

Quản trị cơ sở dữ liệu và Tối ưu hiệu năng

Bài tập

Bài tập 1

- Bảng nhân viên trên ổ đĩa 1, bảng học sinh trên ổ đĩa 2, và file log trên ổ đĩa 2. Bảng học sinh nhỏ hơn so với bảng nhân viên, nhưng được truy cập thường xuyên hơn. Ổ đĩa 2 hỗ trợ nhiều hơn gấp 2 lần tốc độ vào ra (I/O rate) của ổ đĩa 1. Khách hàng sẵn sàng mua một ổ cứng mới.
- Câu hỏi: Bạn nên làm gì với nó?

Bài tập 1

- Giải pháp: Có thể điều tốt nhất để làm là đặt file log vào một đĩa thứ 3. File log sẽ hoạt động tốt hơn trong trường hợp đó (ví dụ: như một công cụ lưu trữ tuần tự), và hệ thống có thể chịu được thất bại của ổ đĩa CSDL nếu cũng có một kết suất CSDL (database dump) trên băng (tape)

Bài tập 2

- Thời gian đáp ứng khá biến đổi. Biết rằng các bảng mới được thêm vào CSDL đồng thời với các giao dịch trực tuyến. Những bảng mới đó không liên quan đến các giao dịch trực tuyến.
- Câu hỏi: Bạn có thể làm gì để tăng thời gian đáp ứng?

Bài tập 2

- Giải pháp: Thêm những bảng mới đó bên ngoài cửa sổ trực tuyến (online window). Lý do là bất kì lệnh DDL nào có xu hướng gây trở ngại với các giao dịch trực tuyến phải truy cập vào thông tin mô tả bảng.

Bài tập 3

- Một thẻ tín dụng mới đưa ra các dòng lớn của tín dụng với các lãi suất thấp.
- Thiết lập một dịch vụ người dùng thẻ mới bao gồm các 3 bước sau:
 - Có được một số khách hàng mới từ một bộ đếm chung (global counter)
 - Hỏi khách hàng các thông tin cụ thể, như : thu nhập, địa chỉ mail.
 - Thêm khách hàng vào bảng khách hàng
- Câu hỏi: Tốc độ giao dịch không hỗ trợ băng thông thêm lớn (large insert traffic).

Hãy giải thích những lý do có thể và cung cấp giải pháp tối ưu tương tranh.

Bài tập 3

- Giải pháp: Cuộc phỏng vấn với khách hàng không nên diễn ra trong khi giữ khóa trên bộ đếm số khách hàng. Thay vào đó, bước 2 nên xuất hiện đầu tiên bên ngoài bối cảnh giao dịch. Vì vậy bước 1 và 3 nên được kết hợp thành một giao dịch duy nhất. Nếu có được số khách hàng gây ra tranh chấp khóa (lock contention), khi đó lấy nó càng muộn càng tốt, hoặc sử dụng cơ chế bộ đếm đặc biệt (special counter facility), ví dụ : các nối tiếp (sequences) trong Oracle hoặc các thuộc tính đồng nhất (identity attributes) trong SQL Server.

Bài tập 4

- Bộ phận kinh doanh thông minh (The business intelligence department) muốn chạy các truy vấn khai phá dữ liệu trên dữ liệu bán hàng. Khi họ chạy các truy vấn, Các cập nhật có thể xuất hiện bất kì. Điều đó làm chậm cả việc cập nhật và truy vấn khai phá dữ liệu.
- Câu hỏi: Những thay đổi nào bạn có thể làm?

Bài tập 4

- Giải pháp: Chia theo thời gian hoặc không gian. Hoặc chạy các truy vấn vào buổi tối hoặc chạy chúng trên một hệ thống CSDL riêng rẽ.
- Bạn không muốn nói cho mọi người bằng lương sẽ có trong ngày cuối cùng, vì ai đó trong bộ phận tiếp thị (marketing) quyết định xem nếu có một mối quan hệ giữa 25 phòng ban...

Bài tập 5

- Một ứng dụng hỗ trợ hàng nghìn hoạt động chèn trên một giờ và hàng trăm nghìn các giao dịch chỉ cập nhật ngắn trên một giờ. Các hoạt động cập nhật được đóng gói như các giao dịch lớn trong mỗi 20 phút hoặc lâu hơn và trong 5 phút cuối cùng. Khi việc chèn được đẩy vào, thời gian phản hồi cập nhật giảm xuống còn 15 đến 30 giây, bế tắc (deadlock) xuất hiện, và một trong số các ổ đĩa trong tình trạng sử dụng rất cao.
- Câu hỏi: Những thay đổi bạn có thể làm để giúp đỡ tình trạng này?

Bài tập 5

- Giải pháp: Các hoạt động chèn xuất hiện làm độc quyền hệ thống, và dữ liệu có vẻ bị phân chia kém. 2 thay đổi trong sự kết hợp hoặc tách biệt có thể giúp.
 - Làm thoáng băng thông chèn (insert traffic) bằng cách chia các giao dịch chèn lớn thành các giao dịch nhỏ (Nếu điều đó là có thể thì tính đúng đắn đồng thời có liên quan) và bắt đầu chúng tại một thời điểm
 - Phân chia dữ liệu để băng thông chèn được dàn trải đến các ổ đĩa khác nhau, nhưng các cập nhật vẫn truy cập nhanh. Một chỉ số phân nhóm (clustering index) trên một key không tuần tự (non-sequential key) sẽ hoạt động tốt

Bài tập 6

- Hệ thống bị chậm do việc sử dụng bộ xử lý quá mức. Một giao dịch quan trọng (liên quan) thực hiện một truy vấn SQL truy cập một bản ghi duy nhất bên trong một vòng lặp của ngôn ngữ lập trình.
- Câu hỏi: Những thay đổi nào bạn có thể thực hiện để cải thiện tình trạng này.

Bài tập 6

- Giải pháp: Thay thế vòng lặp bằng một truy vấn duy nhất truy cập vào các bản ghi và cho phép mã ngôn ngữ lập trình lặp lại một mảng. Điều đó sẽ tiết kiệm thời gian vì ngôn ngữ lập trình tương tác với CSLD có chi phí đắt. Trong khi truy cập mảng chi phí ít hơn.

Bài tập 7

- Một công ty thẻ tín dụng mới được thành lập. Công ty đó thành toán cho khách hàng của họ vào ngày thứ 5 cuối cùng của mỗi tháng. Giao dịch thanh toán phải mất cả đêm, vì vậy hàng loạt các công việc cần thiết không thể hoàn thành trong thời gian đó.
- Câu hỏi: Bạn nên làm gì để cải thiện tình trạng này?

Bài tập 7

- Giải pháp: Câu hỏi đầu tiên để hỏi liệu có hay không ứng dụng phải hoạt động theo cách này. Nếu $1/20$ của hoạt động thanh toán có thể được gửi vào mỗi ngày làm việc. Khi đó ứng dụng thanh toán sẽ tạo các yêu cầu ít hơn trên hệ thống vào mỗi ngày (Phân chia theo thời gian). Các cách tiếp cận khác là chạy việc thanh toán như một loạt các việc lớn nhưng chỉ trong ngày cuối tuần.