

Analysis of ARIN Data on NetHandles with OriginAS

O. Kim, K. Sriram, O. Borchert, P. Gleichmann, and D. Montgomery

Meeting with ARIN Staff
February 26, 2009

Contacts: okim@nist.gov, ksriram@nist.gov, doug@nist.gov

Outline

- Quality analysis of registry information
 - Problem space and goals
 - Routing and addressing registries
 - Registry consistency check algorithm
 - Information quality metrics
- Analysis of ARIN NetHandles with OriginAS
 - Some observations on ARIN NetHandle data
 - WRT route objects in RPSL
 - WRT BGP trace data
 - Aggregation (consolidation) of ARIN NetHandle address space with OriginAS
- Characterization of self-consistency for registry information

What is the Problem?

- Current registry data is considered:
 - Stale, insecure, inaccurate and incomplete
 - Registration is voluntary
 - Routing policy system is not adequately secure
- Despite weaknesses, data is used for:
 - Local route filtering
 - Debugging purposes
- Integrity of registry data is crucial to help maintain global secure routing infrastructure.
- No comprehensive investigations to date
- Improving quality and completeness of routing data could enable new BGP robustness mechanisms

What are NIST's Goals?

- Perform **comprehensive investigations and analysis** of the integrity of the registry data
 - Characterize **correctness and completeness** of existing RIR/IRR databases
 - * Quality measurement of IRR data
 - * Analysis of syntactic correctness of IRR data
 - * Analysis of IRR content changes
 - Characterize **consistency of global IRR data with regard to BGP trace data**
 - * Quality analysis of registry data (route and NetHandle objects) vs. BGP route announcements (Updates)
 - Contribute to the improvement of the quality of global routing information infrastructure
- Understand and gain some insights by analyzing both registry and trace data, to help improve BGP routing robustness

Examples of Registry Data Objects (RPSL)

inetnum

```
inetnum: 129.6.0.0 –  
129.6.255.255  
descr: description stmt  
tech-c: nist-tech-ID  
admin-c: nist-admin-ID  
status: assigned PA  
mnt-by: MNT-NIST  
mnt-routes: iip-bgp-mnt  
source: RIPE
```

route

```
route: 129.6.0.0/24  
descr: NIST/DOC  
origin: AS49  
mnt-by: iip-bgp-mnt  
source: RIPE
```

aut-num

```
aut-num: AS49  
org:  
import:  
export:  
default:  
tech-c: AS49-tech  
mnt-by: MNT-NIST  
mnt-routes: iip-bgp-mnt  
source: RIPE
```

mntner

```
mntner: iip-bgp-mnt  
descr: description  
stmt  
auth: encryp  
mnt-by: MNT-NIST  
source: RIPE
```

Classes for allocation/assignment:

- **inetnum**: IPv4 address space
- **aut-num**: AS number

Class for route information:

- **route**

Classes for security mechanisms:

- **mntner**

Registry Data Object Counts by Source

RIR	route			inetnum (ARIN NetHandle)			aut-num (ARIN ASHandle)		
	06/18/2007	10/18/2008	Incr	06/18/2007	10/18/2008	Incr	06/18/2007	10/18/2008	Incr
ARIN	7,330	8,201	12%	338 (1,618,197)	434 (1,924,454)	28% 19%	758 (18,050)	890 (19,678)	17% 9%
RIPENCC	71,569	89,957	26%	2,044,536	2,458,119	20%	14,106	16,969	20%
APNIC*	23,616	35,515	50%	822,891	1,080,999	31%	4,559	5,347	17%
AFRINIC	0	0		13,948	22,706	63%	342	445	30%
LACNIC**	0	0		45,346	83,036	83%	1,219	1,339	10%
RADB+	345,129	497,124	44%	1	1		3,785	4,643	23%
Total:	447,644	630,797	41%	2,927,060 (1,618,197)	3,645,295 (1,924,454)	25% 19%	24,769 (18,050)	29,633 (19,678)	20% 9%

* Includes TWNIC, JPIRR, JPNIC and APNIC

** RIR only

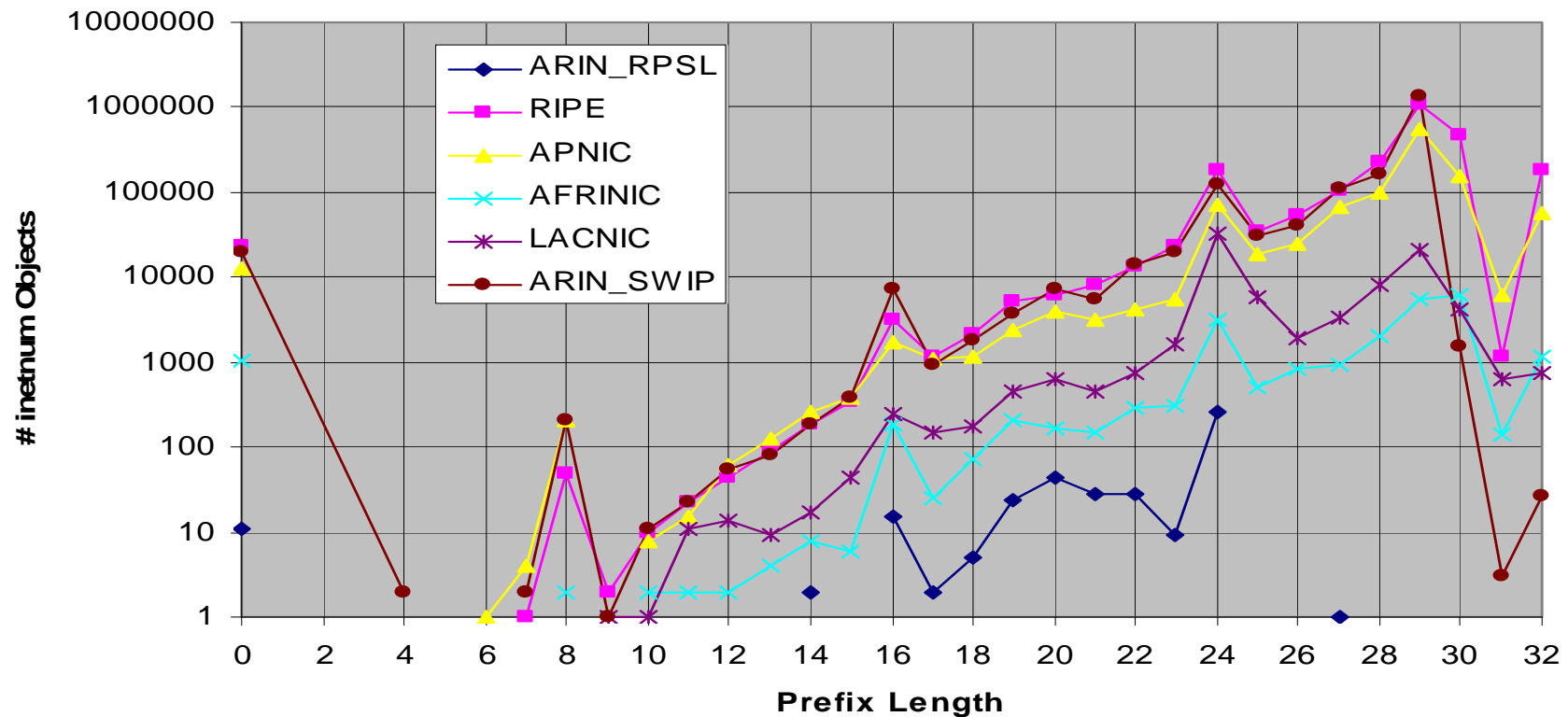
+ includes all mirrored data *EXCLUDING ARIN, APNIC, JPIRR*

Note that route objects can be registered any RIR regardless of where the address spaces are allocated.

Distribution of Prefix Length of inetnum (RPSL) and NetHandle (SWIP)

Registry Data Date: 2008-10-18

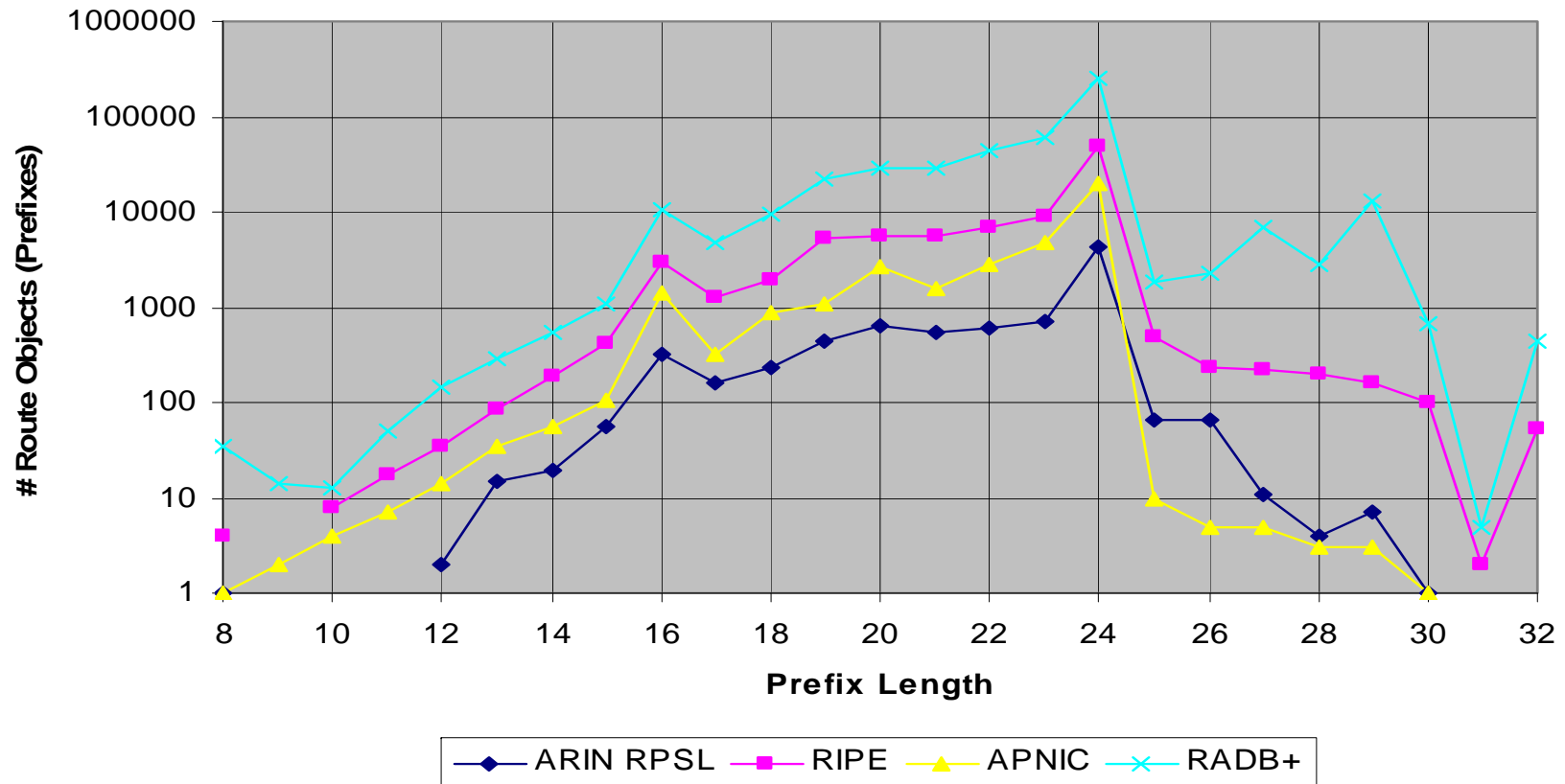
Trustworthy Networking Program



- Length 0 indicates that an address block cannot be represented by a single CIDR
- Length 4 specifies Multicast and Reserved Future Use blocks
- Some Legacy and ERX blocks may be included in one or more RIRs

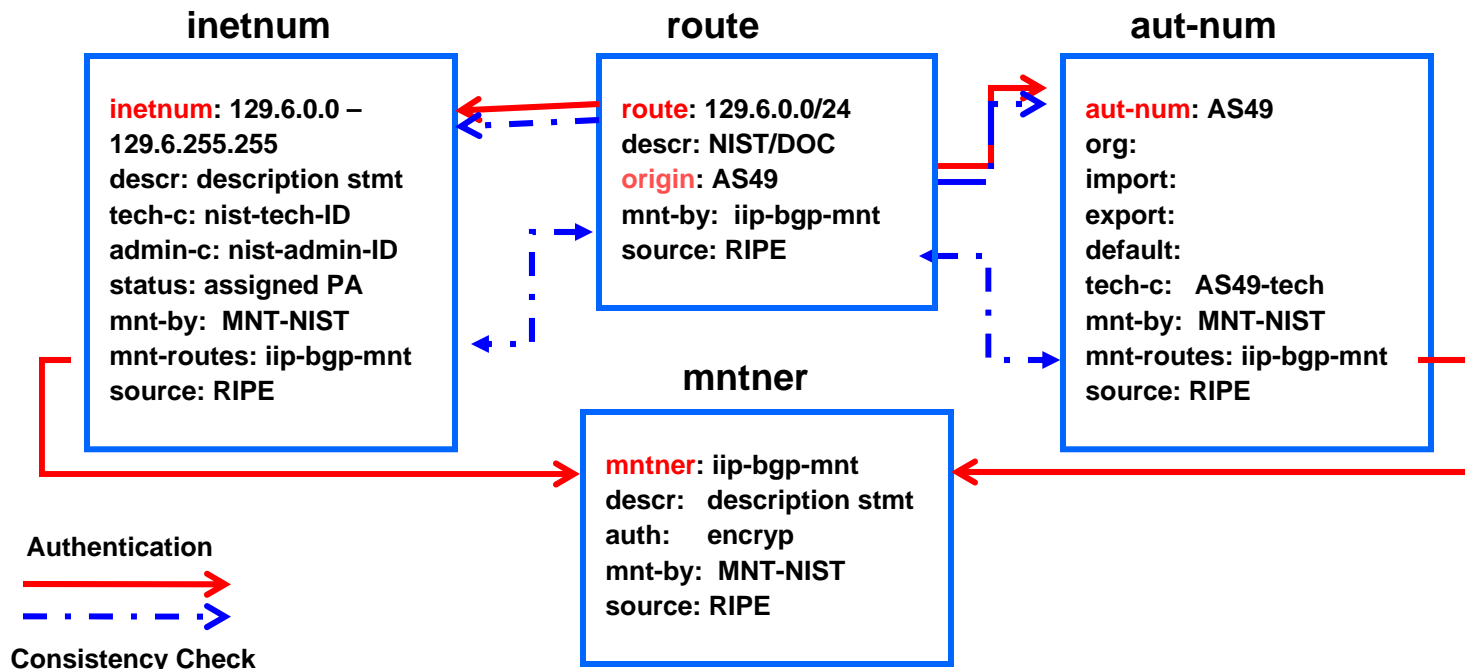
Distribution of Prefix Length of Route Objects in IRR

Registry Data Date: 2008-10-18



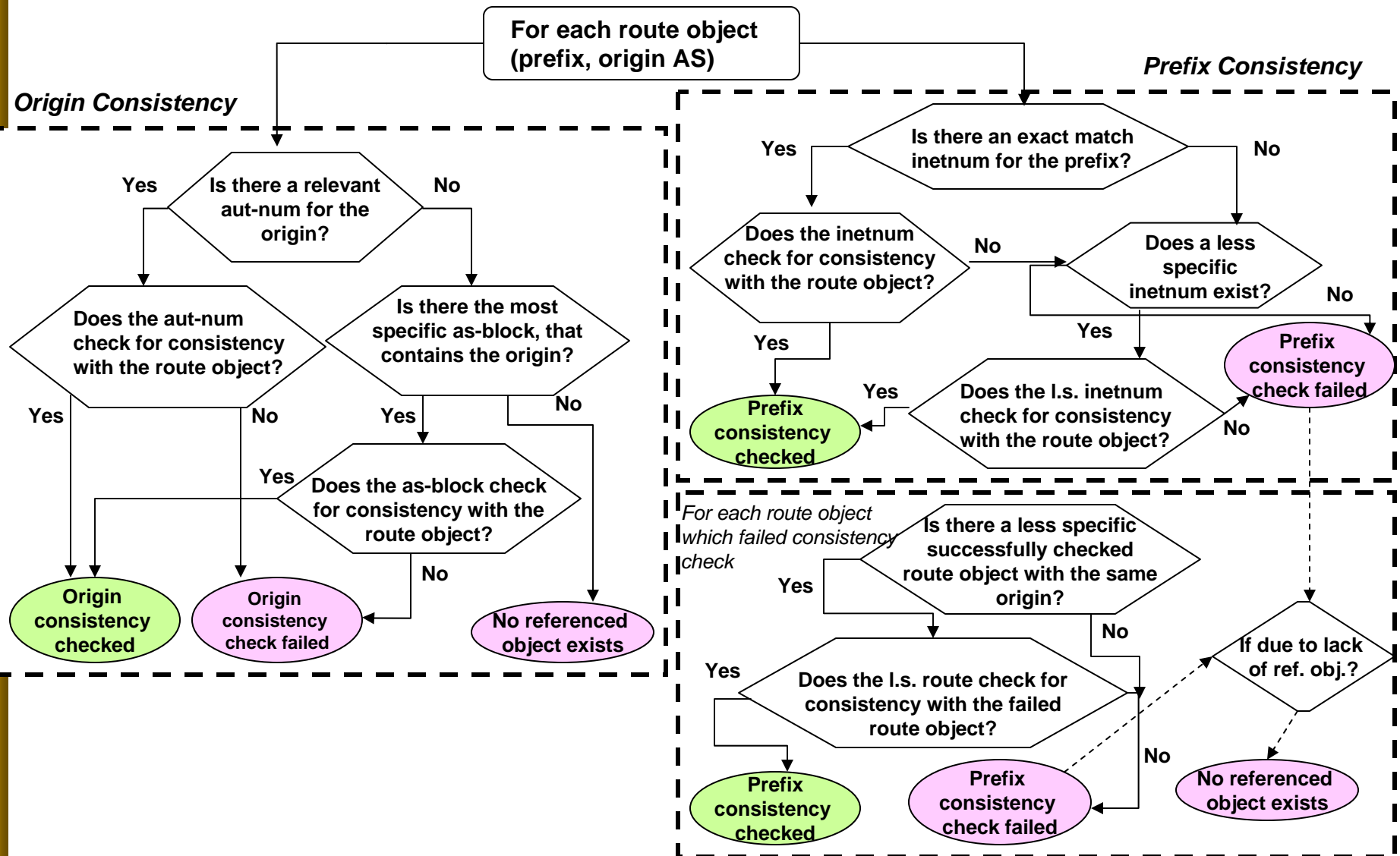
Registry Self-Consistency Check (Quality Analysis Algorithm)

- Self-Consistency check criteria:
 - Check consistency between relevant objects by comparing the following attributes:
 - * 'mntner' related attributes: Used mainly for RPSL
 - * 'orgID' attribute: Used mainly for SWIP
 - * Contact information (i.e., tech-c / admin-c / TechHandle / AbuseHandle)
- A route object is considered as fully consistent if, based on the above criteria, it matches with both of these:
 - ✦ the referenced **aut-num** for the **origin**; and
 - ✦ the referenced **inetnum** for the **prefix**.



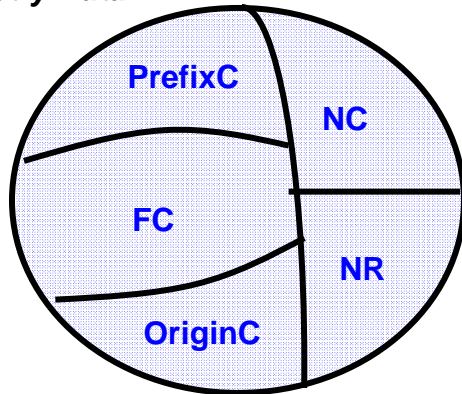
Registry Self-Consistency Check (Detailed Algorithm)

Trustworthy Networking Program



Information Quality Metrics

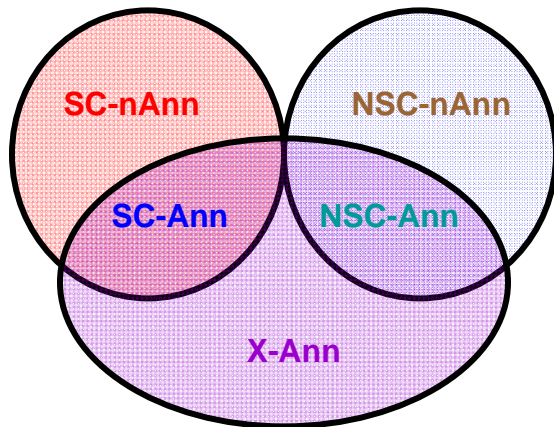
Registry Data



Registry Data Only:

- FC:** Fully (both prefix & origin) Consistent
- PrefixC:** only Prefix Consistent
- OriginC:** only Origin Consistent
- NC:** (referenced object exists, but) Not Consistent
- NR:** No Referenced objects exist

Registry Data



BGP Trace Data

Registry Data WRT BGP Trace Data:

Metrics	Registered route objects in the Registry	(prefix, originAS) Pairs in Trace Data
SC-Ann	Self-consistent	Observed exact match or more or less specific prefixes with same origin
SC-nAnn	Self-consistent	Not observed
NSC-Ann	Not self-consistent	Observed exact match or more or less specific prefixes with same origin
NSC-nAnn	Not self-consistent	Not observed
X-Ann	No registered route objects	Observed (prefix,originAS) pairs

- **SC** includes FC, PrefixC and OriginC
- **NSC** includes NC and NR

Outline

- Overview of registry information quality analysis
 - Problem space and goals
 - Routing and addressing registries
 - Consistency check algorithm
 - Information quality metrics
- **Analysis of ARIN NetHandles with OriginAS**
 - Some observations on ARIN NetHandle data
 - WRT route objects in RPSL
 - WRT BGP trace data
 - Aggregation (consolidation) of ARIN NetHandle address space with OriginAS
- Characterization of self-consistency for registry information

ARIN NetHandle Stats in Comparison to BGP Updates and RIBs

- Raw data
 - ARIN Registry data on 2008-10-18
 - * All NetHandle objects: 1,924,454
 - * Unique (NetHandle, OriginAS) pairs: 73,249 (4%)
 - * Unique (NetRange, OriginAS) pairs: 73,062
 - * Unique OriginASes: 2693
 - BGP Updates & RIB data:
 - * Collector: Oregon from Routeviews
 - * All prefixes from Updates (2008-06-01 to 2008-09-15): 1,159,558,753
 - Unique (prefix,origin) pairs: 362,649
 - * BGP RIBs on 2008-11-3: 283,035
 - unique (prefix,origin) pairs other than those in Updates prefixes above: 14,197
 - **ALL Unique (prefix,origin) pairs from both Updates and RIBs: 376,846**
- Multiple NetHandles that contain the exact same (NetRange, OriginAS) pairs with different allocation types:
 - Allocation types: allocation / reallocation / assignment / reassignment

# of instances with the following:	count
3 NetHandles containing the same (NetRange,OriginAS) pair	2
2 NetHandles containing the same (NetRange,OriginAS) pair	183
NetHandles with unique (NetRange,OriginAS) pair	72,877

Some Observations on ARIN NetHandles with OriginAS

- Two or more NetHandle objects contain the exact same (NetRange, OriginAS) pairs, but different *NetType*:
 - One Example: (66.97.96.0/20, 33125)

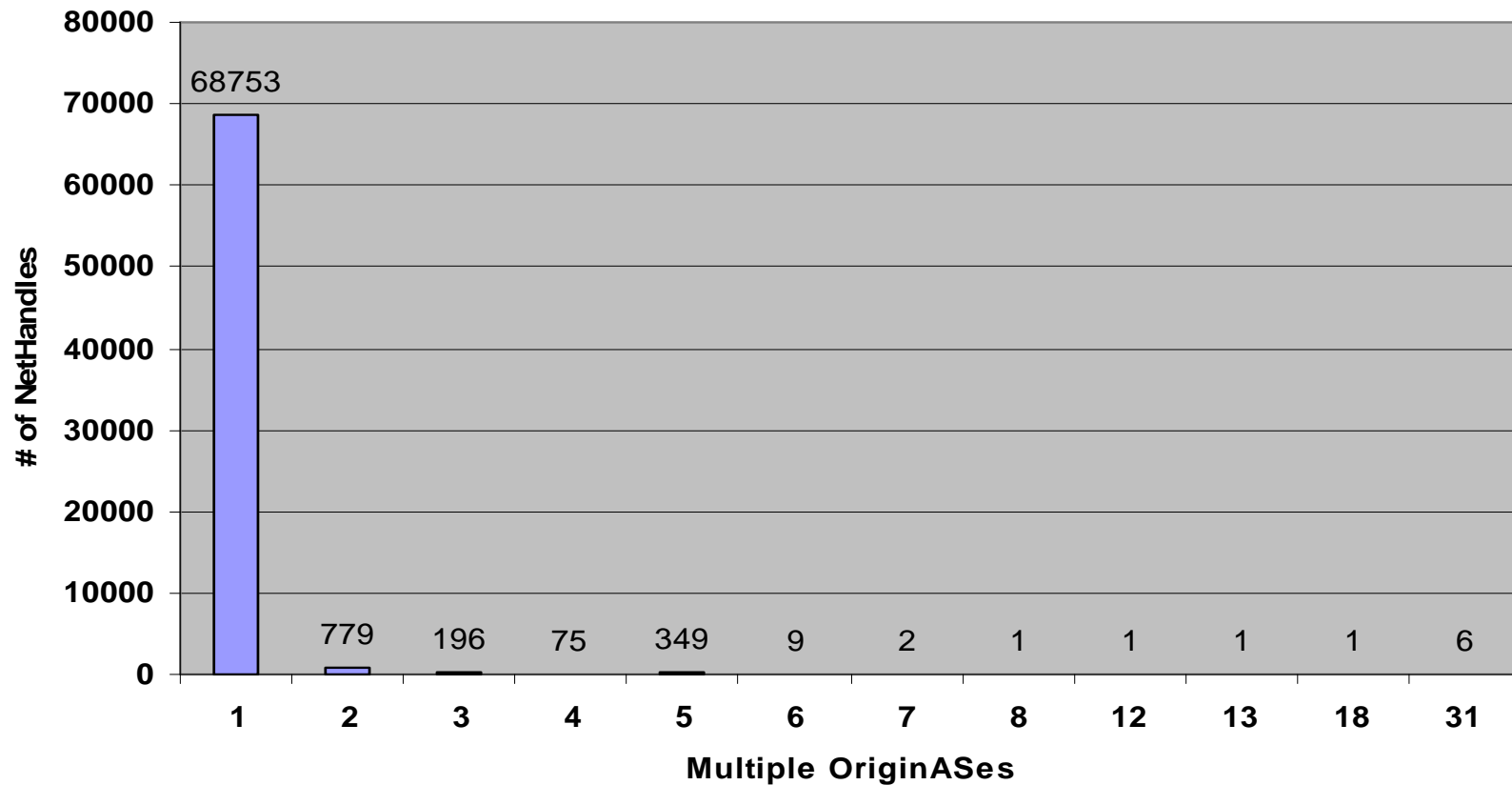
	NetHandle Object 1	NetHandle Object 2
NetHandle:	NET-66-97-96-0-1	NET-66-97-96-0-2
OrgID:	SNL-27	MCB-21
NetRange:	66.97.96.0 - 66.97.111.255	66.97.96.0 – 66.97.111.255
NetType:	<i>Allocation</i>	<i>Reassignment</i>
OriginAS:	AS33125	AS33125
Parent:	NET-66-0-0-0-0	NET-66-97-96-0-1
RegDate:	2006-10-10	2007-06-12
Updated:	2007-06-12	2007-06-12

ARIN NetHandles with OriginAS

Multiple OriginAS (MOAS) Distribution

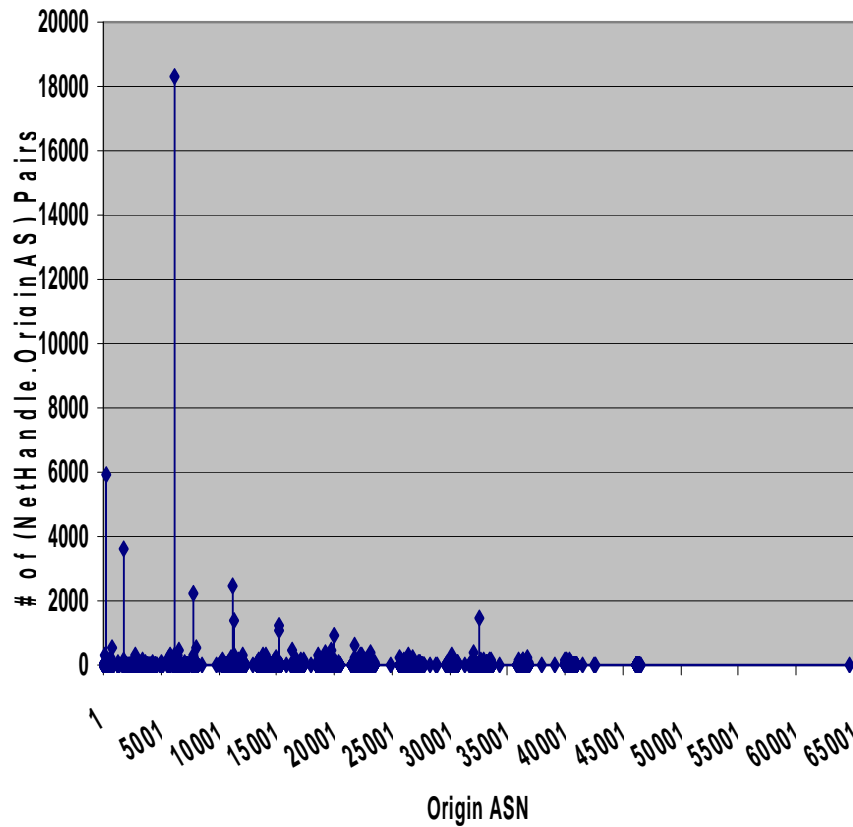
Registry Data Date: 2008-10-18

Trustworthy Networking Program

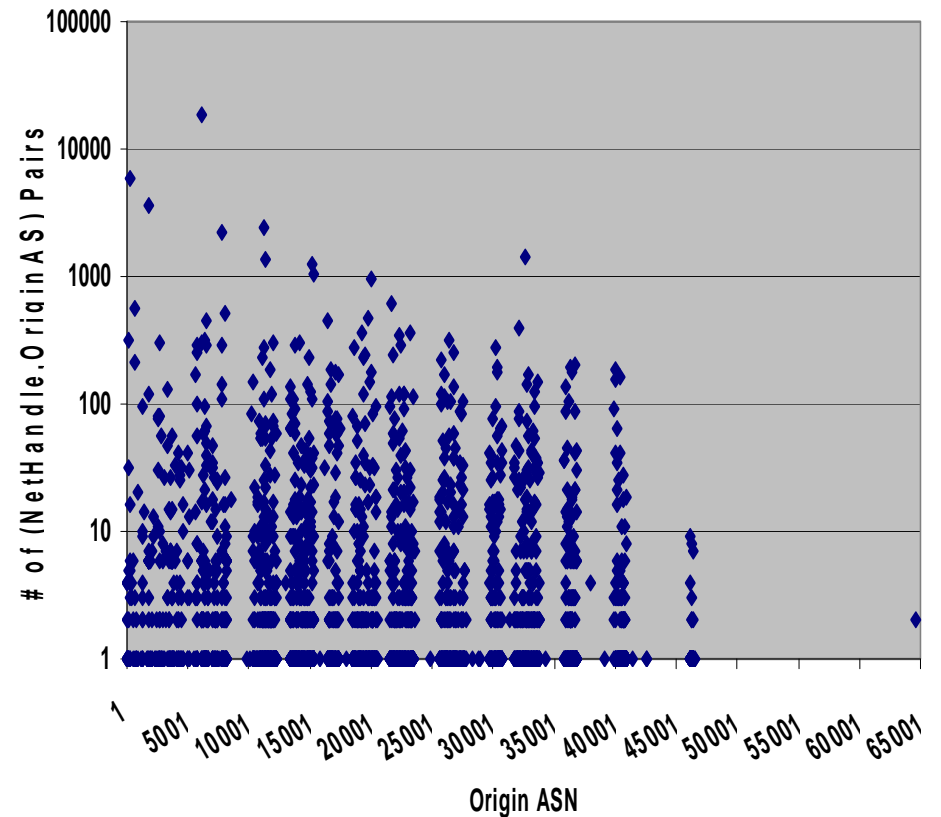


Distribution of NetHandles per Origin AS

Registry Data Date: 2008-10-18



Linear scale

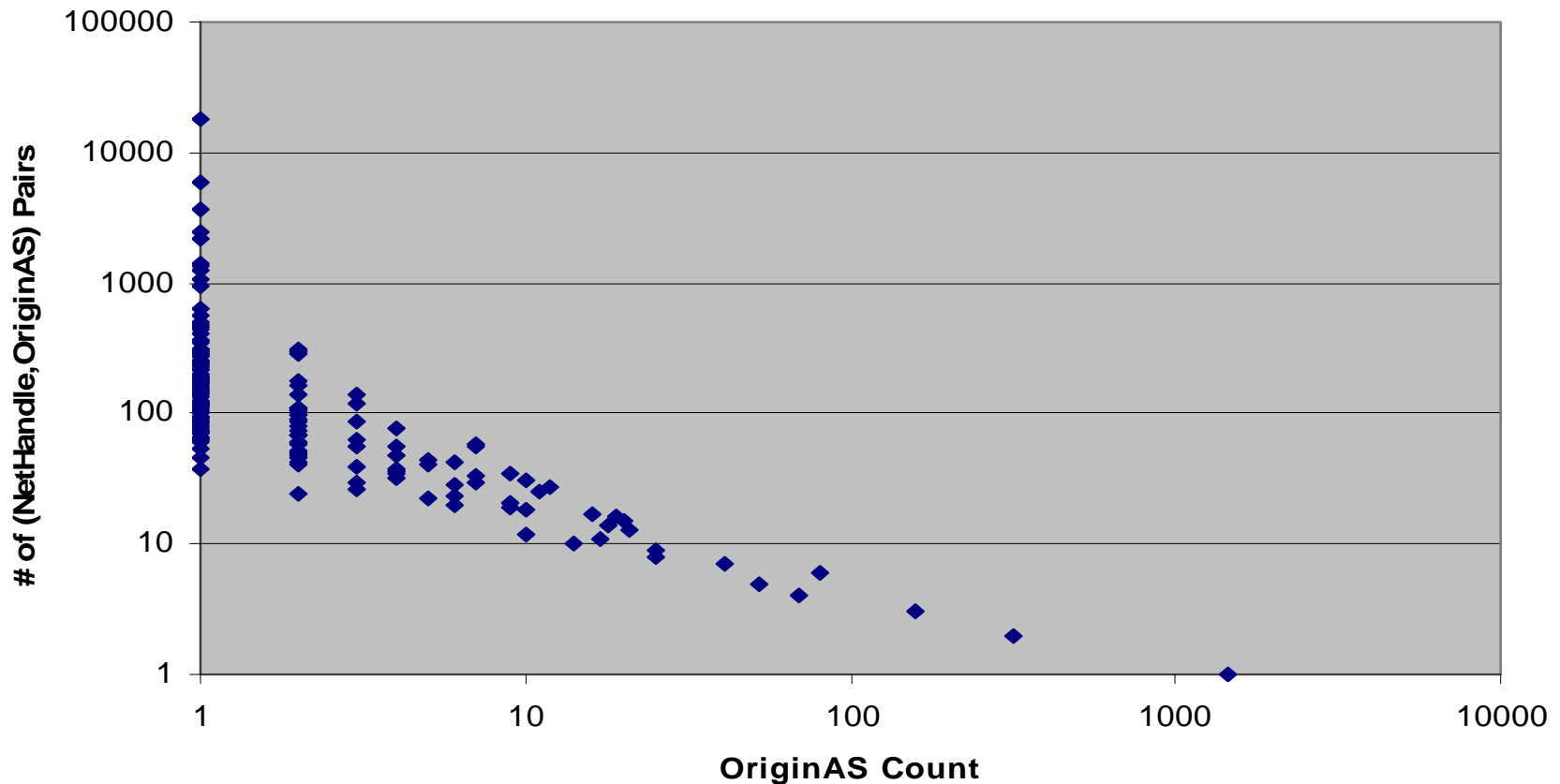


Log scale

Group OriginASes According to # NetHandles the OriginAS is Associated With

Registry Data Date: 2008-10-18

Trustworthy Networking Program

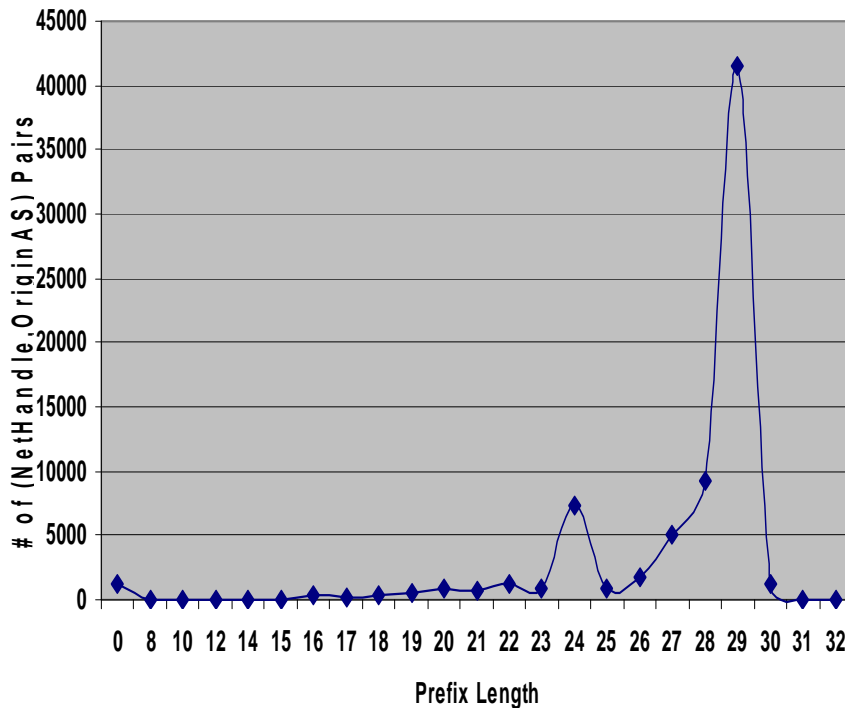


Distribution of # {NetHandle, OriginAS} pairs vs. # OriginASes in such a Group¹⁷

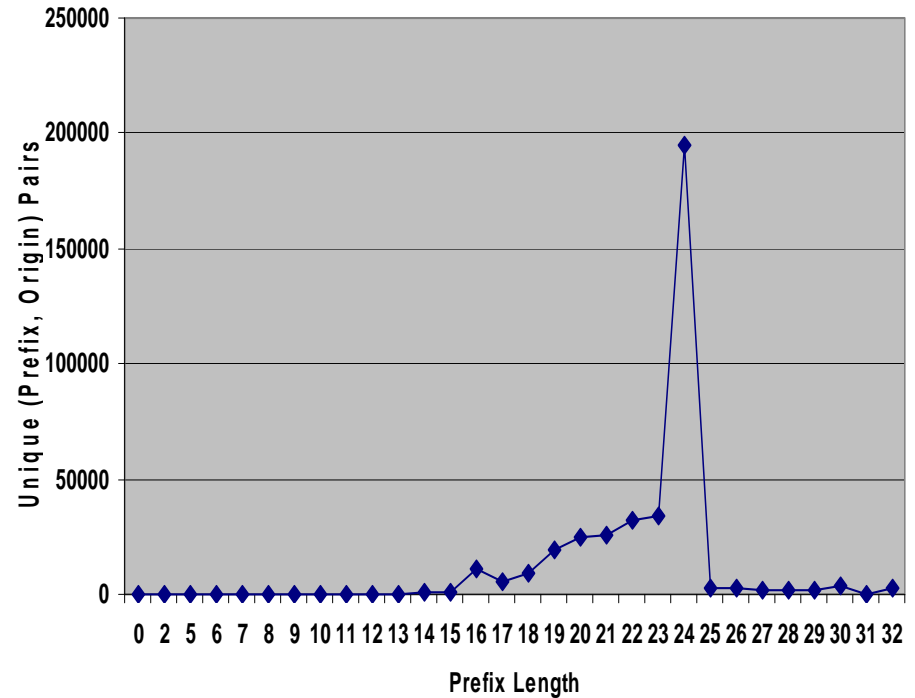
Distribution of Prefix Length of NetHandles w/ OriginAS vs. BGP Trace Data

Registry Data Date: 2008-10-18

BGP Trace Data
from 2008-06-01 to 2008-09-15



Length 0 indicates that NetRange cannot be represented by a single CIDR



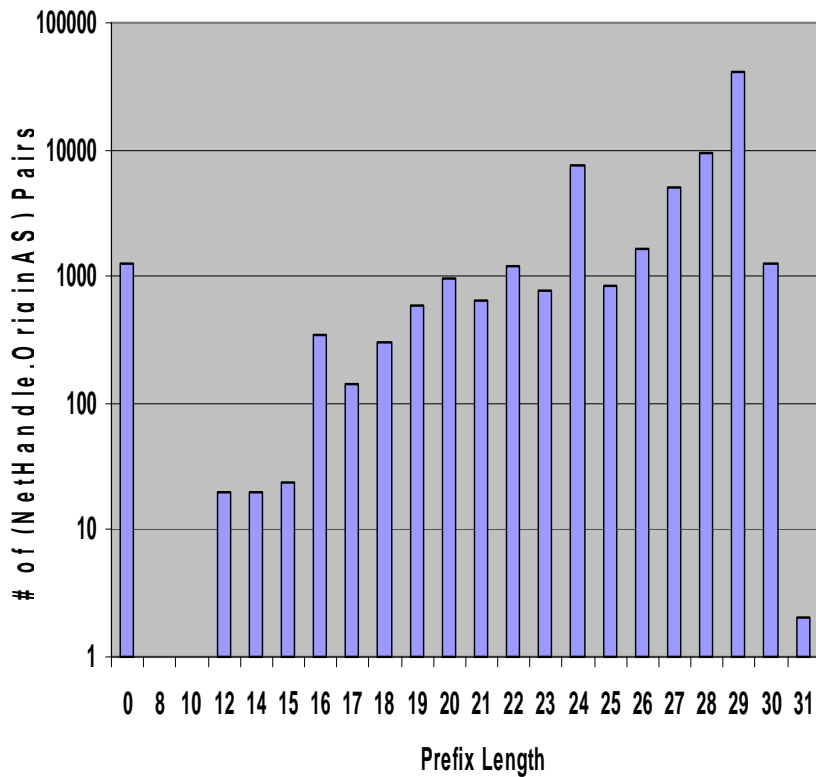
Length 0 indicates prefix 0.0.0.0/0

- Prefix 0.0.0.0/0 is announced by 7 OriginASes (12956, 3561, 19151, 513, 9829, 3130, 293).
- There exist 23 (prefix, origin) pairs with prefix length less than 8, including length 0 above.

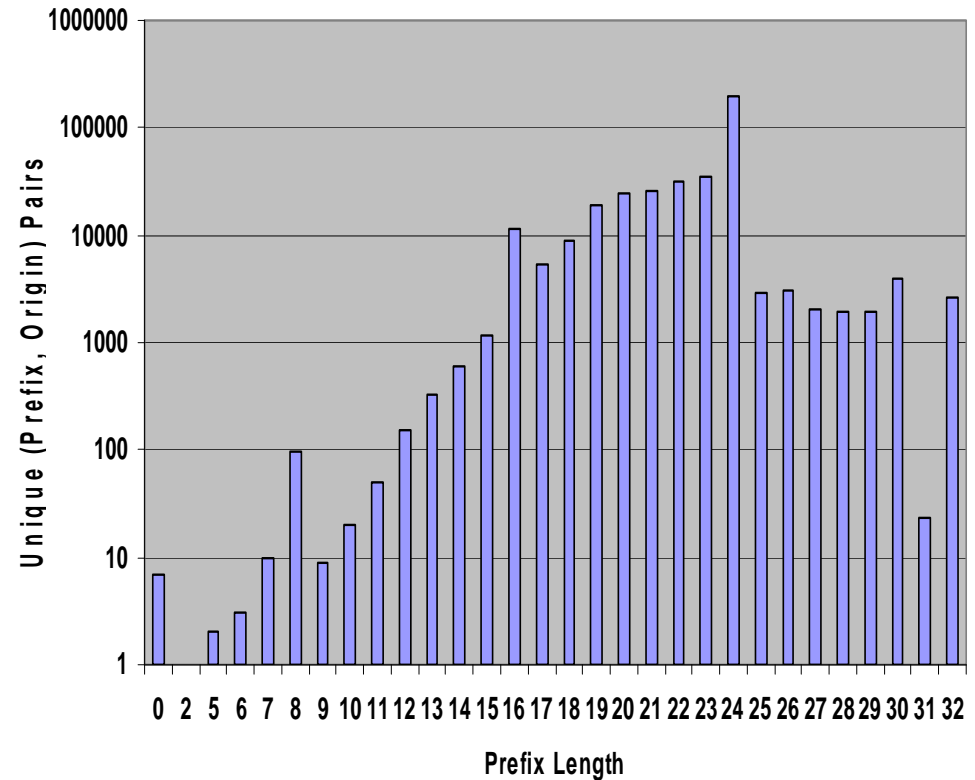
Distribution of Prefix Length of NetHandles w/ OriginAS vs. BGP Trace Data

Registry Data Date: 2008-10-18

BGP Trace Data
from 2008-06-01 to 2008-09-15



Length 0 indicates that NetRange cannot be represented by a single CIDR

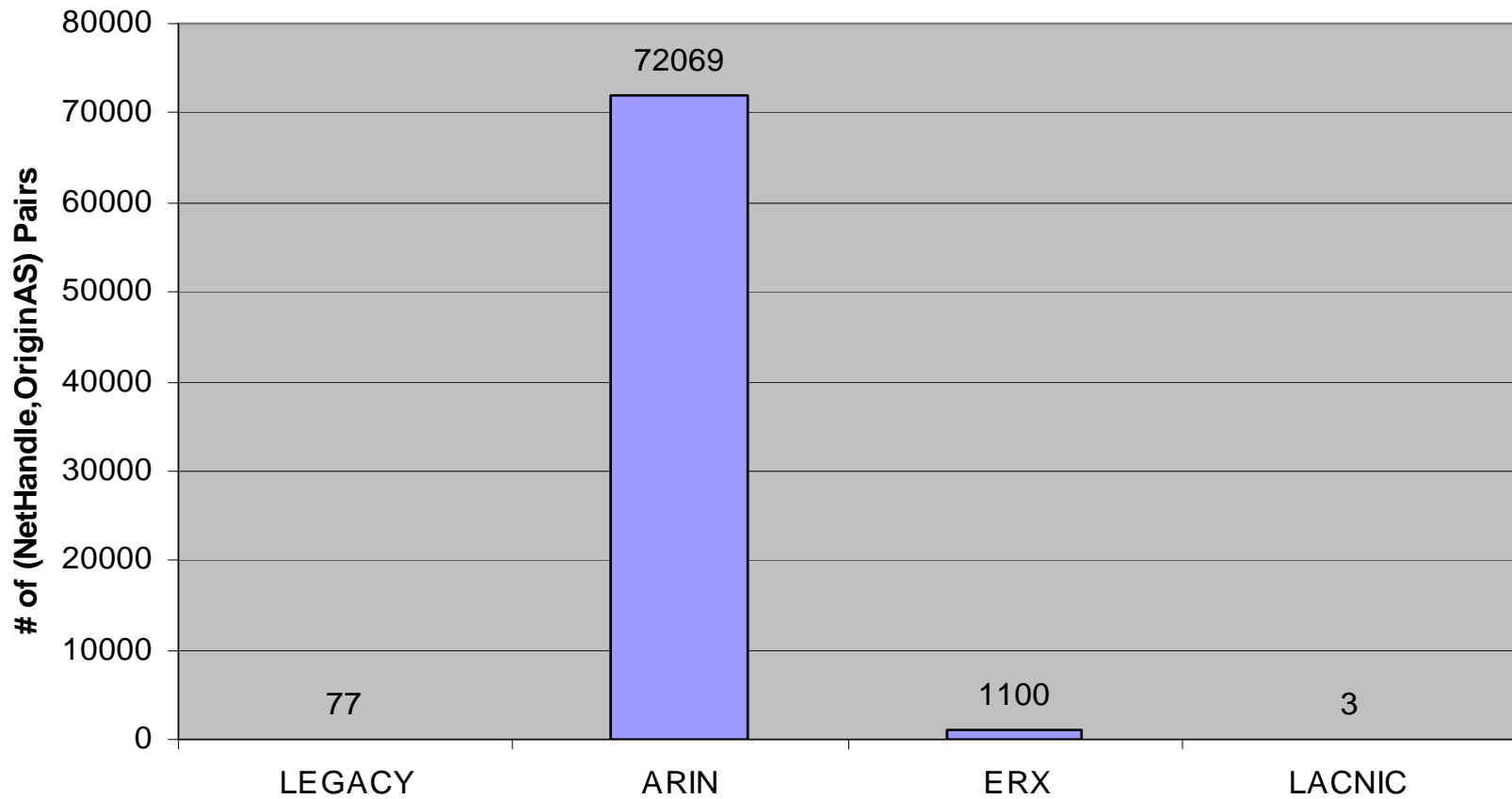


Length 0 indicates prefix 0.0.0.0/0

Distribution of ARIN NetRange Address Block Allocations

Registry Data Date: 2008-10-18

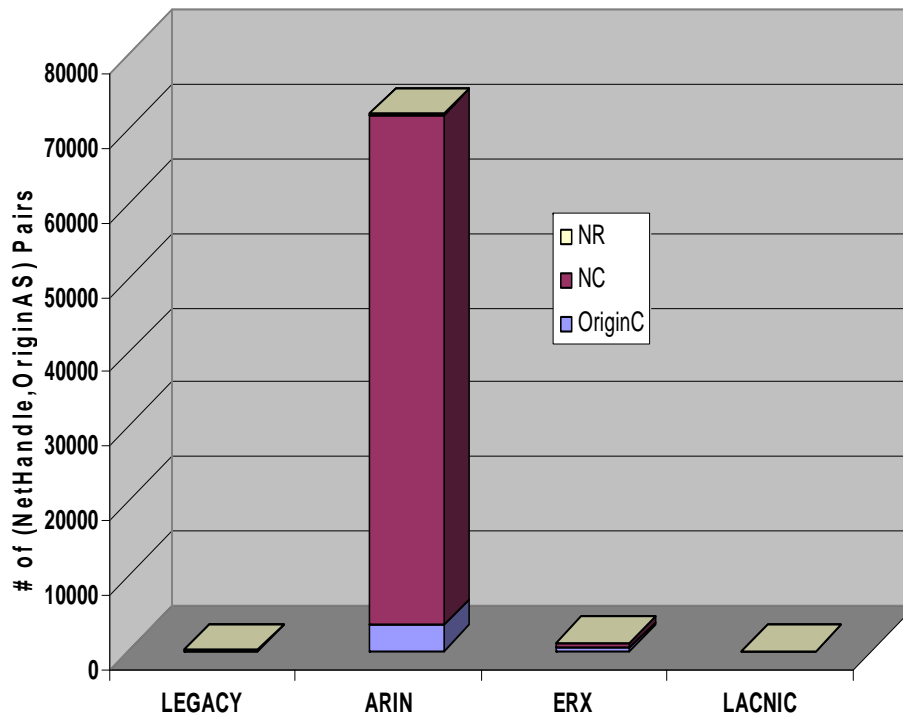
Trustworthy Networking Program



Note: Considering only NetHandles with Origin AS

Consistency Checks for ARIN NetHandles w/ OriginAS

Registry Data Date: 2008-10-18



Region	OriginC	NC	NR	Total
Legacy	3	73	1	77
ARIN	3519	68437	113	72069
ERX	391	697	12	1100
Lacnic	1	2	0	3
Total	3914	69209	126	73249

Scores for Consistency Checks for ARIN
NetHandle w/ OriginAS

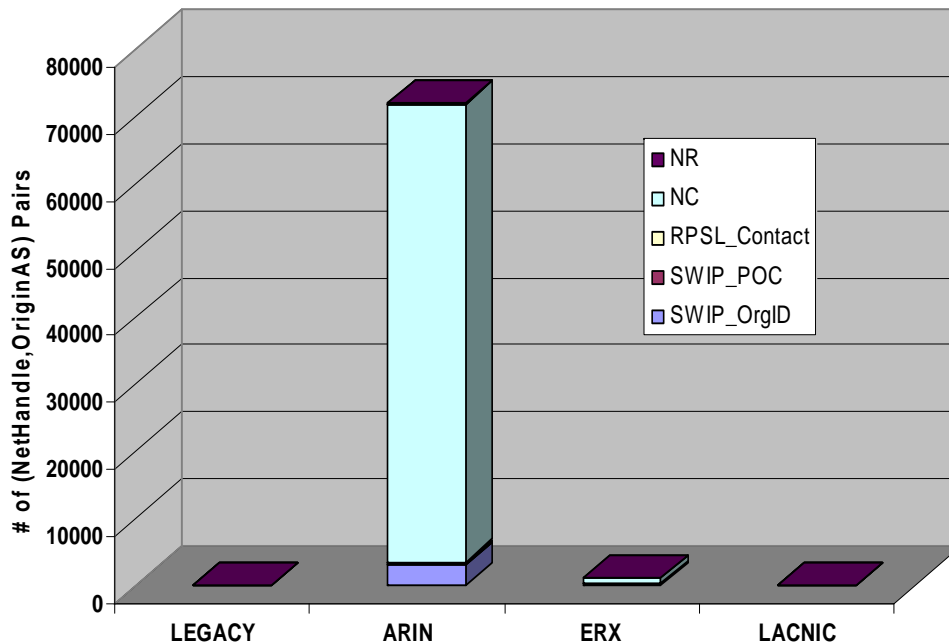
OriginC: Origin Consistent
NC: (referenced objects (e.g., ASHandle or aut-num) exist, but) Not Consistent
NR: No Referenced objects exist

Detailed Consistency Checks for ARIN NetHandles w/ OriginAS

Registry Data Date: 2008-10-18

- **SWIP_OrgID**: Checked with SWIP OrgID
- **SWIP_POC**: Checked with SWIP POC
- **RPSL_Contact**: Checked with RPSL contact info
- **NC**: (referenced objects exist, but) Not Consistent
- **NR**: No Referenced Resource Objects Exist

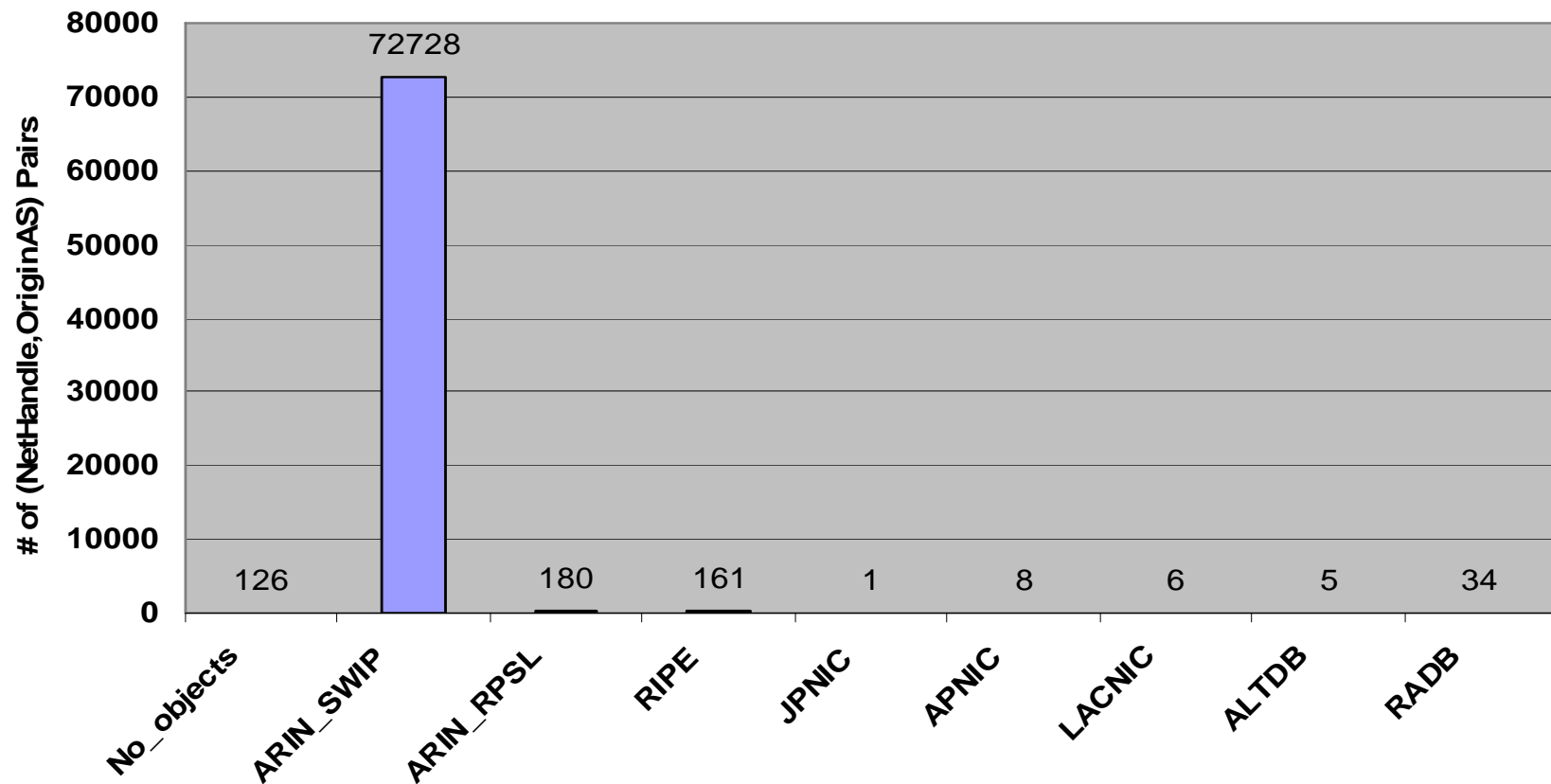
Region	SWIP OrgID	SWIP POC	RPSL Contact	NC	NR	Total
Legacy	3	0	0	73	1	77
ARIN	3149	184	186	68437	113	72069
ERX	318	66	7	697	12	1100
Lacnic	1	0	0	2	0	3
Total	3471	250	193	69209	126	73249



Detailed Scores for Consistency Checks
for ARIN NetHandle w/ OriginAS

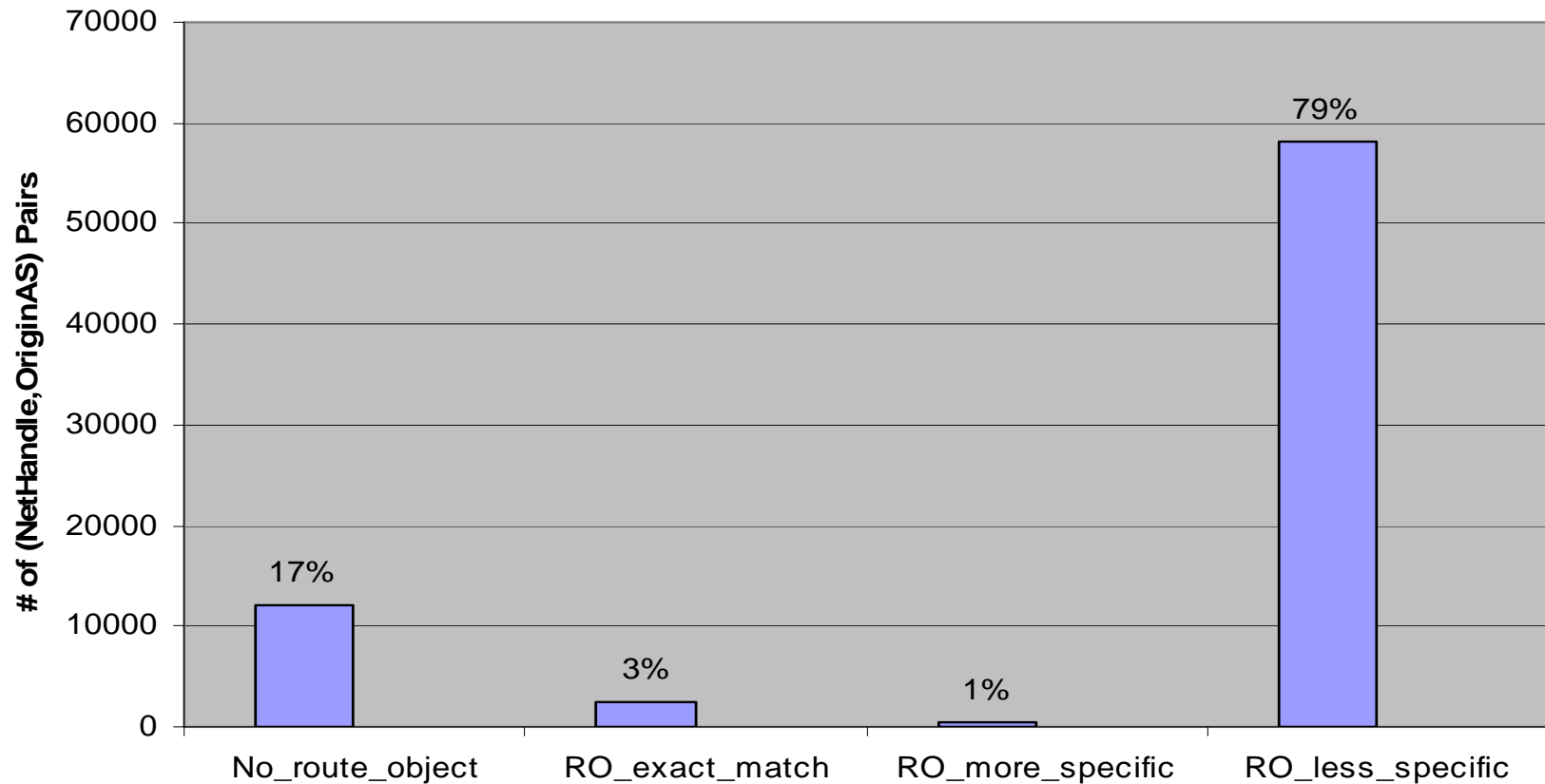
Sources of Referenced Objects (ASHandle or aut-num) for ARIN NetHandles w/ OriginAS

Registry Data Date: 2008-10-18



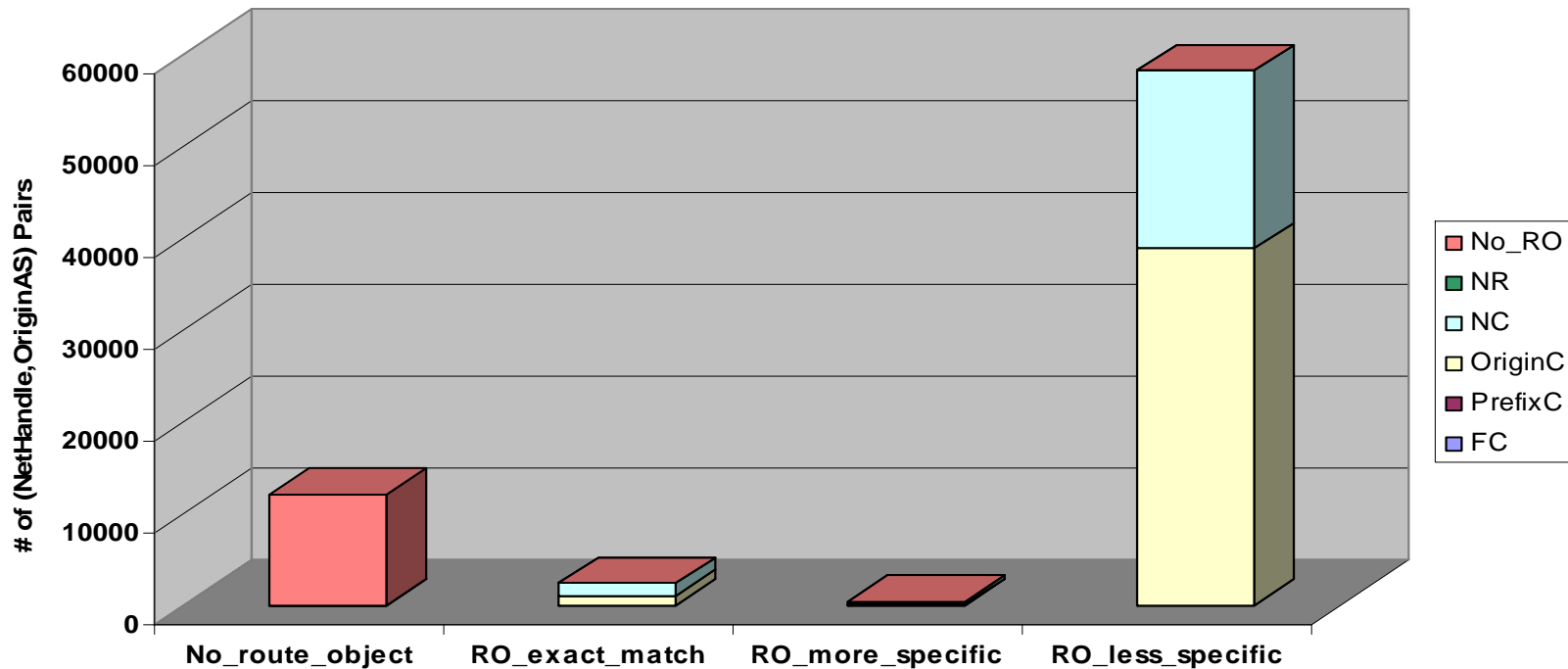
ARIN NetHandles w/ OriginAS and the Existence of Corresponding Route Objects in RPSL

Registry Data Date: 2008-10-18



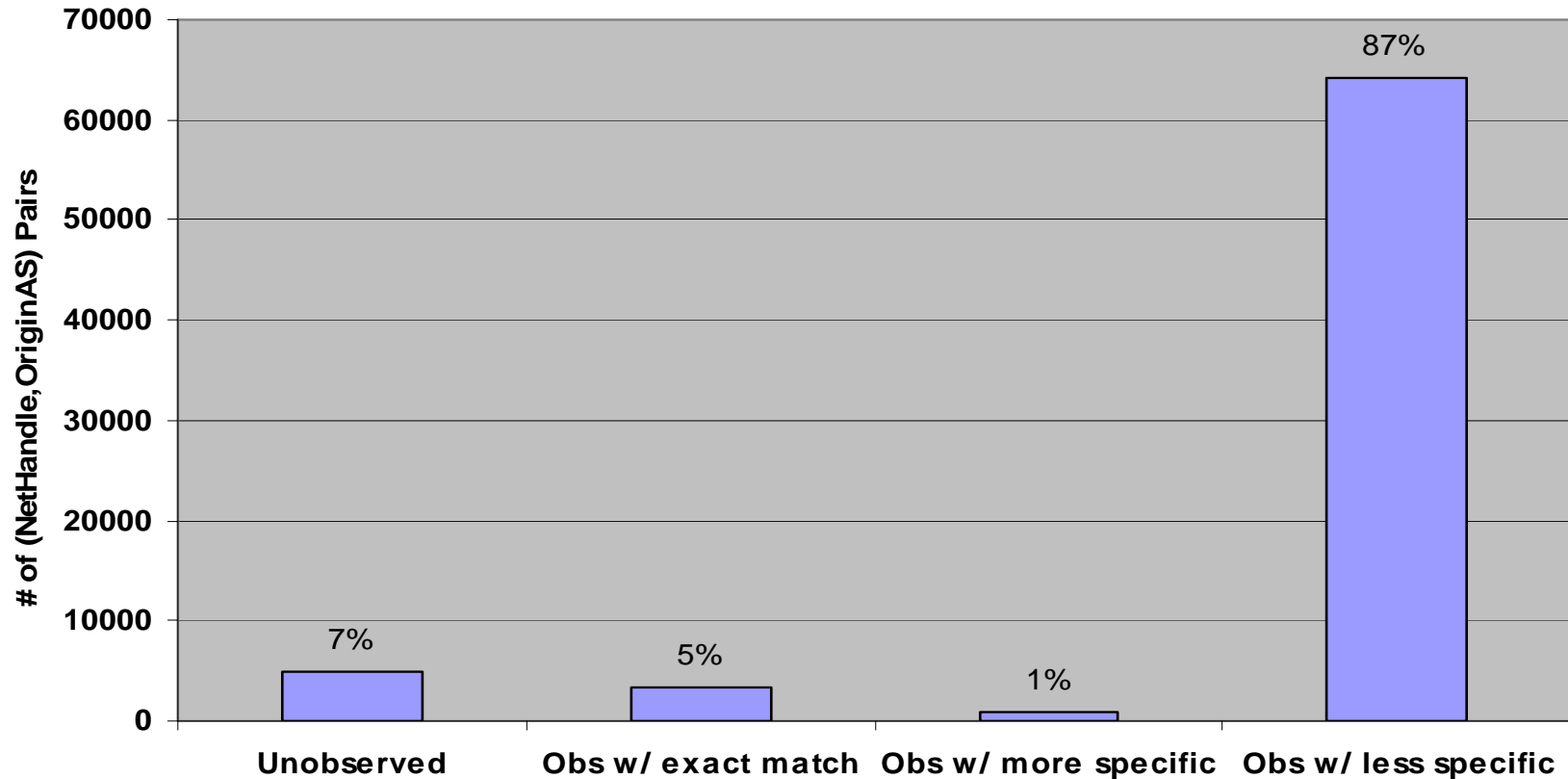
ARIN NetHandles w/ OriginAS and Existence and Quality of Corresponding Route Objects in RPSL

Registry Data Date: 2008-10-18



- **No_RO**: No Route objects exist
- **NR**: No Referenced objects exist (ie., ASHandle or aut-num)
- **NC**: (referenced objects exist, but) Not Consistent
- **FC**: Fully (Prefix & Origin) Consistent
- **PrefixC**: Only Prefix Consistent
- **OriginC**: Only Origin Consistent

ARIN NetHandles w/ OriginAS That are Observed in BGP Trace Data



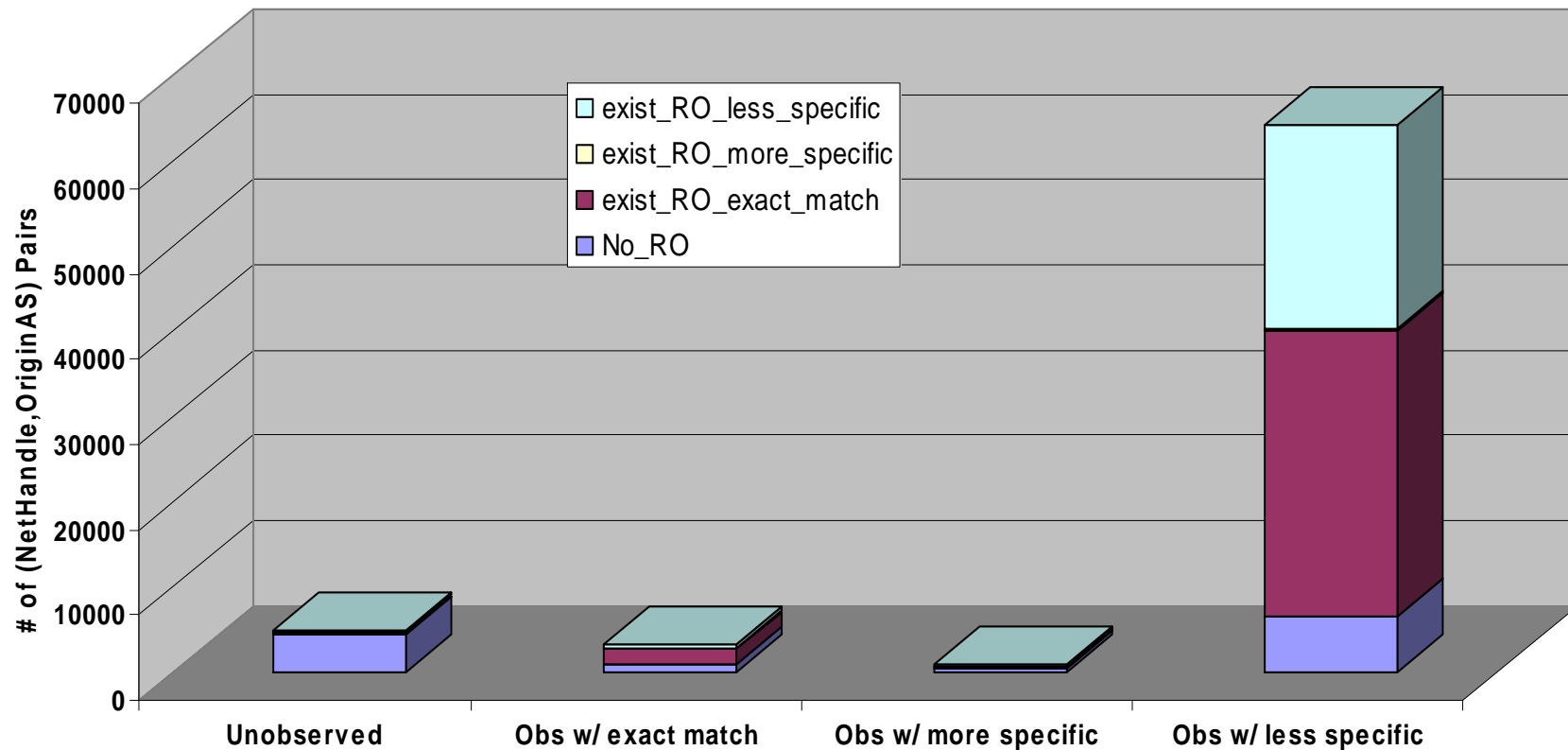
- About 6% of the NetHandles with origin AS are usable for direct verification of origin in BGP update messages; that is less than 5K NetHandles (in Oct. 2008)

Comparison of ARIN NetHandles with OriginAS vs. Announced (p, OAS) Pairs with Prefix Length ≥ 25

	All (p, OAS)	Prefix length ≥ 25	
		# of (p, OAS)	percentage
ARIN NetHandles with OriginAS	73k	60k	81.3%
Announced (p, OAS) that correspond to ARIN Address Space	124k	5.9k	4.8%
Globally announced (p, OAS)	376k	18.3k	4.9%

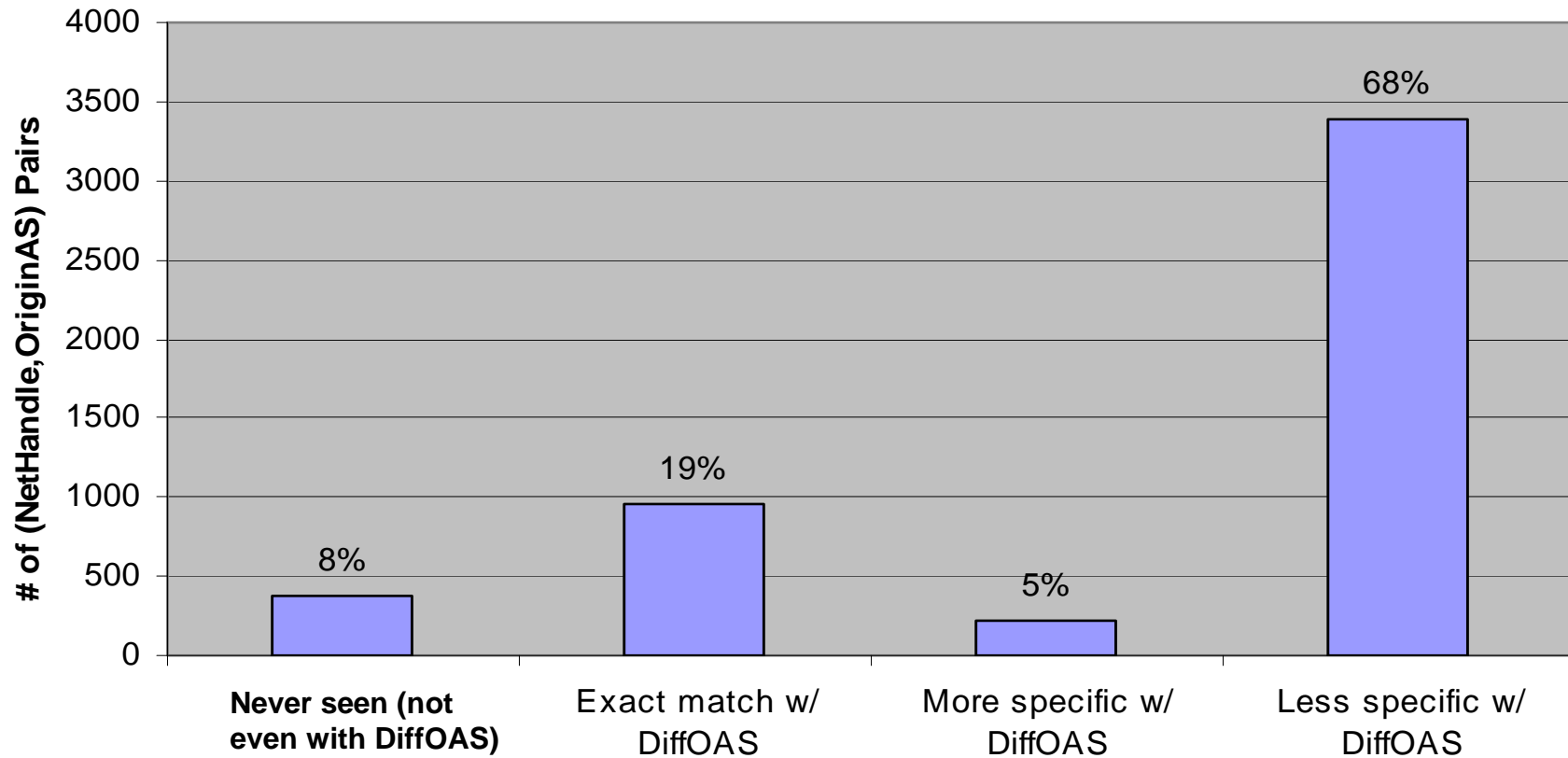
ARIN NetHandles w/ OriginAS That are Observed in BGP Trace Data (w.r.t. Corresponding Route Objects)

Trustworthy Networking Program



Unobserved ARIN NetHandles w/ OriginAS That are Seen with Different OriginAS

Trustworthy Networking Program



of unobserved (NetHandle, OriginAS) Pairs: 4,963

ARIN NetHandle w/ OriginAS Consolidation of (NetHandle, OriginAS) Pairs

On 2008-10-18	Total	# sub-prefixes
All unique (NetRange,OriginAS) Pairs	73,062	
Distinct NH_OAS (NetHandle w/ OriginAS) with no super-prefixes	39,297	
Of these (39,297):		
# of NH_OAS with no sub-prefixes	38,693	0
# of NH_OAS with sub-prefixes (only one level below)	584	16828
# of NH_OAS with sub-prefixes (two levels below)	20	16937

- Note: $38,693 + 584 + 20 + 16828 + 16937 = 73,062$
- Many of the consolidated 39,297 are also subprefixes of what are actually observed

ARIN NetHandle w/ OriginAS

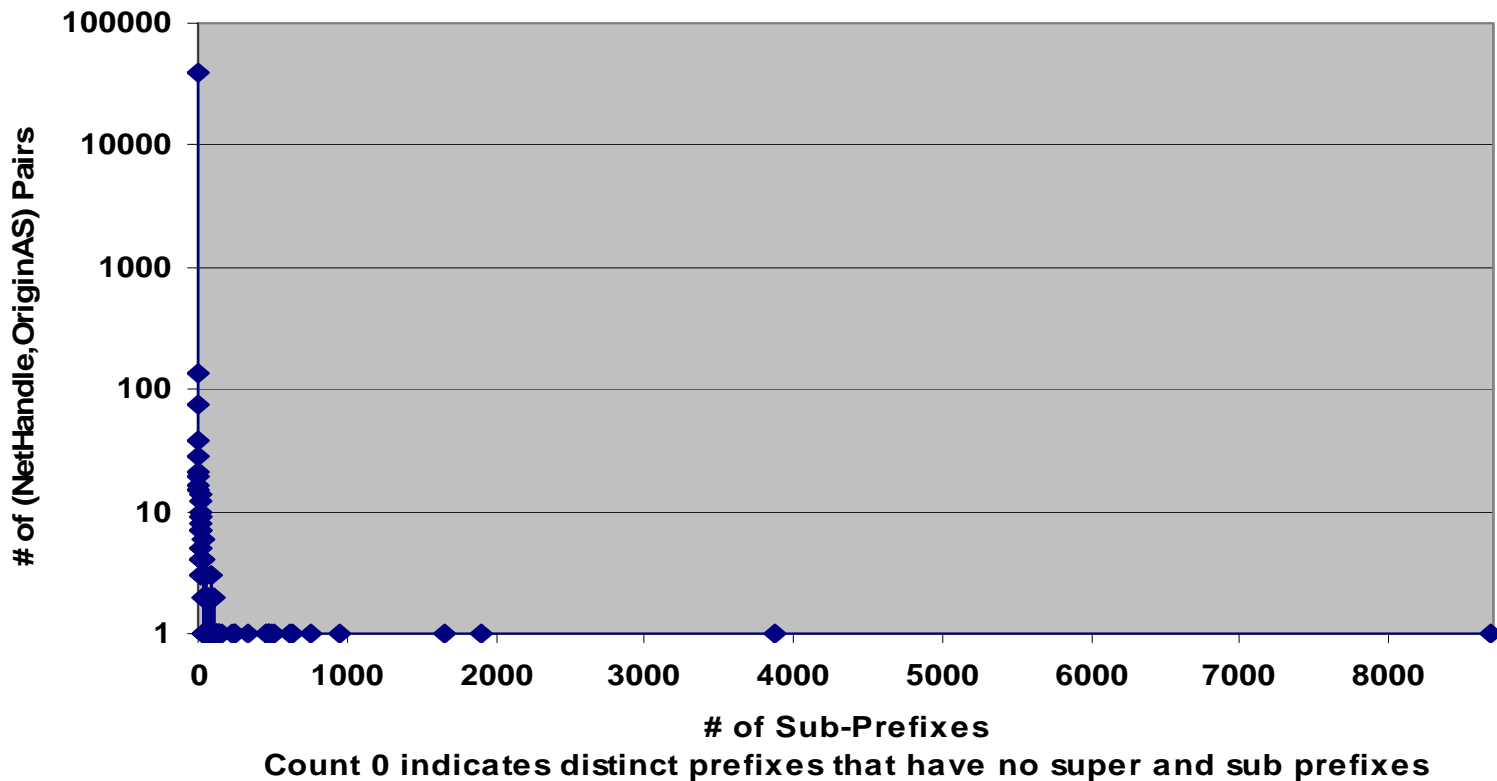
Consolidation of (NetHandle, OriginAS) Pairs

- Example 1:
 - Super prefix: 96.56.0.0/15 (OAS=6128) has **8674** sub-prefixes
 - * First level sub-prefixes: **2600**
 - Prefix lengths ranging from 29 to 20
 - E.g., 96.56.0.0/20, 96.56.32.0/20, 96.57.7.200/29
 - * Second level sub-prefixes: **6074**
 - All prefixes that are sub-prefixes of those in the first level
- Example 2:
 - Super prefix: 64.25.208.0/20 (OAS=20225)
 - First level sub-prefixes: **30**
 - * Prefix lengths ranging from 24 to 29

Distribution of Subprefix Count from Consolidated (NetHandle, OriginAS) Pairs

Registry Data Date: 2008-10-18

Trustworthy Networking Program



ARIN NetHandle w/ OriginAS Consolidation of NetHandle Only

On 2008-10-18	Total	# sub-prefixes
All unique NetHandles	70,173	
Distinct NetHandles with no super-prefixes	35,510	
Of these (35,510):		
# of NetHandles with no sub-prefixes	34,890	0
# of NetHandles with sub-prefixes (only one level below)	594	15,666
# of NetHandles with sub-prefixes (two levels below)	26	18,997

- Here we are examining the cases when a prefix may be consolidated into a superprefix while allowing for the origin AS to be the same or different

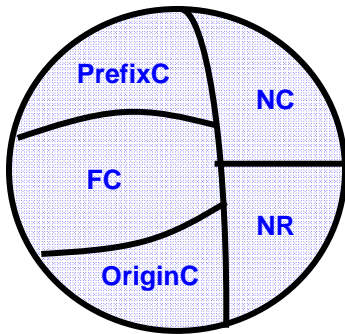
Outline

- Overview of registry information quality analysis
 - Problem space and goals
 - Routing and addressing registries
 - Consistency check algorithm
 - Information quality metrics
- Analysis of ARIN NetHandles with OriginAS
 - Some observations on ARIN NetHandle data
 - WRT route objects in RPSL
 - WRT BGP trace data
 - Aggregation (consolidation) of ARIN NetHandle address space with OriginAS
- Characterization of self-consistency for registry information

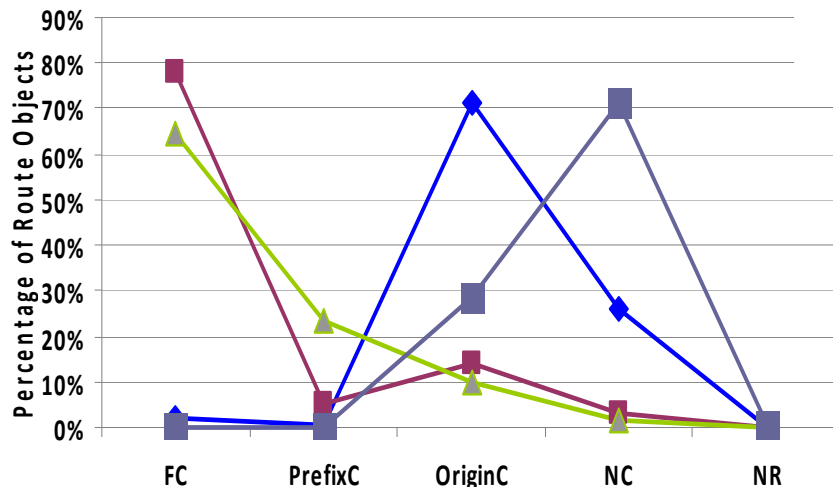
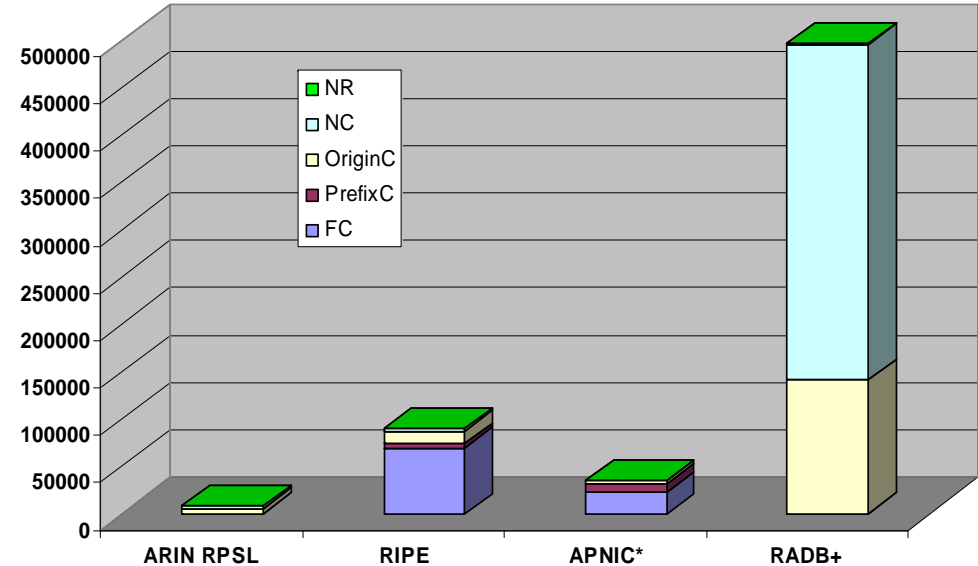
Characterization of IRR Consistency Based on Route Object Registrations

Registry Data Date: 2008-10-18

Registry Data



- **FC:** Fully (Prefix & Origin) Consistent
- **PrefixC:** Only Prefix Consistent
- **OriginC:** Only Origin Consistent
- **NC:** (referenced objects exist, but) Not Consistent
- **NR:** No Referenced Resource Objects Exist

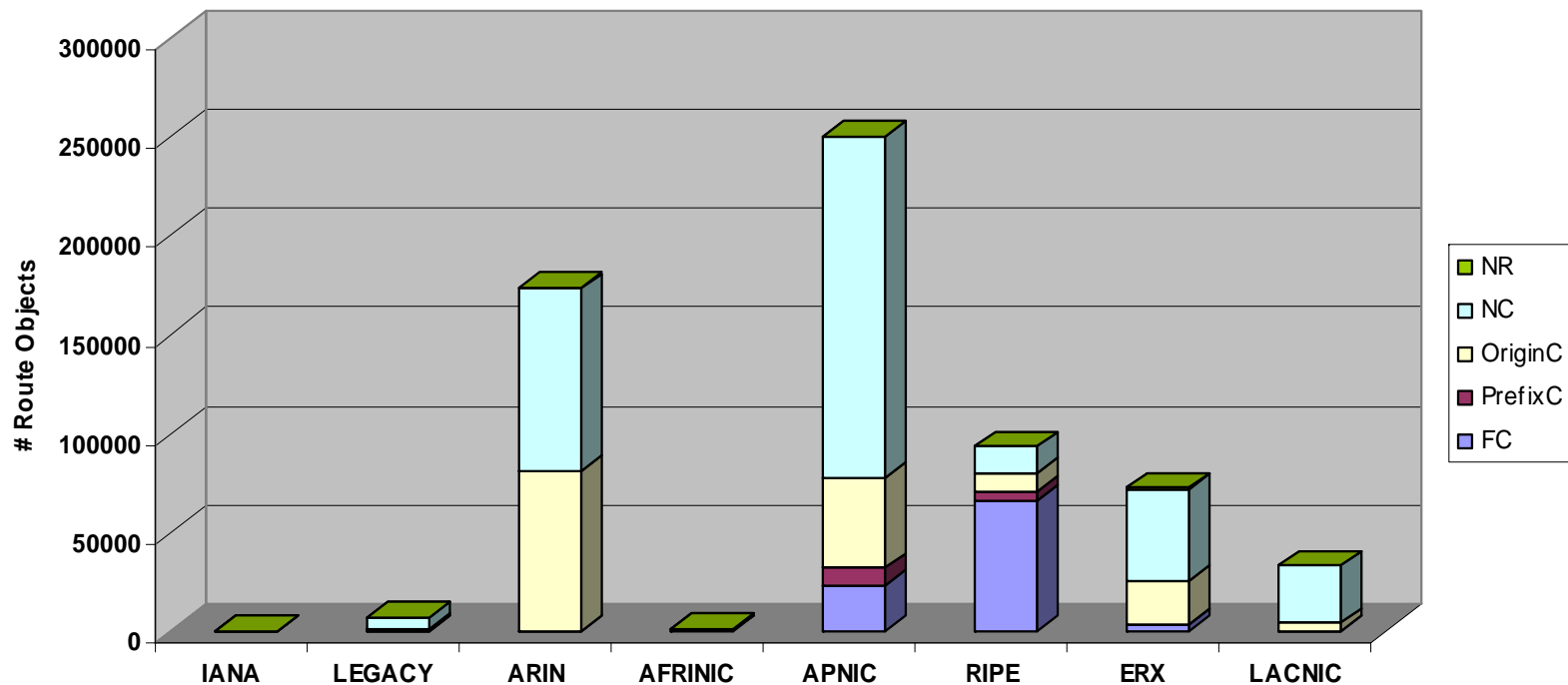


	ARIN RPSL		RIPE		APNIC*		RADB+	
FC	169	2%	70057	78%	22981	65%	534	0%
PrefixC	27	1%	4458	5%	8364	23%	107	0%
OriginC	5845	71%	12627	14%	3562	10%	141323	29%
NC	2147	26%	2815	3%	608	2%	353598	71%
NR	13	0%	0	0	0	0	1534	0%
Total	8201		89957		35515		497096	

Scores for Consistency Checks for IRR Route Objects 35

Characterization of IRR Consistency vs. Source of Prefix Allocations

Registry Data Date: 2008-10-18

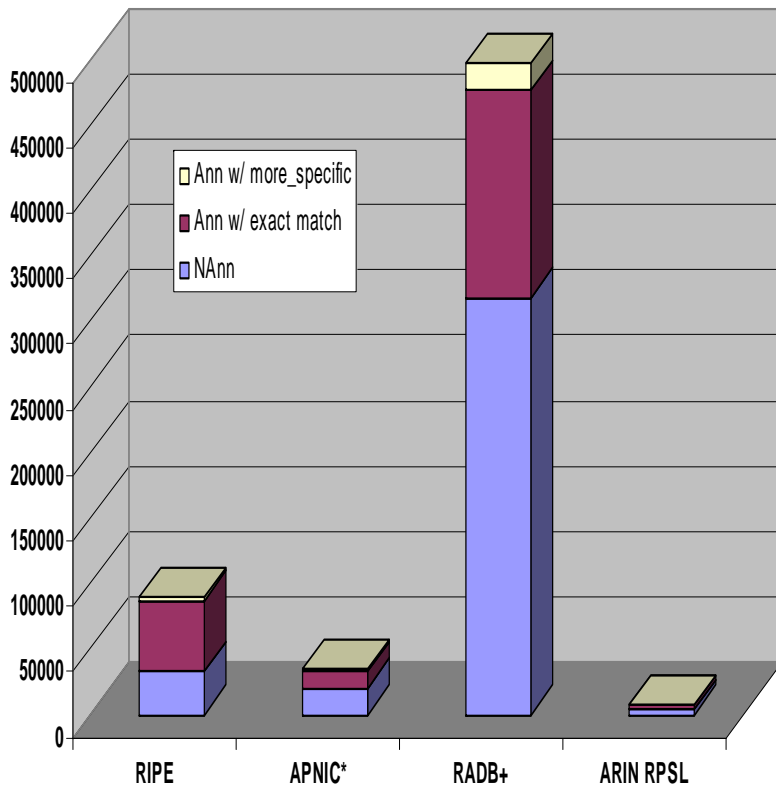


FC: Fully (Prefix & Origin) Consistent
PrefixC: Only Prefix Consistent
OriginC: Only Origin Consistent
NC: (referenced objects exist, but) Not Consistent
NR: No Referenced Resource Objects Exist

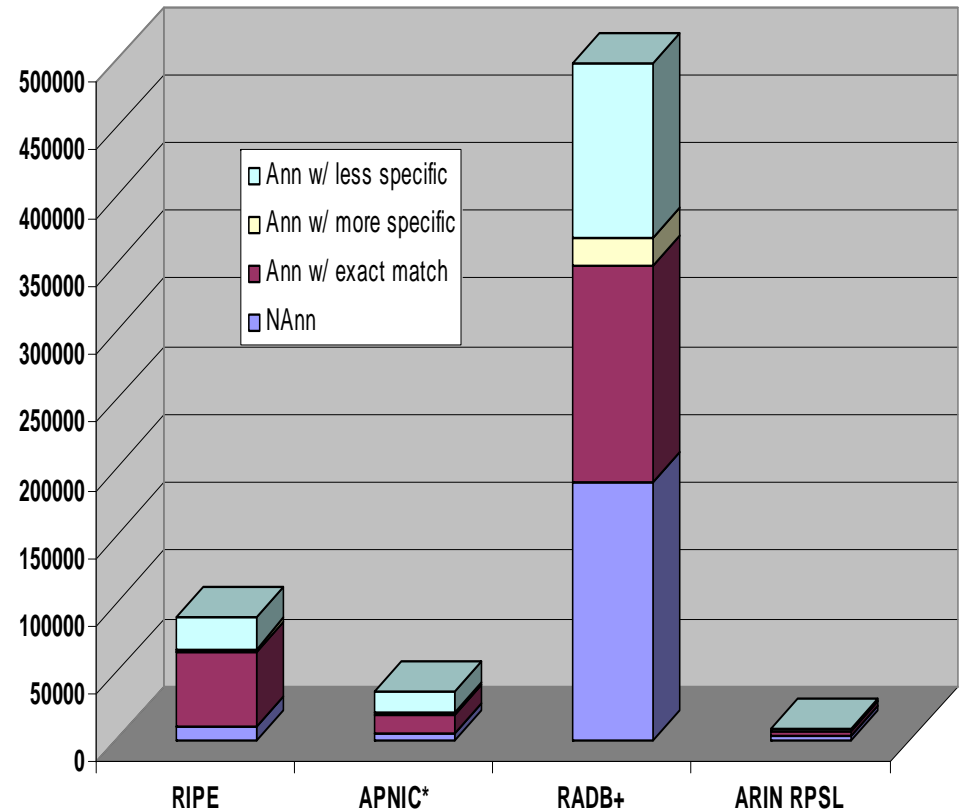
- Note: Many of the route objects are registered in IRRs run by non-RIRs. Most prefixes of these route objects are allocated by either ARIN or APNIC.

IRR Route Objects vs. BGP Trace Data

Route objects are observed with exact match or more specific prefixes with the same originAS in trace data



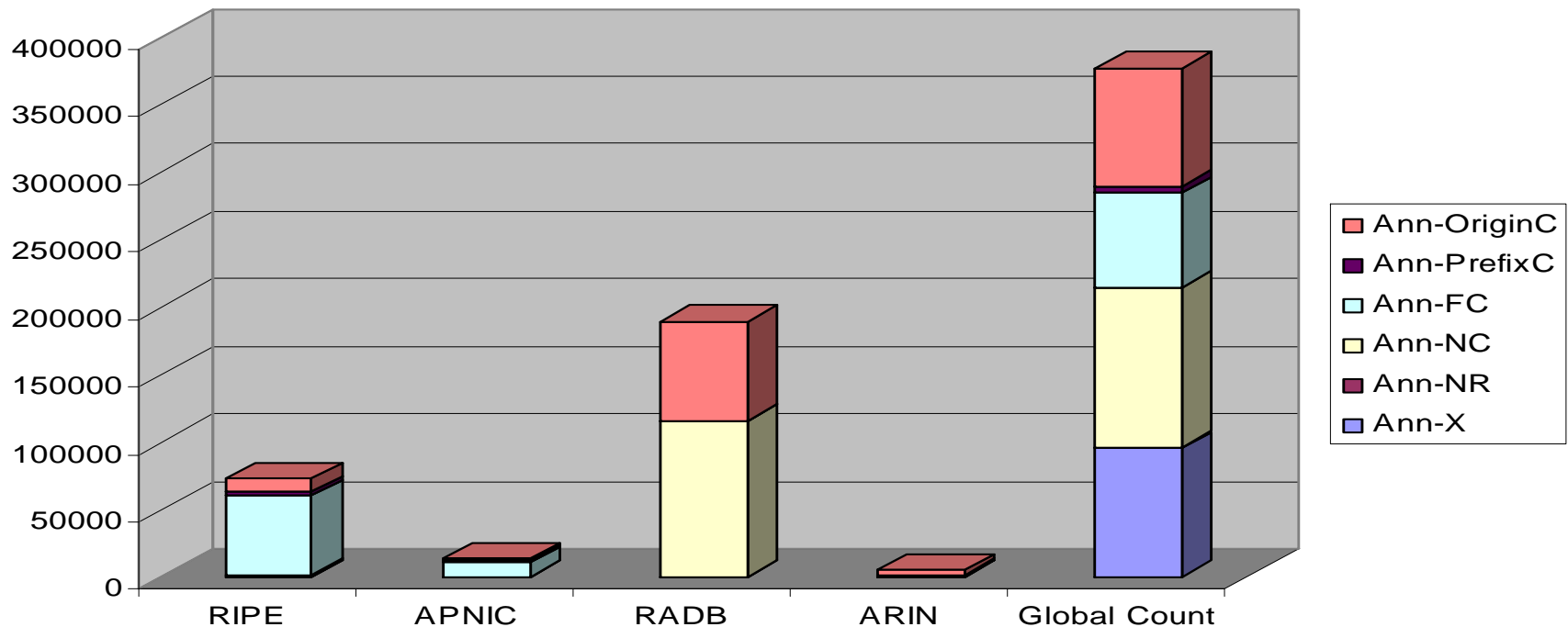
Route objects are observed with exact match, more specific or less specific prefixes with the same originAS in trace data



Registry Data Date: 2008-10-18; Trace Data from 2008-06-01 to 2008-09-15

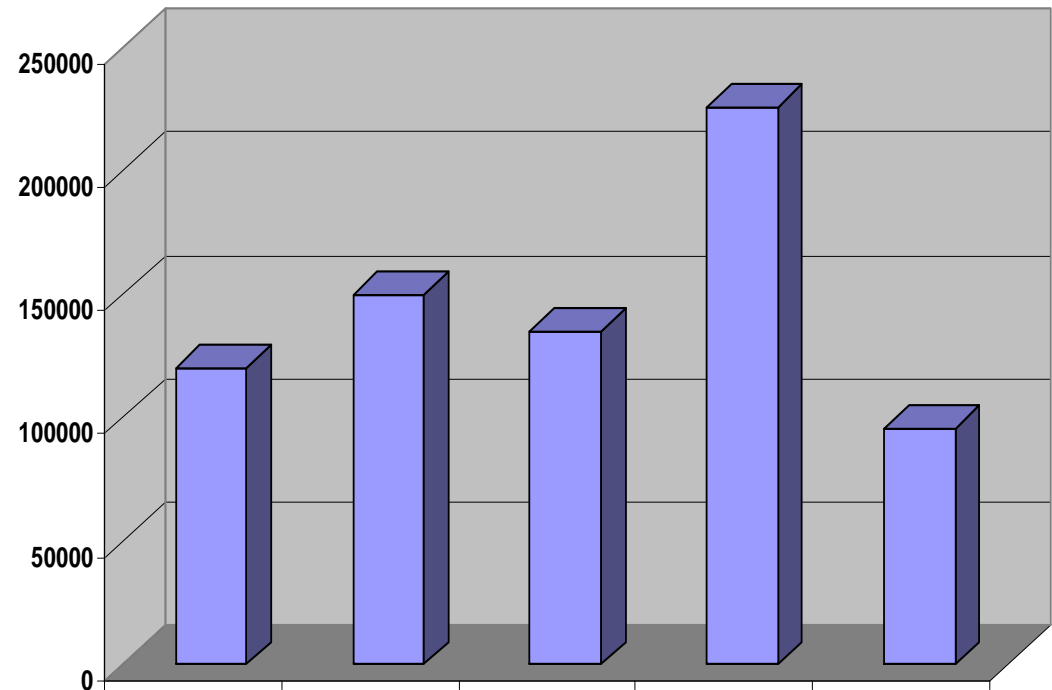
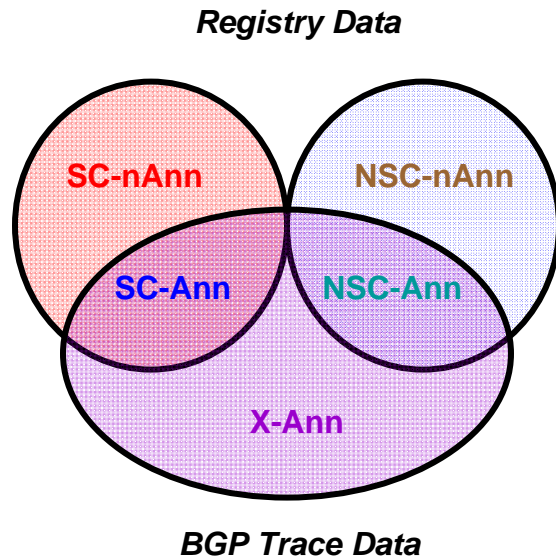
Observed (prefix, originAS) Pairs vs. Registered Route Objects

Registry Data Date: 2008-10-18; Trace Data from 2008-06-01 to 2008-09-15



- **Ann-OriginC:** Announced (prefix, originAS) pairs are only origin consistent with corresponding route objects
- **Ann-PrefixC:** Announced pairs are only prefix consistent with corresponding route objects
- **Ann-FC:** Announced pairs are fully consistent with corresponding route objects
- **Ann-NC:** Announced pairs are not consistent with corresponding route objects
- **Ann-NR:** Announced pairs have corresponding route objects, but no referenced objects in registry data
- **Ann-X:** Announced pairs have no existing corresponding route objects

Measuring Correctness / Completeness of Registry Data WRT BGP Trace Data



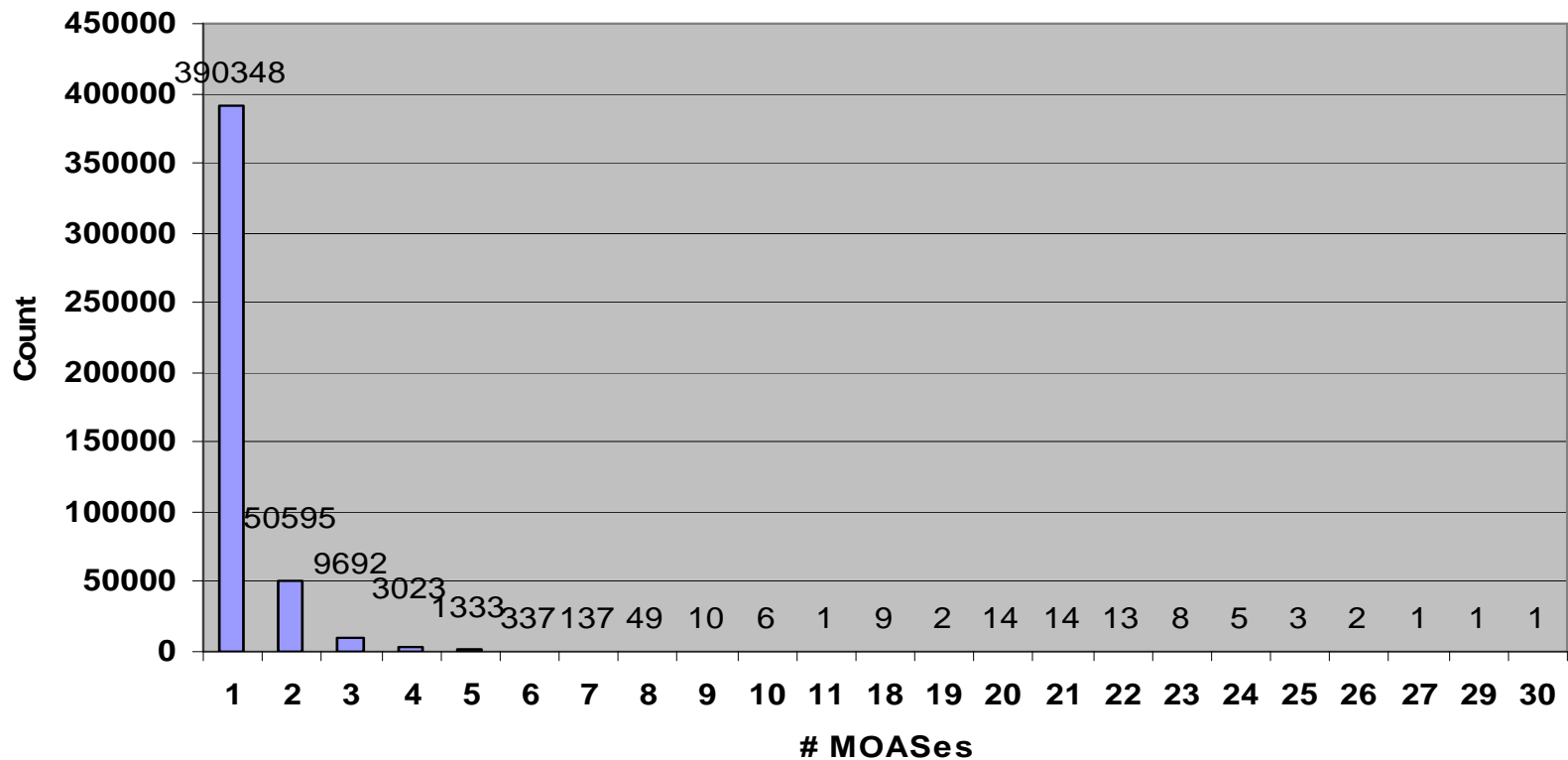
Metrics	Registered route objects in the Registry	(prefix, originAS) Pairs in Trace Data
SC-Ann	Self-consistent	Observed matching exact match or more specific prefixes with same origin
SC-nAnn	Self-consistent	Not observed
NSC-Ann	Not self-consistent	Observed matching exact match or more specific prefixes with same origin
NSC-nAnn	Not self-consistent	Not observed
X-Ann	No registered route objects	Observed (prefix,originAS) pairs

SC_Ann SC_nAnn NSC_Ann NSC_nAnn X_Ann

All global registered route objects: 630,797
All unique observed (prefix, origin) pairs: 376,836

Distribution of Multiple OriginAS (MOAS) in Registered Route Objects

Registry Data Date: 2008-10-18



Conclusions and Future Work

- Presented an overview of quality analysis of registry information
 - Completeness and correctness of existing registry data
 - Characterization of consistency of global IRR data with regard to BGP trace data
 - Analysis of ARIN NetHandles with OriginAS
- Immediate further analysis:
 - Use of extensive BGP trace data from multiple collectors
 - Detailed analysis of registered route objects that are not yet observed
 - Analysis of unobserved ARIN NetHandles with OriginAS whose prefixes are seen with different OriginAS
 - Age distribution of quality of registered route objects
 - * Quality of route objects relative to their longevity in the registry
- Help improve quality and completeness of routing data by providing characterization of the data over time
 - Monthly score cards of completeness and correctness of routing data