

Administração de Bases de Dados

College Scorecard | Tema 7 (Fase 4)

por

Eduardo Romão nº35477,

João Dias nº35476

6 de Maio de 2018

Tabela de Conteúdos

1	Introdução	1
2	Definição de momento de nível 0	2
3	Totais ou Incrementais	3
4	Número de backups em cada nível 0 ou 1	4
5	Definição do armazenamento físico	5
6	Distribuição física dos discos	6
7	Conclusão	7

Introdução

Atualmente vivemos numa era de tecnologia, onde cada vez mais o ser Humano depende dela. Numa base de dados existem diversos riscos como a perda de dados ou até mesmos ataques externos no qual resultam alterações ou roubo de dados. Como tal é necessário a garantia de conseguir reconstruir os dados sem por a empresa em risco.

Para evitar tais situações de risco, é necessário a utilização de backups, que consistem na copia dos dados para um componente diferente ou não.

Neste trabalho iremos estudar o procedimento de fazer um backup para a base de dados com o tema escolhido, College Scorecard, de modo a garantir informação persistente e a não perda de dados.



Definição de momento de nível 0

De forma a definir a melhor estratégia para salvaguardarmos a nossa base de dados é necessário definir o momento de nível zero, que corresponde ao momento em que é feito um backup completo.

Como a base de dados a desenvolver contem dados relativos a estatísticas de escolas, instituições e estudantes a melhor estratégia será fazer um backup de nível zero uma vez por ano.

Fazendo um backup total uma vez por ano a melhor altura irá ser no final do ano letivo ou seja dia 1 de Agosto. Por esta altura a base de dados terá a maior parte dos dados atualizados, relativos a numero de estudantes, classificações escolares ou até mesmo de propinas.

Deste modo, salvaguardamos os principais dados relativos ao ano letivo em questão, descartando apenas algumas atualizações que tenham sido feitas a meio do ano escolar, caso haja algum problema.

Totais ou Incrementais

Após o primeiro backup total a estratégia que melhor se adequa para a base de dados, trata-se de uma estratégia de backups incrementais. Deste modo fica-se com copias sucessivas dos dados que foram alterados desde o ultimo backup de nível zero.

Com uma estratégia de backups incremental, é preciso decidir se se opta por uma incremental cumulativo ou incremental diferencial. Sendo diferencial, qualquer backup de nível 1 é feito com dados atualizados desde o ultimo backup de nível 1, excetuando o backup inicial de nível 1. No cumulativo, qualquer backup de nível 1 é feito com os dados atualizados desde o ultimo backup de nível 0.

Visto que a base de dados em questão é poucas vezes atualizada durante o ano letivo, achamos que uma estratégia de backup incremental cumulativo se adequa, pelo facto de não ser um backup pesado e ser o mais seguro, pois se obtem todos os dados desde o backup completo de uma vez.

De forma a denominar os nossos ficheiros de backups usamos uma configuração que indique a data do backup, o nome da base de dados e o nível do backup, ficando com este aspeto NOMEEDMMAAAANIVEL. Um exemplo de um ficheiro de backup será SchoolDB270920171.bkp.

Número de backups em cada nível 0 ou 1

Perder um ficheiro de backup é um caso sempre possível numa base de dados, daí uma estratégia para melhor salvaguardar a base de dados consistir na redundância de backups, ou seja existirem copias dos ficheiros backups.

Nesta base de dados é optado por criar cinco ficheiros de backups de nível 0, visto ser o backup mais importante pois contem a maior parte dos dados e a totalidade de base de dados. Para o nível 1 é optado por criar três ficheiros de backup pois sendo ficheiros mais pequenos a informação a perder, apesar de também ser bastante importante, não é tão elevada.

A redundância de backups também é abordada a nível de hardware, sendo este ponto abordado no próximo tópico.

Definição do armazenamento físico

Os ficheiros de backups tem de ser armazenados em hardware, mas uma simples solução de armazenar todos os ficheiros numa só unidade não é viável visto que essa unidade pode falhar, fazendo com que se perca todo o backup da base de dados.

Tendo à disposição hardware como disco, volumes, controladores, tapes entre outros achamos que a nossa base de dados fica melhor salvaguardada dividindo os cinco ficheiros de backup de nível 0 por 5 discos diferentes e os ficheiros de nível 1 colocar tres copias em cada um desses cinco discos. Também estas três copias de nível 1 podem ser divididas em cada disco, separando-as por cada volume.

Sempre que um destes componentes falhar ou ficar corrompido tem de se substituir por um novo, copiando os ficheiros que estão nos outros componentes "saudáveis".

Distribuição física dos discos

A ultima fase do processo de salvaguarda da base de dados consiste no armazenamento do hardware, mais uma vez nesta fase não é viável armazenar todos os componentes no mesmo local, visto que pode ocorrer alguma catástrofe natural ou até mesmo um roubo nas instalações, perdendo todos os dados.

Visto que a nossa base de dados contem dados referentes a escolas e instituições dos Estados Unidos da América faz sentido que algum do hardware esteja presente nesse mesmo país. Como no nosso sistema existem cinco discos a estratégia será guardar um disco em Portugal nas instalações da empresa que desenvolveu a base de dados e os outros quatro discos nos Estados Unidos. Um deles ficará no ministério da Educação e os outros três ficarão em estados a designar sendo que pelo menos dois terão de ficar em locais opostos do país.

Uma solução ainda mais segura poderá focar-se no facto de todos os anos mover os três componentes de hardware para estados diferentes do pais, ficando só o disco em Portugal e no ministério da Educação dos EUA com localização fixa.

Conclusão

Este trabalho ajudou a realçar a importância dos backups.

Para o tema escolhido, preferimos determinar o backup de nível 0 no fim do período escolar, pois é nesse período que a escola possui todos os dados relevantes presentes na base de dados.

Optámos por uma estratégia incremental para não ser necessário uma grande quantidade de memória e também por ser uma estratégia mais segura.

Por último definimos os números de componentes de hardware para armazenar os ficheiros de backup e as respectivas distribuições geográficas.

Concluindo assim esta fase o resultado foi um plano de salvaguarda para a nossa base de dados que nos garante a segurança dos dados.
