SQL Server Replication

Ing. Marek Chmel, MSc.
MVP: Data Platform | MCSE: Data Platform | MCT | CEH
marek@datascienceteam.cz
@MarekChmel
Replication Introduction

- Replication allows you to copy and distribute data and objects from one database to another and then continually stream data modifications as they are made.
Types of Replication

- **Snapshot Replication**
  - Distributes data exactly as of a point in time

- **Useful for:**
  - When data changes infrequently
  - Subscriber data allowed to be out-of-date compared to the Publisher
  - Replicating small volumes of data
  - When a large volume of changes occurs over a short period of time

- Commonly used to provide the initial data set for transactional and merge replication
Types of Replication

- Merge Replication
  - Primarily designed for mobile or distributed server scenarios that have possible data conflicts

- Useful for:
  - Exchanging data with mobile users
  - Consumer point-of-sale (POS) applications
  - Integration of data from multiple sites
Types of Replication

- **Transactional Replication**
  - Used in server-to-server scenarios that require high throughput

- Useful for:
  - Improving scalability and availability
  - Data warehousing and reporting
  - Integrating data from multiple sites or heterogeneous data
  - Offloading batch processing
Types of Replication

- Peer-to-Peer Replication
  - Enterprise Edition feature
  - Propagates transactionally consistent changes in near real-time across multiple SQL Server instances
  - Can support bi-directional replication of writes (multi-master)
    - Requires data partitioning and centralized writes are easier to manage
    - Typical solution involves a composite key that includes a value unique to each topology node
- Conflict detection was added in SQL Server 2008
Terminology

- **Publisher**
  - Database instance that makes data available to other locations through replication

- **Distributor**
  - Database instance that acts as a store for replication specific data associated with one or more Publishers. Distribution database stores replication status data, metadata about the publication

- **Subscriber**
  - Database instance that receives replicated data

- **Publication**
  - Collection of one or more articles from one database

- **Article**
  - Database object that is included in a publication

- **Subscription**
  - Copy of a publication that is delivered to a Subscriber
Key Components

- **Snapshot Agent**
  - Prepares snapshot files containing schema and data of publication articles
  - Stores the files in the snapshot folder

- **Log Reader Agent**
  - Analyzes the transaction log of the publication database and looks for transactions marked for replication
  - Copies the transactions marked for replication into the distribution database

- **Distribution Agent**
  - Copies the initial snapshot files from the snapshot folder
  - Copies transactions from the distribution database to one or more subscribers
HANDS ON!
Q&A

Ing. Marek Chmel, MSc.
MVP: Data Platform | MCSE: Data Platform | MCT | CEH
marek@datascienceteam.cz
@MarekChmel