

Introducción a los clústeres



IES Gonzalo Nazareno
CONSEJERÍA DE EDUCACIÓN

Alberto Molina Coballes



Septiembre 2013

Clústeres. Conceptos

- Un clúster es un conjunto de equipos independientes que realizan alguna tarea común en la que se comportan como un solo equipo.
- Ofrece uno o varios de los siguientes servicios:
 - Alto rendimiento
 - Alta disponibilidad
 - Balanceo de carga
 - Escalabilidad
- Surgieron como alternativa al crecimiento vertical de los ordenadores: Al aumentar los requisitos computacionales, en vez de sustituir el equipo por uno nuevo más potente, se añade una máquina más y se crea un clúster entre ambas.
- Con la generalización del uso de tecnologías de virtualización y cloud computing, los clústeres se están convirtiendo en componentes comunes en todo tipo de instalaciones.
- Los equipos dentro de un clúster reciben el nombre de nodos.

Clústeres. Tipos de máquinas

Máquinas físicas La solución tradicional. Salvo para situaciones específicas, está en progresivo desuso por ser la opción más costosa.

Máquinas virtuales Mucho más versátil que el uso de máquinas físicas ya que permite ajustar las características de los equipos a las necesidades reales. Se puede crear un cluster de máquinas virtuales en un solo equipo físico.

Instancias de cloud Añade dos nuevos conceptos que se adaptan muy bien a la tecnología de clústeres: automatización y elasticidad

Obviamente pueden utilizarse distintas combinaciones de las tres opciones anteriores

Tipos de clústeres: Clúster de alto rendimiento

- *High Performance Computing Cluster* o HPC cluster
- Se suelen utilizar equipos físicos
- Imprescindibles para cálculos complejos que requieran gran capacidad computacional
- En la mayor parte de los casos, los grandes ordenadores han sido sustituidos por clústeres de equipos "convencionales" (arquitectura x86)
- Se utilizan principalmente en centros de investigación, ingeniería y otras actividades que requieran cálculos complejos

Visita recomendada: <http://www.top500.org>

Tipos de clústeres: Clúster de alta disponibilidad

- High availability (HA)
- Implementa tolerancia a fallos (failover) para garantizar el servicio en caso de que ocurra algún fallo.
- Utiliza nodos redundantes que sustituyen o complementan a los existentes
- Se suelen implementar con balanceo de carga para repartir las peticiones entre varios nodos
- Ampliamente utilizado en servicios de Internet
- Es el tipo de clúster que vamos a ver en este curso

Balaneo de carga

- Es una propiedad que permite repartir el trabajo entre varios nodos del clúster
- Se utiliza algún algoritmo para el reparto de tarea entre los nodos del clúster: aleatorio, round robin, carga de los nodos, tiempo de respuesta, . . .
- Aunque no es habitual, se puede implementar un clúster de balanceo de carga que no incluya alta disponibilidad

Ejercicio 1 Balanceo de carga con DNS

Escalabilidad. Conceptos

- Es una propiedad que permite modificar el tamaño del clúster en función de las necesidades
- Es difícil y lento implementar un clúster escalable con equipos físicos
- Se puede automatizar bastante la creación de un clúster escalable con máquinas virtuales
- Con instancias de cloud se puede ir aún más allá y se habla algunas veces de elasticidad en lugar de escalabilidad
- Tipos de escalado:
 - Escalado vertical** Varían las características de los nodos (CPU, RAM, espacio de almacenamiento)
 - Escalado horizontal** Varía el número de nodos del clúster

Escalabilidad. Elasticidad

Objetivo Conseguir un clúster escalable de forma automática y dinámica que se adapte en cada momento a las necesidades

- Con el término escalabilidad se asume que el clúster siempre se crece, mientras que con la elasticidad se remarca la adaptación a cada momento (creciendo o decreciendo)
- Amazon EC2: Amazon **Elastic** Compute Cloud
- IaaS: Escalabilidad al alcance de cualquiera porque se paga por uso. Cuesta lo mismo utilizar 10 nodos durante una hora que un nodo durante 10 horas.
- Se utiliza principalmente escalado horizontal
- Vídeo recomendado:
 - Autoscaling on OpenStack using Heat:
<http://www.youtube.com/watch?v=91LSZi6njq4>



Orquestación o automatización

- *Configuration Management Software* (CMS), también se utilizan los términos *orchestration* y *automation*
- ¿Has aprendido por fin a configurar un servicio utilizando el editor? Lástima, porque cada vez se hace menos así :-)
- No se puede pensar en elasticidad sin un sistema totalmente automático de configuración de los nodos
- Un sistema de orquestación o automatización permite crear máquinas virtuales, instancias de cloud, instalar paquetes, configurar servicios o utilizar plantillas de forma totalmente automática
- Los administradores de sistemas se centran cada vez más en crear recetas y plantillas para sus despliegues
- Puppet, chef, salt, ansible, ...

Mascotas vs Ganado

Para representar esta nueva situación respecto a los servidores, se ha sugerido hacer una analogía con las mascotas y el ganado ¹:

Mascotas (Escalado vertical)

- Tienen nombre
- Son únicas
- Se las cuida y mantiene

Ganado (Escalado horizontal)

- Número en vez de nombre
- Indistinguibles entre sí
- Cuando no sirven, se sustituyen

Máquinas físicas o virtuales

Clústeres e instancias de cloud

“Future application architectures should use Cattle but Pets with strong configuration management are viable and still needed”. Tim Bell



Almacenamiento. Conceptos

- Los sistemas de ficheros tradicionales solo pueden montarse en un equipo
- Es muy común compartir ficheros en red en el esquema conocido como NAS (*Network-Attached Storage*), donde una aplicación permite el acceso a ficheros a distintos clientes.
- NAS utiliza los conocidos protocolos NFS, CIFS, FTP, HTTP, ...
- En muchos casos, los clústeres requieren algún sistema de almacenamiento compartido con más propiedades (acceso a nivel de bloque, concurrencia, tolerancia a fallos o mayor escalabilidad, por ejemplo)

Almacenamiento. SAN (*Storage Area Network*)

- Una red de almacenamiento (SAN) es una red dedicada que ofrece almacenamiento a nivel de bloque
- Una SAN está compuesta por nodos de almacenamiento, redes de alta velocidad (Gigabit Ethernet, 10G Ethernet, Myrinet o Infiniband) y el resto de nodos
- Un dispositivo de bloques puede ser un disco, una partición, un volumen lógico, ...
- Los protocolos iSCSI, Fibre Channel (FC) o ATA over Ethernet (AoE) utilizados en SAN, permiten que los nodos manejen un dispositivo de bloques remoto como un dispositivo de bloques local
- Para que varios nodos puedan montar el mismo dispositivo de bloques, deben utilizar un sistema de ficheros para SAN, como Oracle Cluster FileSystem (OCFS) o Red Hat Global FileSystem (GFS).

Almacenamiento. Sistemas de ficheros distribuidos

- No necesitan utilizar los protocolos iSCSI, FC o AoE porque utilizan su propio protocolo para comunicar cliente y servidor.
- El objetivo de un sistema de ficheros distribuidos es proporcionar almacenamiento remoto a sus clientes de forma transparente, ya que estos lo verán como si se tratase de almacenamiento local.
- Los sistemas de ficheros distribuidos deben poseer tolerancia a fallos, gran escalabilidad, utilización en sistemas heterogéneos y control de concurrencia entre otras propiedades
- Algunos de los sistemas de ficheros distribuidos son:
 - Lustre
 - Ceph
 - Google FileSystem (GFS)
 - GlusterFS
 - Windows DFS

Almacenamiento. DRBD

- *Distributed Replicated Block Device*
- Una opción específicamente desarrollada para clústeres de alta disponibilidad
- Se replica un determinado dispositivo de bloques por una red dedicada
- Se denomina por analogía *Network based RAID1*
- Puede utilizarse en modo síncrono o asíncrono
- Incluido en el kernel linux desde la versión 2.6.33

Resumen

- Los clústeres de alta disponibilidad se están convirtiendo en elementos cada vez más comunes en todo tipo de infraestructuras TIC
- La tendencia en el manejo de los nodos del clúster (y en general del resto de los servidores) es mediante aplicaciones de orquestación
- Es muy importante que el clúster sea escalable (o incluso elástico) para adaptarse a diferentes situaciones
- Un clúster requiere de soluciones específicas de almacenamiento