International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

SQL Server Replication Using FTP Server

Amol Gawas[1], Vinod Kochu[2]

[1]Kotak Mahindra Bank, Kotak Securities, HQ, Goregaon, Mumbai, India

[2] Mentor, Kotak Mahindra Bank, Kotak Securities, HQ, Goregaon, Mumbai, India

Abstract: ***Sql server supports multiple high availability solutions. Replication tends to have many advantages but these come with a complex architecture. Sql server allows replication to be configured between two different servers without the other one being of sql server origin. Replication is done by synchronizing data objects. The process is one way as well as two way depending on the business requirement. The usual setup includes sql server at both source and destination, the internals then use replication components to sync data objects between two servers. Replication strictly involves sharing of database logs which needs to be pushed to the other server or the server needs to pull the logs. This criteria requires the two servers to be in same domain so that this sharing takes place with respect to security policies defined by Microsoft Sql server. The real issue arises when the same setup needs to be done between two different domains. In this paper, is demonstrated a solution to the projected issue, which uses file transfer protocol server as mediator between two servers and establishing the syncing process of sql server replication.***

Keywords: Sql Server Replication, Transactional Replication, FTP, IIS server, cross domain

1. Introduction

Organizations have been competing with each other since their existence, but now the competition is of being technologically advanced. This race has introduced the organizations into new cutting edge engineering, the most important venture point is the fault tolerance. The availability of service given any disorder or outage. The business should be up and running minimizing the downtime and potential losses. The availability of business service with most uptime and minimal outage severely impacts positively the organization's popularity and trust in market hence improving business.

Replication is a high availability solution provided by Microsoft Sql server which can sync data object from one server to other making data available at both servers at same time, it is high fault tolerance solution which leads the business to be effectively run from both setups. We will demonstrate the solution to the replication that needs to be configured between two servers from different domains and where trust between these two servers is not possible due to security policies. Replication requires the database changes to be pushed or pull to/from source to destination server. This problem statement can be answered using a mediator which can be used for logs transfer, Based on FTP server we have managed to achieve the replication setup between these servers distinctly different environments.

2. Replication Mechanism in Brief

Microsoft SQL Server presented Replication as an integration of publishing and distributing data and data objects between databases. It makes data available to multiple users over the network. Replication gives the freedom to choose what data objects that need to be synchronized over the network. Unlike other high availability solutions provided this technology lets a parts of database to be made available to distributed users.

Think of it as a newspaper delivery system, the sql server as the press and customer as the secondary target server, Press publishes articles of their newspaper and selects a middleman distributor to handle all the flow to the customer. The customer can select parts or all the articles of newspaper to be delivered. The deal customer has with publisher is called subscription. The example above explains replication logic in a lay man's terms. Here the press is the source database which is called Publisher, which publishes data objects as Articles, the Distributor is another database which can be on same server as publisher is on or it can be different server. The duty of Distributor is to manage the publisher's changes and relay those to the secondary database known as Subscriber, The relation between subscriber and publisher is called Subscription.

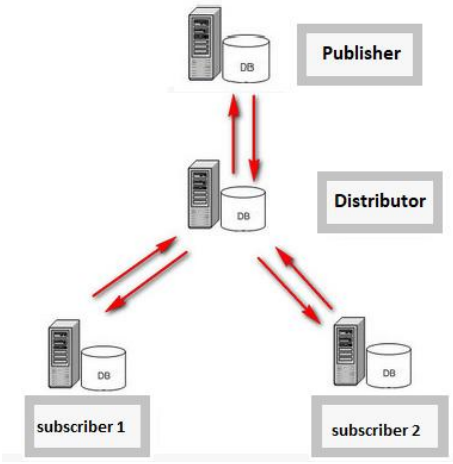

Figure 1

Replication in sql server has four major types, we are going to pertain to Transactional replication for this paper.

Transactional replication in a gist is distributing and applying every single transaction that modifies data in article at publisher. However transactional replication is a one way street, the changes occurring at subscriber end are not replicated back to publisher. All the modifications that

happen on article (data object) at publisher are stored in Distribution database and from there it is relayed to all the subscribers.

The Replication uses two modes to sync data with the subscriber, Pull Subscription and push subscription, here the goal is to make changes that occur on primary database to be available to other subscriber databases over the network. This can be achieved by either the distribution database pushes latest modifications to the subscribers (Push) or the subscribers keep checking the distribution database on regular time intervals for latest changes (Pull).

3. Problem Statement

Replication internally is partially depended on windows for security constraints of the participating server and databases.

- It needs the servers of publisher, distributor and subscriber to be connected via the Hostname for configuration. This means that both the server must resolve each other's hostname and ip.
- Sql server ports must be accessible to and forth from both servers.
- Secondly to transfer and apply the logs of modified data from Server A to Server B, the sql service account of both servers must be under same name or the service account must be known to both servers.
- Also Sql service agent account must be known to both servers.

With respect to above points all the participating servers should belong to same domain so that rights and user/service accounts are recognized which has a huge plus side, when the servers are in same domain the replication is easier and less complex to configure as there is trust between two servers and mainly both belong to same ip segment which nullifies any port that needs to be explicitly given access to for the setup.

The challenge is to configure the above setup on the servers belonging to two different domains, and no trust can be formed between these two servers. This leads to losing all advantages we had when the servers where in same domain.

Let's frame our problem statement in example , Server A has a database Sales, Server B has the same replica of database Sales as on Server A created by backup and restore from initial steps of replication setup. Server A is also going to be the distributor itself. Server A belongs to domain DOMAIN_1 and Server B belongs to domain DOMAIN_2. We need to configure the transactional replication setup as discussed above

4. Discussion

Based on the problem statement, the replication needs to be done between Server A to Server B on database Sales, Replication on the internal uses triggers and log read agent which mark transactions in database log and these are copied to distribution database and stored. Then distributor transports these transaction to subscriber, sql server manages this through number of replication jobs and linked servers.

By default the sql server stores the replication transactions, snapshots to be transferred in a "ReplData\unc" which is in folder path where sql server is installed, This path is provided in publisher > properties, using service account of sql server the subscriber authenticates itself to access this folder and reads and applies the transaction in the files. Here's the catch as the problem statement states that servers are in different domains the subscriber couldn't access the files in repldata folder.

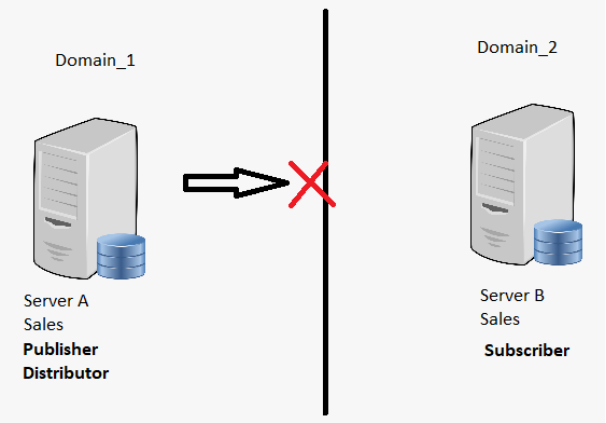

Figure 2

We will be using ftp to get around this domain boundary, ftp because sql server replication supports ftp, as the server B services account could not access the files, we will use ftp credentials on both ends so that the ftp server is accessible from server B.

Before the ftp server is created over the repldata folder make sure to change the default path to some other location as configuring ftp server over the folder path where sql server is installed is a security risk. Once the default location is changed, the ftp server must point the folder till "unc". The folder generated by sql server must not be included in the ftp server. Sql server replication is hard coded internally to look for the replication folder based on publisher name. If we include the sql generated folder the sql server will search for the folder inside the ftp server which will result into failure. Once we configured the ftp server which points to the "repldata" or the path which further has sql server generated folders it needs be to secure with a suitable password. An ftp port 21 needs to be opened between the servers additional to the sql ports. Firstly the ftp server should be open from Server B to test that the folders and files are seen, use Internet explorer for the same. Once the ftp server is accessible from Server B with proper credentials the sql server on B would be able to access the files and apply them. We will need to access the FTP port using rule on firewall of the Server A (publisher). If connection errors still observed despite the firewall changes on Server A, need to check firewall at Server B disable and enable the firewall to check, infrastructure team must now the proper set of rules that are required for the same.

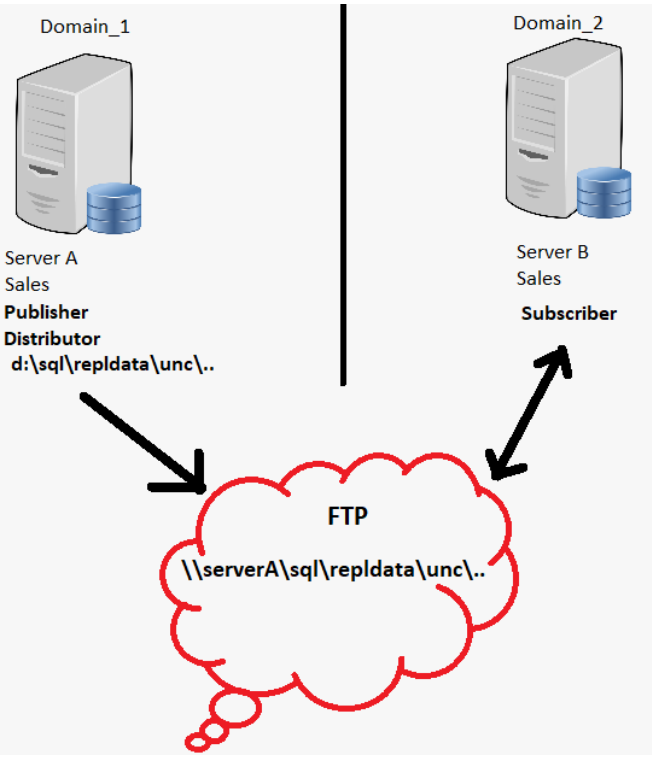

Figure 3

The ftp is going to act as share path to get us around the domain. Now configure this ftp path in replication with valid/same credentials at both the ends. The replication sql jobs will hit the ftp server and search for the sql generated folders and files and syncing of replication will begin. This proposed solution has following strengths:

- This setup does not need to form any trust between two domains
- Ftp can resume the transfer if connection has had any issue.
- More stringent security measures can be applied to ftp server.
- Setup does not violate any security organizational policies followed for distinguishing domains.

5. Conclusion

In this paper, the detailed explanation about sql server Transactional replication is stated and a workable work around to setup a successful sync between two servers, the purpose method in the paper keeps in mind the requirements of the client and the understanding of the domain boundaries and goes further to make use of ftp server as a mediator for the sync between two servers of two different domain with non-similar ip segments.

References

[1] Microsoft documents on sql server replication. https://docs.microsoft.com/en-us/sql/relational-databases/replication/sql-server-replication?view=sql-server-ver15
[2] Paul Randal extensive literature on replication https://www.sqlskills.com/blogs/paul/category/replication/
[3] Johnathan Kehayias and Ted Krueger, Troubleshooting SQL Server, A guide for the accidental DBA, chapter 8, pp. 296 -297

Author Profile

Amol Gawas received the B.E degrees in Information Technology Engineering from Shah and Anchor College in 2017, is certified oracle database administrator and proficiency in managing and installing database management setups, disaster recovery plans, performance tuning, cross database connectivity, server migration and upgradation and also an avid programmer, currently a Team Lead database administrator at a prominent financial organization.