TECNOLOGIA DE GOOGLE

Aparte del '[PageRank](http://google.dirson.com/pagerank.php)' (el sistema de clasificación de la importancia de cada web), una de las claves del éxito de Google es la tecnología de sus servidores.

Google atiende a más de 5,000 millones de búsquedas cada mes (unas dos mil por segundo), y esta cifra aumenta progresivamente cada día. Para atender todas estas peticiones, y buscar entre más de 3,000 millones de documentos, Google optó por la tecnología Linux.

Disponen de un clúster de cerca de 20,000 servidores repartidos en siete 'data centers'. Estos centros de datos se encuentran situados en diversos puntos del planeta, como [Washington D.C.](http://www-dc.google.com) (USA), [Herndon](http://www-va.google.com) (Virginia, USA), [Santa Clara](http://www-ex.google.com) (California, USA) o [Zurich](http://www-zu.google.com) (Suiza). Cada servidor dispone de un solo procesador Intel, y una memoria RAM que va desde 256 Mb a 1 Gb.

Entre estos 'data centers', Google utiliza su propio gestor de tráfico y su propio software de balanceo de cargas, para dirigir cada petición hacia el mejor servidor

En cada PC se encuentran uno o dos discos duros de 40Gb ó 75Gb, de marca IBM. Google prefiere este sistema distribuido de almacenamiento de datos, antes que uno centralizado. Las razones: es mucho más barato y tiene menos posibilidades de fallo.

En cada una de estas máquinas está instalado Linux RedHat. En mayo de 2002, Google llegó a un [acuerdo con RedHat](http://www.redhat.com/about/presscenter/2000/press_google.html) para que esta empresa le proporcionase el software del Sistema Operativo.

La elección de Linux fue sencilla para Google: el menor ratio coste/rendimiento, corre en simples PCs, y la posibilidad de personalizar cualquier parte del Sistema Operativo. De hecho, no solamente usan Linux en sus servidores, sino también en la mayoría de los PCs de los empleados. Una vez instalado Linux en cada equipo, los técnicos de Google, lo personalizan a su antojo. Ellos lo llaman 'Googlelizar' el Sistema Operativo con sus propias aplicaciones.

La mayoría de las aplicaciones están programadas en C++ (aunque también utilizan Python y PHP), y las herramientas que utilizan son 'gcc', 'gdb' y 'gnats' para el desarrollo y 'p4' para el control de las versiones. Google dispone de más de 100 desarrolladores en su plantilla.

Debido a la multitud de servidores de los que se dispone, el tamaño de estos es muy importante. Por ello, la compañía californiana [Rackable Systems](http://www.rackable.com) compacta para Google varios servidores en un espacio realmente reducido.

Referente a la seguridad, la mayoría de las máquinas de Google no están accesibles desde Internet, ya que se encuentran detrás de un router/firewall que lo impide. Las que sí están accesibles (los servidores web) son auditados periódicamente para encontrar posibles agujeros de seguridad.

Más información: [Google & Linux](http://google.dirson.com/google-linux.php).

FUENTE: http://google.dirson.com/tecnologia.php