

BLOQUE I

UT 2

6. BITS Y BYTES. LA MEDIDA DE LA INFORMACION

7. Sistemas de Codificación.

- Sistema de codificación las unidades de almacenamiento.
- Sistema de codificación las unidades de transmisión.

7.1. Sistema de codificación las unidades de almacenamiento.

Nombre	Abrev.	Factor binario	Tamaño en el SI
bytes	B	$2^0 = 1$	$10^0 = 1$
kilo	kb	$2^{10} = 1024$	$10^3 = 1000$
mega	Mb	$2^{20} = 1\,048\,576$	$10^6 = 1\,000\,000$
giga	Gb	$2^{30} = 1\,073\,741\,824$	$10^9 = 1\,000\,000\,000$
tera	Tb	$2^{40} = 1\,099\,511\,627\,776$	$10^{12} = 1\,000\,000\,000\,000$
peta	Pb	$2^{50} = 1\,125\,899\,906\,842\,624$	$10^{15} = 1\,000\,000\,000\,000\,000$
exa	Eb	$2^{60} = 1\,152\,921\,504\,606\,846\,976$	$10^{18} = 1\,000\,000\,000\,000\,000\,000$
zetta	Zb	$2^{70} = 1\,180\,591\,620\,717\,411\,303\,424$	$10^{21} = 1\,000\,000\,000\,000\,000\,000\,000$
yotta	Yb	$2^{80} = 1\,208\,925\,819\,614\,629\,174\,706\,176$	$10^{24} = 1\,000\,000\,000\,000\,000\,000\,000\,000$
Bronta	Bb	$2^{90} = 1237940039285380274899124224$	$10^{27} = 1\,000\,000\,000\,000\,000\,000\,000\,000\,000$
Geopbyte	Geb	$2^{100} = 1267650600228229401496703205376$	$10^{30} = 1\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000$

7.1. Sistema de codificación las unidades de almacenamiento.

- En 1998 fue creado un nuevo sistema de prefijos para denotar múltiplos binarios por la IEC.
- Oficialmente, el padrón IEC especifica que los prefijos del SI son usados solamente para múltiplos en base 10 (Sistema decimal) y nunca base 2 (Sistema binario).
- IEC: International Electrotechnical Commission
- ISO: International Organization for Standardization

CEI

Cantidad de bits

Prefijo SI

Prefijo Binario

Nombre
(Símbolo)

Estandar
SI

Nombre
(Símbolo)

Valor

kilobit (kbit)

10^3

2^{10}

kibibit (Kibit)

2^{10}

megabit (Mbit)

10^6

2^{20}

mebibit (Mibit)

2^{20}

gigabit (Gbit)

10^9

2^{30}

gibibit (Gibit)

2^{30}

terabit (Tbit)

10^{12}

2^{40}

tebibit (Tibit)

2^{40}

petabit (Pbit)

10^{15}

2^{50}

pebibit (Pibit)

2^{50}

exabit (Ebit)

10^{18}

2^{60}

exbibit (Eibit)

2^{60}

zettabit (Zbit)

10^{21}

2^{70}

zebibit (Zibit)

2^{70}

yottabit (Ybit)

10^{24}

2^{80}

yobibit (Yibit)

2^{80}

Prefijos CEI y SI con bit

Nombre	Símbolo	Sistema	Significado
bit	bit		0 ó 1
kibibit	Kibit	CEI	1024 bits
kilobit	kbit	SI	1000 bits
mebibit	Mibit	CEI	1024 kibibits
megabit	Mbit	SI	1000 kilobits
gibibit	Gibit	CEI	1024 mebibits
gigabit	Gbit	SI	1000 megabits
tebibit	Tibit	CEI	1024 gibibits
terabit	Tbit	SI	1000 gigabits
pebibit	Pibit	CEI	1024 tebibits
petabit	Pbit	SI	1000 terabits
exbibit	Eibit	CEI	1024 pebibits
exabit	Ebit	SI	1000 petabits

7.1. Sistema de codificación las unidades de almacenamiento.

- **ASCII**

- *American Standard Code for Information Interchange* (Código Estadounidense Estándar para el Intercambio de Información),
- Es un código de caracteres basado en el alfabeto latino.
- Creado en 1963 por el Comité Estadounidense de Estándares (ASA, conocido desde 1969 como el Instituto Estadounidense de Estándares Nacionales, o ANSI) como una refundición o evolución de los conjuntos de códigos utilizados entonces en telegrafía.
- En 1967, se incluyeron las minúsculas, y se redefinieron algunos códigos de control para formar el código conocido como US-ASCII.
- El código ASCII utiliza 7 bits para representar los caracteres, aunque inicialmente empleaba un bit adicional (bit de paridad) que se usaba para detectar errores en la transmisión.

7.1. Sistema de codificación las unidades de almacenamiento.

- **ASCII Extendido**

- Se denomina ASCII extendido a cualquier juego de caracteres de 8 bits en el cual los códigos 32 a 126 (0x20 a 0x7E) coinciden con los caracteres imprimibles de ASCII, así como los caracteres comúnmente llamados “de espacio”, estos son los códigos de control de 8 a 13 (0x08 a 0x0D), ambos inclusive.
- Las codificaciones de ASCII extendido utilizan además parte o la totalidad de los códigos superiores a 128 para codificar caracteres adicionales a los caracteres imprimibles ASCII.

7.1. Sistema de codificación las unidades de almacenamiento.

- **Unicode**

- Unicode es un estándar industrial cuyo objetivo es proporcionar el medio por el cual un texto en cualquier forma e idioma pueda ser codificado para el uso informático
- Unicode define dos métodos de “mapeo” o de localización de caracteres:
- La codificación UTF (*Unicode Transformation Format*) Formato de Transformación Unicode.
- La codificación UCS (*Universal Character Set*) Juego de Caracteres Universal.

7.1. Sistema de codificación las unidades de almacenamiento.

- **Unicode** , las codificaciones incluyen:
 - **UTF-7:** una codificación relativamente poco popular de 7 bits, a menudo considerada obsoleta.
 - **UTF-8:** una codificación de 8 bits de longitud variable
 - **UCS-2:** una codificación de 16 bits de longitud fija que solamente permite el “mapeo” o la búsqueda en la Plana Básica Multilengüe.
 - **UTF-16:** una codificación de 16 bits de longitud variable.
 - **UCS-4 y UTF-32:** un par de codificaciones de 32 bits de longitud fija que son funcionalmente idénticas.
 - **UTF-EBCDIC:** una codificación poco difundida creada para sistemas basados en EBCDIC.

DEC	HEX	OCT	CHAR	DEC	HEX	OCT	CH	DEC	HEX	OCT	CH	DEC	HEX	OCT	CH
0	0	000	NUL	32	20	040		64	40	100	@	96	60	140	`
1	1	001	SOH	33	21	041	!	65	41	101	A	97	61	141	a
2	2	002	STX	34	22	042	"	66	42	102	B	98	62	142	b
3	3	003	ETX	35	23	043	#	67	43	103	C	99	63	143	c
4	4	004	EOT	36	24	044	\$	68	44	104	D	100	64	144	d
5	5	005	ENQ	37	25	045	%	69	45	105	E	101	65	145	e
6	6	006	ACK	38	26	046	&	70	46	106	F	102	66	146	f
7	7	007	BEL	39	27	047	'	71	47	107	G	103	67	147	g
8	8	010	BS	40	28	050	(72	48	110	H	104	68	150	h
9	9	011	TAB	41	29	051)	73	49	111	I	105	69	151	i
10	A	012	LF	42	2A	052	*	74	4A	112	J	106	6A	152	j
11	B	013	VT	43	2B	053	+	75	4B	113	K	107	6B	153	k
12	C	014	FF	44	2C	054	,	76	4C	114	L	108	6C	154	l
13	D	015	CR	45	2D	055	-	77	4D	115	M	109	6D	155	m
14	E	016	SO	46	2E	056	.	78	4E	116	N	110	6E	156	n
15	F	017	SI	47	2F	057	/	79	4F	117	O	111	6F	157	o
16	10	020	DLE	48	30	060	0	80	50	120	80	112	70	160	p
17	11	021	DC1	49	31	061	1	81	51	121	Q	113	71	161	q
18	12	022	DC2	50	32	062	2	82	52	122	R	114	72	162	r
19	13	023	DC3	51	33	063	3	83	53	123	S	115	73	163	s
20	14	024	DC4	52	34	064	4	84	54	124	T	116	74	164	t
21	15	025	NAK	53	35	065	5	85	55	125	U	117	75	165	u
22	16	026	SYN	54	36	066	6	86	56	126	V	118	76	166	v
23	17	027	ETB	55	37	067	7	87	57	127	W	119	77	167	w
24	18	030	CAN	56	38	070	8	88	58	130	X	120	78	170	x
25	19	031	EM)	57	39	071	9	89	59	131	Y	121	79	171	y
26	1A	032	SUB	58	3A	072	:	90	5A	132	Z	122	7A	172	z
27	1B	033	ESC	59	3B	073	;	91	5B	133	[123	7B	173	{
28	1C	034	FS	60	3C	074	<	92	5C	134	\	124	7C	174	
29	1D	035	GS	61	3D	075	=	93	5D	135]	125	7D	175	}
30	1E	036	RS	62	3E	076	>	94	5E	136	^	126	7E	176	~
31	1F	037	US	63	3F	077	?	95	5F	137	_	127	7F	177	DEL

ASCII Extendido

128	Ç	144	É	160	á	176	☼	193	⊥	209	〒	225	β	241	±
129	ü	145	æ	161	í	177	☼	194	⊥	210	π	226	Γ	242	≥
130	é	146	Æ	162	ó	178	☼	195	⊥	211	ℓ	227	π	243	≤
131	â	147	ô	163	ú	179		196	—	212	ℓ	228	Σ	244	∫
132	ä	148	ö	164	ñ	180	⊥	197	⊥	213	ℓ	229	σ	245	∫
133	à	149	ò	165	Ñ	181	⊥	198	⊥	214	ℓ	230	μ	246	÷
134	â	150	û	166	ª	182	⊥	199	⊥	215	⊥	231	τ	247	≈
135	ç	151	ù	167	º	183	π	200	ℓ	216	⊥	232	Φ	248	◦
136	ê	152	—	168	¿	184	⊥	201	ℓ	217	⊥	233	⊕	249	.
137	ë	153	Ö	169	—	185	⊥	202	⊥	218	⊥	234	Ω	250	.
138	è	154	Û	170	¬	186	⊥	203	〒	219	■	235	δ	251	√
139	ï	156	£	171	½	187	⊥	204	⊥	220	■	236	∞	252	—
140	î	157	¥	172	¼	188	⊥	205	=	221	■	237	φ	253	²
141	ì	158	—	173	¡	189	⊥	206	⊥	222	■	238	ε	254	■
142	Ä	159	ƒ	174	«	190	⊥	207	⊥	223	■	239	∩	255	
143	Å	192	ℓ	175	»	191	⊥	208	⊥	224	α	240	≡		

7.1. Sistema de codificación las unidades de almacenamiento

- **EBCDIC**

- (**E**xtended **B**inary **C**oded **D**ecimal **I**nterchange **C**ode) es un código estándar de 8 bits usado por computadoras *mainframe* IBM.
- IBM adaptó el EBCDIC del código de tarjetas perforadas en los años 1960 y lo promulgó como una *táctica customer-control* cambiando el código estándar [ASCII](#).
- EBCDIC es un código binario que representa caracteres alfanuméricos, controles y signos de puntuación.
- Cada carácter está compuesto por 8 [bits](#) = 1 [byte](#), por eso EBCDIC define un total de 256 caracteres.

7.1. Sistema de codificación las unidades de almacenamiento

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
40			â	ä	à	á	ã	å	ç	ñ	[.	<	(+	!
50	&	é	ê	ë	è	í	î	ï	ì	í]	\$	*)	;	^
60	-	/	Â	Ä	À	Á	Ã	Å	Ç	Ñ		,	%	_	>	?
70	ø	É	Ê	Ë	È	Í	Î	Ï	Ì	Í	:	#	@	'	=	"
80	Ø	a	b	c	d	e	f	g	h	i	«	»	ô	ý	þ	±
90	°	j	k	l	m	n	o	p	q	r	°	°	æ	,	Æ	∞
A0	μ	~	s	t	u	v	w	x	y	z	ı	ç	Đ	Ý	Þ	®
B0	¢	£	¥	·	©	§	¶	¼	½	¾	¬		¯	˘	˙	×
C0	{	A	B	C	D	E	F	G	H	I		ô	ö	ò	ó	õ
D0	}	J	K	L	M	N	O	P	Q	R	¹	û	ü	ù	ú	ÿ
E0	\	÷	S	T	U	V	W	X	Y	Z	²	Ô	Ö	Ò	Ó	Õ
F0	0	1	2	3	4	5	6	7	8	9	³	Û	Ü	Ù	Ú	Ú

Tabla EBCDIC

7.2. Sistema de codificación las unidades de transmisión

Unidades de medida		
	bit	la más pequeña unidad: un valor de 1 o 0, en lenguaje Binario.
	byte	un conjunto de 8 bits.
	bps	bits por segundo.
	Kbps	kilobits por segundo = 1000 bits por segundo.
	Mbps	Megabits por segundo = 1,000,000 bits por segundo.

7.3. bit o Byte

- El IEEE 1541 especifica "b" (minúscula) como el símbolo para [bit](#);
- La [IEC](#) 60027 y el MIXF especifican "bit" (por ejemplo Mbit para megabit), teniendo la máxima desambiguación posible de byte. "b" vs. "B"
- Los países francófonos utilizan una "o" minúscula para "octeto": es posible referirse a estas unidades indistintamente como ko, Mo, o kB, MB.
- Esto no se permite en el [SI](#) por el riesgo de confusión con el cero, aunque esa es la forma empleada en la versión francesa del estándar ISO/IEC 80000-13:2008.

7.4. bytes

- http://www1.bipm.org/en/si/si_brochure/chapter3/prefixes.html
- <http://es.wikipedia.org/wiki/Byte>
- Convertir

<http://www.convertir-unidades.info/convertidor-de-unidades.php?tipo=bytes>

7.4.1. Diferencia entre el SI y CEI

http://es.wikipedia.org/wiki/Prefijos_binarios

8. Representación de la información en el ordenador

8.1. CODIFICACIÓN:

- La codificación consiste en representar los datos por medio de símbolos.
- El código es el conjunto de reglas que especifican la correspondencia entre los datos y su representación.

8. 1. CODIFICACIÓN

- Las cualidades de una representación son:
 - **Simplicidad**
 - **Precisión**
 - **Ser extensible**
 - **Ser significativa**
 - **Ser concisa**

8.2. Configuración Interna

- La codificación interna es la empleada por el ordenador para representar los datos internamente.
- Toda la información que se codifica en un ordenador digital tiene una característica común: Termina convertida en los valores conceptuales 0 y 1 (Sistema Binario).
- El elemento básico de la memoria de un ordenador son los llamados circuitos **biestables** y tienen la capacidad de dejar pasar corriente, o no dejarla Pasar, tal como hace un interruptor.
- Un biestable almacena un bit de información, codificado conceptualmente como 0 o 1.

8.2. Configuración Interna

- Booleanos.
- Numéricos: 0.....9 (Con diferentes rangos. Enteros, decimales, Reales, etc).
- Alfanuméricos:
 - Alfabéticos: A.....Z
 - Numéricos: 0.....9
 - C. Especiales: “ ,) , / , + , *
- Instrucciones para el ordenador.
- Otros significados estandarizados (Formatos gráficos, de audio, etc).
- Estructuras de datos: estáticas y dinámicas.