

Re: Network traffic with transactional replication with immediate updates (SQL 2005)

Source:

<http://www.tech-archive.net/Archive/SQL-Server/microsoft.public.sqlserver.replication/2006-08/msg00130.html>

- *From:* "J.A. García Barceló" <jagbarcelo@xxxxxxxxxxxxxxxx>
 - *Date:* Mon, 24 Jul 2006 14:25:18 +0200
-

Of course I could schedule it to run every 1, 2 or 5 minutes instead of 5 seconds (by default), but in that case the latency of transferred commands will increase accordingly. That will soften the pain but it will not be the cure.

This kind of configuration using 28 publications for a db, a single publisher/distributor running on the main office and a subscriber at the branch has been working without such a big network impact for years with SQL Server 2000. This is a new SQL Server 2005 behaviour, and IMHO is against slow-links network efficiency. If there is not a really important reason for running those commands against the subscriber when there is nothing to replicate, the agent should execute them (maybe the agent just launches those commands in order to be able to calculate average transactions latencies, other statistical reasons).

I think I will open it as a bug at <http://connect.microsoft.com/feedback/default.aspx?SiteID=68> or, at least, as a comment.

--

J.A. García Barceló
<http://jagbarcelo.blogspot.com/>

"Hilary Cotter" <hilary.cotter@xxxxxxxxxx> escribió en el mensaje
news:eFOJOjxrGHA.3324@xxxxxxxxxxxxxxxxxxxxxxxxxxxx

I have confirmed that it does run every 5 seconds. Perhaps if you were to schedule your agents to run every minute it would replicate all transactions and related commands in the queue, go to sleep for 1 minute and then wake up and process the rest. It really depends on what your real time latency requirements are.

--

Hilary Cotter
Director of Text Mining and Database Strategy
RelevantNOISE.Com – Dedicated to mining blogs for business intelligence.

Re: Network traffic with transactional replication with immediate updates (SQL 2005)

This posting is my own and doesn't necessarily represent RelevantNoise's positions, strategies or opinions.

Looking for a SQL Server replication book?

<http://www.nwsu.com/0974973602.html>

Looking for a FAQ on Indexing Services/SQL FTS

<http://www.indexserverfaq.com>

"J.A. García Barceló" <jagbarcelo@xxxxxxxxxxxxxxxx> wrote in message
[news:O\\$7i9fvrGHA.2464@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](news:O$7i9fvrGHA.2464@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

As suggested in microsoft.public.es.sqlserver (where I already posted this problem, but translated), I run profiler on the subscriber:

I do not need to wait until off-hours to run a trace on the subscriber. Right now the database is divided in 28 publications (with 155 tables in total, it is not a good idea to have a huge single publication with the added problems of reinitialisations, schema changes, etc). Some of the publications contain tables seldomly used or almost not updated at all. In particular, I run profiler on the subscriber with a filter:

HostName : publisher name (MYSERVER)

ApplicationName : MYSERVER_mydb_mypublication

And, right now, with a PollingInterval=60 secs, with the trace running filtered by a publication with tables not updated at all, I can read EVERY 60 seconds the following RPC:Completed command:

```
exec sp_executesql N'update MSreplication_subscriptions set
transaction_timestamp = cast(@P1 as binary(15)) +
cast(substring(transaction_timestamp, 16, 1) as binary(1)), "time" = @P2
where UPPER(publisher) = UPPER(@P3) and publisher_db = @P4 and
publication =
@P5 and subscription_type = 0 and (substring(transaction_timestamp, 16,
1) =
0 or datalength(transaction_timestamp) < 16)', N'@P1 varbinary(14),@P2
datetime,@P3 nvarchar(5),@P4 nvarchar(5),@P5 nvarchar(7)',
0x00004F3200000032000100000000, 'Jul 24 2006 9:03:59:000AM',
N'MYSERVER',
N'mydb', N'mypublication'
```

Oviously, if, instead of running this every 60 seconds it was executed every 5 seconds (and taking into account that we have 28 agents doing the same) we would have: $60 \times 28 / 5 = 336$ RPCs/min or more than 5 RPCs/sec 24/7.

If we add legitimate user transactions, usual network traffic (mail, web browsing, dfs transfers) we get the network permanently congested at 300kbps.

Re: Network traffic with transactional replication with immediate updates (SQL 2005)

Can someone else confirm that their subscribers receive these RPC commands every 5 seconds (default PollingInterval for Distribution Agents)? If this is proved to be true, we would be facing a severe problem from the performance (and scalability) point of view when the number of subscribers grow.

Regards.

PS: I forgot to mention: the RPC command sent to subscriber is not always the same: transaction_timestamp and time changes every time ;)

--

J.A. García Barceló
<http://jagbarcelo.blogspot.com/>

--

J.A. García Barceló
<http://jagbarcelo.blogspot.com/>

"Hilary Cotter" <hilary.cotter@xxxxxxxx> escribió en el mensaje
<news:%23jbWYaSrGHA.4016@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>

What autosenses does is if the connection fails the distribution agent will restart if the connection comes back. I am wondering if this isn't pinging the subscriber even if you are running continuously. This behavior should be disabled if you are running at scheduled intervals, ie every 2 minutes or something like that.

--

Hilary Cotter
Director of Text Mining and Database Strategy
RelevantNOISE.Com – Dedicated to mining blogs for business intelligence.

This posting is my own and doesn't necessarily represent RelevantNoise's positions, strategies or opinions.

Looking for a SQL Server replication book?
<http://www.nwsu.com/0974973602.html>

Looking for a FAQ on Indexing Services/SQL FTS
<http://www.indexserverfaq.com>

"J.A. García Barceló" <jagbarcelo@xxxxxxxxxxxxxx>
wrote in message
news:uVHmKNRrGHA.3564@xxxxxxxxxxxxxxxxxxxxxxxxxxxx

Yes, they are running continuously:
[...]-Continuous

Specifies whether the agent attempts to poll replicated transactions continually. If specified, the agent polls replicated transactions from the source at polling intervals, even if there are no transactions pending.[...]

This is by default when you set up a new subscription with transactional replication with immediate updates. Besides, it should not matter how frequently you query the distribution database since the agents are running in the same server as the distribution database is. The right question is: Why is subscriber contacted (traffic) when there is nothing to replicate?

I could try to change things next monday, in the meantime... what is that autosense connection functionality you mentioned? I can't find a word about it now (I don't have BOL in this machine either).

Thanks and regards.

--
J.A. García Barceló
<http://jagbarcelo.blogspot.com/>

"Hilary Cotter" <hilary.cotter@xxxxxxxx>
escribió en el mensaje
news:u1qZ%23GRrGHA.3680@xxxxxxxxxxxxxxxxxxxxxxxxxxxx

This could be the autosense connection functionality.
Are your agents running continuously? What happens if you schedule

them for every 5
minutes or something like
that?

--

Hilary Cotter
Director of Text Mining and
Database Strategy
RelevantNOISE.Com –
Dedicated to mining blogs
for business
intelligence.

This posting is my own and
doesn't necessarily represent
RelevantNoise's positions,
strategies or opinions.

Looking for a SQL Server
replication book?
<http://www.nwsu.com/0974973602.html>

Looking for a FAQ on
Indexing Services/SQL FTS
<http://www.indexserverfaq.com>

"J.A. García Barceló"
<jagbarcelo@xxxxxxxxxxxxxxxx>
wrote in message
<news:eXVOAqOrGHA.4680@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>

We have a VPN set with
two xDSL between our
main office and a branch
site. There are two sql
servers (at each side) with
transactional
replication with immediate
updates. As you might
imagine, bandwidth is
critical (it is not a T1 line,
we have only 300kbps
maximum) and it
must be used for other
purposes also (file
transfer/replication using
Windows DFS, web
browsing, email, etc.)

We have also set up a

network bandwidth
monitoring software
(<http://www.cacti.net/>) to be
able to check the status of
the physical
lines, and VPN, and see
their evolution in time, if
there are
bottlenecks, and so on.

We have been using SQL
Server 2000 for quite a long
time, both at the
publisher and at the
subscriber. One month ago,
more or less, we
installed SQL Server 2005
at the main office as new
publisher (we are
migrating step by step,
subscriber will go next), and
the network
traffic has boosted. We are
now consuming all the
300kbps (uploading
from main office to the
branch site) permanently,
and 95kbp of
downloading (from branch
to main office).

We have confirmed that all
that traffic is caused by SQL
Server, and
not other kind of traffic.
Furthermore, after some
tests, we have
trace it down and guessed
that Distributor Agent is to
blame. Default
agent profile for distributor
agents has as profile
parameters one
called PollingInterval, set by
default to 5 seconds. We
have created a
new user profile based on
the default one, and changed
PollingInterval
from 5 to 60 seconds.

The traffic has been reduced
in somewhat near 12 times
less
 $60/5 = 12$
from 300kbps to something
near 40kbps (being so small,
those 40kbps
can contain other kind of
traffic)

However, I cannot find a
way to explain why
Distributor Agents are to
blame. According
<http://msdn2.microsoft.com/en-us/library/ms147328.aspx>
and regarding
PollingInterval parameter:
[...]Is how often, in seconds,
the distribution database is
queried
for replicated transactions.
The default is 5 seconds.[...]

Both the agents and the
distribution database are at
the same
computer: the publisher at
the main office. Being that
way, in theory,
changing that parameter
from 5 to 60 seconds would
only affect local
traffic. Even more, during
the night, when there is no
activity at
none of the offices, no
matter the frequency the
agent checks the
distribution database, there
should be no pending
transactions to be
applied to subscriber, and
the subscriber should not be
contacted. Our
experience shows that there
is traffic between publisher
and
subscriber, even when there
are no transactions to be
applied, and the
bandwidth of that traffic

depends directly on the value of Distributor Agents' PollingInterval parameter.

I simply cannot understand it.

I attach a pair of images:

In this one we can see how, since the middle of June, the date when SQL Server 2005 started to act as publisher, the traffic reaches 300kbps.

In this capture, we can see the last 24 hours. The valley in traffic around 13:00h belongs to a change of all Distributor Agents' profiles, with a PollingInterval=60 and a restart of the SQL Server Agent service. Half an hour later (more or less), in order to check, the default agent profile is set again (PollingInterval=5) for every Distributor Agent and a new restart of the SQL Server Agent service is done also.

It is very extrange that the traffic stalls permanently around 300kbps, even during the night, when there is no activity (just a few jobs that call to some stored procedures to update a few administrative tables, scheduled to run at 00:00 and 5:00am)

Re: Network traffic with transactional replication with immediate updates (SQL 2005)

Can someone find a reasonable explanation for this? Might it be a SQL Server bug? Should the subscriber be contacted by distributor agents even when there is nothing to replicate?

Regards and thanks a lot for your time.

PS: Publisher is SQL Server 2005 x64 + SP1 running on a Dual Xeon with 4Gb RAM and Windows Server 2003 R2 x64.

--

J.A. García Barceló
<http://jagbarcelo.blogspot.com/>