

RDS for PG(PostgreSQL) 产品介绍

阿里云
digoal

目录

- 什么是 RDS
- RDS 产品形态
- RDS for PG 技术架构
- RDS for PG 产品特性
- 云 VS 自建
- RDS for PG 产品生态
- PG 与 MySQL结合使用解决哪些问题?

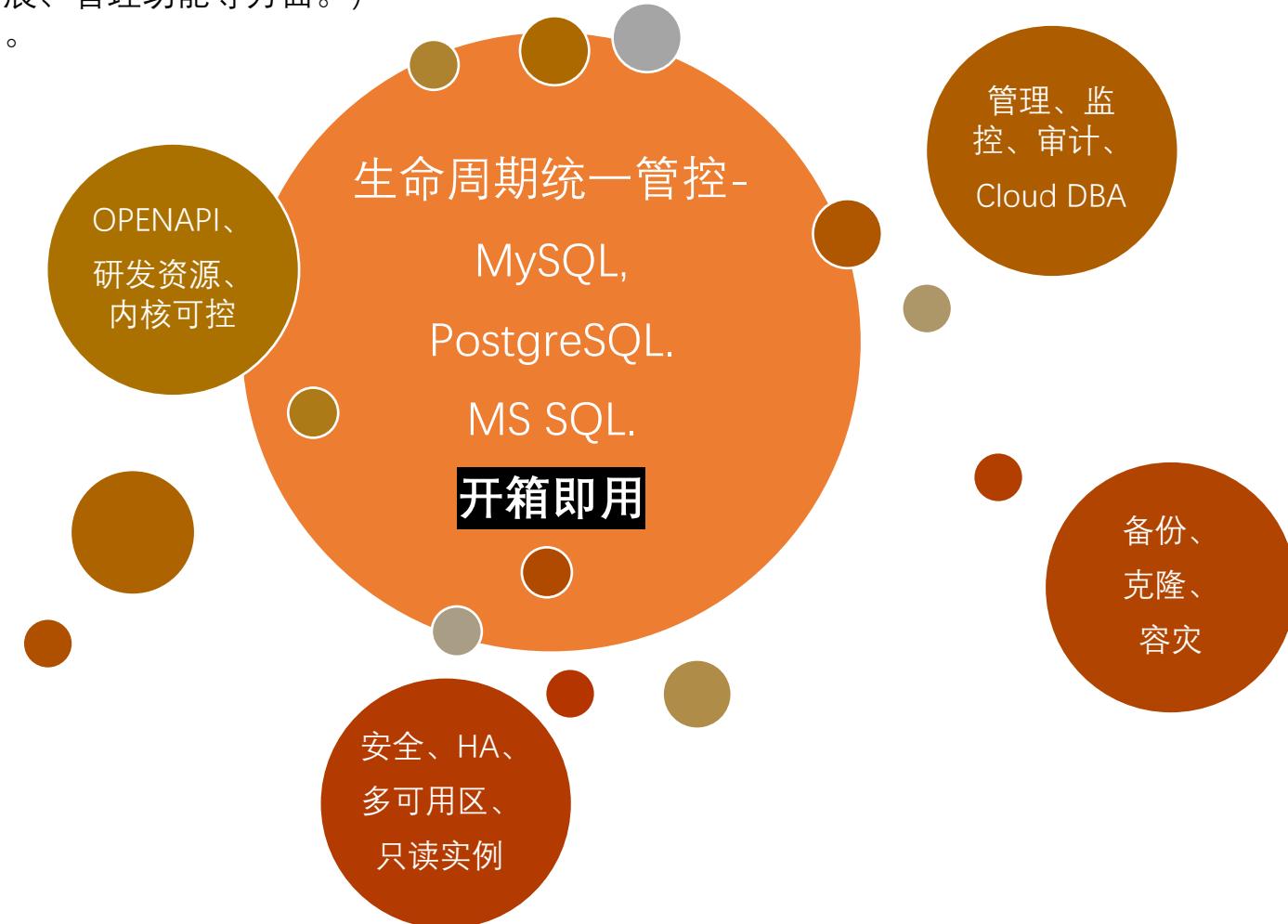
什么是RDS

RDS是阿里云提供的关系型数据库服务，支持MySQL, PostgreSQL, MS SQL。

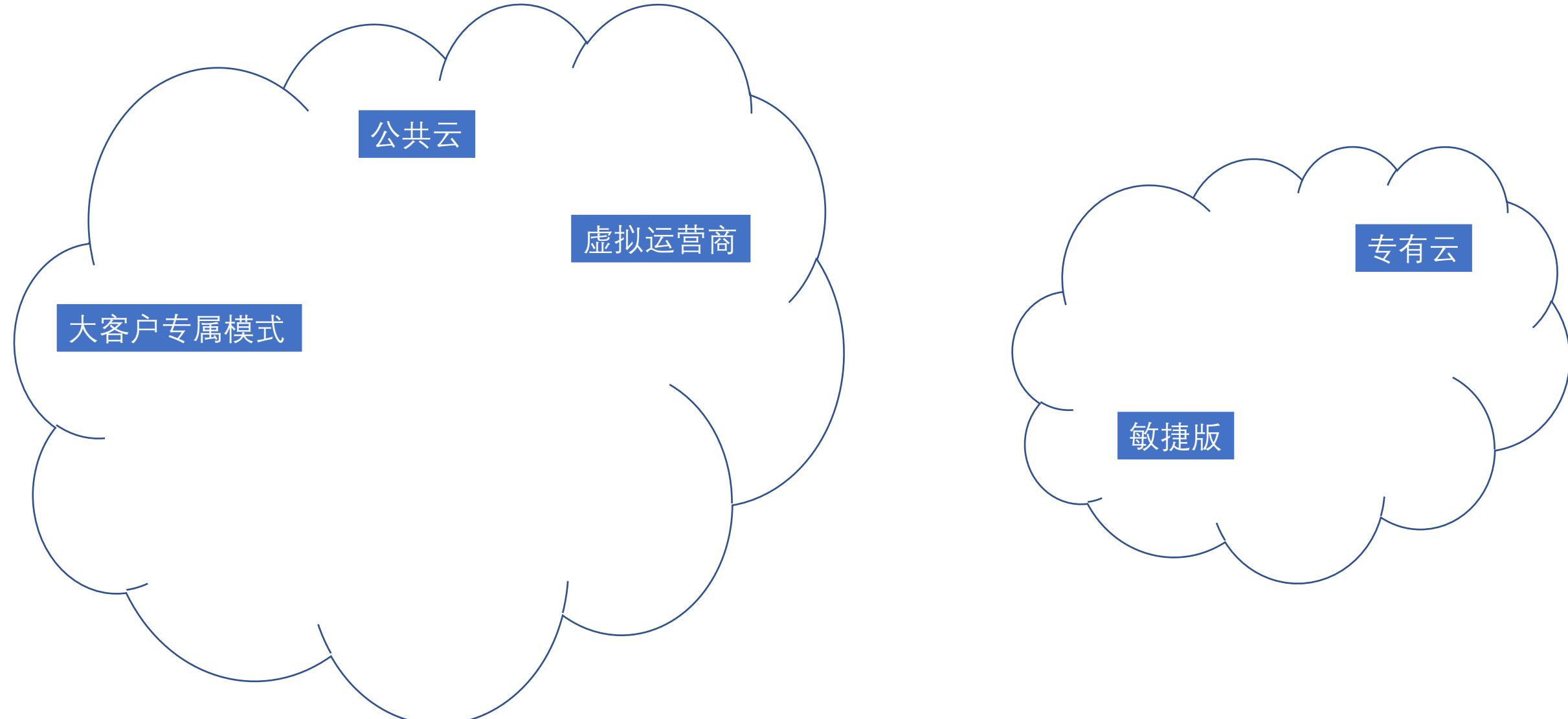
面向企业设计，统一的管控平台，完整支持数据库生命周期管理：

(可靠性、安全、性能检测与优化、扩展、管理功能等方面。)

为企业用户提供开箱即用的数据库能力。

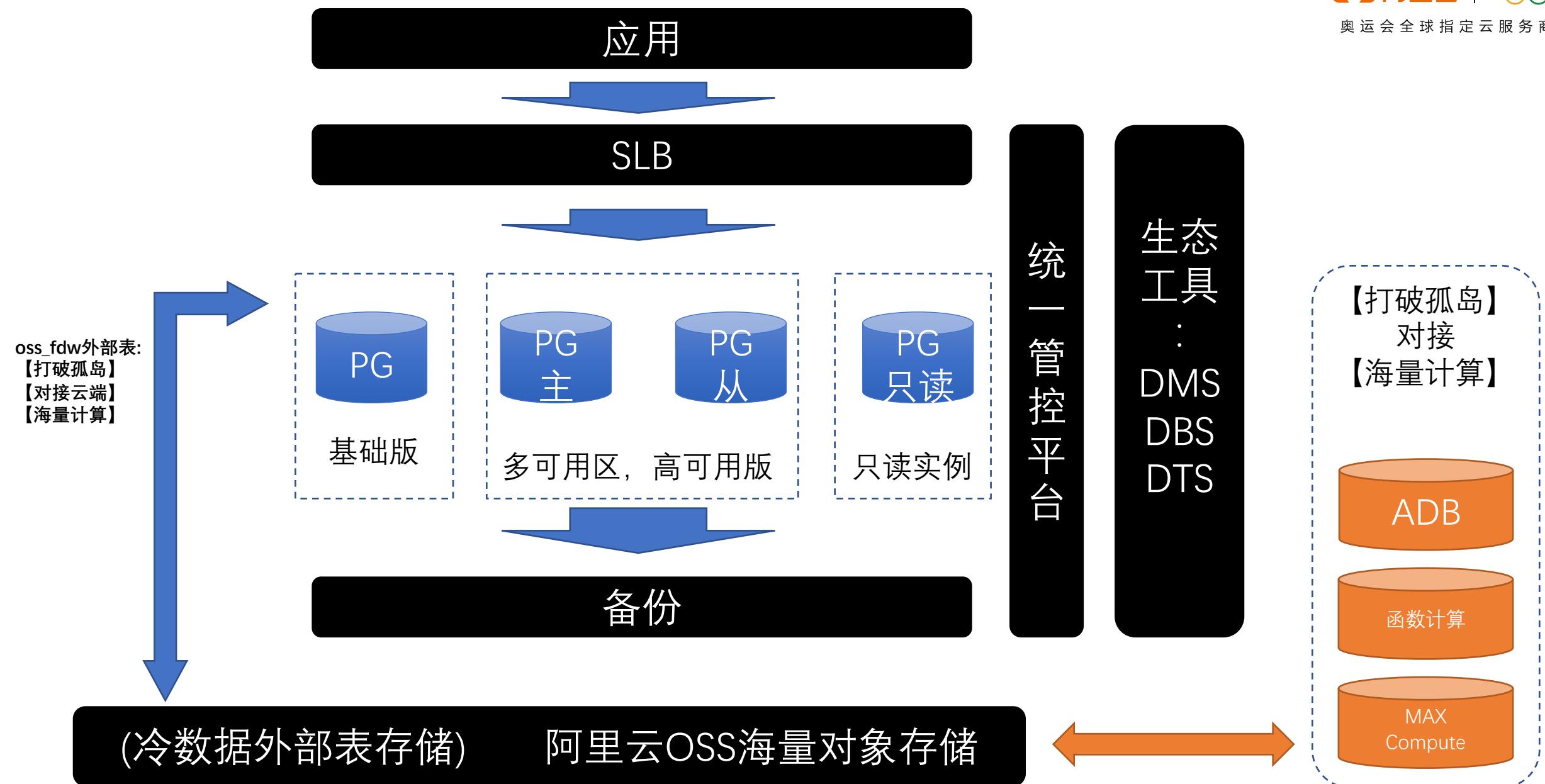


RDS for PG 输出形态



RDS for PG 技术架构

阿里云 | 
奥运会 全球 指定云服务商



RDS for PG 产品特性



- 支持高可用版、基础版
- HA版支持主备多可用区
- 保护模式可选(云盘高可用版)：
 - 最大保护、最高可用、最大性能
- 强同步、半同步、异步

高可用

- ESSD(云盘版、3副本)
- **32T 100W IOPS**
- SSD
 - 6TB 5W IOPS
- 本地盘
 - 6TB 5W IOPS
- 64C 512G(规划104C 768G)

大容量
高性能

- 全量+归档备份
- 克隆实例
 - (时间点、备份集、实例克隆)
- 备份策略管理

备份恢复

- **SSL链路**
- **数据加密**
 - 云盘加密(BYOK,CMK)
- 网络访问安全组
- SQL审计
- SGX全加密版(达摩院合作)
- 9副本(云盘6副本+oss3副本)

安全、可靠

- 只读、容灾实例
- 跨可用区(本地盘版支持, 云盘版2月)
- 跨Region(规划)
- 云盘版只读实例(规划)
- 不限规格

只读、容灾

- **性能洞察**
- SQL洞察
- 一键诊断
- 秒级监控

CloudDBA

- 开放OPENAPI
- 大型企业
- PaaS服务型企业
- 支持Terraform平台

OPENAPI

- 用户、数据库管理
- 版本升级管理
- 参数管理
- HA机制管理
- 性能监控、告警配置
- **SQL翻译(Oracle to PG)规划**
- 插件管理 (规划)

便捷管理

- 冷热分离存储 (OSS外部表)
- 实时中文分词、模糊查询
- 向量相似搜索 (图像搜索)
- 多维搜索
- **用户画像 (实时营销圈人)**
- PostGIS、Ganos时空
- SQL/JSON path语法…其他扩展

应用增强

- GPU加速规格(10基础版)
- CPU并行(10,11,12版)
- **20x+提速**

并行计算加速

RDS for PG 特性解读 - 1 【性能洞察】





高性能

- 平均20倍+提速
- RDS PG 11+



自动并行

- 不需要改写SQL
- 自动根据SQL代价启动并行计算
- 支持自定义配置并行度
- 支持配置最大并行度，防止资源耗尽



覆盖几乎所有复杂查询

- 表扫描
- 索引扫描
- SQL 过滤
- JOIN、子查询
- 聚合、自定义函数
- 分区表、外部表
- 自定义聚合、函数
- 排序、建表、物化、创建索引

RDS for PG 特性解读 - 2 【并行计算，支撑混合负载业务】



场景	数据量	关闭 并行	开启 并行	并 行 度	启 用 并 行 性 能 提 升 倍 数
全表扫描	10亿	53.4 秒	1.8 秒	32	29.7 倍
条件过滤	10亿	53.4 秒	1.87 秒	32	28.6 倍
哈希聚合	10亿	142.3 秒	4.8 秒	30	29.6 倍
分组聚合	10亿	142.3 秒	4.8 秒	30	29.6 倍
select into	10亿	54.5 秒	1.9 秒	32	28.7 倍
create table as	10亿	54.7 秒	2秒	30	27.35 倍
CREATE MATERIALIZED VIEW	10亿	54.7 秒	2秒	30	27.35 倍
create index	10亿	964 秒	252 秒	32	3.83 倍
parallel CREATE INDEX CONCURRENTLY - 不堵塞读写	10亿	509.6 秒	355 秒	16	1.44 倍
排序	10亿	76.9 秒	2.75 秒	32	28倍

RDS for PG 特性解读 - 2 【并行计算，支撑混合负载业务】



自定义并行聚合1(求 distinct 数组字段元素、以及count distinct)	10亿	298.8 秒	8.7 秒	36	34.3 倍
自定义并行聚合2(求 distinct 普通 字段元素、以及count distinct)	10亿	96.5 秒	3.43 秒	36	28 倍
自定义并行函数(UDF)	10亿	456 秒	16.5 秒	30	27.6 倍
普通并行(gather)	10亿	70.2 秒	2.5 秒	30	28.1 倍
归并并行(gather merge)	10亿	78.2 秒	2.76 秒	30	28.3 倍
rc (ud agg count distinct)	10亿	107 秒	3.65 秒	30	29.3 倍
rr (ud agg count distinct)	10亿	107 秒	3.65 秒	30	29.3 倍
parallel OLAP : 中间结果 parallel with unlogged table ; unlogged table 并行求avg case	10亿	73.6 秒	2.5 秒	30	29.44 倍
parallel index scan	10亿	19 秒	1.58 秒	20	12 倍
parallel bitmap scan	10亿	23.98 秒	15.86 秒	20	1.5 倍
parallel index only scan	10亿	8 秒	0.6 秒	20	13.33 倍

RDS for PG 特性解读 - 2 【并行计算，支撑混合负载业务】

parallel nestloop join	10亿 join 10亿 using (i) where t1.i<10000000	14.4 秒	4.6 秒	8	3.13 倍
parallel merge join	10亿 join 10亿 using (i) where t1.i<10000000	3.2 秒	1秒	8	3.2 倍
parallel hash join	10亿 join 10亿 using (i) where t1.i<10000000 and t2.i<10000000	8.1 秒	1秒	20	8.1 倍
parallel hash join	10亿 join 10亿 using (i)	1071 秒	92.3 秒	20	11.6 倍
parallel partition table wise join	10亿 join 10亿 using (i)	1006 秒	76 秒	24	13.2 倍
parallel partition table wise agg	10亿	191 秒	8 秒	24	23.9 倍
parallel append	10亿	70.5 秒	3.16 秒	24	22.3 倍
parallel append merge	10亿	99.4 秒	5.87 秒	24	16.93 倍
parallel union all	10亿	99 秒	5.6 秒	24	17.68 倍
parallel CTE	10亿	65.65 秒	3.33 秒	24	19.7 倍
parallel 递归查询, 树状查询, 异构查询, CTE, recursive CTE, connect by	异构数据1亿, 日志数据10亿	5.14 秒	0.29 秒	24	17.7 倍

RDS for PG 特性解读 - 2 【并行计算，支撑混合负载业务】



parallel scan mult FDW tables (通过继承表方式)	10亿	180 秒	7.8 秒	24	23.1 倍
parallel scan mult FDW tables (通过union all)	10亿	165.6 秒	27.8 秒	5	6 倍
parallel leader process	10亿	186 秒	95 秒	1	2 倍
parallel subquery	20亿	179.7 秒	6.5 秒	28	27.6 倍



时空、 GIS、 时序

- 插件：
- Ganos(自研GIS模块)
 - PostGIS(GIS模块)
 - Timescaledb
 - (时序模块)

- 索引：
- GiST
 - BRIN



图谱、 用户画像、 实时营销圈人

- 插件：
- ltree(树模块)
 - varbitx(位图模块)
 - pg_roaringbitmap
 - (高级位图模块)

- 索引：
- GIN
 - GiST
 - rum
 - bloom



图像、多维 相似分析、 向量搜索

- 插件：
- Imgsqlr(图像相似)
 - cube(高维向量)
 - Smlar(文本相似)
 - Pase(高维向量模块)
 - 向量相似
 - 人脸识别

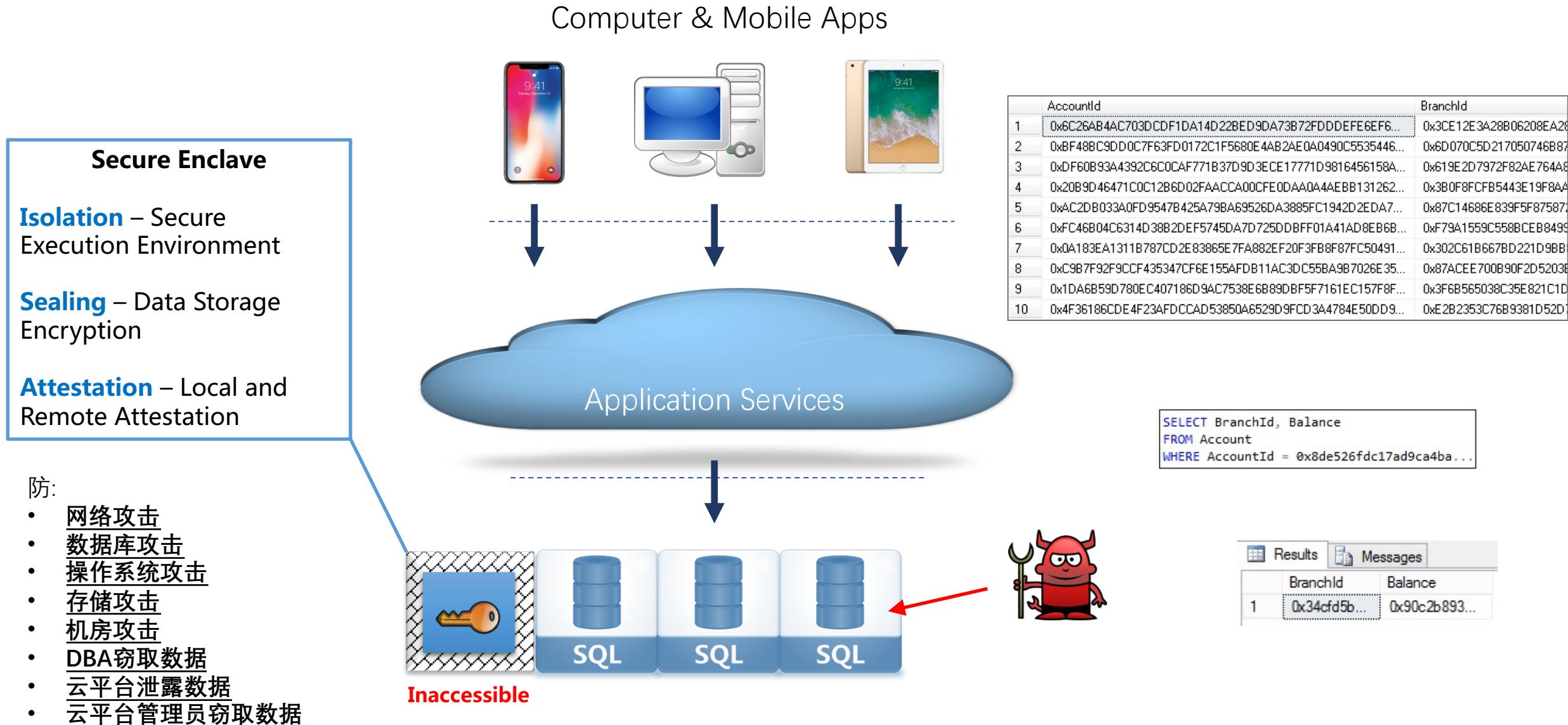
- 索引：
- GIN
 - GiST
 - Pase(hnsw, ivfflat算法)



文本 分析、搜索， 冷热分离

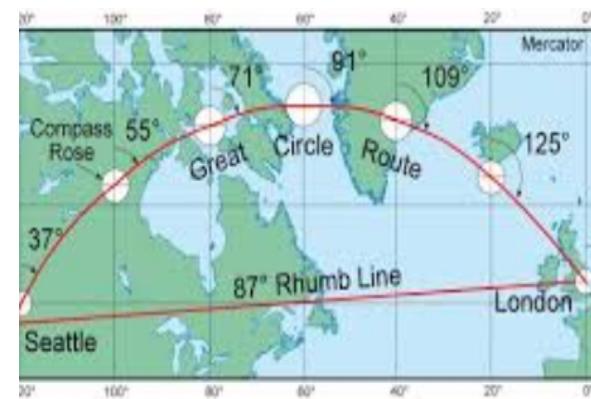
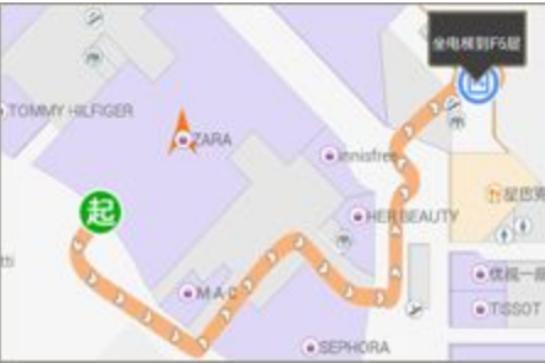
- 插件：
- pg_trgm(模糊查询)
 - Zhparser(中文分词)
 - oss_fdw(冷热分离)
 - JSONB(文档类型)
 - SQL/JSON path语法

- 索引：
- GIN
 - rum



RDS for PG 特性解读 - 5 【专业GIS模块 - Ganos】

阿里云 | 奥运会全球指定云服务商



RDS PG相比自建PG优势



管理优势



安全优势



研发优势

- 开箱即用
 - 多可用区HA、监控、
 - 备份、克隆、性能洞察
- 7*24 专家值班
- 多重HA保护模式可选
 - 最大保护(强同步)、
 - 最高可用(半同步)、
 - 最高性能(异步)
- 多重形态可选
 - 基础版、
 - 高可用版、
 - 只读实例

- SSL、
- 云盘加密、
- SGX全加密数据库规格、
- 安全组

- 持续fix bug
- 持续内核优化
- 专业GIS模块
- 定制、开源插件更丰富
- 持续更新版本
- 顶级数据库专家
- 完全兼容社区版本
 - 不绑定、自由割接

产品生态

阿里云 | 
奥运会 全球 指定云服务商



工具生态



合作伙伴



技术社群



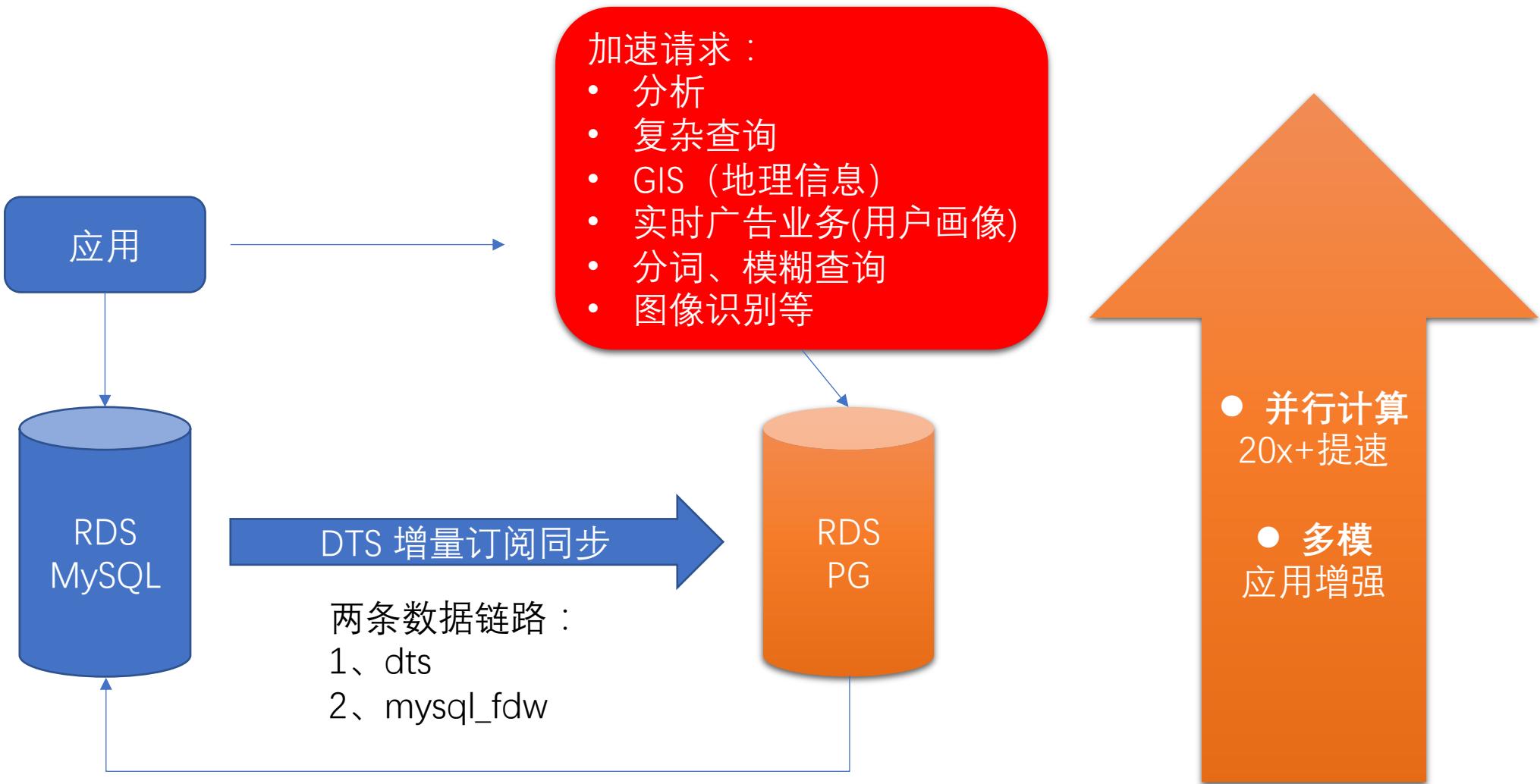
官方支持



官方文档



MySQL+PostgreSQL解决什么问题？



一期开课计划

- - 2019.12.30 19:30 RDS PG产品概览，如何与MySQL结合使用
- - 2019.12.31 19:30 如何连接PG，GUI，CLI的使用
- - 2020.1.3 19:30 如何压测PG数据库、如何瞬间构造海量测试数据
- - 2020.1.6 19:30 MySQL与PG类型、语法、函数等对应关系
- - 2020.1.7 19:30 如何将mysql数据同步到PG (dts)
- - 2020.1.8 19:30 PG外部表妙用 - mysql_fdw, oss_fdw (直接读写MySQL数据、冷热分离)
- - 2020.1.9 19:30 PG应用场景介绍 - 并行计算，实时分析
- - 2020.1.10 19:30 PG应用场景介绍 - GIS
- - 2020.1.13 19:30 PG应用场景介绍 - 用户画像、实时营销系统
- - 2020.1.14 19:30 PG应用场景介绍 - 多维搜索
- - 2020.1.15 19:30 PG应用场景介绍 - 向量计算、图像搜索
- - 2020.1.16 19:30 PG应用场景介绍 - 全文检索、模糊查询
- - 2020.1.17 19:30 pg 数据分析语法介绍
- - 2020.1.18 19:30 pg 更多功能了解：扩展语法、索引、类型、存储过程与函数。如何加入PG技术社群