

Microsoft Modernization Practice

Modernization Viability Assessment

Data Collection Instructions

Updated: 09/17/2025

Table of Contents

Microsoft Modernization Practice	1
Summary	3
Initial Setup	3
Editing the Config File	3
Authentication	4
Required Permissions	5
External Downloads	5
Pre-Staging the DacPac Export Tool	5
Collecting the Data	6
One Time Collection	6
Recurring Collection	7
Multi Region Collection	8
Multi Account Collection	8
Reporting and ETL	8
How to Share Results	8
Environment Clean Up	9
Manual CloudWatch Data Export	10
Data Collection Period	10
Exporting CloudWatch Metrics	10
Additional Resource Information	10
EC2 Metrics	11
EBS Metrics	12
RDS Instances	12
Zip File/Repo Content	14

Summary

This PowerShell script will scan your SQL Server instances (and associated storage) and CloudWatch metrics to collect data points on general configurations and performance statistics and export that into CSV files.

Initial Setup

1. These instructions assume Git is installed, see here on [how to install Git](#)
 - a. If required a ZIP file of these files can be emailed.
2. Open <https://github.com/evolvecloudservicesgit/mva-data-collection>
 - a. Click the green **Code** button
 - b. Copy the **HTTPS** or **SSH** URL
 - i. HTTPS looks like: `https://github.com/username/repo.git`
 - ii. SSH looks like: `git@github.com:username/repo.git`
3. Open a Terminal or Command Prompt, navigate to the folder where you want the repository to be cloned
4. Run the Clone Command
 - a. HTTPS Example


```
git clone https://github.com/evolvecloudservicesgit/mva-data-collection.git
```
 - b. SSH Example


```
git clone git@github.com:evolvecloudservicesgit/mva-data-collection.git
```
5. Navigate to the newly created repository directory `\mva-data-collection`
6. The `config.txt` file will need to be edited to include the resources in scope for this data collection. See section “Editing the config file” for details.

Editing the Config File

The config file is loaded by the PowerShell script to provide the resources in scope for the data collection. Each parameter in the file is of the format “name:value;”.

CustomerName

- This should be your company name or initials with no spaces.
- Example - `customername:AcmeInc;`

Region

- List of which region(s) should be scanned when locating EC2 instances, RDS instances, Storage, and CloudWatch data

- Example – `region:us-east-1;us-east-2;`

Server

- List of servers by domain name using the format `server\instance,port`
- These should be resolvable by DNS from the workstation or server running the script
- If the server is only running a default instance then an entry in the Instances list is sufficient, if there is also named instances then they should be added here. The script will filter for duplicates between Instances and Servers based on IP.
- Example - `server:SQL1.acme.com,1433;SQL2.lab.com;SQL2.lab.com\Prod,51403;`

Instance

- List of InstanceIDs to be collected from
- The IDs will be needed for CloudWatch data collection
- These instances will be mapped to their primary IP and that IP will be used for SQL connectivity.
- Example - `instance:i-0abcdef1234567890;`

Rds

- List of RDS DB Identifiers to be collected from
- Example - `rds:sql-express-2022;sql-prod;`

FSx

- List of FSx file system IDs to be collected from
- Example - `fsx:fs-xxxxxxxxxxxxxxxxxx;`

Database

- List of database names to be **Included** or **Excluded** in scan, if no names provided all user databases will be analyzed. Format is “INCLUDE;” or “EXCLUDE;” followed by database names.
- Example - `database:INCLUDE;StackOverflow2013;WideWorldImporters;`
- Example - `database:EXCLUDE;StackOverflow2013;WideWorldImporters;`

Authentication

For AWS authentication the script can leverage an EC2 associated IAM role, a profile configured locally, SSO login using the referenced profile.

For SQL Server authentication the default will be pass-through authentication where the current domain account will be used. The script, .bat or PS session can be launched using

the “Run As” option to specify an alternate domain account. A SQL login account can also be specified.

Note the same credentials will be used for all resources identified in the config file, if multiple credentials are needed separate data collection runs will be needed for each credential.

Required Permissions

To perform the complete data collection there is a SQL Server access and an AWS account access component. For SQL Server within the download directory there are two files:

- MVA-Sql-Permissions-Create.sql
- MVA-Sql-Permissions-Drop.sql

These scripts will create and drop a sql login named mva_user, (this would need to be run against all SQL instances in scope), this login is “not” required and an equivalent existing account with at least these permissions can be used instead.

For AWS account access within the download directory there is a file:

- MVA-AWS-Policy.json

This file contains the minimum permissions needed to complete the CloudWatch data collection.

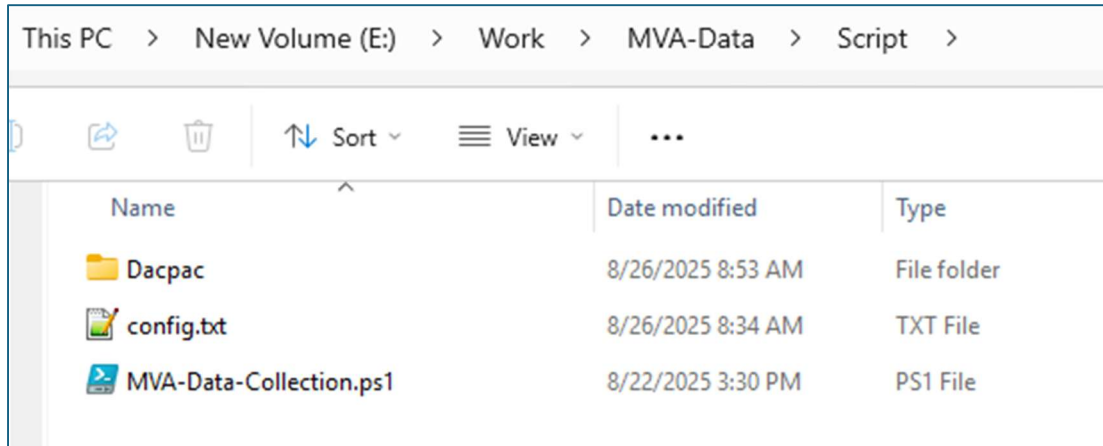
External Downloads

This script will attempt to connect to Microsoft to download the latest version of the [Microsoft DacPac tool](#) (if -ExportDacPacs is used) and will connect to AWS to validate any AWS resources specified. If only **-CollectTsqlData** is used and only **CustomerName:** and **Server:** are populated in the configuration file, then no external resources will be connected to.

Pre-Staging the DacPac Export Tool

For secure networks with no external/internet access the DacPac tool can be manually downloaded and staged for the script to use. To perform this the steps are:

1. Download the [sqlpackage.zip](#) file
2. Unzip the file into the same directory as MVA-Data-Collection.ps1
3. The final path to the sqlpackage.exe in the below example will be *E:\Work\MVA-Data\Script\Dacpac\sqlpackage.exe*



Name	Date modified	Type
Dacpac	8/26/2025 8:53 AM	File folder
config.txt	8/26/2025 8:34 AM	TXT File
MVA-Data-Collection.ps1	8/22/2025 3:30 PM	PS1 File

Collecting the Data

The “Validate Resources” process should be run first to confirm DNS, permissions, and any access issues. When this step is completed without error the data collection step can be started.

*Example Script: **MVA-Data-Collection-ValidateResourcesOnly.bat** – this may need to be edited to meet your environments needs.*

One Time Collection

There are three primary parameters to include in running the data collection script that are needed for a complete evaluation.

- CollectTsqlData - This runs a series of T-SQL scripts against all SQL instances provided and exports them to CSV files.
- ExportDacPacs – This will create a Schema Only DacPac export for all databases in scope. This entails an automated download and install of the latest version of the [SqlPackage software from Microsoft](#) that will be placed in a \Dacpac\ directory alongside the PowerShell script location.
- CollectCloudWatchData – This will collect 30 days of cloud watch data on compute and storage utilization and export to CSV files. This can also be captured manually but is more complex, see section [Manual CloudWatch Data Export](#).

There are also several ancillary parameters that may need to be set dependent on your environment.

- SqlServerConnectionTimeout – This connection timeout defaults to 5 seconds but may need to be increased.

- **SqlServerQueryTimeout** – The query timeout defaults to 300 seconds but may need to be increased.
- **SqlUser & SqlPassword** – The default behavior is Pass Through Authentication, but SQL Authentication can be provided here. The password can also be supplied at run time.
- **AWSProfile** – This can be used to set the environment profile for AWS Authentication.
- **UseSSOLogin** – This will call the SSO login Process, this feature is currently in preview and additional support may be needed to use this.
- **ExportPath** – The Export directory structure defaults to the same directory as the PowerShell script; an alternate Local or remote path can be specified here.
- **UsePublicIPs** – The default is False, but if connectivity needs to use public facing DNS names this can be set.

*Example Script: **MVA-Data-Collection-CollectEnvironmentData.bat** – this may need to be edited to meet your environments needs.*

Recurring Collection

Optionally, to build an application dependency model Evolve recommends using the script to create a SQL Agent Job that will scrape the active connections, create a temporary database **MVA-Data-Collection**, and write them to a **ConnectionInfo** table every five minutes using the **CollectConnectionsOnly** parameter.

The first execution will create the SQL Agent Job and database; a second execution will be needed to collect all data from the **ConnectionInfo** table at the end of the collection period.

The appropriate period for a sufficient sample of connections depends on the characteristics of your workload but Evolve recommends 24hrs as a minimum.

When the connection collection process is no longer needed the job **CollectConnections** can be manually disabled/deleted or the **CleanUpEnvironment** parameter can be used to remove the Job and database from all servers specified in the config file.

Note: Ensure the connections data has been collected and exported to a CSV (using the **CollectConnectionsOnly parameter) prior to clean up as all connections data is permanently deleted.**

*Example Script: **MVA-Data-Collection-CollectConnectionsOnly.bat** – this may need to be edited to meet your environments needs.*

Multi Region Collection

The PowerShell script will process multiple regions to locate the AWS resources added to the config file so within a single account one collection process can handle multiple regions.

Multi Account Collection

The PowerShell script is not configured to connect to multiple accounts, if [AWS Organizations](#) are configured and the account executing the collection script has access across all accounts then a one collection process can be used otherwise its recommended to run a collection effort per account.

When running multiple data collection processes, it's recommended to unzip the scripts into multiple separate directories and keep a separate config file per collection.

Reporting and ETL

Within your SQL Server environment you may have Integration Services (SSIS) and Reporting Services (SSRS) in use. We do scrape the SSISDB and ReportServer databases for basic information on those objects, if those SSIS and SSRS databases reside on a central or separate server from the user databases those servers can also be included in the config (database filtering recommended), collected as a separate data collection run, or info on the SSIS and SSRS objects can be manually submitted separately.

How to Share Results

When the PowerShell script successfully completes there will be a `***-MVA-Export-ALL-mmddyyhhmmss.zip` file created within the `\Export\` directory. This Export directory will either reside in the same directory as the PS script (default behavior) or at the `-ExportPath` location if the override parameter is used.

The CSV files contained within this zip will have metadata such as servernames etc. but will not contain any user data. We highly encourage you to review the contents internally before sharing with Evolve.

Methods to securely send the zip(s) to Evolve will be dependent on your own security policies, common approaches we see are:

- Password protected Zip to email – mva@evolvecloudservices.com
 - *Note: the zip file created in this script is not protected, a third-party utility will be required for this*
- Shared via Amazon S3 bucket using a [Pre-Signed URL](#)

- File sharing solutions such as [DropBox](#), [Box](#) or similar

Environment Clean Up

When all data has been collected and sent to Evolve the final step is to run the script with the CleanUpEnvironment parameter. This will remove all SQL Agent Jobs, databases and DacPac agent install. The final ***-MVA-Export-ALL-mmddyyhhmmss.zip(s) created will not be deleted and will need to be removed manually.

*Example Script: **MVA-Data-Collection-CleanUpEnvironment.bat** – this may need to be edited to meet your environments needs.*

Manual CloudWatch Data Export

Data Collection Period

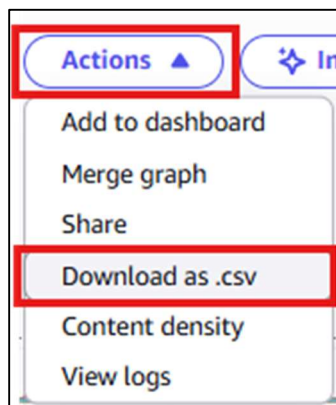
The default retention period before detailed data points is summarized is 30 days, hence we collect 30 days in 5-minute increments.

To configure this within your CloudWatch console goto CloudWatch > Metrics > All Metrics and adjust the time period to Custom and select 30 days as shown below.

The screenshot shows the AWS CloudWatch console's time period configuration interface. At the top, there are tabs for '1h', '3h', '12h', '1d', '3d', '1w', 'Custom (1M)', and 'UTC timezone'. Below these, there are two tabs: 'Absolute' and 'Relative'. Under the 'Relative' tab, there are several rows of buttons for selecting time intervals: Minutes (1, 3, 5, 15, 30, 45), Hours (1, 2, 3, 6, 8, 12), Days (1, 2, 3, 4, 5, 6), Weeks (1, 2, 4, 6), and Months (3, 6, 12, 15). At the bottom, there is a 'Duration' input field with the value '30' and a 'Unit of time' dropdown menu with 'Days' selected. The 'Apply' button is highlighted with a red box.

Exporting CloudWatch Metrics

After the metrics have been added and visualized in the console select Actions > Download as .csv to export this data to a file.



Additional Resource Information







































































To complete the manual data collection in addition to the CloudWatch data we also need information on the configuration of the EC2, EBS and RDS resources. Please complete the attached manual export spreadsheet ***MVA-Manual-Collection-Template-v1.xlsx*** for the in-scope resources.

EC2 Metrics

Note that a combination of average, sum and maximum is collected for the metrics shown in the table below, additionally the CPUUtilization metric will need to be added twice for the average statistic and the maximum statistic, this can be done using the copy icon as shown below.



Metric	Statistic
CPUUtilization Average	Average
CPUUtilization Maximum	Maximum
EBSReadOps	Sum
EBSWriteOps	Sum
EBSReadBytes	Sum
EBSWriteBytes	Sum
NetworkIn	Maximum
NetworkOut	Maximum
NetworkPacketsIn	Maximum
NetworkPacketsOut	Maximum

Browse (20)	Multi source query	Graphed metrics (10)	Options	Source		
<div>Add dynamic label Info</div>						
<input checked="" type="checkbox"/>	Label	Details	Statistic	Period	Y axis	Actions
<input checked="" type="checkbox"/>	<div>EBSReadOps</div>	<div>EC2 • EBSReadOps • InstanceId: i-088cbc...</div>	Sum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>EBSReadBytes</div>	<div>EC2 • EBSReadBytes • InstanceId: i-088cb...</div>	Sum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>EBSWriteOps</div>	<div>EC2 • EBSWriteOps • InstanceId: i-088cbc...</div>	Sum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>EBSWriteBytes</div>	<div>EC2 • EBSWriteBytes • InstanceId: i-088ct...</div>	Sum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>CPUUtilization Average</div>	<div>EC2 • CPUUtilization • InstanceId: i-088ct...</div>	Average	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>NetworkIn</div>	<div>EC2 • NetworkIn • InstanceId: i-088cbc4a...</div>	Maximum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>NetworkOut</div>	<div>EC2 • NetworkOut • InstanceId: i-088cbc4...</div>	Maximum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>NetworkPacketsIn</div>	<div>EC2 • NetworkPacketsIn • InstanceId: i-08...</div>	Maximum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>NetworkPacketsOut</div>	<div>EC2 • NetworkPacketsOut • InstanceId: i-0...</div>	Maximum	5 minu...	< >	<div>      </div>
<input checked="" type="checkbox"/>	<div>CPUUtilization Maximum</div>	<div>EC2 • CPUUtilization • InstanceId: i-088ct...</div>	Maximum	5 minu...	< >	<div>      </div>

This will need to be repeated and exported for each instance in scope for the MVA.

EBS Metrics

This will need to be repeated for all EBS volumes attached to the EC2 instances in scope of this MVA.

Metric	Statistic
VolumeReadOps	Sum
VolumeWriteOps	Sum
VolumeReadBytes	Sum
VolumeWriteBytes	Sum

Browse (16)

Multi source query





























Graphed metrics (4)

Options

Source

Add dynamic label

Info

<input checked="" type="checkbox"/>	Label	Details	Statistic	Period	Y axis	Actions
<input checked="" type="checkbox"/>	<div><div></div><div>VolumeReadOps</div><div>Info</div></div>	<div>EBS • VolumeReadOps • Volumeld: vol-036c2afe</div> <div>Info</div>	Sum ▾	5 minu... ▾	< >	<div><div></div><div></div><div></div><div></div><div></div><div></div></div>
<input checked="" type="checkbox"/>	<div><div></div><div>VolumeWriteOps</div><div>Info</div></div>	<div>EBS • VolumeWriteOps • Volumeld: vol-036c2afe</div> <div>Info</div>	Sum ▾	5 minu... ▾	< >	<div><div></div><div></div><div></div><div></div><div></div><div></div></div>
<input checked="" type="checkbox"/>	<div><div></div><div>VolumeReadBytes</div><div>Info</div></div>	<div>EBS • VolumeReadBytes • Volumeld: vol-036c2af</div> <div>Info</div>	Sum ▾	5 minu... ▾	< >	<div><div></div><div></div><div></div><div></div><div></div><div></div></div>
<input checked="" type="checkbox"/>	<div><div></div><div>VolumeWriteBytes</div><div>Info</div></div>	<div>EBS • VolumeWriteBytes • Volumeld: vol-036c2a</div> <div>Info</div>	Sum ▾	5 minu... ▾	< >	<div><div></div><div></div><div></div><div></div><div></div><div></div></div>

RDS Instances

Note that a combination of average, sum and maximum is collected for the metrics shown in the table below, additionally the CPUUtilization and DatabaseConnections metric will need to be added twice for the average statistic and the maximum statistic, this can be done using the copy icon as shown in the EC2 section.

Metric	Statistic
CPUUtilization Average	Average
CPUUtilization Maximum	Maximum
DatabaseConnections Average	Average
DatabaseConnections Maximum	Maximum
ReadOps	Sum
WriteOps	Sum
ReadThroughput	Sum
WriteThroughput	Sum
ReadLatency	Sum
WriteLatency	Sum
NetworkReceiveThroughput	Maximum
NetworkTransmitThroughput	Maximum

Browse (47)

Multi source query

































































































Graphed metrics (12)

Options

Source

Add dynamic label

Info

<input checked="" type="checkbox"/>	Label	Details	Statistic	Period	Y axis	Actions
<input checked="" type="checkbox"/>	 WriteThroughput 🔗	RDS • WriteThroughput • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 ReadThroughput 🔗	RDS • ReadThroughput • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 ReadIOPS 🔗	RDS • ReadIOPS • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 WriteIOPS 🔗	RDS • WriteIOPS • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 WriteLatency 🔗	RDS • WriteLatency • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 ReadLatency 🔗	RDS • ReadLatency • DBInstanceIdentifier: sql-ex-rds 🔗	Sum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 CPUUtilization Average 🔗	RDS • CPUUtilization • DBInstanceIdentifier: sql-ex-rds 🔗	Average 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 DatabaseConnections Average 🔗	RDS • DatabaseConnections • DBInstanceIdentifier: sql-ex-rds 🔗	Average 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 NetworkReceiveThroughput 🔗	RDS • NetworkReceiveThroughput • DBInstanceIdentifier: sql-ex-rds 🔗	Maximum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 NetworkTransmitThroughput 🔗	RDS • NetworkTransmitThroughput • DBInstanceIdentifier: sql-ex-rds 🔗	Maximum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 CPUUtilization Maximum 🔗	RDS • CPUUtilization • DBInstanceIdentifier: sql-ex-rds 🔗	Maximum 	5 minu... 	< >	    
<input checked="" type="checkbox"/>	 DatabaseConnections Maximum 🔗	RDS • DatabaseConnections • DBInstanceIdentifier: sql-ex-rds 🔗	Maximum 	5 minu... 	< >	    

Zip File/Repo Content

The zip file MVA-Data-Collection.zip should be downloaded onto a server/workstation that has PowerShell 2.0 or higher (PowerShell is standard on Windows 2008 R2 and above) and the SQL Server client tools.

- MVA-Data-Collection.ps1
 - PowerShell Script
- MVA-Data-Collection-CleanUpEnvironment.bat
 - Preconfigured script to remove all objects created by this script
- MVA-Data-Collection-CollectConnectionsOnly.bat
 - Preconfigured script to run connections collection
- MVA-Data-Collection-ValidateResourcesOnly.bat
 - Preconfigured script to validate resources added to config.txt
- MVA-Data-Collection-CollectEnvironmentData.bat
 - Preconfigured script to collect SQL data, Schema only DacPacs and CloudWatch data
- Config.txt
 - Configuration file
- Evolve-MVA-Data-Collection-Instructions.pdf
 - This document
- MVA-Manual-Collection-Template-v1.xlsx
 - Excel file for manual data collection (if needed)
- MVA-Sql-Permissions-Create.sql
 - TSQL script to create a sql login named mva_user
- MVA-Sql-Permissions-Drop.sql
 - TSQL script to drop a sql login named mva_user
- MVA-AWS-Policy.json
 - IAM permissions file