المشروع التجريبي لاسماء النطاقات العربية

لـمحة فنية عامة حول الـمشروع

General Technical Overview

Draft – مـسودة
Version 1.0 - 29 August 2005

Prepared by:
Technical Committee

إعداد:
اللجنة الفنية للمشروع

Arabic Domain Names Pilot Project

1. Introduction

Domain names are used widely by Internet users to locate resources on the Internet in a format that is easy to remember and understand. They consist of alphanumeric strings separated by dots, e.g., www.arabic-domains.org. They are written using Roman characters (ASCII) particularly letters, digits, and hyphen. These names, however, are not required by the network software, but are used for human mnemonic convenience. They are used instead of the numerical addresses that are known as Internet protocol (IP) addresses, which are mainly used by machines to route data packets on the Internet. Name resolution is carried out by the Internet domain name system (DNS) in which domain names are mapped to the actual corresponding IP addresses.

The DNS is basically a distributed database of host information that is organized in a hierarchal tree structure, see Figure (1). Theoretically, there is a "root domain" at the top of the domain name tree which is usually left unnamed. Immediately underneath the root come the top-level domains (TLDs). Basically, there are tow types of TLDs. One is the generic TLDs (gTLDs) such as .com, .org, .net, and .edu. The second one is the country code TLDs (ccTLDs) such as .ae (United Arab Emirates), .bh (Bahrain), .ca (Canada), .de (Germany), .eg (Egypt), .jo (Jordan), .kw (Kuwait), .om (Oman), .qa (Qatar), .sa (Saudi Arabia), and .uk (United Kingdom). There are more than 240 ccTLDs following the two-letter country codes defined in the ISO standard number 3166.
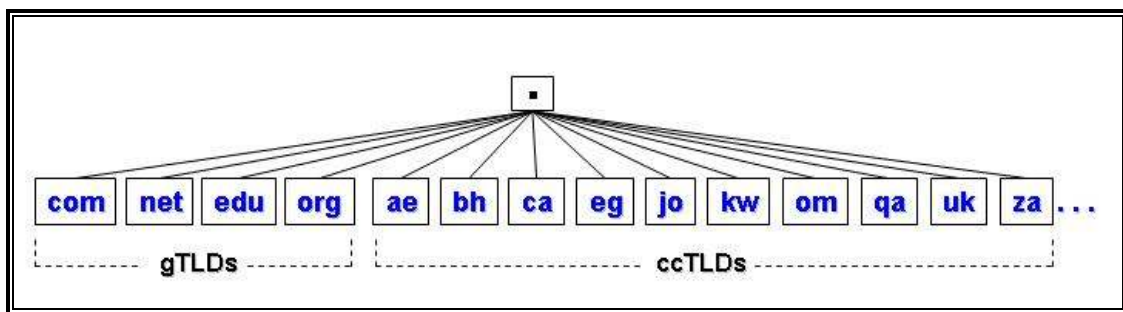
Figure 1: Domain Name Tree Structure

A domain name, whether under a gTLD or ccTLD offers a global presence, which makes sure that the corresponding web site is accessible through the Internet from anywhere. More than 170 millions of such names are estimated to be already stored in the Internet domain name system (DNS).

The Internet has become a global network of most if not all countries of the world with hundreds of millions of users. Currently, it is estimated that more than 60% of the Internet contents are in languages other than English. Also, based on an estimation in year 2003 at least 30% of web users who prefer to do their on-line activities in a language other than English, and that by 2005 only one third of Internet businesses will use English for on-line communication.

Regardless of the worldwide spread of the Internet, the Internet domain name system has not supported other languages to locate resources on the Internet. Users in non-English speaking countries, such as the Arab users, are in disadvantages. Using domain names in a language that is different from the users' native language defeats the main objective of having the domain name in characters rather than just numbers.

The Internet penetration in the Arab world is estimated to be about 1.4% which is indeed very low. One of the obstacles facing the growth of using Internet in the Arab world is the language barrier.

Thus, many countries and nations are encouraging their people to use Internet. Therefore, it is important to make the Internet support the Arabic language not only in web contents but also in their addresses.

Multilingual domain names were first developed in Asia-Pacific countries in 1998, which led later to the creation of a number of non-for profit organizations to supervise and pursuing the deployment of multilingual domain names. Among theses organizations are: the Multilingual Internet Names Consortium (MINC), the Arabic Internet Names Consortium (AINC), the Chinese Domain Name Consortium (CDNC), the International Forum for IT in Tamil (INFITT), and the Japanese Domain Names Association (JDNA). Also, the Internet Corporation for Assigned Names and Numbers (ICANN) established an internal Internationalized Domain Name (IDN) Working Group, and the Internet Engineering Task Force (IETF) created an internationalized DNS group that have been dedicated for exploring the possibility of supporting internationalize Internet.

The IDN group of IETF has issued 3 important RFCs (3490, 3491, 3492.) for Internationalized Domain Names. These new RFC's now make it possible for domain name servers to register non-ASCII domain names and application/client vendors to implement standardized support for handling non-ASCII characters in domain names.

2. How IDN Works?

When a browser sees a host name such as http://www.arabic-domains.org, it passes a request to the DNS resolver service (usually built into an OS), which in turn sends a request to a nearest domain name server to return an IP address that corresponds to the host name. This IP address is then used to connect to the web server in question.

IDN allows host/domain names with non-ASCII characters for user input into a browser's location bar or URL's embedded in web pages. At the network protocol level, there is no change in the restriction that only a subset of ASCII characters be used in URL. If end users input non-ASCII characters as part of a domain name or if a web page contains a link using a non-ASCII domain name, the application must convert such input into a special encoded format using only the usual ASCII subset characters. RFC 3490 (Internationalizing Domain Names in Applications (IDNA)) defines characters used in IDN to be drawn from Unicode Standard 3.2. It also defines how an application should process non-ASCII characters in such a way to conform to existing host name character restrictions.

As an example, an Arabic domain name , "نطاق.السعودية", will look like the following form "xn--mgb5a8an.xn--mgberp4a5d4ar" after converting it from IDN to ASCII format (using "stringprep", "nameprep" and "punycode" operations).

3. Supporting Arabic Domain Names

Supporting the Arabic language in domain names calls for investigating and addressing a number of issues to produce a set of standards that are acceptable by the Internet community in large. These standards should cover several aspects of supporting Arabic domain names in deferent levels, such as:

1. Linguistic issues and the accepted Arabic character set.
2. The Arabic domain name tree structure, i.e., Arabic gTLDs and ccTLDs.
3. Technical solutions to arabize the domain name system
4. The administrative and organizational issues of Arabic root servers.

The 1st and 2nd points have been addressed by the Internet Draft that has been produced during 2003-2004 by the Arabic Domain Names Task Force (ADN-TF) under the auspices of ESCWA, and has since undergone several enhancements and updates. The last of which was performed after thorough review by the first meeting of the LAS Arab Working Group on ADN, held in Damascus on the 31/1-2/2/2005.

The 3rd point is partially addresses by the IETF 3 RFC's (3490, 3491, 3492).

Since there is no indication that ICANN is going to support full IDN in the near future, the Arabic Team agreed on implementing a pilot project among the Arab ccTLD's to address the issue of Arabic root servers (i.e., point 4). This will allow the Arab's ccTLD to early experience the use of Arabic domain names, identify our needs, locate possible problems, and develop tools. In addition, we can use the outcome of the pilot project as an argument with international bodies, e.g., ICANN, to speed up the international recognition of supporting Arabic language in domain names based on our needs.

4. Arabic ccTLD

The new version of the Internet draft (http://www.escwa.org.lb/divisions/ictd/forum/default.asp) suggests the following Arabic ccTLD's for the Arab countries and this project will stick to them:

Country Official Names	Country Code	Short Name	Puny-Code
Hashemite Kingdom of Jordan	jo	الأردن	xn--igbhzh7gpa
United Arab Emirates	ae	الإمارات	xn--kgbdbap4b0ij
Kingdom of Bahrain	bh	البحرين	xn--mgbcpq6gpa1a
Republic of Tunisia	tn	تونس	xn--pgbs0dh
People's Democratic Republic of Algeria	dz	الجزائر	xn--lgbbat1ad8j
Federal and Islamic Republic of Comoros	km	القمر	xn--mgbu4chg
Republic of Djibouti	dj	جيبوتي	xn--ngbee7iid
Kingdom of Saudi Arabia	sa	السعودية	xn--mgberp4a5d4ar
Democratic Republic of Sudan	sd	السودان	xn--mgbaxp8fpl
Syria Arab Republic	sy	سورية	xn--ogbpf8fl
Somalia Democratic Republic	so	الصومال	xn--mgba5b5cceu
Republic of Iraq	iq	العراق	xn--mgba3a5azci

Sultanate of Oman	om	عمان	xn--mgb9awbf
Palestine	ps	فلسطين	xn--ygbi2ammx
State of Qatar	qa	قطر	xn--wgbl6a
Stat of Kuwait	kw	الكويت	xn--mgbg8edvm
Lebanese Republic	lb	لبنان	xn--mgbb7fjb
Socialist People's Libyan Arab Jamahiriya	ly	ليبيا	xn--mgbb7fyab
Arab Republic of Egypt	eg	مصر	xn--wgbh1c
Kingdom of Morocco	ma	المغرب	xn--mgbc0a9azcg
Islamic Republic of Mauritania	mr	موريتانيا	xn--mgbah1a3hjkrd
Yemen Arab Republic	ye	اليمن	xn--mgb2ddes

5. Arabic ccTLD Servers

Each Arab country should setup and runs its own Arabic ccTLD server(s) for the chosen Arabic domain name that represent their country name. Although it is recommended that, each participant (ccTLD administrator) can register some test Arabic domain names such as:

| Example	Arabic test Domain name	Purpose
موقع.السعودية	موقع.<اسم الدولة>	A site for testing Arabic domain name in each ccTLD
دليل.الإمارات	دليل.<اسم الدولة>	A site that list all the registered Arabic test domain names in each ccTLD.
مركز-التسجيل.قطر	مركز-التسجيل.<اسم الدولة>	A site for registering Arabic test domain names in each ccTLD

In addition, each participant can act as secondary for other Arabic ccTLD, which will add more redundancy and localization for the DNS queries.

Note: There is a technical document on how to setup Arabic ccTLD server (see "How to setup Arabic ccTLD server ")

6. Arabic Root Servers

If each Arab country runs its own Arabic ccTLD server, users in other Arab countries would not be able to reach other Arab's Arabic domain names. Therefore, we need a mechanism to address the problem.

The technical committee agreed on using the current Arabic root servers (which are located in AE and SA), which will be responsible for the root zone files for all the Arabic ccTLD (22 countries e.g. (السعودية, قطر، البحرين، الإمارات، عمان، مصر،سوريا). This scenario makes the modification for any Arabic ccTLD zone file very easy and centralized between the root servers. Also this scenario will add one layer above the Arabic ccTLD servers and now any entity that want to resolve Arabic domain name needs to contact them first in order to reach the ccTLD (see figure 2 for more details).
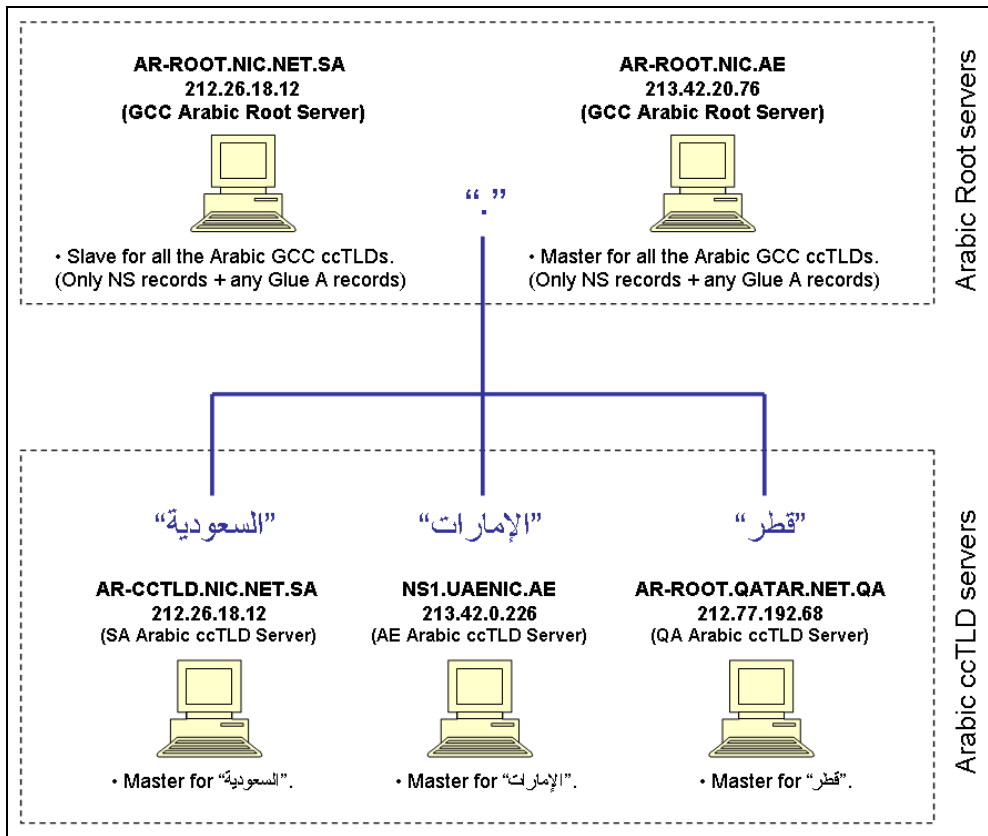
Figure 2: The Arabic Root and ccTLD servers map

Note: There is a technical document on how to setup Arabic root server (see "How to setup Arabic root server ").

7. How to resolve Arabic Domain Names

Since ICANN has not yet supported full Arabic ccTLD any solution for supporting Arabic domain names would be only accessible locally. Therefore, there should be a hacking solution until Arabic domain names is world-wide supported. Hence, any entity needs to reach Arabic Domain names must therefore do some changes in the configuration files of their resolving/caching name servers.

The technical committee has agreed on using the "stub zone" configuration option for any entity that needs to resolve Arabic domain names. A "stub" zone is like a "slave" zone, except that it replicates only the NS records of a master zone instead of the entire zone.

Therefore, if any entity wants to resolve Arabic domain names they should configure all the Arabic ccTLD as "stub" zones in the configuration of their name server and point them to the Arabic root servers.

Here is an example on how to configure stub zone using bind version 8 or higher in the file "named.conf" for all the Arabic ccTLDs:

```
zone "xn--mgberp4a5d4ar" {
        type stub;
        file "idn.sa.zone" ;
```

```
        masters { 213.42.20.76; 212.26.18.12; };
};

zone "xn--kgbdbap4b0ij" {
        type stub;
        file "idn.ae.zone" ;
        masters { 213.42.20.76; 212.26.18.12; };
};
.
.
```

File: named.conf

Hint:

The above stub zones configuration will be kept in a separate file that can be distributed by the admin of Arabic root servers. Then it can be referenced (using the "include" statement) within the bind configuration file "named.conf" in any entity name server that want to resolve Arabic domain names.

This file should not be changed frequently to reduce the participants' overhead on downloading the file each time it get changed. Therefore, this file should contain all the Arab countries in it, even if they are not yet part of the project. This file will be provided to any entity that wants to resolve Arabic domain names (see "Arabic_ccTLD_stub_zone_files.inc").
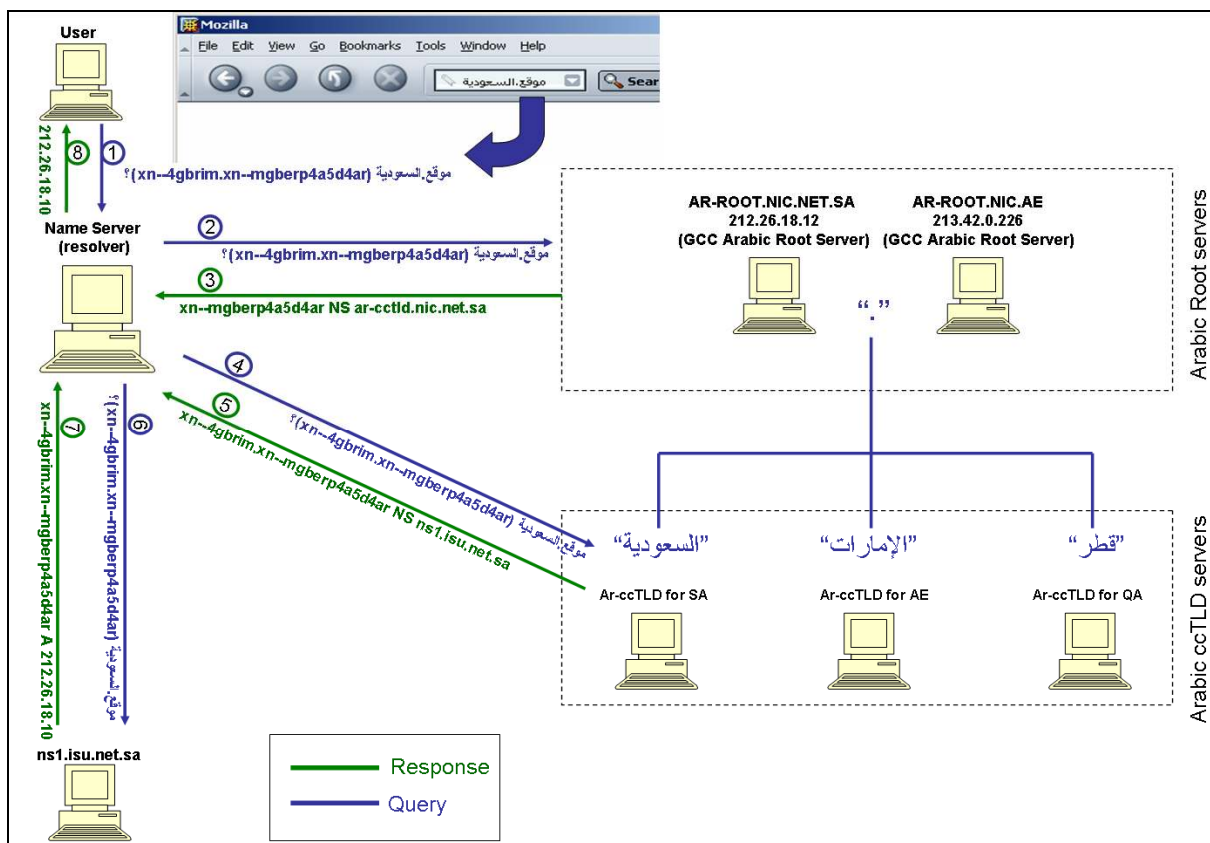

Figure 3: How to resolve Arabic domain names

Note: There is a technical document on how to setup your name server to resolve Arabic domain names (see "How to Resolve Arabic Domain Names") and another one that describes the client requirements to reach Arabic domain names (see "Requirements for Resolving Arabic Domains ").

More information

- The Project website:
    - http://www.arabic-domains.org
- IDNs RFCs:
    - http://www.ietf.org/rfc/rfc3490.txt
    - http://www.ietf.org/rfc/rfc3491.txt
    - http://www.ietf.org/rfc/rfc3492.txt
- News on IDN:
    - http://idn.isc.org/
- GNU Libidn:
    - http://www.gnu.org/software/libidn/