



PostgreSQL中文社区



PostgreSQL中文社区

2021

PostgreSQL China Conference

主办：PostgreSQL 中文社区

第11届 PostgreSQL 中国技术大会

开源论道 × 数据驱动 × 共建数字化未来





RDS PG一键大版本升级技术解密

阿里云 - RDS产品部 | 汪建明 (风移)



CONTENTS 目录

- 1 PG行业位置
- 2 PG版本升级背景
- 3 PG版本升级解密
- 4 PG版本升级成果



PG 行业位置

全球数据库排行

3 次获得年度最佳数据库
开源数据库排名 TOP2

Rank			DBMS
Apr 2021	Mar 2021	Apr 2020	
1.	1.	1.	Oracle +
2.	2.	2.	MySQL +
3.	3.	3.	Microsoft SQL Server +
4.	4.	4.	PostgreSQL +
5.	5.	5.	MongoDB +
6.	6.	6.	IBM Db2 +
7.	7.	8.	Redis +
8.	8.	7.	Elasticsearch +
9.	9.	9.	SQLite +
10.	10.	10.	Microsoft Access

PG年度最佳产品奖

2020 PG亚洲大会上, 阿里云数据库专属集群
MyBase荣膺 “PG年度最佳产品奖”。



1

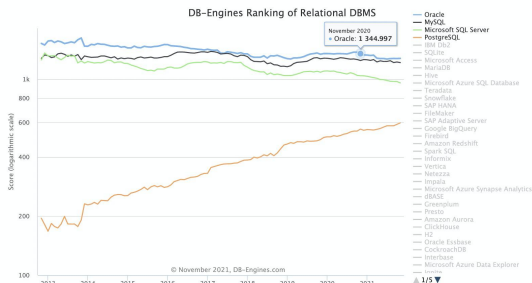
2

3

4

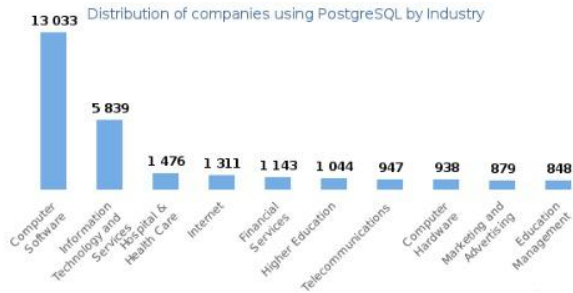
发展潜力

全球流行度趋势 TOP4
2021财年RDS PG实例数
同比增长45%
2021财年RDS PG的营收
同比增长50%



广泛应用于

计算机软件
信息技术及服务
医疗及健康
金融服务
高等教育
通讯服务





RDS PG 优势

RDS PG >> 自建PG

可靠性

- ❑ Logical Slot Failover
- ❑ Standby支持多上游结点
- ❑ 一键大版本升级

安全性

- ❑ 云盘加密
- ❑ SSL自定义证书
- ❑ Intel SGX全加密
- ❑ AD、LDAP访问控制

智能化

- ❑ Metrics: 80+ 指标的秒级监控
- ❑ DAS: 一键诊断、SQL洞察、慢SQL统计和审计等诊断优化功能

丰富插件

- ❑ Ganos: 时空引擎
- ❑ PASE: 高效向量检索
- ❑ oss_fdw: 数据冷热分离

对比结果：RDS PG在可靠性、安全性、智能化、插件丰富度方面优势明显。



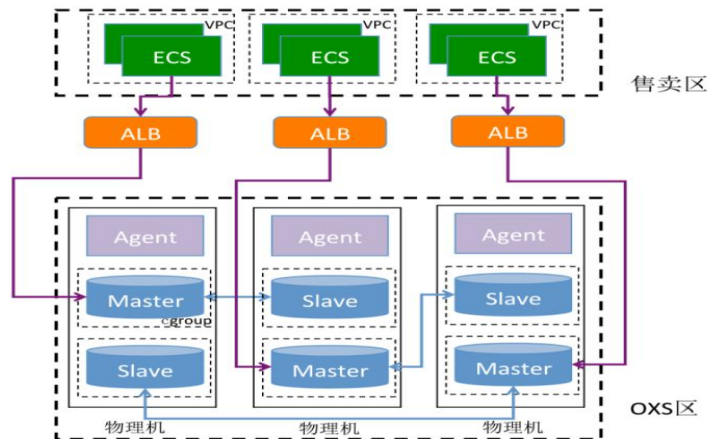
CONTENTS 目录

- 1 PG行业位置
- 2 PG版本升级背景
- 3 PG版本升级解密
- 4 PG版本升级成果



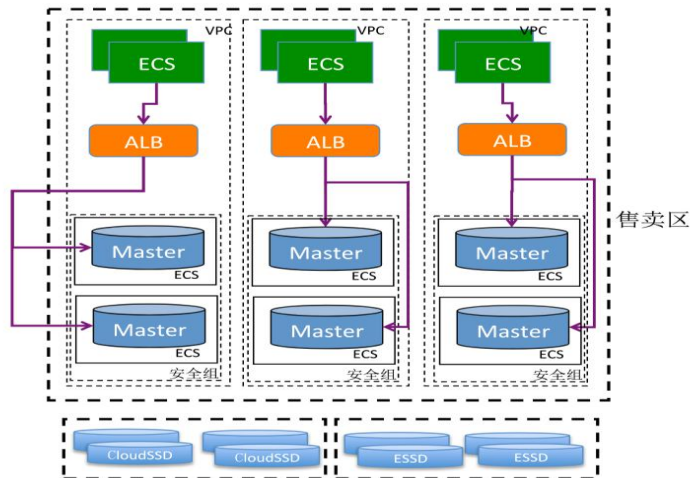
版本与形态

物理机本地盘
RDS PG 9.4-10



- ❑ 资源弹性
- ❑ 资源强隔离
- ❑ 改善供应链
- ❑ 存计算分离

存计分离架构
RDS PG 10-14



RDS PG目前有6个大版本，2种管控架构，低版本物理机本地盘占比达到1/3



背景与挑战

1

低版本难以维护

过时低版本

PG 9.4, 版本过老
6U机型供应链问题
社区不维护, 难以兜底

2

高版本快速迭代

更多新特性

增量排序
并行索引垃圾回收
索引deduplicate能力
分区表、聚合性能提升

3

管控架构升级

新管控红利

秒级快照
极致弹性能力
更大存储空间支持
备份操作无性能损耗

4

如何平滑升级

平滑割接

保证兼容性
可验证、可回滚能力
应用零改动、不停服
一键升级、高度产品化

思考与策略: RDS PG 提供产品化、平滑割接、可预检、可验证和可回滚的大版本升级能力。



CONTENTS 目录

- 1 PG行业位置
- 2 PG版本升级背景
- 3 PG版本升级解密
- 4 PG版本升级成果



设计原则

可验证、可回滚

- 版本回滚：大版本回滚
- DNS地址：连接串回滚
- 可验证：能够做升级演练

验证
回滚

1

限制
要少

2

场景全覆盖

- 可升级到任意高版本
- 高可用和单节点均支持
- 无DDL操作限制
- 无表结构限制
- 无数据类型限制

一键
升级

3

平滑
割接

4

一键升级产品化

- 拒绝升级手册
- 客户一键升级，应用零改动
- 产品侧保证升级成功和兼容性

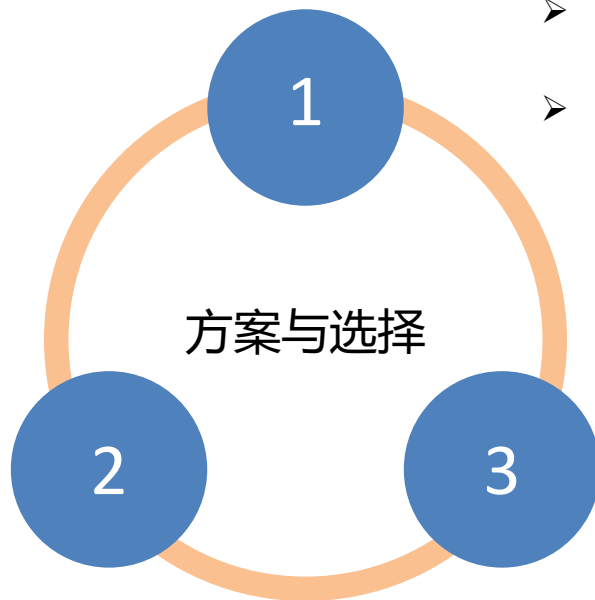
应用最小感知原则

- 升级过程应用不停服
- 升级过程速度快
- 连接地址自动切换

思考与策略：把复杂留给自己，把简单留给用户。



方案选择



2. pg_upgrade

- 优点:
 - 不拷贝数据, 仅元数据升级
 - 效率高, 2TB数据, 升级 < 10s
 - 无逻辑迁移的各种限制
- 缺点:
 - 要考虑验证和回滚逻辑
 - 要考虑产品化和不停服
 - 要考虑参数、插件和内核兼容性
 - 复杂度高、工作量大、挑战大

1. 逻辑复制

- 优点:
 - 兼容性好、平滑割接
- 缺点:
 - 每个库都要创建发布、订阅
 - 每个表必须有PK / UK
 - 不支持DDL、大对象
 - 外键和触发器禁用
 - 可能导致到WAL日志堆积

3. pg_dump

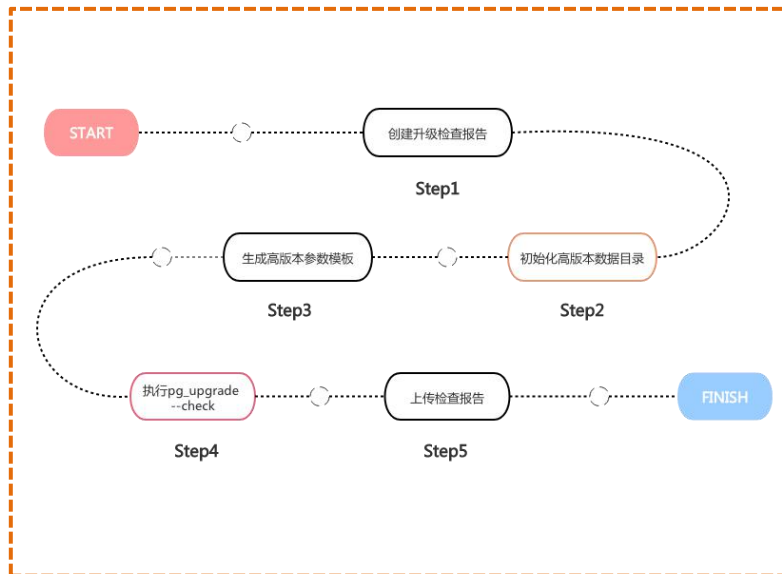
- 优点:
 - 兼容性好
 - 实现简单、工作量小
- 缺点:
 - 仅适用全量迁移
 - 效率低下
 - 应用停服时间长

最终方案：选择限制少、效率高的pg_upgrade方案作为产品化的基础。



大版本升级检查

升级预检查流程



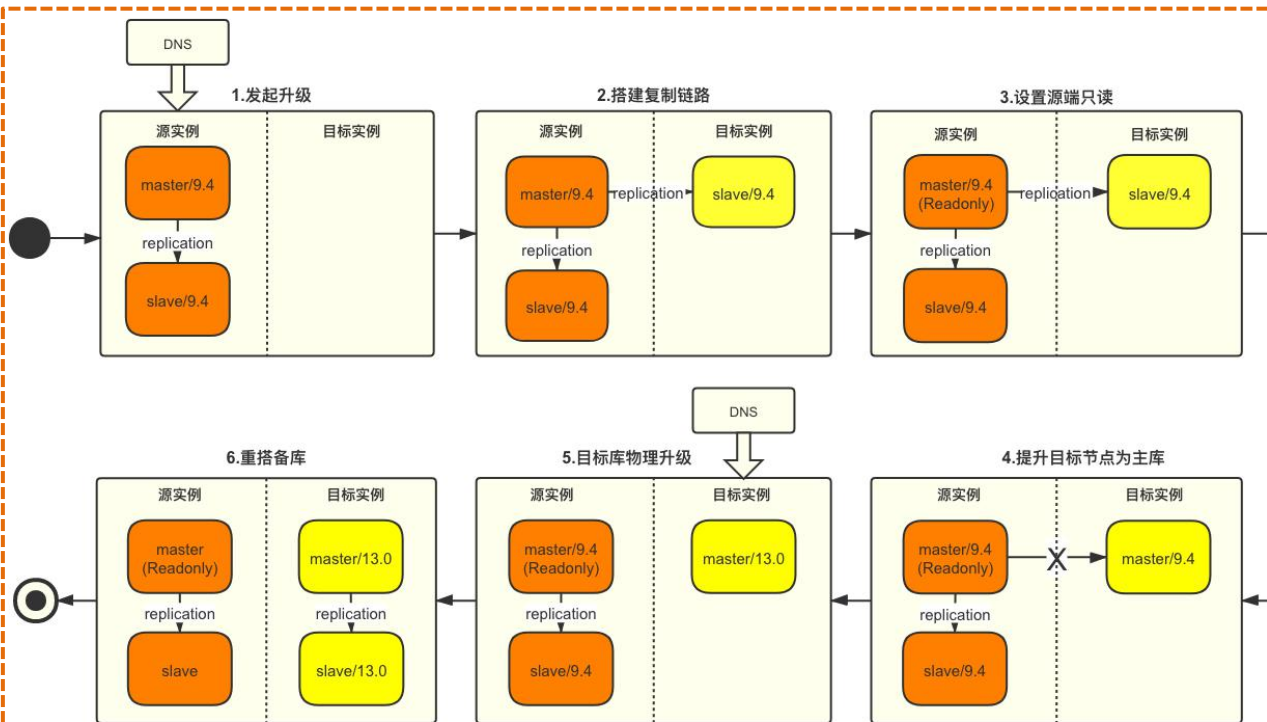
升级预检查结果

```
1 [pg_upgrade_internal_log]Performing Consistency Checks on Old Live Server
2 -----
3 Checking cluster versions                                ok
4 Checking database user is the install user              ok
5 Checking database connection settings                   ok
6 Checking for prepared transactions                      ok
7 Checking for reg* data types in user tables             ok
8 Checking for contrib/isn with bigint-passing mismatch  ok
9 Checking for tables WITH OIDS                           ok
10 Checking for invalid "sql_identifier" user columns     ok
11 Checking for presence of required libraries            ok
12 Checking database user is the install user              ok
13 Checking for prepared transactions                      ok
14
15 *Clusters are compatible*
16
```

预检查结果: 在线检查升级风险, 能否升级, 结果一目了然



升级流程



应用不停服, 仅分钟级的RO

□ 应用不停服

不停服: 全程可读

平滑性: 第5步, 连接地址交换

□ 可验证可回滚

可验证: 克隆模式, 源实例零干预

可回滚: 第5步之前, 零代价回滚, 连接地址随时可回滚

□ 效率高

速度快: 第5步2T10秒

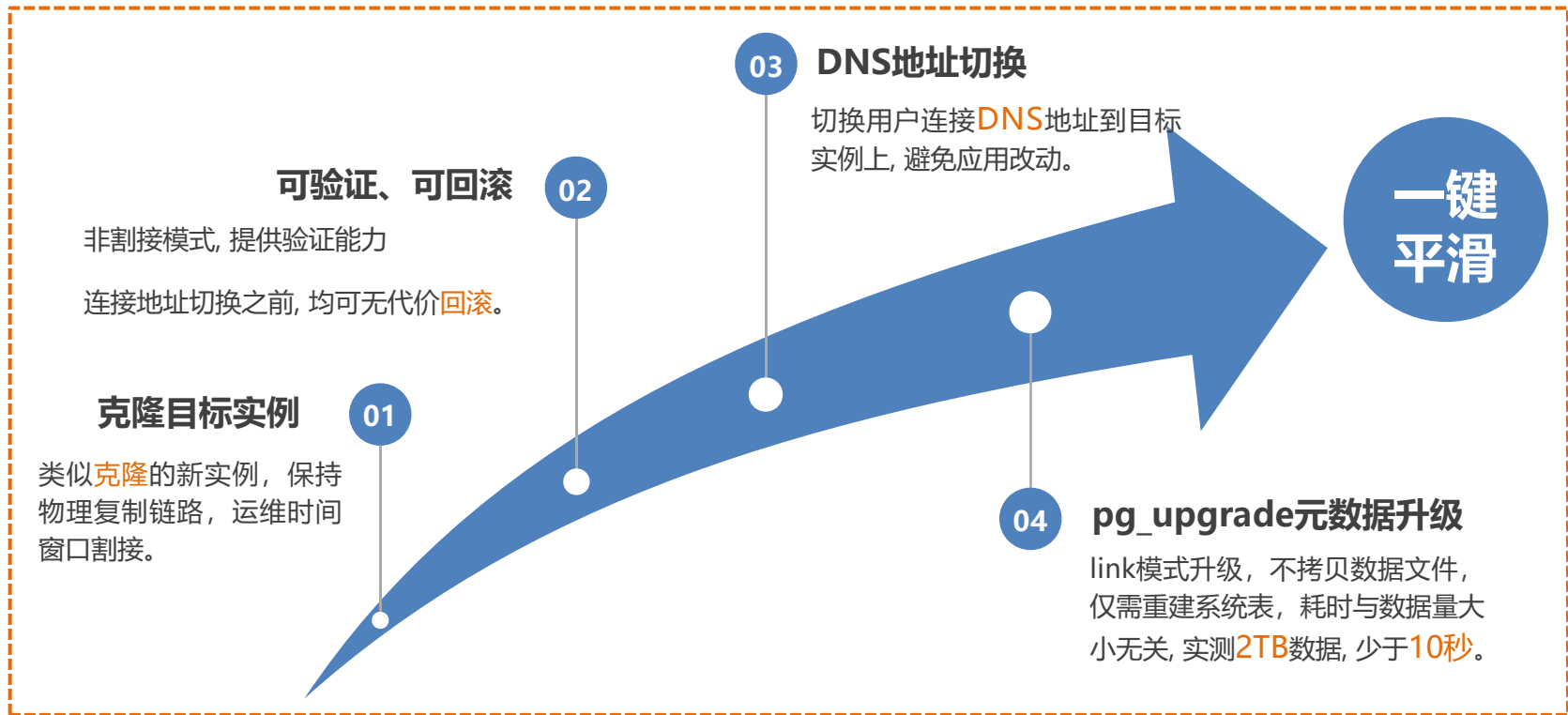
重搭快: 秒级快照, 10分钟左右, 与数据量大小无关

□ 用户影响

第 3-5 步, 仅分钟级RO时间



方案总结





CONTENTS 目录

- 1 PG行业位置
- 2 PG版本升级背景
- 3 PG版本升级解密
- 4 PG版本升级成果

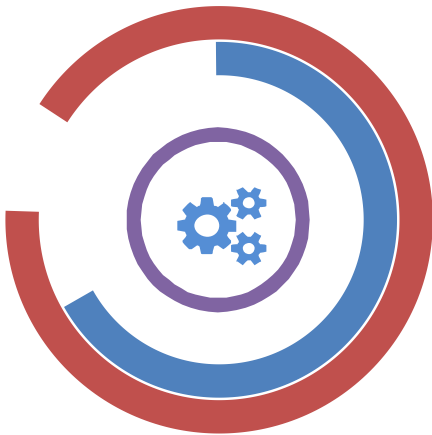


数字成果

8%

升级占比

升级实例占比：超过 **8%** 的实例通过一键大版本升级RDS PG。



21%

评估占比

升级评估占比：超过 **21%** 的实例做了实例升级评估。

2T 10s

升级快

- ❑ 元数据升级，实测**2T 10s**完成。
- ❑ 无论存储大小，备库重建 **20** 分钟内。

329T

升级稳

- ❑ 大版本升级实例总数据量达**329TB**。
- ❑ 应用**零**停机，分钟级RO。

0

安全

- ❑ 可验证，可回滚，DNS地址交换
- ❑ 应用**零**停机，**零**故障。

专利



成果展示

版本全覆盖
15版本

PG 9.4 向上5个版本
PG 10 向上4个版本
PG 13 向上1个版本

应用不停服
零宕机

应用无需停服,零宕机
用户连接串平滑割接
DNS交换仅分钟级RO

升级速度快
2T10秒

仅元数据升级
备库重搭10分钟
实测2TB数据10秒完成

割接克隆
两模式

割接模式, 平滑割接
克隆模式, 升级验证
连接地址交换, 可回滚

一举两得
弹性

底层架构一键升级
PG引擎一键大版本升级
具备弹性、存计算分离

取得成果: RDS PG大版本升级在覆盖面、可用性、效率、可验证可回滚能力方面取得显著成果。



行业对比



升级能力	阿里云	国际友商A	国际友商A2	国内友商A	国内友商B	优点
跨版本升级	✓	✓	✗	✗	✗	新特性&稳定性
零宕机割接	✓	✗	✗	✗	✗	分钟级切换
一键产品化	✓	✗	✗	✗	✗	提升用户体验
兼容性分析	✓	✗	✗	✗	✗	降低失败率
可验证可回滚	✓	✗	✗	✗	✗	降低业务风险

行业对比：阿里云RDS PG一键大版本升级在产品化、用户体验、可验证可回滚能力上引领行业。



前后对比

升级后



计算存储分离

- 更快：备库重搭、只读实例和备份更快，秒级快照
- 更稳：备份和重搭不消耗IO
- 更大容量：最大 32TB



ECS独享 + ESSD

- 更快：性能平稳、减少毛刺
- 更稳：资源独占、不存在争抢
- 安全：VPC、SG、云盘加密、AD LDAP访问控制、SGX全加密



运维习惯

- IOPS放开：不具备条件、升级ESSD Level (4G/S)、实例规格
- 插件：更多、更新版本插件支持
- 性能指标：更多、更全、5秒粒度

>>

大版本升级对比

升级前



物理机 + 本地盘

- 备库、只读重搭，物理COPY
- 备份：IOPS消耗严重、性能抖动、指标毛刺
- 容量：最大 6TB



主机共享 + 资源隔离

- 毛刺：性能毛刺且无法根本解决
- 资源：CPU Cores、亲和调度
- 安全：无法彻底隔离、阉割功能
- 供应链：6U机型过保、下线



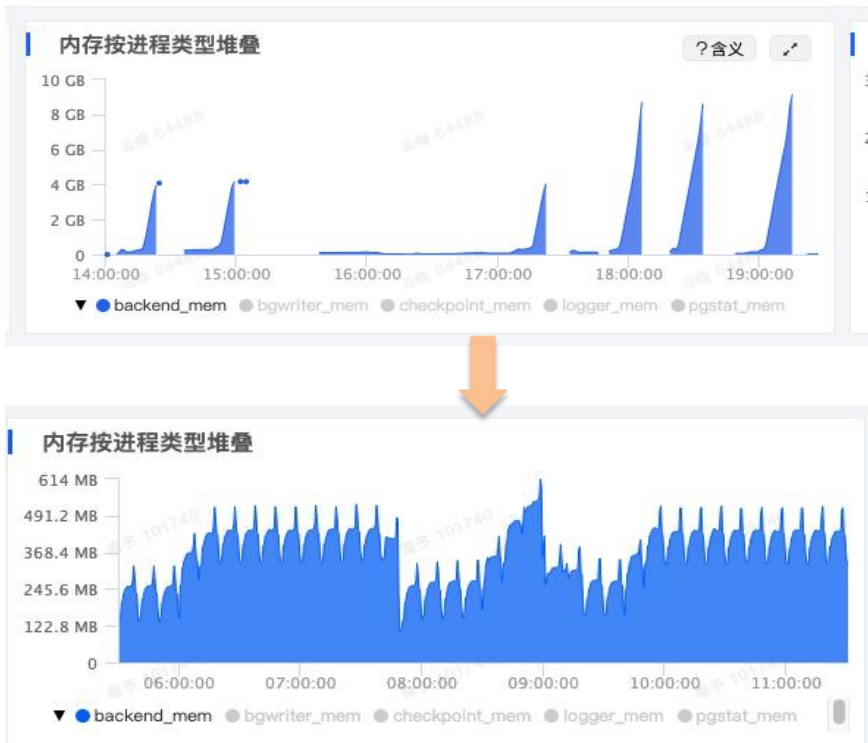
运维习惯

- IOPS放开：阿里云手动放开
- 插件：阿里云后台支持
- 性能指标：受限、相对较少



客户案例

客户	某大客户
时间	2021/03/15
问题	RDS PG12频繁OOM
现象	PG backend 进程内存快速上涨至8G
原因	用户SQL执行计划使用Hash Aggregation, 中间对象始终放在内存, 容易OOM。
解法	联系用户自助升级至 PG13 , 该版本Hash Aggregation使用内存超过一定大小会落盘。
效果	内存占用由8GB降低到491M。





友情提醒

0
1

同主机多实例升级

避免对同主机上多个实例同时升级，可能会导致数据备份过程中会打满磁盘IO，影响实例稳定性和高效性。

0
2

IOPS不支持临时放开

升级后实例资源独享、不具备资源共享特性，所以不具备IOPS临时放开的先决条件。可通过：

- ❑ ESSD PL等级升级，PL3最大吞吐达4G/S
- ❑ 实例规格升级、增大ECS VM吞吐

0
3

自建数据逻辑订阅

- ❑ 用户云下备库需要重搭
- ❑ ECS、IDC数据逻辑订阅重建、与大版本升级适配

风险
提醒



2021 PostgreSQL China Conference
第 11 届 PostgreSQL 中国技术大会



PostgreSQL 中文社区

THANKS

谢谢观看

开源论道 × 数据驱动 × 共建数字化未来