Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 18, 2013

M. Richardson
SSW

November 14, 2012

Secret Gardens are Better than Walled Gardens
draft-richardson-homenet-secret-gardens-01

Abstract

This document explains a few use cases where operators would like to introduce so-called "walled gardens" into home-networks, including distribution of new DNS anchors. This document proposes an alternative solution involving DNS delegations to access controlled DNS servers. The results are much more scalable, and can be deployed today, using existing operating systems and existing DNS infrastructure.

Table of Contents

1. Introduction: A brief history of split-horizon DNS

DNS settings have been a regular concern since the very early 1990s, when the first firewalls began to partition the Internet. In the earliest works, Cheswick and Bellovin describe the use of split-DNS in an enterprise: all internal machines (including the inside interface of the firewall, and possibly machines on the DMZ/ Service-Network) would use an internal recursive DNS server for names, and if the name was external, the internal recursive DNS server would ask the world.

The above split-horizon configuration survives to today, but it has become very complicated. The first complication was remote access (VPN) to the enterprise. For the computer at home to be able access things, the internal DNS server had to be used. Often short internal names ("smtp", "web", "wiki","printer") would be used, depending upon the fact that all internal machines had the same search path. As the VPN could go on, and off, if it was off, and the end user entered the word "wiki" on the browser, instead of going to the internal resource (which in IPv4 space, had an RFC1918 address), it would either go to an external resource, or cause a search.

Worse, some computer systems originally needed to be rebooted in order to change their DNS settings, as this was really the only way to convince all application to flush name to IP address mappings that they had cached.

Particularly gruesome is the case of the contractor or consulting, who works at enterprise A, and then visits enterprise B. While on the network of Enterprise B (where they may be located for some months), in order to do simple things like reach the printer (using the name "printer"), they need to use the DNS settings for Enterprise B. But, in order to fetch their email, they must have the DNS settings (and VPN) for their home base, Enterprise A. There have been regular reports in the VPN/Remote Access community of situations where a worker needs to have VPNs up with two remote locations, while residing at a third.

The VPN situation is tragic for the technical user, for the non-technical user, it is impossible. For the technical user, typing in longer names, setting up multiple search paths, and running a local recursive name server are possible. For the less technical user, typing in IP addresses rather than names works as long as HTTP Virtual Hosting is not involved (for which the name is important). But the degree to which these things occur has been limited in part due to the inevitable conflict of RFC1918 addresses, which means, that, even if one uses IP addresses, the routing doesn't work.

DNSSEC has put a new twist into things: the best way to run DNSSEC is to have a secure recursive resolver. Now, this recursive resolver needs to be taught about split-horizon names, to talk to the internal name server when it is reachable (and probably to trust it), and to talk directly to the Internet when not reachable. A second problem is that the local recursive name server may not always get flushed. If a query for "smtp.example.com" should result in the internal mail relay when at the office, or connected via VPN, but the same name should resolve in the external address when not connected (possibly, the same machine with multiple interfaces, but not always), then the caching presents a problem. This problem is acute if one tries to run a caching name server on a mobile, multiple-interface devices, such as a stock smartphone that has 3G and wifi interfaces, which may operate concurrently.

Some organizations, where there are multiple data centres, and the network is distinctly non-convex, have solved the caching problem by dispensing with a true split-horizon DNS, and have simply put all external names under "example.com", with internal names under "internal.example.com". Whether or not externally made requests for foo.internal.example.com are resolved by externally facing authoritative name servers is now a security policy question, rather than a routing or caching concern.

There have been further abuses of split-horizon DNS. One of these is by hotels and other places with "captive portals". In order to authenticate the user using a web interface, they must make the portal visible to the user. This involves making the user "captive": no packets may leave the local network until the user has sufficiently authenticated (possibly, involving a financial transaction). Some captive portals have decided that they will intercept all DNS requests, and no matter what the user asks for, they will answer with their own IP. This can fail for three reasons: 1) the user does not use the DNS servers provided, instead uses their own (perhaps intending to be reached through a VPN) or or the root name servers directly (perhaps in order to do DNSSEC), 2) the names used in the web interfaces are unqualified ones, and the user has not accepted the search path from the DHCP offer (ref: some BCP on this), 3) even if all of this goes well, the user's browser has now cached an incorrect mapping for the desired web site.

1.1. Requirements Language

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 [RFC2119].

2. Use Cases

This section lists a few use cases which homenet believes are within scope. Only the business goals are listed of the use cases.

2.1. IPv6 TV

[SERVER-SELECTION] Section 3.1 "CPU Deployment Scenario" describes a similar situation to the IPv6 TV below.

A seperate internet connection is provided by the IP TV provider. This can be on a seperate physical wire, or a seperate logical wire. Historically, only the TV set itself is placed on this network, the home owner may be unaware that the wires are other than analog cable TV wires.

The TV set does DNS lookups for the content servers, and is expected to receive answers that are within the IP TV provider's network. The TV set can access the provider's servers as long as it originates connections from the IPv6 prefix that the IP TV provider provided.

The provider does not publish the addresses of the servers publically, and does not want to. Even if the servers can be accessed from the Internet by IP address, they will not serve any TV content that way. It is more likely that the content servers are completely inaccessible from the Internet via firewall rules and/or routing horizons.

2.2. Corporate VPN

[SERVER-SELECTION] Section 3.2 "Cellular Network Scenario" provides more details.

The corporate VPN use case involves an employee of an enterprise connecting to the enterprise's network via an IP over IP tunnel. IPsec is a typical technology, and there are various kinds of SSL VPNs, but this use case concerns itself with scenarios where IPv6 pakets will be routed through the tunnel. (Some SSL VPNs provide TCP, usually HTTP-only services. They are not relevant)

As described in the introduction, the Corporate VPN is almost the original walled garden: there is a seperate routing domain (RFC1918 in IPv4, End-User Assigned IPv6, Non-Connected Network or ULA for IPv6), and there is an addition a desire to have an island of DNS.

Some corporate VPNs have a policy that all traffic from the employees' computing device must go through the tunnel, even traffic not addressed to the Enterprise. (That is, the default route on the

laptop must point through the tunnel. A host route for the VPN tunnel end point is usually created to permit the encapsulated traffic to travel, or another kind of source-address sensitive policy route is used to create multiple routing tables). These types of VPNs are sometimes called an extruded IP VPN: if the enterprise' network is thought of as a fungible sponge, then the laptop user at home as an IP address, from the corporate cloud/sponge, extruded, rather like a sea urchin (??) extends a pseudopod.

It is not unusual for the networks of enterprises to grow to be very complex: multiple sites with multiple (sometimes overlapping IPv4 RFC1918) address spaces. With IPv6, enterprises are likely to have fewer blocks of addresses, and none will overlap, but due to acquisitions the number of blocks will still be greater than one.

Frequent teleworkers often have an entire home office connected to the enterprise network. This can include a selection of desktop computers, laptops, tablets, and smartphones (on wifi). The home office network might have several subnets (at least one wired and one wireless). These teleworkers usually have a VPN router as their home gateway, and often these devices are controlled by the Enterprise IT. Just the same, access to the rest of the homenet is desired such that the home worker(s) can share things like printers, talk to home automation from their desk, and to avoid having large-bandwidth cross the corporate network unnecessarily.

More about DNS expectations of enterprises. ActiveDirectory is relevant here.

2.3. Multi-homed to 3G/LTE

[SERVER-SELECTION] Section 3.2 "Cellular Network Scenario" provides more details.

There is a problem here: so far it sounds like it's just IP TV over again.

3. Walled Garden DNS

[SERVER-SELECTION] proposes a solution where ISP/connectivity providers provide a suffix, such as "mytv.example." or even unregistered names like ".internal" (common with ActiveDirectory) into the homenet.

Along with suffix are provided one or more DNS servers, accessible via the semi-private connection, which would resolve names in that zone.

4. Using Secret-Gardens to accomplish goals

As [SERVER-SELECTION] essentially mandates that DNSSEC be done, it has essentially mandated that a recursive cacheing validating name server be run on each host. Many have suggested that, as the days of 20Mhz, 8MB workstations are long gone, that having a caching, validating name server on every host is not a problem. Once one assumes this, new solutions become possible.

The fundamental problem that [SERVER-SELECTION] tries to solve is that of mapping DNS suffixes to name servers. [RFC1034] provides a way to do this: the NS resource record. [DNSSEC] provides a way to do this securely by adding the DS resource record.

So what was stopping enterprises such as IP TV, or corporate VPNs from putting in-accessible IP addresses into a reachable DNS server? Usually two things: the perception that the security through obscurity benefit of split-horizon DNS provided was greater than the operational complexity and inconvenience of using a single namespace. The second thing was that putting RFC1918 addresses in a publically visible DNS would be confusing.

IPv6 addresses removes both problems. First, since almost every walled-garden situation insists that any machine talking to it use an address in the prefix that the walled-garden provides. This means that if one puts the DNS servers for the walled garden names in the walled gareden then can be queried only from machines that are already in the garden.

Second: since IPv6 addresses are all unique, there is no confusion when an IPv6 address is listed. It is therefore possible to unambiguously indicate in an NS that that mytv.example.com is reachable via DOCPREFIX:1234:0001. Just because the entire Internet can see the (possibly secure) delegation doesn't meant that they can query it.

4.1. Example 1: Corporate VPN

A company EXAMPLE has been given the assignment 2001:DB8:0F00:/48. It has the forward name example.com. In addition, due to an acquisition, it also owns example.org and assignment 2001:DB8:0123:/48. Further, assume the VPN concentrator for example.com gives /128 tunnel addresses to clients from the 2001:DB8:0F00:ff91::/64 range. In this example, the client has been assigned 2001:DB8:0F00:ff91::1234.

Routes are established through the VPN tunnel for the above two prefixes:

```
% netstat -rn -f inet6
Routing tables

Internet6:
Destination           Gateway                     Flags     Interface
...
2001:db8:0f00::/48    2001:db8:0f00:ff91::1       UGRS      ipsec0
2001:db8:0123::/48    2001:db8:0f00:ff91::1       UGRS      ipsec0
...
```

Figure 1

Rather than create a split-horizon DNS server on the inside that answers (differently) for example.com, the enterprise instead creates a new zone internal.example.com:

```
% dig internal.example.com. ns
;; QUESTION SECTION:
;internal.example.com.             IN      NS

;; ANSWER SECTION:
internal.example.com.  7200    IN    NS     ns1.internal.example.com.
internal.example.com.  7200    IN    NS     ns2.internal.example.com.

;; ADDITIONAL SECTION:
ns1.internal.example.com.  7200 IN   AAAA   2001:db8:0f00:0001::8888
ns2.internal.example.com.  7200 IN   AAAA   2001:db8:0f00:0002::8888
```

Figure 2

Note that these records are in the public example.com zone, and are visible to the Internet.

The ns1 and ns2 servers SHOULD be configured such that they will not answer queries for internal.example.com, unless the query comes from 2001:db8:0f00::/48.

What is the effect here? A host with it's VPN active performs a query for thing.internal.example.com. The local recursive DNS server chases the NS pointers from the root, and it gets the above NS records. It then performs a query to ns1.internal.example.com for the name thing.internal.example.com. Since the route for this IP address goes through the ipsec0 device, the source address for the packet will be 2001:DB8:0F00:ff91::1234. That source address fits into the policy for the ns1 server, so it answers.

There are a number of other affects: 1) the result of this lookup is cacheable, and does not need to be flushed if the node moves. 2) if the VPN is not active, then the result of performing the lookup for thing.internal.example.com causes a packet to be sent to the 2001: db8:0f00::/48 network. If policies are setup right, then this packet may well cause the VPN to be started (or even restarted, if it failed).

5. How does this differ from SERVER-SELECTION?

run through use cases, show how this is equivalent or better.

6. Security Considerations

7. Other Related Protocols

8. IANA Considerations

9. Acknowledgements

10. References

10.1. Informative References

10.2. Normative References

11. Normative references

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

Author's Address

Michael C. Richardson
Sandelman Software Works
470 Dawson Avenue
Ottawa, ON K1Z 5V7
CA

Email: mcr+ietf@sandelman.ca
URI: http://www.sandelman.ca/