BAG (Basic Availability Group) Setup for MS SQL in AWS instances

Introduction, background and scope:

MS SQL natively provides BAG feature for DR(Disaster Recovery) and HA(High Availability). For a detailed workflow and insights we will be trying to install MS SQL in AWS instances and then perform the setup. The goal is to survive a server or data center failure with near-zero data loss, which requires low-latency, synchronous communication. Although for DR scenarios we should be doing for different regions, for a small scale HA setup we are trying the setup for the same region but different Availability Zones.

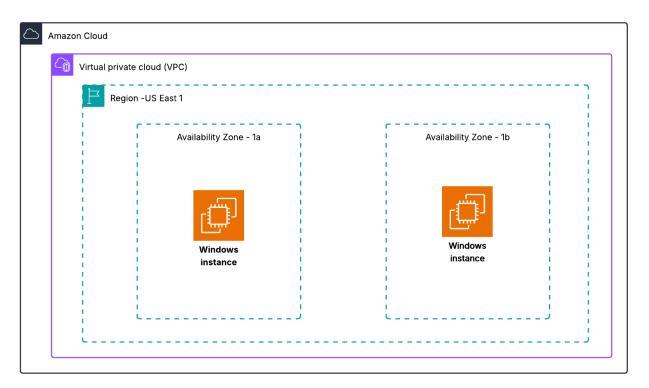


Figure: Simple Architecture Diagram for our setup

First, let's clarify the limitations of a Basic Availability Group (BAG) in SQL Server Standard Edition:

One Database per Group:

We can only have one user database in a BAG. If we need to protect multiple databases, you must create multiple BAGs.

Two Replicas Only:

A BAG is limited to a primary replica and one secondary replica.

No Read Access on Secondary:

The secondary replica cannot be used for reading data or offloading backups. It is purely a hot standby.

Asynchronous-Commit Recommended:

While synchronous-commit is possible, it's generally recommended to use asynchronous commit for performance.

Manual Failover Only:

There is no support for automatic failover. If the primary goes down, you must manually fail over to the secondary.

Phase 1: Network Infrastructure Setup

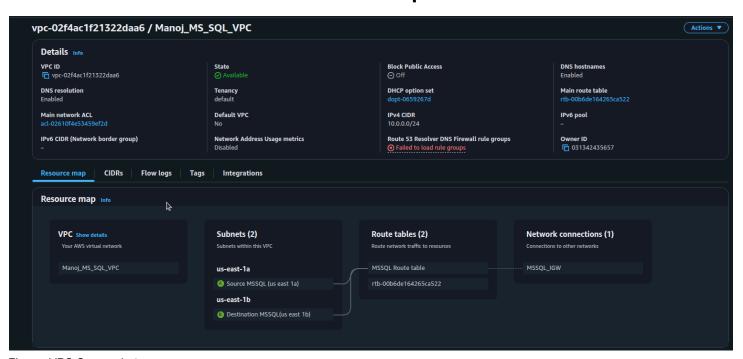


Figure: VPC Screenshot

Summary: The infrastructure is set up in US EAST - 1. With the subnets in us-east-1a and us-east-1b. Both have the same route table redirecting the traffic internally and to the internet as shown in the snip below.

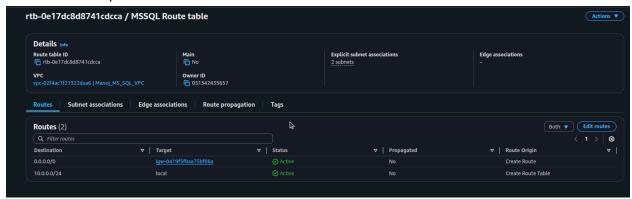


Figure: Route table for the VPC

Phase 2: Instance Setup:

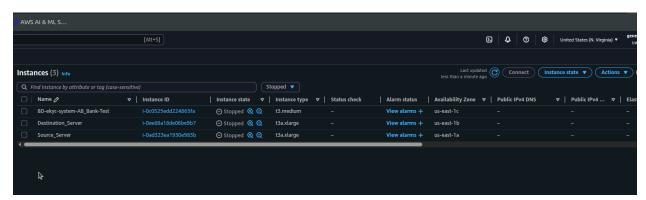


Figure: Source and destination MSSQL servers

It's a critical note that the two instances that are supposed to be in the Instances must be started with the same AMI as we have different checks running while making a cluster, 2 nodes are eligible to be inside a cluster only when they have the same version of operating system. It was a recurring problem during the R&D as shown in the snip below:

The Instances are to be in different subnets for now.

The security group is set up as below:



These rules are nothing less than a mess. It was assumed that the vague ports like 3389, 5022 135 were taking traffic through the public internet but it is not all that necessary. Most likely, we'll only need RDS and rule no 2. **allows all types of traffic (TCP, UDP, ICMP, etc.) on all ports** from any **EC2 instance** that is part of the security group. That is our only requirement since MSSQL server (When used for failover clustering) is known for using weird ports that are not well documented.

That's all for the EC2 setup, other configurations are carried out normally.

Phase 3: Clustering, Failover Tools, MSSQL Setup

According to this Document, also in the snippets:

0	To create a new cluster or to add nodes to the cluster, a local account needs to be provisioned on all nodes of the cluster (as well as the no		
	from which the operation is invoked) with the following requirements:		

- 1. Create a local 'User' account on each node in the cluster
- 1. The username and password of the account must be the same on all nodes
- 1. The account is a member of the local 'Administrators' group on each node
- 1. When using a non-builtin local administrator account to create the cluster, set the LocalAccountTokenFilterPolicy registry policy to 1, on all the nodes of the cluster. Builtin administrator accounts include the 'Administrator' account. You can set the LocalAccountTokenFilterPolicy registry policy as follows:

We have to make new **administrator** accounts with the same passwords in each of the instances that we are treating as nodes. Additionally, they should be assigned to the **Remote Desktop Users** group as sometimes just administrator privilege is not enough for RDP access. Another challenge is:

SQL Server Management Studio (SSMS)

By default, only the built-in local administrator account can access a SQL Server instance launched from an AWS Windows AMI. You can use SQL Server Management Studio (SSMS) to add domain users so that they can access and manage SQL Server.

Perform the following steps to access a SQL Server instance on Amazon EC2 as a domain user.

So we should follow this doc for access to new administrator users.

Now for the messy part:

- RDP through the Source Server and Destination server through the **new Administrative user** that we have just created and then install the failover clustering feature using this CLI command. (Note: Also can be done through GUI but CLI faster)

Install-WindowsFeature -Name Failover-Clustering -IncludeManagementTools

Administrator Windows Powerbhell

Reply from 10.0.0.13: bytes=32 timeclms TTL=128
Reply from 10.0.0.13: bytes=32 timeclms TTL=128

Collecting data...

100

(0000000000

PS C:\Hindows\system32> ping DestHSSQL.test-workgroup.local

Pinging DestHSSQL.test-workgroup.local (Fe80::5db4:34b0:3fd3:e4c7%8) with 32 bytes of data:

Reply from fe80::5db4:34b0:3fd3:e4c7%8: timeclms
Reply from fe80::5db4:34b0:3fd3:

It recommends for a restart but hold for a bit

- Also do:

```
new-itemproperty -path
HKLM:\SOFTWARE\Microsoft\Windows\CurrentVersion\Policies\System -Name
LocalAccountTokenFilterPolicy -Value 1
```

Otherwise it will give error:

- Additionally, You have the option to turn off the firewall as whole or run these commands:

```
Enable-NetFirewallRule -DisplayName "File and Printer Sharing (Echo Request - ICMPv4-In)"

Enable-NetFirewallRule -DisplayGroup "Windows Management Instrumentation (WMI)"
```

- Now this is the right time to restart both instances.

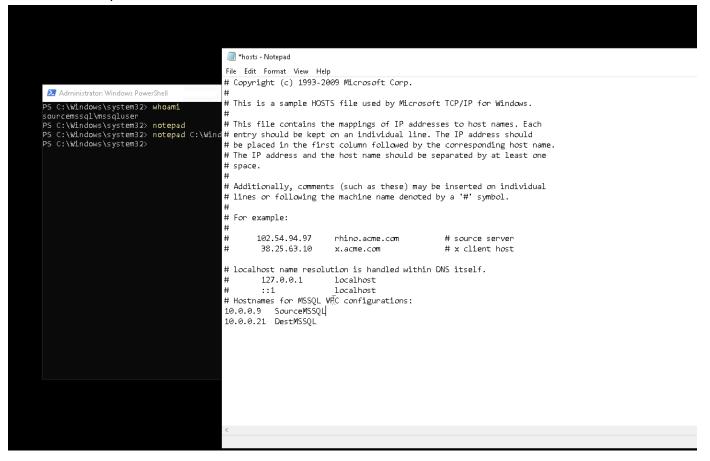
A slight reminder that we are trying to do this setup without setting up an AD, An AD would have been more easier and seamlessly connected to the SQL Server Management Studio (SSMS).

We are trying for resolving hostnames through the hostnames files in the windows path:

```
C:\Windows\System32\Drivers\etc\hosts
```

Simply, in both the instances go to the file and write the private ip address of the two instances, assigned according to the VPC and the subnet's CIDR.

For example:



- Try to ping the source server with the destination server and vice versa. This is our crucial step for making the cluster and assigning nodes.
- If everything is completed till this step, we are ready for making failover clusters.

Clusters:

Our main goal for a successful implementation of BAG (Basic Availability Groups) is making clusters and assigning nodes.

- First, we check the compatibility of nodes to make a cluster. We do this by using the following command:

```
Test-Cluster -Node hostname1, hostname2
```

```
➤ Select Administrator: Windows PowerShell

>> Enable - NetFirewallRule - DisplayGroup "Windows Defender Firewall Remote Management"

PS C:\Windows\system32> Test-Cluster - Node SourceMSSQL, DestMSSQL

>>

WARNING: System Configuration - Validate Active Directory Configuration: The test reported some warnings..

WARNING: Network - Validate Network Communication: The test reported some warnings..

WARNING:

Test Result:

HadUnselectedTests, ClusterConditionallyApproved

Testing has completed for the tests you selected. You should review the warnings in the Report. A cluster solution is supported by Microsoft only if you run all cluster validation tests, and all tests succeed (with or without warnings).

Test report file path: C:\Users\MSSQLUser\AppData\Local\Temp\Validation Report 2025.08.06 At 10.08.06.htm
```

Above command runs several specific checks and makes a detailed report:





Back to Summar Back to Top

It is recommended to check the warnings thoroughly and then only continue, otherwise it will result in a **broken cluster**.

- We now continue with the configuration and make a cluster using powershell. Remember to:
 - 1. Run the powershell as administrator
 - 2. Make sure that you are doing RDS with the new local administrator account, which we created.
 - 3. All above steps are completely cleared without issues.
- For making cluster, we use the command (syntax):

```
New-Cluster -Name MSSQLCluster -Node SourceMSSQL,DestMSSQL -AdministrativeAccessPoint Dns -StaticAddress 10.0.1.100 -NoStorage
```

Syntax Breakdown

Parameter	Description
New-Cluster	The command that creates the failover cluster.
-Name	Specifies the network name for the cluster.
-Node	A comma-separated list of servers to add to the cluster.
-AdministrativeAccessPoint	Dns creates the cluster for a workgroup (no Active Directory).
-StaticAddress	Assigns a static IP address to the cluster's network name.
-NoStorage	Creates the cluster without shared storage, ideal for SQL Always On groups.

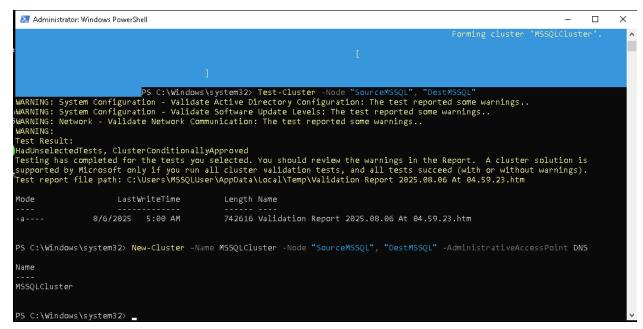


Figure: Running the New-Cluster command

Also to check if the nodes are assigned in the cluster we can run the command:

Get-Cluster

These are to be run in both nodes.

For more details on resources like ips and storage we use:

Get-ClusterResource

If both IPs are offline in the cluster, remote registry service fails sometimes, run these individually:

```
Get-Service -Name RemoteRegistry
Start-Service -Name RemoteRegistry
Set-Service -Name RemoteRegistry -StartupType Automatic
```

Now we add the cluster name that we just made to the hostname file:

C:\Windows\System32\Drivers\etc\hosts



Then we ping the cluster. In my case the command is, ping MSSQLCluster

```
PS C:\Windows\system32> Test-Cluster -Node "SourceMSSQL", "DestMSSQL"

WARNING: System Configuration - Validate Active Directory Configuration: The test reported some warnings..

WARNING: System Configuration - Validate Software Update Levels: The test reported some warnings..

WARNING: Network - Validate Network Communication: The test reported some warnings..

WARNING:

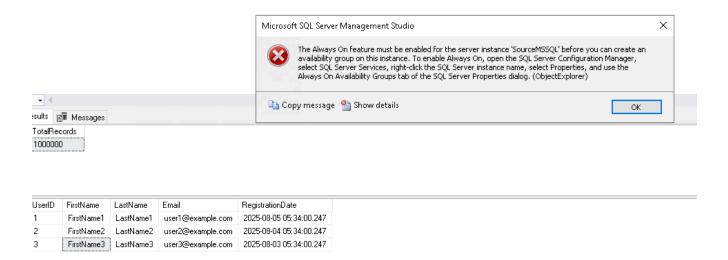
W
```

Important Note: In my above screenshot that is running the New-Cluster line, i do not have any static IP, but the IP can be seen using the Get-ClusterResource command in the powershell.

Cluster is successfully configured.

Phase 4: Making Availability Groups

Have a look at this SSMS screenshots:



So we go and enable availability groups

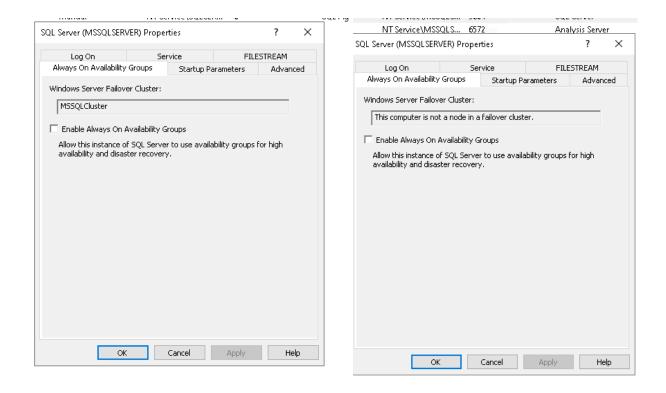
We go to the path: Server Configuration Manager --> SQL Server Services --> SQL Server name and Right click to properties

Then we just tick the option: Enable Always on Availability Groups

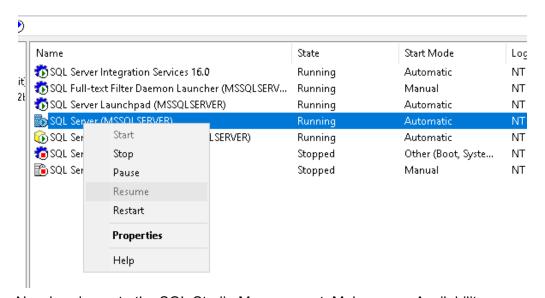
If the node is not in the cluster, It will give the error like below:

"The computer is not in a failover cluster"

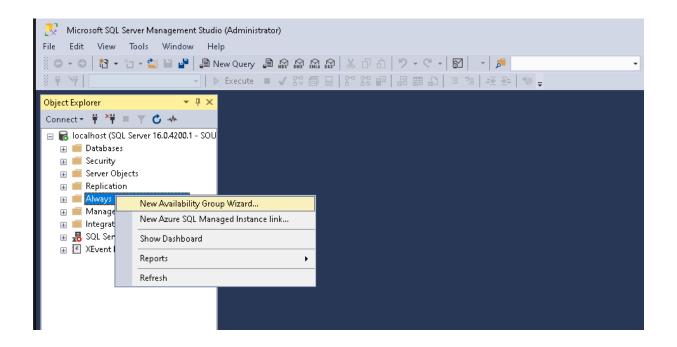
If the cluster is not broken, this will not happen



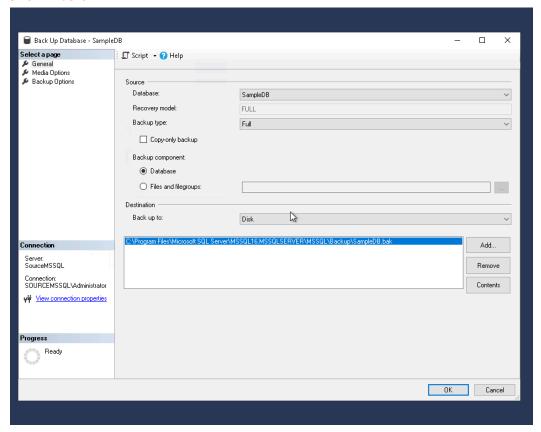
Then Restart the MSSQL server.

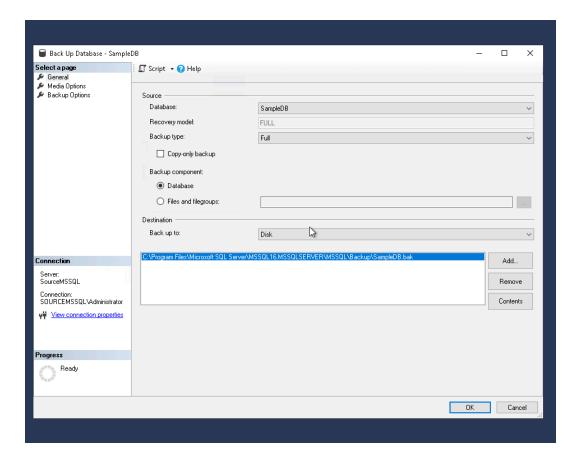


Now head over to the SQL Studio Management. Make a new Availability group using the GUI.

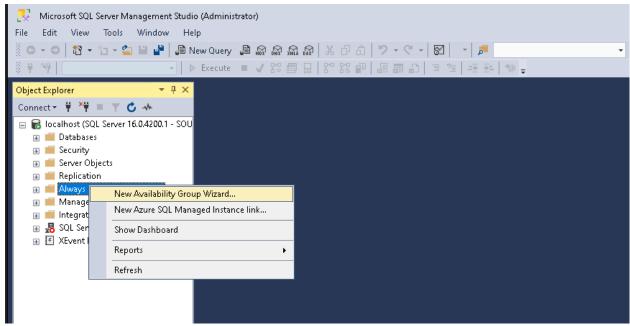


Critical note here: By default the BAG supports only one database in the availability group. For the database to be used like that, It MUST be fully backed up. The process is shown below.

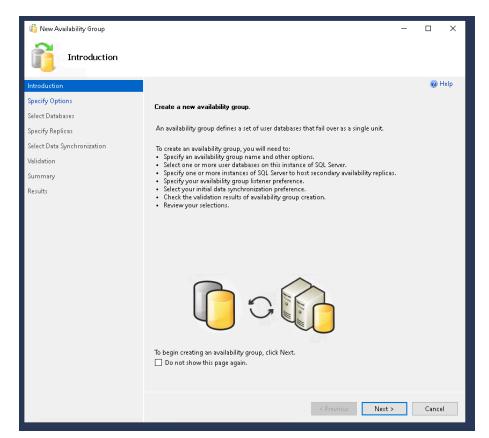




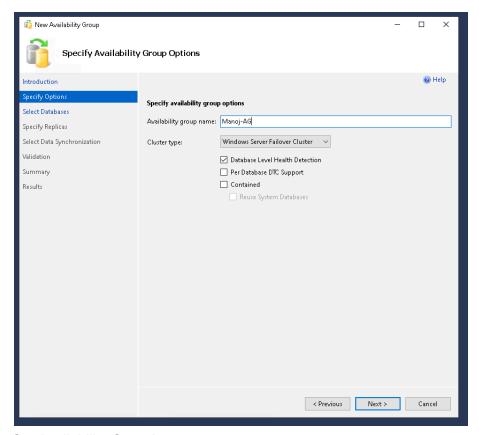
If the database is backed up, We'll continue here:



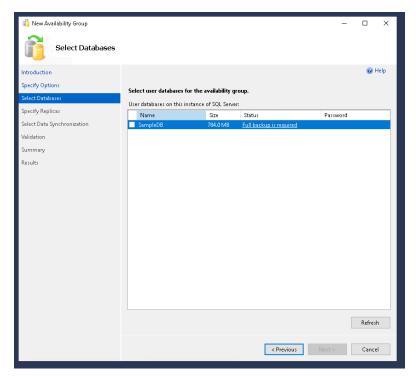
Select the new availability group wizard



Click Next



Set Availability Group's name



Select the Database

This completes the BAG setup for our scenario!

Refer to the next doc (AWS Hybrid AD setup) for the next part. (o ^o)



Additional References:

https://www.youtube.com/watch?v=CbXtHGBVBjU

https://techcommunity.microsoft.com/blog/failoverclustering/workgroup-and-multi-domain-cluster s-in-windows-server-2016/372059

https://docs.aws.amazon.com/sql-server-ec2/latest/userguide/connect-sql-server-on-ec2-instanc e.html