

Arquitectura de Clústers



Dept. Arquitectura de Computadores
Universidad de Málaga

Curso 2016/17

Tendencias Tecnológicas en el Pasado

- **Procesador**

- **Velocidad:** 1.35x por año (antes de 1985); 1.58x por año (después)
- **Coste:** Se reduce a la mitad por año (1983-90)

- **Memoria (DRAM)**

- **Capacidad:** 1.6x por año
- **Tiempo de ciclo:** 1/3 en 10 años
- **Coste:** se reduce en un factor 2.7 por año

- **Disco magnético**

- **Capacidad:** 1.25x por año (antes de 1990); 1.5x por año (después)
- **Tiempo de acceso:** se reduce a la mitad en 10 años (1980-90)
- **Coste:** se reduce en 1/3 por año (1986-92) y 1/2 por año (1992-95)

Tendencias Tecnológicas en el Pasado

- **Tecnologías de interconexión**

- **Bus de sistema:** de 10 MB/s (Multibus-I) a 1 GB/s (Sun SMP)
 - » **Una mejora de un factor 100**
- **Red de interconexión:** de 1 MB/s (CM-2) a 300 MB/s (Cray T3D/E)
 - » **Una mejora de un factor 300**
- **Red de área local (LAN):**
 - » Ethernet (1974): 10 Mb/s (1.25 MB/s)
 - » Token Ring (1980): 4/16 Mb/s (0.5/2 MB/s)
 - » FDDI (1987): 100 Mb/s (12.5 MB/s)
 - » Fast Ethernet (1993): 100 Mb/s (12.5 MB/s)
 - » Actual 40GbE y 100GbE
 - » **Una mejora de un factor 10 en entre 1974 y 1993**

Impacto en el Diseño de Computadores

- Los computadores se diseñaban ignorando las LAN
 - Mainframes (terminales)
 - SMPs (bus o switches entre CPUs y memoria principal)
 - MPPs (redes ad-hoc)
 - PCs (conexión a Internet)
- LANs se estaban usando para:
 - Proporcionar conectividad entre diferentes sistemas
 - Proporcionar transferencia de ficheros entre diferentes computadores (cliente-servidor) y recursos (discos, impresoras ...)
 - Proporcionar transacciones entre sistemas

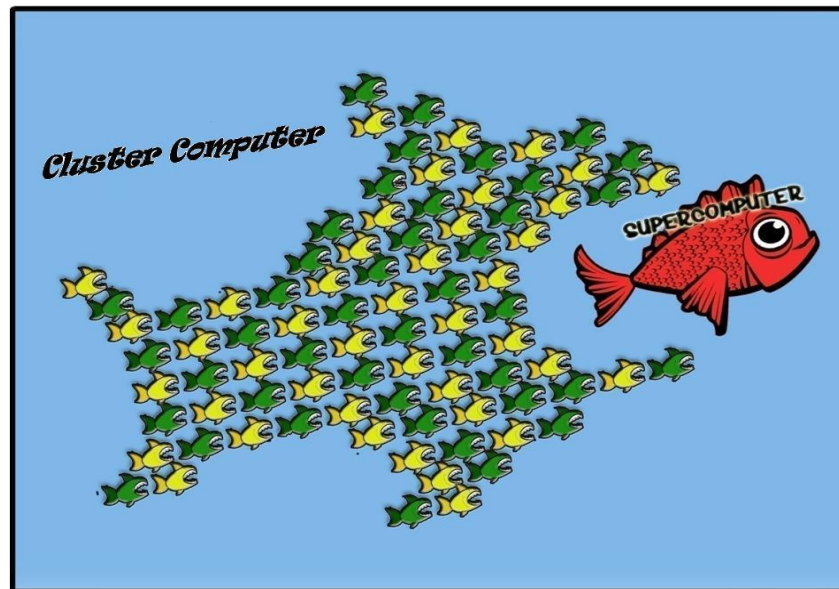
Tendencias Tecnológicas Actuales y Futuras

- Avances fundamentales en tres líneas:
 - Tecnologías de interconexión
 - » Una mejora de 1000 entre 1993 y 2015 (40GbE, 100GbE)
 - » Otros estándares: MiryNet, Infiniband, FibreChannel.
 - Tecnologías de almacenamiento
 - » Disco duros con 10000 veces más capacidad de almacenamiento que en 1990
 - Sistemas multicore y SMP potentes con gran cantidad de memoria
 - Desarrollo de las tecnologías Web
- Estos avances facilitan el desarrollo de sistemas grandes (muchos nodos), muy descentralizados y con una alta y compleja compartición de información y datos
- Computación en red (network computing)

Cluster

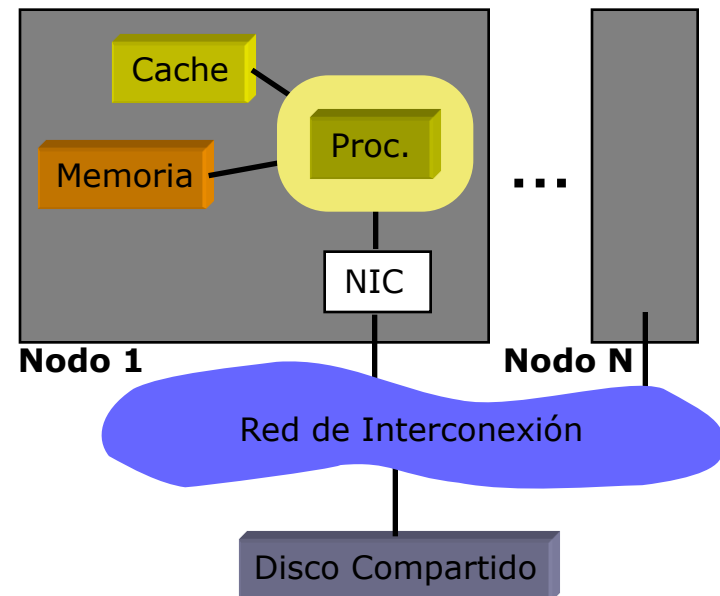
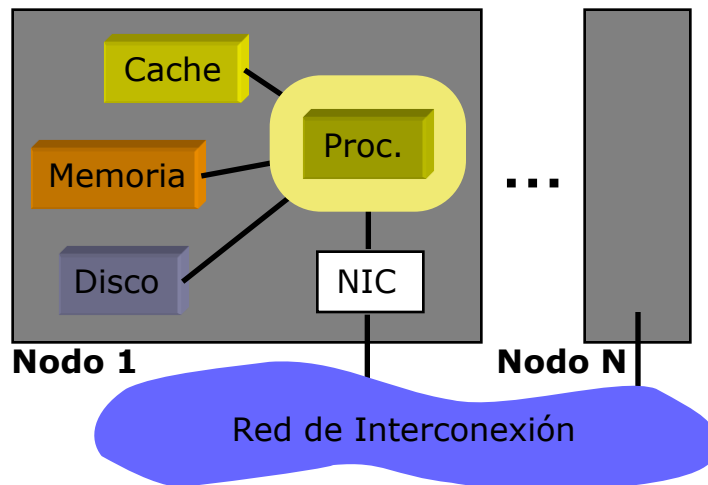
• Definición

- Un cluster es una tipo de arquitectura paralela distribuida que consiste de un conjunto de computadores independientes interconectados operando de forma conjunta como un único recurso computacional
- Sin embargo, cada computador puede utilizarse de forma independiente
- Los elementos que componen el clúster son estándares:
 - » procesadores y memoria
 - » red de interconexión
 - » almacenamiento



Compartición

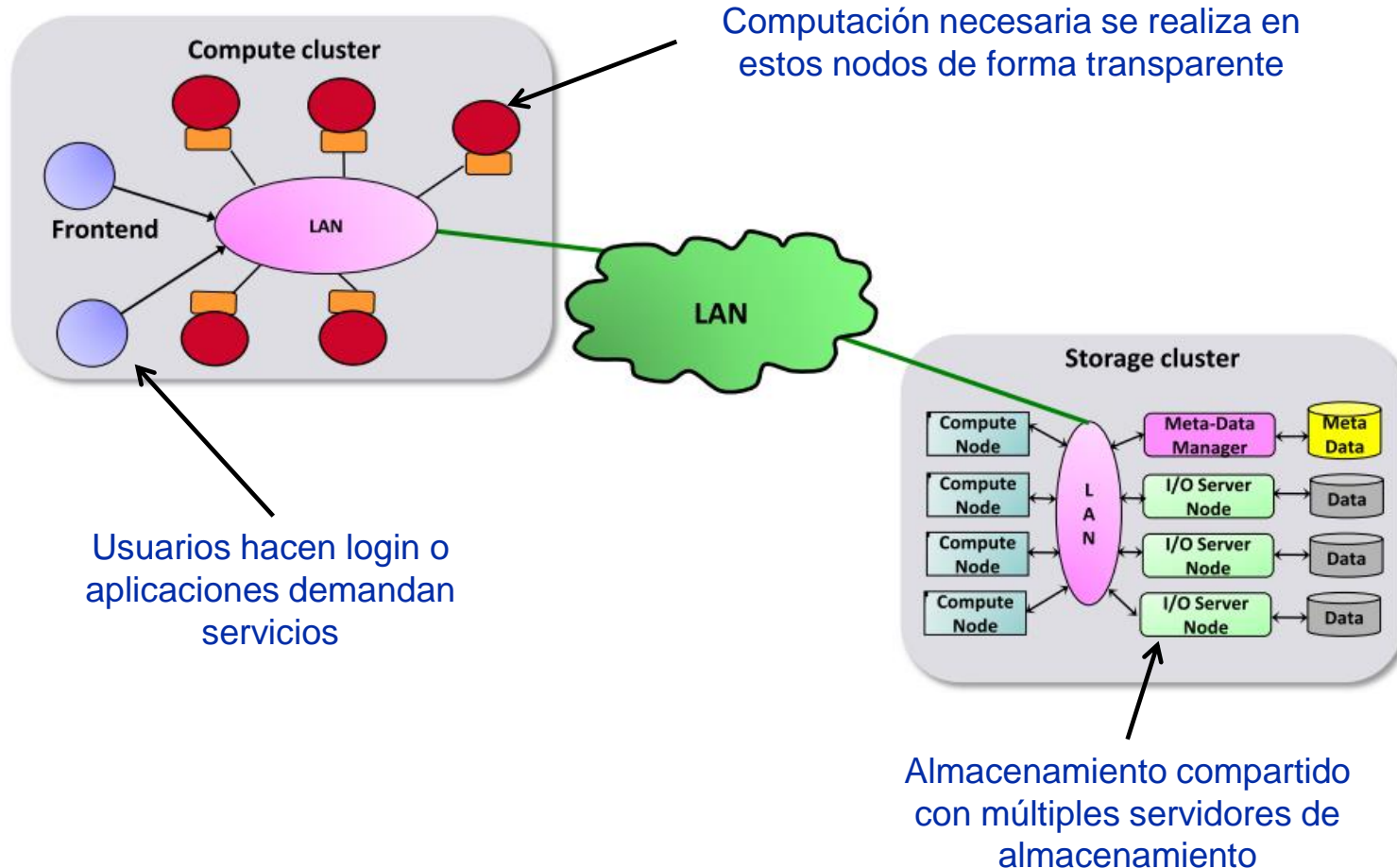
- Compartición de la red de interconexión y, posiblemente, de la memoria secundaria



Ejemplos redes de interconexión

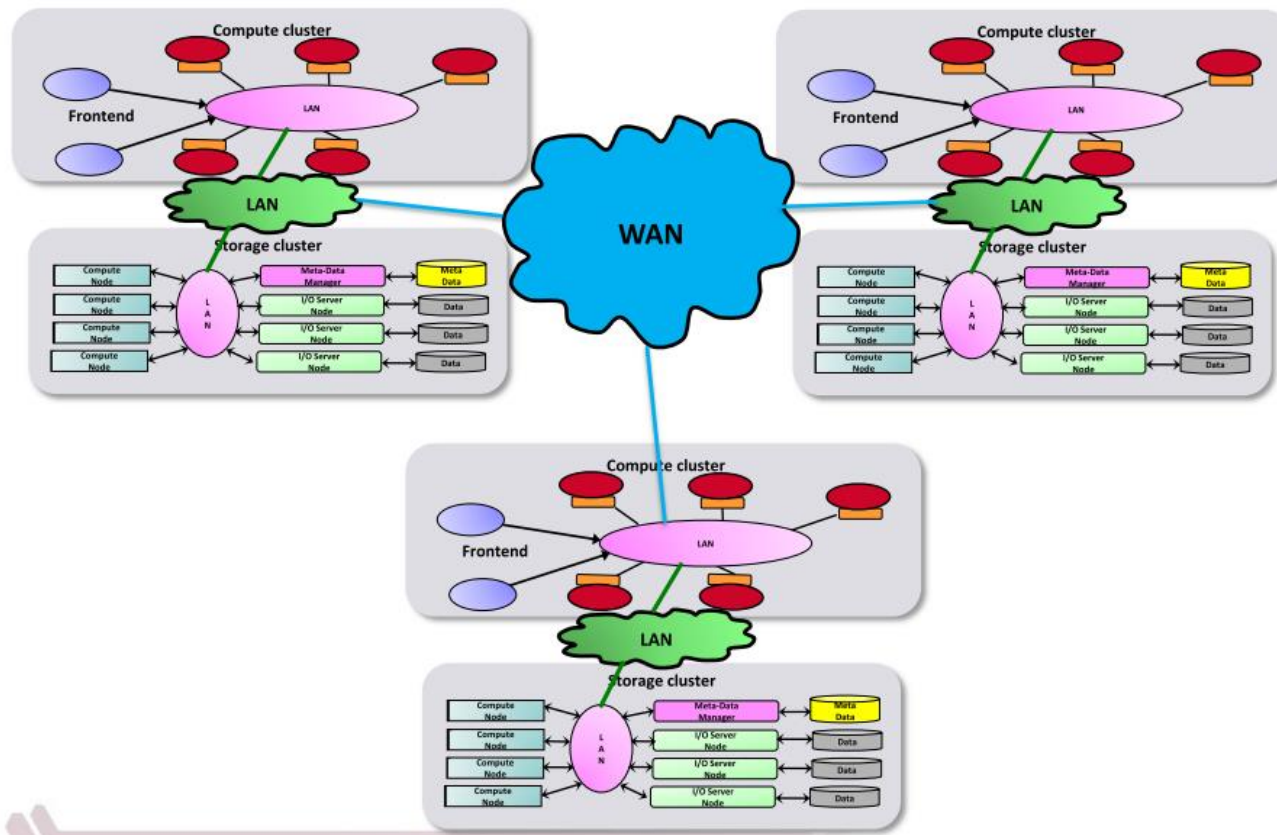
- » LAN (Local Area Network)
- » SAN (Storage Area Network)
- » WAN (Wide Area Network)

Computación en clústeres



Computación Grid

- Infraestructura de clústeres físicamente alejados interconectados a través de una WAN



Servicios de un clúster

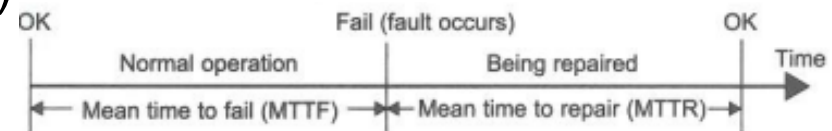
- **Alto rendimiento**

- Disminuir latencia de aplicaciones paralelas
- Alta productividad: ejecución de mayor número de tareas (procesos) por unidad de tiempo.

- **Alta disponibilidad**

- $\text{Availability} = \text{MTTF} / (\text{MTTF} + \text{MTTR})$

- » MTTF mean time to failure
- » MTTR mean time to repair



- **Balanceo de carga**

- Repartir trabajo entre recursos (computadores, enlaces de interconexión, discos, ...) para obtener mayor utilización de los recursos y disminuir el tiempo de operación.

- **Escalabilidad**

Servicios de un clúster

- Arquitecturas Escalables

- Un computador, incluyendo hardware y software, se dice **escalable** si podemos aumentar sus recursos para soportar una mayor demanda de rendimiento y funcionalidad, y/o disminuir sus recursos para reducir costes.

- Implicaciones

- **Funcionalidad y rendimiento**

- » La potencia computacional del sistema debe incrementarse de forma proporcional al aumento en recursos. De forma ideal, debiera observarse un aumento de capacidad computacional próximo a un factor N cuando se mejoran los recursos en ese mismo factor

- **Escalado de coste**

- » El coste de escalado debe ser razonable. Típicamente, si el escalado se produce en un factor N , se espera que su coste sea un factor entre N y $N \log N$

- **Compatibilidad**

- » Los componentes ya existentes, hardware y software, deben poderse reusar cuando se proceda al escalado del sistema (*escalabilidad incremental*).

- Un clúster es una sistema altamente escalable

Clasificación de clústeres

- **Alto rendimiento:**

- Ejecutan tareas que requieren de gran capacidad computacional, grandes cantidades de memoria, o ambos a la vez.
 - » El llevar a cabo estas tareas puede comprometer los recursos del clúster por largos periodos de tiempo.

- **Alta disponibilidad**

- Proveen de disponibilidad y confiabilidad. Estos clústeres tratan de brindar la máxima disponibilidad de los servicios que ofrecen.
 - » La confiabilidad se provee mediante software que detecta fallos y permite recuperarse frente a los mismos, mientras que en hardware se evita tener un único punto de fallos.

- **Alta eficiencia**

- Son clústeres cuyo objetivo de diseño es el ejecutar la mayor cantidad de tareas en el menor tiempo posible.
 - » Existe independencia de datos entre las tareas individuales. El retardo entre los nodos del clúster no es considerado un gran problema.

Clasificación en función del uso

- Clústeres comerciales

- Alta disponibilidad y alta eficiencia

- Clústeres científicos

- Alto rendimiento

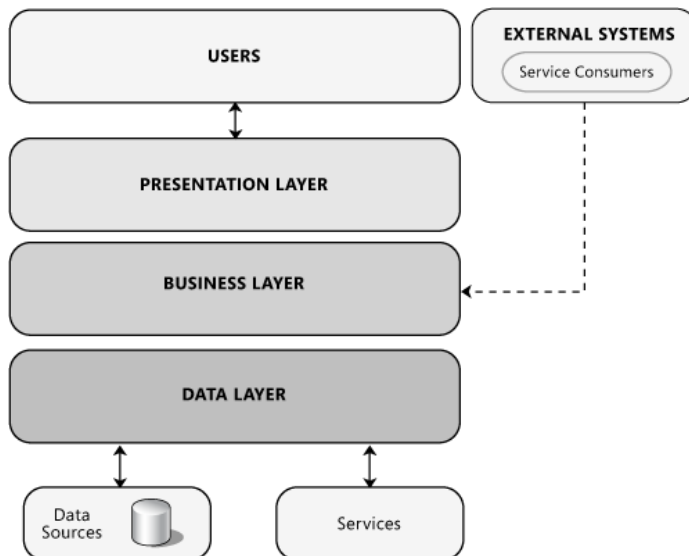
- Clúster de almacenamiento

- Nodos servidores conectados a diferentes data appliances a través de una red de interconexión
 - » Ofrecen almacenamiento incrementando el rendimiento y la disponibilidad

Diseño de configuraciones clúster

- Se sigue el modelo estándar para soluciones de infraestructura:
 - Los módulos de las aplicaciones se estructuran en capas (**layers**).
 - Cada capa se ejecuta en una ubicación determinada de la infraestructura hardware (**tier**) diseñada en función de la carga que genera.
 - El diseño de la plataforma hardware está guiado por el proceso que asocia cada capa software (**layer**) a una ubicación hardware (**tier**).

Microsoft Application Architecture Guide



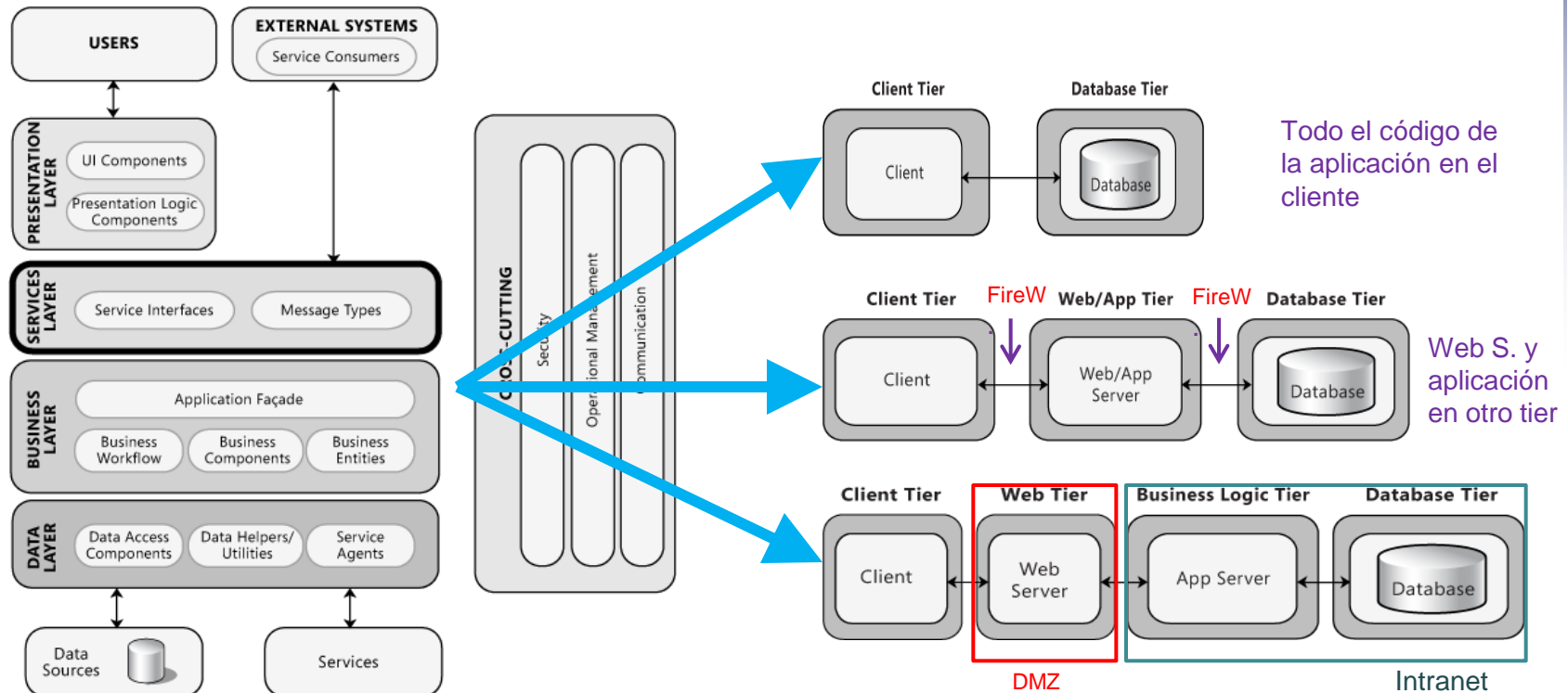
Funcionalidad orientada al usuario que usa el sistema.

Lógica del negocio. Se puede añadir una capa de servicios que ofrece la funcionalidad del negocio a aplicaciones

Organización y acceso a los datos del sistema

Diseño de configuraciones clúster

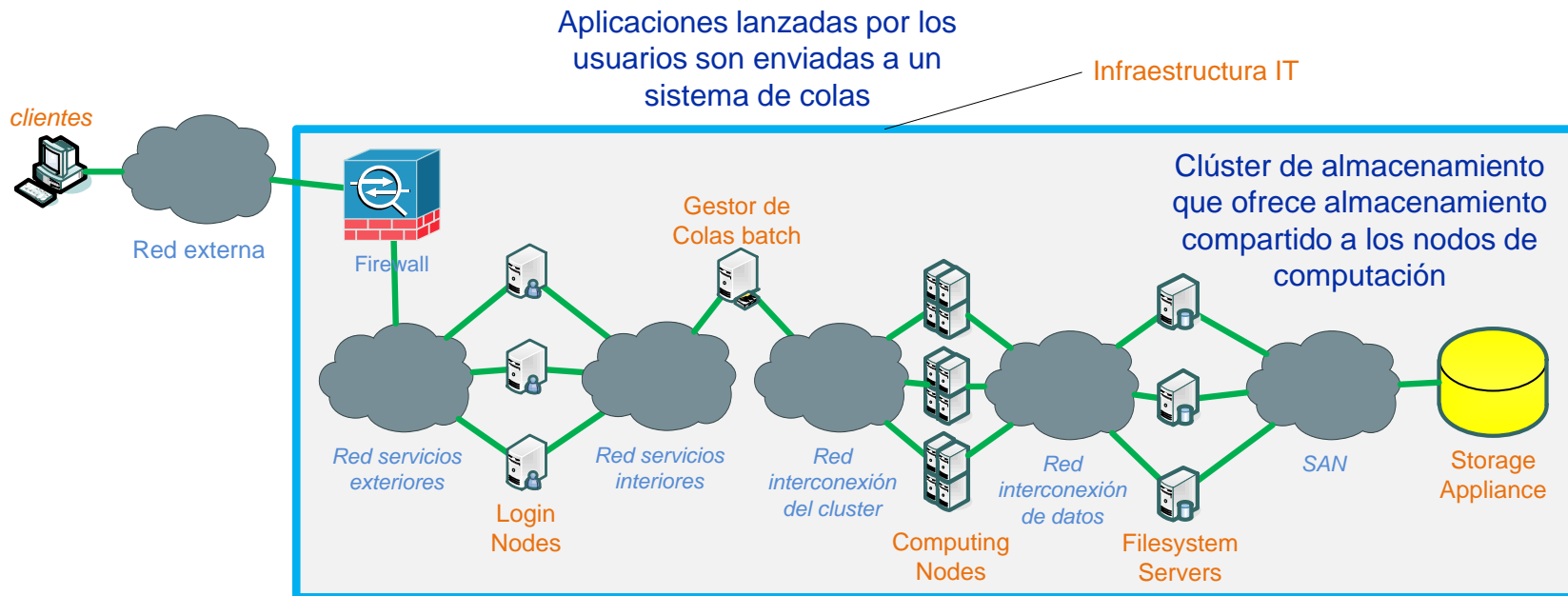
- Layers de arquitectura software son mapeados a tiers hardware



Posibles alternativas de mapeo

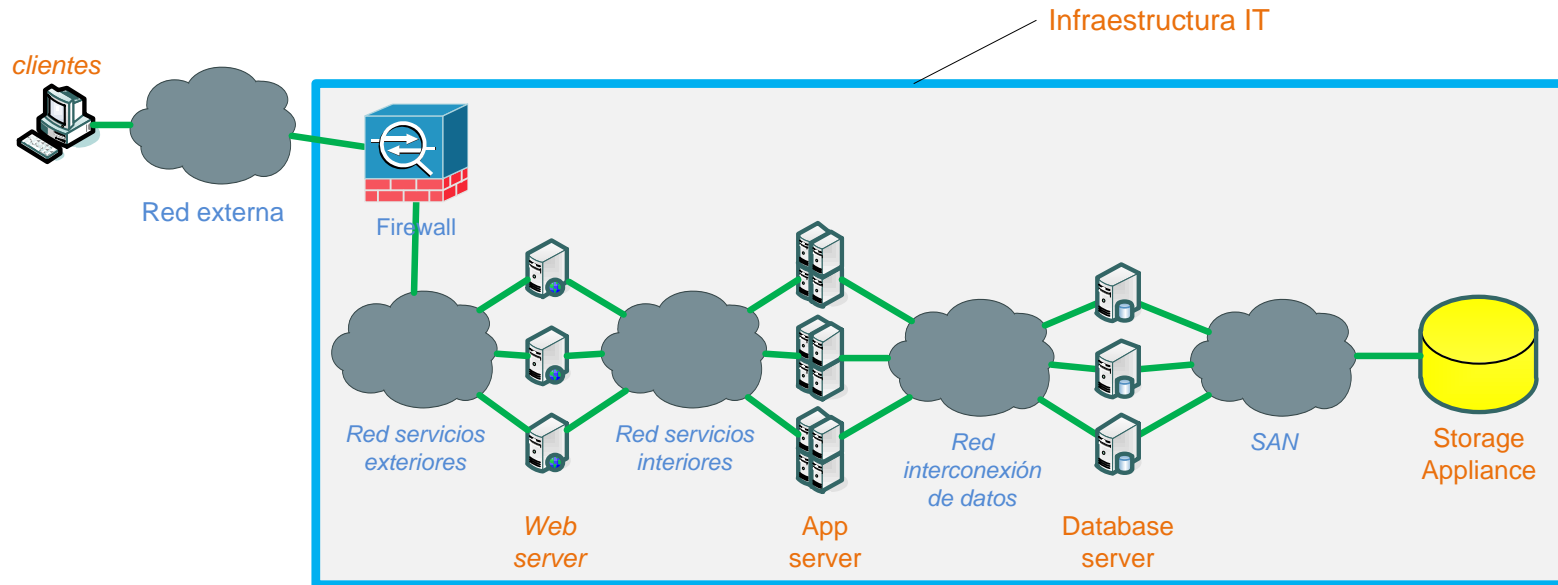
- Distintos tiers pueden ser implementados con clústeres por necesidades de rendimiento, productividad y/o disponibilidad

Ejemplo: Clúster de computación



Aplicaciones son ejecutadas en los nodos de computación

Ejemplo: Clúster de base de datos

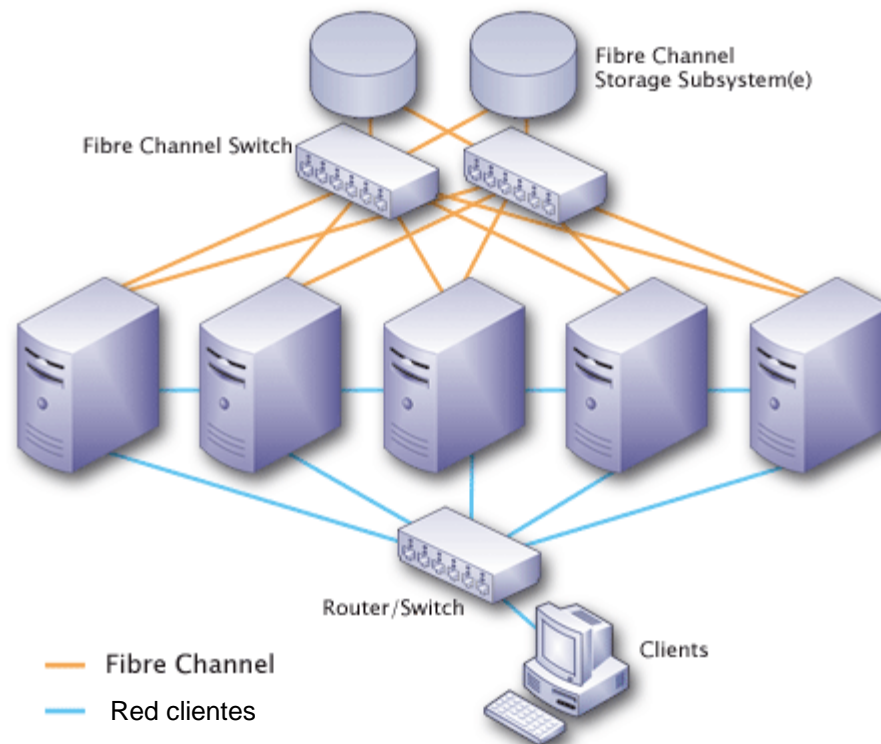


Acceso por HTTP y HTTPS desde el exterior.
 Procesan código fuente empotrado en páginas HTML.
 Realizan llamadas remotas (RPC) a los servicios de los nodos de aplicación.

Ejecutan subrutinas iniciadas por las llamadas remotas (RPC) a los servicios.
 También ejecutan procesos de cálculo intensivo periódicos para procesar datos.
 Realizan llamadas remotas (RPC) a los servicios de los nodos de **base de datos**.

Comparten el acceso a los mismos discos mediante una SAN y una base de datos en cluster (Oracle, DB2, MySQL, SQL Server, etc). Usan bastante RAM como caché de las tablas.
 Procesan transacciones de base de datos encargadas por los nodos de aplicación a través del protocolo de red del SGBD.

Ejemplo: Clúster de almacenamiento



Ejemplo cluster real: Picasso

