

ebook

8 lưu ý quan trọng khi sao lưu & phục hồi SQL Server

Microsoft®
SQL Server®

Backup Academy Team
zBackup | Dedicated for Your Backup

Trong bất kỳ hệ thống IT nào có sử dụng SQL Server, SQL Server luôn được xem là thành phần quan trọng nhất. Không có gì khó hiểu bởi khi đó hầu như tất cả dữ liệu sản xuất – kinh doanh quan trọng của công ty đều được lưu trong các database SQL Server. Với tầm quan trọng đó, công tác sao lưu và đảm bảo an toàn dữ liệu SQL Server là nhiệm vụ hàng đầu của người IT. Nếu database gặp sự cố và không thể phục hồi, cái giá phải trả là thiệt hại nặng nề về tài chính và uy tín của công ty. Cùng với đó có thể là công việc của chính người IT.

Ebook ngắn này tập hợp 8 vấn đề quan trọng bạn cần lưu ý trong quá trình quản lý công tác sao lưu và phục hồi SQL Server. Đọc kỹ và áp dụng những lưu ý này sẽ giúp bạn nâng cao đáng kể tính an toàn của hệ thống SQL Server. Cũng đồng nghĩa đảm bảo tính an toàn cho hoạt động của công ty và cho công việc của chính bạn.



Microsoft®
SQL Server®

Lưu Log file và Data file trên các ổ đĩa/đĩa cứng khác nhau

Khi bạn tạo mới một database, SQL Server thường mặc định tạo Data file .MDF và Log file .LDF trong cùng thư mục (VD: Với SQL Server 2008 là C:\Program Files\Microsoft SQL Server\MS-SQL10\MSSQLSERVER\MSSQL\DATA). Thông thường, chúng ta cứ để thiết lập mặc định này mà không thay đổi. Chính điều này gây ảnh hưởng đến tốc độ truy xuất và khả năng phục hồi database.

Nếu có thể, bạn nên lưu trữ Data file và Log file trên đĩa cứng (hoặc ít ra là ổ đĩa) khác nhau vì 2 lý do sau:

- **Tốc độ truy xuất**

Trong khi Data file được truy xuất theo cơ chế ngẫu nhiên (random) thì Log file lại được truy xuất tuần tự (serial). Vì thế, việc đặt 2 loại dữ liệu này trên các đĩa cứng khác nhau sẽ giúp tối ưu tốc độ truy xuất database.

- **Sao lưu Tail log**

Có thể bạn chưa biết: Nếu Data file bị lỗi nhưng Log file vẫn truy xuất được thì bạn vẫn có thể sao lưu Log file bằng lệnh BACKUP LOG với tùy chọn NO_TRUNCATE (gọi là sao lưu Tail log). Khi phục hồi, kết hợp bản sao lưu Tail log này với các bản sao lưu Full/Differential và Transaction log trước đó, bạn có thể khôi phục database trở về thời điểm ngay trước khi sự cố xảy ra (chứ không chỉ phục hồi về thời điểm ở bản sao lưu Transaction log gần nhất).

Tóm lại, việc lưu trữ Log file ở ổ đĩa/đĩa cứng khác với Data file mang đến cho bạn cơ hội phục hồi dữ liệu mới phát sinh sau lần sao lưu Transaction log gần nhất. Điều này giúp bạn phục hồi nguyên vẹn SQL Server, không mất mát bất kỳ dữ liệu nào.

Lựa chọn đúng Recovery model

SQL Server có 3 giá trị Recovery model: SIMPLE, BULK-LOGGED và FULL. Thao tác điều chỉnh giá trị này là hoàn toàn đơn giản. Nhưng việc lựa chọn sử dụng giá trị nào thì phải cân nhắc kỹ dựa trên yêu cầu của công ty đối với hệ thống SQL Server (Service Level Agreement - SLA). Xét về kỹ thuật, Recovery model ảnh hưởng đến cách bạn sao lưu và khả năng phục hồi dữ liệu. Xét trên khía cạnh hoạt động của công ty, Recovery model ảnh hưởng đến lượng dữ liệu mà công ty chấp nhận mất mát khi sự cố xảy ra.

Nếu công ty bạn chấp nhận mất mát lượng dữ liệu tương đương 24 giờ (*) thì chính sách sao lưu Full mỗi ngày một lần là đủ (thường tiến hành vào buổi tối). Nếu là 12 giờ thì bạn có thể sao lưu Full 2 lần/ngày vào buổi tối và giờ nghỉ trưa, hoặc kết hợp sao lưu Full vào buổi tối và Differential vào giờ nghỉ trưa. Với yêu cầu về lượng dữ liệu mất mát và cách sao lưu như vậy, chọn Recovery model có giá trị SIMPLE là phù hợp.

() Lượng dữ liệu tối đa chấp nhận mất mát khi sự cố xảy ra được gọi là Recovery Point Objective – RPO. Giá trị này thường được định lượng bằng đơn vị giờ.*

Lưu ý:

Với Recovery model BULK-LOGGED, khi có những thao tác xử lý dữ liệu hàng loạt như BULK INSERT, CREATE INDEX, SELECT INTO,... thì log của từng thao tác không được ghi vào file Transaction log. Do đó, bạn không thể tiến hành phục hồi Point-in-Time.

Nếu dữ liệu SQL Server là cực kỳ quan trọng và công ty chỉ chấp nhận mất mát lượng dữ liệu tối đa là 1 giờ, bạn cần sao lưu liên tục theo tần suất 1 giờ/lần. Trên lý thuyết, yêu cầu này có thể đáp ứng bằng cách sao lưu Full hoặc Differential cứ 1 giờ/lần. Tuy nhiên, điều này là không thực tế vì thời gian sao lưu Full (hay thậm chí Differential) có thể lâu hơn 1 giờ, khiến bạn không thể đáp ứng RPO đặt ra. Bên cạnh đó, cách làm này khiến SQL Server hầu như luôn sao lưu 24/24, gây ảnh hưởng nhiều đến hiệu năng và truy xuất của người dùng. Ở tình huống này, bạn cần kết hợp sao lưu Full, Differential và Transaction log. Cụ thể, bạn có thể sao lưu Full vào mỗi tối, Differential vào giờ nghỉ trưa, và Transaction log cứ 1 giờ/lần. Lượng dữ liệu Transaction log phát sinh trong 1 giờ là rất ít so với dữ liệu database.

Với cách sao lưu Transaction log này, Recovery model SIMPLE không thể đáp ứng được. Bởi với SIMPLE, các record trong Transaction log được cắt bỏ (Log truncation) ngay khi xuất hiện CHECKPOINT. Điều này khiến các Log record không liên tiếp nhau. Vì thế, SQL Server không cho phép bạn sao lưu Transaction log khi Recovery model là SIMPLE. Chỉ với Recovery model là FULL hoặc BULK-LOGGED, bạn mới có thể sao lưu Transaction log.

Tóm lại, nếu muốn sao lưu Transaction log, bạn cần chọn Recovery model là FULL hoặc BULK-LOGGED.

Sao lưu Tail log giúp phục hồi về thời điểm ngay trước khi sự cố xảy ra

Tail log là phần Transaction log chưa được sao lưu kể từ lần sao lưu Transaction log gần nhất. Nói rõ hơn đây là tất cả các Log phát sinh từ lần sao lưu Transaction log gần nhất đến thời điểm hiện tại. Khi Data file bị lỗi nhưng Log file vẫn sử dụng được, việc bạn cần tiến hành ngay là sao lưu phần Tail log này. Nhờ đó, khi tiến hành phục hồi, bạn có thể kết hợp bản sao lưu Tail log này cùng các bản sao lưu Full/Differential và Transaction log trước đó để khôi phục database trở về thời điểm ngay trước khi sự cố xảy ra.

Bạn có thể sao lưu Transaction log ngay cả khi Data file bị lỗi. Lệnh BACKUP LOG với tùy chọn NO_TRUNCATE sẽ giúp bạn thực hiện việc này:

```
BACKUP LOG ERP  
TO DISK = 'C:\SQLBackupData\TAIL_LOG_ERP.bak'  
WITH NO_TRUNCATE  
GO
```

Lưu ý:

Khả năng sao lưu Tail log chính là lý do bạn nên lưu trữ Log file trên đĩa cứng khác với Data file. Nếu có lỗi gì xảy ra với đĩa cứng chứa Data file, Log file của bạn vẫn an toàn và có thể truy xuất để sao lưu.

Kết hợp sao lưu Full, Differential và Transaction log

Sử dụng bản sao lưu Full là cách đơn giản nhất để phục hồi database. Tuy nhiên, sao lưu Full lại chiếm dung lượng lớn và mất nhiều thời gian (gây ảnh hưởng đến hiệu năng SQL Server). Điều này khiến bạn khó có thể sao lưu theo mật độ hàng giờ. Do đó, cần kết hợp sao lưu Full với Differential và Transaction log.

Sao lưu theo cơ chế Differential sẽ sao lưu những dữ liệu thay đổi trong database kể từ lần sao lưu Full gần nhất. Trong khi đó, sao lưu Transaction log sẽ sao lưu các Log thay đổi từ lần sao lưu Transaction log ngay trước đó.

Khi phục hồi, sử dụng kết hợp Full, Differential và Transaction log cho phép bạn khôi phục database trở về bất kỳ thời điểm nào (Point-in-Time Recovery). Đặc biệt, khi sử dụng thêm bản sao lưu Tail log (Transaction log hiện tại của database), bạn có thể khôi phục database trở về thời điểm ngay trước khi mất dữ liệu.

Lưu ý:

Bạn chỉ có thể sao lưu Transaction log khi Recovery model là FULL hoặc BULK-LOGGED.

Việc sử dụng kết hợp bản sao lưu Full, Differential và Transaction log mang lại cho bạn khả năng phục hồi với lượng dữ liệu mất mát là ít nhất. Trong khi đó, thời gian sao lưu và dung lượng lưu trữ cũng giảm rất nhiều khi so với sao lưu Full thường xuyên.

Chính sách tối ưu nhất là sao lưu Full vào mỗi tối, sao lưu Differential vào giờ nghỉ trưa, và sao lưu Transaction log theo chu kỳ 1 giờ/lần (hoặc thậm chí sao lưu cứ 15-30 phút/lần).

Luôn lưu trữ offsite cách xa

Một nguyên tắc cơ bản (và quan trọng) trong sao lưu là dữ liệu sao lưu phải được lưu trữ cách xa dữ liệu gốc. Có thể bạn thắc mắc là với những lỗi như hư ổ cứng, virus, xóa nhầm thì chỉ cần lưu trữ dữ liệu sao lưu tách biệt là đã đủ; đâu cần phải lưu trữ cách xa. Đúng là như vậy. Nhưng vấn đề sao lưu là câu chuyện về sự cố, và bạn không thể biết được khi nào và sự cố gì sẽ xảy ra. Do đó, lưu trữ offsite là cách tốt nhất để đảm bảo 100% khả năng phục hồi dù sự cố là gì.

Với SQL Server, rất nhiều người quản lý IT mắc lỗi trong vấn đề sao lưu offsite này. Thông thường, dữ liệu được sao lưu thành file .bak và lưu trữ tại chính máy chủ SQL Server mà không sao chép sang thiết bị khác. Tình huống tốt hơn là sao lưu vào Tape/External HDD/USB nhưng vẫn đặt gần hoặc cùng vị trí với máy chủ. Với những sự cố như cháy nổ, trộm cắp, phá hoại,... nhiều khả năng bạn mất cả dữ liệu gốc lẫn dữ liệu sao lưu.

Nếu không đủ nhân lực và thời gian để lưu trữ offsite hàng ngày, ít nhất bạn cũng nên đảm bảo điều này hàng tuần để tránh tình trạng “mất cả chì lẫn chài” – đặc biệt với dữ liệu quan trọng như SQL Server.

Xây dựng DR Plan và duy trì testing định kỳ

Sao lưu là vô nghĩa nếu bạn không thể phục hồi. Vậy điều gì giúp bạn chắc chắn có thể phục hồi khi bị mất dữ liệu? Sẽ là rủi ro lớn nếu bạn chờ đến lúc sự cố thật sự xảy ra mới có thể trả lời câu hỏi này. Bởi khi đó có thể đã quá muộn, bạn không còn cơ hội để sửa chữa sai lầm.

Giải pháp duy nhất cho vấn đề này là bạn cần testing quá trình phục hồi định kỳ và thường xuyên. Đồng thời, xây dựng được tài liệu DR Plan hướng dẫn cách phục hồi cho từng sự cố có thể xảy ra. Việc làm này nhằm 4 mục đích:

- Bạn kiểm tra được dữ liệu sao lưu SQL Server trong Tape/Disk/Cloud không có vấn đề gì và có thể phục hồi.
- Nắm bắt thuần thục từng tình huống phục hồi (phục hồi từ bản sao lưu Full, Differential, Transaction log, Tai log).
- Kiểm tra thời gian tiến hành phục hồi từng tình huống lỗi. Từ đó, có thể cam kết với công ty về thời gian phục hồi SQL Server khi sự cố xảy ra.
- Tài liệu hóa thành DR Plan với chi tiết từng bước tiến hành và phân công nhân sự cụ thể. Nhờ đó, khi mất dữ liệu, bạn biết được ai phải làm gì để tiến hành khôi phục.

Bạn không cần hiểu DR Testing là việc gì đó phức tạp khiến bạn không thể tiến hành. Nếu không thể testing hàng tháng, bạn có thể testing hàng quý. Nếu không thể testing cho toàn bộ database trong SQL Server, bạn có thể testing với 1-2 database ngẫu nhiên nào đó. Hoặc chỉ cần sử dụng những lệnh RESTORE VERIFYONLY để kiểm tra dữ liệu (cách sử dụng lệnh này được trình bày ở mục tiếp theo).

Kiểm tra bản sao lưu bằng lệnh RESTORE VERIFYONLY

Để kiểm tra dữ liệu sao lưu, cách tốt nhất là tiến hành khôi phục database. Tuy nhiên, nếu dung lượng database quá lớn hoặc lo ngại quá trình phục hồi thực tế sẽ ảnh hưởng đến SQL Server đang chạy, bạn có thể sử dụng lệnh RESTORE VERIFYONLY để tiến hành việc kiểm tra này.

Lệnh RESTORE VERIFYONLY đọc dữ liệu trong file để kiểm tra xem file có bị lỗi gì không, có thể phục hồi được hay không. Lệnh này hoạt động giống như một lệnh phục hồi thực sự. Điểm khác biệt duy nhất là dữ liệu chỉ được kiểm tra chứ không được phục hồi thật sự. Bạn chạy lệnh này như sau:

```
RESTORE VERIFYONLY  
FROM DISK = 'C:\SQLRestoreData\DATABASE_ERP.bak'  
GO
```

Nếu bản sao lưu không có lỗi và có thể phục hồi, SQL Server sẽ trả về kết quả như bên dưới:

```
The backup set on file 1 is valid
```

Nếu bản sao lưu bị lỗi, SQL Server sẽ thông báo lỗi như bên dưới:

```
Msg 3242, Level 16, State 2, Line 1
```

```
The file on device 'C:\SQLRestoreData\DATABASE_ERP.bak'  
is not a valid Microsoft Tape Format backup set.
```

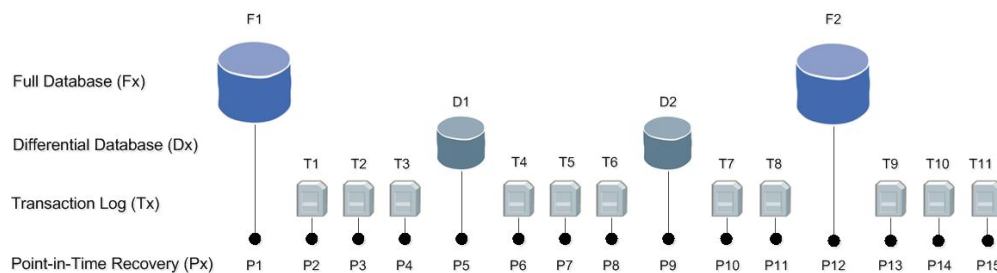
```
Msg 3013, Level 16, State 1, Line 1
```

```
VERIFY DATABASE is terminating abnormally.
```

Vì phải đọc qua toàn bộ file để kiểm tra dữ liệu, lệnh này sẽ chạy tương đối lâu. Tuy nhiên, đây vẫn là việc nên tiến hành thường xuyên bởi nó giúp bạn chắc chắn dữ liệu sao lưu không bị lỗi. Nếu phát hiện có lỗi, việc bạn cần làm là lập tức tiến hành sao lưu database.

Phục hồi Point-in-time theo nhiều cách kết hợp khác nhau

Khi có các bản sao lưu Full, Differential và Transaction log, để phục hồi database trở về một thời điểm P nào đó, bạn có thể kết hợp các bản sao lưu này theo nhiều cách khác nhau. Tính năng này của SQL Server mang đến cho bạn khả năng phục hồi khi một bản sao lưu nào đó bị mất hoặc lỗi. Tham khảo ví dụ bên dưới để nắm rõ hơn về cách phục hồi này.



Trong minh họa trên, có 3 loại dữ liệu sao lưu được ký hiệu như sau:

- Full Database: Ký hiệu Fx
- Differential Database: Ký hiệu Dx
- Transaction Log: Ký hiệu Tx

Ví dụ 1 – Phục hồi về thời điểm P10

Để phục hồi về thời điểm P10, bạn có thể sử dụng 3 cách kết hợp:

- Cách 1: F1 > D2 > T7
- Cách 2: F1 > D1 > T4 > T5 > T6 > T7
- Cách 3: F1 > T1 > T2 > T3 > T4 > T5 > T6 > T7

Trong 3 cách trên, phục hồi theo cách 1 là nhanh nhất. Nhưng nếu bản sao lưu D2 có lỗi thì bạn không thể sử dụng cách này. Khi đó, bạn có thể phục hồi theo cách 2 hoặc cách 3.

Ví dụ 2 – Phục hồi về thời điểm P14

Để phục hồi về thời điểm P14, bạn có thể sử dụng 3 cách kết hợp:

- Cách 1: F2 > T9 > T10
- Cách 2: F1 > D2 > T7 > T8 > T9 > T10
- Cách 3: F1 > D1 > T4 > T5 > T6 > T7 > T8 > T9 > T10
- Cách 4: F1 > T1 > T2 > T3 > T4 > T5 > T6 > T7 > T8 > T9 > T10

Cách 1 là nhanh nhất. Nhưng nếu F2 bị mất hoặc lỗi, bạn vẫn có thể phục hồi từ F1 theo 3 cách còn lại.

Kiến thức:

Bạn có thể phục hồi database bằng cách kết hợp Full/Differential và Transaction log theo những cách như vậy là do bản sao lưu Full và Differential không làm ảnh hưởng đến tính liên tục của bản sao lưu Transaction log. Do đó, nếu không thể bắt đầu từ bản sao lưu Full/Differential liền kề trước, bạn có thể sử dụng bản sao lưu Full/Differential trước đó nữa. Miễn là các bản sao lưu Transaction log kế tiếp đó phải liên tục cho đến thời điểm bạn muốn phục hồi.

About Us

Về zBackup

zBackup là dịch vụ sao lưu dữ liệu theo mô hình Điện toán đám mây, mang đến cho các doanh nghiệp một giải pháp sao lưu toàn diện, bảo mật và hiệu năng. Dữ liệu được mã hóa AES, sao lưu tự động và lưu trữ offsite trên hệ thống zBackup đặt tại các Data Center tiêu chuẩn: Viettel, VDC và SuperData. Hỗ trợ hầu hết các ứng dụng phổ biến trong doanh nghiệp, zBackup giúp bạn dễ dàng sao lưu dữ liệu File Server, Active Directory, SQL Server, Exchange Server, MySQL, Oracle, VMware, Hyper-V,...

Được cung cấp theo mô hình dịch vụ, zBackup giúp doanh nghiệp dễ dàng tiếp cận và sử dụng mà không tốn chi phí đầu tư thiết bị, phần mềm, không bận tâm công tác vận hành, bảo trì, nâng cấp. Hiện zBackup đang giúp hơn 200 doanh nghiệp trên cả nước bảo vệ những dữ liệu kinh doanh quan trọng.

Về Backup Academy Team

Backup Academy Team là đội ngũ các chuyên viên Backup Recovery đang công tác tại zBackup. Các chuyên viên này tập trung nghiên cứu các vấn đề Backup Restore, Disaster Recovery của các dữ liệu, ứng dụng thường được sử dụng trong doanh nghiệp. Từ đó, xây dựng thành các tài liệu hữu ích dưới dạng Article, White paper, Ebook, Presentation,... nhằm chia sẻ đến cộng đồng IT thông qua trang <http://backupacademy.zbackup.vn>. Mục tiêu của Backup Academy Team là giúp bộ phận IT trong các doanh nghiệp dễ dàng nắm bắt kiến thức Backup Recovery để bảo vệ an toàn các dữ liệu quan trọng trong các hệ thống server.

Liên hệ với chúng tôi

Bạn có bất kỳ thắc mắc hay cần hỗ trợ, vui lòng liên hệ chúng tôi:

- zBackup Team: <http://www.zbackup.vn> | support@zbackup.vn | Hotline 09760-05430

- Backup Academy: <http://backupacademy.zbackup.vn> | ba@zbackup.vn