ORACLE

Continuous Availability

数据库与应用持续可用性设计

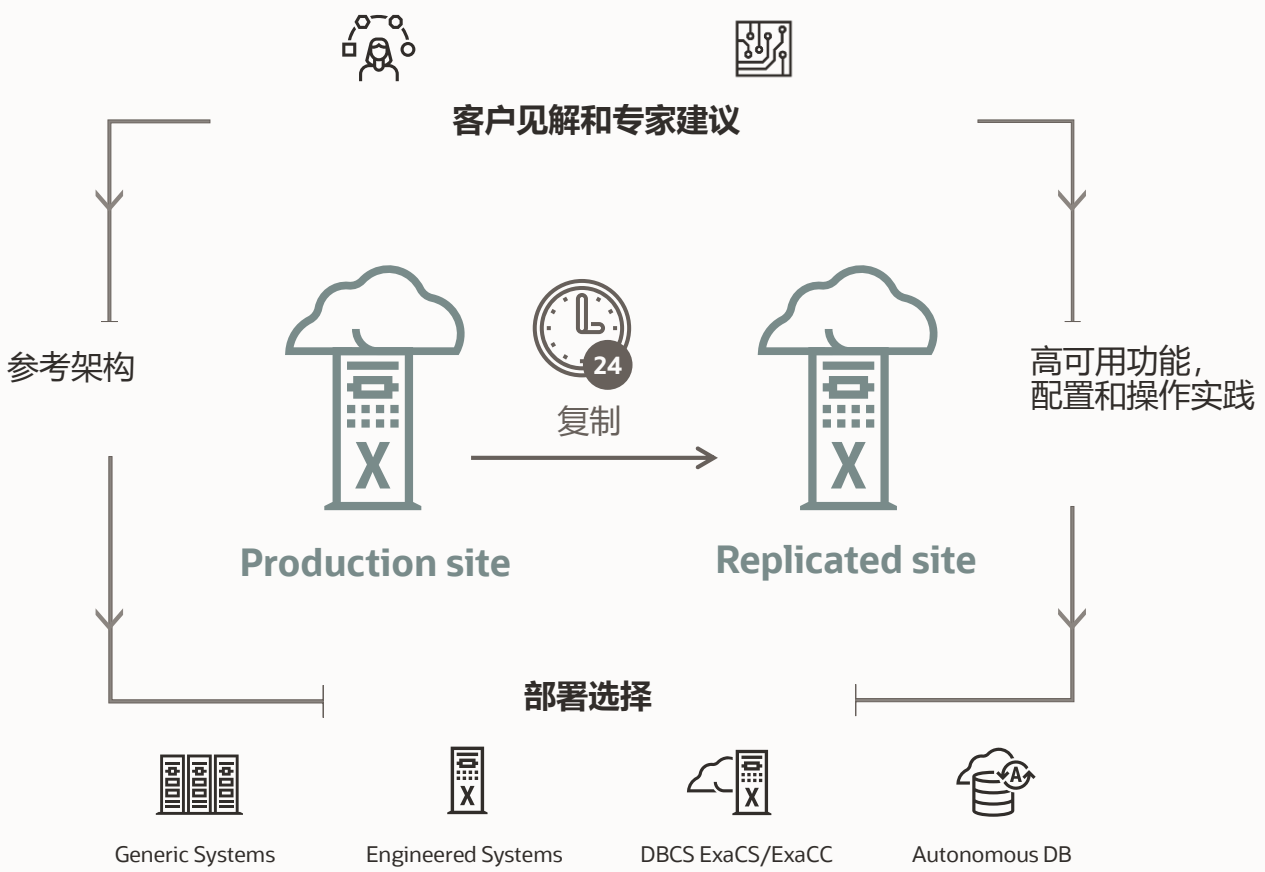
张 灿
资深解决方案工程师

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.

Oracle 高可用架构(MAA)

铂金
黄金
白银
青铜



Oracle 多租户

Continuous availability
- Application Continuity
- Global data services

Data protection
- Flashback
- RMAN + ZDLRA

Active replication
- Active Data Guard
- Golden Gate

Scale out
- RAC
- ASM
- Sharding

MAA参考架构
高可用服务级别

青铜

开发, 测试, 生产

单实例数据库

可重启

备份/还原

白银

生产

青铜 +

Database HA with RAC

Application continuity

黄金

Business critical

白银 +

DB replication with active data guard

铂金

Mission critical

黄金 +

Golden Gate

Edition based redefinition

All tiers exist with on-premises and cloud. However, platinum currently must be configured manually while bronze to gold are covered with cloud tool automation

什么是持续可用（Continuous Availability）？

持续可用着重于用户体验

- 计划内维护，超时，计划外停机被隐藏
- 正在进行的工作会继续进行
- 性能是可预测的
- 用户不受影响！

如今很容易实现连续可用

Fast Application Notification (FAN)

The dead thing cannot tell you it is dead

- ✓ 当数据库、节点和网络发生停机的时候，FAN提供对客户端的立即通知响应。
- ✓ FAN对在故障时解决客户端TCP/IP超时问题至关重要。
- ✓ FAN启动数据库会话的排空，因此客户端在计划维护期间不会出现中断。当资源再次可用时，FAN会立即通知客户端。

All Oracle uses ONS
JDBC Universal Connection Pool
OCI/OCCI driver
ODP.NET Unmanaged Provider (OCI)
ODP.NET Managed Provider (C#)
OCI Session Pool
WebLogic Active GridLink
Tuxedo
JDBC Thin Driver **(new 12.2)**
CMAN in Traffic Manager mode (new 18c)

Auto-Configured

```
DESCRIPTION =
  (CONNECT_TIMEOUT=90)
      (RETRY_COUNT=20)(RETRY_DELAY=3)
      (TRANSPORT_CONNECT_TIMEOUT=3)
  (ADDRESS_LIST =
          (LOAD_BALANCE=on)
          ( ADDRESS = (PROTOCOL = TCP)
          (HOST=primary-scan) (PORT=1521)))
      (ADDRESS_LIST =
          (LOAD_BALANCE=on)
          ( ADDRESS = (PROTOCOL = TCP) (HOST=
          scan)(PORT=1521)))
(CONNECT_DATA=(SERVICE_NAME=gold)))
```

ONS Node Set 1

ONS Node Set 2

新概念：请求界限（Request boundaries）

描绘工作单位

Oracle Pools 12c+, Java Standard (JDK9+), Transparent Application Continuity (TAC)

```
PoolDataSource pds = GetPoolDataSource();
Connection conn = pds.getConnection();
PreparedStatement pstmt = …

        …
        SQL, PL/SQL, local calls,
        …

conn.commit();
conn.close();
```

我们如何使用请求?

- 排空（Draining）
- 负载均衡（Load Balancing）
- 恢复进行时工作
- 计划的故障转移
- 重置会话状态
- 性能指标
- 防护等级

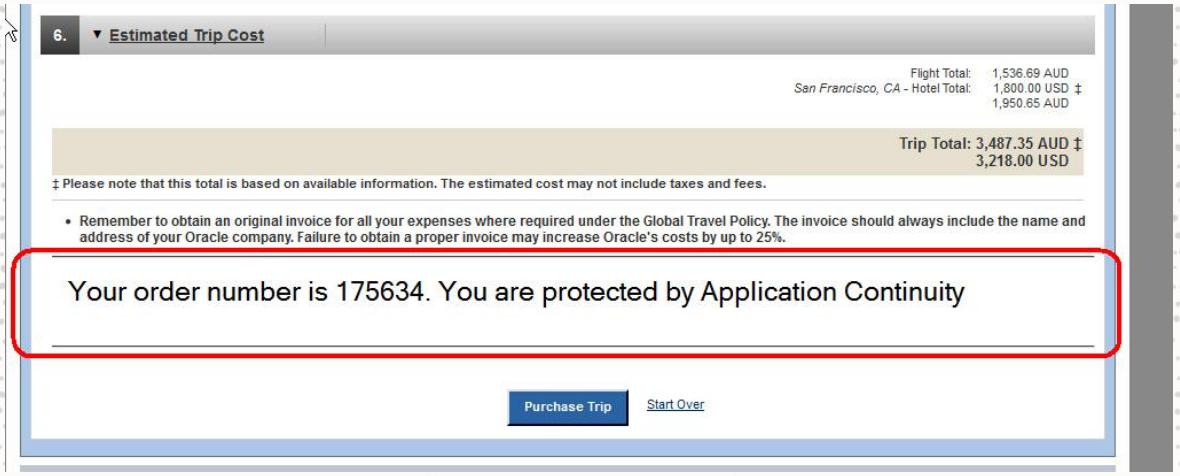
开始请求

请求正文通常以COMMIT结尾

结束请求

Application continuity

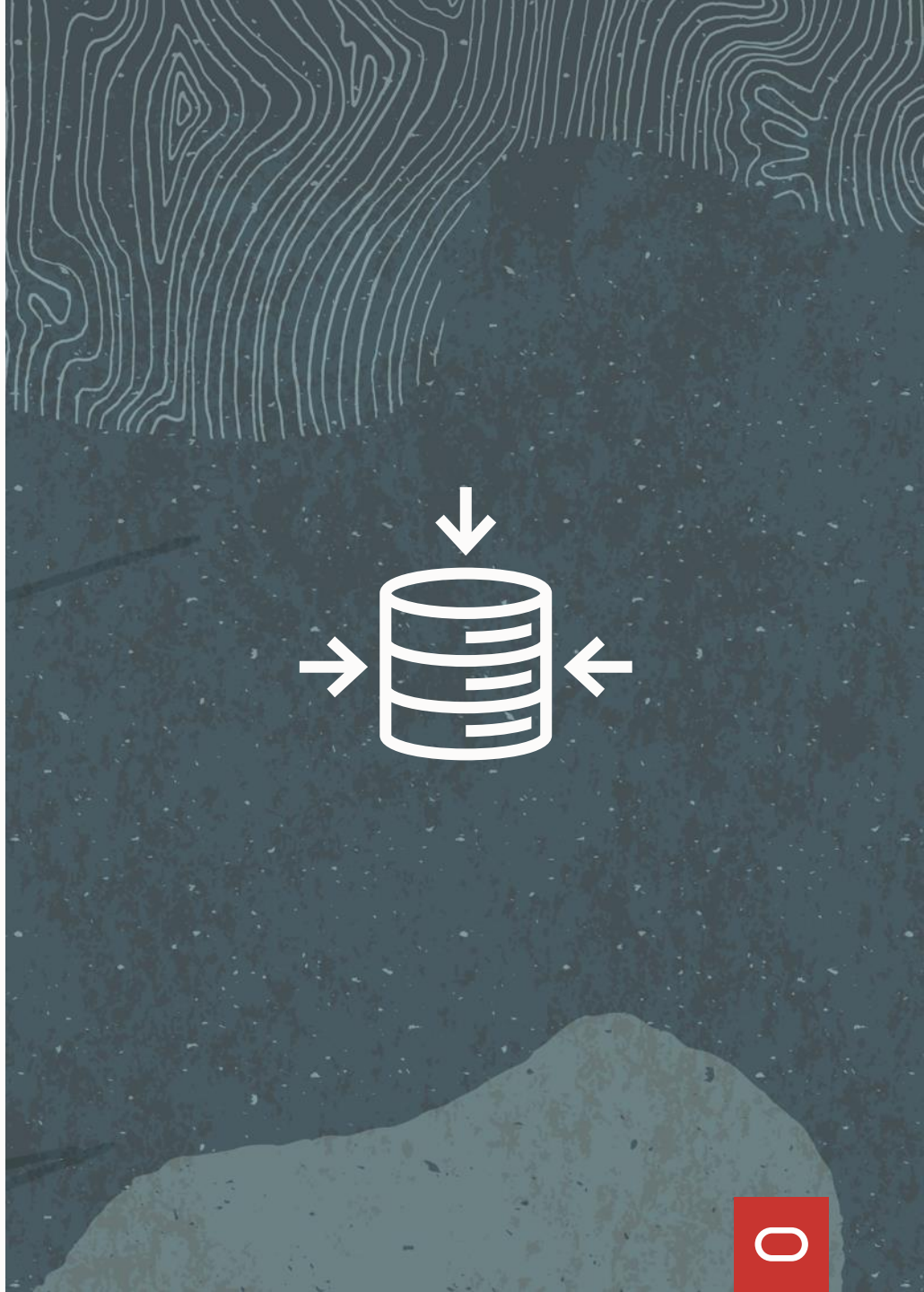
- 发生错误时可重放正在进行中的可恢复的工作
- 屏蔽硬件，软件，网络，存储，会话错误和超时
- 12.1 JDBC-Thin, UCP, WebLogic Server, 3rd Party Java application servers
- 12.2 OCI, ODP.NET unmanaged, JDBC Thin on XA data source
- Transparent Application Continuity (19c)
- TAC 在ADB Dedicated上为默认配置



计划内维护

维护前的排空工作

- 数据库服务已重定位（relocate）
- 自动的, 用户的工作排空
- 很多阶段可以排空:
  - 归还连接给连接池
  - 连接测试
  - 注销
  - 计划的故障转移

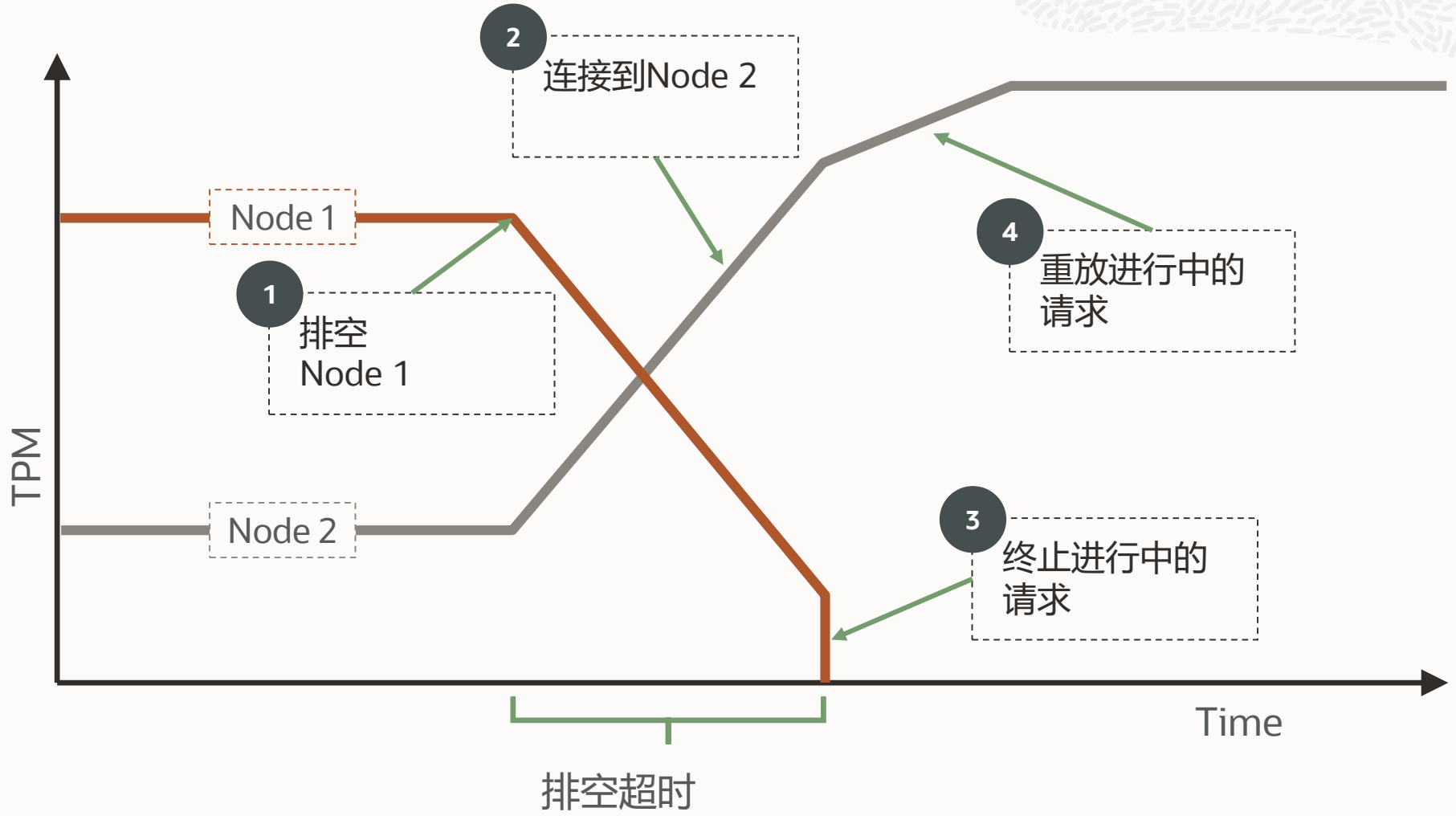


有哪些先决条件?

1. **设计良好的应用程序**
2. **启用的客户端堆栈**(driver, pool, or app server)
   - Oracle或启用的第三方堆栈
   - 正确配置网络连接
3. **正确配置的数据库和集群架构**
   - 根据MAA最佳实践进行配置
4. **停机期间可以满足业务SLA**
   - 恢复时间性能目标
   - 维护窗口

您的部署必须满足这些要求!

排空... 连接... 故障转移



DBA commands: Group commands (12.2 and later)

指定数据库/节点/PDB，迁移所有服务

- srvctl **relocate service -database -instance** -drain_timeout.. –stopoption [immediate|transactional]
- srvctl **relocate service -node** . -drain_timeout.. –stopoption
  srvctl **stop service –pdb** . -drain_timeout.. –stopoption

停止单个节点的服务或数据库实例

- srvctl **stop service -node <node_name>** -drain_timeout.. –stopoption
  srvctl **stop database -node <node_name>** -drain_timeout.. –stopoption

Data Guard 切换

- **Switchover** to <db_resource_name> **[wait [xx]];**

Use Oracle Pools with FAN

Full Lifecycle - Drain and Rebalance

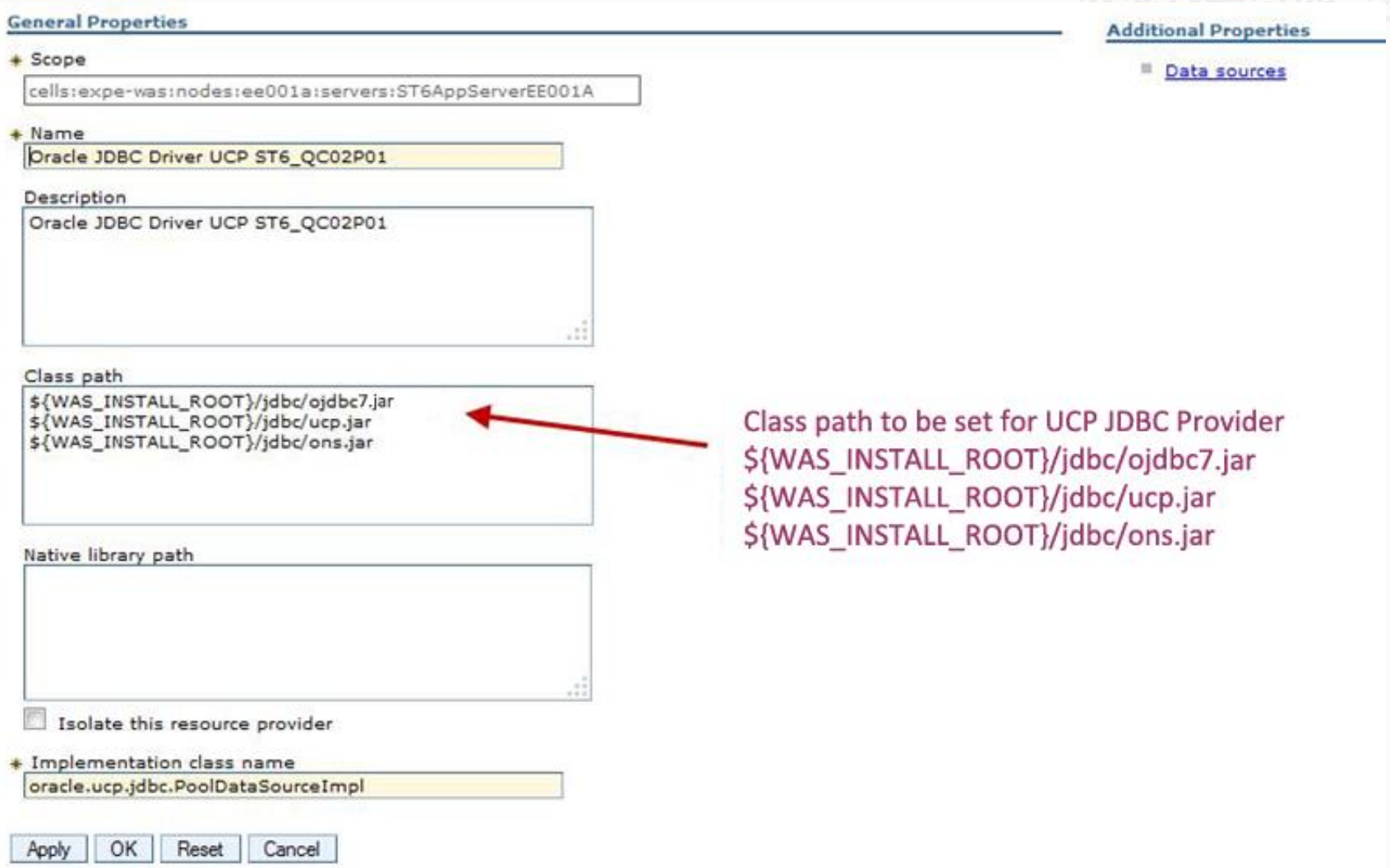
应用程序使用…	**Oracle – WebLogic Active GridLink, UCP, ODP.NET managed and unmanaged, OCI Session Pool, Tuxedo** 使用UCP的第三方App Server: IBM WebSphere and Liberty, Apache Tomcat, NEC WebOTX, Red Hat WildFly (JBoss), Spring
FPP or srvctl …	**srvctl [relocate|stop] service -drain_timeout**
排空会话	Immediately new work is redirected Gradually idle are drained **Active sessions are released when returned to pools**

FAN Planned

Drain & rebalance

UCP与其他基于Java的应用程序服务器

A simple data source replacement



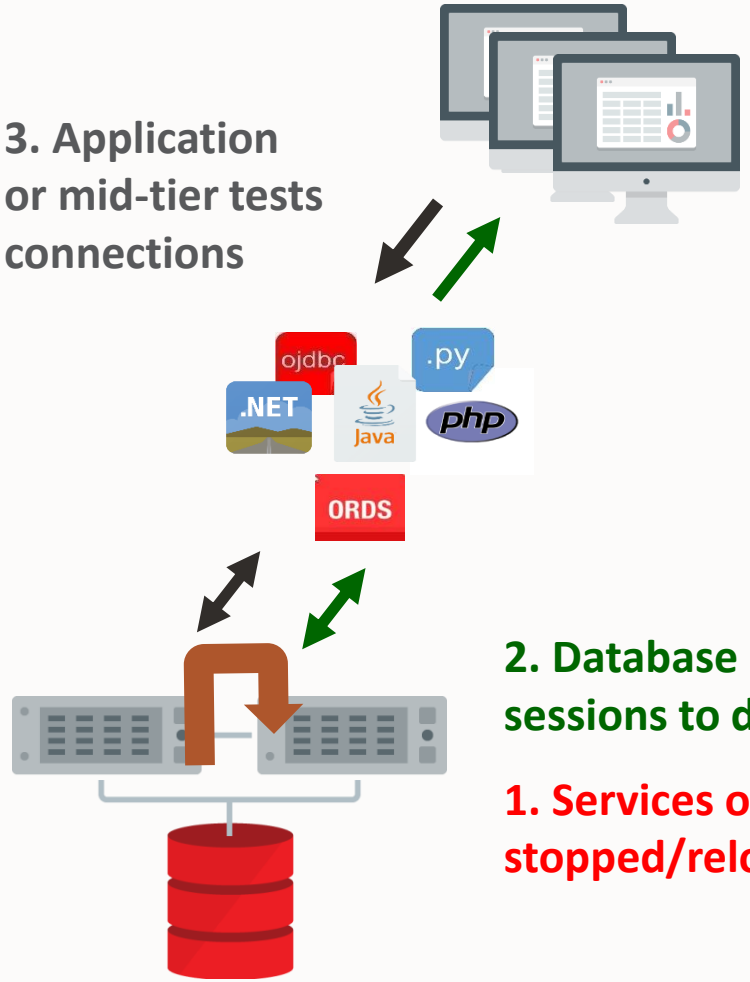
- IBM WebSphere
- IBM Liberty
- Apache Tomcat
- NEC WebOTX
- Red Hat WildFly (JBoss)
- Hibernate
- Spring
- Your own

Tip: Excludes TP Managed 2PC XA

由Oracle Database 来排空

Connection Tests, All Drivers

NEW IN 19c



Plan B

- 应用程序“测试”连接
- 数据库响应连接错误
- 新的工作在另一个新连接上继续

Tip: enable more with DBMS_APP_CONT_ADMIN view in DBA_CONNECTION_TESTS

DBA_CONNECTION_TESTS

Database Side ONLY. CDB, PDB, Service Levels

```
select connection_test_type, sql_connection_test, enabled from
dba_connection_tests;
```

CONNECTION_TEST	SQL_CONNECTION_TEST	ENABLED
SQL_TEST	SELECT 1 FROM DUAL	Y
SQL_TEST	SELECT COUNT(*) FROM DUAL	Y
SQL_TEST	SELECT 1	Y
SQL_TEST	BEGIN NULL;END	Y
PING_TEST	NA	N
ENDREQUEST_TEST	NA	N

部署您的应用程序

Easy steps to success

1. **配置应用程序 –（TAC 无代码更改 ）**
   - 使用推荐的连接串
   - 使用最新的客户端版本
2. **为您的工作负载配置数据库端服务(DBA)**
   - 为特定工作负载创建数据库服务
   - 确保连接字符串使用此数据库服务
3. **测试和评估重放保护**
   - 检查AWR以获取保护统计信息
   - 使用ACCHK实用程序识别重放异常

应用程序开发人员和所有者，以及DBA。

应用程序部署说明
FAN Configuration

Fast Application Notification (FAN) 是排空，负载均衡等操作的先决解件。

对于诸如节点和网络故障之类的中断，如果客户端没有被FAN中断，则应用程序不会快速故障转移。

配置FAN时，请使用ONS的自动配置。 严格使用推荐的数据库链接串格式。

FAN for JDBC

12c or later JDBC clients and database

- Use recommended URL for auto-ons
- ons.jar (plus Wallet jars, osdt_cert.jar, osdt_core.jar, oraclepki.jar) on the CLASSPATH
- UCP pool property **fastConnectionFailoverEnabled=true**
- Third party pools, UCP is recommended
- Open port 6200 for ONS

Before 12c clients

- Set oracle.ons.nodes =mysun05:6200,mysun06:6200, mysun07:6200,mysun08:6200

FAN for OCI

12c or later OCI clients and database

- ATP-D presets aq_ha_notifications
- Use Recommended TNS for auto-ons
- AutoConfig, Events (plus Wallet) in oraaccess.xml
- Link the app with client thread O/S library
- Open port 6200 for ONS

```
<oraaccess>
<default_parameters>
<events>true</events>
 <ons>
<auto_config>true</auto_config>
<wallet_location>/my_host_path/
wallets/</wallet_location>
</ons>
</default_parameters>
</oraaccess>
```

Before 12c OCI clients

- Enable OCI_EVENTS at environment creation OCIEnvCreate(..)
- Link the app with the client thread o/s library.

FAN for ODP.NET

1. ATP-D presets aq_ha_notifications
2. Use Recommended TNS for auto-ons
3. AutoConfig, Events (plus Wallet) in oraaccess.xml
4. Open port 6200 for ONS
5. FAN, in the connection string -

- **"user id=oracle; password=oracle; data source=HA; pooling=true; HA events=true;“**

6. Runtime Load Balancing, also in the connection string -

- **"user id=oracle; password=oracle; data source=HA; pooling=true; HA events=true; load balancing=true;“**

```
<default_parameters>
<ons>
   <auto_config>true</auto_config>
   <debug>true</debug>
   <wallet_location>C:\ADE\jesu_win18\oracle\
work</wallet_location>
</ons>
</default_parameters>
</oraaccess>
```

应用程序部署说明
维护之前先排空会话

在开始维护之前，请从要维护的实例或节点上排空会话。

启用FAN。

带有FAN的Oracle连接池是最好的解决方案，因为连接池提供了移动会话的整个生命周期。排空会话和负载均衡已集成到了运维流程中。

事务结束后，将您的连接归还到连接池中。

如果您的连接测试不是标准测试，请使用DBMS_APP_CONT_ADMIN将连接测试添加到数据库中。在DRAIN_TIMEOUT内未排空的会话将进行故障转移。

应用程序部署说明
故障转移解决方案

	TAC	AC	TAF	Draining
I don't know how the application is implemented	Yes	No	No	Yes
My application does transactions	Yes	Yes	No for unplanned Transactional disconnect only	Yes
My application uses Oracle state (temp lobs, PL/SQL, temp tables.)	Yes	Yes No for static mode	No	Yes
My application does not use connection pools	Yes	No	Yes	Yes
My application has side effects (such as file transfers)	Yes Side effects are not replayed	Customizable	No	Yes
My app needs Initial State Restored	Yes and custom	Yes and custom	Yes and custom	Yes
Future proofed for application changes	Yes	No	No	Yes

会话故障转移的标准解决方案是Transparent Application Continuity（TAC）。

如果您使用的是Oracle Database 12c第2版，或者要在事务中使用回调进行自定义，需要使用Application Continuity（AC）。

如果您的应用程序是只读的，并且在初始设置后未更改会话中的Oracle会话状态，请使用Transparent Application Failover （TAF）。

应用程序部署说明
使用服务

使用非默认数据库服务（默认服务与数据库或PDB具有相同的名称）。 您创建的服务提供了位置透明性和HA功能。

For example:

Transparent Application Continuity

$ srvctl add service -db mydb -service GOLD -preferred inst1,inst2 - failover_restore **AUTO** -failoverretry 1 -failoverdelay 3 - commit_outcome **TRUE** -failovertype **AUTO** -replay_init_time 600 -retention 86400 -notification TRUE -drain_timeout 300 - stopoption IMMEDIATE

Application Continuity

$ srvctl add service -db mydb -service SILVER -preferred inst1 -available inst2 -failover_restore **LEVEL1** -failoverretry 1 -failoverdelay 3 - commit_outcome **TRUE** -failovertype **TRANSACTION** -replay_init_time 1800 - retention 86400 -notification TRUE - drain_timeout 300 -stopoption IMMEDIATE

Transparent Application Failover

$ srvctl add service -db mydb -service BRONZE -preferred inst1 -available inst2 -failover_restore **LEVEL1** -failoverretry 1 - failoverdelay 3 - commit_outcome TRUE -failovertype **SELECT** -retention 86400 -notification TRUE -drain_timeout 300 -stopoption IMMEDIATE

To add with the Data Guard role, here is the TAC example:

$ srvctl add service -db mydb -service GOLD -preferred inst1 -available inst2 -failover_restore AUTO -failoverretry 1 -failoverdelay 3 - commit_outcome TRUE -failovertype AUTO -replay_init_time 600 -retention 86400 -notification TRUE -role PHYSICAL_STANDBY - drain_timeout 300 - stopoption IMMEDIATE

应用程序部署说明
服务例子

Service Name	Description	Draining	FAN	TAC
TPURGENT	OLTP Highest Priority	Yes	Yes	Yes
TP	OLTP General Priority Recommended to be used as main service	Yes	Yes	Yes
HIGH	Reporting or Batch Highest Priority	Yes	Yes	
MEDIUM	Reporting or Batch	Yes	Yes	

应用程序部署说明
使用推荐的连接串

将推荐的连接串与内置的超时，重试和延迟一起使用，以便传入的连接在中断期间不会看到错误。

Use this Connection String for ALL Oracle driver version 12.2 or higher:

```
Alias (or URL) =
  (DESCRIPTION =
   (CONNECT_TIMEOUT= 90)(RETRY_COUNT=50)(RETRY_DELAY=3)(TRANSPORT_CONNECT_TIMEOUT=3)
   (ADDRESS_LIST =
     (LOAD_BALANCE=on)
     (ADDRESS = (PROTOCOL = TCP)(HOST=primary-scan)(PORT=1521)))
   (ADDRESS_LIST =
     (LOAD_BALANCE=on)
     (ADDRESS = (PROTOCOL = TCP)(HOST=standby-scan)(PORT=1521)))
  (CONNECT_DATA=(SERVICE_NAME = YOUR SERVICE)))
```

Always Know Your Protection Level

AWR, system, session, service statistics

Your application is fully protected when

cumulative user calls in request = cumulative user calls protected

Statistic	Total	per Second	per Trans
cumulative begin requests	1,500,000	14,192.9	2.4
cumulative end requests	1,500,000	14,192.9	2.4
cumulative user calls in request	6,672,566	63,135.2	10.8
cumulative user calls protected	6,672,566	63,135.2	10.8

Always Know Your Protection Level

```
SELECT a.name,
a.value as "Total",
b.name, b.value as "AC calls",
(b.value/NULLIF(a.value,0) * 100) as "Protected%"
FROM v$sysstat a, v$sysstat b
WHERE a.name = 'cumulative user calls in requests'
and b.name =
'cumulative user calls protected by Application Continuity'
```

NAME	TOTAL	NAME	AC calls	Protected%
cumulative user calls in requests	1247074	cumulative user calls protected by Application Continuity	1246395	99.9455525

Killing Sessions
Extended

DBA Command	Replays
srvctl stop service -db orcl -instance orcl2 -force	**YES**
srvctl stop service -db orcl -node rws3 -force	**YES**
srvctl stop service -db orcl -instance orcl2 **–noreplay** -force	
srvctl stop service -db orcl -node rws3 **–noreplay** -force	
alter system kill session … immediate	**YES**
alter system kill session … **noreplay**	
dbms_service.disconnect_session([service], dbms_service. **noreplay**)	

重放 – 我可以自己做吗?

我可以重新连接吗?

- 自动的?
- 用户会注意到还是必须采取措施?
- 我可以重建会话状态吗?
- 错误发生的地点或时间?

我应该重放吗?

- 可以或应该重放什么?
- 哪些错误或异常仍应报告给用户?
- 数据库升级会改变这种逻辑吗?
- 这种逻辑有多复杂?

我可以重放吗?

- 我是否捕获了足够的信息以重播整个数据库请求?
- 我的事务在什么时候失败的，在提交或回滚之前或之后?
- 如何使我的重放返回完全相同的结果?
- 如果重放失败怎么办?

其他资源

Application Continuity product page -> https://www.oracle.com/goto/ac

Technical briefs (how to/explanations/details) ->

Deployment Checklist (for developers, dba's and application owners)

- https://www.oracle.com/technetwork/database/clustering/checklist-ac-6676160.pdf

Fast Application Notification (FAN) – everything anyone would want to know about FAN

- https://www.oracle.com/technetwork/database/options/clustering/applicationcontinuity/learnmore/fastapplicationnotification12c-2538999.pdf


ORACLE