IDN Variant TLD Implementation: Recommendations and Analysis

25 January 2019

Contents

1 Background

The current report is part of the six documents finalized and published after the public comment:

A. IDN Variant TLD Implementation – Executive Summary
B. IDN Variant TLD Implementation – Motivation, Premises and Framework
C. IDN Variant TLD Implementation – Recommendations and Analysis
D. IDN Variant TLD Implementation – Rationale for RZ-LGR
E. IDN Variant TLD Implementation – Risks and their Mitigation
F. IDN Variant TLD Implementation – Appendices (A: Glossary, B: Use of ROID, C: Limiting Allocated Variant TLDs)

2 Recommendations

This report proceeds from the starting premises and the findings of Integrated Issues Report (IIR). In addition, the report considers the practical facts that IDN variant Top-Level Domains (TLDs) labels will possibly be delegated for the first time so there is limited operational experience, and that as these variant labels are designated for the root zone, there is maximum impact of the decisions down the tree at other levels of the DNS. These considerations call for a conservative analysis at the outset, which could be reviewed over time as more experience is gained by the community.

Specifically, the core recommendations are as follows:

1. **Root Zone Label Generation Rules (RZ-LGR) the only source for valid TLDs and their variant labels.** All TLDs and their variant labels must be defined by the RZ-LGR. TLDs and their variant labels in the scripts not integrated in the RZ-LGR are undefined and, therefore, cannot be applied for or allocated. Any script that has no variant labels under the RZ-LGR has, by definition, no variants under this recommendation.

2. **IDN variant TLDs {t1, t1v1, …} allocated to the same entity.** For IDN variant TLDs that arise from an application and the RZ-LGR, all allocatable IDN variant TLD labels in the set must be allocated to the same entity or withheld for possible allocation only to that entity. In other words, for a top-level label *t1* allocated to Entity X, its allocatable variant label *t1v1* must only be allocated to Entity X or else withheld for possible allocation only to Entity X.

3. **Same second level label under IDN variant TLDs s1.{t1, t1v1, …} registered to the same entity.** For each allocated IDN variant TLD, a given second level label beneath the TLD must only be allocated to the same entity/registrant, or else withheld for possible allocation only to that entity. In other words, *s1* under *{t1, t1v1, …},* e.g., *s1.t1* and *s1.t1v1,* must be allocated to Entity Y or else withheld for possible allocation only to Entity Y.

There are additional recommendations to reduce end-user confusion, address end-user security and manage the variant labels in a stable manner. These recommendations, related to variant labels at the second level, also align with the recommendations in the latest version 4.0 of the IDN Implementation Guidelines, as has been revised by the community based working group.

4. **Second-level variant labels under IDN variant TLDs {s1, s1v1, …}.{t1, t1v1, …} registered to the same entity.** According to the IDN Implementation Guidelines, for second-level IDN variant labels that arise from a registration based on a second-level IDN table, all allocatable IDN variant labels in the set must only be allocated to the same entity or withheld for possible allocation only to that entity. This implies that all allocatable second-level labels *{s1, s1v1, …}* under all allocated variant TLD labels *{t1, t1v1, …}* must be allocated to Entity Z or else withheld for possible allocation only to Entity Z.

5. **Second-level IDN tables offered under IDN variant TLDs harmonized.** Second-level IDN tables applicable for an IDN variant TLD set must be mutually coherent but not necessarily identical. For two second-level variant labels *s1* and *s1v1* under any TLD *t1* generated using the applicable IDN table for *t1,* these must also be variant labels under TLD *t1v1* if generated by the applicable IDN table for *t1v1.* This also implies that the complete set of second-level variant labels may not all be valid under all variant TLDs. For example, for the second level label *s1v2,* the domain name *s1v2.t1* may be valid, but due to difference in IDN tables for variant TLDs, *s1v2.t1v1* may not be valid.

6. **Second-level variant label allocatable or activated under IDN variant TLDs not necessarily same.** The set of allocatable or activated second-level variant labels may not be identical across the activated IDN variant TLDs. For two variant labels *s1* and *s1v1* which are allocatable under the active IDN variant TLDs *t1* and *t1v1,* the label *s1.t1* may be allocated or activated but *s1.t1v1* may not be allocated or activated. Similarly, if *s1v1.t1* is allocated or activated, *s1v1.t1v1* may not be allocated or activated.

7. **Same registry service provider for IDN variant TLDs.** For feasible and consistent implementation of these requirements, the same back-end registry service provider[1], if applicable, must be employed for operating all the activated IDN variant TLDs by the registry operator.

The following additional recommendations also apply:

8. **Update existing policies and associated procedures to accommodate the recommendations for IDN variant TLDs.**   Existing policies and associated procedures must be adjusted to ensure that the recommendations above remain true under the functioning of gTLD and ccTLD policy and procedures.

[1] In the gTLD space, a registry service provider (RSP) is an organization providing one or more registry services (e.g., DNS, DNSSEC, RDDS, EPP) for a registry operator (the organization that signed a registry agreement with ICANN).

9. **All other existing TLD policies and procedures apply to IDN variant TLDs, unless otherwise identified.** Unless adjusted due to recommendation 9 above or other reasons identified and agreed by the community, because each IDN variant TLD is also another TLD, all existing TLD policies and procedures for allocation and delegation remain applicable for IDN variant TLDs as well.

It is worth emphasizing that subordinate zones are expected under IDNA2008 to contain additional policies for effectively managing IDN registrations.

2.1 A Re-Organized View of these Recommendations

These recommendations can be re-organized and viewed at Administrative, Policy and Technical Implementation levels based on the framework discussed in the document: IDN Variant TLD Implementation - Motivation Premises and Framework.

- Root Zone
    - Administrative
        - **IDN variant TLDs {t1, t1v1, …} allocated to the same entity**
        - **Same registry service provider for IDN variant TLDs**
    - Policy
        - **Root Zone Label Generation Rules (RZ-LGR) the only source for valid TLDs and their variant labels**
    - Technical Implementation
        - **None**
- Second Level
    - Administrative
        - **Same second-level label under IDN variant TLDs s1.{t1, t1v1, …} registered to the same entity**
        - **Second-level variant labels under IDN variant TLDs {s1, s1v1, …}.{t1, t1v1, …} registered to the same entity**
    - Policy
        - **Second-level IDN tables offered under IDN variant TLDs harmonized**
        - **Second-level variant label allocatable or activated under IDN variant TLDs not necessarily same**
    - Technical Implementation
        - **None**
- Additional Root Zone and Second Level
    - Administrative, Policy and Technical Implementation
        - **Update existing policies and associated procedures to accommodate the recommendations for IDN variant TLDs**
        - **All other existing TLD policies and procedures apply to IDN variant TLDs, unless otherwise identified**

2.2 Summary of Recommended Management Mechanism

Based on the discussion above, for the different dimensions and levels identified, the degree of influence on the management of IDN variant TLDs being proposed can be summarized as in the table below.

	Administrative	Policy	Implementation
Root Zone	***Intermediate***	***Intermediate***	***None***
Second Level	***Intermediate***	***Intermediate***	***None***
Subordinate Zones	***None***	***None***	***None***

The table indicates that for the root zone and the second level, an intermediate degree of influence is suggested for admirative and policy levels, whereas a minimal level of influence is suggested for the technical implementation level at the root zone. This relates to the position EF, as discussed in the document: IDN Variant TLD Implementation - Motivation Premises and Framework.

3 Analysis of Recommendations

3.1 Using RZ-LGR for covering all TLDs

Before IDN TLDs were possible, TLDs were limited to alphabetic labels[2] formed by only using letters a-z of Latin script. Latest version of the Maximal Starting Repertoire (MSR-3) identifies supporting 28 scripts[3], including letters a-z and other letters of Latin script, from the Unicode standard for inclusion in the RZ-LGR for representing the TLDs. Therefore, it is possible to review current and future TLDs using the RZ-LGR. This provides the additional advantage of capturing relevant cross-script variants, e.g., between Latin, Armenian, Cyrillic and Greek scripts. Segregating the evaluation of TLDs formed using letters a-z from the IDN TLD analysis would prevent discovering such cross-script variants, e.g. “epic” (Latin – U+0065 U+0070 U+0069 U+0063) vs. “epic” (Cyrillic - U+0435 U+0440 U+0456 U+0441). Therefore, evaluating all TLDs using RZ-LGR allows for a consistent approach.

3.2 What does “same entity” mean?

As explained in Section 2.3 of SAC 060 there is a “failure mode” when dealing with variant names that presents security implications. That is when two given variant names have two

[2] As per RFC 1123.

[3] MSR is limited to modern use scripts. See the MSR *Overview and Rationale* document for the complete list of scripts covered. MSR may be updated in the future to include additional scripts, based on input from the community.

different name holders (usually called “registry operators” at the TLD level, and “registrants” at any other level). In order to minimize this security risk, variant names that ICANN delegates[4] must have the “same entity” as the name holder. It is worth noting that this does not eliminate the risk, because that entity could itself be exploiting the failure mode; but it reduces the number of actors who can present the risk that the variants are managed inconsistently.

At the top level, having the same entity for two variant TLDs can be achieved by ensuring that the Registry Operator is the same. In practical terms, this is achieved by ensuring the sponsoring organization (including the name and address) for the two variant TLDs is the same, and that is reflected in the root zone Registration Data Directory Services (RDDS) operated by IANA.

3.2.1 Policy about the same label beneath variant TLDs

At the second level and below, ensuring the same label beneath all variant labels is allocated to the same entity could be achieved by some or all of the following: 1) having the same ROID for the registrant of both names; 2) having all the registrant fields be the same (without considering the ROID) for both names; 3) having a core subset of the registrant fields be the same (without considering the ROID) for both names; or 4) requiring a cryptographic probe that both registrants are indeed the same.

Option 4) above appears complex to implement since it would require, at least, two exchanges (i.e., at least one exchange for each variant names) between the Registry Operator (or the Registrar or a third party in its behalf) and the registrant(s). This option also has the drawback of not readily allowing a third party that has no relation to the Registry Operator to assess the sameness of the registrants, e.g., by looking at RDDS, assuming that the RDDS did not indicate this.

Options 1), 2) and 3) are similar, in that the only difference among them is the number of fields that have to be the same. These options also share the ability to be verified by a third party having access to RDDS under current RDDS policies. Multi-field validation of human-generated data is sometimes prone to error. Option 1) is the simplest by only requiring comparing one field (which implies the rest of the registrant fields are the same). In addition, because the ROID is generated by the repository, it is guaranteed to refer to the same contact object in the registry. Therefore, this report recommends Option 1). See Appendix 2 for further details on use of ROID in RDDS.

There is, however, a practical consideration with Option 1): some registrars in practice may not reuse contact objects for different registrations. If this option were to be used, registrars would need to support the requirement. ccTLDs that do not implement EPP must identify the "same entity" by some other mechanism. It is important to note that depending on heuristic matches of data fields generated by humans tends to be subject to errors introduced by those humans, so it

[4] Normally this is “from the root zone,” but the principle is stated this way to accommodate any case where ICANN delegates from some other zone.

is better to identify common data based on a unique identifier of some sort (i.e. something functionally equivalent to the use of contact object ROID in EPP).

1. *Consideration for both ccNSO and GNSO: Update policies and procedures to ensure that the definition of variant TLDs depends exclusively on the RZ-LGR.*
2. *Consideration for both ccNSO and GNSO: Update policies and procedures to incorporate the "same entity" rule for a given label beneath two variant TLDs.*
3. *Consideration for gNSO: Update policies and procedures to use ROID for the definition of "same entity". In case ROID is not an option finalized, then define an alternate functional definition for "same entity".*
4. *Consideration for ccNSO: Update policies and procedures to set a functional definition for "same entity" in the absence of EPP contact objects or associated ROIDs. (The action might be that this is not a ccNSO responsibility, but one taken by each registry instead.)*
5. *Consideration for PTI: Create mechanism for receiving RZ-LGR.*
6. *Consideration for PTI: Create "same entity" root zone registry rules.*

3.3 Enabling IDN Variant TLD Delegations

Because of the way the LGR works, some scripts can generate large numbers of variants. The RZ-LGR Procedure manages such numbers by minimizing allocatable variant labels and maximizing blocked variant labels. However, though this approach is optimal in most cases, the outcome may be worse for a specific label in some cases. For example, the table below shows the results of testing done on 476 valid[5] test labels given by the Arabic Generation Panel (GP) using the Arabic script LGR for the root zone. The table shows that though on average a label may have two variants, a particular label could have as many as 23 allocatable variants.

Original 476 Labels	**# Variant Labels Generated**	**Average # of Variant Labels**	**Maximum # of Variant Labels**
Allocatable	1033	2	23
Blocked due to variant mapping	100844	203	12388
Invalid due to WLE rule[6]	7368	15	1200
Total	**109721**	**221**	

[5] I.e. IDNA protocol-valid labels

[6] WLE: Whole Label Evaluation rule: this is an LGR rule that applies to whole labels instead of individual code points. This is necessary in some cases where for example, some codepoints cannot be at the beginning of a word in some scripts.

Based on the premises, the goal would be to minimize allocation of variant labels. RFC 5891, Internationalized Domain Names in Applications (IDNA): Protocol, Section 4.3, explicitly says that registries might create policies that restrict labels, and RFC 5894, Internationalized Domain Names for Applications (IDNA): Background, Explanation, and Rationale, Section 3.2 recommends being restrictive (though it also notes that variants may be helpful). Moreover, application of the principles in RFC 6912 would appear to suggest that fewer labels are better. Therefore, it might be argued that it is important to limit the "over-produced" labels: those that are allocatable but are still redundant. Recommendation 8 in [SAC060] states that "A conservative process needs to be developed to activate variants from allocatable variants in LGR." It explains:

> A string that is allocatable does not imply automatic activation; rather that it can be allocated … a clear process needs to be developed to avoid ad hoc treatment of new gTLD applications.
>
> The user experience report recommends that ICANN must implement a well-defined and conservative variant TLD allocation process. The SSAC agrees with the recommendations below:
>
> - The approval of a variant TLD must not be automatic, but initiated upon the request of a TLD applicant, explicitly specifying … (3) the need for the variant (e.g., motivated by linguistic, security, usability and/or other considerations) ….

The Integration Panel also noted on the release of LGR-1 that "it is a useful reminder that having a label that is 'allocatable' neither means that it will necessarily be delegated, nor that it necessarily should be delegated."

This mechanism could be different for gTLDs vs. ccTLDs, as they can have differing selection criteria. We consider that possibility in what follows; but recommend that, for the purposes of the DNS, each allocated and delegated variant is just another TLD. It follows from this that those existing policies for TLD allocation and delegation that are not explicitly altered continue to apply.

3.3.1 IDN gTLDs

For gTLDs, there is a significant application fee, which means that there is already a barrier to application for a large number of labels, since each application will be filtered through a business requirement analysis done by the applicant, and the same-entity requirement prevents the need for defensive registration. As this study recommends that the application process (and thus the application fee) for a variant will be same as for the gTLD label, this should encourage conservative allocation of variants[7].

[7] Alternate mechanisms have also been proposed for application evaluation process and fees in the public comments. The process and fees will be finalized in consultation with the community based on the policy and procedures finalized by the GNSO and ccNSO for IDN variant TLDs.

If any subsequent gTLD application process takes the current applicant guidebook as the starting point, then the process and guidebook should be updated considering the requirements set in this document. Specifically:

- the IDN tables for the second level domains should be submitted using the LGR format given in RFC 7940, Representing Label Generation Rulesets Using XML, in line with IDN Guidelines 4.0.
- the string review process should be updated to add the requirement that the applied-for TLD string should be verified against the latest root zone LGR; the output must be allocatable.
- each variant label in a set should be a separate TLD application.
- reserved names and the strings ineligible for delegation should be revisited to include any possible variant labels.

There are additional considerations around manageability and usability which still remain applicable to the TLD application review process, especially in cases where large number of allocatable variant labels may be allocated. Therefore, additional measures may be considered for limiting such cases. See Appendix 3 for a more detailed discussion on limiting the IDN variant domain names with the delegation of IDN variant TLDs.

3.3.2 IDN ccTLDs

IDN ccTLDs are currently allocated and delegated as per the IDN ccTLD Fast Track Process, Final Implementation Plan (FIP). The FIP limits the IDN ccTLDs applications to “[o]ne string per official language or script per country or territory.”   Moreover, at least the following linguistic requirements, in addition to the security and stability constraints, must also be fulfilled:

1. The label must be in an official language or script of the country or territory.
2. The label proposed must be a meaningful representation of the country or territory.
3. The label applied for must be supported by the community from the relevant country or territory.

Because a variant ccTLD is just another ccTLD, the same constraints should apply to it. However, these constraints may entail that in some cases no variants of an IDN ccTLD can be delegated, because, for example, the variant may be in the same language as the original IDN ccTLD or may not be in the official language of the country.

In Section 3.4 of FIP it is suggested that “The community is expected to continue working on more clear definitions of variants, solutions or methods for delegation of variants, and any necessary dispute mechanisms related to disagreement regarding desired and non-desired variants. For the purpose of including new development in the Fast Track Process, it is scheduled for revision.”

The revised process should minimally meet recommendations in Section 4 (above). Any revision should also strive to minimize allocation of variants as discussed above in this section and do so in a way which can be applied as part of the application evaluation, e.g. by developing usability-based criteria vetted by the local community.

Sections 3 and 4 of FIP need to be appropriately updated. The ccNSO and the community should also consider adjusting the IDN ccPDP accordingly.

It is worth observing that treating each variant label as being just another ccTLD, and applying the existing policy without any revision, could be at once too broad and too narrow:

- It may be too broad in that it could permit a very large number of allocatable (and therefore delegatable) variants in the case where a ccTLD operator is sufficiently insistent and the LGR is sufficiently permissive. This runs directly counter both to the recommendations of IIR and of advice from the IAB in RFC 6912 about administration of public infrastructure zones.

- It may be too narrow in that it could reject a defensible variant label on the grounds that it is not well-formed under the policy, when the ill-formation is precisely the point of the variant. (An example of this sort of case would be the use of a "foreign language" YEH character – e.g. U+064A ARABIC LETTER YEH vs. ARABIC LETTER FARSI YEH U+06CC – in a candidate string, which could be introduced precisely to solve issues with different encodings of the "same letter". This is by no means the only example of this sort of issue.)

Absent additional policies from the ccNSO, every new variant ccTLD should be treated just as any other ccTLD. If the ccNSO decides to produce new policies, however, it should attend to the fact that, in the DNS, there is at present no way to detect that names are related to one another and no reliable way either in the DNS or applications to make them work together automatically. Therefore, it may be desirable to leave the policies alone, and to treat the cases where they do not work well as an indicator that such cases are unlikely to work reliably anyway. It is nevertheless worth noting that the existing policy would not allow, for instance, a variant label in Persian language for an Arabic language IDN ccTLD, if Persian were not an official language of the country or territory in question.

There are additional considerations around manageability and usability which still remain applicable, especially in cases where a large number of allocatable variant labels may be allocated. Therefore, additional measures may be considered for limiting such cases. See Appendix 3 for a more detailed discussion on limiting the IDN variant domain names with the delegation of IDN variant TLDs.

7. *Consideration for ccNSO: Update Final Implementation Plan (FIP) of the Fast Track Process and subsequent IDN ccPDP, including update of Sections 3 and 4 of FIP (see discussion above); may include new draft letter between ccTLD managers and ICANN*

8. *Consideration for GNSO: Update policies and procedures for string review to ensure that every candidate label for allocation is allocatable.*
9. *Consideration for GNSO and ccNSO: Update domain transfer and update process to reflect inter-TLD linkages due to variants and the need to enforce the "same entity" rule (e.g. that s1.t1 and s1.v1t1 may have the same contact ROID after a <domainUpdate>).*
10. *Consideration for GNSO and ccNSO: Update policies and procedures to allow the lists of reserved names and the strings for inappropriate delegation to reflect any variants.*
11. *Consideration for ccNSO: Update ccTLD redelegation policy to reflect "same entity" constraint on variant TLDs.*

3.4 TLD Label States

A given label in an IDL set may be in one of several states[8]:

- Delegated
- Withheld-same-entity
- Blocked
- Allocated
- Rejected

To clarify the states defined in the IIR and copied in the Glossary, two other states are hereby defined:

Withheld-same-entity: A Withheld label is set aside for possible allocation only to the same entity of the other labels in the variant set. This is a special case of "withheld", with the condition made explicit. Note that this status does not guarantee that the label in question will in fact be allocated (because the label is also subject to other application conditions).

Rejected: This is a new proposed stated, beyond those identified in IIR. A Rejected label is set aside on administrative grounds outside the ordinary LGR procedures. In the gTLD application states (https://newgtlds.icann.org/en/applicants/advisories/application-contention-set-14mar14-en) this state encompasses both "Not Approved" and "Will Not Proceed". Labels that cannot be allocated on visual confusability grounds, based on the string similarity review step in the TLD application process, are also Rejected. If a single label in an IDL set is Rejected, it can return to Withheld-same-entity, but the condition is only satisfied if the Rejected status can be removed.

As an example, suppose that among the variant set (*t1*, *t1v1*, *t1v2*, *t1v3*, *t1v4*), an applicant has applied for *t1*, *t1v3* and *t1v4*. The labels of the set could then have the following state:

- *t1*: delegated
- *t1v1*: withheld-same-entity

[8] This list is not exhaustive: there may be some other states defined by other policies, but they are not relevant to the current report.

- *t1v2*: blocked
- *t1v3*: delegated
- *t1v4*: rejected (and withheld-same-entity)

In the example above, *t1* and *t1v3* would have been requested by the applicant, were allocatable by the LGR, then passed all other checks, and were therefore delegated. *T1v2* would have been blocked by the RZ-LGR and therefore could not have been delegated. *T1v1* would be withheld-same-entity, which means that it was allocatable by the RZ-LGR but was not requested by the applicant. The label *t1v1* could later be delegated to the same entity by the means of the normal process for requesting new labels. Also in this scenario, *t1v4* would be rejected by the evaluation process – for example, because of visual similarity with another already delegated label *t2*. Therefore, it is given a state “rejected” (as well as withheld-same-entity).

The table below details possible state transformations for individual labels in an IDL set (if the IDL set is one label large, then it is just a single label and the ordinary policies remain unchanged). Note that, from the point of DNS, the results of many of these will be the same, but the administrative basis for that outcome differs.

Initial state	**State may change to**	**Remarks**
Withheld-same-entity	Allocated	Allocation only to the same entity as another label in the IDL set. This change happens if a variant was not initially requested for allocation and later is.
Blocked	Withheld-same-entity	A later LGR may broaden the available labels in the IDL set. Such possible labels automatically become Withheld-same-entity.
Allocated	Delegated	Happens when name servers are added. (Not new.)
Delegated	Allocated	If a domain is removed from the DNS, the allocation can remain in place anyway. Rare in the root zone, but not new.
Rejected	Withheld-same-entity	Every Rejected label is automatically Withheld-same-entity as well. If the Rejected status comes off, the label can be handled as any other Withheld-same-entity label.

Note that an allocated or withheld-same-entity label cannot become blocked unless a new version of the LGR makes this possible. Under the LGR procedures[9], that should never happen, and if it did it would trigger a re-evaluation of the procedures.

[9] ICANN, “Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels”, March 2013. Available from https://www.icann.org/en/system/files/files/lgr-procedure-20mar13-en.pdf. Last visited August 2016.

12. *Consideration for ccNSO and GNSO: Update policies and procedures to incorporate variant label states and transitions between them.*

3.5 Registry Services Impacts

3.5.1 Harmonized IDN Tables for the Second-Level

A registry may offer registrations using different IDN tables to support different languages for different scripts. In case multiple IDN tables are offered, the latest (draft) version of the IDN Implementation Guidelines requires that these be harmonized to produce a consistent set of second-level variant labels. The current recommendation extends this requirement for such IDN tables to be consistent across the IDN tables used for the TLD variant labels as well.

Therefore, a registry will be required to harmonize all the relevant IDN tables for all the variant TLDs. The set of IDN tables do not need to be exactly the same under the TLD variants, as long as they do not produce a conflicting set of second-level variant labels. This means that if *s1v1 and s1v2* are second level variant labels under *t1* then these labels must not be non-variants under *t1v1*. One way to manage this would be to maintain a common set of harmonized second-level IDN tables for all IDN variant TLDs and then (a) choose all these IDN tables to offer for all IDN variant TLDs, or (b) choose a relevant different subset of IDN tables to offer for each different IDN variant TLD.

For example, harmonize IDN tables A, B and C. Then either offer all A, B and C under *t1* and *t1v1* or possibly offer different sets such as A and B for *t1* and B and C for *t1v1*. But IDN table D is not offered under any *t1* or *t1v1* before first harmonizing it with A, B, and C. Harmonization of IDN tables could be more easily managed if IDN tables are represented in machine readable format specified in RFC 7940.

3.5.2 Withheld Second-Level Variant Labels

Moreover, in principle, given the recommendations in this report, each variant domain name resulting from a second-level and/or a top-level variant label (e.g. *s1.t1*, *s1.v1t1*, *s1v1.t1*, *s1v1.v1t1*, and so on) can possibly be registered through different registrars. This could make inter-registrar transfer harder. So a registry may require all the variants to be registered through the same registrar. In all cases, all resulting domain names from the IDL set must either belong to the same entity, or be unallocated or blocked. By extension, if the policy is restricted to a single registrar, then any registrar transfer should involve all variants at the same time[10].

In the resulting set of variant domain names, the labels that have withheld-same-entity status cannot be allocated to another entity, because of the same entity rule. Therefore, upon querying the RDDS for those labels, one should receive the information that this label is not activated, but also that it is not available because of being part of a variant set. The EPP response message

[10] Note that this is entirely a registry policy; the only proposed recommendation is that the registrant must be the same for all the registrations of *s1* beneath all the variants of the TLD's IDL set.

should be enhanced to disclose the reason of the unavailability. Since all member labels of a variant set can be found by the RZ-LGR and the second level IDN tables, then there is no need to disclose the specific allocated variant in this case.

For the EPP <check> command, the response for a withheld-same-entity label should be <avail=”0”> and the reason should be: “unavailable because a variant of an allocated label.”

When a change in registrant of one of these variant domain names is started, the registry should wait until all allocated variants are in the pendingUpdate state for the same change before finalizing the change. The reason is to ensure that all variants of a set belong to the same entity.

13. *Consideration for GNSO and ccNSO: Update policies and procedures for filing IDN tables using the LGR format specified in RFC 7940 as per IDN Guidelines 4.0.*
14. *Consideration for PTI: Create IANA registry for registry-operator-submitted IDN tables (i.e. registry-specific LGRs) in RFC 7940 format.*
15. *Consideration for members of GNSO and ccNSO: Update policies and procedures to require harmonized IDN tables across IDN variant TLDs to produce a consistent set of second-level variant labels. Also, require second level variant labels to be allocated to the same registrant under all variant TLDs.*
16. *Consideration for GNSO and ccNSO: Those TLDs using EPP may need to create an enhancement (either a protocol modification, a standard message, or a standard extension) that permits expressing response messages for unavailability of an unallocated label due to variants. Work with the technical community to make this enhancement.*

3.6 Adjustments in Registry Agreement

In order to implement the recommendations of this report, the following changes will be needed to agreements prior to allocation of variant TLDs.

Every gTLD is subject to a Registry Agreement with ICANN. In the case of IDN variant TLDs (*t1, t1v1, t1v2,* and so on), each would be the subject of a (separate) Registry Agreement with ICANN, as each variant TLD is, in effect, a TLD. In accordance with discussion above, ICANN would execute each of these Registry Agreements with the same entity. To ensure that the requirement for IDN variant TLDs to be allocated to the same entity is maintained, each of these registry agreements (including the existing TLD contract for *t1*) must contain provisions requiring all the labels in the IDL set to follow the same process in the event of any registry transition via a Registry Transition Process or Change of Control. In this event, the composition of the allocated and delegated set of variant TLDs should not be allowed to change at the same time as the change of the operator. (This follows from the “stable transfer plan” requirement in existing IANA redelegation procedures.) Otherwise, such changes will follow the existing rules and procedures for (re)assignments.

The Registry Agreement must also include a requirement for subordinate names (i.e., domain names at the second and lower levels) allocated in the TLD to be treated as an atomic set (i.e. *s1.t1*, *s1.t1v1, s1v1.t1* and *s1v1.t1v1* are always either allocated to the same entity or withheld for possible assignment to that entity, if not blocked). This is true irrespective of whether any of the names is actually activated in the DNS, and whether any of the variants is actually registered. For every label *s1* allocated under each of the variant TLDs, if the label *s1* is allocated under any other of the variant TLDs, it is allocated to the same entity; otherwise it is withheld-same-entity. Similarly, for every label *s1v1* allocated under each of the variant TLDs, if the label *s1v1* is allocated under any other of the variant TLDs, it is allocated to the same entity; otherwise it is withheld-same-entity.

The registry agreement must also require all the different IDN tables being used by the IDN TLD and its variant TLDs to be harmonized, meaning if two labels *s1* and *s1v1* are variant labels under *t1*, then they cannot be non-variant labels under *t1v1*.

The Registry Agreement will also require that all the IDN variant TLDs are implemented through the same registry service provider, to promote a consistent and stable implementation across all such variant TLDs.[11]

In the event that a TLD's operations are transitioned to any Emergency Back-End Registry Operator (EBERO), an emergency transition of a TLD to an EBERO must trigger emergency transition of all variant TLDs to the EBERO. This is a continued application of the requirement for the TLDs to be operated by the same entity.

In the event of an emergency transition, an EBERO will provide continuation of a set of five critical functions for a TLD registry, but does not perform any billable transactions (i.e., new registrations, renewals, or inter-registrar transfers). Although it would be technically possible for an EBERO to be operating *t1* and the original contracted provider to continue operating *t1v1*, the latter could make changes to registration data whereas the former could not, and thus the requirement for subordinate names to be allocated to the same entity in both TLDs could not be maintained.

17. *Consideration for GNSO: If applicable, for the registry agreement propose changes for Registry Transition Process or Change of Control ensuring "same entity" rule is maintained.*
18. *Consideration for GNSO: Update EBERO provisions to ensure all names in an IDL set remain under unified control during EBERO.*
19. *Consideration for GNSO: Update registry agreement documents to ensure the label under variants TLDs (e.g. s1.t1*, *s1.t1v1, s1v1.t1* and *s1v1.t1v1) follow the "same entity" rule.*

[11] The initial set of recommendations proposed for public comment also required that the IDN variant TLDs be implemented using the same nameservers, unless otherwise justified. However, the recommendation is now removed based on the feedback received by the community asking for more operational flexibility in the implementation of IDN variant TLDs.

3.7 Adjustments in Registration Dispute Resolution Procedures and Trademark Protection Mechanisms

In order to implement the recommendations of this report, the following changes will be needed to dispute resolution procedures prior to allocation of variants.

Because of the possibility for a domain label and its variant labels (*s1* and *s1v1*) to be allocated to a registrant in multiple TLDs as a group (e.g., *s1.t1*, *s1.t1v1, s1v1.t1 and s1v1.t1v1*), decisions under certain dispute resolution procedures that result in changes to one domain name registration may also have variant domain name implications.

The Uniform Domain Name Dispute Resolution Policy (UDRP) provides an expedited administrative procedure for resolution of complaints alleging abusive domain name registrations (for example, cybersquatting). As the UDRP is a consensus policy, the GNSO is responsible for reviewing and recommending any changes to the policy. Accordingly, the issue of how domain name registrations in variant TLDs should be handled in the UDRP process should be brought to the attention of the GNSO working group examining the current Rights Protection Mechanisms (RPMs), including the UDRP.

There are at least two points in the process where variant TLD considerations could arise, and each presents a possible mechanism for addressing them. The first is in the UDRP itself. To address the variant TLD case, the UDRP must incorporate a provision that the “same entity” principle described above cannot be broken by a UDRP decision. Panels reviewing UDRP complaints would then abide by this principle such that if a panel decides to uphold a complaint and order transfer of a domain name registration to a new entity, this would result in transfer of all linked domain names (i.e., if a complaint is filed on *s1.t1* and succeeds, *s1.t1*, *s1.t1v1, s1v1.t1 and s1v1.t1v* would all be transferred). It would also be advisable for registrars to include provisions in the registration agreement to advise registrants of this possibility during the registration process. In this scenario, the complaint filing process would need to incorporate a mechanism to provide the UDRP panel with all relevant domain names to the case so that these could be taken account in the deliberations. This could be part of the information submitted with a complaint, subject to verification by the relevant registry/registrar.

The other mechanism is for the application of decisions to variant TLDs to occur at the point of implementation. In this scenario, no changes would be needed to the policy. The panel would consider the case and issue a decision based on the domain name(s) listed as the source of the complaint, and the registrar would implement the decision as occurs today. Where the complaint is successful, the UDRP decision would order transfer of *s1.t1* to the complainant and, given the registry’s obligation to ensure that *s1.t1*, *s1.t1v1, s1v1.t1 and s1v1.t1v* are allocated to the same entity, any related allocated names would be moved together. This should happen automatically according to the registry procedures – implementation of a UDRP decision would need to include steps to meet the same-entity requirement, but that follows from the general application of the same-entity rule. This can be enforced independently of any discussions about changes to the policy itself. The dispositions of the labels in question are always unaffected by the move:

those are controlled by the registrant, so if a registrant receives multiple names and does not want all of them activated, he or she can indicate that via the registrar. This may be an unintended consequence from the viewpoint of the complainant and the registrant and therefore should be well communicated.

The Uniform Rapid Suspension system is a complement to the UDRP for similar purposes; however, there is no apparent interaction with the same entity principle since the registrant data does not change as a result of a URS proceeding. Each name may be subject to URS independently.

The Trademark Post-Delegation Dispute Resolution Procedure (PDDRP) generally addresses a Registry Operator's complicity in trademark infringement on the first or second level of a new gTLD. Because each TLD registry can be treated independently in this procedure, each can be subject to decisions independently. In the case where a Registry Agreement is terminated as a result of a PDDRP determination, this would trigger the Registry Transition Procedure and various outcomes could apply. In the case of a reassignment of the TLD, the same entity rule would continue to apply so that *t1* and *t1v1* would be assigned to the same entity together, as described above.

Trademark Clearinghouse (TMCH) mechanism functions include authenticating information from rights holders and providing this information to registries and registrars. Recording a trademark with the Clearinghouse provides a rights holder with access to Sunrise registration periods in new gTLD registries and generates notifications from the Clearinghouse when a domain matching the authenticated trademark has been registered. The TMCH mechanism may need to be enhanced to include variants of trademarks *s1, s1v1* under TLD labels and their variants *t1, t1v1*. Further, a broader set of variants may need to be considered for TMCH instead of the registry-specific IDN tables. See SAC 60 for more detailed discussion.

20. *Consideration for GNSO: Update UDRP for UDRP application in the face of "same entity" restrictions. (See discussion.)*
21. *Consideration for GNSO: Possibly recommend updates to TMCH mechanism to include second level labels and their variants s1, s1v1 under TLD labels and their variants t1, t1v1 and broader calculation of variant labels.*

3.8 Adjustments in String Similarity Process

This section outlines a number of potential issues and modifications to the string similarity process that need to be considered in order to facilitate the deployment of TLDs containing IDN variant characters. While variant labels of a TLD that are eligible for inclusion in the root zone will be determined using the RZ-LGR, , the output of the RZ-LGR Procedure does not check for visual similarities. Thus, a string similarity review is required in order to prevent user confusion in the DNS resulting from delegation of similar strings.

New gTLD Applications. In the initial evaluation for new gTLD applications, a proposed applied-for TLD is checked against (a) existing TLDs and reserved names; (b) other applied-for strings; (c) strings requested as IDN ccTLDs; and (d) applied-for 2-character IDN gTLD strings against every other single character and any other 2-character ASCII string.[12]

Fast Track Requests. In the case of Fast Track requests for IDN ccTLDs, a string similarity assessment is made during the DNS Stability Evaluation. Requested IDN ccTLD strings are checked against (a) existing TLDs and reserved names; (b) other strings requested as IDN ccTLDs; and (c) applied-for gTLD strings.

An entity either already operating an existing TLD or applying for a new one may also apply for one or more allocatable variants (i.e. those with a status of withheld-same-entity) from the IDL set. Such TLDs will also be assessed for similarities to other existing or applied-for strings via the appropriate string similarity check (i.e. for gTLDs or ccTLDs according to the case). If an applied-for variant string fails the string similarity review, it would not pass the string evaluation step and this visually similar variant would get rejected.

In addition, the string similarity process must be modified to compare strings under consideration not just against all allocated or applied-for strings, but also all variants of those strings (including allocatable, withheld-same-entity, and blocked). The reason for this restriction is conservatism. If a string is merely withheld-same-entity and a second string is visually similar, then allocating the second string undermines the predictability of the outcome of variant processing from the RZ-LGR. Similarly, if a string is blocked under the RZ-LGR, but a visually similar string is allocatable, then the second (visually similar) string might become a “work around” for the blocked string. This approach is maximally conservative. It is worth remembering that, for such a conflict to occur, the strings as wholes must be visually confusable with one another. It is nevertheless worth noting that this expands considerably the number of strings that might need to be considered; the entire similarity review process will consequently probably become more expensive to operate.

This requirement does not mean that labels in the IDL set that are not being applied for must be evaluated to determine the status of the application. Rather, the evaluation needs to happen only at the time of allocation of a TLD label. But, as discussed above, it must happen against unallocated variants (including blocked ones) and not just against those actually allocated.

3.8.1 Resolving Possible Conflicts

Under the *Fast Track process*, the applicant of a requested IDN ccTLD string that is found to be confusingly similar by the panel may request for an external and independent “Extended Process Similarity Review Panel” to conduct a final confusing similarity assessment. There is, as noted, no established process to determine what to do in the event that two applications for two visually similar strings happen contemporaneously with one another (generally, the first-come-first-served rule prevails).

[12] The evaluations here are outlined in *gTLD Applicant Guidebook*, version 2012-06-04, section 2.2.1.1.1.

In the evaluation process for new gTLD applications, a string similarity panel creates contention sets that may be used in later stages of evaluation.

Contention situations between Fast Track requests and new gTLD applications appear unlikely to happen because of existing policies; however, should a conflict be identified, the application that has completed its respective evaluation process will prevail over a newly-filed string request. In the case where neither application has completed its respective process, where the gTLD application does not have the required approval from the relevant government or public authority, a validated request for an IDN ccTLD will prevail and the gTLD application will not be approved.

As shown in the table below, a contention set may occur in cases where two or more applied-for strings are (a) visually similar to one another; (b) variants of each other; or (c) both visually similar and variants of each other.

		Visually Similar?	
		Yes	No
Variants?	Yes	Same entity: Allowed Different entities: Contention Set	Same entity: Allowed Different entities: Contention Set
	No	Contention Set	Allowed

However, as per the current assumption, there is an exception to the string similarity process. In the event that two or more applied-for variant strings are visually similar they may only be allocated if they are associated with the same variant set and are being requested by the same entity. Regardless of whether applied-for top-level variants *t1v1* and *t1v2* of top-level label *t1* are visually similar to one another or are merely variants, they cannot be allocated to any other entity except to the entity that operates *t1*. In case of such conflicts across variants, the entire IDL set gets processed as one contention set; if one of the labels is already allocated, the contention is resolved in favor of the current operator.

Owing to the automatic way that variants are generated by the relevant LGR, it is necessary to perform the visual similarity checks for every requested-to-be-allocated variant in any given set against *all* the possible variants in every other set. (This is because such an available variant could be requested at any time in the future.) So, for example, suppose there is an applicant for *t1* that generates variants *t1v1, t1v2, …, t1v*n. Moreover, suppose that there is an applicant for *t2* that generates variants *t2v1, t2v2, …, t2v*n. Suppose in addition that the first applicant requests allocation of *t1v1* and *t1v2*, but the second applicant requests allocation only of *t2v1*. In this case, *t1v1* and *t1v2* must be tested against *t2v1, …, t2v*n anyway, and any variants found to be visually similar must place the entire IDL set into the Contention Set; alternatively, if both parties agree, the variants may both be marked “rejected”.

3.8.2 String Contention Resolution in gTLDs

Under current procedures, resolution of string contention for applied for gTLD strings may include components such as a *settlement between the parties* resulting in the withdrawal of one or more applications or a *community priority evaluation* if a community-based applicant in a contention set elects this option. If the contention is not resolved by community priority evaluation or agreement between the parties, an *auction* will take place.

For contention issues that involve the same entity, the following resolution options may be considered:

1. When the applied-for variant strings are placed in a contention set for later evaluation, the applicant is notified of the contention is and has the opportunity to establish that both applications are from the same entity.
2. It may be more efficient to establish early on in the string similarity review that the applied-for variant strings are being requested by the same entity prior to reaching the contention phase.

In both cases, the string similarity issue may be resolved directly with the applicant, which may result in the allocation or voluntary withdrawal of one or more applications.

For an applied-for variant string that is rejected as a result of a contention resolution, the following outcomes may be considered:

1. *Only* the applied-for variant string is rejected. For example, the applied-for variant *t1v2* of top-level label *t1* will get rejected while *t1v1* and *t1v3* from the same variant set continue to remain allocatable; or
2. The entire variant set including the applied-for variant string is rejected. For example, applied-for variant *t1v2* of top-level label *t1* will get rejected including *t1v1* and *t1v3* from the same variant set as *t1v2*. This outcome appears to be difficult to justify, though an applicant could decide that, if it cannot receive *t1v2* then it does not wish to proceed with the application. This option does not appear to be entirely in line with the rest of the recommendations in this report, but is included here for completeness.

22. *Consideration for GNSO and ccNSO: Update the string similarity guidelines for TLDs and their variant labels.*

3.9 Unaffected Policies and Procedures

3.9.1 No metadata sync requirement

Because the “same entity” recommendation is satisfied by using the same contact ROID (or functional equivalent), registrant metadata is automatically the same for all registrants of every allocated variant. As a practical matter, it is likely that it will be easiest for registry operators to keep a given name and its variants in the same registry database, but such an operational method is not a requirement. This also means that issues around privacy and proxy services are addressed, because the privacy or proxy service must still generate a contact ROID (or its functional equivalent) for the registrant. In registration systems that do not use contact objects, there is reason to suppose that a requirement about metadata syncing will be needed. That suggests that such registration systems are poor candidates for variants overall, because of the difficulty of maintaining metadata synchronization.

3.9.2 Data escrow

Because each variant of the IDL set is just another registration, data escrow policies for TLDs apply individually to each. Therefore, escrow requirements are automatically satisfied.

3.9.3 RZ-LGR Procedure

There are two ways that the RZ-LGR Procedure and these recommendations could interact. As new versions of the RZ-LGR come out (because of new versions of Unicode, or because of new scripts being added, or because of additional language support within a script), there may be additional variants that become allocatable for an existing allocation. In that case, the RZ-LGR has worked, and the new variants become eligible for allocation.

It is possible, however, that a new RZ-LGR would create a new set of variants such that a conflict arose between existing allocated variants. If that were to happen, however, it would be covered under the rules in B.6 of the “Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels”, version 2013-03-20b: “If an iteration of the process causes a subsequent repertoire to remove a code point that was in an earlier repertoire or to change an existing variant rule, all operation of the procedure must halt.” In effect, this would be a case where the RZ-LGR procedure turned out not to be as conservative as had been hoped, and would have failed to deliver a suitably stable RZ-LGR. (It might be inferred from this that the RZ-LGR Integration Panel should check a candidate RZ-LGR against the root zone before publishing it to ensure this situation does not occur.)

3.9.4 Domain name life cycles

The cycle of domain name registrations, either in the root zone or at levels beneath the root, is unchanged by this recommendation. Since each variant allocation is simply a different registration, it follows that names can be created and can expire at different times. The only restriction is the “same-entity rule”.

3.9.5 “Two-character” rule

The ccTLD labels in the root depend on an external registry (ISO 3166) that allocates alphabetic codes to countries. In order to ensure that no conflicts with future assignments by ISO can happen, ICANN has traditionally also maintained a restriction against the use of two-letter TLDs that do not correspond to ISO 3166 entries.

This restriction should be maintained for all Latin script letters. No variants should be generated for ccTLDs even if the Latin letters are included in the RZ-LGR, because the basis for the ccTLD labels lies in the external source (the ISO 3166 code); therefore, only that code should be allocated. For reasons of conservatism, however, the lack of “decorated” ISO 3166 two-character codes should not be regarded as a reason to permit such registrations.

23. *Consideration for ccNSO and GNSO: Review string similarity procedure to address decorated two-character Latin labels.*

3.9.6 Objections to gTLD applications

Under the rules of the most recent gTLD application round, there are four criteria for objections to a string (see “gTLD Applicant Guidebook”, version 2012-06-04, section 3.2.1). Under the recommendations of this report, there is no need to adjust those criteria, because each possible variant to be actually allocated is a separate application. Therefore, each such application is subject to the same rules.

3.9.7 Other Security and Stability guidelines

The “same entity” rule is intended to provide the maximal security and stability protection available given the way the DNS functions and the general-purpose utility of variants.

3.10 A Need for Policies at Every Level

It is important to emphasize that the recommendation in this report is in support of variants for *Top Level Domains* (more exactly. domains delegated directly by IANA under instruction from ICANN), and is not intended to suggest either that variants will be generally useful throughout the DNS, or that a variant system can be made completely and reliably safe for all users regardless of their sophistication. It is necessary that every zone operator makes policies for the affected zone. This issue is covered at length in RFC 5894, Section 3.2.

In most cases, user experience will be improved (and users will be better protected from bad actors) if other zones in the DNS (apart from the root zone) that are intended to be variants of one another ensure all of the following:

1. The LGR for each zone covers the same repertoire.
2. The LGR for each zone either generates the same IDL set or maximizes blocking such that divergent IDL sets are not possible

3. The LGR for each zone minimizes the number of code points permitted for registration at all.

Without such policies, it is possible that variants between two zones may result in user confusion or, worse, attacks on users via phishing.

4 Next Steps

These recommendations and the associated analysis were released for public comment. Based on the community input, this report on the implementation of IDN variant TLDs has been finalized. The final set of recommendations and its analysis will be presented to the ICANN Board for further consideration, in the context of its resolution in 2010, which calls for “an issues report identifying what needs to be done with the evaluation, possible delegation, allocation and operation of gTLDs containing variant characters IDNs".

This report sets the stage to initiate the discussion on implementing IDN variant TLDs. Until a reasonably comprehensive, cohesive and technically secure and stable solution has been agreed by the community for all the TLDs and adopted by the ICANN Board, the existing restriction on the delegation of IDN variant TLDs will continue to apply.