

Servidores Web de Grandes Empresas

1. Introducción

Generamos cada día datos y más datos que necesitan un montón de espacio para guardarse. Por ello los servidores que los conservan y distribuyen cada vez son mayores y necesitan espacios adecuados para su tamaño y necesidades.

Normalmente los servidores se alojan en los sótanos de las empresas, en zonas poco frecuentadas y en lugares mal acondicionados para las personas.

2. Facebook.

2.1. Infraestructura.

Facebook opera con 9 Centros de Datos (granjas con miles de servidores) en ambas costas de Estados Unidos, la mayoría alquilados excepto el Centro de Datos Oregon. Estos proveedores de infraestructura y terrenos, construyen los Centros de Datos y luego alquilan todo exclusivamente a Facebook.

Normalmente estos edificios tienen la apariencia de una nave industrial y pasan desapercibidos para los viandantes. A no ser que seas Facebook, que entonces lo señalizas convenientemente. Aunque la elección de un desierto no parece la más apropiada para un data center, Facebook lo ubicó aquí para aprovechar la energía solar y con ella suministrar los sistemas de refrigeración.

En cada pasillo del data center de Facebook en Prineville hay una docena de servidores que almacenan los datos de los usuarios. Todos ellos son open source y predomina el color azul porque emplean iluminación LED, que consume menos.

El sistema de refrigeración de este data center de Facebook es diferente al que emplean otras compañías porque no se basa en aire acondicionado, de modo que es más silencioso. Para enfriar los servidores, el edificio cuenta con pasillos fríos y calientes; en los primeros está la parte delantera de los servidores, donde entra el aire frío, y la parte trasera da al pasillo caliente, donde se expulsa el aire cuando se calienta. Después los ventiladores se encargan de expulsarlo al exterior.

Además de los servidores tradicionales, el data center de Facebook también aloja servidores que funcionan con machine learning, que son los que luego permiten identificar rostros en las fotos para que los etiquetes o adivinar tus gustos y sugerirte determinadas publicaciones. Son los llamados servidores Big Sur, con ocho procesadores gráficos de alta potencia conocidos como GPU (fabricado por Nvidia).

El alto consumo energético de los GPU obliga a diseñar una infraestructura diseñada para la refrigeración del edificio. Además de los pasillos fríos y calientes, este data center de Facebook cuenta con salas de aire y de agua. Cuando fuera hace demasiado frío, el edificio está programado para mezclarlo con el aire caliente que desprende las salas de los servidores. En el edificio que se ve en esta fotografía, Facebook almacena datos e información de sus usuarios que han quedado en desuso; por ejemplo, fotografías o vídeos que hace mucho tiempo que no reciben interacciones de ningún tipo. Allí se lleva a cabo el "almacenamiento en frío", pues los servidores que contiene están dormidos, almacenando esa información, hasta que alguien vuelve a recurrirla.

Fuera del edificio principal encontramos también potentes y gigantescos generadores de energía, para suministrar al data center en caso de corte eléctrico y que así los servidores sigan funcionando con total normalidad.

Si volvemos al edificio destinado al "almacenamiento en frío" vemos que allí también hay un espacio destinado a pruebas de laboratorio. Básicamente se trata de varias decenas de bastidores autónomos con dispositivos móviles, donde se van testando las últimas versiones y actualizaciones que luego se implementan en la app de Facebook.

En cada Centro de Datos donde funciona Facebook, generalmente se consume de 2.25 megawatts y 6 megawatts de energía.

2.2. Hardware.

El crecimiento vertiginoso que tuvo Facebook presentó un gran desafío técnico a los ingenieros de la empresa. Estos debían de diseñar toda una nueva arquitectura de hardware que pudiese ser escalable de manera rápida y fácil.

El proyecto Open Compute (Computación Abierta) es una iniciativa lanzada por Facebook con el objetivo de administrar servidores de una manera mucho más eficiente, mejorando de manera sustancial la forma en que se diseñan los centros de datos en las empresas.

De manera general, el proyecto Open Compute plantea los siguientes lineamientos al momento de diseñar un centro de datos eficiente:

- Las placas base (mother board) deben de soportar doble procesador.
- Las fuentes de energía (Power Supply) de los servidores deben de ser de alta eficiencia.
- La placa madre debe de poder adherirse al chasis con sólo un tornillo.
- Los Racks deben de estar diseñados para albergar servidores con dimensiones de 42U
- Cada gabinete debe de tener unas baterías de 48 voltios DC y un sistema integrado de corriente DC/AC

2.3. Software.

Facebook fue desarrollado con software de código abierto, estos servidores corren con una versión modificada de Linux. El sitio esta escrito primordialmente en PHP y sus bases de datos están en MySQL. Para acelerar el sitio, el equipo de ingenieros de Facebook desarrolló un programa llamado HipHop, que se encarga de transformar el código PHP en C++ ganando así sus beneficios de rendimiento.

2.4. Seguridad.

Respecto a seguridad Facebook compró PrivateCore, un acuerdo que fue anunciado el 7 de agosto de 2014. PrivateCore desarrolla software para proteger los datos del servidor a través de certificación de servidor y encriptado de memoria. La tecnología de encriptado de la empresa llena un vacío existente entre el encriptado de datos en movimiento (cifrado de datos en movimiento) y el encriptado de datos en reposo (encriptado de disco, encriptado de cintas) mediante la protección de "datos en uso" (memoria de acceso aleatorio). La tecnología de cifrado de memoria PrivateCore protege contra amenazas a servidores tales como ataques de arranque en frío, amenazas persistentes avanzadas de hardware, rootkits / bootkits, ataques de cadena de suministro de hardware del equipo y amenazas físicas a servidores de parte de personas internas.

Bibliografía de Facebook:

<http://computerhoy.com/noticias/software/mirada-interior-centro-datos-facebook-56238>

<http://tlife.guru/profesional/donde-tiene-servidores-facebook/>

<https://ornitorrincodigital.com/2011/01/06/el-corazn-de-facebook-sus-60000-servidores/>
<http://www.aeromental.com/2011/01/05/los-servidores-de-facebook/>

3. Twitter.

3.1. Infraestructura.

Twitter tiene varios data center repartidos por el mundo. En Bengaluru, donde se encuentra su principal data center que alberga decenas de miles de servidores.

3.2. Hardware.

A medida que escalaba sus mayores cargas de trabajo (Mesos, Hadoop, Manhattan, MySQL) se puso de manifiesto las ofertas de mercado disponibles que no acababa de satisfacer las necesidades. Los componentes tomados fuera de estante vienen con las características de la empresa, al igual que los controladores RAID y fuentes de alimentación de intercambio en caliente. Estos componentes mejoran la fiabilidad a pequeña escala, pero a menudo disminuyen el rendimiento y aumento de los costos; por ejemplo, algunos controladores RAID interfirieron con el rendimiento de los SSD y podrían ser de un tercio del coste del sistema.

Los principales cambios que ha hecho twitter en su hardware a lo largo de la historia son estos:

- 2012 - Los SSD se convierten en el soporte de almacenamiento primario para nuestra MySQL y bases de datos clave / valor.
- 2013 – Su primera solución personalizada para las cargas de trabajo de Hadoop se desarrolla y se convierte en la solución de almacenamiento a granel primaria.
- 2013 - Solución personalizada se desarrolla para los Mesos, TFE, y cargas de trabajo de caché.
- 2014 - La llave de la SSD de encargo / servidor de valor completa de desarrollo.
- 2015 – Se desarrolla base de datos personalizada.
- 2016 – Desarrollo de sistemas de GPU para la inferencia y la formación de modelos de aprendizaje automático.

3.3. Software.

Los servidores de twitter están programados en Scala , un lenguaje multiparadigma e incluye una API abierta para desarrolladores con el fin de integrar Twitter en otras aplicaciones.

3.4. Seguridad.

A nivel físico los datacenter de twitter cuenta con guardias de seguridad las 24 horas del día ,sistemas de vigilancia por circuito cerrado de televisión y en cuanto a mantenimiento y prevención de incendios usan el sistema de climatización airdata y el sistema de extinción basado en NFPA-750.

En cuanto a software se refiere twitter mantiene la información privilegiada encriptada con software externo el cual no detallan

Bibliografía de Twitter:

<https://careers.twitter.com/en/teams/infrastructure-operations.html>

<https://blog.twitter.com/2016/the-infrastructure-behind-twitter-efficiency-and-optimization>

<https://blog.twitter.com/2017/the-infrastructure-behind-twitter-scale>

4. Google.

4.1. Infraestructura.

Google hace sus propios servidores a medida desde hace unos cuantos años y monta sus datacenters. Decidió montárselo por su cuenta y ahora tiene 8 datacenters propios : 6 en Estados Unidos y 2 en Europa. Hay planeado construir 2 más en Asia y otro en Europa.

Los servidores que albergan los centros de datos de Google están concentrados en containers.

Cada container puede contener 1160 servidores. Los containers se apilan de dos en dos y son totalmente independientes.

Se estima que actualmente Google dispondría de una cantidad que rondaría los 2 millones y medio de servidores.

4.2. Hardware.

No llevan carcasa, y llevan el hardware justo para poder trabajar.

Cada uno de estos servidores, que ocupa un par de unidades en un rack, utiliza placas base de Gigabyte, un par de procesadores x86, indistintamente de Intel y AMD, dos discos duros y ocho ranuras de memoria.

Cada servidor usado es de bajo rendimiento que contiene el hardware mínimo necesario para realizar su trabajo. En el 2009 el servidor promedio era doble procesador de doble núcleo con 16GB de RAM y 2 TB de disco duro.

Pero donde más se apartan de lo habitual es al incorporar una batería de 12 voltios.

Google ha optado por incorporar una batería a cada ordenador, en aras de conseguir una mayor eficiencia energética. Según ellos, un SAI consigue entre un 92 y un 95% de eficiencia, mientras su solución obtiene un 99.9% de eficiencia.

4.3. Software.

Funcionan con una distribución de Linux totalmente personalizada por Google

uso de máquinas virtuales, en concreto de una versión personalizada del hipervisor KVM. De hecho, Google asegura que "la mayoría de envíos de código para corrección de errores son para el hipervisor KVM de Linux".

4.4. Seguridad.

Aseguran tener un margen de fallo tan ínfimamente pequeño (un 99,9% de servicio garantizado).

Las medidas de seguridad usadas son: cámaras, detectores de calor, detectores de metales, escáners de iris, control de acceso (el personal de Google que no esté autorizado no puede entrar en las secciones que no le pertenecen).

Diseñan sus propios chips, incluyendo uno de "seguridad por hardware que actualmente se está desplegando tanto en servidores como en periféricos". Estos chips permiten a la empresa identificar de forma segura y autenticar dispositivos de Google a nivel físico.

Dicho chip funciona junto con firmas criptográficas empleadas en "componentes a bajo nivel" como la BIOS, el bootloader, el kernel y una imagen base de un sistema operativo. Las firmas se pueden validar durante cada arranque o actualización, y los componentes los controla Google

directamente. Además, la magnitud de nuestra infraestructura permite a Google simplemente absorber muchos ataques de denegación de servicio. a través de varias capas de hardware y software de equilibrio de carga. Estos equilibradores de carga informan de información sobre el tráfico entrante a un servicio central de DoS que se ejecuta en la infraestructura. Cuando el servicio de centro de Dos detecta que un ataque DoS está llevando a cabo, se puede configurar los equilibradores de carga de la gota o el tráfico regulador asociado con el ataque.

Bibliografía de Google:

<https://cloud.google.com/security/security-design/>

<https://diarioti.com/google-explica-la-seguridad-de-su-infraestructura-que-incluye-hardware-propio/102721>

<http://www.guatemwireless.org/hardware/el-hardware-de-los-servidores-de-google.htm>

l