

Extreme performance issues (SQL Server 2000/ADO.NET/C#)

Source:

<http://www.tech-archive.net/Archive/DotNet/microsoft.public.dotnet.framework.adonet/2004-09/0732.html>

From: Jon Rista (*jrista_at_hotmail.com*)

Date: 09/14/04

Date: Mon, 13 Sep 2004 19:50:04 -0600

I'm using ADO.NET in a windows service application to perform a process on SQL Server 2000. This process runs very quickly if run through Query Analyser or Enterprise Manager, but takes an excessively long time when run through my application. To be more precise, executing stored procedures and views through Query Analyser take between 10 and 20 seconds to complete. The same exact stored procedures and views, run in the same exact order, through my program, take anywhere from 30 minutes to 2 hours to complete, and the system that runs SQL Server (a 4-cpu Xeons system with 2gigs of physical ram) is pegged at 25% cpu usage (the query uses 100% of a single cpu's worth of processing power). I am at a complete loss as to why such a vast difference in execution time would occur, but here are some details.

The windows service executes on a workstation.

SQL Server 2000 executes on a server different from the workstation through a 100mbps ethernet network.

Query Analyser/Enterprise Manager run on the same workstation as the windows service.

The process is as follows:

- 1) Run a stored procedure to clear temp tables.
- 2) Import raw text data into a SQL Server table (Reconciliation).
- 3) Import data from a Microsoft Access database into 3 SQL Server tables (Accounts, HistoricalPayments, CurrentPayments).

(This takes about 10 – 15minutes to import 70,000 – 100,000 records from an access database, housed on a network share on a different server.)

- 4) "Bucketize" the imported data. This process gathers data from the 4 tables stated so far (Reconciliation, Accounts, HistoricalPayments, CurrentPayments, and places records into another table (Buckets) and assigned a primary category number to each record through a stored procedure.

- 5) Sort buckets of data into subcategories, updating each record in (Buckets) and assigning a sub category number, through another stored procedure.

- 6) Retrieve a summary of the data in (Buckets) (this summary is a count of

rows and summation of monetary values), grouped by the primary category number. This is a view.

7) Retrieve a summary of the data in (Buckets), grouped by both the primary and sub category numbers. This is a view.

When I execute these steps manually through query analyser, (save step 3), each query takes anywhere from 1 second to 20 seconds. The views, surprisingly, take more time than the fairly complex stored procedures of step 4 and 5.

When I execute these steps automatically using my windows service (written in .NET, C#, using ADO.NET), the simple stored procedures like clearing tables and whatnot execute quickly, but the stored procedures and views from steps 4–7 take an extremely long time. The stored procedures take at a minimum 30 minutes to complete, and sometimes nearly an hour. The views are the worst of all, taking no less than 1 hour to run, and often two hours (probably longer, actually, since my CommandTimeout is set to 7200 seconds, or two hours). I have never seen such a drastic difference between the execution of a query or stored procedure between query analyser and an application. There should be little or no difference at all, considering that everything is stored procedures (even the views...I wrap all the views in a simple stored procedure that calls the view using a SELECT), and as such executes on the server. Not only that, but Query Analyser is running on the same exact box that the application is running on, and is connecting to the same SQL Server.

I doubt this is a network bandwidth issue, as after calling the stored procedure from code, there is no network activity except mssql keep-alive messages, until the procedure completes and returns its result set or return value (if any), and then its only a momentary blip as the data is sent across.

I've followed proper practice when using views and stored procedures. When I select, I always explicitly name the columns I wish to retrieve. I have appropriate indexes on the columns in the 4 data tables. The queries that execute in the stored procedures are fairly complex, involving summations, count(), group by, and order by. I can understand a moderate difference in performance between query analyser and an ADO.NET application due to ADO.NETs extra overhead, but a difference between 20 seconds and 1 hour is more than can be attributed to .NET overhead.

I greatly appreciate anyone who might have some insight to this offering some help. I've scanned the net looking for similar situations, but searching for them is somewhat difficult, considering the nature and volume of factors. Thanks.

— Jon Rista