DADDS
(DCS) Administrative and Data Distribution System
Operations & Maintenance Training

PREPARED FOR

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AN ATMOSPHERIC ADMINISTRATION
NATIONAL ENVIRONMENTAL SATELLITE, DATA AND INFORMATION SERVICE

March 2019

MICROCOM DESIGN, INC.
10948 BEAVER DAM ROAD
HUNT VALLEY, MD 21030
410-771-1070 Fax: 410-771-0018
Support@MicrocomDesign.com

Database Training Topics

1. Databases
2. SQL Manager
3. Replication
4. Database Jobs

Databases

Review

Databases - Overview

- DADDS uses Microsoft SQL Server 2016 for database servers, management and replication.
- The DADDS real-time and web applications interact with server databases to process and disseminate DCS message data or manage the DCS system.
- Due to the distributed nature of DADDS, with processing spread over multiple servers, databases share data for various purposes using Microsoft SQL replication technology.
- Most functions, once setup, are automated and left to run, though periodic status checks are recommended.
- Operators will not be expected to work with DADDS databases but they should be familiar with the topics.

Databases - Types

- DADDS utilizes two main database types to separate management and DCP message data, and a sub-type of the management database containing PII (personal identifying information) used on archive servers.
- Both types are used by both the DADDS real-time and web applications in different roles.
- DaddsMgmtCS2 contains data for platforms, users, channels, groups, ect. used to manage DCS as a whole, and process received messages
- DaddsMsgsCS2 contains DCS message data for processing, dissemination and web queries and data download.

Databases - Servers

- Real-time servers have two databases:
  - DaddsMgmtCS2 for DCS message processing.
  - DaddsMsgsCS2 for DADDS process message storage.
- Databases used by the DADDS application to receive, process and disseminate DCS message data.
- Message and statistics data are copied to respective archive server database instance.
- Archive servers have two databases:
  - DaddsMgmtCS2 for DCS system management.
  - DaddsMsgsCS2 for website DCS message archive.
- Databases used by the web application to manage the DCS system and query for, and download message data.

Databases - Instances

- The DADDS system contains a total of 16 databases, deployed to 8 servers over 2 locations; WCDA & NSOF
- Data is shared between databases by extensive use of Microsoft SQL replication technology.
- WCDA Database Instances
  - WRTA: DaddsMgmtCS2, DaddsMsgsCS2
  - WRTB: DaddsMgmtCS2, DaddsMsgsCS2
  - WARCA: DaddsMgmtCS2_PII, DaddsMsgsCS2
  - WARCB: DaddsMgmtCS2_PII, DaddsMsgsCS2
- NSOF Database Instances
  - NRTA: DaddsMgmtCS2, DaddsMsgsCS2
  - NRTB: DaddsMgmtCS2, DaddsMsgsCS2
  - NARCA: DaddsMgmtCS2_PII, DaddsMsgsCS2
  - NARCB: DaddsMgmtCS2_PII, DaddsMsgsCS2

SQL Manager

Review

SQL Manager - Introduction

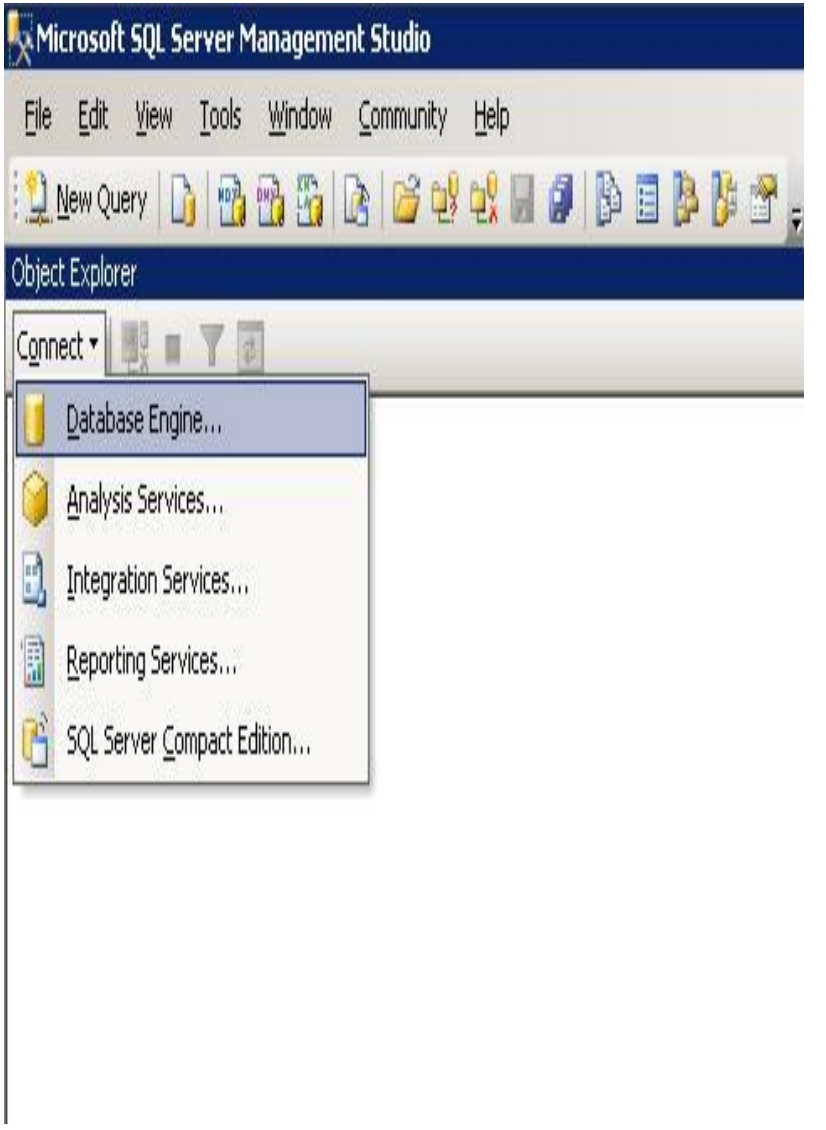
- SQL Manager is used to perform database operations:
  - Analysis
  - Maintenance
  - Configuration
- Training goals include:
  - Introduce operators to SQL Manager
  - Review replication and monitoring
  - Review database jobs
- Operators will not normally use SQL Manager but they should be aware of its existence and purpose.

SQL Manager - Opening

- Shortcuts to SQL Management Studio are found on the desktop and quick launch bar of each DADDS database server instance
- In the event a shortcut is not present, SQL Management Studio can be accessed from the Start menu.
  - Start -> Programs -> Development -> SQL Microsoft Server 2017
- After opening SQL Manager, a database connection dialog will automatically be opened.

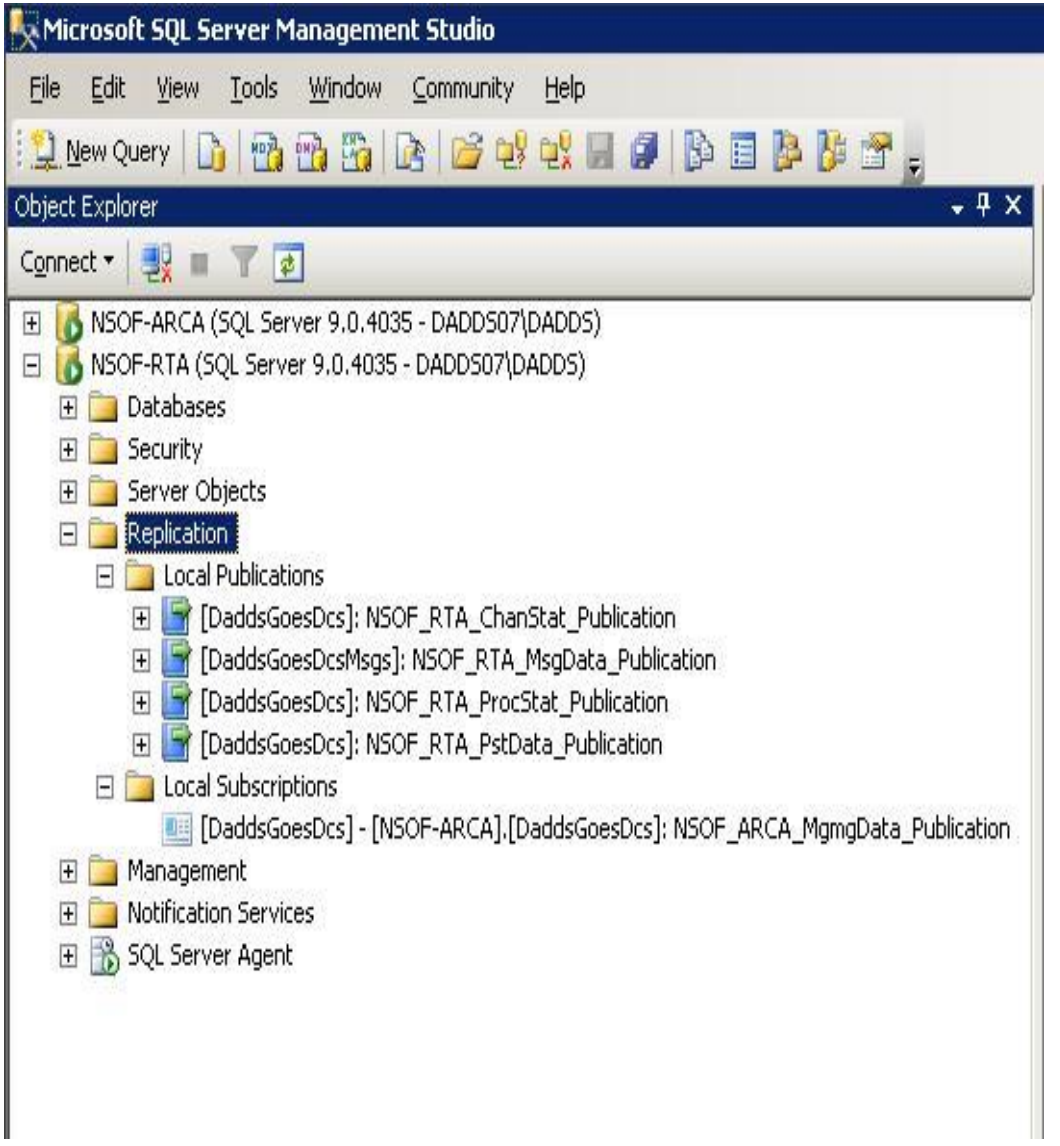
SQL Manager - Connections

- After opening SQL Manager, a database connection must be made.
- In the ‘Server Name’ field enter the name of the database instance being connected:
  - Server IP Address
  - Logical Server Name
- Default authentication method will always be set to ‘Windows Authentication’
- Change to method to ‘SQL Authentication’ and enter appropriate account information.
- Click ‘Connect’ button to initiate a connection to the SQL instance.



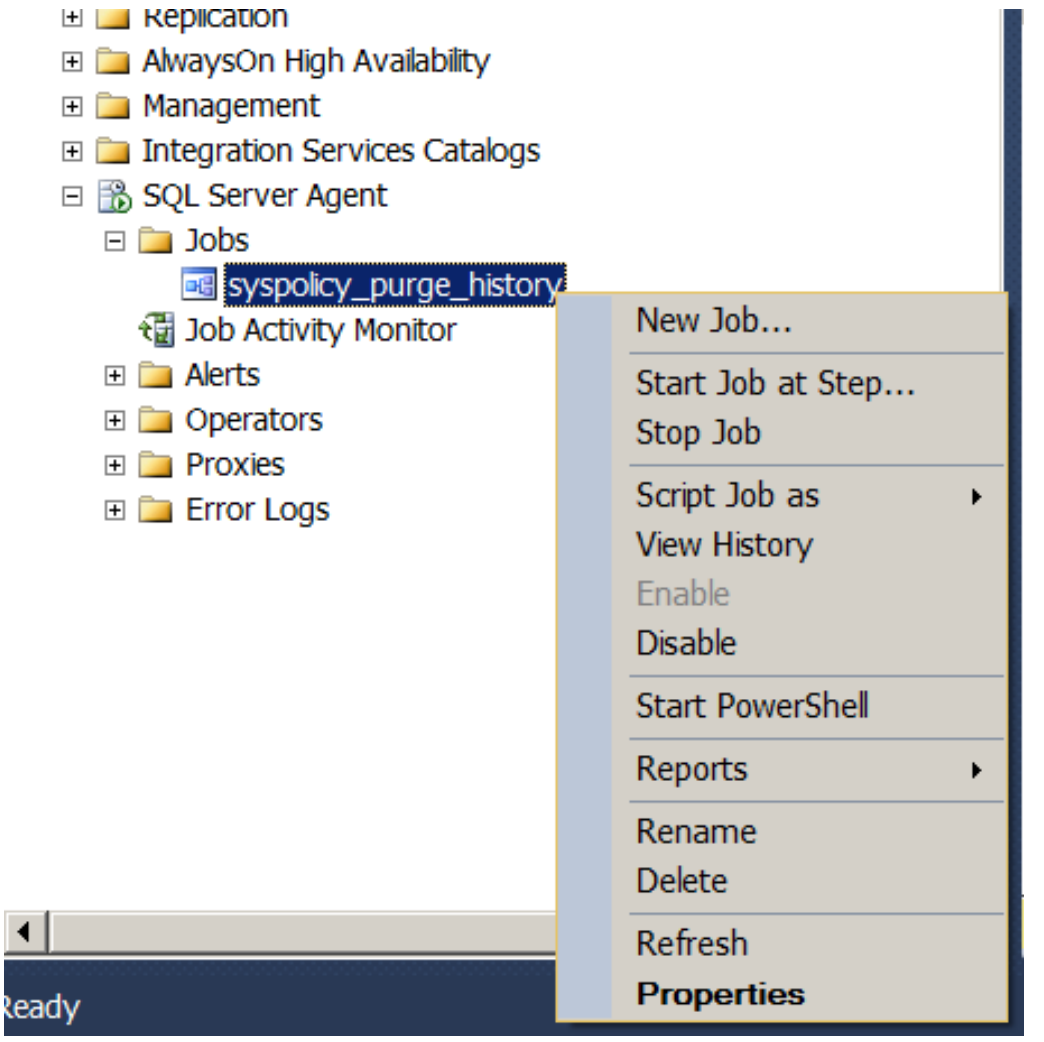
SQL Manager - Replication

- Contains all replication links configured on an SQL database node.
- ‘Local Publications’ folder contains all replication ‘senders’ configured on an SQL node.
- Individual publications can be expanded to show subscribers to the publication data.
- ‘Local Subscriptions’ contains all replication ‘receivers’ configured on an SQL node.
- To open the replication monitor, right click the ‘Replication’ folder and select the monitor option from the popup list.



SQL Manager - Jobs

- The ‘Jobs folder contains the database jobs defined on a server.
- Located in the SQL Server Agent section of the object explorer.
- Right clicking a job displays a popup list with various options.
  - Properties
  - Start / Stop
  - Enable / Disable
  - Rename
  - Delete
- Selecting ‘View History’ opens a dialog containing the job status and execution history.



Replication

Review & Monitoring

Replication - Introduction

- SQL server replication is designed to copy data from one SQL instance to another.
- It allows database developers to share data between any number of databases in a system architecture.
- The tasks of copying and distributing data is automated once a replication has been configured.
- Both tasks are carried out via internal SQL processing allowing replication to be both stable and recoverable, should an error occur.
- In error situations where data fails to replicate, SQL server continues to prepare data for distribution and back logged data is sent once the error is resolved.

Replication - Terminology

- Replication

  Transfer of data from one database node to another.

- Publisher

  Source database of the data to be transferred via replication

- Publication

  Subset of data from the source database, the sending side of a replication link

- Subscriber

  Destination database of the data to be transferred via replication.

- Subscription

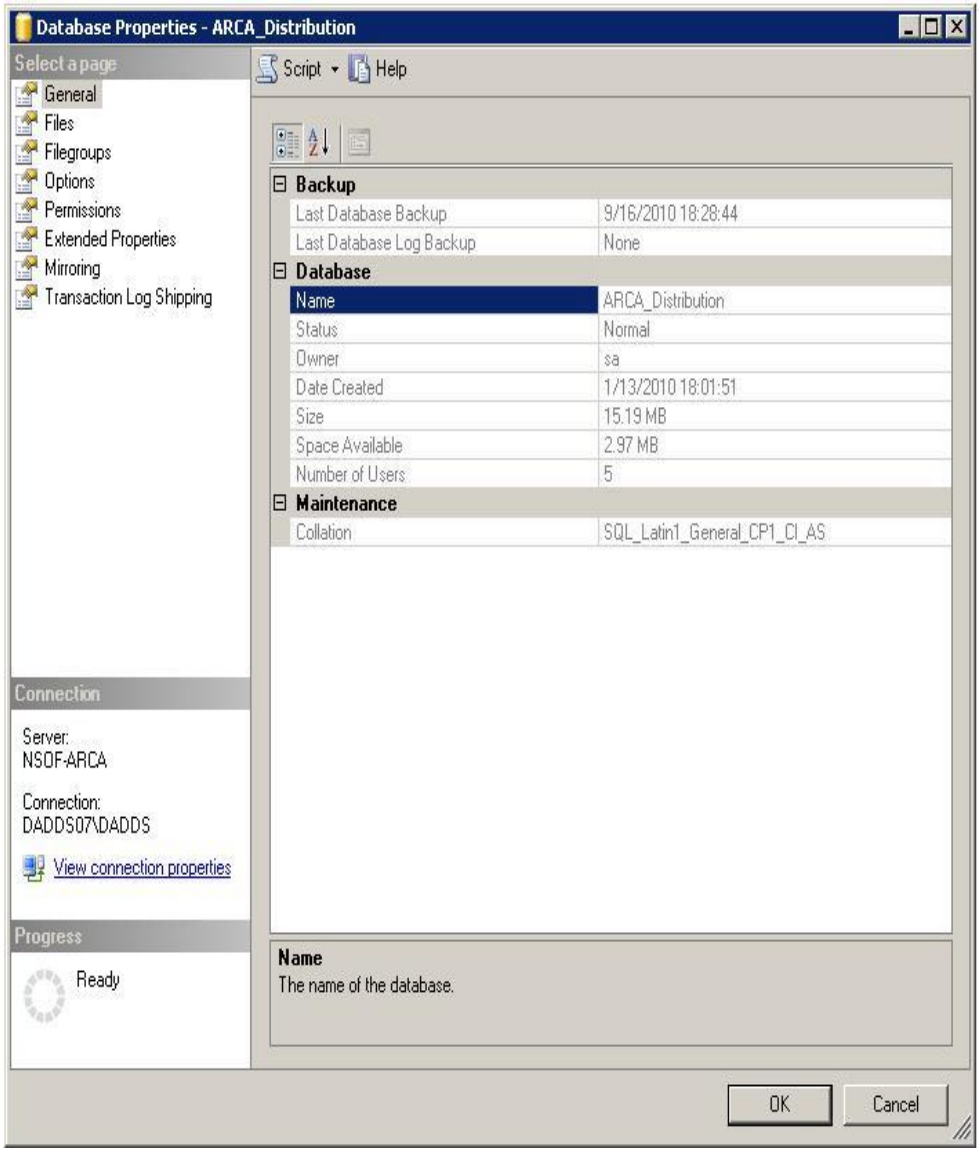
  Receiving side of a replication link

- Distributer

  Database that mediates the transfer of data between the publisher (source) and subscriber (destination) databases.

Replication - Distribution

- The distribution component of replication is responsible for temporarily storing changes to a published table, and sending them to subscribing databases.
- The distribution database contains changes to a source database that are waiting to be sent to the destination database.
- Distribution health, specifically database size, are a good indicator of replication efficiency and health.
- The ‘Size’ property on the ‘General’ properties page should have a value less than 50MB.
- A very large distribution database size can indicate:
  - Replication links are down
  - SQL jobs are not running
  - Expired subscriptions
  - Data consistency errors



Replication - Publications

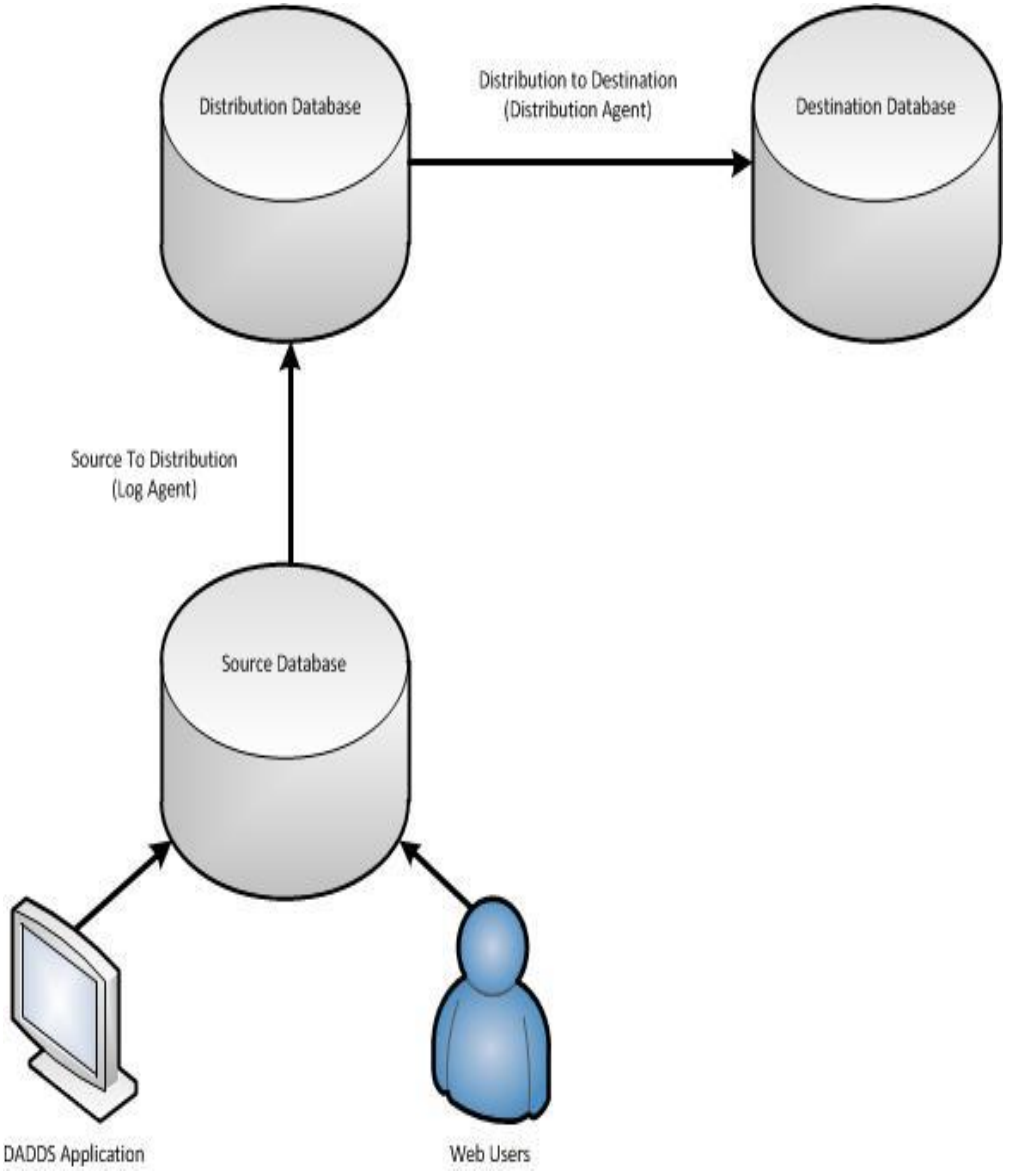
- Publishers are systems that provide data to be replicated, and consist of and SQL server instance with the database containing data to be replicated.
- Keep track of what data has changed in the database so it can be replicated to subscribers.
- Created to ‘publish’ data to other databases via subscriptions to the publication.
- Defined as a set of database tables that describe a replication link.
- Once created, publications and their data are configured and maintained for distribution to other systems automatically by SQL server.
-

Replication - Subscriptions

- Subscribers are database servers that store replicated data and receive updates from publishers.
- A subscription is connected to a publication and receives changes made to the tables it contains.
- Multiple subscriptions can exist on a database node acting as a subscriber.
- Subscriptions define what published data is sent to the subscriber by a publication, and how and when it is transferred.
- Once created, subscriptions handle the replication of data to the destination database via an SQL job dedicated to the task.

Replication – Link Detail

- A single replication link involves three separate databases:
  - Source Database
  - Distribution Database
  - Destination Database
- Web users or the DADDS application modifies the source database. SQL processes transfer changes from the source to the distribution database.
- Replication database job for the subscription copies changes stored in the distribution database to the destination database and marks them as ‘sent’.



Replication – Real-time Links

- Real-time publications replicate data generated at the real-time servers by the DADDS application.
- Subscriptions to the publications facilitate the transfer of data to archive server databases for use by the DADDS website.
- Three sets of data are replicated to archive server databases in separate publications:
  - **MsgData:**

    Complete DCP message stream generated and stored by the message processor for dissemination.
  - **PstData:**

    PST data contains the status of all platforms (active/inactive) in the system, and their last report times.
  - **StatData:**

    Statistics data contains a historical record of real-time process message counts for various types, and message counts on DCS channels.

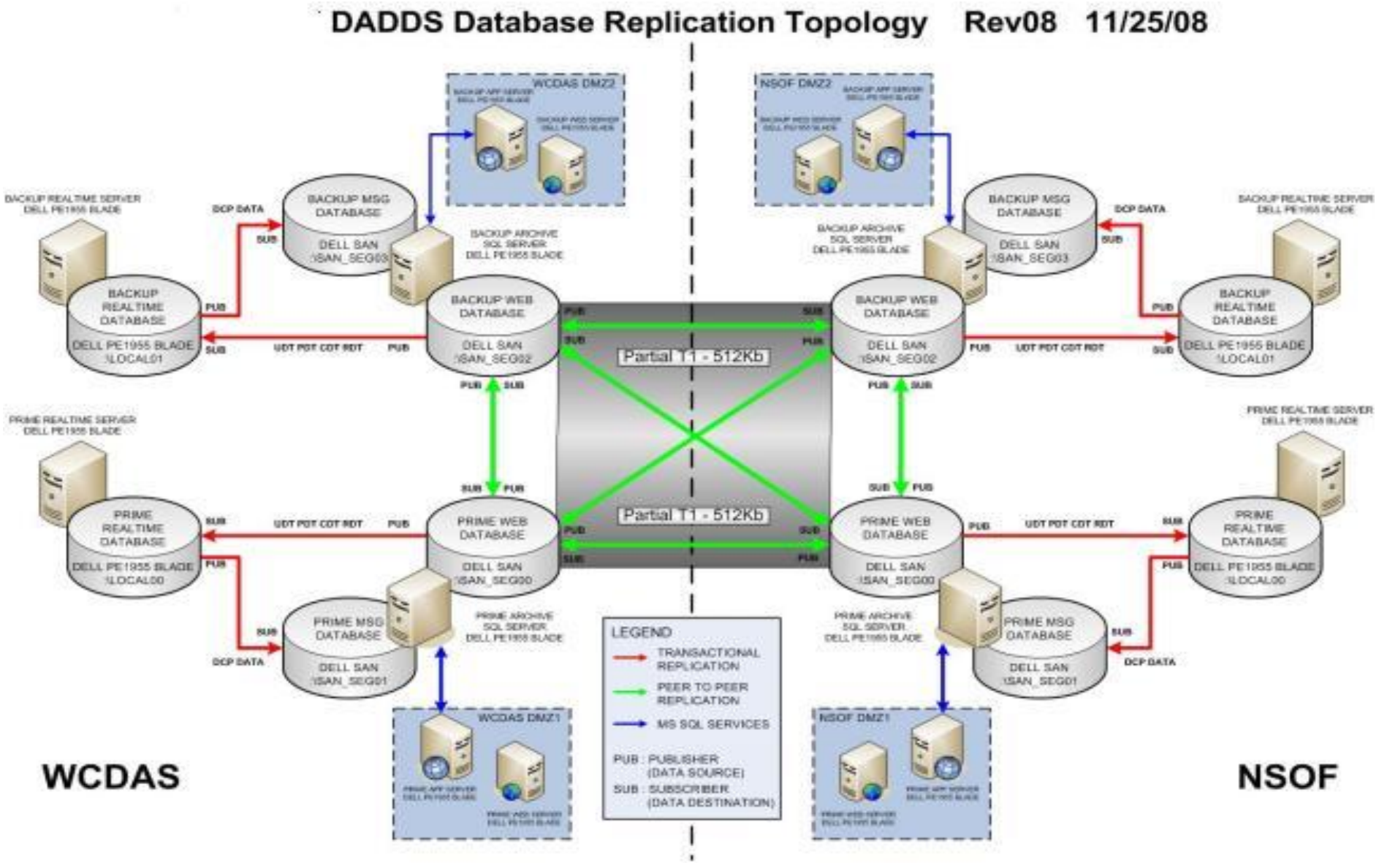
Replication – Archive Links

- Archive publications replicate data generated at the archive servers by the DADDS web application.
- Subscriptions to the publications facilitate the transfer of management data to real-time server databases and peer data to other archive server database nodes.
- Two sets of data are replicated to real-time and archive server databases in separate publications:
  - **MgmtData:**

    Management data is defined as the PDT, CDT, RDT and Group tables in the database. Changes to these records are applied to real-time databases for use by the DADDS application when processing messages.
  - **PeerData:**

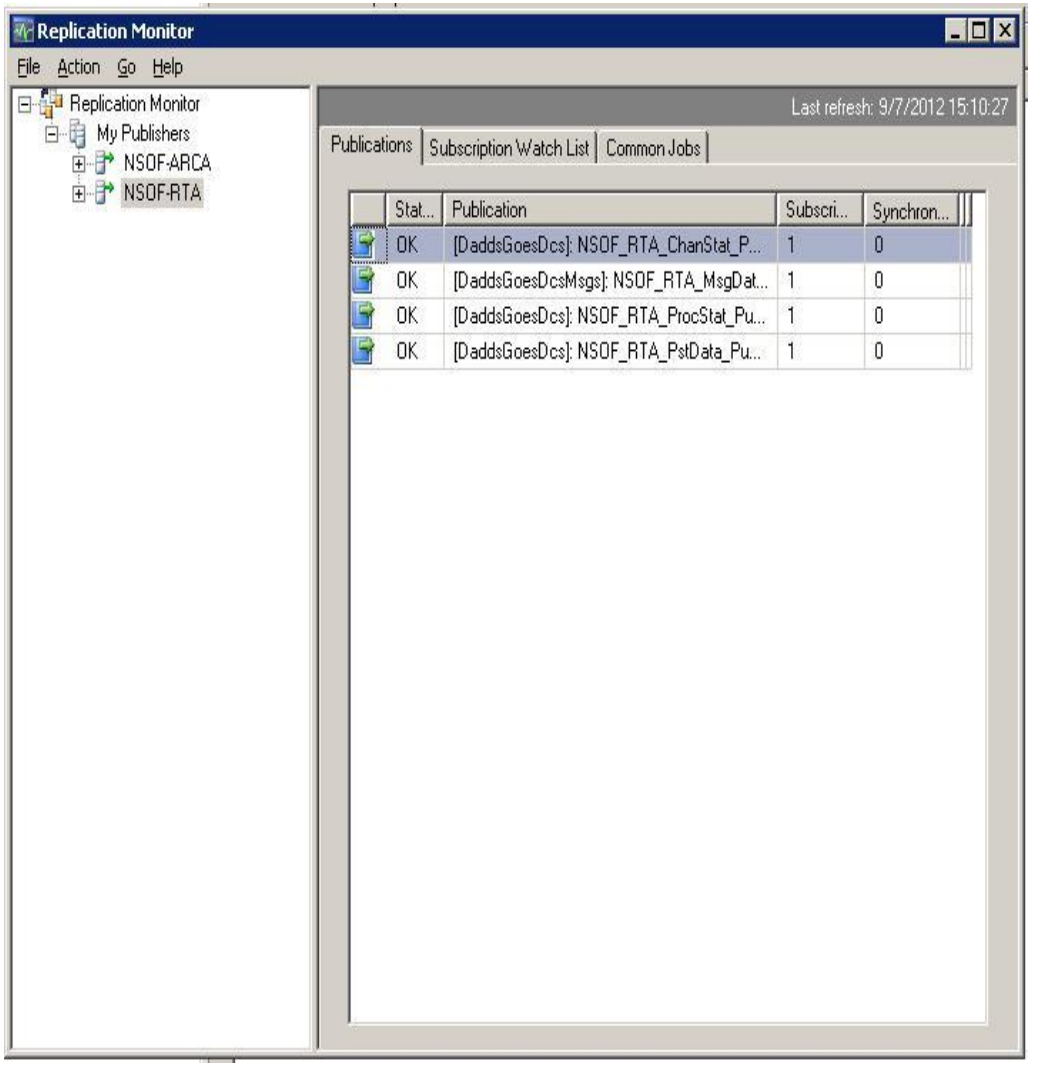
    Peer data consists of management, user, system use and other tables used by the DADDS web application. It must be shared and synchronized between all four DADDS archive server database nodes so that data shown on all DCS websites match.

Replication - Architecture



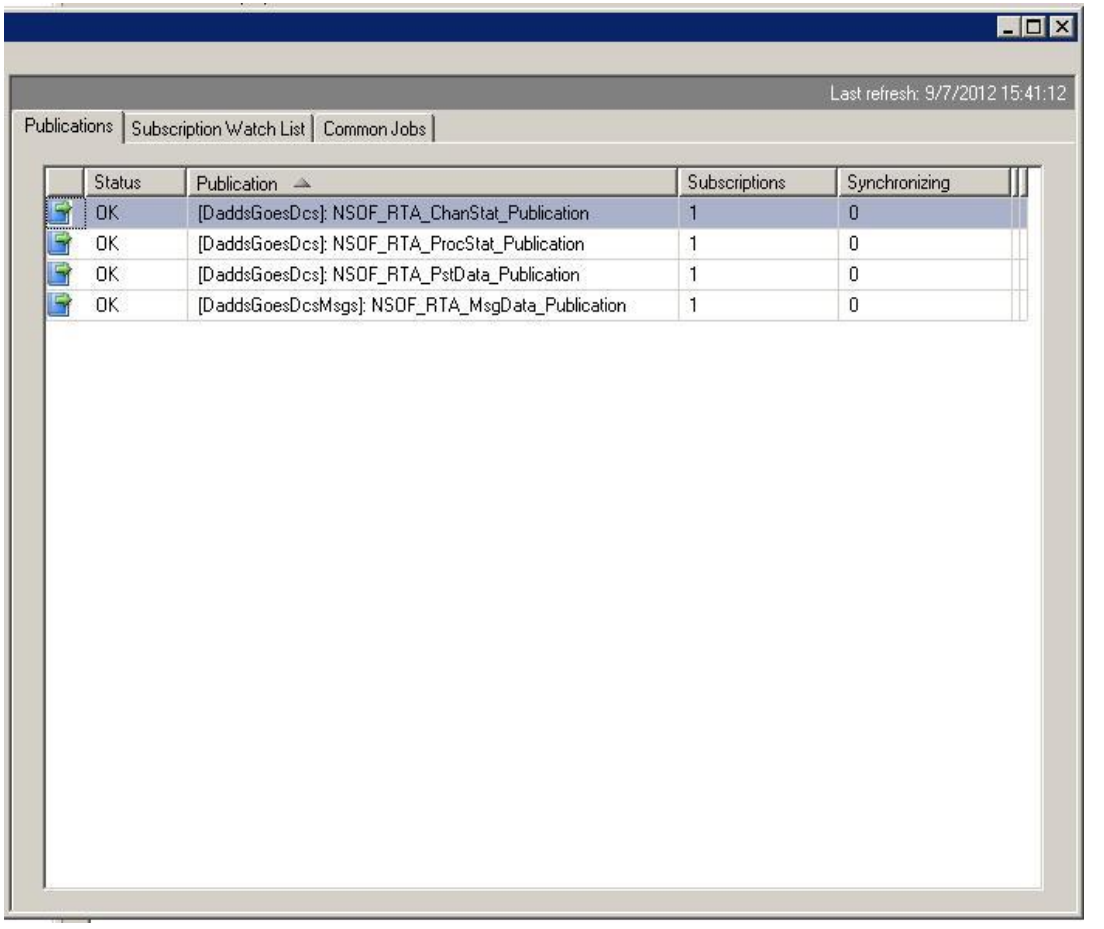
Replication - Monitoring Utility

- SQL utility used to efficiently monitor and manage replication links
- SQL server instances currently loaded on the monitor are found in the left window.
- Detailed information concerning the configuration and status of replication on a given server are located in the right window.



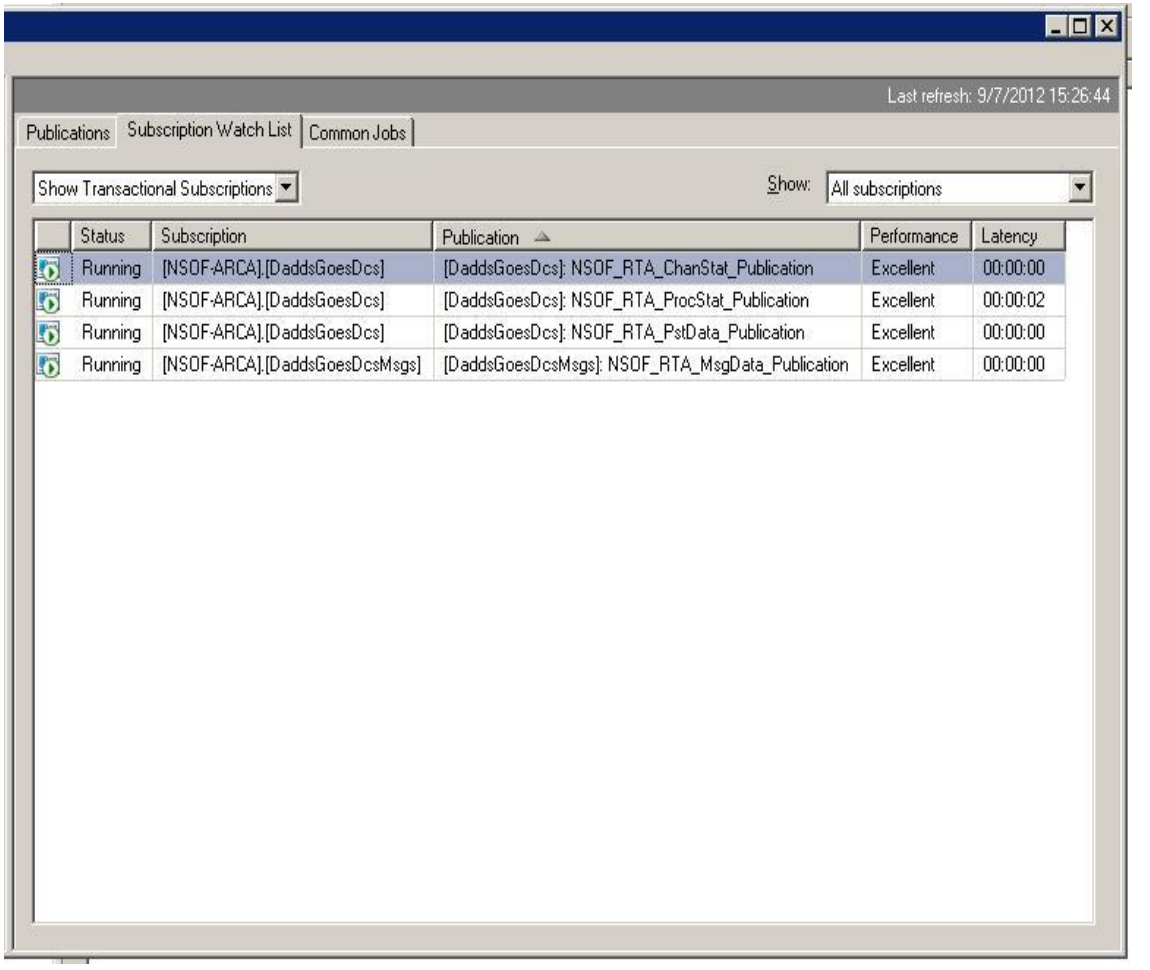
Replication- Publications Tab

- Found on the first tab of the right window of the monitor utility.
- Displays information and status on all publications (data senders) that exist on an SQL server instance.
- Publication column displays the sending server, database and associated link data in the form:
  - [DATABASE].[SERVER]_[DATA]
- Status column displays the health of the publication (data sender), and will always be ‘OK’ for a properly operating link.



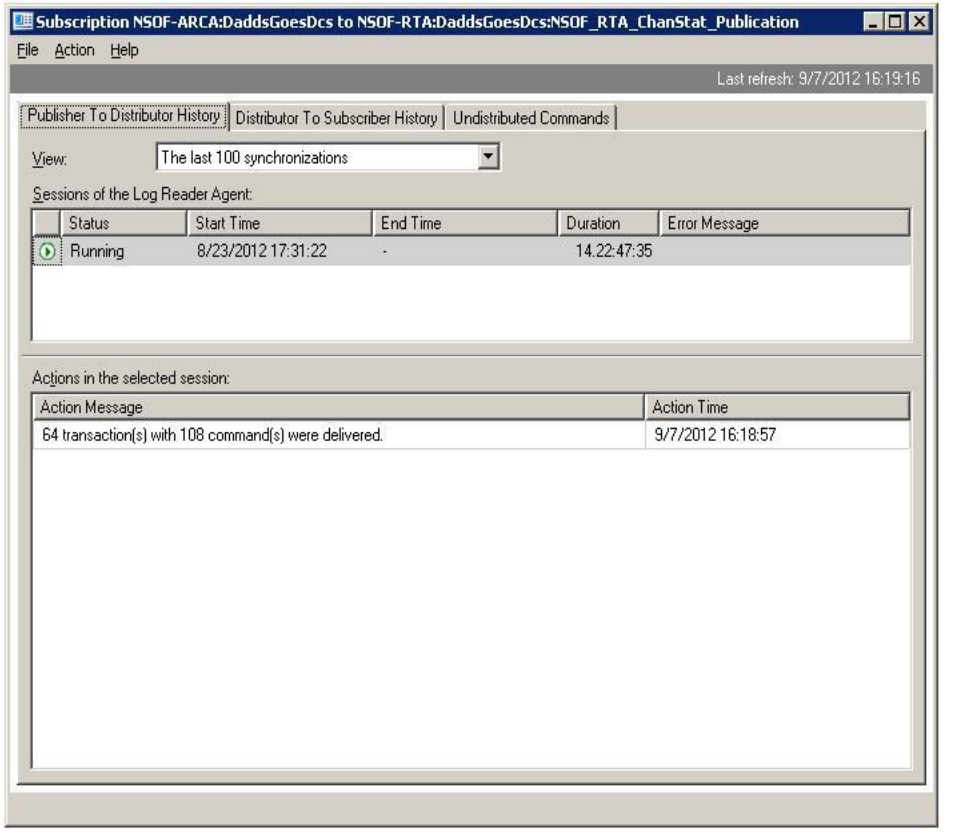
Replication - Subscriptions Tab

- Found on the second tab of the right window of the monitor utility.
- Lists information and status on all subscriptions that exist on an SQL server instance.
- Subscription and publication columns identify the source and destination of replicated data.
- Status, Performance and Latency columns reflect link health.
  - Status column will always show ‘Running’ for a properly working link.
  - Performance values other than ‘Good’ and ‘Excellent’ indicates problems on a link.
  - Latency column displays the average time for a replication cycle to complete. High latency times indicates problems on a link.



Replication - Subscriptions Details

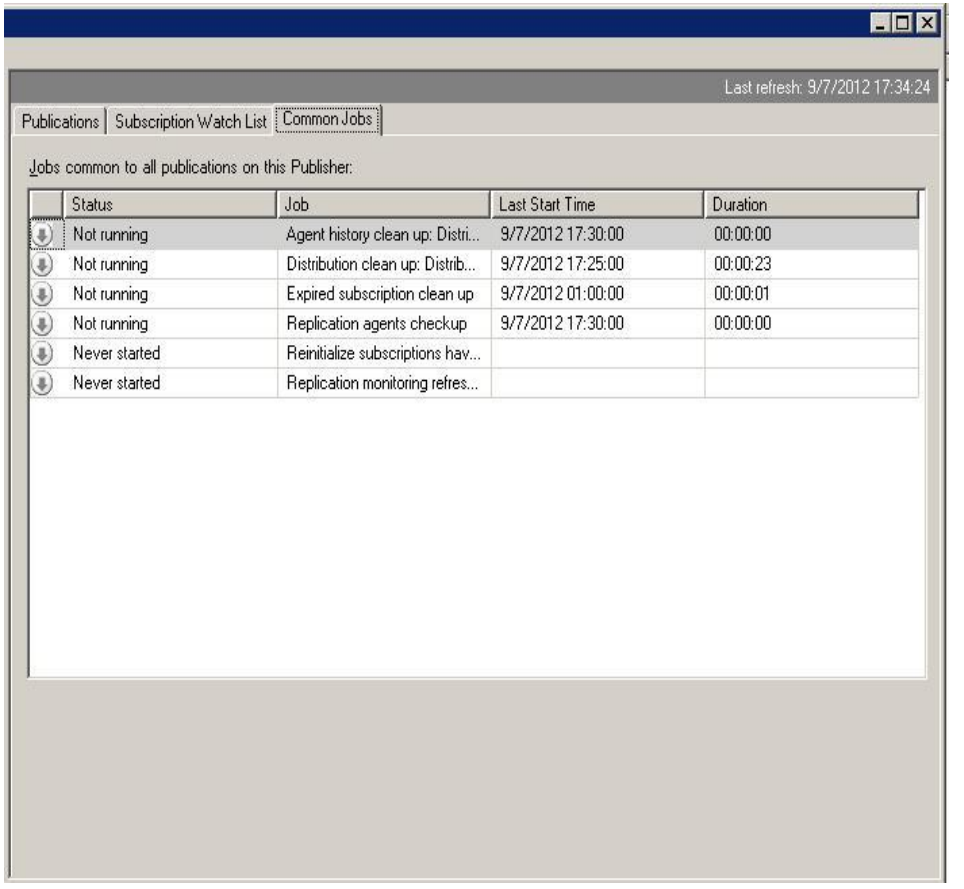
- To open, right click the desired subscription from the list and select ‘Details’.
- Publisher to Distributor History provides information concerning the transfer of data from the source database to the distribution database for replication.
- Distributor to Subscriber History provides information concerning the transfer of data from the distribution database to the receiving database
- Undistributed Commands tab displays the number of commands in the distribution database that are waiting to be applied, and the estimated time for that to occur.



- ‘View’ drop down box contains different filter options for the link history
  - ‘Synchronizations with Errors’ options provides a quick look at any unresolved link errors
  - Remaining view options are time based
- ‘Status’ column of the top grid will display link health
  - ‘Running’ if SQL replication is actively sending data for distribution
  - ‘Error’ if SQL replication can not send data for distribution

Replication - Common Jobs

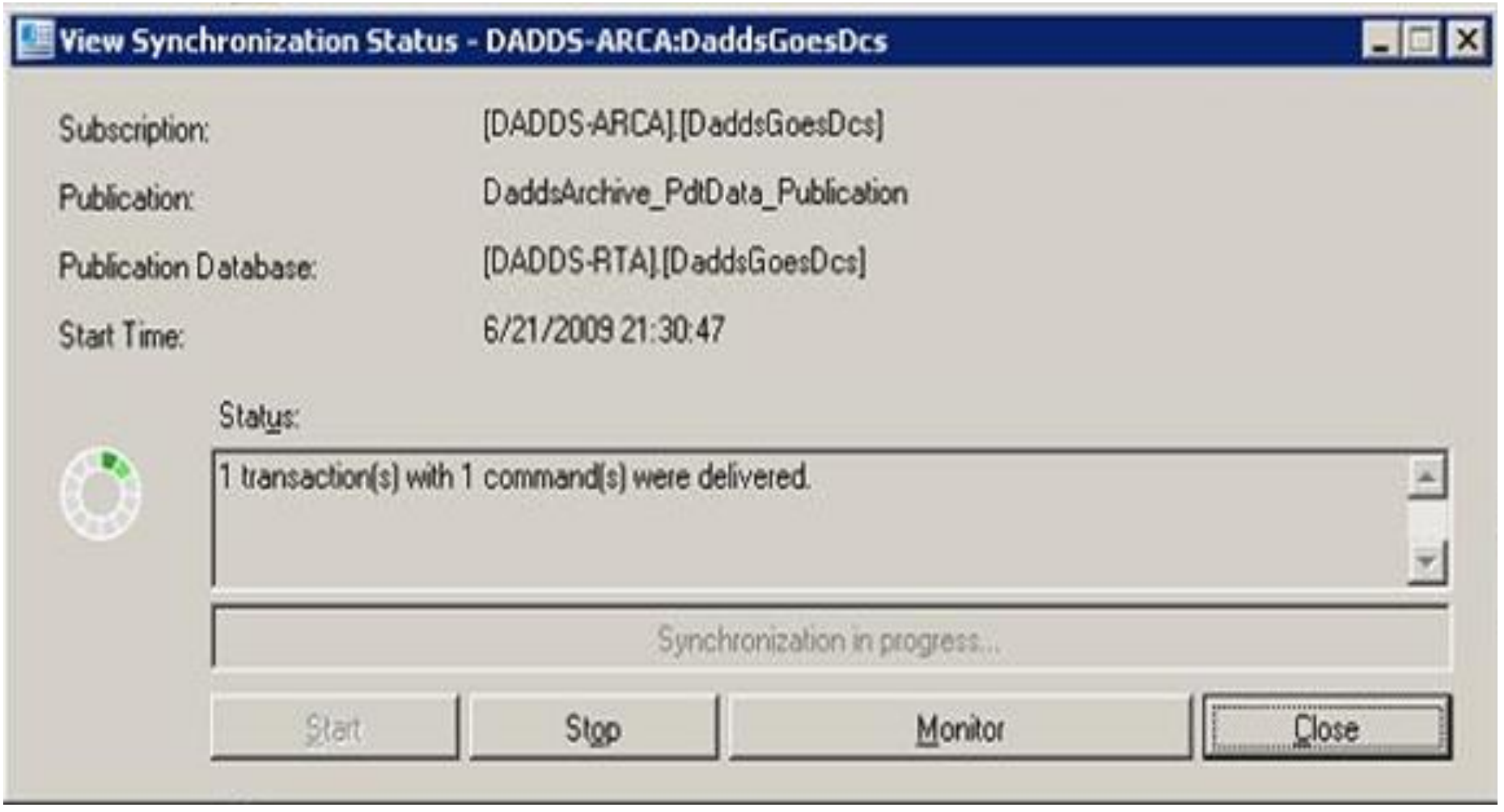
- Replication jobs are SQL processes that perform tasks related to replication clean up and reporting.
- Tab contains information concerning the SQL processes:
  - ‘Status’ will indicates if a job is active or idle, and reports ‘Error’ if the previous execution failed.
  - The ‘Last Start Time’ is the most recent execution of the process.
  - ‘Duration’ is how long the process took to execute and should always be less than 30 seconds.
  - To view history, right click the job and select ‘View History’ from the popup list.



Replication - Subscription Status

- An alternative to using the replication monitoring utility involves manually checking the status of each subscription.
- While this method is more time consuming, it is a more direct indication of the subscription status, whereas the monitoring utility must be continually refreshed so that data does not become stale.
- To check subscription status manually, drill down into a publication and right click the desired subscription. Select ‘View Synchronization Status’ from the popup list to open the synchronization status dialog.

Replication - Subscription Status



Database Jobs

Review

Database Jobs - Overview

- Database jobs are SQL scripts or maintenance tasks ran by the SQL agent process.
- Can be triggered by some condition, execute continuously, or execute on a predefined schedule to perform an SQL operation.
- Multiple steps, and corresponding logic to advance between them, can be defined within a single job, allowing for advanced automation capabilities.
- Within DADDS, database jobs are used extensively to perform database maintenance, backups and replication.

Database Jobs - Prune

- The database prune job trims old data from the message and statistics data tables so they do not grow without bounds and affect performance.
- Runs every minute and removes a maximum of 2,500 records for each execution.
- Message data is pruned differently on real-time and archive server databases:
  - 2 day depth on real-time databases to ensure DADDS application processing is fast and efficient.
  - 30 day depth on archive databases to provide users with a month's worth of DCS messages that can be searched and downloaded via the DADDS web application.
- Process and channel statistics data are pruned at 60 days on both the real-time and archive server databases.

Database Jobs - Backup

- Database backups protect the DADDS system from failures and data corruption by ensuring that recent copies of management databases are available for restoration in emergency situations.
- The job creates database backups for DaddsMgmtCS2 on both real-time and archive servers on a daily basis.
- Can be executed more frequently if required by NOAA; DADDS Backup & Restore manual describes more advanced backup strategies and plans.
- Has the ability to store backup files locally or on a remote server dedicated to storing the backups. Currently backup files are stored locally on the servers.

Database Jobs - Rebuild Indexes

- Rebuilding indexes ensures that operations and queries on a database table remain efficient.
- For large sets of data, such as days worth of DCP messages, rebuilding indexes to re-organize the tables layout in memory becomes very important.
- While indexes are used throughout the DADDS databases, they are only regularly rebuilt on DCP message data tables.
- Archive servers rebuild indexes on TblDcpData01 used by the web application for message data.
- Real-time servers rebuild indexes on TblDcpData01 and TblDcpData02 for the Message processor and DAMS-NT clients respectively.
- Executed daily at 3:00AM to prevent slowing down the system during normal use.

Database Jobs - Replication

- Replication uses database jobs to perform the actual functions in the replication system.
- Created automatically by SQL server when configuring publications and subscriptions.
- Replication links consist of three SQL database jobs:
  - **Snapshot Job**

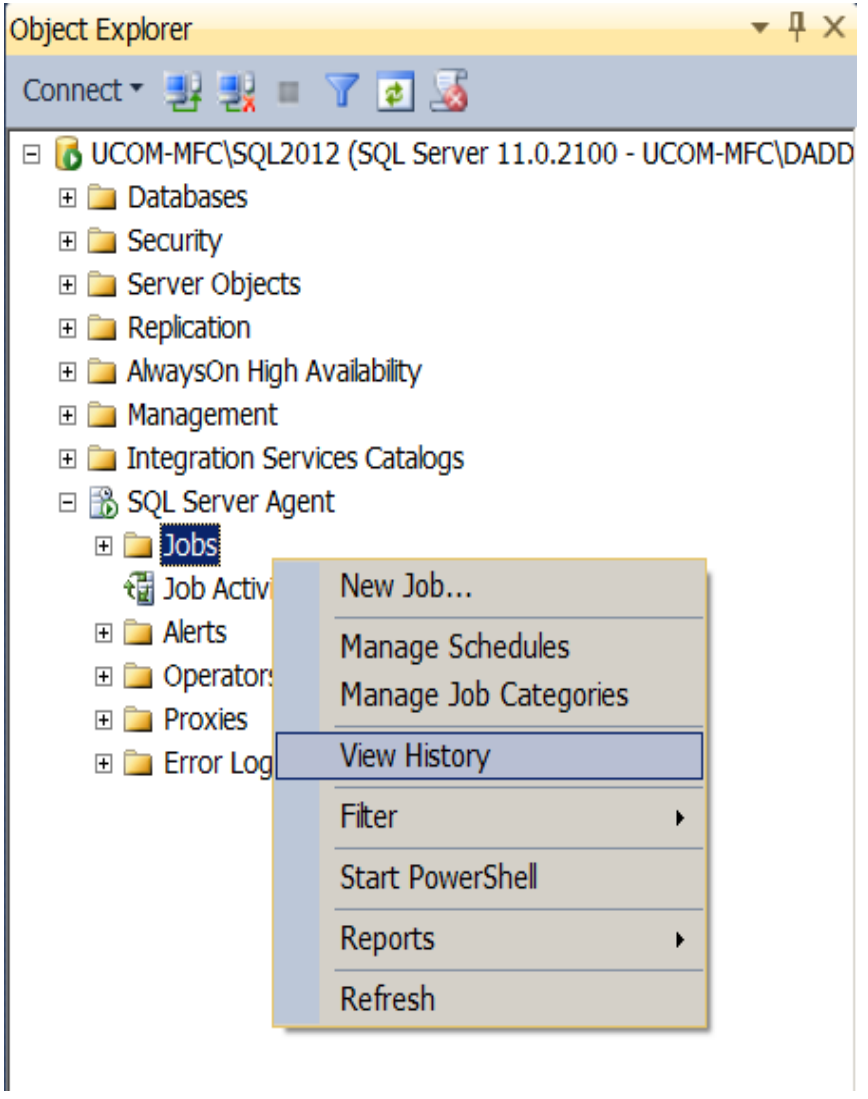
    Records a copy of an publication's entire data set for use when initializing subscriptions. This ensures that the publisher and subscriber have the exact same copy of data before starting replication
  - **Log Reader Job**

    Created by the publisher to monitor the changes made to a published database, and copy them to the distribution database for replication to subscribers.
  - **Subscription Job**

    Replicate data from the distribution database to the destination database of the subscription. Each subscription to a publication will have its own associated job to transfer data.

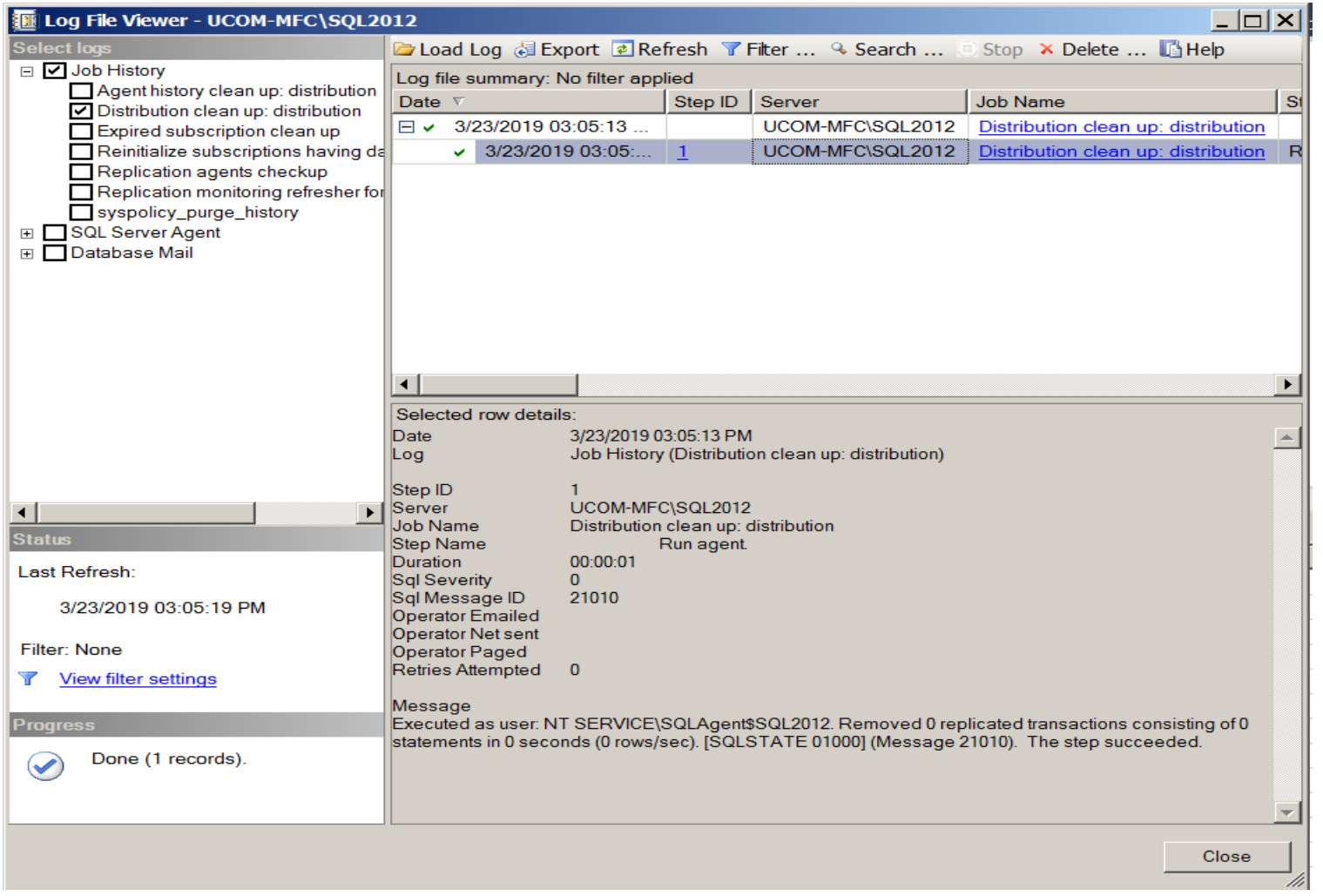
Database Jobs - History

- Each database job retains a history of it's previous executions.
- Execution records contain job result messages, run times and any error information.
- To open the history for all jobs, right click the 'Jobs' folder and select 'View History' from the popup list.
- To open the history for a single job, drill down into the jobs folder, right click the desired job and select 'View History' from the popup list.



Database Jobs - Status

- View history dialog contains the job status and most recent execution's results



- End Database Training -