Taller de Sistemas Empresariales

Análisis de performance y escalamiento

Grupo 16

Federico Bentancor

Facundo Laborde

Gonzalo Santa María

Federico Sierra

Michael Rodriguez

Contents

1. Introducción

Este documento pretende mostrar los resultados de las pruebas de performance sobre el sistema vacunas.uy, un sistema de gestión crítico del cual debemos identificar su punto de quiebre, así como verificar que los tiempos de respuesta no aumenten conforme aumenta la cantidad de request enviados para la prueba.

2. Herramientas

Para las pruebas usaremos las siguientes herramientas:

- JMeter 5.4.1
- Java 11
- Monitor del sistema app.elasticloud.uy

3. Servicios monitoreados

Identificamos 2 servicios críticos del sistema en momentos pico. Estos son:

- Monitor del sistema (http://node1340-vacunasuyg16.web.elasticloud.uy/monitor).
- Consulta de agendas (https://node1340-vacunasuyg16.web.elasticloud.uy/schedule).

Según estimaciones realizadas por el equipo, es necesario que el sistema sea capaz de responder a 3000 request HTTP por minuto. La realidad contemplada es ante casos de notificación pública de apertura de nuevas agendas y ante consultas masivas del monitor de vacunación.

4. Condiciones iniciales y pruebas

Se realizan 5 pruebas diferentes para poder determinar el punto de quiebre del sistema, iniciando con 1, 4, 100, 150 y 200 request HTTP por segundo.

Notamos que los tiempos de respuesta son similares, pero a partir de las 150 Rps el servidor de aplicaciones comienza a devolver errores HTTP 503 (Service unavailable).

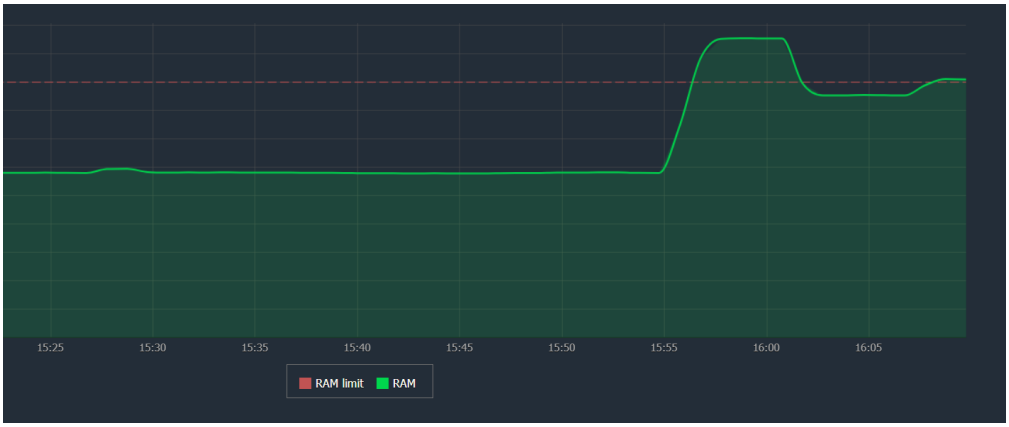
Requests Per Second	ResponseTime	%ERROR
1	279ms	0%
4	200	0%
100	216	0%
150	259	0%
200	1724	21%

5. Punto de quiebre del sistema

Encontramos el punto de quiebre del sistema sobre los 200 request por segundo, valor que comienza a devolver un alto porcentaje de peticiones fallidas (HTTP code 503).

6. Cuello de botella

Podemos determinar que el cuello de botella de nuestra aplicación se encuentra en la memoria RAM de las instancias Wildfly del servicio Cloud Antel Minube.
Según podemos ver en la gráfica, a los 200 request por segundo el sistema tiene picos de memoria RAM, por encima del límite



7. Tiempos de respuesta

Si bien los tiempos de respuesta aumentaron en las pruebas con 0% de errores, consideramos que son tiempos despreciables que no impactan en la experiencia de usuario ni en ningún tiempo de procesamiento a nivel de backend.

Podemos determinar entonces que no tenemos degradación de servicio en las condiciones de prueba ideales (previo al punto de quiebre del sistema)

8. Conclusiones

Podemos afirmar que el sistema se comporta dentro de los parámetros comprometidos por los interesados del sistema.

No se detectan problemas graves de performance y no existe degradación considerable de los tiempos de respuesta de los servicios a los cuales se le realizaron las pruebas de estrés.

9. Acciones preventivas.

Para poder superar este punto de quiebra se resuelve crear un plan de escalamiento horizontal con las herramientas dispuestas por el servicio Cloud del proveedor Antel, Minube.

Debido a que el sistema no guarda ningún tipo de estado a nivel de servidor de aplicaciones, podemos crear un plan de escalamiento horizontal de las instancias de Wildfly 22 que nos permitirá sobrellevar picos de carga que estén por encima de lo máximo esperado y comprometido.