SkyCluster™ Deployment Guide
for Oracle RAC on Azure Cloud

rev. 20.06-2020.07.16

Table of Contents

1 Introduction

SkyCluster is an engineered cloud system that enables active-active database high availability infrastructure in public clouds. This guide provides step-by-step instructions for system and database administrators deploying SkyCluster with Oracle RAC on Azure cloud.

Key components of SkyCluster 20.06 for Azure:

- FlashGrid Storage Fabric: ver. 19.06
- FlashGrid Cloud Area Network: ver. 19.10
- FlashGrid Diagnostics: ver. 20.02
- SkyCluster Health Checker ver. 20.04
- Oracle Database: ver. 19c, 18c, 12.2.0.1, 12.1.0.2, or 11.2.0.4.
- Oracle Grid Infrastructure: ver. 19c. (Versions 18c and 12.2.0.1 available on request)
- Operating System: Oracle Linux 7, Red Hat Enterprise Linux 7
- Azure VMs: DSv2, DSv3, ESv3, M, GS, Ls_v2.
- Disks: Premium SSD or local NVMe SSD

SkyCluster is delivered as Azure Resource Manager templates that automate configuration of multiple components required for a database cluster. SkyCluster Launcher is an online tool that simplifies the deployment process by guiding through the cluster configuration parameters and generating Azure Resource Manager templates.

Additional information about the SkyCluster architecture is available in the following white paper: *"Mission-Critical Databases in the Cloud. Oracle RAC in Microsoft Azure Enabled by FlashGrid®."*

2 Prerequisites

The following prerequisites are required for automated deployment of an Oracle RAC cluster in Azure using SkyCluster Launcher:

- **Azure Storage Blob Container** with Oracle installation files that will be downloaded to the cluster nodes during cluster initialization. The list of files that must be placed in the Storage Container will be shown in SkyCluster Launcher. The corresponding storage account must have access for *'All networks'* enabled in *'Firewall and virtual networks'* settings.
- **Enabled Service Endpoints** when deploying in an existing VNet. Enabling service endpoints allows access to the storage container from the VMs. If Service Endpoints are disabled and public IPs not assigned then cluster initialization will fail because downloading Oracle files from the VMs will not be possible.
- **Azure subscription with sufficient quotas** for creating the required number and type of VMs and sufficient number and size of Premium Managed Disks.
- **SSH key pair** that will be used for accessing the VMs. Use of passwords instead of the key pair is not supported. To create a new key pair use *ssh-keygen* in Linux or *puttygen* in Windows. In the SkyCluster Launcher tool you will need to provide the public key that will be placed on the VMs. Example of a valid public key pair format:

  ```
  ssh-rsa <PublicKeyBody>
  ```

- **Properly configured Network Security Group (NSG)** when deploying in an existing VNet. You have a choice of attaching an NSG to the VMs or using the NSG attached to the subnet. In either case, the following ports must be open between the cluster node VMs: UDP 4801, 4802, 4803 and TCP 3260. FlashGrid recommends configuring the NSG rules by using an Application Security Group (ASG) for the cluster node VMs. You can configure one ASG per cluster or a separate ASG for each cluster.

3 Deploying a Cluster

The FlashGrid SkyCluster Launcher tool simplifies provisioning of Oracle RAC clusters in Azure by automating the following tasks:

- Creating cloud infrastructure: VMs, storage, and optionally network
- Installing and configuring FlashGrid Cloud Area Network
- Installing and configuring FlashGrid Storage Fabric
- Installing, configuring, and patching Oracle Grid Infrastructure
- Installing and patching Oracle Database software
- Creating ASM disk groups

To create a cluster

1. Open FlashGrid SkyCluster Launcher tool:
    - Start with one of the standard configurations at https://www.flashgrid.io/skycluster-for-azure
    - or, if you have a custom configuration file, upload it at https://2006.cloudprov.flashgrid.io
2. Configure parameters of the cluster
3. Click *Validate Configuration* button
4. If verification passes then click *Launch Cluster* button, which will take you to Azure Resource Manager
5. Select *Resource group -> Create new*. By having the cluster in a separate resource group you can later delete the entire cluster by simply deleting the resource group.
6. Enter a name for the new resource group that will contain the cluster. A name matching the cluster name is recommended.
7. Select your target location (region)
8. Check *'I agree to the terms and conditions state above'*
9. Click *Purchase*
10. Open list of Notifications (bell icon) and click *'Deployment in progress…'*
11. Wait until the deployment status changes to *Succeeded*
12. If the deployment fails:
    a) Check for the cause of the failure in the *Operation details*
    b) Correct the cause of the error
    c) Delete the failed resource group
    d) Repeat the steps for creating a new resource group
13. SSH to the first (as it was specified on the cluster configuration page) cluster node VM as user ***az-admin@***
14. The welcome message will show the current initialization status of the cluster: in progress, failed, or completed.
15. If initialization is still in progress then wait for it to complete (this includes Oracle software installation and configuration). You will receive a broadcast message when initialization completes or fails. Cluster initialization takes approximately 1 to 2 hours depending on configuration.

4 After Deploying a Cluster

4.1 Verifying cluster status

On any of the cluster nodes run `flashgrid-cluster` command to verify that the cluster status is *Good* and all checks are passing.

```
[fg@rac1 ~]$ flashgrid-cluster
FlashGrid 18.07.15.48564 #95f2b5603f206af26482ac82386b1268b283fc3c
License: via Marketplace Subscription
Support plan: 24x7
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
FlashGrid running: OK
Clocks check: OK
Configuration check: OK
Network check: OK

Querying nodes: quorum, rac1, rac2 ...

Cluster Name: myrac
Cluster status: Good
-------------------------------------------------------------
Node    Status  ASM_Node  Storage_Node  Quorum_Node  Failgroup
-------------------------------------------------------------
rac1    Good    Yes       Yes           No           RAC1
rac2    Good    Yes       Yes           No           RAC2
racq    Good    No        No            Yes          QUORUM
-------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
GroupName  Status  Mounted   Type    TotalMiB  FreeMiB  OfflineDisks  LostDisks  Resync  ReadLocal  Vote
-------------------------------------------------------------------------------------------------------------
GRID       Good    AllNodes  NORMAL  12588     3376     0             0          No      Enabled    3/3
DATA       Good    AllNodes  NORMAL  2048000   2048000  0             0          No      Enabled    None
FRA        Good    AllNodes  NORMAL  1024000   1024000  0             0          No      Enabled    None
-------------------------------------------------------------------------------------------------------------
```

4.2 Verifying synchronization of clocks

Chrony service is used for synchronizing cluster node clocks with external NTP servers. Without active clock synchronization service the clocks are likely to get out of sync. Oracle CTSS service synchronizes system clocks while CRS is running. However, it cannot synchronize clocks before CRS is started or on quorum nodes.

To check the current clock difference between the cluster nodes

```
$ flashgrid-cluster verify
```

To check status of CHRONYD service

```
$ chronyc sources
```

Example:

```
[fg@rac1 ~]$ chronyc sources
210 Number of sources = 6
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^+ time1.google.com              1   9   377   197  +4820us[+4820us] +/-   23ms
^* time2.google.com              1  10   377   838  -3017us[-3363us] +/-   14ms
^- time3.google.com              1   9   377   313  -5459us[-5459us] +/-   60ms
^- time4.google.com              1   9   377    41    +13ms[  +13ms] +/-  100ms
=- rac2-ext.example.com          2   9   367   398   +109us[ +109us] +/-   14ms
=- racq-ext.example.com          2  10   357   263   -341us[ -341us] +/-   14ms
```

Note that the '*' character shows which NTP server is currently used for synchronization. Normally this should be one of the external NTP servers. If it shows that one of the cluster nodes is used for synchronization then this means that the external NTP servers are not accessible.

The cluster nodes are configured as peers for the Chrony service to provide synchronization between the nodes even when the external NTP servers are not accessible. However, this is only a temporary solution for cases when external NTP servers become inaccessible. In production environments the peer nodes should not be relied upon as a permanent clock synchronization solution.

Public NTP servers, e.g. *timeX.google.com*, can be used only if public IPs are enabled on the VMs (not recommended in production use for security reasons) or if NAT is configured in the network. If needed, the list of NTP servers can be modified in `/etc/chrony.conf` after the cluster is configured.

4.3 OS user accounts

During cluster initialization the following OS user accounts are created:

- *az-admin* - the user account used to SSH to the VMs with the SSH key that was selected when creating the cluster configuration. The user has sudo rights.
- *fg* - can be used for running FlashGrid Storage Fabric or FlashGrid Cloud Area Network utilities. The user has sudo rights. The user has key-based SSH configured between *all* nodes of the cluster.
- *grid* - Grid Infrastructure owner. GI environment variables are preconfigured. The user has key-based SSH configured between all *database* nodes of the cluster.
- *oracle* - Database home owner. Database environment variables, except ORACLE_SID and ORACLE_UNQNAME, are preconfigured. After creating a database you can configure ORACLE_SID and ORACLE_UNQNAME by editing `/home/oracle/.bashrc` file on each database node. The user has key-based SSH configured between all *database* nodes of the cluster.

Note that no passwords are configured for any users. Also password-based SSH authentication is disabled in `/etc/ssh/sshd_config`. Key-based authentication is recommended for better security. Creating passwords for any user is not recommended.

Users *az-admin* and *fg* have sudo rights and allows switching to any other user without requiring a password (which is not configured by default). Example:

```
$ sudo su - grid
```

Users *fg*, *grid*, and *oracle* have key-based SSH access configured between the nodes of the cluster. The corresponding key pairs are generated automatically during cluster initialization. For example, if you are logged in to *node1* as user *fg* then you can SSH into *node2* by simply running `'ssh node2'` without entering a password or providing a key.

4.4 Finalizing cluster configuration

See knowledge base articles for performing the following steps:

1. Changing temporary ASM passwords: https://kb.flashgrid.io/asm-password
2. Creating a database: https://kb.flashgrid.io/createdb
3. Connecting clients to a database: https://kb.flashgrid.io/connect-clients

4.5 Adding a protection lock for the cluster

It is strongly recommended to add a lock to the cluster resource group to protect it against accidental deletion or modification.

4.6 Installing database software (standalone or additional RAC db home)

In most cases manual installation of database software is not required. However, if you need a standalone (non-RAC) database or an additional RAC database home then follow Oracle Database documentation for installing the database software.

4.7 Use of anti-virus software

If anti-virus software has to be used then it is recommended to configure it in a way that avoids putting any files in quarantine. Automatic quarantine of files creates risk of the cluster downtime in case of a false positive detection on a critical system file on multiple nodes of the cluster.

4.8 Use of automatic configuration tools

Automatic configuration tools (e.g. Ansible, Salt, etc.) must be used with extra care. Incorrect modification of a critical system file (e.g. /etc/resolv.conf) on multiple cluster nodes may cause cluster downtime. Note that many critical system configuration files are protected with immutable attribute and have warnings in them. Do not remove the immutable attribute or allow automatic modification of such files unless absolutely necessary.

4.9 Security hardening

Cluster nodes are deployed using RHEL 7 or Oracle Linux 7 images that have main security best practices implemented by default. The following steps are recommended, in case additional security hardening is required:

1) Request FlashGrid support to review the list of required changes.
2) Back up all cluster nodes: https://kb.flashgrid.io/backup-restore/backup-and-restore-in-azure
3) Implement the required changes on all nodes.
4) Restart the entire cluster: https://kb.flashgrid.io/maintenance/maintenance-azure#restarting-the-entire-cluster
5) Verify health of the cluster: `$ sudo skycluster-health-check`
6) In case of errors, roll back the changes or restore the nodes from backup.

5 Monitoring Cluster Health

The following methods of monitoring cluster health are available:

- *skycluster-health-check* utility checks multiple items including database configuration, storage, OS kernel, config file modifications, errors in the logs, and other items that may affect health of the cluster or could help with troubleshooting. It is recommended for manual checks only.
- *flashgrid-cluster* utility displays status of the storage subsystem (FlashGrid Storage Fabric and ASM) and its main components. The utility can be used in monitoring scripts. It returns a non-zero value if status of the cluster is *Warning* or *Critical*.
- Alerts about failures are recorded in system log and can be analyzed by 3rd-party tools
- Email alerts can be sent to one or several email addresses
- ASM disk group monitoring and alerting via Oracle Enterprise Manager

To test email alerts

1. On all nodes (including quorum node) run

   ```
   $ flashgrid-node test-alerts
   ```

2. Check that test alert emails were received from all cluster nodes at each of the configured email addresses.

To modify the list of email alert recipients

As user fg@ on any database node run

```
$ flashgrid-cluster set-email-alerts name1@host1 name2@host2 ...
```

Note that by default the *From* address is set to *flashgrid@localhost.localdomain*. This will ensure that delivery failure notifications are sent to root's mailbox on the originating node, which can help with troubleshooting delivery issues. It is recommended to add this address to the whitelist of senders on the receiving email server and in the email clients.

6 Before Going Live

Before switching the cluster to live use:

1. Verify health of the cluster: `$ sudo skycluster-health-check`
2. Confirm that email alerts are configured and delivered: `$ flashgrid-node test-alerts`
3. Upload diags to FlashGrid support: `$ sudo flashgrid-diags upload-all`
4. Stop the cluster and back up all cluster nodes: https://kb.flashgrid.io/backup-restore/backup-and-restore-in-azure
5. Start the cluster and do final check of the cluster health: `$ sudo skycluster-health-check`

7 Deleting a cluster

To delete a cluster

1. Delete any protection lock(s) for the resource group
2. Delete the resource group corresponding to the cluster

8 Additional Documentation

Maintenance Tasks in Azure: https://www.kb.flashgrid.io/maintenance/maintenance-azure

Backup and Restore Best Practices in Azure: https://www.kb.flashgrid.io/backup-restore/backup-and-restore-in-azure

Troubleshooting: https://www.kb.flashgrid.io/troubleshooting

FlashGrid Storage Fabric CLI Reference Guide: https://www.kb.flashgrid.io/cli-ref/sf-cli

FlashGrid Cloud Area Network CLI Reference Guide: https://www.kb.flashgrid.io/cli-ref/clan-cli

9 Contacting Technical Support

For technical help with SkyCluster please open a support request at https://www.flashgrid.io/support/

To expedite troubleshooting please also collect and upload diagnostic data to the secure storage used by FlashGrid support by running the following command:

```
$ sudo flashgrid-diags upload-all
```

For reporting emergency type of issues that require immediate attention please also use the 24/7 telephone hotline: +1-650-641-2421 ext 7. Please note that use of the 24/7 hotline is reserved for emergency situations only.