

Base64 Encoding

[RFC 2045](#) (sectie 6.8: Base64 Content-Transfer-Encoding)

[How Base64 encoding works](#)

[Base64 Encoding Table](#)

Base64 is een algoritme om binaire data te encoderen als ASCII tekst. Dit is noodzakelijk aangezien e-mail enkel kan omgaan met 7-bit ASCII karakters.

Door het feit dat Base64 3-bytes van de originele data indeelt in 4 ASCII karakters, zal het uiteindelijke ge-encodeerde resultaat ongeveer een derde ($\pm 33\%$) groter zijn.

Het algoritme:

Er worden telkens drie bytes (1byte = 8bits) van de binaire data genomen, deze 3 bytes worden na het algoritme weergegeven als 4 ASCII karakters.

Alleereerst worden de 3-bytes omgezet naar 4 6-bits nummers. Elk ASCII karakter bestaat uit 7-bits, waarvan Base64 er enkel 6 gebruikt ($2^6 = 64$ karakters in Base64 – $2^7 = 128$ karakters in ASCII), dit om te verzekeren dat de ge-encodeerde data universeel (in elke karakter) lees –en weergeefbaar is. De 64 karakters gebruikt in Base64 zijn a-z, A-Z, 0-9, / en +. Om precies te zijn is er ook nog een 65ste karakter het '=' teken, dit teken heeft een speciale functie.

Om te verzekeren dat de ge-encodeerde data geen mailserver's maximum lijngroote overschrijft wordt de lijnlengte na 76 karakters be-eindigd. (newline CRLF).

Niet elk bestand heeft een grootte die een veelvoud is van drie. In zulk geval zal op het einde voor het encodersproces de bytes die te weinig zijn vervangen worden door '0' bytes (00000000). Per ontbrekende byte wordt dan het padding teken gebruikt '='.

Voorbeeld:

De eerste 3-bytes van een jpg file zijn:

- Hex: FF D8 FF
- Decimaal: 377 330 377
- Binair: 11111111 11011000 11111111
- ASCII: ÿ Ø ÿ (y met 2 puntjes, doorstreept O en weer y met 2 puntjes)

De volgende stap in het omzetten is: deze 3 bytes omzetten in 4 groepjes van 6-bits.

- 111111 111101 100011 111111

Nu moeten deze bits omgezet worden naar decimale waarden

- 63 61 35 63

Na vergelijking van deze waarden met het Base64 alfabet (zie bijlage x) krijgen we:

- / 9 j /

PHP beschikt over ingebouwde functies die Base64 encoding en decoding voor hun rekening nemen nl. [base64_encode\(\)](#) en [base64_decode\(\)](#). Om compatibel te blijven met de 76 karakter lijnlengte (opgelegd in RFC2045) is er de [chunk_split\(\)](#) functie.

Value	Char	Value	Char	Value	Char	Value	Char
0	A	16	Q	32	g	48	w
1	B	17	R	33	h	49	x
2	C	18	S	34	i	50	y
3	D	19	T	35	j	51	z
4	E	20	U	36	k	52	0
5	F	21	V	37	l	53	1
6	G	22	W	38	m	54	2
7	H	23	X	39	n	55	3
8	I	24	Y	40	o	56	4
9	J	25	Z	41	p	57	5
10	K	26	a	42	q	58	6
11	L	27	b	43	r	59	7
12	M	28	c	44	s	60	8
13	N	29	d	45	t	61	9
14	O	30	e	46	u	62	+
15	P	31	f	47	v	63	/