Internet Society

The ISP Column
A monthly column on things Internet

December 2009

Geoff Huston

NXDOMAIN?

Sometimes its not just the words that create meaning in a conversation, but the gaps in between the words, or the white spaces that can also convey meaning. Sometimes it's not just what is said, but also the manner of what is not said that conveys meaning and has value. Oddly enough, this observation about the value of the spaces between the words in human communications carries through to the Domain Name System, or DNS.

It might seem a little strange, but in the current economics of the market in registration of DNS names it appears that the set of names that are not "visible," or at least not associated with any dedicated network service point, represents a far larger set, and has a far higher total value to the DNS name registration industry, than the set of network-visible service endpoint domain names. In other words, there appears to be a larger and more valuable market for names that do not exist than for names that do. Dedicated cynics may well say that this is nothing new, and that the ongoing exercises in mapping telephone numbers into domain names (ENUM), and the work with Internationalized Domain Names (IDNs) are just expressions of the same desire on the name of the name registry business to create massive numbers of new and as-yet unregistered domain names in order to bring more value into the name registration business. I must admit that this observation is perhaps a little too cynical for my tastes, as I'd like to think that the IDN work in particular does indeed create a useful means of mapping the DNS into non-roman script environments. However, the observation still holds that in an industry that charges the same fixed annual wholesale fee for each and every registered name, then business self-interest on the part of every registry lies the encouragement of more names to register, not less. And if its possible to register names that are not associated with visible service endpoint on the Internet, then this is just fine. Whether the name exists or not as a service point, the registration fee remains the same.

It appears that its not just the name registry business that sees value in names that don't exist. For an ISP providing DNS name resolution for it's customers can be seen as an overhead and a cost burden, rather than as a revenue-generating activity. Every ISP has found that its customers expect it to operate a DNS name resolver system. Such name resolvers often become points of vulnerability for DOS attacks and similar, so its also the case that the ISP often has to spend some time, effort and of course money in setting up a DNS resolver architecture that is strongly resilient to attack. As the attacks become more virulent in terms of the imposed query load, the ISP is often forced to invest ever increasing sums in larger platforms and pay license costs for high performance DNS resolver software systems that go beyond the standard public domain DNS resolver code offerings. But ISPs are businesses, and investments in infrastructure services are like any other business investment - they should generate a financial return to the business. But making this investment in DNS resolvers generate a financial return is not easy. It's not as if the DNS is constructed with a transaction-based charging model in mind. Indeed, its quite the opposite. So it becomes a problem for the ISP. The ISP has to make continual investment in DNS infrastructure, but do so without any offset in incremental revenue. The DNS becomes in effect an off-the-top cost that plays its part in eroding the overall revenue margins for the ISP business.

So, at some point, it should not come as a surprise when a product manager within an ISP enterprise looks at the expense line for DNS infrastructure and wonders "How can I generate revenue from this service?" And then comes that little spark: "If we can't sell DNS resolution of names that already exist in the DNS to our customers or to the name holders, let's see if we can sell the blank spaces between the DNS names!"

Who would buy non-existent DNS names? Well, it should come as no surprise that in a world where there is already a large and valuable market for selling DNS names that are not Internet-visible as service endpoints, there is also a valuable market in identifying yet more names that users are using in their applications that are not even visible to the DNS. There is value in catching the NXDOMAIN responses from a DNS resolver and substituting a page impression. There is value in the so-called practice of "typosquatting".

Wildcards and the DNS

One of the most direct ways of "filling" the white spaces in a DNS zone is via a wildcard entry in the DNS Zone file. The domain name administrator can add a wildcard entry to the zone file, and return a common response for all names that do not have an explicit entry in the zone file. In this case a query for any name that does not have an explicit entry in the DNS zone file generates the wildcard entry in response.

```
*.example.com 3600 IN A 192.0.2.1
```

In this case a query for an A record for the name "www.example.com" would provide the response 192.0.2.1. Interestingly, the wildcard entry works to arbitrary depths, so that a query for an A record for the name "w.w.w.example.com" would also provide the same response of 192.0.2.1. However, the wildcard is also selective, in that if a second record exists in the example.com zone, such as

```
test.example.com 3600 IN AAAA 2001:DB8::1
```

then a query for www.test.example.com will not refer to the original wildcard, but instead return a "no such domain", or MXDOMAIN response.

However, this action is generally only a realistic option for zones that exist at lower levels in the DNS name hierarchy. Some notable efforts some years back by Verisign to insert a wildcard entry in the .com zone file generated considerable angst and disquiet from the technical and name registrar communities, and the exercise was unravelled relatively quickly.

Why was there so much of a reaction? If its ok to do this at lower levels in the DNS name hierarchy, why is it not ok to do the same at the upper levels of the name space?

> A good summary history of the wildcard entry in the .com domain, or the "sitefinder episode" is in Wikipedia. [http://en.wikipedia.org/wiki/Site_Finder].

One of the issues raised at the time was the concept of "typosquatting", where the wildcard entry implicitly created a redirection of the user's original query to the address of some other service point, as directed by the wildcard entry. [http://en.wikipedia.org/wiki/Typosquatting]

> One of the better summaries of the issues surrounding the use of wildcards in the DNS was written by the Internet Architecture Board. The IAB document is an interesting read, and one part I found to be helpful was the section in architectural principles, describing the "Principle of Least Astonishment"

> Principles To Keep In Mind
>
> In reading the rest of this document, it may be helpful to bear in mind two basic principles of architectural design which have served the Internet well for many years:
>
> The Robustness Principle: "Be conservative in what you do, be liberal in what you accept from others." [Jon Postel, RFC 793]
>
> The Principle Of Least Astonishment: A program should always respond in the way that is least likely to astonish the user. [Traditional, original source unknown]
>
> [http://www.iab.org/documents/docs/2003-09-20-dns-wildcards.html]

On the other hand, a self-inflicted wildcard can be useful for dynamic content. I've played around with such a construct in the domain "potaroo.net", where any RFC can be retrieved using the name form rfc<*number*>.potaroo.net, such as http://rfc4592.potaroo.net.

Wildcards are a way that a domain name zone administrator can "cover" all the possible names in a DNS zone, and with some limitations, also cover all possible subdomains of that zone.

But is there a way to "cover" all non-existent names in all zones? Here's where NXDOMAIN comes into play.

NXDOMAIN Substitution

Conventionally, when a query is made to return a resource record associated with a non-existent DNS name, the DNS system will return a "NXDOMAIN" response, indicating "no such domain name exists". This is different from the response where the name exists, but the requested resource record is not defined in the DNS zone file. In the case of NXDOMAIN this is a DNS response code (RCODE) value 3, indicating that the name itself is not defined in the DNS.

This leads to an observation that at the resolver level it is possible to take a NXDOMAIN response from the authoritative server, substitute a different, and positive, response, and pretend that this is an authoritative answer. This form of "typosquatting" is a a substitution in the middle, where the actions of a DNS resolver are explicitly altered, without the knowledge, let alone the informed agreement, of either the DNS zone administrator, nor that of the end user and the end user's applications.

Here's a recent example from a large Australian ISP:

> **BigPond redirects typos to 'unethical' branded search page**
>
> By Josh Mehlman
> Nov 20, 2009

> Telstra has begun redirecting failed domain name lookups on its BigPond broadband network to a Telstra-branded search page which includes advertisements and pay-per-click links.
>
> Telstra has already rolled out this system for BigPond broadband customers in Queensland with other states to follow by mid-December. Optus had adopted a similar strategy last month, much to the ire of some of its subscribers.
>
> In an email sent to customers, Telstra emphasised the convenience of the new service.
>
> "BigPond Broadband customers from today will experience less keyboard anguish with our new DNS redirection application. Typos can happen at any time. It's frustrating when you get an MS Internet Explorer Live Search page with suggestions from a mistyped URL. It's even more annoying when you get an unfriendly "404" page error message," the letter said.
>
> "The BigPond-branded landing page provides BigPond customers with organic search results, sponsored links, display advertisements and intelligent recommendations, all derived from the invalid domain input - much more helpful and friendly than a nasty 404 page error."
>
> [http://www.crn.com.au/News/160923,bigpond-redirects-typos-to-unethical-branded-search-page.aspx]

There are many possible responses to this: Good? Useful? Harmless? Just annoying? Harmful? Unethical? Insecure?

A Brief Digression into the DNS

If you took an extreme technology-centric view of the networked world, names are completely irrelevant. In the IP world routers don't understand domain names in order to forward packets. More generally, the task of packet forwarding or even circuit establishment requires only knowledge of network addresses, not names. At the level of pushing the bits around, if I want to communicate with you I need to know your network address, and if you want to reply to me then you need to know my network address.

One could stop right here and what you would have is the telephone network. Within an address-only network users need to know numbers to call, not names. The association of names to numbers was traditionally held in a manual lookaside list, otherwise known as a telephone directory.

Computer networks started out using the same paradigm of making the numerical value of the addresses visible to the user, but quickly veered off into a world of names, where names were mapped into a numerical address via a local alias files.

> In the same way that we carry traces of our distant ancestral forebears in our DNS, many computers still carry the vestiges of this earliest form of name to number mapping in their file system.
>
> For example, on MAC OSX I see:

```
$ more /etc/hosts
##
# Host Database
#
# localhost is used to configure the loopback interface
# when the system is booting.  Do not change this entry.
##
127.0.0.1       localhost
255.255.255.255 broadcasthost
::1             localhost
fe80::1%lo0     localhost

[http://en.wikipedia.org/wiki/Hosts_file]
```

The problems with having every computer maintain its own complete copy of name to number mapping are forbidding. Ensuring that everyone has the same information, and that new updates are unique, and that old information is flushed out of each and every locally maintained alias file poses a logistical nightmare! Consequently, its little wonder that the DNS appeared at a very early stage in the evolution of the Internet. The DNS is a hierarchical, distributed database that straddles the entire namespace of the Internet.

In terms of a single digital artefact, the DNS is a truly impressive system. It has millions of discrete elements, updates are being performed continuously, yet it manages to provide accurate responses to each and every transaction on the internet, generally within unnoticeable factions of a second. Looking up a web page, sending mail, or performing any other transaction on a network that starts with a name, then the DNS is invoked.

The fact that it works at all is remarkable. The fact that it works reliably, quickly and efficiently for most of the time across hundreds of millions of individual DNS transactions going on all the time is, simply, amazing. The fact that the commercialisation of the DNS has generated a revenue stream in the name registration industry that is now well into billions of dollars each and every year is unbelievably prodigious! To quote Paul Vixie:

> "DNS is, in a word, glorious"

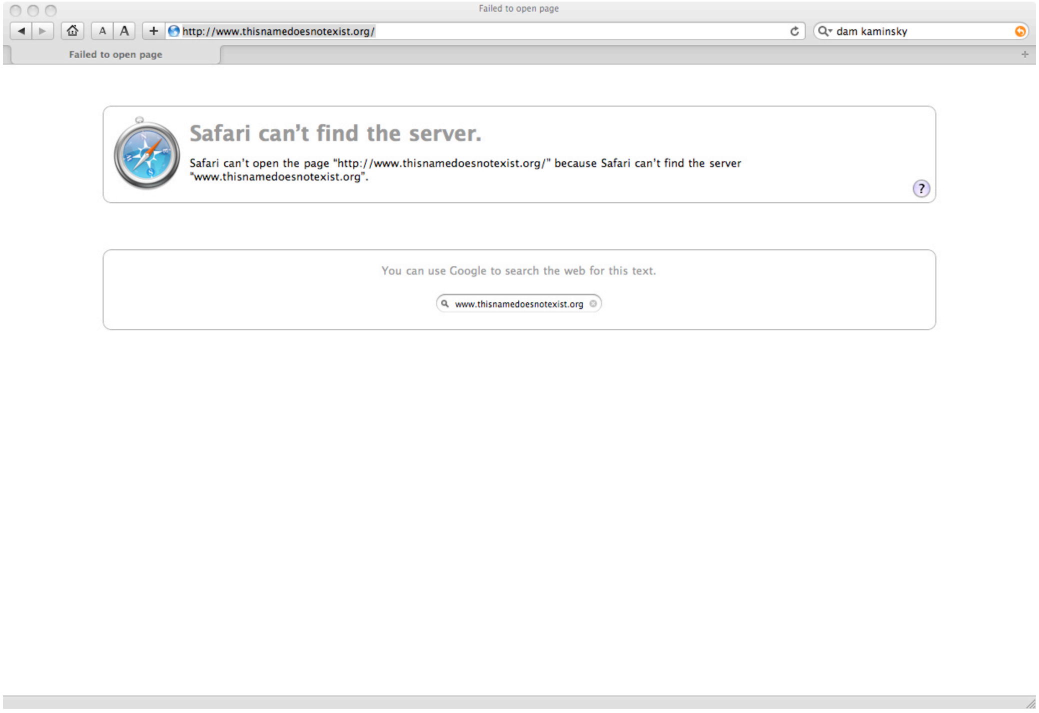
NXDOMAIN Substitution - Good or Evil?

A DNS query is of the form of a domain name and a Resource Record Type. When a DNS zone's name server is queried for a name that is not defined in the zone file, and there is no wildcard record in the zone, then the server should send back a negative response code 3, or NXDOMAIN. The application that generated the query is then able to respond to the user as appropriate. If this is a mail delivery application, then the mail application can reject the mail with a notice that the mail address is invalid. If this is a web browser, then the browser will display a "page not found" response. Every well structured application that uses the DNS has to cope with these negative responses.

The substitution approach is via a modified DNS resolver, or intermediary, who, when served with an NXDOMAIN response from an authoritative nameserver will substitute a local resource record response in its place. The intention is to "capture" the web browser and replace the standard "no such domain" browser page with a page that is filled with their own funded ads. In a world of funded impressions, eyeballs and clickthroughs this substitution represents a new income stream from the DNS.

And a number of ISPs, large and small, have taken up this ability to deploy DNS resolvers that include NXDOMAIN substitution over the past couple of years, and captured any DNS impression revenue that can be gained from this list of non-existent domains.

The argument could be made that this practice is at worst harmless, and could possibly be considered as a benefit for the end user. Harmless, in that the substitution of a browser's "I can't find a server for this URL" with its typical inclusion of a link to a search engine, with an ISP's search page with a web provided by the ISP is not exactly veering off in an unusual direction.



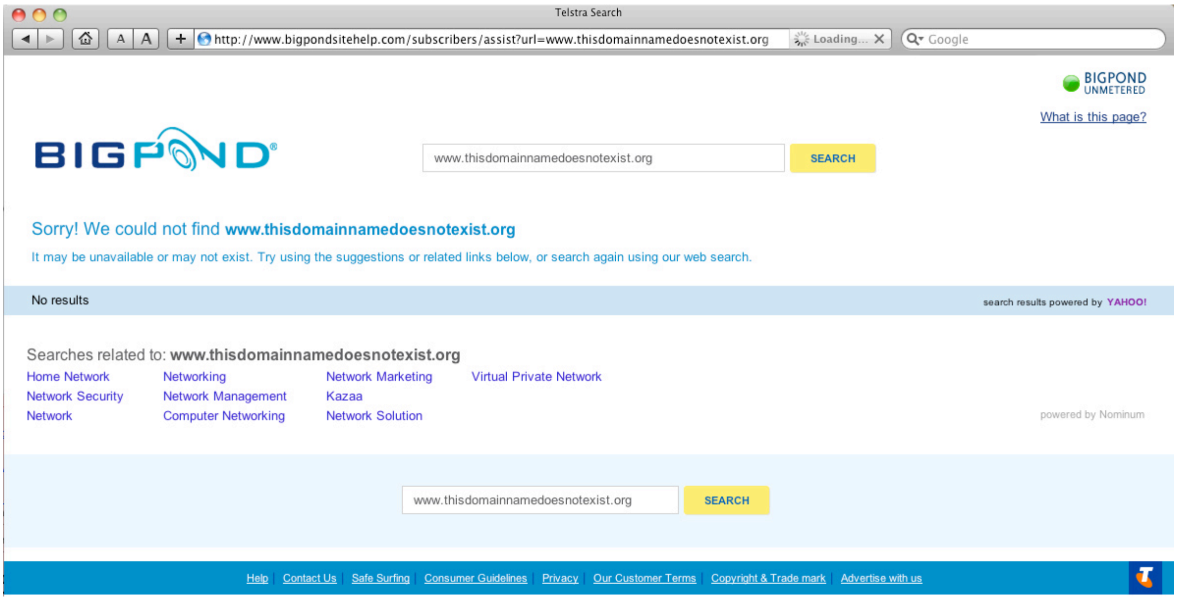
Safari can't find the server.

Safari can't open the page "http://www.thisnamedoesnotexist.org/" because Safari can't find the server "www.thisnamedoesnotexist.org".

You can use Google to search the web for this text.

www.thisnamedoesnotexist.org

Safari's "NXDOMAIN" page


Telstra Search
http://www.bigpondsitehelp.com/subscribers/assist?url=www.thisdomainnamedoesnotexist.org
Loading...
Google
BIGPOND UNMETERED
What is this page?
BIGPOND
www.thisdomainnamedoesnotexist.org
SEARCH

Sorry! We could not find **www.thisdomainnamedoesnotexist.org**

It may be unavailable or may not exist. Try using the suggestions or related links below, or search again using our web search.

No results

search results powered by YAHOO!

Searches related to: **www.thisdomainnamedoesnotexist.org**

Home Network	Networking	Network Marketing	Virtual Private Network
Network Security	Network Management	Kazaa	
Network	Computer Networking	Network Solution	

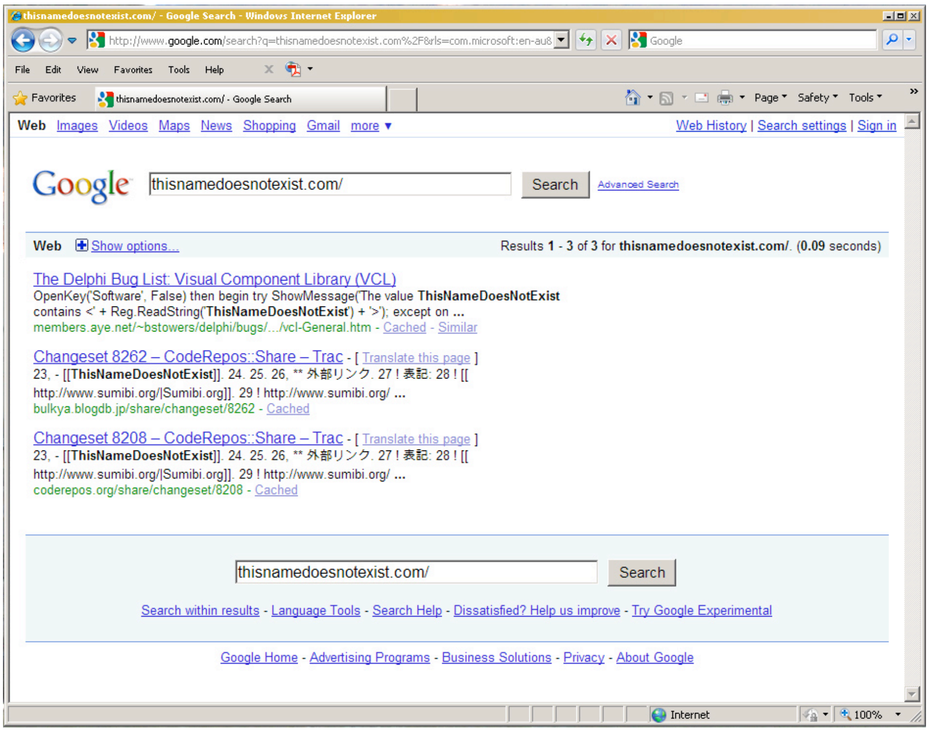
powered by Nominum

www.thisdomainnamedoesnotexist.org
SEARCH
Help | Contact Us | Safe Surfing | Consumer Guidelines | Privacy | Our Customer Terms | Copyright & Trade mark | Advertise with us

Telstra's ISP service NXDOMAIN page

Aside from the window dressing there is very little that is different here between these two pages.

As an aside, I thought that for completeness it would be good to reproduce the actions that Windows Explorer takes on an NXDOMAIN response. If you just enter a non-existent domain name into the address bar Explorer will provide a search results page based on the search provider you've configured into Explorer.



thisnamedoesnotexist.com/ - Google Search - Windows Internet Explorer

http://www.google.com/search?q=thisnamedoesnotexist.com%2F&rls=com.microsoft:en-au8

thisnamedoesnotexist.com/

Results **1** - **3** of **3** for **thisnamedoesnotexist.com/**. (**0.09** seconds)

The Delphi Bug List: Visual Component Library (VCL)
OpenKey('Software', False) then begin try ShowMessage('The value **ThisNameDoesNotExist** contains <' + Reg.ReadString('**ThisNameDoesNotExist**') + '>'); except on **...**
members.aye.net/~bstowers/delphi/bugs/.../vcl-General.htm - Cached - Similar

Changeset 8262 – CodeRepos::Share – Trac - [Translate this page]
23, - [[**ThisNameDoesNotExist**]]. 24. 25. 26, ** 外部リンク. 27 ! 表記: 28 ! [[http://www.sumibi.org/|Sumibi.org]]. 29 ! http://www.sumibi.org/ **...**
bulkya.blogdb.jp/share/changeset/8262 - Cached

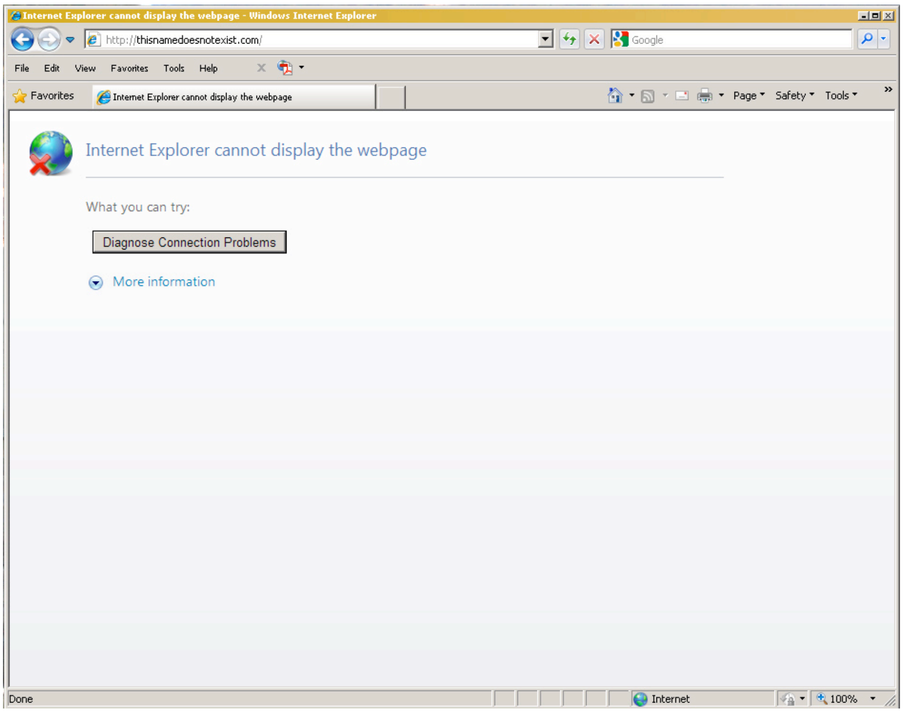
Changeset 8208 – CodeRepos::Share – Trac - [Translate this page]
23, - [[**ThisNameDoesNotExist**]]. 24. 25. 26, ** 外部リンク. 27 ! 表記: 28 ! [[http://www.sumibi.org/|Sumibi.org]]. 29 ! http://www.sumibi.org/ **...**
coderepos.org/share/changeset/8208 - Cached

thisnamedoesnotexist.com/

Search
Search within results - Language Tools - Search Help - Dissatisfied? Help us improve - Try Google Experimental
Google Home - Advertising Programs - Business Solutions - Privacy - About Google
Internet
100%

Internet Explorer's NXDOMAIN page

But if you enter a well formed URL into Explorer, such as http://thisnamedoesnotexist.com, then you get a somewhat more cryptic screen, where the set of hints shown in "more information" is not completely helpful.



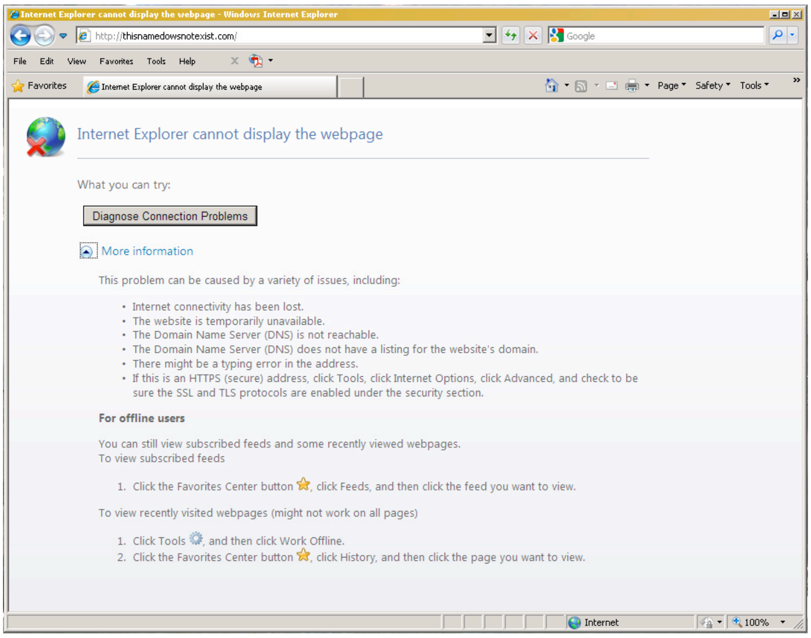
Internet Explorer cannot display the webpage - Windows Internet Explorer

http://thisnamedoesnotexist.com/

Internet Explorer cannot display the webpage

What you can try:

Internet Explorer's more cryptic NXDOMAIN page


Internet Explorer cannot display the webpage - Windows Internet Explorer
http://thisnamedowsnotexist.com/
Google
File
Edit
View
Favorites
Tools
Help
Favorites
Internet Explorer cannot display the webpage
Page
Safety
Tools
Internet
100%

Internet Explorer cannot display the webpage

What you can try:

Diagnose Connection Problems

More information

This problem can be caused by a variety of issues, including:

- Internet connectivity has been lost.
- The website is temporarily unavailable.
- The Domain Name Server (DNS) is not reachable.
- The Domain Name Server (DNS) does not have a listing for the website's domain.
- There might be a typing error in the address.
- If this is an HTTPS (secure) address, click Tools, click Internet Options, click Advanced, and check to be sure the SSL and TLS protocols are enabled under the security section.

For offline users

You can still view subscribed feeds and some recently viewed webpages.
To view subscribed feeds

1. Click the Favorites Center button ☆, click Feeds, and then click the feed you want to view.

To view recently visited webpages (might not work on all pages)

1. Click Tools ⚙, and then click Work Offline.
2. Click the Favorites Center button ☆, click History, and then click the page you want to view.

Internet Explorer's NXDOMAIN explanation

From this perspective it does not seem that intercepting a DNS NXDOMAIN response and substituting a splash page is exactly a harmful act to the Internet and its users, even with the ISP gathering of page impressions of DNS typos.

But is this really as benign and as harmless as it seems?

As Paul Vixie has pointed out in his recent article on "What the DNS is Not" (published by the ACM in its Queue Column [http://queue.acm.org/detail.cfm?id=1647302]): "NXDOMAIN wasn't designed to be a revenue hook—many applications depend on accurate error signals from DNS." The best example of this was shown by Dan Kaminsky and Jason Larsen when they announced to the Toorcon conference in 2008 things are not always as straightforward as they seem with NXDOMAIN substitution. The problem lies in the combination of the "same origin trust model" used by Web cookies and NXDOMAIN substitution. Back to Paul Vixie for an excellent description of the potential problem that Dan and Jason reported:

> "If you're holding a cookie for GOOGLE.COM and you can be fooled into following a link to KJHSDFKJHSKJHMJHER.GOOGLE.COM, and the resulting NXDOMAIN response is remapped into a positive answer to some advertising server, then you're going to send your cookie to that advertising server when you send your HTTP GET request there. Not such a bad thing for a GOOGLE.COM cookie, but a real problem for a BANKOFAMERICA.COM cookie."

The problem here is that this form of name substitution has the potential to break some of the assumptions behind the security models of the DNS. For example, if you have administer a family of services under a common domain name, such as for example, ftp.example.com, www.example.com, mail.example.com, than you may assume a security model where referrals from one subdomain of example.com to another may have the ability to transfer information about the content of the user between such servers. The intrusion of name substitution allows third parties to populate the same zone with third party servers, who can therefore enter into the same security domain of the common namespace.

NXDOMAIN and DNS Lies

I suppose that at the heart of this is the entire issue about infrastructure integrity. Aside from all the risks of viruses and worms and similar exploits of vulnerable applications on end systems, there are two major points of vulnerability in today's Internet that are insidious. Even when the end points are functioning perfectly well and are not compromised in any way, if the integrity of routing is compromised, or if the integrity of the DNS is compromised, then users can be lead astray quite unwittingly.

Lies is the DNS can be used to turn the Internet into a hostile place and turn users into unwitting victims. It seems to be bordering on hypocrisy for an ISP to claim that network infrastructure security is important and valuable, while at the same time deploy a name server system that deliberately generates lies in the DNS. Or are we supposed to be able to be able draw a distinction between "good lies" in the DNS, "harmless lies" and "evil lies"?

If the ISP really wants to be treated as a common carrier, and not be held responsible for the actions of users and the content that users may exchange, then it should treat all user traffic with respect, and should refrain from acts of selective falsification of the DNS.

Further comments on NXDOMAIN and Typosquatting

- Paul Vixie on "What the DNS is Not"
  http://queue.acm.org/detail.cfm?id=1647302

- Comment on the typosquatting sctions by Telstra
  http://www.crn.com.au/News/160923,bigpond-redirects-typos-to-unethical-branded-search-page.aspx

- Earlier Washington Post report on typosquatting
  http://blog.washingtonpost.com/securityfix/2008/04/when_monetizing_isp_traffic_go.html

- Wikipedia on Typosquatting
  http://en.wikipedia.org/wiki/Typosquatting

- ICANN on the same topic
  http://www.icann.org/en/topics/new-gtlds/nxdomain-substitution-harms-24nov09-en.pdf

- Google's Public DNS service - an open DNS resolver that does not perform NXDOMAIN substitution
  http://code.google.com/speed/public-dns/

Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre, nor the Internet Society.

About the Author

GEOFF HUSTON is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. he graduated from the Australian National University with a B.Sc, and M.Sc. in Computer Science. He has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector. He is author of a number of Internet-related books, and was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of Trustees of the Internet Society from 1992 until 2001.

http://www.potaroo.net