NIST National Institute of Standards and Technology

INFORMATION TECHNOLOGY LABORATORY

Trustworthy Networking Program

Characterization of the "Size and Shape" of Static RPKI

Okhee Kim, K. Sriram, Oliver Borchert, and Doug Montgomery

Contact: {okim, ksriram, borchert, dougm}@nist.gov

December 7, 2010

Outline

- Goals
- Methodology
- Quantitative analysis of the current resource allocations
- Characterization of static RPKI
- Conclusions

NIST's Goals

- Develop models of the "size and shape" of a potential global RPKI structure from existing RIR/IRR databases.
- Provide quantitative analyses of the scalability and the potential performance impact of global-scale deployed RPKI on routing dynamics.
- Study the potential future changes in routing information infrastructure.
- Evaluate how such issues as IPv4 address exhaustion will impact on the deployed RPKI.
- Assess the potential load and weaknesses of the "moving parts" of the proposed RPKI infrastructure.

Methodology (1 of 2)

- **Use NIST TERRAIN DB data:**
  - Global bulk Whois databases:
    - * 5 RIRs and IRRs from the RADB site.
  - BGP trace data:
    - * RIPE NCC and Route Views.
- **Develop models of the potential global RPKI infrastructure:**
  - Select all distinctively registered objects.
  - For multiple registrations across RIRs:
    - * Select one from a RIR where the resource is allocated to, if exists.
    - * If not, select one arbitrarily among RIRs/IRRs.
    - * *For APNIC, the same resource may be registered in different registries such as RIR and/or NIR. In this case, select one that contains the "status:" attribute.*
  - Build number resources (IPv4 and ASN) structures describing allocation chains.
  - Classify selected objects per region based on IANA allocation registries:
    - * ARIN / RIPE / APNIC / AFRINIC / LACNIC / LEGACY / ERX.

Methodology (2 of 2)

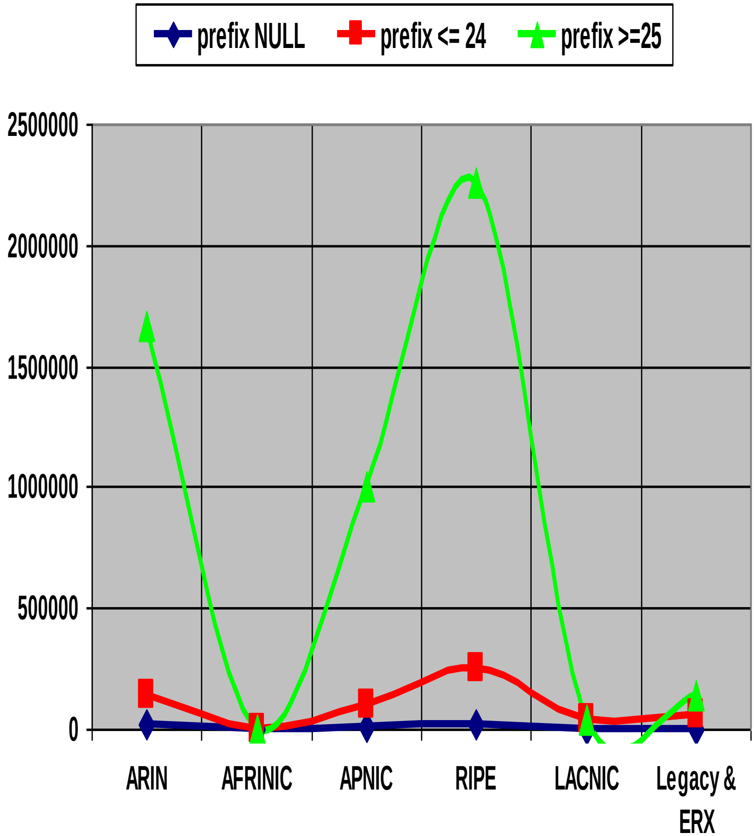
- Details of building number resources structures:
  - ASNs:
    - For SWIP:
      - Distinct ASHandles.
      - Distinct ASNs (aut-nums) registered in RPSL (i.e., aut-num), which are assigned to ARIN but not registered in SWIP (as either a single ASN or AS range).
    - For RPSL:
      - Unique aut-nums.
      - as-block objects that contain a range of ASNs in RPSL. *Note that some as-blocks contain a single ASN (e.g., ASn – ASn), most of which have corresponding either aut-num or ASHandle objects.*
  - IPv4 addresses:
    - Globally distinct inetnums in RPSL and NetRanges in SWIP.
    - For multiple registrations, select one from a RIR where the resource is allocated to, if exists.
    - If not, select one arbitrarily among RIRs/IRRs.
    - *Partial registrations from a /8 block may be found in other RIRs but they are considered to belong to the same RIR where the /8 is allocated*
    - *Exceptions in LEGACY/ERX IP address space:*
      - *The LEGACY/ERX blocks may contain a large number of cross-RIR partial allocations, especially between RPSL and SWIP. These partial allocations are combined before processing.*
      - *Example: If 129.1/16 registered in RIPE (RPSL) and 129.2/16 registered in ARIN (SWIP), then both 129.1/16 and 129.2/16 are considered as LEGACY/ERX.*

ERX Partial Allocations Examples

- 129/8: currently administered by ARIN:
  - Partial allocations in SWIP: 396
  - Partial allocations in RPSL: 592
  - Multi registrations in both SWIP and RPSL: 30
- 151/8: currently administered by RIPE NCC:
  - Partial allocations in RPSL: 6,999
  - Partial allocations in SWIP: 2,084
  - Multi registrations in both SWIP and RPSL: 15
- 198/8: currently administered by ARIN
  - Partial allocations in RPSL: 320
  - Partial allocations in SWIP: 15,760
  - Multi registration in both SWIP and RPSL: 63

Distribution of Registry IPv4 Address Allocations/Assignments

Registry data date: 2009-02-18



Unit: 1k objects

RIR	# of /8 blocks	# objects p_len=NULL*	# objects p_len <= 24	# objects p_len >= 25	Total # objects
ARIN	31	17	145	1,667	1,829
RIPE	28	24	248	2,262	2,534
APNIC	30	13	100	1,004	1,117
AfriNIC	2	0	1	5	6
LACNIC+	6	0	36	40	76
LEGACY/ ERX+	92	4	59	144	207
Total	189	58	589	5,122	5,769

* Prefix Length NULL indicates that an address block cannot be represented by a single CIDR.
+ from both RPSL and SWIP except duplicates.

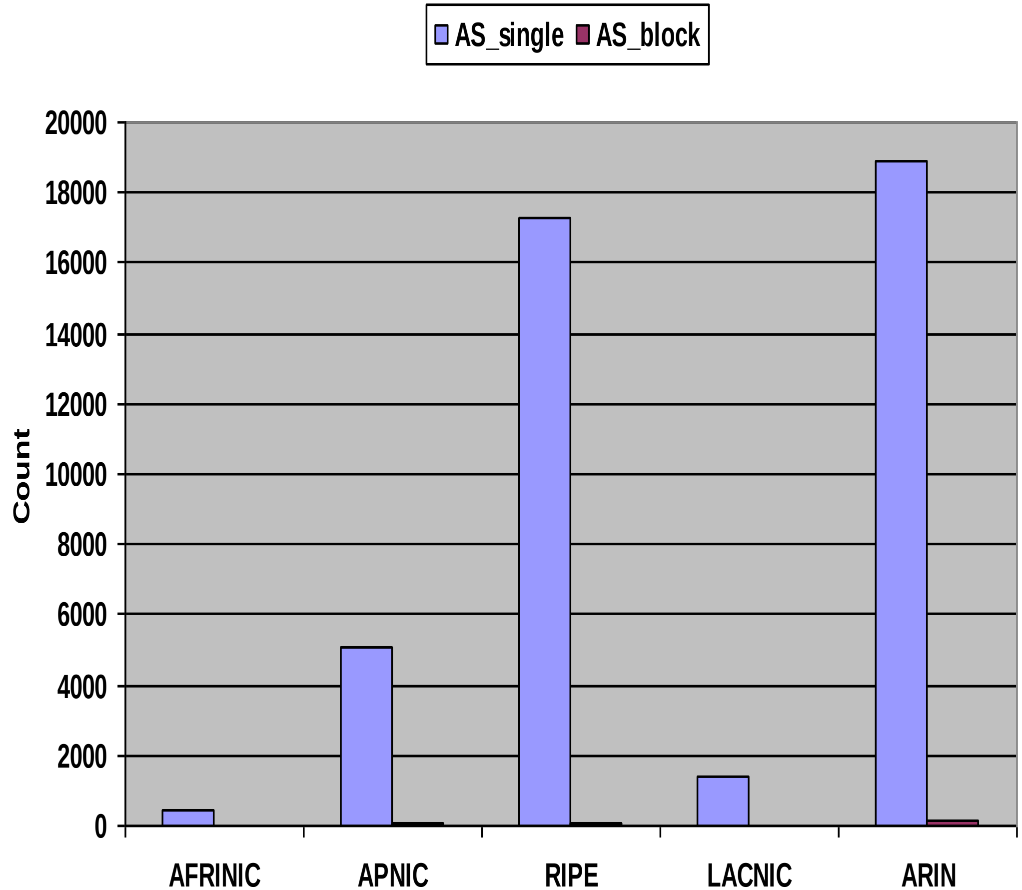
As of August 2010, 14 /8 blocks are unallocated.

Distribution of Global ASN Assignment

Based on IANA and RIR/IRR Datasets

Registry data date: 2009-02-18

RIR	AS single	AS block
ARIN	18,862	137
RIPE	17,280	59
APNIC	5,082	70
AfriNIC	406	4
LACNIC	1,391	2
Total	43,021	272

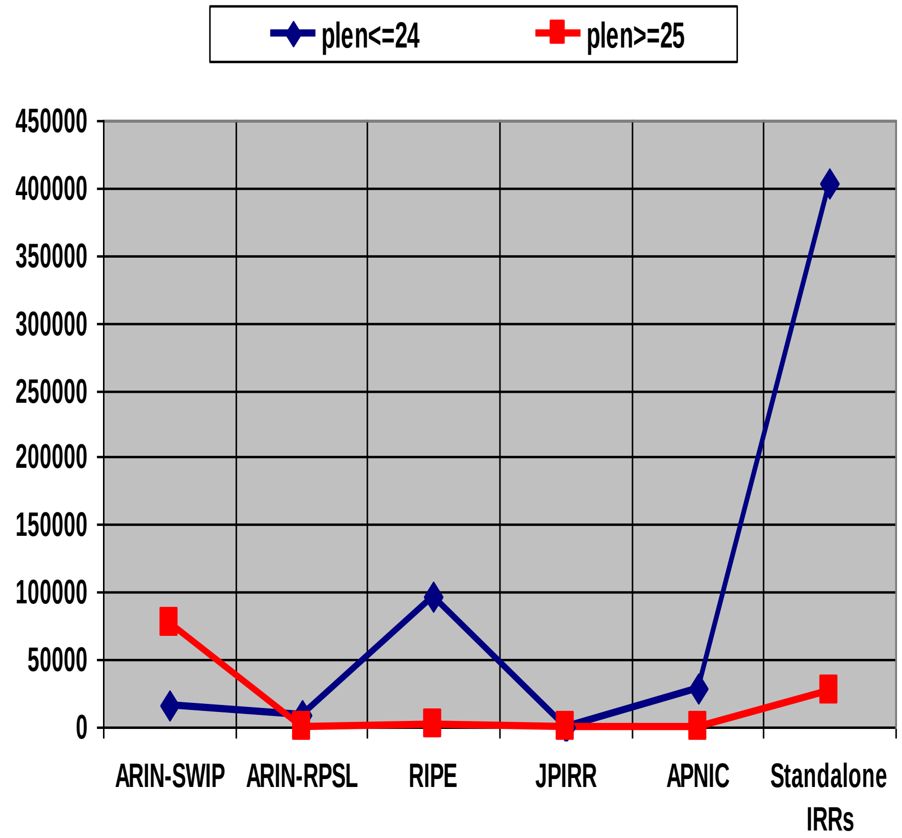


Distribution of Potential ROAs
Based on Route Object Registrations

Registry data date: 2009-02-18

Unit: 1k objects

RIR	# objects p_len <= 24	# objects p_len >= 25	Total # objects
ARIN-SWIP	15.5	77	92.6
ARIN-RPSL	8.2	0.2	8.4
RIPE	96.4	1.5	97.9
JPIRR	0.6		0.6
APNIC	28	0	28
Standalone IRRs +	403	27.3	430.3
Total	551.7	106	657.8



+ Standalone IRRs includes all individual IRRs mirrored from the RADB site.

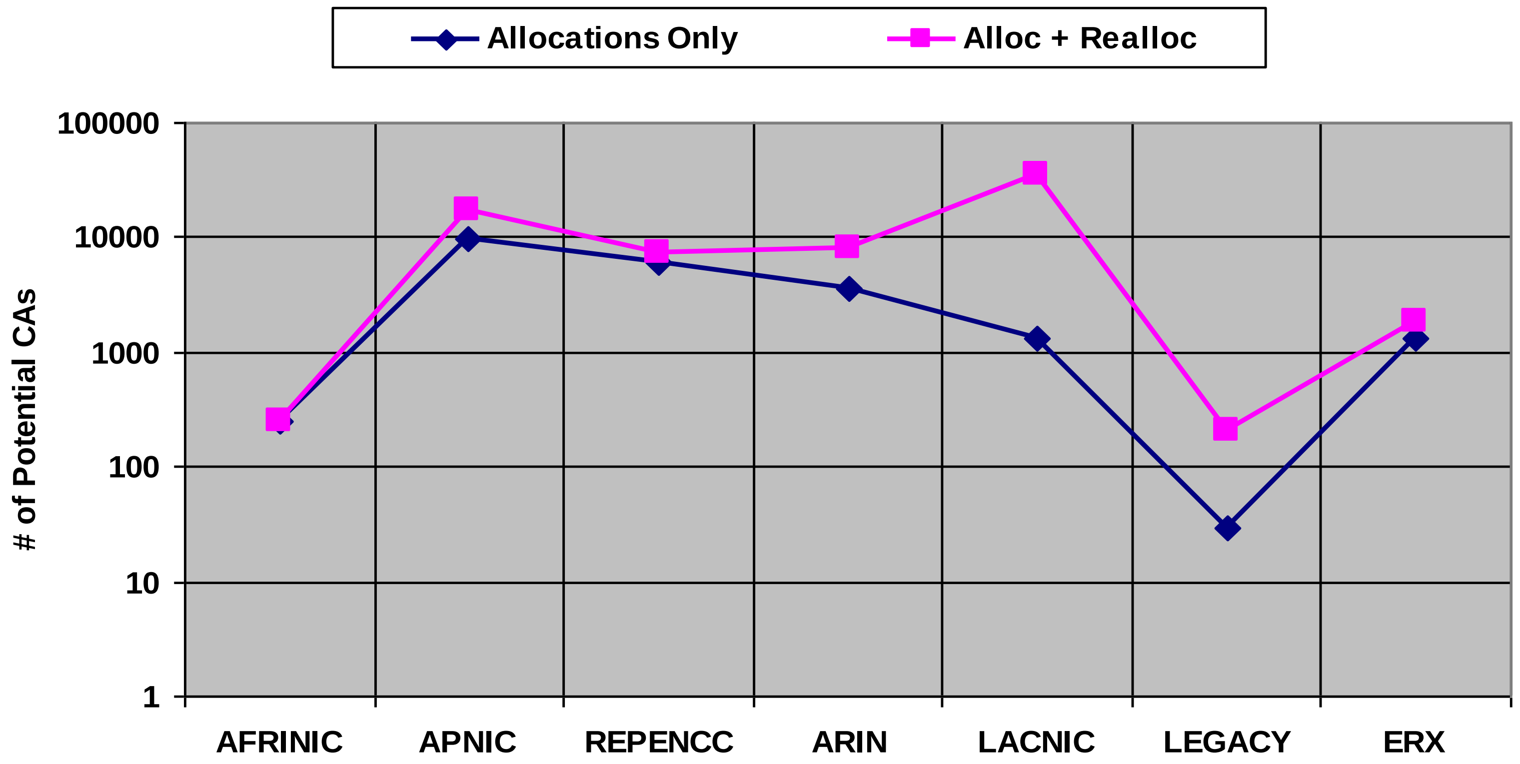
Characterization of the of static RPKI

- Analysis of potential CAs:
  - Distribution of potential CAs per RIR
  - Distribution of CA path depths per RIR
- Analysis of IPv4 certificates:
  - Full deployment vs. optimized deployment
  - IPv4 prefix lengths vs. IPv4 certification path depths
- Analysis of ROAs:
  - The cost estimate of ROA verifications in terms of certification path lengths
  - Distribution of PI address space.
  - Analysis of MOASes of potential ROAs

Potential CAs

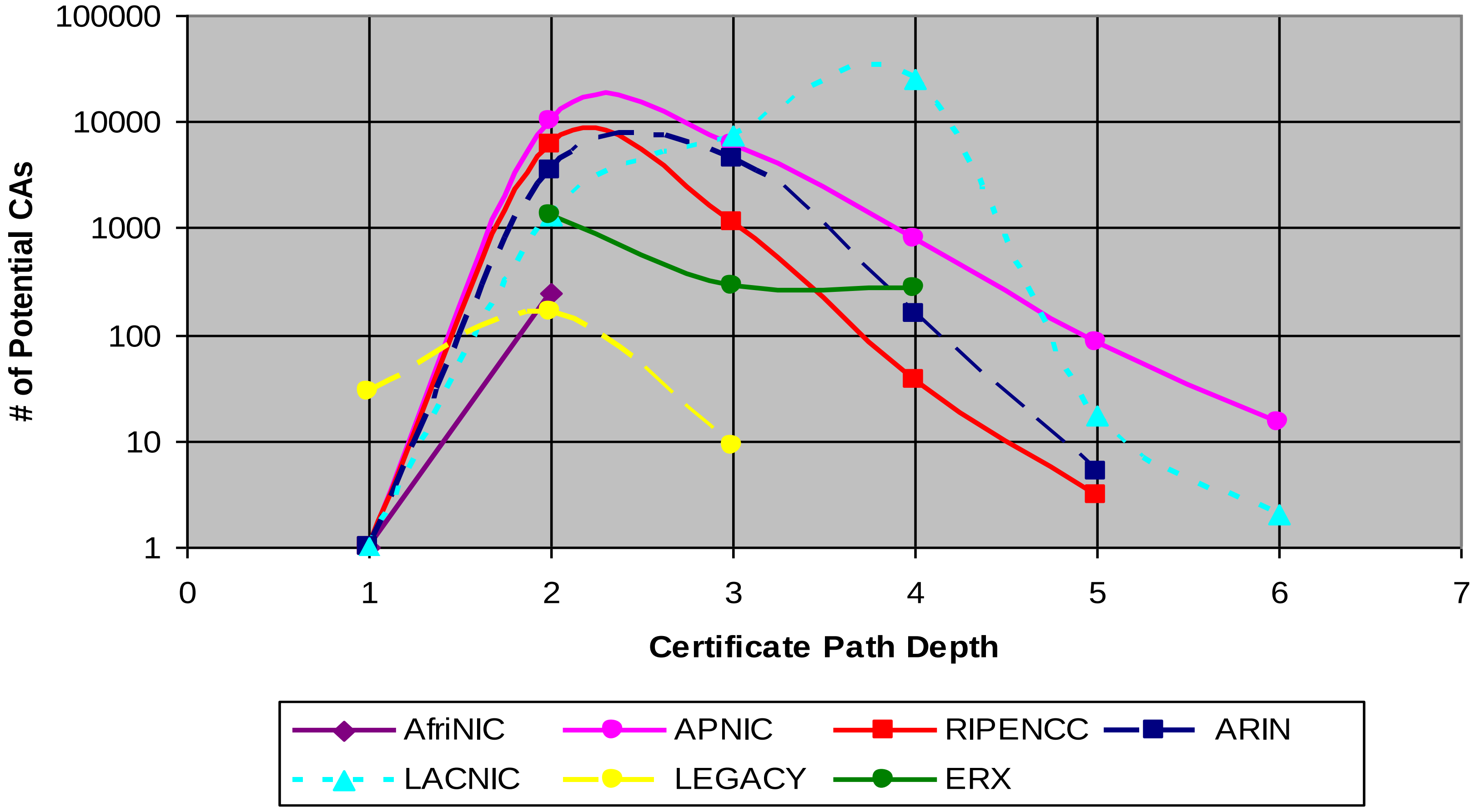
- Selection criteria:
  - Resource allocation objects:
    - * inetnums in RPSL.
    - * NetHandles in SWIP.
  - Attributes contained in an object to identify the allocation type:
    - * "status:" in inetnum.
    - * "NetType:" in NetHandle.
  - Status/NetType Attribute values: Allocation, Re-allocation
    - * First consider "Allocation" ONLY (including both PA and PI)
    - * Then consider "Allocation" and "Re-allocation"
  - Five top level CAs: ARIN, RIPE NCC, APNIC, LACNIC, AfriNIC in addition to IANA
    - * For blocks with prefix length **<=** 8, the certificates are created by the RIRs
    - * For these blocks, the RIRs are the CAs
  - Eliminate also objects whose size < 255 (i.e, more specific than /24)
- Algorithm for selecting potential CAs:
  - Legacy:
    - * If Org of an object is uniquely defined and the object is either
      - Direct assignment (/8) to an organization; OR
      - Allocation to an ISP under Legacy space (e.g., 4/8 and 8/8 are allocated to Level 3 Comm).
  - Regular allocations and ERX:
    - * If Org of an object is uniquely defined AND the object is allocation (or, reallocation) regardless of the allocation depths

Distribution of Potential CAs per RIR



- # of potential global CAs (allocations only): ~22.4K
- # of potential global CAs (alloc + realloc): ~69.1K
- Note that AfriNIC, APNIC and RIPE NCC do not have the value “re-allocation”. Hence, the first level of direct allocations by these RIR is considered as “Allocation Only”.
- Note also that some objects do not contain “org:” attribute, especially for the regions such as RIPE NCC and APNIC.

Distribution of Certificate Path Depths of potential CAs (Alloc + Realloc)



- LACNIC, LEGACY and ERX Data are selected from both RPSL and SWIP excluding duplicates.
- Certification path depth "1" indicates the top-level allocations by IANA to RIRs, i.e., address blocks >= /8.

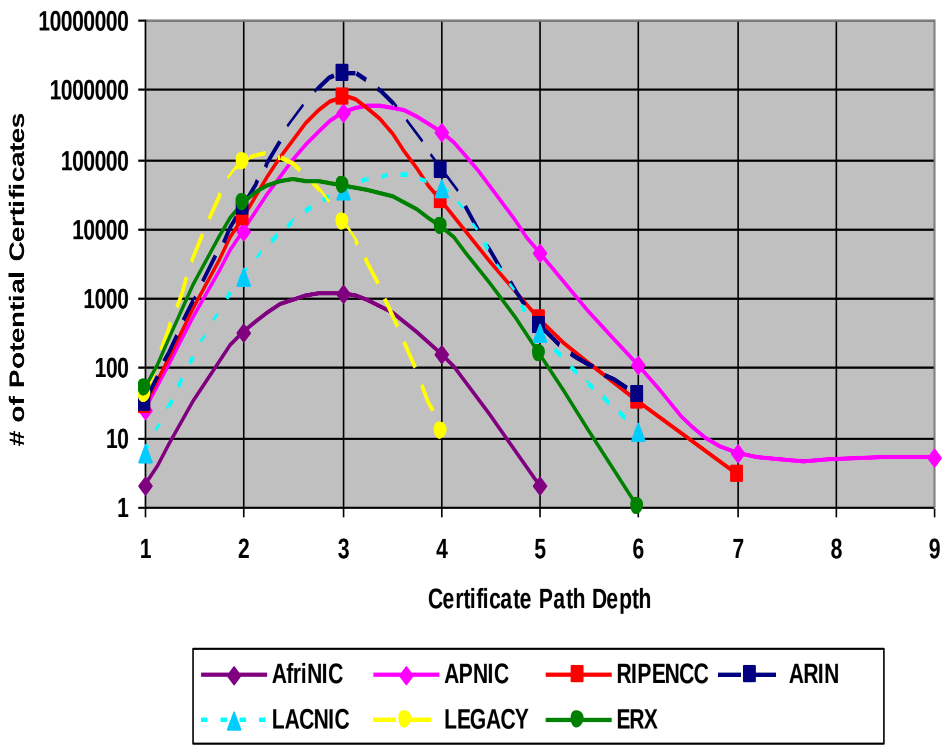
Analysis of IPv4 certificates

Trustworthy Networking Program

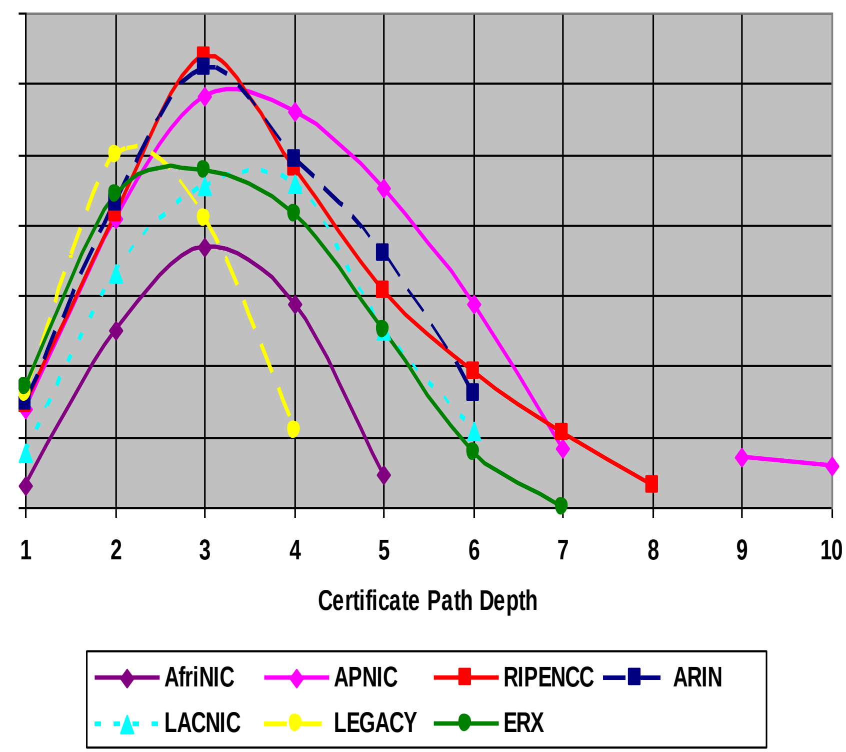
- Full deployment vs. optimized deployment:
  - Full deployment: if it was currently deployed based on the registry allocation data.
  - Optimized deployment after IPv4 prefix optimization:
    - Aggregation of adjacent equal length prefixes
- Algorithm for IPv4 prefix optimization:
  - For every possible aggregate (i.e., two adjacent, equal sized, aggregatable prefixes), check the following attributes:
  - If organizations in the two objects are defined and the same, aggregate the two.
  - Else if *organizations* in the two objects are defined but different, do not aggregate the two.
  - Else if both or either one of the two contain no organization, then:
    - If both *country code* and *status (e.g., PI vs. PA and allocation vs. assignment)* between the two are the same:
      - Check ***mntner-related attributes*** (i.e., mnt-by, mnt-lower, mnt-routes) between the two.
      - If check passes, then aggregate the two.
- Create a new aggregate, if no existing prefix for the aggregate exists, as follows:
  - Aggregated by org:
    - generate a new aggregate with the *org/status* values of the first prefix without mnt values.
  - Aggregated by mnt:
    - Generate a new aggregate with the *country/status/mnt* values of the first prefix excluding org.

Distribution of IPv4 Certificate Path Depths

Optimized deployment



Full deployment



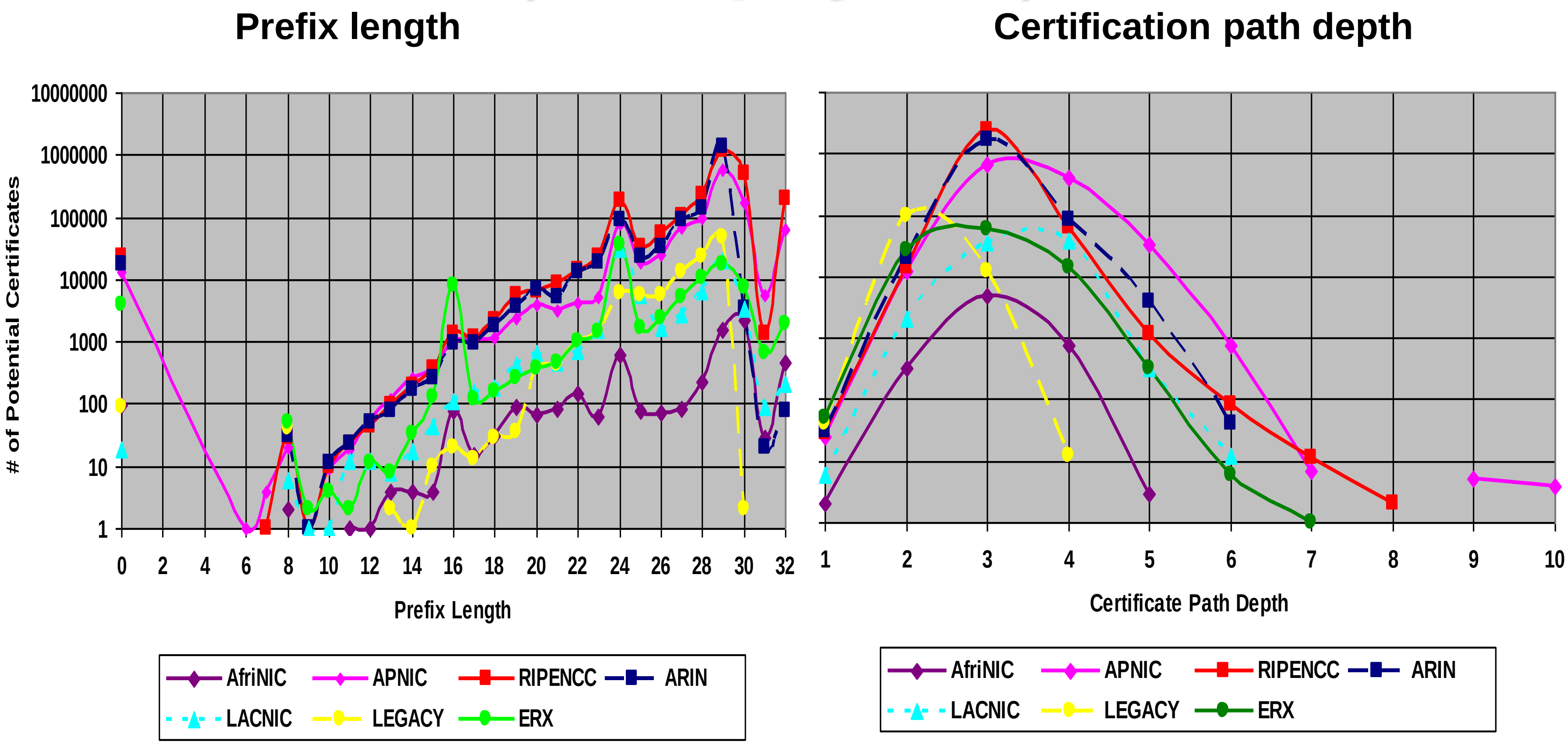
- **LACNIC, LEGACY and ERX data are selected from both RPSL and SWIP excluding duplicates.**
- **Prefix length "0" indicates that an address block cannot be represented by a single CIDR prefix.**
- **Certification path depth "1" indicates the top-level allocations to RIRs by IANA, i.e., address blocks >= /8. Each ">= /8" block is counted separately.**

Improvement from optimization for IPv4 Certificates

	All objects			# objects with prefix length <= /24			# objects with prefix length >= /25		
	Full deployme nt	Optimized deployment	Reduction rate	Full deployment	Optimized deployment	Reduction rate	Full deployment	Optimized deployment	Reduction rate
RPSL	3,733K	1,598K	57%	385K	245K	36%	3,311K	1,316K	60%
SWIP	1,829K	1,816K	0.7%	145K	137K	6%	1,667K	1,662K	0.3%
LEGAC Y/ERX	207k	178K	14%	59K	48K	19%	144K	126K	13%
Global	5,769K	3,592K	38%	589K	430K	27%	5,122K	3,104K	39%

Prefixes with prefix length NULL are not included in this table.

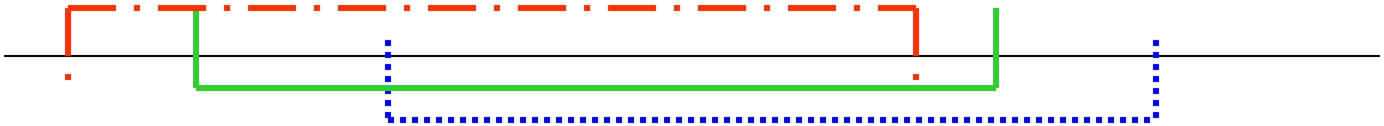
Distribution of Prefix Lengths vs. Certificate Path Depths of IPv4 (full deployment)



- LACNIC, LEGACY and ERX data are selected from both RPSL and SWIP excluding duplicates.
- Prefix length "0" indicates that an address block cannot be represented by a single CIDR prefix.
- Certification path depth "1" indicates the top-level allocations to RIRs by IANA, i.e., address blocks >= /8. Each ">= /8" block is counted separately.

IPv4 Non-contiguous (Overlapping) Sub-allocations in RPSL (examples)

- RIPE:
  - 62.128.192.0 – 62.128.207.255
  - 62.128.195.0 – 62.128.223.255
- APNIC:
  - 211.100.249.184 – 211.100.250.191
  - 211.100.249.192 – 211.100.250.199
  - 211.100.249.200 – 211.100.250.207
  - 211.100.249.208 – 211.100.250.215
  - 211.100.250.216 – 211.100.250.225



Analysis of ROAs

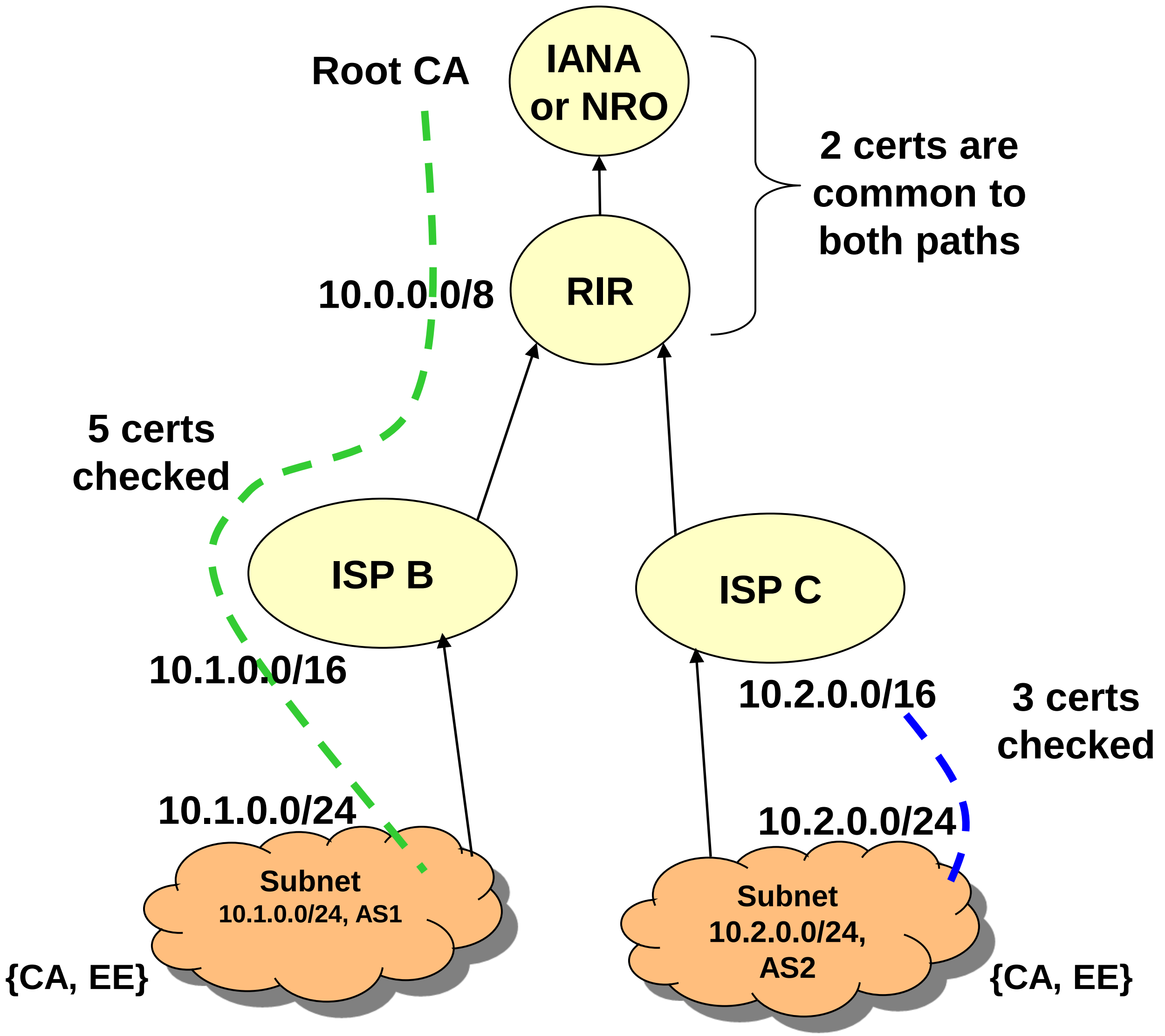
- The cost estimate of ROA verifications in terms of certification path lengths
- Distribution of PI address space
- Distribution of MOASes of potential ROAs

Trustworthy Networking Program

Analysis of ROAs

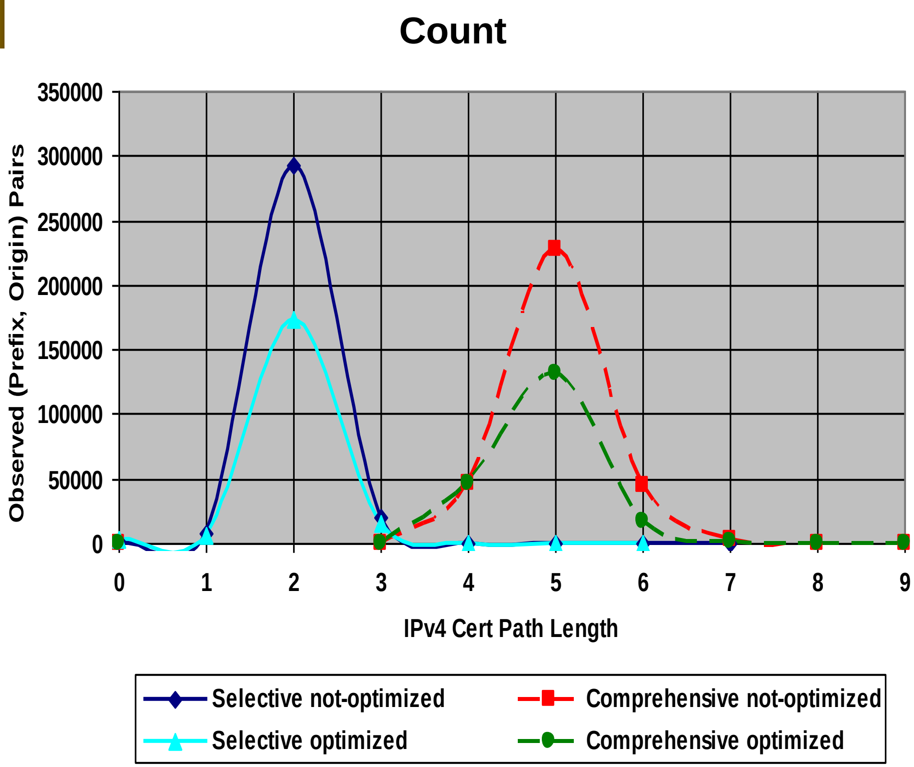
- **ROA analysis techniques:**
  - ROA prefix optimization with the same AS:
    - * *Not optimized:* full-scale
    - * *Optimized*: Aggregation of adjacent equal length prefixes with the same Origin AS
  - ROA prefix verification optimization:
    - * *Comprehensive*:
      - Check every single resource certificate in a certification path including a root.
    - * *Selective*:
      - Use "validation state" of a certificate to avoid redundant checks on the certificates that have already been checked.
- **Categorization** of the ROA verification:
  - *Comprehensive and not-optimized*
  - *Comprehensive and optimized*
  - *Selective and not-optimized*
  - *Selective and optimized*
- **Method** for computing the length of a certification path:
  - Does the prefix of a potential ROA have an exact match resource allocation record?
    - * If yes, then that object is considered as a CA and assume an EE for the prefix is created.
    - * If not, then assume both a CA and an EE for the prefix are created.
  - Assume also that routes with prefix length >= 25 have only the corresponding EEs, not CAs.
  - Compute the number of certificates included in a particular certification path for the EE including a root certificate and a target EE.
  - *IANA or NRO (the top-level entity) is assumed to be a single trust anchor for this analysis.*

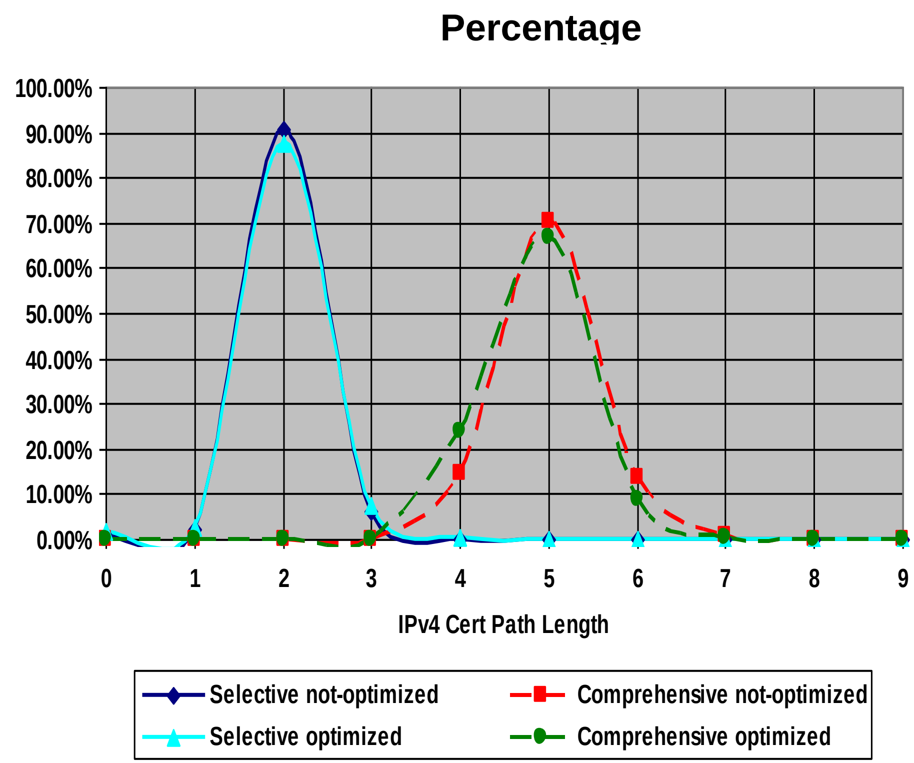
Optimization in ROA Prefix Validation: Selective Method



Distribution of Certification Path Lengths for ROA Prefix Validation

BGP trace data date: February, 2009





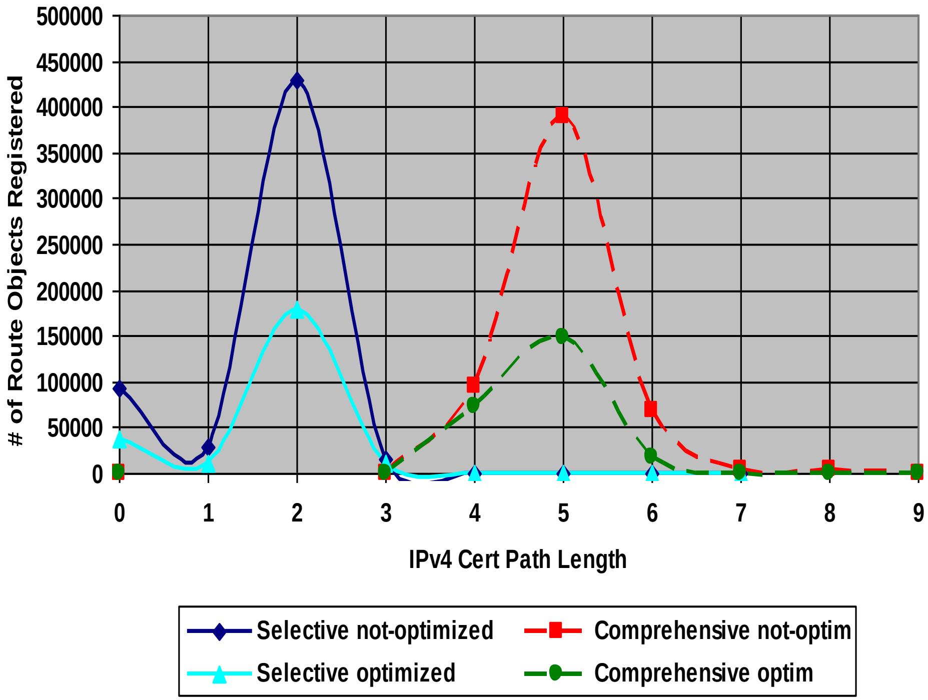
In the case "*Selective and not-optimized*", a realistic scenario for the global-scale deployed RPKI, the average cert. path length for IPv4 address is ~2.03. About 93.6% of observed (P,O) pairs need to verify about two or less IPv4 address certificates for the prefix of a route.

Trustworthy Networking Program

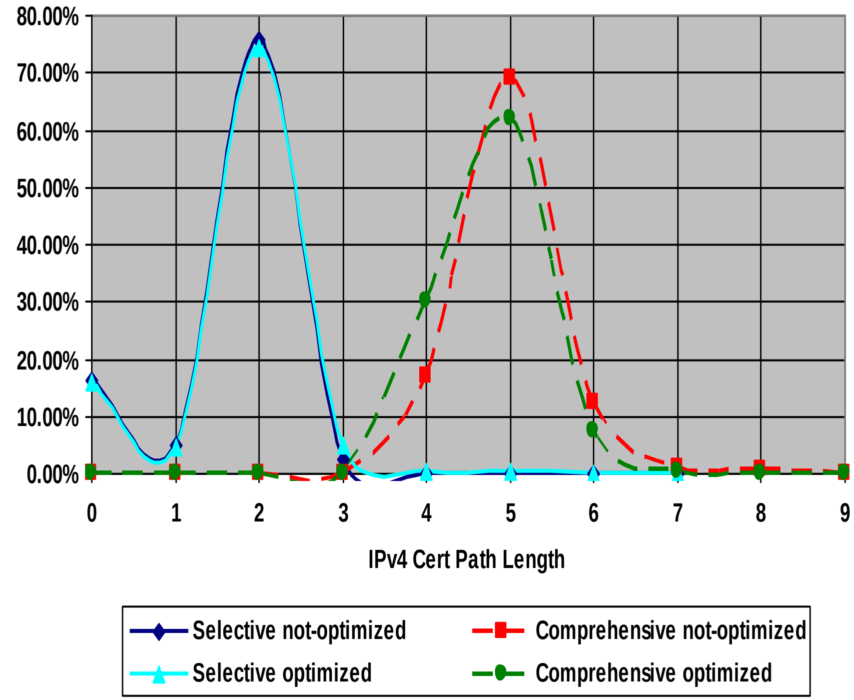
Distribution of Certification Path Lengths for ROA Prefix Validation

RPSL data date: 2009-02-18

Count



Percentage

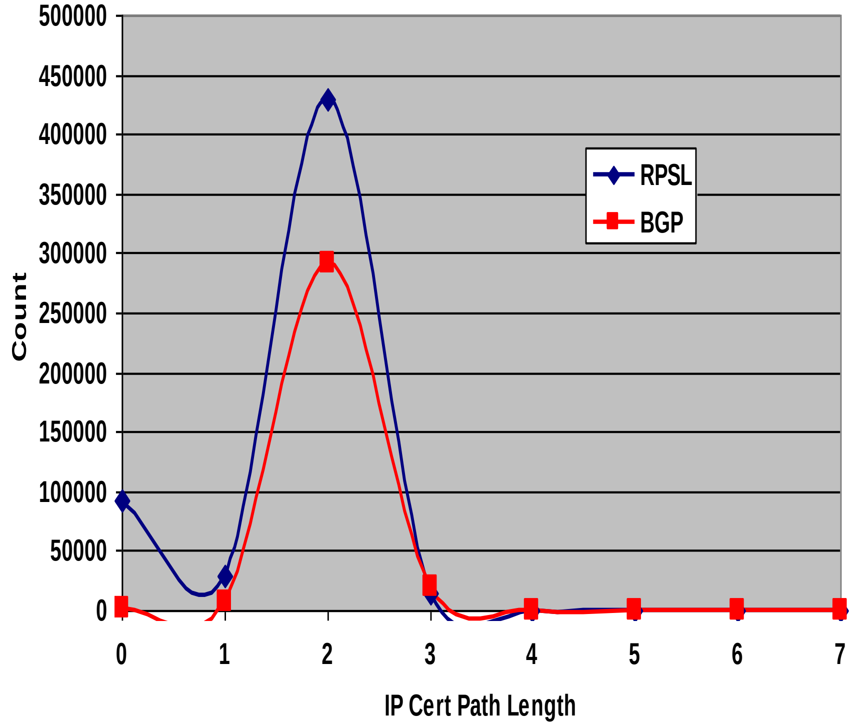


In the case "*Selective and not-optimized*", the average cert. path length for IPv4 address is ~1.7. About 81% of registered route objects need to verify two or less IPv4 address certificates for the prefix and about 16% need not verify the prefix of a route at all (due to multi-homed prefixes).

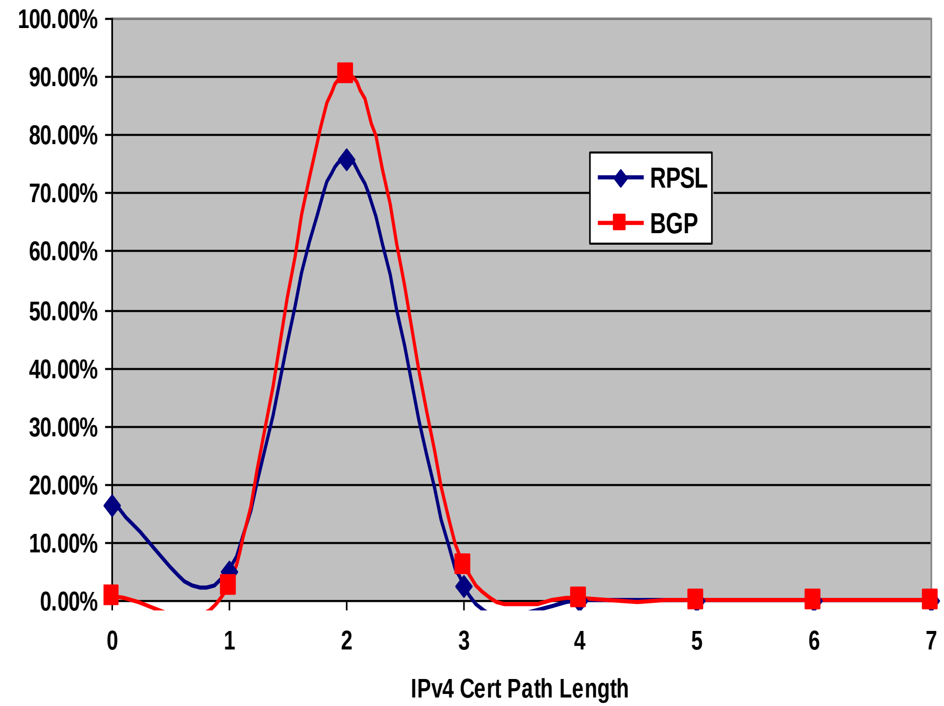
Distribution of Certification Path Lengths for ROA Prefix Validation

selective and not-optimized (RPSL vs. BGP Trace)

Count



Percentage



These graphs depict that the two data sources show similar behavior, i.e., the majority of ROAs (94% for BGP and 97% for RPSL) need to check only 2 or less IPv4 address certificates for ROA validation.

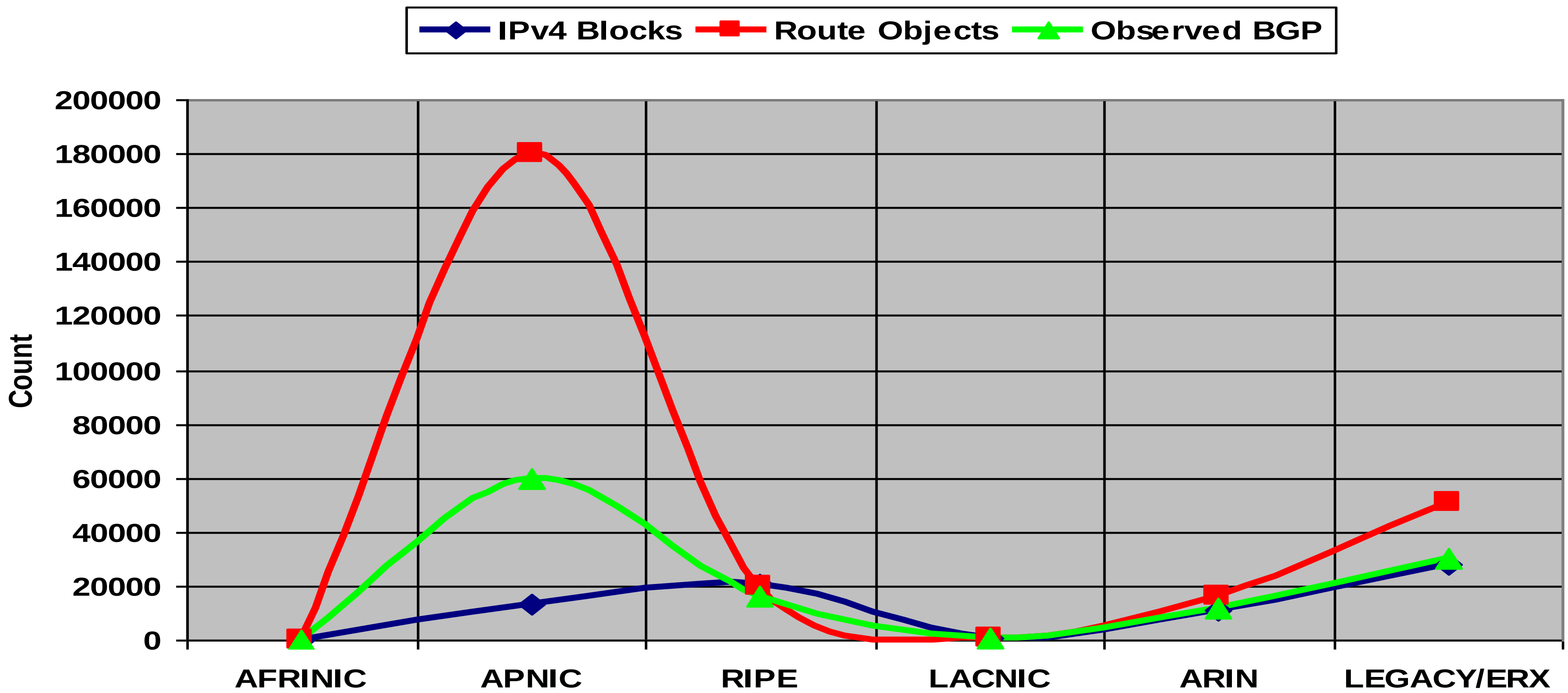
Analysis of PI Space in RPKI Issues

- Attributes "status:" in inetnum and "NetType:" in NetHandle:
  - Specify the type of address range represented by the address allocation object.
- No globally defined values of these attributes across RIRs. The defined values for PI blocks are as follows:
  - RIPE / AFRINIC:
    - *ALLOCATED PI / ASSIGNED PI / LIR PARTITIONED PI.*
  - APNIC:
    - *ALLOCATED PORTABLE / ASSIGNED PORTABLE.*
    - All /8 blocks are defined as ALLOCATED PORTABLE.
  - Some LEGACY blocks are contained in both RPSL and SWIP.
- The LEGACY/ERX blocks are generally assumed to be PI. However, some LEGACY/ERX blocks are specifically defined as PA. These specifically defined PA blocks are excluded for PI analysis.
- Some inetnum objects (in RPSL) do not contain "status:" attribute at all:
  - *# of inetnums with no "status:": 490,661.*
  - *Almost all of these came from JPNIC (one of NIRs under APNIC): 490,559*

Analysis of PI Space in RPKI Methodology

- Select IP resource allocation objects with PI specification.
- Adapt a different approach to each RIR:
  - RIPE / AFRINIC:
    - *All inetnum objects with the locally defined values for PI (ALLOCATED PI, ASSIGNED PI, LIR PARTITIONED PI).*
    - */8 blocks are defined as ALLOCATED UNSPECIFIED.*
  - APNIC:
    - *All inetnum objects with the locally defined values for PI (ALLOCATED PORTABLE, ASSIGNED PORTABLE).*
    - */8 blocks are defined as ALLOCATED PORTABLE, which are excluded.*
  - ARIN / LACNIC:
    - All objects that are directly "ASSIGNED" to an organization by the RIR.
  - LEGACY/ERX:
    - First, select all NetHandle objects with PI from SWIP, which belong to LEGACY/ERX.
    - Then, select all the LEGACY/ERX inetnum objects with PI from RPSL, which are not included in SWIP.
- Classify these PI blocks based on IANA allocation registry.

Distribution of PI Address Blocks based on Allocation Source



The graph depicts that APNIC-allocated PI address blocks are heavily sub-allocated to both route objects and advertised BGP updates.

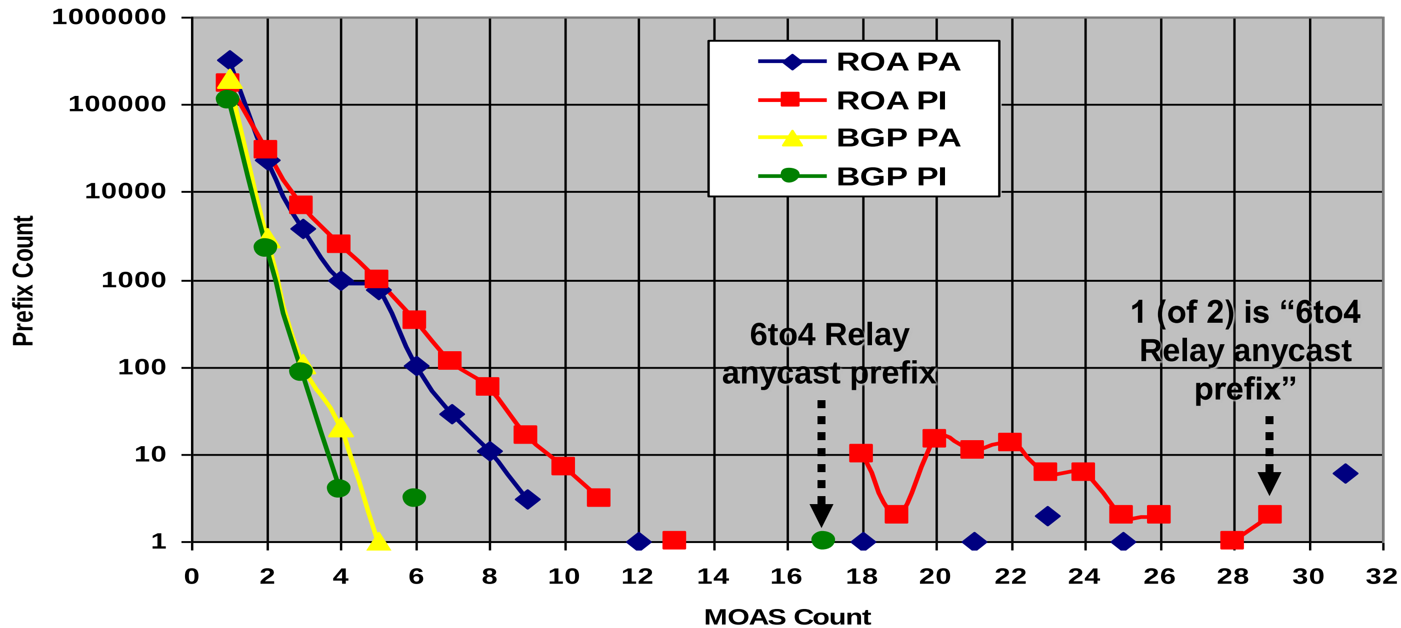
- **# IPv4 blocks with the valid "status": 5,281K**
- **# IPv4 blocks with NULL "status": 491K**
- **# IPv4 blocks with PI: 74K (~1.4%)**

- **# route objects (RPSL + SWIP): 654K**
- **# route objects with PI: 268K (~41%)**
- **# objects in both RPSL and SWIP: 4K**
- **There may be many proxy-registered route objects.**

- **# observed (P, O) pairs: 322K**
- **# observed (P, O) pairs with PI: 118K (~37%)**

Distribution of MOASes of Route Objects (ROA) and Observed BGP Updates (BGP) with PI vs. PA

Registry data date: 2009-2-18



- Here PA means the rest of address blocks other than PI space in the registry.
- PI address blocks tend to have more MOASes, especially in route objects. Does this indicate that many of them could be proxy-registered route objects or stale objects?

- # globally unique route objects (RPSL + SWIP): 654K
- # globally unique route objects with PI: 268K (~41%)
- # multi registrations between RPSL and SWIP: 4K
- There may be many proxy-registered route objects.

- # of observed unique (P, O) pairs: 322K
- # of observed unique (P, O) pairs with PI: 118K (~37%)

Conclusions

- We performed quantitative analysis of potential deployed RPKI and compared two possible deployment scenarios: full vs. optimized deployments
  - The total number of IPv4 certificates can be significantly reduced with prefix aggregation.
  - The global reduction rate of the total number of IPv4 certificates is ~38%, and ~26% on the certificates with prefix length <= /24.
- ROA validation in RPKI may not be a big performance issue:
  - About 89% of the total number of IPv4 address certificates (as of 2/18/2010) are address blocks with prefix length >= /25, which may not call for ROA creation.
  - The performance of ROA verifications can be significantly improved by the use of the cached "validation state" of certificates being verified.
    - * About 933K IPv4 certificates among total of more than 5.8M need to be verified for ROA verification when used with existing route objects.
- Handling of partial allocations across multiple RIRs?
  - Who would be responsible for creating resCerts for LEGACY/ERX address blocks?
- **Future tasks:**
  - Analysis of RPKI growth over time
  - Potential impact of RPKI on global BGP dynamics:
    - * The effect of creation, expiration or revocation of resource certificates and ROAs
  - The models can help generate synthetic RPKI workload models for routers for origin / path validation

Trustworthy Networking Program