edix
Digital Workers

UF-3 CLOUDS, VIRTUALIZACIÓN Y DOCKERS

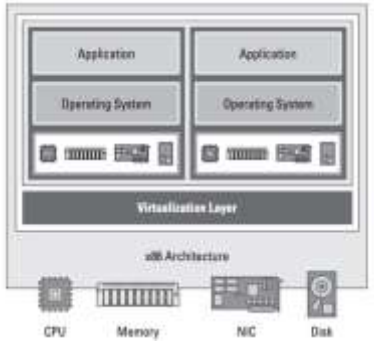
Profesor Raúl Salgado Vilas

- Virtualización: Un paso más allá a la hora de utilizar los recursos físicos es el de la virtualización: aprovechar los recursos disponibles en un ordenador, o en una red de ordenadores, para generar ‘máquinas virtuales’ que los aíslen y generen un entorno de ejecución o procesamiento.

- La virtualización tiene varios usos comunes y lo cierto es que todos ellos giran en torno al concepto de que su tecnología representa una abstracción de los recursos físicos.

- Hay dos tipos de virtualización basada en hipervisores:

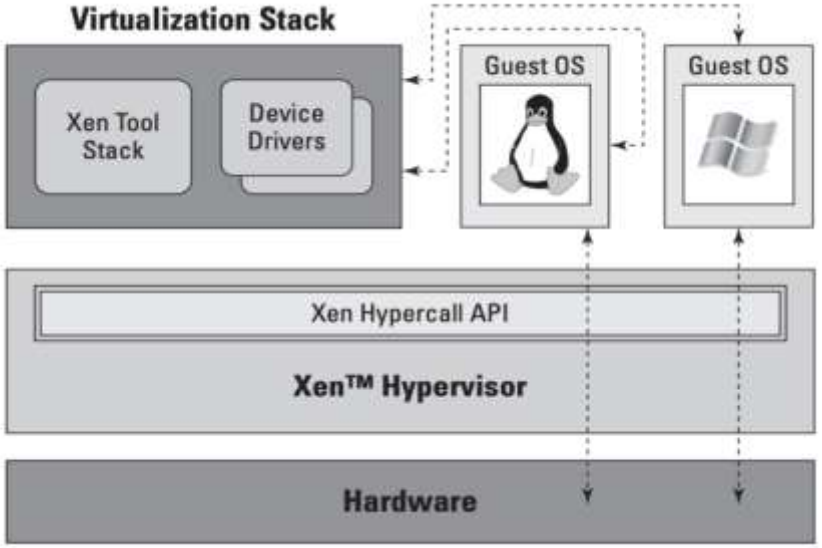
  - ✓ Imitación o emulación del hardware: En este caso, el software de virtualización (hipervisor) crea una máquina virtual que imita todo el entorno de hardware. El sistema operativo que está cargado en una máquina virtual es un producto estándar no modificado. Cuando realiza llamadas para recursos del sistema, el software de emulación de hardware captura la llamada del sistema y la redirige para que pueda gestionar estructuras de datos proporcionadas por el hipervisor. Es el propio hipervisor el que realiza las llamadas al hardware físico real, subyacente a toda la aglomeración de software:

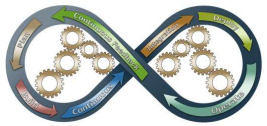


- Hay dos tipos de virtualización basada en hipervisores:

  - ✓ La emulación o imitación de hardware también es conocida como virtualización de metal desnudo (del inglés, bare metal virtualization), para simbolizar el hecho de que ningún software se encuentra entre el hipervisor y el ‘metal’ del servidor. Como hemos mencionado, el hipervisor intercepta las llamadas del sistema desde la máquina virtual huésped y coordina el acceso al hardware subyacente directamente.

  - ✓ Paravirtualización La paravirtualización no intenta emular un entorno de hardware en software, sino que un hipervisor de paravirtualización coordina (o multiplexa) el acceso a los recursos de hardware subyacentes del servidor. La arquitectura de la paravirtualización son los entornos Xen:





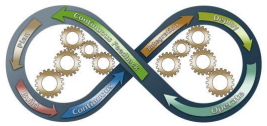
- Hay dos tipos de virtualización basada en hipervisores:

  - ✓ En la paravirtualización, el hipervisor reside en el hardware y, por lo tanto, esta se puede concebir como una arquitectura de virtualización de metal desnudo. Uno o más sistemas operativos huésped (equivalente a máquinas virtuales en virtualización de emulación del hardware) se ejecutan sobre el hipervisor. Un huésped privilegiado se ejecuta como una máquina virtual huésped, pero tiene privilegios que le permiten acceder directamente a ciertos recursos en el hardware subyacente.

- Proveedores de virtualización: VMware, Citrix y Microsoft son los principales proveedores de virtualización de servidores x86 en entornos profesionales, más adelante, veremos otras opciones susceptibles de ser usadas también en entornos personales y de prueba.

  - ✓ VMware Es el proveedor de virtualización de servidores más extendido y afianzado en el mercado. La plataforma insignia de VMware, vSphere, utiliza la tecnología de emulación de hardware.

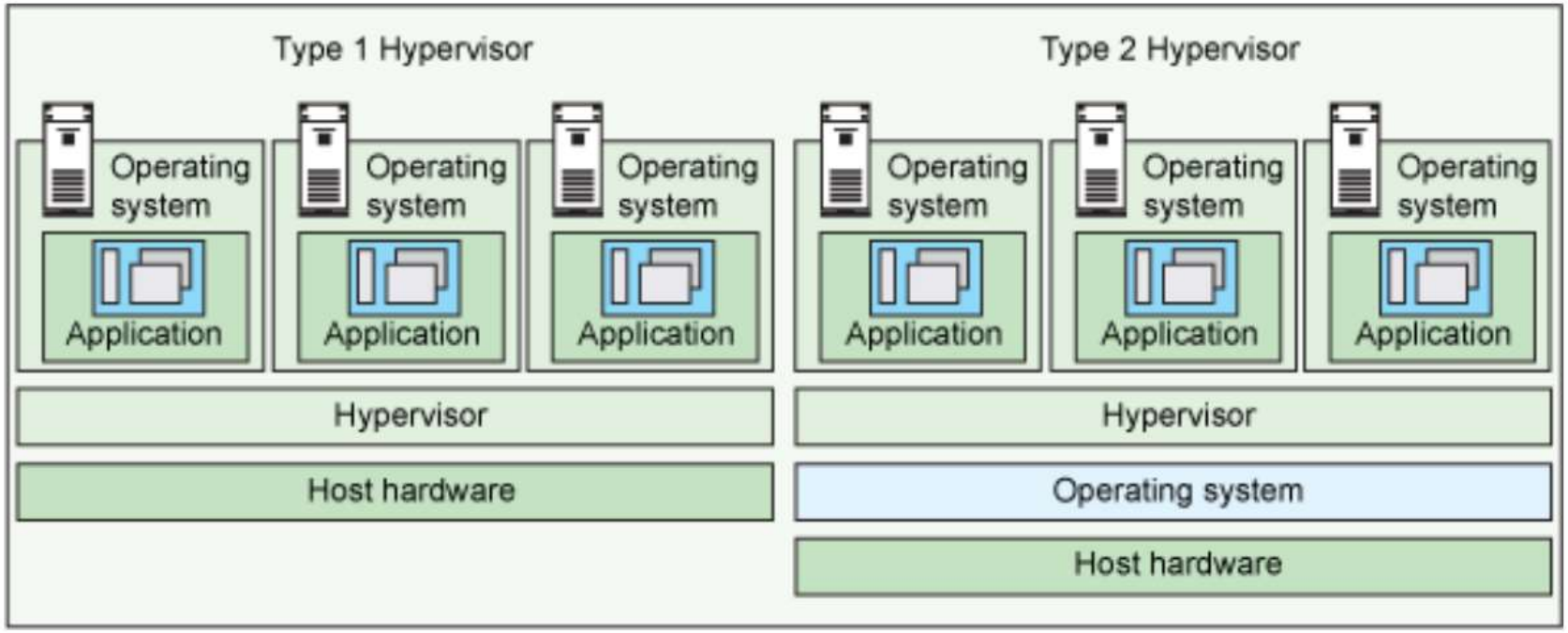
  - ✓ Citrix Ofrece un producto de virtualización de servidor llamado XenServer basado en paravirtualización. El huésped privilegiado (llamado control domain en lenguaje Xen) y el hipervisor Xen trabajan en equipo para permitir que las máquinas virtuales huésped interactúen con el hardware subyacente.

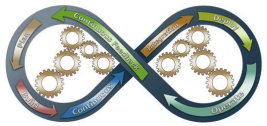


- ✓ Microsoft Hyper-V, el producto de virtualización del servidor de Microsoft tiene una arquitectura muy similar a la de Xen. En lugar de usar el término control domain para referirse a las máquinas virtuales huésped, Hyper-V se refiere a ellas como particiones y a la contraparte del control domain de Xen se la denomina partición principal

- ❑ Tipos de hipervisores Una definición sencilla de hipervisor podría ser: la parte de la nube privada que gestiona las máquinas virtuales, es decir, es la parte (programa) que permite que múltiples sistemas operativos compartan el mismo hardware. Cada sistema operativo podría usar todo el hardware (procesador, memoria) si no hay otro sistema operativo encendido. Ese es el hardware máximo disponible para un sistema operativo en la nube. Sin embargo, el hipervisor es el que controla y asigna qué parte de los recursos de hardware debe obtener cada sistema operativo, para que cada uno obtenga lo que necesita y no se interrumpa entre sí.

- Hay dos tipos de hipervisores:
  - ✓ Hipervisor de tipo 1: Los hipervisores se ejecutan directamente en el hardware del sistema: un hipervisor integrado ‘básico’.
  - ✓ Hipervisor tipo 2: Los hipervisores se ejecutan en un sistema operativo host que proporciona servicios de virtualización, como soporte de dispositivos de E / S y administración de memoria.





- Hipervisores tipo 1:

- ✓ VMware ESX y ESXi Estos hipervisores ofrecen funciones avanzadas y escalabilidad , pero requieren licencia, por lo que los costos son más altos. Su producto vSphere / ESXi está disponible en una edición gratuita y 5 ediciones comerciales.

- ✓ Microsoft Hyper-V El hipervisor de Microsoft, Hyper-V, no ofrece muchas de las funciones avanzadas que ofrecen los productos de VMware. Sin embargo, con XenServer y vSphere, Hyper-V es uno de los 3 principales hipervisores tipo 1. Actualmente, las nuevas versiones de supervisores de Microsoft están íntimamente relacionadas a sus productos de cloud y se les conoce como 'Azure Stack'.

- ✓ Citrix XenServer Comenzó como un proyecto de código abierto. La tecnología principal del hipervisor es gratuita, pero al igual que ESXi gratuito de VMware, casi no tiene características avanzadas. Xen es un hipervisor de tipo desnudo de tipo 1 y así como Red Hat Enterprise Virtualization usa KVM, Citrix usa Xen en el XenServer comercial.

- ✓ Oracle VM El hipervisor Oracle se basa en el código abierto Xen. Sin embargo, si necesita soporte de hipervisor y actualizaciones de productos, le costará. Oracle VM carece de muchas de las características avanzadas que se encuentran en otros hipervisores de virtualización de metal desnudo.

❑ Hipervisores tipo 2:

- ✓ VMware Workstation / Fusion / Player VMware Player es un hipervisor de virtualización gratuito. Está destinado a ejecutar solo una máquina virtual (VM) y no permite crear máquinas virtuales. VMware Workstation es un hipervisor más robusto con algunas características avanzadas, como grabación, reproducción y compatibilidad con instantáneas de VM. VMware Workstation tiene tres casos de uso principales: 1. Para ejecutar múltiples sistemas operativos. 2. Para ejecutar versiones diferentes de un sistema operativo en un escritorio. 3. Para desarrolladores que necesitan entornos de sandbox e instantáneas o para laboratorios y con fines de demostración.

- ✓ Servidor VMware VMware Server es un hipervisor de virtualización alojado gratuito que es muy similar a VMware Workstation. VMware ha detenido el desarrollo en el servidor desde 2009.

- ✓ Microsoft Virtual PC Esta es la última versión de Microsoft de esta tecnología de hipervisor, Windows Virtual PC y solo se ejecuta en Windows 7 y solo es compatible con los sistemas operativos Windows que se ejecutan en él.

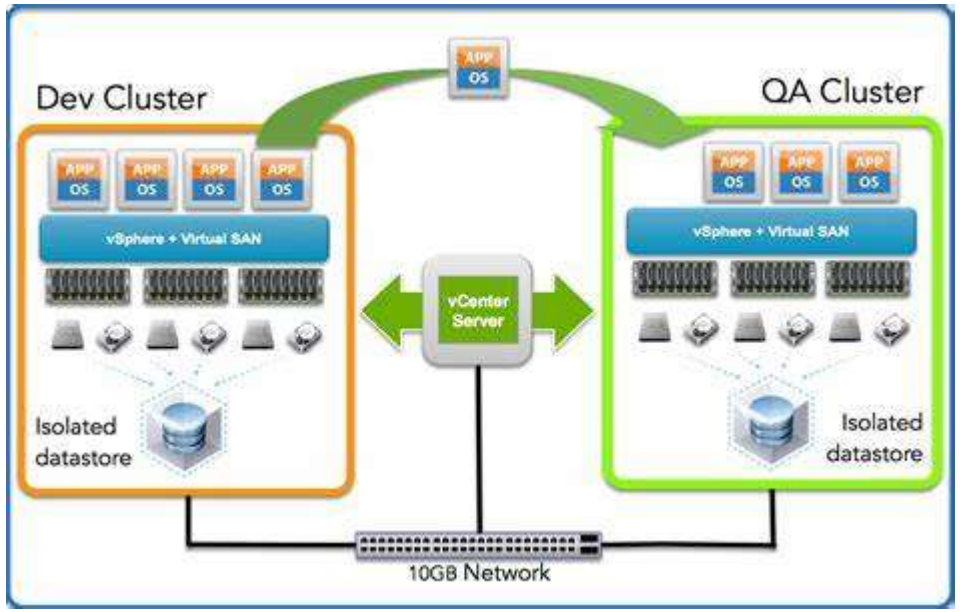
- ✓ Oracle VM VirtualBox La tecnología de hipervisor VirtualBox proporciona un rendimiento y características razonables si desea virtualizar con un presupuesto limitado. A pesar de ser un producto alojado gratuito con una huella muy pequeña, VirtualBox comparte muchas características con VMware vSphere y Microsoft Hyper-V.

- Hipervisores tipo 2:

- ✓ Red Hat Enterprise Virtualization La máquina virtual basada en el kernel (KVM) de Red Hat tiene cualidades tanto de un hipervisor de virtualización alojado como virtual. Puede convertir el núcleo de Linux en un hipervisor para que las máquinas virtuales tengan acceso directo al hardware físico.

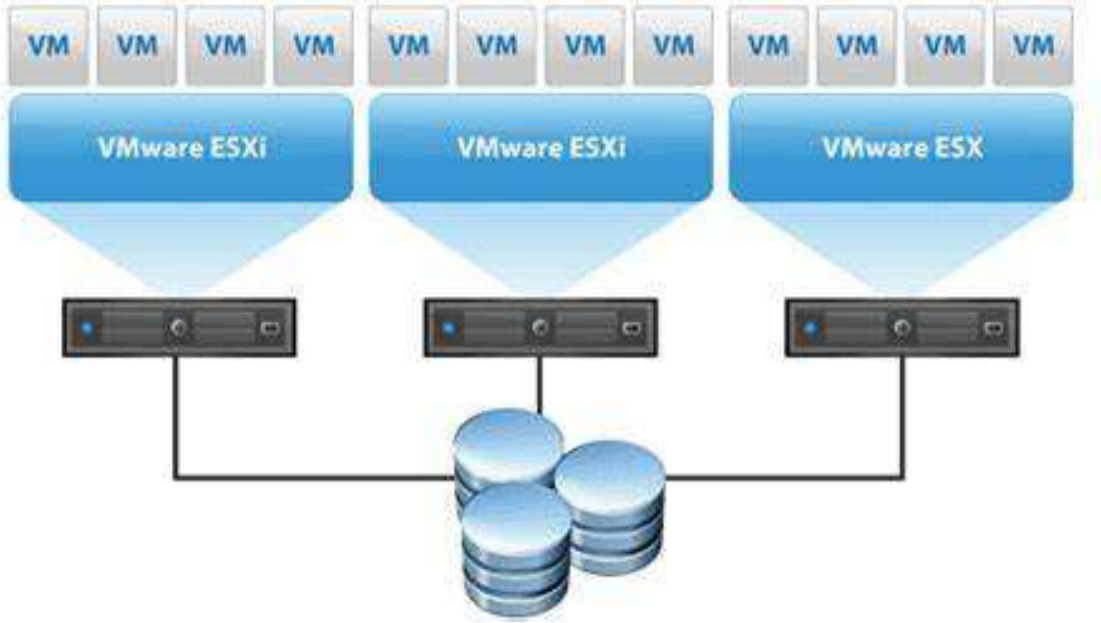
- ✓ KVM Esta es una infraestructura de virtualización para el kernel de Linux. Admite la virtualización nativa en procesadores con extensiones de virtualización de hardware. El KVM de código abierto (o máquina virtual basada en el núcleo) es un hipervisor tipo 1 basado en Linux que se puede agregar a la mayoría de los sistemas operativos Linux, incluidos Ubuntu, Debian, SUSE y Red Hat Enterprise Linux, pero también Solaris y Windows.

- Alta disponibilidad (HA): extiende el concepto de conmutación por error al incorporar un servidor de hardware adicional. Es decir, que en el caso de que la máquina virtual fallase, esta no se iniciaría en la misma pieza de hardware, sino que se iniciaría en un servidor diferente, evitando así el problema de conmutación por error de virtualización por adición del hardware:



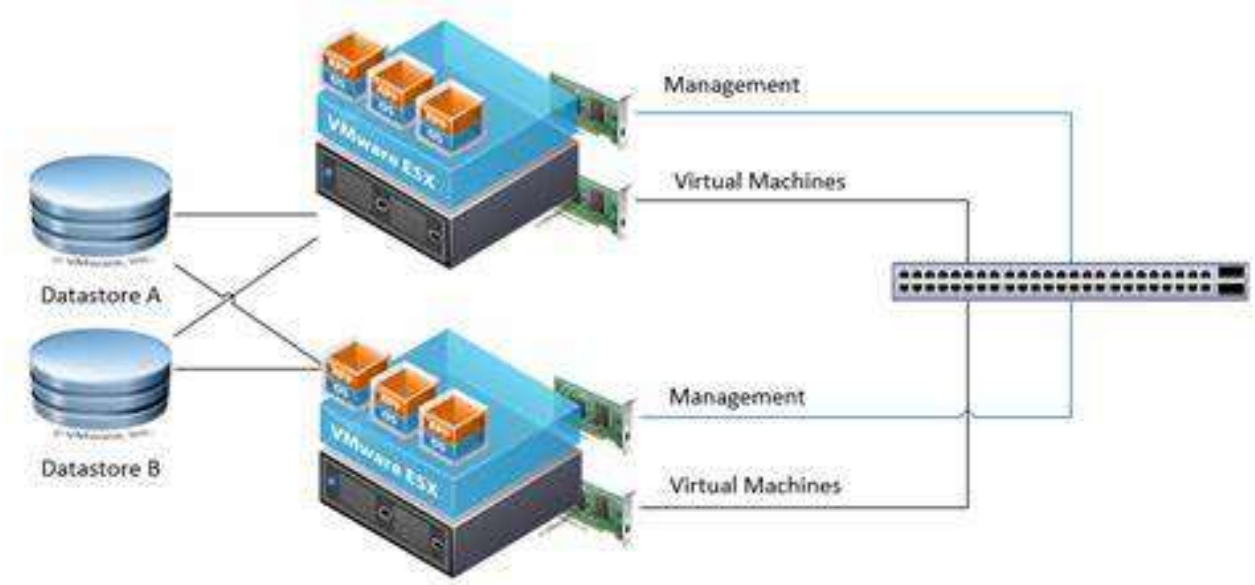
- Agrupamiento (clustering) El agrupamiento está diseñado para garantizar que no se pierdan datos en caso de que haya un fallo de software o de hardware. Hay un gasto extra que se produce por la necesidad de contar con hardware adicional, con el sistema en espejo en espera (en modo standby.

- El sistema SB está listo para asumir el control si fallase el sistema primario; sin embargo, si en el sistema se operan transacciones de millones de euros, mantener un servidor redundante que esté listo para operar cuando haya un fallo, puede ser una inversión que valga la pena.



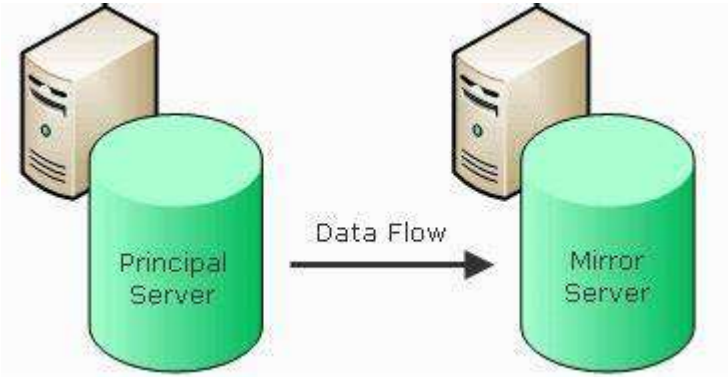
- Esencialmente, el software de coordinación de virtualización ejecuta dos máquinas virtuales en máquinas separadas. Las máquinas virtuales son idénticas en cuanto al sistema operativo y la configuración de la aplicación, pero difieren, naturalmente, en los detalles de sus conexiones de red y hardware local. El supervisor de virtualización se comunica constantemente con las máquinas virtuales en el clúster para confirmar que están trabajando (heartbeat).

- Una VM (máquina virtual) es el servidor primario y es el sistema con el que los usuarios interactúan. La segunda VM sirve como backup (copia de seguridad), lista para actuar en caso de que el servidor primario se caiga:

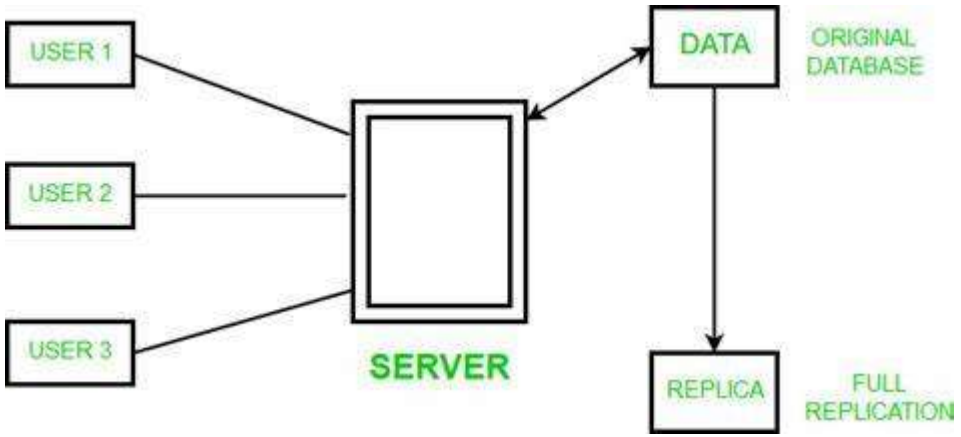


- **Duplicación de datos (data mirroring):** Una forma de mantener los datos disponibles es a través de la duplicación. Como el nombre implica, esta duplicación o reflejo significa que los datos existentes en un sitio son reflejados en otro y ambas contienen la misma información. La duplicación permite la consistencia en tiempo real entre dos fuentes de datos.

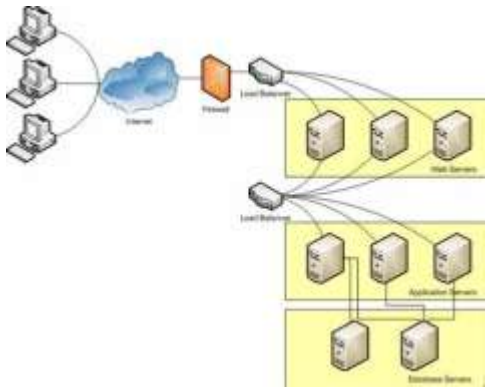
- Esto posibilita el cambio inmediato entre un sistema y otro, es decir, conectando el segundo sistema a la duplicación o el reflejo de los datos del sistema primario:



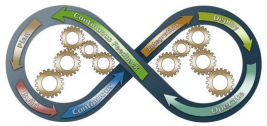
- **Réplica de datos:** La replicación es otro servicio orientado a mejorar la calidad del servicio de datos. A diferencia de la duplicación (data mirroring) que se enfoca en cómo mantener copias de datos consistentes en tiempo real, la replicación aborda la necesidad de mantener copias completas de los datos para que puedan ser utilizados en la reconstrucción del sistema. Esto se logra enviando copias de datos a un almacenamiento centralizado, lo que permite a una organización tener la seguridad de que en caso de que necesite acceder a los datos críticos por algún motivo, estos están almacenados de forma segura y disponibles en caso de ser necesarios. La eficiencia es vital para la replicación, es decir, que no debemos pensar que porque los datos se están moviendo a una ubicación de almacenamiento hay que olvidarse de que los datos deben fluir correctamente. Un software de replicación inteligente mantiene los cambios, minuto a minuto fluyendo a la ubicación central, asegurando así que una organización de TI pueda localizar rápidamente los datos y usarlos para reconstruir el sistema en caso de fallos:



- **Balanceo de carga (load balancing):** El equilibrio o balanceo de carga protege a un sistema de la vulnerabilidad contra cualquier condición de error dada al implementar la denominada redundancia. Esta se logra a través de la ejecución de una o más copias de una máquina virtual en servidores separados. Cuando se ejecutan dos instancias de una máquina virtual y una de ellas se bloquea, la otra continúa funcionando. Si el hardware que da soporte a una de las máquinas virtuales falla, la otra máquina sigue funcionando. De esta manera, se evita que la aplicación sufra una interrupción. El balanceo de carga también hace un mejor uso de los recursos de la máquina. Esto es así porque en lugar de que la segunda máquina virtual esté inactiva y no realice ningún trabajo útil, aunque esté siendo actualizada por la máquina principal, la segunda VM lleva la mitad de la carga y esto hace que al menos la mitad de sus recursos se utilicen. El uso de recursos duplicados puede extenderse más allá de las máquinas virtuales en sí mismas. Las organizaciones que luchan por alcanzar altos niveles de disponibilidad, a menudo, implementan redes duplicadas con cada servidor físico con conexión cruzada con el resto de la red, lo que garantiza que las máquinas virtuales continuarán siendo capaces de comunicarse incluso si parte de la red se cae.

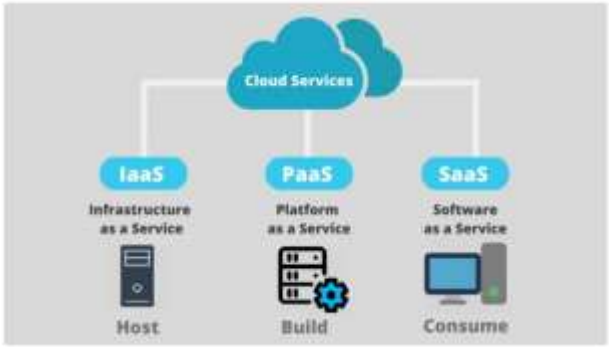


- **Cloud computing:** El paradigma de la nube introduce un cambio en la visualización del sistema y los datos que son propiedad de una empresa. Además, el uso compartido de servicios o recursos, tales como el almacenamiento, hardware y aplicaciones de cloud computing, de una manera totalmente diferente ha facilitado la coherencia de los recursos y las economías de escala a través de su modelo de negocio de pago por uso. Ya no se trata de un conjunto de dispositivos en una ubicación física que ejecutan un programa de software específico con todos los datos y los recursos presentes en un lugar físico, sino que es un sistema que se distribuye geográficamente, involucrando tanto a la aplicación como a los datos:

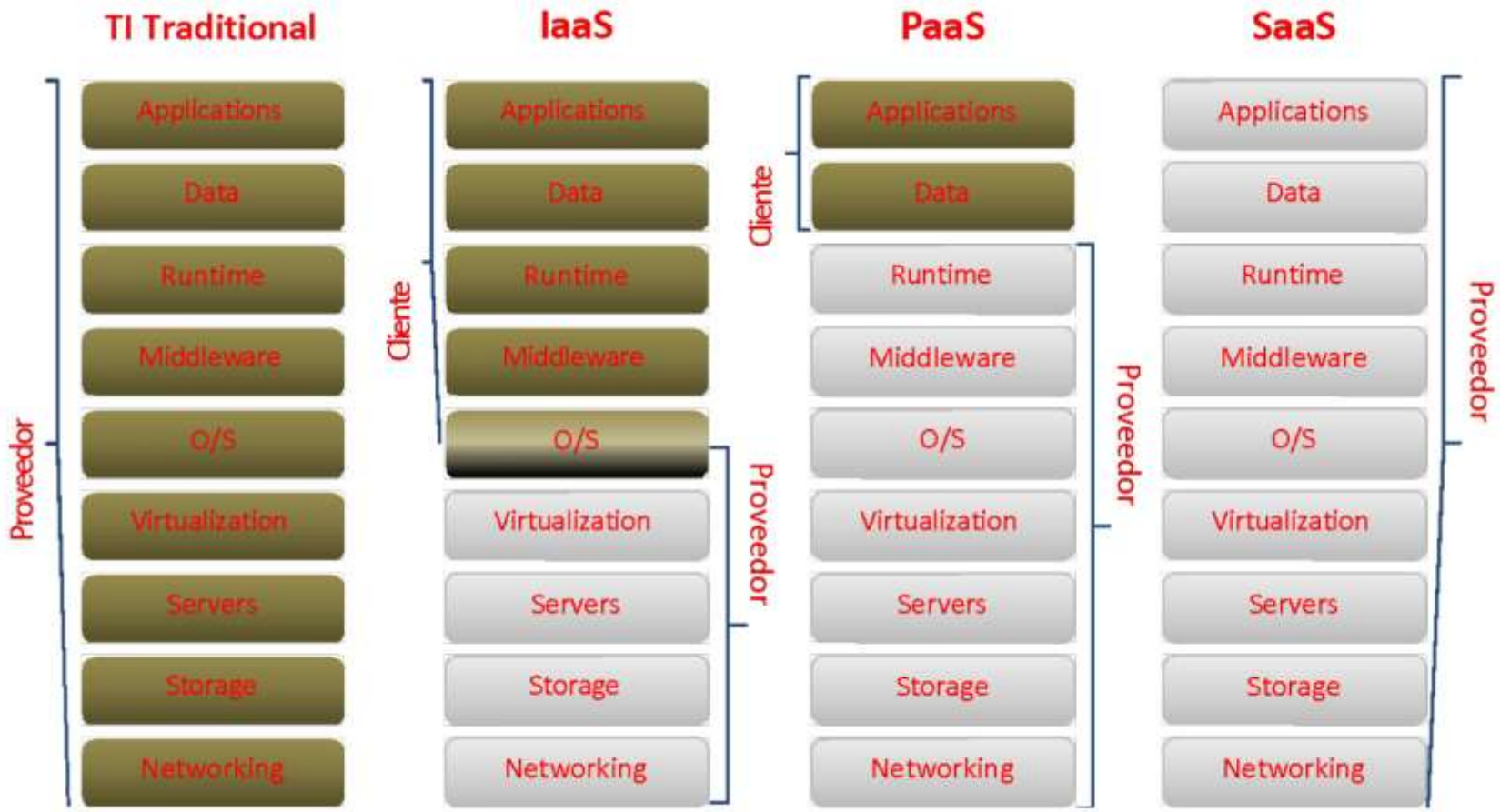


- Cloud computing:
- Hay 3 tipos de nube El hecho de que la información resida de forma temporal o definitiva en servidores de nube da como resultado que dichos servicios ofrezcan distintos formatos de privacidad que cada usuario puede elegir, según sus necesidades:
  - ✓ Nube pública: Los usuarios acceden a los servicios de manera compartida sin que exista un exhaustivo control sobre la ubicación de la información, que reside en los servidores del proveedor. Es importante resaltar que el hecho de que sean públicas no es un sinónimo de que sean inseguras, pero la realidad es que suelen ser más vulnerables a los ataques. Cuando hablamos de nube pública, queremos decir que toda la infraestructura de computación se encuentra en las instalaciones de una empresa de cloud computing que ofrece el servicio en la nube. La ubicación permanece, por lo tanto, separada del cliente y este no tiene control físico sobre la infraestructura.
  - ✓ Nube privada: Nube privada significa usar una infraestructura en la nube (red) por cada cliente u organización. Si bien no se comparte con otros, se encuentra remotamente localizada. Las empresas tienen la opción de elegir una nube privada en la propia sede, que es más cara, pero tiene la ventaja de que así se puede tener control físico sobre la infraestructura. Resulta evidente que el nivel de seguridad y control es más alto cuando se utiliza una red privada que una red pública. Sin embargo, la reducción de costes puede ser mínima si la empresa necesita invertir en una infraestructura en la nube on premise.
  - ✓ Nube híbrida: Nubes híbridas Combinan características de las dos anteriores, de manera que parte del servicio se puede ofrecer de manera privada (por ejemplo, la infraestructura) y otra parte de manera compartida

- **Infraestructura como servicio (IaaS):** También conocida como HaaS, del inglés hardware as a service, esta arquitectura se basa en el modelo de dotar de forma externalizada a sus usuarios / empresas del hardware necesario. IaaS proporciona hardware, almacenamiento, servidores y espacio de centro de datos o componentes de red. Como inconveniente podemos destacar que se requiere de los mismos conocimientos informáticos en sistemas operativos y redes informáticas que necesitábamos con una arquitectura tradicional.
- **Plataforma como servicio (PaaS):** Se trata de un modelo en el que se proporciona un servicio de plataforma con todo lo necesario para dar soporte al ciclo de diseño, desarrollo y puesta en marcha de aplicaciones y servicios web a través de esta. El proveedor es el encargado de escalar los recursos en caso de que la aplicación lo requiera, de que la plataforma tenga un rendimiento óptimo, de la seguridad de acceso, etc. Para desarrollar software se necesitan bases de datos, herramientas de desarrollo y en ocasiones servidores y redes. Con PaaS, el cliente únicamente se enfoca en desarrollar, depurar y probar ya que las herramientas necesarias para el desarrollo de software son ofrecidas a través de internet, lo que teóricamente permite aumentar la productividad de los equipos de desarrollo gracias a que abstrae del hardware físico al cliente.
- **Software como servicio (SaaS):** Consiste en la entrega de aplicaciones completas como un servicio, permitiendo la abstracción completa, no solo del hardware subyacente, sino incluso de la plataforma, permitiendo al usuario (cliente) dedicarse únicamen



❑ Comparativa de los niveles de servicio en la nube:

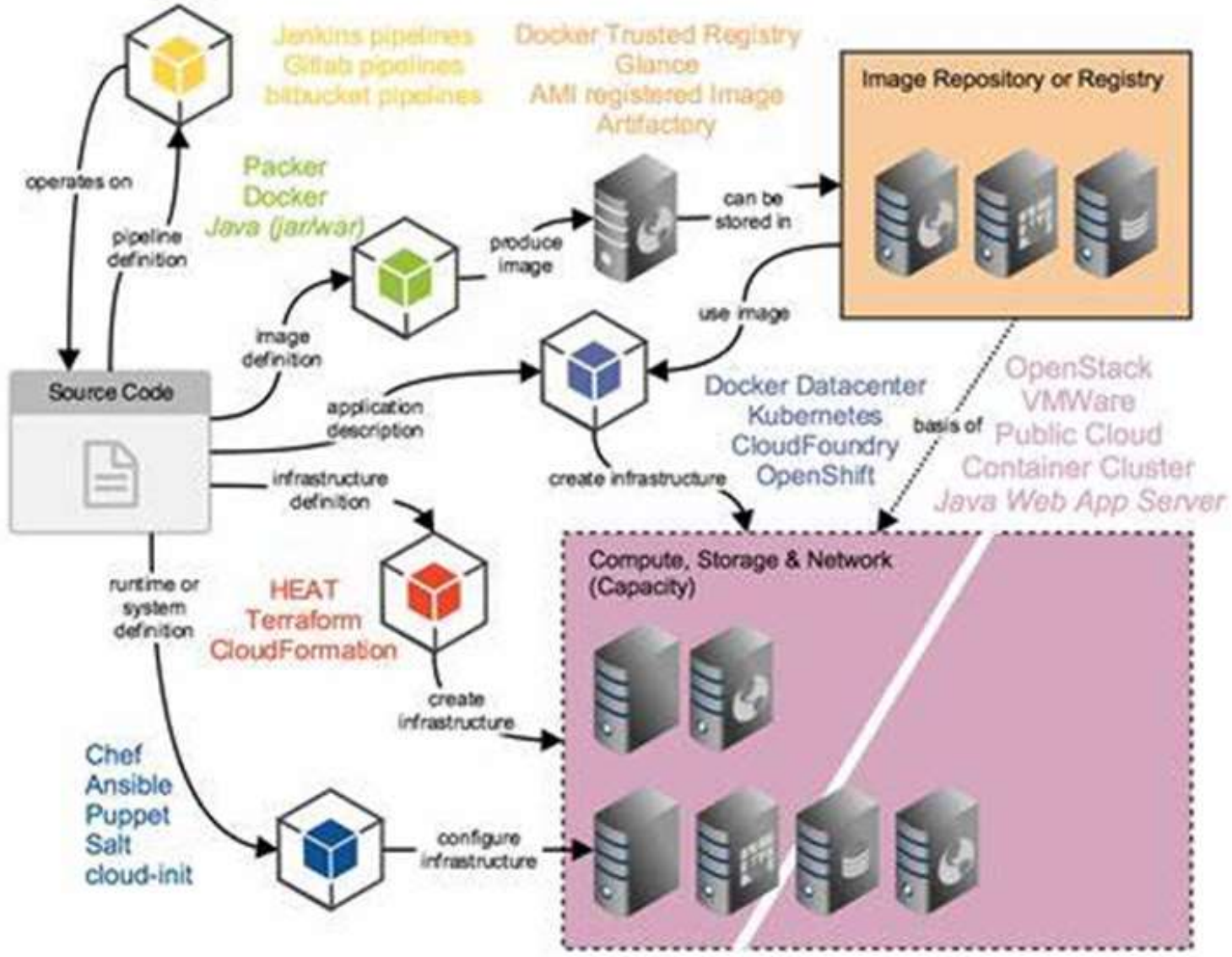


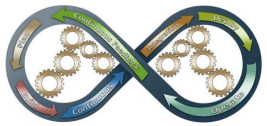
Pie: Comparativa de los niveles de servicio en la nube.

- **Funciones como servicio (FaaS):** Las funciones como servicio son un concepto relativamente moderno, introducido en el mundo cloud por Amazon Web Services con AWS Lambdas en el 2014. Se trata de un tipo de servicio de computación que ejecuta código en respuesta a eventos, pero sin que los usuarios se tengan que preocupar de la infraestructura que sería necesaria para albergar un servicio de estas características: esto permite a los desarrolladores enfocarse únicamente en la creación de código fuente funcional. FaaS es un concepto que se suele asociar, o incluso equiparar, con otro paradigma de computación cloud: serverless.
- **El concepto serverless es el de un PaaS para aplicaciones orientadas a internet o web, en el que el coste es por llamada a la aplicación**. Esto significa que, si no tiene llamadas, el servicio se apaga y queda en espera automáticamente. Y, en caso de tenerlas, el servicio escala automáticamente y de manera transparente para atender la demanda. Pero además, serverless sí que incluye toda la tecnología subyacente para dar servicio a las FaaS, por lo que deben considerarse estas como un subtipo dentro del mundo serverless.

- **Infraestructura como código (IaC):** La respuesta que ha dado la informática a estos problemas es la optimización y automatización del despliegue de aplicaciones y la configuración de los servidores. Actualmente, es posible configurar servidores a través de la programación y sin intervención de los administradores. Esta nueva forma de administración, que considera a la infraestructura como un aplicativo más a ser gestionado de manera análoga al software, se las conoce como infraestructura programable o infraestructura como código, más conocida como IaC (Huttermann, M., 2012).
- La IaC ofrece las siguientes ventajas frente a la configuración manual tradicional:
  - ✓ Alta eficiencia: Automatiza la mayor parte de la administración de los recursos, lo que lleva a optimizar el ciclo de vida de desarrollo SW.
  - ✓ Reutilización: Una vez se haya descrito una infraestructura como código, esta se puede ejecutar en cualquier momento, todas las veces que se desee, de forma idempotente.
  - ✓ Control de versiones: Al ser código fuente, lo natural es que se almacene en un repositorio, lo que lleva a poder revisar los cambios a lo largo del tiempo.
  - ✓ Minimiza costes y esfuerzo: Al automatizar trabajo tedioso y tendiente a errores.

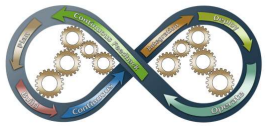
❑ Infraestructura como código (IaC):





❑ Existen dos grandes grupos de herramientas de IaC:

1. Herramientas de proveedor Son las herramientas presentes en los grandes proveedores de la nube, y permiten describir y provisionar únicamente sus propias tecnologías. Los tres ejemplos más conocidos son:
   - ✓ **AWS CloudFormation CloudFormation:** es la herramienta interna de IaC de Amazon Web Services (AWS) y, como tal, es prácticamente imprescindible para cualquiera que trabaje con productos de AWS como ELB, S3 o EFS. Utilizarla no conlleva ningún coste adicional, tan solo hay que pagar por los recursos reservados.
   - ✓ **Azure Resource Manager:** Servicio de Azure, plataforma en la nube de Microsoft, que proporciona la administración de infraestructuras mediante plantillas, de forma que implementa y supervisa todos los recursos. Esto da coherencia a la hora de reimplementar recursos ya existentes y permite definir las dependencias de estos.
   - ✓ **Google Cloud Deployment Manager:** Deployment Manager es para la plataforma Google Cloud lo que CloudFormation es para AWS. Con esta herramienta gratuita, los usuarios de recursos de IaaS de Google pueden administrarlos fácilmente mediante archivos de configuración central en el lenguaje de marcado YAML.



2. Las herramientas multiproveedor más empleadas actualmente son:
   - ✓ **Terraform:** es una herramienta que se utiliza para la construcción, el cambio y el versionado de infraestructura, de manera segura y eficiente. Esta puede administrar tanto servicios existentes como nubes públicas o soluciones internas personalizadas.
   - ✓ **Heat:** Implementa un motor de orquestación para poder lanzar múltiples aplicaciones en la nube basadas en plantillas, proporcionando una compatibilidad total con las plantillas de AWS CloudFormation. Proporciona, a su vez, una API nativa y una API compatible con AWS CloudFormation. Heat es el proyecto más ambicioso de OpenStack.
   - ✓ **Chef Infra:** la solución de IaC de la empresa estadounidense Chef, está disponible desde abril de 2019 bajo la licencia gratuita Apache 2.0 y es utilizado por Facebook, entre otras empresas. Entre las plataformas compatibles se incluyen Google Cloud, Microsoft Azure, Amazon EC2 y OpenStack.
   - ✓ **Puppet:** Herramienta Open Source de gestión desarrollada en Ruby para la administración de sistemas de forma declarativa, la cual al estar basada en modelos no requiere un alto conocimiento de programación para su uso.
   - ✓ **Red Hat Ansible Tower:** La herramienta de infraestructura como código de Ansible forma parte del catálogo de desarrollo de software Red Hat desde el año 2015. Ofrece un panel de control, su propia línea de comandos y una potentísima API REST. En este caso, ambos paquetes disponibles, tanto el estándar como el extendido, son de pago.

2. Las herramientas multiproveedor más empleadas actualmente son:
   - ✓ **Terraform:** es una herramienta que se utiliza para la construcción, el cambio y el versionado de infraestructura, de manera segura y eficiente. Esta puede administrar tanto servicios existentes como nubes públicas o soluciones internas personalizadas.
   - ✓ **Heat:** Implementa un motor de orquestación para poder lanzar múltiples aplicaciones en la nube basadas en plantillas, proporcionando una compatibilidad total con las plantillas de AWS CloudFormation. Proporciona, a su vez, una API nativa y una API compatible con AWS CloudFormation. Heat es el proyecto más ambicioso de OpenStack.
   - ✓ **Chef Infra:** la solución de IaC de la empresa estadounidense Chef, está disponible desde abril de 2019 bajo la licencia gratuita Apache 2.0 y es utilizado por Facebook, entre otras empresas. Entre las plataformas compatibles se incluyen Google Cloud, Microsoft Azure, Amazon EC2 y OpenStack.
   - ✓ **Puppet:** Herramienta Open Source de gestión desarrollada en Ruby para la administración de sistemas de forma declarativa, la cual al estar basada en modelos no requiere un alto conocimiento de programación para su uso.
   - ✓ **Red Hat Ansible Tower:** La herramienta de infraestructura como código de Ansible forma parte del catálogo de desarrollo de software Red Hat desde el año 2015. Ofrece un panel de control, su propia línea de comandos y una potentísima API REST. En este caso, ambos paquetes disponibles, tanto el estándar como el extendido, son de pago.