

程式設計 (一)

CH09.Q Base64

Ming-Hung Wang 王銘宏

tonymhwang@cs.ccu.edu.tw

Department of Computer Science and Information Engineering
National Chung Cheng University

Fall Semester, 2022

目錄

1. Base64 介紹
2. 轉換規則
3. 問題

Base64 介紹

Base64 是一種基於 64 個可列印字元來表示二進位資料的表示方法。常用於在通常處理文字資料的場合，表示、傳輸、儲存一些二進位資料。

Base64 常用於電子郵件中二進位資料的傳輸，或是於網頁的文本檔案表示圖片時使用，將圖片資料直接寫在文本檔案中，以減少 client 端請求資料的次數。

Base64 介紹

Base64 使用 0-9、A-Z、a-z、+、/ 共 64 個可視字元來表示二進位資料，因為 $\log_2 64 = 6$ ，每個字元可以表示 6 個位元。

Base64 以 4 個字元為一組，來表示二進位資料中的 3 個位元組。因此，轉換成 Base64 之後，資料長度會比原本多出 33% 以上。

例如 I see, I come, I conquer. (共 25 個位元組)，轉換為 Base64 為 SSBzZWUslEkgY29tZSwgSSBjb25xdWVyLg== (共 36 個字元)

轉換規則

Base64 索引表

Index	Binary	Char	Index	Binary	Char	Index	Binary	Char	Index	Binary	Char
0	000000	A	16	010000	Q	32	100000	g	48	110000	w
1	000001	B	17	010001	R	33	100001	h	49	110001	x
2	000010	C	18	010010	S	34	100010	i	50	110010	y
3	000011	D	19	010011	T	35	100011	j	51	110011	z
4	000100	E	20	010100	U	36	100100	k	52	110100	0
5	000101	F	21	010101	V	37	100101	l	53	110101	1
6	000110	G	22	010110	W	38	100110	m	54	110110	2
7	000111	H	23	010111	X	39	100111	n	55	110111	3
8	001000	I	24	011000	Y	40	101000	o	56	111000	4
9	001001	J	25	011001	Z	41	101001	p	57	111001	5
10	001010	K	26	011010	a	42	101010	q	58	111010	6
11	001011	L	27	011011	b	43	101011	r	59	111011	7
12	001100	M	28	011100	c	44	101100	s	60	111100	8
13	001101	N	29	011101	d	45	101101	t	61	111101	9
14	001110	O	30	011110	e	46	101110	u	62	111110	+
15	001111	P	31	011111	f	47	101111	v	63	111111	/
Padding		=									

轉換規則

將二進位資料轉換為 Base64 時，會以 3 個字元組為一組進行轉換。
為了方便說明，這裡使用字串作為二進位資料。

經過下表所示的轉換，字串 Man 的 Base64 為 TWFu。

Source	Text (ASCII)	M								a								n							
	Octets	77 (0x4d)								97 (0x61)								110 (0x6e)							
Bits		0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	1	0	1	1	0	1	1	1	0
Base64 encoded	Sextets	19								22								5							
	Character	T								W								F							
	Octets	84 (0x54)								87 (0x57)								70 (0x46)							

轉換規則

如果要編碼的位元組數不能被 3 整除，需先使用 0 位元組值在末尾補足，並加上一個或兩個 = 號，代表補足的位元組數。

Source	Text (ASCII)	M								a															
	Octets	77 (0x4d)								97 (0x61)															
Bits		0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	1	0	0						
Base64 encoded	Sextets	19								22								4							
	Character	T								W								E							
	Octets	84 (0x54)								87 (0x57)								69 (0x45)							

Source	Text (ASCII)	M																							
	Octets	77 (0x4d)																							
Bits		0	1	0	0	1	1	0	1	0	0	0	0												
Base64 encoded	Sextets	19								16								Padding							
	Character	T								Q								=							
	Octets	84 (0x54)								81 (0x51)								61 (0x3D)							

問題

請參考 Base64 轉換規則的說明，將輸入的字串轉換為 Base64 輸出。

- 範例輸入 1: CSIE CCU
- 範例輸出 1: Q1NJRSBDQ1U=
- 範例輸入 2: I see, I come, I conquer.
- 範例輸出 2: SSBzZWUsIEkgY29tZSwgSSBjb25xdWVyLg==