Exadata稀疏克隆及ACFS 快照

公益讲座11：00准时开始，请大家先浏览云技术微信公众号技术文章资料会在各群同步发布，已入群客户请勿重复入群！

20-19

数据库和云讲座群

甲骨文云技术公众号

ORACLE

Exadata稀疏克隆及ACFS 快照

用于开发和测试的数据库克隆

金哲男

Oct 2022

Oracle SE Hub Database Management

Exadata数据库克隆

- 许多组织使用 Exadata 进行生产、灾难恢复和开发/测试
- 适用于所有生产和测试/开发数据库用例的单一解决方案平台
- Exadata 是运行 Oracle 数据库的最佳平台

测试/开发案例	Oracle解决方案
完整的端到端性能测试	Non-Sparse Exadata 与主要系统相同或类似的系统
使用简单的快照和 Exadata 智能功能进行测试	Exadata Sparse Clone
类似于第三方写时复制的高级快照功能，但无法使用 Exadata off-load 功能	ACFS Snapshot on Exadata

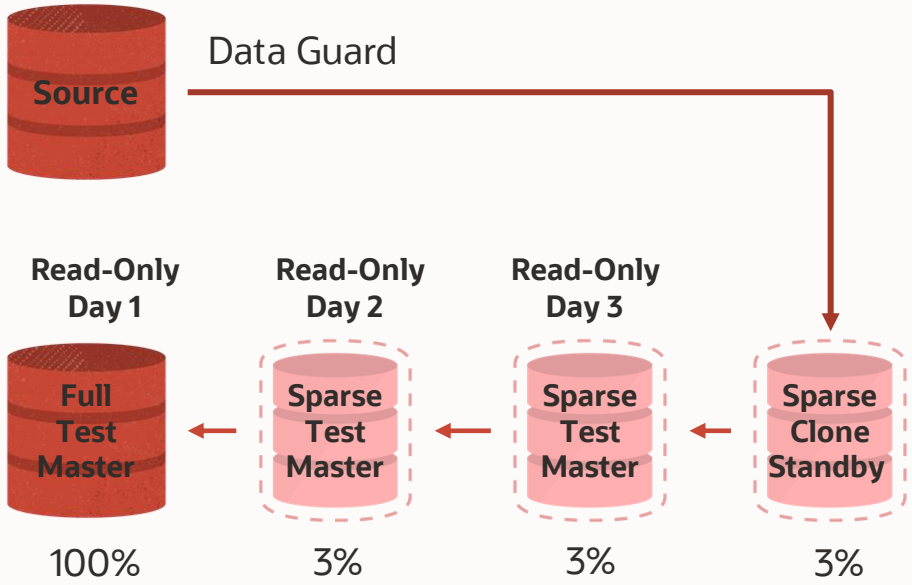
Sparse Clones vs. Storage Snapshots

Sparse Clone (copy-on-write)

Sparse clone 母版是 read-only

Sparse clones 保留变化的Blocks

块随着克隆的变化而累积

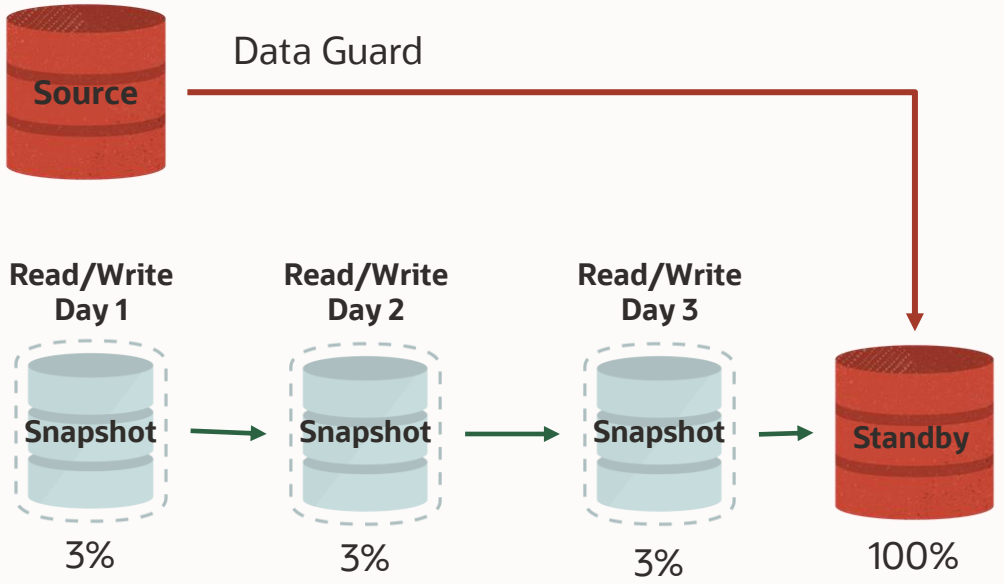


Snapshot (preserve prior block versions)

Snapshot 母版 read/write

Snapshot 保留较旧的块版本

块随着主节点的变化而累积



Exadata Sparse Clones

Exadata 的组成部分

- 与 Exadata 存储功能（SQL 卸载、I/O 优先级等）完全兼容

节省空间的稀疏克隆(Sparse Clone)

- 在内部使用写时复制技术(Copy-on-write technology)
- 每个快照仅包含基于主副本的数据更改，包括稀疏主副本

支持数据范围时间线

- 来自 Data Guard
- 选择数据更改时间线中的任意点（小时、日、月、年）

节省空间的稀疏备份

- 备份也节省空间，仅备份稀疏克隆中的更改

数据混淆能力

- Oracle Data Masking and Subsetting, Data Redaction, Virtual Private Database

ACFS Snapshots on Exadata

节省空间的文件系统快照

在内部使用写时复制技术

快照包含从父快照所做的更改

支持多个时间线

来自 Data Guard 源

选择数据更改时间线中的任意点（小时、日、月、年）

每个文件系统最多支持 1023 个快照

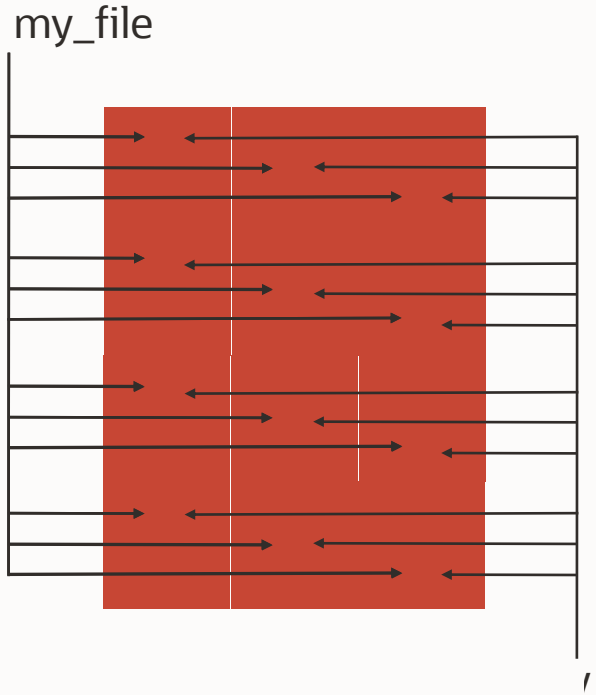
数据混淆能力

Oracle Data Masking and Subsetting, Data Redaction, Virtual Private Database

Exadata 功能支持

仅支持 Exadata 智能闪存缓存

Copy-on-write Concept



数据文件包含两部分即：

数据块（blocks）

数据文件首部(header)

仅拷贝数据文件首部（header）

- 存储使用更加高效
- 极低的拷贝IO消耗
- 仅需数秒即可完成

仅在发生写操作时拷贝数据块

```
cp my_file my_file_copy
```

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

ACFS Snapshots

Summary

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

ACFS Snapshots

Summary

Exadata Sparse Snapshots

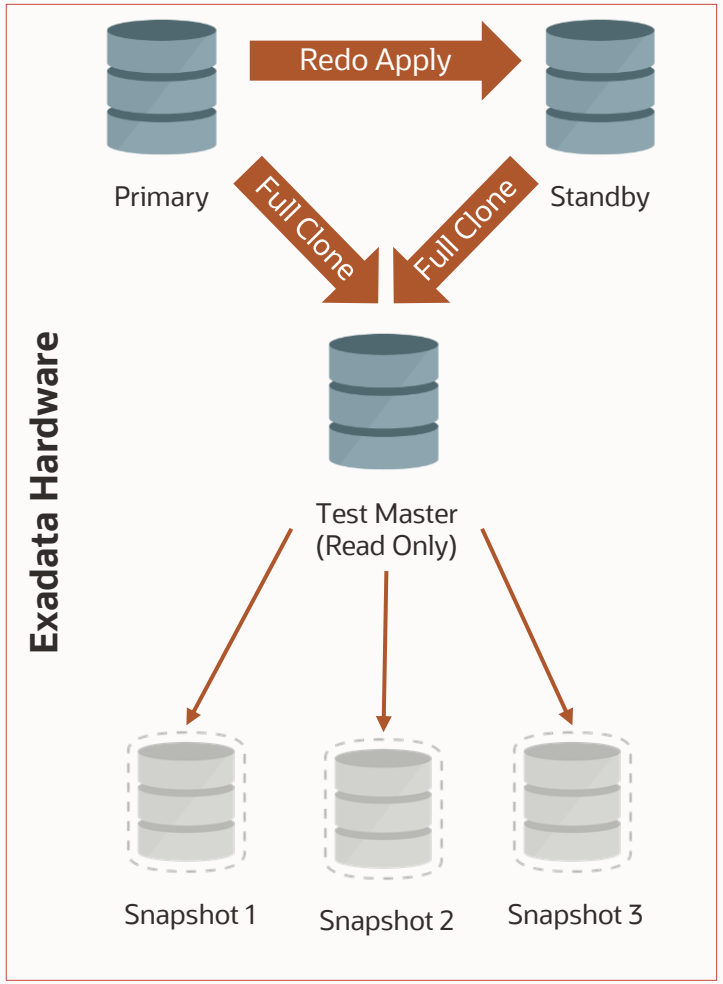
- 能够在 Exadata 上快速创建节省空间的测试/开发数据库
- 稀疏快照测试/开发数据库可以使用所有 Exadata 智能特性，包括Smart Offload，因此应用程序可以使用 Exadata 特性进行评估
- 稀疏快照测试/开发数据库不是完整的副本，从而节省了存储空间和成本
- HCC 存储压缩透明地工作，提供额外的存储节省

术语

- Sparse Griddisk
  - Virtual size
  - Physical size
- Sparse Diskgroup
- Sparse Data Files
- Exadata Snapshots
- Exadata Test Master
- Exadata Sparse Test Master

Test/Dev Exadata Snapshots 生命周期

1. 生产数据库在 Exadata 上运行
2. 从 Exadata 上的备用或生产创建的完整克隆测试主数据库
   - 可选择屏蔽 Test Master 中的敏感数据
   - 使用 PDB 的一个命令从此处创建空间高效的测试/开发数据库
   - 快照上可用的 Exadata 智能功能（query offload; storage index; smart log; smart flash cache; HCC; etc）
   - 重要 – 刷新测试主机会使现有快照失效；必须创建新的完整 Test Master 才能创建新的刷新的 Exadata 快照



SIZE

Full

Full

Sparse

Sparse Database/File/GridDisk

Sparse Database

- 只有稀疏数据库中的数据文件是稀疏的
- 控制文件/在线重做日志/临时文件； 等并不稀疏

Sparse File

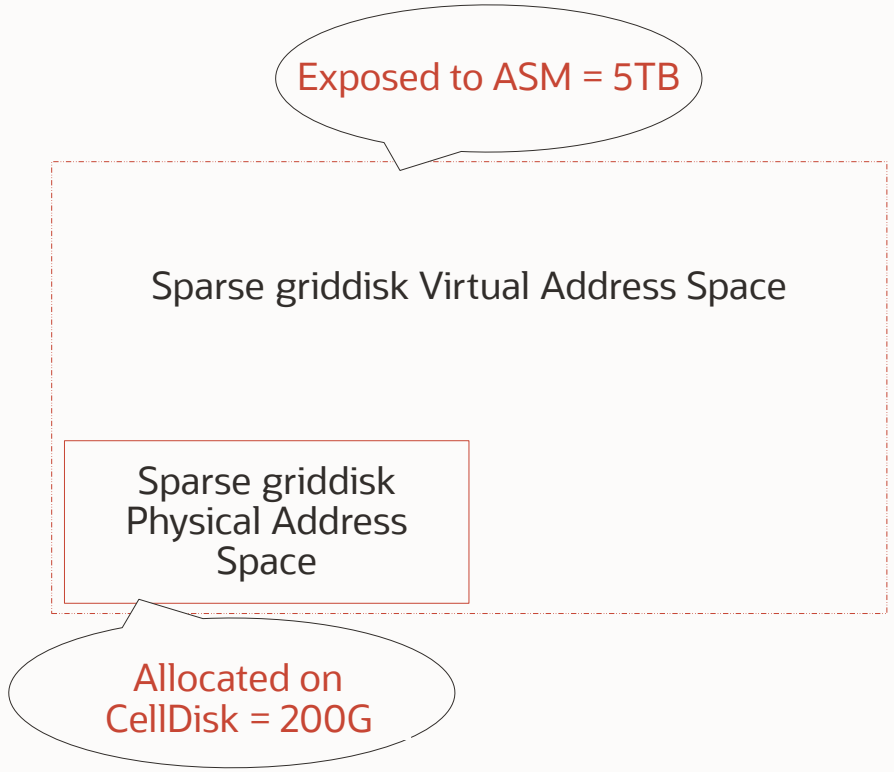
- 稀疏数据文件指向测试主数据库数据文件
- 仅在写入期间按需分配块

Sparse Griddisk

- 除了物理尺寸外，还显示虚拟尺寸
- 最大限度。 允许的虚拟大小/磁盘 = 100TB
- 最大限度。 允许的聚合物理大小/磁盘 = 4TB

Example:

```
Cellcli> create griddisk all harddisk
prefix=SPARSE size=200G,
virtualsize=5TB
```



Sparse Database/File/GridDisk

10.8.1.1 Calculating the Physical Size for Grid Disks

You can use the following formula to get a rough estimate of the total physical space to set aside for a sparse ASM disk group:

```
Total physical space =
   (SUM(size of all test masters in the sparse ASM disk group) +
    SUM(approximate size of all updates to the snapshot databases))
   * ASM Redundancy
```

Copy

In the formula above, ASM redundancy takes into account ASM mirroring of extents. Exadata requires ASM redundancy set to either normal redundancy (double mirror the extents) or high redundancy (triple mirror the extents). If the sparse ASM disk group will use normal redundancy, expect to double the space used. If using high redundancy, expect to triple the space used.

For example, if you want 2 test masters in the sparse ASM disk group created with normal redundancy with a combined total space of 500 GB (250 GB each) and each test master will have 5 Exadata snapshots with the expectation that each snapshot will modify 20% of the blocks, then the total physical space that will be needed can be calculated as follows:

```
Space for 2 test masters:   2 * 250 GB =                          500 GB
Space for 5 snapshots per test master, for a total of 10 snapshots:
    10 * 250 GB * 20% =                                           500 GB
Subtotal                                                         1000 GB
Normal redundancy: 2 * 1000 GB =                                 2000 GB
```

Copy

https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-9BF24CF3-E191-4417-9B98-E0BC82C4B02A

Sparse Database/File/GridDisk

10.8.1.2 Calculating the Virtual Size for Grid Disks

You can use the following formula to get a rough estimate of the virtual size to assign for a sparse ASM disk group:

```
Virtual size required for sparse disks =
   (SUM(full virtual size of all Exadata snapshots) + Physical space allocated)
   * ASM Redundancy
```

Copy

To continue with the example from the previous section, you have 10 Exadata snapshots. If they were full copies of the test master, they would be 250 GB each.

The following shows the calculation for the total virtual space:

```
Full size for 5 snapshots per test master, for a total of 10 snapshots:
    10 * 250 GB =                                                2500 GB
Size of the 2 test masters: 2 * 250 GB =                          500 GB
Subtotal                                                         3000 GB
Normal redundancy: 2 * 3000 GB =                                 6000 GB
```

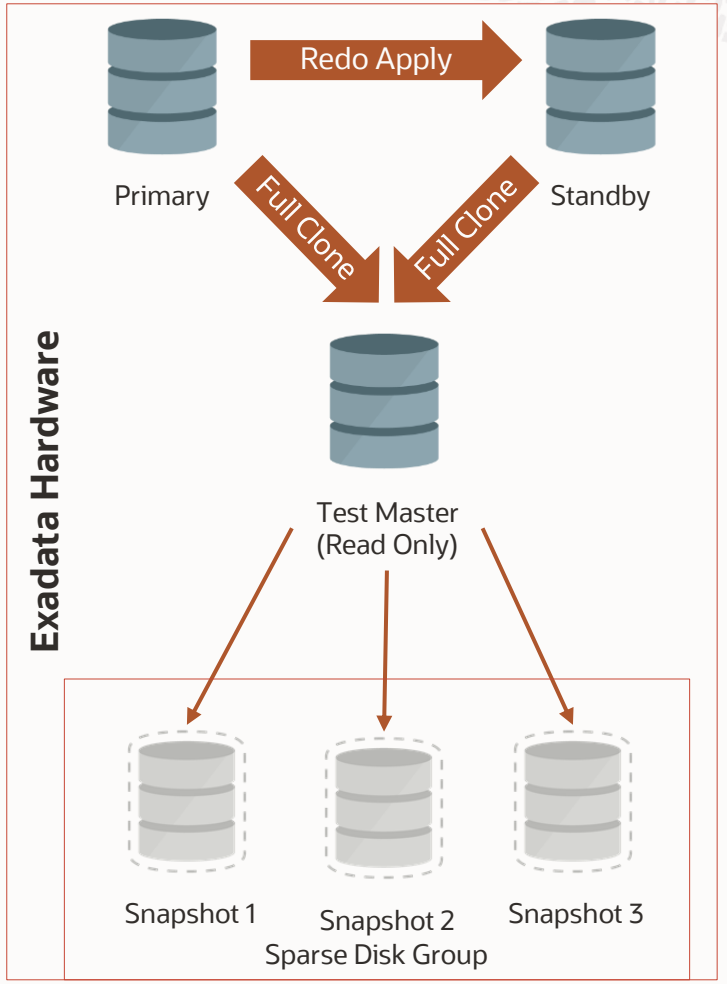
Copy

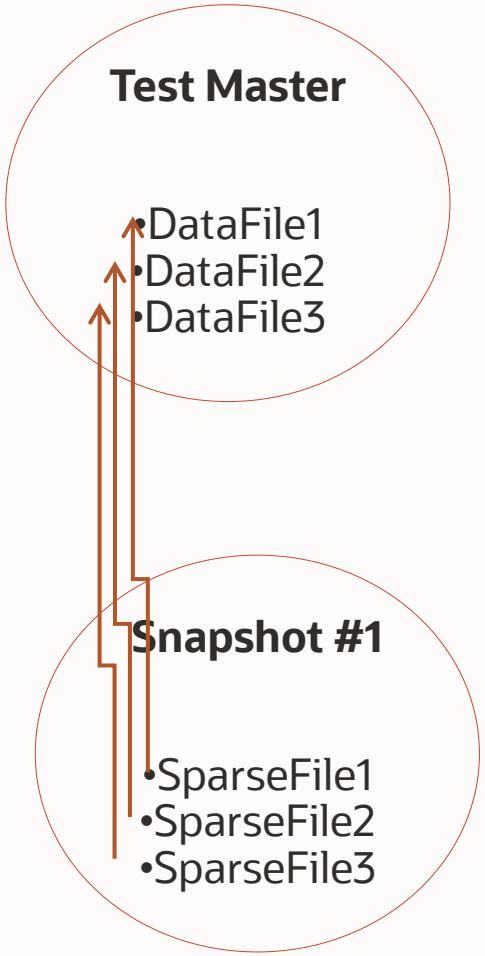
https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-9BF24CF3-E191-4417-9B98-E0BC82C4B02A

Sparse Diskgroup

- Sparse Files 只能在 sparse diskgroup 被创建
  - cell.sparse_dg =allsparse
  - 必须将 compatible.asm 和 compatible.rdbms 设置为 12.1.0.2 或更高版本
  - 对 4M AU 使用 16X 范围大小； 每个extent是 64M

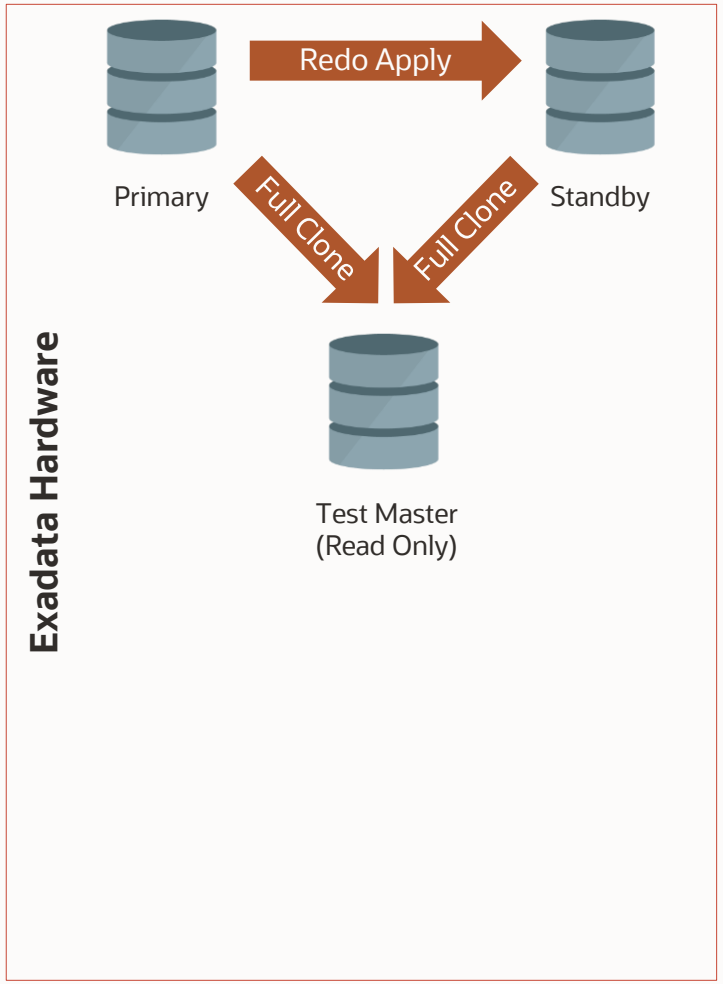
```
SQL> create diskgroup SPARSEDG
  normal redundancy
  disk 'o/*/SPARSE_*'
  attribute
    'compatible.asm'          = '19.0.0.0',
    'compatible.rdbms'        = '12.2.0.2',
    'cell.smart_scan_capable' = 'true',
    'cell.sparse_dg'          = 'allsparse',
    'au_size'                 = '4M';
```





Prepare Exadata for Snapshot Database

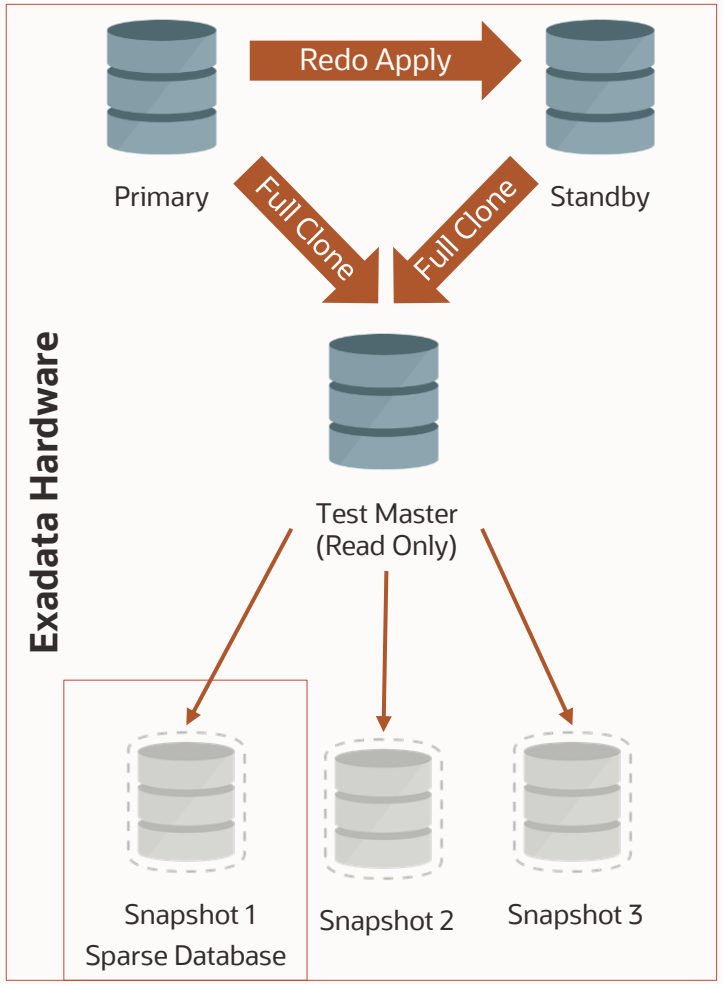
- 在存储节点创建sparse griddisks
- 在ASM instance 创建sparse diskgroup
- 配置 Test Master Database
  - Test Master的 Diskgroup 启用 ASM ACCESS_CONTROL
  - 创建完整克隆; 或者
  - 将 Exadata 上现有的完整克隆转换为 Test Master； 或者
  - 将备用数据库转换为测试主机



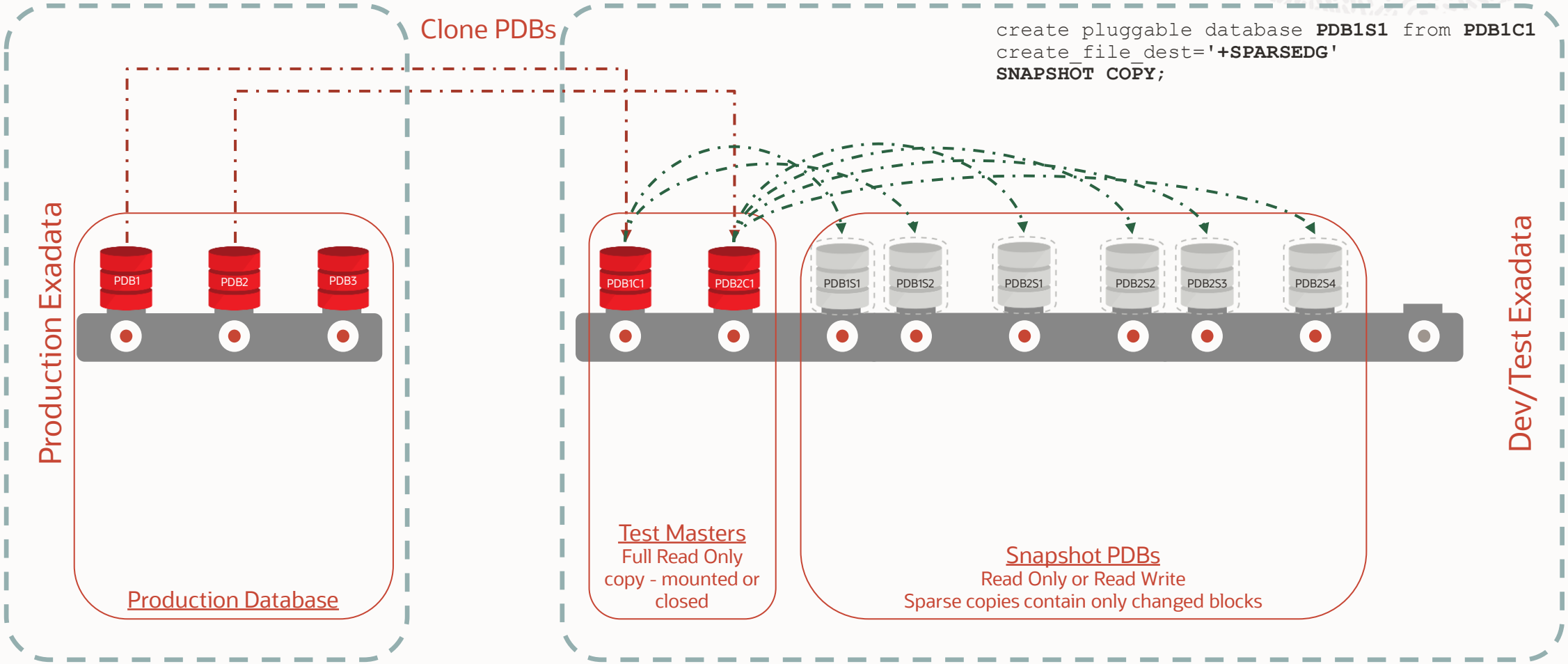
创建 Exadata Snapshot Database

- 创建 Sparse/Snapshot Database
  - **create_file_dest** 配置为 SPARSE diskgroup
  - Pluggable database.. 使用 SNAPSHOT COPY
  - Non-container database… using DBMS_DNFS.CLONEDB_RENAMEFILE

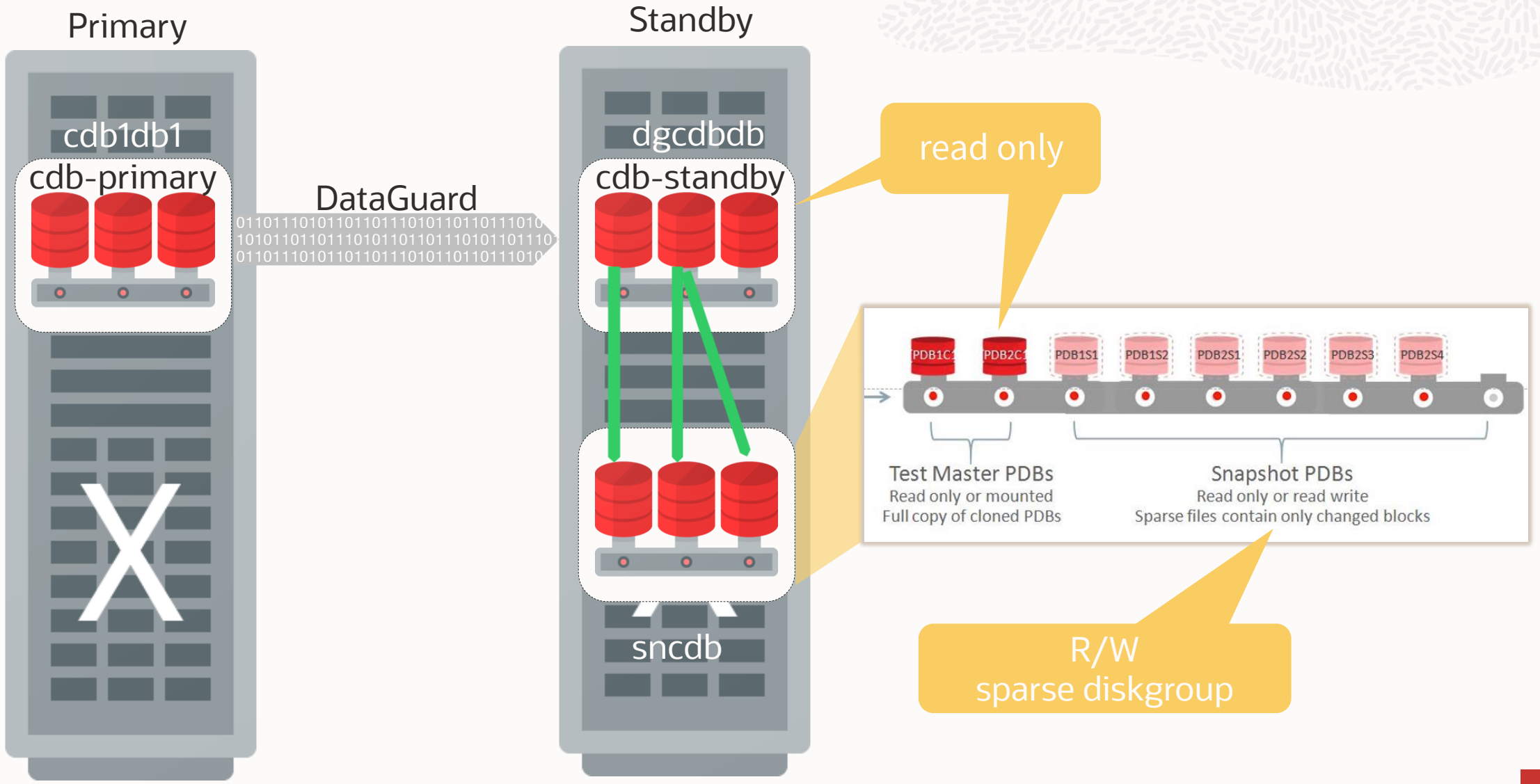
```
SQL> alter pluggable database
TestMaster open read only;

SQL> create pluggable database
JohnTest from TestMaster
create_file_dest='+SPARSEDG'
SNAPSHOT COPY;
```



Exadata Snapshot PDBs



Exadata Snapshot PDBs



Exadata Snapshot PDBs

测试内容

远程克隆EX：create pluggable database PDBCLONE from PDB@DBLINK;

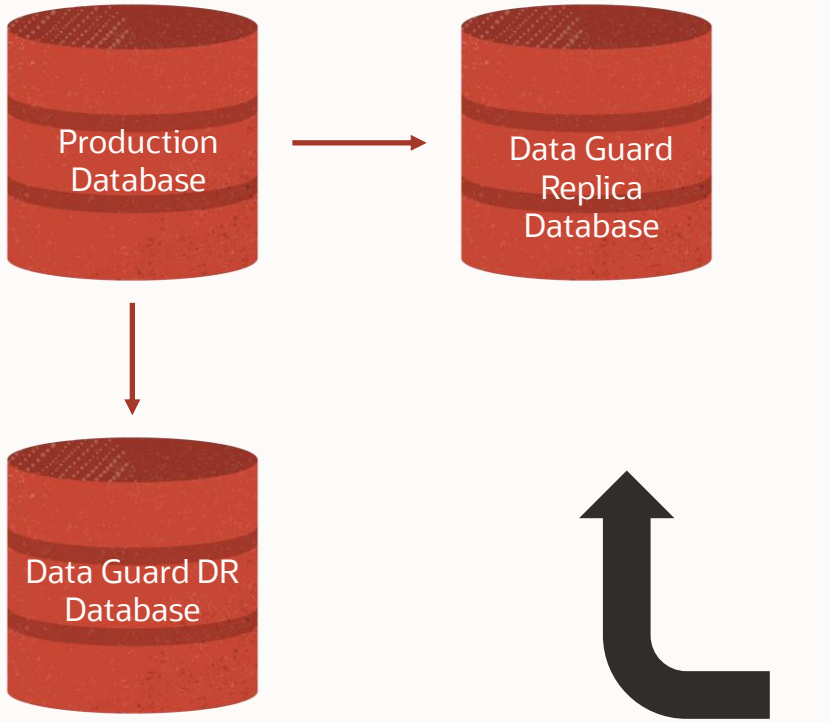
场景一	100GB数据文件PDB远程克隆 - 无实际数据	39.01秒
场景二	100GB数据文件PDB远程克隆 - 76GB数据	49.12秒
场景三	500GB数据文件PDB远程克隆 - 76GB数据	3分8.5秒
场景四	500GB数据文件PDB远程克隆 - 412GB数据	3分15.6秒
场景五	1TB数据文件PDB远程克隆 - 412GB数据	6分4.4秒
场景六	1TB数据文件PDB Sparse Snapshot 克隆 - 412GB数据	13.03秒

https://chulnamkim.notion.site/Exadata-PDB-Clone-8a3e5dc71970492f9b299850dfd9f6f3

Use Standby Database as Test Master

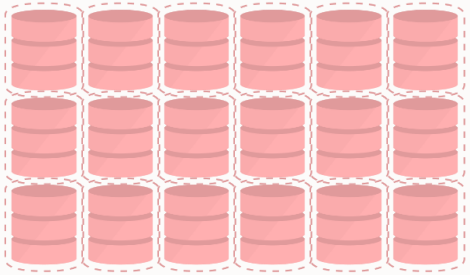
- Standby Database 作为 Test Master 时不可启用 redo apply
- 必须配置 ASM ACCESS_CONTROL enabled + ownership
- 定期刷新数据时需如下：
  - Drop 所有的 Snapshots 包括 datafiles
  - 使 Test Master Datafiles 变成 Read Write
  - 从生产库同步 Test Master 可以通过以下
    - DATAGUARD REDO APPLY; OR
    - RMAN RECOVER … FROM SERVICE
  - 关闭 Test Master(Standby Database),把TM的 所有 Datafiles转成 Read Only
  - 创建新的可读写的 Snapshots 进行测试

Refresh Standby Test Master Database



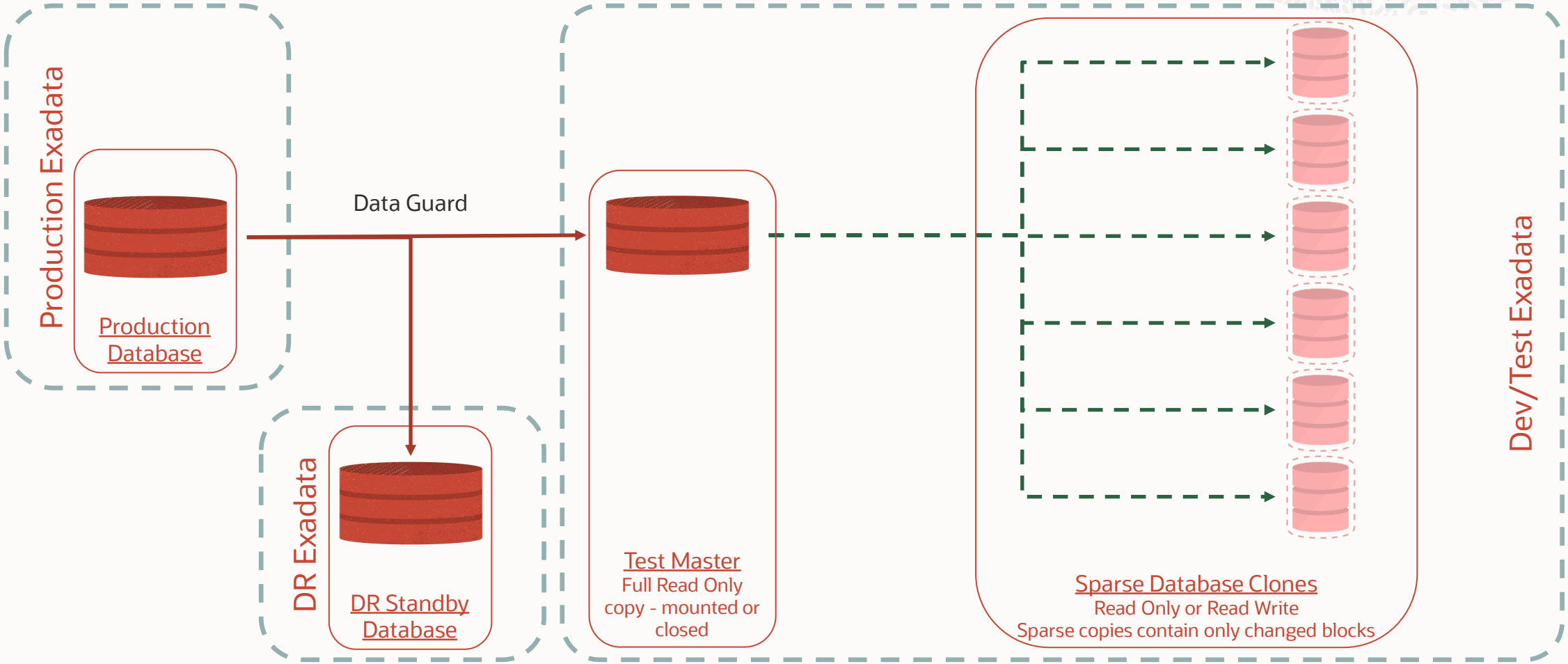
步骤1: DG转换为测试数据库(Test Master)
- 停止 Redo Apply
- 转换到 snapshot standby(可选)
- 准备成为测试主数据库
- 关闭数据库并打开只读

步骤2: 创建Sparse Clone 测试数据库
- 可以创建多个Sparse Clone测数据库
- 所有的Sparse Clone 是稀疏的，节省空间的



步骤3:刷新测试主数据库
- 删除 Sparse Clone
- 将数据库转换回原始状态 Data Guard
- 重新启用 Redo Apply 完成 Data Guard 同步
- 重复步骤1

Exadata Snapshot Databases



Exadata Snapshot Databases for CDB

DBMS_DNFS.CLONEDB_RENAMEFILE

This procedure is used to rename datafiles that were pointing to the backup set to the actual file name in cloned database.

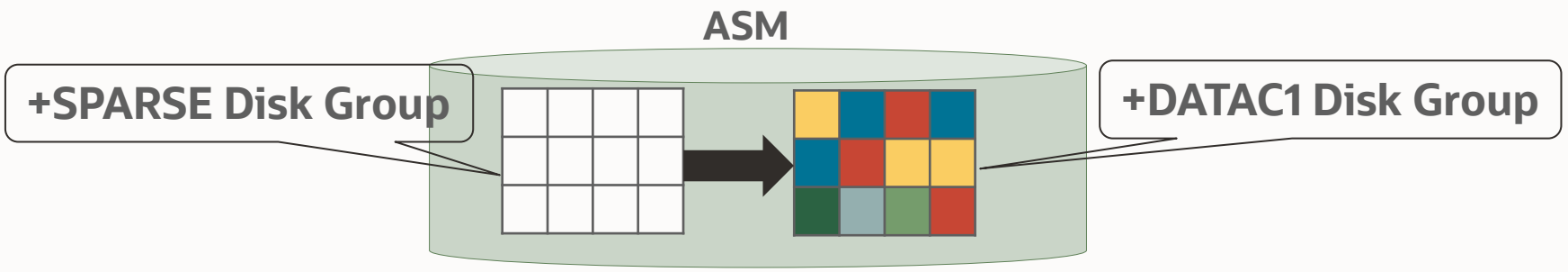
The **srcfile** is the file name that represents the data file in the backup image copy or a read-only storage snapshot.

The **destfile** destination file path must point to a NFS volume where cloneDB datafiles will be created. When the procedure is run successfully, the control file record is updated with the new datafile name.

Syntax

```
DBMS_DNFS.CLONEDB_RENAMEFILE (
   srcfile              IN          VARCHAR2,
   destfile             IN          VARCHAR2);
```

```
EXECUTE DBMS_DNFS.CLONEDB_RENAMEFILE
('+DATAC1/DGCDBDB/DATAFILE/tbs_5tb_1_352_1104239471','+SPARS
E/SNAPDB01/DATAFILE/tbs_5tb_1_352_1104239471');
```



https://docs.oracle.com/en/database/oracle/oracle-database/19/arpls/DBMS_DNFS.html#GUID-63D23FA1-3C95-4B49-B42A-830D0ED5132E

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

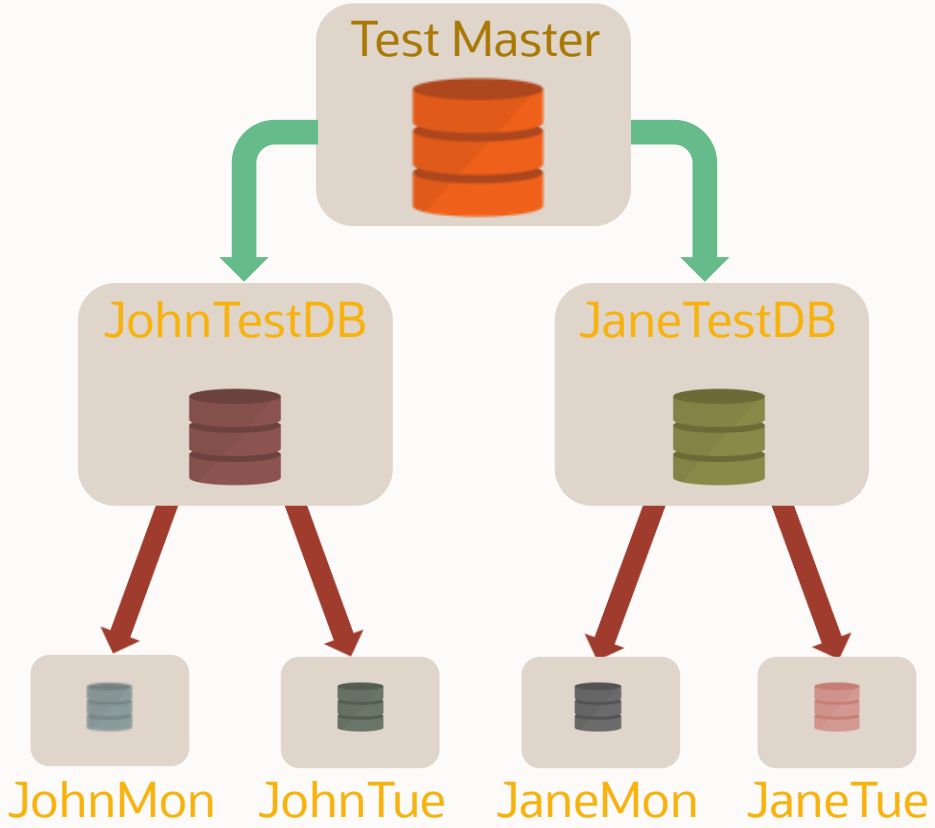
ACFS Snapshots

Summary

分层快照
Architecture

- 从以前创建的快照创建数据库的快照

  CREATE PLUGGABLE DATABASE JOHNMon from JOHNTESTDB create_file_dest= ‘+SPARSE’ SNAPSHOT COPY;
- 语法和技术保持不变
- 适用于可插拔和不可插拔数据库
- 深度要<10，为了性能
- 确保上层快照为只读



Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

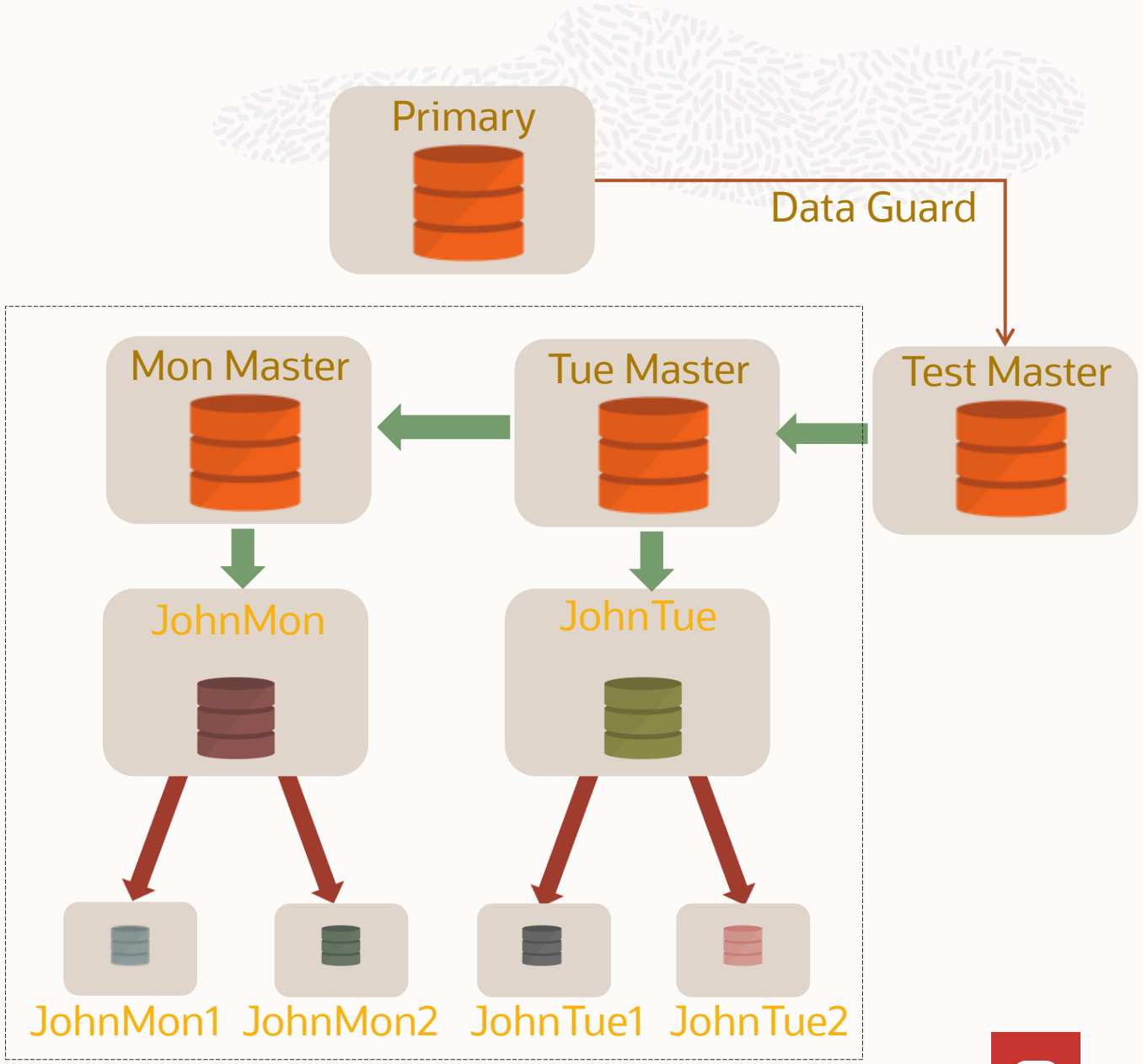
ACFS Snapshots

Summary

Sparse Test Masters

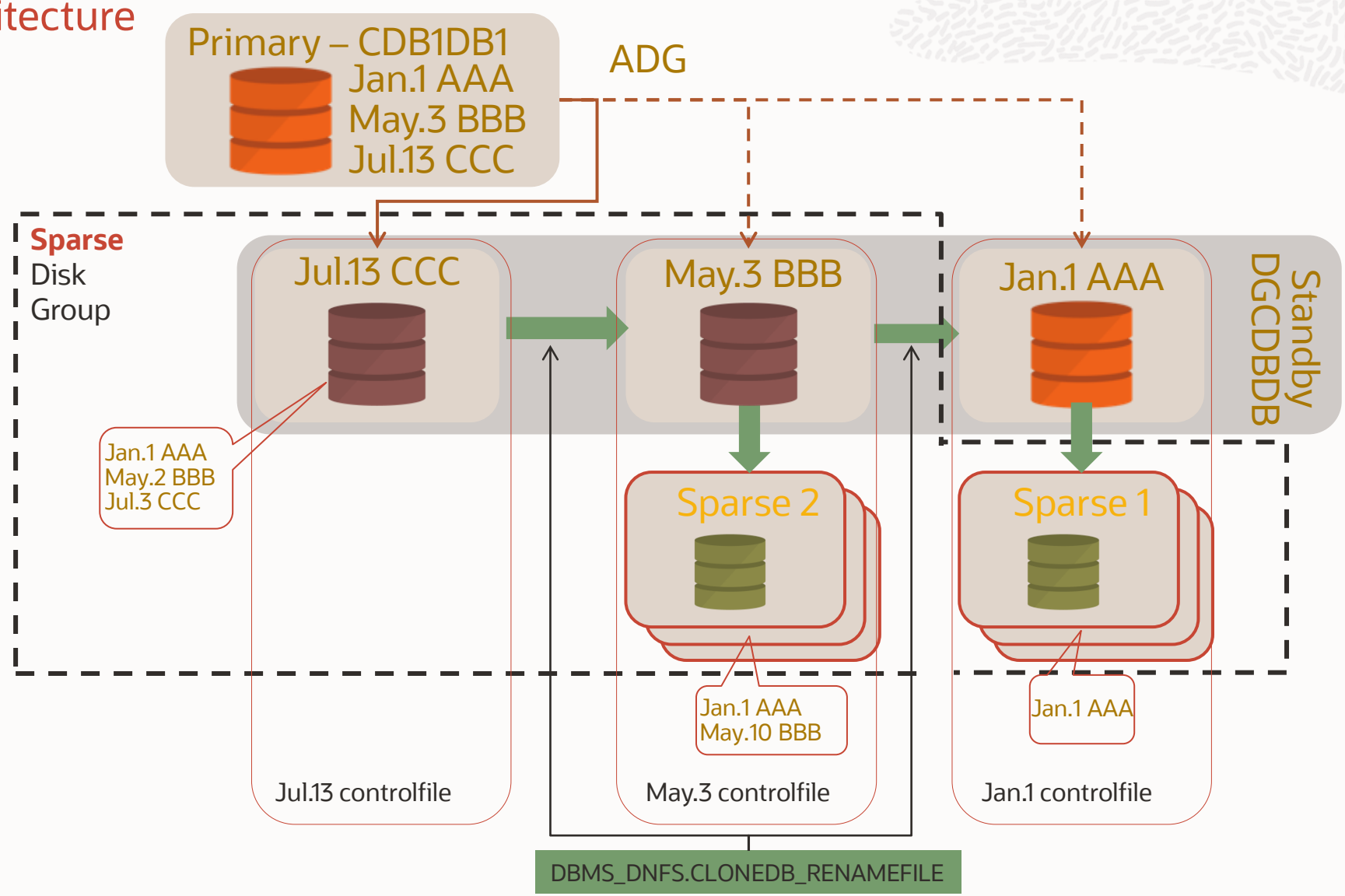
Architecture

- Use cases
  - 能够在不同时间点创建快照；没有在每个时间点用全部空间。
  - Test Master 可以是可写的 Data Guard Target
- Steps
  - 在Test Master的任何时候；停止redo apply应用；并创建一个“Mon Master”。
  - 创建一个新的Test Master “Mon Master”是空间节省的快照
  - “Mon Master”现在是只读的，可以被视为创建其他测试/开发快照的母版。
  - 重复 Test Master 上的步骤以创建“Tue Master”（节省空间）。
  - ‘Tue Master’ 是一个稀疏的 Test Master



Sparse Test Masters

Architecture

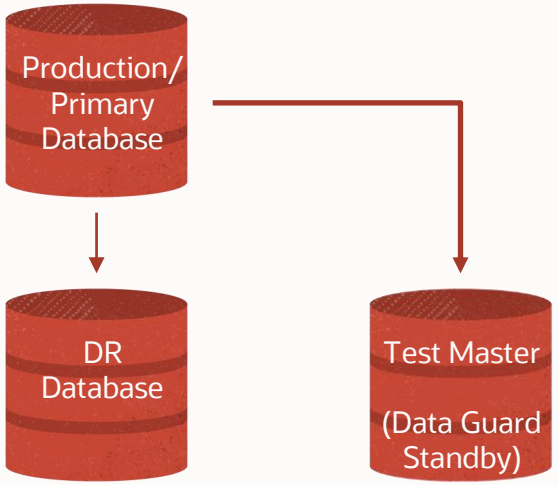


Sparse Test Masters

Lifecycle Start (Create Test Master)

创建新的Data Guard 或者 RMAN 全量的 Test Master (TM) – 可以在sparse 或者 non-sparse diskgroup

- Sparse Test Master 的文件必须在 sparse diskgroup

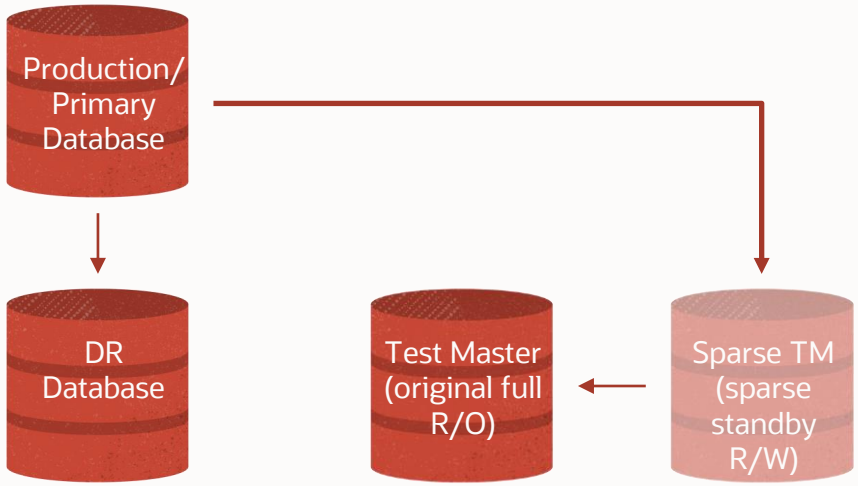


停止 Data Guard Redo Apply

- 转变为 Read Only Test Master

创建新的 Sparse Test Master (for Standby)

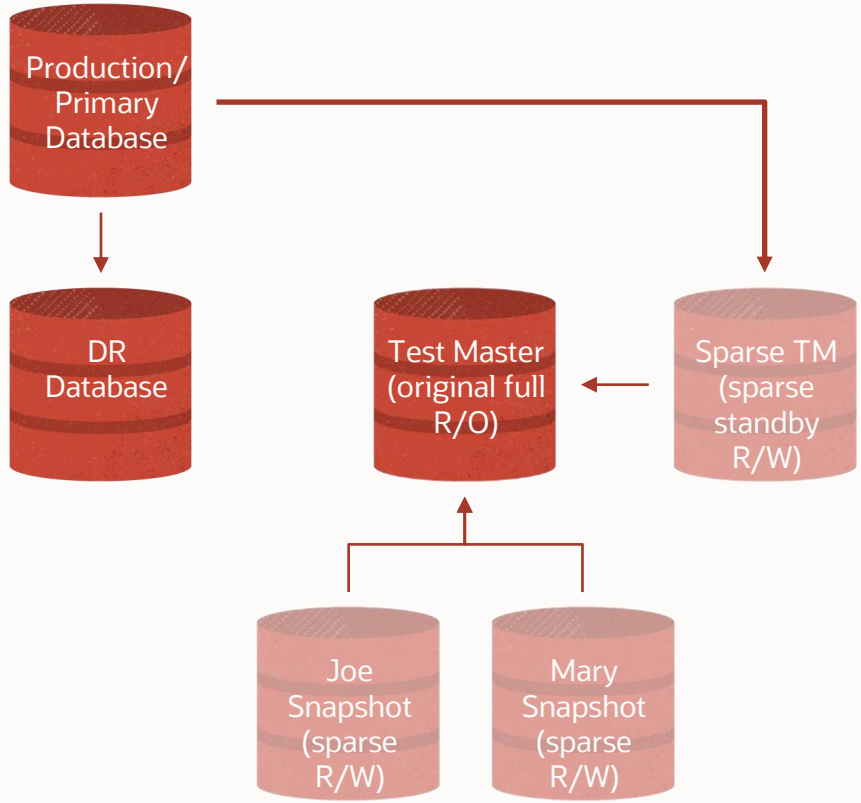
- 开始 Data Guard Redo Apply， 同步数据



Sparse Test Masters

Create Snapshots

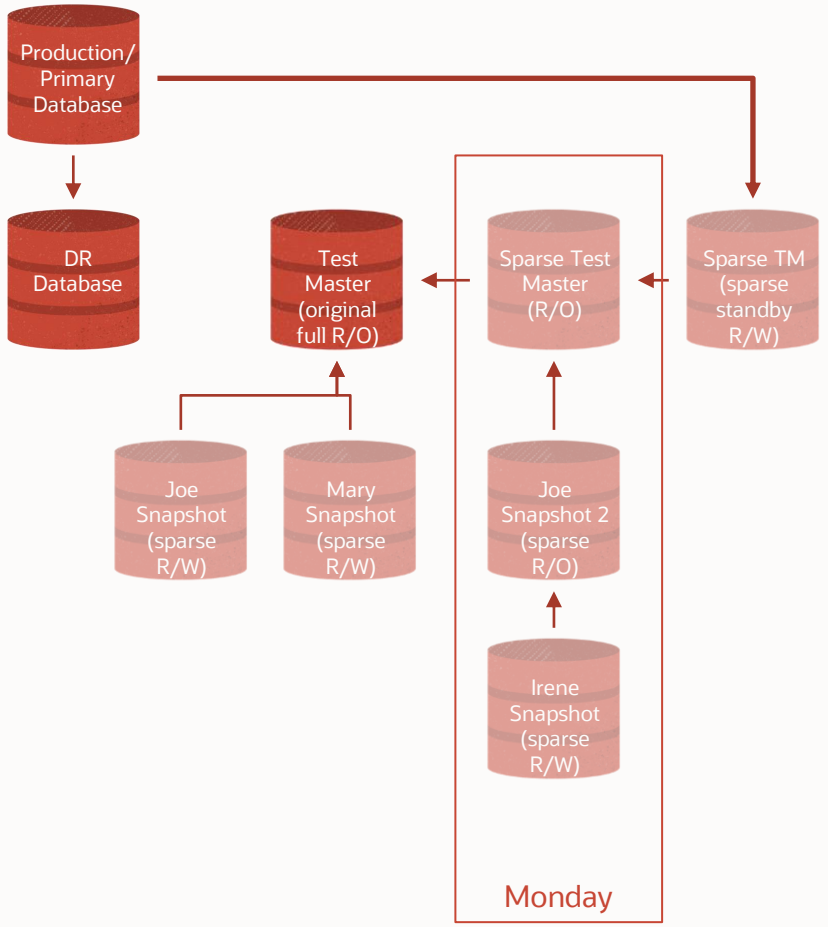
- 初始Test Master 是一个 Read Only 的 Test Master
- 创建 Exadata snapshots
  - Joe Snap
  - Mary Snap
- Sparse Test Master for Standby receives redo from primary and writes to sparse datafiles
- Sparse Test Master for Standby 从主服务器接收重做并写入sparse datafiles



Sparse Test Masters

刷新 Test Master + 创建新的 Snapshots

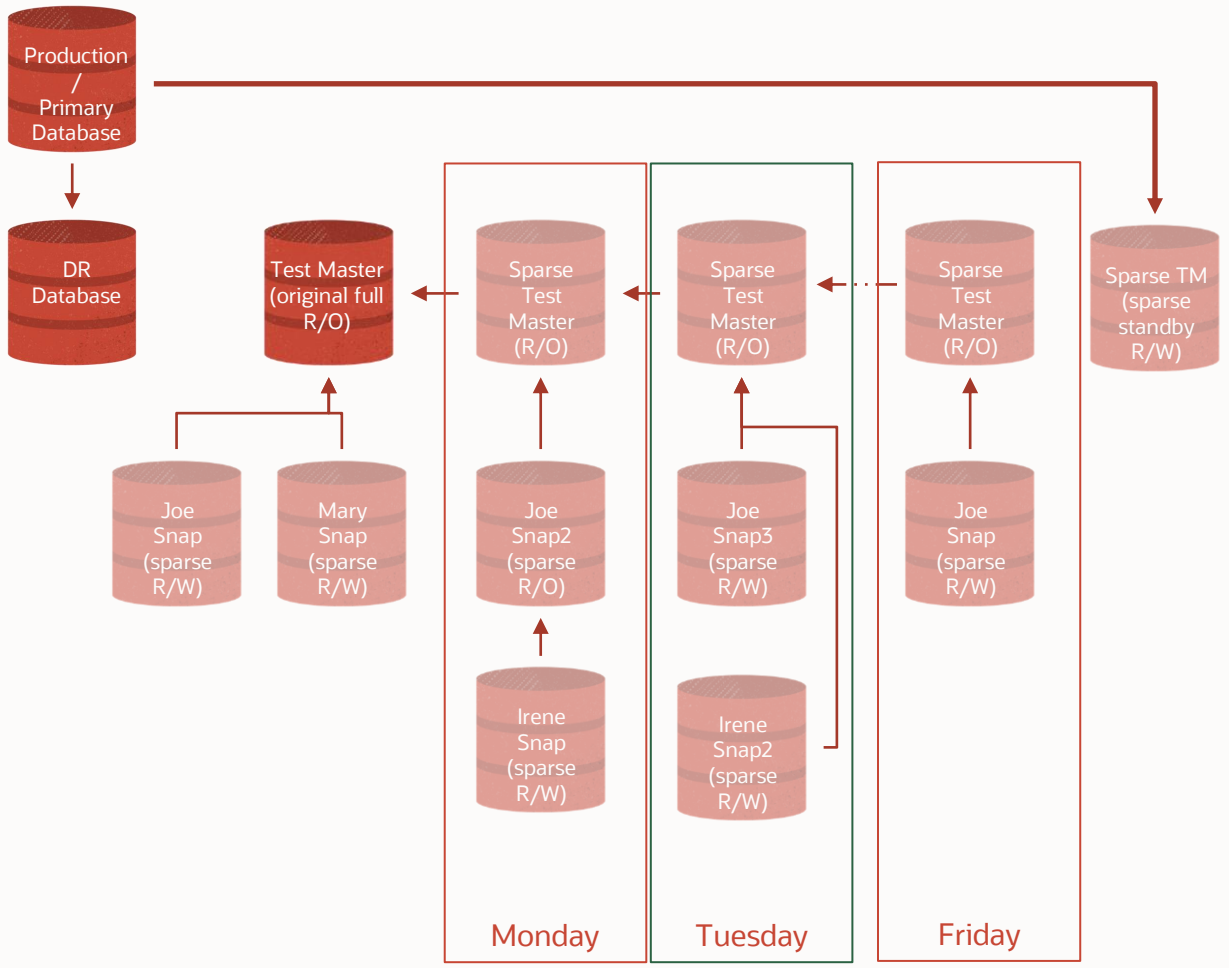
- 重复创建新的Exadata snapshots
- 重复流程以创建新的 Exadata Snapshots， 同时保留之前的 Exadata 快照
- 所有 Sparse Test Masters 和 Snapshots 都是稀疏的



Sparse Test Masters

Additional Test Master Refreshes

- 建议最多 10 个级别 - 9 个Sparse Test Masters
- 出于性能原因，我们建议限制分层快照的深度
- 达到限制时可以
  - 删除所有快照并同步 TM - 或 -
  - 空间允许创建新的完整的TM



Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

ACFS Snapshots

Summary

Monitoring and Statistics

- Sparse IO stats in RDBMS
- Wait events in RDBMS
- v$ views
    - v$asm_disk_sparse
    - v$asm_diskgroup_sparse
    - v$clonedfile

RDBMS stats (v$sysstat; v$mystat)

New stats

Name	Meaning
Physical read snap IO request no data	No physical IO done for these (i.e. wasted roundtrips)
Physical read snap IO request base	Number of physical IOs on base level
Physical read snap IO request copy	Number of physical IOs on any snap hierarchy
Physical read snap bytes base	Number of bytes read from the base
Physical read snap bytes copy	Number of bytes read from the snap

RDBMS stats (v$sysstat; v$mystat) contd…

Updated stats

Name	Meaning
Physical read total IO requests	Number of physical IOs submitted by user
Physical read total multi block requests	Number of multi block IOs submitted by user

- Only the IOs that lead to a real physical IO will be counted here, using the same logic as described in wait events to omit completely sparse IOs.

RDBMS wait events

- Following wait events are monitored for 0 byte reads returned; i.e. sparse buffers
  - cell single block physical read
  - cell multi block physical read
  - cell list of blocks physical read
- List of blocks wait events are also tracked
- Then, we change the wait event to "**cell sparse block physical read**"
  - this wait event is significantly faster since there is no IO involved and if the request is large in size, then even network transfer is significantly faster because of packing

ASM Sparse Disk (v$asm_disk_sparse)

Name	Meaning
GROUP_NUMBER	Number of the diskgroup containing the disk
DISK_NUMBER	Number assigned to the disk within this diskgroup
INCARNATION	Incarnation number for the disk
ALLOCATED_MAT_MB	Total used physical capacity on this disk
TOTAL_MAT_MB	Total physical capacity on this disk
SPARSE_READS	Number of read requests on sparse regions of this disk
SPARSE_BYTES_READ	Bytes read from sparse regions of this disk
SPARSE_READ_TIME	Time taken by sparse read IOs

v$asm_disk_sparse

```
SQL> select
     DISK_NUMBER           dsk_num,
     ALLOCATED_MAT_MB      alloc,
     TOTAL_MAT_MB          total
from V$ASM_DISK_SPARSE
where GROUP_NUMBER = 5;

DSK_NUM    ALLOC      TOTAL
---------- ---------- ----------
         0       5536     204774
         1       5424     204774
         2       5532     204774
         3       5424     204774
         4       5424     204774
```

ASM Sparse Diskgroup (v$asm_diskgroup_sparse)

Name	Meaning
GROUP_NUMBER	Cluster wide number assigned to the diskgroup
ALLOCATED_MAT_MB	Total used physical capacity of the diskgroup
TOTAL_MAT_MB	Total physical capacity of the diskgroup

```
SQL> select
       ALLOCATED_MAT_MB    alloc,
       TOTAL_MAT_MB        total
from V$ASM_DISKGROUP_SPARSE
where GROUP_NUMBER = 5;

ALLOC      TOTAL
---------- ----------
    197208    7371864
```


v$clonedfile

- Only works on mounted databases/files
- Can be run in either database instance or in ASM
  - In snapshot instance will display parent files for that snapshot
  - In ASM instance, possible to see parent/child relationships for all open/mounted snapshots

```
SQL> select FILENUMBER
     , SNAPSHOTFILENAME
     , CLONEFILENAME
  from V$CLONEDFILE;

FILENUMBER SNAPSHOTFILENAME
---------- --------------------------
CLONEFILENAME
--------------------------------------------------------------------------------
16         +DATA/TESTMASTER/09D05108AB70216BE053D6CBF00AA040/DATAFILE/system.257.865863315
+SPARSEDG/JOHNTEST/09D05108AB70216BE053D6CBF00AA041/DATAFILE/system.257.865863315
17         +DATA/TESTMASTER/09D05108AB70216BE053D6CBF00AA040/DATAFILE/sysaux.258.865863317
+SPARSEDG/JOHNTEST/09D05108AB70216BE053D6CBF00AA041/DATAFILE/sysaux.258.865863317
```

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

ACFS Snapshots

Summary

Resize operations

Resizing

- 可以改变稀疏的虚拟或物理空间
  - 注意每个磁盘 100Tb 的虚拟大小限制和 4Tb 的物理大小限制
- 更改虚拟空间
  - 增大
    - First alter the cell disks
      CellCLI> alter griddisk SPARSE_CD_00_CELL01,SPARSE_CD_01_CELL01,....,SPARSE_CD_11_CELL01 virtualSize=newBiggerSize;
    - Then alter the disk group in an ASM instance
      SQL> alter diskgroup SPARSE resize all size newBiggerSize;
  - 减小
    - Ensure you have the free space to reduce virtual size
    - First alter the diskgroup in ASM
      SQL> alter diskgroup SPARSE resize all size newSmallerSize;
    - Then alter the cell disks
      CellCLI> alter griddisk SPARSE_CD_00_CELL01,SPARSE_CD_01_CELL01,....,SPARSE_CD_11_CELL01 virtualSize=newSmallerSize;

Release

All Exadata Snapshot features

- Available
  - Database software → 12.2.0.1
  - Grid software → 12.2.0.1
  - Cell software → 12.2.1.1.0
  - RMAN sparse backups → 18.1.0.0
- Recommended 19c or later

Documentation

- Exadata Storage Server Software User's Guide
  - Chapter 10 → Setting up Oracle Exadata Storage Snapshots
- https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-78F67DD0-93C8-4944-A8F0-900D910A06A0

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

ACFS Snapshots

Summary

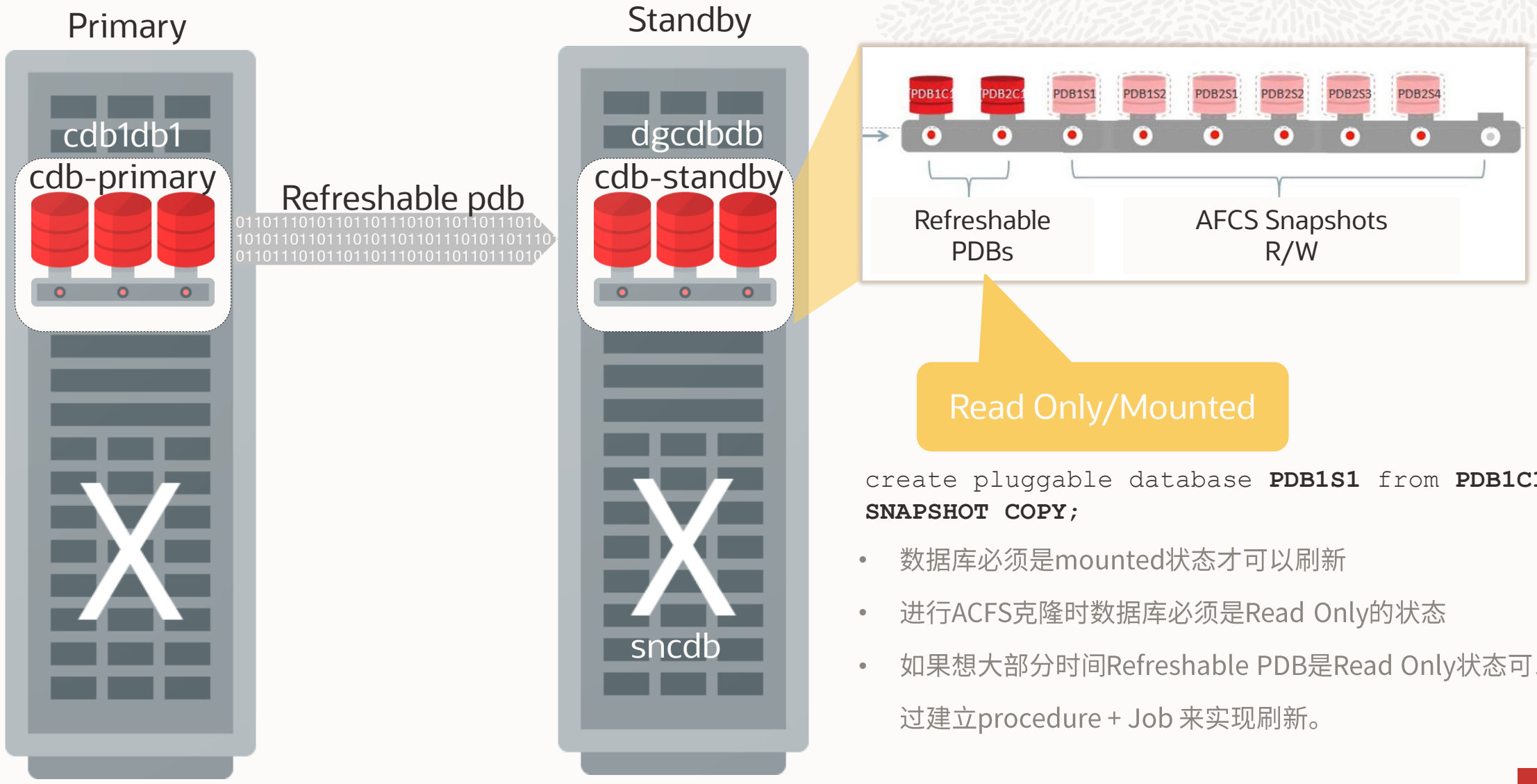
ACFS Snapshots

- 先决条件
  - 19.10 (plus performance & stability patches) or later Grid Infrastructure
  - COMPATIBLE.ASM 和 COMPATIBLE.ADVM 匹配GI版本的最新功能访问
- 优势
  - 解决方案具有类似于第三方写时拷贝快照的特性和功能
  - 支持数据库11gR2 (11.2.0.4) 或更高版本
- Limitations
  - 除了Flash Cache，没有Exadata智能卸载功能
  - ACFS加密Oracle数据库文件不支持

ACFS Exadata Snapshot Use Cases

- 支持的数据库
  - Singleton PDBs
  - Full non-CDBs
  - Full CDBs
- 可拥有多个时间轴的读写测试主机
  - 源数据的单个完整副本(例如，物理备用数据库)
  - 定期创建只读快照
    - 作为时间轴的测试模板
  - 完整复制源数据将继续更新
  - 能够创建多达1023个快照

Exadata ACFS Database Snapshots for PDB

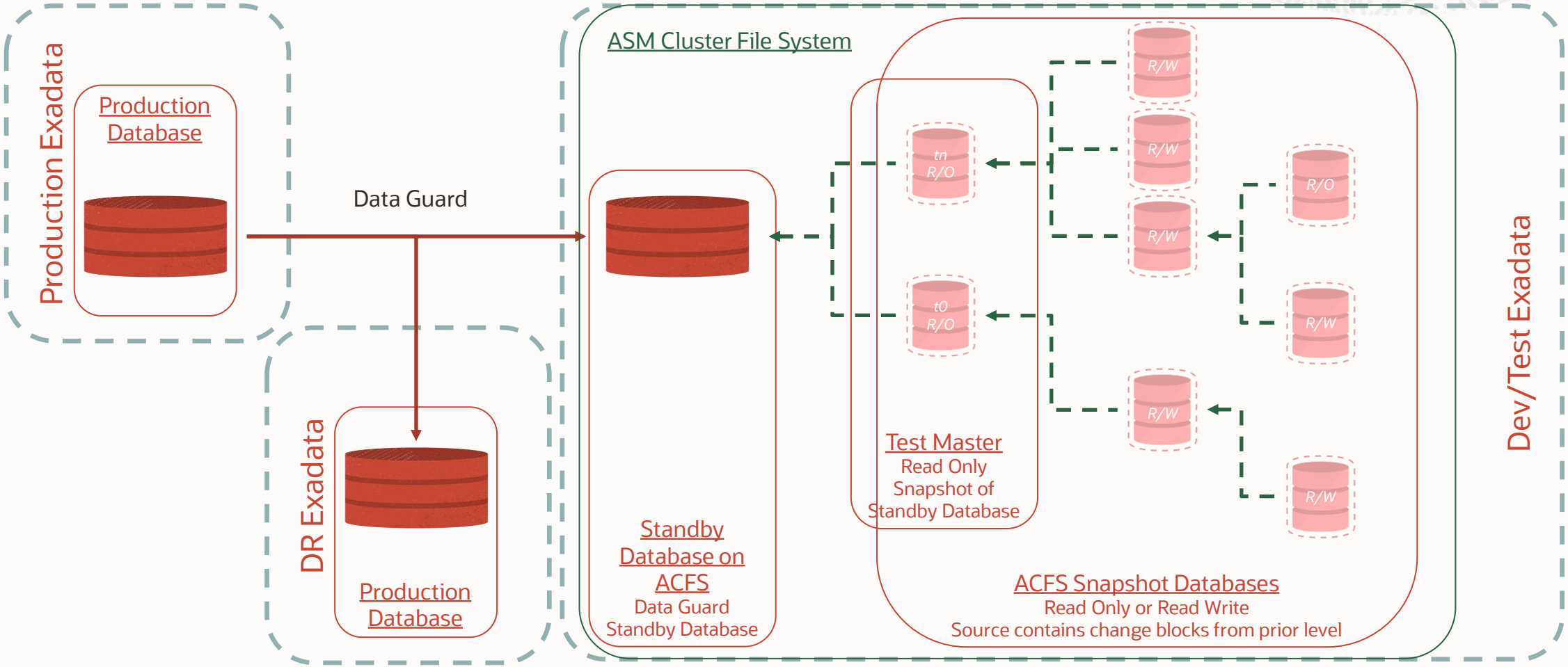



```
create pluggable database PDB1S1 from PDB1C1
SNAPSHOT COPY;
```

- 数据库必须是mounted状态才可以刷新
- 进行ACFS克隆时数据库必须是Read Only的状态
- 如果想大部分时间Refreshable PDB是Read Only状态可以通过建立procedure + Job 来实现刷新。

Exadata ACFS Database Snapshots For CDB

概念视图



Exadata ACFS Database Snapshots For CDB

创建快照

1. 备份控制文件
2. 创建PFILE
3. 关闭数据库创建Snapshot
4. 修改pfile & 控制文件
5. 拷贝密码文件
6. 启动ACFS Snapshot数据库

Exadata ACFS Database Snapshots For CDB

创建快照

```
##创建名称为clone的ACFS Snapshot，-w参数为读写，-r为只读
/sbin/acfsutil snap create -w clone /mnt/dbvol/
##查看ACFS Snapshot
/sbin/acfsutil snap info clone /mnt/dbvol/
```

```
[oracle@x8efdbadm01 ~]$ /sbin/acfsutil snap create -w clone /mnt/dbvol/
acfsutil snap create: Snapshot operation is complete.
[oracle@x8efdbadm01 ~]$ /sbin/acfsutil snap info clone /mnt/dbvol/
snapshot name:               clone
snapshot location:           /mnt/dbvol/.ACFS/snaps/clone
RO snapshot or RW snapshot:  RW
parent name:                 /mnt/dbvol/
snapshot creation time:      Thu May 12 21:40:08 2022
file entry table allocation: 17170432   (  16.38 MB )
storage added to snapshot:   17170432   (  16.38 MB )


[oracle@x8efdbadm01 ~]$
```

Lifecycle Operations

高可用性和文件放置

- 仅对测试/开发数据库使用Exadata上的ACFS
- 测试主数据库(备数据库)与灾备数据库不能相同
- 始终建议使用高冗余磁盘组以获得最佳的存储保护和高可用性
- ACFS“DATA”文件系统将创建在+DATA上，应该包含您的数据库文件，在线日志，控制文件和spfile
- ACFS“RECO”文件系统将创建在+RECO上，应该包含您的存档文件和闪回日志
- 不要将Exadata ACFS用于Oracle Home、诊断目标、审计目标或安全/加密钱包
- 不要将测试/开发数据库和生产数据库放在同一个RAC或VM集群上

Lifecycle Operations

备份/恢复

- 对于RMAN来说，ACFS快照看起来像是一个普通的数据库文件，而不是一个“稀疏”的数据库。RMAN不支持ACFS快照。

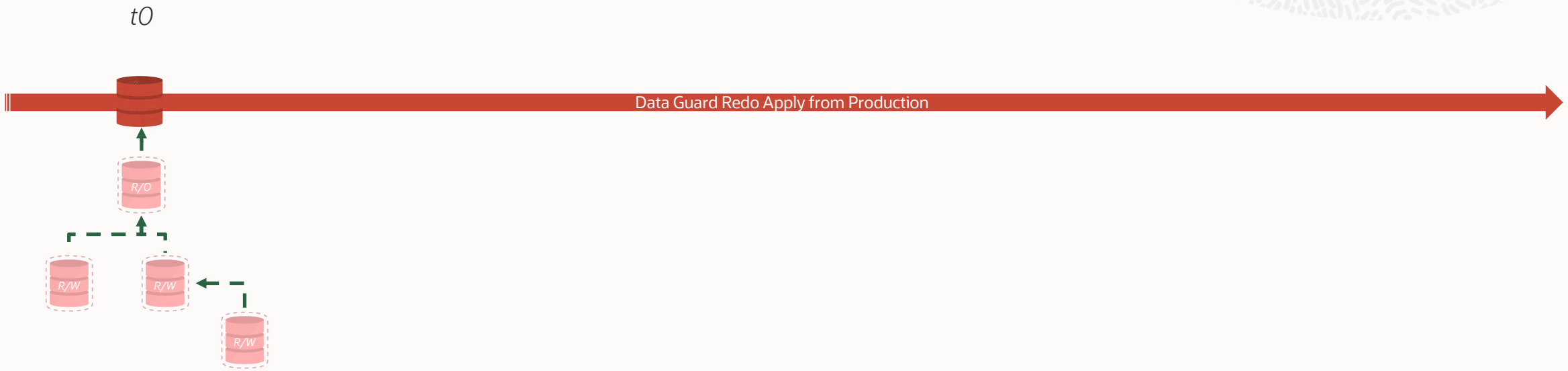
软件更新

- 软件和数据库更新的选项与非ACFS数据库相同
- 创建、调整和删除ACFS文件系统非常简单。请参阅文档ID 2761360.1

ACFS Read/Write Test Master

- 创建 ACFS 文件系统
- 创建Physical standby 数据库在 ACFS
  - 使用 RMAN Duplicate for Standby – 或者 –
  - gDBClone

ACFS Read/Write Test Master

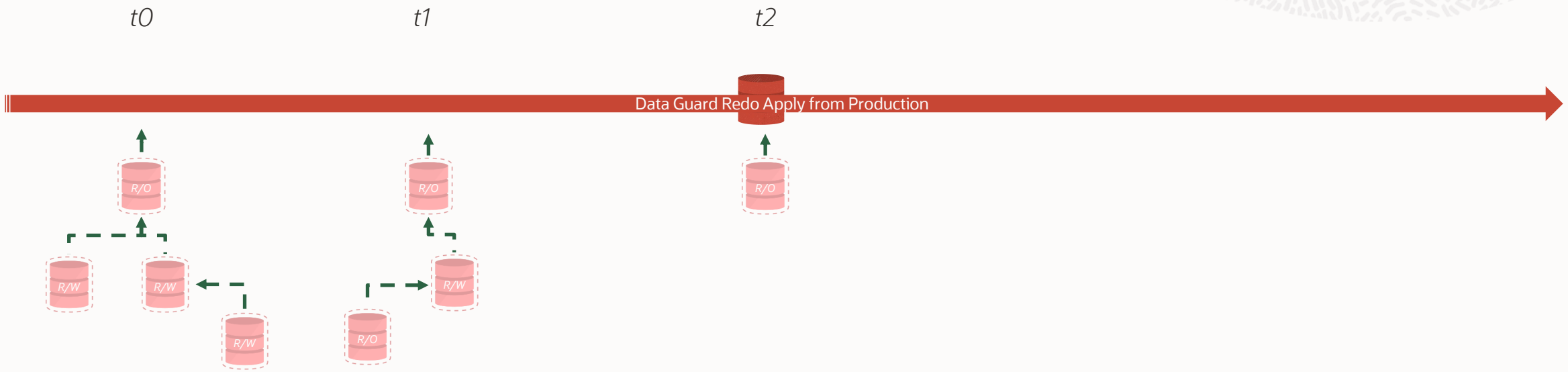


- 在t0时间线创建
  - 停止 redo apply
  - 创建 Read-Only ACFS snapshot 为母版
  - 重启 Redo Apply
  - 根据测试用例的需要创建额外的读写或只读ACFS快照

ACFS Read/Write Test Master

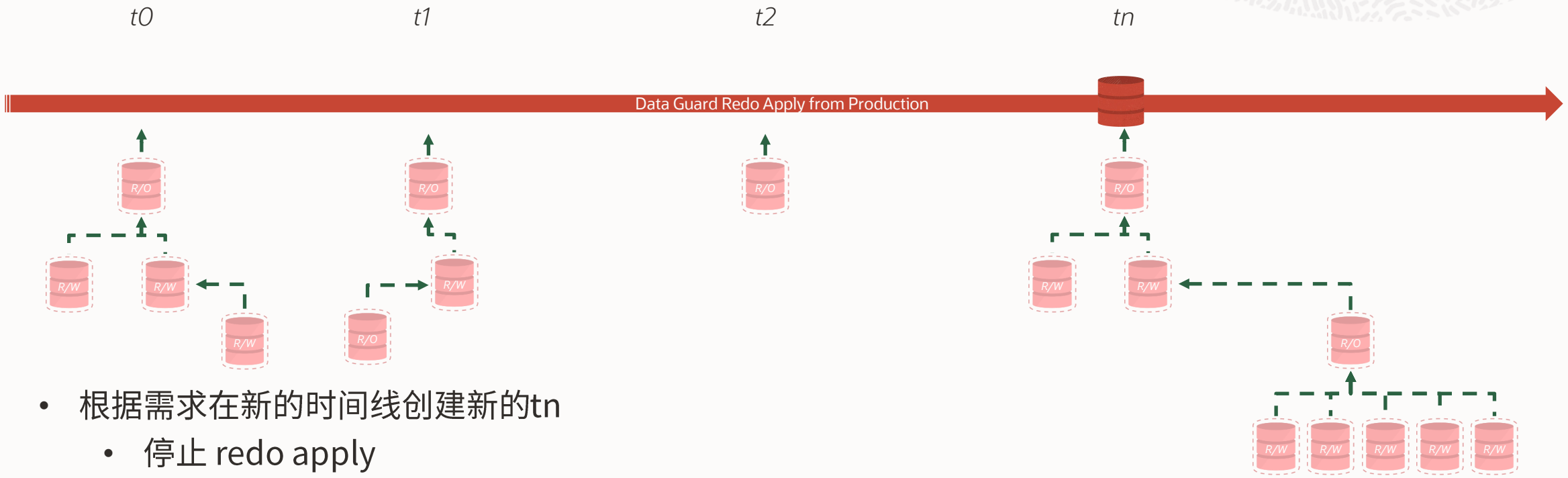
- 根据需求在新的时间线创建新的t1
  - 停止 redo apply
  - 创建 Read-Only ACFS snapshot 为母版
  - 重启 Redo Apply
  - 根据测试用例的需要创建额外的读写或只读ACFS快照

ACFS Read/Write Test Master



- 根据需求在新的时间线创建新的t2
  - 停止 redo apply
  - 创建 Read-Only ACFS snapshot 为母版
  - 重启 Redo Apply
  - 根据测试用例的需要创建额外的读写或只读ACFS快照

ACFS Read/Write Test Master



- 根据需求在新的时间线创建新的tn
  - 停止 redo apply
  - 创建 Read-Only ACFS snapshot 为母版
  - 重启 Redo Apply
  - 根据测试用例的需要创建额外的读写或只读ACFS快照

ACFS Snapshot References

- Oracle ACFS Support on Oracle Exadata Database Machine (Linux only) (Doc ID 1929629.1)
- Oracle ACFS Snapshot Use Cases on Exadata (Doc ID 2761360.1)
- Oracle Automatic Storage Management Cluster File System - Administrator's Guide (https://docs.oracle.com/en/database/oracle/oracle-database/19/ostmg/index.html)
  - Creating an Oracle ACFS File
  - Managing Oracle ACFS Snapshots
  - How to Clone a Master Database with ACFS Snapshots
- Oracle System Software – User's Guide for Exadata Sparse to compare (https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/index.html)
  - Setting up Oracle Exadata Storage Snapshots

Agenda

Exadata Sparse Clones

- 特性
- 分层快照
- Sparse Test Masters
- 监控与统计
- 资源

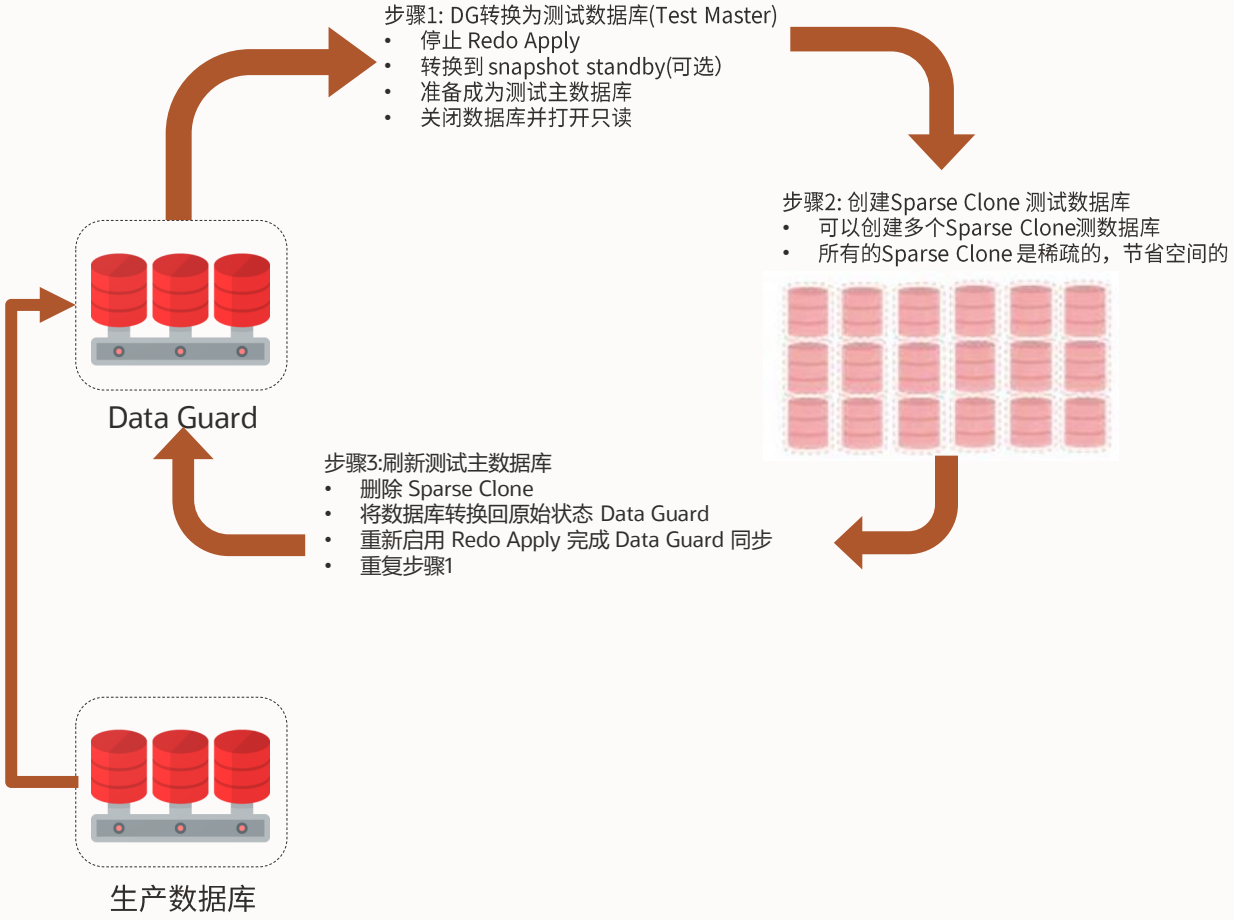
ACFS Snapshots

Summary

Exadata Database Sparse Clones & ACFS Snapshots

Feature/Requirement	Exadata Sparse Clones	ACFS Snapshots on Exadata
Exadata 特性	All Exadata Features available	Exadata Smart Flash Cache
CDB & PDB	Yes	Yes
Non-CDB	Yes	Yes
Enterprise Manager	Yes	No
快照的Source状态(母版)	read-only	read/write
快照的Source位置(母版)	ASM任意Disk Group	ACFS
快照存储位置	ASM +SPARSE Disk Group	ACFS

Exadata Sparse Clones - CDB



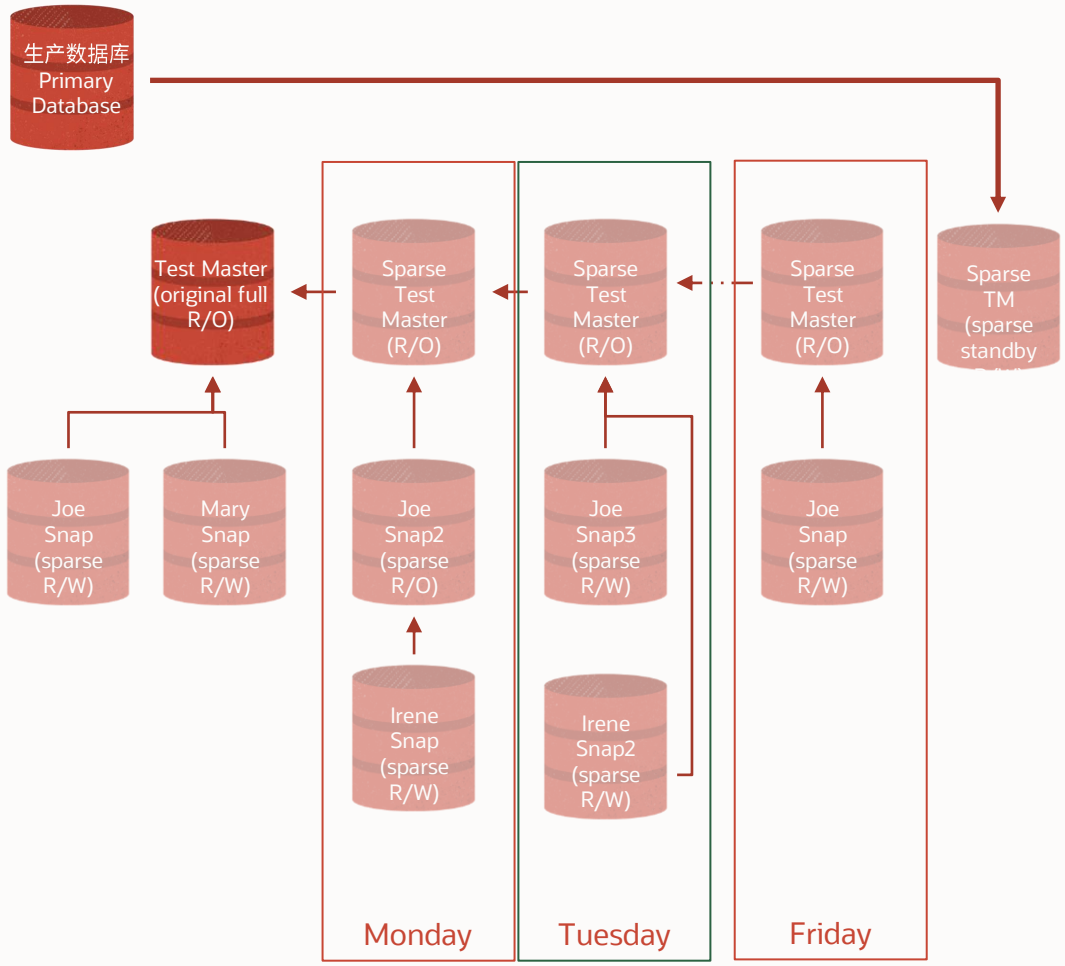
优点：
1. 以一定的步骤快速完成
2. 可以创建多个空间节省的R/W 测试环境（Sparse Snapshot）

缺点：
1. DG同步数据时必须删除所有的测试数据库
2. 只有一个时间点的数据

时间：
1. 取决于创建空间节省的R/W 测试环境的时间
2. 不涉及数据拷贝或复制

https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-D15BAA38-218A-4A61-A4E8-594A7528B204

Exadata Sparse Clones + Sparse Test Master- CDB



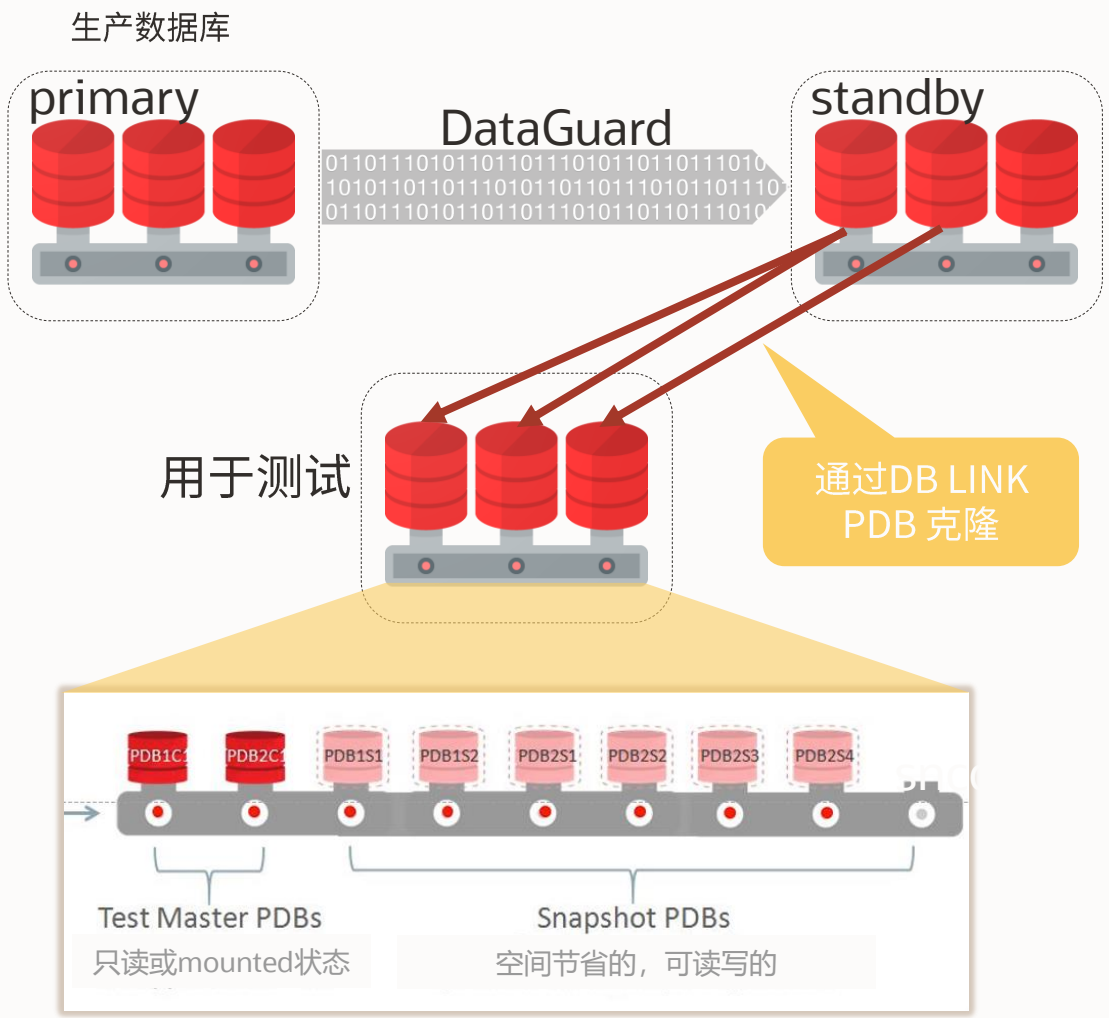
优点：
1. 以一定的步骤快速完成
2. 可以在任何时间点创建Sparse Test Master（空间节省的，不需要全量复制）
3. 可以创建多个空间节省的R/W 测试环境（Sparse Snapshot）

缺点：
1. 出于性能考虑最多能创建9个Sparse Test Master，达到9个后后需要删除所有快照同步 Test Master - 或 - 空间允许创建新的完整的TM
2. 基于前一个 Sparse Test Master 创建新的Sparse Test Master，不能删除中间的Sparse Test Master。

时间：
1. 取决于创建Sparse Test Master +空间节省的R/W Sparse Clone 测试环境的时间
2. 不涉及数据拷贝或复制

https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-4E58569A-BB16-462F-BFC6-9F67FAAC64EA

Exadata Sparse Clones - PDB



优点：
1. 非常简单快速的以PDB提供测试环境
2. 可以在任何时间点克隆创建Test Master PDB
3. 可以创建多个空间节省的R/W 测试环境（Sparse Snapshot PDBs）

缺点：
1. Test Master PDB是全量的克隆
2. 根据PDB的数据文件的总大小克隆的时间不同

时间：
1. 取决克隆PDB的时间（可以配合使用parallel_pdb_creation_clause加速克隆时间）

测试内容 Exadata X8-2 EF

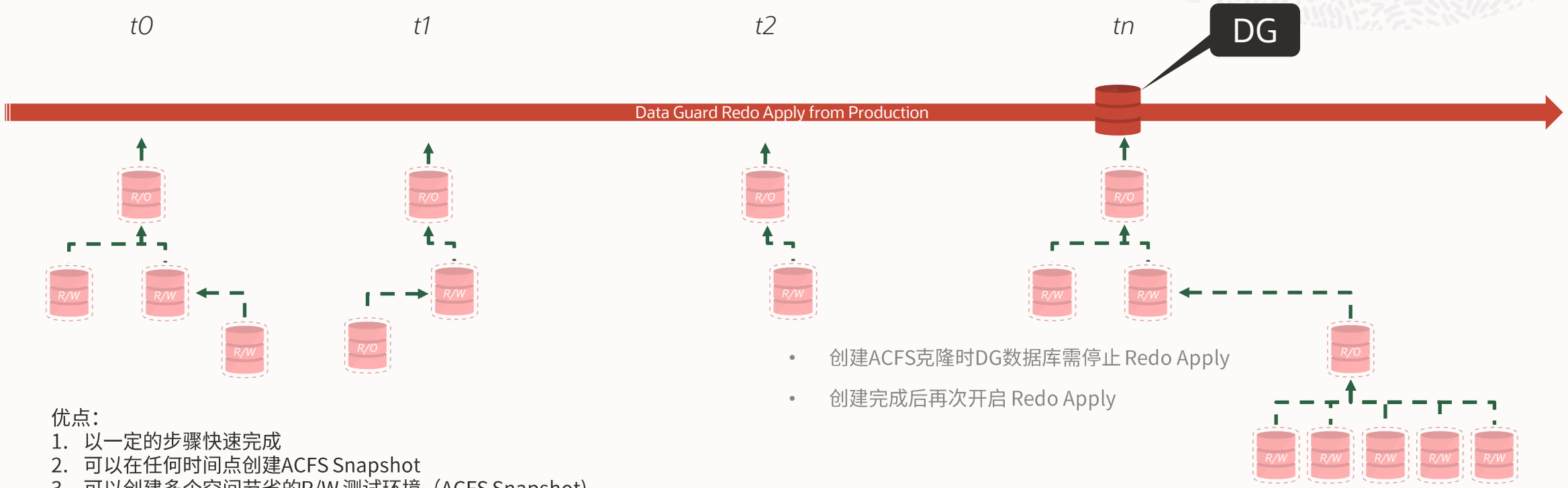
Table + Add view

场景	说明	用时
场景一	100GB数据文件PDB远程克龙 - 无实际数据	39.01秒
场景二	100GB数据文件PDB远程克龙 - 76GB数据	49.12秒
场景三	500GB数据文件PDB远程克龙 - 76GB数据	3分8.5秒
场景四	500GB数据文件PDB远程克龙 - 412GB数据	3分15.6秒
场景五	1TB数据文件PDB远程克龙 - 412GB数据	6分4.4秒
场景六	1TB数据文件PDB Sparse Snapshot 克龙 - 412GB数据	13.03秒

```
SQL> create pluggable database PDB1S1 from
PDB1C1
create_file_dest='+SPARSEDG'
SNAPSHOT COPY;
```

https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-F85C2D2F-F807-43F0-9ED0-E4A9743DFAFD

Exadata ACFS Snapshots - CDB



- 创建ACFS克隆时DG数据库需停止 Redo Apply
- 创建完成后再次开启 Redo Apply

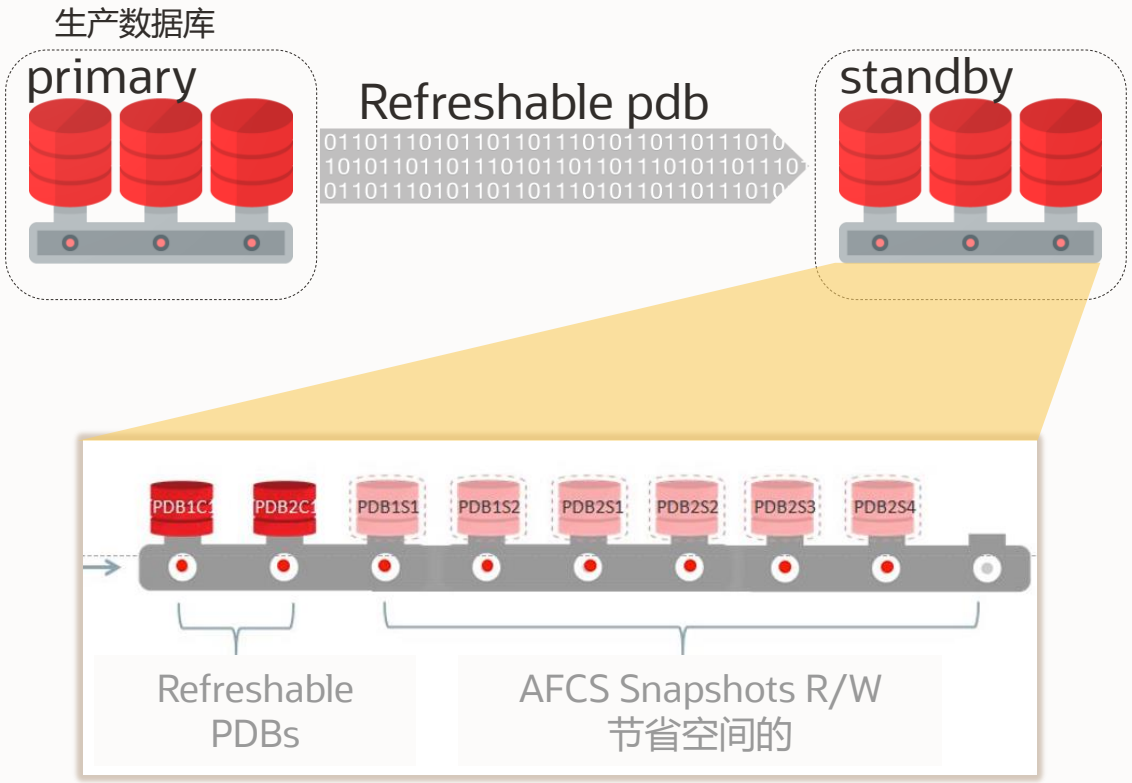
优点：
1. 以一定的步骤快速完成
2. 可以在任何时间点创建ACFS Snapshot
3. 可以创建多个空间节省的R/W 测试环境（ACFS Snapshot)

缺点：
1. 无法使用Exadata的智能功能
2. 所有数据库需在ACFS上。

时间：
1. 取决于创建空间节省的R/W 的ACFS Snapshot 测试环境的时间
2. 不涉及数据拷贝或复制

https://support.oracle.com/epmos/faces/DocumentDisplay?_afrLoop=603805633789089&id=2761360.1&_afrWindowMode=0&_adf.ctrl-state=kgn8xzbnk_4

Exadata ACFS Snapshots - PDB



- 数据库必须是closed状态才可以刷新
- 进行ACFS克隆时数据库必须是Read Only的状态

优点：
1. 非常简单快速的以PDB提供测试环境
2. 可以在任何时间点创建ACFS Snapshot
3. 可以创建多个空间节省的R/W 测试环境（ACFS Snapshot PDBs)

缺点：
1. 无法使用Exadata的智能功能

时间：
1. 取决于刷新PDB的时间（定时刷新后无需再次手动刷新时可忽略）

PS: Exadata Sparse Clone 也可以使用Refreshable PDB，但是刷新前需要把所有的Sparse Snapshot PDB删除后在刷新。

```
SQL> create pluggable database JohnTest from TestMaster SNAPSHOT COPY;
```


https://support.oracle.com/epmos/faces/DocumentDisplay?_afrLoop=603805633789089&id=2761360.1&_afrWindowMode=0&_adf.ctrl-state=kgn8xzbnk_4

附加内容

从DG端进行 Refreshable PDB(不可行)

现阶段不支持原端为 physical standby database进行 Refreshing PDB 。
通过dblink进行PDB的远程克隆原端是physical standby database是支持的。

ORA-65345 When Refreshing PDB Sourced From Active Data Guard Standby (Doc ID 2765472.1)

```
alter pluggable database <pdb name> refresh
*
ERROR at line 1:
ORA-65345: cannot refresh pluggable database
```


https://database-heartbeat.com/2021/05/18/how-to-clone-a-pdb-from-a-standby-database-in-a-data-guard-environment/

资源

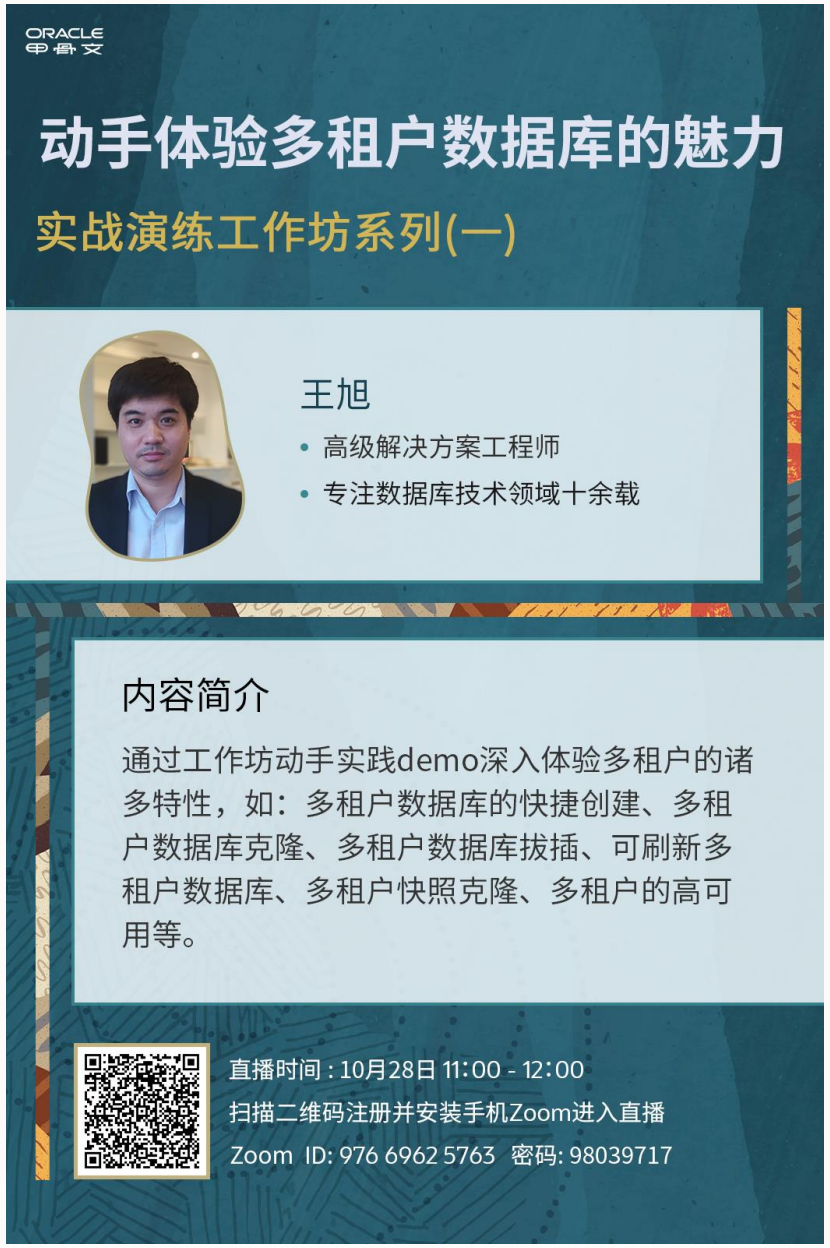
Sparse Clone

- Exadata Storage Server Software User's Guide
  - Chapter 10 → Setting up Oracle Exadata Storage Snapshots
- https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/exadata-storage-server-snapshots.html#GUID-78F67DD0-93C8-4944-A8F0-900D910A06A0

ACFS Snapshot

- Oracle ACFS Support on Oracle Exadata Database Machine (Linux only) (Doc ID 1929629.1)
- Oracle ACFS Snapshot Use Cases on Exadata (Doc ID 2761360.1)
- Oracle Automatic Storage Management Cluster File System - Administrator's Guide (https://docs.oracle.com/en/database/oracle/oracle-database/19/ostmg/index.html)
  - Creating an Oracle ACFS File
  - Managing Oracle ACFS Snapshots
  - How to Clone a Master Database with ACFS Snapshots
- Oracle System Software – User's Guide for Exadata Sparse to compare (https://docs.oracle.com/en/engineered-systems/exadata-database-machine/sagug/index.html)
  - Setting up Oracle Exadata Storage Snapshots

https://chulnamkim.notion.site/Snapshot-65325b93d6c1485b8b727945da67921b



数据库和云讲座群

20-19

甲骨文云技术公众号

技术专家1V1深入交流