

**57. VIRTUALIZACIÓN DE
SERVIDORES. VIRTUALIZACIÓN
DEL ALMACENAMIENTO.
VIRTUALIZACIÓN DEL PUESTO
CLIENTE. COMPUTACIÓN
BASADA EN SERVIDOR (SBC).
GRID COMPUTING. CLOUD
COMPUTING. GREEN IT Y
EFICIENCIA ENERGÉTICA.**

Tema 57: Virtualización de servidores. Virtualización del almacenamiento. Virtualización del puesto cliente. Computación basada en servidor (SBC). Grid Computing. Cloud computing. Green IT y eficiencia energética.

57.1 Virtualización de servidores.....	1
57.2 Virtualización del almacenamiento.....	5
57.3 Virtualización del puesto cliente.....	10
57.4 Computación basada en servidor.....	12
57.5 Grid Computing.....	15
57.6 Cloud computing.....	19
57.7 Green IT e eficiencia energética.....	21
57.8 Bibliografía.....	29

57.1 VIRTUALIZACIÓN DE SERVIDORES

Podemos definir virtualización como la técnica que consiste básicamente en agrupar diferentes aplicaciones y servicios de sistemas heterogéneos dentro de un mismo hardware, de forma que los usuarios y el propio sistema los vean como máquinas independientes dedicadas. Para ello, el sistema operativo virtualizado debe ver el hardware de la máquina real como un conjunto normalizado de recursos independientemente de los componentes reales que lo formen.

De esta forma, para virtualizar un sistema de servidores, los administradores deben, básicamente, optimizar los recursos disponibles, incluyendo el número y la identidad de los servidores físicos individuales, procesadores, y sistemas operativos, con el objetivo de producir una mejora tanto en la gestión como en el manejo de sistemas informáticos complejos. El administrador del sistema virtual utilizará un software para la

división del servidor físico en entornos virtuales aislados. Estos entornos son lo que se conoce técnicamente como servidores privados virtuales, pero también se pueden encontrar referencias como particiones, instancias, contenedores o emulaciones de sistemas.

En concreto, podemos decir que un servidor privado virtual es un término de marketing utilizado por los servicios de hosting para referirse a una máquina virtual para el uso exclusivo de un cliente individual del servicio. El término se utiliza para enfatizar que la máquina virtual, a pesar de ejecutarse en el mismo equipo físico que las máquinas virtuales de otros clientes, es funcionalmente equivalente a un equipo físico independiente, está dedicado a las necesidades individuales del cliente y puede ser configurado para ejecutarse como un servidor de internet (es decir, para ejecutar software de servidor). El término VDS o Virtual Dedicated Server (Servidor Virtual Dedicado) para el mismo concepto.

Cada servidor virtual puede ejecutar su propio sistema operativo y ser reiniciado de modo independiente.

57.1.1 Funcionamiento

El servidor físico realiza una abstracción de los recursos que se denomina Hypervisor o VMM (Virtual Machine Monitor), elemento software que se instala en la máquina donde se va a llevar a cabo la virtualización y sobre la que se configuran las máquinas virtuales que es donde van a residir las aplicaciones. Es el encargado de gestionar los recursos de los sistemas operativos “alojados” (guest) o máquinas virtuales.

Desde un punto de vista lógico, el usuario percibe que son máquinas independientes y aisladas entre sí, pero desde una perspectiva física, todas las máquinas virtuales residen en un único servidor. A estas máquinas virtuales se les asigna un porcentaje de los recursos del servidor físico, que serán los únicos que el cliente conozca.

Se pueden encontrar tres modelos de virtualización: el modelo de máquina virtual o virtualización completa, el modelo paravirtual o virtualización parcial; y la virtualización a nivel de sistema operativo.

57.1.1.1 Virtualización completa

El modelo de máquina virtual está basado en la arquitectura cliente/servidor, donde cada cliente funciona como una imagen virtual de la capa hardware. Este modelo permite que el sistema operativo cliente funcione sin modificaciones. Además permite al administrador crear diferentes sistemas cliente con sistemas operativos independientes entre sí. La ventaja principal de este modelo radica en el desconocimiento por parte de los sistemas huésped del sistema hardware real sobre el que está instalado. Sin embargo, realmente todos los sistemas virtuales hacen uso de recursos hardware físicos. Estos recursos son administrados por un el hypervisor que coordina las instrucciones CPU, convirtiendo las peticiones del sistema invitado en las solicitudes de recursos apropiados en el host, lo que implica una sobrecarga considerable. Casi todos los sistemas pueden ser virtualizados utilizando este método, ya que no requiere ninguna modificación del sistema operativo. A pesar de esto, es necesaria una virtualización de la CPU como apoyo para la mayoría de los hypervisores que llevan a cabo la virtualización completa.

Ejemplos típicos de sistemas de servidores virtuales son VMware Workstation, VMware Server, VirtualBox, Parallels Desktop, Virtual Iron, Adeos, Mac-on-Linux, Win4BSD, Win4Lin Pro, y z/VM, openvz, Oracle VM, XenServer, Microsoft Virtual, PC 2007 y Hyper-V.

57.1.1.2 Paravirtualización

El modelo de máquina paravirtual (PVM) o virtualización parcial se basa, como el modelo anterior, en la arquitectura cliente/servidor, incluyendo también la necesidad de contar con un sistema monitor. Sin embargo, en este caso, el VMM accede y modifica el código del sistema operativo del sistema huésped. Esta modificación se conoce como porting. El porting

sirve de soporte al VMM para que pueda realizar llamadas al sistema directamente. Al igual que las máquinas virtuales, los sistemas paravirtuales son capaces de soportar diferentes sistemas operativos instalados en el hardware real. Esta técnica se utiliza con intención de reducir la porción de tiempo de ejecución empleada por el huésped empleado en realizar las operaciones que son mucho más difíciles de ejecutar en un entorno virtual en comparación con un entorno no virtualizado. Así se permite que el invitado(s) y el huésped soliciten y reconozcan estas tareas, que de otro modo serían ejecutados en el dominio virtual (donde el rendimiento de ejecución es peor). Una plataforma paravirtualizada exitosamente puede permitir que el VMM sea menos complejo (por la reubicación de la ejecución de las tareas críticas del dominio virtual en el dominio del servidor), y/o reducir la degradación del rendimiento global de la máquina virtual durante la ejecución de invitado.

UML, XEN, Xen, Virtuozzo , Vserver y OpenVZ (que es el código abierto y la versión de desarrollo de Parallels Virtuozzo Containers) son modelos de máquinas paravirtuales.

57.1.1.3 Virtualización por S.O.

La virtualización a nivel de sistema operativo se diferencia de las anteriores en que, en este caso, no existe un sistema cliente/servidor propiamente dicho. En este modelo el sistema principal exporta la funcionalidad del sistema operativo desde su propio núcleo. Por esta razón, los sistemas virtuales usan el mismo sistema operativo que el nativo (aunque en la mayoría de los casos pueden instalar distintas distribuciones). Esta arquitectura elimina las llamadas del sistema entre capas, lo que favorece una reducción importante en el uso de CPU. Además, al compartir los ficheros binarios y librerías comunes del sistema en la misma máquina, la posibilidad de escalado es mucho mayor, permitiendo que un mismo servidor virtual sea capaz de dar servicio a un gran número de clientes al mismo tiempo.

La Virtualización de SO mejora el rendimiento, gestión y eficiencia. Podemos entenderlo como un sistema en capas. En la base reside un sistema operativo huésped estándar. A continuación encontramos la capa de virtualización, con un sistema de archivos propietario y una capa de abstracción de servicio de kernel que garantiza el aislamiento y seguridad de los recursos entre distintos contenedores. La capa de virtualización hace que cada uno de los contenedores aparezca como servidor autónomo. Finalmente, el contenedor aloja la aplicación o carga de trabajo.

Ejemplos de sistemas que usan virtualización a nivel de sistema operativo son Virtuozzo y Solaris.

57.2 VIRTUALIZACIÓN DEL ALMACENAMIENTO

Este tipo de virtualización permite una mayor funcionalidad y características avanzadas en el sistema de almacenamiento. Consiste en abstraer el almacenamiento lógico del almacenamiento físico y suele usarse en SANs (Storage Area Network, Red de área de almacenamiento).

Este sistema de almacenamiento también se conoce como “storage pool”, matriz de almacenamiento, matriz de disco o servidor de archivos. Estos sistemas suelen usar hardware y software especializado, junto con unidades de disco con el fin de proporcionar un almacenamiento muy rápido y fiable para el acceso a datos. Son sistemas complejos, y pueden ser considerados como un ordenador de propósito especial diseñado para proporcionar capacidad de almacenamiento junto con funciones avanzadas de protección de datos. Las unidades de disco son sólo un elemento dentro del sistema de almacenamiento, junto con el hardware y el software de propósito especial incorporado en el sistema.

Los sistemas de almacenamiento pueden ser de acceso a nivel de bloque, o acceso a nivel de ficheros. El acceso por bloques suele llevarse a cabo por medio de Fibre Channel , iSCSI , SAS , FICON u otros protocolos. Para el acceso a nivel de archivo se usan los protocolos NFS o CIFS.

Dentro de este contexto nos podemos encontrar con dos tipos principales de virtualización: la virtualización por bloques y la virtualización por archivos.

57.2.1 *Virtualización por bloques*

Este tipo de virtualización se basa en la abstracción (diferenciación) entre el almacenamiento lógico y el almacenamiento físico, consiguiendo que el acceso no tenga en cuenta el almacenamiento físico o estructura heterogénea.

Existen tres tipos de virtualización por bloques: basada en host, basada en dispositivos de almacenamiento, basada en red.

57.2.1.1 Virtualización basada en host

Esta virtualización requiere software adicional que se ejecuta en el host. En algunos casos la administración de volúmenes está integrada en el sistema operativo, y en otros casos se ofrece como un producto separado. Los volúmenes (LUN) disponibles en el sistema son manejados por un controlador de dispositivos físicos tradicional. Por encima de este controlador se encuentra una capa software (el gestor de volúmenes) que intercepta las peticiones de E / S, y proporciona la búsqueda de meta-datos y mapeos de E / S.

Los sistemas operativos más modernos tienen algún tipo de gestor de volúmenes lógicos integrado (MVI en UNIX / Linux, o Administrador de discos lógicos o LDM en Windows), que realiza tareas de virtualización.

Existen varias tecnologías que implementan este tipo de virtualización, como pueden ser la gestión de volúmenes lógicos (Logical Volume Management, LVM), los sistemas de archivos (CIFS, NFS) o el montaje automático (autofs)

57.2.1.2 Virtualización basada en dispositivos de almacenamiento

Se puede llevar a cabo la virtualización basada en medios de almacenamiento masivo utilizando un controlador de almacenamiento primario que proporcione los servicios de virtualización y permita conexión directa de los controladores de almacenamiento. En función de la implementación es posible usar modelos de distintos fabricantes.

El controlador primario proporcionará la puesta en común y los meta-datos de servicio de gestión. También puede ofrecer servicios de replicación y migración a través de los controladores que se virtualizan.

Una nueva generación de controladores de serie del disco permite la inserción posterior de los dispositivos de almacenamiento.

Los sistemas RAID pueden ser un ejemplo de esta técnica. Estos sistemas combinan varios discos en una sola matriz.

Las matrices avanzadas de disco, cuentan a menudo con clonación, instantáneas y replicación remota. En general, estos dispositivos no ofrecen los beneficios de la migración de datos o de replicación a través de almacenamiento heterogéneo, ya que cada fabricante tiende a utilizar sus propios protocolos propietarios.

57.2.1.3 Virtualización basada en red

Esta es una virtualización de almacenamiento operando en un dispositivo basado en red (por lo general un servidor estándar o un smart switch) y el uso de redes iSCSI o FC de Fibre Channel para conectar como SAN (Storage Area Network). Este es el tipo de virtualización de almacenamiento más común.

El dispositivo de virtualización se encuentra en la SAN y proporciona la capa de abstracción entre los host, que permiten la entrada/salida, y los controladores de almacenamiento, que proporcionan capacidad de almacenamiento.

Hoy en día existen dos implementaciones distintas, la basada en el **dispositivo** y la basada en **conmutación**. Ambos modelos proporcionan los mismos servicios: gestión de discos, búsqueda de meta-datos, migración y replicación de datos. Igualmente, ambos modelos necesitan de un hardware específico que permita ofrecer dichos servicios.

La basada en dispositivos consiste en establecer el hardware especializado entre los hosts y la parte de almacenamiento. Las solicitudes de entrada/salida se redirigen al dispositivo, que realiza la asignación de meta-datos, mediante el envío de sus propias órdenes de E/S a la solicitud de almacenamiento subyacente. El hardware usado también puede proporcionar almacenamiento de datos en caché, y la mayoría de las implementaciones proporcionan algún tipo de agrupación de cada uno de los dispositivos para mantener un punto de vista atómico tanto de los meta-datos como de los datos de la caché.

Este tipo de almacenamiento también puede clasificarse en in-band (simétrica) o out-of-band (asimétrica).

57.2.1.3.1 In-band (simétrica)

En este caso los dispositivos de virtualización se asientan entre el host y el almacenamiento. Todas las peticiones de E/S y datos pasan a través del dispositivo. Los host nunca interactúan con el dispositivo de almacenamiento sino con el dispositivo de virtualización.

57.2.1.3.2 Out-of-band (asimétrica)

Los dispositivos usados en este tipo de virtualización también son llamados servidores de meta-datos. La única finalidad de estos dispositivos es proporcionar la asignación de meta-datos. Esto implica el uso de software adicional en el host, que es conocedor de la ubicación real de los datos. De este modo, se intercepta la petición antes de que salga del host, se solicita una búsqueda de meta-datos en el servidor (puede ser a través de una interfaz que no sea SAN) y se devuelve la ubicación real de los datos solicitados por el host. Finalmente se recupera la información a través de

una solicitud de E/S común al dispositivo de almacenamiento. No se puede dar un almacenamiento en caché ya que los datos nunca pasan a través del dispositivo de virtualización.

57.2.2 *Virtualización a nivel de archivo*

Con este tipo de virtualización se pretende eliminar las dependencias entre el acceso a datos a nivel de archivo y la ubicación física de los mismos. Esta técnica, conocida como NAS (Network-Attached Storage) o almacenamiento conectado a red, suele ser un equipo especializado pensado exclusivamente para almacenar y servir ficheros. Los equipos que funcionan como dispositivo NAS suelen incluir un sistema operativo específico para el propósito, como puede ser FreeNAS o FreeBSD.

Estos sistemas pueden contener uno o más discos duros, dispuestos a menudo en contenedores lógicos redundantes o arrays RAID.

NAS utiliza protocolos basados en archivos como NFS (sistemas UNIX), SMB / CIFS (Server Message Block/Common Internet File System) (sistemas MS Windows), o AFP (Apple Filing Protocol, sistemas Apple Macintosh). Las unidades NAS no suelen limitar a los clientes a un único protocolo. FTP, SFTP, HTTP, UPnP, rsync y AFS (Andrew File System) también lo soportan.

De este modo se consigue optimizar la utilización del almacenamiento y las migraciones de archivos sin interrupciones.

57.2.3 *Diferencias entre NAS y SAN*

NAS proporciona almacenamiento y un sistema de archivos, lo que suele contrastar con SAN, que solamente proporciona almacenamiento basado en bloques y deja del lado del cliente la gestión del sistema de archivos.

NAS aparece en el sistema cliente como un servidor de archivos (se pueden asignar unidades de red a las acciones del servidor) mientras que un disco a través de una SAN se presenta al cliente como un disco más del sistema operativo, que podemos montar, desmontar, formatear...

	NAS	SAN
Tipo de datos	Archivos compartidos	Datos a nivel de bloque, por ejemplo, bases de datos.
Cableado utilizado	Ethernet LAN	Fibre Channel dedicado
Clientes principales	Usuarios finales	Servidores de aplicaciones
Acceso a disco	A través del dispositivo NAS (IP propia)	Acceso directo
	NAS	
		SAN

57.3 VIRTUALIZACIÓN DEL PUESTO CLIENTE

Esta técnica consiste en la separación del entorno de usuario de un ordenador personal de la máquina física con el modelo cliente-servidor. El modelo que sigue un servidor para implementar dicha característica se

denomina VDI (Virtual Desktop Infrastructure, NAS Infraestructura de Escritorio SAN Virtual), también llamada Interfaz de Escritorio Virtual.

La mayoría de implementaciones comerciales de esta tecnología usan un servidor central remoto para llevar a cabo la “virtualización” del escritorio del cliente, en lugar de usar el almacenamiento local del cliente remoto. Esto implica que todas las aplicaciones, procesos, configuraciones y datos del cliente están almacenadas en el servidor y se ejecutan de forma centralizada.

El sistema cliente puede utilizar una arquitectura de hardware completamente diferente de la utilizada por el entorno de escritorio

proyectado, y también puede estar basada en un sistema operativo completamente diferente.

El modelo de virtualización del puesto cliente permite el uso de máquinas virtuales para que múltiples suscriptores de red puedan mantener escritorios individuales en un solo ordenador, el servidor central. Este servidor central puede operar en una residencia, negocio o centro de datos. Los usuarios pueden estar geográficamente dispersos, pero todos están conectados a la máquina central por una red de área local, una red de área amplia, o Internet.

57.3.1 *Modos de Operación de VDI*

Básicamente existen cuatro modelos de operación VDI:

- Alojado (como servicio). Suelen contratarse a proveedores comerciales y normalmente proporciona una configuración del sistema operativo del puesto cliente administrado. Los principales suministradores son CITRIX, VMware y Microsoft.
- Centralizado. En este caso todas las instancias VDI están alojadas en uno o más servidores centralizados, los datos están en sistemas de almacenamiento conectados a estos. Este modelo a su vez puede distinguir dos tipos:
 - o VDI estático o persistente. Existe una única imagen de escritorio asignado por cliente y estos deben ser gestionados y mantenidos.
 - o VDI dinámico o no persistente. Existe una imagen maestra común para todos los clientes que se clona y personaliza en el momento de la petición con los datos y aplicaciones particulares de cada cliente.
- Remoto (o sin ataduras). Tiene como base el concepto de VDI centralizado pero permite trabajar sin la conexión a un servidor

central o a Internet. Se copia una imagen al sistema local y se ejecuta sin necesidad de más conexión. Las imágenes tienen un cierto periodo de vida y se actualizan periódicamente. Esta imagen se ejecuta en el sistema local que necesita un sistema operativo y un hipervisor (que ejecuta la instancia VDI). Esto implica que el dispositivo cliente tenga mayores necesidades de memoria, espacio en disco, CPU... La ventaja es la menor dependencia de conexión.

Los modelos alojado y centralizado necesitan de una red que conecte con el servidor donde se ejecuta la instancia VDI. El concepto base de este modelo es similar al de clientes ligeros debido a que el cliente sólo tiene que mostrar el escritorio virtual.

En el caso del modelo remoto, se permite a los usuarios copiar la instancia VDI en el sistema y luego se ejecutará el escritorio virtual sin necesidad de ningún tipo de conexión.

57.4 COMPUTACIÓN BASADA EN SERVIDOR

También conocida como SBC del inglés Server Based Computing, consiste en la separación del procesamiento de ciertas tareas como la gestión de datos que será realizado en un servidor central y otras tareas de procesamiento, como la presentación de aplicaciones de usuario e impresión de datos en el cliente. Lo único transmitido entre servidor y cliente son las pantallas de información. Esta arquitectura puede dar solución a los principales problemas que aparecen cuando se ejecutan aplicaciones en los clientes. Además simplifica procesos como pueden ser los entornos hardware, actualizaciones de software, despliegue de aplicaciones, soporte técnico, almacenamiento y respaldo de datos. Se centraliza la gestión de todos estos procesos en un único servidor.

Los clientes que actúan en esta arquitectura suelen llamarse thin clients, o clientes ligeros, este es un término general para dispositivos que se basan en un servidor para operar. El thin client proporciona pantalla, teclado,

ratón y un procesador básico que interactúa con el servidor. Los thin client no almacenan ningún dato localmente y requiere de pocos recursos de procesamiento. La característica más destacada de estos terminales es la reducción de costes asociados con el mantenimiento, administración, soporte, seguridad e instalación de aplicaciones comparándolo con un PC tradicional.

Esta tecnología está compuesta por tres componentes principales:

- Sistemas operativos multi-usuario que permiten el acceso y ejecución de modo concurrente, usando aplicaciones diferentes y con sesiones de usuario protegidas. Ejemplos de algunas terminales de servicio son: 2x Terminal Server para Linux, Microsoft Windows Terminal Server (Windows NT/2000), Microsoft Windows Terminal Services (Windows 2003), Citrix Presentation Server, Citrix XenApp Server, AppliDis Fusion, 2X Application Server, HOblink, Propalms TSE (antes Tarantella), Jethro cabina, GraphOn GO-Global, VMware View..
- El thin client se puede ejecutar con una cantidad mínima de software pero necesita al menos un programa de conexión a servicios de terminal. El thin client y el programa de servicios de terminal pueden ser ejecutados en sistemas operativos completamente diferentes.
- Un protocolo que permita al programa de servicios de terminal y al thin client comunicarse y enviar las pulsaciones de teclado, de ratón y las actualizaciones de pantalla a través de la red. Los protocolos más populares son RDP3 (Remote Desktop protocol), ICA y NX.

Entre las ventajas de la computación basada en servidor se puede nombrar:

- Reducción de los costes de administración. La gestión de clientes ligeros está casi en su totalidad centralizada en el servidor.

- Reducción de costes de hardware. El hardware en los clientes ligeros es generalmente más barato porque no es necesario tener memoria para las aplicaciones o un procesador de gran alcance.
- Seguridad. Puede ser controlada centralmente.
- Menor consumo de energía. El hardware especializado en el cliente ligero tiene un consumo mucho menor de energía que los tradicionales.
- Reducción de la carga de red. El tráfico de red que generan los terminales ligeros sólo es el de los movimientos del ratón, teclado e información de pantalla desde / hacia el usuario. En el caso de que un cliente pesado abriese y guardase un documento ya implicaría el paso de este dos veces por la red. Usando protocolos eficientes de red tales como ICA y NX ya es posible usar esta tecnología en un ancho de banda de 28,8 Kbps.
- Actualización de hardware simple. Si el uso está por encima de un límite predefinido, es relativamente sencillo solucionar el problema, bastaría con un disco nuevo en un rack de servidores, aumentando así el número de recursos, exactamente la cantidad necesaria. Si ocurriese esto con clientes pesados habría que reemplazar un PC completo, lo que acarrearía tanto costes económicos como de recursos humanos.

A pesar de lo anterior, esta tecnología también presenta ciertos inconvenientes:

- Altos requerimientos de servidor. Al centrarse la carga de trabajo en el servidor, el sistema de clientes ligeros implica mayor consumo de recursos en los servidores, incluso es habitual que se use un gran número de servidores, lo que se denomina “granja de servidores”.

- Pobre rendimiento multimedia. El envío de datos de audio y video requieren mucho ancho de banda, por lo que estos sistemas son menos útiles para aplicaciones multimedia.
- Menos flexibilidad. No todos los productos software del mercado pueden funcionar correctamente en un cliente ligero.

57.5 GRID COMPUTING

Arquitectura distribuida y paralela, de ámbito extenso geográficamente, en la que se premia la distribución, y a continuación la paralelización. Sus creadores fueron Ian Foster y Carl Kesselman. Su nombre proviene del paradigma de la red eléctrica (power grid).

Se basa en la compartición, selección y agregación de forma dinámica y en tiempo de ejecución de recursos autónomos, distribuidos geográficamente, dependiendo de criterios como la disponibilidad del hardware, la capacidad transaccional, el rendimiento que se pueda aportar a la solución final, el coste y los criterios de calidad del servicio que el demandante pueda proporcionar y exigir.

La red está formada por un conjunto de ordenadores independientes e interconectados que ponen a disposición del grid los excedentes de su procesamiento individual, es decir, los ciclos de reloj de sus CPUs no aprovechados por ellos mismos, sin poder superar un determinado porcentaje de dedicación configurado individualmente en cada nodo. A partir del porcentaje proporcionado por cada nodo, se virtualiza un recurso computacional único.

Los sistemas basados en grid computing están indicados para atender productividades sostenidas y sostenibles, sin poder nunca superar un determinado umbral. En estos sistemas se garantiza la escalabilidad como un criterio parametrizable. Es posible definir con qué criterio añadimos cada nuevo nodo a la solución final.

Actualmente, el único criterio que se tiene en cuenta es la capacidad de procesamiento (transaccionalidad), pero en el futuro, será posible tener en cuenta criterios más finos, referidos a la calidad del servicio.

Además, estos sistemas están dotados de un comportamiento dinámico, según el cual, un determinado programa en ejecución en el sistema, puede modificar en tiempo real el dimensionamiento de la grid para adaptarlo a sus necesidades.

57.5.1 Características:

- Podemos conseguir un máximo aprovechamiento de los nodos (100% de utilización de la CPU).
- Los nodos no tienen que estar dedicados. Además, al contrario que en el caso del cluster, nos aseguramos que la aportación al Grid no va a sobrepasar un determinado porcentaje de tiempo de procesamiento en cada nodo.
- Son sistemas heterogéneos, en los que podemos encontrar diversos HW y SW.
- La escalabilidad parametrizable es la característica más potente de esta arquitectura.

57.5.2 Funcionalidades:

- Localización dinámica de recursos (máquinas con excedente).
- Optimización del acceso a datos, mapeando las estructuras de datos en cachés temporales locales (directorios).
- Autenticación del usuario (usr/pwd, certificados...).
- Monitorización de tareas y procesos desde cualquier nodo de la red, siempre que el usuario tenga permisos.
- Las máquinas se encuentran en situación paritaria.

- Si es posible, se paraleliza. Lo fundamental es la distribución de procesos débilmente acoplados.

57.5.3 *Arquitectura Grid*

Habitualmente se describe la arquitectura del Grid en términos de "capas", ejecutando cada una de ellas una determinada función. Como es habitual en este tipo de enfoque, las capas más altas están más cerca del usuario, en tanto que las capas inferiores lo están de las redes de comunicación.

Empezando por los cimientos, nos encontramos con la capa de red, responsable de asegurar la conexión entre los recursos que forman el Grid.

En la parte más alta está la capa de recursos, constituida por los dispositivos que forman parte del Grid: ordenadores, sistemas de almacenamiento, catálogos electrónicos de datos e incluso sensores que se conecten directamente a la red.

En la zona intermedia está la capa "middleware", encargada de proporcionar las herramientas que permiten que los distintos elementos (servidores, almacenes de datos, redes, etc.) participen de forma coordinada en un entorno Grid unificado. Esta capa es la encargada de las siguientes funciones:

Encontrar el lugar conveniente para ejecutar la tarea solicitada por el usuario.

- Optimiza el uso de recursos, que pueden estar muy dispersos.
- Organiza el acceso eficiente a los datos.
- Se encarga de la autenticación de los diferentes elementos.
- Se ocupa de las políticas de asignación de recursos.
- Ejecuta las tareas.
- Monitoriza el progreso de los trabajos en ejecución.

- Gestiona la recuperación frente a fallos.
- Avisa cuando se haya terminado la tarea y devuelve los resultados.

El ingrediente fundamental del middleware son los metadatos (datos sobre los datos), que contienen, entre otras cosas, toda la información sobre el formato de los datos y dónde se almacenan (a veces en varios sitios distintos).

El middleware está formado por muchos programas software. Algunos de esos programas actúan como agentes y otros como intermediarios, negociando entre sí, de forma automática, en representación de los usuarios del Grid y de los proveedores de recursos. Los agentes individuales presentan los metadatos referidos a los usuarios, datos y recursos. Los intermediarios se encargan de las negociaciones entre máquinas (M2M) para la autenticación y autorización de los usuarios y se encargan de definir los acuerdos de acceso a los datos y recursos y, en su caso, el pago por los mismos. Cuando queda establecido el acuerdo, un intermediario planifica las tareas de cómputo y supervisa las transferencias de datos necesarias para acometer cada trabajo concreto. Al mismo tiempo, una serie de agentes supervisores especiales optimizan las rutas a través de la red y monitorizan la calidad del servicio.

En la capa superior de este esquema está la capa de aplicación donde se incluyen todas las aplicaciones de los usuarios, portales y herramientas de desarrollo que soportan esas aplicaciones. Esta es la capa que ve el usuario.

Además, en las arquitecturas más comunes del Grid, la capa de aplicación proporciona el llamado "serviceware", que recoge las funciones generales de gestión tales como la contabilidad del uso del Grid que hace cada usuario.

Para poder hacer todo lo anterior, las aplicaciones que se desarrollen para ser ejecutadas en un PC concreto, tendrán que adaptarse para poder invocar los servicios adecuados y utilizar los protocolos correctos. Igual que las aplicaciones que inicialmente se crearon para funcionar aisladamente se adaptan para poder ser ejecutadas en un navegador Web, el Grid requerirá que los usuarios dediquen cierto esfuerzo a "GRIDizar" sus aplicaciones.

Sin embargo, una vez adaptadas al Grid, miles de usuarios podrán usar las mismas aplicaciones, utilizando las capas de middleware para adaptarse a los posibles cambios en el tejido del Grid.

57.6 CLOUD COMPUTING

Modelo que permite acceso a un conjunto compartido de recursos informáticos configurables a través de la red (por ejemplo, redes, servidores, almacenamiento, aplicaciones y servicios) que pueden ser desarrollados y desplegados rápidamente con mínimo esfuerzo de gestión o interacción con el proveedor de servicios.

Este término se refiere a la utilización y el acceso de múltiples recursos basados en servidores a través de una red. Los usuarios de la "nube" pueden acceder a los recursos del servidor utilizando un ordenador, netbook, pad computer, smart phone u otro dispositivo. En el cloud computing, el servidor presenta y gestiona las aplicaciones; los datos también se almacenan de forma remota en la configuración de la nube. Los usuarios no descargan ni instalan aplicaciones en su sistema, todo el procesamiento y almacenamiento se mantiene por el servidor. Los servicios on-line pueden ser ofrecidos a partir de un "proveedor de la nube" o por una organización privada.

57.6.1 Arquitectura

Normalmente la arquitectura de los sistemas software implicados en el desarrollo de cloud computing incluyen múltiples componentes

denominados “componentes cloud” que se comunican mediante mecanismos de bajo acoplamiento, tales como las colas de mensajes.

Los dos componentes más significativos de la arquitectura cloud computing se conocen como el front-end y el back-end. El front-end es la parte vista por el cliente, es decir, el usuario del PC. Esto incluye la red del cliente y las aplicaciones utilizadas para acceder a la nube a través de una interfaz de usuario, como un navegador web. El back-end de la arquitectura es la propia nube, que comprende varios ordenadores, servidores y dispositivos de almacenamiento de datos.

Dentro de esta arquitectura se pueden distinguir las siguientes capas:

- Proveedor: Empresa responsable de proporcionar el servicio en la “nube”.
- Cliente: Serán el hardware y software diseñados para cloud computing, que permiten interactuar con los servicios remotos.
- Aplicación: Son los servicios en la “nube” o “Software as a Service” (SaaS), el software se proporciona a través de internet como si de un servicio se tratase. De este modo se evita la necesidad de instalar y ejecutar en el equipo del cliente la aplicación. Se reducen así el mantenimiento y el apoyo.
- Plataforma: Son los servicios de plataforma en la “nube”, también conocidos como “Platform as Service” (PaaS), proporcionan una plataforma de procesamiento y una pila de soluciones como un servicio, constituyen la base e infraestructura de las aplicaciones de la nube. Facilita el desarrollo de aplicaciones evitando el coste y la complejidad de comprar y mantener el hardware y las capas de software de base.
- Infraestructura. Servicios de infraestructura, también conocidos como “Infrastructure as a Service” (IaaS), proporciona la infraestructura

como un servicio, suele ser una plataforma virtualizada. En lugar de comprar servidores, software, centro de datos especiales o equipos de red, los clientes adquieren dichos recursos de servicios externos. La IaaS ha evolucionado a partir de las ofertas de servidores virtuales privados.

57.6.2 Modelos de implementación

- Nube pública o external cloud: Es el concepto tradicional donde los recursos se presentan a través de internet en función de la demanda, a través de aplicaciones o servicios web.
- Nube de la comunidad: Se da cuando varias organizaciones con las mismas necesidades comparten recursos. En este caso existen menos usuarios que en la nube pública y se ofrece mayor privacidad y seguridad. Un ejemplo puede ser el Google's "Gov Cloud".
- Nube híbrida. Es común que una empresa use tanto la nube pública como desarrollos privados para satisfacer sus necesidades con respecto a las TI. Existen varias empresas como HP, IBM, Oracle and VMware que ofrecen tecnologías para manejar la complejidad de mantenimiento, seguridad y privacidad consecuencia del uso del conjunto de estos servicios.
- Nube combinada. Se denomina al conjunto formado varios servicios cloud de distintos proveedores.
- Nube privada. Es trasladar el concepto de nube pública a una red de uso privado. Es decir, el uso de la nube única y exclusivamente dentro de la red de una empresa.

57.7 GREEN IT E EFICIENCIA ENERGÉTICA

El término Green Computing se acuñó posiblemente por primera vez tras el inicio del programa Energy Star en 1992, promocionado por el gobierno estadounidense.

Tenía por objetivo etiquetar monitores y equipamiento electrónico caracterizados por su eficiencia energética. El término quedó registrado ya en 1992 en un grupo de noticias. Hoy en día el programa Energy Star es el motor de la eficiencia energética en los sistemas electrónicos (no sólo de procesamiento de la información, sino también del equipamiento electrónico doméstico).

La adopción de productos y aproximaciones más eficientes pueden permitir más equipamiento dentro del mismo gasto energético, lo que se denomina huella energética, o energy footprint. Las regulaciones se están multiplicando y podrían limitar seriamente a las empresas a la hora de construir centros de procesamiento de datos, ya que el efecto de las redes de suministro eléctrico, las emisiones de carbono por el incremento de uso y otros impactos medioambientales están siendo investigadas. Por tanto, las organizaciones deben considerar las regulaciones y tener planes alternativos para el crecimiento de sus centros de procesamiento de datos y de su capacidad.

Con el paso de los años, el número de servidores existentes en todo el mundo crece de forma casi exponencial. Consecuencia de esto es el creciente gasto energético para la refrigeración y gestión de los equipos. Hoy en día ya se están empezando a plantear soluciones que optimicen este gasto energético.

Este consumo energético no es el único problema ambiental relacionado con las TI. La etapa de fabricación de equipos presenta serios problemas relacionados con el medio ambiente: materiales de desecho tóxicos, producción de gases contaminantes, etc. La tendencia actual es la de minimizar el impacto contaminante (carbon footprint) presente en las tecnologías de fabricación de los sistemas electrónicos.

Finalmente, también tiene un impacto inmediato la eliminación de equipos para las TI, caracterizados por un tiempo de vida increíblemente breve de unos dos o tres años. Si no se reciclan de forma eficiente, terminan tirados

en vertederos, y debido a la presencia de componentes tóxicos, son una fuente de contaminación terrestre y de las aguas. Todos estos aspectos deben ser considerados de manera global por los fabricantes y usuarios de equipos TI. La concienciación de la existencia de este problema ha llevado a la elaboración de numerosas y rígidas normativas a todos los niveles, lo que empieza a obtener algunos resultados.

GreenPeace Internacional realiza un ranking con los 18 principales fabricantes del sector electrónico (ordenadores personales, teléfonos móviles etc.) de acuerdo con sus políticas de reducción de emisiones tóxicas, reciclado o minimización de impacto en el cambio climático, y lo publica en su Guía para la Electrónica Verde (Guide to Greener Electronics), de publicación trimestral. Como se puede ver en los resultados de Diciembre de 2010, las empresas del sector obtienen unas calificaciones realmente bajas, siendo la mejor Nokia con un 7,5 sobre 10.

La mitad de estas 18 empresas suspenden un estudio que busca que las empresas analizadas:

- Limpian sus productos al eliminar sustancias peligrosas. Los productos químicos peligrosos con riesgo impiden el posterior reciclado de los equipos.
- Reciclen de equipos/productos bajo su responsabilidad una vez quedan obsoletos.
- Reduzcan el impacto climático debido a sus operaciones y productos.

Por todo lo expuesto, la resolución efectiva del impacto ambiental de las tecnologías TI requiere un enfoque holístico del problema que englobe las cuatro vías:

- Utilización ecológica: principalmente a través de la reducción del consumo energético. La producción de energía eléctrica es la principal fuente de generación de gases de efecto invernadero.
- Diseño ecológico o eco-diseño: incluye diseño de equipos más eficientes energéticamente y respetuosos con el medio ambiente.
- Fabricación ecológica: eliminando completamente o minimizando el impacto del proceso de fabricación en el medio ambiente (emisiones, materiales de desecho, etc.).
- Eliminación ecológica: una vez finalizado el período de utilización de un equipo se deben poner en marcha las estrategias denominadas tres R: reutilización y renovación de equipos y, si no son aprovechables, reciclado.

La idea principal del enfoque holístico es que se cierre el ciclo de vida de los equipos TI de forma que no se perjudique el medio ambiente, lo que permitiría conseguir una mejora sustancial de cara al desarrollo sostenible.

57.7.1 *Tecnologías verdes*

Hoy en día existen distintos enfoques tecnológicos que se acercan a un desarrollo sostenible de las TI.

- Monitores LCD. Con el paso de los años los monitores pasaron de ser CRT a LCD, este cambio no es sólo estético o de tamaño, sino que los niveles de consumo han disminuido notablemente. Un monitor CRT medio requiere 85W si está activo, frente a los 15W de uno LCD, 5W en modo bajo consumo para un CRT mientras que un LCD consumiría 1,5W. Apagados ambos consumirían 0,5W. En los últimos años se ha revolucionado el mercado de las pantallas de ordenador con la aparición de la tecnología OLED (Organic Light Emitting Diode), basadas en la utilización de diodos LED cuya capa electro-luminiscente se hace con un compuesto orgánico (un polímero que se

ilumina al aplicarle un voltaje). La ventaja principal de este tipo de pantallas frente a las tradicionales de cristal líquido (LCD) es que los diodos OLED no necesitan retro-iluminación, por lo que el consumo de energía que requieren es muy inferior.

- **Discos duros.** El consumo de los discos duros no es para nada despreciable, sobre todo en el arranque del sistema. Por ejemplo, el disco Seagate Barracuda 7200.8 requiere hasta 2,5 A de la línea de alimentación de 12 V. Si a esto le sumamos 3W que extrae desde la línea de +5 V se puede llegar a un consumo de pico en el arranque de 33 W. Si en lugar de sólo un disco duro hablamos de un equipo con dos o más empezamos a hablar de cifras muy comprometidas. Esto ha hecho que los fabricantes de discos duros comiencen a tener en cuenta el consumo en sus productos, creando casi todos una nueva gama denominada “verde” o “ecológica”. Por ejemplo Así, Western Digital con “Caviar Green”, Samsung con Eco Green, o Hitachi con eco-friendly Deskstar y Travelstar. Como alternativa a los discos tradicionales aparecen los discos en estado sólido (SSD), que presentan menores consumos de energía y es la tecnología a la que se espera evolucionen los sistemas de almacenamiento.
- **CPDs.** Aquí es donde se aloja toda la infraestructura de soporte a los diversos servicios computacionales, y una estructura adecuada permitirá buenos ahorros de energía, de espacio y de costos a mediano y/o largo plazo. Buscando la reducción de energía se puede empezar por la acción más simple que es apagar el equipo que no se esté utilizando, la reducción del hardware estudiando necesidades reales, o actuaciones específicas en función de la actividad de la empresa.
- **Virtualización.** La virtualización de servidores permite el funcionamiento de múltiples servidores en un único servidor físico. Esto ayuda a reducir la huella de carbono del centro de datos al

disminuir el número de servidores físicos y consolidar múltiples aplicaciones en un único servidor con lo cual se consume menos energía y se requiere menos enfriamiento. Además se logra un mayor índice de utilización de recursos y ahorro de espacio.

- Cliente/Servidor. Estos sistemas mantienen el software, las aplicaciones y los datos en el servidor. Se puede tener acceso a la información desde cualquier ubicación y el cliente no requiere mucha memoria o almacenamiento. Este ambiente consume menos energía y enfriamiento.
- Cloud computing. Esto proporciona a sus usuarios la posibilidad de utilizar una amplia gama de recursos en red para completar su trabajo. Al utilizar computación en nube las empresas se vuelven más ecológicas porque disminuyen su consumo de energía al incrementar su capacidad sin necesidad de invertir en más infraestructura.
- Tele trabajo. Definido por Merrian-Webster como el trabajo en casa con el uso de un enlace electrónico con la oficina central. Al no desplazarse el empleado la contaminación es menor.

57.7.2 *Actividades relacionadas con Green IT*

Existen varias actividades que promocionan e intentan solventar las cuestiones expuestas anteriormente. Estas actividades están patrocinadas bien desde administraciones públicas, bien desde empresas, que están entendiendo que Green IT, además de una necesidad, puede ser un negocio, desde el punto de vista de consultoría y servicios, o bien por consorcios de empresas.

The Green Grid (<http://www.thegreengrid.org>) es un consorcio global dedicado a avanzar en la eficiencia energética de los centros de procesamiento de datos y en ecosistemas de computación de negocio. En cumplimiento de su misión, The Green Grid se centra en:

- Definir métricas y modelos significativos y centrados en el usuario.
- Desarrollar estándares, métodos de medida, procesos y nuevas tecnologías para mejorar el rendimiento de los centros de procesamiento de datos frente a las métricas definidas.
- Promocionar la adopción de estándares, procesos, medidas y tecnologías energéticamente eficientes.

El comité de directores de The Green Grid está compuesto por las siguientes compañías miembros: AMD, APC, Dell, HP, IBM, Intel, Microsoft, Rackable Systems, Sun Microsystems y VMware.

Climate Savers. Iniciada por Google e Intel en 2007, Climate Savers Computing Initiative (www.climatesaverscomputing.org) es un grupo sin ánimo de lucro de consumidores y negocios con conciencia ecológica y organizaciones conservacionistas. La iniciativa se inició bajo el espíritu del programa Climate Savers de WWF (<http://www.worldwildlife.org/climate/projects/climateSavers.cfm>), que ha movilizado a una docena de compañías desde 1999 a recortar las emisiones de dióxido de carbono, demostrando que reducir las emisiones es bueno para el negocio. Su objetivo es promover el desarrollo, despliegue y adopción de tecnologías inteligentes que puedan mejorar la eficiencia de uso de la energía del computador y reducir su consumo cuando el computador se encuentra inactivo.

SNIA (Storage Networking Industry Association, <http://www.snia.org>) es una organización global sin ánimo de lucro compuesta por unas compañías de la industria del almacenamiento. SNIA Green Storage Initiative (<http://www.snia.org/green>) está llevando a cabo una iniciativa para avanzar en el desarrollo de soluciones energéticamente eficientes para el almacenamiento en red, incluyendo la promoción de métricas estándares, la formación y el desarrollo de buenas prácticas energéticas o el establecimiento de alianzas con organizaciones como The Green Grid.

Energy Star. En 1992 la Agencia de Protección Medioambiental de EEUU (U.S. Environmental Protection Agency) lanzó el programa Energy Star, que se planificó para promocionar y reconocer eficiencia energética en monitores, equipos de climatización y otras tecnologías. Aunque de carácter voluntario inicialmente, resultó pronto de amplia aceptación, pasando a ser un hecho la presencia de un modo de descanso (sleep mode) en la electrónica de consumo.

Directiva Europea de Eco-Diseño. Siguiendo la misma línea que la iniciativa Energy Star de EEUU, la Unión Europea aprobó la directiva 2005/32/EC para el eco-diseño, nuevo concepto creado para reducir el consumo de energía de productos que la requieren, tales como los dispositivos eléctricos y electrónicos o electrodomésticos. La información relacionada con las prestaciones medioambientales de un producto debe ser visible de forma que el consumidor pueda comparar antes de comprar, lo cual está regulado por la Directiva de Etiquetado de la Energía (Energy Labelling Directive). Los productos a los que se conceda la Eco-etiqueta serán considerados como cumplidores con la implementación de las medidas, de forma muy similar a la etiqueta de Energy Star.

El Código de Conducta de la Unión Europea para Centros de Datos está siendo creado como respuesta al creciente consumo de energía en centros de datos y a la necesidad de reducir el impacto ambiental, económico y de seguridad de abastecimiento energético relacionado. El objetivo es informar y estimular a los operadores o propietarios de los centros de datos a que reduzcan el consumo de energía de una forma rentable sin dificultar su funcionamiento. Este código de conducta quiere conseguir esto mediante la mejora de la comprensión de la demanda de energía dentro del centro de datos, aumentando la concienciación, y mediante la recomendación de prácticas y objetivos energéticamente eficientes.

Grupo de trabajo sobre Green IT de la plataforma INES (Iniciativa Española de Software y Servicios, <http://www.ines.org.es>) es la Plataforma

Tecnológica Española en el área de los Sistemas y Servicios Software y constituye una red de cooperación científico-tecnológica integrada por los agentes tecnológicos relevantes de este ámbito (empresas, universidades, centros tecnológicos, etc.).

Según la Agenda Estratégica de Investigación de INES, el plan de dinamización para el Grupo de Trabajo de Green IT consiste en las siguientes acciones:

- Análisis de la influencia e importancia de las soluciones de Green IT.
- Difusión de las informaciones, noticias y existencia de este grupo de trabajo por Internet.
- Fomentar el interés y apoyar el desarrollo bajo Green IT.

Big Green Innovations (<http://www.ibm.com/technology/greeninnovations/>), programa de IBM. Dentro de este programa, y con fines educativos, IBM ha presentado un centro de datos virtual ecológico denominado Virtual Green Data Center.

La lista Green500 (<http://www.green500.org>) proporciona una clasificación de los supercomputadores más eficientes energéticamente del mundo, sirviendo como una visión complementaria a la lista Top500 (<http://www.top500.org>).

Otras empresas, como Google, Dell o Symantec, están desarrollando programas de eficiencia energética, tanto para sus propios procesos de TI como para los de sus clientes.

57.8 BIBLIOGRAFÍA

- Windows Server 2008 Hyper-V : kit de recursos. Larson, Robert. Anaya, D.L. 2009

- Grid computing : experiment management, tool integration, and scientific workflows. Prodan, Radu Berlin: Springer, cop. 2007
- Virtualización na Wikipedia: <http://es.wikipedia.org/wiki/Virtualizaci%C3%B3n>
- Green IT: Tecnologías para la Eficiencia Energética en Sistemas TI. Marisa López-Vallejo, Eduardo Huedo Cuesta y Juan Garbajosa Sopeña.
- Dot-cloud : the 21st century business platform built on cloud computing. Fingar, Peter Tampa (FL) : Meghan-Kiffer Press, cop. 2009

Autor: Francisco Javier Rodriguez Martínez

Subdirector de Sistemas Escola Superior Enxeñaría Informática Ourense

Colegiado del CPEIG