ORACLE

NOSQL DATABASE

Deploying Oracle NoSQL Database on the Oracle Bare Metal Cloud Platform

Quick Start White Paper | February 2017 | Version 1.0

ORACLE WHITE PAPER | FEBRUARY 2017

ORACLE

Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Questions or comments on this paper? Please email: oraclenosql-info_ww@oracle.com.

Table of Contents

"Oracle's NoSQL offers value to customers looking at ACID transactions; geodistributed data; application security with authentication and session-level SSL encryption; and integration with Oracle Database, Oracle Wallet, and Hadoop."

"Oracle NoSQL is a key-value database that delivers good performance, scale, security, and high availability capabilities."

SOURCE: THE FORRESTER WAVE™: BIG DATA NOSQL, Q3 2016

Oracle NoSQL Database on Oracle Bare Metal Cloud Overview

Many software engineering organizations are faced with the challenge of building systems to handle extremely-high throughput (10s of thousands writes/sec) while maintaining low latency (< 10 msec). The Oracle NoSQL Database running on the Oracle Bare Metal Cloud Service (BMCS) easily handles these types of workloads in a secure, highly-available environment.

The Oracle NoSQL Database is a best-in-class NoSQL database that provides:

- High-performance distributed read/write capability using share-nothing architecture
- Linear scalability with transparent load rebalancing when new nodes are added
- Kerberos authentication, table-level authorization, and secure client/server and server/server communication
- Highly-configurable ACID transaction model
- Table model with SQL-like query capability

The Oracle BMCS offers hourly metered Bare Metal instances; by eliminating the hypervisor Oracle can deliver better performance at a lower cost than traditional IaaS providers. In addition to compute unencumbered by a hypervisor, Oracle BMCS offers instances with up to 28TBs of locally attached NVMe storage. Each 28TB instance (9 NVMe storage units) is capable of over 3 million 4K IOPs/sec.

Instances in the Oracle BMCS are attached via a 10Gb non-blocking network with no oversubscription. Each node has access to the full performance of the hardware, there are no "noisy neighbors" or hypervisors to share resources with. Instances in the same region are always less than 1ms from each other.

This Quick Start white paper is designed as a reference guide for deploying the Oracle NoSQL Database on the Oracle Bare Metal Cloud platform. The following sections describe the preliminary setup of the BMC environment and then how to run the NoSQL cluster install scripts.

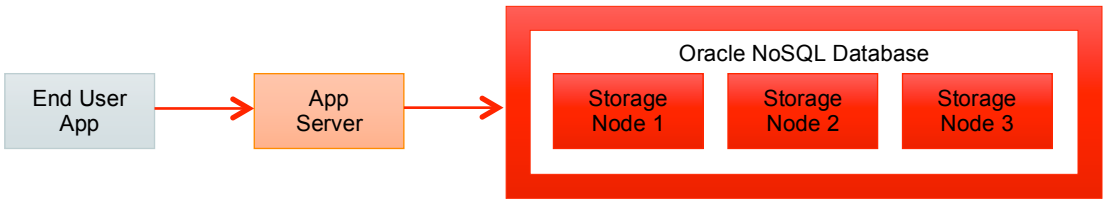
Assumptions

Consumers of this document should –

- Be familiar with the fundamentals of the Oracle BMCS
  - https://docs.us-phoenix-1.oraclecloud.com
- The BMCS walkthrough is highly recommended if this is the first time you have used the platform
  - https://docs.us-phoenix-1.oraclecloud.com/Content/GSG/Reference/overviewworkflow.htm
- Have a basic understanding of Oracle NoSQL
  - Oracle NoSQL Database Concepts Manual
    - http://docs.oracle.com/cd/NOSQL/html/ConceptsManual/Oracle-NoSQLDB-Concepts.pdf

Planning an Oracle NoSQL Database Deployment on BMCS

A minimal configuration for the Oracle NoSQL Database typically consists of a 3-node server cluster with a middle-tier application server between the end-user application and the Oracle NoSQL Database.



The middle tier can be an application server or web server that does not need much storage, but requires significant compute, memory and network resources. The backend database on the other hand usually requires large amounts of persistent storage.

There are 3 instance shapes for Oracle BMCS –

Oracle BMCS Instance Shapes

BM.StandardIO1.36	BM.HighIO1.36	BM.DenseIO1.36
32GB local disk storage 36 cores 256GB memory	12.8TB local NVMe storage 36 cores 512GB memory	28.8TB local NVMe storage 36 cores 512GB memory

Standard instances are suitable for middle-tier services (e.g. web servers); High and Dense shapes have local direct attached storage and are recommended for high-performance database servers.

For this quick start guide we will walk through the steps of allocating a single Standard Shape middle-tier application server and 3 DenseIO NoSQL database server nodes. The database server nodes will be set up in 3 different Availability Domains (ADs) for enhanced reliability.

Outline of Steps to Get Oracle NoSQL Database Up and Running on BMC

The following steps outline the process to get Oracle NoSQL Database up and running on Oracle Bare Metal Cloud.

1. Set up BMC network and provision BMC instances

2. Acquire Oracle NoSQL Database License
3. Copy the Oracle NoSQL Database install scripts, Oracle JDK and Oracle NoSQL Database software to your local machine
4. Run the Oracle NoSQL Database install script and verify the install was successful
5. Run the server install script to install Oracle NoSQL Database components on an application server that will access the the Oracle NoSQL Database.
6. Verify the application server can communicate with the Oracle NoSQL Database.

The details for these steps are contained in the following sections.

Create the Required BMC Network Resources

For this part of the process you will need to go to Oracle Bare Metal Cloud Service at https://console.us-phoenix-1.oraclecloud.com to set up an account, if not already done, then navigate the BMC UI to do the following steps. The Internet Gateway, Route Table, and Subnet configuration described below should be equivalent to the default settings when you create a new Virtual Cloud Network.

1. Create a new Virtual Cloud Network
    a. Name – Oracle_NoSQL_VCN
    b. CIDR block – 10.0.0.0/16
2. Create a new Internet Gateway
    a. Name - Oracle_NoSQL_IG
3. Create a new Route Table
    a. Name - Oracle_NoSQL_RT
    b. CIDR block – 0.0.0.0/0
    c. Target – Oracle_NoSQL_IG
4. Create Subnet 1
    a. Name – AD1_ Oracle_NoSQL_private
    b. AD – PHX-AD-1
    c. CIDR block – 10.0.0.0/24
    d. Route table – Oracle_NoSQL_RT
5. Create Subnet 2
    a. Name – AD2_ Oracle_NoSQL_private
    b. AD – PHX-AD-2
    c. CIDR block – 10.0.1.0/24
    d. Route table – Oracle_NoSQL_RT
6. Create Subnet 3
    a. Name – AD3_ Oracle_NoSQL_private

b. AD – PHX-AD-3

c. CIDR block – 10.0.2.0/24

d. Route table – Oracle_NoSQL_RT

7. Open Oracle NoSQL Ports

a. Go to Networking > Virtual Cloud Networks for your compartment

b. Click on your Virtual Cloud Network (Oracle_NoSQL_VCN)

c. Click Security Lists then click Default Security List for Oracle_NoSQL_VCN

d. Click Edit All Rules and add a rule with:

- Source CIDR: 10.0.0.0/16
- IP PROTOCOL: TCP
- SOURCE PORT RANGE: All
- DESTINATION PORT RANGE: 5000-5050

Create the Required BMC Compute Instances

1. Create BM.DenseIO compute instance 1 to run Oracle NoSQL DB storage node 1

a. Name – Oracle_NoSQL_DB_AD1_0

b. Image – Oracle-Linux-6.8-2017.01.09-0

c. Shape – BM.DenseIO1.512

d. AD – PHX-AD-1

e. Cloud Network – Oracle_NoSQL_VCN

f. Subnet – AD1_Oracle_NoSQL_private

g. SSH Key – <public half of your key pair>

2. Create BM.DenseIO compute instance 2 to run Oracle NoSQL DB storage node 2

a. Name – Oracle_NoSQL_DB_AD2_0

b. Image – Oracle-Linux-6.8-2017.01.09-0

c. Shape – BM.DenseIO1.512

d. AD – PHX-AD-2

e. Cloud Network – Oracle_NoSQL_VCN

f. Subnet – AD2_Oracle_NoSQL_private

g. SSH Key – <public half of your key pair>

3. Create BM.DenseIO compute instance 3 to run Oracle NoSQL DB storage node 3

a. Name – Oracle_NoSQL_DB_AD3_0

b. Image – Oracle-Linux-6.8-2017.01.09-0

c. Shape – BM.DenseIO1.512

d. AD – PHX-AD-3

e. Cloud Network – Oracle_NoSQL_VCN

f. Subnet – AD3_Oracle_NoSQL_private

g. SSH Key – <public half of your key pair>

4. Create one new “Standard Shape” compute instance to run the app server that will interface with your Oracle NoSQL Database

   a. Name – Oracle_NoSQL_AS_AD1_0

   b. Image – Oracle-Linux-6.8-2017.01.09-0

   c. Shape – BM.StandardO1.512

   d. AD – PHX-AD-1

   e. Cloud Network – Oracle_NoSQL_VCN

   f. Subnet – AD1_Oracle_NoSQL_private

   g. SSH Key – <public half of your key pair>

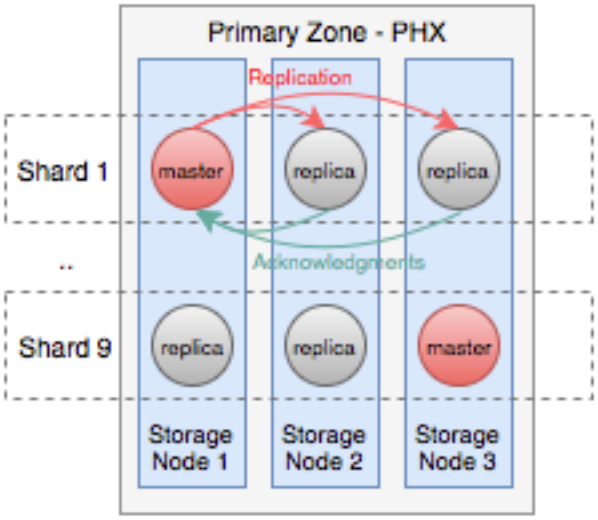
For each instance note the public and private (RFC1918) IP addresses in the table below. These IPs can be found in the BMC UI on the Compute > Instances > Instance Details page.

Instance	Public IP	Private IP
Oracle_NoSQL_DB_AD1_0		
Oracle_NoSQL_DB_AD2_0		
Oracle_NoSQL_DB_AD3_0		
Oracle_NoSQL_AS_AD1_0		

NOTE: All communication between Oracle NoSQL Database nodes and associated app servers should be across the private address of the instances, not the public IPs. Using the public IP adds latency to the connection and limits the bandwidth. Using the RFC1918 private IP guarantees access to the full network bandwidth and the lowest possible latency. The public IPs can be used for external access, e.g. to ssh to the machines or as parameters to the Oracle NoSQL Database install scripts.

Oracle NoSQL Database Zone and Replication Architecture

The preceeding compute instance setup has 3 BMC DenseIO instances (Storage Nodes), each having 9 3TB NVMe drives. By default the Oracle NoSQL Database uses a replication factor (RF) of 3 resulting in the following single-zone layout consisting of 9 shards.



In this case there is a 1-to-1 mapping between shards and the number of NVMe drives on each machine.

Installing the Oracle NoSQL Database on BMCS

The following subsections describe the steps needed to install the Oracle NoSQL database on BMCS.

Download Oracle NoSQL Database Software and Oracle JDK

Depending on which software license you own, you can download either Oracle NoSQL Database Enterprise Edition, Basic Edition or Community Edition as follows:

1) **Enterprise Edition** – Requires a commercial license from Oracle. This version can be downloaded for commercial use from the Oracle Software Download Cloud (edelivery.oracle.com).
2) **Basic Edition** – If you have an Oracle Database Enterprise Edition license you may use the Basic Edition at no additional cost, provided you are in compliance with the terms of the Oracle Database Enterprise Edition license agreement. You can download the Basic Edition from the Oracle Software Download Cloud (edelivery.oracle.com)
3) **Community Edition** – This version can be downloaded from the Oracle Technology Network (OTN) downloads page at:

   http://www.oracle.com/technetwork/database/database-technologies/nosqldb/downloads/index.html

As an example, the Oracle NoSQL Database Community Edition can be downloaded to your local machine using the following command.

```
$ wget http://download.oracle.com/otn-pub/otn_software/nosql-database/kv-ce-4.3.11.tar.gz
```

Later, during the install process described below, this tar file will be uploaded to the BMC compute nodes.

Download the latest 64-bit JDK for RPM-based Linux Platforms (Note: Oracle NoSQL Database requires Java version 1.8 or greater). For example, go to

http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html

and click on the link to download jdk-8u121-linux-x64.rpm to download the .rpm to your local machine.

Both the NoSQL .tar.gz file and the JDK .rpm files should be placed in the same directory as your NoSQL cluster install scripts so that the install scripts can copy them to the BMC compute nodes.

Scripts to Install and Configure the Oracle NoSQL Database

The install scripts allow you to install and configure the Oracle NoSQL Database from your local machine. The scripts can be run from a BASH or OS X shell.

These scripts will automatically:

1. Install the NVMe file system on BMC nodes
2. Configure the BMC network and firewall ports for NoSQL storage nodes
3. Install operating system utilities and services needed for the Oracle NoSQL database, including
    a. Oracle JDK
    b. NTP service
4. Install the Oracle NoSQL Database
5. Configure and start up the Oracle NoSQL Database

The Oracle NoSQL Database cluster install script takes as input a list of IP addresses and runs the node install on each node in the IP list. The NoSQL cluster install script has the following form.

```
$ ./ons_cluster_install.sh --zone <zoneid> --store <dbname> <ipaddrs>
```

The script takes a zone name (e.g. PHX), a store value, which is the name of the Oracle NoSQL Database (e.g. ONSDB), and a list of IP addresses, which are the Public IP Addresses obtained from the BMCS Compute > Instances > Instance Details page as noted above.

ons_cluster_install.sh parameters:

parameter	description
--zone	name of the zone where the cluster will be installed
--store	name of the database that will be installed on the cluster
--capacity	number of Replication Nodes a Storage Node supports
--partitions	Number of partitions per shard
--username	admin user for the cluster when security is enabled
--passphrase	admin user's password

The ons_cluster_install.sh script runs on a machine that has public key access to the database nodes; for example the machine with the public key that was provided during BMC compute instance creation above.

If any of the ons_cluster_install.sh parameters are missing the install program will prompt the user if necessary, as with the following example.

```
./ons_cluster_install.sh --zone PHX --store kvstore 129.146.1.1 129.146.2.2 129.146.3.3
Enter a passphrase to create a secure database.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Enter username (admin):
Enter capacity (9): 3
Enter partitions (90): 30
```

The cluster install script copies the local Oracle JDK .rpm and Oracle NoSQL Database .tar.gz to each node in the cluster. It then copies the ons_node_install.sh script to each node, runs the node install scripts in parallel and waits for them to complete. Once the installs are complete the admin node (first node in the list) is given a deployment plan that it uses to deploy the database to all storage nodes in the cluster.

The NoSQL install directory on your local machine should include the following files:

```
ons_cluster_install.sh     ons_node_install.sh        ons_server_install.sh
```

Your local install directory should also include some variation of the following:

```
jdk-8u121-linux-x64.rpm    kv-ce-4.3.11.tar.gz
```

The install scripts will automatically pick up the jdk and kv files based on their prefix/suffix patterns.

The following subsections provide some code snippets from the ons_node_install.sh script. These snippets are meant to give insight into the BMC system configuration.

Script Details – File System Setup

The BMCS is delivered as a raw, bare-metal machine with Oracle Linux installed. It is necessary to do some initial hardware configuration, including formatting and mounting drives. The node install script runs the following file-system installation procedure on each node in the cluster.

```
#!/bin/bash

>/tmp/fstab

for nvme in `sudo fdisk -l | grep nvme | sort | cut -f 2 -d ' ' | sed 's/://g'`; do
  echo -e "n\np\n1\n\n\nw" | sudo fdisk -u -c $nvme
  sudo mkfs.ext4 $nvme
  mount_dir=`echo $nvme | sed 's/dev/ons/g'`
  sudo mkdir -p $mount_dir
  sudo mount $nvme $mount_dir
  sudo chown opc:opc $mount_dir
  printf "$nvme\t$mount_dir\text4\tdefaults\t0\t0\n" >> /tmp/fstab
done

sudo su -c "cat /tmp/fstab >> /etc/fstab"

exit 0
```

Once the node install script completes you should see the following mounted volumes on each of the NoSQL nodes.

```
$ df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda3        38G  2.4G   34G   7% /
tmpfs           252G     0  252G   0% /dev/shm
/dev/sda1       543M  280K  543M   1% /boot/efi
/dev/nvme0n1    2.9T   72M  2.8T   1% /ons/nvme0n1
/dev/nvme1n1    2.9T   72M  2.8T   1% /ons/nvme1n1
/dev/nvme2n1    2.9T   72M  2.8T   1% /ons/nvme2n1
/dev/nvme3n1    2.9T   72M  2.8T   1% /ons/nvme3n1
/dev/nvme4n1    2.9T   72M  2.8T   1% /ons/nvme4n1
/dev/nvme5n1    2.9T   72M  2.8T   1% /ons/nvme5n1
/dev/nvme6n1    2.9T   72M  2.8T   1% /ons/nvme6n1
/dev/nvme7n1    2.9T   72M  2.8T   1% /ons/nvme7n1
/dev/nvme8n1    2.9T   72M  2.8T   1% /ons/nvme8n1
```

Script Details – Network Configuration

The Oracle NoSQL database needs to open ports for client/server and server/server communication. These port numbers are typically in the range 5000+ and will fit in the range 5000-5050 for a 9-shard cluster. Each 9-NVMe node requires 20 ports: 1 admin port, 1 client port and 2 ports per shard for inter-server communication. The node install script opens up ports as follows.

For Oracle-Linux release versions less than 7.0 it uses:

```
$ sudo /sbin/iptables -D FORWARD -j REJECT --reject-with icmp-host-prohibited
$ sudo /sbin/iptables -D INPUT -j REJECT --reject-with icmp-host-prohibited
```

For Oracle-Linux release versions 7.0 and greater it uses:

```
$ sudo firewall-cmd --permanent --zone=public --add-rich-rule='rule family="ipv4"
source address="10.0.0.0/27" port protocol="tcp" port="5000-5050" accept'
```

Script Details – NTP Installation

NTP is installed to keep the clocks sync'd on the NoSQL database nodes. This is required for NoSQL replication.

Commands to install NTP and add ntpd to the boot configuration:

```
$ sudo yum -y install ntp
$ sudo service ntpd start
$ sudo chkconfig ntpd on
```

Verifying the Oracle NoSQL Database Installation was Successful

Once the install has completed you can log on to any of the storage nodes in the Oracle NoSQL Databaes Cluster (using ssh opc@<database IP>) and verify the install was successful by issuing the following command :

```
$ java -jar $KVHOME/lib/kvstore.jar ping -host `hostname` -port 5000
```

This should produce something along the lines of the following:

```
Pinging components of store ONSDB based upon topology sequence #336
270 partitions and 3 storage nodes
Time: 2017-01-19 17:55:09 UTC   Version: 12.1.4.3.11
Shard Status: healthy:9 writable-degraded:0 read-only:0 offline:0
Admin Status: healthy
Zone [name=PHX id=zn1 type=PRIMARY allowArbiters=false]   RN Status: online:24 offline:0
maxDelayMillis:1 maxCatchupTimeSecs:0
Storage Node [sn1] on Oracle_NoSQL_DB_AD1_0:5000    Zone: [name=PHX id=zn1 type=PRIMARY
Storage Node [sn2] on Oracle_NoSQL_DB_AD2_0:5000    Zone: [name=PHX id=zn1 type=PRIMARY
Storage Node [sn3] on Oracle_NoSQL_DB_AD3_0:5000    Zone: [name=PHX id=zn1 type=PRIMARY
```

Installing Oracle NoSQL Database Components on an Application Server

As described above an application server is a middle-tier server that sits between between a client application and the Oracle NoSQL Database. An application server needs java components from the Oracle NoSQL Database to communicate with the Oracle NoSQL Database storage nodes. The ons_server_install.sh script installs these components and has the following form:

```
$ ./ons_server_install.sh --server <ipaddr> --dbnodes <ipaddrs>
```

The server parameter is the server address where NoSQL server software will be installed. In the example configuration it would be the IP address of the Oracle_NoSQL_AS_AD1_0 server from the Creating Required BMC Compute Instances section above. The database nodes are the database public IP addresses from the Oracle NoSQL Database cluster, e.g. the IPs for Oracle_NoSQL_DB_* from above.

Connecting to the Oracle NoSQL Database from the Application Server

Once the application server is installed you should be able to log in (using account 'opc') and verify that it can communicate with the Oracle NoSQL Database.

Here's an example test that can be run on the app server to verify things are working as expected:

```
$ ssh opc@<appserver IP>
$ printf "1\n2\n3">/tmp/t.dat
$ java -Xmx256m -Xms256m -jar $KVHOME/lib/sql.jar -helper-hosts
Oracle_NoSQL_DB_AD1_0:5000 -store ONSDB
sql-> create table t(i INTEGER, PRIMARY KEY (i));
Statement completed successfully
sql-> import -table t -file /tmp/t.dat CSV
Loaded 3 rows to t.
sql-> select i from t order by i;
 +---+
 | i |
 +---+
 | 1 |
 | 2 |
 | 3 |
 +---+
```

Adding Multiple NoSQL Data Stores to BMC Compute Instances

By default the Oracle NoSQL Database install scripts create a single data store on a set of BMC compute instances. In some cases this might not be the best use of these compute resources. For instance, during an intitial development phase you may want separate data stores for development, test and production that run on the same BMC hardware. These custom configurations can be implemented by running the ons_cluster_compile.sh script. For example, to install 3 databases – dev, test, and prod – on a single cluster run the following command:

```
./ons_cluster_compile.sh --zone PHX --store dev,test,prod -capacity 3 -partitions 30 -P
"" 129.146.1.1 129.146.2.2 129.146.3.3
```

Appendix A - Running Oracle NoSQL Database on BMC VMs

It is possible to install Oracle NoSQL Database on BMC VMs. The following gives an example of how to create a 3-node Oracle NoSQL Database with each node having 3 block volumes attached.

1. Create the Required Block Volumes

   From the BMC UI (Storage>Block Volumes) create 9 block volumes, 3 in each Availability Domain (AD) :

   Oracle_NoSQL_BV_AD1_0, Oracle_NoSQL_BV_AD1_1, Oracle_NoSQL_BV_AD1_2
   Oracle_NoSQL_BV_AD2_0, Oracle_NoSQL_BV_AD2_1, Oracle_NoSQL_BV_AD2_2
   Oracle_NoSQL_BV_AD3_0, Oracle_NoSQL_BV_AD3_1, Oracle_NoSQL_BV_AD3_2

   Each block volume should have 256GB of storage.

2. Create the Compute Instances

   From the BMC UI (Compute>Instances) create 3 VMs. Follow the instructions from the 'Create the Required BMC Compute Instances' section above, except use VM.Standard1.4 shapes instead of BM.DenseIO shapes.

3. Attach the Block Volumes to the Compute Instances

   Click on each virtual compute instance and add all three block volumes for the associated availability domain.

4. Run the install program

   Run ons_install_cluster as described previously, e.g.:

```
./ons_cluster_install.sh --zone PHX --store ONSDB 129.146.1.1 129.146.2.2 129.146.3.3
```

Further Information about the Oracle NoSQL Database

Check out the following links for more information about the Oracle NoSQL Database

- http://www.oracle.com/technetwork/database/database-technologies/nosqldb/overview/index.html
- http://www.oracle.com/technetwork/database/database-technologies/nosqldb/documentation/index.html

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

ORACLE

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Deploying Oracle NoSQL on the Oracle Bare Metal Cloud Service
January 2017
Author: Rick George (rick.george@oracle.com)
Contributing Authors: Craig Carl (craig.carl@oracle.com)