ORACLE

Zero Data Loss Recovery Appliance X8 Performance Proof of Concept

February 17, 2021 | Version 1.0

Public

PURPOSE STATEMENT

This technical brief details the setup and results of an Oracle Database backup and recovery performance Proof of Concept (PoC) conducted by the Oracle Solution Center with Oracle Exadata X8-2 and Zero Data Loss Recovery Appliance X8.

INTENDED AUDIENCE

This document is intended for the Database Administrator and Recovery Appliance Administrator. There should be a high-level understanding of Zero Data Loss Recovery Appliance and Oracle Database backup and recovery via Recovery Manager (RMAN).

DISCLAIMER

TABLE OF CONTENTS

INTRODUCTION

Oracle's Zero Data Loss Recovery Appliance (ZDLRA) is the premier data protection system designed for the Oracle Database. The appliance eliminates data loss exposure and dramatically reduces backup overhead on production servers through innovative real-time transaction protection. Built on the industry-leading Oracle Exadata platform with high performance and availability, the Recovery Appliance scales out to handle backup and recovery needs for all sizes and volume of databases.

This technical brief highlights the results of a restore performance proof of concept (PoC) conducted by the Oracle Solution Center at the request of a large financial services customer.

The customer's PoC objective was to achieve restoration of a 120 TB database within 10 hours (12 TB/hour) using the Recovery Appliance.

With the test environment described in this brief, the restore from Recovery Appliance completed in less than 4 hours, with average 35 TB/hour restore rate, achieving more than twice the performance speed versus the PoC objective.

TEST ENVIRONMENT

Recovery Appliance

Recovery Appliance X8 Base Rack + 6 Storage Servers (9 storage servers total)
See Recovery Appliance Configuration (B) for details

The ingest/backup network on each compute server consists of two 25 Gb Ethernet ports configured with active-active bonding via Link Aggregation Communication Protocol (LACP). This configuration allowed a total Recovery Appliance system ingest bandwidth of 4x25 Gb (100 Gb/sec) via the ingest network scan IP address, which distributes network traffic between the two Recovery Appliance compute nodes.

Test Database

An Oracle 19c Database (service name: `db19c`) was created as an 8-node Real Application Clusters (RAC) on Exadata X8-2 Full Rack High Capacity system with 120 TB of data. Recovery Manager (RMAN) was used for backup & restore operations.
See Exadata Configuration (C) for details

Database Backup & Restore Network

Each Exadata database node was configured with 25Gb active-active bonding via LACP, i.e. bondeth0, with scan IP address setup for the RAC Test Database.

25 Gb Network Switch

Arista DCS-7504N
Hardware version: 13:00
See Network Switch Configuration (D) for details

Performance Monitoring

Oracle Enterprise Manager (OEM) 13.4 and Linux OS SAR performance tool was used to monitor backup, restore, and system performance.

TEST CASE

Prior to the restore test, an initial RMAN full (Level 0) backup of the test database was taken to the appliance.

Note: the appliance supports incremental forever strategy, which means that after the initial full backup, only incremental backups are required thereafter. Incremental backups received by the appliance are virtualized as new full backups that can be restored as normal. See the product data sheet under Additional Resources for more information.

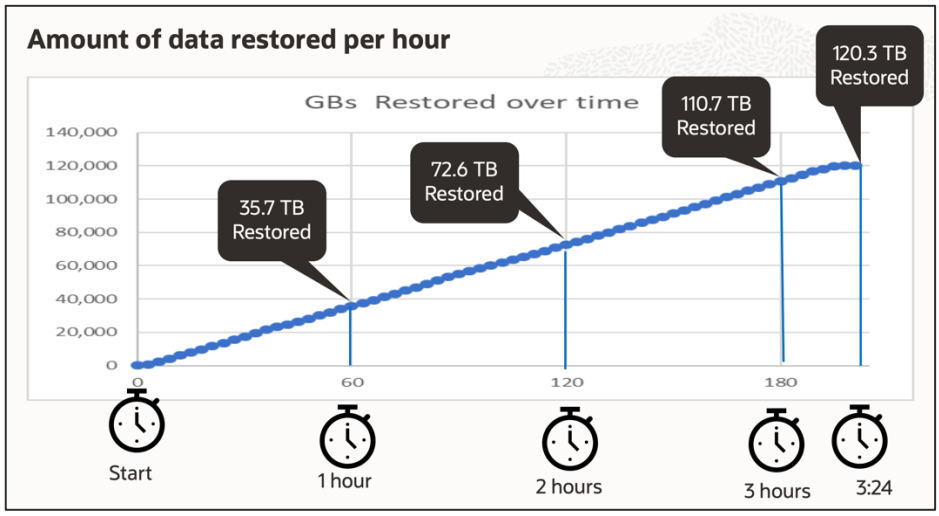
To conduct the restore test, RMAN connects to the RAC database service name and dynamically allocates 64 channels across all nodes so that the restore operation can leverage all available resources.

```
> RMAN TARGET <SYS or SYSBACKUP user>/<pwd>@<prot DB service name, e.g. db19c>
       CATALOG <VPC user/<pwd>@<RA service name, e.g. zdl1>

RMAN> CONFIGURE DEVICE TYPE sbt_tape PARALLELISM 64
  PARMS='SBT_LIBRARY=/u01/app/oracle/product/19.0.0.0/dbhome_1/lib/libra.so,
  ENV=(RA_WALLET=location=file:/u01/app/oracle/product/19.0.0.0/dbhome_1/dbs/zdlra
  credential_alias=zdl1ingest-scan:1521/zdl1:dedicated)' FORMAT '%U_%d' ;

RMAN> RUN {
  RESTORE DATABASE;
  RECOVER DATABASE;
 }
```

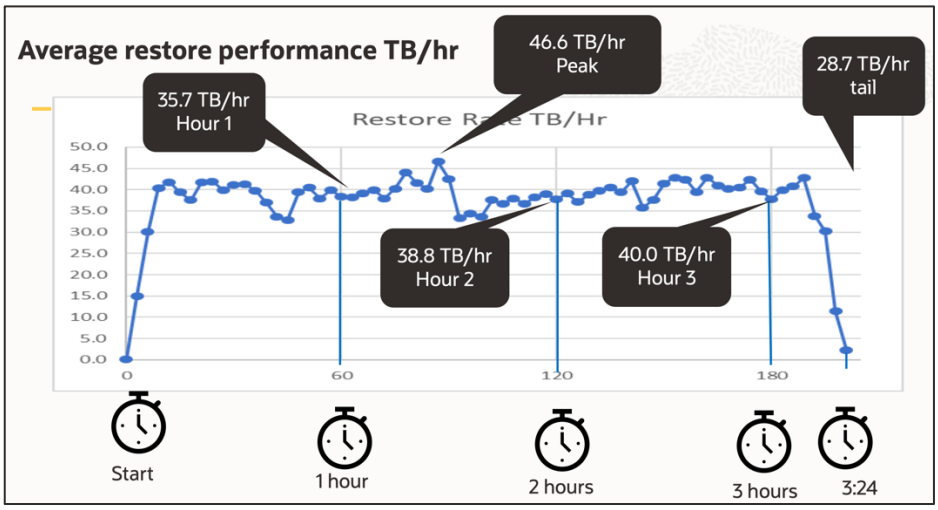
For documentation on procedural details, see RAC Database Restore and Recovery Steps (A).

With this configuration, the RESTORE operation for the 120 TB database completed in 3 hours and 24 min, yielding 35 TB/hour overall restore rate with peak performance measured at 46 TB/hr.

The below graphs show additional details on Data Restored per Hour and Average Restore Performance within each of the first 3 hours, along with the Peak rate achieved.



Data Restored Per Hour



Average Restore Performance Per Hour

For comparison, a second test was conducted where the restore was performed using RMAN channels allocated on a single RAC node versus across all 8 nodes as previously described. The single node test yielded lower performance, as node resources became the overall bottleneck, not the network. Net-net, this shows that leveraging multiple nodes for restore effectively maximizes network consumption versus single node. Results of both tests are shown in Restore Test Cases (E).

CONCLUSION

This Proof of Concept demonstrated how to achieve 35 TB/hour sustained restore rate of 120 TB database running on Exadata X8-2 Full Rack High Capacity Rack from a Recovery Appliance X8 system with 9 Storage Servers over 4x25 Gb Ethernet network.

These results were achieved without any special tuning or configuration on the Recovery Appliance.

ADDITIONAL RESOURCES

- **Zero Data Loss Recovery Appliance Product Central**
  (https://www.oracle.com/zdlra)
- **Product Data Sheet**
  (https://www.oracle.com/technetwork/database/availability/recovery-appliance-ds-2297776.pdf)
- **Technical Resources & Customer Case Studies**
  (https://tinyurl.com/zdlraresources)
- **Oracle Solution Center**
  (https://www.oracle.com/osc)

APPENDIX: CONFIGURATION AND TEST DETAILS

(A) RAC Database Restore and Recovery Steps:

Improve RMAN restore database performance by using Multi-Instance Database Restore (MOS Note 2710709.1)

(B) *Recovery Appliance Configuration*

- Recovery Appliance X8 Base Rack + 6 Storage Servers
- 9 Storage servers total
- 2 compute nodes: zdl1db01 and zdl1db02
- ZDLRA SW release: ra_automation-19.2.1.1.1.202003-31138109.x86_64
- 25 Gb Ingest Network Interface: bondeth0

(C) Exadata Configuration

- Exadata X8-2 Full Rack High Capacity with nodename: exa11db[01 – 08]
- 25Gb Client Network interface used for backup: bondeth0
- 14 Storage servers total
- Oracle Database 19c Enterprise Edition (19.8.0.0)
- Database configured as 8-node RAC (database name: db19c and 8 instance name: db19c[1 – 8])
- Database sized at 120 TB, as follows:

Tablespace Used	Used MB	Free MB	Total MB	Pct. Free
SYSTEM	988	1,060	2,048	52
UNDOTBS2	88	3,687	3,775	98
UNDOTBS8	115	3,260	3,375	97
SYSAUX	2,655	3,489	6,144	57
UNDOTBS1	115	3,560	3,675	97
TBS_DATA3	30,332,010	76,687	30,408,697	0
UNDOTBS5	106	3,069	3,175	97
UNDOTBS6	135	3,615	3,750	0
TBS_DATA1	30,331,715	76,982	30,408,697	0
TBS_DATA2	30,332,423	76,274	30,408,697	0
TBS_DATA4	30,332,356	76,341	30,408,697	0
USERS	3	2,045	2,048	100
UNDOTBS3	85	3,690	3,775	98
UNDOTBS7	118	3,207	3,325	96
UNDOTBS4	126	3,399	3,525	96

(D) 25 Gb Network Switch Configuration

Arista DCS-7504N
Hardware version: 13.00

Port configuration to Recovery Appliance

```
aristacore1(s1)#sh inter status | include zdl1

Et3/8/1   lacp_zdl1_ingest1_pc109      connected   in Po109  full   25G   100GBASE-SR4
Et3/8/2   lacp_zdl1_ingest1_pc109      connected   in Po109  full   25G   100GBASE-SR4
Et3/8/3   lacp_zdl1_ingest2_pc110      connected   in Po110  full   25G   100GBASE-SR4
Et3/8/4   lacp_zdl1_ingest2_pc110      connected   in Po110  full   25G   100GBASE-SR4

Et3/9/1   lacp_zdl1_repl1_pc111        connected   in Po111  full   10G   40GBASE-SR4
Et3/9/2   lacp_zdl1_repl1_pc111        connected   in Po111  full   10G   40GBASE-SR4
Et3/9/3   lacp_zdl1_repl2_pc112        connected   in Po112  full   10G   40GBASE-SR4
Et3/9/4   lacp_zdl1_repl2_pc112        connected   in Po112  full   10G   40GBASE-SR4

Po109     lacp_zdl1_pc109_core1_3/8/1-2 connected   1302     full   50G   N/A
Po110     lacp_zdl1_pc110_core1_3/8/3-4 connected   1302     full   50G   N/A
Po111     lacp_zdl1_pc111_core1_3/9/1-2 connected   1501     full   20G   N/A
Po112     lacp_zdl1_pc112_core1_3/9/3-4 connected   1501     full   20G   N/A

aristacore1(s1)#sh run inter po 109
interface Port-Channel109
  description lacp_zdl1_pc109_core1_3/8/1-2
  switchport access vlan 1302

aristacore1(s1)#sh run inter po 110
interface Port-Channel110
  description lacp_zdl1_pc110_core1_3/8/3-4
  switchport access vlan 1302

aristacore1(s1)#sh run inter po 111
interface Port-Channel111
  description lacp_zdl1_pc111_core1_3/9/1-2
  switchport access vlan 1501

aristacore1(s1)#sh run inter po 112
interface Port-Channel112
  description lacp_zdl1_pc112_core1_3/9/3-4
  switchport access vlan 1501

interface Ethernet3/8/1
  description lacp_zdl1_ingest1_pc109
  speed forced 25gfull
  channel-group 109 mode active
interface Ethernet3/8/2
  description lacp_zdl1_ingest1_pc109
```

```
  speed forced 25gfull
  channel-group 109 mode active

interface Ethernet3/8/3
  description lacp_zdl1_ingest2_pc110
  speed forced 25gfull
  channel-group 110 mode active
interface Ethernet3/8/4
  description lacp_zdl1_ingest2_pc110
  speed forced 25gfull
  channel-group 110 mode active

interface Ethernet3/9/1
  description lacp_zdl1_repl1_pc111
  speed forced 10000full
  channel-group 111 mode active
interface Ethernet3/9/2
  description lacp_zdl1_repl1_pc111
  speed forced 10000full
  channel-group 111 mode active

interface Ethernet3/9/3
  description lacp_zdl1_repl2_pc112
  speed forced 10000full
  channel-group 112 mode active
interface Ethernet3/9/4
  description lacp_zdl1_repl2_pc112
  speed forced 10000full
  channel-group 112 mode active
```

Port Configuration to Exadata

```
aristacore1(s1)#sh inter status | include exa11
Et3/19/1   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/19/2   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/19/3   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/19/4   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/20/1   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/20/2   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/20/3   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/20/4   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/21/1   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/21/2   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/21/3   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/21/4   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
Et3/22/1   exa11-1to16              connected    1302     full   25G    100GBASE-SR4
```

Et3/22/2	exa11-1to16	connected	1302	full	25G	100GBASE-SR4
Et3/22/3	exa11-1to16	connected	1302	full	25G	100GBASE-SR4
Et3/22/4	exa11-1to16	connected	1302	full	25G	100GBASE-SR4

(E) ***Restore Test Cases***

Multiple (8) node restore (`RESTORE DATABASE`) - 64 channels – 3h 24m

120 TB / 204 min = 588.23 GB/min = 35.29 TB / hour

Single node restore (`RESTORE DATABASE`) - 32 channels – 11h 57min 43s

120 TB / 718 min = 167.13 GB / min = 10.03 TB / hour

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com.
Outside North America, find your local office at oracle.com/contact.

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