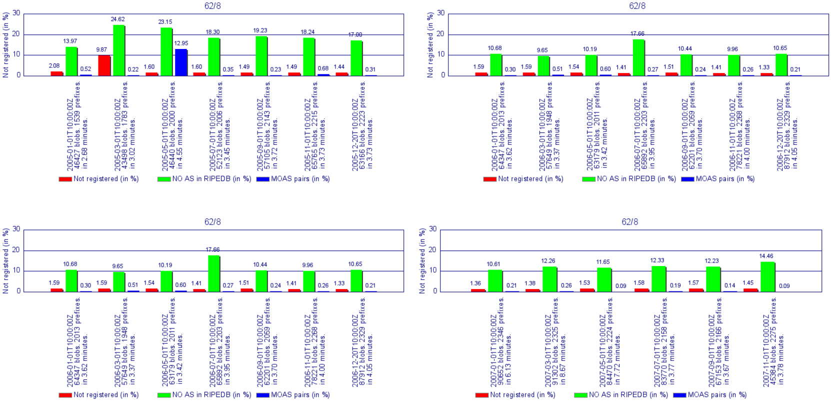


Figure: Sample Overview 62/8 2004-2007 (see report)

Results

Measurements / samples: historical RIS data of 62/8 in 2005

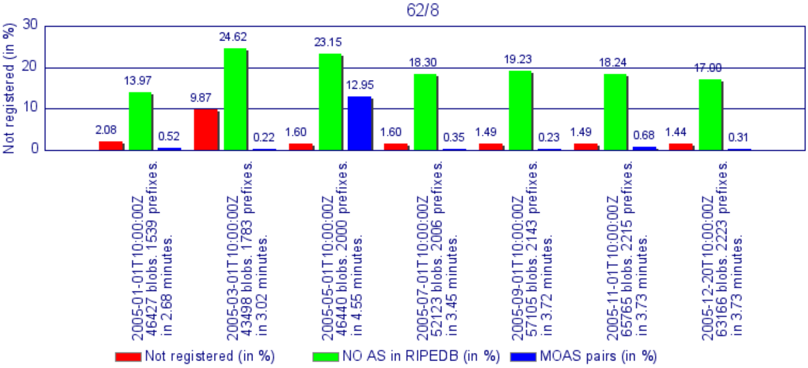


Figure: Other reports [22] show similar MOAS cases during the Google 2005 Outage

Results

Measurements / samples all /8 RIPE NCC

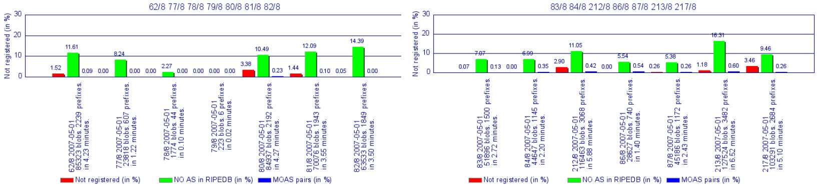


Figure: One hour samples all /8 from RIPE NCC

- legitimate cases of MOAS are anycast addresses, private links, specific cases of multihoming (long uptime)
- MOAS can also be caused by misconfiguration. Repeating MOAS for multiple prefixes by the same AS are suspicious

Comments on Results

four year overview of historical RIS data of 62/8 and one hour samples of all /8

- Anomalies: 62.9.0.0/16 with 839 route objects in RIPEDB, some registrations weirdly listed
- RIS RIB table growth 62/8 from 1495 unique prefixes in 2004 to 2166 unique prefixes in 2007
- Positive result for RIPEDB: 88% of the RIS entries have matching RIPEDB Origins (averaged)
- 80% of MOAS keeps coming back every month, the remaining MOAS are unique and suspicious.
- The MOAS that come back could be anycast addresses, private links, specific cases of multihoming
- 2 % of all unique prefixes are used without being registered including weird cases and bogus addresses.

Future work and conclusions

- Conclusion: correlating these MOAS cases with listings in RIPEDB and registration data is not enough to determine if they are hijacked or not
- check for bogus prefixes and bogus ASN (use filters)
- examining if the resulting 20% repeats MOAS behaviour (repeatedly hijacking more prefixes), filter MOAS on bogons
- Future work: overlap detection(**radix trees**), timeframe processing and resource certification validation.
- all organization(s) should deploy resource certification!

Time for discussion (and questions)

- Questions.

An Analysis of BGP Multiple Origin AS (MOAS) Conflicts
Xiaoliang Zhao, Dan Pei, Lan Wang, Dan Massey, Allison Mankin, S. Felix Wu, Lixia Zhang, 2001
`http://www.imconf.net/imw-2001/imw2001-papers/88.pdf`

Analyzing BGP Policies: Methodology and Tool
Proceedings of IEEE INFOCOM, Hong Kong, China, March 2004.
`http://www.cs.ucr.edu/~siganos/papers/Nemecis.pdf`

A study of prefix hijacking an Interception in the Internet
Hitesh Ballani, Paul Francis, Xinyang Zhang, Cornell University, 2007
`http://www.cs.cornell.edu/People/francis/sigcomm07-interception.pdf`

Analysis of BGP Prefix Origins During Googles May 2005 Outage
Tao Wan Paul, C. van Oorschot, Carleton University, 2005
`http://www.scs.carleton.ca/~paulv/papers/ssn06-fine.pdf`

Beware of BGP Attacks)
Ola Nordstrom, Constantinos Dovrolis, College of Computing
`http://www.cc.gatech.edu/~dovrolis/Papers/ccr-bgp.pdf`

Detecting Bogus BGP Route Information: Going Beyond Prefix Hijacking
Jian Qiu, Lixin Gao *et al*, Department of ECE, Univ. of Massachusetts, 2007
`http://www.ece.rice.edu/~sranjan/publications/securecomm07-hijacking.pdf`

How prevalent is prefix hijacking on the internet?
Peter Boothe, James Hiebert, Randy Bush
`http://rip.psg.com/~randy/030603.nanog-sxbgp.pdf`

RIPE NCC Science Group
`http://www.ripe.net/info/ncc/staff/science_grp.html`

More see report `http://staff.science.uva.nl/~delaat/sne-2007-2008/p02/report.pdf`