

Il sistema di conversione Base64 prende il suo nome dal fatto che 64 è la maggiore potenza di due che si può ottenere con il set di caratteri stampabili ASCII (ovvero 2^7).

Questo sistema di codifica è molto utilizzato in varie tecniche di trasmissione o memorizzazione dei dati specialmente per quanto riguarda alcuni ben noti servizi Internet come la posta elettronica ed il Web. Ad esempio, lo standard MIME (Multipurpose Internet Mail Extensions) utilizzato per l'invio di posta elettronica tramite il protocollo SMTP, prevede alcune possibili codifiche del messaggio (testo + allegati) tra cui la più comune è proprio la Base64.

Lo schema Base64 codifica sequenze di byte in altre sequenze binarie secondo un rapporto 4:3, ovvero per ogni 3 byte in ingresso la codifica ne produce 4. Il meccanismo è il seguente:

- vengono prelevati i primi 3 ottetti della sequenza da codificare
- viene creato un buffer della dimensione di 24 bit (3 byte)
- il primo ottetto viene posto nel buffer a partire dalla posizione più significativa e a questo seguono il secondo ed il terzo ottetto
- nel caso ci siano meno di 3 byte da codificare, si procede ad un riempimento del buffer con una serie di zeri (zero-fill)
- dal buffer vengono prelevate 4 sequenze binarie di 6 bit, che genereranno i 4 byte in uscita
- il processo si ripete per i secondi 3 ottetti e così via in modo iterativo fino alla fine della sequenza binaria
- se l'ultimo gruppo è composto da 1 o 2 byte, vengono aggiunti rispettivamente 1 o 2 segni di '='

NB. I caratteri ASCII considerati sono le lettere dell'alfabeto romano A-Z/a-z, i numeri 0-9 ed i due simboli + e /.

'A','B','C','D','E','F','G','H',
 'I','J','K','L','M','N','O','P',
 'Q','R','S','T','U','V','W','X',
 'Y','Z','a','b','c','d','e','f',
 'g','h','i','j','k','l','m','n',
 'o','p','q','r','s','t','u','v',
 'w','x','y','z','0','1','2','3',
 '4','5','6','7','8','9','+','/'

Esempio n.1

Testo decodificato: **Per**
 Testo codificato: **UGVy**

Codice ASCII di 'P'	80 (decimale)	=	01010000 (binario)
Codice ASCII di 'e'	101 (decimale)	=	01100101 (binario)
Codice ASCII di 'r'	114 (decimale)	=	01110010 (binario)

Buffer24bit = 01010000 01100101 01110010

1° gruppo di 6 bit	00 010100	=	20 (decimale)	=	U
2° gruppo di 6 bit	00 000110	=	6 (decimale)	=	G
3° gruppo di 6 bit	00 010101	=	21 (decimale)	=	V
4° gruppo di 6 bit	00 110010	=	50 (decimale)	=	y

Esempio n.2

Testo decodificato: **Pera**
 Testo codificato: **UGVyYQ==**

Codice ASCII di 'P'	80 (decimale)	=	01010000 (binario)
Codice ASCII di 'e'	101 (decimale)	=	01100101 (binario)
Codice ASCII di 'r'	114 (decimale)	=	01110010 (binario)
Codice ASCII di 'a'	97 (decimale)	=	01100001 (binario)

1° Buffer24bit = 01010000 01100101 01110010

1° gruppo di 6 bit	00 010100	=	20 (decimale)	=	U
2° gruppo di 6 bit	00 000110	=	6 (decimale)	=	G
3° gruppo di 6 bit	00 010101	=	21 (decimale)	=	V
4° gruppo di 6 bit	00 110010	=	50 (decimale)	=	y

2° Buffer24bit = 01100001 00000000 00000000

1° gruppo di 6 bit	00 011000	=	24 (decimale)	=	Y
2° gruppo di 6 bit	00 010000	=	16 (decimale)	=	Q
3° gruppo di 6 bit	00 000000	=	0 (decimale)	=	'='
4° gruppo di 6 bit	00 000000	=	0 (decimale)	=	'='