



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

Draft Policy Document For INTERNATIONALIZED DOMAIN NAMES

Language: MALAYALAM



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

RECORD OF CHANGES

***A - ADDED M - MODIFIED D - DELETED**

VERSION NUMBER	DATE	PAGES AFFECTED	A * M D	TITLE OR BRIEF DESCRIPTION	COMPLIANCE VERSION OF MAIN POLICY DOCUMENT
1.0	20/11/09	Whole Document	M	Language Specific Policy Document for MALAYALAM	1.5
1.1	22/11/2010	Page No 11, 24	A	Restriction rule added , ccTLD added	1.6



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF).....	4
1.1 Naming of Variables:.....	4
1.2 ABNF Operators.....	4
1.3 The Vowel Sequence.....	4
1.4 Consonant Sequence.....	5
1.5 ABNF Applied to MALAYALAM IDN.....	8
2. RESTRICTION RULES:.....	9
3. EXAMPLES.....	12
4. LANGUAGE TABLE : MALAYALAM	14
5. NOMENCLATURAL DESCRIPTION TABLE OF MALAYALAM LANGUAGE TABLE	15
6. VARIANT TABLE.....	22
7. EXPERTS/BODIES CONSULTED.....	23
8. PROPOSED ccTLD FOR MALAYALAM	24

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF)

1.1 Naming of Variables:

Dash	→ Hyphen -
Digit	→ Indo-Arabic digits [0-9]
C	→ Consonant
V	→ Vowel
M	→ Matra
D	→ Anusvara
X	→ Visarga
H	→ Halant/Chandrakala/Virama
Y	→ Avagraha/Praslesham
L	→ Chillu

1.2 ABNF Operators

S. No.	Symbols	Functions
1	"/"	Alternative
2	"[]"	Optional
3	"*"	Variable Repetition
4	"()"	Sequence Group

In what follows the Vowel Sequence and the Consonant Sequence pertinent to Malayalam are given. To facilitate use , equivalents in Devanagari are provided.

1.3 The Vowel Sequence

A vowel sequence is made up of a single vowel. It may be followed but not necessarily (optionally) by an Anusvara (D) or a Visarga (X) The number of D or X which can follow a V in Malayalam may be restricted to one.

The vowel sequence in Malayalam is therefore V [D | X]

Examples :

V	അ	अ
V[D]	അം	अं
V[X]	അഃ	अः

1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C or L)

Example:

C	ക	क
L	ണ	

2. A consonant optionally followed by dependent vowel sign[M] or Anusvara[D] or visarga[X] or Halanta [H]

C[M|D|X|H]

Example:

C[M]	കി	कि
C[D]	കം	कं
C[X]	കഃ	कः

C[H] ക് क् (Pure Consonant)

2.a. A CM sequence can be optionally followed by D or X
 (CM)[D|X]

Example:

CM[D] കിറ की

CM[X] വീഃ वीः

3. A sequence of consonants (up to 4) joined by Halanta *3(CH)C

Example:

CHCHCHC [ന്ത്ര്യ] → ന + ് + ത + ് + ര + ് + യ

न + ं + त + ं + र + ं + य

Subsets

3.a. The combination may be followed by H, M , D or X

*3(CH)C[H|M|D|X]

Example:

CHC[M] ക്കി ക്ക ് ക്ക ീ क्की

CHC[D] ക്കിറ ക്ക ് ക്ക ീറ क्कं

CHC[X] ക്കിഃ ക്ക ് ക്ക ീഃ क्कः

CHC[H] ക്കി ക്ക ് ക്ക ീ क्क्

3.b. *3(CH)CM may be followed by a D or X

*3(CH)CM[D|X]

Example:

CHCM[D]	ക്കിം	ക ്ക ിം	क्कीं
CHCM[X]	ക്കിഃ	ക ്ക ിഃ	क्कीः

4. A Chillu character may be followed by a Virama and consonant.

LHC

Example:

LHC ന്റ¹

5. A LHC can be optionally followed by a Anusvara[D], Matra[M] or Virama[H]

LHC[D|H|M]

Example:

LHC[D] ന്റം

LHC[H] ന്റ്

LHC[M] ന്റി

6. A LHCM can be optionally followed by a Anusvara[D]

LHCM[D]

Example:

LHCM[D] ന്റിം

1. Combination only exist where L is ന്(0D7B) and the following consonant must be

റ(0D31).



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

1.5 ABNF Applied to MALAYALAM IDN

Consonant Sequence \rightarrow $*3(\text{CH}) \text{ C } [\text{H}|\text{D}|\text{X}|\text{M}[\text{D}|\text{X}]] \mid \text{L}[\text{HC}[\text{D}|\text{H}|\text{M}[\text{D}]]]$

Vowel Sequence \rightarrow $\text{V}[\text{D}|\text{X}]$

Sequence \rightarrow $\text{Consonant Sequence}[\text{Y}] \mid \text{Vowel Sequence}[\text{Y}]$

IDN-Label \rightarrow $(\text{Sequence} \mid \text{digit})^*([\text{dash}](\text{Sequence} \mid \text{digit}))$

2. RESTRICTION RULES:

The ABNF is generic in nature and when applied to a specific language/script certain restriction rules apply. In other words, in a given language some of the Formalism structures do not necessarily apply. To take care of such cases restriction rules are set in place. These restrictions will help to fine-tune the ABNF. In the case of Malayalam the following rules apply:

1. H | D | X | M can not occur in the beginning of an IDN domain name

Example:

ഓ	ക്
ിക	ക
ക	ക
ഃക	ക

As can be seen they will result automatically in a “gol” marking an invalid character. This is an intrinsic property of the Indic syllable and is quasi automatically applied wherever supported by the OS.

2. H is not permitted after V, D, X, M, digit and dash

Example:

അ	ക്
ക	ക്
കഃ	ക്
കി	ക്
1	ക്
-	ക്



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

3. Number of D or X permitted after consonant-sequence or vowel-sequence or M is restricted to one

Example:

കറററ	കറ
കറഃ	കഃ
കീററ	കീ
കീഃ	കീഃ
അററ	അ
അഃ	അഃ

4. Number of M permitted after consonant-sequence is restricted to one

Example:

കീീ	കീ
-----	----

5. M is not permitted after V

Example:

ഇവറ	ഇവ
-----	----

6. The combination of Anusvara+Visarga and vice-versa is not permissible

Example:

കറഃ	കഃ
കഃറ	കഃ

7. “H” is permitted after only one “L” ie. ററ(0D7B) and the following consonant must be

റ(0D31).



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

8. A consonant syllable that is intended to end with Halant [H] can only be followed by Hyphen, digit or Avagraha [Y].



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

3. EXAMPLES

Following are the some examples of MALAYALAM IDN Label.

C	ക	കട
CH	ക്	അഴക്
CM	കി	കിണർ
CD	കം	കംഗാരു
CX	നഃ	പുനഃസംഘടന
CMD	കിം	കിംവദന്തി
CMX	ദുഃ	ദുഃഖം
CHC	ക്ക	നടക്കുക
CHCHC	ഇജ	വഇജം
V	അ	അവധി
VD	അം	അംഗീകാരം



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

L	ൽ	പാൽ
LHC	ൻറ	അൻറാർട്ടിക്ക
LHCD	ൻറം	ഫാൻറം
LHCM	ൻറി	ആൻറി



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

4. LANGUAGE TABLE : MALAYALAM¹

	Malayalam							
	0D00	0D01	0D02	0D03	0D04	0D05	0D06	0D07
0		എ	ഓ	ര	ീ		ജ	ധ
1			ഡ	റ	ൂ		ഞ	ന
2	ഌ	ഒ	ഡ	ല	ൂ		ി	നം
3	ഋ	കാ	ണ	ള	്യ		്	റ
4		ൗ	ത	ഴ	ൂ			ൾ
5	അ	ക	ഥ	വ				ൻ
6	ആ	ഖ	ഭ	ശ	ഐ		ഓ	
7	ഇ	ഗ	ധ	ഷ	ഓ	ൗ	ഫ	
8	ഈ	ഘ	ന	സ	ഐ		ഊ	
9	ഉ	ഘ		ഹ			ന	നം
A	ഊ	ച	വ		ഐ		ർ	ൻ
B	ഋ	ഛ	ഫ		ഐ		രി	ൻ
C	ൺ	ജ	ബ		ഐ		ന	ർ
D		ഡ	ഭ	ഛ	ഐ		ഊ	ൽ
E	എ	ഞ	മ	ഓ			വ	ൾ
F	എ	ട	യ	ി			ൻ	ക

¹ Characters marked in yellow are not applicable to the language.



5. NOMENCLATURAL DESCRIPTION TABLE OF MALAYALAM LANGUAGE TABLE

Unicode Value	Character	Character Name
ANUSVARA (D)		
0D02	◌̣	MALAYALAM SIGN ANUSVARA
VISARGA (X)		
0D03	◌̣̣	MALAYALAM SIGN VISARGA
VOWEL LETTERS (V)		
0D05	അ	MALAYALAM LETTER A
0D06	ആ	MALAYALAM LETTER AA
0D07	ഇ	MALAYALAM LETTER I
0D08	ഈ	MALAYALAM LETTER II
0D09	ഉ	MALAYALAM LETTER U
0D0A	ഊ	MALAYALAM LETTER UU
0D0B	ഋ	MALAYALAM LETTER VOCALIC R



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D0C	ൻ	MALAYALAM LETTER VOCALIC L
0D0E	എ	MALAYALAM LETTER E
0D0F	ഏ	MALAYALAM LETTER EE
0D10	ഐ	MALAYALAM LETTER AI
0D12	ഒ	MALAYALAM LETTER O
0D13	ഓ	MALAYALAM LETTER OO
0D14	ഔ	MALAYALAM LETTER AU
CONSONANTS (C)		
0D15	ക	MALAYALAM LETTER KA
0D16	ഖ	MALAYALAM LETTER KHA
0D17	ഗ	MALAYALAM LETTER GA
0D18	ഘ	MALAYALAM LETTER GHA
0D19	ങ	MALAYALAM LETTER NGA



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D1A	ച	MALAYALAM LETTER CA
0D1B	ചറ	MALAYALAM LETTER CHA
0D1C	ജ	MALAYALAM LETTER JA
0D1D	ജധ	MALAYALAM LETTER JHA
0D1E	ഞ	MALAYALAM LETTER NYA
0D1F	ട	MALAYALAM LETTER TTA
0D20	ഠ	MALAYALAM LETTER TTHA
0D21	ഡ	MALAYALAM LETTER DDA
0D22	ഡറ	MALAYALAM LETTER DDHA
0D23	ണ	MALAYALAM LETTER NNA
0D24	ത	MALAYALAM LETTER TA
0D25	ഥ	MALAYALAM LETTER THA



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D26	ദ	MALAYALAM LETTER DA
0D27	ധ	MALAYALAM LETTER DHA
0D28	ന	MALAYALAM LETTER NA
0D2A	പ	MALAYALAM LETTER PA
0D2B	ഫ	MALAYALAM LETTER PHA
0D2C	ബ	MALAYALAM LETTER BA
0D2D	ഭ	MALAYALAM LETTER BHA
0D2E	മ	MALAYALAM LETTER MA
0D2F	യ	MALAYALAM LETTER YA
0D30	ര	MALAYALAM LETTER RA
0D31	റ	MALAYALAM LETTER RRA
0D32	ല	MALAYALAM LETTER LA



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D33	ള	MALAYALAM LETTER LLA
0D34	ഴ	MALAYALAM LETTER LLLA
0D35	വ	MALAYALAM LETTER VA
0D36	ശ	MALAYALAM LETTER SHA
0D37	ഷ	MALAYALAM LETTER SSA
0D38	സ	MALAYALAM LETTER SA
0D39	ഹ	MALAYALAM LETTER HA
VOWEL SIGNS (MATRAS) (M)		
0D3E	ാ	MALAYALAM VOWEL SIGN AA
0D3F	ി	MALAYALAM VOWEL SIGN I
0D40	ീ	MALAYALAM VOWEL SIGN II
0D41	ു	MALAYALAM VOWEL SIGN U
0D42	ൂ	MALAYALAM VOWEL SIGN UU



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D43	ു	MALAYALAM VOWEL SIGN VOCALIC R
0D46	െ	MALAYALAM VOWEL SIGN E
0D47	േ	MALAYALAM VOWEL SIGN EE
0D48	ൈ	MALAYALAM VOWEL SIGN AI
0D4A	ൊ	MALAYALAM VOWEL SIGN O
0D4B	ോ	MALAYALAM VOWEL SIGN OO
0D57	ൗ	MALAYALAM AU LENGTH MARK
0D62	ൌ	MALAYALAM VOWEL SIGN VOCALIC L
VIRAMA (H)		
0D4D	്	MALAYALAM SIGN VIRAMA
AVAGRAHA (Y)		
0D3D	/	MALAYALAM SIGN AVAGRAHA
CHILLU LETTERS (L)		
0D7A	ൺ	MALAYALAM LETTER CHILLU NN



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

0D7B	ന	MALAYALAM LETTER CHILLU N
0D7C	ര	MALAYALAM LETTER CHILLU RR
0D7D	ൽ	MALAYALAM LETTER CHILLU L
0D7E	ൾ	MALAYALAM LETTER CHILLU LL



6. VARIANT TABLE

VARIANTS	
<p>ഉള</p> <p>0D33 + 0D33</p>	<p>ഉള</p> <p>0D33 + 0D4D + 0D33</p>
<p>സ</p> <p>0D24 + 0D4D + 0D38</p>	<p>ഡ</p> <p>0D1D</p>
<p>ഗ</p> <p>0D17 + 0D4D + 0D17</p>	<p>ഗ</p> <p>0D17 + 0D4D + 0D32</p>
<p>ന</p> <p>0D28 + 0D4D + 0D24</p>	<p>ന</p> <p>0D28 + 0D4D + 0D28</p>
<p>ശ</p> <p>0D36 + 0D4D + 0D32</p>	<p>ശ</p> <p>0D36 + 0D4D + 0D36</p>
<p>സ</p> <p>0D38 + 0D4D + 0D38</p>	<p>ഡ</p> <p>0D21 + 0D4D + 0D21</p>
<p>ബ</p> <p>0D2C + 0D4D + 0D27</p>	<p>ബ</p> <p>0D2C + 0D4D + 0D26</p>



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

7. EXPERTS/BODIES CONSULTED

Expertise provided by C-DAC Thiruvananthapuram.



Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

8. PROPOSED ccTLD FOR MALAYALAM

India (Bhārat) localized in Malayalam - ഭാരതം

Note: You can send your feedbacks to ids-feedback@cdac.in