



An analysis of the NET Top Level Domain

 The article was obtained at the following URL: <http://www.kfwebs.net/articles/article/38>
 The article might be distributed further as long as it is provided as it is, with the credits stated.
 The Article was written and first published by KF Webs, at <http://www.kfwebs.net>
 #####

KF Webs recently got access to the gTLD zone-files, of which we decided to have a closer look on the NET zone-file.
Added: 2006-05-21 17:18:55 - Modified: 2006-05-22 12:30:20 - Level: Beginner

Introduction

KF Webs recently got access to the gTLD zone-files, of which we decided to have a closer look on the NET zone-file. We previously performed [an analysis of the COM zone-file](#). KF Webs performs domain name services through the [interface at passive12.net](#), a fully automated service for easy and quick maintenance and new registration of domain names.

All the work has been performed on a system purchased in 2001. It is a Dual PIII 1000MHz with a gigabyte of RAM and a terabyte of storage. Not the most powerful box in the world, but it gets the job done. We used MySQL 5.0.20 with MyISAM tables to perform the analysis. The zonfile was copied 2006-05-21

What is a zone-file

DNS is the abbreviated for of Domain Name System (DNS). When you visit a website such as KFWEB.S.NET your computer sends a request to translate the domain name, kfwebs.net., from a human readable form into a computer readable form, referred to as an IP-address. In the time of writing the IP address of kfwebs.net looks like 213.161.224.2.

The DNS is hierarchical, both in its form and its function. Considering a domain such as kfwebs.net , the Top Level Domain (TLD) would be NET, while the second level domain name would be KFWEB.S.

An interesting note is that the full domain name is really "KFWEB.S.NET.". Notice the final dot and feel free to try to enter it in your Web User Agent's address-bar. Although this is technically the full name, the final dot is usually omitted, but it shows the hierarchical structure of the domain.

A ZONE-file is used on DNS servers in order to delegate control and to store information about domains. In the COM zone-file there are the two records:

```
SECURE-MY-INTERNET NS NS3.KFWEB.S.NET.
SECURE-MY-INTERNET NS NS4.KFWEB.S.NET.
```



These lines tells where to get more information about the SECURE-MY-INTERNET second level domain, and hence delegates control to the mentioned nameservers.

NS3.KFWEBS.NET. and NS4.KFWEBS.NET. again contains the information required for the computer to turn the human-readable form of the name into a computer readable form using a so-called A record. This zone-file again looks like:

SECURE-MY-INTERNET.COM. A 72.29.83.156

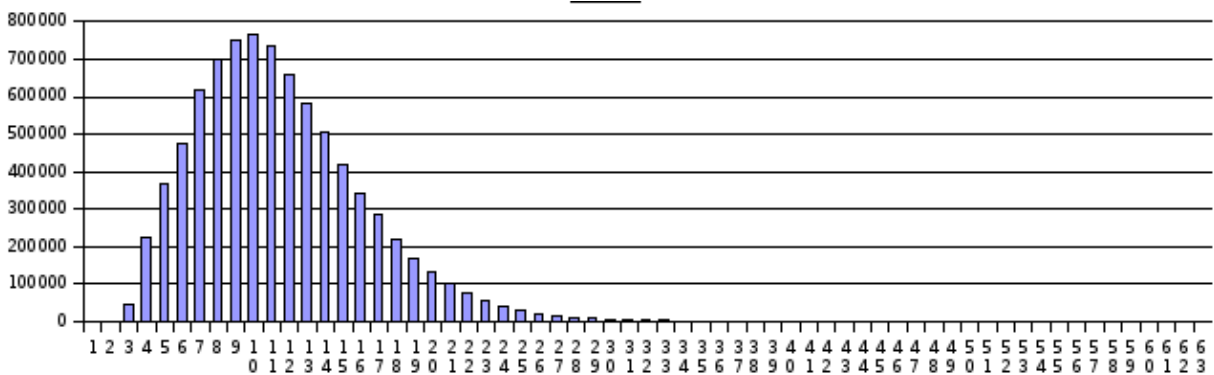
For more in-dept information [this wikipedia entry](#) is recommended

Getting started

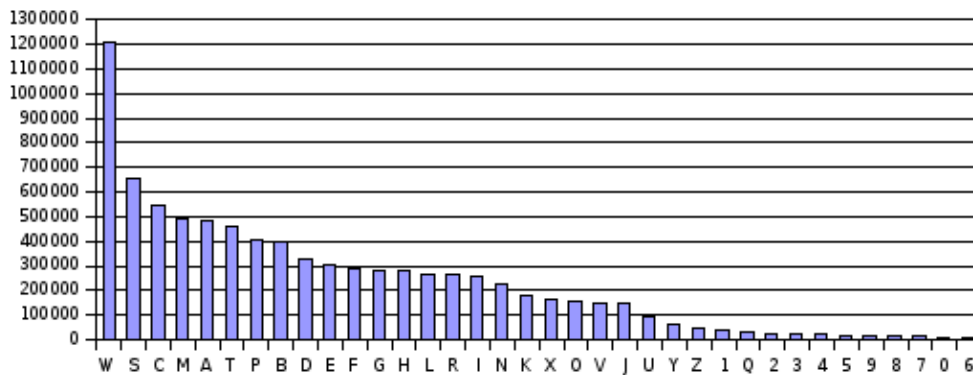
The NET zone-file is 130 MiB compressed, 676 MiB uncompressed. It currently holds 8,365,033 domain names.

Distribution of names

Length of domain names

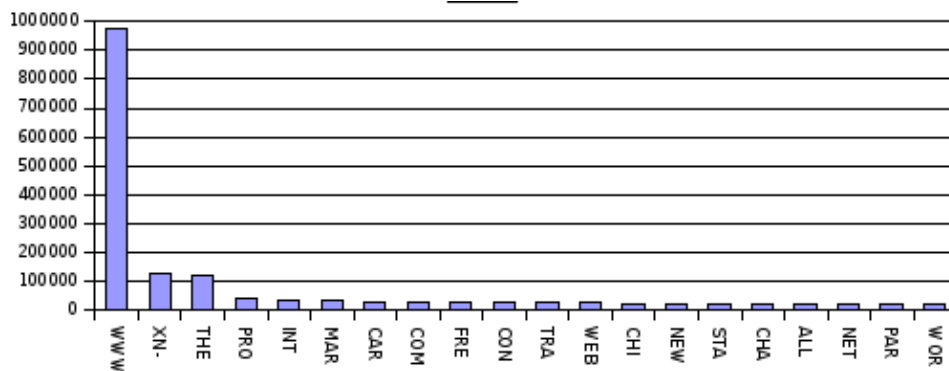


Domain names by first character



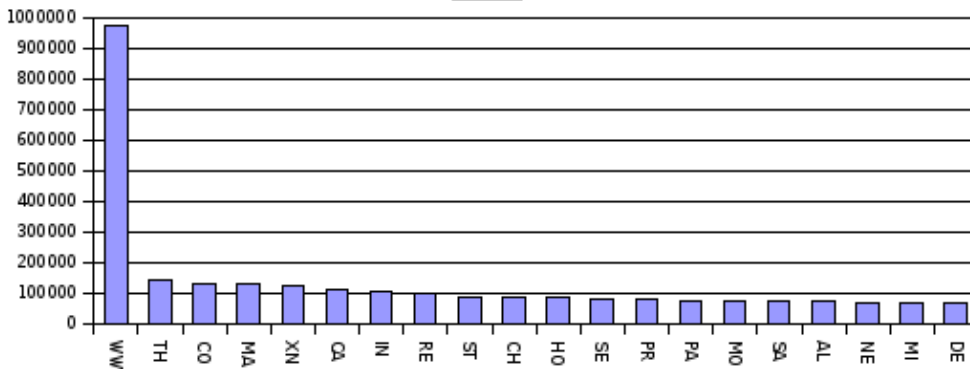
As with the COM top level domain the four characters S,C,M and A are occurring often, the reason for W being the top on the list is detailed in the section about first three characters. M occurs more often than A in the NET zone file, while these letters were shifted in the COM zonefile.

Domain names by first three characters



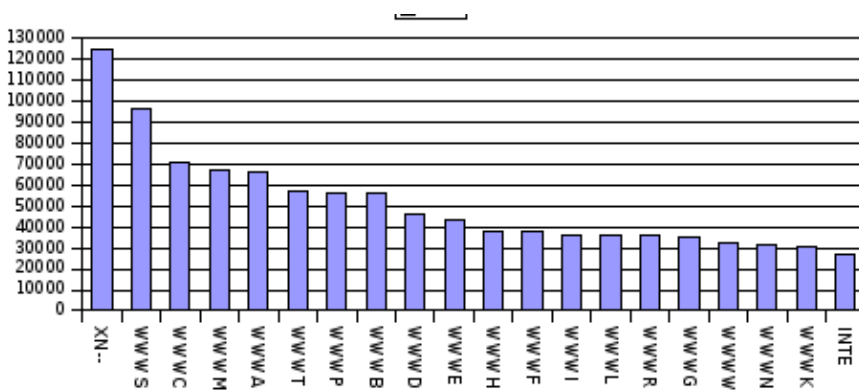
11.7% of the total domain NET domain names in the zone-file begins with the letter combination "www". This explains the high number of occurrences in the first character for W. The following 4 characters follows the distribution found when looking directly at the first character.

Domain names by first two characters



51,405 NET domain names starts with the phrase "go".

Domain names by first four characters



International domain names (xn--)

International domain names are stored as punycode. There are presently 124,942 domain names stored with international domain name support.

Punycode, defined in RFC 3492, is the self-proclaimed "bootstring encoding" of Unicode strings into the limited character set supported by the Domain Name System. The encoding is used as part of IDNA, which is a system enabling the use of internationalized domain names in all languages that are supported by Unicode, where the burden of translation lies entirely with the user application (a web browser for example).

The encoding is applied separately to each component of a domain name which is not represented solely within the ASCII character set, and a reserved prefix 'xn--' is added to the translated Punycode string. For example, bÄ¼cher becomes bcher-kva in Punycode, and therefore the domain name bÄ¼cher.ch would be represented as xn--bcher-kva.ch in IDNA.



Generics

90,546 NET domain names contains the phrase "web" of which 14,234 consists of the phrase "webs", including KFWEBBS.

35,836 domain names contains the phrase "host" and 15,654 NET domain names contains the phrase "hosting". As the NET top level domain is ordinarily used by infrastructure companies, domain registrars and hosters this doesn't strike as odd.

12,685 names contains "domain" and 1836 contains "register", whereby 38,166 contains the phrase "free".

The high occurrence of domain names is as with the COM zone-file probably squatters, hoping that people will mis-type wwwcompany.net and get their advertisement instead of the actual site they intended going to.

Related articles: