

# FlashGrid® SkyCluster™

## Deployment Guide

### for Oracle RAC in Azure Cloud

rev. 18.10-2018.11.17



# Table of Contents

- 1 Introduction ..... 3
- 2 Prerequisites ..... 3
- 3 Deploying a Cluster ..... 4
- 4 After Deploying a Cluster ..... 5
  - 4.1 Verifying cluster health ..... 5
  - 4.2 Verifying synchronization of clocks..... 5
  - 4.3 OS User Accounts ..... 6
  - 4.4 Finalizing Cluster Configuration ..... 6
  - 4.5 Adding a protection lock for the cluster ..... 6
  - 4.6 Deleting a cluster ..... 7
- 5 Measuring Performance ..... 7
- 6 Monitoring Cluster Health ..... 8
- 7 Additional Documentation..... 8
- 8 Contacting FlashGrid Technical Support..... 8

# 1 Introduction

FlashGrid 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 FlashGrid SkyCluster with Oracle RAC in Azure cloud.

Key components of FlashGrid SkyCluster 18.10 for Azure:

- FlashGrid Storage Fabric: ver. 18.10
- FlashGrid Cloud Area Network: ver. 18.07
- FlashGrid Cluster Monitoring and Diagnostics: ver. 18.10
- Oracle Database: ver. 12.2.0.1, 12.1.0.2, or 11.2.0.4
- Oracle Grid Infrastructure: ver. 18.3.0 or 12.2.0.1
- Operating System: Oracle Linux 7, Red Hat Enterprise Linux 7
- Azure VMs: DSv2, DSv3, ESv3, M, GS.
- Disks: Premium SSD

FlashGrid SkyCluster is delivered as Azure Resource Manager templates that automate configuration of multiple components required for a database cluster. FlashGrid 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 FlashGrid 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 provisioning of an Oracle RAC cluster in Azure using FlashGrid SkyCluster Launcher:

- **Azure Storage Blob Container** with Oracle installation files that will be downloaded to the cluster nodes during provisioning. The list of files that must be placed in the Storage Container will be shown in SkyCluster Launcher.
- **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 (must be a single line):

```
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://1810.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. Get IP address of the newly created VM corresponding to the first node
14. SSH to the first (as it was specified on the cluster configuration page) cluster node VM as user `fg@`
15. If the cluster initialization is still in progress or failed then you will see a corresponding welcome message. If there is no welcome message then the cluster initialization has completed successfully.
16. Wait for cluster initialization (including Oracle software installation and configuration) to complete. You will receive a broadcast message when initialization completes or fails. Cluster initialization takes approximately 90 minutes.

# 4 After Deploying a Cluster

## 4.1 Verifying cluster health

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. Clock synchronization is important for FlashGrid Storage Fabric. Clocks difference of 30 seconds or more will cause warnings. Clocks difference of 60 seconds or more will cause disconnect of remotely attached disks.

**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:

- *fg* - the user account used to SSH to the VMs with the SSH key that was selected when creating the cluster configuration. It can also be used for running FlashGrid Storage Fabric or FlashGrid Cloud Area Network utilities. The user *fg* has sudo rights.
- *grid* - Grid Infrastructure owner. GI environment variables are preconfigured.
- *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.

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.

User *fg* has 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 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

# 5 Measuring Performance

DBMS\_RESOURCE\_MANAGER.CALIBRATE\_IO procedure provides an easy way for measuring storage performance including maximum bandwidth, random IOPS, and latency. The CALIBRATE\_IO procedure generates I/O through the database stack on actual database files. The test is read-only and it is safe to run it on any existing database. It is also a good tool for directly comparing performance of two storage systems because the CALIBRATE\_IO results do not depend on any non-storage factors, such as memory size or the number of CPU cores.

### To measure storage performance with CALIBRATE\_IO

1. Create or load a database on the corresponding ASM disk group
2. Make sure the total size of the database files is larger than 5 GB per disk. If needed, create an additional large table space / data file.
3. Customize the first parameter in the SQL code below with the number of disks corresponding to your storage setup. Keep the second parameter (max latency) with the minimum allowed value of 10 milliseconds.
4. Connect to the database with sqlplus and run the customized SQL code.
5. Wait for the CALIBRATE\_IO to complete. This may take 10 to 30 minutes.

### Example of running CALIBRATE\_IO on a 2-node cluster with E64s\_v3 VMs and 16 x 1024 GB Premium disks per node

```
SQL> SET SERVEROUTPUT ON;
DECLARE
lat INTEGER;
iops INTEGER;
mbps INTEGER;
BEGIN DBMS_RESOURCE_MANAGER.CALIBRATE_IO (32, 10, iops, mbps, lat);
DBMS_OUTPUT.PUT_LINE ('Max_IOPS = ' || iops);
DBMS_OUTPUT.PUT_LINE ('Latency = ' || lat);
DBMS_OUTPUT.PUT_LINE ('Max_MB/s = ' || mbps);
end;
/
Max_IOPS = 140605
Latency = 1
Max_MB/s = 3250

PL/SQL procedure successfully completed.
```

## 6 Monitoring Cluster Health

The following methods of monitoring cluster health are available:

- flashgrid-cluster utility displays status of the cluster and its main components
- Alerts about failures are recorded in system log and can be analyzed by 3<sup>rd</sup>-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.

## 7 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>

## 8 Contacting FlashGrid Technical Support

For help with troubleshooting an issue on an existing FlashGrid cluster please use Technical Support Request form located at <https://www.flashgrid.io/support/>

To expedite troubleshooting please also collect diagnostic data by running 'flashgrid-diags upload-all' command on any node and upload it using a secure upload form provided to your company by FlashGrid technical support.

Customers with Mission-Critical SLA subscription may also use the 24x7 telephone hotline for reporting *emergency* type of issues that require immediate attention: +1-650-641-2421 ext 7. Please note that use of the 24x7 hotline is reserved for emergency situations only.



Copyright © 2016-2018 FlashGrid Inc. All rights reserved.

This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document.

FlashGrid is a registered trademark of FlashGrid Inc. Amazon and Amazon Web Services are registered trademarks of Amazon.com Inc. and Amazon Web Services Inc. Oracle and Java are registered trademarks of Oracle and/or its affiliates. Red Hat is a registered trademark of Red Hat Inc. Other names may be trademarks of their respective owners.