

How Complete and Accurate is the Internet Routing Registry (IRR)?

Dec 5th 2011

4th CAIDA-WIDE-CASFI Joint Measurement Workshop

Akmal Khan, Hyun-chul Kim, Ted "Taekyoung" Kwon

Seoul National University

raoakhan@mmlab.snu.ac.kr

Internet Routing Registry (IRR)

- ▶ Globally distributed repository of routing policy databases,
 - ▶ Established in 1995, <http://www.irr.net>, with the purpose to
 - ▶ ensure stability and consistency of global Internet routing [RFC 2622]
 - ▶ troubleshoot routing problems, look up peering agreements [RFC 2650]
 - ▶ automatically configure backbone routers [RFC 2650]
- ▶ 36 Routing Registries (RRs) across the Internet (Jan. 1st 2011)
 - ▶ Operated by organizations such as
 - ▶ Regional Internet Registries (RIRs): RIPE NCC, ARIN, APNIC
 - ▶ ISPs: NTT, Level3, SAVVIS, etc .
 - ▶ Commercial Service : RADb

IRR : a reliable reference ?

- ▶ Common claims (with little or no empirical basis.)
 - ▶ **IRR may not be complete** [Butler@IEEE Comm. Surveys'10]
 - ▶ **IRR may not be accurate** [Nanog'08 , Arbor'09]
 - ▶ Lack of incentives for ASs to maintain up-to-date routing policies
 - ▶ Error-prone manual data entry
 - ▶ **Some RRs can be more complete/accurate than others** [ENISA'10]

How true are these negative claims ?

- ▶ **As they have resulted in**
 - ▶ Little or Limited usage by the research/operational community
- ▶ **While proposed applications highlight its importance to**
 - ▶ help mitigate the BGP robustness problem [Siganos@INFOCOM07]
 - ▶ extract AS topology information [He@ToN09]
 - ▶ analyze business relationship between ASes [Siganos@INFOCOM04]
 - ▶ generate AS-to-organization mapping [Cai@IMCI0]
- ▶ **per-AS Completeness/Accuracy**
 - ▶ How many ASes have started using the IRR?

Key Questions

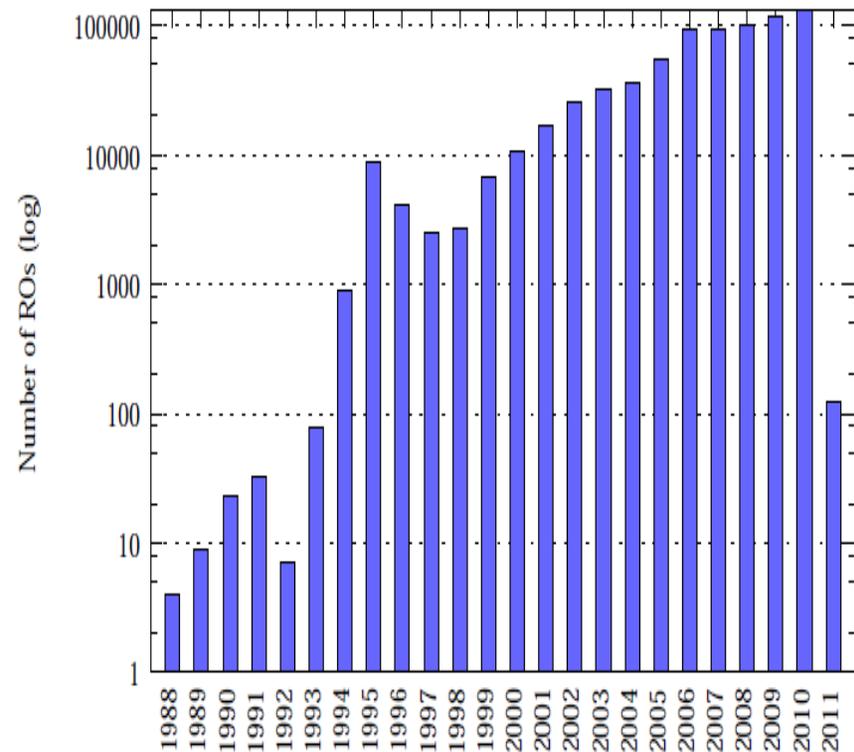
- ▶ **How Complete is the IRR?**
 - ▶ How many prefixes in **IRR** match with **BGP**?
 - ▶ How many prefixes in **BGP** match with **IRR**?
 - ▶ per-AS Completeness

- ▶ **How Accurate is the IRR?**
 - ▶ How many prefixes + Origin AS in **IRR** match with **BGP**?
 - ▶ How many prefixes + Origin AS in **BGP** match with **IRR**?
 - ▶ per-AS Accuracy

Methodology

Establish a reference point to evaluate IRR Route Objects (ROs)

- ▶ IRR data [1st Jan'11]
 - ▶ 736 K ROs
- ▶ BGP traces from UCLA [Oct'08~ 1st Jan'11]
 - ▶ after cleaning 857 K BGP announcements
- ▶ AS links data
 - ▶ UCLA BGP, CAIDA AS relationship
 - ▶ DIMES , Ono, IXP Mapping project
 - ▶ IRR
- ▶ RIR prefix allocation records
- ▶ UCLA AS types (stub, small ISPs, large ISPs, Tier1s)

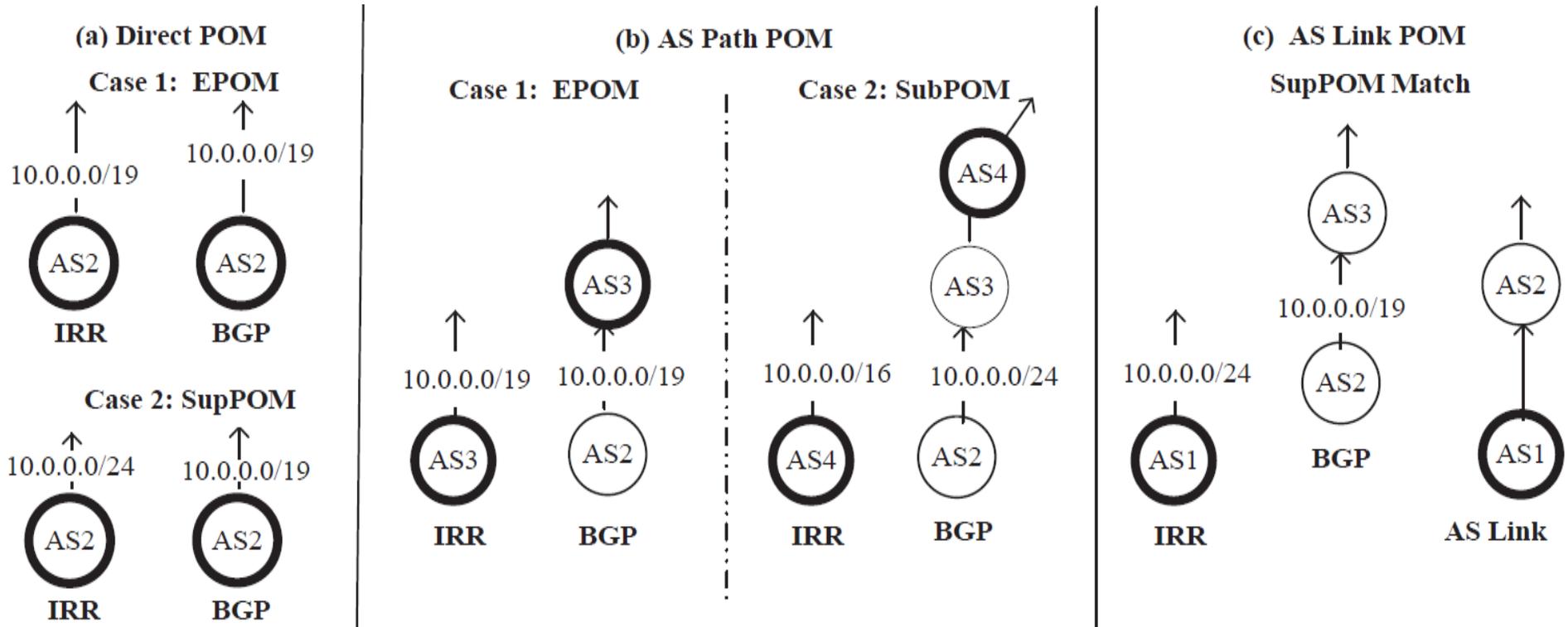


IRR ROs based on latest update date

AS link datasets

Dataset	Duration	ASes	Unique ASes	AS links	Unique AS links
IRR	Jan 1, '11	32,098	3,104	182,211	72,835
BGP-UCLA	Apr'04-Jan'11	41,232	3,037	364,752	142,998
CAIDA	Mar'08-Jan'11	45,827	210	266,205	70,326
DIMES	Jan'07-Dec'10	34,401	37	622,465	454,005
Ono	Dec'07-Sep'08	31,847	35	143,384	24,671
IXPMap	Apr'09	3,757	11	51,990	13,791

Methodology: Accuracy Checks



- ▶ **(a) Direct POM:** Origin AS is the same between BGP-UCLA and the IRR
- ▶ **(b) AS Path POM:** Another AS (other than the origin AS) registers the IP prefix (proxy registration)
- ▶ **(c) AS Link POM:** Whether the link between two ASes can be verified by AS link information.
 - ▶ Possible reasons for origin AS change between the IRR and BGP
 - ▶ **Route Aggregation, Route Splitting, Static Routing**

Results Preview

▶ How Complete is the IRR?

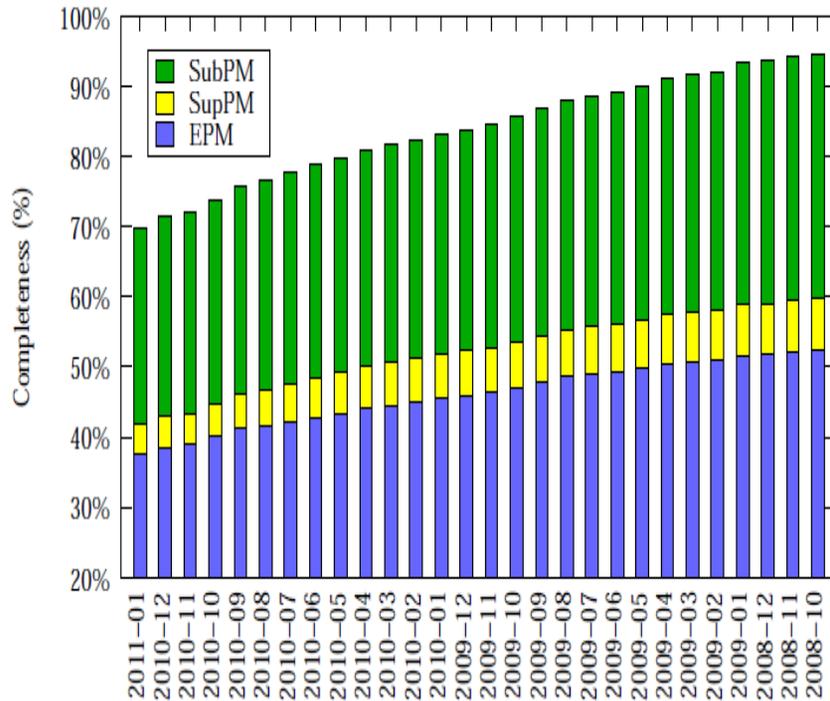
- ▶ How many prefixes in **IRR** match with **BGP**?
 - ▶ About **95%** of prefixes in the IRR match in over 2 years of BGP.
 - Not matched (**5%**) may be stale, yet to be announced or wrong.

▶ How Accurate is the IRR?

- ▶ How many prefixes + Origin AS in **IRR** match with **BGP**?
 - ▶ About **88%**.

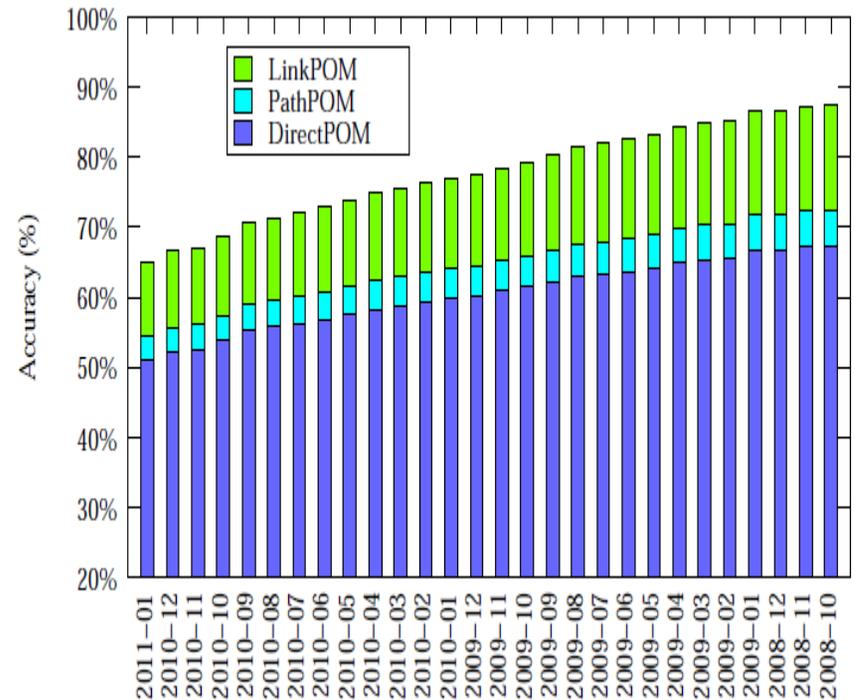
How complete/accurate is the IRR?

How much old is gold ?



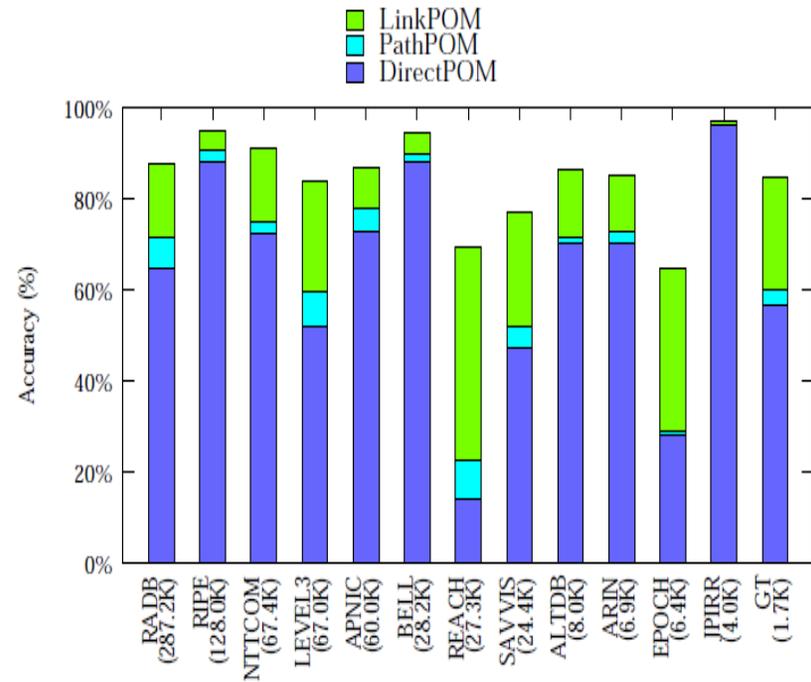
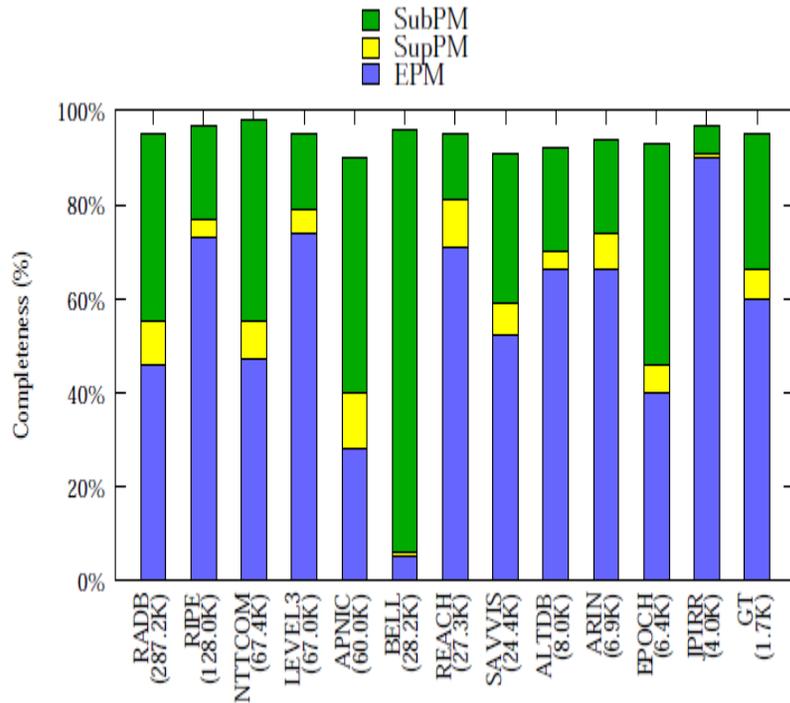
Prefixes (%) in IRR ROs (1st Jan 2011) matched in BGP-UCLA

Accuracy checking methodology can be further improved



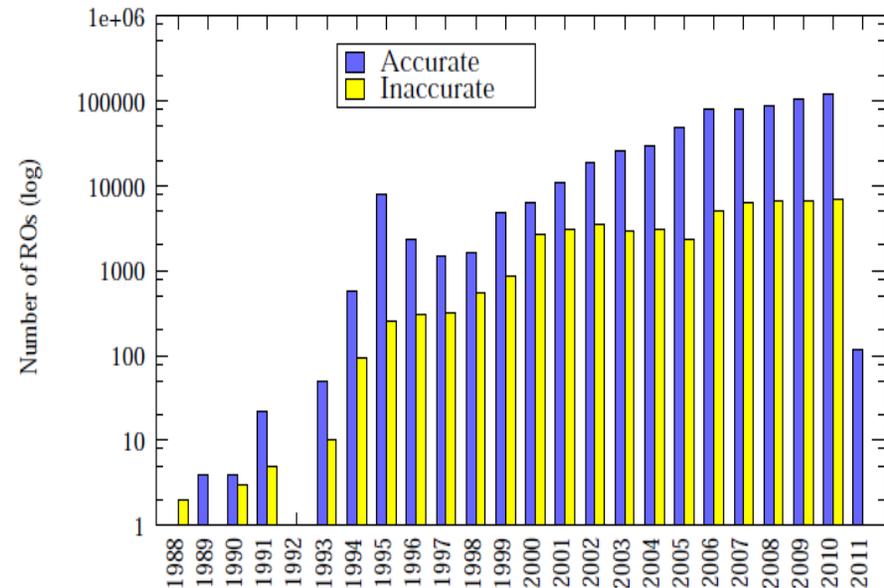
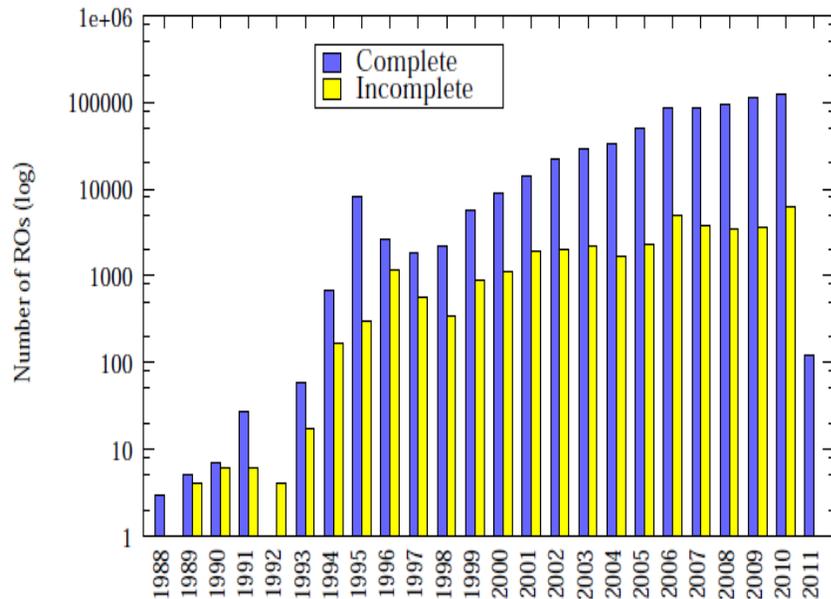
Prefixes (%) in IRR ROs whose origin ASes are verified by our accuracy checking methodology

Are some RRs more complete/accurate ?



- ▶ Large RRs [Number of registered ROs > 1 K ROs]
- ▶ Based on RR's Accuracy
 - ▶ Best RRs (91~97%): RIPE, JPIRR, BELL
 - ▶ Good RRs (81~90%): NTTCOM, RADb, APNIC, etc.
 - ▶ Average RRs (61~80%): REACH, SAVVIS, EPOCH

Possible Reasons for the Incompleteness/Inaccuracy



- ▶ Most popular assumed reason is staleness of the IRR
 - ▶ However, staleness is not the only reason
 - ▶ Recently entered information is also incomplete /inaccurate.
- ▶ Possible reasons:
 - ▶ Incompleteness/Inaccuracy of the BGP traces / AS link datasets
 - ▶ Limited knowledge about routing policies between ASes

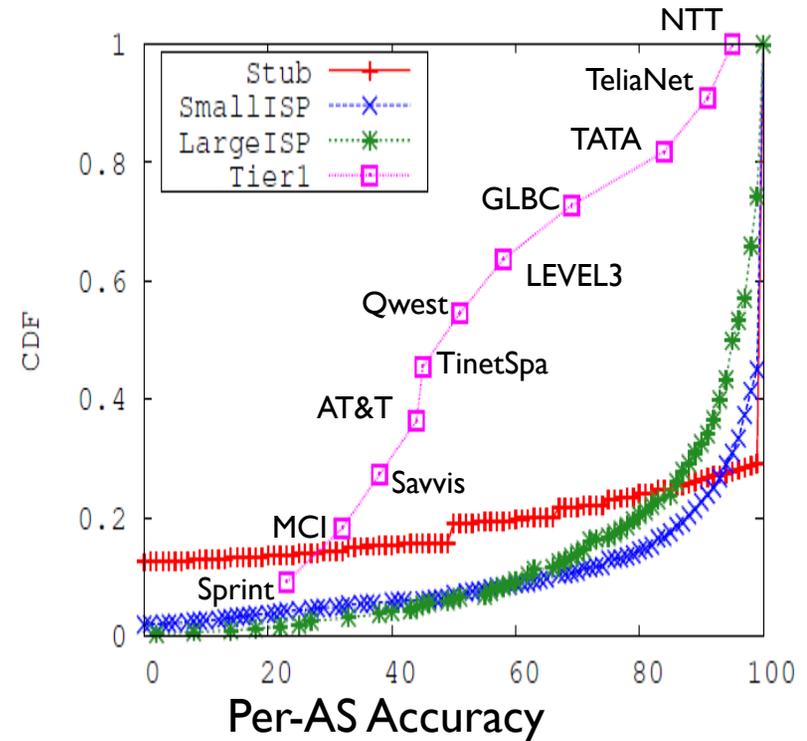
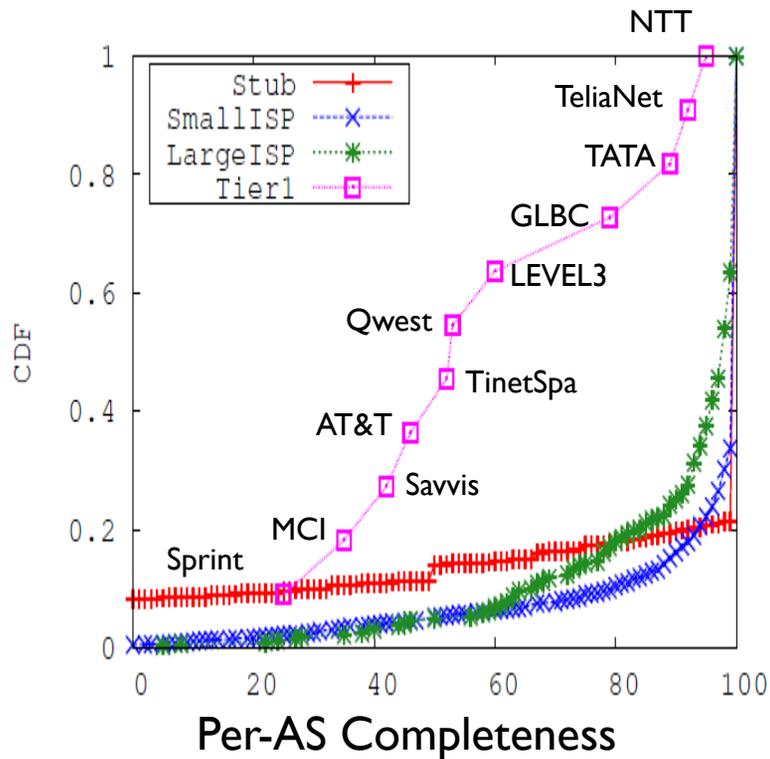
Results Preview

- ▶ How Complete is the IRR?
 - ▶ How many prefixes in IRR match with BGP?
 - ▶ How many prefixes in **BGP** match with **IRR**?
 - ▶ About 84%
 - In progress analysis expected to increase this to 88~90%
 - ▶ Incomplete prefixes (16%) can be attributed to
 - Misconfigured BGP announcements / ASes not using the IRR
- ▶ How Accurate is the IRR?
 - ▶ How many prefixes + Origin AS in IRR match with BGP?
 - ▶ How many prefixes, OAS in **BGP** match with **IRR**?
 - ▶ About 75%

Per-AS Completeness/Accuracy Types

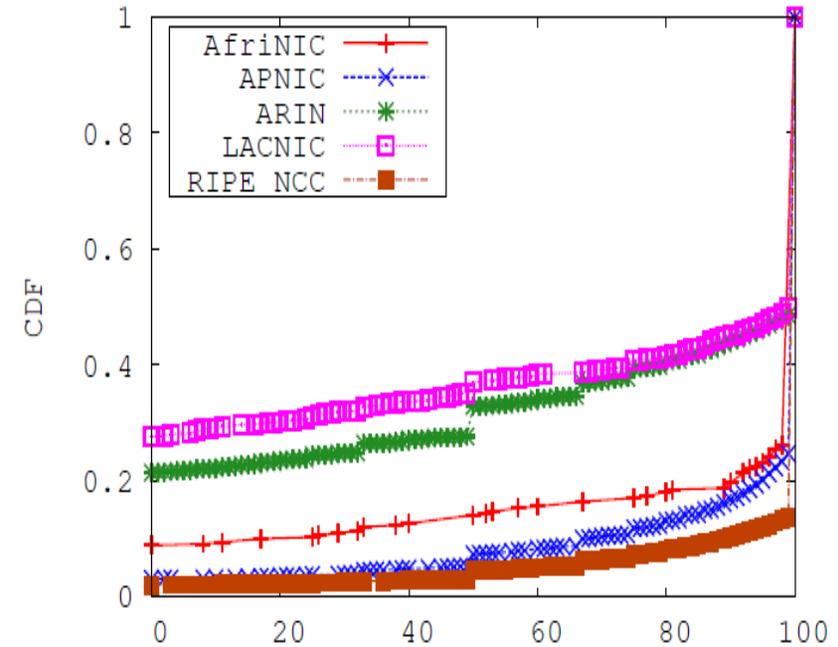
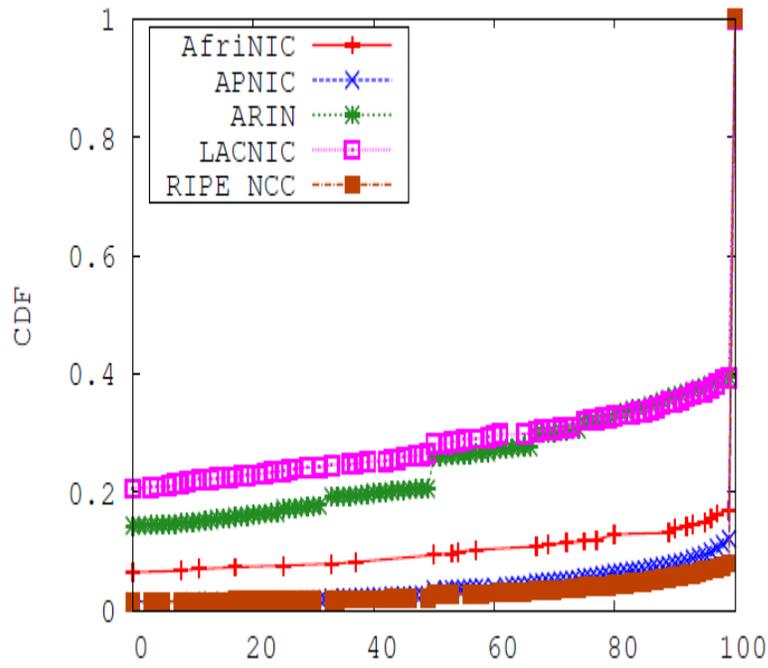
- ▶ Calculate per-AS Completeness/Accuracy
 - ▶ Per-AS Completeness =
 - ▶ Number of prefixes by an AS in BGP (over two years) matched in IRR
 - ▶ Divided by total number of prefixes by that AS in BGP * 100
- ▶ Three types of per-AS Completeness
 - ▶ **Full Completeness:** All BGP announcements in IRR
 - ▶ **Partial Completeness:** BGP announcements missing in the IRR
 - ▶ **No Completeness:** AS is not using the IRR
- ▶ Similar calculation and classification for per-AS Accuracy

Per-AS Completeness/Accuracy (AS types)



- ▶ 92 % ASes found in the BGP (1st Jan 2011 ~ Oct 2008) use the IRR
 - ▶ with varying levels of completeness/accuracy across different AS types
 - ▶ 8 % not using the IRR at all
 - ▶ Single-homed customers of few tier 1s

Per-AS Completeness/Accuracy (RIR wise)



► In different RIR regions

- more active usage of IRR by ASes in the RIPE NCC, APNIC, AfrinIC regions than ARIN and LACNIC regions.

Conclusion

- ▶ IRR is used by 92% of the ASes
 - ▶ across different AS types and RIR regions
 - ▶ with varying levels of completeness/accuracy
- ▶ **Work-in-progress**
 - ▶ Use of IRR to mitigate BGP Robustness problem
 - ▶ Misconfigured announcements / Prefix hijacking
 - ▶ Validation
 - ▶ Accuracy checking methodology
 - ▶ How to accurately declare some IRR information as stale?

THANK YOU!

Questions ?