



金融行业运维实践-上海站



Oracle12cR2数据库资源池实践

杨欣捷

- 4年宇宙行DBA运维经验
- Oracle 11G OCM
- 浦发银行总行架构处架构师
- 目前主要负责：
 - 传统数据库架构体系优化（资源池）
 - 开源和国产数据库引入
 - 大数据和机器学习
 - 数据容灾
 -



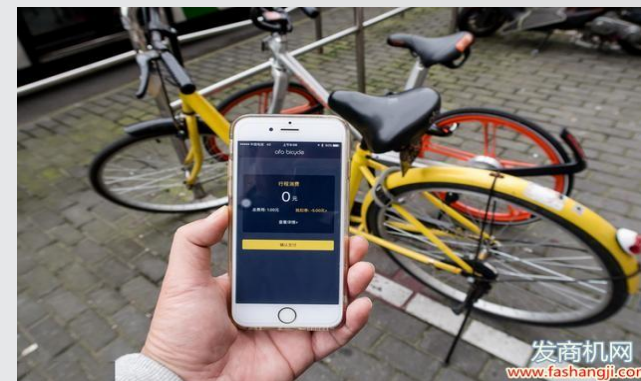
什么是资源池



- 共享



- 标准化



- 按需、快速获取



- BaaS (Bike as a Service)

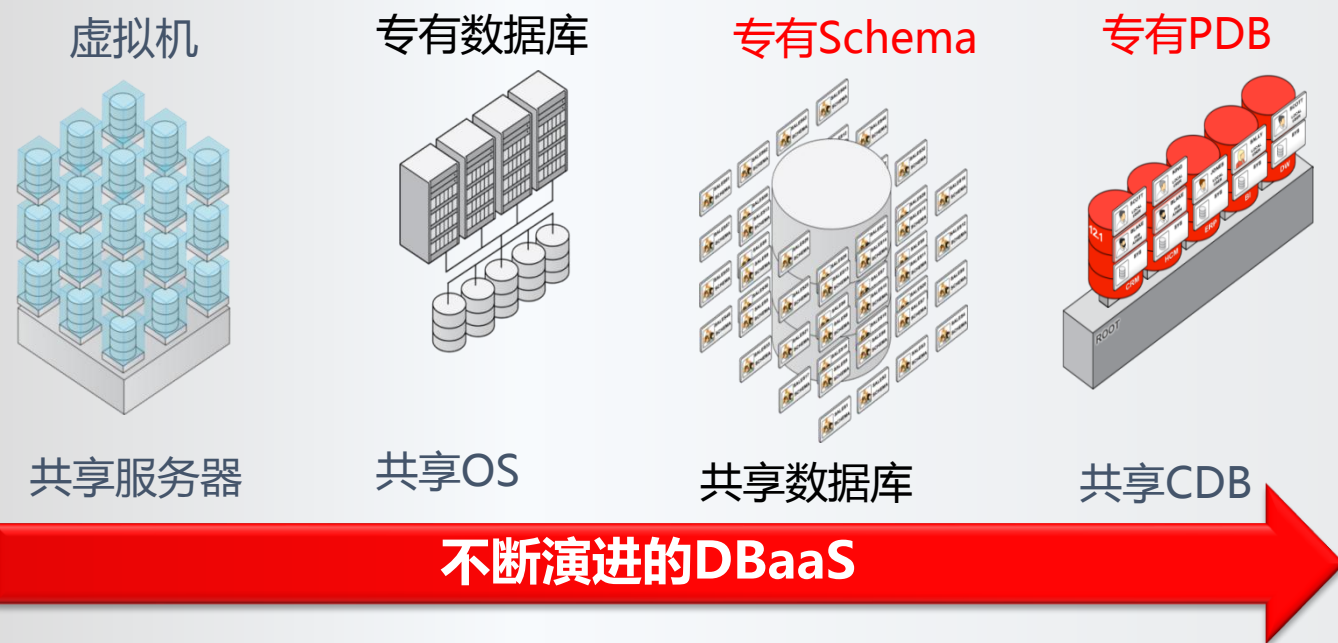
对于一家提供Oracle支持服务的公司：

- 给工程师配私家车：竖井是部署
- 和租车公司合作：购买租车公司的IaaS
- 报销出租车费：购买了出租车公司的PaaS
- 用别家公司工程师去现场：购买别家公司的SaaS

对城市建设者：

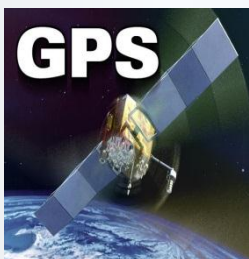
- 拼车软件=软件定义出行

- 为什么要基于Oracle12c构建数据库资源池？



- 虚拟机&共享OS：整合度太低、无横向扩展能力
- Schema方式：隔离性有限，应用相关、无原生的快速发布接口，自服务能力差
- Oracle12c
 - 命名空间隔离
 - 权限控制
 - 更好的资源隔离
 - 灵活的service
 -

- 为什么共享单车2016年火了起来？



全面DBaaS的时机已经成熟

DBaaS

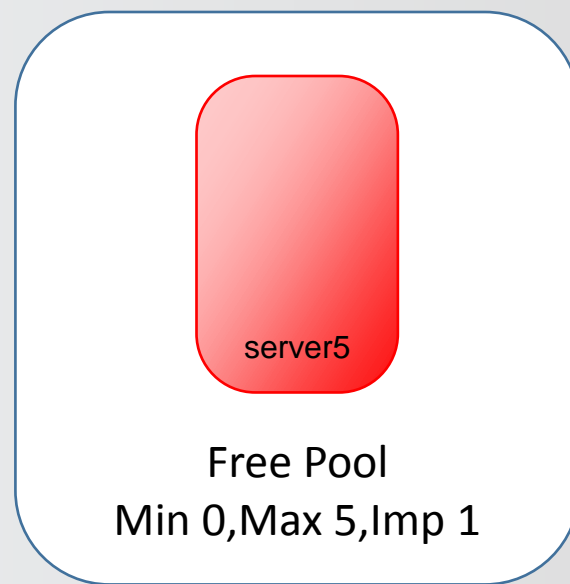
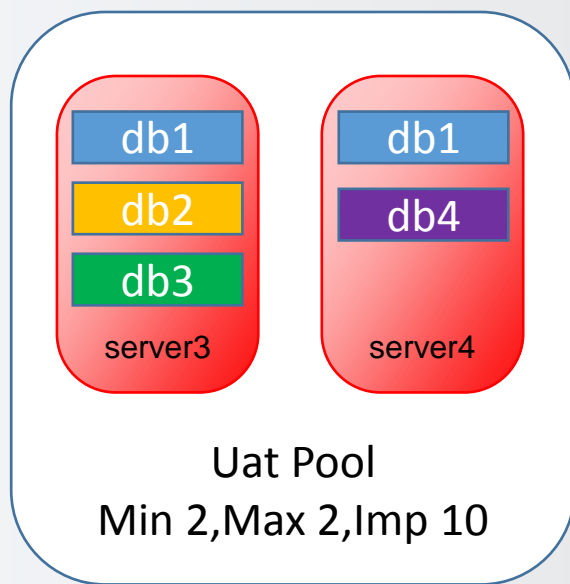
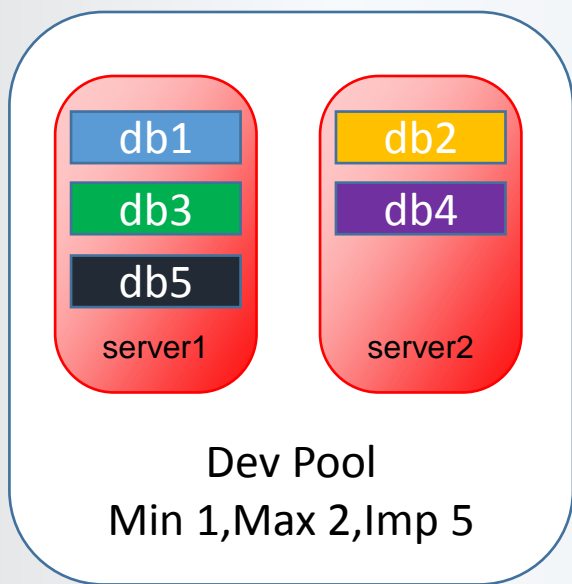
- 时机已经成熟，合理的规划及技术使用，是资源池成败的关键



- 开发环境的特点：
 - 数据库数量多
 - 性能需求小
 - 申请回收频繁
 - 自服务需求高
- 11gR2的年代，已经以资源池方式运作
 - 使用技术：RAC OneNode、Policy-managed database

自服务门户

(DB name, 字符集, 用户信息, 表空间需求, 环境类型, 是否RAC)



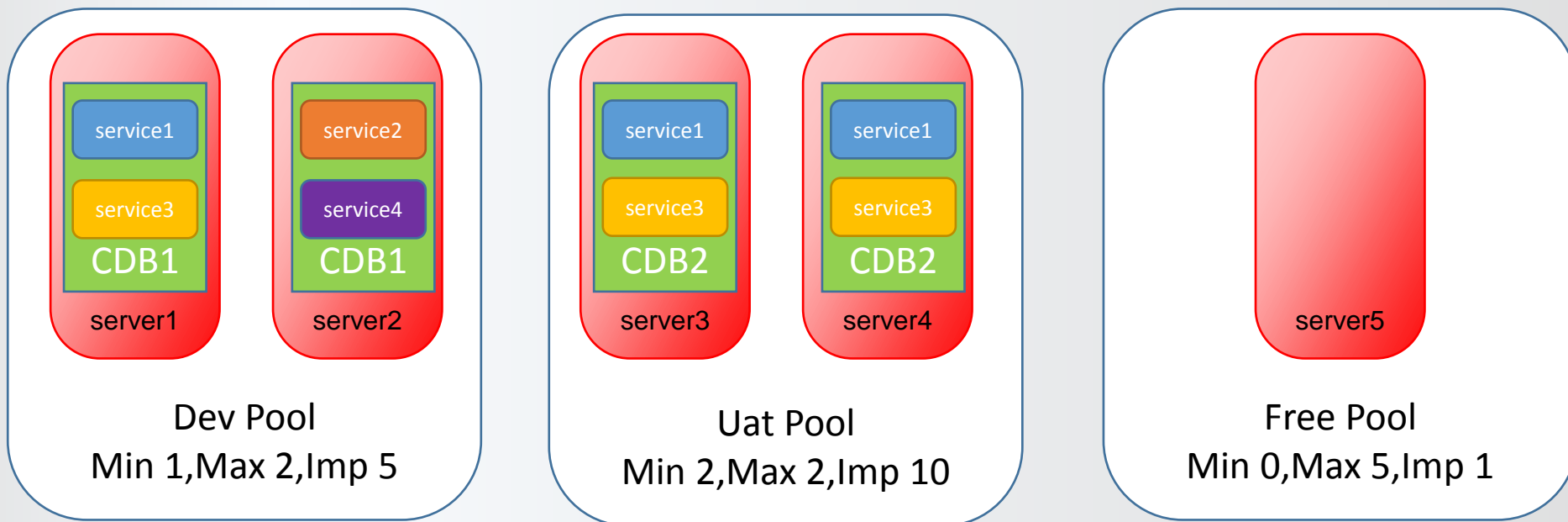
Oracle GI+ASM

缺点: DBA角色不能放出, 否则磁盘使用不可控。内存不共享, 整合度不高。

新建数据库服务

```
dbca -silent -createDatabase -templateName General_Purpose.dbc \  
-gdbname testdb -sid testdb -responseFile NO_VALUE \  
-RACOneNode -RACOneNodeServiceName test1 \  
-characterSet AL32UTF8 -policyManaged -serverPoolName UATPOOL \  
-sysPassword oracle -systemPassword oracle \  
-emConfiguration NONE -dbsnmpPassword oracle \  
-sysmanPassword oracle -storageType ASM \  
-asmsnmpPassword oracle -diskGroupName DATA \  
-sampleSchema false \  
-redoLogFileSize 300 \  
-automaticMemoryManagement false \  
-initParams open_cursors=500 -totalMemory 1024 \  
-databaseType OLTP
```

自服务门户
(PDB name, 服务名, 服务类型)



Oracle GI+ASM

改进: PDB_DBA角色放出可以解放大量管理工作量, 单PDB磁盘使用量可控。整合程度更高, 数据库发布更快

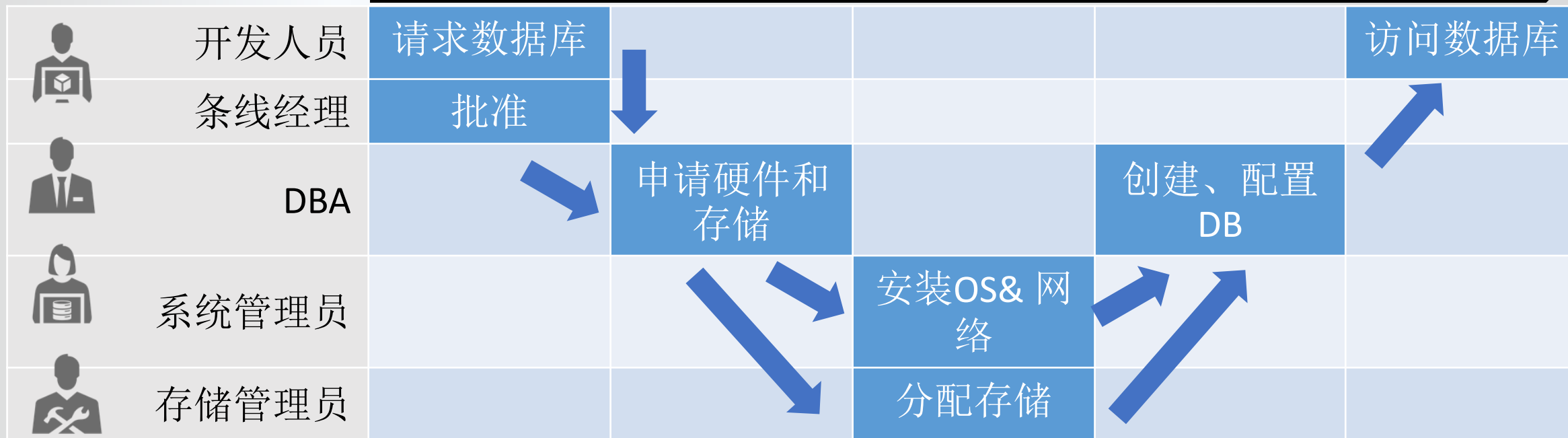
创建PDB:

```
CREATE PLUGGABLE DATABASE pdb1  
  ADMIN USER pdbadm IDENTIFIED BY oracle  
  ROLES = (dba)  
  DEFAULT TABLESPACE users datafile SIZE 1g AUTOEXTEND ON  
  STORAGE (MAXSIZE 50G);
```

创建service:

```
Srvctl add service -s pdv1_srv -d cdb1 -pdb pdb1 -cardinality  
singleton
```

天到周。。。



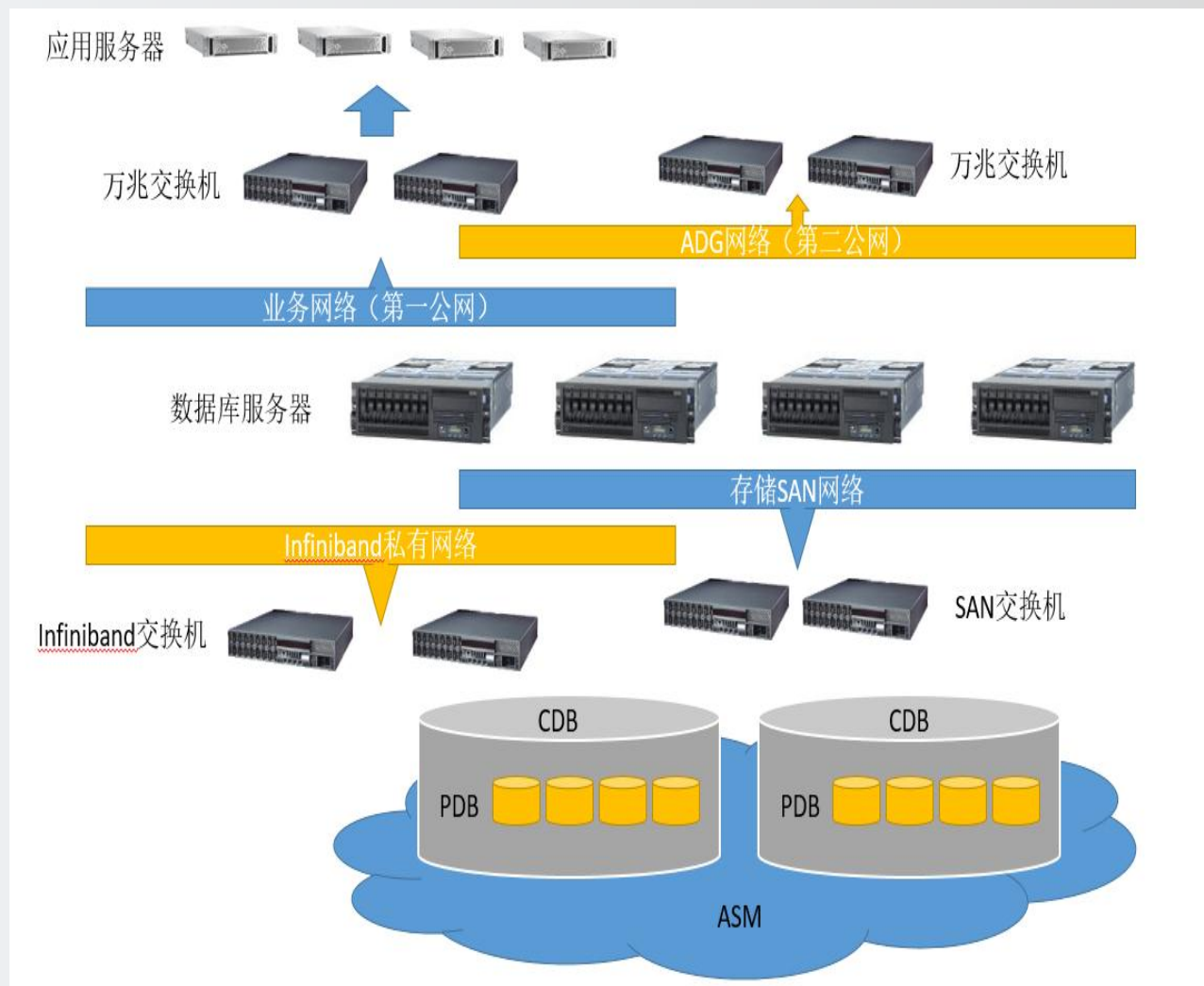
	非资源池	资源池
交付时间	3小时	5分钟
维护工作量	大	小
存储	30个数据库，约9TB。	3TB
资源共享程度	低	高
标准化程度	低	高
备份	无备份。数据丢失只能重新造数据，影响进度。	有备份，可以及时恢复数据。
监控	项目组发现问题，请求技术支持。	DBA主动发现问题

- 生产环境的特点：
 - 数据库数量比开发少
 - 申请回收不频繁
 - 专职DBA精细化管理
 - 尚无自服务需求
 - 性能要求高
 - 资源隔离要求高
 - 安全性要求高
 - 可扩展性要求高
 - 不容有失

需求不同于开发环境
架构不同于开发环境

- 单集群不超过4节点
- X86架构
- Infiniband (EDR 100GB) 实现私网
- 组建第二公网用于ADG异地容灾
- 配置NVme SSD卡用于DB二级缓存
- 操作系统使用OEL 7.3
- 4块16GB HBA卡

没有存储节点的未来版Exadata



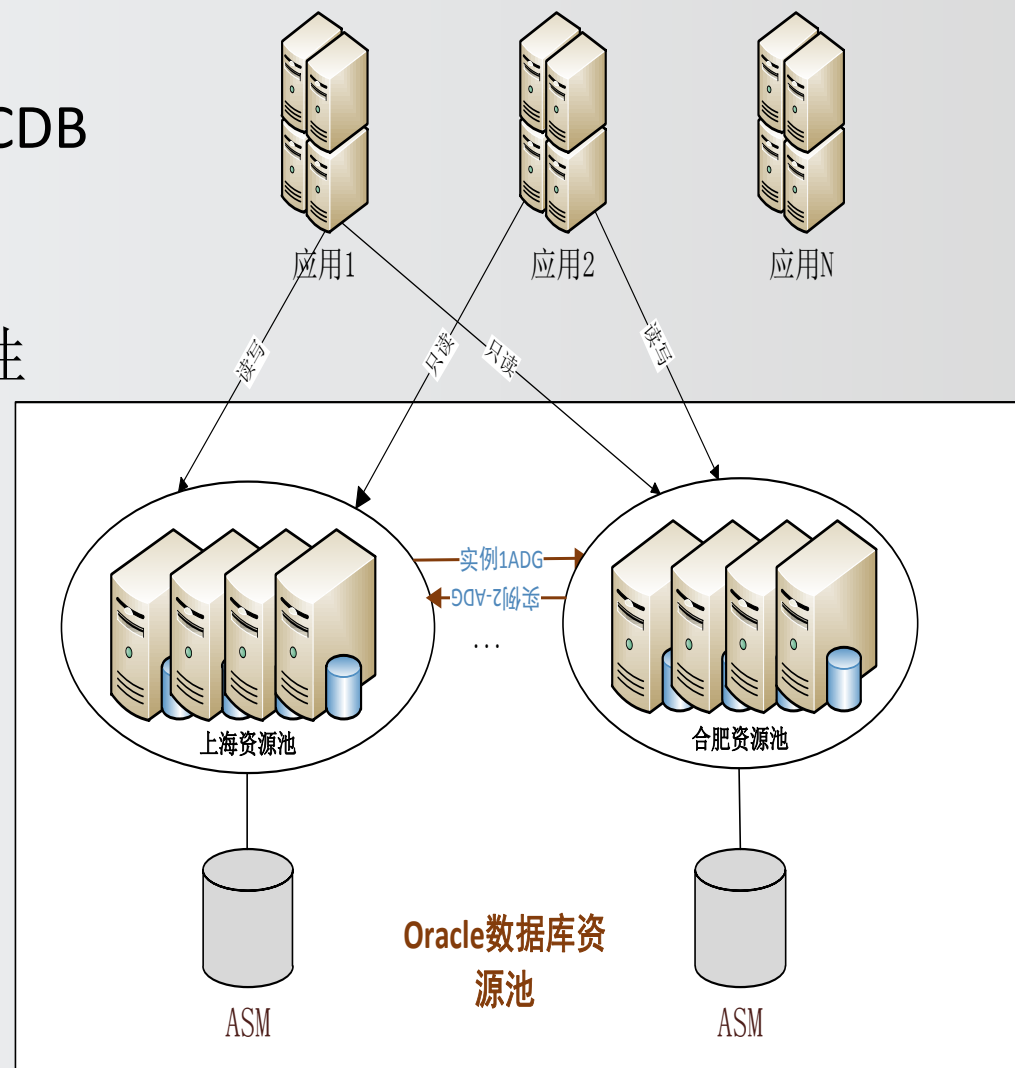
	第一路公网-业务网段		第二路公网-ADG网段	
节点1	业务网段IP	4个 业务网段VIP	ADG网段IP	4个 ADG网段VIP
节点2	业务网段IP		ADG网段IP	
节点3	业务网段IP		ADG网段IP	
节点4	业务网段IP		ADG网段IP	
SCANIP	1个或3个业务网段IP		不提供	

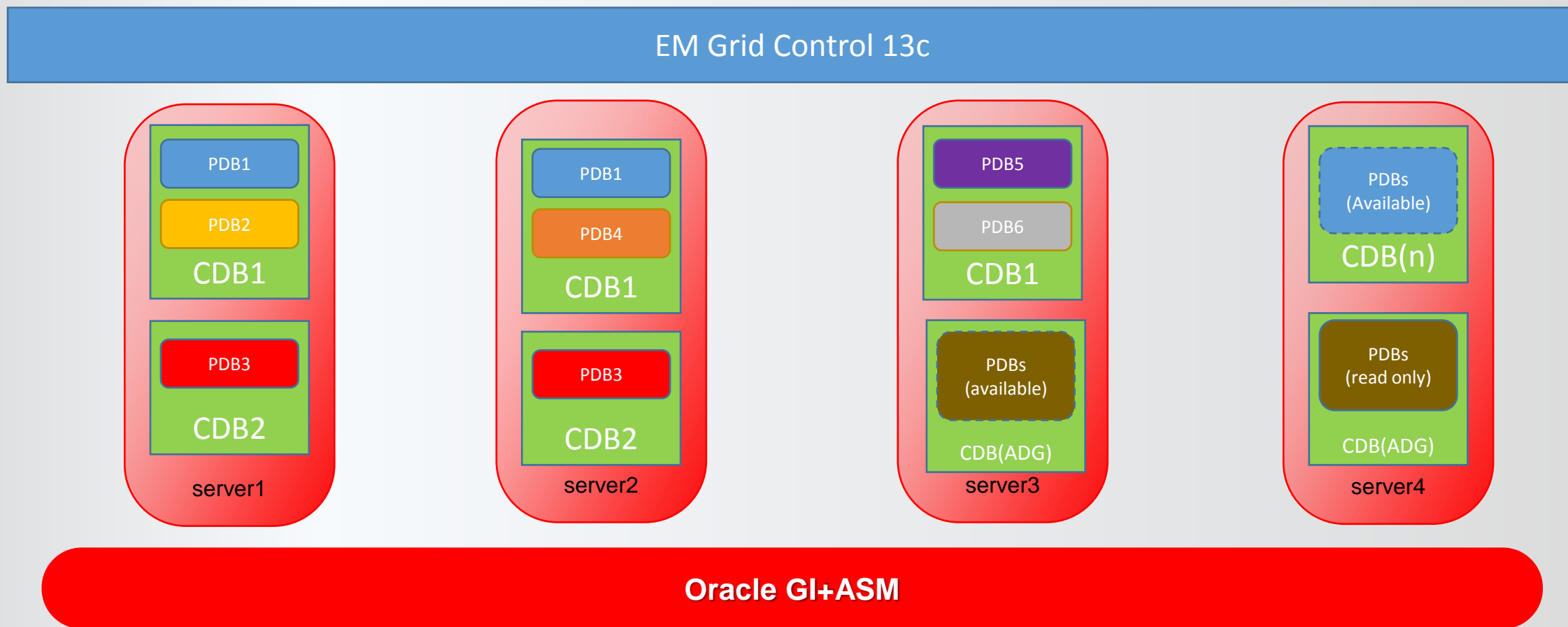
- 集群私网
 - 2个Infiniband交换机实现内联
 - 目前使用IPoIB，未来使用RDS传输协议。
 - HAIP实现网络冗余
- 对外网络
 - 三个SCANIP供DNS轮询。应用通过DNS+服务名方式访问数据库
- 日志传输网络
 - ADG网络为第二公网
 - 配置ADG专用的VIP地址和监听器
 - 使用VIP保障ADG传输在节点故障时的连续性

- ASM规划
 - 使用ASMFD管理磁盘，防止误操作损坏数据
 - 将MGMT库从默认的OCR所在磁盘组中迁出，独立磁盘组
 - 每个CDB使用独立的一组DATADG、REDODG和ARCHDG
- 权限控制
 - 使用SYS/SYSTEM用户管理CDB和PDB数据库。
 - 不提供PDB级别的DBA管理账户，涉及到数据库管理方面操作，统一由CDB DBA操作
- 使用EM 13c管理CDB和PDB

- ADG的保护单位是CDB
- 双中心的资源池分别有作为主库和ADG库的CDB
- 根据容灾一致性切换组来分配CDB
- 应用通过DNS+ServiceName访问数据库
- Service Name对应用屏蔽了数据库层的复杂性
- 简化灾备切换

资源池的角度没有“主备”之分





数据库服务的Prefer节点为1, 2, 3. 节点4为各服务的available节点。同时节点4还是异地主库的ADG备节点, 节点3是ADG备节点的available节点。这样最大程度上共享资源, 同时在节点故障时仍然尽可能保障计算能力。

主要技术	Prior To 12cR2	12cR2	使用	备注
Service affinity	✓	✓	广泛	最基本的隔离方式，通过service让不同应用优先连接不同节点。12c的PDB的open也可以跟着service。
Instance Caging	✓	✓	广泛	12cR2加强，可在PDB层面设置CPU_COUNT
Resource Manager	✓	✓	选择	可以在CDB内控制不同的PDB资源使用，也可以在PDB内部细粒度控制
Memory Control		✓	尝试	PDB级别设置更细粒度的内存控制例如设置参数"SGA_MIN_SIZE"
IO Control		✓	观望	12cR2加强,CDB/PDB级别设置参数"MAX_IOPS"/"MAX_MBPS"

性能目录

编号	应用类型	等级	TPCC支持范围
OTP01	联机交易为主	顶配	大于400万
OTP02	联机交易为主	高配	300万-400万
OTP03	联机交易为主	次高配	200万-300万
OTP04	联机交易为主	中配	100万-200万
OTP05	联机交易为主	低配	小于100万
OAP01	分析批量为主	N/A	N/A

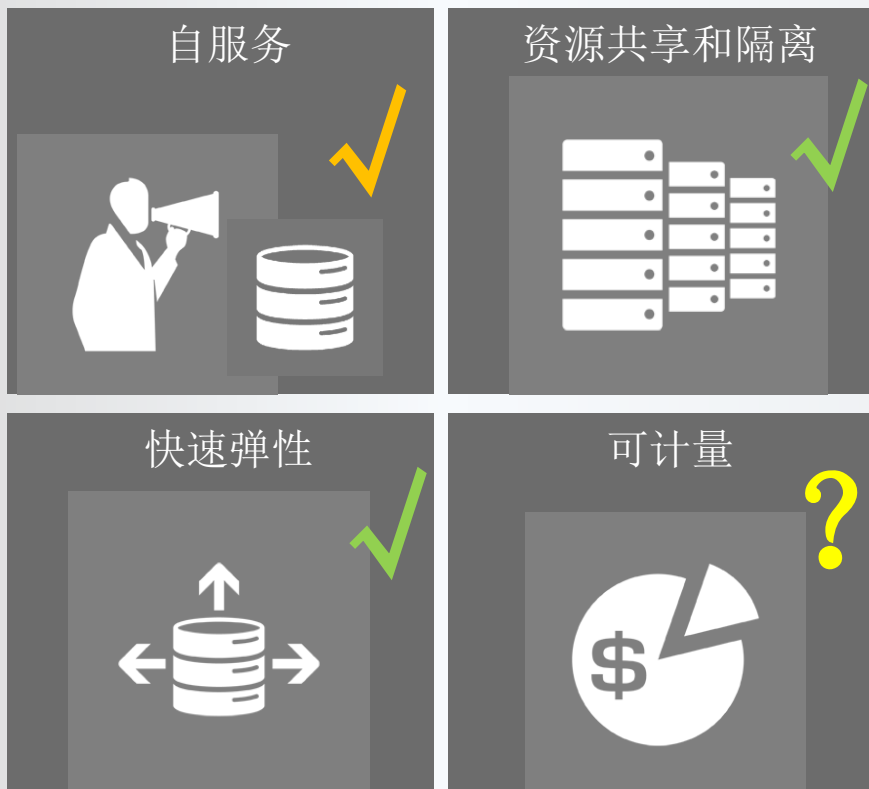
资源分配策略

编号	首选连接节点数	单节点处理器
OTP01	2	54
OTP02	1	54
OTP03	1	24
OTP04	1	16
OTP05	1	8
OAP01	1	54

应用维护等级

级别	适用场景
A	适用于直接面向客户的A类信息系统，对性能要求极高的信息系统。
B	适用于A类信息系统及对外服务的B类信息系统，以及性能要求较高的信息系统。
C	适用于非实时、非对外服务的B类信息系统及对性能要求不高的C类信息系统

- 性能目录编号+应用维护等级+是否需要灾备保护(P)，形成服务目录选取结果。如OTP02BP（高配服务于B类业务且需要异地容灾）
- 资源池管理员可根据编号，选择相对应的资源。服务目录对应用屏蔽了数据库的复杂度。
- 简化了技术方案设计
- 服务目录只是第一步，关键是运维部门在实践中积累经验和灵活调整能力



- 开发测试环境深化自服务
- 实践中优化服务目录
- 改进性能测试方法
- 探索资源的计量和后评估机制
- 结合Far Sync，实现RPO=0
- 同城双活架构

The logo for DBAplus features the letters 'DBA' in a bold, sans-serif font. The 'D' is red, the 'B' is blue, and the 'A' is orange. The word 'plus' is in a green, lowercase, sans-serif font. The entire logo is centered on a blue background with a faint, stylized city skyline and geometric patterns.

DBAplus

www.dbaplus.cn

THANK YOU



限时优惠，扫码抢票



APAC OTN TOUR 2017

The APAC OTN Tour 2017 will be running from November 20th until December 9th visiting 4 countries/7 Cities in the Asia Pacific Region. Bellow you can find more information regarding the events that are part of this year tour:

Dates:

- Wellington, NZ : November 20th
- Auckland, NZ : November 22nd
- Sydney, Australia: November 24th
- Melbourne, Australia: November 27th
- Perth, Australia: November 29th
- Shanghai, China: December 3rd
- Hyderabad, India: December 8 and 9

12月3日 D+ Day 欢迎来撩：
讲师、茶歇、场地、赞助，统统可以。

预计150人规模。
微信来撩：boypoo