

Proyecto: estudio de red de distribución de contenidos

Febrero de 2019

Una empresa dedicada a la distribución de contenidos mediante acceso `http` tiene organizada su red de servidores por grupos según cada una de S áreas geográficas, además de un servidor central:

- En cada área geográfica hay:
 - un grupo de N_1 servidores de acceso directo (con direcciones ANYCAST para favorecer acceso por proximidad geográfica) incorporados en los servidores DNS dando acceso a las direcciones lógicas comunes usadas por los usuarios;
 - un grupo de N_2 servidores de respaldo, que, como el servidor central, no son accesibles directamente a los usuarios mediante DNS, sino que las peticiones son redirigidas a ellos mediante comandos `http redirect` desde los servidores de acceso directo del mismo grupo geográfico, y desde los servidores de respaldo al servidor central con el mismo método.

La redirección de peticiones se llevará a cabo en función de la carga de conexiones en curso en el servidor implicado:

- Si el número de estas es igual a m_1 en un servidor de acceso directo, una nueva petición entrante es redirigida a un servidor de respaldo de la misma área geográfica.
- Si el número de estas es igual a m_2 en un servidor de respaldo, una nueva petición entrante es redirigida al servidor central;
- Si el número de estas es igual a m_3 en el servidor central, la petición es simplemente rechazada con mensaje de aviso de congestión del servicio.

Entre las direcciones ANYCAST y los servidores DNS se consigue que las peticiones sean distribuidas homogéneamente entre los servidores de acceso directo, con las siguientes características de las que recibe cada uno de ellos:

- Llegadas según un proceso de Poisson de 3 peticiones/s.
- Descargas de tamaño medio $15 \cdot 10^6$ octetos.

Se ha decidido fijar N_1 a 2 y N_2 a 1, y evaluar el incremento de áreas geográficas S según se expanda el servicio.

Supondremos que las peticiones de los clientes se atienden a su máxima velocidad de descarga en su acceso a Internet, $20 \cdot 10^6$ bit/s, desde los servidores del grupo geográfico (acceso directo y respaldo), adecuando para ello las capacidades de sus enlaces para acceso a Internet, C_1 y C_2 .

Se pide:

- a) Selecciónese m_1 para que los servidores de la primera capa atiendan algo más del 80 % de las peticiones. Fije C_1 con una holgura del 10 %.
- b) Selecciónese mediante técnica analítica m_2 para que el servidor central reciba algo menos del 5 % de las peticiones.
- c) Afínese la selección de m_2 , de ser posible, con estudio mediante simulación, suponiendo distribución exponencial de la longitud de las demandas. Fije C_2 con una holgura del 10 %.

Fijando m_2 del anterior apartado (valor afinado) y aún suponiendo distribución exponencial de la longitud de las demandas, determine para $S \in \{1, 2, 4, 8, \dots\}$:

- d) El valor de m_3 según análisis para que se rechacen globalmente algo menos del 1 % de peticiones.
- e) El valor de m_3 cumpliendo el requisito anterior según estudio de simulación. Obtenga este resultado hasta al menos $S = 8$.¹
- f) Gráfica representando la evolución de m_3/S del sistema estudiado, así como de la evolución de m_3/S si el tráfico ofrecido al servidor central fuese de Poisson. ¿Cabe esperar algún comportamiento asintótico en ambas?

Las simulaciones computarán intervalos de confianza del 95 % y como mucho ancho del ± 10 % (más estrecho cuando sea necesario o se considere que las simulaciones son suficientemente rápidas).

¹La simulación de una topología perfectamente simétrica —todos los tráficos ven la misma probabilidad de bloqueo— permite utilizar la opción “-A” en el simulador: se estima únicamente la probabilidad de bloqueo conjunta, con ganancia en velocidad que puede llegar a varios órdenes de magnitud.