**Chỉ mục (index) trong mysql**

**1. Khái niệm về chỉ mục (index) trong mysql**

* Chỉ mục (Index) là bảng tra cứu đặc biệt mà Database Search Engine có thể sử dụng để tăng nhanh thời gian và hiệu suất thu thập dữ liệu. Hiểu đơn giản, một chỉ mục là một con trỏ tới dữ liệu trong một bảng. Một chỉ mục trong một Database là tương tự như một chỉ mục trong Mục lục của cuốn sách.
* Ví dụ, nếu bạn muốn tham chiếu tất cả các trang trong một cuốn sách về một chủ đề nào đó, đầu tiên bạn nghĩ ngay đến mục lục của nó, mà liệt kê tất cả các chương, chủ đề theo thứ tự và sau đó được tham chiếu tới một hoặc nhiều trang cụ thể. Khi đã có mục lục của cuốn sách, việc tìm kiếm trang sách đó sẽ nhanh chóng và đỡ tốn công hơn rất nhiều so với việc bạn phải lật từng trang.

**2. Khi nào cần đánh chỉ mục cho bảng**

* Bất cứ khi nào bạn thay đổi cấu trúc bảng dữ liệu bạn đều cần Add Index lại, giống như khi bạn thay đổi nội dung quyển sách, bạn cần phải làm lại mục lục. Vậy hãy cân nhắc, nếu Database của bạn sử dụng INSERT hay UPDATE nhiều hơn là SELECT thì Index chỉ làm chậm thêm mà thôi.
* Lưu ý: Index giống như các tab. Chúng làm cho nó nhanh hơn để bạn có thể tìm thấy những dữ liệu bạn cần phù hợp với một số tiêu chí nhất định. Để tạo chỉ mục tốt bạn cần phải suy nghĩ về cách bạn sẽ được truy vấn dữ liệu như thế nào. Khi bạn có thông tin này, bạn sẽ có thể tạo index và nó sẽ tăng tốc độ truy vấn của bạn ngay lập tức.

**3. Các loại index trong mysql**

Ở các post trước, các tác giả đã nói cụ thể về cú pháp sử dụng index nên trong post này, tôi xin nói thêm về các loại index trong mysql

MySQL cung cấp 3 kiểu index khác nhau cho data đó là B-Tree, Hash và R-Tree index. Hiểu rõ về các kiểu index này sẽ giúp chúng ta tạo ra các loại index trong bảng một cách hiệu quả hơn. Do R-Tree được sử dụng cho các loại dữ liệu không gian spacial data và thường ít khi gặp phải nên bài viết này chúng ta sẽ tập chung so sánh 2 loại index là B-Tree và Hash:

1. B-Tree Index

B-Tree index có các đặc điểm sau:

* Dữ liệu index trong B-Tree được tổ chức và lưu trữ theo dạng cây(tree), tức là có root, branch, leaf. Giá trị của các node được tổ chức tăng dần từ trái qua phải.
* Khi truy vấn dữ liệu thì việc tìm kiếm trong B-Tree là 1 quá trình đệ quy, bắt đầu từ root node và tìm kiếm tới branch và leaf, đến khi tìm được tất cả dữ liệu - thỏa mãn với điều kiện truy vấn thì mới dùng lại.
* B-Tree index được sử dụng cho những column trong bảng khi muốn tìm kiếm 1 giá trị nằm trong khoảng nào đó. Ví dụ: tìm kiếm những sinh viên có điểm Toán từ 5-9.

1. Hash Index

Hash index có các đặc điểm sau:

* Hash index có một vài đặc điểm khác biệt so với Btree index.
* Dữ liệu index được tổ chức theo dạng Key - Value được liên kết với nhau.
* Khác với B-Tree, thì Hash index chỉ nên sử dụng trong các biểu thức toán tử là = và <>. Không sử dụng cho toán từ tìm kiếm 1 khoảng giá trị như > hay < .
* Không thể tối ưu hóa toán tử ORDER BY bằng việc sử dụng Hash index bởi vì nó không thể tìm kiếm được phần từ tiếp theo trong Order.
* Toàn bộ nội dung của Key được sử dụng để tìm kiếm giá trị records, khác với B-Tree một phần của node cũng có thể được sử dụng để tìm kiếm.
* Hash có tốc độ nhanh hơn kiểu Btree.

=> Việc chọn index theo kiểu B-Tree hay Hash ngoài yếu tố về mục đích sử dụng index thì nó còn phụ thuộc vào việc Storage Engine có hỗ trợ loại index đó hay không. Ví dụ MyISAM, InnoDB hay Archive chỉ hỗ trợ B-Tree, trong khi MyISAM lại hỗ trợ cho cả 2.

**4. Cách sử dụng index**

* Tạo index trên một bảng:

CREATE UNIQUE INDEX index\_name

ON table\_name;

Unique index được sử dụng không chỉ để tăng hiệu suất, mà còn cho mục đích toàn vẹn dữ liệu. Một Unique index không cho phép bất kỳ bản sao giá trị nào được chèn vào trong bảng.

* Tạo index cho cột được chỉ rõ trong bảng:

CREATE UNIQUE INDEX index\_name

ON table\_name(column 1, column 2,...);

Lưu ý: Trong trường hợp không chỉ rõ tên column được đánh index trong bảng, mysql sẽ mặc định đánh index trên primary key hoặc các ràng buộc mang tính duy nhất trong bảng databases.

* Các lệnh thêm index:

1. Thêm index cho bảng, với index mang tính duy nhất và không thể null

ALTER TABLE tbl\_name ADD PRIMARY KEY (olumn 1, column 2,..)

1. Thêm index cho bảng, với index mang tính duy nhất nhưng có thể null

ALTER TABLE tbl\_name ADD UNIQUE index\_name (column 1, column 2,..)

1. Thêm index cho bảng, có thể trùng lặp

ALTER TABLE tbl\_name ADD INDEX index\_name (column 1, column 2,..)

1. Thêm index đặc biệt thường cho tìm kiếm văn bản

ALTER TABLE tbl\_name ADD FULLTEXT index\_name (column 1, column 2,..)

* Các lệnh xóa index:

Một chỉ mục có thể bị xóa bởi sử dụng lệnh DROP INDEX trong SQL. Bạn nên cẩn thận trong khi xóa một chỉ mục, bởi vì khi đó hiệu suất có thể chậm hơn hoặc không được cải thiện.

Cú pháp cơ bản của lệnh DROP INDEX:

DROP INDEX index\_name;

**5. Một số lưu ý khi sử dụng chỉ mục**

Mặc dù các chỉ mục nhằm mục đích để nâng cao hiệu suất của Database, nhưng đôi khi, bạn nên tránh dùng chúng. Dưới đây là một số điểm bạn cần xem xét để quyết định có nên sử dụng chỉ mục hay không:

* Các chỉ mục không nên được sử dụng trong các bảng nhỏ.
* Bảng mà thường xuyên có các hoạt động update, insert.
* Các chỉ mục không nên được sử dụng trên các cột mà chứa một số lượng lớn giá trị NULL.
* Không nên dùng chỉ mục trên các cột mà thường xuyên bị sửa đổi.