

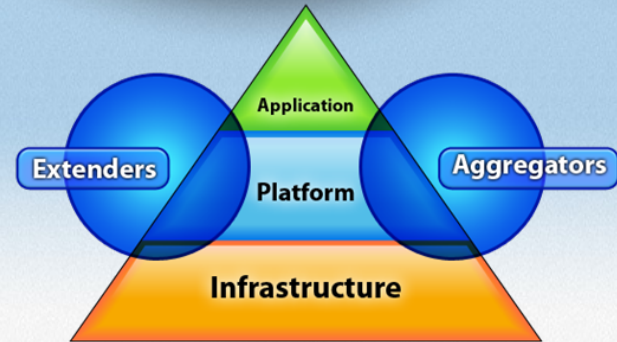
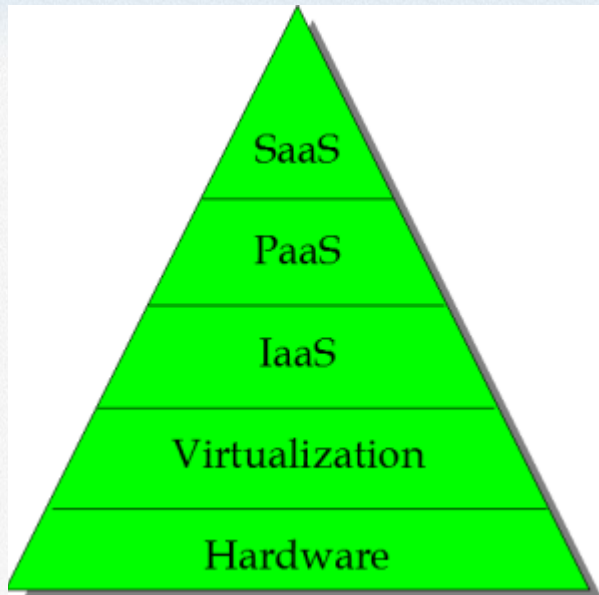
Big Data y Cloud Computing

T1.2. Infraestructura como Servicio

Contenido

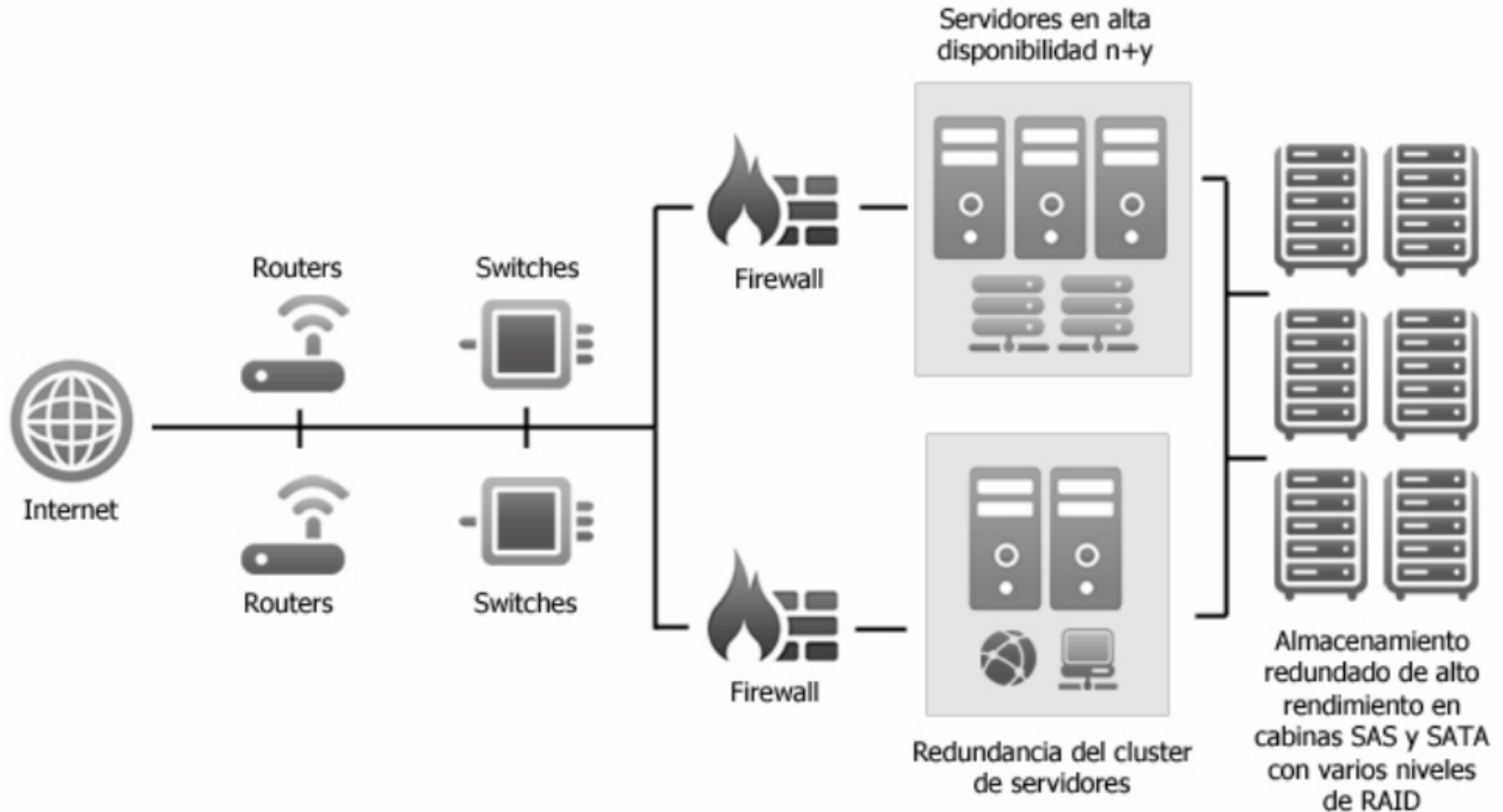
- Virtualización
- Almacenamiento
- IaaS OpenSource
- IaaS comerciales

Arquitectura de Servicios



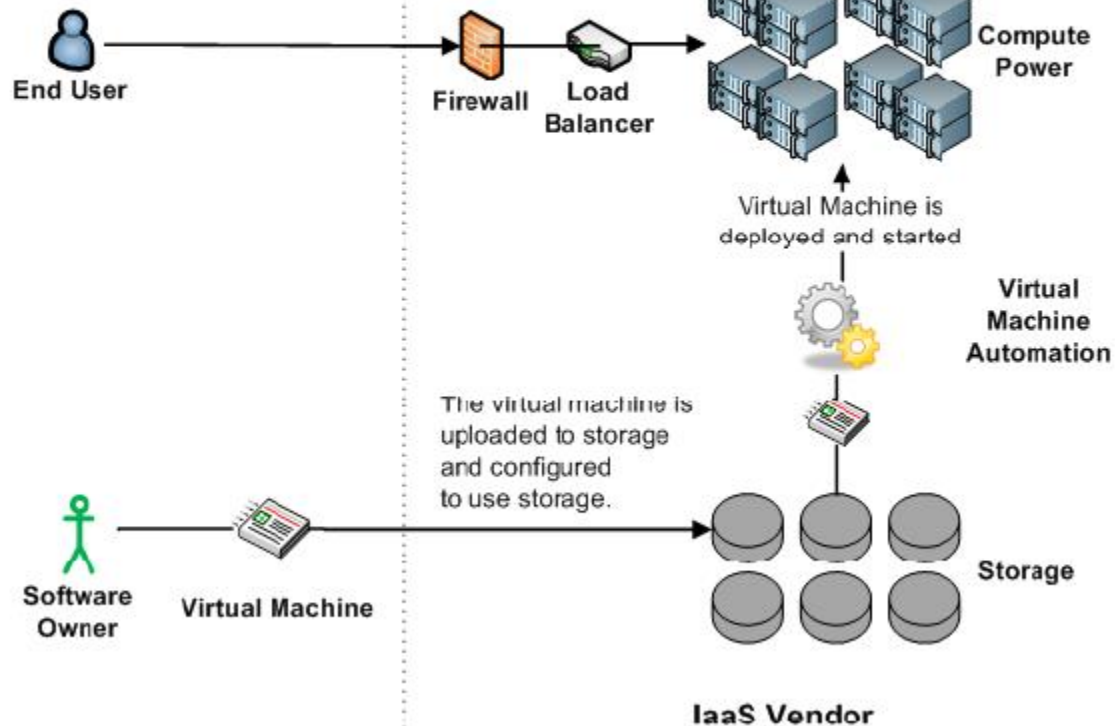
- Computación
- Almacenamiento
- Comunicación

¿Qué compone la infraestructura?



Escenario de uso

The end user sees a finished application



¿Dónde se ubica?

- Formados por “granjas” o “*datacenters*” de ordenadores interconectados
- Dispersión geográfica (mundial)
- Ejemplo: [AWS DataCenters](#)



Características esperadas

- Alta disponibilidad
- Capacidades ilimitadas
- Sin problemas reales (no se rompen, no se “caen”, ...)
- Robustez y redundancia

Máquina Virtual

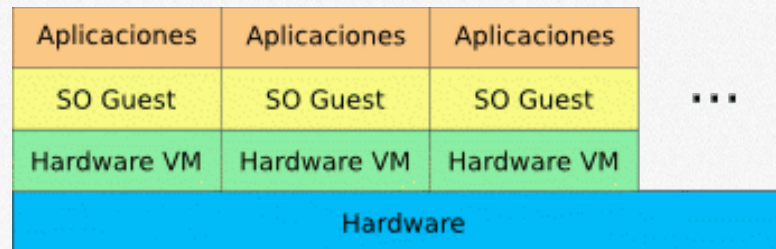
- Software que **simula** una computadora y puede ejecutar programas como si fuese una computadora real
- “With this technique you can ‘partition’ a single computer to act as if it were several independent computers, allowing the system to run several operating systems at the same time”

Máquina Virtual (2)

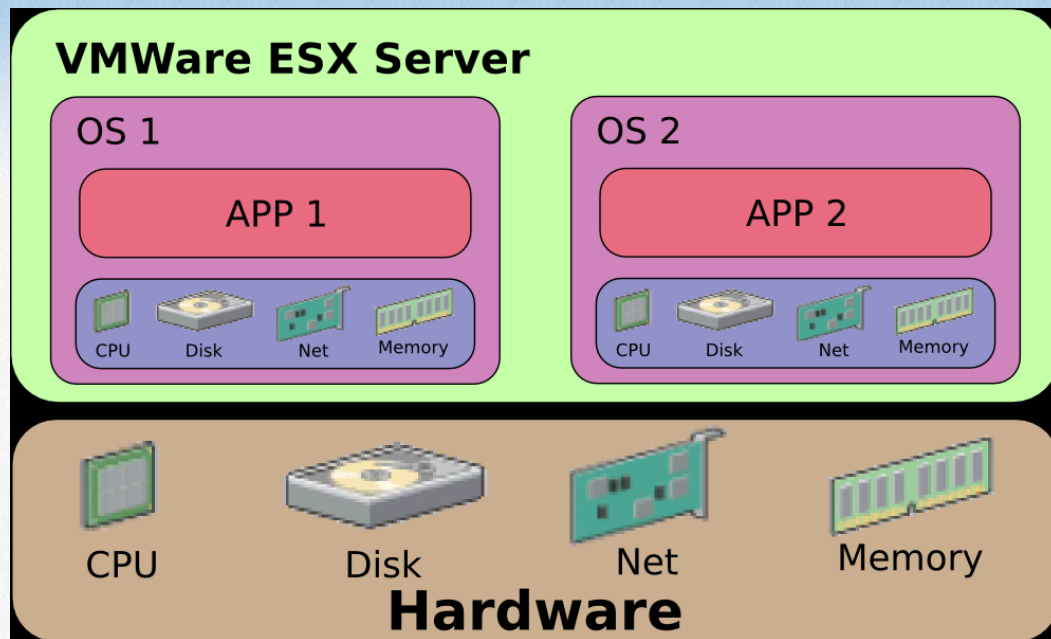
- Abstracción de los recursos de una máquina, de **modo transparente** al usuario, ocultando los detalles técnicos, mediante la encapsulación de los recursos
- La simulación no suele ser al 100% y hay ciertas limitaciones asociadas a los recursos disponibles y abstracciones proporcionadas

Conceptos básicos

- *Host*: máquina física que alberga todos los procesos
- *Guest(s)*: sistema operativo virtualizado
- *Hipervisor*: abstracción del hardware y de los S.O.



VMWare



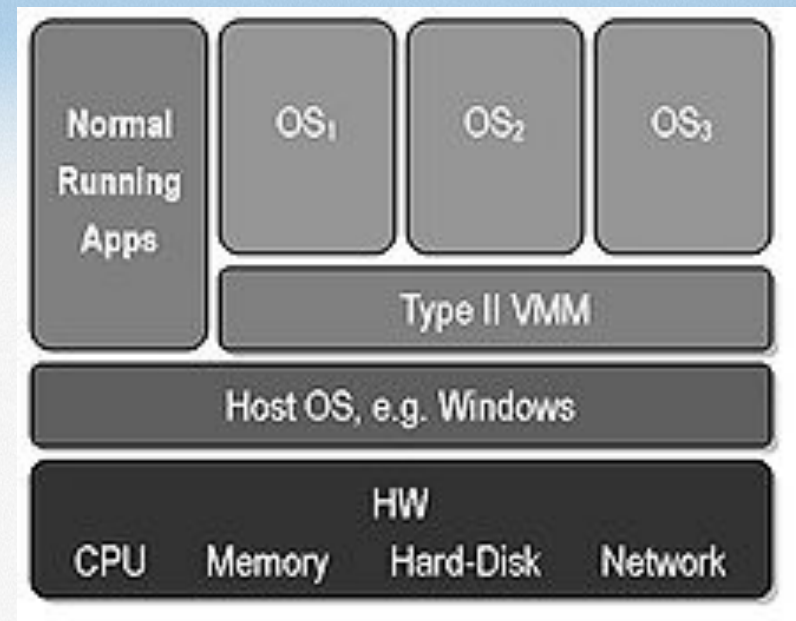
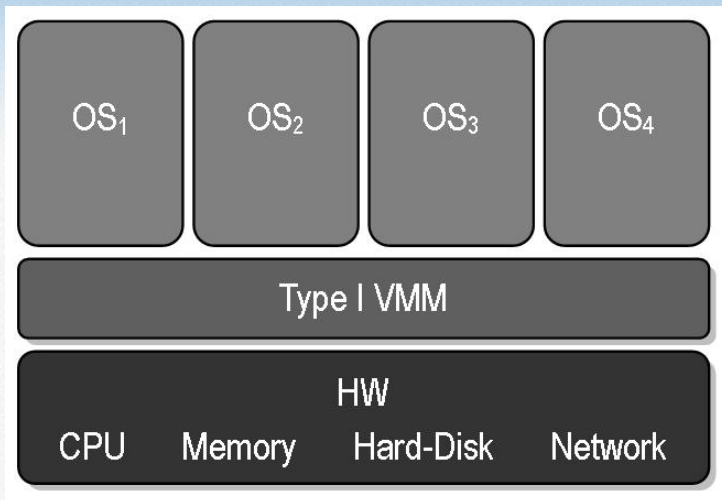
Máquina virtual de proceso

- Es un proceso normal y soporta sólo un proceso. No hay soporte para S.O., ni otros procesos
- Ej. Máquina virtual de Java
 - Procesador
 - Registros virtuales
 - Gestión de memoria (física, virtual)

Hypervisor

- Monitor de máquina virtual: explota distintas técnicas de virtualización para utilizar, al mismo tiempo, distintos S.O. en un mismo ordenador

Hipervisores nativos y alojados



Desacoplamiento entre hardware y software

- **Particionamiento**: Separación de los recursos físicos, compartidos entre diversos S.O. y aplicaciones
- **Individual**: cada MV es independiente del host físico. Si cae una MV, no afecta al resto
- **Encapsulamiento**: Una MV se gestiona como un único fichero.

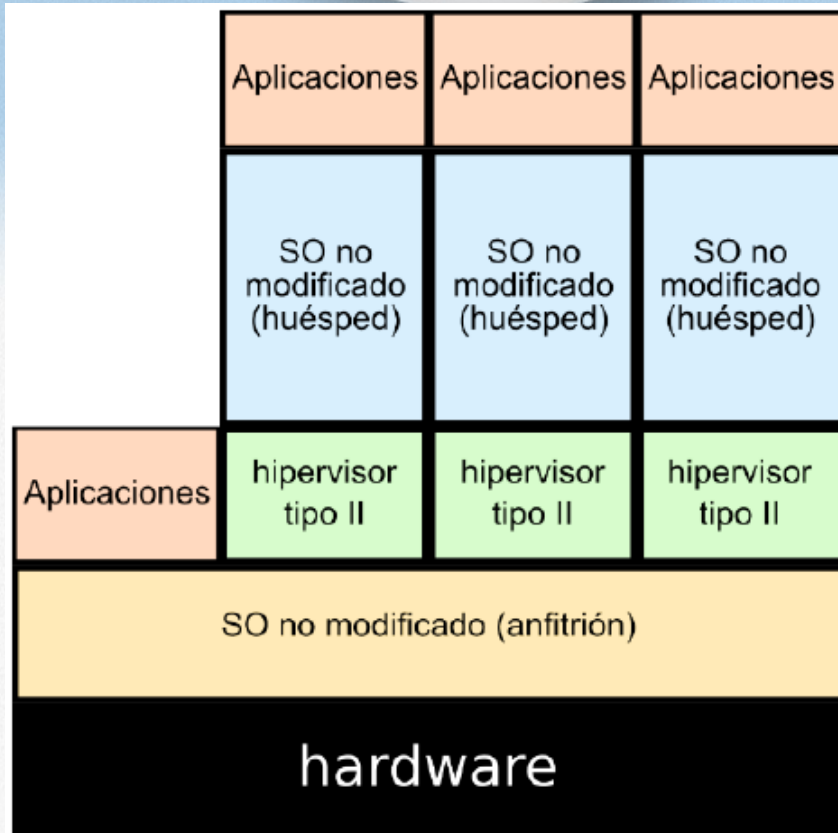
Principales hipervisores

- Xen
- Qemu
- VirtualBox
- VMWare
- Virtual PC
- Oracle VM

Tipos de virtualización

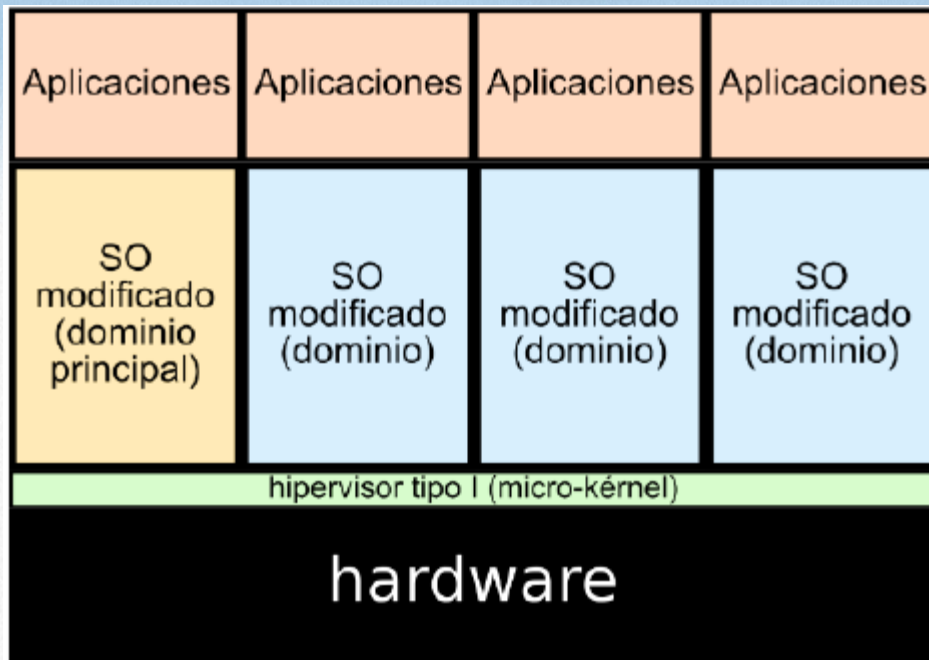
- Completa
- Paravirtualización
- A nivel de S.O.

Virtualización completa



- Ventajas:
 - No modifica el S.O. huésped
- Inconveniente:
 - Peor rendimiento
- Ejmplos:
 - VMWare Server
 - VirtualBox
 - Qemu

Paravirtualización



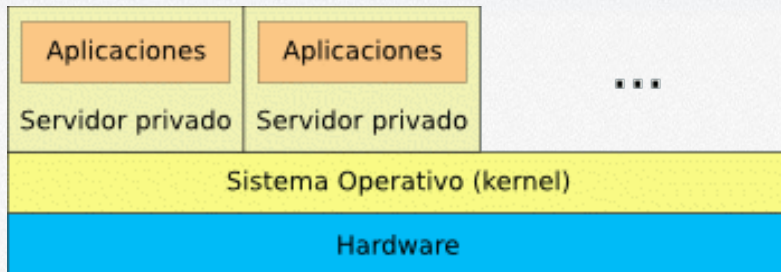
- Ventajas:
 - Mejor rendimiento
 - Facilita la migración
- Inconvenientes:
 - S.O. modificados
- Ejemplos:
 - Xen
 - Hyper V
 - VMWare ESX

Emulación del hardware

- Ventajas:
 - Emular distintas plataformas hardware:
p.ej. X86 sobre SPARC
- Inconveniente:
 - Alto coste de traducción de cada una de las operaciones



Virtualización a nivel de S.O.



- Ventajas:
 - Separación de los procesos de usuario sin pérdida de rendimiento (apenas)
- Inconvenientes:
 - Comparten todos el mismo kernel: No se virtualiza el hardware
- Ejemplos:
 - Docker
 - OpenVZ
 - Virtuozzo

Almacenamiento en la nube

- Debemos distinguir dos tipos:
 - Servicios puros de almacenamiento de datos
P.ej: partición virtual
 - Datos generados dentro de un sistema cloud

Seguridad sobre almacenamiento

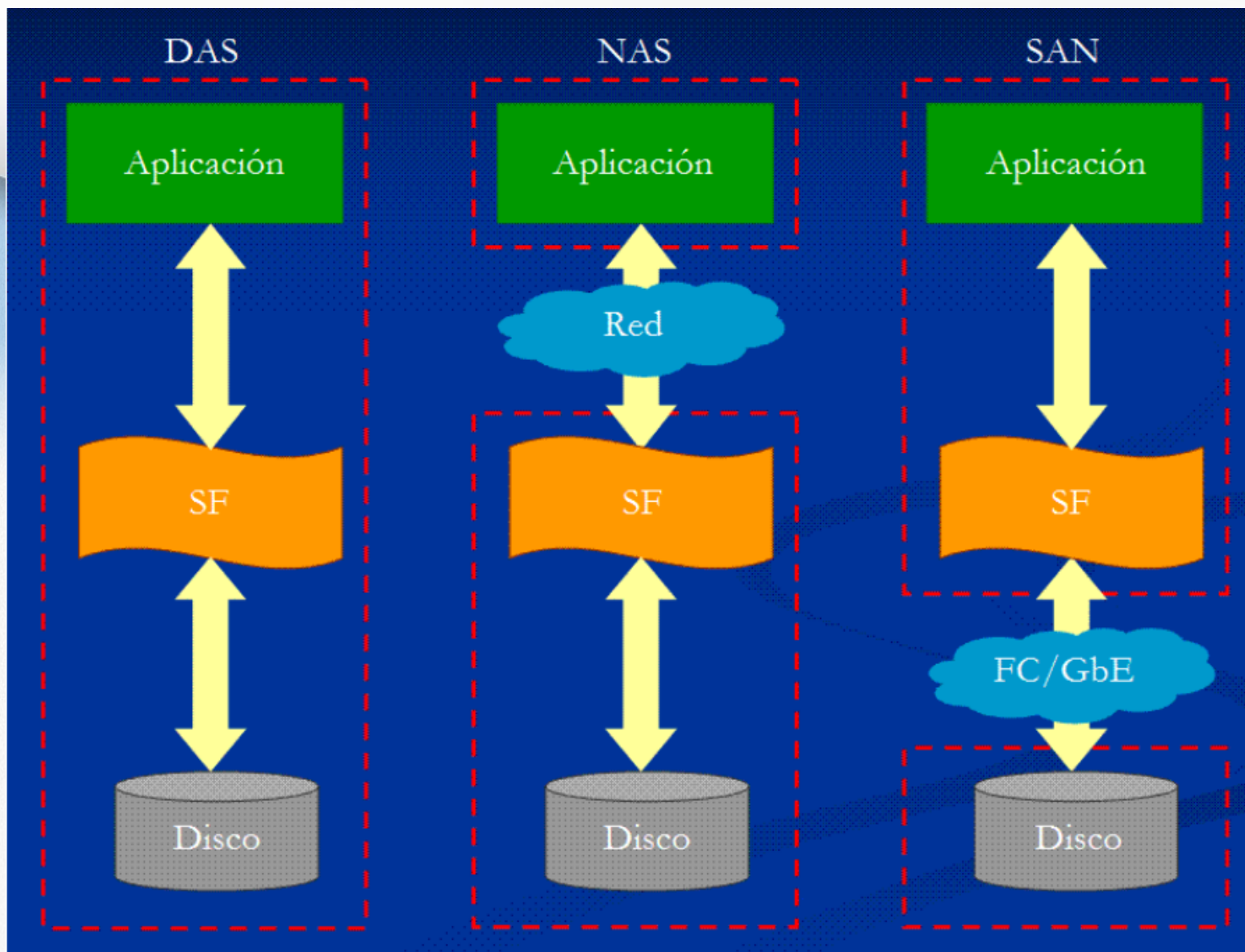
- Se centra en:
 - Localización
 - Control
 - Transferencia
- Los proveedores debe garantizar la seguridad y privacidad, aunque el responsable último es el usuario

Tipo de almacenamiento

- A. **por volúmenes**: grupo de bloques asignados.
- A. **por objetos**: Se estructura como un sistema híbrido de base de datos/archivo
- En ambos casos el cliente final trabaja de forma transparente como si se tratase de un HDD normal

Tecnologías para almacenamiento en la nube

- **Network Attached Storage (NAS)**: Almacenamiento a nivel de archivo compartido mediante un servidor dedicado que está conectado a una red. Acceso vía NFS, Samba, AFP
- **Storage Area Network (SAN)**: Almacenamiento a nivel de bloque, el dispositivo está directamente compartido en la red. P.ej. FS lustre
- **Direct Attached Storage (DAS)**: HDD conectado directamente al ordenador, sin mediar red. Peticiones a nivel de bloque y permite acceso concurrente



Servidores y micros para cloud

- **Servidores estándar:** ordenador completo con factor de forma rack
- **Servidores blade:** Servidores diseñados para aprovechar espacio, reducir el consumo y simplificar explotación. Elementos hardware compartidos. Suelen tener sólo microprocesador, memoria y buses

Cluster de servidores estándar

- Ventajas:
 - Los fallos hardware sólo afecta al servidor en cuestión
 - Mayor flexibilidad
- Inconvenientes:
 - Amplio cableado
 - Mayor coste económico



Servidores blade



- Ventajas:
 - Más barato
 - Menor espacio y menor cableado
- Inconvenientes:
 - Los errores afectan a todos los blades
 - Menor flexibilidad

Microprocesadores

- Los fabricantes aportan esfuerzos para el diseño orientado a cloud computing
- Consumo energético minimizado
- Soporte en hardware para la virtualización

Tipos de micros

- Intel: La familia Xeon regula el consumo y soporta VT FlexMigration
- AMD: La familia Opteron contiene tecnología AMD-P y AMD-V
- Tiler: La serie TILE-Gx 3100 es de bajo consumo y baja frecuencia
- En linux, buscar xvm o svm en /proc/cpuinfo

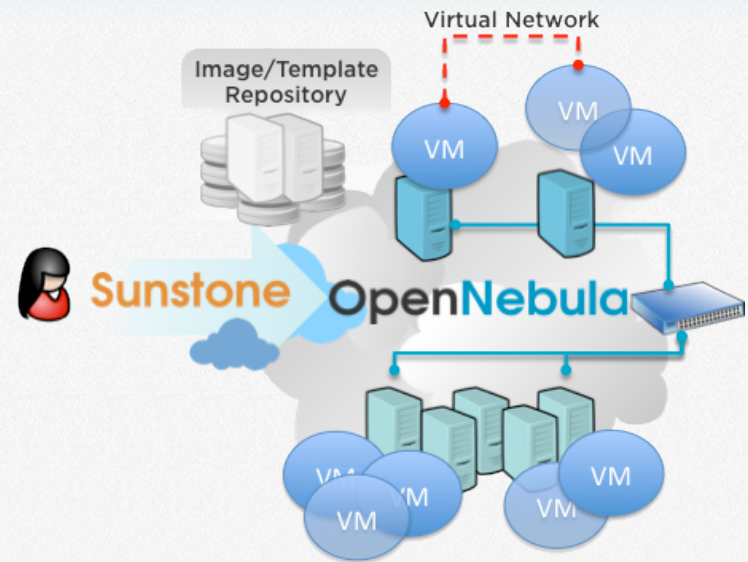
Proveedores de IaaS

- Amazon Elastic Computer Cloud (EC2)
- Google Compute Engine
- Azure
- RackSpace



OpenNebula

- Software open-source para proporcionar un servicio de IaaS



OpenStack



- Software para construir clouds privados y públicos
- Pretende ofrecer soluciones para todo tipo de clouds siendo sencillo de implementar, masivamente escalable y con muchas características
- Respaldado por más de 200 compañías: RackSpace, Dell, AMD, Intel, HP, Cisco, RedHat, Yahoo!, Nasa, Fujitsu, ...

Soporte para OpenStack

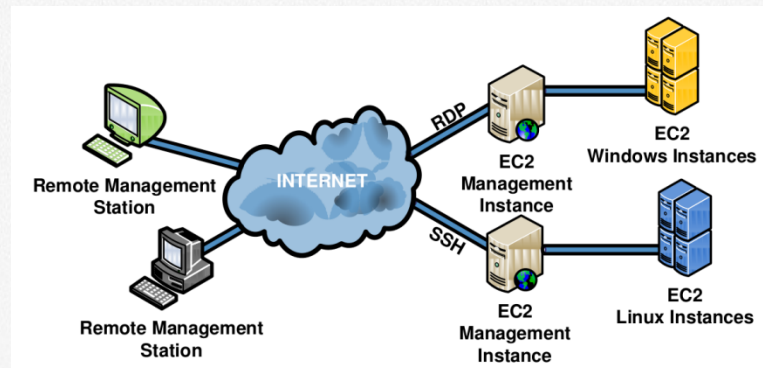
- Plataformas que soportan OpenStack:
 - Ubuntu
 - Red Hat (Linux Enterprise)
 - Fedora
 - Oracle Linux
 - Oracle Solaris
 - SuSE
 - VMWare
 - Rackspace

Amazon Elastic Compute Cloud (Amazon EC2)

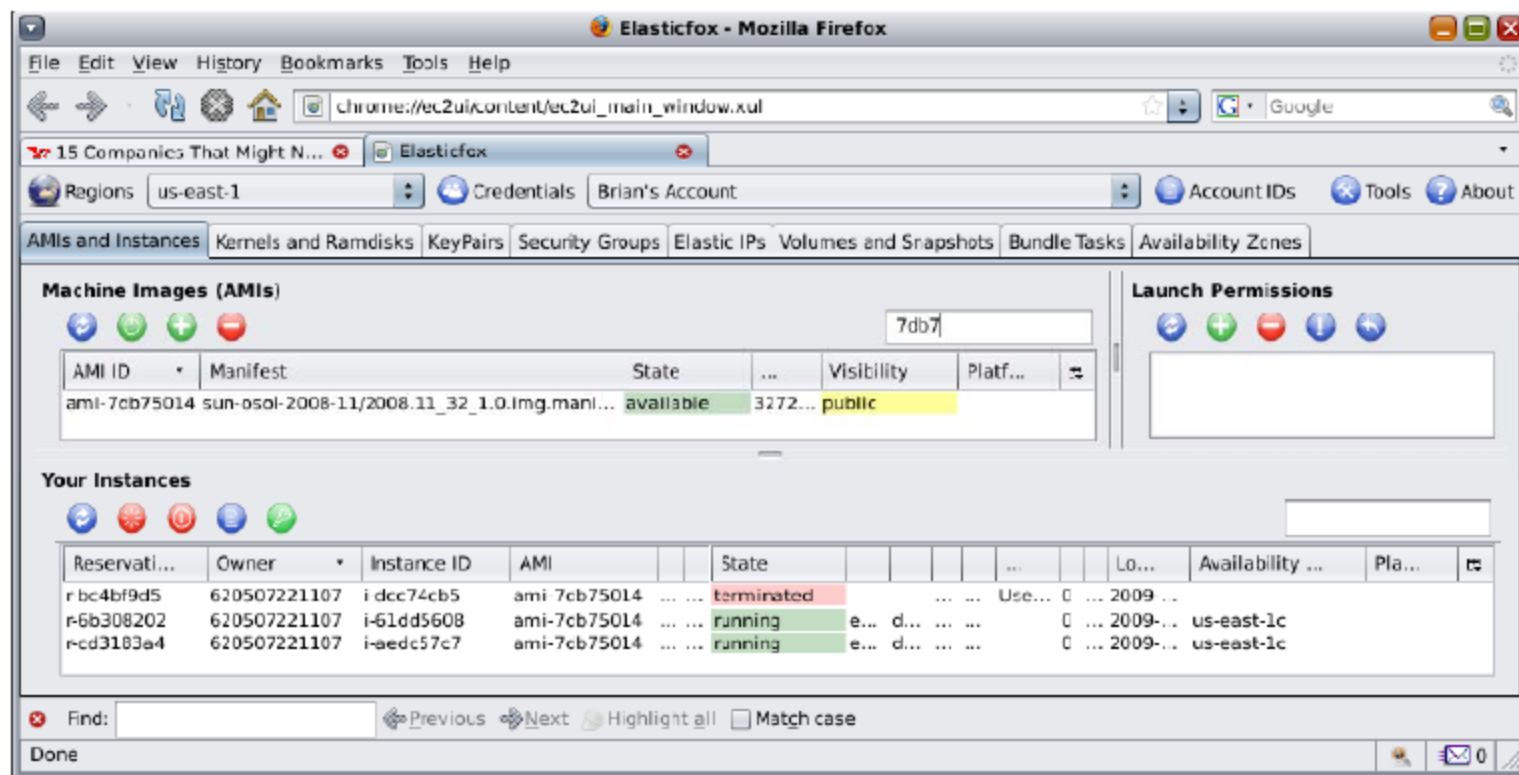
- Amazon EC2 se presenta como un servicio web que ofrece:
 - Virtualización bajo demanda
 - Escalabilidad
- Su infraestructura distribuida geográficamente
- Factura por hora de uso y ancho de banda consumido
- Permite VMs basadas en Linux, OpenSolaris, Windows, ...
- Caso de uso: aws.amazon.com
 - Arquitectura
 - Tarifas

Amazon EC2

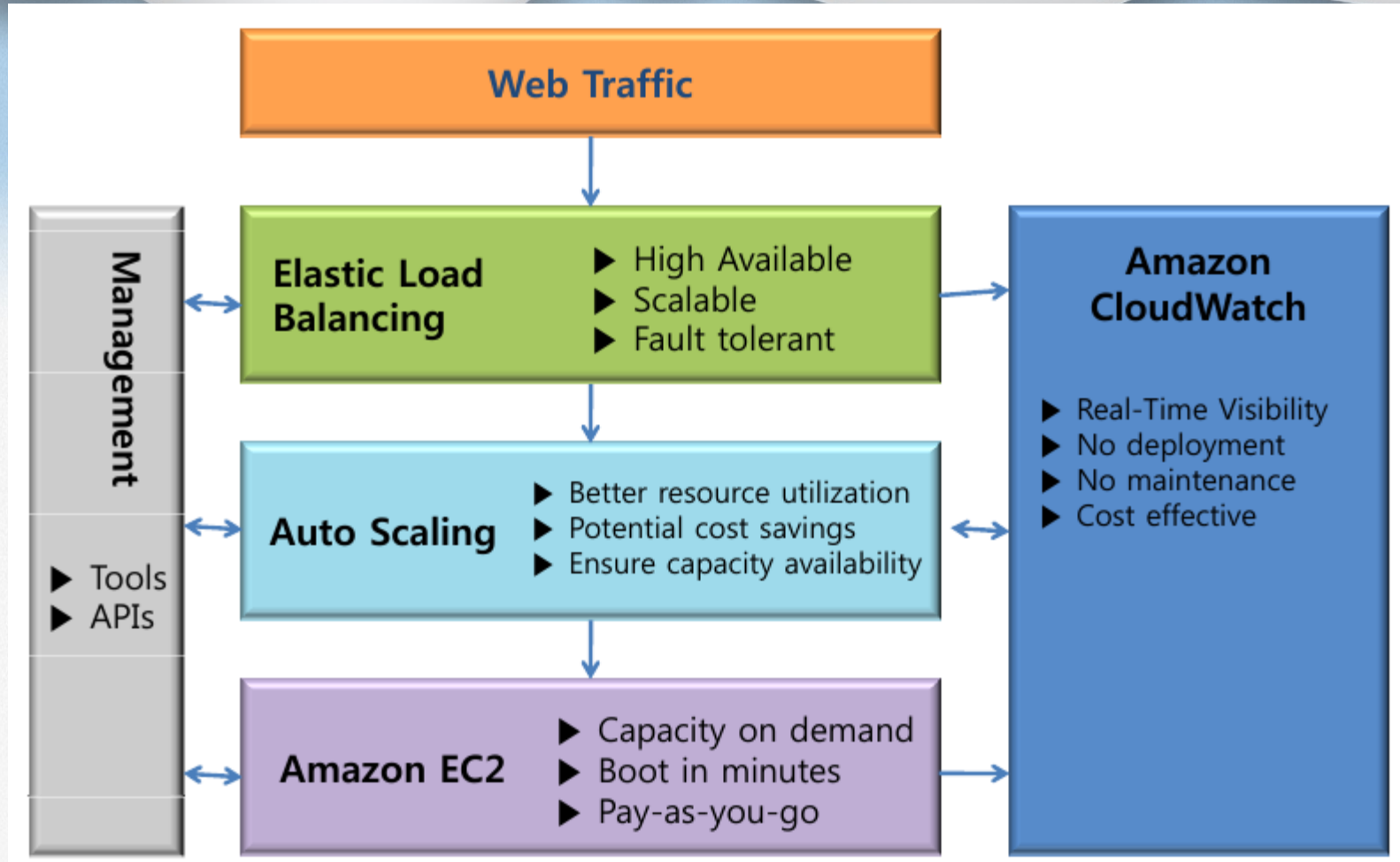
- Usa paravirtualización con Xen
- Dispone de MVs preconfiguradas llamadas Amazon Machine Images, accesibles por ssh o escritorio remoto
- Se conecta con Amazon S3 para acceso a grandes volúmenes de datos



Interfaz de usuario

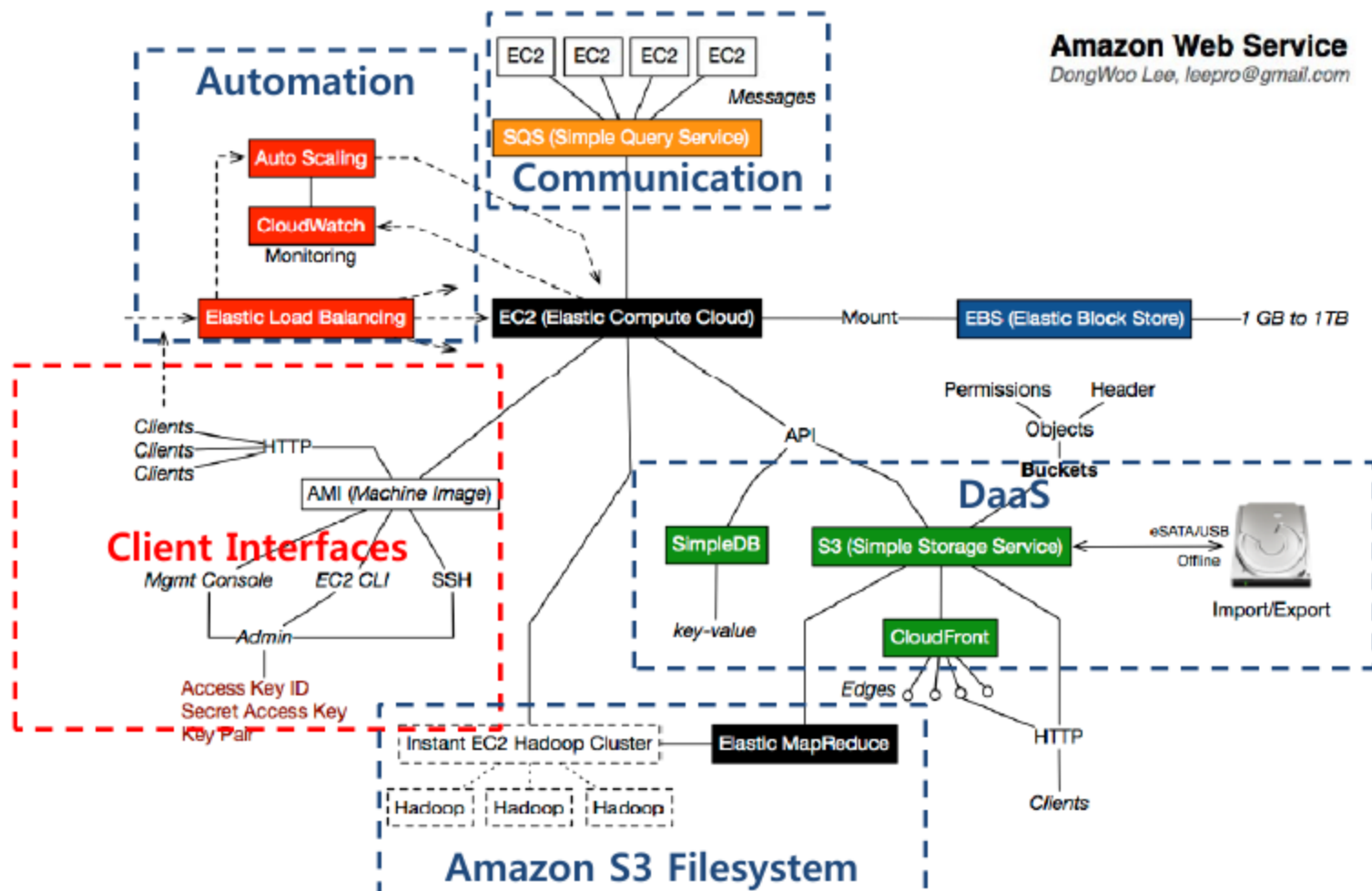


Esquema de gestión



Amazon Web Service

DongWoo Lee, leepro@gmail.com



Otros IaaS comerciales



Google Cloud Platform

- <https://cloud.google.com/free-trial/>



- <http://azure.microsoft.com/es-es/pricing/free-trial/>