

-Si un módulo DRAM tiene chips en ambas caras se denomina módulo con 2 rangos.

Falso. El rango es el número de palabras completas que devuelve la memoria DRAM.

-Un módulo de DRAM DDR4 con ECC 2Rx4 tiene 36 chips.

Verdadero. Cada palabra tiene 72 bits (hay que sabérselo de memoria) genera 2 palabras y por cada chip 4 bits. Por tanto, $N^{\circ}\text{Chips} = (2 \times 72) / 4 = 36$.

-Las tarjetas que usaban las ranuras ISA tenían que ser pre-configuradas para asignarles la interrupción que correspondía a cada una.

Verdadero. Las tarjetas que utilizaban las ranuras ISA (Industry Standard Architecture) a menudo requerían pre-configuración manual para asignarles la interrupción (IRQ) adecuada, así como las direcciones de E/S (Entrada/Salida) (Cortesía de Chat GPT)

-Podemos mejorar la escalabilidad de un servidor utilizando RAID 1.

Falso. RAID 1 genera redundancia, lo cual es muy útil ante pérdida de datos, no obstante, si queremos aumentar la escalabilidad buscamos añadir nuevas funcionalidades al servidor lo que no está relacionado con RAID 1.

-Si una interfaz es Plug and Play podemos cambiar componentes en caliente (sin reiniciar).

Falso.

Diferencia principal:

Plug and Play (PnP)

Definición: Plug and Play es una característica que permite a un sistema operativo detectar y configurar automáticamente dispositivos de hardware sin necesidad de intervención manual del usuario, como configurar direcciones de E/S, IRQs, o canales DMA.

Básicamente conectar un USB y que el ordenador nos lo detecte sin tener que hacer un mount manualmente.

Hot Swapping (Cambio en Caliente)

Definición: El cambio en caliente, o hot swapping, se refiere a la capacidad de añadir, remover o reemplazar componentes de hardware sin tener que apagar o reiniciar el sistema. Esta característica es crucial en entornos donde la disponibilidad continua es esencial, como servidores y equipos de red.

Esta es la característica a la que se refiere el enunciado. IMPORTANTE: No confundir.

-Un SSD con MLC es menos fiable que uno SLC.

Verdadero.

SLC(Single level cell): Almacenan un bit por celda.

MLC(Multiple level cell): Almacenan múltiples bits por celda.

El profesor comentó en clase que estos últimos tienen más escrituras por celda, lo que aumenta el desgaste considerablemente, y que en muchas ocasiones había fabricantes que

vendían discos con más celdas de la cuenta, porque se presuponía que un porcentaje elevado (10% aprox) se iban a romper.

-Un disco SAS usa un único conector que unifica tanto los cables de alimentación procedentes de fuente de alimentación como los de datos, procedentes de la placa base.

Verdadero. Como noción general SATA está más orientado al público general y SAS para Servidores. SAS es compatible con SATA pero no al revés.

-Los chips de la serie AMD Opteron X son SoC.


En teoría verdadero. No sé para qué quiere saber este tipo de cosas.

-AGP tiene más ancho de banda que PCI.

Verdadero.

Interfaces PCI y AGP

- **PCI** (*Peripheral Component Interconnect*). Intel.
 - Bus paralelo de 32 o 64 bits. Las líneas se comparten entre todos los dispositivos PCI.
 - Half-duplex.
 - *Plug and Play*.
 - Ancho de banda:
 - 33MHz, 32b (4B) → 133MBps.
 - 66MHz, 32b (4B) → 266MBps.
 - 66MHz, 64b (8B) → 533MBps.
 - Versión PCI-X → SERVIDORES::
 - 64b (8B), 133MHz ⇒ \cong 1GBps.
- **AGP** (*Accelerated Graphics Port*). Intel.
 - Bus paralelo de 32 bits.
 - Half-duplex.
 - Uso: tarjeta gráfica.
 - Ancho de banda (AGP 8x): 2GBps.



-Con la Ley de Amdahl podemos estimar la ganancia en velocidad de la ejecución de un único trabajo en un computador después de mejorar k veces un componente.

Verdadero. Es la aplicación directa que tiene la ley.

-Las prestaciones de un servidor no se ven afectadas por la configuración de la memoria virtual.

~~Verdadero.~~ Viene escondido en la diapositiva 20 del tema 1 y además suena coherente.

Falso

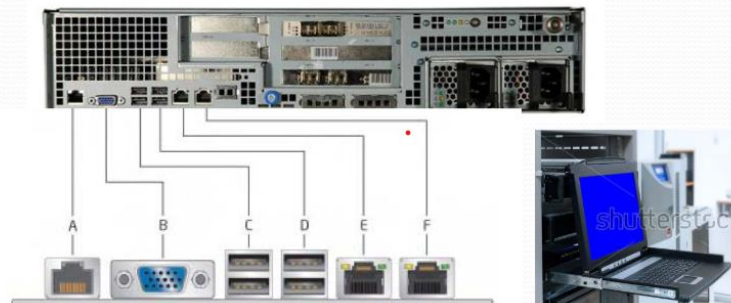
-La ganancia en velocidad siempre es un número comprendido entre 0 y 2, ambos inclusive.

Entiendo que quiere que pongas Falso. Ya que es un número entre 0 y 1 pero un número entre 0 y 1 es un número entre 0 y 2. Debería formularse mejor la pregunta.

-Es poco frecuente que las placas de los servidores tengan conectores VGA en el panel trasero.

Entiendo que es Falso. Es cierto que en servidores actuales se usan menos los VGA pero en las diapositivas del tema 2 página 45 hay un servidor Intel® Server Board S5520UR. Una imagen en la que se muestra que tiene un conector VGA. Este servidor se comercializó a lo largo de 2014-2016. En consecuencia, todos los anteriores deberían de tener VGA, por lo que debería de ser frecuente encontrar placas de este tipo.

Conectores del panel trasero de un servidor (Intel® Server Board S5520UR)



A	Serial Port A	D	Dual USB Port Connector
B	Video	E	NIC Port 1 (1 Gb)
C	Dual USB Port Connector	F	NIC Port 2 (1 Gb)

Dos puertos Ethernet (uno para el interior de la empresa usado mantenimiento y otro para Internet)

45

-El protocolo USB 2.0 es half-duplex

Verdadero. Viene en la diapositiva 47 del tema 2.

-El puente sur del chipset es el encargado de las transferencias entre el puente norte y el resto de periféricos con menores exigencias de velocidad de la placa.

Verdadero. Diapositiva 50 tema 2.

Juego de chips (chipset)

- El chipset es el conjunto de circuitos integrados (chips) de la placa base encargados de controlar la comunicación entre los diferentes componentes de la placa base.
- Un chipset se suele diseñar para una familia específica de microprocesadores.

El juego de chips suele estar distribuido en dos componentes principales:

- El **puente norte** (north bridge), encargado de las transferencias de mayor velocidad (principalmente con el microprocesador, la memoria, la tarjeta gráfica y el puente sur).
- El **puente sur** (south bridge), encargado de las transferencias entre el puente norte y el resto de periféricos con menores exigencias de velocidad de la placa.



-En un test t, el valor-p equivale esencialmente a la probabilidad de que la hipótesis nula sea falsa.

Falso. Es un valor que se toma de referencia. Básicamente, es el valor de tolerancia estadística cuanto mayor p-valor, menos probabilidad tenemos de que nuestra hipótesis (sea verdadera o falsa) sea lo contrario en el caso real.

-La hipótesis de partida de un test ANOVA es que el factor que se está estudiando influye en el rendimiento.

Falso. Se presupone que el factor no influye en el rendimiento.

-Con “sar -u”, sin más argumentos, iremos obteniendo la información de la utilización global del procesador desde el momento actual en adelante.

Falso. Para obtener datos en tiempo real se han de proporcionar argumentos adicionales a la orden. Por ejemplo, sar -u 1 5 mostrará estadísticas de utilización del procesador cada segundo durante cinco segundos.

-Un microbenchmark también suele denominarse benchmark de sistema completo.

Falso.

- Microbenchmarks o benchmarks para **componentes**: estresan componentes o agrupaciones de componentes concretos del sistema: procesador, caché, memoria, discos, red, procesador+caché, procesador+compilador+memoria virtual, etc.
- Macrobenchmarks o benchmarks de sistema **completo** o de **aplicación real**: la carga intenta imitar situaciones reales (normalmente servidores con muchos clientes) típicas de algún área. P.ej. comercio electrónico, servidores web, servidores de ficheros, servidores de bases de datos, sistemas de ayuda a la decisión, paquetes ofimáticos + correo electrónico + navegación, etc.

-Si la suma de las utilizaciones de los dispositivos de un servidor es mayor que 1, el servidor está saturado.

Falso. En la diapositiva 64 del tema 5 tenemos un ejemplo donde la suma es mayor que 1 y el servidor no está saturado. En general, no tiene que ver puede ocurrir que todos los recursos del servidor se utilicen durante un periodo elevado de tiempo. Pero que el servidor no esté saturado.

a)

Recurso	V_i	S_i (s)	D_i (s)	U_i
CPU	9	0,010	0,09	0,45
DISCO	3	0,020	0,06	0,30
RED	5	0,016	0,08	0,40

-Decimos que un servidor es fiable si se encuentra operativo.

Falso. Un servidor está disponible si se encuentra operativo.

-Podemos aumentar la escalabilidad de un servidor añadiendo fuentes de alimentación redundantes.

Falso. Hay que diferenciar dos conceptos escalabilidad y expansibilidad. Escalabilidad es la capacidad de ampliar las funcionalidades de un servidor, expansibilidad es la capacidad de añadirle cosas CPU'S, RAM'S, fuentes de alimentación. Para poder lograr la escalabilidad frecuentemente necesitamos expansibilidad, pero expansibilidad no implica escalabilidad.

-Los procesadores para servidores suelen tener más canales de memoria que los de CPU en la misma generación.

Verdadero. Es de sentido común, los procesadores de los servidores procesan más datos, muchas veces en paralelo, para procesar más datos tienen que llegar más datos, para que te lleguen más datos necesitas más canales.



-Las unidades de estado sólido (SSD) son capaces de alcanzar anchos de banda superiores a los que el protocolo SATA-3 puede proporcionar.

Verdadero. Hoy en día hay dos modelos de discos SSD unos para SATA y otros para PCIe. Los de SATA se adaptan estrictamente al ancho de banda de SATA-3 mientras que los de PCIe he leído que pueden llegar a 3500MB/s.

	SATA I	SATA II	SATA III
Frecuencia	1500 MHz	3000 MHz	6000MHz
Bits/clock	1	1	1
Codificación 8b10b	80%	80%	80%
bits/Byte	8	8	8
Velocidad real	150 MB/s	300 MB/s	600 MB/s

-El primer procesador con un conjunto de instrucciones de 64 bits lo diseñó AMD a principios de los 2000.

Falso. Tenemos en las diapositivas (bastante escondido he de decir) que POWER de IBM diseñó una arquitectura de 64 bits en 1996.

→ Valen 2M €


- POWER = *Performance Optimization With Enhanced RISC*. Resultado del trabajo conjunto entre Apple, IBM y Motorola para servidores de gama alta de muy altas prestaciones por vatio, disponibilidad y fiabilidad (**mainframes**).
Esos servidores los usaban por ejemplo las empresas de la NSA...
la siempre encierran en un cuarto que no dan errores

Major POWER® Innovation

- 1990 RISC Architecture
- 1994 SMP
- 1995 Out of Order Execution
- 1996 64 Bit Enterprise Architecture
- 1997 Hardware Multi-Threading

→ NO es X86-64 de Intel (la arquitectura propia de IBM de 64 bits)

- Ejemplo: Power 10 (2020)
 - Núcleos (cores): 15.
 - Cada núcleo puede ejecutar hasta 8 hilos en paralelo.
 - 16 micros máx / placa base



-La 'S' del acrónimo SRAM es por Synchronous.

Falso. Significa static, no obstante, no confundir con SDRAM que ahí sí significa Synchronous.

-Un módulo de RAM con chips a los dos lados tiene dos rangos.


Falso. Ya explicada, el rango es las palabras de 64 bits que puede generar la memoria.

-Puedo conectar una tarjeta PCIe 2.0 en un conector PCIe 4.0 de la placa base, pero no al revés.

Falso. PCIe está diseñado para que sea perfectamente compatible entre versiones, eso sí siempre tendremos la velocidad del más lento, ya que solo es compatible.

-PCIe 2.0 x4 tiene un ancho de banda de 2GB/s en cada sentido.

Verdadero.

- Características (continuación):
 - Codificación: 8b/10b (1.x 2.x), y 128b/130b (3.0, 4.0, 5.0).
 - Versiones y velocidades (por cada sentido de cada LANE):
 - PCIe 1.1: hasta 2,5GT/s (250 MBps),
 - PCIe 2.0: hasta 5GT/s (500 MBps)
 - PCIe 3.0: hasta 8GT/s (~1GBps)
 - PCIe 4.0: hasta 16GT/s (~2GBps)
 - PCIe 5.0: hasta 32GT/s (~4GBps)
 - Velocidad negociable: Se puede usar una tarjeta con una versión de PCIe diferente a la de la placa base.
 - Fácilmente escalable: x1, x2, x4, x8, x16. 
 - PCIe x16: uso en tarjetas gráficas:
 - PCIe x16 (4.0) : hasta 32GBps en cada sentido.
 - PCIe x16 (5.0) : hasta 64GBps en cada sentido.
 - Puedo usar una tarjeta con más o menos LANES que el conector PCIe de la placa base.

Tenemos que sabernos esos datos, no es difícil, ya que el primero es 250 MBps y el resto es ir multiplicando por 2. Como es PCIe 2.0 x4 tiene 4 lanes es decir una velocidad total de

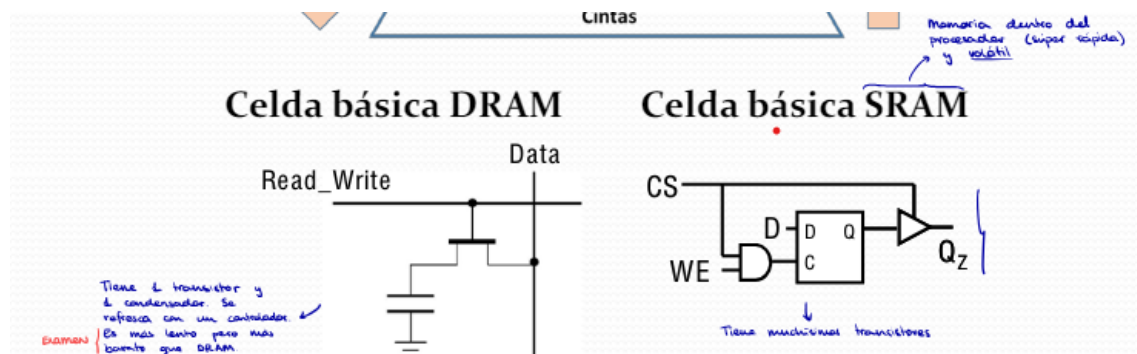
$500 \text{ MB/s} * 4 = 2 \text{ GB/s}$.

-Puedo conectar un disco SATA usando un controlador SAS de la placa, pero no un disco SAS usando el controlador SATA de la placa.

Verdadero. Un disco SATA se puede conectar con SAS pero no al revés por la forma que tiene el conector.

-Una celda básica SRAM tiene más componentes electrónicos que una DRAM.

Verdadero. Una DRAM tiene un condensador y un transistor. La SRAM tiene un biestable que tiene muchos más componentes electrónicos.



-AMD ofrece los procesadores EPYC para el mercado de sobremesa y OPTERON para servidor.

Falso. Opteron es la denominación antigua, mientras que EPCY es la denominación actual.

Procesadores de AMD para servidores

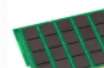
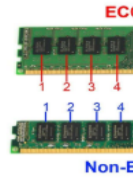
- La familia de procesadores de AMD para servidores fue inicialmente denominada *Opteron*.
 - El primer *Opteron*, presentado en 2003, fue el primer procesador con el conjunto de instrucciones AMD x86-64.
 - En 2004, los *Opteron* fueron los primeros procesadores x86 con 2 núcleos.
- Recientemente, AMD ha modificado el nombre de sus procesadores para servidores, denominándolos EPYC.

-SO-DIMM es un tipo de memoria pensada para servidores.

Falso. Es un tipo para PC hay U-DIMM y SO-DIMM la SO es más pequeña.

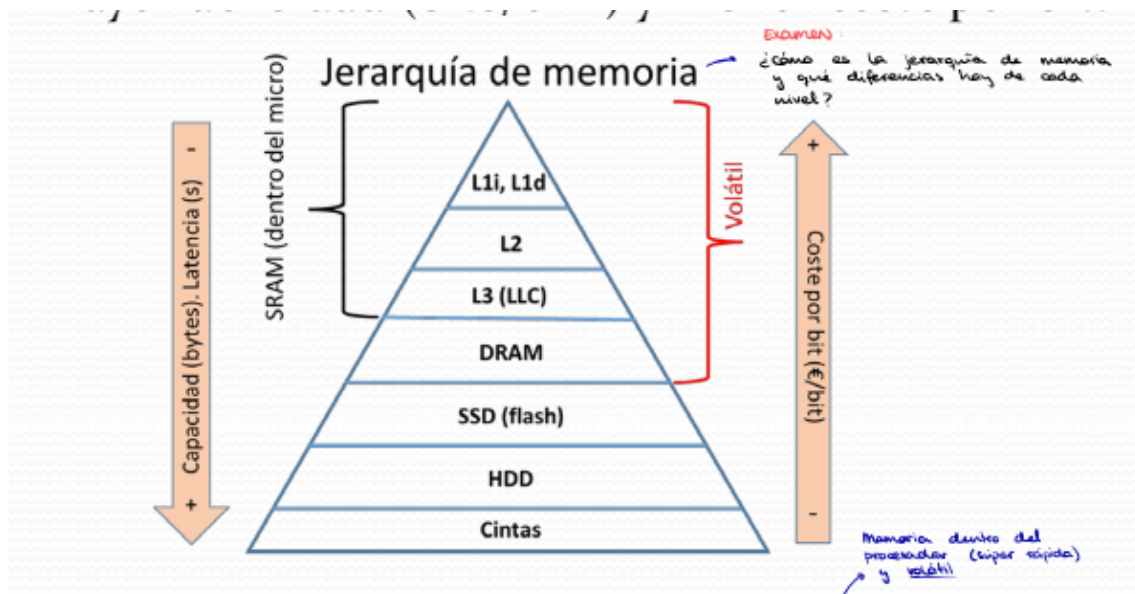
Tipos de DIMM para una tecnología dada

- Para PC y portátiles:
 - DIMM ó U-DIMM: Unbuffered (ó Unregistered) DIMM.
 - SO-DIMM: Small Outline DIMM. Tamaño más reducido para equipos portátiles (tienen menos contactos).
- Para servidores:
 - EU-DIMM: U-DIMM con Error Correcting Code, ECC (mayor fiabilidad).
 - R-DIMM: Registered DIMM. Hay un registro que almacena las señales de control (operación a realizar, líneas de dirección...). Mayor latencia que EU-DIMM pero permiten módulos de mayor tamaño. Tienen ECC.
 - LR-DIMM: Load Reduced DIMM. Hay un buffer que almacena tanto las señales de control como los datos a leer/escribir. Mayor latencia que R-DIMM, pero son las que permiten los módulos con mayor tamaño. Tienen ECC.



-En la jerarquía de memoria de un computador, la memoria LLC tiene menor latencia que la DRAM.

Verdadero. Hay que saberse la pirámide de las jerarquías, hay que recordar que LLC es la caché de nivel 3 es decir la L3.



-El conjunto de instrucciones que ejecutan el auto-test de arranque (power on self-test) se encuentran almacenadas en la ROM/Flash BIOS de la placa.

Verdadero. La parte donde viene es random.

ROM/Flash BIOS (Basic I/O System)

- Almacena el código de arranque (*boot*) del computador. Este código se encarga de identificar los dispositivos instalados, instalar drivers básicos para acceder a los mismos, realizar el **Power**-on self-test (POST) del sistema e iniciar el S.O.

-Es el propio microprocesador de muchas placas bases actuales el que realiza la función de puente norte del chipset.

Verdadero. No viene escrito en ninguna parte, pero buscando en internet puedes encontrar que la mayoría de los dispositivos actuales usan SoC (System on a chip) tecnología donde sucede lo que comenta el enunciado.

-Para diagnosticar si una memoria RAM funciona correctamente debemos utilizar un benchmark específico para memorias RAM.

Falso. Los benchmark están diseñados para medir las prestaciones, no si un dispositivo funciona correctamente.

-Si $N_T \gg N_T^*$ el servidor dejará de estar en equilibrio de flujo (para el caso de redes cerradas).

Falso. Una red cerrada siempre está en equilibrio de flujo.

Ley del tiempo de respuesta interactivo

- Un servidor en una red cerrada siempre está en equilibrio de flujo (siempre supondremos que el tamaño de las colas es suficientemente grande, en este caso, $\geq N_T$).
- Al ser una red cerrada, el número total de trabajos (=clientes) en la red ($N_T = N_Z + N_0$), es constante.

-Las principales medidas de prestaciones se basan en tiempos de respuesta (o latencias) y en productividades (o anchos de banda).

Verdadero. Ya que cumplen varios de los criterios de ser una buena medida para cuantificar prestaciones.

-Todos los sistemas escalables son extensibles pero no a la inversa.

Verdadero. Ya comentada, escalable ampliar funcionalidades para ampliar funcionalidades se necesita ser extensible. Pero al revés no.

-Xeon es una familia de procesadores de IBM especialmente dirigida a servidores.

Verdadero. Como novedad viene en una esquina de una diapositiva perdida del tema 2.

Falso

Principales fabricantes de microprocesadores para servidores: Intel, AMD e IBM.

Ejemplo: Intel **Xeon** para servidores vs Intel **Core** para PC (desktop) (<http://ark.intel.com>)

-La familia "AMD Opteron X Series" usa microprocesadores ARM

Falso. No sé que tipo de microprocesadores usa pero ARM es de 32 bits y AMD es de 64 así que no.

-La celda básica de una SRAM es mayor que la de una DRAM.

Verdadero. El hecho de tener más componentes hace que sea más grande.

-El voltaje que usa un módulo DDR4 es menor que el de un módulo DDR3.

Verdadero. En las memorias RAM los voltajes son decrecientes conforme se aumenta de versión

pero no se pueden cambiar en la vida

	Nº contactos	Voltaje (V)	Bus datos (half-duplex)	Ancho de banda típico (GB/s)
SDRAM	168	3,3	32b	1,3
DDR	184	2,5	32b	3,2
DDR2	240	1,8	64b	8,5
DDR3	240	1,5	64b	17,1
DDR4	288	1,2	64b	25,6
DDR5	288	1,1	32b+32b ↳ la divide en 2 porque tiene 2	38,4

-Las latencias de las unidades de cinta suelen ser muy bajas, ya que hay que rebobinar la cinta hasta que el cabezal se encuentre en la posición deseada.

Falso. Las latencias serán muy altas ya que precisamente hay que rebobinar la cinta.

Unidades de cinta (tape drives)



- Almacenan la información de forma permanente (no volátil) a través de una cinta recubierta de material magnético que se enrolla por medio de carretes.
- Las latencias suelen ser muy altas ya que hay que rebobinar la cinta hasta que el cabezal se encuentre en la posición deseada.
- Es el medio con la mayor densidad de bits para un área dada. Actualmente, permiten almacenamiento de decenas de TB por cinta y velocidades de lectura secuencial en torno a 150 MBps.
- Es el medio de almacenamiento masivo más barato.
- Se usan normalmente como almacenamiento de respaldo (*backup*) y archivado.

-AHCI es una interfaz para facilitar la conexión SSD a través de PCIe.

Falso. AHCI es una interfaz de SATA, me he leído 1000 veces el tema 2 y siempre salen cosas nuevas, como esta.

-SAS es full-duplex.

Verdadero.

- **SAS (Serial Attached SCSI).**
 - Conexión **serie** punto a punto.
 - Reloj embebido en los datos.
 - Full-duplex. *Hot-plug*. Codificación 8b/10b.
 - Frecuencias de 3, 6 y 12 GHz: hasta 1200 MB/s (1,2 Gpbs).
 - Compatible con discos SATA.
 - Mayores voltajes que SATA. Longitud del cable máxima: 10m.

-A través del "system panel" se puede conectar el altavoz del chasis a la placa base.

Verdadero. Esto sí que creo que no viene en ninguna parte, para explicar un poco hay una serie de pines en el system panel. Los más comunes son:

- **PWR SW (Power Switch):**

- **Función:** Conecta el botón de encendido del chasis. Permite encender y apagar la computadora.
- **Ubicación típica:** Dos pines etiquetados como "PWR SW" o "PW".

- **RESET SW (Reset Switch):**

- **Función:** Conecta el botón de reinicio del chasis. Permite reiniciar la computadora.
- **Ubicación típica:** Dos pines etiquetados como "RESET SW" o "RST".

- **PWR LED (Power LED):**

- **Función:** Conecta los LEDs que indican si la computadora está encendida.
- **Ubicación típica:** Dos o tres pines etiquetados como "PWR LED" o "PLED". A veces puede haber un pin para el LED positivo (+) y otro para el negativo (-).

- **HDD LED (Hard Drive Activity LED):**

- **Función:** Conecta el LED que indica la actividad del disco duro.
- **Ubicación típica:** Dos pines etiquetados como "HDD LED" o "HDLED". Normalmente, un pin es positivo (+) y el otro es negativo (-).

- **SPEAKER (PC Speaker):**

- **Función:** Conecta el altavoz del chasis, que emite pitidos para indicar el estado del hardware y errores durante el arranque.
- **Ubicación típica:** Cuatro pines, pero usualmente solo se utilizan dos de ellos (uno para positivo y otro para negativo), etiquetados como "SPEAKER" o "SPK".

-La pila que hay en la placa base sirve, entre otras cosas, para tener el reloj en tiempo real de dicha placa actualizado.

Verdadero. También sirve para guardar la configuración del BIOS/UEFI.

-Existen servidores con fuentes de alimentación reemplazables en caliente (hot swappable).

Verdadero.

-El puente sur del chipset se encarga de la comunicación con la DRAM.

Falso. Regla general, cosas rápidas RAM, CPU, GPU con el puente norte, cosas lentas con el puente sur, se incluyen PCIe y todo lo que sea más lento que él (prácticamente todo).

-En saturación, el cuello de botella está al máximo de su productividad.

Verdadero. Si no estuviera al máximo significa que puede incrementar la productividad, entonces, no estaría saturado, ya que se podría aumentar la carga del servidor.

-Si la suma de todas las utilizaciones es mayor que 1 el servidor está saturado.

Falso. Ya explicada, hay ejemplos en las diapositivas en los que pasa esa situación.

- $R_0 = R_1 + R_2 + \dots + R_k$.

Falso. Viene escrito en las diapositivas.

Nótese que, en general: $R_0 \neq R_1 + R_2 + \dots + R_K = \sum_{i=1}^K R_i$

-La demanda media de servicio de un dispositivo en el seno de un servidor nunca puede ser menor que su tiempo de servicio.

Falso. $D_i < S_i$. Nos llevaría que $V_i < 1$. Puede parecer cierta, pero un mismo trabajo puede necesitar varios accesos al mismo servicio por tanto V_i puede ser mayor que 1, de hecho, tenemos ejemplos.

Dispositivo	V_i	S_i (s)	D_i (s)
CPU	5	0,02	0,1
DISCO A	4	0,02	0,08
DISCO B	5	0,01	0,05
CPU 2	5	0,02	0,1

-Si ejecutamos la línea “sar -d” en un servidor con sar instalado, iremos obteniendo información sobre las transferencias de cada disco del servidor de forma interactiva.

Falso. Nos proporciona la información que estamos buscando pero no es interactiva para ello tendríamos que escribir algo del tipo: `cat -d 1 10`. Nos da la actividad del disco durante 10 segundos.

-C0 es un valor medio calculado durante un periodo de monitorización T.

Falso. C0 no es una media es un valor absoluto. El número de trabajos completados por el servidor.

-gprof basa su información en lo que obtiene leyendo de /proc

Falso. Gprof en función del programa que compile y ejecute se crea sus propios archivos y saca la información de ahí.

-En una placa de un PC es habitual encontrar tanto conectores SATA como SAS.

Falso. En un PC lo habitual es SATA, ya que está enfocado al público en general. En un servidor lo habitual es SAS

-En un test t, el valor p obtenido depende del nivel de confianza seleccionado.

Falso. El valor no depende del nivel de confianza seleccionado, la aceptación o la refutación de la hipótesis sí.

-La versión serie de la interfaz SCSI se llama M.2, también conocida como NGFF.

Falso. La interfaz M.2 es conocida como NGFF pero no es de SCSI es un NVMe para conectar con PCIe adjunto la diapositiva.

NVMe: Non-Volatile Memory Express

- Es un protocolo para el acceso a SSD conectadas a través de PCIe.
- PCIe x4, 4GBps
- M.2 NVMe (NGFF). Usa internamente PCIe x4, 4GBps
- U.2. Usa internamente PCIe x4, 4GBps
- SATA 3.2 (SATA Express). Combina PCIe y SATA, 2GB/s



-El conjunto de instrucciones que ejecutan el auto-test de arranque (power On self-test) se encuentran almacenadas en la ROM/Flash BIOS de la placa.

Verdadero. Ya vista.

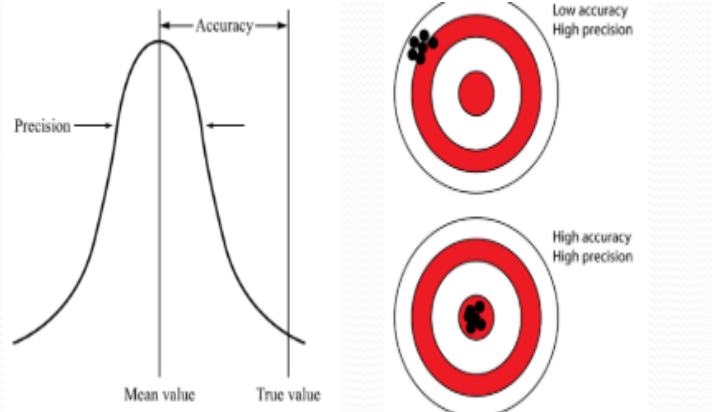
-La **precisión** de un sensor se refiere esencialmente a cómo se aproximan las medidas que toman al valor verdadero.

Falso.

Exactitud: Capacidad de acertar al objetivo. Lo que se desliza la media de los valores tomados del valor real.

Precisión: la diferencia entre dos valores tomados es muy pequeña. Mirar diapositiva.

- **Exactitud** del sensor (*Accuracy, offset*): ¿Cómo se aleja el valor medido del valor real que se quiere medir?
- **Precisión** del sensor: Cuando se mide varias veces el mismo valor real, ¿se mide siempre lo mismo?, ¿cuál es la dispersión de las medidas?



- **Resolución** del sensor: ¿Cuánto tiene que cambiar el valor a medir para detectar un cambio?

-Tanto PCI como PCIe definen un protocolo de comunicación serie punto a punto.

Falso. PCI es un Bus que se comparte entre todos los dispositivos PCI. PCIe sí tiene comunicación en serie punto a punto.

- **PCI** (*Peripheral Component Interconnect*). Intel.
 - Bus paralelo de 32 o 64 bits. Las líneas se comparten entre todos los dispositivos **PCI**.
 - Half-duplex.
 - *Plug and Play*.
 - Ancho de banda:
 - 33MHz, 32b (4B) → 133MBps.
 - 66MHz, 32b (4B) → 266MBps.
 - 66MHz, 64b (8B) → 533MBps.
 - Versión PCI-X → SERVIDORES:
 - 64b (8B), 133MHz ⇒ ≈ 1GBps.



Interfaz PCI-Express (PCIe)

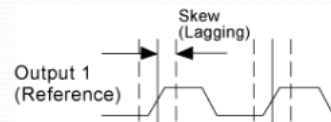
- Características:
 - Conexión serie punto a punto (no es un bus con líneas compartidas) por medio de varias "LANES".
 - Cada LANE está compuesta por 4 cables, 2 por cada sentido de la transmisión. Full-Duplex.
 - *Hot plug*.
 - Transmisión SÍNCRONA estando el [reloj embebido en los datos](#).

-Los protocolos de comunicación serie pueden ser half-duplex

Verdadero. Le sucede a USB 2.0.

Ventajas de usar una interfaz **serie** con reloj embebido con respecto a una paralela con reloj común

- Mayor frecuencia de reloj ya que evita el “desfase” entre las señales (*timing skew*): no todas las pistas recorren la misma distancia por lo que algunas señales llegan antes que otras.
- Menor nº de pistas para un rendimiento similar.
- Mayor facilidad para obtener conexiones full duplex. Pero ¡ojo!, no todas las interfaces serie con reloj embebido son full dúplex (USB 2.0 NO es full duplex).



Universal Serial Bus (USB)

- **USB 2.0**
 - Conexión serie. Reloj embebido en los datos.
 - 4 pines: 2 datos (diferencial), alimentación y tierra.
 - Ancho de banda: hasta **480Mbps** (60MBps).
 - Hasta 127 dispositivos. Hot plug. Half-duplex.

-Cuando nos referimos al índice de prestaciones que se calcula según el benchmark SPEC CPU2017, el SPEC pico (peak) nunca puede ser menor que el SPEC base.

Verdadero. Los nombres en este caso son intuitivos el peak es siempre el más alto. En consecuencia, más alto que el base.

Índices de prestaciones de SPECspeed®2017

- También llamados, de forma genérica, **índices SPEC**:
 - Aritmética entera: CPU2017IntegerSpeed_**peak**, CPU2017IntegerSpeed_**base**.
 - Aritmética en coma flotante: CPU2017FP_Speed_**peak**, CPU2017FP_Speed_**base**.
- Significado de “base” y “peak”:
 - Base: Compilación en modo conservador: todos los programas escritos en el mismo lenguaje usan las mismas opciones de compilación.
 - Peak: Rendimiento pico, permitiendo que cada uno escoja las opciones de compilación óptimas para cada programa.

-Tanto las SRAM como las DRAM son volátiles, pero solo las DRAM necesitan refresco.

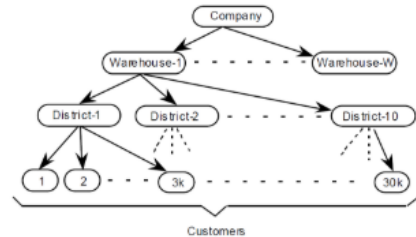
Verdadero. Las DRAM necesitan ser refrescadas (volver a escribir los bits almacenados en la misma posición) mientras que las SRAM no lo requieren.

-En los benchmarks TPC-C, TPC-H, el índice de prestaciones se calcula a partir de la media geométrica de las ganancias de velocidad con respecto a una máquina de referencia.

Falso.

Ejemplo: Benchmark TPC-C

- Es de tipo **OLTP** (*on-line transaction processing*). Simula una gran compañía que vende 100.000 productos y que tiene varios almacenes (configurable), cada uno a cargo de 10 zonas, con 3000 clientes/zona.
- Las peticiones (nuevos pedidos, pagos, envíos, estado de un pedido, existencias) involucran acceso a las bases de datos tanto locales como distribuidas (a veces el producto no está en el almacén más cercano), ejecución simultánea de consultas y acceso no uniforme a las bases de datos.
- **Índice de rendimiento utilizado (métrica):** transacciones que de media se han procesado por segundo, superando unos ciertos requisitos de tiempos de respuesta (el 90% deben tener un tiempo de respuesta inferior a 5s). También suelen proporcionar tanto el consumo de potencia como el coste por transacción procesada (incluido mantenimiento de 3 años).



Ejemplo: Benchmark TPC-H


- Es de tipo **DS** (*decision support*). Contiene un total de 22 tipos de consultas diferentes que requieren examinar grandes volúmenes de datos para poder contestar a preguntas complejas. También incluyen modificaciones concurrentes de los datos. Algunos de estos tipos de consultas son:
 - Realizar un informe detallado sobre las ventas en un determinado periodo de tiempo.
 - Buscar el proveedor más adecuado según un conjunto determinado de criterios.
 - Cuantificar el aumento de ingresos que habría resultado de eliminar ciertos descuentos en toda la empresa en un rango de fechas determinado.
- Aunque es escalable, la base de datos cuenta con la información de un mínimo de 10000 proveedores con un mínimo de 10 millones de filas.
- **Índice de rendimiento utilizado (métrica):** *TPC-H Composite Query-per-Hour Performance Metric (QphH@Size)*, donde *Size* es el tamaño de la base de datos utilizada (factor de escala). Es una fórmula compleja que tiene en cuenta principalmente el número medio de consultas atendidas por unidad de tiempo. También suelen proporcionar tanto el consumo de potencia como el coste por *QphH@Size* (incluido mantenimiento de 3 años).

- $Q_i = X_0 * W_i$.

Falso. Es $Q_i = X_i * W_i$

-La transmisión de información entre un módulo de memoria de tipo DDR4 y la CPU es half-duplex.

Verdadero. La DRAM es half-duplex.



En cada ciclo de reloj se lee o se escribe para no se produzcan conflictos en la Vb2

	Nº contactos	Voltaje (V)	Bus datos (half-duplex)	Ancho de banda típico (GB/s)
SDRAM	168	3,3	32b	1,3
DDR	184	2,5	32b	3,2
DDR2	240	1,8	64b	8,5
DDR3	240	1,5	64b	17,1
DDR4	288	1,2	64b	25,6
DDR5	288	1,1	32b+32b	38,4

↳ se divide en 2 porque tiene 2 canales, 18Gb

-La expresión $W_i = N_i \cdot S_i$ es una ley operacional.

Falso. Carece de sentido simplemente.

-Es la media aritmética y no el índice SPEC el que nos ayuda a saber qué computadora ejecuta el conjunto total de programas benchmark en menos tiempo.

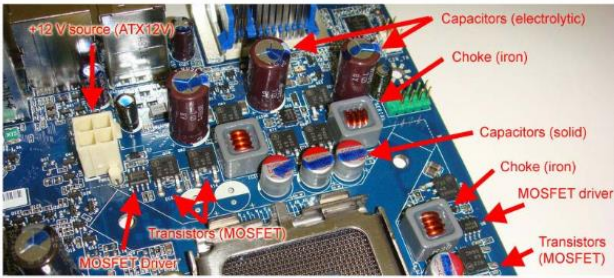
Verdadero. Tiene sentido, si su suma es la menor la media será la menor.

-El módulo regulador de voltaje, entre otras cosas, convierte la corriente alterna en corriente continua.

Falso. Adapta corrientes continuas.

El módulo regulador de voltaje (VRM)

- VRM (Voltage Regulator Module): Adapta la tensión continua de la fuente de alimentación (5V, 12V) a las tensiones continuas menores que necesitan los diferentes elementos de un computador (CPU, memoria, chipset, etc.), dándoles también estabilidad.



Labels in the image: +12 V source (ATX 12V), Capacitors (electrolytic), Choke (iron), Capacitors (solid), Choke (iron), MOSFET driver, Transistors (MOSFET), MOSFET driver, Transistors (MOSFET).

-Si el número total de usuarios de un servidor modelado mediante una red de colas cerrada interactiva es superior a N_T^* , entonces el servidor está saturado.

Falso. A partir de este punto no se puede conseguir el tiempo de respuesta mínimo, pero no significa que el servidor esté en saturación.

Punto teórico de “saturación” (knee point)

- Es el valor de N_T en donde las asíntotas coinciden:

$$D = D_b \times N_T^* - Z \Rightarrow N_T^* = \frac{D + Z}{D_b}$$

- Propiedades del punto teórico de “saturación” N_T^* :
 - Para un número total de trabajos $N_T > N_T^*$, los límites en las prestaciones vienen impuestos únicamente por el cuello de botella del servidor.
 - A partir de N_T^* trabajos ya no se puede conseguir el tiempo de respuesta mínimo ya que se empiezan a formar colas de espera en, al menos, el dispositivo cuello de botella (en la práctica, esto sucede bastante antes).
 - En principio, podría parecer el número ideal de trabajos en la red ya que, al menos teóricamente, para $N_T = N_T^*$ se podría conseguir la productividad máxima y el tiempo de respuesta mínimo absolutos del servidor (en la práctica esto nunca se puede conseguir de forma simultánea): $N_T^* = \underbrace{X_o^{max}}_{\Delta / o_o} \times \underbrace{(R_o^{min} + Z)}_0 = \frac{D+Z}{D_b}$

-Las pistas en una placa base están hechas de una resina no conductora y no inflamable.

Falso. Importante están hechas de un material no conductor y no inflamable, pero las pistas son de cobre.

¿Qué es una placa base?

- Una **placa base** (o **placa madre**, *motherboard*, *mainboard*) es la tarjeta de circuito impreso (PCB, *Printed Circuit Board*) principal de un computador.
- En ella se conectan los componentes hardware del computador y contiene diversos conectores para añadir distintos periféricos adicionales.
- Una PCB, en general, está hecha de una lámina de un substrato no conductor (normalmente fibra de vidrio con una resina no inflamable) sobre la que se extienden pistas de cobre (material conductor).



-La expresión $Q_i = \lambda_{dai} * W_i$ es válida, aunque el servidor esté saturado.

Verdadero. La tasa de entrada de trabajos por el tiempo de espera de cada trabajo nos dará el número de trabajos en espera. No influye el hecho de que esté saturado.

-Jugando con las distintas configuraciones de un RAID se puede conseguir más fiabilidad o más disponibilidad pero no mayores prestaciones.

Falso. Los raids pueden aportar más fiabilidad y disponibilidad, ya que al tener raid 5 por ejemplo podemos recuperar los datos si hay algún problema. Y mayores prestaciones, ya que

cambiar de raid puede acelerar las lecturas o escrituras en disco a cambio de perder por otro lado. Si no recuerdo mal Raid 5 era más rápido que raid 1 pero menos fiable.

-SATA-3 usa codificación 128b/130b.

Falso. Este tipo de preguntas me parecen criminales. Usa Codificación 8b/10b.

-Sar es un monitor software por eventos.

Falso. Va tomando muestras cada x tiempo, no toma muestras únicamente cuando hay un evento.

-La pasta térmica es un aislante que evita que el calor salga de la CPU.

Falso. Si se lee detenidamente la frase es absurda, siempre quieres que la CPU tenga la temperatura baja, ¿para qué querías que el calor se quede dentro?

-El fenómeno llamado "timing skew" motivó la aparición de protocolos paralelos como P-ATA (también llamado IDE) o PCI.

Falso. Explicación el timing skew.

El término "timing skew" (desfase de temporización o sesgo de temporización) se refiere a la diferencia en el tiempo de llegada de dos o más señales que deberían llegar simultáneamente. Este fenómeno es especialmente relevante en sistemas digitales y redes de comunicación, donde la sincronización precisa de las señales es crucial.

Para solucionarlo se implementa la temporización en serie, los protocolos PCI o P-ATA no la implementan, en consecuencia, no tiene sentido que sea verdadera.

-Las primeras instrucciones que ejecuta un procesador en el arranque proceden de las primeras direcciones de la DRAM.

Falso. La DRAM es volátil, por tanto, la información que contiene tras el encendido es basura.

-Una de las líneas del protocolo PCIe se usa en exclusiva para transmitir la señal del reloj.

Falso.

En los sistemas PCIe, no hay una línea dedicada exclusivamente para transmitir la señal del reloj, como en algunos otros sistemas de interconexión más antiguos. En lugar de ello, PCIe utiliza una técnica llamada "clock data recovery" (CDR), que permite que el reloj sea extraído directamente de las señales de datos transmitidas.

-eDonkey es un ejemplo de Arquitectura Cliente/Servidor de varios niveles.

Falso.

La información sobre eDonkey como un ejemplo de arquitectura Cliente/Servidor de varios niveles no es del todo correcta. eDonkey, junto con su red eDonkey2000, es en realidad un ejemplo de arquitectura de red P2P (Peer-to-Peer), que es diferente de la arquitectura Cliente/Servidor. (chat-gpt4)

5.- (1 punto) Cuestiones.

- ¿Para qué se usa una "rack unit" (1U) en el mundo de los servidores? (0,25 puntos)
- ¿Qué diferencia hay entre los conceptos de precisión y exactitud cuando hablamos de la medida realizada por un sensor? (0,25 puntos)
- ¿Cómo obtiene `gprof` información sobre el número de veces que se ha ejecutado cada función de un programa? ¿Es un valor estimado o exacto? (0,25 puntos)
- Indique las principales características de `perf` (0,25 puntos).

Ejercicio 5:

a) Diapositiva 55 del tema 2. Es una unidad de medida utilizada para estandarizar la distancia entre estantes en un armario (rack) de computadores. De esa forma, los chasis de un computador se suelen diseñar para tener un número entero de "rack units" y así facilitar que puedan acoplarse a dichos armarios.

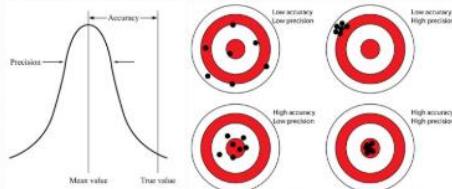
b) Diapositiva 12 del tema 3.

Atributos que caracterizan a un sensor/monitor

- **Exactitud de la medida (Accuracy, offset):** ¿Cómo se aleja el valor medido del valor real que se quiere medir?

- **Precisión (Precision):** ¿Cuál es la dispersión de las medidas?

- **Resolución del sensor:** ¿Cuánto tiene que cambiar el valor a medir para detectar un cambio?



- **Tasa Máxima de Entrada (Max Input Rate):** ¿Cuál es la frecuencia máxima de ocurrencia de los eventos que el monitor puede observar? (monitores por eventos)

- **Anchura de Entrada (Input Width):** ¿Cuánta información (p.ej. nº de bytes) se almacena por cada medida que toma el monitor?

c) Diapositiva 37 del tema 3. Importante matizar que `gprof` instrumenta el programa en la compilación para que cada vez que se ejecute una función del programa se ejecuten unas instrucciones que incrementan el contador asignado a dicha función. Es un valor exacto. Otra cosa bien distinta es cómo obtiene `gprof` el tiempo de cpu de cada función (usando un temporizador...), pero eso no se pregunta aquí.

d) Diapositiva 43 del tema 3.

Otros profilers: **Perf**

- `Perf` es un conjunto de herramientas para el análisis de rendimiento en Linux basadas en eventos software y hardware (hacen uso de contadores hardware disponibles en los últimos microprocesadores de Intel y AMD). Permiten analizar el rendimiento de **a)** un hilo individual, **b)** un proceso + sus hijos, **c)** todos los procesos que se ejecutan en una CPU concreta, **d)** todos los procesos que se ejecutan en el sistema. Algunos de los comandos que proporciona:

6.- (0,75 puntos) Responda brevemente a las siguientes cuestiones sobre el benchmark CPU 2017 que ha desarrollado el consorcio SPEC:

- ¿Qué componentes del sistema informático evalúa? (0,25 puntos)
- Indique cómo se calcula el índice CPU2017IntegerSpeed_peak (tanto de palabra como poniendo la fórmula). El método de cálculo empleado, ¿satisface todas las exigencias de un buen índice de prestaciones? Razone la respuesta. (0,5 puntos)

- ¿Qué componentes se evalúan? Procesador (enteros o coma flotante según el caso), sistema de memoria y compilador (C, Fortran y C++).
- Cálculo: Cada programa del benchmark se ejecuta 3 veces y se escoge el resultado intermedio (se descartan los 2 extremos). El índice SPEC es la media geométrica de las ganancias en velocidad con respecto a una máquina de referencia (en SPEC CPU2017 una Sun Fire V490).
- Ejemplo: Si llamamos t_i al tiempo que tarda la máquina a evaluar en ejecutar el programa de benchmark i -ésimo y t_i^{REF} lo que tardaría la máquina de referencia para ese programa (y hay 10 programas en el benchmark):

$$\text{índice SPEC} = \sqrt[10]{\frac{t_1^{REF}}{t_1} \times \frac{t_2^{REF}}{t_2} \times \dots \times \frac{t_{10}^{REF}}{t_{10}}}$$

A esta respuesta hay que añadir que el test verifica:

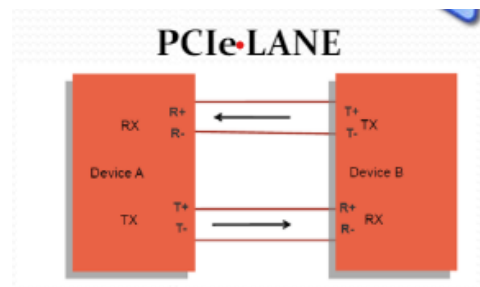
- Repetibilidad: Sí, siempre que se realice se obtienen los mismos resultados.
- Representabilidad: sí y no, ya que solo se evalúan el procesador, el sistema de memoria y compilador, probablemente en estas componentes sea mejor pero puede que en el resto empeore.
- Consistencia: Debería, ya que es un benchmark de referencia.
- Facilidad de medición: Supongo.
- Linealidad: Supongo.

7.- (0,75 puntos) Cuestiones (0,25 puntos cada una).

- ¿A qué nos referimos por "LANE" cuando hablamos de la interfaz PCIe?
- ¿Para qué sirve el system panel en una placa base?

Interfaz PCI-Express (PCIe)

- Características:
 - Conexión serie punto a punto (no es un bus con líneas compartidas) por medio de varias "LANES".
 - Cada LANE está compuesta por 4 cables, 2 por cada sentido de la transmisión. Full-Duplex.
 - Hot plug.
 - Transmisión SÍNCRONA estando el [reloj embebido en los datos](#).



El system panel es una parte crucial de la interacción entre el usuario y el hardware del sistema. Proporciona una interfaz física a través de la cual el usuario puede controlar y monitorear el funcionamiento del sistema sin necesidad de acceder directamente a la placa base o al interior del chasis. Esto facilita la operación y el mantenimiento del sistema, así como la resolución de problemas durante el arranque y el funcionamiento del sistema.

- ¿Qué es una "rack unit", 1U? ¿Para qué se usa? (0,25 puntos)
- ¿Por qué no se ha seguido subiendo en los últimos años la frecuencia de la CPU a pesar de las mejoras en la tecnología? (0,25 puntos)
- ¿Qué son las memorias LR-DIMM? (0,25 puntos)
- Demuestre que el tiempo medio de respuesta de un servidor en baja carga se aproxima a D. (0,25 puntos)
- Demuestre que la asíntota de alta carga del tiempo medio de respuesta de un servidor modelado mediante una red de colas cerrada interactiva viene dada por $D_b \cdot NT \cdot Z$. (0,5 puntos)
- Suponga que usted es el ingeniero jefe de una importante entidad pública y que quiere instalar una sala de servidores con las características que usted considere necesarias para poder proporcionar servicios de streaming de vídeo. Indique las principales secciones que debe contener el pliego de prescripciones técnicas para licitar un contrato para dicha instalación junto con una frase explicativa del tipo de información que debe contener cada una de dichas secciones principales. (0,5 puntos)

Se x_0 es máx $\left(\frac{1}{\rho_b}\right) \rightarrow R_0 \xrightarrow{\text{trazda}} D_b \cdot NT - Z$