

[MS-DPREP]:

Replication Data Portability Overview

Intellectual Property Rights Notice for Open Specifications Documentation

- **Technical Documentation.** Microsoft publishes Open Specifications documentation (“this documentation”) for protocols, file formats, data portability, computer languages, and standards support. Additionally, overview documents cover inter-protocol relationships and interactions.
- **Copyrights.** This documentation is covered by Microsoft copyrights. Regardless of any other terms that are contained in the terms of use for the Microsoft website that hosts this documentation, you can make copies of it in order to develop implementations of the technologies that are described in this documentation and can distribute portions of it in your implementations that use these technologies or in your documentation as necessary to properly document the implementation. You can also distribute in your implementation, with or without modification, any schemas, IDLs, or code samples that are included in the documentation. This permission also applies to any documents that are referenced in the Open Specifications documentation.
- **No Trade Secrets.** Microsoft does not claim any trade secret rights in this documentation.
- **Patents.** Microsoft has patents that might cover your implementations of the technologies described in the Open Specifications documentation. Neither this notice nor Microsoft's delivery of this documentation grants any licenses under those patents or any other Microsoft patents. However, a given Open Specifications document might be covered by the Microsoft [Open Specifications Promise](#) or the [Microsoft Community Promise](#). If you would prefer a written license, or if the technologies described in this documentation are not covered by the Open Specifications Promise or Community Promise, as applicable, patent licenses are available by contacting iplg@microsoft.com.
- **License Programs.** To see all of the protocols in scope under a specific license program and the associated patents, visit the [Patent Map](#).
- **Trademarks.** The names of companies and products contained in this documentation might be covered by trademarks or similar intellectual property rights. This notice does not grant any licenses under those rights. For a list of Microsoft trademarks, visit www.microsoft.com/trademarks.
- **Fictitious Names.** The example companies, organizations, products, domain names, email addresses, logos, people, places, and events that are depicted in this documentation are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

Reservation of Rights. All other rights are reserved, and this notice does not grant any rights other than as specifically described above, whether by implication, estoppel, or otherwise.

Tools. The Open Specifications documentation does not require the use of Microsoft programming tools or programming environments in order for you to develop an implementation. If you have access to Microsoft programming tools and environments, you are free to take advantage of them. Certain Open Specifications documents are intended for use in conjunction with publicly available standards specifications and network programming art and, as such, assume that the reader either is familiar with the aforementioned material or has immediate access to it.

Support. For questions and support, please contact dochelp@microsoft.com.

Revision Summary

Date	Revision History	Revision Class	Comments
2/9/2011	0.1	New	Release new document
7/7/2011	0.1	None	No changes to the meaning, language, or formatting of the technical content.
11/3/2011	0.1	None	No changes to the meaning, language, or formatting of the technical content.
1/19/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
2/23/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
3/27/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
5/24/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
6/29/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
7/16/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
10/8/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
10/23/2012	0.1	None	No changes to the meaning, language, or formatting of the technical content.
3/26/2013	0.1	None	No changes to the meaning, language, or formatting of the technical content.
6/11/2013	1.0	Major	Updated and revised the technical content.
8/8/2013	2.0	Major	Updated and revised the technical content.
12/5/2013	2.0	None	No changes to the meaning, language, or formatting of the technical content.
2/11/2014	2.0	None	No changes to the meaning, language, or formatting of the technical content.
5/20/2014	2.0	None	No changes to the meaning, language, or formatting of the technical content.
5/10/2016	2.0	None	No changes to the meaning, language, or formatting of the technical content.
8/16/2017	2.1	Minor	Clarified the meaning of the technical content.

Table of Contents

1	Introduction	4
1.1	Glossary	4
1.2	References	5
2	Data Portability Scenario.....	7
2.1	Retrieve Intellectual Property from a Replication Topology	7
2.1.1	Data Description.....	7
2.1.2	Format and Protocol Summary	8
2.1.3	Data Portability Methodology.....	8
2.1.3.1	Preconditions.....	8
2.1.3.2	Versioning.....	8
2.1.3.3	Error Handling	9
2.1.3.4	Coherency Requirements	9
2.1.3.5	Additional Considerations	9
3	Change Tracking.....	10
4	Index.....	11

1 Introduction

The Replication Data Portability Overview document provides an overview of the components and methodologies that are used for data portability with the SQL Server Replication system.

In this document, replication refers to logical replication of data by using Microsoft SQL Server.

Replication is a set of technologies for copying and distributing data and database objects from one database to another and then synchronizing between databases to maintain consistency. By using replication, you can distribute data to different locations and to remote or mobile users over local and wide area networks, dial-up connections, wireless connections, and the Internet.

Transactional replication is typically used in server-to-server scenarios that require high throughput scalability and availability. It is used to feed data warehouse and reporting systems, integrate data from multiple sites, integrate heterogeneous data, and offload batch processing. Merge replication is primarily designed for mobile applications or distributed server applications that have possible data conflicts. Common scenarios include exchanging data with mobile users, consumer point of sale (POS) applications, and integration of data from multiple sites. Snapshot replication is used to provide the initial data set for transactional and merge replication; it can also be used when complete refreshes of data are appropriate. With these three types of replication, SQL Server provides a powerful and flexible system for synchronizing data across your enterprise.

Replication is implemented by using the following two-step process after the replication topology is set up:

- **Step 1: Initial synchronization.** Synchronization through which SQL Server delivers a snapshot to the destination. During this step, SQL Server ensures that the destination has the initial schema and data so that it can send only the subsequent changes during the next synchronizations.
- **Step 2: Subsequent synchronization.** Synchronization that occurs after the snapshot is delivered. In this type of synchronization, only the data that corresponds to changes that occurred since the last synchronization is delivered to the destination.

For more information about what replication is and how it works, see SQL Server Replication [\[MSDN-RepMain\]](#).

This document provides a high-level overview of the following items:

- The location where the user data is stored and how to access it.
- The details of the replication topology and how to access it.
- Information about which user changes need to be replicated and how to view them.

1.1 Glossary

This document uses the following terms:

article: A database object, such as a table, view, or stored procedure, that is included in a publication. For more information, see [\[MSDN-RepPub\]](#).

Distributor database instance: A database instance that acts as a store for replication-specific data that is associated with one or more Publishers. For more information, see [\[MSDN-RepPub\]](#).

Log Reader Agent: A replication agent that moves transactions marked for replication from the transaction log on the Publisher to the distribution database.

publication: A collection of one or more articles from one database. For more information, see [\[MSDN-RepPub\]](#).

Publisher database instance: A database instance that makes data available to other locations through replication. A Publisher can have one or more **publications**, each defining a logically related set of objects and data to replicate. For more information, see [\[MSDN-RepPub\]](#).

SQL Server Agent: A replication agent that hosts and schedules the agents used in replication and provides an easy way to run replication agents. For more information, see [\[MSDN-RepAgent\]](#).

Subscriber database instance: A database instance that receives replicated data. A Subscriber can receive data from multiple Publishers and publications. Depending on the type of replication chosen, the Subscriber can also pass data changes back to the Publisher or republish the data to other Subscribers. For more information, see [\[MSDN-RepPub\]](#).

subscription: A request for a copy of a publication to be delivered to a subscriber. For more information, see [\[MSDN-RepPub\]](#).

1.2 References

Links to a document in the Microsoft Open Specifications library point to the correct section in the most recently published version of the referenced document. However, because individual documents in the library are not updated at the same time, the section numbers in the documents may not match. You can confirm the correct section numbering by checking the [Errata](#).

[MSDN-BrwseRepCmd] Microsoft Corporation, "sp_browsereplcmds (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms176109.aspx>

[MSDN-HlpArtcle] Microsoft Corporation, "sp_helparticle (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms187741.aspx>

[MSDN-HlpDist] Microsoft Corporation, "sp_helpdistributor (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms177504.aspx>

[MSDN-HlpMrgeArtcle] Microsoft Corporation, "sp_helpmergearticle (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms174278.aspx>

[MSDN-HlpMrgeFltr] Microsoft Corporation, "sp_helpmergefilter (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms190294.aspx>

[MSDN-HlpMrgePub] Microsoft Corporation, "sp_helpmergepublication (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms189475.aspx>

[MSDN-HlpMrgePullSub] Microsoft Corporation, "sp_helpmergepullsubscription (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms186319.aspx>

[MSDN-HlpMrgeSub] Microsoft Corporation, "sp_helpmergesubscription (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms189437.aspx>

[MSDN-HlpPub] Microsoft Corporation, "sp_helppublication (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms189782.aspx>

[MSDN-HlpPullSub] Microsoft Corporation, "sp_helppullsubscription (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms187714.aspx>

[MSDN-HlpSrvr] Microsoft Corporation, "sp_helpserver (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms189804.aspx>

[MSDN-HlpSubProp] Microsoft Corporation, "sp_helpsubscription_properties (Transact-SQL)", <http://msdn.microsoft.com/en-us/library/ms186254.aspx>

[MSDN-HlpSub] Microsoft Corporation, "sp_helpsubscription (Transact-SQL)",
<http://msdn.microsoft.com/en-us/library/ms190493.aspx>

[MSDN-RepErr] Microsoft Corporation, "Errors and Events Reference (Replication)",
<http://msdn.microsoft.com/en-us/library/ms152467.aspx>

[MSDN-RepMain] Microsoft Corporation, "SQL Server Replication", <http://msdn.microsoft.com/en-us/library/ms151198.aspx>

[MSDN-RepShowCmd] Microsoft Corporation, "sp_replshowcmds (Transact-SQL)",
<http://msdn.microsoft.com/en-us/library/ms175114.aspx>

[MSDN-ShoChnge] Microsoft Corporation, "sp_showpendingchanges (Transact-SQL)",
<http://msdn.microsoft.com/en-us/library/ms186795.aspx>

2 Data Portability Scenario

2.1 Retrieve Intellectual Property from a Replication Topology

This scenario describes extracting the replication topology information from a Microsoft SQL Server database. The information is retrieved by SQL Server stored procedures that are executed from SQL Server Management Studio. To export them, the user can save the result set from SQL Server Management Studio in any format supported by that tool, such as text or CSV.

2.1.1 Data Description

Infrastructure information

Which servers are participating in a replication topology can be found by running the **sp_helpserver** stored procedure [\[MSDN-HlpSrvr\]](#) on the master database of any server.

This stored procedure reports information about a particular remote server or replication server, or about all servers of both types. It provides the server name, the network name of the server, the replication status of the server, the identification number of the server, and the collation name. It also provides time-out values for connecting to, or running queries against, linked servers. Replication status is reported by the stored procedure as follows:

- **pub:** A **Publisher database instance**.
- **dist:** A **Distributor database instance**.
- **sub:** A **Subscriber database instance**.

Source and destination object information

More information about a Distributor is available by running the **sp_helpdistributor** stored procedure [\[MSDN-HlpDist\]](#) at the Publisher on the **publication** database or any database. This stored procedure lists information about the Distributor, distribution database, working directory, and **SQL Server Agent** user account.

A list of which objects are published is available through the list of publications, which can be obtained from the system by executing the following stored procedures on the Publishers:

- **sp_helppublication** [\[MSDN-HlpPub\]](#) for snapshot and transactional replication.
- **sp_helpmergepublication** [\[MSDN-HlpMrgePub\]](#) for merge replication.

To return the list of the objects from the source database that are published and the names of the destination objects, use either of the following stored procedures on the Publishers:

- **sp_helparticle** [\[MSDN-HlpArticle\]](#)
- **sp_helpmergearticle** [\[MSDN-HlpMrgeArticle\]](#)

If a filter has been defined between two **articles** that are participating to a merge publication, the definition of this filter can be found by running the **sp_helpmergefilter** stored procedure [\[MSDN-HlpMrgeFltr\]](#).

The list of **subscriptions** as well as details about the subscriptions can be found by running the following stored procedures:

- **sp_helpsubscription** [\[MSDN-HlpSub\]](#)
- **sp_helppullsubscription** [\[MSDN-HlpPullSub\]](#)

- **sp_helpmergesubscription** [\[MSDN-HlpMrgeSub\]](#)
- **sp_helpmergepullsubscription** [\[MSDN-HlpMrgePullSub\]](#)

Among other things, these stored procedures return the name of the Subscribers and the names of the subscribing databases.

More generic information about a subscription can be found by running the **sp_helpsubscription_properties** stored procedure [\[MSDN-HlpSubProp\]](#).

2.1.2 Format and Protocol Summary

The information that is retrieved from SQL Server is exported in text format.

2.1.3 Data Portability Methodology

A user can extract the data from Management Studio by executing the stored procedures that are shared in this document, and then saving the results to a file on disk.

2.1.3.1 Preconditions

- Servers are required to be online.
- Databases are required to have been created.
- Replication is required to have been set up.
- At least one **publication** is required to have been created.
- At least one valid **article** is required to have been created for this publication.
- At least one **subscription** is required to have been created.

2.1.3.2 Versioning

This scenario applies to the following versions, including released service packs:

- Windows XP operating system
- Windows Server 2003 operating system
- Windows Server 2003 R2 operating system
- Windows Server 2008 operating system
- Windows 7 operating system
- Windows Server 2008 R2 operating system
- Windows 8 operating system
- Windows Server 2012 operating system
- Windows 8.1 operating system
- Windows Server 2012 R2 operating system

2.1.3.3 Error Handling

The stored procedures referred to in this document have built-in error handling and raise specific messages depending on the error conditions. For more information, see [\[MSDN-RepErr\]](#).

2.1.3.4 Coherency Requirements

There are no special coherency requirements.

2.1.3.5 Additional Considerations

In the case of transactional replication, SQL Server propagates transactions. Transactions and their commands are harvested from the transaction log of the published database. The user can view the transactions that are waiting for the **Log Reader Agent** to propagate them to the distribution database by executing the **sp_replshowcmds** stored procedure [\[MSDN-RepShowCmd\]](#).

The actual Transact-SQL commands within those transactions are then stored in the distribution database. The user can view the list of those commands and transactions by executing the **sp_browsereplcmds** stored procedure [\[MSDN-BrwseRepCmd\]](#).

In the case of merge replication, SQL Server does not propagate the commands of a transaction but replicates net changes. In merge replication, each row of any table participating in merge replication is identified by a **uniqueidentifier**. To determine which changes are required to be replicated, the user can execute the **sp_showpendingchanges** stored procedure [\[MSDN-ShoChnge\]](#).

By using the rowguids (**uniqueidentifiers**) that are returned by this procedure, the user can query the user table to identify which row is required to be replicated.

3 Change Tracking

This section identifies changes that were made to this document since the last release. Changes are classified as Major, Minor, or None.

The revision class **Major** means that the technical content in the document was significantly revised. Major changes affect protocol interoperability or implementation. Examples of major changes are:

- A document revision that incorporates changes to interoperability requirements.
- A document revision that captures changes to protocol functionality.

The revision class **Minor** means that the meaning of the technical content was clarified. Minor changes do not affect protocol interoperability or implementation. Examples of minor changes are updates to clarify ambiguity at the sentence, paragraph, or table level.

The revision class **None** means that no new technical changes were introduced. Minor editorial and formatting changes may have been made, but the relevant technical content is identical to the last released version.

The changes made to this document are listed in the following table. For more information, please contact dochelp@microsoft.com.

Section	Description	Revision class
1 Introduction	Clarified that replication is not limited to the use of a specific version of SQL Server.	Minor

4 Index

C

[Change tracking](#) 10

G

[Glossary](#) 4

I

[Informative references](#) 5

[Introduction](#) 4

R

[References](#) 5

T

[Tracking changes](#) 10