

Concepto de Rendimiento - Autorizador de Pagos

1. Resumen del Análisis

Se realizaron pruebas de performance sobre el microservicio **Autorizador de Pagos**, para evaluar su desempeño después del cambio “*un nuevo método de pago aceptado por el autorizador el cual permite realizarlo con el subAccountId, el cual hace referencia a el mismo CDUSUARIO/WALLETID de la cuenta en cuestión*” bajo diferentes escenarios de pruebas (carga y estrés, tolerancia a fallos con escalado automático mediante HPA).

Los resultados obtenidos reflejan lo siguiente:

APDEX - Satisfacción del Usuario

- **Carga:** APDEX promedio de **0.54** en carga y estrés, categorizado como **Condicionado**.
- **Transacciones Críticas:**
 - *autorizacionConSubAccountId*: APDEX **0.98** → **Satisfactorio**.
 - *autorización*: APDEX **0.52 - 0.57**.

Tiempo de Respuesta

- **Tiempos máximos esperados:** 900ms.
- **Valores observados:**
 - *preAutorizacionPlacasNoFlypass*: **2605 ms** (muy por encima del umbral).
 - *autorización (Carga 05)*: **2595 ms**.
 - *autorización (Estrés 03)*: **21927 ms**.

Infraestructura y Comportamiento del HPA

- **Uso de memoria:**
 - Se observó que la memoria no se libera tras la carga, lo que llevó a la creación de nuevos pods.
 - Consumo de memoria estable en **51-52%**, pero acumulativo tras múltiples pruebas.
- **Uso de CPU:**
 - Bajo carga: **7.98 - 10.5%** (normal).
 - Bajo estrés: **21.5%** (alta utilización, pero sin saturación).
- **Réplicas:**
 - Se generaron **nuevas réplicas debido a la acumulación de memoria**, sin que el uso de CPU fuera un factor clave.

Hallazgos

1. Bajo carga normal, el rendimiento es aceptable, pero con áreas críticas que requieren optimización.
2. Bajo estrés, los tiempos de respuesta en algunas transacciones que superan los 20 segundos.
3. El HPA está escalando por memoria y no por CPU, lo que sugiere posible fuga de memoria o procesos en segundo plano que no se liberan.
4. Durante la ejecución de la prueba, se eliminaron dos pods intencionalmente. Posteriormente, el sistema desplegó dos nuevas réplicas para reemplazarlos; sin embargo, las réplicas existentes no se apagaron, lo que resultó en una cantidad de pods mayor a la esperada.
5. Las transacciones de autorización funcionan correctamente en términos de volumen, pero algunas tienen tiempos excesivos.
6. Algunos endpoints no están optimizados para alta concurrencia, lo que genera latencias elevadas y potenciales bloqueos.

Recomendaciones

A. Optimización del Código

- **Diagnóstico y corrección de procesos en segundo plano:**
 - Identificar procesos que continúan ejecutándose después de la carga.
 - Asegurar que todos los hilos y conexiones a bases de datos sean cerrados adecuadamente.
 - Monitorear Garbage Collector para validar si está limpiando la memoria correctamente.
- **Reducción de tiempos de respuesta:**
 - Revisar la transacción *preAutorizacionPlacasNoFlypass*, ya que su tiempo de respuesta es **2605 ms**, más de **2.5 veces el umbral aceptable**.
 - Implementar estrategias de caching o paralelización para reducir los tiempos en consultas de autorización.

B. Ajustes en la Infraestructura

- **Optimización del HPA:**

- Verificar la configuración de umbrales para que el escalado se active antes de llegar a niveles críticos de memoria.
- Evaluar si se deben configurar otra alerta de auto escalado relacionada con la carga transaccional .
- Analizar si se requiere más RAM por pod en lugar de escalar horizontalmente.

C. Pruebas de Performance Continuas: Siempre realizar pruebas de performance a los cambios relacionados con funcionalidades o nuevos componentes de arquitectura

Conclusión

El sistema tiene un rendimiento **Condicionado**, con áreas críticas a mejorar:

- Se recomienda **optimizar los endpoints con tiempos de respuesta elevados.**
- Se debe **ajustar la configuración del HPA** para también impulsar el auto escalado por carga transaccional
- **Es fundamental investigar los procesos en segundo plano** que impiden la liberación de memoria y causan escalado innecesario.

Con las optimizaciones propuestas, se espera que el sistema pueda mejorar el APDEX a niveles satisfactorios y garantizar estabilidad bajo carga y estrés.