1 Introduction

As the Resource Public Key Infrastructure (RPKI) [12] has begun its early deployment trials, estimating how large a global RPKI deployment must be to certify today's Internet has become increasingly relevant. In this document, we present a "back-of-the-envelope" style evaluation of the number of objects required to fully deploy a global RPKI in today's Internet, the performance characteristics of the RPKI under potential usage models (i.e., the degree of "hosting" employed, cache replication strategies, etc.), and how deployments of the RPKI under these different deployment models might scale. This work is intended to be a first stab at addressing these questions, and the authors welcome feedback, corrections, comments, etc. Note this document does not try to estimate the increasing size of the RPKI system due to massive new connectivity expected as IPv6 grows. With these caveats in mind, we begin from the perspective of a thought exercise, and pose the following questions:

- What would be the size of a fully deployed RPKI in today's Internet?
- What is the complexity of finding/learning/gathering a fully deployed RPKI?
- How long might such a gathering process actually take?

To answer these questions, we begin by estimating the size of the sum of the RPKI's various objects contained across the global set of repositories. These object types (described in [12]) are: CA certificates, Route Origin Authorizations (ROAs), Manifests, Certificate Revocation Lists (CRLs), and ghostbuster records. Objects like the ROA and ghostbuster record contain an End Entity (EE) certificate in their ASN.1 structure. Since each certificate is maintained in a repository, the caching system must download all objects from all repositories before validation computation can begin. While there are currently 5 RIRs running repositories, it is not clear if that will always be the case (the so-called "hosted" approach), if each Autonomous System (AS) ultimately ends up running its own repository, or some approach in between these two extremes eventually becomes the normal deployment model. To account for each of these cases, our analysis considers a range of cases where there are: 5 (the current set of RIRs), 10, 100, 1000, 10000, 42000 (the number of Autonomous System Numbers, ASNs, routed today) repositories.

What we find below is that with a lower-bound estimate of about 601,337 objects (assuming no BGPsec deployment) and 2,601,377 objects when RPKI is used to support a global BGPsec deployment. From these numbers, a globally deployed RPKI (in the worst cases) would likely take between 15 days to over 30 days to synchronize the local cache at every Relying Party (RP) in the world. While the methodology and measurements in this document are meant to be clear and concise, they are simply submitted as candidates of what a globally deployed RPKI might look like. Corrections are welcome, and alternate methodologies and measurements are very much encouraged. This work is intended to be a living document, which should evolve as operational experience is gained with the RPKI.
2 Methodology

The methodology used here is to examine the types of objects that must be generated to create a complete RPKI, then to estimate the size or number of each of these objects, and finally to combine these two estimates to provide a projected RPKI size in terms of individual replicated objects. Each section below represents an individual part of this process.

2.1 Object Definitions

In our candidate methodology, we enumerate the different types of objects according to the naming scheme in Table 1. One important clarification is that the RPKI architecture requires all objects to have verifying End Entity (EE) Certificates associated with them. However, as an optimization, objects are always transferred with their EE certs, as if they were one object. For example, one would not retrieve a crl object alone. Rather, one would transfer an o\textsubscript{CRL} tuple. While this reduces transfer costs, the cryptographic material in the c\textsubscript{EE} certificate in o\textsubscript{CRL} has potentially very different operational complexities associated with it. Certificates’ private portions may need to be managed in HSMs, the creation of new ROAs may (or may not) require new cryptographic certificate generation, etc. Thus, it is critical to account for their separate natures in different operational contexts.

According to the architecture of the RPKI, each ca must contain certain constituent objects: an organization-level CA cert, a CRL tuple, and a ghost busters tuple:

\[ ca_i = \{ c_{CA}^i, o_{CRL}^i, o_{GB}^i \} \]

This set may include a manifest object o\textsuperscript{M}, if it is hosted in a hierarchical repository (i.e. not a flat repo), or in its own separate repository.

An organization must manage its own ca and the allocation resource objects and its routing objects. That is:

\[ org_i = \{ ca_i, \{ o_{ROA}^{ROA_1}, o_{ROA}^{ROA_2}, \ldots o_{ROA}^{ROA_n} \}, \{ c^{EE}_{rtr_1}, c^{EE}_{rtr_2}, \ldots, c^{EE}_{rtr_m} \} \} \]

We generalize ca\textsubscript{i} set to be called InfObjs, and the set of allocation objects (ROA tuples) to be called AllocObjs and the set of routing keys (EE certs for routers) to be called RoutingObjs, and (finally) the repository provisioning/structure dictates if an organization needs its own o\textsuperscript{M} object (standalone or hierarchical repositories need this, flat structures do not) and we call this set (that might be empty) StructObjs. That makes org\textsubscript{i} = \{InfObjs, AllocObjs\textsubscript{i}, RoutingObjs\textsubscript{i}, StructObjs\textsubscript{i}\}. At this point we note the cardinalities of these sets. Each organization will normally have a 1:1 cardinality with its InfObjs, 1:n cardinality with its AllocObjs\textsubscript{i}, and 1:m cardinality with its RoutingObjs\textsubscript{i}.

<table>
<thead>
<tr>
<th>o\textsubscript{CA}</th>
<th>A CA certificate</th>
</tr>
</thead>
<tbody>
<tr>
<td>o\textsubscript{EE}</td>
<td>An End Entity certificate</td>
</tr>
<tr>
<td>m</td>
<td>A manifest object</td>
</tr>
<tr>
<td>o\textsuperscript{M} = (M, C\textsubscript{EE})</td>
<td>A manifest object with its EE cert</td>
</tr>
<tr>
<td>crl</td>
<td>A Certificate Revocation List (CRL)</td>
</tr>
<tr>
<td>o\textsuperscript{CRL} = (CRL, C\textsubscript{EE})</td>
<td>A CRL object with its EE cert</td>
</tr>
<tr>
<td>gb</td>
<td>A Ghost Busters record</td>
</tr>
<tr>
<td>o\textsuperscript{GB} = (GB, C\textsubscript{EE})</td>
<td>A Ghost Busters record and its EE cert</td>
</tr>
<tr>
<td>roa</td>
<td>A Route Origin Authorization (ROA)</td>
</tr>
<tr>
<td>o\textsuperscript{ROA} = (ROA, C\textsubscript{EE})</td>
<td>A ROA and its EE cert</td>
</tr>
<tr>
<td>ca</td>
<td>A Certificate Authority (not to be confused with a CA certificate)</td>
</tr>
<tr>
<td>rir</td>
<td>A Regional Internet Registry CA</td>
</tr>
<tr>
<td>org</td>
<td>An organization in the routing system (an ISP, an enterprise, a service provider, etc.)</td>
</tr>
<tr>
<td>RIR</td>
<td>The set of all RIRs</td>
</tr>
<tr>
<td>ORG</td>
<td>The set of all organizations responsible to routing in the Internet (ISPs, enterprises, service providers, etc.)</td>
</tr>
</tbody>
</table>

Table 1: Object definitions
In addition, each rir is currently planning to run 5 cas internally (one for each of the other RIRs, and one for ICANN):

\[ \text{rir}_a = \{ \text{ca}_RIR, \text{ca}_RIR, \text{ca}_RIR, \text{ca}_RIR, \text{ca}_RIR \} \]

The rationale here is to certify transfer blocks between RIRs, and to account for ERX/transfers (legacy numbers). A detailed explanation is beyond the scope of this sizing document.

With a few simple assumptions, we can model the expected size of a global deployment of RPKI from the above definitions. If we consider the total downloadable object count to be the set of all RIRs and all the objects needed to attest to all allocations and routers in the Internet today, we arrive at the following expression.

\[ O^{\text{Total}} = \{ \text{RIR, ORG} \} \]

With the above expansion of \( \text{rir} \), we can expand \( O^{\text{Total}} \) as follows:

\[ O^{\text{Total}} = \bigcup_{i=1}^{5} \{ (\text{ca}_RIR, \text{ca}_{RIR +1}, \text{ca}_{RIR +2}, \text{ca}_{RIR +3}, \text{ca}_{RIR +4}), \text{ORG} \} \]

Next, we make the observation that since each object in the RPKI is intended to come from a single root (i.e. no object can have multiple roots in its cryptographic delegation chain), then the intersection between any two object sets (infrastructure, allocation, or routing objects) will necessarily be \( \emptyset \). For example, consider two organizations \( A \) and \( B \). When comparing the infrastructure object sets for these two organizations, we will always be able to say \( \text{InfraObjs}_A \cap \text{InfraObjs}_B = \emptyset \). This observation has the powerful side effect that it allows us to describe the aggregate sets that amalgamate all sets of each type of object into corresponding supersets:

\[
\begin{align*}
\text{InfraObjs} & = \bigcup_{i=1}^{\text{|ORG|}} \text{InfraObjs}_i \\
\text{AllocObjs} & = \bigcup_{i=1}^{\text{|ORG|}} \text{AllocObjs}_i \\
\text{RoutingObjs} & = \bigcup_{i=1}^{\text{|ORG|}} \text{RoutingObjs}_i \\
\text{StructObjs} & = \bigcup_{i=1}^{\text{|ORG|}} \text{StructObjs}_i
\end{align*}
\]

These amalgamations allow us to further expand \( O^{\text{Total}} \) as follows:

\[ O^{\text{Total}} = \bigcup_{i=1}^{5} \{ (\text{ca}_RIR, \text{ca}_{RIR +1}, \text{ca}_{RIR +2}, \text{ca}_{RIR +3}, \text{ca}_{RIR +4}), \text{InfraObjs}, \text{AllocObjs}, \text{RoutingObjs}, \text{StructObjs} \} \]

Next, consider the relative sizes of these sets as they would be in a fully deployed RPKI system, today. To do this, we consider the size of what’s deployed in today’s Internet, today: 5 RIRs, 42,000 ASNs, and 431,000 Prefixes. A reported estimate of the number of eBGP speaking routers is 1 million \[6\] (this makes roughly \( \frac{42,000}{1,000,000} = 23.81 \) eBGP routers per AS).\(^1\)

We note that the number of prefixes here does not include dark prefixes, or rather publicly usable address blocks allocated for private use. Whether or not these prefixes are ever advertised on the public Internet, they are still often used in private peering arrangements, and the organizations that own them would certainly still want certificates generated and distributed for these blocks. Currently, neither the literature nor conventional operational wisdom describe ways to measure or estimate the number of dark prefixes that currently exist but which are not advertised. However, we could take \( 2^{416} \) as the probable number of address allocations that will ultimately have certificates distributed through any given origin authentication system.

When a new organization receives an allocation from an RIR, this will result in three necessary downloadable objects: \{\text{CA}, \text{GB}, \text{CRL}\}, and depending on the repository provisioning/structure it may have \( \sigma^M \). In addition,

\(^1\)While the authors appreciate the estimation, we are still investigating a more systematic method of estimating this number.
Table 2: Approximation of object set sizes

<table>
<thead>
<tr>
<th>Set</th>
<th>Derivation</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>RIR</td>
<td>\rangle$</td>
</tr>
<tr>
<td>$</td>
<td>InfraObjs</td>
<td>\rangle$</td>
</tr>
<tr>
<td>$</td>
<td>AllocObjs</td>
<td>\rangle$</td>
</tr>
<tr>
<td>$</td>
<td>RoutingObjs</td>
<td>\rangle$</td>
</tr>
<tr>
<td>$</td>
<td>StructObjs</td>
<td>\rangle$</td>
</tr>
</tbody>
</table>

at least one additional $o^{ROA}$ will be needed for each RIR that an organization gets an allocation from (IP or ASN). That means each $InfraObjs$ set will account for at least three objects. We, therefore, make the general estimate that the number of organizations today is the sum of RIR members plus the number of ERX legacy blocks (as those must be represented separately, even for existing organizations) plus the number of suballocations from RIR members times three. More formally:

$$|InfraObjs| = (|RIR|^{mem} + |ERX| + |SubAlloc|) \times 3$$

Where $RIR^{mem}$ is the set of all members of all RIRs, ERX is the set of all legacy ERX transfers, and $SubAlloc$ is the set of all suballocations. This addition clearly does not include any sub-suballocations. Indeed, in [10], on slides 12 and 13, the authors show measurements that (in particular) LACNIC has over an order of magnitude more sub-suballocations than suballocations. We, therefore, treat this measure as a pronounced lower bound.

Today there are $3,370 + 4,277 + 2,705 + 1,276 + 8,618 = 20,246$ members of RIRs [4, 5, 3, 11, 14], since the number of ERX allocations is hard to know we estimate it at 10\% of the RIR members 2,025, and with 61,505 suballocations [9], we get a total count of $|InfraObjs| = 20,246 + 2,025 + 61,505 = 83,776$.

2.2 Estimating the Number of Objects

Next we consider $AllocObjs$. There are many cases in which an organization will need multiple $o^{ROA}$ objects. For example, each routed allocation needs at least one ROA, and also if an organization wants to announce a prefix from multiple origins (MOAS), etc. It is for this reason that $AllocObjs$ will be loosely related to the way in which allocations are routed. We use this intuition to estimate that the number of allocation objects could roughly match an optimized view of the routing system today (i.e. a view in which announcements were maximally aggregated). For this, we use the origin report [9]. This site can be used to collapse the routing table entries down to the smallest set possible, and results in 273,592, or $|AllocObjs| = 273,592$. We note that in [10], on slide 9, the authors calculate that RIPE alone would result in 100,000 ROA objects. Rather than use that inflated estimate and invent an arbitrary multiplier (and to avoid any potential systemic bias), we choose the smaller and more systematic measure that is based on the routing table.

As mentioned earlier, the total number of eBGP speaking routers throughout the Internet was estimated as roughly 23.8 per ASN. We conservatively double this (to represent standby keys) as advised by a lead on the BGPsec design team [6]. This leads to $RoutingObjs = 2,000,000$.

Finally, the set of $StructObjs$ depends entirely on the RPKI deployment. If there are just a few flat repositories, then $StructObjs$ would be very small. If, on the other hand, every organization were to run their own repository, this number would grow to $|StructObjs| = |ORG|$. As a result, we use this as an independent variable in Section 3 when calculating object counts.

With the above, Table 2 maps our object types to size estimates based on the size of the Internet today. Therefore, we estimate that

$$O^{Total} = 100 + 251,328 + 273,592 + 2,000,040 + 83,776 + |StructObjs| = 2,608,836 + |StructObjs|$$

\(^2\)depends on how many of an RIR’s members have historical resources or resources transferred to them from another RIR.
### 3 Evaluation

While 2.6 million objects could be considered to be a large corpus (especially as a lower bound estimate), it will certainly grow whenever any keys are being rolled over and changed, it will necessarily almost double during cryptographic algorithm rollovers, and it doesn’t include those objects needed for the actual allocation hierarchy. So, again, we submit that this is a lower bound.

Next, to estimate the amount of time it takes for caches to actually gather a fully deployed RPKI (at today’s Internet’s scale), let’s examine the algorithm. We consider the set of interconnected repositories to be a directed graph, in which the set of vertices \( V \) are discovered by gathering (i.e., an repository contains certificates who have SIA and AIA URI pointers to potentially external repositories), and to traverse all of it, we use a depth-first search approach. The complexity of this approach is well known to be: \( O(|V| + |E|) \), where \( V \) is the set of vertices, and \( E \) is the set of edges.

In the RPKI, \( V \) is the set of repositories, and \( E \) is the set of unique pairs of vertices in \( V \) that come from certificates in each repository and which point to external SIA or AIA URIs (for example, certificates chaining to their CAs). From this understanding, we need to estimate how long it would actually take to follow this algorithm (with the above set of objects) in order for a cache to discover all of the repositories in the Internet (on each running). To map this complexity to actual deployment numbers, we consider the presentation [1]. In this presentation, we see 10 different performance graphs for RPKI repositories. These graphs each explicitly show the average number of objects and the average sync time. Table 3 lists these values and the computed aggregate statistics across all of them.

With these numbers, Table 4 lists estimates for how long it would take to download from a varying number of repositories containing objects for the Internet, if RPKI were fully deployed, today. We note that in our evaluation, we should have modeled a latency factor for traversing edges in \( E \) (for example, DNS takes time, timeouts, etc.). However, to continue to focus on a lower bound, we ignored the \( E \) cost in our calculations.

Figure 1 illustrates that as the number of repositories grows, so too does the sync time for the system. In addition to these numbers, [1] appears to analyze the caching behavior of a large ISP with an internal hierarchical caching system. This would mean that after a master cache (for a single ISP) has gathered all of the objects...
3.1 Deadline Crawling

A repository system that relies on RPs having a consistent/fresh view of all objects in the entire global repository before validation can be assured begins with an architectural handicap. Recall that in the RPKI, caches must fully enumerate, and known, and support $O^{\text{Total}}$ objects, spread out across $|\text{StructObjs}|$ repositories. If there

| Number of Repositories (|StructObjs|) | Number of Objects (with router $c^{EE}$) | Time to Gather | Days to Gather |
|-------------------------------|----------------------------------------|----------------|----------------|
| 5                             | 2,608,841                              | 1,638,355      | 18.96245       |
| 10                            | 2,608,846                              | 1,638,362      | 18.96252       |
| 100                           | 2,608,936                              | 1,638,475      | 18.96383       |
| 1,000                         | 2,609,836                              | 1,650,909      | 19.10774       |
| 10,000                        | 2,618,836                              | 1,691,101      | 19.57293       |
| 42,000                        | 2,650,836                              |                |                |

Table 5: Gathering times as the number of repositories grows (with router certs).

Figure 1: Between 4.3 and 5 days to sync (19 to 20 with router certs). That means that operators have to count on 9 to 10 days (or 38 to 40 with router certs) after making a change that they want RPs to pick up (twice the polling period).

from across the global RPKI, it must begin replicating this to other internal “tiers” of caches (inside the ISP). Therefore, for such ISPs, we must also add the numbers from [1] to this (as internal caches must replicate). In [1], the replication time was between 70 and 80 minutes. However, while the authors describe their internal caching testbed as being, “fairly large scale,” their simulated RPKI deployment only had 14,000 objects (vs. our estimated 2.6 million, in Table 4). As a result, one might worry that with strictly linear scaling, this replication time for internal caches could become $70 \times \frac{2,650,836}{14,000} = 13,254.18$ more minutes (or $9.204292$ more days). This could make estimates inflate to $19.51868 + 9.178375 = 28.697055$ days. Admittedly, each ISP may choose different caching strategies, but this hierarchical model provides a lot of operational robustness, though it complicates convergence time significantly. One final note, the above analysis does not add in time for potential system failures, latency as load increases in repository systems, etc.

3.1 Deadline Crawling
3.1 Deadline Crawling

Figure 2: This Figure illustrates the coefficient needed to accelerate the downloading of RPKI objects as the repository grows.

Figure 3: This Figure illustrates the the required average object/second rate required to meet a global deadlines.

becomes a requirement to provide a certain degree of freshness (i.e. caches must be able to learn new objects in less than \( t \) seconds), then we must use the above formulation to drive requirements on the synchronization time.

That is, if there are \( |O_{Total}| \) objects in the RPKI, then if our deadline is \( t \), we can determine how long synchronizing them must take. The above calculations assume a linear fetch time as object counts grow. On the other hand, if we consider a deadline of \( t \), and time coefficient of \( z(t, n) \) (where \( n \) is an object count, and \( t \) is the deadline) for the objects/second synchronization rate \( r \), then we can cast the overall synchronization time as:

\[
t = (|O_{Total}| \times r) \times z(t, |O_{Total}|)
\]

If we fix our synchronization time (deadline), we can solve for \( z(t, n) \) as:

\[
z(t, |O_{Total}|) = \frac{t}{|O_{Total}| \times r}
\]
Thus, if we (for example) want to set a deadline of one day (86,400 seconds), we would need a coefficient of:

$$\frac{86,400}{0.0519} = 0.0519.$$  Using the measurements from Tables 4 and 5, Figure 2 shows the value of the coefficient needed to meet the deadline of one day. That is, what the sublinear scaling would need to look like in order to have a global system pick up changes in one day. Based on caching, this could actually be double, but for simplicity, we leave the measurements as presented.

Switching this around, if the goal is to have a full synchronization of the global RPKI in one hour (such as might be needed by DDoS service providers [13]), then Figure 3(a) shows the global average transfer rate required for a deadline of 1 hour. Figure 3(b) shows the global average transfer rate required for a deadline of 1 day.

4 Discussion

So far, we have dealt with the size of the database of objects required to build a complete RPKI, if the system is in a stable state. Many factors will require regular updates to the objects discussed so far, including key rollover and duplication of objects for systemic resiliency. The rate at which key rollovers occur within the system is undetermined at this time, so it is difficult to know precisely how key rollovers will impact system scaling. Duplication of information [8] for system resiliency, however, is likely to be fairly common, and therefore worth investigating.

5 Conclusion

The calculations in this document, while just candidate representations of RPKI, project a very long lower-bound on the potential time needed for RP caches to gather all objects in the RPKI. Furthermore, when repositories in the RPKI undergo key rollovers, repositories and caches must store old and new objects for about a month (the polling period + potential cache replication times) before they can safely expunge them. This, in turn, will directly inflate the above numbers, and cause longer delays.

For example, if an algorithm is rolled over [7], the system will have roughly twice as many objects as it normally does, because every signed object in the entire global RPKI will necessarily have to be signed with an outdated algorithm, and then again with a new one. Moreover, this state is envisioned to persist for on the order of “months or years” at a time. This may almost double the sync time, and cause the RPKI to be even less responsiveness to key rollovers (possibly almost twice the 30 day estimate).

Also, if rsync scales sublinearly, then it is likely that the actual time needed to sync will grow beyond these conservative estimates. Again, even if rsync periodically has to restart because objects being transferred have changed [2], this underestimate will get worse. In addition to this, active CRLs will not (by design) include expired data. Therefore, as soon as operators (following advice in existing drafts) begin to use expired data, Relying Parties (RPs) will necessarily be forced to keep track of old data everywhere (thus increasing the object count, again).

However, as we saw in Section 3, if we consider deadlines for global synchronization, we find that object synchronization rates must be superlinear (not sublinear). That is, the system must get faster as it gets more loaded. What is worrisome about these findings is that even the modest goal of 24-hour synchronization times would require a sublinear scaling coefficient (and thereby superlinear performance), which would need to scale roughly two orders of magnitude to meet today’s scale. One reason to begin considering deadline scheduling was illustrated during mail list discussions [13].

The participants of this email thread outlined the need to have tangible deadlines for the routing system to be able to reflect changes in routed origins. The explanation focused on the service provider DDoS market, in which new customers that are being impacted by DDoS attacks may ask service providers to protect them through new advertisements. Today, this can be done in minutes (and normal routing can be returned to customers in minutes too), and that is a business-case deadline.

Prior to this document, only publications such as [1] reported scaling measurement. While this presentation (and its relative incarnations) purport to be “fairly large scale” they appear to have estimated object counts roughly two orders of magnitude too small. We mention this simply to motivate the importance of having a living document such as this one. Finally, this analysis does not even take into account churn or systemic dependencies. Such issues should be discussed in similar (but separate) technical notes.

©2012 Verisign, Inc. All rights reserved.
References


