

RocksDB - Leveled Compaction

Trong RocksDB, Leveled Compaction là một cơ chế quan trọng để quản lý và tổ chức dữ liệu trong cơ sở dữ liệu. Nó được sử dụng để tối ưu hóa việc đọc và ghi dữ liệu bằng cách chia dữ liệu thành các cấp độ (levels) và thực hiện compaction (tái cấu trúc) dữ liệu giữa các cấp độ.

Một cấp độ (level) trong Leveled Compaction bao gồm một tập hợp các tệp SSTable (Sorted String Table) được sắp xếp theo thứ tự khóa. Các cấp độ có kích thước khác nhau và được sắp xếp theo mức độ độc lập với nhau. Cấp độ cao hơn chứa các tệp SSTable lớn hơn và cung cấp tốc độ đọc tốt hơn, trong khi cấp độ thấp hơn có kích thước nhỏ hơn và cung cấp tốc độ ghi tốt hơn.

Khi dữ liệu được ghi vào RocksDB, nó sẽ được ghi vào Memtable trước. Khi Memtable trở nên quá lớn, RocksDB sẽ chuyển đổi dữ liệu từ Memtable vào một cấp độ (level) thấp nhất. Khi một cấp độ (level) trở nên quá lớn, RocksDB sẽ thực hiện quá trình compaction để tái cấu trúc dữ liệu và giảm kích thước của cấp độ đó.

Quá trình compaction trong Leveled Compaction bao gồm hai giai đoạn chính: minor compaction và major compaction.

- **Minor compaction:** Trong minor compaction, RocksDB sẽ chọn một số tệp SSTable từ cấp độ thấp hơn và hợp nhất chúng thành một tệp SSTable mới ở cấp độ cao hơn. Quá trình này giúp giảm số lượng tệp SSTable và tăng tốc độ đọc dữ liệu. Minor compaction thường được thực hiện thường xuyên để duy trì cân bằng kích thước của các cấp độ.
- **Major compaction:** Trong major compaction, RocksDB sẽ hợp nhất tất cả các tệp SSTable từ một cấp độ thành một tệp SSTable duy nhất ở cấp độ cao nhất. Quá trình này giúp giảm tải công việc đọc và tăng tốc độ truy vấn dữ liệu. Major compaction thường được thực hiện ít thường xuyên hơn và yêu cầu nhiều tài nguyên hơn so với minor compaction.

Leveled Compaction giúp tăng hiệu suất đọc và ghi dữ liệu trong RocksDB bằng cách tối ưu hóa cấu trúc dữ liệu. Nó cũng giúp giảm kích thước của dữ liệu và sử dụng bộ nhớ hiệu quả. Tuy nhiên, việc lựa chọn loại Memtable phù hợp cũng là một yếu tố quan trọng, phụ thuộc vào loại dữ liệu, kích thước bộ nhớ và yêu cầu của ứng dụng cụ thể.