Daniel Montezuma Sevillano – A00382231

Diego Fernando Mueses Zúñiga – A00382021

ATRIBUTOS DE CALIDAD

**Latency**: El análisis de latencia en nuestro sistema se llevó a cabo mediante un enfoque estadístico utilizando un conjunto de 9 muestras de tiempos de respuesta en milisegundos para solicitudes individuales. Los resultados de esta evaluación se presentan a continuación:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Num | 1 | 107374182 | 214748364 | 1498 | 654321 | 9999999 | 1234321 | 777889 | 123198 |
| ms | 9 | 0 | 1 | 0 | 7 | 1 | 1 | 1 | 0 |

El análisis revela una variedad de tiempos de respuesta, que oscilan entre 0 y 9 milisegundos. Los valores más bajos indican un rendimiento óptimo, mientras que los valores más altos sugieren cierta variabilidad en la latencia. En conclusión, el sistema demuestra un buen rendimiento en términos de latencia en la mayoría de las solicitudes, con tiempos de respuesta generalmente bajos.

**Throughput**: En este contexto, hemos realizado una prueba de rendimiento en la que se enviaron un total de 1,000,000 de solicitudes al servidor. Cada solicitud representaba un número único dentro del rango de 1 a 1,000,000, con el objetivo de obtener los números primos correspondientes. Durante esta evaluación, se registró un tiempo total de respuesta de 103,900 ms (milisegundos).

El hecho de que todas las solicitudes se hayan atendido correctamente en un tiempo total de 103,900 ms es un indicador de que el sistema demostró una capacidad eficiente para manejar una carga significativa en un período relativamente corto. Para cuantificar el throughput en términos comprensibles, podemos calcularlo de la siguiente manera:

Throughput = (Número de solicitudes) / (Tiempo de respuesta total)

Throughput = 1,000,000 / 103,900 ≈ 9.62 solicitudes por milisegundo

Para una interpretación más convencional, podemos convertir el throughput a solicitudes por segundo:

Throughput = 9.62 solicitudes por milisegundo \* 1000 (para convertir a solicitudes por segundo) ≈ 9620.48 solicitudes por segundo

**Deadline**: En nuestro sistema, no se establece un "deadline" específico de antemano. Esto se debe a que es el cliente quien tiene el control sobre cuándo finalizar la conexión y detener el envió de solicitudes hacia el servidor. En lugar de imponer un límite de tiempo predefinido desde el servidor, se le otorga al cliente la libertad de cerrar la conexión en el momento que considere oportuno.

**Varianza de latencia**: Tomando los datos de “Latency”, podemos analizar la variación de esta según los siguientes pasos:

1. Calcular la media:

Se suma todos los datos y se divide entre el número de muestras:

Media = (9 + 0 + 1 + 0 + 7 + 1 + 1 + 1 + 0) / 9 = 2.222...

1. Calcular la Diferencia Cuadrada para Cada Muestra:
2. (9 - 2.222...)^2 = 45,941284
3. (0 - 2.222...)^2 = 4,937284
4. (1 - 2.222...)^2 = 1,493284
5. (0 - 2.222...)^2 = 4,937284
6. (7 - 2.222...)^2 = 22,829284
7. (1 - 2.222...)^2 = 1,493284
8. (1 - 2.222...)^2 = 1,493284
9. (1 - 2.222...)^2 = 1,493284
10. (0 - 2.222...)^2 = 4,937284
11. Calcular la varianza

Varianza = (45.94 + 4.93 + 1.49 + 4.93 + 22.82 + 1.49 + 1.49 + 1.49 + 4.93)/9 = 9,945…

La anterior varianza nos indica que algunas de las solicitudes están siendo procesadas mucho más rápido o mucho más lento que la media general. Es posible que el sistema este sujeto a posibles fluctuaciones en el rendimiento del sistema.

**Miss Rate:** En nuestro caso, no observamos ningún caso de Miss Rate, ya que todas las solicitudes que enviamos efectivamente llegan y son procesadas por el sistema sin contratiempos.

**Process Rate:** En este escenario, cada solicitud que el cliente envió fue exitosamente procesada por el servidor, lo que significa que el servidor proporcionó una respuesta a cada una de estas solicitudes. Por lo tanto, no hubo ninguna solicitud que no se haya respondido.

**Query and Reporting Times**: En nuestro caso, donde el tiempo requerido para que el cliente se conectara al servidor fue de 15 segundos, esto implica que hubo una demora de 15 segundos antes de que la comunicación entre el cliente y el servidor se estableciera completamente. Esta demora puede deberse a diferentes aspectos, como configuraciones iniciales del servidor o la necesidad de cargar recursos al iniciar.