

Re: Generating one table with a terabyte of data

Source:

<http://www.tech-archive.net/Archive/SQL-Server/microsoft.public.sqlserver.dts/2005-05/msg00201.html>

- *From:* "Kevin3NF" <KHill@xx>
 - *Date:* Mon, 9 May 2005 14:04:26 -0500
-

I've locked SQL server to 1/3 physical memory usage. DB growth is 512Meg chunks

You need to free it up to fly, and pre-allocate the db size. In-process file growth will kill performance

—
Kevin Hill
President
3NF Consulting

www.3nf-inc.com/NewsGroups.htm

www.DallasDBAs.com/forum – new DB forum for Dallas/Ft. Worth area DBAs.

www.experts-exchange.com – experts compete for points to answer your questions

"lee" <lee@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx> wrote in message
news:B772A8C2-F009-4E29-816C-5B6C2F45151E@xxxxxxxxxxxxxxxx

> Hi All,

>

> I need ideas on how to improve a process that can generate a terabyte of
> data in a week. Our company produces software that analyzes code and one
> of

> the things we do is Code Duplication. We're processing hundreds of
> millions

> of lines against each other and generate gigabytes of results. In reality
> we

> generate over 20 gigabytes of data a day. We are noticing major slow downs
> in

> the system. Here's how we're going about it

>

> 1. Non-Indexed table residing on it's own FileGroup (which is not on the
> same drive as the SQL server).

> 2. A process that generates Text files that we use BULK INSERT to load the
> data (any other way of doing inserts is just to slow).

Re: Generating one table with a terabyte of data

>
> This process should generate 10 million rows in the table per hour. Only
> problem is that we're noticing a considerable slow down as we pummel the
> Database. We're making sure our Bulk Insert has exclusive table locks and
> nobody else accesses the system while this is happening. We're also
> noticing
> a strange behavior at the OS level. The system just slows down to a crawl.
> Even mouse movement becomes jagged (Win2K Server).
>
> I should also mention that once we're done inserting we create indexes at
> the end. Average time to create 1 indexed field (on a different filegroup)
> on
> 100 million rows is about 3 to 4 hours. Seems a little excessive to me,
> must
> be doing something wrong.
>
> I've gone over this 100 times and I can't see anything else to improve
> performance. I have a friend that is an ORACLE DBA and he keeps telling me
> that SQL server just can't handle this type of volume and that I should
> move
> to ORACLE. I refuse to believe ORACLE can do more then SQL Server. He
> actually took one of our 500 meg text files and loaded it in 1/10 the time
> it
> took SQL Server. My answer to that is "I must be configuring it
> incorrectly".
>
> I've locked SQL server to 1/3 physical memory usage. DB growth is 512Meg
> chunks. We don't need any transaction logs, so we're using RECOVERY MODEL
> =
> SIMPLE (any other mode kills SQL server anyway, LOG fills to fast with our
> data creation speed). I've thrown everything at this except the kitchen
> sink.
>
> Any suggestions on how I can improve my Bulk Inserts and how to improve my
> Index creations? Should I consider SQL Server 2005? Will it handle my
> needs
> better?
>

• **Follow-Ups:**

◆ **Re: Generating one table with a terabyte of data**

◇ From: Jim Vierra

• **References:**

◆ **Generating one table with a terabyte of data**

◇ From: lee

Re: Generating one table with a terabyte of data

- Prev by Date: [*Re: Unzip file using DTS*](#)
- Next by Date: [*ActiveX task hangs in DTS*](#)
- Previous by thread: [*Re: Generating one table with a terabyte of data*](#)
- Next by thread: [*Re: Generating one table with a terabyte of data*](#)
- Index(es):
 - ◆ [*Date*](#)
 - ◆ [*Thread*](#)