Libidn2 Reference Manual

COLLABORATORS			
	TITLE : Libidn2 Reference Manual		
ACTION	*NAME*	*DATE*	*SIGNATURE*
WRITTEN BY		March 28, 2017	

REVISION HISTORY			
NUMBER	DATE	DESCRIPTION	NAME

Contents

Chapter 1

Libidn2 Overview

Libidn2 is a free software implementation of IDNA2008 and TR46.

1.1 idn2

idn2 —

Functions

int	idn2_lookup_u8 ()
int	idn2_register_u8 ()
int	idn2_lookup_ul ()
int	idn2_register_ul ()
const char *	idn2_strerror ()

Types and Values

#define	IDN2_VERSION
#define	IDN2_VERSION_NUMBER
#define	IDN2_LABEL_MAX_LENGTH
#define	IDN2_DOMAIN_MAX_LENGTH
enum	idn2_flags
enum	idn2_rc

Description

Functions

idn2_lookup_u8 ()

```
int
idn2_lookup_u8 (const uint8_t *src,
                uint8_t **lookupname,
                int flags);
```

Perform IDNA2008 lookup string conversion on domain name *src* , as described in section 5 of RFC 5891. Note that the input string must be encoded in UTF-8 and be in Unicode NFC form.

Pass IDN2_NFC_INPUT in *flags* to convert input to NFC form before further processing. Pass IDN2_ALABEL_ROUNDTRIP in *flags* to convert any input A-labels to U-labels and perform additional testing. Pass IDN2_TRANSITIONAL to enable Unicode TR46 transitional processing, and IDN2_NONTRANSITIONAL to enable Unicode TR46 non-transitional processing. Multiple flags may be specified by binary or:ing them together, for example IDN2_NFC_INPUT | IDN2_NONTRANSITIONAL.

After version 0.11: *lookupname* may be NULL to test lookup of *src* without allocating memory.

Parameters

src	input zero-terminated UTF-8 string in Unicode NFC normalized form.	
lookupname	newly allocated output variable with name to lookup in DNS.	
flags	optional idn2_flags to modify behaviour.	

Returns

On successful conversion IDN2_OK is returned, if the output domain or any label would have been too long IDN2_TOO_BIG_DOMAIN or IDN2_TOO_BIG_LABEL is returned, or another error code is returned.

Since: 0.1

idn2_register_u8 ()

```
int
idn2_register_u8 (const uint8_t *ulabel,
                  const uint8_t *alabel,
                  uint8_t **insertname,
                  int flags);
```

Perform IDNA2008 register string conversion on domain label *ulabel* and *alabel* , as described in section 4 of RFC 5891. Note that the input *ulabel* must be encoded in UTF-8 and be in Unicode NFC form.

Pass IDN2_NFC_INPUT in *flags* to convert input *ulabel* to NFC form before further processing.

It is recommended to supply both *ulabel* and *alabel* for better error checking, but supplying just one of them will work. Passing in only *alabel* is better than only *ulabel* . See RFC 5891 section 4 for more information.

After version 0.11: *insertname* may be NULL to test conversion of *src* without allocating memory.

Parameters

ulabel	input zero-terminated UTF-8 and Unicode NFC string, or NULL.	
alabel	input zero-terminated ACE encoded string (xn--), or NULL.	
insertname	newly allocated output variable with name to register in DNS.	
flags	optional idn2_flags to modify behaviour.	

Returns

On successful conversion IDN2_OK is returned, when the given *ulabel* and *alabel* does not match each other IDN2_UALABEL_MIS is returned, when either of the input labels are too long IDN2_TOO_BIG_LABEL is returned, when *alabel* does does not appear to be a proper A-label IDN2_INVALID_ALABEL is returned, or another error code is returned.

idn2_lookup_ul ()

```
int
idn2_lookup_ul (const char *src,
                char **lookupname,
                int flags);
```

Perform IDNA2008 lookup string conversion on domain name *src* , as described in section 5 of RFC 5891. Note that the input is assumed to be encoded in the locale's default coding system, and will be transcoded to UTF-8 and NFC normalized by this function.

Pass IDN2_ALABEL_ROUNDTRIP in *flags* to convert any input A-labels to U-labels and perform additional testing. Pass IDN2_TRANSITIONAL to enable Unicode TR46 transitional processing, and IDN2_NONTRANSITIONAL to enable Unicode TR46 non-transitional processing. Multiple flags may be specified by binary or:ing them together, for example IDN2_ALABEL_ROUND | IDN2_NONTRANSITIONAL. The IDN2_NFC_INPUT in *flags* is always enabled in this function.

After version 0.11: *lookupname* may be NULL to test lookup of *src* without allocating memory.

Parameters

src	input zero-terminated locale encoded string.	
lookupname	newly allocated output variable with name to lookup in DNS.	
flags	optional idn2_flags to modify behaviour.	

Returns

On successful conversion IDN2_OK is returned, if conversion from locale to UTF-8 fails then IDN2_ICONV_FAIL is returned, if the output domain or any label would have been too long IDN2_TOO_BIG_DOMAIN or IDN2_TOO_BIG_LABEL is returned, or another error code is returned.

Since: 0.1

idn2_register_ul ()

```
int
idn2_register_ul (const char *ulabel,
                  const char *alabel,
                  char **insertname,
                  int flags);
```

Perform IDNA2008 register string conversion on domain label *ulabel* and *alabel* , as described in section 4 of RFC 5891. Note that the input *ulabel* is assumed to be encoded in the locale's default coding system, and will be transcoded to UTF-8 and NFC normalized by this function.

It is recommended to supply both *ulabel* and *alabel* for better error checking, but supplying just one of them will work. Passing in only *alabel* is better than only *ulabel* . See RFC 5891 section 4 for more information.

After version 0.11: *insertname* may be NULL to test conversion of *src* without allocating memory.

Parameters

ulabel	input zero-terminated locale encoded string, or NULL.	
alabel	input zero-terminated ACE encoded string (xn--), or NULL.	
insertname	newly allocated output variable with name to register in DNS.	
flags	optional idn2_flags to modify behaviour.	

Returns

On successful conversion IDN2_OK is returned, when the given *ulabel* and *alabel* does not match each other IDN2_UALABEL_MIS is returned, when either of the input labels are too long IDN2_TOO_BIG_LABEL is returned, when *alabel* does does not appear to be a proper A-label IDN2_INVALID_ALABEL is returned, or another error code is returned.

idn2_strerror ()

```
const char~*
idn2_strerror ();
```

Convert internal libidn2 error code to a humanly readable string. The returned pointer must not be de-allocated by the caller.

Returns

A humanly readable string describing error.

Types and Values

IDN2_VERSION

```
#define IDN2_VERSION "2.0.0"
```

Pre-processor symbol with a string that describe the header file version number. Used together with idn2_check_version() to verify header file and run-time library consistency.

IDN2_VERSION_NUMBER

```
#define IDN2_VERSION_NUMBER 0x02000000
```

Pre-processor symbol with a hexadecimal value describing the header file version number. For example, when the header version is 1.2.4711 this symbol will have the value 0x01021267. The last four digits are used to enumerate development snapshots, but for all public releases they will be 0000.

IDN2_LABEL_MAX_LENGTH

```
#define IDN2_LABEL_MAX_LENGTH 63
```

Constant specifying the maximum length of a DNS label to 63 characters, as specified in RFC 1034.

IDN2_DOMAIN_MAX_LENGTH

```
#define IDN2_DOMAIN_MAX_LENGTH 255
```

Constant specifying the maximum size of the wire encoding of a DNS domain to 255 characters, as specified in RFC 1034. Note that the usual printed representation of a domain name is limited to 253 characters if it does not end with a period, or 254 characters if it ends with a period.

enum idn2_flags

Flags to IDNA2008 functions, to be binary or:ed together. Specify only 0 if you want the default behaviour.

Members

IDN2_NFC_INPUT	Normalize input string using normalization form C.
IDN2_ALABEL_ROUNDTRIP	Perform optional IDNA2008 lookup roundtrip check.
IDN2_TRANSITIONAL	Perform Unicode TR46 transitional processing.
IDN2_NONTRANSITIONAL	Perform Unicode TR46 non-transitional processing.

IDN2_ALLOW_UNASSIGNED	Libidn compatibility flag, unused.
IDN2_USE_STD3_ASCII_RULES	Libidn compatibility flag, unused.

enum idn2_rc

Return codes for IDN2 functions. All return codes are negative except for the successful code IDN2_OK which are guaranteed to be

1. Positive values are reserved for non-error return codes.

Note that the idn2_rc enumeration may be extended at a later date to include new return codes.

Members

IDN2_OK	Successful return.
IDN2_MALLOC	Memory allocation error.
IDN2_NO_CODESET	Could not determine locale string encoding format.

IDN2_ICONV_FAIL	Could not transcode locale string to UTF-8.
IDN2_ENCODING_ERROR	Unicode data encoding error.
IDN2_NFC	Error normalizing string.
IDN2_PUNYCODE_BAD_INPUT	Punycode invalid input.
IDN2_PUNYCODE_BIG_OUTPUT	Punycode output buffer too small.
IDN2_PUNYCODE_OVERFLOW	Punycode conversion would overflow.
IDN2_TOO_BIG_DOMAIN	Domain name longer than 255 characters.
IDN2_TOO_BIG_LABEL	Domain label longer than 63 characters.

IDN2_INVALID_ALABEL	Input A-label is not valid.
IDN2_UALABEL_MISMATCH	Input A-label and U-label does not match.
IDN2_INVALID_FLAGS	Invalid com-bi-na-tion of flags.
IDN2_NOT_NFC	String is not NFC.
IDN2_2HYPHEN	String has for-bid-den two hy-phens.
IDN2_HYPHEN_STARTEND	String has for-bid-den start-ing/end-ing hy-phen.

IDN2_LEADING_COMBINING	String has forbidden leading combining character.
IDN2_DISALLOWED	String has disallowed character.
IDN2_CONTEXTJ	String has forbidden context-j character.
IDN2_CONTEXTJ_NO_RULE	String has context-j character with no rull.
IDN2_CONTEXTO	String has forbidden context-o character.

IDN2_CONTEXTO_NO_RULE	String has context-o char-ac-ter with no rull.
IDN2_UNASSIGNED	String has for-bid-den unas-signed char-ac-ter.
IDN2_BIDI	String has for-bid-den bi-directional prop-er-ties.
IDN2_DOT_IN_LABEL	Label has for-bid-den dot (TR46).
IDN2_INVALID_TRANSITIONAL	Label has char-ac-ter for-bid-den in tran-si-tional mode (TR46).

IDN2_INVALID_NONTRANSITIONAL	Label has character forbidden in nontransitional mode (TR46).

Chapter 2

Index