

# Tema1:

## Introducción a los centros de datos

David Fernández  
david@dit.upm.es - B-216  
Curso 2018/2019



---

## Contenido

- ▶ Introducción
  - Definición, componentes, aplicaciones, requisitos.
- ▶ Normativa sobre centros de datos
- ▶ Eficiencia energética
- ▶ Ejemplos de centros de datos
- ▶ Sistemas operativos, virtualización y almacenamiento
- ▶ Computación en la nube

# INTRODUCCIÓN

## Aplicaciones telemáticas actuales

- ▶ Características:
  - Gran complejidad
  - Predominio de aplicaciones con tecnologías web
  - Gran número de usuarios y de cantidad de información
  - Las comunicaciones son básicas
- ▶ Los usuarios requieren alta calidad:
  - Comportamiento eficiente, fiable y seguro
  - Prestaciones suficientes
- ▶ Necesidad de grandes infraestructuras para darles soporte:
  - Redes de comunicaciones
  - Centros de datos

# Centro de Datos (CdD)

## ► Definiciones:

- “A data center is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression) and security devices.” (Wikipedia)
- “A data center is a special facility conceived to house, manage and support computing resources that are considered critical for one or more organizations.” (Data Center Virt. Fundamentals, Cisco Press)



## ► También llamados Centros de Proceso de Datos (CPD)

# Necesidad de Centros de Datos

## ► Múltiples razones:

- Las aplicaciones demandan cada vez mayor cantidad de recursos de computación, red y almacenamiento
- Creciente tendencia a mover aplicaciones y equipamiento desde centros de datos corporativos de tamaño pequeño/medio a grandes centros de datos

## ► Los centros de datos proporcionan:

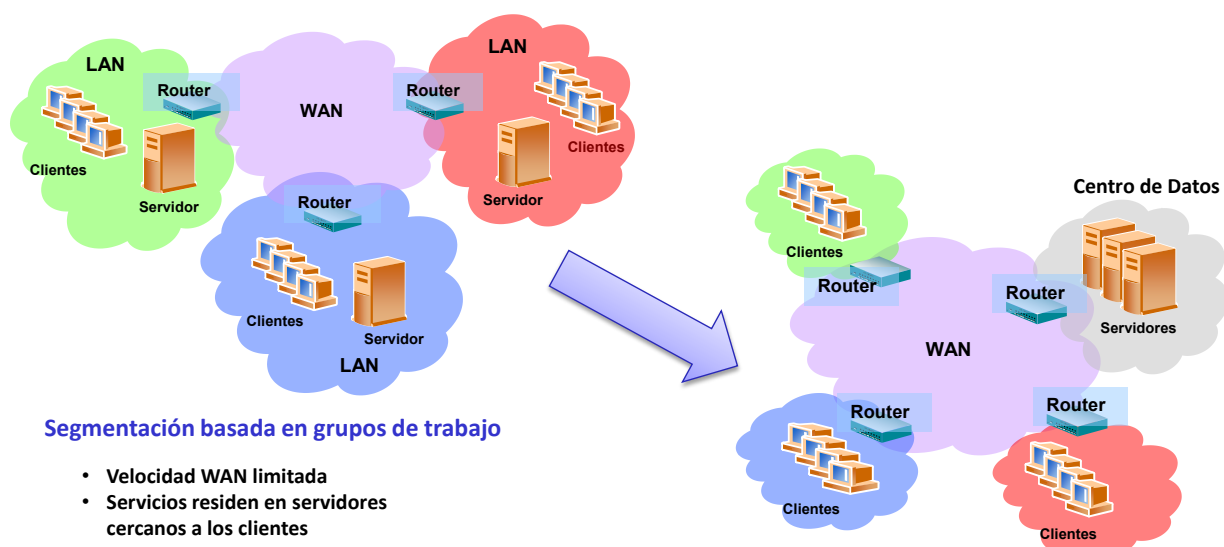
- Mayor escalabilidad
- Economía de escala
- Mayor fiabilidad
- Mejor conectividad con proveedores

# Evolución TI



Fuente: Bret Piatt. OpenStack Tutorial. IEEE CloudCom 2010

## Evolución TI (II)



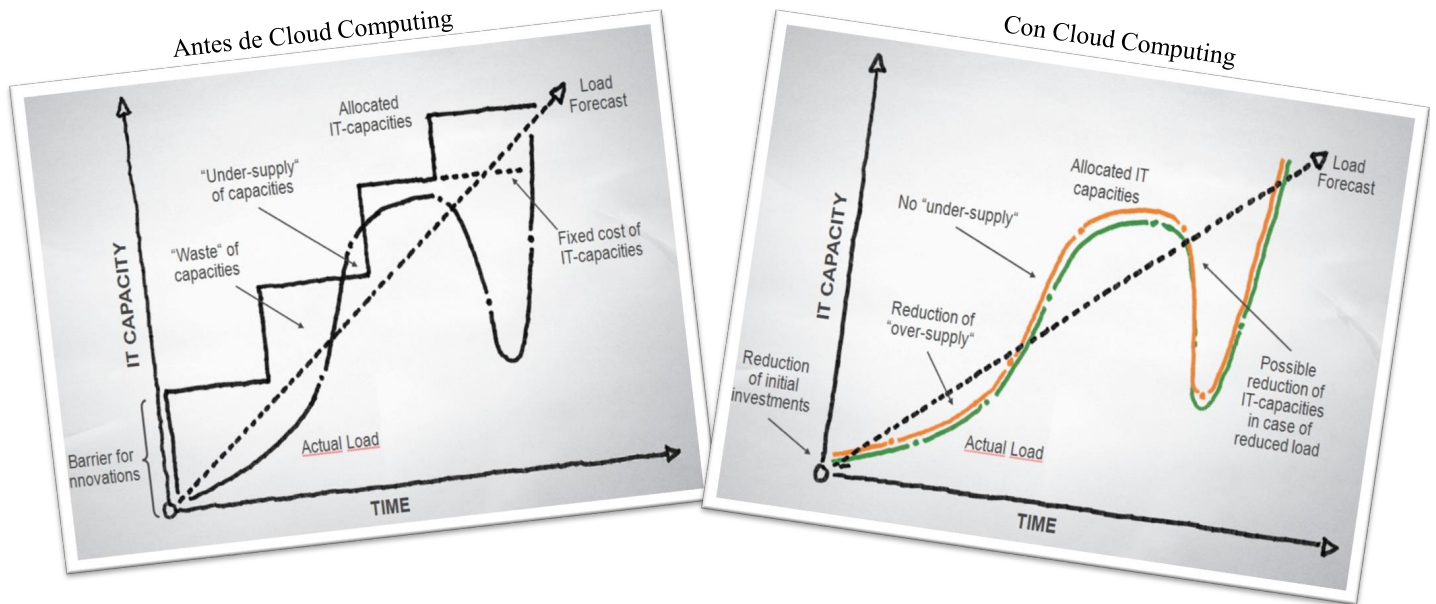
### Segmentación basada en grupos de trabajo

- Velocidad WAN limitada
- Servicios residen en servidores cercanos a los clientes

### Centralización de servidores

- Mejoras en la velocidad de la WAN permiten el acceso remoto a servidores centralizados
- Centros de datos privados de una organización o compartidos entre varias

# Ventajas CdD compartidos

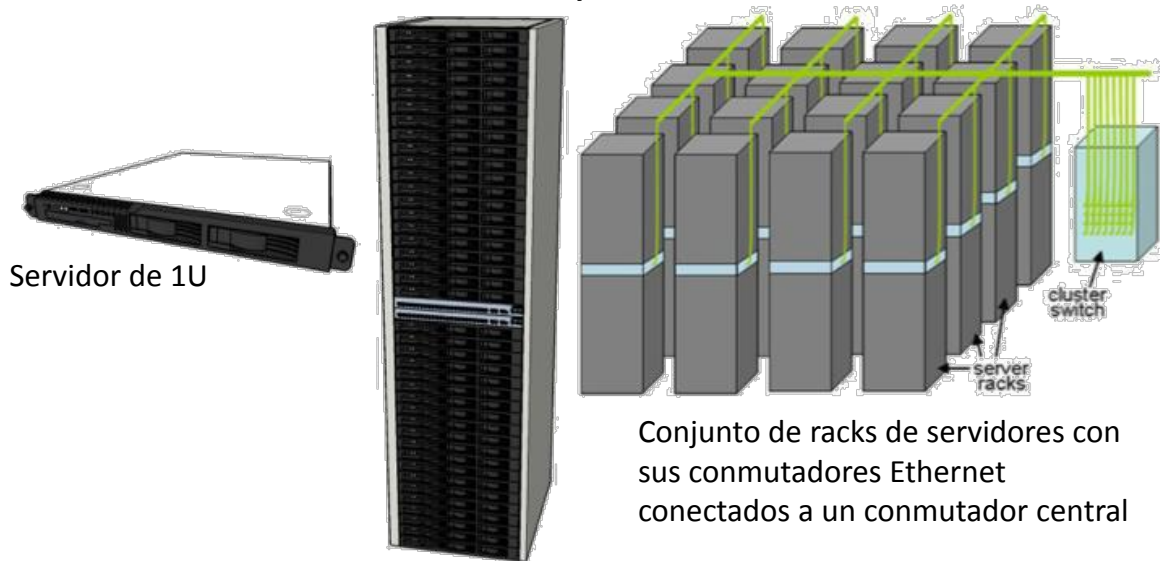


## Componentes de un CdD

- ▶ Equipos:
  - Servidores, almacenamiento, nodos de comunicaciones, etc.
- ▶ Infraestructuras básicas para el funcionamiento de los equipos:
  - Electricidad, refrigeración, sistemas antiincendios, etc.
- ▶ El funcionamiento correcto de un CdD requiere una inversión equilibrada en ambos aspectos

# Elementos de un CdD (I)

## ► Servidores, conmutadores y racks



# Elementos de un CdD (II)

## ► Servidores "blade":

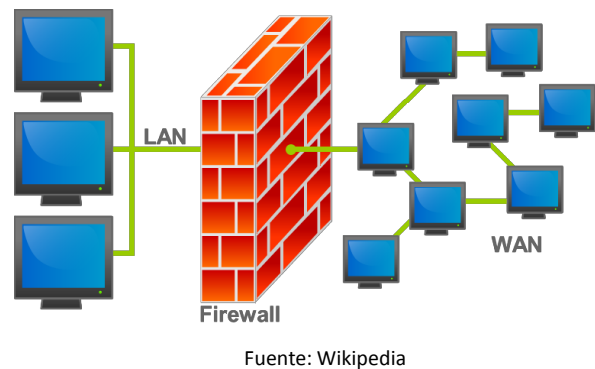
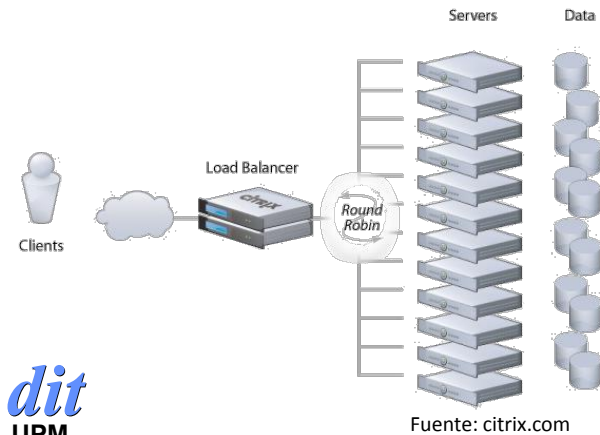
- Diseño modular: chasis + placas procesadoras (blade servers)
- Reducen espacio y consumo de energía
- Fuentes de alimentación y otros componentes compartidos (ej: switches Ethernet)





## Elementos de un CdD (III)

- ▶ Equipos de comunicaciones:
  - Nivel 2: Conmutadores (switches)
  - Nivel 3: Encaminadores (routers)
  - Nivel  $\geq 4$ : Firewalls, Balanceadores de tráfico
- ▶ Costosos debido a las velocidades de proceso (10 Gbps)



dit  
UPM

CDPS 18/19

Tema 1: Introducción a los Centros de Datos

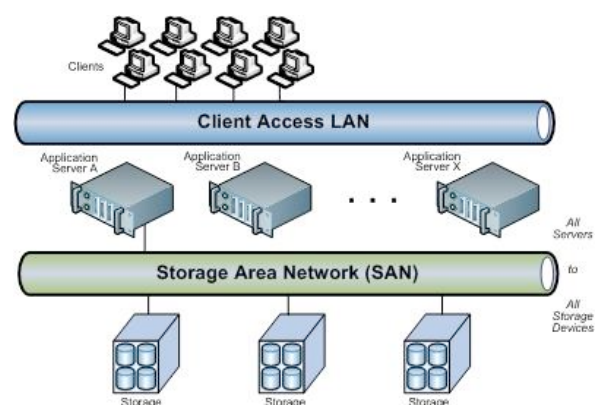
13

## Elementos de un CdD (IV)

- ▶ Almacenamiento:
  - Discos locales (Direct Attached Storage, DAS)
  - Servidores de discos externos (Network Attached Storage, NAS)



- Storage Area Networks (SAN)



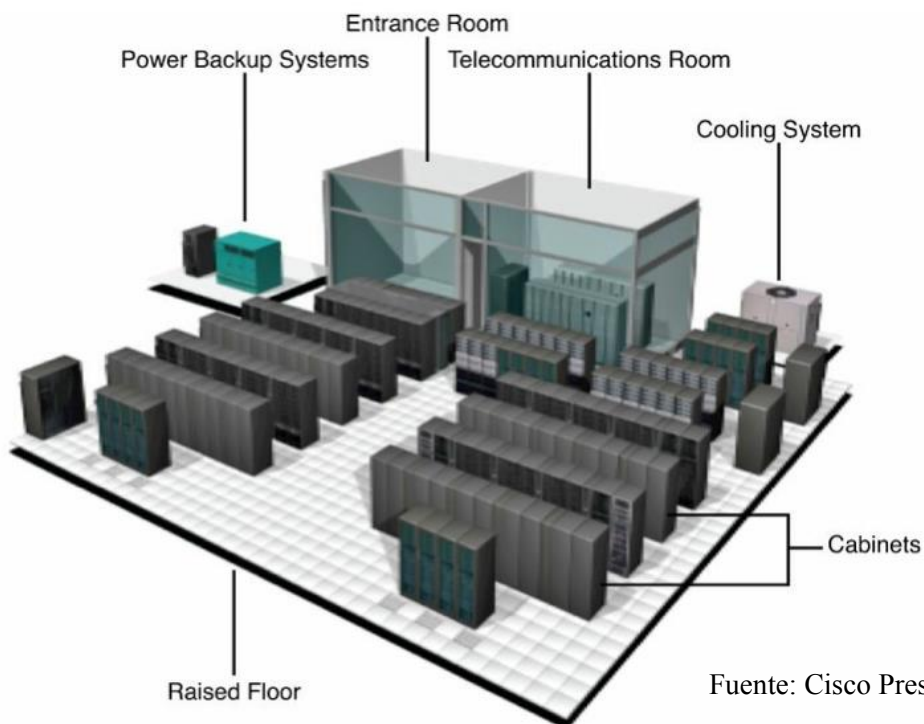
dit  
UPM

CDPS 18/19

Tema 1: Introducción a los Centros de Datos

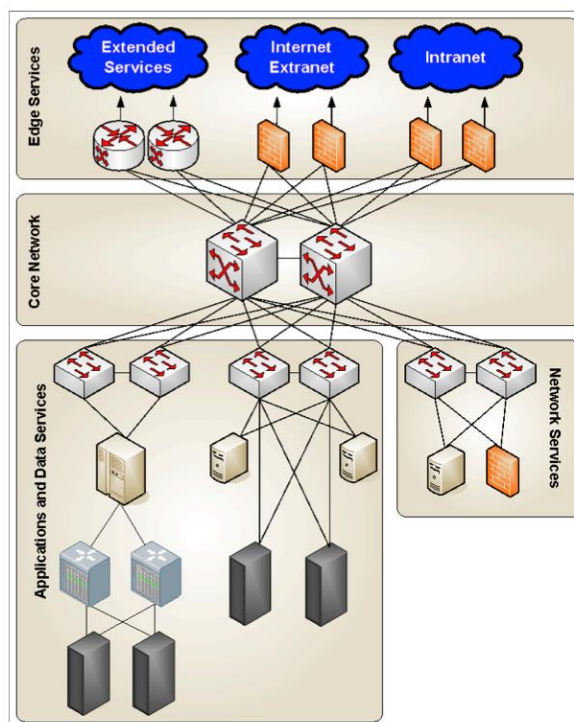
14

# Estructura física de un CdD



## Arquitectura de red

- Uso de topologías altamente redundantes para dar fiabilidad
- Problemas de escalabilidad de los protocolos de nivel 2 en grandes CdD
- Tecnologías: Ethernet (1/10 G), Infiniband (10 G), Fiber Channel (2-16 G)





# Aplicaciones de los CdD

- ▶ Aplicaciones de negocios:
  - Ej: Enterprise resource planning (ERP), Customer relationship management (CRM), etc
- ▶ Aplicaciones en la nube
- ▶ Recuperación frente a desastres
- ▶ Copias de seguridad
- ▶ Escritorios remotos (VDI)
- ▶ ...

## Requisitos funcionales de un CdD

- ▶ Espacio físico (localización)
- ▶ Alimentación
- ▶ Control de temperatura (refrigeracion)
- ▶ Cableado
- ▶ Seguridad
- ▶ Sistemas protección contra incendios
- ▶ Racks
- ▶ Etiquetado
- ▶ Acceso a consolas

# Cableado

Bajo suelo técnico



Por el techo



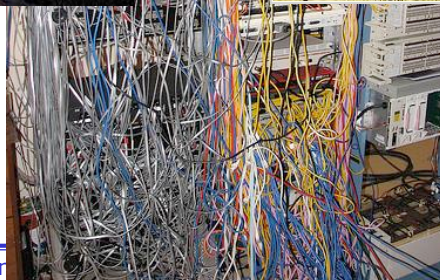
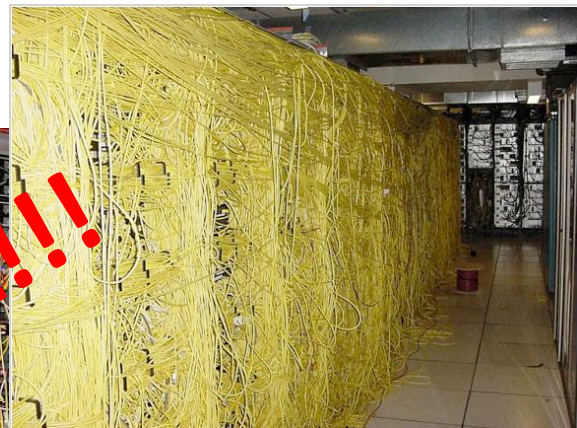
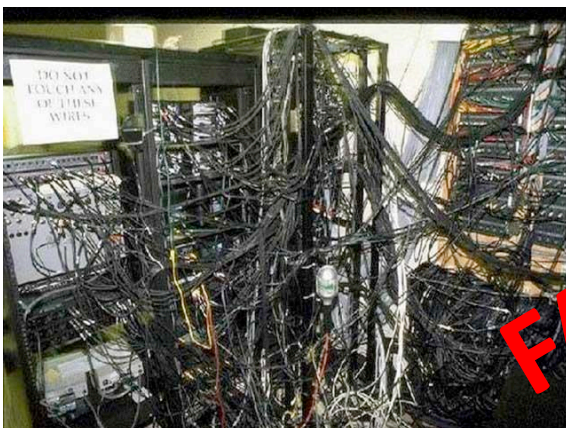
UPM



Fuente: Douglas Alger, The Art of the Data Center, Prentice Hall, 2012

## Cableado de un CdD

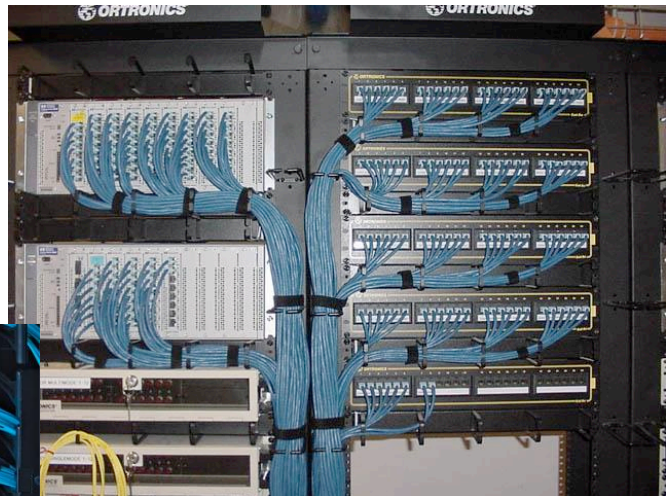
- El orden es muy importante



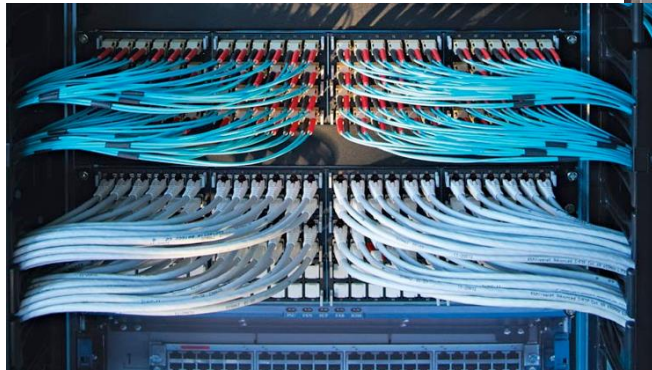
UPM



# Cableado



Fuente: <http://www.dciq.com>



**dit**  
UPM

# Generadores



**dit**  
UPM

# NORMATIVA SOBRE CENTROS DE DATOS

## Normativa sobre Centros de Datos

- ▶ TIA 942, *Telecommunications Infrastructure Standard for Data Centers*
  - estándar publicado por la *Telecommunications Industry Association* (TIA) en abril de 2005
  - serie de guías y pautas para el diseño y construcción de CPDs
  - Trata sobre:
    - *Data center telecommunications spaces and related topologies*
    - *Data center cabling systems*
    - *Data center cabling pathways*
    - *Data center redundancy*

# TIA-942

## ► Infraestructura de soporte de CDs dividida en cuatro subsistemas:

- Telecomunicaciones
- Arquitectura
- Sistema eléctrico
- Sistema mecánico

Cuadro 1.			
Telecomunicaciones	Arquitectura	Eléctrica	Mecánica
Cableado de racks	Selección del sitio	Cantidad de accesos	Sistemas de climatización
Accesos redundantes	Tipo de construcción	Puntos únicos de falla	Presión positiva
Cuarto de entrada	Protección ignífuga	Cargas críticas	Cañerías y drenajes
Área de distribución	Requerimientos NFPA 75	Redundancia de UPS	Chillers
Backbone	Barrera de vapor	Topología de UPS	CRAC's y condensadores
Cableado horizontal	Techos y pisos	PDU's	Control de HVAC
Elementos activos redundantes	Área de oficinas	Puesta a tierra	Detección de incendio
Alimentación redundante	NOC	EPO (Emergency Power Off)	Sprinklers
Patch panels	Sala de UPS y baterías	Baterías	Extinción por agente limpio (NFPA 2001)
Patch cords	Sala de generador	Monitoreo	Detección por aspiración (ASD)
Documentación	Control de acceso	Generadores	Detección de líquidos
	CCTV	Transfer switch	

## Tipos de Centros de Datos

### ► TIER I Data center, **Basic**:

- disponibilidad del 99.671 (máx. 28.82 horas/año sin servicio)
- dispone de climatización y una adecuada distribución de líneas de alimentación.
- No es necesario que disponga de suelo técnico, SAI o grupo electrógeno.
- El fallo o mantenimiento del servicio causa la detención del mismo.



# Tipos de Centros de Datos

## ► *TIER II Data center, **Redundant Components**:*

- disponibilidad del 99.741 (máx. 22.68 horas/año sin servicio)
- Todos los componentes están redundados (duplicados).
- Se dispone de suelo técnico, SAI y grupos electrógeno, pero únicamente tiene una acometida de alimentación.
- El mantenimiento no requiere detención del servicio (salvo que implique la acometida eléctrica).

# Tipos de Centros de Datos

## ► *TIER III Data center, **Concurrently Maintainable**:*

- disponibilidad del 99.982 (máx. 1:57 horas/año sin servicio)
- TIER II + línea de distribución de alimentación adicional, aunque únicamente una de ellas está activa.
- Cualquier mantenimiento no implica la detención del servicio.

## ► *TIER IV Data center, **Fault Tolerant**:*

- disponibilidad del 99.995% (máx. 52.56 minutos/año sin servicio)
- TIER III + múltiples líneas de alimentación activas y ambas con componentes redundados para cada línea

# Otras clasificaciones

## ► Según la dedicación

- Single tenant: dedicado a una única organización (privado)
- Multi tenant: compartido por múltiples organizaciones (público)

## ► Wikipedia:

- “**Multitenancy** refers to a principle in software architecture where a single instance of the software runs on a server, serving multiple client-organizations (tenants)”

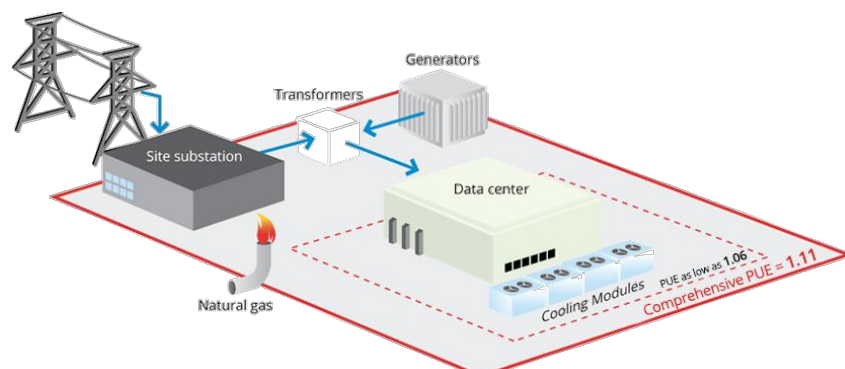
# EFICIENCIA ENERGÉTICA

# Eficiencia energética

- ▶ El coste de la energía es uno de los principales
- ▶ Energías alternativas:
  - Solar, geotérmica, eólica, etc
- ▶ Equipamiento de bajo consumo:
  - Iluminación LED,...
- ▶ Curiosidades:
  - Generadores eólicos que aprovechan la circulación de aire dentro de edificios (ver AISO Data Center)

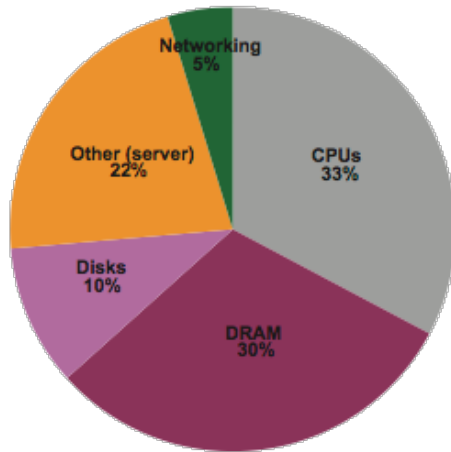
## Medida de la eficiencia energética

- ▶ Eficiencia en el uso de energía (Power Usage Effectiveness, PUE)
  - Un valor PUE de 2,0 significa que por cada vatio de potencia de TI, se consume un vatio adicional para refrigerar y distribuir energía al equipo de TI.
  - Un valor PUE cercano a 1,0 significa que casi toda la energía se utiliza para los ordenadores.



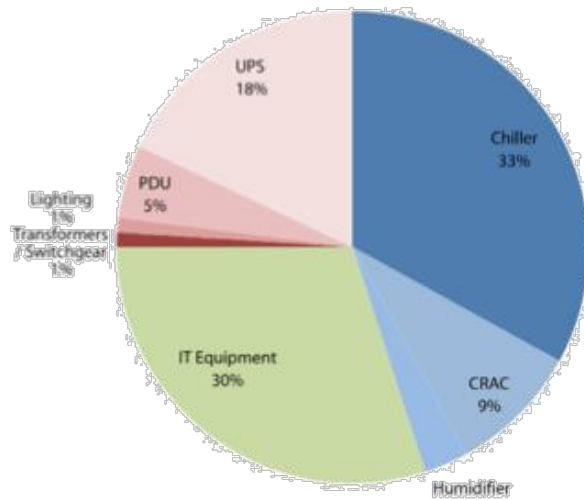
# Distribución consumo de energía

## ► De un servidor



Fuente: Google, 2007

## ► De un Centro de Datos



Fuente: American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)

# Distribución del coste de un Centro de Datos

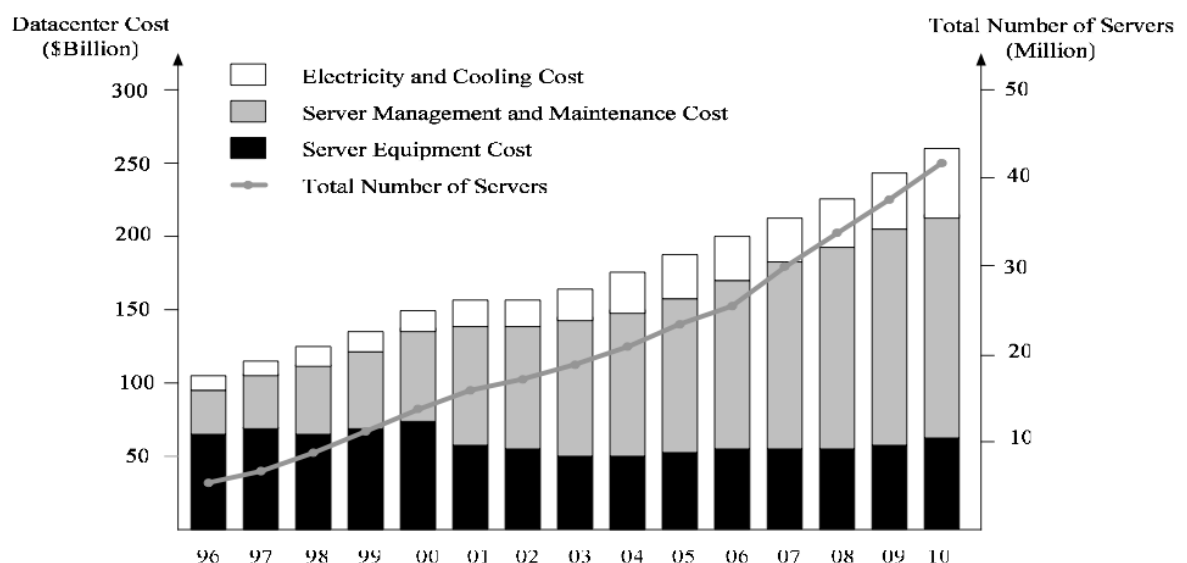


Figure 7.7 Growth and cost breakdown of datacenters over the years (Source: IDC Report 2009).

# Refrigeración de un Centro de Datos

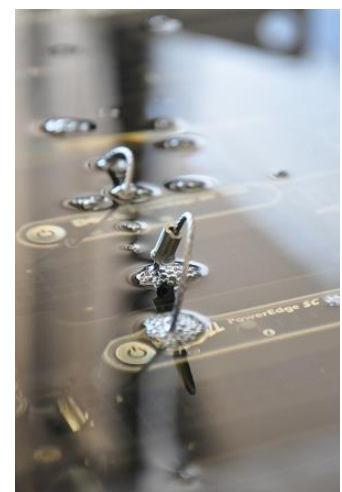
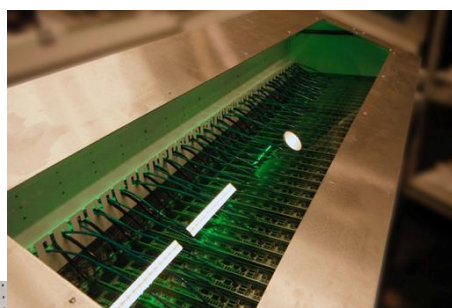
- ▶ Tradicionalmente basada en sistemas de aire acondicionado
- ▶ Creciente uso de refrigeración por agua tomada del mar o de ríos
  - Por cuestiones medioambientales, el agua debe enfriarse en estanques o torres de refrigeración antes de devolverla
  - O utilizarse para cogeneración (ej.: calefacciones domésticas)
- ▶ Temperatura recomendada por Google -> 27º
- ▶ Necesidad de controlar la humedad ambiente -> ideal entre 45 y 55 %
  - Poca humedad -> descargas de estática



Mucha humedad -> condensación

## Ejemplos eficiencia energética

- ▶ Midas Networks (<http://www.grcooling.com/>)
  - Refrigeración por inmersión en fluido dieléctrico (aceite)
  - Reduce la energía de refrigeración en hasta un 95%
  - Reduce el coste de construcción hasta un 60%





# EJEMPLOS DE CENTROS DE DATOS

## Ejemplos de Centros de Datos

### ► Centros de datos de Google:

- Fotografías: <http://www.google.com/about/datacenters/gallery/>
- Vídeo y Street View:  
<http://www.google.com/about/datacenters/inside/streetview/>

### ► Muchos ejemplos interesantes en:

- Douglas Alger, “The Art of the Data Center: A Look Inside the World's Most Innovative and Compelling Computing Environments”, Prentice Hall 2012.
- Disponible en Safari:  
<http://www.etsit.upm.es/biblioteca/recursos-electronicos.html>

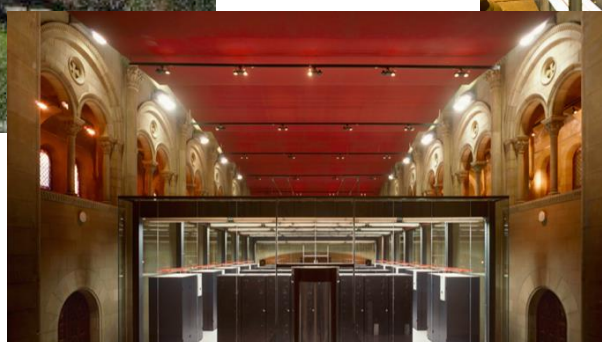
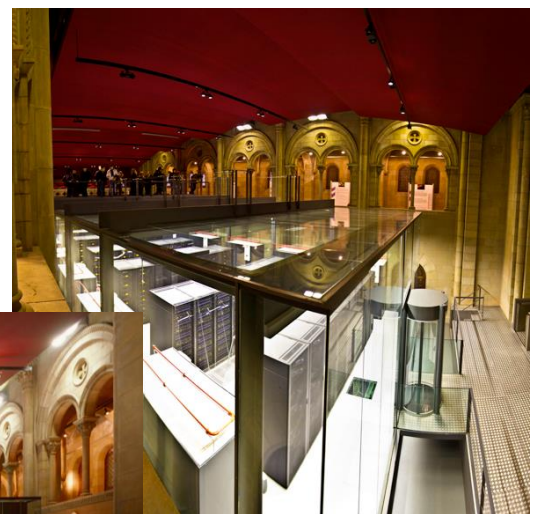
- ▶ Centro de Supercomputación y Visualización de Madrid (UPM)
- ▶ Aloja el superordenador Magerit
  - clúster formado por 245 nodos de cómputo eServer BladeCenter PS702
  - cada nodo tiene 16 procesadores PPC de 3'3 GHz (294 GFlops) con 32 GB de RAM
  - Situado en campus Montegancedo
  - Utilizado para cálculos masivos de proyectos científicos:
    - <http://www.cesvima.upm.es/projects.html>



Procesadores	3.920,0
Memoria	7.840,0 GB
Potencia	72.030,0 GFLOPS

## Barcelona Supercomputing Center

- ┆ MareNostrum: 48,896 procesadores Intel Sandy Bridge en 3,056 nodes, 96.6 TB de memoria, 2 PB de disco. Número 29 en el [TOP500](#)



# Catedral Uspenski

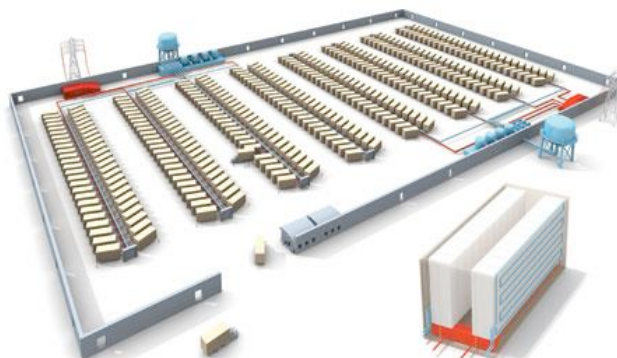
- ▶ Centro de Datos bajo una catedral en Helsinki, Finlandia
- ▶ Da calefacción a 500 hogares y empresas aledañas (para el 2015 esperan triplicar la cantidad).



dit  
UPM

## Centros de datos en contenedores

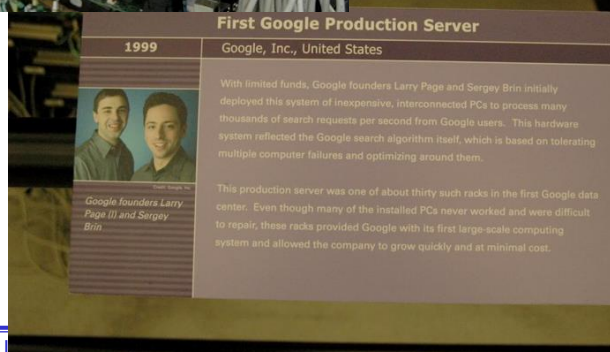
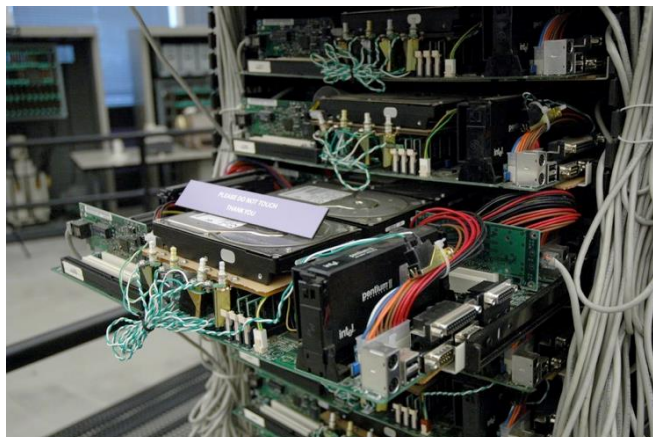
- ▶ CdD autocontenidos, modulares y transportables
  - Ventaja: despliegue rápido de CdD, tanto temporales como permanentes





# Curiosidades

- Primeros servidores de Google (<http://goo.gl/OuB28r>)



**dit**  
UPM

CDPS 18/19

Tema 1: Introducción a

## SISTEMAS OPERATIVOS Y VIRTUALIZACIÓN

**dit**  
UPM

CDPS 18/19

Tema 1: Introducción a los Centros de Datos

49

# Software en un CdD

- ▶ Software propio de equipos
  - Servidores: firmware, sistemas operativos, hipervisores de virtualización
  - Equipos de red: firmware y sistemas operativos
- ▶ Software de gestión de infraestructura
  - Software de configuración, control y monitorización, sistemas de ficheros distribuidos, planificación, gestión de inventarios, etc.
- ▶ Software de soporte a aplicaciones
  - Servidores de aplicaciones, modelos de programación (Ej: MapReduce)
- ▶ Software de aplicaciones
  - Servicios ofrecidos a usuarios

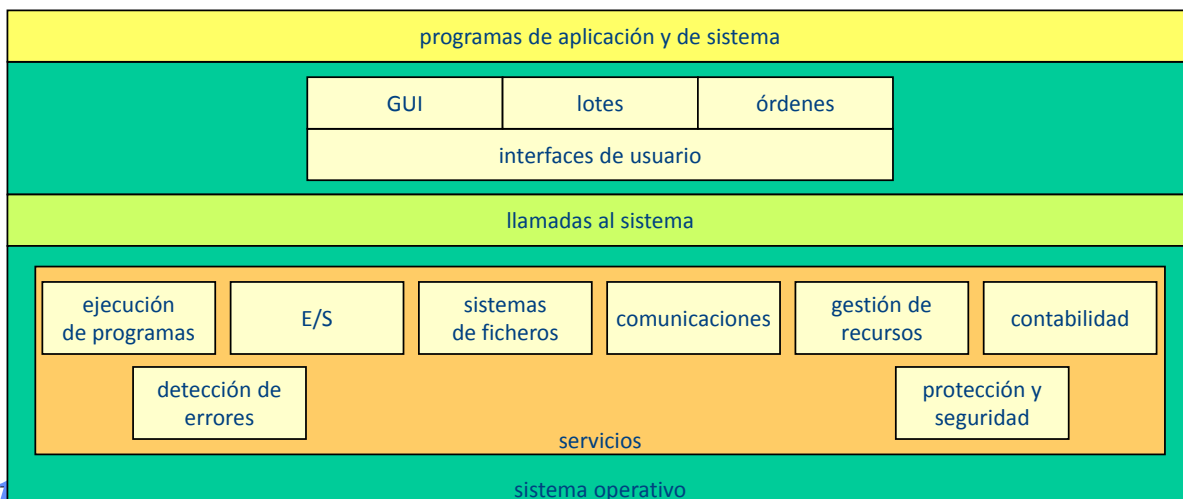
## Sistema Operativo

- ▶ Forma parte de la plataforma básica de un CdD
- ▶ Conocimiento de SO facilita:
  - Instalación de aplicaciones y servicios
  - Integración en redes
  - Configuración hardware, para rendimiento adecuado
  - Entender y configurar núcleo de virtualización
- ▶ Funciones de un sistema operativo:
  - Facilitar el uso del ordenador: abstracciones adecuadas
  - Utilizar el ordenador eficientemente



# Sistema operativo

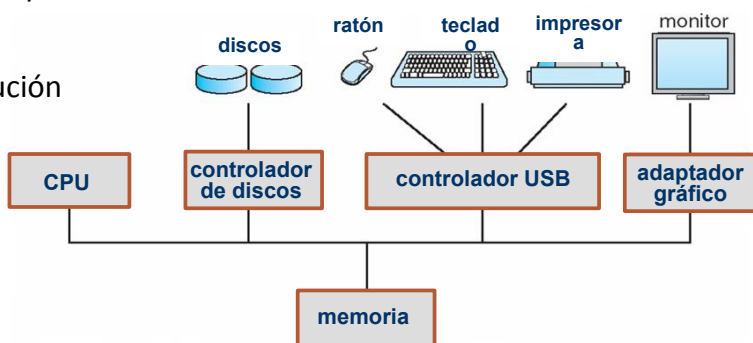
- Objetivo: gestionar los recursos hardware de un ordenador y proporcionar un entorno para la ejecución de programas



dit  
UPM

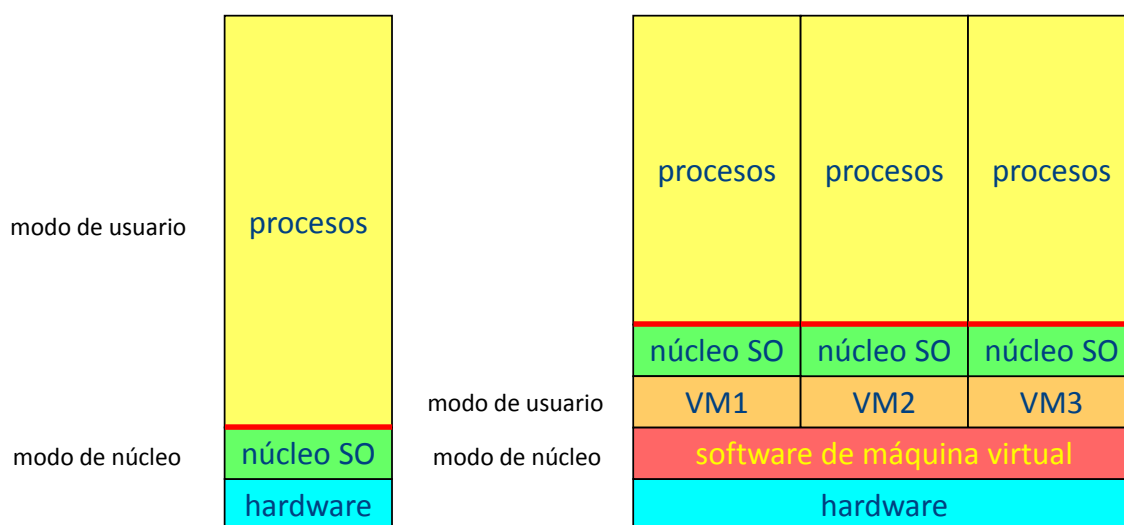
## Componentes de un Sistema Operativo

- Gestión de procesos
  - Ejecución de programas, programación concurrente
- Gestión de memoria
  - Asignación de memoria a procesos
- Gestión de almacenamiento
  - Sistemas de ficheros y gestión de dispositivos de almacenamiento
- Gestión de dispositivos de E/S
  - Interacción con dispositivos de E/S
- Protección y seguridad
  - Protección de los datos y ejecución segura de procesos
  - Evitar accesos externos indebidos

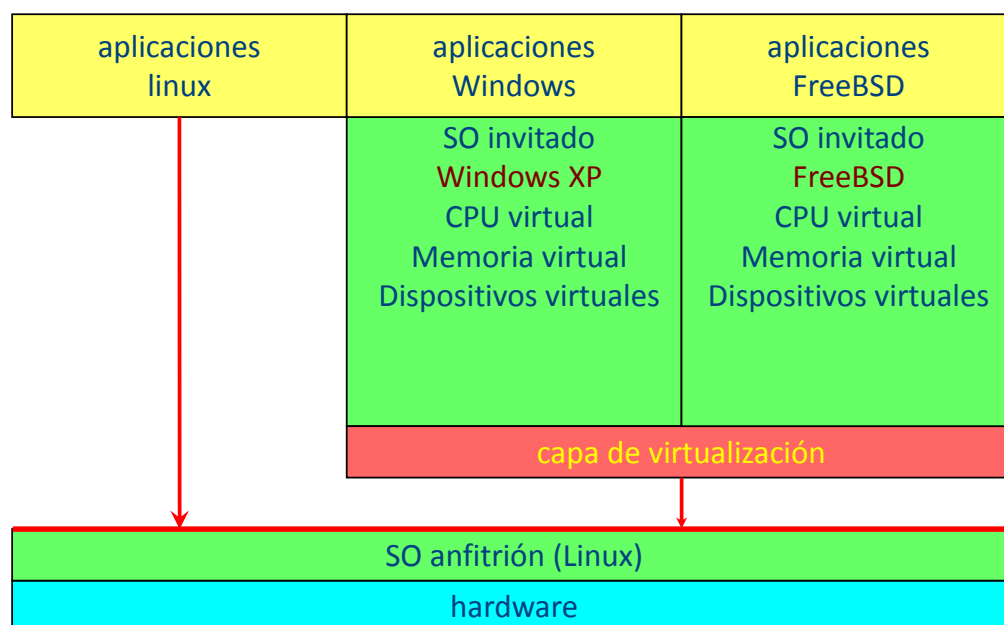


dit  
UPM

# Virtualización de Servidores

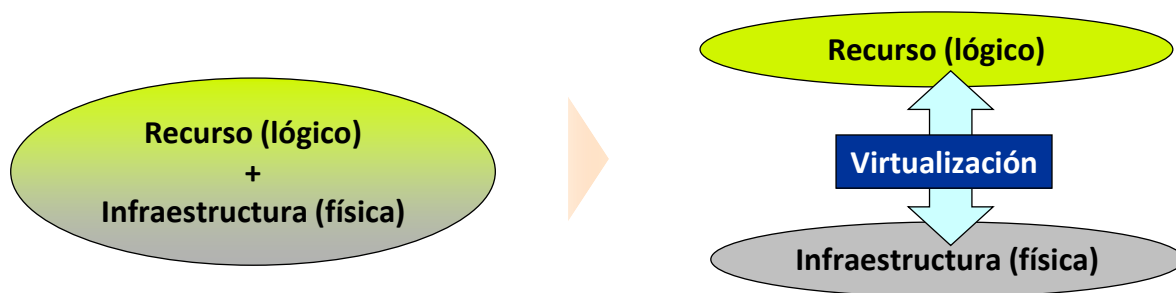


# Virtualización de Servidores



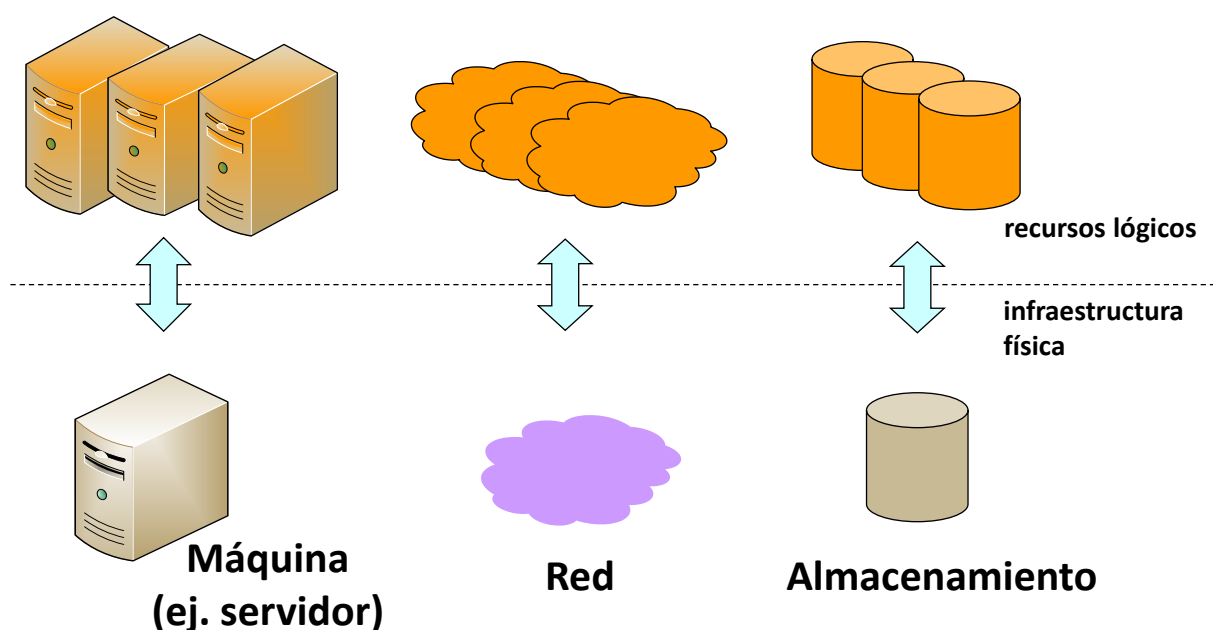
# Virtualización: Definición

Por virtualización se entiende cualquier tecnología que permita desacoplar o abstraer los recursos lógicos de la infraestructura física que los proporciona

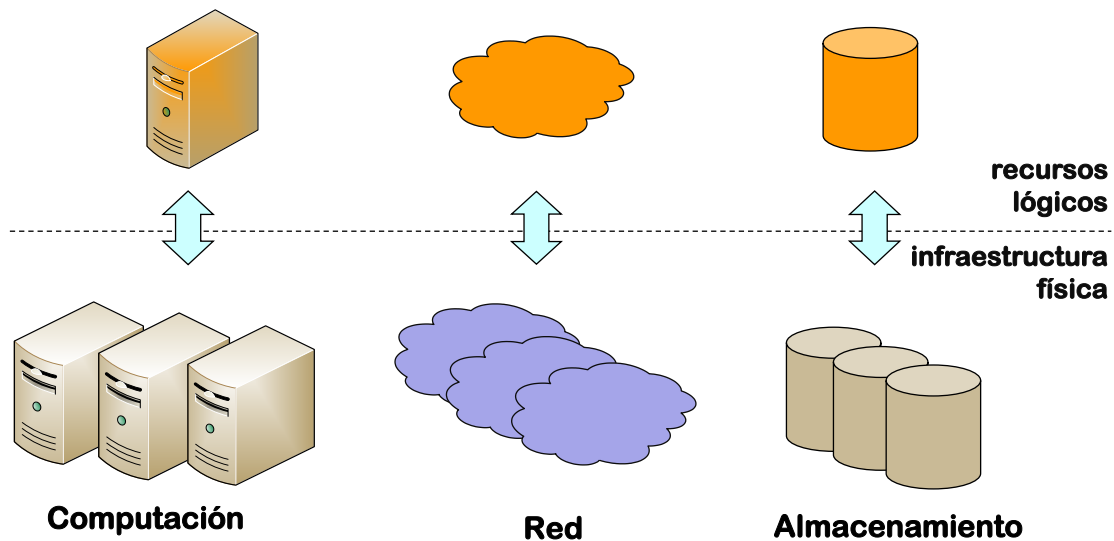


Desacoplar  $\Rightarrow$  multiplexación, portabilidad, movilidad, ...

## Virtualización = Abstracción de Recursos

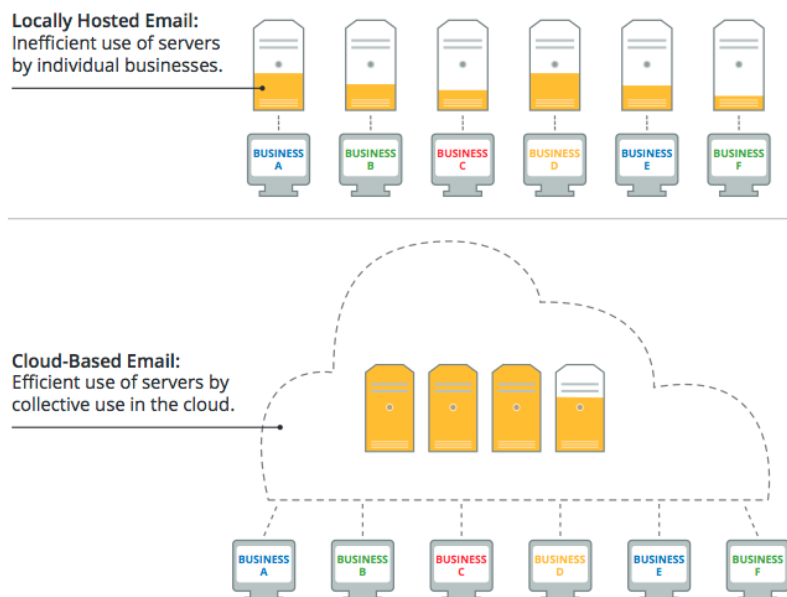


# Virtualización = Abstracción de Recursos

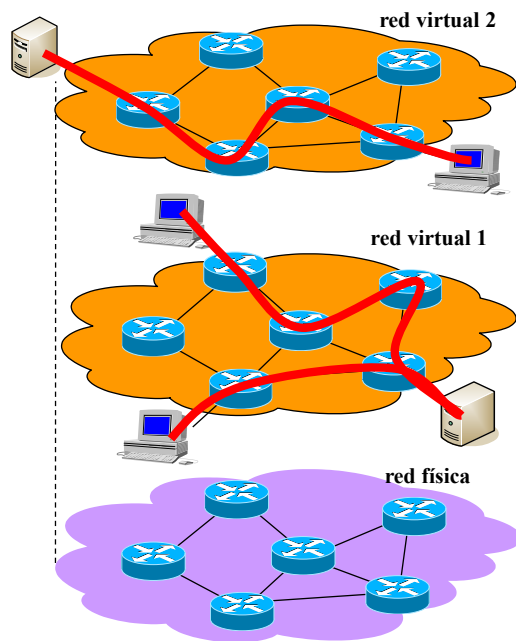


## Virtualización de Servidores

- Ventajas principal: ahorro económico en equipos, energía, etc.



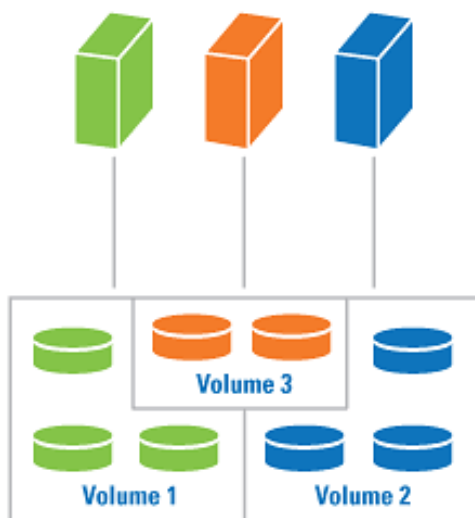
# Virtualización de Red



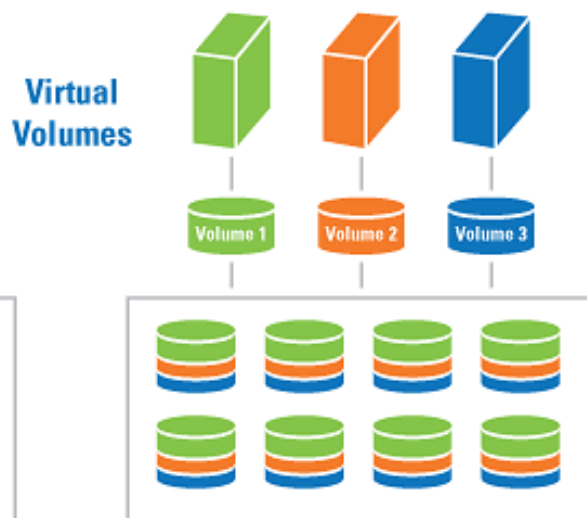
- ▶ Creación de redes virtuales (overlay) basados en redes físicas.
- ▶ Tecnologías:
  - | Basadas en etiquetado: 802.1q VLAN, MPLS
  - | Basadas en túneles: GRE, IPSec, IPv6
  - | Software Defined Networks (SDN)
- ▶ Aplicaciones
  - | VPN, redes privadas corporativas sobre redes públicas (añadiendo cifrado a los túneles)
  - | Evolución del troncal de red

# Virtualización del Almacenamiento

## Traditional Disk Mapping



## Virtualized Storage Disk Mapping





# COMPUTACIÓN EN LA NUBE

## Servicios en la nube

- ▶ Visión de los usuarios:
  - Servicios a los que se accede de forma sencilla a través de la red sin necesidad de instalar aplicaciones complejas
- ▶ Se implementan utilizando la computación en la nube
  - El usuario desconoce dónde se ejecutan los programas o dónde se almacenan sus datos (transparencia)
  - El prestador del servicio se encarga de todo (recursos, mantenimiento, actualizaciones, etc.)
- ▶ Proporcionan movilidad y alta disponibilidad:
  - accesibles desde cualquier dispositivo conectado a Internet
- ▶ Ejemplos:
  - Google, Twitter, Facebook, Youtube, Dropbox, etc.



# Computación en la nube (Cloud computing)

- ▶ No existe consenso en su definición
- ▶ Según el NIST:

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction”

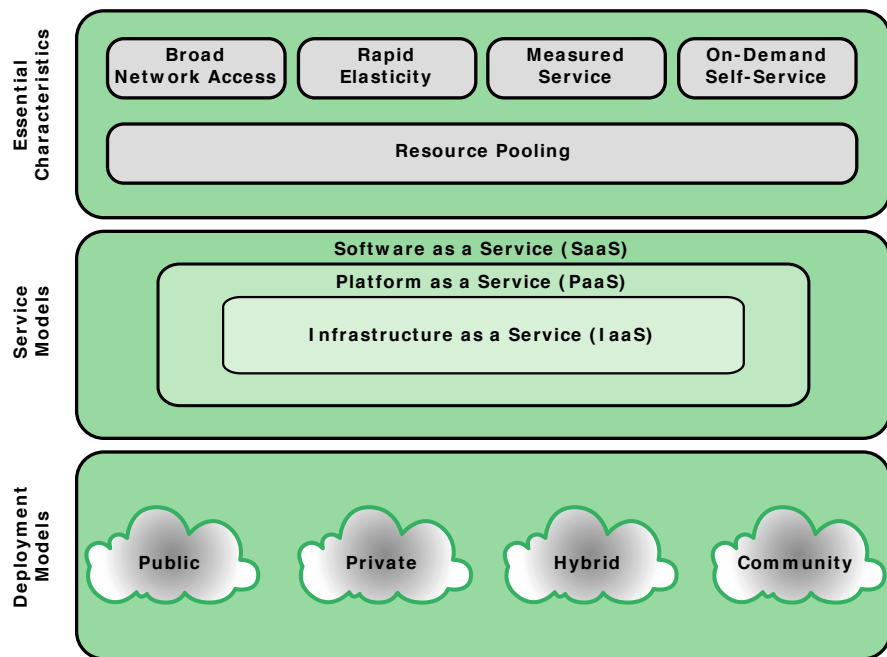
- ▶ La computación en nube proporciona las TI como un servicio más (*computing as a utility*)

## Computación en la nube

- ▶ Otras definiciones:

- A style of computing where massively scalable IT-related capabilities are provided ‘as a service’ using Internet technologies to multiple external customers (Gartner)
- Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. (Armbrust)
- A computing Cloud is a set of network enabled services, providing scalable, QoS guaranteed, normally personalized, inexpensive computing platforms on demand, which could be accessed in a simple and pervasive way (Wang)

# Elementos Computación en Nube



Fuente: Stallings, "Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud"

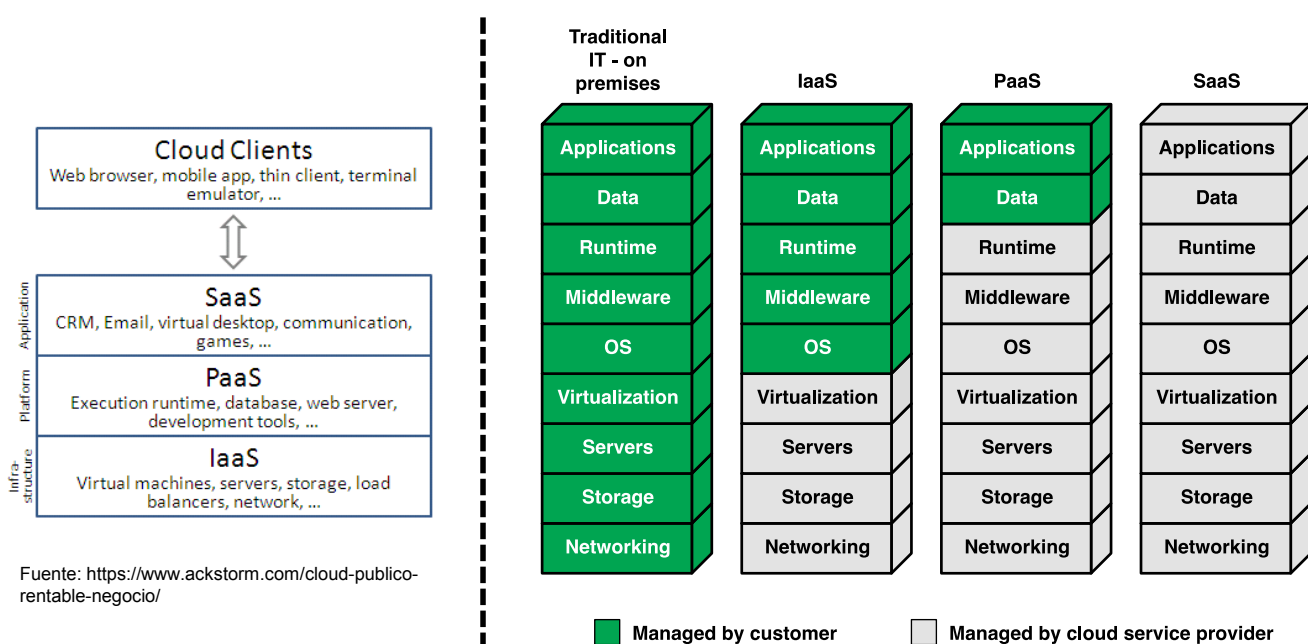
## Computación en la nube: Características

- ▶ **Servicio bajo demanda (autoservicio)**
  - Servicios proporcionados automáticamente sin interacción humana con el proveedor
- ▶ **Acceso a través de la red**
  - Desde múltiples tipos de dispositivos
- ▶ **Recursos comunes (Resource pooling)**
  - Proporcionados para dar servicio a múltiples usuarios simultáneamente (multi-tenant)
- ▶ **Elasticidad rápida**
  - Adaptación rápida a la demanda
- ▶ **Servicio medido**
  - Uso de recursos monitorizado, controlado e informado, tanto al cliente como al proveedor. Permite el pago por uso (pay-per-use)

# Ventajas de la Computación en la Nube

- ▶ Ahorro económico por compartición de recursos entre múltiples organizaciones (economía de escala)
  - | Los clientes cambian CAPEX (inversión en equipamiento) por OPEX (inversión en operación), pagando solo por el uso
- ▶ Eficiencia en la utilización de los recursos
  - | Flexibilidad a la hora de reasignar recursos según las necesidades
- ▶ Rapidez y simplicidad en el despliegue de nuevos servicios
  - | Se elimina la necesidad de provisionar equipamiento

## Responsabilidades operación en la nube

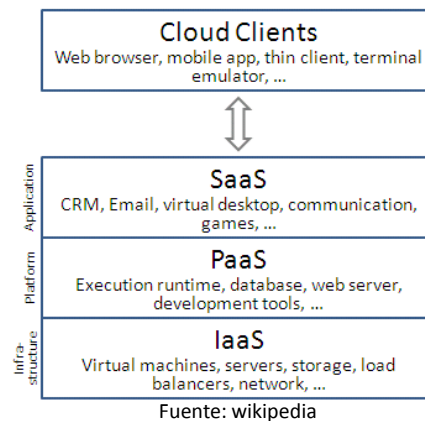


Fuente: <https://www.ackstorm.com/cloud-publico-rentable-negocio/>

Fuente: Stallings, "Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud"

# Modelos de Servicio

- ▶ Infraestructure as a Service (IaaS)
- ▶ Platform as a Service (PaaS)
- ▶ Software as a Service (SaaS)
- ▶ En la asignatura nos centramos principalmente en IaaS



## Infraestructure as a Service (IaaS)

- ▶ Oferta de recursos de computación, almacenamiento y comunicaciones bajo demanda a través de la red.
- ▶ Ejemplo: servicio EC2 Amazon proporciona máquinas virtuales (VM) con un sistema operativo y software instalado.
  - La gestión de la VM (arranque parada, instalación o actualización de sw, etc.) la realiza el usuario.
- ▶ Otros: Flexiscale, Joyent, Rackspace

# Platform as a Service (PaaS)

- ▶ Oferta de entornos en los que los desarrolladores pueden crear y desplegar aplicaciones sin preocuparse de los recursos (procesos, memoria, etc.) que estas utilizan.
- ▶ Los entornos ofrecen bibliotecas de servicios adicionales (acceso a datos, autenticación, pagos, etc).
- ▶ Ejemplos:
  - AppEngine de Google para desarrollo y alojamiento de aplicaciones web
  - Microsoft Azure
  - Heroku

# Software as a Service (SaaS)

- ▶ Oferta de servicios de aplicaciones
  - Aplicaciones con funcionalidades similares a las tradicionales (procesadores de texto, hojas de calculo, aplicaciones de gestión, etc.) pero ejecutadas y mantenidas por el proveedor
  - Acceso mediante web típicamente.
- ▶ Ejemplo: Google Docs



# Modelos de despliegue

- ▶ Nube privada
- ▶ Nube de comunidad
- ▶ Nube pública
- ▶ Nube híbrida

## Gestión de infraestructuras en la nube

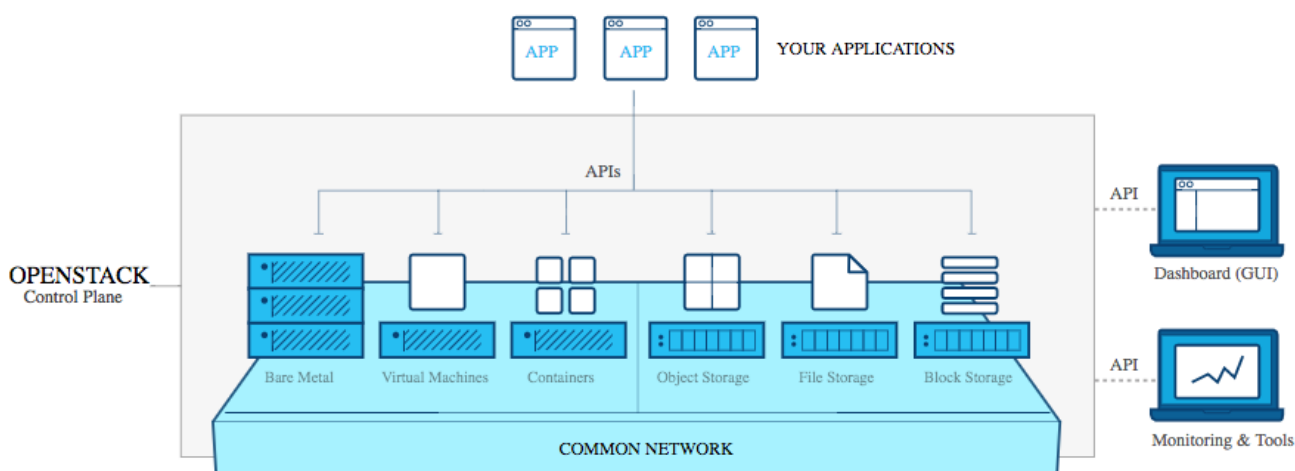
- ▶ Necesidad de herramientas que permitan gestionar de forma integrada todos los recursos físicos y lógicos de un proveedor IaaS
  - Alto nivel de automatización
- ▶ Virtual Infrastructure Manager (VIM): software responsable de coordinar todos los recursos de un IaaS para ofrecer una visión uniforme a usuarios y aplicaciones.
- ▶ Cloud toolkit: interfaz remoto y seguro para la creación, control y monitorización de VMs.
- ▶ Ejemplos: Apache VCL, OpenStack, VMware vSphere y vCloud, OpenNebula, CloudStack

# OpenStack

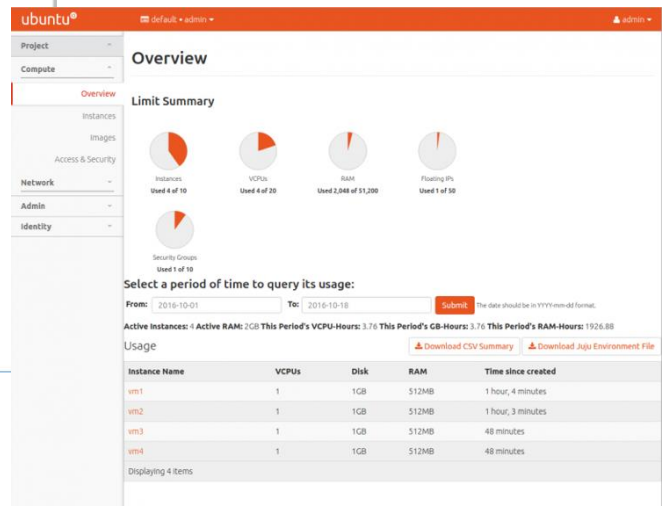
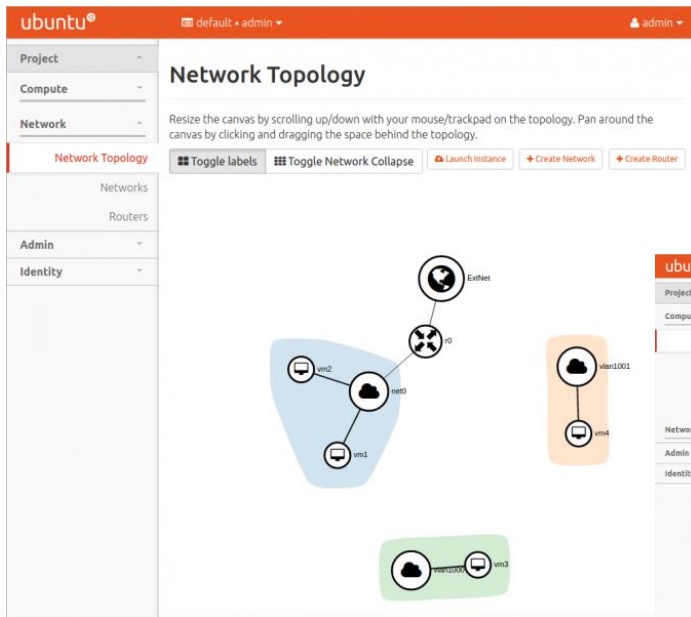
- ▶ Principal plataforma de computación en la nube de código abierto
- ▶ Desarrollado originalmente por Rackspace y la NASA
  - Soportado actualmente por la Fundación OpenStack en la que participan las principales empresas del sector (más de 200)
- ▶ Conjunto de proyectos software interrelacionados con el objetivo de controlar, automatizar y asignar eficientemente los recursos de un IaaS (computación, almacenamiento y comunicaciones)
  - ▶ OpenStack = Massively scalable cloud operating system
- ▶ Desarrollo de nuevas versiones cada 6 meses

## ¿Qué es OpenStack?

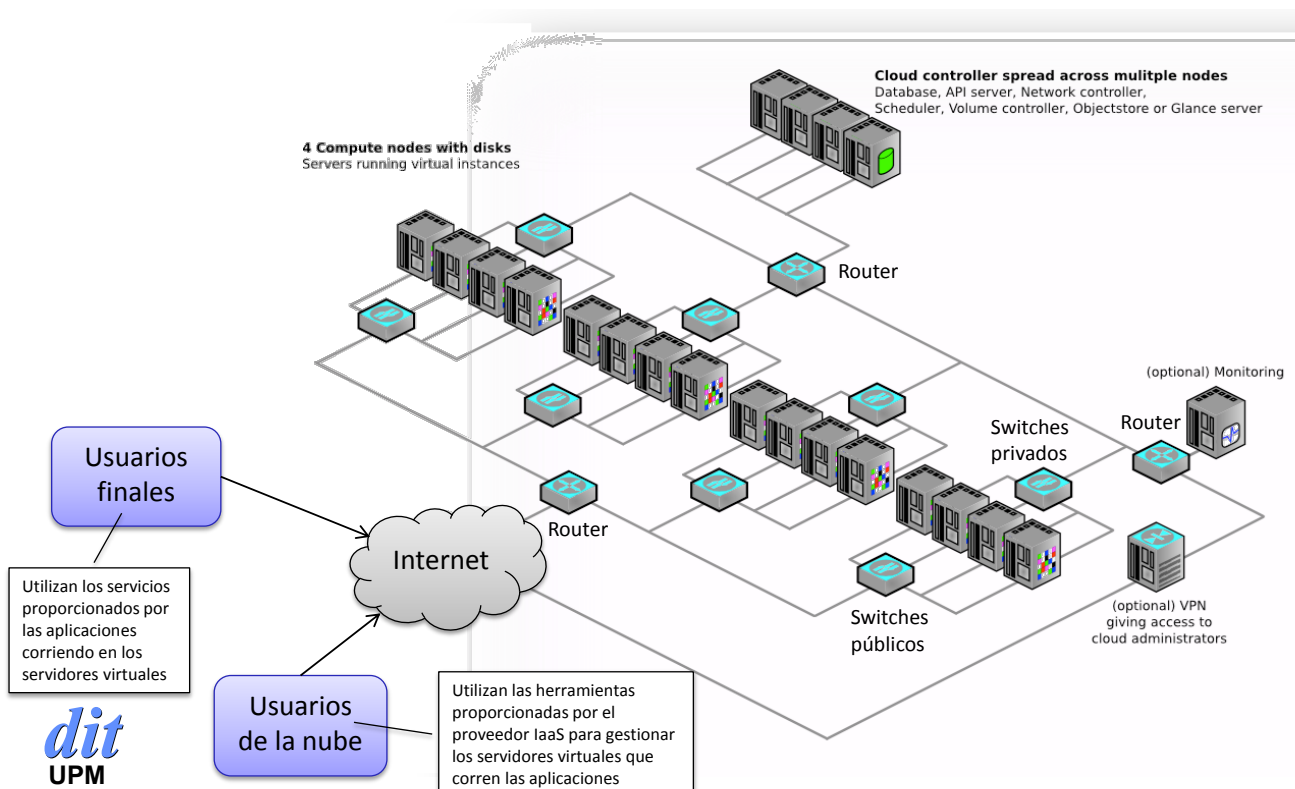
- ▶ Sistema Operativo de la nube



# Gestión Openstack

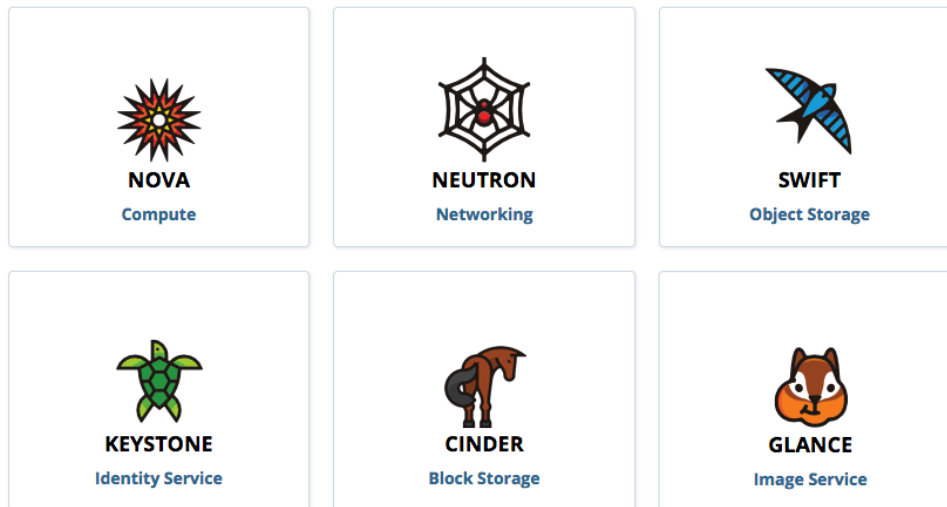


## Ejemplo Instalación Openstack



# Proyectos OpenStack

## ► Principales proyectos:



## ► Lista completa de proyectos en:

<https://www.openstack.org/software/project-navigator>

## Referencias

- Cap. 6 de “The Practice of System and Network Administration”, Addison-Wesley, 2007
- Cap. 1 de “Cloud Computing. Automating the Virtualized Data Center”, Cisco Press, 2012
- The NIST Definition of Cloud Computing
- Above the Clouds: A Berkeley View of Cloud Computing
- Google’s Green Computing: Efficiency at Scale
- Douglas Alger, “The Art of the Data Center: A Look Inside the World's Most Innovative and Compelling Computing Environments”, Prentice Hall 2012.