



RD1 Información y datos digitales



Información Digital

- **RD1 Información y datos digitales.**
- **RD2 Textos digitales.**
- **RD3 Audio digital.**
- **RD4 Imágenes digitales.**
- **RD5 Vídeo digital.**
- **RD6 Conceptos de compresión de datos.**

Información Digital

- **RD1 Información y datos digitales.**
 - RD1.1 Capacidad de información y caudal de datos.
 - RD1.2 Patrones de bits y códigos binarios.
 - RD1.3 Detección automática de errores.
 - RD1.4 Tipos de Información y archivos.
 - RD1.5 Concepto de compresión de datos.
- **RD2 Textos digitales.**
- **RD3 Audio digital.**
- **RD4 Imágenes digitales.**
- **RD5 Vídeo digital.**
- **RD6 Conceptos de compresión de datos.**

3

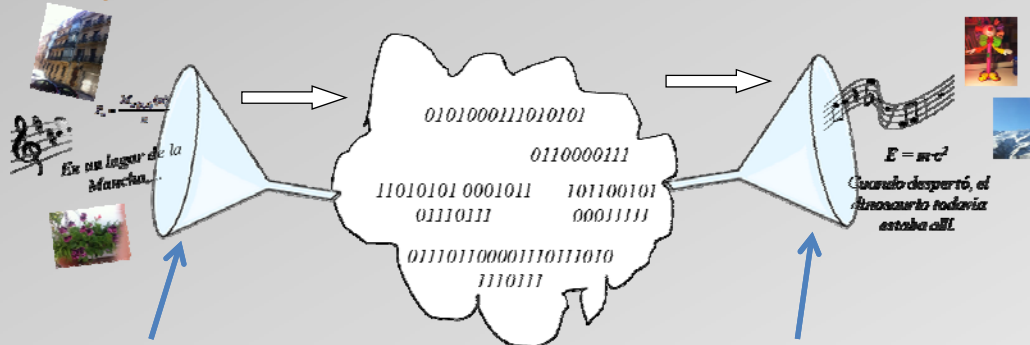


Información y datos digitales

RD1.1 CAPACIDAD DE INFORMACIÓN Y CAUDAL DE DATOS.



La información dentro de un sistema digital se codifica con un **código binario**; es decir, con **dos símbolos** que frecuentemente se representan por **cero y uno**.



En las **unidades de entrada**:

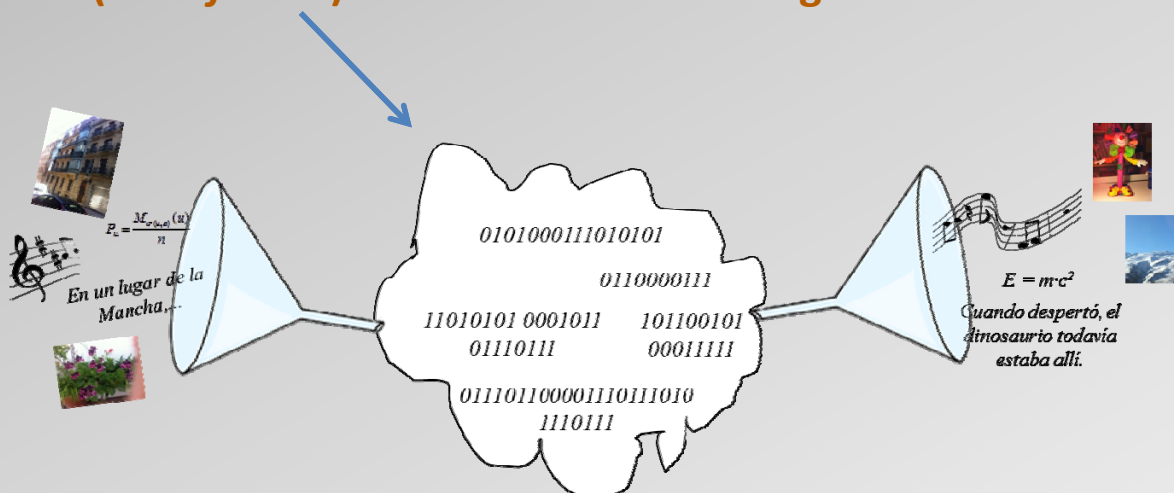
- Se transforma la información en señales eléctricas.
- Se digitaliza la información (codifica en binario).

En las **unidades de salida**:

- Se transforma la información digitalizada en información inteligible de acuerdo con un determinado formato.

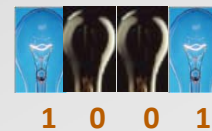
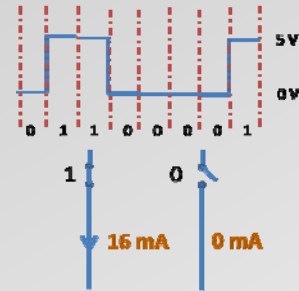


El sistema digital **almacena, transfiere y procesa (transforma)** la información en código binario



{0,1} no son números, sino símbolos que representan uno entre dos estados posibles, y dependiendo del soporte de información pueden representar:

- Un estado de magnetización: Norte o Sur (o cambio de estado): {N, S}
- Un nivel de tensión eléctrica: Alto o bajo: {H,L}
- Un valor de corriente, por ejemplo {16 mA, 0 mA}
- Un nivel de luz: {presencia de luz, ausencia de luz}



Unidades de información en sistemas digitales

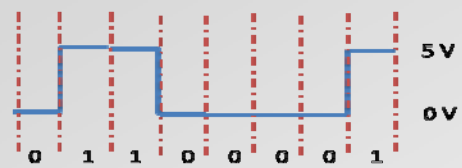
- **Bit** →
 - Unidad más elemental o capacidad mínima de información.
 - Es una posición o variable que toma el valor 0 ó 1.
- **Byte** →
 - En la actualidad se considera sinónimo de grupo de 8 bits.
 - (Históricamente: nº de bits necesarios para almacenar un carácter)



Ejemplos de bit y byte

- **Texto: Hay 24**

H → 0100 1000
a → 0110 0001
y → 0111 1001
SP → 0010 0000
2 → 0011 0010
4 → 0011 0100



Código ASCII



Múltiplos para capacidad de información (Bytes o bits) según la IEC - International Electrotechnical Commission

Prefijo	Prefijos binarios (IEC)
Exa (E)	2^{60}
Peta (P)	2^{50}
Tera (T)	2^{40}
Giga (G)	2^{30}
Mega (M)	2^{20}
Kilo (K)	$2^{10}=1.024$



Múltiplos según la SI (*International System of Units*)

Prefijo	Valor (SI)	Prefijos binarios (IEC)	
Exa (E)	10^{18}	2^{60}	Exbi (Ei)
Peta (P)	10^{15}	2^{50}	Pebi (Pi)
Tera (T)	10^{12}	2^{40}	Tebi (Ti)
Giga (G)	10^9	2^{30}	Gibi (Gi)
Mega (M)	10^6	2^{20}	Mebi (Mi)
Kilo (K)	10^3	$2^{10}=1.024$	Kibi (Ki)

— *Kilo byte binario (bi: binario): KibiByte*



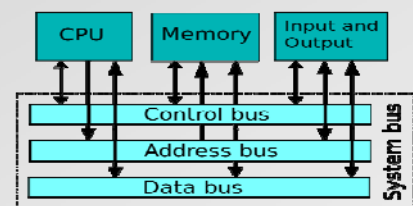
Ejemplos de capacidades (*usuales*)

- **Capacidad de la memoria principal**
 - Decenas de GB.
- **Capacidad de discos magnéticos y de estado sólido (SSD)**
 - Cientos GB, TB.
- **Tarjetas de memoria flash:**
 - Decenas y cientos de GB.
 - Tarjetas SD (foto/video)
 - MicroSD (dispositivos móviles)
 - CompactFlash (profesionales foto/video)
- **Flash USB**
 - Decenas y cientos de GB.



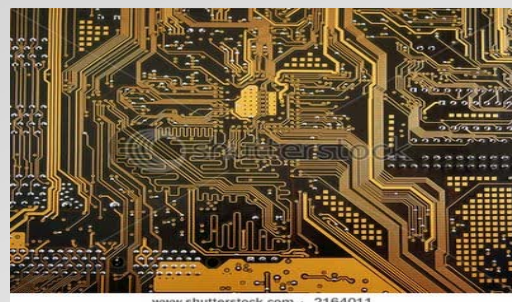
La información se transmite a través de “canales” de comunicación

- Usualmente (conexiones *alámbricas*) un canal está constituido por un conjunto de líneas (cables) o pistas conductoras de la electricidad que se denomina genéricamente **bus**.
- Los buses **interconectan** las distintas unidades del sistema.



Un bus paralelo está compuesto de varias líneas conductoras que transmiten simultáneamente varios bits

- Usualmente los buses paralelos son de 8, 16, 32, 64 o 128 líneas.
 - Es decir, transmiten a la vez, 1, 2, 4, 8 o 16 Bytes.
 - El nº de líneas de un bus se conoce como “**ancho del bus**”

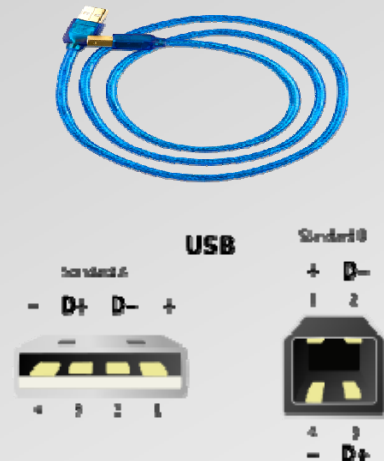


www.shutterstock.com · 2164011



Un bus serie transmite la información a través de una única línea, transfiriéndose los bits uno a uno.

- Hay buses serie que, en realidad contienen dos líneas con las que se transmiten los bits, uno a uno, en modo diferencial.
 - Con este modo se reducen considerablemente los efectos de las interferencias (ruido electromagnético)
- Ejemplo: Bus USB
 - Contiene 4 conexiones. Dos para la alimentación (5 voltios, y tierra), y dos para transmitir la información (bits) → transmisión **half-duplex**.



¿Cómo se mide la rapidez en la transmisión?

- Ancho de banda, caudal, tasa o velocidad de transferencia:
 - Cantidad de información que discurre en un determinado lugar por unidad de tiempo.
 - Ejemplos:
 - El ancho de banda del bus es de 8 MB/s.
 - Mi línea ADSL es de 50Mbits/s
 - En las conexiones serie se suele dar en bits/segundo
- El nombre más riguroso en español sería: **caudal de transferencia**



Conceptos a aprender en esta lección:

- **Código binario**
- **Funciones de un sistema digital:**
 - Almacenar, procesar y transferir información en un código binario.
- **Funciones de las unidades de entrada/salida**
- **Unidades de información:**
 - Bit y Byte
- **Múltiplos de capacidad de información:**
 - Kilo, Mega, Giga, Tera, etc.
- **Canal de información y bus.**
- **Ancho de banda, caudal de datos o tasa de transferencia**



Información y datos digitales

RD1.2 Patrones de bits y códigos binarios.



Conceptos básicos

- Un computador es una máquina que procesa, memoriza y transmite **información**.
- **Datos**
 - son conjuntos de símbolos utilizados para expresar o representar un valor numérico, un hecho, un objeto o una idea; en la forma adecuada para ser objeto de tratamiento.
- La información se utiliza principalmente bajo las formas de:
 - Textos
 - Sonidos
 - Imágenes
 - Valores numéricos
- La información se representa en el interior de la máquina de acuerdo con un **código binario**.

19



- **Patrón de bits:**
 - El código que representa un carácter.
 - La amplitud de una muestra de una señal de audio (voz, música, etc.)
 - La información de un punto de una imagen (píxel): código de un color
 - Una instrucción máquina, etc.

20



Con frecuencia se presenta la necesidad de codificar un conjunto α de m objetos o símbolos con un n° fijo, n , de bits

- ¿Cuántos bits necesitamos para codificar los m objetos de α ?
 - Con $n=2$ bits podemos hacer 4 combinaciones
 - $\beta \equiv \{00, 01, 10, 11\} = \{0,1\}^2$
 - Es decir, se pueden codificar hasta $m=4$ símbolos.
 - Con $n=3$ bits podemos hacer 8 combinaciones
 - $\beta \equiv \{000, 001, 010, 011, 100, 101, 110, 111\} = \{0,1\}^3$
 - Es decir, se pueden codificar hasta $m=8$ símbolos.
 - Con $n=4$ bits podemos hacer 16 combinaciones
 - $\beta \equiv \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111\} = \{0,1\}^4$
 - Es decir, se pueden codificar hasta $m=16$ símbolos.
- Generalizando
 - Con n bits podemos hacer m combinaciones.
 - $m = 2^n$
 - Siendo m el n° de elementos del conjunto α de objetos a codificar y $\beta \equiv \{0,1\}^n$ el conjunto de códigos posibles (combinaciones con repetición de dos elementos tomados de n en n).

21



Nº de bits mínimo de un código

- Obviamente para ocupar poco espacio interesa que el n° de bits, n , para codificar los m objetos sea el menor posible.
 - Como $m = 2^n$
 - n debe ser el **menor** n° entero que verifique:
 - $n \geq \log_2 m = 3.32 \log(m)$
- De esta forma la información introducida en el computador:
 - Ocupará el menor espacio posible

22



Información y datos digitales

RD1.3 Detección automática de errores.



Redundancias

- No obstante a que al definir un código se busque que tenga el menor nº de bits para ocupar el menor espacio posible, con frecuencia al patrón de bits que codifica un objeto o que compone un mensaje se le añaden bits adicionales (**bits redundantes**), con objeto de poder:
 - Detectar posibles errores en el **almacenamiento** o la **transmisión** (interferencias), e incluso,
 - Corregir posibles errores (Un ejemplo es el “Código Hamming”).

Patrón de n bits	k bits redundantes
--------------------	----------------------

Mensaje de n bits	k bits redundantes
---------------------	----------------------

24



Ejemplo de redundancia: letra de DNI

- Hacer la división entera del número de DNI por 23
- La letra asociada es la que corresponde al resto de la división anterior en la siguiente tabla

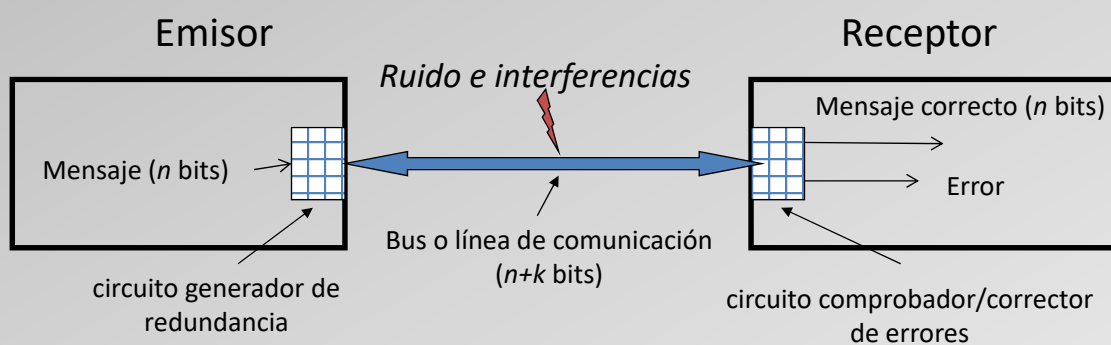
CÓDIGO PARA LA LETRA DEL D.N.I. O DEL N.I.F.																							
RESTO	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
LETRA	T	R	W	A	G	M	Y	F	P	D	X	B	N	J	Z	S	Q	V	H	L	C	K	E

- Con este sistema se detectan automáticamente:
 - El 100% de los errores simples
 - El 100% de los errores por transposición.
 - No se detectan los errores si la diferencia entre el nº original y el erróneo es múltiplo de 23 (probabilidad 0,043). Es decir, se detecta el 95,7% de posibles errores

25



Detección de errores en grabación o transmisión



Ejemplos: Línea de transmisión de datos o comunicación entre dos unidades del computador (CPU y disco; CPU e impresora)

26



Bit de paridad se añade un bit (0 ó 1) de forma tal que el número total de unos de código que resulte sea par o impar.

Mensaje o código inicial

100 0001 →
101 1011 →
101 0000 →
110 1000 →

Bit de paridad
(criterio par)

Mensaje con redundancia

0 100 0001
1 101 1011
0 101 0000
1 110 1000

Bit de paridad
(criterio impar)

1 100 0001 →
0 101 1011 →
1 101 0000 →
0 110 1000 →

27



Otros procedimientos de detección e incluso corrección de errores

☐ Tipos de errores:

- Errores en bits aislados
- Errores en bits múltiples
- Errores en ráfagas

☐ Métodos de detección de errores

- Comprobación de redundancia vertical (VCR, Vertical Redundancy Checking)
- Comprobación de redundancia horizontal (LCR, Longitudinal Redundancy Checking)
- Suma de comprobación (Checksum)
- Redundancias cíclicas (CRC, Cyclic Redundancy Checking)

• **Página 678 del libro de texto**

28



Detección de errores CRC-CITT (*Comité Consultatif International Téléphonique et Télégraphique*)

- Se demuestra que utilizando el polinomio generador en una codificación polinómica (CRC)
 - CRC-CCITT $\rightarrow G(x) \equiv x^{16} + x^{12} + x^5 + 1$
- se detectan:
 - todos los errores de 1 bit individual,
 - todos los errores que alteren dos bits,
 - todos los errores que alteren un número impar de bits,
 - todos los errores de ráfagas de 16 o menos bits,
 - el 99.997% de errores en ráfagas de 17 bits, y
 - el 99.998% de errores en ráfagas de 18 o más bits.

29



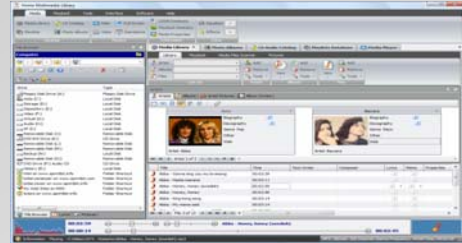
Información y datos digitales

RD1.4 Tipos de Información y archivos.



La información que se proporciona a un computador se suele dar bajo las formas de:

- Textos
- Sonidos
- Imágenes y videos
- Valores numéricos



- Los sistemas multimedia combinan textos, sonidos e imágenes.

31



Un archivo (*file*) o fichero es un conjunto de información sobre un mismo tema, tratada como un todo desde el punto de vista del sistema operativo.

- Son las unidades o contenedores de información para almacenar datos o programas de forma permanente en dispositivos de memoria externa o para enviar a través de líneas de datos.
- Ejemplos:
 - Una carta o documento.
 - Archivo de Alumnos de clase (DNI, nombres, apellidos, calificaciones, etc.)
 - Una foto.
 - Un video.
 - Un programa.
 - Movimientos de una cta. cte. bancaria etc

32



Información y datos digitales

RD1.5 Concepto de compresión de datos.



COMPRESIÓN DE DATOS

- **Diversas aplicaciones (multimedia, etc.) requieren utilizar archivos de gran capacidad. Lo que implica elevados:**
 - volumen para su almacenamiento.
 - tiempo de transmisión a través de buses de comunicación de datos.
- **Reducción: transformación denominada compresión de datos.**
 - El archivo, antes de ser almacenado o transmitido se comprime mediante un **algoritmo de compresión de datos**, y
 - cuando se recupera para procesarlo o visualizarlo se aplica la técnica inversa para descomprimirlo.



34



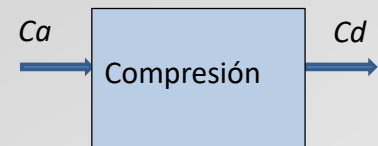
Factor de compresión:

- Denominando C_a y C_d a las capacidades del archivo antes y después de comprimirlo,

- factor de compresión como:

$$f_c = \frac{C_a}{C_d}$$

- “factor de compresión de $f_c : 1$ ”.



- porcentaje de compresión: tanto por ciento que queda de la capacidad original:

$$p_c = \frac{C_d}{C_a} \cdot 100 \%$$

.

D / 35



Supongamos que un archivo de música de 14 MB después de comprimirse ocupa 5 MB.

- Factor de compresión:

$$f_c = \frac{C_a}{C_d} = \frac{14 \text{ MB}}{5 \text{ MB}} = 2,8$$

- Factor de compresión de 2,8 a 1 \rightarrow 2,8:1

- Porcentaje de compresión:

$$p_c = \frac{C_d}{C_a} \cdot 100 = \frac{5 \text{ MB}}{14 \text{ MB}} \cdot 100 = 35,7\%$$

36



Tipos y algoritmos de compresión de datos

- **Sin perdidas:**
 - En la descompresión se puede recuperar exactamente el documento original
 - Codificación por secuencias (RLE)
 - Codificación relativa o incremental
 - Codificación dependiente de la frecuencia
 - Codificación con diccionario adaptativo
 - Codificación Lempel-Ziv LZ77
- **Con perdidas**
 - En la descompresión no se puede recuperar exactamente el documento original
 - Compresión MP3 (sonidos)
 - Compresión GIF (imágenes)
 - Compresión JPEG (imágenes)
 - Compresión MPEG (imágenes)
- **Mas detalles:**
 - págs. 124 a 128 del libro de texto

37



Resumen y conclusiones

- **Nociones básicas sobre representación de la información**
 - Datos, patrones de bits y archivos.
 - Detección automática de errores.
 - Formas de información.
 - Concepto de compresión de datos.

38

