

## Tema 9. Centros de datos: computación en nube y organización física



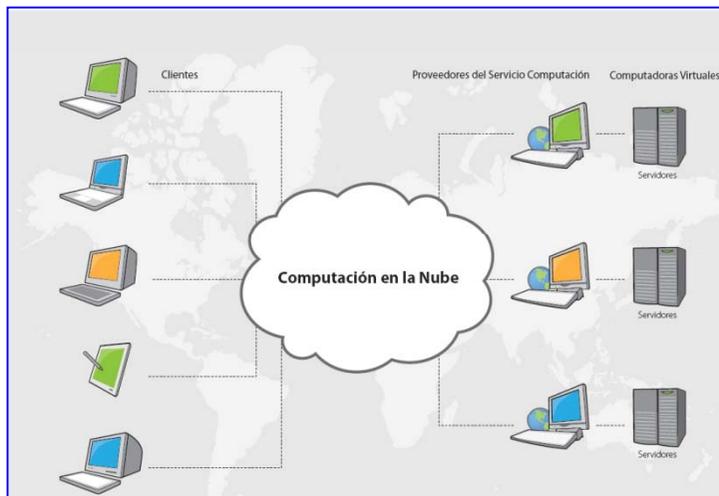
### Indice

1. Introducción a la computación en nube (*cloud computing*)
2. Virtualización de recursos: consolidación de servidores
3. Arquitectura orientada a servicios (SOA): Servicios web
4. Recursos informáticos como servicios
5. Tipos de nubes: públicas, privadas e híbridas
6. Arquitectura de los *Data Center*

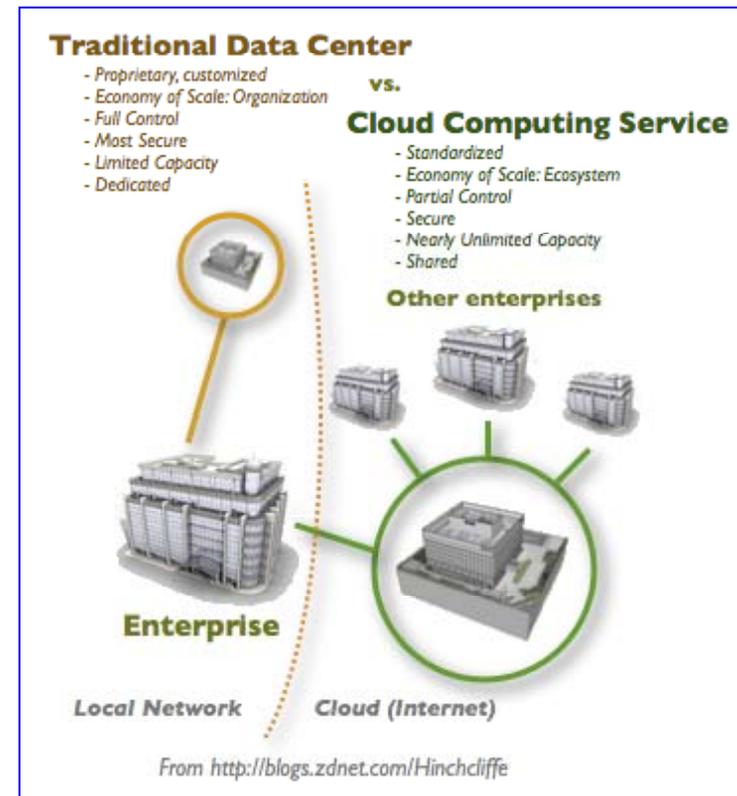
# 1.Introducción a la computación en nube



- La computación en nube (*Cloud computing*) es un paradigma que permite ofrecer servicios de computación a través de Internet. La *nube* es una metáfora de Internet.
- En este tipo de computación todo lo que puede ofrecer un sistema informático se ofrece como **servicio**.
- La nube proporciona aplicaciones comerciales *on-line* accesibles desde un navegador web, con el software y los datos almacenados en servidores en Internet.



- Ejemplos de computación en nube:
  - Amazon EC2
  - Google Apps
  - Windows Azure
  - Ubuntu One



# 1.Introducción a la computación en nube



## Ventajas de la computación en nube

- **El usuario no necesita instalar ningún tipo de hardware.**

Requiere mucha menor inversión para empezar a trabajar que la alternativa convencional.

- **Las actualizaciones son automáticas.**

Conserva las personalizaciones e integraciones automáticamente durante la actualización.

- **Independencia del dispositivo y la ubicación.**

Permite que los usuarios puedan acceder a los sistemas usando un navegador de Internet independiente de su ubicación geográfica, sistema operativo o computador.

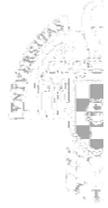
- **Elevada capacidad de adaptación para demandas variables**

Utiliza una infraestructura tecnológica dinámica que se caracteriza por un alto grado de *elasticidad*: permite aumentar o disminuir rápidamente los recursos dependiendo de las necesidades cambiantes de los usuarios.

- **Servicio tarifado**

Cada recurso que consume el usuario es medido no sólo para fines de tarificación sino también de control. El usuario paga únicamente por el tiempo de uso del servicio.

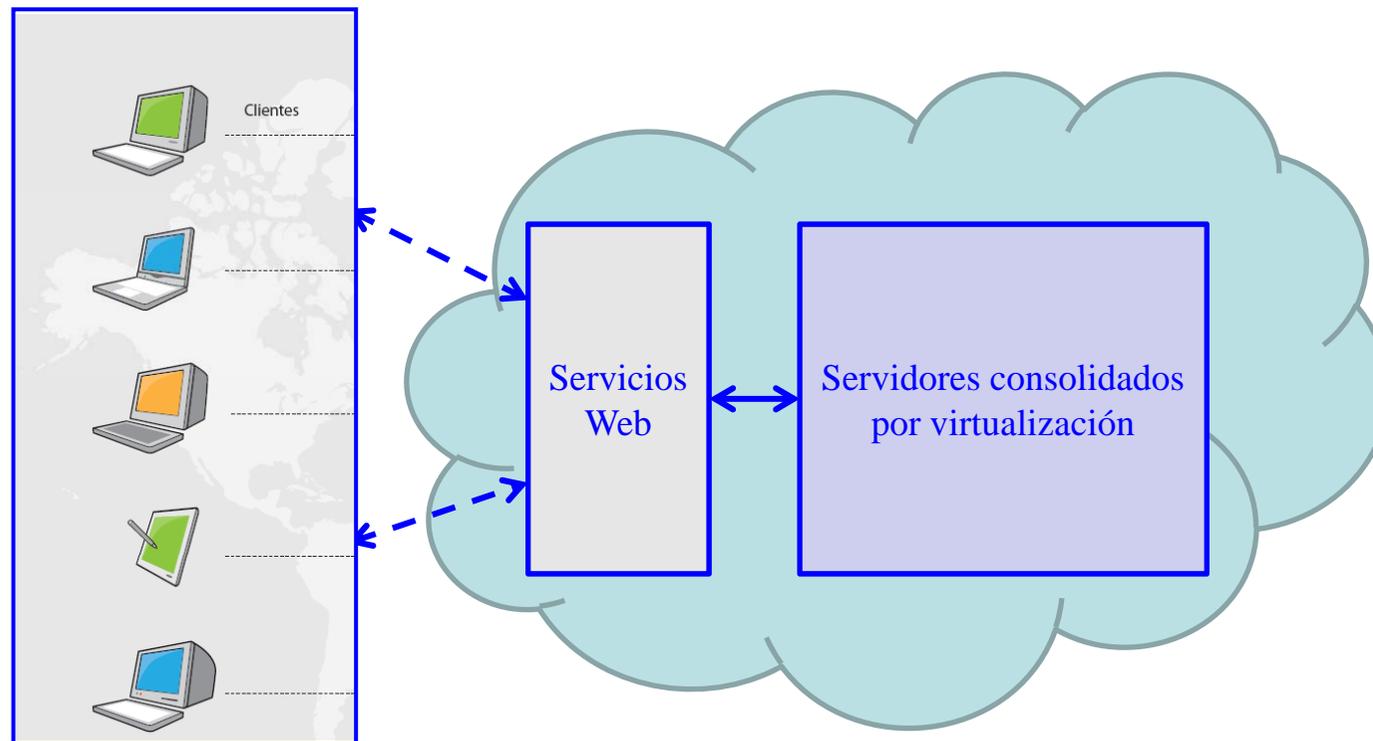
# 1.Introducción a la computación en nube



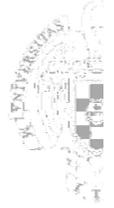
## Tecnologías de soporte para la computación en nube

- Virtualización de recursos
- Servicios Web: SOA (*Service Oriented Architecture*)
- Utilización bajo demanda.

demanda



## 2. Virtualización de recursos



- Es el uso del software para la creación de una versión virtual de algún recurso informático: servidor, sistema operativo, software de servidor, dispositivo de almacenamiento, recursos de red o aplicaciones.
- Se trata de la creación de una capa de abstracción entre el hardware de la máquina física (host) y el sistema operativo de la máquina virtual (guest), que permite dividir el recurso hardware en uno o más entornos de ejecución.
- Esta capa de software gestiona los cuatro recursos principales de una computadora: CPU, Memoria, Almacenamiento y Conexiones de Red; repartiendo dinámicamente dichos recursos entre todas las máquinas virtuales definidas.
- Se pueden tener varios computadores virtuales ejecutándose en el mismo computador físico.
- Puede crear una interfaz externa que encapsula una implementación subyacente mediante la combinación de recursos en localizaciones físicas diferentes.
- Permite la consolidación de servidores: en lugar de tener múltiples servidores físicos que no suelen operar a pleno rendimiento, cada uno dedicado a una carga de trabajo específica, la virtualización del servidor hace posible que esas cargas de trabajo se vean **consolidadas** en un número menor de máquinas pero a pleno rendimiento.

## 2. Virtualización de recursos



### Ventajas de la virtualización

- **Disminución de costes**

Se puede disponer de un solo servidor más potente y no tener que comprar muchos servidores: éstos se crearían como máquinas virtuales.

- **Crecimiento más flexible**

Instalar un servidor virtual es más sencillo y rápido que un servidor físico.

- **Administración simplificada**

Desde la consola del gestor de máquinas virtuales se pueden aumentar o disminuir los recursos asignados a una determinada máquina, reiniciarla, actualizarla o eliminarla.

- **Aprovechamiento de aplicaciones antiguas**

Posibilidad de conservar aplicaciones que funcionan en sistemas antiguos y aun así modernizar la infraestructura informática de la empresa.

- **Centralización de tareas de mantenimiento**

Se pueden realizar copias de seguridad de todas las máquinas, programar actualizaciones y otras actividades desde el gestor de máquinas virtuales.

- **Disminución de los tiempos de parada**

Se puede clonar una máquina y seguir dando servicio mientras se realiza mantenimiento de la máquina virtual de producción como actualizaciones.

- **Balanceo de recursos**

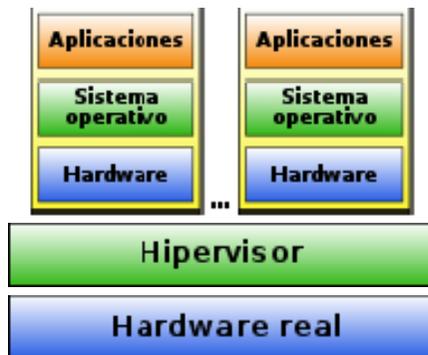
Es posible utilizar una aplicación para proporcionar recursos a las máquinas virtuales que lo necesiten, haciendo un balanceo de los mismos.

## 2. Virtualización de recursos



Un **hipervisor** o **monitor de máquina virtual** (*virtual machine monitor*) es una plataforma que permite aplicar diversas técnicas de control de virtualización. Es una extensión del término *supervisor* que se aplicaba al kernel de un sistema operativo. Esta capa de software (VMM) maneja, gestiona y arbitra los cuatro recursos principales de una computadora (CPU, Memoria, Almacenamiento y Conexiones de Red), repartiendo dinámicamente estos recursos entre todas las máquinas virtuales definidas en el computador central. Pueden clasificarse en dos tipos:

**Hipervisor tipo 1:** También denominado nativo, unhosted o sobre el metal desnudo (bare metal), es un software que se ejecuta directamente sobre el hardware, para ofrecer la funcionalidad descrita.



Algunos de los hipervisores tipo 1 más conocidos son los siguientes:

- VMware ESXi (gratis)
- [VMware ESX](#) (de pago),
- [Xen](#) (libre)
- Citrix XenServer (gratis)
- Microsoft [Hyper-V](#) Server (gratis).

**Hipervisor tipo 2:** También denominado hosted, es un software que se ejecuta sobre un sistema operativo para ofrecer la funcionalidad descrita.



Algunos de los hipervisores tipo 2 más utilizados son los siguientes:

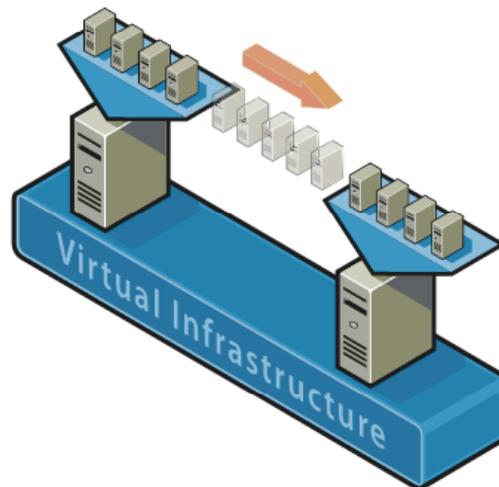
- Oracle: VirtualBox (gratis),
- VirtualBox OSE (libre),
- VMware: Workstation (de pago), Server (gratis), Player (gratis), QEMU (libre),
- Microsoft: Virtual PC, Virtual Server.

## 2. Virtualización de recursos



### Consolidación de servidores

- Es un planteamiento para el uso eficiente de los recursos de un servidor a fin de reducir el número total de servidores de una organización
- Consiste en combinar cargas de trabajo de máquinas diferentes o aplicaciones sobre un número más pequeño de servidores.
- Existen varias formas de consolidación:
  - Trasladar cargas de trabajo heterogéneas de múltiples servidores a un único servidor de mayor capacidad.
  - Combinar bajo un único SO múltiples cargas de trabajo, reduciendo las imágenes de SO.
- La virtualización es la principal tecnología para la consolidación de servidores.

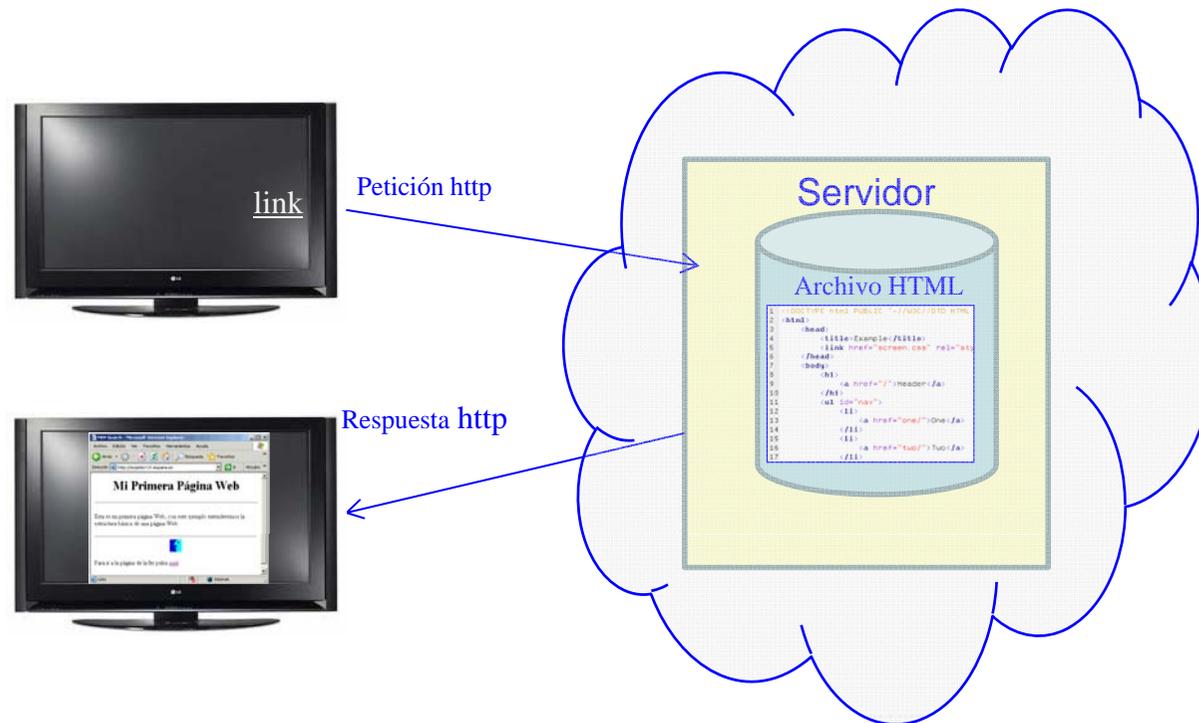


### 3. Servicios WEB



#### Antecedentes de los servicios WEB: páginas estáticas

- La Web se inició como un servicio de páginas estáticas de información multimedia soportado por el lenguaje de representación HTML y el protocolo de Internet HTTP.
- En este esquema un usuario (cliente) se limita a proporcionar al servidor el URL (*Uniform Resource Locator*) de la página que quiere visualizar, y el navegador del cliente transforma la página HTML recibida del servidor en información legible en pantalla.

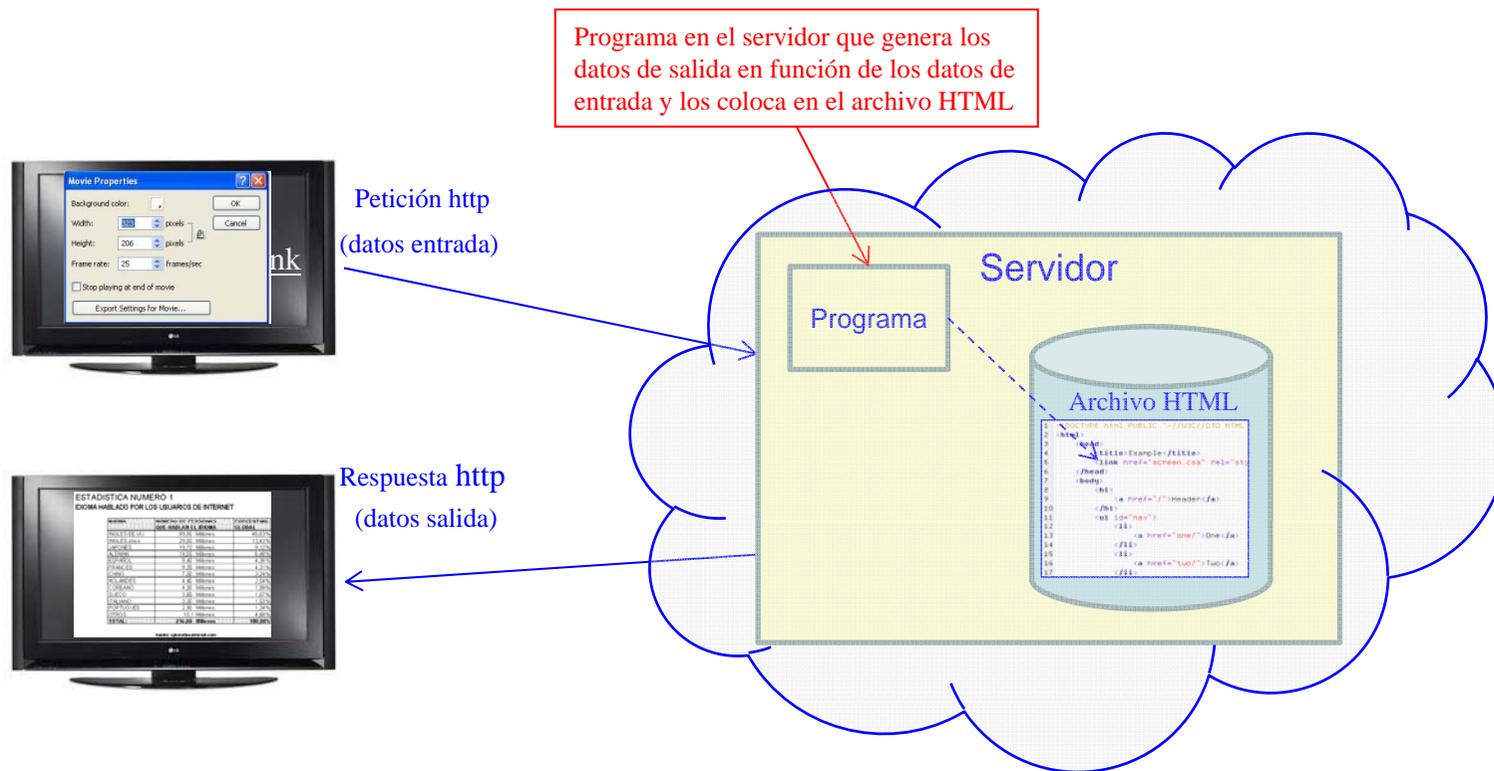


## 3. Servicios WEB



### Antecedentes de los servicios WEB: páginas dinámicas

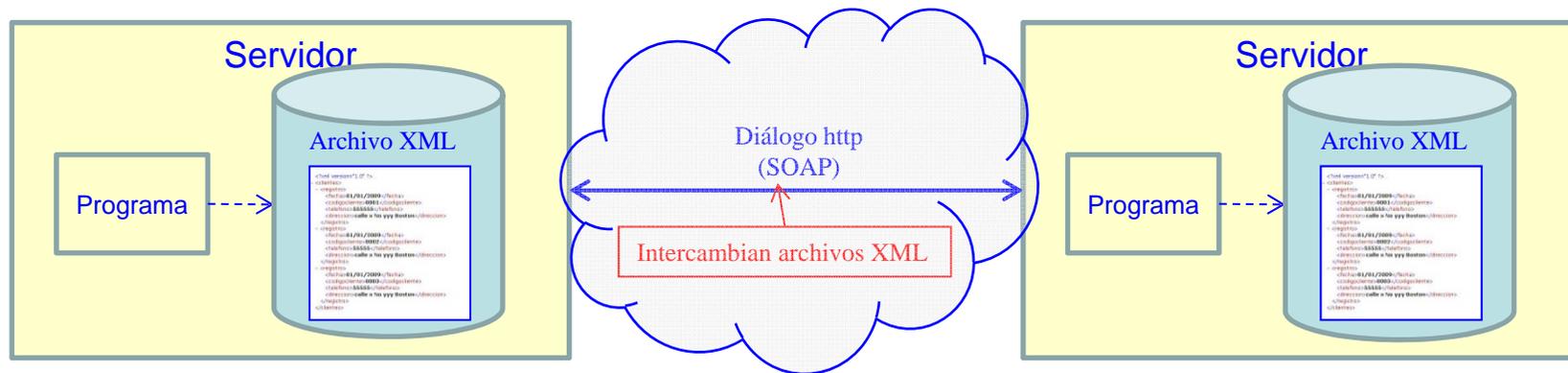
- Posteriormente se añadió la posibilidad de que el servidor generase dinámicamente parte de las páginas (ASP, JSP, PHP) en función de información pasada previamente por el cliente.
- La generación dinámica de contenidos amplía las posibilidades de la Web, al permitir la ejecución remota, a través de un navegador, de programas residentes en el servidor.
- Sin embargo, con esta ampliación el cliente sigue recibiendo una respuesta HTML que sólo es válida para ser visualizada en pantalla.





## Servicios Web

- Los Servicios Web dan un paso más al permitir una comunicación directa entre máquinas clientes y servidores, sin intervención del usuario, a través de la infraestructura de la Web.
- Para ello se sustituye el lenguaje HTML por XML, que permite expresar no sólo el formato de representación de las páginas, sino el significado de los datos que contiene.



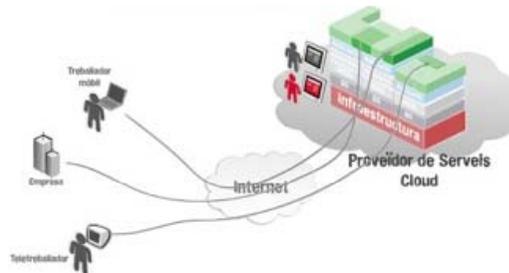
- Los Servicios Web permiten realizar transferencias de información entre programas que se ejecutan en computadores conectados a Internet, utilizando XML como lenguaje común de intercambio y HTTP como vehículo de comunicación.
- De esta forma, los Servicios Web permiten interoperar a las aplicaciones con independencia del lenguaje de programación y plataforma de ejecución utilizados.
- Unas aplicaciones, los *proveedores*, ofrecen unos servicios en forma de procedimientos remotos, y otras, los *clientes*, utilizan estos servicios llamando a los procedimientos a través de la Web.

## 4. Recursos informáticos como servicios



### Computación bajo demanda: suministro de computación (*utility computing*)

- La computación *on-demand* o el modelo *utility* constituye el tercer pilar de la computación en nube.
- Consiste en el tratamiento de los recursos de computación, almacenamiento y redes como servicios cuyo consumo se mide y se paga por la cantidad consumida.
- El procesamiento de información se equipara a los tradicionales suministros públicos: electricidad, agua, gas natural o telefonía
- Este planteamiento tiene la ventaja de un costo bajo o nulo para adquirir los recursos de computación
- Los recursos son alquilados, convirtiendo la necesidad de comprar productos (hardware, software y ancho de banda) en un servicio.
- Las empresas (clientes) acceden a los recursos informáticos a medida que los necesitan, en lugar de a tiempo completo.
- La informática bajo demanda también reduce los gastos generales de gestión.
- Reduce el uso de la energía disminuyendo la huella ecológica de una organización.



## 4. Recursos informáticos como servicios



### Tipos de servicios

- **Software como un Servicio (SaaS)**

Constituye la capa más alta de la nube y se caracteriza porque una aplicación completa es ofrecida como un servicio bajo demanda. Una sola instancia del software corre en la infraestructura del proveedor y sirve a múltiples organizaciones de clientes.

- **Plataforma como un Servicio (PaaS)**

Constituye la capa intermedia de la nube, y se caracteriza porque lo que se ofrece es un entorno de desarrollo y el empaquetamiento de un carga de servicios. Las ofertas de PaaS pueden dar servicio a todas las fases del ciclo de desarrollo y pruebas del software, o pueden estar especializadas en cualquier área en particular, tal como la administración del contenido.

- **Infraestructura como un Servicio (IaaS)**

Constituye la capa inferior de la nube y se caracteriza por ofertar almacenamiento básico y capacidades de computación como servicios estandarizados en la red. Servidores, sistemas de almacenamiento, conexiones, enrutadores, y otros sistemas se concentran para manejar tipos específicos de cargas de.



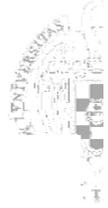
## 4. Recursos informáticos como servicios



### Software como un Servicio (*SaaS*)

TIPO DE SERVICIO	EJEMPLO
Aplicaciones como sitios Web	Box.net (Box.net ), Microsoft Office Live (Microsoft), Facebook (Facebook, Inc.), LinkedIn (LinkedIn Corporation), Twitter (Twitter, Inc.), MySpace (MySpace.com), Zillow (Zillow.com), Google Maps (Google).
Colaboración y aplicaciones de oficina	Cisco WebEx Weboffice (Cisco Systems, Inc.), Google Docs (Google), Google Talk (Google), IBM BlueHouse (IBM, Corp.), Microsoft Exchange Online (Microsoft), RightNow (RightNow Technologies, Inc.), Gmail (Google), Microsoft Hotmail (Microsoft Hotmail), Yahoo! Mail (Yahoo! Inc.).
Servicios de pago	Amazon Flexible Payments Service (Amazon FPS) (Amazon Web Services, LLC), Amazon DevPay (Amazon Web Services, LLC).
Software basado en Web integrable a otras aplicaciones	Flickr Application Programming Interface (API) (Flickr, LLC), Google Calendar API (Google), Salesforce.com's AppExchange (Salesforce.com, Inc.), Yahoo! Maps API (Yahoo! Inc.), Zembly (Sun Microsystems, Inc.).

## 4. Recursos informáticos como servicios



### Plataforma como un Servicio (*PaaS*)

TIPO DE SERVICIO	EJEMPLO
Plataformas de desarrollo	Amazon Simple Queue Service (Amazon SQS) (Amazon Web Services, Amazon Simple Queue Service (Amazon SQS)), Amazon Simple Storage Service (Amazon S3) (Amazon Web Services, LLC), Google App Engine (Google), GRIDS Lab Aneka (Vecchiol, Chu, & Buyya, 2009).
Bases de datos	Amazon SimpleDB (Amazon Web Services, Amazon SimpleDB), Big Table (Chang, y otros, noviembre de 2006), Microsoft SQL Azure Database (Microsoft).
Cola de mensajes	Amazon Simple Queue Service (Amazon SQS) (Amazon Web Services, Amazon Simple Queue Service (Amazon SQS)).
Servidores de aplicaciones	NetSuite Business Operating System (NS-BOS) (NetSuite, Inc.).

## 4. Recursos informáticos como servicios



### Infraestructura como un Servicio (IaaS)

TIPO DE SERVICIO	EJEMPLO
Procesamiento	Amazon Elastic Compute Cloud (Amazon EC2) (Amazon Web Services, LLC), Sun Network.com (Sun Grid) (SUN Microsystems, Inc.), ElasticHost (ElasticHosts Ltd.), Eucalyptus (Nurmi, y otros, 2009), Nimbus (Alliance), OpenNebula (Grupo de Arquitectura Distribuida), Enomaly (Enomaly, Inc.).
Distribución de contenido a través de servidores virtuales	Akamai (Technologies), Amazon CloudFront Beta (Amazon Web Services, LLC).
Almacenamiento	Amazon Simple Storage Service (Amazon S3) (Amazon Web Services, LLC), Amazon SimpleDB (Amazon Web Services, Amazon SimpleDB), Amazon Elastic Block Store (Amazon Web Services, Amazon Elastic Block Store (EBS)), Microsoft SkyDrive (Microsoft Corporation), Flickr (Flickr, LLC), Youtube (YouTube, LLC), Nirvanix Storage Delivery Network (Nirvanix), Microsoft Live Mesh Beta (Microsoft Corporation, 2009), Flickr (Flickr, LLC).
Administración de sistemas	Elastra (Elastra Corporation), Engine Yard (Engine Yard, Inc.), FlexiScalable (XCalibre Communications), Grid Layer (Layered Technologies, Inc.), Joyent (Joyent, Inc.), Mosso (Rackspace, US Inc.), Savvis Virtual Intelligent Hosting (Savvis, Inc.).
Administración de alojamiento	Digital Realty Trust (Digital Realty Trust, Inc.), GoDaddy.com (GoDaddy.com, Inc.), Layered Technology (Layered Technologies, Inc.).

## 5. Tipos de nubes



- Existen tres tipos de nubes:
  - Nubes públicas: las utilizan muchos usuarios compartiendo la misma infraestructura.
  - Nubes privadas: no comparten recursos con ninguna otra organización.
  - Nubes híbridas: permiten un espacio privado para unas aplicaciones y otro público para el resto.
- **Nubes públicas**
  - Son gestionadas por terceras partes
  - Los trabajos de los clientes pueden estar mezclados en los servidores, los sistemas de almacenamiento y otras infraestructuras de la nube.
  - Los usuarios finales no conocen qué trabajos de otros clientes pueden estar corriendo con los suyos propios.
- **Nubes privadas**
  - Son una buena opción para las compañías que necesitan alta protección de datos.
  - Operan en una infraestructura bajo demanda gestionada por un solo cliente que controla las aplicaciones que ejecuta y dónde las ejecuta.
  - Son propietarios del servidor, la red, y los discos, y pueden decidir los usuarios autorizados para utilizar la infraestructura.
- **Nubes híbridas**
  - Combinan los modelos de nubes públicas y privadas.
  - El usuario es propietario de unas partes y comparte otras.

## 6. Arquitectura de los centros de datos



### Centros de datos para *cloud computing*

- Las plataformas hardware que tienen que dar soporte a la computación en nube necesitan unos recursos muy superiores a los centros de datos convencionales.
- Estas instalaciones no pueden diseñarse como una simple colección de servidores ya que el hardware y software tiene que trabajar de manera coordinada
- Estos centros de datos hay que tratarlos como un gran computador del tamaño de un gran almacén (WSC: *warehouse-scale computer* ).
- Los centros de datos tipo WSC dan soporte actualmente a los servicios *online* ofrecidos por compañías tales como Google, Amazon, Yahoo y Microsoft
- Se diferencian de los centros de datos tradicionales en los siguientes hechos:
  - Pertenecen a una misma organización
  - Utilizan una plataforma hardware y software relativamente homogénea
  - Comparten una capa de gestión común del sistema
  - Ejecutan un número más pequeño de aplicaciones muy grandes

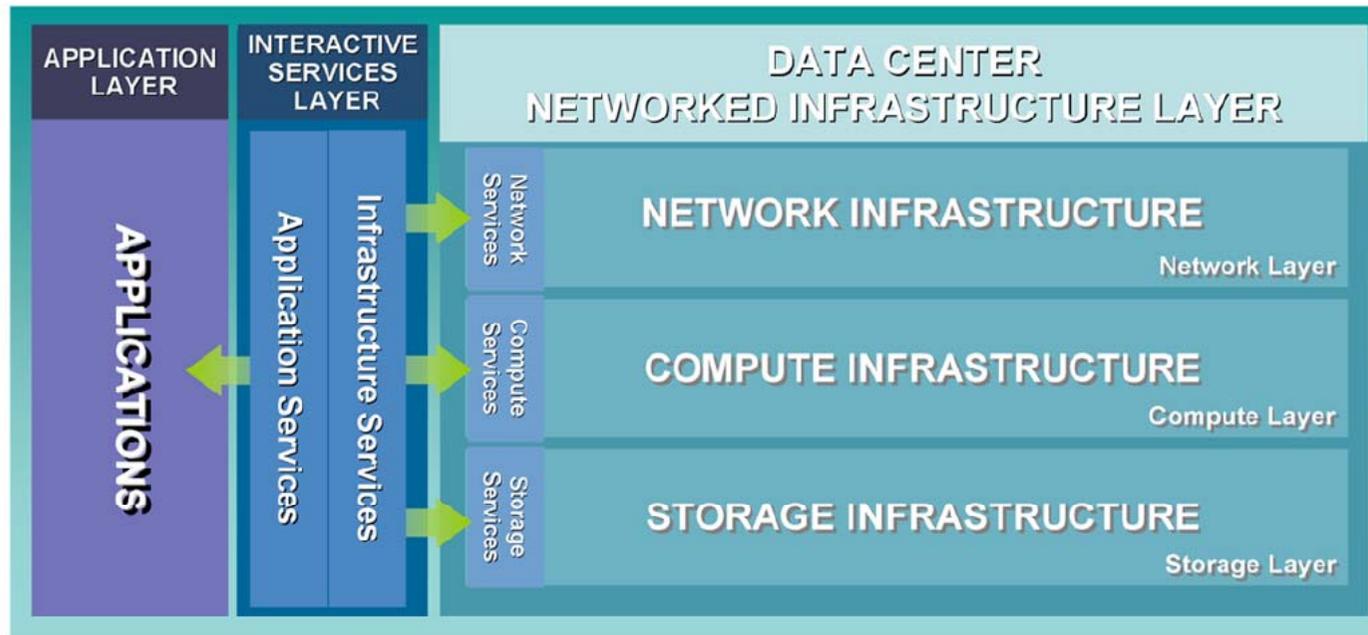


## 6. Arquitectura de los centros de datos



### Arquitectura de un Centro de Datos de tres capas

- La arquitectura de los centros de datos tradicionales constan de tres niveles: red computación y almacenamiento.
  - El nivel de red proporciona acceso seguro y fiable de los usuarios.
  - El nivel de computación proporciona los recursos de proceso.
  - El nivel de almacenamiento proporciona los servicios de bases de datos.

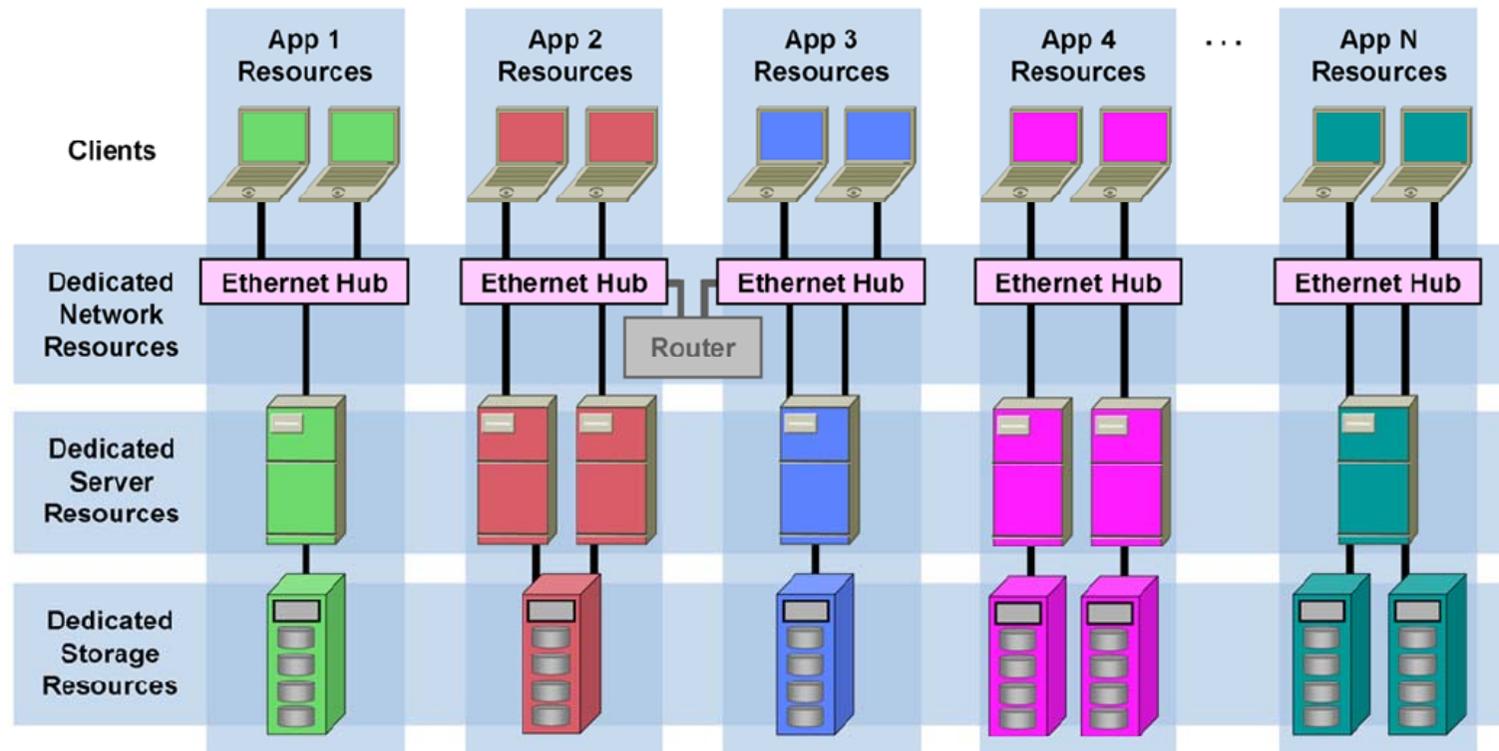


- Los centros de datos para *cloud computing* han seguido un proceso de virtualización de estos tres niveles

## 6. Arquitectura de los centros de datos



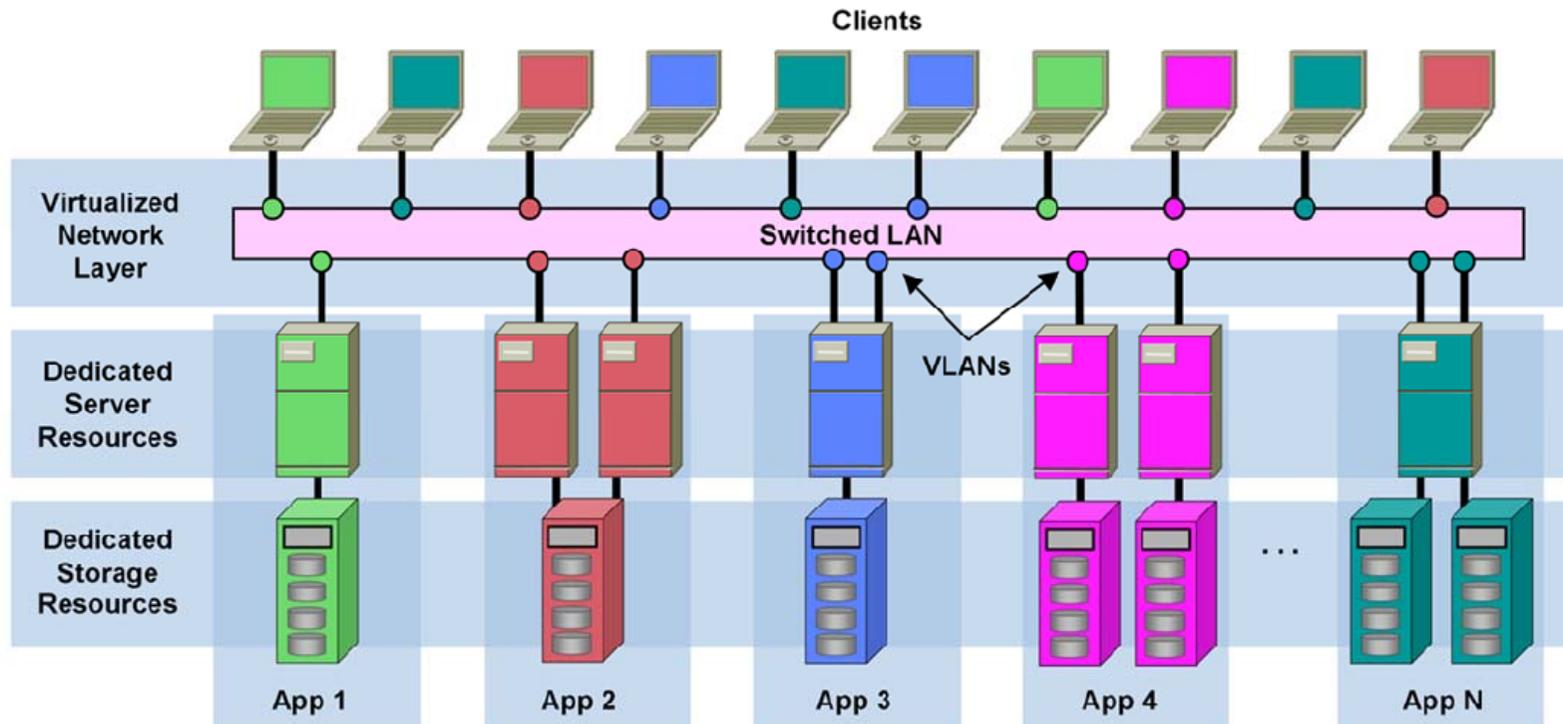
### Arquitectura convencional de un Centros de datos: recursos dedicados



## 6. Arquitectura de los centros de datos



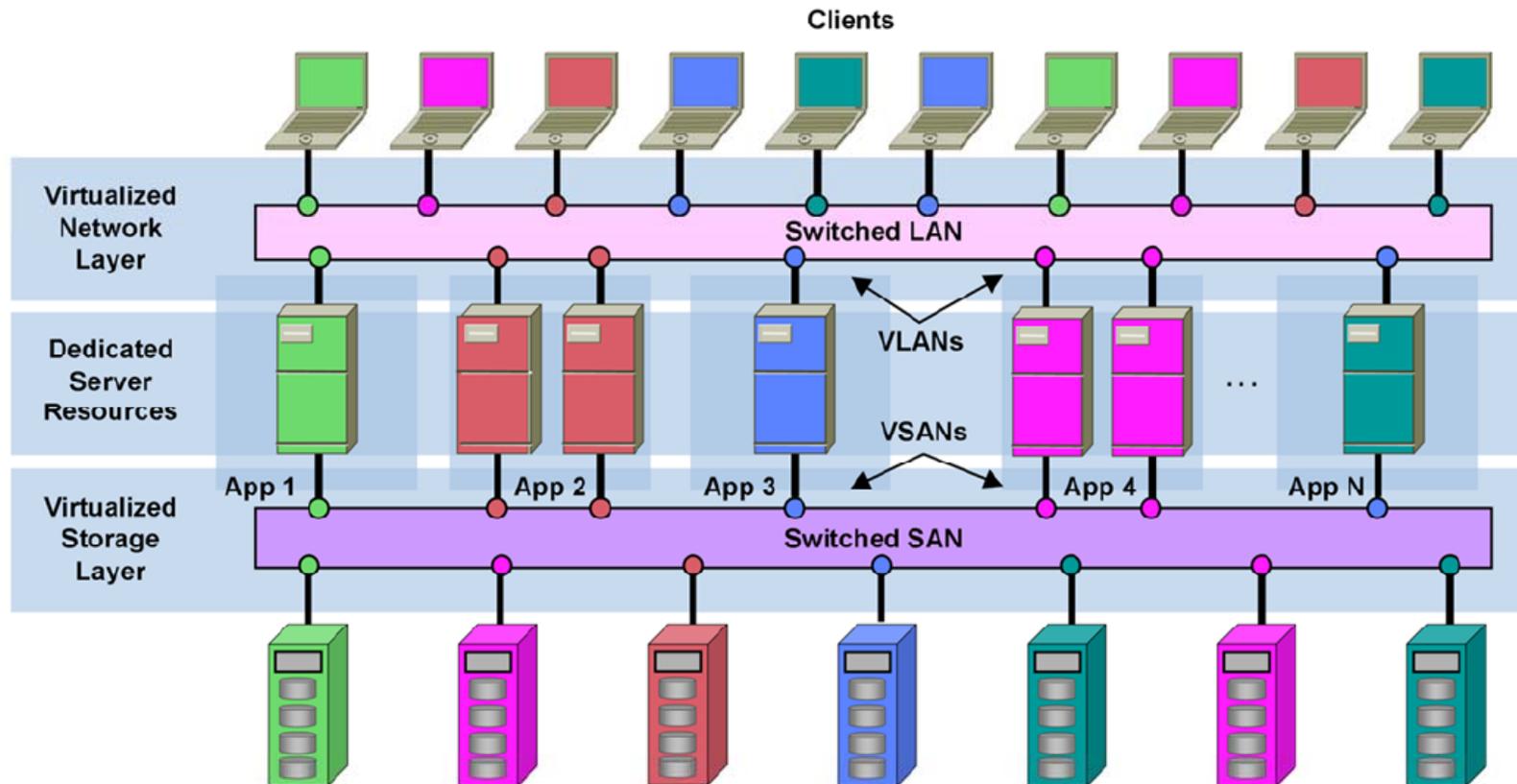
### Centros de datos con la capa de red virtualizada



## 6. Arquitectura de los centros de datos



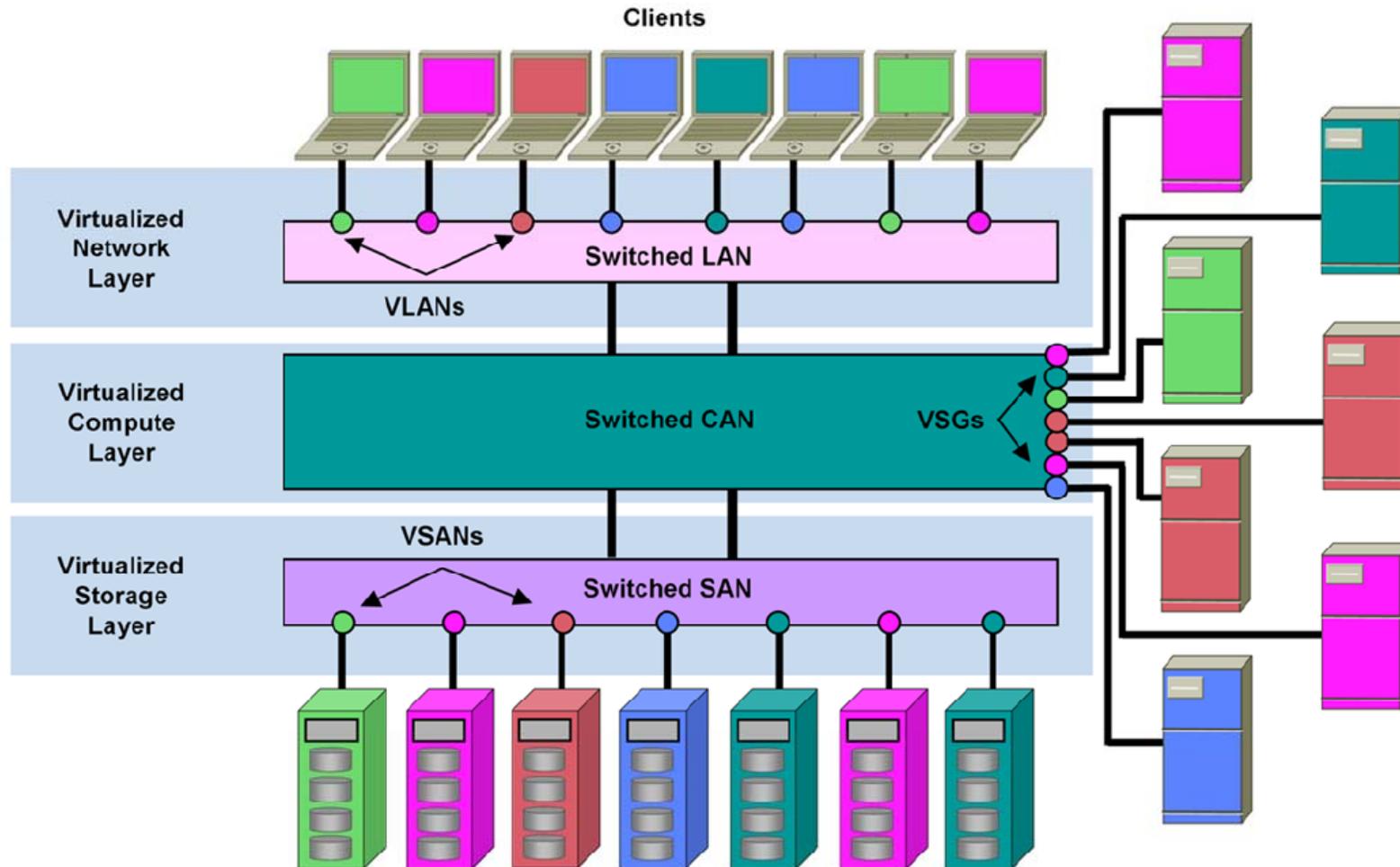
### Centros de datos con capa de red y almacenamiento virtualizados



## 6. Arquitectura de los centros de datos



### Centro de datos virtualizado punto-a-punto



## 6. Arquitectura de los centros de datos



### Nivel de fiabilidad y disponibilidad de un centro de datos: *Tier*

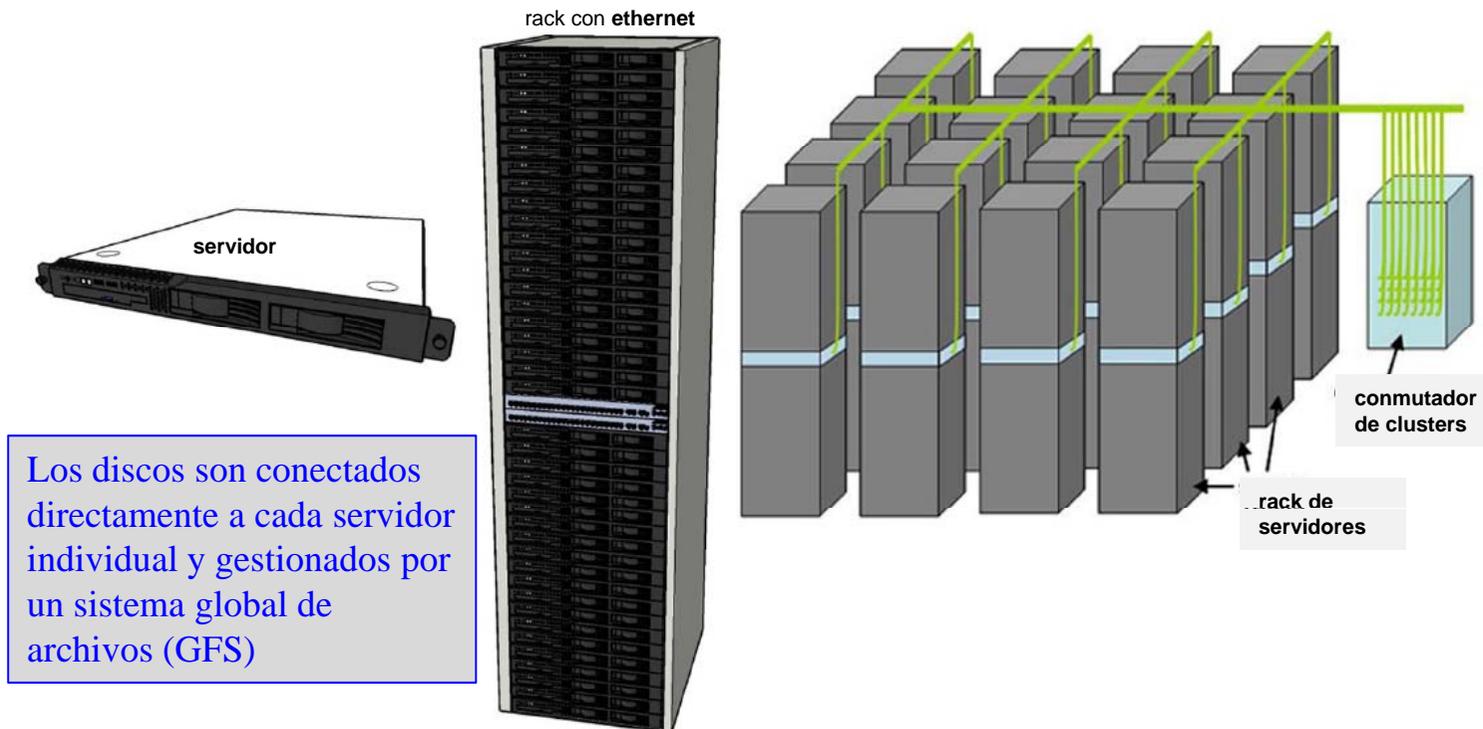
- Se han definido cuatro niveles de fiabilidad y disponibilidad (*Tier*) para los centros de datos:
  - **Tier 1: Centro de datos Básico: Disponibilidad del 99.671%.**
    - El servicio puede interrumpirse por actividades planificadas o no planificadas.
    - No hay componentes redundantes en la distribución eléctrica y refrigeración.
    - Puede disponer de generadores auxiliares o UPS.
    - Deberá estar fuera de servicio al menos una vez al año por mantenimiento y/o reparaciones.
  - **Tier 2: Centro de datos Redundante: Disponibilidad del 99.741%.**
    - Menos susceptible a interrupciones por actividades planificadas o no planificadas.
    - Componentes redundantes (N+1)
    - Dispone de generadores auxiliares o UPS.
    - Conectados a una única línea de distribución eléctrica y de refrigeración.
    - El mantenimiento de esta línea de distribución requiere una interrupción del servicio.
  - **Tier 3: Centro de datos Concurrentemente Manténibles: Disponibilidad del 99.982%.**
    - Permite planificar actividades de mantenimiento sin afectar al servicio de computación.
    - Componentes redundantes (N+1)
    - Conectados a más de una línea eléctrica y de refrigeración, pero únicamente con una activa.
    - Se pueden llevar a cabo tareas de mantenimiento en una línea mientras se da servicio por otras.
  - **Tier 4: Centro de datos Tolerante a fallos: Disponibilidad del 99.995%.**
    - Permite planificar actividades de mantenimiento sin afectar al servicio de computación críticos.
    - Conectados múltiples líneas de distribución eléctrica y de refrigeración.
    - Múltiples componentes redundantes (2 (N+1) significa 2 UPS con redundancia N+1).

## 6. Arquitectura de los centros de datos



### Estructura física de un Centro de Datos para *cloud computing*

- La implementación hardware de un WSC difiere significativamente de una instalación a la siguiente.
- Incluso dentro de una misma organización como puede ser Google los sistemas desplegados en años diferentes utilizan elementos básicos diferentes, reflejando las mejoras de hardware proporcionadas por la industria.
- Sin embargo la organización arquitectónica de estos sistemas ha sido relativamente estable en los últimos 5 años.



## 6. Arquitectura de los centros de datos

- El sistema de refrigeración es una parte muy significativa de un SWC
- Los modernos centros de datos disponen de sistemas adaptativos de control para la refrigeración de la instalación.

