INTERNET-DRAFT

Danny McPherson
Ryan Donnelly
Frank Scalzo
Verisign, Inc.

Expires: January 2012

July 2, 2011

Intended Status: Best Current Practice

Unique Per-Node Origin ASNs for Globally Anycasted Services

<draft-ietf-grow-unique-origin-as-01.txt>

Status of this Memo

Abstract

This document makes recommendations regarding the use of unique origin autonomous system numbers per node for globally anycasted critical infrastructure services in order to provide routing system discriminators for a given anycasted prefix. Network management and monitoring techniques, or other operational mechanisms may employ this new discriminator in whatever manner best accommodates their operating environment.

Table of Contents

1. Terminology

This document employs much of the following terminology, which was taken in full from Section 2 of [RFC 4786].

Anycast: the practice of making a particular Service Address available in multiple, discrete, autonomous locations, such that datagrams sent are routed to one of several available locations.

Anycast Node: an internally-connected collection of hosts and routers that together provide service for an anycast Service Address. An Anycast Node might be as simple as a single host participating in a routing system with adjacent routers, or it might include a number of hosts connected in some more elaborate fashion; in either case, to the routing system across which the service is being anycast, each Anycast Node presents a unique path to the Service Address. The entire anycast system for the service consists of two or more separate Anycast Nodes.

Catchment: in physical geography, an area drained by a river, also known as a drainage basin. By analogy, as used in this document, the topological region of a network within which packets directed at an Anycast Address are routed to one particular node.

Local-Scope Anycast: reachability information for the anycast Service Address is propagated through a routing system in such a way that a particular anycast node is only visible to a subset of the whole routing system.

Local Node: an Anycast Node providing service using a Local-Scope Anycast Address.

Global Node: an Anycast Node providing service using a Global-Scope Anycast Address.

Global-Scope Anycast: reachability information for the anycast Service Address is propagated through a routing system in such a way that a particular anycast node is potentially visible to the whole routing system.

Service Address: an IP address associated with a particular service (e.g., the destination address used by DNS resolvers to reach a particular authority server).

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119].

2. Introduction

IP anycasting [RFC 4786] has been deployed for an array of network services since the early 1990s. It provides a mechanism for a given network resource to be available in a more distributed manner, locally and/or globally, with a more robust and resilient footprint, commonly yielding better localization and absorption of systemic query loads, as well as better protections in the face of DDoS attacks, network partitions, and other similar incidents. A large part of the Internet root DNS infrastructure, as well as many other resources, has been anycasted for nearly a decade.

While the benefits realized by anycasting network services is proven, some issues do emerge with asserting routing system reachability for a common network identifier from multiple locations. Specifically, anycasting in BGP requires injection of reachability information in the routing system for a common IP address prefix from multiple locations. These anycasted prefixes and network services have traditionally employed a common origin autonomous system number (ASN) in order to preserve historically scarce 16-bit AS number space utilized by BGP for routing domain identifiers in the global routing system. Additionally, a common origin AS number was used in order to ease management overhead of resource operations associated with acquiring and maintaining multiple discrete AS numbers, as well as to avoid triggering various operations- oriented reporting functions

aimed at identifying "inconsistent origin AS announcements" observed in the routing system. As a result, the representation of routing system path attributes associated with those service instances, and that anycasted prefix itself, typically bear no per-instance discriminators in the routing system (i.e., within the network control plane itself).

Service level query capabilities may or may not provide a mechanism to identify which anycast node responded to a particular query, although this is likely both service (e.g., DNS or NTP) and implementation dependent. For example, NSD, Unbound, and BIND all provide 'hostname.bind or hostname.id' [HNAME] query support that enables service-level identification of a given server. Tools such as traceroute are also used to determine which location a given query is being routed to, although it may not reveal local-scope anycast instances, or if there are multiple servers within a given anycast node, which of the servers responded to a given query, in particular when multiple servers within an anycast node are connected to a single IP router. When utilizing these service level capabilities, query responses are typically both deterministic and inherently topology-dependent, however, these service level identifiers at the data plane provide no control plane (routing system) uniqueness.

As more services are globally anycasted, and existing anycasted services realize wider deployment of anycast nodes for a given service address in order to accommodate growing system loads, the difficulty of providing safeguards and controls to better protect those resources expands. Intuitively, the more widely distributed a given anycasted service address is, the more difficult it becomes for network operators to detect operational and security issues that affect that service. Some examples of such security and operational issues include BGP route leaks affecting the anycasted service, rogue anycast nodes appearing for the service, or the emergence of other aberrant behavior in either the routing system, the forward query datapath, or query response datapath. Diagnosis of the routing system issues is complicated by the fact that no unique discriminators exist in the routing system to identify a given local or global anycast node. Furthermore, both datapath and routing system problem identification is compounded by the fact that these incident types can be topologically-dependent, and only observable between a given client-server set.

Additionally, while it goes without saying that many anycasted services strive for exact synchronization across all instances of an anycasted service address, if local policies or data plane response manipulation techniques were to "influence" responses within a given region in such a way that those responses are no longer authentic or that they diverge from what other nodes within an anycasted service were providing, then it should be an absolute necessity that those modified resources only be utilized by service consumers within that region or influencer's jurisdiction.

Mechanisms should exist at both the network and service layer to make it abundantly apparent to operators and users alike whether any of the query responses are not authentic. For DNS, DNSSEC [RFC 4033] provides this capability at the service layer with object level integrity, assuming validation is being performed by recursive name servers, and DNSSEC deployment at the root and top level domain (TLD) levels is well underway [DNSSEC-DEPLOY]. Furthermore, control plane discriminators should exist to enable operators to know toward which of a given set of instances a query is being directed, and to enable detection and alerting capabilities when this changes. Such discriminators may also be employed to enable anycast node preference or filtering keys, should local operational policy require it.

3. Recommendation for Unique Origin ASNs

In order to be able to better detect changes to routing information associated with critical anycasted resources, globally anycasted services with partitioned origin ASNs SHOULD utilize a unique origin ASN per node where possible, if appropriate in their operating environment and service model.

Discrete origin ASNs per node provide a discriminator in the routing system that would enable detection of leaked or hijacked instances more quickly, and would also enable operators that so choose to

proactively develop routing policies that express preferences or avoidance for a given node or set of nodes associated with an anycasted service. This is particularly useful when it is observed that local policy or known issues exist with the performance or authenticity of responses returned from a specific anycast node, or that enacted policies meant to affect service within a particular region are affecting users outside of that region as a result of a given anycast catchment expanding beyond its intended scope.

Furthermore, inconsistent origin AS announcements associated with anycasted services for critical infrastructure SHOULD NOT be deemed undesirable by routing system reporting functions, but should instead be embraced in order to better identify the connectedness and footprint of a given anycasted service.

While namespace conservation and reasonable use of AS number resources should always be a goal, the introduction of 32-bit ASNs significantly lessens concerns in this space. Globally anycasted resources, in particular those associated with critical infrastructure-enabling services such as root and TLD name servers, SHOULD warrant special consideration with regard to AS number allocation practices during policy development by the constituents of those responsible organizations (e.g., the Regional Internet Registries). Additionally, defining precisely what constitutes "critical infrastructure services" or "special consideration" (e.g., some small range of 32-bit AS numbers might be provided) is left to the constituents of those organizations. Additionally, critical infrastructure employment of 32-bit ASNs for new nodes might well help to foster more rapid adoption of native 32-bit ASN support by network operators.

One additional benefit of unique origin AS numbers per anycast node is that Resource PKI (RPKI) Secure Inter-domain Routing [SIDR] machinery, and in particular, that of Route Origin Authorizations (ROAs), and routing policies that may be derived based on those ROAs, can be employed with per anycast node resolution, rather than relying

on a single ROA and common origin AS to cover all instantiations of an anycasted prefix (possibly hundreds) within the global routing system. For example, deployments that incorporate partitioned ASN anycast models that have a single ASN bound to all nodes but cross organizational or political boundaries, a situation may arise where

nobody would be deemed appropriate to hold the key for the ROA. Additionally, a globally anycasted service within a given IP prefix that shares a common ASN might be taken totally offline because of the revocation of a ROA for that origin ASN. The RPKI model today already inherently accommodates issuance of multiple ROAs with unique origins for a given prefix.

4. Additional Recommendations for Globally Anycasted Services

Two additional recommendations for globally anycasted critical infrastructure services are related to publication of information associated with a given node's physical location, and which adjacent upstream ASNs an origin AS interconnects with. The former would allow operators to better define and optimize preferences associated with a given node to align with local policy and service optimizations. The latter would allow expression through policy such as Routing Policy Specification Language [RFC 4012] specified in Internet Routing Registries (IRRs) in a manner that illustrates a discrete set of upstream ASNs for each anycast node, rather than the current model where all upstream ASNs associated with a common origin AS may or may not be expressed. This information would provide an additional level of static routing policy or monitoring and detection models by network operators, and perhaps explicit network layer source address validation in the datapath.

5. Security Considerations

The recommendations made in this memo aim to provide more flexibility for network operators hoping to better monitor and prevent issues related to globally anycasted critical infrastructure resources. Anycast itself provides considerable benefit in the face of certain attacks, yet if a given instance of a service can appear at many points in the routing system and legitimate instances are difficult to distinguish from malicious ones, then anycast expands the service's attack surface rather than reducing it.

The recommendations made in this document are expressed to assist with visibility and policy specification capabilities in order to improve the availability of critical Internet resources. Use cases where the recommendations outlined in this memo may have helped to more easily detect or scope the impact of a particular incident are illustrated in [RENESYS-BLOG].

Furthermore, while application layer protection mechanisms such as DNSSEC provide object level integrity and authentication, they often do so at the cost of introducing more failure conditions. For example, if a recursive name server is performing DNSSEC validator functions and receives a bogus response to a given query as a result of a man-in-the-middle (MITM) or injected spoofed response packet such as a cache poisoning attempt, the possibility might exist that the response packet is processed by the server and results in some temporal or persistent DoS condition on the recursive name server and for its client set. The unique origin AS mechanism outlined in this document provides the capability for network operators to expressly avoid anycast node catchments known to regularly elicit bogus responses, while allowing the anycasted service address to remain available otherwise.

6. Deployment Considerations

Maintenance of unique ASNs for each node within an anycasted service may be challenging for some critical infrastructure service operators initially, but for globally anycasted resources there needs to be some type of per-node discriminator in the control plane to enable detection, remediation, and optimally, preventative controls for dealing with routing system anomalies that are intensified by the application of IP anycasting. Additionally, this technique sets the stage to employ RPKI-enabled machinery and more secure and explicit routing policies, which all network operators should be considering.

The granularity of data publication related to anycast node location should be left to the devises of each services operator, and the value of this mechanism in each operators unique environment, but some reasonable level of detail to enable operators and service consumers to make informed decisions that align with their security and operational objectives as outlined herein should be provided by each critical services operator.

Adjacent AS information for a given origin AS can be obtained through careful routing system analysis already when prefixes are advertised

via a given set of AS adjacencies, and therefore should present no new threat. However, network interconnection and peering policies may well present some challenges in this area. For example, if a technique such as unique origin AS per node is employed then a single organizaton may no longer have a single AS for interconnection at each location, and interconnection policies should expressly consider this. That said, interconnection with networks that provide critical infrastructure services should certainly be given due consideration as such by network operators when evaluating interconnection strategies.

Some root and TLD operators today identify erroneous anycast prefix announcements by detecting prefix announcements with an origin AS other than the common origin AS shared via all nodes. This detection model would need to be expanded to account for unique origin ASNs per node if a given service operators chooses to employ such a model, and given that AS paths are trivial to manipulate in the current system, the above technique would only assist in the event of unintentional configuration errors that reoriginate the route (e.g., it doesn't even detect leaks that preserve the initial path elements). In that case, work underway on routing security origin and path validation in the SIDR working group and beyond should be consulted.

While local policy based on any BGP attributes, to include AS path information, can influence policy within a local administrative domain and possibly downstream, there exists a possibly that upstream nodes continue to use a route deemed undesirable by the local admin once data packets reach that network. Network operators must understand the implications of this property in their operating environment, as it is inherent in all Interent routing.

Finally, anycast node presence at exchange points that employ route servers may make enumeration of adjacent ASNs for a given node challenging. While this is understood, service operators should make every effort to enumerate the set of adjacent ASNs associated with a given anycast node's origin AS. Without express understanding of legitimate AS interconnection and authorized origin AS information, more secure routing is difficult to achieve.

7. Acknowledgements

Thanks to David Conrad, Steve Kent, Mark Kosters, Andrei Robachevsky, Paul Vixie, Brad Verd, Andrew Herrmann, Gaurab Raj Upadhaya, Joe Abley, Benson Schliesser, Shane Amante, Hugo Salgado, and Randy Bush

for review and comments on this concept.

8. IANA Considerations

This document requires no direct IANA actions, although it does provide general guidance to number resource allocation and policy development organizations, and in particular Regional Internet Registries, regarding allocation of AS numbers for globally anycasted services.

9. References

9.1. Normative References

[RFC 2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC 4786] Abley, J., and Lindqvist, K., "Operation of Anycast Services", RFC 4786, BCP 126, December 2006.

9.2. Informative References

[RFC 4012] Blunk, et al., "Routing Policy Specification Language

next generation (RPSLng)", RFC 4012, March 2005.

[RFC 4033] Arends, et al., "DNS Security Introduction and Requirements", RFC 4033, March 2005.

[DNSSEC-DEPLOY] "Root DNSSEC", <http://www.root-dnssec.org/>

[HNAME] ISC, "Which F-root node am I using?" <http://www.isc.org/community/f-root/which_node>

[RENESYS-BLOG] Zmijewski, E., "Accidentally Importing Censorship", Renesys Blog, March 30, 2010. <http://www.renesys.com/blog/2010/03/fouling-the-global-nest.shtml>

[SIDR] Lepinski, M., Kent, S., "An Infrastructure to Support Secure Internet Routing", October 2009, Internet-Draft, "Work in Progress".

10. Authors' Addresses

Danny McPherson
Verisign, Inc.
21345 Ridgetop Circle
Dulles, VA USA 20166
Phone: +1 703.948.3200

Email: dmcpherson@verisign.com

Ryan Donnelly
Verisign, Inc.
21345 Ridgetop Circle
Dulles, VA USA 20166
Phone: +1 703.948.3200

Email: rdonnelly@verisign.com

Frank Scalzo
Verisign, Inc.
21345 Ridgetop Circle

Dulles, VA USA 20166
Phone: +1 703.948.3200

Email: fscalzo@verisign.com

Copyright Statement