

Network Working Group	M. Blanchet Ed.	
Internet-Draft	Viagenie	
Intended status: Informational	March 01, 2010	
Expires: September 2, 2010		

[TOC](#)

NewPrep Problem Statement

draft-blanchet-newprep-problem-statement-00.txt

Abstract

Using Unicode codepoints in protocol strings requires preparation of the string. Internationalized Domain Names(idn) initial work defined and used Stringprep and Nameprep. Other protocols have defined Stringprep profiles. New approach different from Stringprep/Nameprep is used for a revision of IDN. This document summarize the characteristics of both approach and provides guidance to protocol designers for handling internationalized strings.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on September 2, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such

materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

- [1.](#) Introduction
- [2.](#) Key Characteristics of Stringprep
- [3.](#) Key Characteristics of IDNAbis
- [4.](#) Protocols already using Stringprep
- [5.](#) Protocols looking for preparing internationalized strings
- [6.](#) Generalizing IDNAbis for other protocols
- [7.](#) Summary
- [8.](#) Security Considerations
- [9.](#) IANA Considerations
- [10.](#) Discussion home for this draft
- [11.](#) Informative References
- [§](#) Author's Address

1. Introduction

[TOC](#)

As part of the Internationalized Domain Names(idn) initial work [[RFC3490](#)] ([Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications \(IDNA\)," March 2003.](#)) [[RFC3491](#)] ([Hoffman, P. and M. Blanchet, "Nameprep: A Stringprep Profile for Internationalized Domain Names \(IDN\)," March 2003.](#)) [[RFC3492](#)] ([Costello, A., "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications \(IDNA\)," March 2003.](#)), the unicode-based strings needed to be prepared and normalized to enable their use in the DNS with exact match mechanism. The method, called Nameprep [[RFC3491](#)] ([Hoffman, P. and M. Blanchet, "Nameprep: A Stringprep Profile for Internationalized Domain Names \(IDN\)," March 2003.](#)), was specific to idn, but was generalized as Stringprep [[RFC3454](#)] ([Hoffman, P. and M. Blanchet, "Preparation of Internationalized Strings \("stringprep"\)," December 2002.](#)), to help other protocols with similar needs, but with different constraints than idn.

Stringprep defines a framework where protocols define their Stringprep profiles. Known IETF specifications using Stringprep are:

*The Nameprep profile [[RFC3490](#)] ([Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications \(IDNA\)," March 2003.](#)) for use in Internationalized Domain Names (IDNs)

*The iSCSI profile [[RFC3722](#)] ([Bakke, M., "String Profile for Internet Small Computer Systems Interface \(iSCSI\) Names," April 2004.](#)) for use in Internet Small Computer Systems Interface (iSCSI) Names

- *The Noderep and Resourceprep profiles [\[RFC3920\] \(Saint-Andre, P., Ed., "Extensible Messaging and Presence Protocol \(XMPP\): Core," October 2004.\)](#) for use in the Extensible Messaging and Presence Protocol (XMPP)
- *The Policy MIB profile [\[RFC4011\] \(Waldbusser, S., Saperia, J., and T. Hongal, "Policy Based Management MIB," March 2005.\)](#) for use in the Simple Network Management Protocol (SNMP)
- *The SASLprep profile [\[RFC4013\] \(Zeilenga, K., "SASLprep: Stringprep Profile for User Names and Passwords," February 2005.\)](#) for use in the Simple Authentication and Security Layer (SASL)
- *The trace profile [\[RFC4505\] \(Zeilenga, K., "Anonymous Simple Authentication and Security Layer \(SASL\) Mechanism," June 2006.\)](#) for use with the SASL ANONYMOUS mechanism
- *The LDAP profile [\[RFC4518\] \(Zeilenga, K., "Lightweight Directory Access Protocol \(LDAP\): Internationalized String Preparation," June 2006.\)](#) for use with LDAP

Based on findings [\[RFC4690\] \(Klensin, J., Faltstrom, P., Karp, C., and IAB, "Review and Recommendations for Internationalized Domain Names \(IDNs\)," September 2006.\)](#) on early deployments of idn, IDNs specifications have been updated /* note to add ref to idnabis RFCs*/ and do not use stringprep/nameprep anymore. Instead, a complete table of the supported Unicode codepoints is defined. This table is based on some properties of the codepoints as defined by Unicode and some specific exceptions. It also defines an algorithm to update the table with new versions of the Unicode codepoints database. Do other protocols, either current users of Stringprep or other protocols, would benefit from the new approach for idns? The aim of this document is to provide advice on the preparation of internationalized strings to protocol designers of current stringprep profiles as well as for new protocols.

2. Key Characteristics of Stringprep

[TOC](#)

Stringprep is used to prepare a string and output a string or an error. The output string is normalized so that it could be compared with others in a protocol. The functions and operations done by Stringprep are: mapping, normalization, prohibition and bidirectional string processing. Many functions use or may use Unicode specifications such as the normalization rules /* note to add ref to unicode normalization rules docs*/.

Functions provided by Stringprep are typically needed for many network protocols that handles internationalized strings. Stringprep has the following key characteristics:

- *Based on Unicode 3.2
- *Defines a list of codepoints that are specially processed, such as removed

*Codepoints not specifically listed are only normalized by Unicode normalization rules

3. Key Characteristics of IDNAbis

[TOC](#)

IDNAbis method has the following key characteristics:

- *Based on Unicode 5.2
- *Each Unicode codepoint is tagged
- *Basic table is created algorithmically using some properties of the Unicode codepoint database.
- *Some exceptions are specially handled
- *An exception table(currently empty) is provisionned in case a future version of Unicode introduces a new character or a change in a character that creates a compatibility issue with previous tables.
- *IANA registers the whole codepoint table for each Unicode database version.

When Unicode releases a new version of the codepoint database, the table has to be regenerated and registered by IANA. Exceptions may appear in the new table. In this case, an expert will have to manually resolve the exceptions.

4. Protocols already using Stringprep

[TOC](#)

Possible options are:

- *Leave it as is.
 - *Update Stringprep to new version of Unicode. Update protocol profile to new version of Stringprep.
 - *Define a generalized version of IDNAbis for protocols. Update protocol to use the new method.
-

5. Protocols looking for preparing internationalized strings

[TOC](#)

Possible options are:

- *Use Stringprep as is. Define a profile for it

*Use IDNAbis.

*Define a generalized version of IDNAbis for protocols. Define protocol to use the new method.

6. Generalizing IDNAbis for other protocols

[TOC](#)

A generalized version of IDNAbis to be used for other protocols might be interesting to do. Similar to Stringprep but using the IDNAbis approach.

7. Summary

[TOC](#)

When this document matures enough, then appropriate summary and conclusions will be written...

8. Security Considerations

[TOC](#)

TBD

9. IANA Considerations

[TOC](#)

This document has no actions for IANA.

10. Discussion home for this draft

[TOC](#)

This document is intended to define the problem space discussed in the newprep@ietf.org mailing list.

11. Informative References

[TOC](#)

[RFC3454]	Hoffman, P. and M. Blanchet, " Preparation of Internationalized Strings ("stringprep") ," RFC 3454, December 2002 (TXT).
[RFC3490]	Faltstrom, P., Hoffman, P., and A. Costello, " Internationalizing Domain Names in Applications (IDNA) ," RFC 3490, March 2003 (TXT).
[RFC3491]	

	Hoffman, P. and M. Blanchet, " Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN) ," RFC 3491, March 2003 (TXT).
[RFC3492]	Costello, A., " Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA) ," RFC 3492, March 2003 (TXT).
[RFC3722]	Bakke, M., " String Profile for Internet Small Computer Systems Interface (iSCSI) Names ," RFC 3722, April 2004 (TXT).
[RFC3920]	Saint-Andre, P., Ed., " Extensible Messaging and Presence Protocol (XMPP): Core ," RFC 3920, October 2004 (TXT, HTML, XML).
[RFC4011]	Waldbusser, S., Saperia, J., and T. Hongal, " Policy Based Management MIB ," RFC 4011, March 2005 (TXT).
[RFC4013]	Zeilenga, K., " SASLprep: Stringprep Profile for User Names and Passwords ," RFC 4013, February 2005 (TXT).
[RFC4505]	Zeilenga, K., " Anonymous Simple Authentication and Security Layer (SASL) Mechanism ," RFC 4505, June 2006 (TXT).
[RFC4518]	Zeilenga, K., " Lightweight Directory Access Protocol (LDAP): Internationalized String Preparation ," RFC 4518, June 2006 (TXT).
[RFC4690]	Klensin, J., Faltstrom, P., Karp, C., and IAB, " Review and Recommendations for Internationalized Domain Names (IDNs) ," RFC 4690, September 2006 (TXT).

Author's Address

[TOC](#)

	Marc Blanchet
	Viagenie
	2600 boul. Laurier, suite 625
	Quebec, QC G1V 4W1
	Canada
	Email: Marc.Blanchet@viagenie.ca
	URI: http://www.viagenie.ca