

新 gTLD 计划：IDN 变体

2010 年 2 月

引言

新 TLD 中变体字符的管理和实施会影响计划使用 IDN 和变体字符的语言群体。无论是实施 gTLD 还是 ccTLD 都会面临这种影响。本文件的目标是为 IDN 表的制定和使用以及变体 TLD 的分配和管理制定一套统一的程序。

定义：

1. IDN 表提供了根据注册政策可用于注册域名的字符的列表，其中包含可能存在的任何变体字符。
2. 变体字符是具有两种或多种表示形式（看上去可能彼此相似，容易引起混淆）的字符。
3. 变体 TLD是指彼此完全一致但字符串内部替代使用变体字符的 TLD 字符串。

机构群体已在多个场合对顶级字符串 IDN 表和变体字符的管理进行了讨论。在 2009 年墨西哥城会议和悉尼会议讨论之后，ICANN 建立了一支独立的实施工作组，负责调查这些问题。该工作组的部分工作是研究 TLD 字符串中的变体管理课题并提出可能的解决方案。工作组包括来自各个语言群体的语言和技术专家，由两名在 IDN 和 DNS 领域造诣精深的 ICANN 理事共同领导。

工作组建议目前不宜授权使用变体字符串作为 TLD，但可以对其机制进行测试，以便将来能够使用变体 TLD。特别建议，对于使用 DNAME 作为变体授权机制应进行系统化的测试，以便制定出一套适当的解决方案，确保在 DNS 中部署变体域名时保持一致性。工作组认为：如果将来要授权使用变体字符串，则必须满足 ICANN 与申请人之间达成的约定（即协议）所规定的特定条件。而后，IETF（互联网工程任务组）通过《互联网草案》(Internet Drafts) 提出将新资源记录（称之为 BNAME）作为管理变体 TLD 的解决方案。ICANN 将对 DNAME 和 BNAME 以及其他变体解决方案进行测试，并比较测试结果。

另外建议中还申明，TLD 申请人应对由相关 IDN 表生成的任何 TLD 变体字符串进行确认，如果申请人“想要”使用这些变体字符串，则这些字符串会分配给申请人并保持“待定状态”（申请人暂时无法使用），直至完成相应的映射机制。同时该工作组建议，作为一种预防措施，申请人应确定“不想要”的变体字符串，通过阻止使用这些变体字符串来避免申请人错误提交这些字符串的申请。

您可以访问以下网址来查看实施工作组的报告：

<http://www.icann.org/en/announcements/announcement-2-03dec09-en.htm>。

有关本主题的其他背景信息，请参见 <http://www.icann.org/en/topics/idn/fast-track/proposed-implementation-details-idn-tables-revision-1-clean-29may09-en.pdf>。

提议讨论的条款

截至目前，《gTLD 申请人指南》草案未对在顶级域名中使用 IDN 变体字符串的可能性给予说明。在提议纳入指南的新增内容中提到，待相应的机制成型并且制定好相关流程之后，可能会允许授权使用变体 TLD。

下面列出了可能纳入到《申请人指南》最终版本的文本内容，这些内容基于工作组提出的建议，有助于为讨论提供参考依据。相关内容出现在《申请人指南》的第 1 单元；该单元的完整内容请参见 <http://icann.org/en/topics/new-gtlds/intro-04oct09-en.pdf>。第 1 单元介绍了 gTLD 申请和评估流程，其中包括了有关对 IDN gTLD 申请人的特定要求的信息。

按照此处提议的方法，ICANN 会从所有新 gTLD 申请人处收集变体 TLD 字符串列表。按照新 gTLD 申请流程（如 IDN ccTLD 快速通道一样），申请人在申请相关 TLD 字符串时需要提交 IDN 表。相关语言的变体字符串通过申请人的 IDN 表来生成。

申请时，申请人必须确定想要的变体字符串（即申请人希望在相应机制成型后授权作为 TLD 使用的字符串）以及不想要的变体字符串（即申请人不希望使用而应被阻止掉的那些字符串）。但是，对于申请而言，不会授权批准在根区域中使用变体字符串。而且只有在制定好相应机制，完成测试，并且申请人证明能够按照所接受的标准实施机制的情况下，才能对变体字符串授权。此方式与 IDN ccTLD 快速通道中采用的方式一致。在快速通道流程中，申请中根据 IDN 表指明变体字符串，并同时指定想要的变体字符串。正如《IDN ccTLD 快速通道参与者的申请人手册》中所述，这并不表示会在 DNS 根区域中授权使用这些申请人想要使用的变体 TLD。

此处讨论的实施方法需要从申请人处收集变体信息，并且要考虑随着相关工作的进展，有关该程序需要确认的其他细节信息。

ICANN 欢迎就此处使用的文字提出意见和建议。这些文字仅供讨论之用，尚未合并到《申请人指南》中。《申请人指南》第 4 版完整草案中将会考虑这些意见，该版本预计在 2010 年 6 月发布。

1.3 IDN (国际化域名) 申请人的信息

可以预计到,某些申请的 g TLD 字符串属于国际化域名 (IDN),需要在 DNS 根区域中插入 IDN 编解码的 A 标签。IDN 是指包含以非基本拉丁字母表 (欧洲阿拉伯数字 (٠) 和连字符 -)(书写且在本地语言中使用的字符串的域名。

1.3.1 IDN 的特定要求

IDN 字符串申请人必须提供相应的信息,以表明符合 IDNA 协议和其他要求。可参见 <http://icann.org/en/topics/idn/rfcs.htm>。

申请人必须同时以 U 标签和 A 标签的格式提交所申请的 g TLD 字符串。

A 标签是 ASCII 形式的 IDN 标签。A 标签均以 IDNA ACE 前缀“xn--”开头,其后为通过 Punycode 算法获得的有效输出字符串,因此此类标签的最大长度为 59 个 ASCII 字符。前缀以及字符串都必须满足对可以存储在 DNS 中的标签的所有要求,包括符合 RFC 1034、RFC 1123 和其他相关文档中描述的 LDH (主机名) 规则。

U 标签是 Unicode 形式的 IDN 标签,也是用户希望实际显示的域名标签。

例如,以西里尔文字显示当前的 IDN 测试字符串,即 <xn--80akhbyknj4f>,而 A 标签为 <xn--80akhbyknj4f>。A 标签与 U 标签之间必须可以通过互转换来互相生成。

IDN gTLD 申请人在申请时还需要提供以下信息:

1. 字符串的短写形式 (英文)。申请人需要提供字符串含义或所代表对象的简要英文说明。
2. 标签语言 (ISO 639)。申请人应遵从 ISO 有关语言名称的表示规范,用英文指明所申请的 TLD 字符串的语言。
3. 标签文字 (ISO 15924)。申请人应遵从 ISO 有关文字名称的表示规范,用英文指明所申请的 TLD 字符串使用的文字。
4. Unicode 代码点。申请人应按照 Unicode 格式列出 U 标签中包含的所有代码点。
5. IDN 表。IDN 表提供了根据注册政策可用于注册域名的字符的列表。该字符列表包含注册二级域名时任何可视为彼此“相同”的多个字符 (“变体字符”)。IDN 表一旦由某个现有的 TLD 注册管理机构使用,即会存入 IANA 的 IDN 实践方法库中。有关其他信息,请参见 <http://iana.org/domains/idn-tables/> 处的现有表格以及 <http://iana.org/procedures/idn-repository.html> 处的提交指南。

申请时,至少要提交所申请的 g TLD 字符串使用的语言或文字的相关 IDN 表。对于希望提供相应的二级域名 IDN 注册的每一种语言和文字,申请人都要提供相应的 IDN 表。

应促使申请人在明确变体字符的工作中考虑语言和书写文字问题，此外，申请人应与其他提供相同或视觉上相似的字符的域名注册服务的 TLD 运营机构也可以将申请人提交的 IDN 表与已提交的其他同类语言的 IDN 表进行比较，并就任何不一致的地方进行询问。

6. 申请人必须进一步证实他们已合理采取措施来确保编码的 IDN 字符串不会引起任何显示或运行问题。例如，找到了在数字叫号路径分隔符（即一个点）的情况下字符串中混合了从右至左和从左至右两种方向的字符。如果申请人所申请的字符串存在已知的问题，则在申请中应记录将采取哪些措施来消除这些问题。虽然并不可能确保避免所有的显示问题，但重要的是应提尽早可能多地发现显示问题，并使注册管理机构运营商联系这些问题。申请人可通过以下方式熟悉这些问题：理解 IDNA 协议，尤其是所提议的新版 IDNA 协议（请参见 <http://www.icann.org/en/topics/idn/ufcs.htm>）；积极参与 IDN 维客活动（请参见 <http://idn.icann.org/>），其中说明了一些显示问题。
7. **[可选]** - 标签的音标字母表示法。对于所申请的 g TLD 字符串，申请人可以选择提供按照国际音标字母表示法 (<http://www.langsci.ucl.ac.uk/ipa/>) 标注的字符串音标。请注意，ICANN 不会对此类音标信息做任何评估或评价。如果申请人提供了音标信息，则 ICANN 在公开演示中答复关于相应域名申请的询问或者说出所申请的域名时会参考这些信息。

1.3.2 IDN 变体 TLD

每个申请包含一个申请的 gTLD 字符串。对于申请的字符串，申请人也可以指定任何现有的变体字符串。变体字符串是使用变体字符替换所申请的 g TLD 字符串中的一个或多个字符后所获得的字符。此外，列出的每个变体字符串必须符合第 2.1.1.3.2 节中的字符串要求。

对于列出的每个变体字符串，申请人必须指定其是否为：

- 想要的变体字符串。这表示申请人希望之后将该字符串用作 TLD 或者
- 不想要的变体字符串。这表示申请人在任何时候都不打算使用该字符串。

请注意，两个变体字符之间在视觉上并不一定相似。因此，没有对变体字符串之间的视觉相似性做任何要求。视觉相似性取决于其中使用的文字。

TLD 字符串的翻译或音译不会被视作变体。GNSO 在其 g TLD 政策制定工作中建议，新字符串在顶级域名的优先权不应由现有的字符串决定。

ICANN 将记录申请人列出的各个变体字符串作为参考信息，但在此阶段不会对变体字符串进行评估。

如果申请成功，则只授权所申请的 gTLD 字符串作为具有在制定好变体 TLD 的管理机制并经过 ICANN 测试后，才可对变体字符串进行授权。同时，ICANN 会要求申请人提交包括变体 TLD 管理机制实施细节在内的信息，并参与随后的评估流程。预计增加的费用以及审核步骤仍待确定。

1.3 国际化域名申请人的信息 (其中红线处显示的是相对《申请人指南》第3版所作的变动)

可以预计到,某些申请的 g TLD 字符串属于国际化域名 (IDN),需要在 DNS 根区域中插入 IDN 编码的 A 标签。IDN 是指包含以非基本拉丁字母表 (欧洲阿拉伯数字 (٠) 和连字符 -)(书写且在本地语言中使用的字符串)的域名。

1.3.1 IDN 的特定要求

IDN 字符串申请人必须提供相应的信息,以表明符合 IDNA 协议和其他要求。目前, IDNA 协议尚在修订之中,协议文档可参见 <http://icann.org/en/topics/idn/rfcs.htm>。

申请人必须同时以 U 标签和 A 标签的格式提交所申请的 g TLD 字符串。

A 标签是 ASCII 形式的 IDN 标签。A 标签均以 IDNA ACE 前缀“xn--”开头,其后为通过 Punycode 算法获得的有效输出字符串,因此此类标签的最大长度为 59 个 ASCII 字符。前缀以及字符串都必须满足对可以存储在 DNS 中的要求,包括符合 RFC 1034、RFC 1123 和其他相关文档中描述的 LDH (主机名)规则。

U 标签是 Unicode 形式的 IDN 标签,也是用户希望实际显示的域名标签。

例如,以西里尔文字显示当前的 IDN 测试字符串 `xn--80akhbyknj4f`,而 A 标签为 `xn--80akhbyknj4f`。A 标签与 U 标签之间必须可以通过相互转换来互相生成。

IDN gTLD 申请人在申请时还需要提供以下信息:

1. 字符串的短写形式 (英文)。申请人需要提供字符串含义或所代表对象的简要英文说明。
2. 标签语言 (ISO 639)。申请人应遵从 ISO 有关语言名称的表示规范,用英文指明所申请的 TLD 字符串的语言。
3. 标签文字 (ISO 15924)。申请人应遵从 ISO 有关文字名称的表示规范,用英文指明所申请的 TLD 字符串使用的文字。
4. Unicode 代码点。申请人应按照 Unicode 格式列出 U 标签中包含的所有代码点。
5. IDN 表。IDN 表提供了根据注册政策可用于注册域名的字符的列表。该字符列表包含注册二级域名时任何可视为彼此“相同”的多个字符(“变体字符”)。IDN 表一旦由某个现有的 TLD 注册管理机构使用,即会存入 IANA 的 IDN 实践方法库中有关其他信息,请参见 <http://iana.org/domains/idn-tables/> 处的现有表格以及 <http://iana.org/procedures/idn-repository.html> 处的提交指南。

申请时，至少要是交所申请的 g TLD 字符串使用的语言或文字的相关表。对于希望提供相应的二级域名 IDN 注册的每一种语言和文字，申请人都要提供相应的 IDN 表。

应促使申请人在明确变体字符的工作中考虑语言和书写文字问题，此外，申请人应与其他提供相同或视觉上相似的字符的域名注册服务的 TLD 运营机构也将申请人提交的 IDN 表与已提交的其他同类语言的 IDN 表进行比较，并就任何不一致的地方进行询问。

6. 申请人必须进一步证实他们已合理采取措施来确保编码的 IDN 字符串不会引起任何显示或运行问题。例如，找到了在数字叫号路径分隔符（即一个点）的情况下字符串中混合了从右至左和从左至右两种方向的字符。如果申请人所申请的字符串存在已知的问题，则在申请中应记录将采取哪些措施来消除这些问题。虽然并不可能确保避免所有的显示问题，但重要的是应提尽早尽可能多地发现显示问题，并使注册管理机构运营者知晓这些问题。申请人可通过以下方式熟悉这些问题：理解 IDNA 协议，尤其是所提议的新版 IDNA 协议（请参见 <http://www.icann.org/en/topics/idn/rfcs.htm>）；积极参与 IDN 维客活动（请参见 <http://idn.icann.org/>），其中说明了一些显示问题。
7. **[可选]** - 标签的音标字母表示法。对于所申请的 g TLD 字符串，申请人可以选择提供按照国际音标字母表示法 (<http://www.langsci.ucl.ac.uk/ipa/>) 标注的字符串音标。请注意，ICANN 不会对此类音标信息做任何评估或评价。如果申请人提供了音标信息，则 ICANN 在公开演示中答复关于相应域名申请的询问或者说出所申请的域名时会参考这些信息。

1.3.2 IDN 变体 TLD

每个申请包含一个申请的 g TLD。字符串申请的字符串，申请人也可以指定任何现有的变体字符串。变体字符串是使用变体字符替换所申请的 g TLD 字符串中的一个或多个字符后所获得的字符。此外，列出的每个变体字符串必须符合第 2.1.1.3.2 节中的字符串要求。

对于列出的每个变体字符串，申请人必须指定其是否为：

- 想要的变体字符串。这表示申请人希望之后将该字符串用作 。 TLD
- 或者
- 不想要的变体字符串。这表示申请人在任何时候都不打算使用该字符串。

请注意，两个变体字符之间在视觉上并不一定相似。因此，没有对变体字符串之间的视觉相似性做任何要求。视觉相似性取决于其中使用的文字。

TLD 字符串的翻译或音译不会被视作变体。GNSO 在其 g TLD 政策制定工作中建议，新字符串在顶级域名的优先权不应由现有的字符串决定。

ICANN 将记录申请人列出的各个变体字符串作为参考信息,但在此阶段不会对变体字符串进行评估。

如果申请成功,则只授权所申请的 g TLD 字符串作为顶层在制定好变体 TLD 的管理机制并经过 ICANN 测试后,才可对变体字符串进行授权。此时, ICANN 会要求申请人提交包括变体 TLD 管理机制实施细节在内的其他信息,并参与随后的评估流程。计划增加的费用以及审核步骤仍待确定。