Adriel Araya Vargas 2019312845 Instituto Tecnológico de Costa Rica Resumen 8 Bases de Datos 2 IC-4302

Bigtable: Un sistema de almacenamiento distribuido para datos estructurados.

Modelo de Datos

Bigtable es un mapa multidimensional ordenado, distribuido y persistente, que está esparsamente poblado. El mapa se indexa mediante una clave de fila, una clave de columna y una marca de tiempo; cada valor en el mapa es una matriz de bytes sin interpretar.

Filas

Las claves de fila en una tabla son cadenas arbitrarias que pueden tener un tamaño de hasta 64KB. Cada operación de lectura o escritura bajo una misma clave de fila es atómica, lo cual facilita el manejo de actualizaciones concurrentes en la misma fila. Bigtable organiza los datos en orden lexicográfico por clave de fila y las filas se dividen dinámicamente en rangos llamados "tablets" para distribución y equilibrio de carga. Esto permite que las lecturas de rangos cortos de filas sean eficientes y requieran la comunicación con un número reducido de máquinas.

Columnas

Las columnas se agrupan en conjuntos llamados familias de columnas, que son la unidad básica de control de acceso. Cada familia de columnas almacena datos del mismo tipo y se comprimen juntos. Antes de almacenar datos en una familia de columnas, esta debe ser creada, y luego se pueden utilizar múltiples claves de columna dentro de esa familia. Se recomienda tener un número pequeño de familias de columnas en una tabla, mientras que el número de columnas puede ser ilimitado. Cada clave de columna se compone de una familia y un calificador, siendo los nombres de las familias de columnas imprimibles y los calificadores pueden ser cadenas arbitrarias.

Marcas de tiempo

Cada celda puede contener múltiples versiones de los mismos datos, indexadas por marcas de tiempo. Las marcas de tiempo pueden ser asignadas automáticamente por Bigtable o por las aplicaciones cliente. Las versiones de una celda se almacenan en orden descendente de marcas de tiempo para facilitar la lectura de las versiones más recientes.

API

La API de Bigtable proporciona funciones para crear y eliminar tablas y familias de columnas. También ofrece funciones para modificar metadatos de clústeres, tablas y familias de columnas, como los derechos de control de acceso. Las aplicaciones cliente pueden escribir o eliminar valores en Bigtable, buscar valores en filas individuales o iterar sobre un subconjunto de los datos de una tabla. Se muestra código en C++ que utiliza una abstracción de RowMutation para realizar una serie de actualizaciones. También se muestra código en C++ que utiliza una abstracción de Scanner para iterar

sobre todos los anclajes en una fila específica. Bigtable admite transacciones a nivel de una sola fila, contadores enteros en las celdas y la ejecución de scripts proporcionados por el cliente en los servidores. Además, se menciona que Bigtable se puede utilizar con MapReduce, un marco para ejecutar cálculos en paralelo a gran escala.

Bloques de construcción.

Big Table se apoya en la infraestructura de Google File System (GFS) para almacenar archivos y registros. Utiliza un clúster compartido de máquinas que ejecutan diversas aplicaciones distribuidas. Bigtable depende de un sistema de gestión de clústeres para programar trabajos, administrar recursos y manejar fallas. Utiliza el formato de archivo interno SSTable y se apoya en Chubby, un servicio de bloqueo distribuido, para garantizar la consistencia y gestionar tareas como el control de acceso y el almacenamiento del esquema. Aunque la indisponibilidad de Chubby puede afectar la disponibilidad de Bigtable, en pruebas recientes se encontró que el impacto fue mínimo, con un porcentaje promedio muy bajo de horas de no disponibilidad.

Implementación

Bigtable tiene tres componentes principales: una biblioteca para los clientes, un servidor maestro y múltiples servidores de tabletas. Los servidores de tabletas pueden agregarse o eliminarse según sea necesario. El servidor maestro se encarga de asignar y equilibrar las tabletas, además de gestionar los cambios de esquema. Cada servidor de tabletas administra varias tabletas y se encarga de las solicitudes de lectura y escritura, así como de dividir las tabletas grandes. Los clientes se comunican directamente con los servidores de tabletas, lo que alivia la carga del servidor maestro. Las tablas de Bigtable son conjuntos de tabletas, que se dividen automáticamente a medida que crecen.

El sistema Bigtable utiliza una jerarquía de tres niveles para almacenar la información de ubicación de las tabletas. Hay un archivo en Chubby que contiene la ubicación de la tableta raíz, que a su vez contiene la ubicación de todas las tabletas en la tabla METADATA. Cada tableta METADATA almacena la ubicación de un conjunto de tabletas de usuario. La tabla METADATA también almacena otros datos secundarios, como un registro de eventos relacionados con cada tableta. La biblioteca del cliente almacena en caché las ubicaciones de las tabletas para mejorar el rendimiento. Si la información en caché está desactualizada, el cliente realiza una serie de consultas para obtener la ubicación correcta. Esto se logra a través de una combinación de lecturas desde Chubby y consultas recursivas en la jerarquía de ubicación de las tabletas. Cada tableta se asigna a un servidor de tabletas a la vez. El maestro supervisa los servidores de tabletas activos y el estado de asignación de las tabletas. Cuando una tableta no tiene asignado un servidor y hay un servidor disponible con suficiente espacio, se le asigna la tableta enviando una solicitud de carga al servidor de tabletas. El maestro utiliza Chubby para rastrear los servidores de tabletas. Los cambios en las tabletas, como la creación, eliminación, fusión y división, son gestionados por el maestro. El maestro registra los cambios y realiza las asignaciones correspondientes.

Bigtable también utiliza refinamientos como grupos de localidad, compresión, caché, filtros de Bloom, registro de transacciones e inmutabilidad para mejorar el rendimiento y la confiabilidad del sistema. Entre sus aplicaciones en el mundo real se encuentra Google Analytics, Google Earth y búsquedas personalizadas de Google.