Zero Data Loss Recovery Appliance: Insider's Guide to Architecture & Best Practices

Tim Chien
Director of Product Management
Oracle Development

Jony Safi
Senior Manager
Oracle Development

Stefan Reiners
DBA
METRO-nom GmBH

October 24, 2018

ORACLE
OPEN
WORLD

ORACLE®

Safe Harbor Statement

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, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Agenda

1 Introduction & Architecture

2 Operational & Monitoring Best Practices

3 METRONOM Case Study

4 Resources & Next Steps

ORACLE®

Agenda

1 Introduction & Architecture

2 Operational & Monitoring Best Practices

3 METRONOM Case Study

4 Resources & Next Steps

ORACLE®

Traditional Backup Solutions Are Not Designed for Database

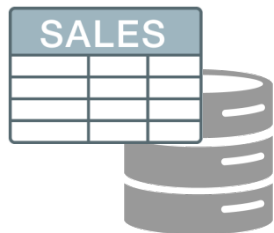**Treat Databases as Just Files to Periodically Copy**

Data Loss Exposure

Lose all data since last backup – no validation

Daily Backup Window

Large performance impact on production

Poor Database Recoverability

Many files are copied but protection state of database is unknown

Many Systems to Manage

Scale by deploying more backup appliances

ORACLE®

Zero Data Loss Recovery Appliance

Engineered Data Protection
For Complete Recoverability
of the Oracle Database

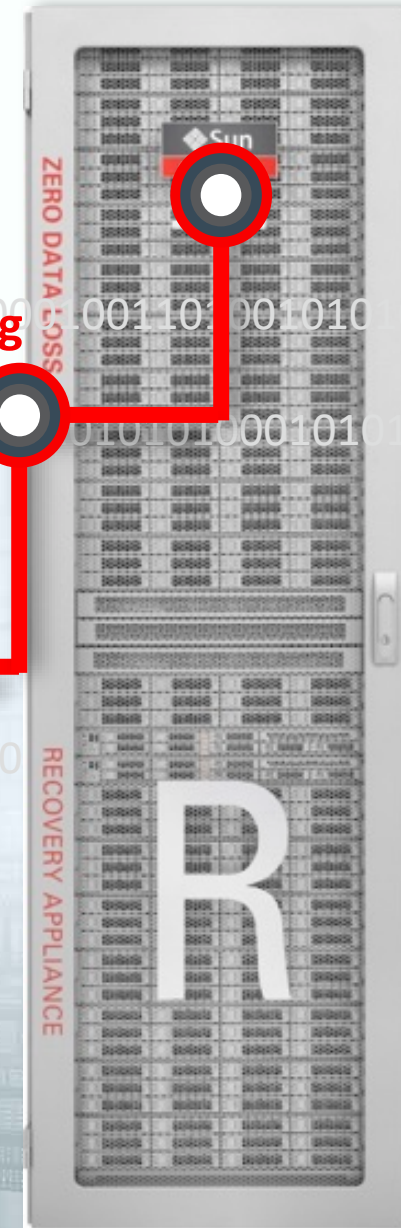Enterprise Databases & Applications

End-to-End
Data Protection
Control & Monitoring

Scale-Out
Hardware & Storage
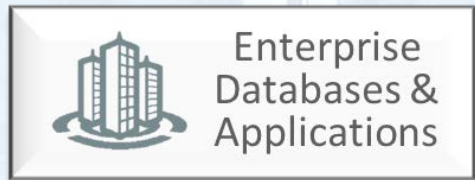
Eliminate Long
Backup Windows

Continually Validates
Recovery Status

Archive Backup to
Cloud Storage

Reliable & Complete
Recovery

ORACLE

Recovery Appliance Architecture

Protected DBs

Compressed Delta Store

Delta Push

Changed blocks and Real-Time Redo (no full backups)

Day 0 — Full

Day 1 — Incremental

Day N — Incremental

Day N — Virtual Full

Virtual Full

Data validation on receipt, copy, restore, periodically
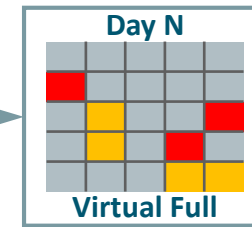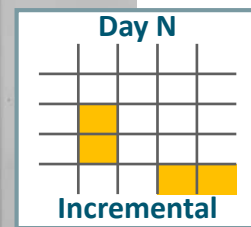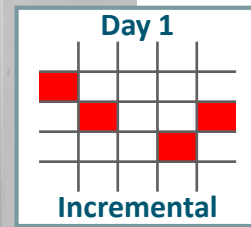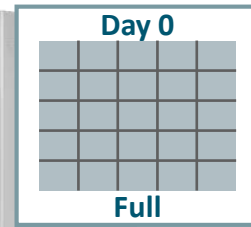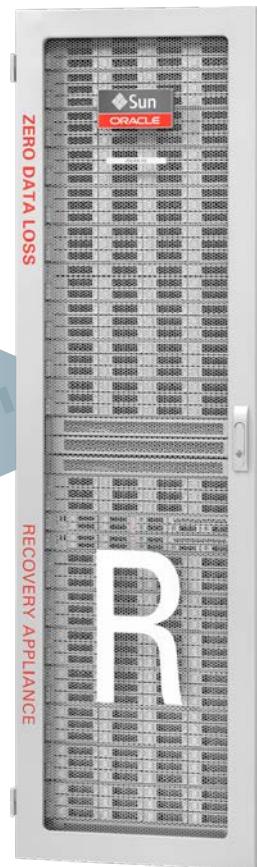
DB Backup Cloud Service — Cloud Archive

Tape Backup

Remote Replica

Agenda

1 Introduction & Architecture

2 Operational & Monitoring Best Practices

3 METRONOM Case Study

4 Resources & Next Steps

ORACLE®

MAA Recommendations

- **Do not make any changes to the Recovery Appliance**

- One[1] Recovery Appliance (RA) system per data center[2]

- Backup primary and standby databases to their respective local RA

- No RA replication for any database with a remote standby

  - Restore operation can use any RA in any location

[1] It can be an interconnected ZDLRA configuration

[2] The definition of a Data Center also includes "fire rooms"

RA Replica: Data Loss Protection from Site Disasters

Local Data Center

Remote Data Center

One Way

Bi-Directional

Hub & Spoke

Tape Library
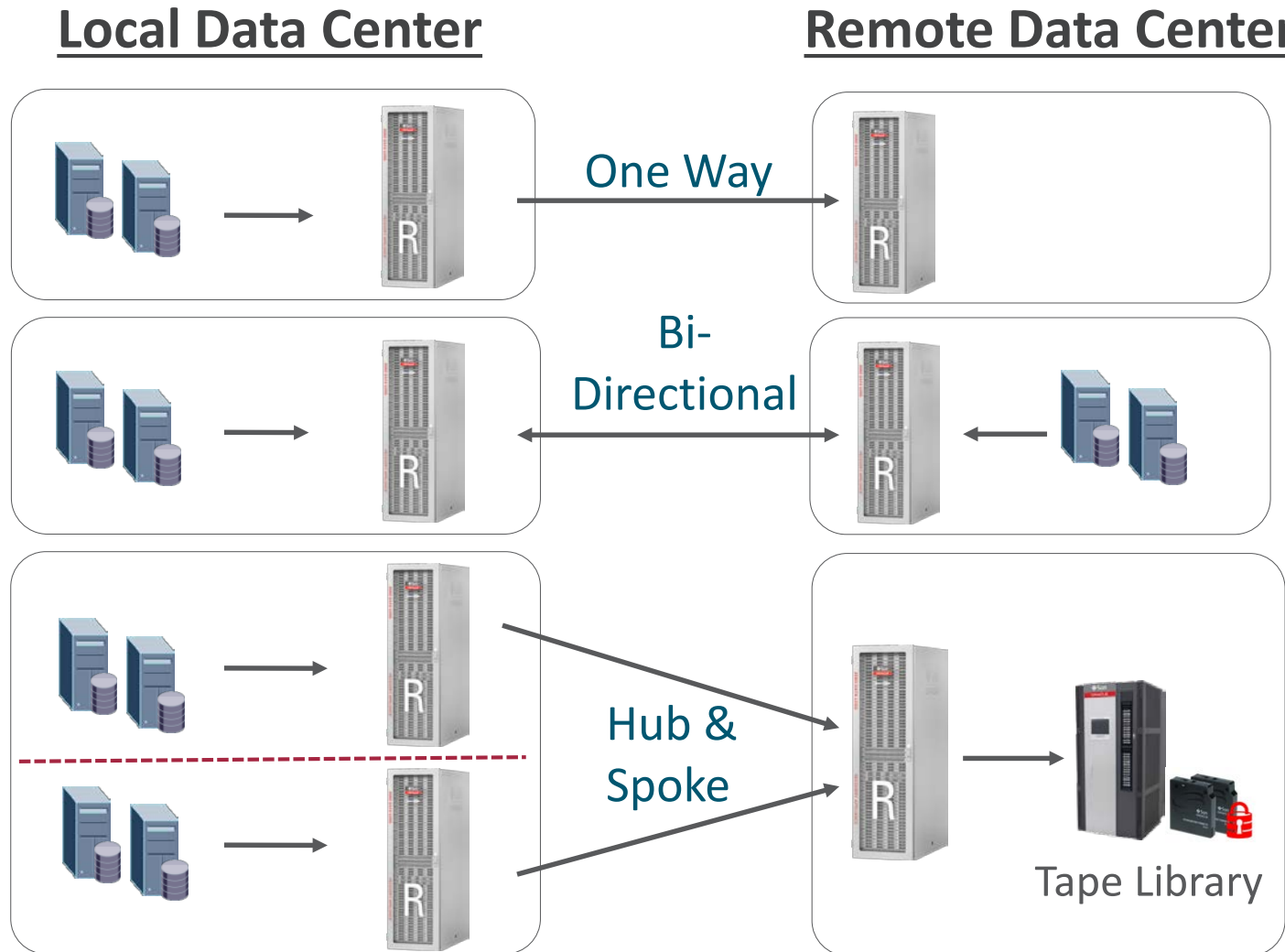
BENEFITS

- **Replication to Remote Appliance protects data from disasters or site failures**

- **Automated restore from Local Appliance or directly from Remote Appliance**

ORACLE®

Leading North America Healthcare Company

1000+ Protected DBs, 3 Data Centers, Bi-Directional + Hub & Spoke Replication

East Coast DC #1 (1000 DBs)

PROD

NON-PROD

X

X

Primary DBs
->Local RA

Standalone DBs
->Local RA
->Replicated

R R

East Coast DC #2 "Hub" (200 DBs)

DR

NON-PROD

X

X

Standby DBs
->Local RA
->DB Clone

Standalone
->Local RA
->Replicated

R R

Midwest DC (50 DBs)

PROD, NON-PROD

X

Standalone
->Local RA
->Replicated
->DB Clone

R

Replication

Replication

ORACLE®
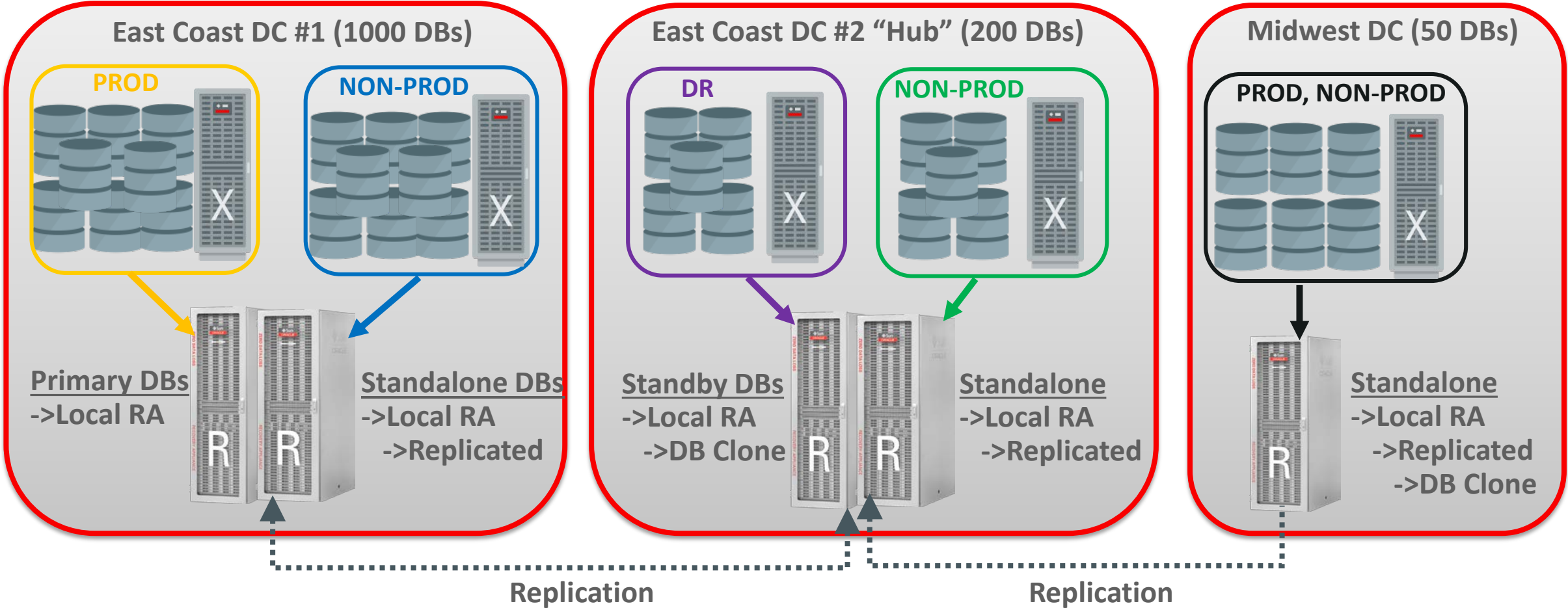
NEW: ZDLRA High Availability for Backup & Recovery

Preserve Backup & Recovery Continuity during Planned or Unplanned ZDLRA Outages

- Database backup and redo operations **automatically failover** to downstream replica appliance.

- ZDL and point-in-time recovery operations from downstream appliance are **fully supported.**

- Virtual full backups on primary appliance are **synchronized to current state** when back online

ORACLE®
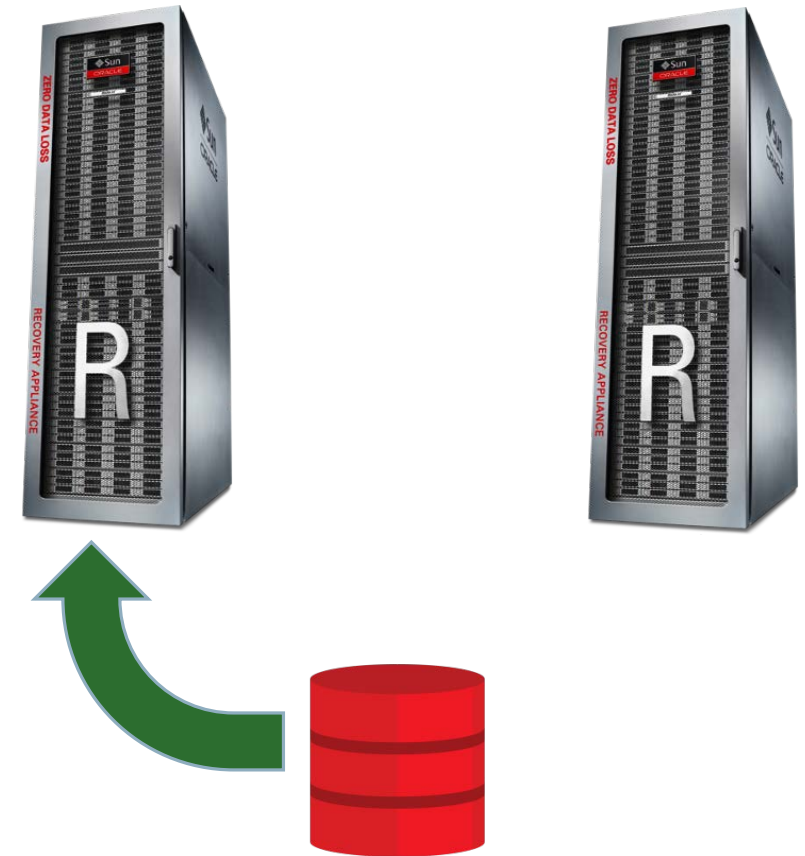
ZDLRA High Availability for Backup & Recovery

Backup Failover to Alternate Appliance

- Incrementals and Redo normally sent to Primary RA

- Alternate RA serves as backup staging area when primary RA is unavailable, then syncs with primary RA afterwards

  – No virtual fulls created on alternate, hence recoverability not supported

  – Space sized for 'n' incrementals and archived log backups during primary downtime period

- **Benefits:**

  – **Preserves backup and redo shipment continuity during planned maintenance / upgrades**

  – **Prevents local Fast Recovery Areas from filling up with archived logs**

  – **Incremental forever backups continue**

RA Boston Primary Appliance

Incrementals & Redo

RA NYC Alternate Appliance

Alternate Appliance takes over backups and redo transport
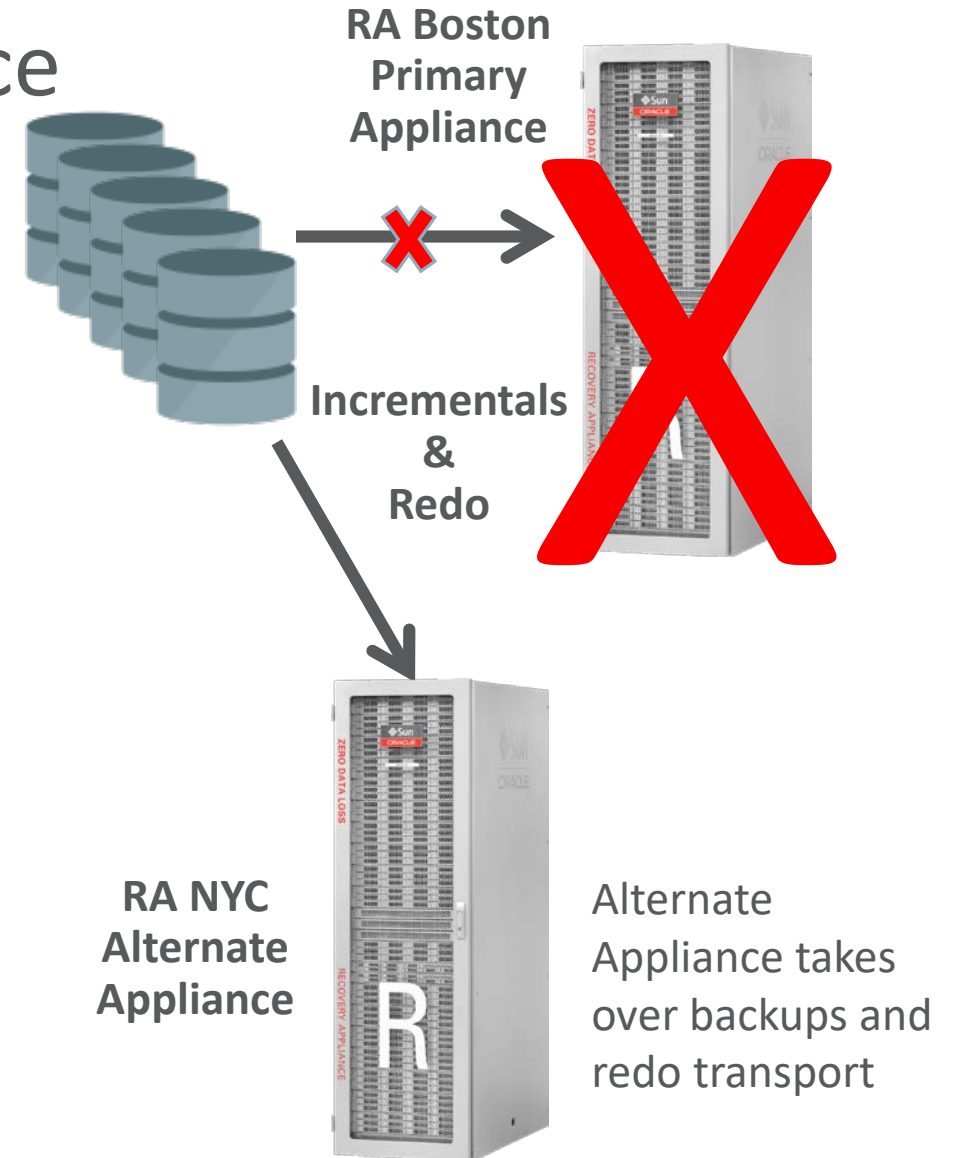
"BF_FORWARD" Policy on *Alternate RA*: STORE_AND_FORWARD = 'YES'

ORACLE®

Global Financial Services Company

8000+ Protected DBs, Global Data Centers, Backup Failover to Alternate Appliance

DB+RA POD 1 (600-700 DBs)

DB Group 1

DB Group 2

DB Group 1
->RA1
  ->Tape
Failover to:
->RA2

DB Group 2
->RA2
  ->Tape
Failover to:
->RA1

DB+RA POD 2 (600-700 DBs)

DB Group 3

DB Group 4

DB Group 3
->RA3
  ->Tape
Failover to:
->RA4

Standby DBs
->Local RA
Failover to:
->RA3
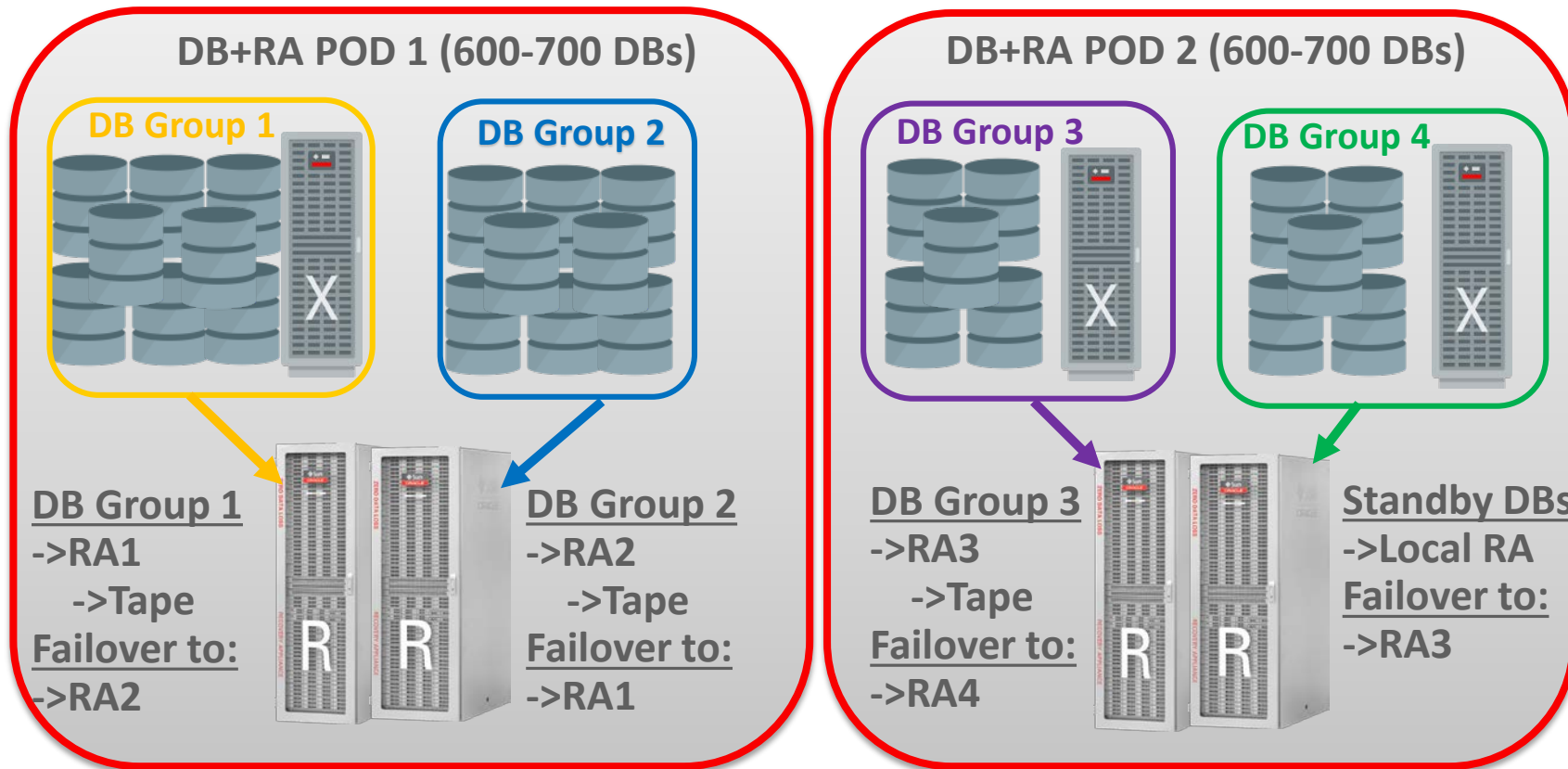
Each RA in a pod is configured as failover for the other.

- **Space is reserved for normal & failover backups**

Standardized DB On-Boarding:

- **Data Guard or Standalone?**
- **Failover Config Required?**
- **DB Size + Change % + Retention -> Reserved Space**
- **Reserved Space periodically adjusted per space needs**

NEW: Zero Data Loss Cross-Platform Database Migration
Simple Database Migration to On-Premise or Cloud@Customer using ZDLRA

- Significant reduction in downtime – less than 2 hour read-only downtime, regardless of DB size.

- New automation tool simplifies platform migration steps - especially useful for large databases

- Supports same and cross-endian platform migration activities
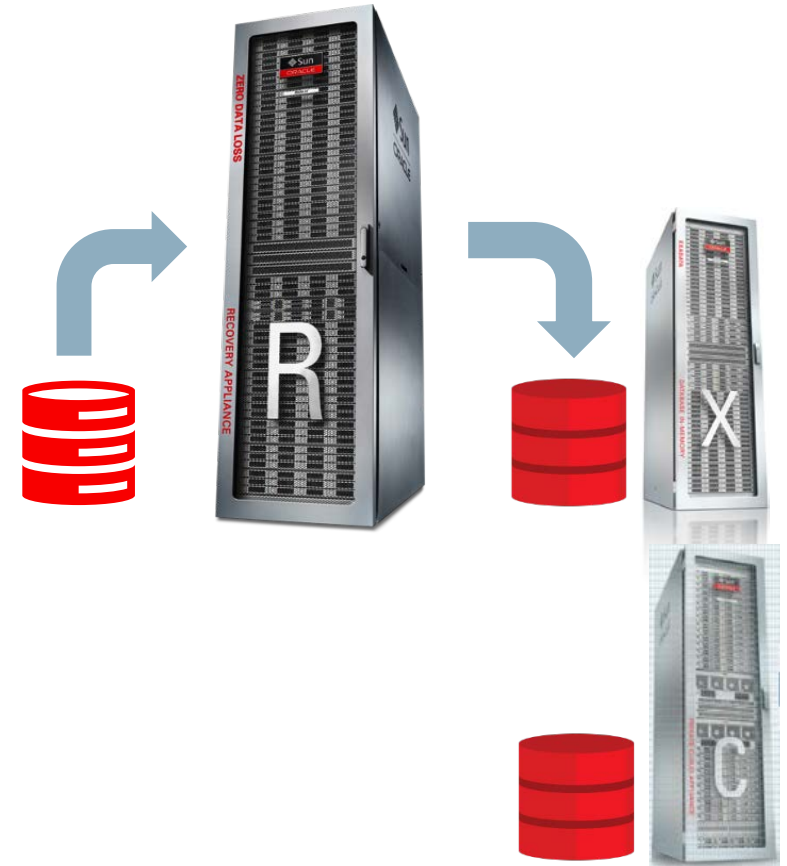
Zero Data Loss Cross-Platform Database Migration

Dramatically Reduce Migration Time – From Hours/Days to Few Hours or Less
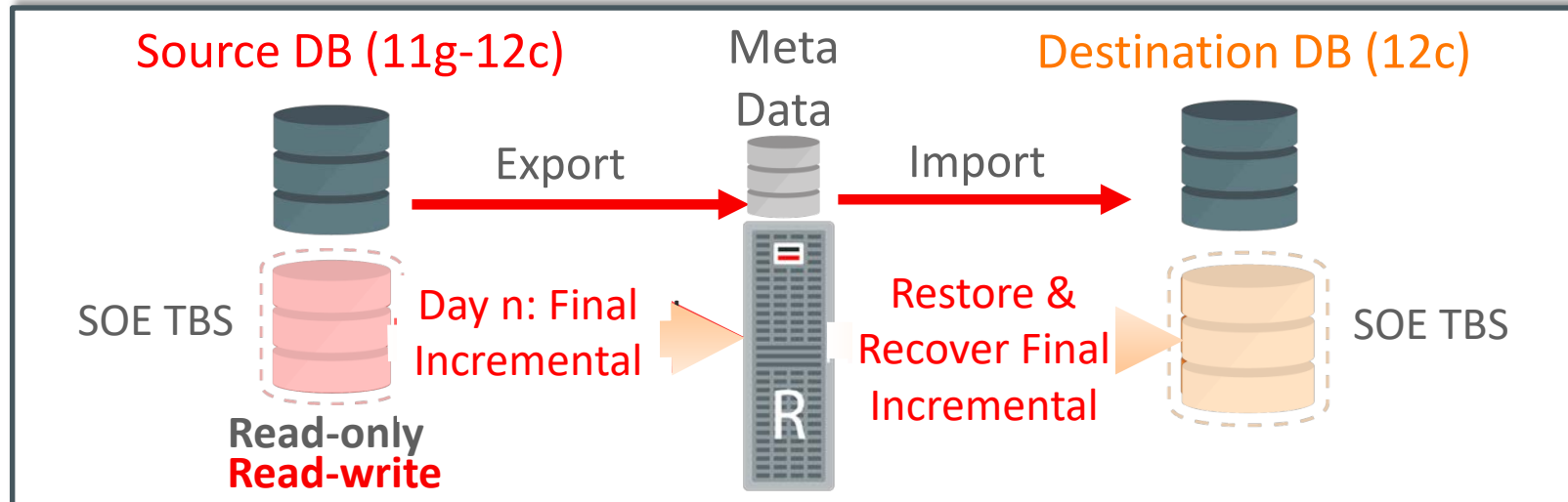
- Centralized Recovery Appliance "migration engine" + minimal downtime (short read-only at end)
  - Daily incremental backups -> virtual full backups on Recovery Appliance
- At destination, restore latest virtual full backup, prior to migration window
  - RESTORE FROM PLATFORM XXX FOREIGN DATAFILE YYY
- When ready to switchover: At source, take final incremental and metadata tablespace export in read-only
  - RECOVER FROM PLATFORM XXX FOREIGN DATAFILECOPY YYY at destination
  - IMPORT Data Pump export file at destination
- Destination open in read-only to verify migrated data, then open read-write for business
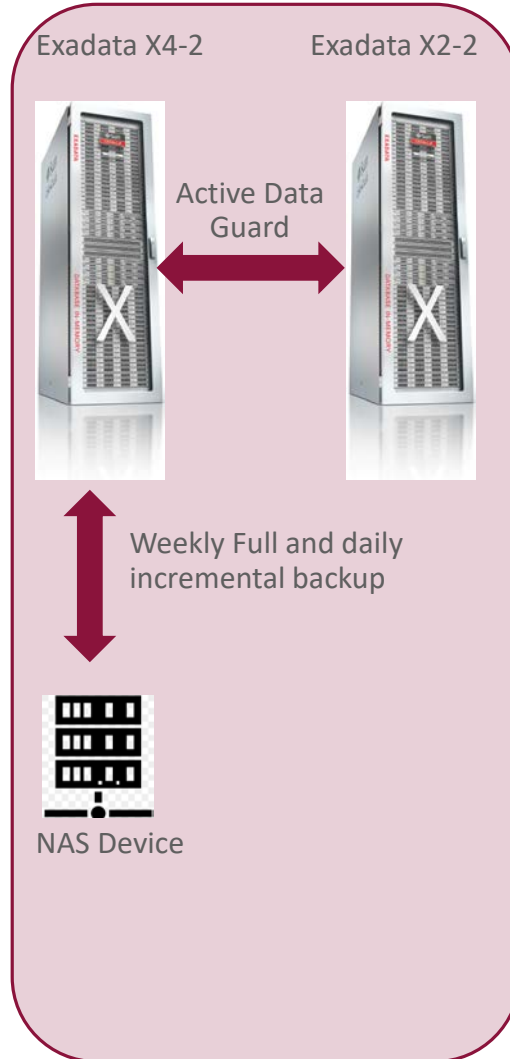
Leading Global Semiconductor Manufacturer

Business Needs

- **Accelerate Growth**
- **Drive Operational Excellence**
  - **Customer Experience**
  - **Operational Efficiency**
- **Grow organizational capabilities – optimize innovation**
- **Address current & planned business growth objectives**

Solution Needs

- **Stability**
- **Zero Preventable Outages**
- **Focus on Business Ops**
- **Increase IT agility, self-service and alignment to business drivers**

Legacy Architecture

Exadata X4-2   Exadata X2-2

Active Data Guard

Weekly Full and daily incremental backup

NAS Device

New Architecture

Exadata X4-2   Exadata X2-2

Cascade Standby

Active Data Guard

Incremental forever backups
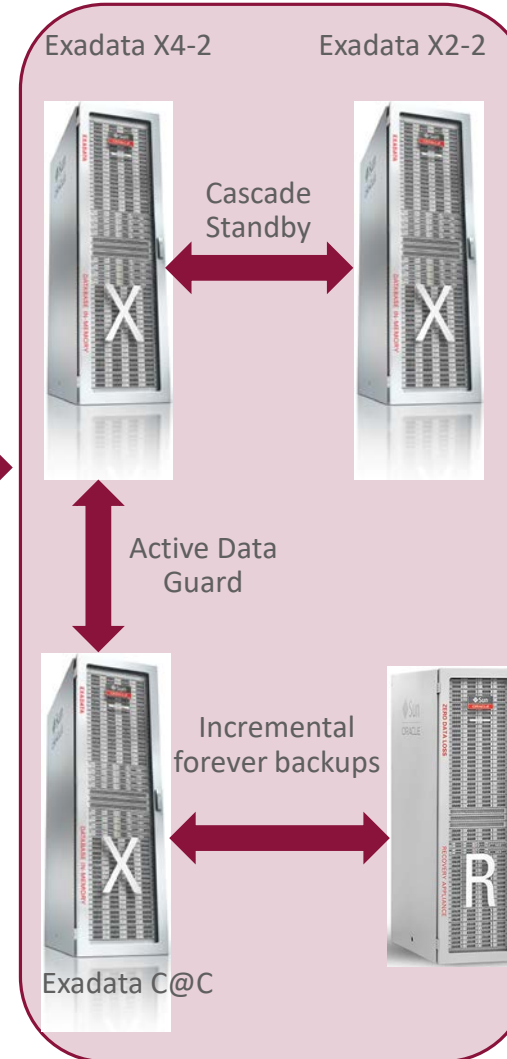
Exadata C@C

Results Achieved

- **Consolidate and standardize**
  - Consolidated several database servers
  - Compatible with Exadata
  - Multi-Tenant option
- **Reliable, Scalable and High Performing**
- **Improved the time and cost to build and maintain Analytics platform**
- **Near zero downtime migration using ZDLRA – RMAN DUPLICATE.**
- **Deliver exceptional service to business users**
- **Eliminated full backups**
- **Improved RTO by 4X**
- **Reduced backup windows by 2X**

ORACLE®

NEW: RA System Activity Report in BI Publisher

Tracks Recovery Appliance Activity and Highlights Areas of Action

- Displays information for the following:
  - State of Protected Databases
  - RA Space Utilization
  - State of the Recovery Appliance
  - Tasks & Task History

TASK TYPE	TASK PRIORITY	TASK STATE	TASK COUNT	TASK - LAST EXECUTE TIME	TASK - WORK TYPE	TASK - MIN CREATION
INDEX_BACKUP	140	ORDERING_WAIT	346		Work	11-MAY-2018
VALIDATE	330	RUNNING	1	15-MAY-2018 19:36:52	Maintenance	11-MAY-2018
PURGE_DUP	60	RUNNING	6	16-MAY-2018 06:32:54	Work	11-MAY-2018
INDEX_BACKUP	140	RUNNING	1	16-MAY-2018 07:32:13	Work	16-MAY-2018
OPT_DF	310	TASK_WAIT	4		Maintenance	12-APR-2018
PURGE_DF	80	TASK_WAIT	6		Work	16-MAY-2018
VALIDATE	330	TASK_WAIT	1		Maintenance	12-APR-2018
PURGE_DUP	60	TASK_WAIT	109		Work	11-MAY-2018
INDEX_BACKUP	140	TASK_WAIT	25		Work	14-MAY-2018
CROSSCHECK_DB	350	TASK_WAIT	2		Maintenance	01-MAY-2018

ORACLE

RA System Activity Report: Healthy?

- Items to watch for: Bad report

VERSION	NAME	CURRENT_TIME
27-11-2017 15:21:40 ZDLRA_12.1.1.1.8.201711_LINUX.X64_RELEASE	ZDLRA1	09-DEC-2017 19:41:04

TASK_TYPE	STATE	CURRENT_COUNT	LAST_EXECUTE_TIME	WORK_TYPE	MIN_CREATION
CROSSCHECK_DB	EXECUTABLE	1		Maintenance	08-DEC-2017
PLAN_DF	EXECUTABLE	498,959		Maintenance	22-NOV-2017
VALIDATE	EXECUTABLE	228		Maintenance	02-SEP-2017
REBUILD_INDEX	EXECUTABLE	805		Maintenance	21-OCT-2017
OPTIMIZE	EXECUTABLE	224		Maintenance	28-NOV-2017
OPT_DF	EXECUTABLE	98		Maintenance	22-NOV-2017
RESTORE_RANGE_REFRESH	EXECUTABLE	203		Maintenance	08-DEC-2017
DB_STATS_REFRESH	EXECUTABLE	1		Maintenance	08-DEC-2017
RM_INC_FILES	EXECUTABLE	1		Work	28-NOV-2017
OBSOLETE_SBT	EXECUTABLE	1		SBT	28-NOV-2017
PURGE_DUP	EXECUTABLE	213		Work	13-NOV-2017
INDEX_BACKUP	EXECUTABLE	179,451		Work	01-DEC-2017
CROSSCHECK_DB	EXECUTABLE	10		Work	29-NOV-2017
BACKUP_ARCH	EXECUTABLE	1,072		Work	09-DEC-2017
PURGE_DF	EXECUTABLE	208,782		Work	09-DEC-2017
INDEX_BACKUP	ORDERING_WAIT	112		Work	13-NOV-2017
PURGE_DUP	RUNNING	1	09-DEC-2017 18:47:01	Work	13-NOV-2017
BACKUP_ARCH	RUNNING	2	09-DEC-2017 19:38:15	Work	09-DEC-2017
PURGE	RUNNING	1	09-DEC-2017 06:36:12	Work	08-DEC-2017
PURGE_DF	RUNNING	94	09-DEC-2017 19:38:11	Work	09-DEC-2017
DEFERRED_DEL	RUNNING	4	09-DEC-2017 10:40:04	Work	09-DEC-2017
PURGE_DUP	STALL_WHEN_WAIT	172,675		Work	31-OCT-2017
CHECK_FILES	TASK_WAIT	1		Maintenance	19-NOV-2017
VALIDATE	TASK_WAIT	1		Maintenance	12-JUN-2017
OPT_DF	TASK_WAIT	293		Maintenance	09-OCT-2017
CROSSCHECK_DB	TASK_WAIT	218		Work	21-NOV-2017
PURGE_DF	TASK_WAIT	3		Work	09-DEC-2017

Watch for :

If there are tasks of WORK TYPE in RUNNING state and created a day earlier then investigate.

If there are tasks of MAINTENANCE or SBT TYPE present and their creation time is older than one week then it should be investigated.

A large number of tasks in EXECUTABLE state for the same TASK_TYPE

The same TASK_TYPE is increasing in the number of jobs (CURRENT_COUNT) over time.

A system with ordering waits older than 1 day.

A system with tasks in STALL_WHEN_WAIT. This should only be seen if Oracle Support is troubleshooting the system.

ORACLE®

RA System Activity Report: Healthy?

VERSION		NAME	CURRENT_TIME
27-11-2017 15:21:40 ZDLRA_12.1.1.1.8.201711_LINUX.X64_RELEASE		ZDLRA1	11-JAN-2018 19:58:14

The Good:

Minimum creation time for active tasks is within the last 24 hours for work tasks .

Task history state should indicate work is being completed.

Review other sections

TASK_TYPE	STATE	CNT	MIN_COMPLETION_TIME	MAX_COMPLETION_TIME
BACKUP_ARCH	COMPLETED	11,046	10-JAN-2018 19:58:16	11-JAN-2018 19:57:40
DB_STATS_REFRESH	COMPLETED	131	10-JAN-2018 20:06:01	11-JAN-2018 19:55:40
DEFERRED_DEL	COMPLETED	24,602	10-JAN-2018 19:58:14	11-JAN-2018 19:58:09
HISTOGRAM	COMPLETED	8	10-JAN-2018 20:11:02	11-JAN-2018 17:12:22
INDEX_BACKUP	COMPLETED	24,599	10-JAN-2018 19:58:14	11-JAN-2018 19:58:06
OBSOLETE_SBT	COMPLETED	1	11-JAN-2018 05:10:31	11-JAN-2018 05:10:31
PURGE_DUP	COMPLETED	266	11-JAN-2018 00:29:46	11-JAN-2018 19:58:11
RESTORE_RANGE_REFRESH	COMPLETED	29,737	10-JAN-2018 20:06:02	11-JAN-2018 19:55:52
RM_INC_FILES	COMPLETED	131	10-JAN-2018 20:06:00	11-JAN-2018 19:55:37

Backup Best Practices

- Use Transparent Data Encryption (TDE) instead of RMAN encryption
  - RMAN encryption will prevent ZDLRA from creating Virtual Full Backups (VB$).

- TDE backups will not get compressed on the ZDLRA
  - Encrypted backups don't get compressed by the storage in general – A sizing exercise must be (re)-conducted when preparing for TDE
    - **The incremental forever strategy still applies to TDE backups**

- Use RA built-in compression instead of RMAN compression
  - RMAN compression incurs DB CPU utilization & backups are decompressed + recompressed on RA

  - **$ rman target \<target string\> catalog \<catalog string\>**
    backup device type sbt
      cumulative incremental level 1 filesperset 1 section size 64g database
       plus archivelog not backed up
        filesperset 32;

Use Recommended RA Software

Very Important

- Subscribe to MOS alerts and refer periodically to the following notes:
  - Recovery Appliance Critical Issues MOS note for critical issues alerts
  - Recovery Appliance Supported Versions MOS note for latest software update

- Use Recommended RA Software to avoid known critical issues
  - Number 1 method to avoid problems: Upgrade to recommended software release
  - Zero Data Loss Recovery Appliance Supported Versions (Doc ID 1927416.1)
  - Zero Data Loss Recovery Appliance Upgrade and Patching (Doc ID 2028931.1)
  - Patches are cumulative and include
    - Bug fixes (Avoid a bug before it happens)
    - Enhancements (New features, optimizations, etc.)

- Coordinate with Platinum Patching
  - Schedule early & Open Proactive SR

Important things to avoid

1. **Don't** ignore incidents: Resolve and understand
2. **Don't** neglect the RA: Monitor the system
3. **Don't** make modifications to the RA configuration
   i. Zero Data Loss Recovery Appliance - Installing Third-Party Software and Modifying Internal Appliance Software (Doc ID 2014361.1)
4. **Don't** take periodic Level 0: Virtual Level 0 requires only one level 0 followed by level 1s
5. **Don't** backup to another media: Switching to another media can impact past and future backups made to the RA

ORACLE®

4 important things to do

1. **Monitor the RA**

   i. OEM Unified Management Dashboard: Review twice daily and setup alerts and notifications

   ii. Run the System Activity Report daily and monitor trends:
   Zero Data Loss Recovery Appliance System Activity Script (Doc ID 2275176.1)

   iii. Run Exachk Monthly and review findings:
   How To update exachk outside ZDLRA Install, Patching and Upgrade (Doc ID 2399688.1)

   iv. Review Capacity Planning Report Monthly or Bi-Monthly

2. **Use Multi Section:** Set section size to 64GB, this allows for efficient processing in ZDLRA's flash cache

   i. Large datafiles > 16TB will automatically see an increase in section size - section size = sizeof(datafile) / 256

   ii. Small datafiles < 64GB will not have sections

   iii. Forces FILESPERSET to 1

3. **Use the latest libra** (RMAN client sbt library that supports ZDLRA):
   ZDLRA: Where to download new sbt library (libra.so module) (Doc ID 2219812.1)

4. **Validate the backup/restore network first:**

   a. Use Zero Data Loss Recovery Appliance Network Test Throughput script (Doc ID 2022086.1) –
   requires qperf, supported on specific OSes

   b. Use How to measure network performance from RMAN for ZDLRA or Cloud Backups (Doc ID 2371860.1) –
   uses RMAN's "NETTEST" option, OS agnostic

ORACLE®

Agenda

1 ▸ Introduction & Architecture

2 ▸ Operational & Monitoring Best Practices

3 ▸ METRONOM Case Study

4 ▸ Resources & Next Steps

ORACLE®

ZDLRA @ METRONOM

10.24.2018

Agenda

METRO|NOM

Introduction

Introduction

FOR ALL TO WHOM LOVE IS THE MOST IMPORTANT INGREDIENT.

From banana to physalis: METRO offers everything to give your creation the final touch.

PARA TODOS AQUELLOS QUE NO TIENEN SOLO CLIENTES, SINO AMIGOS.

YOU & METRO

THE BIGGEST SOFTWARE COMPANY YOU NEVER HEARD ABOUT

METRONOM – SETTING THE PACE IN FOOD AND TECHNOLOGY

150.000 Employees

€37 Billion

25 Countries

2000 Employees

IT-Services

IT-Solutions

METRO|NOM

Introduction

Team of 13 DBAs

- ~ 2100 Oracle databases

- AIX, Linux, Windows - Servers

- Database features including:

  - RAC

  - Data Guard

  - GoldenGate for Minimal Downtime Maintenance (MDM)

METRO|NOM
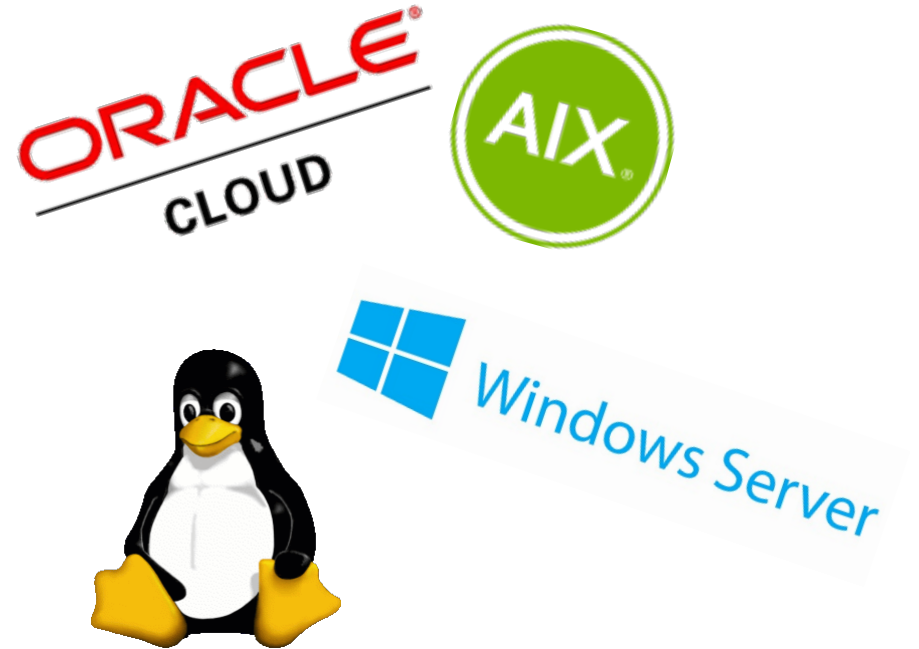
Mission

Mission

- **Started POC in 2016**

- **Replication between two DCs**

- **Easy Migration**

- **Go Live April 2017**

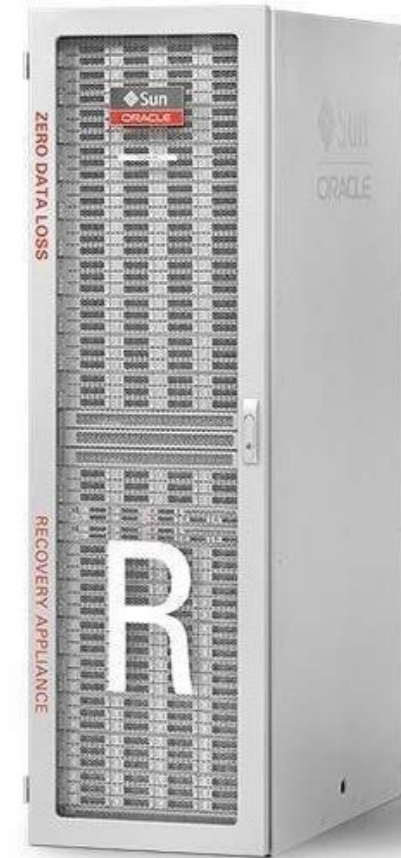- **Backup of more than 350 productive databases by end of 2017**

- **Zero Downtime**

METRO|NOM

Best Practices

METRO|NOM
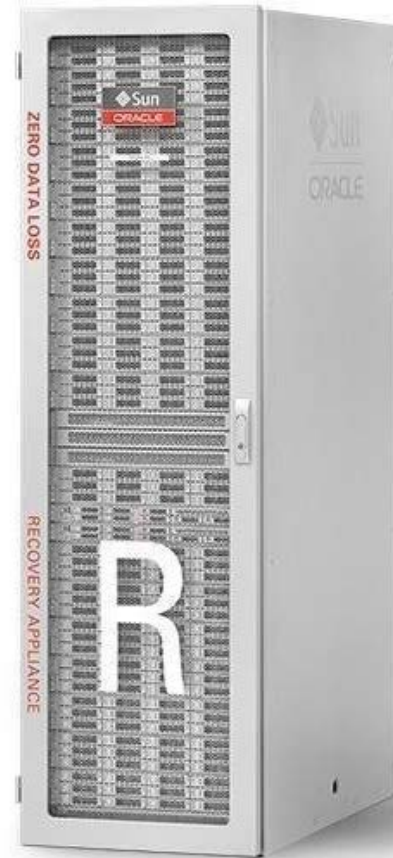
Best Practices

Easy migration

`addzdlra.sh -d testdb -p gold`

METRO|NOM

Best Practices

MAX_RETENTION_WINDOW

- Set it, but not too aggressive

Don't Submit multiple requests to Delete Databases

- A delete can take a lot of time

- If database delete does not progress for some time, contact Support

Don't Neglect the RA

- Monitor and adjust the System

- System Activity Report

  - Clarify Findings with Support

METRO|NOM
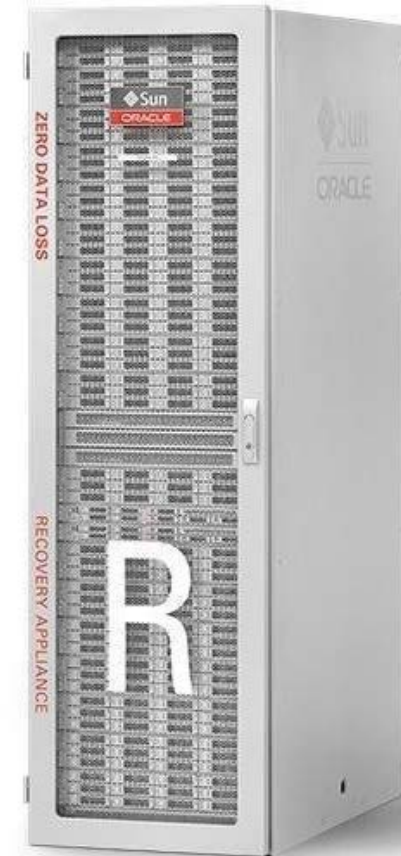
Target

Target

- Use ZDLRA to Backup ~900 databases within the next 12 months

- Add more storage to ZDLRA for growth and replication needs

- Implement Minimal Downtime Maintenance patching solution via new RA High Availability for Backup & Recovery procedure (aka Auto-Failover to Replica RA)

- Outsource Backups to Oracle Cloud

METRO|NOM

RA-based Migration of IBM AIX to ExaCC/OCC

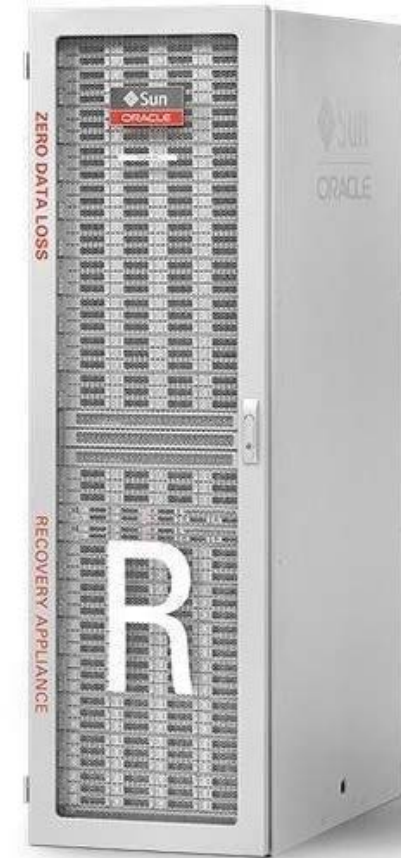ATTEND OUR MIGRATION SESSION [TRN4033] TOMORROW, 9 AM, MOSCONE WEST 3004
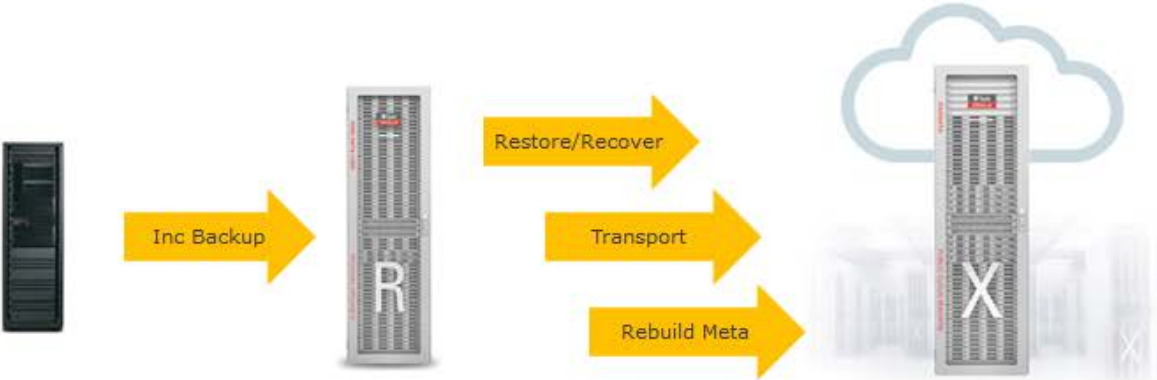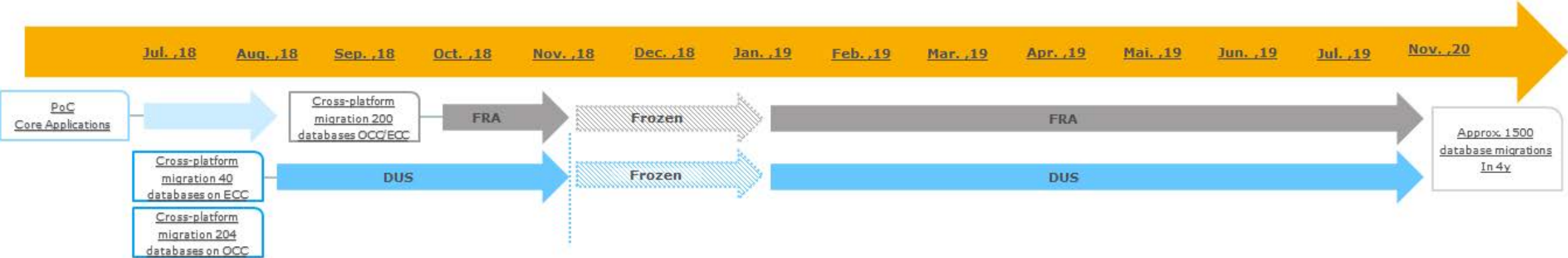
Summary

Summary

- **Less time for Backups**

- **Fast restore and recovery**

- **Easy Migration due to automated processes**

- **Scalable for data growth**

METRO|NOM

Thank you for your attention!

Stefan Reiners

METRO|NOM

Agenda

1. Introduction & Architecture

2. Operational & Monitoring Best Practices

3. METRONOM Case Study

4. **Resources & Next Steps**

ORACLE®

Resources & Next Steps

- Recovery Appliance Product Central
  - www.oracle.com/recoveryappliance

- Recovery Appliance MAA Best Practices Central
  - www.oracle.com/goto/maa -> Zero Data Loss Recovery Appliance

- OpenWorld MAA Presentations
  - www.oracle.com/goto/maa -> Presentations

- AskTOM Backup & Recovery Sessions
  - asktom.oracle.com -> Office Hours -> Search for 'backup'

- Get our business card for follow up questions!