

Re: on-line and near-line databases

Source:

<http://www.tech-archive.net/Archive/SQL-Server/microsoft.public.sqlserver.replication/2007-03/msg00116.html>

- *From:* ishmael.tabaha@xxxxxxxxxx
 - *Date:* 8 Mar 2007 21:59:41 -0800
-

On Mar 8, 10:29 pm, kipdayn...@xxxxxxxxxx wrote:

Thanks for your help, Paul. I'm another member of the team and had a few questions/comments:

ishmael.tab...@xxxxxxxxxx wrote:

We would like to implement some sort of approach where we archive user data older than some delta, say 30 days, into another database (the 'B' database) and only keep users from the last 30 days in the production ('A') database. If an old user does return to the site, we would then retrieve their data from the archive database and bring in back to the 'current' production database.

A slight correction, Ish... I think what we want is to remove user records from our production DB which haven't been accessed within some period. i.e. we'd delete records with the 'datelastaccessed' timestamp older than x days. This also means that restored records should again be deleted if not accessed within our archive period. We need to make sure changes/insertions affecting/related to restored records are replicated to the archive as well.

Agreed, we'd do it by lastaccessed. I think though it would be by user – i.e. we would remove all user data for users with lastaccessed > x days which of course would involve a number of tables.

"Paul Ibison" <Paul.Ibi...@xxxxxxxxxxxxxxxx> wrote:

Perhaps the best method would be to archive off the old records to a separate table then remove them from the replicated table. This way adding them back in won't cause a problem. Views can be used to union the records together to present a complete record if necessary.

Re: on-line and near-line databases

Paul, I want to be sure I understand what you are suggesting for the data restoration approach. Here's what I understood you meant by 'archive off...to a separate table', please confirm:

1. Copy the records in the subscriber database which are to be restored to separate 'archive' tables within the subscriber DB (or non-replicated tables within publisher)
2. Delete the records copied in step 1 from the replicated tables
3. Insert those records back to the publisher DB. (at which point they would be replicated back to the subscriber in their original place)
4. Delete the records restored from the 'archive' tables

What need would there be for views to union the records together? ...views to records in which tables and at which point are you referring to?

ishmael.tab...@xxxxxxxxx wrote:

Is it perhaps possible to filter the A->B replication to only do it for rows where for example the creationdate/timestamp > some setting (say, today-30)?

Ish, Do you mean this as an alternative to Paul's intermediate table suggestion? I'm not sure this would work since we would want any changes made to old (restored) records (or new records associated with restored users) replicated immediately as well. For example, when a user record is restored a new session row is generated, datelastaccessed updated – this insert and update need to be replicated back to the archive.

Yes, I was hoping this could be an alternative to the intermediate table solution. My understanding is we would only be suppressing DELETES, and that too only on the very specific subset of tables that hold user data – UPDATES and INSERTs would still propagate to the subscriber database. So the new session row and any changes to the user's info would be replicated. We'd have to investigate whether we are using deletes anywhere to perform an update operation (e.g. deleting a set of rows and reinserting new rows) on user data. I can see how this might be a problem

Hilary suggested an approach that might solve this:

do your batch delete's through a proc and replicate the execution of the proc. The proc on the subscriber can do nothing. Regular deletes (ones which are not part of the archive process) which occur on the publisher table will be replicated

Re: on-line and near-line databases

I'm not quite sure how this would work but it sounds promising. I assume this means we have an sp which deletes data that is archival-ready (lastaccessed > x days) and we 'replicate the execution of the proc' on the subscriber, but with a dummy do-nothing sp. But wouldn't the deletes performed by the execution of the SP on the publisher also replicate? Or does a replicated SP happen on the SP level on both publisher and subscriber, and not on the table level?

.