VeriSign's Site Finder Implementation

VeriSign Naming and Directory Services
VeriSign, Inc.

27 August 2003

TABLE OF CONTENTS

1 Introduction

This document describes aspects of VeriSign's implementation of its Site Finder service and its use of "wildcard" functionality as specified in the DNS protocol.

VeriSign's Site Finder service improves the user web browsing experience when the user has submitted a query for a nonexistent second-level domain name in the *.com* and *.net* top-level domains. Before this service was implemented, when a user entered a URL containing a nonexistent (e.g., unregistered) domain name ending in *.com* or *.net*, his or her web browser returned an error message that contained no useful information. With the rollout of Site Finder, in the same situation users now receive a helpful web page offering links to possible intended destinations and allowing an Internet search.

VeriSign refers users to the Site Finder web site through the use of a wildcard address (A) record entry in the *.com* and *.net* zones. As explained more fully below, VeriSign's processing of queries for nonexistent domain names is in full compliance with provisions of the DNS protocol that address wildcards as well as the operational best practices described in the document entitled *Domain Name System Wildcards in Top-Level Domain Zones* ("the Guidelines") [1].

2 Aspects of Wildcard Implementation

2.1 Time to Live (TTL) of Wildcard Entries

The wildcard A record entries in the *.com* and *.net* zones have a TTL of 15 minutes. This value is large enough to provide some caching benefit and prevent a busy recursive name server's cache from being overloaded by synthesized wildcard responses. This TTL value was purposefully chosen to be of the same magnitude as a typical negative caching timeout value that would be experienced without Site Finder.[1]

2.2 General Responses to Application Protocols

While HTTP is the main application protocol supported by Site Finder, VeriSign recognizes the importance of having its response server provide appropriate responses to connections initiated through other protocols.[2] VeriSign has carefully researched the application protocols that make connection attempts to VeriSign's response servers. In accordance with applicable Guidelines, VeriSign has configured its response servers to respond to connection attempts on both Transmission Control Protocol (TCP) and User

[1] This TTL value is consistent with Guideline G1.1, which provides that: "Zone administrators should consider choosing an appropriate time to live (TTL) value for wildcard entries. A TTL value of no longer than 15 minutes for any wildcard records in TLD zones is suggested."

[2] The term "response server" refers to the server whose IP address is specified in the RDATA of a wildcard A record. The response server receives the traffic from various clients supporting various application protocols because of the presence of the wildcard A record.

Datagram Protocol (UDP) ports assigned to various application protocols.[3] Additionally, the response server responds to all TCP connection attempts and UDP packets (with the exception of a short list of application protocols). See Section 2.3 below for a detailed explanation.

Before Site Finder was deployed, VeriSign conducted research to determine the most popular application protocols, anticipating that the response server would receive connection attempts from applications supporting these protocols. During pre-production testing, actual network traffic received by the response server was analyzed to refine this list of application protocols. VeriSign analyzed the semantics of popular application protocols and the behavior of popular implementations of these protocols to determine the best response in each case. The specific response server behavior is described in more detail below in Section 2.3.[4]

2.3 Responses to Specific Application Protocols

This section describes how the Site Finder response server handles specific application protocols.

2.3.1 HTTP

A user reaches the Site Finder web page after he or she enters a URL containing a nonexistent *com* or *net* domain name. For the initial version of Site Finder, this web page is not localized for language, but VeriSign is actively researching localization options that would allow the service to accept search parameters and display responses in a user's native language. The text of the Site Finder web page makes it clear to a user that he or she did not reach the intended web page, but instead reached the Site Finder page as a result of special processing. [5]

Site Finder follows the Guidelines regarding the Robots Exclusion Protocol [2], including a *robots.txt* file that instructs a robot to not access any part of the site[6]:

---

[3] See Guideline G1.2 ("A response server should be configured to respond to connection attempts on both Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) ports assigned to various application protocols. The number of TCP and UDP ports that must be managed on a response server is large (but finite), with a small number of application protocols expected to produce a majority of the connection attempts.")

[4] See Guidelines G1.3 through G1.6.

[5] This approach is consistent with Guideline G1.7, which states:

> An HTTP response server should provide a web page with localized human-readable text to clearly explain that the page is being displayed as a result of wildcard processes along with additional information that might be useful to the user. The user can then make informed decisions about their options for subsequent actions.

[6] This instruction is in accordance with Guideline G1.8, which provides that: "An HTTP response server should instruct World Wide Web robots to not access or archive any of the web pages found on the server."

```
User-agent: *
Disallow: /
```

2.3.2 SMTP

The Site Finder response server runs a limited SMTP server that returns an SMTP 550 error response for any specified destination (i.e., any RCPT TO value). The error message includes explanatory text explaining that the message did not reach the intended recipient because the destination domain does not exist.[7]

2.3.3 Other Application Protocols

Regarding other application protocols, the VeriSign response server actively returns an appropriate transport protocol error response to discourage further connection attempts.[8] TCP connection attempts to any ports other than 25 (SMTP) and 80 (HTTP) result in the connection being reset, and any UDP packets cause the return of an ICMP port unreachable message.[9]

The vast majority of network traffic to the Site Finder web site receives a response.[10] The exception is a list of protocols that are not allowed into VeriSign's network for security reasons. This following list application protocols is filtered at the network border and no response is returned because of this filtering:

- UDP port 161 (SNMP)
- UDP port 162 (SNMP traps)
- UDP port 514 (syslog)
- UDP port 1985 (Cisco HSRP)
- TCP port 23 (Telnet)
- TCP port 79 (Finger)
- TCP port 161 (SNMP)
- TCP port 162 (SNMP traps)

---

[7] This approach is in accordance with Guideline G1.9, which states:

> A response server should provide a limited SMTP server that returns an SMTP 550 error response for any destination address to indicate that mail delivery is not possible, with text provided to explain why the message is being rejected. This active error response will allow the sending server to take appropriate action for all undeliverable messages.

[8] See Guideline G1.10.

[9] See Guideline G1.11 ("The response server should exhibit standard behavior in response to connection attempts to TCP and UDP ports without corresponding applications listening. For TCP-based protocols, the response server should immediately reset the connection (i.e., in response to a client's initial TCP packet with the SYN flag set, the server should return a TCP packet with the RST flag set). For UDP-based protocols, the response server should immediately respond with an ICMP port unreachable message. These are 'hard errors' and should cause properly written applications to cease further connection attempts.); and Guideline G1.12 ("Response servers should not return other ICMP messages, such as host or network unreachable, which is interpreted by many TCP/IP implementations as a 'soft error' and will not stop connection attempts.)

[10] See Guideline G1.13 ("Likewise, the response server (or network devices in front of it) should not simply drop packets corresponding to connection attempts from unsupported protocols.")

- TCP port 514 (BSD remote shell)
- TCP port 1985 (Cisco HSRP)
- TCP port 4750 (Network Shell)

2.4 Monitoring and Communication

VeriSign actively monitors all traffic associated with Site Finder, including DNS queries matching the wildcard entries in *.com* and *.net* and associated responses, and all traffic sent to the response server.[11] This traffic is correlated and monitored in real time, 24 hours a day, seven days a week, by VeriSign's Network Operations Center. Anomalous events are escalated to engineering staff for analysis. Several hours of the complete traffic stream to the *.com* and *.net* name servers and the response server, as well as rolled up statistics, are stored for analysis.

VeriSign plans to adjust response server behavior as necessary in response to results obtained from this comprehensive monitoring program.[12]

While this monitoring data will not be publicly available at the launch of Site Finder, VeriSign is considering making this information available in the future.[13]

By long-established convention, a zone's administrator can be contacted by sending an email to the address published in the RNAME field of the zone's SOA record. VeriSign publishes the email address *nstld@verisign-grs.com* in the RNAME field of the *.com* and *.net* zone SOA records and actively reads and responds to legitimate messages sent to this address.[14]

The response server supporting Site Finder is named *sitefinder-idn.verisign.com* and this domain name has a single A record that corresponds to the response server's IP address.[15] This address can also be obtained by querying for a wildcard A record in *.com* or *.net*.

---

[11] See Guideline G1.14 ("Each TLD likely has unique characteristics in terms of the applications that will produce DNS queries and attempt response server connections, so an active DNS query and response server port monitoring program should be in place to remain aware of usage trends as they develop and change over time.")

[12] See Guideline G1.15 ("As usage trends develop, the response server should be reconfigured as necessary to provide appropriate services. An active monitoring program is also recommended to detect and respond to denial of service attacks.")

[13] See Guideline G1.16 ("Publishing the results of the DNS query and response server port monitoring program in a public place, such as on the TLD zone administrator's web site, will allow the user community to see the benefits, including added user efficiencies, provided by the new service.")

[14] See Guideline G1.17 ("Zone administrators should provide a means to allow contact from individuals and organizations that experience operational issues with a wildcard response server. Support for community-based monitoring increases the likelihood that the zone administrator will be made aware of operational issues in a timely manner.")

[15] See Guideline G1.18 ("The response server and the services it provides should be visible using a well-known host name and Internet Protocol (IP) address. This will allow applications to recognize DNS responses directing them to the response server, and give them an opportunity to take appropriate actions (such as notifying the user) if the wildcard synthesis is unanticipated.")

For example, using the standard DNS query tool *dig*, the wildcard A record for *.com* could be obtained with this invocation:

```
dig @a.gtld-servers.net a *.com.
```

The text of the Site Finder web page makes it clear to a user that he or she did not reach the intended web page, but instead reached the Site Finder page as a result of special processing. Further, when a user's web browser connects to the response server using HTTP, the browser is redirected to a URL containing the domain name *sitefinder.verisign.com*. This redirection causes the user's web browser to change the displayed URL, which clearly indicates that the user has not reached the intended site.[16]

The Site Finder response server is actually a combination of multiple servers and network load balancing devices located at two physically distinct sites. Without revealing details that would aid an attacker, the overall processing capacity and network bandwidth is over-provisioned to withstand denial of service attacks. Network traffic to all of the *.com* and *.net* name servers and the response servers is continuously monitored for anomalies, including using any of these resources as a "reflector" for a denial of service attack against another party.[17]

3 References

1. Hollenbeck, S. and Larson, M., "Domain Name System Wildcards in Top-Level Domain Zones", http://www.verisign.com/resources/gd/sitefinder/bestpractices.pdf, August 2003.
2. Koster, M, "A Standard for Robot Exclusion", http://www.robotstxt.org/wc/norobots.html, June 1994.

[16] See Guideline G1.19 ("A response server should return information to indicate that the provided services were reached as a result of a wildcard entry. In some cases it may be appropriate to not provide services at all, using appropriate application or network error responses to inform the client that services are unavailable.")

[17] See Guideline G1.20 ("With or without wildcards, if queries contain a spoofed source address the DNS responses can be used to mount a flooding attack against the machine associated with the spoofed address. Response server operators should anticipate the possibility of packet flooding "denial of service" attacks. TLD name servers receive many queries for nonexistent domain names as a result of application errors and user data entry errors. Consistent with all packet flooding "denial of service" attacks, if the TLD name server is flooded with queries for non-existent domain names, the DNS wildcard responses returned for these queries may result in applications attempting to contact the response server.")