

1 Consideration of reducing allocatable variant labels

2 (Evaluation report of CJK GP's LGR-alpha1)

3 2015-10-16 JGP

4 5 1. Recapitulation of previous meeting

- 6
- 7 • JGP explained that the biggest number of allocatable labels for one applied label
- 8 was 20736 amongst 125628 existing Japanese JP domain names as of Oct 2014.
- 9 • IP members required JGP to consider reducing allocatable variant labels.

10 11 2. JGP's investigation

- 12
- 13 • JGP used C-LGR-alpha (mss-var-20150615), J-LGR-alpha (2015-03-23) and K-
- 14 LGR-alpha (v0.1) for input.
- 15 • JGP used CJK integration algorithm 0.4.
- 16 • The biggest number of allocatable labels was not 20736.
 - 17 ➤ The number was 'The biggest number of calculated variant labels for one
 - 18 applied label', which means that the number includes 'blocked (out-of-
 - 19 repertoire-var)' variant labels.
- 20 • Real biggest number of allocatable labels for 125628 existing Japanese JP
- 21 domain names was 540 (Table 1-1, 1-2).
 - 22 ➤ 540 is still big number, but it was extremely rare case
 - 23 ➤ Most of (95% of) existing Japanese JP domain names has six or less
 - 24 allocatable variant labels.
- 25 • The biggest number of really registered 'variant labels for one registrant' was
- 26 four (there were only two cases).
- 27 • Average label length of existing Japanese JP domain name is about six
- 28 characters (Table 2), and the biggest number of allocatable variant labels for six
- 29 or less characters was 32 (Table 3-1, 3-2).
 - 30 ➤ It is hardly suppose that seven or more characters' Japanese label will be
 - 31 applied for TLD.
 - 32 ➤ In practice, Japanese applicant will select only one or two variant labels
 - 33 including applied-for label.
- 34

35 3. JGP's proposal

36

- 37 • For the variants marked as 'review', assign variant type (a) allocatable if the
38 variant is in original repertoire, or (b) blocked if the variant is out of original
39 repertoire.

40

Table 1-1: Calculated allocatable variant labels

Domains	Allocatable variant labels	Accumulation ratio	Domains	Allocatable variant labels	Accumulation ratio
64133	1	51.05%	8	60	99.90%
38931	2	82.04%	16	64	99.91%
3912	3	85.15%	37	72	99.94%
8765	4	92.13%	1	80	99.94%
183	5	92.28%	2	81	99.94%
3594	6	95.14%	31	96	99.97%
2550	8	97.17%	4	108	99.97%
399	9	97.48%	3	120	99.97%
151	10	97.60%	3	128	99.98%
1243	12	98.59%	9	144	99.98%
11	15	98.60%	3	162	99.99%
505	16	99.00%	6	192	99.99%
318	18	99.26%	2	216	99.99%
49	20	99.30%	1	240	99.99%
351	24	99.58%	1	243	99.99%
17	27	99.59%	1	256	99.99%
23	30	99.61%	1	288	100.00%
100	32	99.69%	1	324	100.00%
138	36	99.80%	1	384	100.00%
11	40	99.81%	1	432	100.00%
1	45	99.81%	1	486	100.00%
88	48	99.88%	1	512	100.00%
20	54	99.89%	1	540	100.00%

43

Table 1-2: Components of many variant labels

Allocatable variant labels	Total variant labels	Components of label	
540	972	66776767	# 0 LDH
512	4608	666766666766767	# 6 Level-1 Kanji (without new/old form)
486	972	7666BB66B6677	# 7 Level-1 Kanji (new form)
432	972	66776766676	# 8 Level-1 Kanji (old form)
384	20736	767666676606667	# 9 Level-2 Kanji (without new/old form)
324	1728	7676767667	# A Level-2 kanji (new form)
			# B Level-2 Kanji (old form)

44

45

Table 2: Distribution of Japanese JP domain label length

Domains	Label length	Accumulation ratio
626	1	0.50%
8521	2	7.28%
13781	3	18.25%
21809	4	35.61%
19393	5	51.05%
16986	6	64.57%
13596	7	75.39%
10696	8	83.90%
7255	9	89.68%
4792	10	93.49%
3057	11	95.93%
2101	12	97.60%
1413	13	98.72%
945	14	99.48%
657	15	100.00%

Label length average: 5.88

46

47

Table 3-1: Calculated allocatable variant labels for six or less characters

Domains	Allocatable variant labels	Accumulation ratio
48051	1	59.24%
28553	2	94.44%
8	3	94.45%
4043	4	99.43%
1	6	99.43%
423	8	99.95%
36	16	100.00%
1	32	100.00%

48

49

Table 3-2: Components of many allocatable variant labels for six or less characters

Allocatable variant labels	Total variant labels	Components of label	Allocatable variant labels	Total variant labels	Components of label	
32	162	777767	16	108	77767	# 0 LDH
16	1024	7767B6	16	108	7777	# 1 Hiragana
16	540	777667	16	108	7777	# 2 Katakana
16	540	667777	16	90	67777	# 3 Middle dot
16	288	777667	16	81	767767	# 4 Prolonged sound mark
16	288	777676	16	81	77767	# 5 Quasi Kanji
16	288	677767	16	81	77277	# 6 Level-1 Kanji
16	216	677677	16	81	77277	(without new/old form)
16	216	766777	16	72	77776	# 7 Level-1 Kanji (new form)
16	216	67B677	16	72	77776	# 8 Level-1 Kanji (old form)
16	216	777766	16	72	77677	# 9 Level-2 Kanji
16	162	767767	16	72	7777	(without new/old form)
16	144	77776	16	54	776677	# A Level-2 kanji (new form)
16	144	67777	16	54	767677	# B Level-2 Kanji (old form)
16	144	7777	16	54	76777	
16	144	7777	16	54	7777	
16	108	677677	16	48	67777	
16	108	767767	16	36	7777	
16	108	677767				

50