

Ingeniería del caos



Raúl UO

Taso Rodriguez Meana UO251434

¿Problemas de los sistemas distribuidos:

Los sistemas distribuidos ganan cada vez más peso a medida que el tráfico en internet fue creciendo a lo largo de los años y resultaba inviable tener un solo gran servidor para sostener los servicios que se querían ofrecer a través de la nube. Estos sistemas traen sin duda muchas ventajas, como la posibilidad de escalar con mucha facilidad a medida que el uso de estos servicios iba creciendo, y también pueden llegar a ofrecer una buena resiliencia al fallo, pues es más raro que fallen a la vez 2000 servidores que 1 solo. No obstante, con casi todas las nuevas tecnologías vienen las desventajas, los sistemas distribuidos, debido a su escala y a las interacciones que ocurren entre ellos, son intrínsecamente más complejos, y la posibilidad de predecir y evitar cualquier fallo (cosa ya casi imposible con un solo servidor) se convierte en una tarea posible solo en la imaginación.

Para poder construir sistemas distribuidos muy seguros y resistentes al fallo, se ha de preparar nuestros servicios para que sean capaces de seguir funcionando, aunque ocurra un fallo en parte de los servidores de nuestro sistema. Para conseguirlo, seguiremos un enfoque empírico que estudie este caos inherente a los sistemas distribuidos, para preparar los mismos ante el fallo y asegurar que siguen funcionando correctamente tras el mismo, aumentando su resistencia. Llamamos a esto ingeniería del caos.

“La Ingeniería del Caos es la disciplina de experimentar en un Sistema, con la finalidad de generar confianza en la capacidad del Sistema para soportar condiciones turbulentas en producción.” -Principios de la ingeniería del caos

El Caos en la práctica:

La base de la ingeniería del caos es la realización de experimentos con el objetivo de descubrir las debilidades que tiene un sistema en líneas generales estos experimentos siguen estos cuatro pasos:

- 1- La Definición del estado inicial de nuestro sistema, este es el estado en el que está funcionando, y debería de incorporar algún tipo de métrica que nos permita valorar si el sistema está funcionando correctamente o no. Esta métrica puede ser algo como el número de peticiones por segundo, el conjunto de peticiones respondidas, el tráfico en nuestro servicio... En general se ha de intentar buscar una métrica con efecto real en el sistema y perceptible por el usuario final, antes que métricas que tengan más efecto en el estado interno del mismo.

- 2- Hipotetizar sobre el futuro comportamiento del sistema, como evolucionara la métrica previamente definida a lo largo del tiempo, tanto en el grupo de control, un sistema que funciona correctamente, como en el experimental, en el que introduciremos los fallos.
- 3- Introducir los fallos en el sistema experimental que nos permitan comparar los dos grupos que estamos siguiendo, la idea es buscar fallos (o eventos que pongan en peligro el sistema como un aumento del tráfico) que reflejen los fallos más comunes/con más impacto. Cuantos más eventos posibles se consideren y se prueben más seguros estaremos de que las conclusiones sacadas del experimento y del posterior endurecimiento mejoraran la resiliencia de nuestro sistema
- 4- Estudiar el sistema experimental para buscar diferencias de funcionamiento con el sistema de control, cuanto más difícil es modificar la métrica en el grupo experimental, más resistente es el sistema a los posibles fallos que puedan suceder. Si se encuentran debilidades, hay que intentar mejorar el sistema antes de que esos comportamientos se manifiesten en el sistema real.

Otras consideraciones:

Ejecutar los experimentos en producción: la ingeniería del caos busca replicar lo mas cerca posible una experiencia real, por tanto, es recomendable ejecutar parte de los experimentos en producción, para asegurar que las consecuencias que sacamos son lo más reales posibles.

Minimizar el radio de explosión: si se realizan los experimentos en producción existe el riesgo de causar daño real al producto/cliente, es responsabilidad del ingeniero del caos de asegurarse que este daño es mínimo y nunca mayor de lo necesario.

Automatizar experimentos: al igual que el resto de las pruebas, los experimentos del caos deberían automatizarse para que se puedan ejecutar continuamente.

Ejemplo de Netflix:

El equipo encargado de Netflix tiene una sección dedicada a la ingeniería del caos. Originalmente empezaron con un servicio llamado “chaos monkey” que semi-aleatoriamente mataba servidores para probar que el sistema podía aguantar una caída. A medida que fueron ahondando en esta metodología acabaron aumentando la escala, creando “chaos Kong”, que mataba una región entera de ASW. Usando otras regiones para “salvar” a la región caída eran capaces de mantener el servicio en producción de Netflix sin que la experiencia del usuario apenas se viene comprometida.

En general el equipo de Netflix destaca lo útil que les ha resultado la ingeniería del caos a la hora de mantener un servicio consistente con un tráfico de cientos de millones de usuarios en todo el mundo

