

Planeación de recursos

La respuesta corta a " *¿qué recursos y configuración necesitaré para llevar a Zeebe a producción?* "Es: depende.

Si bien no podemos decirle exactamente lo que necesita, más allá de *lo que depende* , podemos explicar qué depende, de qué depende y cómo depende de ello.

Espacio del disco

Todos los Brokers en una partición usan espacio en disco para almacenar:

- El registro de eventos para cada partición en la que participan. Por defecto, esto es un mínimo de 512 MB para cada partición, incrementándose en segmentos de 512 MB. El registro de eventos se trunca en un agente dado cuando los datos han sido procesados y exportados con éxito por todos los exportadores cargados.
- Una o más instantáneas periódicas del estado de ejecución (datos en vuelo) de cada partición (sin límites, en función del trabajo en vuelo). El número de instantáneas para retener, y su frecuencia es configurable. De forma predeterminada, se establece en tres instantáneas retenidas y una instantánea [cada 15 minutos](#) .

Además, el líder de una partición también usa espacio en disco para almacenar:

- Una proyección del estado de ejecución de la partición en RocksDB. (ilimitado, basado en el trabajo en vuelo)

Por defecto, estos datos se almacenan en

- **segments** - los datos del registro se dividen en segmentos. El registro solo se agrega, hasta que se trunca al alcanzar el recuento máximo de instantáneas.
- **state** - El estado activo. Flujos de trabajo desplegados, instancias de flujo de trabajo activas, etc. Se eliminan las instancias o trabajos de flujo de trabajo completados.
- **snapshot** - los datos de un estado en un determinado momento

Si desea una receta para explotar el uso de espacio en disco, aquí hay algunas maneras de hacerlo:

- Cree una gran cantidad de instantáneas con un largo período entre ellas.
- Cargue un exportador, como el exportador de depuración, que no avanza su posición de registro.

Registro de eventos

El registro de eventos para cada partición está segmentado. Por defecto, el tamaño del segmento es de 512 MB. Esto se puede cambiar en el archivo `zeebe.cfg.toml` con la configuración `logSegmentSize`.

Un segmento de registro de eventos puede eliminarse una vez que todos los eventos que contiene hayan sido procesados por los exportadores, replicados a otros intermediarios y procesados. Tres cosas pueden hacer que el registro de eventos no se trunque:

- Un clúster pierde su quórum, en cuyo caso los eventos se ponen en cola pero no se procesan.
- Un exportador no avanza su posición de lectura en el registro de eventos.
- El número máximo de instantáneas no se ha escrito.

Un segmento de registro de eventos no se elimina hasta que todos los eventos en él hayan sido exportados por todos los exportadores configurados. Esto significa que los exportadores que confían en los efectos secundarios, realizan cálculos intensivos o experimentan contrapresión del almacenamiento externo harán que el uso del disco aumente, ya que retrasan la eliminación de segmentos de registro de eventos.

La exportación solo se realiza en el líder de la partición, pero los seguidores de la partición no eliminan segmentos en su réplica de la partición hasta que el líder marca todos los eventos en ella como innecesarios por los exportadores.

No se eliminan segmentos de registro de eventos hasta que se alcanza el número máximo de instantáneas. Cuando se alcanza el número máximo de instantáneas, el registro de eventos se trunca hasta la instantánea más antigua.

Instantáneas

El estado de ejecución de la partición se captura periódicamente en el líder en una instantánea. Por defecto, este período es `cada 15 minutos`. Esto se puede cambiar en el archivo `zeebe.cfg.toml`. El número de instantáneas válidas para retener se configura en `zeebe.cfg.toml`. Por defecto, el líder y los seguidores retienen solo la última instantánea válida.

Una instantánea es una proyección de todos los eventos que representan el estado actual de los flujos de trabajo que se ejecutan en la partición. Contiene todos los datos activos, por ejemplo, flujos de trabajo implementados, instancias de flujo de trabajo activo y trabajos aún no completados.

Cuando el intermediario tiene tantas instantáneas como lo configuró el parámetro `maxSnapshots`, elimina todos los datos del registro que se escribieron antes de la instantánea más antigua.

RocksDB

En el intermediario principal de una partición, el estado de ejecución actual se mantiene en la memoria y en el disco en RocksDB. En nuestros experimentos, esto crece a 2 GB bajo una gran carga de procesos de larga duración. Las instantáneas que se replican a los seguidores son instantáneas de RocksDB.

Efecto de los exportadores y fallas del sistema externo

Si un sistema externo en el que confía un exportador falla, por ejemplo, si está exportando datos a Elasticsearch y falla la conexión al clúster Elasticsearch, el exportador no avanzará su posición en el registro de eventos y los intermediarios no podrán truncar sus registros. El registro de eventos del agente crecerá hasta que el exportador pueda restablecer la conexión y exportar los datos. Para asegurarse de que sus agentes sean resistentes en caso de falla del sistema externo, debes suficiente espacio en disco para continuar operando sin truncar el registro de eventos hasta que se restablezca la conexión al sistema externo.

Efecto en los exportadores de falla de nodo

Solo el líder de una partición exporta eventos. Solo los eventos confirmados (eventos que se han replicado) se pasan a los exportadores. El exportador actualizará su posición de lectura. La posición de lectura del exportador solo se replica entre los intermediarios en la instantánea. No está escrito en el registro de eventos. Esto significa que *la posición actual de un exportador no se puede reconstruir desde el registro de eventos replicado, solo desde una instantánea*.

Cuando una partición falla a un nuevo líder, el nuevo líder puede construir el estado actual de la partición proyectando el registro de eventos desde el punto de la última instantánea. La posición de los exportadores no se puede reconstruir desde el registro de eventos, por lo que se establece en la última instantánea. Esto significa que un exportador puede ver los mismos eventos dos veces en caso de falla.

Debe asignar identificadores idempotentes a eventos en su exportador si esto es un problema para su sistema. La combinación de la posición del registro y la identificación de la partición es confiable como una identificación única para un evento.

Efecto de la pérdida de quórum.

Si una partición cae en quórum (por ejemplo: si dos nodos en un clúster de tres nodos caen), el líder de la partición continuará aceptando solicitudes, pero estas solicitudes no se replicarán y no se marcarán como confirmadas. En este caso, no se pueden truncar. Esto hace que el registro de eventos crezca. La cantidad de espacio en disco necesaria para continuar operando en este escenario es una función del rendimiento del intermediario y la

cantidad de tiempo para restaurar el quórum. Debe asegurarse de que sus nodos tengan suficiente espacio en disco para manejar este modo de falla.