

1. Executive Summary

The IDN Implementation Working Team studied the management of variant labels as defined in Section 3.1 below, and the requirement of a three character minimum length for gTLD labels. The following recommendations are made on the basis of that work:

Variant Management

- A TLD applicant may request variant forms of a base label and must classify each variant as 'desired' or 'undesired';
- Desired variants can be allocated to the applicant;
- Desired variants can be delegated to the applicant if the applicant agrees to conform to the Rules outlined under Section 3.3;
- Undesired variants will be reserved and neither allocated nor delegated.

Three-character Restriction in gTLDs

- Two-character gTLD strings can be allowed subject to conformance to the Rules outlined under Section 4.1.1;
- The team does not recommend the banning of the one-character gTLDs. This issue is a matter of policy and must therefore be addressed by bodies such as the ccNSO and GNSO;
- See Rules in Section 4.1.2.

2. History

The IDN-Implementation working team was established after discussions during the ICANN meetings in Mexico City (1-6 Mar, 2009) and Sydney (21-26 Jun, 2009). These discussions were related to the implementation of new gTLDs and IDN ccTLDs (Fast Track Process), and focused on two topics:

1. Variant management in TLD strings
2. Three-character restriction in gTLD strings

The scope of the working team was to study these topics and to propose potential solutions to the problems under each topic, in order to facilitate further discussion and consultation with the community. The working team was asked to report its recommendations to the ICANN Board and community in time for the ICANN meeting in Seoul (25-30 Oct, 2009).

The IDN-Implementation working team was composed of linguistic and technical experts from different language communities, and co-chaired by two ICANN Board Directors who are experts in the fields of DNS and IDNs.

Team co-chairs

Harald Alvestrand (IDNA author and ICANN Board)

Ram Mohan (.info and ICANN Board)

Team members

Raed Al-Fayez (SaudiNIC)

Sarmad Hussain (member of technical committee of Ministry of IT, Pakistan)

Cary Karp (IDNA author and .museum)

Guanghao Li (CNNIC)

Siavash Shahshahani (Advisor, IRNIC)

Yang Yu (CONAC, China)

ICANN staff support

Kurt Pritz

Tina Dam

Baher Esmat

Working method

The team held weekly meetings in telephonic conference from 2 Sep to 21 Oct 2009. The discussions at these meetings were noted and summarized by ICANN staff, reviewed by the co-chairs, and distributed to and approved by all members. All summaries are included in Annex-2 of this report. Eight such meetings were held.

The team acknowledged that its task was not to develop policies, but rather to consider the two problems of variant management at the top-level, and the required three character minimum length for gTLD strings, and to provide recommendations for further consultation with community.

3. Variant Management

Language communities that use variant characters are affected by decisions about how variants are managed and implemented in new TLDs.

This is of concern in both IDN gTLD and IDN ccTLD implementations. Variant TLDs will be identified in direct reference to the IDN Tables that are required attachments to a TLD request/application.

The team defined various categories or types of variants before discussing technical solutions to deal with them at the top-level of the domain name system. The intention was to devise a clear definition of the different types of variants, and propose for short-term solutions on that basis, particularly for the IDN Fast Track process, with intent to avoid long-term problems.

Variant characters (as defined in RFC 3743) occur where a single conceptual

character has two or more graphic representations, which may or may be visually similar. Variant TLDs contain one or more characters that have such variants. Allowing variant TLDs may result in user confusion, while excluding them may disenfranchise communities that use the characters in the excluded TLD strings.

Note: See Example 1 in Annex-2.

3.1 Definitions

Base Label: A Base IDN TLD Label is the primary form used in an IDN TLD request, in a specified language.

Variant Label: A Variant Label is resulting from the substitution of one or more characters in a Base IDN TLD Label with variant characters.

Variant Types: Variants may be classified in two types to differentiate between those that are graphemically identical (visually indistinguishable), and variants that are of dissimilar appearance but have the same meaning (semantically equivalent):

- Type 1 Variants: Variants that are visually identical to the base label.
- Type 2 Variants: Variants that are not visually identical but are considered equivalent to the base label in some orthographic sense.

TLD Allocation: A Base IDN TLD Label and its Desired Variant Label(s) must be indicated by the applicant. Once a Base IDN TLD Label is assigned to an applicant, the Desired Variant Label(s) will be reserved solely for the use of the applicant. This process is called “Allocation”.

TLD Delegation: The addition of the Allocated TLD Label (Base IDN TLD Label or Desired Variant Label(s)) to the DNS root zone.

An applicant who requests a Base Label and either Type 1 or Type 2 Variant Labels must indicate each variant as either a Desired Variant or an Undesired Variant. Desired variants are intended for use interchangeably with the base IDN TLD label. All other variants are not desired.

3.2 Managing variants at the top-level

Based on the above definitions, there was agreement on the following principles:

1. Desired variants must be indicated by the requester (applicant);
2. Desired variants will be either allocated or delegated to the applicant;
3. Undesired variants would neither be allocated nor delegated, and will be blocked.

The team members had dissimilar views with regard to desired variants and whether or not they should be delegated. Some were of the view that desired variants should be allocated for the applicant but placed on hold. Others thought

that allocating variants without delegating them would cause confusion to end users who will fail to reach certain TLDs because their keyboards do not support all the characters in the TLD label.

In discussing solutions for variant management, some members confirmed that the delegation of variants has proved to be successful on the second level and recommended adoption of a similar practice at the top-level. On the other side, it was argued that while enforced variant delegation has worked with some registries on the second level, the behavior of applications is beyond the control of the registries, and there is no mechanism by which the registry can enforce behavior at the end user's side. For example, major server-side applications such as mail servers and Web servers receive and process domain names exactly as they are provided by the user, entirely independently of what is in the DNS or the way in which the client-side application invokes the DNS.

The team recommends that DNAME as a mechanism for variant delegation be systematically tested to formulate an appropriate solution for ensuring consistency when deploying variants in the DNS.

The team concluded that if desired variants are to be delegated, certain conditions must be fulfilled (as detailed in the proposal below). These conditions would be included in an arrangement between ICANN and the TLD registry (i.e., gTLD contract, IDN ccTLD agreement, exchange of letters, application form terms and conditions).

3.3 Proposal for management at the top-level of both Type 1 and Type 2 Variants

1. In addition to the Base IDN TLD Label, a TLD applicant may identify a number of Variant Labels.
2. The applicant must classify each Variant Labels as 'Desired' or 'undesired'.
3. Desired Variant Labels will be allocated, and may also be delegated, to the applicant.
4. Undesired Variant Labels will neither be allocated nor delegated, but will be blocked.
5. Desired Variant Labels may be delegated to the (same) applicant provided the following conditions are met:
 - A. The applicant requests the Variant Label.
 - B. The domains directly below the variant TLD are to be mapped exactly to the corresponding domains under the Base TLD.

- C. The applicant must demonstrate how this mapping is to be achieved, preferably by presenting earlier work on a testbed.
- D. The applicant must provide information on how it intends to inform its registrars/resellers/registrants about the correct set-up of name servers, and all other application servers that process domain names directly, so that end-user confusion can ideally be prevented, and at least held to an acceptable minimum. (This is discussed in further detail in Annex 3.)
- E. ICANN shall conduct periodic checks to ensure that the said mapping is in place.
- F. The delegation document for the Base IDN TLD Label (whatever form it takes: contract, MoU, accountability framework, ...) will commit the registry to the points above.
6. It is understood that that technical or other considerations may necessitate changes in the items under 5 above.

4. Three-Character Restriction in gTLD Strings

In discussing the issue of the three-character restriction in gTLD strings, team members recognized that in some languages many meaningful words are composed of one or two characters. The desire to have gTLD strings with fewer than three characters appeared legitimate, and addresses the needs of the affected language communities.

The team agreed on the general principle of relaxing the three-character restriction in gTLD strings. The team also acknowledged the statement made by the [GAC during the Sydney meeting](#):

“The GAC expresses concern about the three-character requirement for the new IDN gTLDs, for example in Chinese, Japanese, Indian and Korean scripts, where many one or two characters have meaning. It should be taken into account developing the next version of DAG.”

Note: See Example 2 in Annex-2.

The team agreed to devise a set of rules that could be used as a guideline for further consultation with community on the possible implementation of one-character and two-character TLDs. In its deliberations about those rules, team members recognized the following:

- **The three-character restriction is only related to the new gTLD process.**

In the IDN ccTLD Fast Track Process, strings with two or more characters have been allowed, and strings would only be ruled out on the basis of confusability. This means that the distinction between the ccTLD space and gTLD space is blurring, and the desire to keep the distinction therefore does not seem to be a valid reason for disallowing strings that are less than three-characters in the new gTLD process.

- **Measuring the string's length.** It has been suggested that the number of characters in a string should be determined by the number of keystrokes. The team found this not to be the case because:
 - A technological change in keyboards could also change the number of keystrokes required.
 - The number of keystrokes may vary for the same input, as Unicode is redundant and in some instances, the same "character" can be entered with a single keystroke (pre-composed form) or with two keystrokes (combining form). Many of these equivalent characters (at least in case of Arabic script) are not automatically dealt with through the Unicode normalization process.

As a result, the team suggests that ICANN adopt a more strictly defined measure of string length. The term "glyph" was considered and rejected as problematic. The team suggests using the term "grapheme cluster" where a combining sequence of a base character and combining mark(s) appears to be a single character, using the definition of an "extended grapheme cluster" from section 3 of Unicode Standard Annex #29.

- **Protecting two-character ASCII codes is essential.** All possible two-character ASCII codes should be set aside for use by the ISO 3166 Maintenance Agency, and neither those codes nor any other codes that are visually confusable with them should be allowed in the gTLD space. Two-character codes that are not visually confusable with the ASCII 3166 codes should not be restricted. Rules should be formulated to ensure consistency with established rules and procedures, to protect the two-letter ASCII 3166 codes. ICANN must communicate with the TC46/ISO 3166 Maintenance Agency to achieve better and mutual understanding of how each organization operates in areas pertaining to the domain name space.
- **Exercise caution with one-character strings.** There seem to be no technical reasons for restricting one-character IDN TLD labels. What works for one language or writing system may not work for the other, hence limiting IDN TLD labels to a minimum of two-characters eliminates a large number of meaningful words in some languages such as Chinese in which almost every single character is a meaningful word.

However, there are significant economic considerations associated with the introduction of one-character TLDs. There are also policy aspects involved, so the decision to allow them must be made by policy bodies such as ccNSO and GNSO.

4.1 Rules for relaxing the three-character restriction in IDN TLD Strings

4.1.1 Two-character rules in IDN TLD labels

Rule 1: No visual confusion is allowed with two-character ASCII strings.

1a. All possible two-character ASCII codes are set aside for use as ccTLDs by territories identified by ISO 3166 -1 codes.

1b. Both current and potential future codes need to be protected against visually similar strings.

Rule 2: No visual confusion is allowed with existing or applied-for TLDs.

2a. This rule is the same for all TLD applications.

Rule 3: No TLD may be labeled with a string that is a meaningful identifier of a country or territory listed in ISO 3166-1 without the express consent of the government or properly constituted authority of the designated country or territory

3a. This is the same rule for all TLD applications.

Rule 4: Two-character strings that are visually confusable with one-character strings, for example: a string composed of one character + one or more combining mark, should follow the rules for one-character IDN TLD labels.

Note: A uniform treatment of all ASCII strings is easy to understand and easy to practice.

4.1.2 One-character rules in IDN TLD labels

Rule 1: One-character IDN TLD labels should be restricted pending further analysis.

1a. Analysis is required of the potential impact of relaxing the restriction on the allocation of one-character IDN TLD labels.

1b. Until such analysis is conducted, the impact understood, and appropriate reviews completed, one-character IDN TLD labels should be restricted.

Rule 2: The preceding rule should apply to labels containing two or more characters that are visually confusable with a single-character (for example: a label composed of one character + one or more combining marks).

Recommendations and the Way Forward

Recommendation 1

- 1.1 Desired Variant Labels must be indicated by the applicant;
- 1.2 Desired Variant Labels will be allocated, and may also be delegated, to the applicant;
- 1.3 The delegation of Desired Variant Labels, where otherwise appropriate, is contingent upon the applicant agreeing to conform to the Rules outlined in Section 3.3;
- 1.4 Undesired Variant Labels will be neither allocated nor delegated, and will be blocked.

Recommendation 2

Two-character gTLD labels can be allowed subject to conformance with the Rules outlined in Section 4.1.1.

Recommendation 3

- 3.1 The team does not recommend the banning of one-character gTLDs.
- 3.2 The team recommends that further ramifications of this issue be addressed by policy bodies such as the ccNSO and GNSO.

Recommendation 4

The team cannot make a recommendation about the adoption of DNAME at the top-level. Further technical study is needed, and ICANN should fund this if necessary.

Recommendation 5

A process is needed for managing the delegation of desired variants, and the allocation and delegation of additional desired variants that may be identified and requested at a later time. The team regards this work as external to its own mandate.

Annex-1: Background

Variant Management

On 18 Feb 2009, ICANN published a draft paper titled "[IDN ccTLD Fast Track Program Proposed Implementation Details Regarding: Development and Use of IDN Tables and Character Variants for Second and Top Level Strings](#)", in which it explained in details what an IDN Table is and purpose it is needed for, and proposed recommendations for the process of developing IDN Tables and in particular definition and use of variant characters, in both second and top-level strings.

There were discussions during the ICANN meeting in Mexico City, (1-6 Mar, 2009) on that paper, and public comments were received with regard to the proposed definition of variants and the way they could be managed at the top-level.

Based on that input, ICANN drafted a [second version](#) of the paper and posted it on 29 May 2009 for public comments, and for further discussion with community during the ICANN meeting in Sydney (21-26 Jun, 2009). The paper stated the following in relation to the issue of variant management:

There will be situations in which an IDN ccTLD requester may have reasonable grounds for wishing to have more than one label for the requested domain, due to use of variant characters and resulting in "variant strings".

Previously ICANN proposed that variant strings could be either allocated or reserved for registration. In order to be allocated it was proposed that the string (i) fulfilled all the string requirements in the Fast Track Process and (ii) that while the strings was inserted as a separate delegation in the root, they needed to be treated as aliased in order to avoid confusion. All other variant string was proposed to be blocked for allocated as they otherwise would introduce a confusable situation for users.

For the purpose of this paper, aliasing means that, say there are two variant string ".variant" and ".variànt". Under aliasing, if a registrant registers 'example.variant' then 'example.variànt' would also resolve to the same address, i.e. the two TLDs are considered the same or replaceable.

However, comments on that proposal have indicated that allocation of variant strings would cause technical stability problems for the name space.

The resource record DNAME was originally expected to enable the aliasing functionality in the root zone, as DNAME is being used for this purpose at the second level under various TLDs, however, analysis to date shows that DNAME does not function at the root level. The proposal made by ICANN in the previous version of this

paper (as mentioned above) was to delegate the variant strings separately and then require that the TLD manager ensures duplication of the multiple zones. However, the technical complication with this proposal is that while a registry manager can duplicate zone immediately under a TLD, this will not function at lower levels. This would put a requirement upon the registrants (and their sub-domains) to duplicate zone contents at lower levels as well. There is no mechanism to ensuring that this takes place. Unless a technically sound solution is demonstrated to successfully demonstrate aliasing or duplication functionality the variant strings cannot be allocated at this time.

ICANN understands the need expressed in the community for enabling allocation of variant strings, in particular for locations where some users will key in one string and other users will key in the variant string when accessing for example a website. ICANN urges the community to continue to discuss and develop a technical solution that will enable the allocation of variant strings in the root zone in a stable manner. Until then IDN ccTLD Fast Track requesters will need to select one string per script or language only or alternatively wait until a technical solution has been found.

In order to reserve the possibility of allocating variant strings to the appropriate entities, ICANN will ensure that all variant strings are reserved or blocked for allocation for now.

Blocked strings will be considered as “existing strings” when incoming requests are checked for conflicts with existing TLDs. Therefore, any later request for the same string will be denied.

Three-Character Restriction in gTLDs Strings

In the process of introducing new gTLDs, the [Applicant Guidebook \(version 2.0\)](#) states that “Applied-for strings must be composed of three or more visually distinct letters or characters in the script, as appropriate”. Several comments received express the need for allowing gTLDs consisting of less than three characters.

On 30 May 2009, ICANN posted an [Explanatory Memoranda](#) on discussions about three-character string requirements, in which it summarized the different views that are with and against relaxing the requirement of a minimum of three characters in gTLD strings. In this paper ICANN stated that “In order for a solution to be acceptable it will need to set clearly understandable guidelines for when strings less than 3 characters long will be allowed as TLDs. The solution must be technical feasible and address valid concerns such as potential for user confusion”.

The feedback ICANN had received before and during the Sydney meeting on the two topics of variant management and the three-character restriction demonstrated the need to create this working team to address the two issues and suggest guidelines that could serve as base for feasible solutions to be implemented.

Annex-2: Examples

Example 1: Need for delegating the Desired Variant Labels to the same applicant of the Base TLD Label:

1. Without delegation of Variant Labels users will be confused by having their domain name not resolvable from many parts around the world (ccTLD or gTLD).

Example: ایران (Iran with Arabic YEH) & ایران (Iran with Farsi YEH)

2. If the reason for not delegating Variant Labels is to avoid the risk of the ISP or the hosting provider messing up the connection between the same domain name under the Base IDN Label and its Variant Label, the same risk still exists if the registry decided to bundle variants under second-level and third-level domain names.

Example: كل.مصر (Eat.Egypt with Arabic KAF) & كل.مصر (Eat.Egypt with Arabic KEHEH)

The IETF IDNA-WG suggests that TLD registries should handle variants, a case can be made to do the same at the TLD level.

Awareness is required at various levels (ICANN, Registry, Registrar, Registrant, Users) in any case and this will make it happen.

Example 2: Need for 1-character TLDs

There are about 70,000 CJK characters in Unicode 5.0 whereby each is assigned a unique code point. All of these CJK characters, with some exceptions of components of ideographs like strokes and radicals, have a unique meaning.

The single character Chinese word is the most generic word since it is represent a concept without further elaboration. For example, the generic word for shop in Chinese is “店”. One may add other Chinese character to the word to further elaborate or sub-categorize it:

药店 - Medicine Shop

商店 - Department Store

家具店 - Furniture Shop

便利店 - Convenient Store

宠物店 - Pet Shop

礼品店 - Gift Shop

专卖店 - Brand Shop
etc.

Therefore, we could have one single TLD like “店” and sub-categorize the rest as 2LD, or we could have all the possible categories in the top level, hundreds if not thousands of different variations in the top level.

Annex-3

This Appendix reviews basic concepts that should be included in the documentation provided by the operator of an IDN TLD that has variant labels, to the administrators of autonomous subdomains. To ease its inclusion in stand-alone texts, some material in the main body of the document is repeated below.

The term “variants” refers to alternate labels that designate the same domain but with differences in their encoding and graphic representation that may or may not be apparent under common display conditions. Variants accommodate differences in the way a script is used by the language communities that share it. They are also used by language communities in which alternate forms of some elements of its script are prevalent, and one segment of the community cannot be expected to adopt the practice of another.

A label provides no indication, in itself, about the possible existence of variant forms or what they might be. It is therefore essential that a registry supporting variants provides information to the administrators of all delegated zones about those variants, particularly in cases where the alternate forms are displayed in an identical or confusingly similar manner. (Although this is an intrinsically subjective matter, there are situations where differences will not be perceptible even to experienced users of a script, as well as those where differences can be obvious to a neophyte.) This information should include guidelines about the configuration of application servers that operate directly on domain names.

The intention with variants is for them to share the same target. The DNS can provide part of that linkage by returning the same IP address for all related variants. If the labels "tld1" and "tld2" are equivalent to each other and they both point to, say, 123.456.789.012, the potential of the DNS for indicating their equality will be exhausted. In addition, however, important services that run on Internet hosts – such as HTTP (the Web) and SMTP (e-mail) – process domain names precisely as they are typed by the end users of those services. The servers must also be configured to handle variant labels in those names.

A Web server, for example, will typically use a single IP address for a large number of "virtual hosts". It determines which of them the user intends to access by looking at the domain name that was entered into the address line of the user's browser. What the server then does is determined by its own configuration, independently of the DNS, and is therefore beyond the control of the TLD registry. Corresponding conditions pertain to e-mail. The destination domain name in an e-mail address is used to determine the IP address of the machine that receives mail for that domain (as indicated in a zone file by MX – mail transfer – records). The e-mail is forwarded

to that host without changing the domain name in the address as provided by the user. The mail server can be configured to perform numerous operations on the basis of that address but, again, this is beyond the scope of the DNS and control through the registry.

The coherent and consistent behavior of a domain with variant labels requires every single server that examines and operates on domain names, in each and every sub-domain, to be configured to treat all variants identically. This, in turn, depends on the operators of those servers recognizing the necessary additional detail, which fortunately can be implemented with familiar configuration procedures.

This will be illustrated with configuration files from the server for ICANN's IDNwiki at <http://idn.icann.org/>, and the e-mail server on the same host. The crucial underlying concept is that, despite any similarities in the way variant labels are displayed, from the perspective of the DNS, each is a separate and distinct entity. IDN labels are stored in the DNS in A-label form (“xn--encodedstuff”), which is also used in HTTP and SMTP configuration files.

Eleven IDN TLDs appear in the IDNwiki. One of them includes a variant second-level label, and two have variants on the fourth level. The section of the configuration file (for an Apache server) that locks the second- and fourth-level variants to the same document hierarchy is:

```
<VirtualHost idn.icann.org:80>
    DocumentRoot /var/www/wiki/yi/
    ServerName xn--fdbk5d8ap9b8a8d.xn--deba0ad
    ServerAlias xn--5dbqaap0c8a.xn--deba0ad
    ServerAlias xn--cdb6dqac0h.idn.icann.org
    ServerAlias xn--7dblab4f.idn.icann.org
</VirtualHost>
```

The other domain including a fourth-level variant is similarly supported by:

```
<VirtualHost idn.icann.org:80>
    DocumentRoot /var/www/wiki/ur/
    ServerName xn--hhbbbh02d.xn--hgbk6aj7f53bba
    ServerAlias xn--mgbqf7g.idn.icann.org
    ServerAlias xn--mgbqf7g6a.idn.icann.org
</VirtualHost>
```

The distinction between the ServerName and the ServerAlias corresponds directly to the distinction made elsewhere between the base form of label that has variants, and the variants themselves.

It is obvious that variants may also be deliberately associated with separate document hierarchies, for example, when the intention is to separate material

presented in alternative orthographies. This approach is also illustrated in the IDNwiki, where Simplified and Traditional Chinese are supported as separate entities, with parallel facets of the wiki maintained for each. This harnesses, rather than obscures, the fact that every variant is a unique entity in the DNS. (The extent to which the concept of variants can be applied to segregated document hierarchies is debatable, but the present text is focused on the narrower sense of the term.)

Until fully internationalized e-mail addresses are available, the configuration of an SMTP server to process IDN-labeled domains reduces to little more than providing a list of all the A-labels it needs to recognize. The host for the IDNwiki supports e-mail testing as described at [http://idn.icann.org/E-mail test](http://idn.icann.org/E-mail_test), using the following configuration (for SENDMAIL):

```
# local-host-names - include all aliases for your machine here.
idn.icann.org
xn--mgbh0fb.xn--kgbechtv
xn--hxajbheg2az3al.xn--jxalpdlp
xn--p1b6ci4b4b3a.xn--11b5bs3a9aj6g
xn--r8jz45g.xn--zckzah
xn--9n2bp8q.xn--9t4b11yi5a
xn--mgbh0fb.xn--hgbk6aj7f53bba
xn--e1afmkfd.xn--80akhbyknj4f
xn--zkc6cc5bi7f6e.xn--hlcj6aya9esc7a
xn--fdbk5d8ap9b8a8d.xn--deba0ad
xn--5dbqaap0c8a.xn--deba0ad
xn--fsqu00a.xn--0zwm56d
xn--fsqu00a.xn--g6w251d
```

Here again, the critical point is the explicit listing of the second-level variant for the TLD that includes one. The MX records in the DNS zone file must also cover all variants that are intended to receive e-mail.

In summary, as long as it is realized that client software uses the DNS to determine the IP addresses of the hosts to which the Web and e-mail service requests are sent, but the designated hosts process those requests on the basis of the domain names as they are supplied by the user, the configuration of the servers should not be problematic. It must be understood, however, that there is only a marginal likelihood of every requisite server in every subdomain being correctly configured. The potential confusion that can result from the use of variants is therefore intrinsically greater in a TLD that includes them, than it would be otherwise. The introduction of TLD variants into the root zone of the DNS extends this uncertainty into the entire namespace. It is therefore vitally important for a TLD registry participating in this process to implement special measures to reduce the potential for negative side effects.

Annex-4: Minutes of Meetings

IDN Implementation Support Team First Meeting: Conference Call Wednesday 2 Sep, 2009

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)
Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Summary

General introduction:

- Brief introduction by each member of the team, his/her affiliation, and what he/she wants to achieve in here.
- Purpose of this team is not to develop policies, it is rather to look into two specific problems and try to propose solutions for further discussion with community:
 1. Managing variants at the top-level
 2. Explore lifting the current restriction of the 3-character gTLD string
- This team is dealing with problems that are part of many language groups, so trying to find a universal solution may not be practical - the team should try to define a couple of generic rules or guidelines, agree to them and see how to apply them to solve the problems.

Discussion on variant management:

- In defining variants, it is important to realize that there are several combinations of the following two different cases of variants: 1) characters that look the same but have different code points (Arabic), 2) characters that do not look the same but they have the same meaning (Chinese).
- In other words, variants could mean either visual or semantic similarity.

- Both cases exist in Chinese, and it does not matter whether they look the same or not, the important part is if they mean the same, and it is up to the community to define that.
- It is difficult to have one definition that addresses both cases. The current definition has to be changed because it does not reflect the Chinese/Japanese/Korean (CJK) case. It may be more important to create a set the rules rather than trying to come up with a universal definition.
- ICANN's proposed definition of variants for the purpose of the Fast Track process was developed to enable allocation of all strings that would not jeopardize the stability of the Internet. As such only confusable strings were suggested to be placed on hold until a stable solution for introducing them was developed and implemented.
- Different views on how to manage variants were discussed as follows:
 - There is no mechanism to allow two TLDs to be synchronized in the root, i.e. no resource record can be used to make them map at the root.
 - The counter view is that this is not a problem. If a TLD can be represented with two strings (variants), then both strings should be delegated in the root, and leave it to the registry to figure out a solution to deal with aliasing.
 - There could be an agreement between ICANN, the registry, and possibly the government stating how this aliasing/mapping is going to work at the registry level.
 - Aliasing or mapping should not be used because it's proven not to be working - alternatively, variants should be delegated to the registry that can make bundles by using different records not resource records.
- A note is made to differentiate between language and script while talking about variants.
- Participants will continue the work before next call by sending proposals for solutions to deal with variants to the mailing list. This with a focus on the technical implementation of such solutions.

Discussion on the 3-character restriction in gTLD strings:

- Should the U-label or the A-label be used as a reference for measuring a string's length?
 - For the purpose of this discussion the U-label length is used.
 - A suggestion is made to use A-label instead.
- The issue of restricting the strength length to 3-characters has to do with semantic alphabets, and it has nothing to do with phonetic alphabet (i.e. CJK).
 - One proposed solution is to specify languages that have problems with this restriction and treat them separately.
 - However, a global solution would be difficult.

IDN-Implementation Working Team – Final Report

- Discussion on what needs to be protected in the 2-character space so that a solution does not infringe with 2-letter ASCII strings or create confusion with ASCII country code names.
- The 2-character ASCII space is reserved, so perhaps it is fair to reserve less-than-3-character TLDs in those scripts that are visually confusable with ASCII.
- It is considered beneficial to understand how TC 46 and ISO-3166 maintenance agency function and what communication may be necessary with them to expand the gTLD space.
- There is an early GAC statement to protect ccTLD namespace and whether it should be limited to ASCII, and not extend into IDNs.

Tasks:

- Participants to send to the list their proposed solutions for variants with specific examples and technical details on how these solutions work
- ICANN staff to provide information on how TC 46 and ISO-3166 maintenance agency function operates
- ICANN staff to provide reference to GAC statement on ccTLDs protection
- ICANN staff to provide summary of discussions for public posting

**IDN Implementation Support Team
Second Meeting: Conference Call
Wednesday 9 Sep, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)
Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Summary

Time allocation for Variants vs. 3-character:

- A point was raised that variants have been dominating the team's discussions since last week's call while less attention has been given to the 3-char gTLD problem. There was a homework on variants given to team members who were asked to send their proposed solutions to the list, so there was more material on variants available for further discussion, hence more momentum.

Posting of meeting minutes:

- The aim is to have summary of teams' discussions published in order to maintain a level of transparency. Minutes should be posted in the same week as the call.

Discussion on variants:

- Variants based on script systems: A question on whether or not there is a technical solution that can deal with variants at the root level led to a deeper discussion on what variants are, the various categories, and the need to deal with variants on a case-by-case basis, and based on the case / category, try to look for a technical solution. Short-term vs. Long-term: Another question was raised on whether the team should propose short term solutions as far as the IDN ccTLD Fast Track Process is concerned, or long term solution that could also feed in the new gTLD Program and the IDN ccTLD PDP process. There were different views on this point as follows:
 - It would be more appropriate to look for short-term solutions particularly for countries that have no choice but to use variant TLD or otherwise a portion of their population would be disfranchised.

- If a solution exists today, it could be the long-term solution for a specific case or category of variants.
- This team should come up with solutions that are short-term, but without creating any long-term problems.

It was also noted that it would be unfair for those who are ready to introduce IDNs today, to wait until a universal variant solution is available.

- **Enforced Bundling:** Discussion on bundling showed that while enforcing bundling may work in some cases, it may not work across the board. For instance the Chinese registry has been testing enforced bundling for a while and the user's experience and feedback has been positive. However, it was argued that the behavior of applications is beyond the control of any registry, and while registries can enforce bundles they won't be able to control the applications.

Discussions continued on what to do with variants, bundle them or block them. The counter argument to blocking was that it would not solve the problem as people who have no other way but to type the variant characters will be disfranchised. A question was raised if some client software could be used to deal with that problem (users are not able to type the primary TLD characters on their keyboard). It was stated that previous experience showed that relying on client software is not a good solution. A point was raised to differentiate between cases where variants look the same (fx. Arabic script) and cases where variants are distinct (fx. Chinese case).

- **The need for variants:** There was a proposal to gather a description of the problems that may arise if variants are not allowed at the top level. The purpose being (i) documenting the need for variants, (ii) discussing other solutions to the listed problems, and (iii) using the material for user education and outreach efforts. It was agreed to gather this material on the email list while the call should be dedicated to discussing solutions for variant delegation in the DNS root.
- **Reserving Preferred Variants:** It was noted that what ICANN had proposed in the IDN ccTLD Fast Track Draft Implementation Plan is a middle road between bundling and blocking. ICANN proposed to reserve the preferred variant (to put them on hold so they do not go to the wrong party), and to block the non-preferred variants. The intent being that ICANN considered reserving variants as an interim solution that would eventually secure the preferred variants to go to the right registry. If ICANN received a subsequent request for delegating a string that is blocked or reserved, that request would be declined.

Action items on variants:

- Cary to detail why enforced bundling is not feasible
- Howard to write down his proposal on variant handling list
- Tina to send out current preferred versus non-preferred proposal to list
- Tina to compile what has been said regarding problems arise due to absence of variants and send to list in a base document that can be used to compile the team's recommendations/results of their work.
- Harald to list the threshold "can this be enforced/will this work" questions to list

Discussion on 3-character restriction in gTLD strings:

- Discussion on confusability with ISO 3166 2-character codes, and there seemed to be general agreement that all possible 2-character ASCII codes should be set aside for use by the ISO 3166 MA, and neither those codes nor any codes that are visually confusable with them should be allocated through the gTLD process. The team also agreed that there are 2-character strings that are not confusable with the 3166 codes, and that such strings should not be restricted. Otherwise a lot of Chinese names for instance will be ruled out. The question was how to formulate that in a way that do not cause problems with procedures already established in the ICANN community, for example with the GAC, ISO MA, TC46, and others.
- The team also agreed that the problem they were looking at (3-character restriction) is only related to the gTLD process. In the IDN ccTLD Fast Track Process the 2-character strings are allowed, and strings could only be ruled out based on the confusability.. The point was raised that since strings that are longer than 2-character are allowed in the Fast Track, then the distinction between the ccTLD and gTLD space is blurred already and the desire to keep that distinction does not appear to be a valid argument for disallowed string with less than 3 characters in the gTLD process.
- It was noted that in the gTLD process, any applications for strings that are meaningful representation of country names should be endorsed by an approval of that country's government and that governments have an option for objecting to strings that resemble a country or territory name without adequate support.
- There was a discussion on confusability and mechanisms proposed by ICANN to deal with confusability. There was a comment that the issue of confusability would require some additional work.
- There was a discussion that there is no need for rules preventing any 2-character strings other than ASCII (or those similar to ASCII) from being allowed through the gTLD process.

Action items on 3-character restriction in gTLD strings:

- Harald to send his understanding on the agreements among the team on the 3-character topic, so that this can be agreed on the mailing list or next weeks call.

IDN-Implementation Working Team – Final Report

Next week's call:

- On Wednesday 16th at the same time
- Will spend an hour on the 3-char issue
- Agenda will be distributed beforehand

**IDN Implementation Support Team
Third Meeting: Conference Call
Wednesday 16 Sep, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)
Baher Esmat (ICANN Staff)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Absents:

Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)

Summary

Posting of the circulated summaries of call 1 & 2:

- Agreement on having the summaries of the first two calls posted.

Exploring relaxing the 3-char gTLDs restriction:

- There was a general agreement on relaxing the 3-character restriction for gTLDs. The team noted that in the IDN ccTLD Fast Track Process, the 2-character restriction for ccTLDs has been relaxed, so symmetrically the 3-character restriction should be relaxed for gTLDs.
- The team also noted that certain precautions must be considered while relaxing the 3-character gTLDs restriction. Some team members highlighted different precautions such as avoiding visual confusion with existing and potential ccTLDs, dissimilarity in contractual relationships between ICANN from one side and ccTLDs and gTLDs from the other side, objection process by governments, and need to get community or users to decide whether or not there is confusion. Participants were asked to send their proposed and detailed precautions to the list.
- Some suggested to try to take a conservative approach by limiting the number of languages that can have this relaxation. Although there is not a clear criteria on how to identify this list of languages, the aim is to try to take a gradual approach by starting with languages that really need to have the relaxation and continue to work on the issue as the need arises for other languages.

- There was a proposal to avoid the 1-character TLDs because of the higher opportunity of human error in typing a single character on the keyboard, which may lead the user to an undesired destination. Some argued this might not be the case in ideographic languages where multiple keystrokes may be required to generate a single character. Others thought it would be better to take a conservative approach and make the relaxation gradually from 3-character to 2-character, then see if the community raises any issues in relation to 1-character TLDs.

Action items:

- All team members to send their proposed list of precautions on the relaxing of the 3-character restriction.
- Howard to draft a proposal for 1-character TLDs.

Variant Management

- In reviewing the three statements on variants from last call where there was general agreement, it seemed that the terminology used with “desired” and “undesired” variants (delegated and blocked) was somehow confusing. It was not clear whether the term “delegation” meant “NS records in the root” or “available for putting NS records in the root”. Likewise, it was not clear if the term “blocked” meant that the string is not going to be available for any party for allocation, and what if another applicant (registry) disputed an undesired variant, could it be taken off the blocked list? However, the three statements were reasonably accepted by the team, though the terminology would need to be refined particularly the terms “delegated” and “blocked”.
- There was a strong opinion that the team should not go deeper with dispute mechanisms as these are policy issues that should be dealt with elsewhere.
- Regarding IDN tables, it was mentioned that the intention of ICANN’s staff was to use the tables submitted by registries, but also to try to cross reference them though there is no particular method to do that other than precautions like preventing confusion in the root, for which the SWORD algorithm is being implemented. It was also noted that for the IDN ccTLD Fast Track Process, a technical panel (stability panel) would do the visual similarity check. ICANN staff was asked to provide more information on the proposed algorithm.
- On the question of whether a universal solution is required or is possible for variants, it seemed that the team was more inclined towards the opinion of different solutions for the different cases. It was noted though that there has to be some text describing the cases and their solutions.
- On the issue of enforcing bundling and based on the “web servers” example that was sent on the mailing list, the team realized the difficulty of enforcing bundling, because there is not a mechanism to enforce certain behavior at the end user’s side. However, there was also an opinion that the problem caused by reserving the desired variants is bigger than that of how the web server

behaves. It was also noted that registries might have solutions to deal with bundling.

Action items:

- Tina to provide information about the SOWRD algorithm.
- Raed to share information about the bundling prototype that was proposed for the Arabic script.

Next week's call:

- On Wednesday 23rd at the same time
- Agenda will be distributed beforehand

**IDN Implementation Support Team
Forth Meeting: Conference Call
Wednesday 23 Sep, 2009**

Participants:

Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Absents:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)

Summary

Future calls:

- Agreed to continue with the weekly calls until 14 Oct 2009.

Posting of the circulated summary of the 3rd call:

- Agreed to have it posted.

Precautions with the relaxing of the 3-char gTLDs restriction:

- From last call, there were some precautions suggested by team members, and there was a discussion about whether an open or a conservative approach should be taken in dealing with that issue (relaxing of the 3-char restriction). In this call the team was trying to converge on the list of precautions, and in doing so the discussion was kicked off around the question of what a gTLD is and what a ccTLD is. However, the purpose of the question was not to define or redefine gTLDs and ccTLDs, it was rather to try to list and agree on some rules to help make the relaxation of the 3-char restriction implementable. Some members noted that the team should avoid making any policy statements, and rather try to focus on the implementation reality, and if there is any policy angle or implication, the team could make it clear in its recommendations that there is a policy aspect involved in this area, hence it should be dealt with in another fora.
- The team proposed three rules for 2-char strings:
 - IDN TLDs must not have visual confusion with existing ASCII TLDs
 - IDN TLDs must not have visual confusion with potential ASCII ccTLDs

- 2-char ASCII strings must only be used for ccTLDs
- A question was raised on reasons to give this kind of reservation to ccTLDs (bullets 2 and 3 above), and the answer was that there is no compelling reason why this restriction (3-char gTLD) should be removed on the ASCII side, whereas there is a compelling reason for other languages.
- The team seemed to agree that keystrokes should not be considered as a way to determine number of characters in a string. It was noted though that it would be useful to document the reasons behind this conclusion, so that the group can go back to the text when needed. There is a Section at the end of this report that incorporates input from team members on this subject.
- On the single character issue, there has been a proposal on the table to allow single character TLDs as long as they meet the following criteria:
 - The single character must be a meaningful word on its own within the language it from
 - The character must not be confusingly similar with ASCII letter/character
- There was a discussion on why the string must be meaningful if it is not confusingly similar with ASCII TLDs, and concerns were raised about the difficulty in defining what “meaningful” actually means for the different languages, etc. Some raised the point that in all previous IDN work (guidelines, protocol, etc.), invoking the concept of languages has always been avoided because DNS has nothing to do with languages, and it was rather the script that has been considered in such effort.
- It was also mentioned that the technical community had raised concerns in the past about single character TLDs. Given the technical and policy concerns, one suggestion was that this team would not recommend either lifting the ban on single character TLDs, or banning TLDs, but rather recommend to take this discussion to another forum. The question was what would this forum be? For example, is there a place where the Chinese example could be presented then lead to the creation of a new policy that may lead to a formal action? On this particular question, ICANN staff would check with the policy team if this could be addressed in GNSO or ccNSO.
- Some suggested that the team should make a stronger recommendation such as single character TLDs may be allowed if there are no technical or policy reasons to ban them. It was noted that the team would need to have better understanding as to the concerns of the technical community in order to be able to advocate for any counter proposals.
- ICANN staff also noted that the [Reserved Names Working Group Final Report](#) stated that *“Single and two-character U-labels on the top level and second level of a domain name should not be restricted in general. At the top level, requested strings should be analyzed on a case by case basis in the new gTLD process depending on the script and language used in order to determine whether the string should be granted for allocation in the DNS.”*

IDN-Implementation Working Team – Final Report

- There was a request to ICANN staff to liaise with members of technical community and get something in written about their reservations with regard to single character TLDs.
- Another point was raised with regard to visual confusion and the importance of comparing new strings with existing TLD strings, both ASCII and IDN. ICANN staff indicated that in the new gTLD process, the proposed algorithm does that. ICANN staff will come back with more information about the similarity check in the IDN ccTLD Fast Track Process.

Action items:

- ICANN staff to check with policy team on forums where the issue of single character TLDs can be discussed
- ICANN staff to liaise with technical community and ask for clarification on why single character TLDs should not be allowed
- ICANN staff to provide more details regarding similarity check mechanisms in the IDN ccTLD Fast Track Process
- Ram to refine the 3 rules suggested for the 2-char strings and to send his proposal to the list

Variant Management

- There was yet a disagreement among members as regards desired variants and whether they should be reserved or allocated/delegated. One of the problems is how to enforce behavior should desired variants be delegated. The discussion about DNAME demonstrated different views as some team members said they tried it on 2nd level and it didn't work, while others said it did. There also seemed to be contradicting messages from technical community on whether or not DNAME could be used in the root. Team members who had tried DNAME were asked to share their experience and to send some text to the list. The aim is to reach conclusion on DNAME and to decide if it is a viable solution for handling variants at the root level.
- The team recognized that terms like "delegation" and "allocation" still cause confusion, and agreed to refine the language and eliminate confusion.
- On desired variants the team seemed to have agreement on some parts and disagreement on others. There was an agreement on: 1. desired variants must be indicated by the requester (applicant); 2. desired variants must be reserved for the applicant; 3. undesired variants must be blocked until further notice. There was a proposal to prioritize the variants and to make their delegation dependent upon the analysis of the impact of the delegation of the initial string, and contingent of having a viable technology for mapping. Although the concept of prioritizing or ranking seemed to be accepted, there seemed to be disagreement on how to handle variants at the root level.

Action items:

- Howard and Siavash to provide text regarding their experience with DNAME

IDN-Implementation Working Team – Final Report

- Question to Raed (was not on the call) if .sa has any experience with DNAME and if so to share with the team
- ICANN staff to get feedback from .gr on its DNAME experience
- Ram to refine the definition of desired variants particularly the terms “delegation” and “allocation”

Next week’s call:

- On Wednesday 30th at the same time
- Aim of next call is to close down the 3-char issue
- Agenda will be distributed beforehand

Number of keystrokes as a criteria for determining number of characters in a string:

- Doesn't sound like a stable requirement as technology change in keyboards could change that.
- Number of keystrokes may vary for the same input, as Unicode is redundant and some same "characters" can be entered with a single keystroke (combined form) or with two keystrokes (bare character + combining mark). Thus the keystrokes would depend on the keyboard design as well, not just on Unicode encoding. And, many of these equivalent characters (in case of Arabic script at least) are not automatically dealt through the Unicode normalization process.

**IDN Implementation Support Team
Fifth Meeting: Conference Call
Thursday 1 Oct, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)
Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Summary

Posting of the circulated summary notes of the 4th call:

- Agreed.

Proposal regarding 2-character restriction in IDN TLD labels:

- The team discussed and agreed¹ on the following proposal:

Rule 1: No visual confusion is allowed with 2-letter strings drawn from A-Z (ASCII strings).

1a. All possible 2-character ASCII codes are set aside for use as ccTLDs by territories identified by ISO 3166 -1 codes.

1b. Both current and potential future codes need to be protected against visually similar strings.

1c. An uniform treatment of all ASCII strings is easy to understand and easy to practice.

Rule 2: No visual confusion is allowed with existing or applied-for TLDs.

2a. This is the same rule as for all other TLD applications.

Rule 3: No TLD is allowed with any string that is a meaningful identifier of a country or territory listed in ISO 3166-1 without the express consent of the government of the designated country or territory

3a. This is the same rule as for all other TLD applications.

¹ In discussing the 1-character rules below, the team agreed to add another rule to the 2-character rules above that deal with combining marks.

Proposal to restrict 1-character IDN TLD labels:

- Proposal suggested on the mailing list and discussed during the call:

1-character IDN TLD labels should be restricted pending further analysis: In its discussions, the IDN-Implementation team did not find implementation reasons to restrict 1-character IDN TLD labels. However, further analysis is required regarding the impact of the potential relaxation of the restriction on the allocation of 1-character IDN TLD labels. Until such analysis is conducted, the impact understood and appropriate review completed, 1-character IDN TLD labels should be restricted.

Notes:

The Reserved Names Working Group Final Report

<<http://syd.icann.org/node/6655>> stated that “Single and two-character U-labels on the top level and second level of a domain name should not be restricted in general. At the top level, requested strings should be analyzed on a case by case basis in the new gTLD process depending on the script and language used in order to determine whether the string should be granted for allocation in the DNS.”

Next Steps:

The IDN-Implementation team recommends that ICANN staff engage in further discussion on the topic of allowing 1-character IDN TLD labels with the ccNSO and GNSO Councils and consult with the technical community before proceeding further.

- It was noted that the rationale behind the above proposal is to take a conservative approach, since the team has not studied the details of the problems that may exist with 1-character TLDs. The proposal does not recommend the banning of the 1-character TLDs, rather it says that the issue should be addressed by policy bodies such as ccNSO and GNSO, because there are policy aspects involved and this team is not chartered to set policies.
- A concern was raised that the issue of 1-character TLDs mainly concerns the CJK community, and as far as the Chinese language is concerned there is no technical issue with having 1-character TLDs, thus one suggestion could be to allow 1-character TLDs for certain languages where there is a community need, and no evidence of any technical problem whatsoever. It was debated though that having this team decide on languages for which 1-character TLDs may or may not be allowed, will get the team into policy decisions it is not mandated to take.
- There was a suggestion to let the 2-character TLDs go first and hold up the 1-character TLDs to the next round, otherwise there is a risk of having both proposals delayed.

- Another concern was raised as to the lack of any potential timeline for ccNSO and GNSO to address this issue. The team recognized that the CJK case is essential and should therefore be addressed quickly.
- There was an agreement on passing on the issue of the 1-character TLDs to ccNSO and GNSO so that they make their determination, in addition to including some language in the final report that demonstrates the need for 1-character TLDs, so to instill in the policy making bodies the urgency in discussing the matter.
- There was a discussion on the likelihood of land-rush for 1-character TLDs (ASCII and IDNs), and the fact that it might be difficult to predict if a land-rush is going to happen, because although some believe that 1-character TLDs are valuable, there are application aspects that need to be taken into account (i.e. to make applications recognize 1-character TLDs). The team realized that there are significant economic considerations associated with the introduction of 1-character TLDs, and agreed to address those considerations in its final report.
- Some raised the point that in some languages/scripts there are certain characters that look as if they were one character, where in fact they are one character plus a combining mark (like for example letter 'a' with umlaut). The suggestion here was to include in the rules of the 2-character TLDs a rule to ban any 2-characters that are visually confusable with 1-characters, as long as there is no policy for implementing 1-character TLDs. There was an agreement on that proposal and on including a similar rule under 1-character TLDs too.

Action items:

- Howard to provide text that can be included in the final report about urgency for 1-character TLDs in the Chinese language
- Ram and Harald to put out modified 1-character and 2-character rules
- Team to review modified 1-character and 2-character rules and comment

Variant Management

- *Definitions proposed on the mailing list for Variants*
 - Base IDN TLD Label:** The applied for TLD label, in a specified language, represented in a defined Unicode script
 - Variant:** A Base IDN TLD Label combined with IDN Tables can generate several variant labels. Each variant label might have the same meaning or might be visually confusable with the Base Label.
 - Desired Variant(s):** A Variant that is commonly used interchangeably with the Base IDN TLD Label in the community that uses the Base Label.
 - Blocked Variant(s):** All Variants that are not Desired Variant(s).
 - Allocation:** A Base IDN TLD Label and its Desired Variant(s) must be indicated by the requestor/applicant. Once a Base IDN TLD Label is assigned to a requestor, then the Desired Variant(s) of the Base IDN TLD Label will be reserved solely for the use of the requestor. This process is called "Allocation".

Delegation: *The addition of the allocated TLD label (Base IDN TLD Label or Desired Variant) to the DNS root zone.*

- There were many comments and suggestions on the definitions above. One comment was that the definition proposed for variants brought in again the issue of “meaning” though in previous discussion the team seemed to have realized that meaning is language specific whereas visual is script specific. It was mentioned that the reason the concept of meaning was in there is that to bring in the Chinese case in which variants are not visually confusable. It was then argued whether the Chinese case would be considered variants or not, as some said that it is not, and suggested that variants should only deal with visually confusable strings. There did not seem to be an agreement on that proposal, but there was a thought that the team may want to address homoglyphic and synoglyphic variants separately and have different rules for each type. Team members also realized that they should be careful not to confuse the community with the definition, so they should try to have a definition for each of the different cases of variants, and to separate the definition part from the implementation part. There was also a stress that the team should not refrain from discussing the implementation part, and try to suggest solutions for how the different variants should be dealt with on the root level.
- The team then discussed the above definition in the context of homoglyphic variants only. A point was made that if two variant strings are to be delegated there must be a way to ensure that the two strings would not grow away from each other as TLDs and cause confusion to the end user. In discussing how to ensure that, there was the same kind of proposals that some had suggested in previous discussions. So there was a discussion on including terms and conditions in the arrangement the IDN ccTLD manager would have with ICANN. It was noted though that in the Fast Track process there would be very limited contractual relationships between ICANN and the IDN ccTLD manager, if any at all, so it would be unfeasible to include any terms and conditions in relation to variants in such arrangement, though it should be part of the longer-term PDP discussions.
- The suggestion on the table was to assign the variants for the same requestor/applicant but not to delegate them. Some argued that disallowing the delegation of variants will not solve the problem of end user confusability as in some cases users may fail to reach to certain TLDs because they type the wrong characters on their keyboards. It was also mentioned that the Urdu Internet community in the Middle East represents almost 10% of that in Pakistan, which is a significant figure.
- One of the issues the team had discussed in the last call was about the behavior after delegation, and the homework was to look into DNAME as to whether or not it could be a viable solution for variants on the root level. ICANN is still seeking input from one root server operator on that issue.

IDN-Implementation Working Team – Final Report

Other team members said they got reports on why DNAMES do not work and would send them to the team.

Action items:

- Team to review variant definitions and propose alternatives
- Team to review Harald's variant proposal and propose alternatives
- Howard and Leo to send to the list why DNAMES do not work
- Tina to send to the list why DNAMES do not work from root operator perspective

Next week's call:

- On Wednesday 7 Oct at the same time
- Call details and agenda will be provided beforehand

**IDN Implementation Support Team
Sixth Meeting: Conference Call
Wednesday 7 Oct, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Tina Dam (ICANN Staff)
Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Summary

Posting of the circulated summary notes of the 4th call:

- Agreed.

2-character restriction in IDN TLD labels:

- Team continued discussing the rules proposed from the last call. There seemed to be ambiguity with regard to some of the rules and team agreed that ICANN staff would work with the Chairs on refining the rules and send them back to the list.
- A question was raised as to whether the team intends to only list the rules or also propose means to implement them. It was noted that the team should try to get agreement on rules first, as this would be seen as a preliminary success, then either this team or another group would look into the implementation part.

1-character restriction in IDN TLD labels:

- There was a suggestion to call upon policy bodies like ccNSO and GNSO to pay attention to the concept of “glyphs” which may help with the CJK case in sorting out the issue of 1-character TLDs. Some expressed concerns towards using terminology that may not be well defined. It was mentioned that UNICODE got a definition for the term “glyph” though it seemed to be a bit ambiguous. There was agreement that whatever rules the team is going to propose, should use clear and well defined terminology.

Action items:

- ICANN staff to point out the group to the definition of “visual confusability” as defined in the new gTLD draft applicant guidebook.
- ICANN staff to point out the group to the definition of “meaningfulness” as defined in the IDN ccTLD Fast Track implementation plan.
- ICANN staff with Chairs of the group to refine the rules proposed to restrict 1-character and 2-character IDN TLD labels, and send them out to the group for final discussion and agreement.

Variant Management

- Distinction between two types of variants: Type 1: the variant is visually identical with the base label; and Type 2: the variant is not visually identical but is considered equivalent to the base label in some other sense. Team agreed that this distinction between the two types is necessary, though some suggested that the definition of Type 2 should be more specified.
- Some raised the point that a process would be needed to deal with variants after IDN TLD has been delegated, taking into account the basic principle that a delegated variant label should point to the same resource as the base TLD label.
- A discussion on expected behavior of variants followed. It was noted that 2 methods were suggested but were never agreed upon: 1. DNAME; and 2. Delegating the preferred variants to the applicant with terms and conditions set forth in a contract. There was yet no agreement on any of the two suggestions. It was noted that the team is yet to dig for more information as regards DNAME. Some also said that there must be clear and detailed documentation that explains the behavior of variants, should they be delegated, and how user confusion could be avoided.

Action items:

- Sarmad to propose text illustrating process to deal with variants after IDN TLD has been delegated (includes Ram’s suggested additions)
- Chairs to review text and suggest rules for the group
- Raed to make a proposal on how to make the definition of Type 2 variants more specific. The proposed definition thus far is: “variants that are not visually identical but are considered equivalent to the base label in some other sense.”
- Team members who have specific examples to sufficiently constraint the definition of Type 2 variants to send them to the list.
- Cary to provide a text on the argument that introducing potential confusion on the lower levels of the DNS is less significant than on the top level.

Next week’s call:

- On Wednesday 14 Oct at the same time
- Call details and agenda will be provided beforehand

**IDN Implementation Support Team
Seventh Meeting: Conference Call
Wednesday 14 Oct, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Baher Esmat (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Absents:

Tina Dam (ICANN Staff)

Summary:

Posting of the circulated summary notes of last call:

- Agreed.

Variant management:

2.1 Proposed text on Type 1 and Type 2 variants

- Discussion was focused on Type 2 variants. There was a proposal from last call defining Type 2 variants as “variants that are not visually identical but are considered equivalent to the base label in some other sense”. This definition seemed to be a bit broad, so there was another proposal sent to the list: “variants that are not visually identical but are considered equivalent to the base label from visual or sound in a way that might confuse many users”.
- Some were of the view of not including phonetic or acoustic similarity in the definition. It was also noted that phonetic variants is something that has been studied in the past and rejected, and if team members want to include it in the definition, they have to look into what have already been written in this area.
- There was a suggestion to include the term “orthographic” in the definition so it would read: “variants that are not visually identical but are considered equivalent to the base label in some orthographic sense”. There seemed to be agreement on including the term “orthographic”, and Cary was asked to amend the definition to include the term “orthographic”.

2.2 What is the intended behavior of variants and base TLD labels, means to avoid user confusion

- There was an agreement that the desired behavior of a variant is that both the base label and the variant give the same result for the end user, and in order to achieve that, it is necessary, but not sufficient, that both domain names have to resolve to the same resource records.
- Team also agreed that the final report should elaborate on cases where source of user confusion would lie entirely outside the DNS. Examples of HTTP and SMTP servers that operate on domain names in the form submitted by client software.
- The following proposal on how to handle Type 1 variants was discussed:
 - *1. An authorized TLD applicant may identify in addition to the main TLD string, a number of variant strings that are visually identical to the main string but bear different ASCII coding.*
 - 2. These variants are classified as 'desired' and 'undesired'.*
 - 3. Undesired variants are blocked; desired variants are reserved for the main string applicant.*
 - 4. A desired variant will be delegated to the (same) applicant provided conditions below are met:*
 - 4A. The applicant requests it.*
 - 4B. The domains under the variant are to be mapped exactly to corresponding domains of the main string.*
 - 4C. The applicant has to demonstrate how this mapping is to be achieved, preferably by presenting earlier work on a testbed.*
 - 4D. The applicant has to provide information on how it intends to inform its registrars/resellers/registrants about the correct set-up of name servers, and all other application servers that process domain names directly, so that end-user confusion can ideally be prevented, and at least held to an acceptable minimum.*
 - 4E. ICANN can conduct periodic checks to ensure that the said mapping is in place.*
 - 4F. The delegation document for the main domain (whatever form it takes: contract, MoU, accountability framework, ...) will bear an appendix committing the registry to the operation of the mapping.*
 - 5. It is understood that that technical or other considerations may necessitate changes in the items under 4 above.*
- There seemed to be an agreement on the above proposal. Also, there was a discussion about those who may act in a bad faith, and what measures should be taken to prevent end user confusion. It was noted that safeguards must be included in gTLD contracts, and in whatever kind of arrangement between the IDN ccTLD registry and ICANN (application form, exchange of letters, agreement), exactly as other terms and conditions are included.

2.3 Proposed process to deal with variants requested after a TLD string has been delegated

- Discussion here was focused on gTLDs, and fees associated with new gTLD applications. There was a question on the kind of fee one would expect to pay for the case of adding variant(s) after the base TLD string has been applied for and delegated. The team would request ICANN to work out a fee structure for that process.

Before winding up the discussion on variants, a point was raised with regard to the lack of technical details the team had received on DNAMES. For inclusion in its final report, the team would welcome input from Registries who have tried DNAMES on second level. However, team would not be able to comment on DNAMES (neither approve nor disapprove) at the top level pending technical details.

3. Final proposals on one-character and two-character rules in IDN TLD labels:

- Discussion was primarily focused on one-character rules as some thought the text proposed so far would give negative impression about the subject. There was an agreement to refine the language without getting into policy rules that lie outside the team's mandate.
- Team agreed to document discussions and rules on both one-character and two-character TLDs in the final report. Howard was asked to provide text with regard to the Chinese (or CJK) case.

4. Final report:

- There was an agreement that the team is going to make its final report ready by the ICANN meeting in Seoul. The report would simply document the progress the team had made with regard to the two key topics of variants and one/two-char TLDs, in addition to any recommendations for the way forward.
- ICANN staff would make a draft and pass it on to the Chairs by next Monday. The team would close the report down by next Wednesday to get it ready for Seoul.
- The report should target the ICANN Board and community members who are interested in the matter.

5. Any other business:

- Some indicated that more discussion is required on variants. There was an agreement to have another call next Wednesday.
- Some raised new agenda items such as the application form for new gTLDs and IDN ccTLDs, where applicants may not be aware of issues in relation to variants. Some said that whether this team or another group will have to look into these issues. It was noted that for this team to get its charter extended beyond Seoul, the team ought to identify the extra work that needs to be achieved.

IDN-Implementation Working Team – Final Report

- Team also agreed to have a face-to-face meeting in Seoul. ICANN staff should propose dates for team members to agree on one.

Action items:

- Cary to amend the definition of Type 2 variants including the term “orthographic”.
- Cary to provide text with regard to application servers such as HTTP and SMTP that operate on domain names in the form submitted by client software.
- Chairs to finalize the rules proposed to restrict one-character and two-character IDN TLD labels, and send them out to the group for agreement.
- Howard to provide text with regard to the Chinese (or CJK) case of one-character TLDs
- ICANN staff to make first draft of the report and send it to Chairs.
- ICANN staff to set up a meeting for the team in Seoul.

Next week’s call:

- On Wednesday 21 Oct at the same time
- Call details and agenda will be provided beforehand

**IDN Implementation Support Team
Eighth Meeting: Conference Call
Wednesday 21 Oct, 2009**

Participants:

Raed Al-Fayez (.sa)
Harald Alvestrand (IDNA author and ICANN Board)
Baher Esmat (ICANN Staff)
Cary Karp (IDNA author and .museum)
Guanghao Li (CNNIC)
Ram Mohan (.info and ICANN Board)
Kurt Pritz (ICANN Staff)
Siavash Shahshahani (.ir)
Yang Yu (CONAC, China)

Absents:

Tina Dam (ICANN Staff)
Sarmad Hussain (member of technical committee of Ministry of IT, Govt. of Pakistan)

Summary:

Posting of the circulated summary notes of last call

- Agreed.

Variant management

- There was a discussion Type 1 and Type 2 variants and whether they should be dealt with differently in the application process. Some said the distinction between the two types is without a difference, and that they should be handled in a similar way (using same rules), with a few exceptions.
- No objections or concerns raised with regard to the 5 rules that were proposed and discussed during the last call:
 - *1. An authorized TLD applicant may identify in addition to the main TLD string, a number of variant strings that are visually identical to the main string but bear different ASCII coding.*
 - *2. These variants are classified as 'desired' and 'undesired'.*
 - *3. Undesired variants are blocked; desired variants are reserved for the main string applicant.*
 - *4. A desired variant will be delegated to the (same) applicant provided conditions below are met:*
 - A. The applicant requests the variant TLD.*
 - B. The domains under the variant are to be mapped exactly to corresponding domains of the main string.*
 - C. The applicant has to demonstrate how this mapping is to be achieved, preferably by presenting earlier work on a testbed.*

D. The applicant has to provide information on how it intends to inform its registrars/resellers/registrants about the correct set-up of name servers, and all other application servers that process domain names directly, so that end-user confusion can ideally be prevented, and at least held to an acceptable minimum.

E. ICANN can conduct periodic checks to ensure that the said mapping is in place.

F. The delegation document for the main domain (whatever form it takes: contract, MoU, accountability framework, ...) will bear an appendix committing the registry to the operation of the mapping.

○ 5. It is understood that that technical or other considerations may necessitate changes in the items under 4 above.

- Agreement to add the issue of handling variants after an operator has been allocated TLD labels/variants to list of items that would require further work.

Final report

- A draft report was sent to team members right before the call. Team members didn't have the chance to look into it, but the co-chairs reviewed the report, and team members agreed to send their comments and edits by Friday 23 Oct.
- Team is intended to finalize the report by the Seoul meeting and make it available to the ICANN Board, Chairs of Supporting Organizations and Advisory Committees, and the ICANN community for public comments.

Team feedback

- Team agreed to disband after Seoul.
- General feedback on what the team has achieved was quite positive. Some specific remarks:
 - Better and more detailed mandate
 - Boundary between policy & implementation blurred
 - Inadequate attention given in policy bodies to details
 - More technical people needed for some topics
 - Better if discussions were more focused/narrow
 - Need to give more time to other items/new issues
 - Clear results achieved – 2 character restriction resolved; 1-character direction defined; variant problem better understood Very good for
 - Face-to-face time needed

Action items:

- Team to send comments on the draft report by Friday 23 Oct.
- Team to agree on date/time for face-to-face meeting in Seoul.

