

Lectura #2 Solución

Kevin Córdoba Chevez - 2020100920

1. ¿En qué consisten los datos timeseries?

En la lectura se menciona que times series consiste en guardar los cambios sobre el tiempo. También se dice que es una secuencia de medidas ordenadas en el tiempo.

2. ¿Qué son métricas?

Según lo mencionado en la lectura, las métricas son medidas numéricas y estas pueden ayudar a entender por qué una aplicación funciona de cierta manera.

3. ¿En qué consiste la Observabilidad?

Según lo mencionado en la lectura, la observabilidad ayuda a entender el estado de un sistema de modo que los problemas puedan ser identificados rápidamente y sean abordados proactivamente. Se basa en 3 pilares los cuales son: métricas, logs y traces.

4. Explique el concepto de dimensiones en datos timeseries

Según lo mencionado en la lectura, con los datos time series frecuentemente los datos contienen más de una serie y es un conjunto de múltiples time series.

Se ejecuta una sola consulta para una medición con una o más propiedades adicionales conocidas como dimensiones.

Y con esto lo que se conoce como dimensiones en los datos timeseries son propiedades adicionales que poseen este tipo de datos.

5. ¿Por qué los tags en métricas permiten generar mejores gráficos en Grafana?

Con base a lo mencionado en la lectura, a Grafana le benefician los tags en métricas debido a que Grafana se beneficia de las métricas que se graban repetidamente sobre el tiempo para hacer comparaciones, y los tags son una de las características que posean las bases de datos time series, estas pueden añadir información de contexto como por ejemplo en dónde fue tomada la medición. Debido a la gran cantidad de contexto que los tags ofrecen, grafana puede crear mejores métricas ya que tiene más datos para realizar las comparaciones.

6. Suponiendo que se están recolectando datos IoT (Internet of Things) de miles de dispositivos, los mismos generan una métrica cada 15 segundos con el consumo de energía y temperatura, explique:

• ¿Por qué una base de datos relacional no es una buena opción para almacenar esta información?

Porque tomando en cuenta la lectura, se menciona que aunque una base de datos relacional regular en efecto es capaz de almacenar los datos time series, no es la opción más eficiente, una base de datos time series, abreviada TSDB permite mejores opciones de optimización. Existen datos que cambian muy poco con el tiempo, especialmente tomando en cuenta que en este caso se generan métricas cada 15 segundos, muchos de los datos generados tendrán valores idénticos, por lo que almacenarlos sería un mal uso del espacio en disco, al utilizar un TSDB se puede almacenar cada marca de tiempo subsecuente en forma de tiempo, por ejemplo +10, en lugar de datos repetidos.

- **Dada la naturaleza de datos timeseries, ¿de qué forma la localidad puede ayudarnos a ahorrar dinero?**

Basado en lo que comprendí de los datos time series en la lectura, considero que en un caso como este donde se recolectan datos del IoT de miles de dispositivos cada 15 segundos, y pensando en el colector que se debe instalar en la máquina a monitorear, considero que la localidad puede ahorrar dinero especialmente al ahorrarse costo de almacenamiento en línea, también en la lectura se menciona que las TSDB puede almacenar más eficientemente datos cuidando el espacio en disco por lo que no habría mucho problema con espacio local. Otro punto sería que si no fuera local habría que invertir en costos de internet para tener mayor velocidad en el momento de compartir datos. En general creo que se ahorraría dinero al depender en menor medida de los servicios externos.