Replication (SQL Server 2000)
How Snapshot Replication Works

Snapshot replication is implemented by the Snapshot Agent and the Distribution Agent. The Snapshot Agent prepares snapshot files containing schema and data of published tables and database objects, stores the files in the snapshot folder, and records synchronization jobs in the distribution database on the Distributor. By default, the snapshot folder is located on the Distributor, but you can specify an alternate location instead of or in addition to the default. For more information, see Alternate Snapshot Locations [http://msdn2.microsoft.com/en-us/library/aa237130(SQL.80,printer) .aspx] .

The Distribution Agent moves the snapshot held in the distribution database tables to the destination tables at the Subscribers. The distribution database is used only by replication and does not contain any user tables.
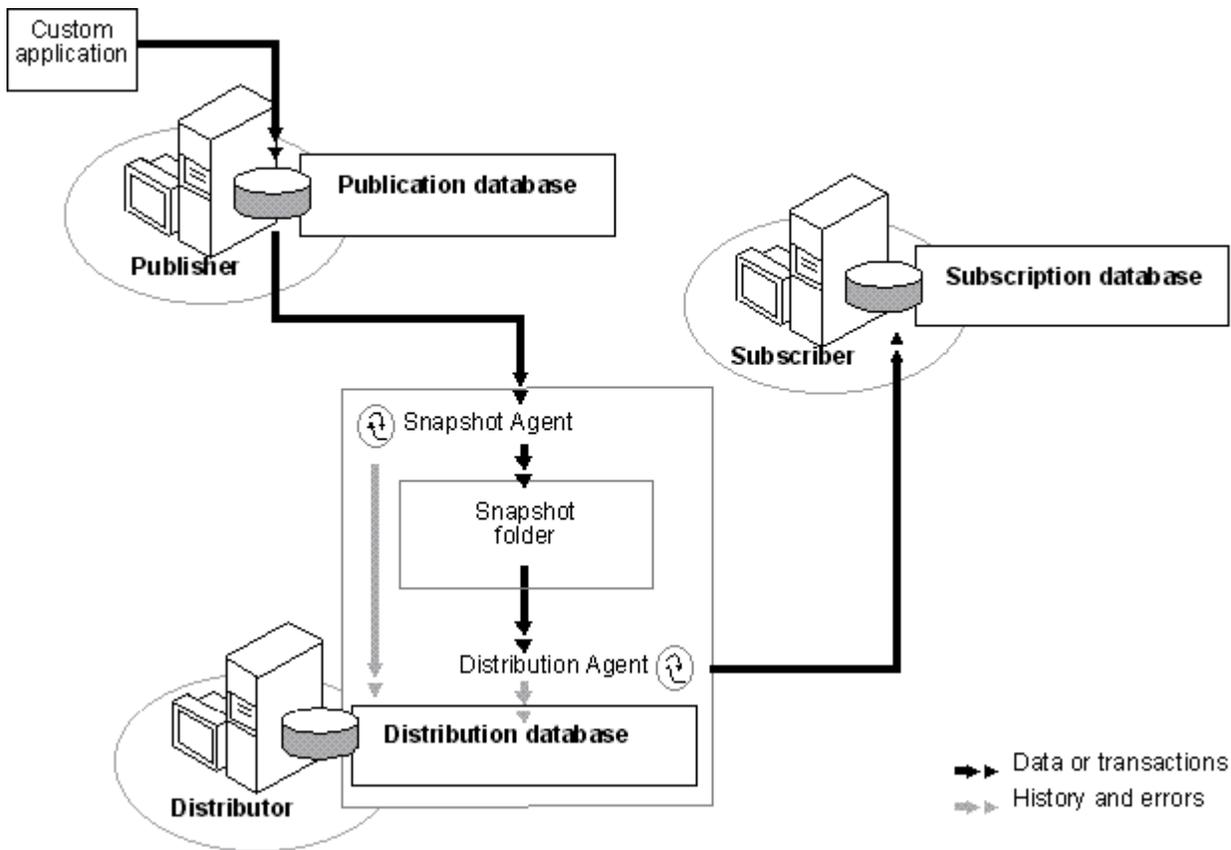
Snapshot Agent

Each time the Snapshot Agent runs, it checks to see if any new subscriptions have been added. If there are no new subscriptions, no new scripts or data files are created. If the publication is created with the option to create the first snapshot immediately enabled, new schema and data files are created each time the Snapshot Agent runs. All schema and data files are stored in the snapshot folder and then either the Distribution Agent or Merge Agent transfers them to Subscriber or you can transfer them manually. The Snapshot Agent performs the following steps:

1. Establishes a connection from the Distributor to the Publisher and sets a share-lock on all tables included in the publication. The share-lock ensures a consistent snapshot of data. Because the locks prevent all other users from updating the tables, the Snapshot Agent should be scheduled to execute during off-peak database activity.

2. Establishes a connection from the Publisher to the Distributor and writes a copy of the table schema for each article to an .sch file. If you request that indexes and declarative referential integrity be included, the agent scripts out the selected indexes to an .idx file. Other database objects, such as stored procedures, views, user-defined functions, and others, can also be published as part of replication.

3. Copies the data in the published table on the Publisher and writes the data to the snapshot folder. If all Subscribers are instances of Microsoft® SQL Server™ 2000, the snapshot is stored as a native bulk copy program file. If one or more Subscribers is a heterogeneous data source, the snapshot is stored as a character mode file. The files are the synchronization set that represents the table at one point in time. There is a synchronization set for each article within a publication.

4. Appends rows to the **MSrepl_commands** and **MSrepl_transactions** tables in the distribution database. The entries in the **MSrepl_commands** tables are commands indicating the location of the synchronization set (.sch and .bcp files) and references to any specified pre-creation scripts. The entries in the **MSrepl_transactions** table are commands referencing the Subscriber synchronization task.

5. Releases the share-locks on each published table and finishes writing the log history tables.

After the snapshot files are generated, you can view them in the Snapshot Folder using the Snapshot Explorer. In SQL Server Enterprise Manager, expand the Replication and Publications folders, right-click a publication, and then click **Explore the Latest Snapshot Folder**. For more information, see Exploring Snapshots [http://msdn2.microsoft.com/ en-us/library/aa237140(SQL.80,printer).aspx] .

Distribution Agent

Each time the Distribution Agent runs for a snapshot publication, it moves the schema and data to Subscribers. The Distribution Agent performs the following steps:

1. Establishes a connection from the server where the agent is located to the Distributor. For push subscriptions, the Distribution Agent is usually run on the Distributor, and for pull subscriptions, the Distribution Agent is usually run on the Subscriber.

2. Examines the **MSrepl_commands** and **MSrepl_transactions** tables in the distribution database on the Distributor. The agent reads the location of the synchronization set from the first table and the Subscriber synchronization commands from both tables.

3. Applies the schema and commands to the subscription database. If the Subscriber is not an instance of Microsoft SQL Server 2000, the agent converts the data types as necessary. All articles of a publication are synchronized, preserving transactional and referential integrity between the underlying tables (presuming the subscription database, if not SQL Server, has the transactional capabilities to do so).

When handling a large number of Subscribers, running the Distribution Agent at the Subscriber, either by using pull subscriptions or by using remote agent activation, can save processing resources on the Distributor. With remote agent activation, you can choose to run the Distribution Agent at the Subscriber for push subscriptions or at the Distributor for pull subscriptions. For more information, see Remote Agent Activation [http://msdn2.microsoft.com/en-us/library/ aa237177(SQL.80,printer).aspx] .

Snapshots can be applied either when the subscription is created or according to a schedule set at the time the publication is created.

Note For agents running at the Distributor, scheduled synchronization is based on the date and time at the Distributor (not the date and time at the Subscribers). Otherwise, the schedule is based on the date and time at the Subscriber.

Because automatic synchronization of databases or individual tables requires increased system overhead, a benefit of scheduling automatic synchronization for less frequent intervals is that it allows the initial snapshot to be scheduled for a period of low activity on the Publisher.

The Snapshot Agent is usually run by SQL Server Agent and can be administered directly by using SQL Server Enterprise Manager. The Snapshot Agent and Distribution Agent can also be embedded into applications by using Microsoft ActiveX® controls. The Snapshot Agent executes on the Distributor. The Distribution Agent usually executes on the Distributor for

push subscriptions, or on Subscribers for pull subscriptions, but remote agent activation can be used to offload Distribution Agent processing to another server.

Cleaning Up Snapshot Replication

When the distribution database is created, SQL Server 2000 adds the following tasks at the Distributor:

- Agent checkup

- Transaction cleanup

- History cleanup

These tasks help replication to function effectively in a long-running environment. After the snapshot is applied at all Subscribers, replication cleanup deletes the associated .bcp file for the initial snapshots automatically.

If the publication is enabled for anonymous subscriptions or with the option to create the first snapshot immediately, at least one copy of the snapshot files are kept in the snapshot location. This ensures that if a Subscriber with an anonymous subscription to a snapshot publication synchronizes with the Publisher, the most recent snapshot will be available.

See Also

Planning for Snapshot Replication [http://msdn2.microsoft.com/en-us/library/aa237481(SQL.80,printer).aspx]

Replication Options [http://msdn2.microsoft.com/en-us/library/aa237252(SQL.80,printer).aspx]