

# 阿里云RDS PG最佳实践

云栖社区在线培训

digoal , 阿里云数据库内核高级技术专家

- 上云实践
- RDS PG使用最佳实践
- 阿里云RDS相关周边组件用法
- 数据同步



- 如何评估 RDS 规格
- RDS PG 性能评测
- RDS PG数据迁移
  - 同构全量迁移
  - 同构增量迁移
  - 异构全量迁移
  - 异构增量迁移



## • 可用区

- 尽量与应用服务器在同一可用区。
- 否则只能通过公网地址访问。

## • 数据库版本

- 根据业务需求选择，建议选择新版本。

## • 网络

- 与应用服务器在同一VPC，或者都在经典网络。
- 否则只能通过公网地址访问。

## • 容量

- 根据实际的需求购买，最后留有足够余量（用于存储XLOG，已经运行SQL时可能产生的临时文件）。余量建议不低于1GB、同时不低于实际数据空间的20%

## • 内存

- 建议大于活跃数据的实际占用空间

## • CPU

- 根据购买的内存规格设定

## • IOPS

- 根据购买的内存规格设定

# RDS PG 性能评测与瓶颈分析

- TPC-B
- TPC-C
- 自定义模型评测



- TPC-B 性能评测，测试模型

```
BEGIN;
UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid;
SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
UPDATE pgbench_tellers SET tbalance = tbalance + :delta WHERE tid = :tid;
UPDATE pgbench_branches SET bbalance = bbalance + :delta WHERE bid = :bid;
INSERT INTO pgbench_history (tid, bid, aid, delta, mtime) VALUES (:tid, :bid, :aid, :delta,
CURRENT_TIMESTAMP);
END;
```

- 在同VPC或同经典网络的ECS，安装PostgreSQL最新稳定版的客户端
- 配置白名单，允许该ECS访问被测试的RDS PG



- 生成测试数据(1个scale代表 10万记录，约15MB)。如下表所示，1亿数据, 15GB：

```
pgbench -i -q --foreign-keys -s 1000
```

- 测试(256连接)

```
export PGPASSWORD=$RDS_PG_USER_PWD  
pgbench -M prepared -n -r -P 5 -c 256 -j 8 -T 600 -h $RDS_PG_IP -p $RDS_PG_PORT -U $RDS_PG_USER
```

- 分析瓶颈
  - 测试结果是否满足业务需求？
  - 分析当前的RDS资源消耗指标，找到可以提升的指标，升级规格：
  - 登陆 RDS管理控制台 -> 监控与报警 -> 监控 -> CPU与内存、IOPS

- 首先判断，TPS是否满足业务场景需求？

```

progress: 2425.0 s, 1797.9 tps, lat 142.550 ms stddev 79.681
progress: 2426.0 s, 1534.1 tps, lat 150.261 ms stddev 78.176
progress: 2427.0 s, 1521.1 tps, lat 186.646 ms stddev 131.803
progress: 2428.0 s, 1827.3 tps, lat 138.590 ms stddev 83.922
progress: 2429.0 s, 1688.0 tps, lat 152.734 ms stddev 80.126
progress: 2430.0 s, 1762.0 tps, lat 144.072 ms stddev 77.251
progress: 2431.0 s, 1789.0 tps, lat 145.597 ms stddev 77.115
progress: 2432.0 s, 1688.0 tps, lat 149.240 ms stddev 69.960
progress: 2433.0 s, 1617.0 tps, lat 156.999 ms stddev 73.309
progress: 2434.0 s, 1782.0 tps, lat 142.088 ms stddev 84.638
progress: 2435.0 s, 1805.9 tps, lat 141.142 ms stddev 111.987
    
```

- 如果不能满足需求，接下来看实例诊断

SRE | RDS 技术·保障奇迹 数据库管理平台

首页 实例管理 主机管理 任务管理 资源管理

实例链路信息 实例性能信息 实例卡慢诊断 执行SQL

查询区间: 最近3小时 最近6小时 最近12小时 最近1天 最

Y-values

100k  
50k  
0k

15:15 15:30

磁盘空间  
磁盘空间详情  
disk\_usage  
cpu利用率  
iops  
IOPS利用率  
物理内存(MB)  
mem\_usage  
连接数利用率  
mem\_cache  
cpu\_throttled\_averag...  
conn\_count  
ms\_latency

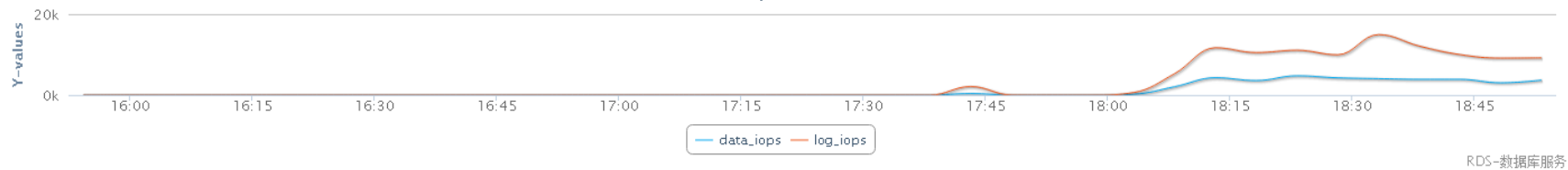


## 实例诊断

abc cpu利用率(单位:%)



abc iops(单位:int)



abc IOPS利用率(单位:%)

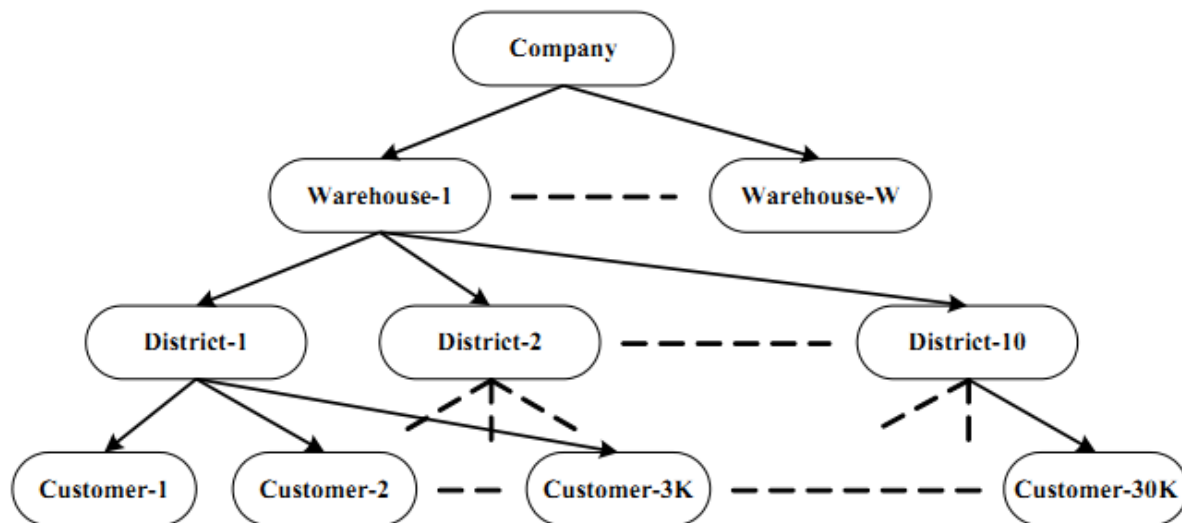


abc 物理内存(MB)(单位:MB)

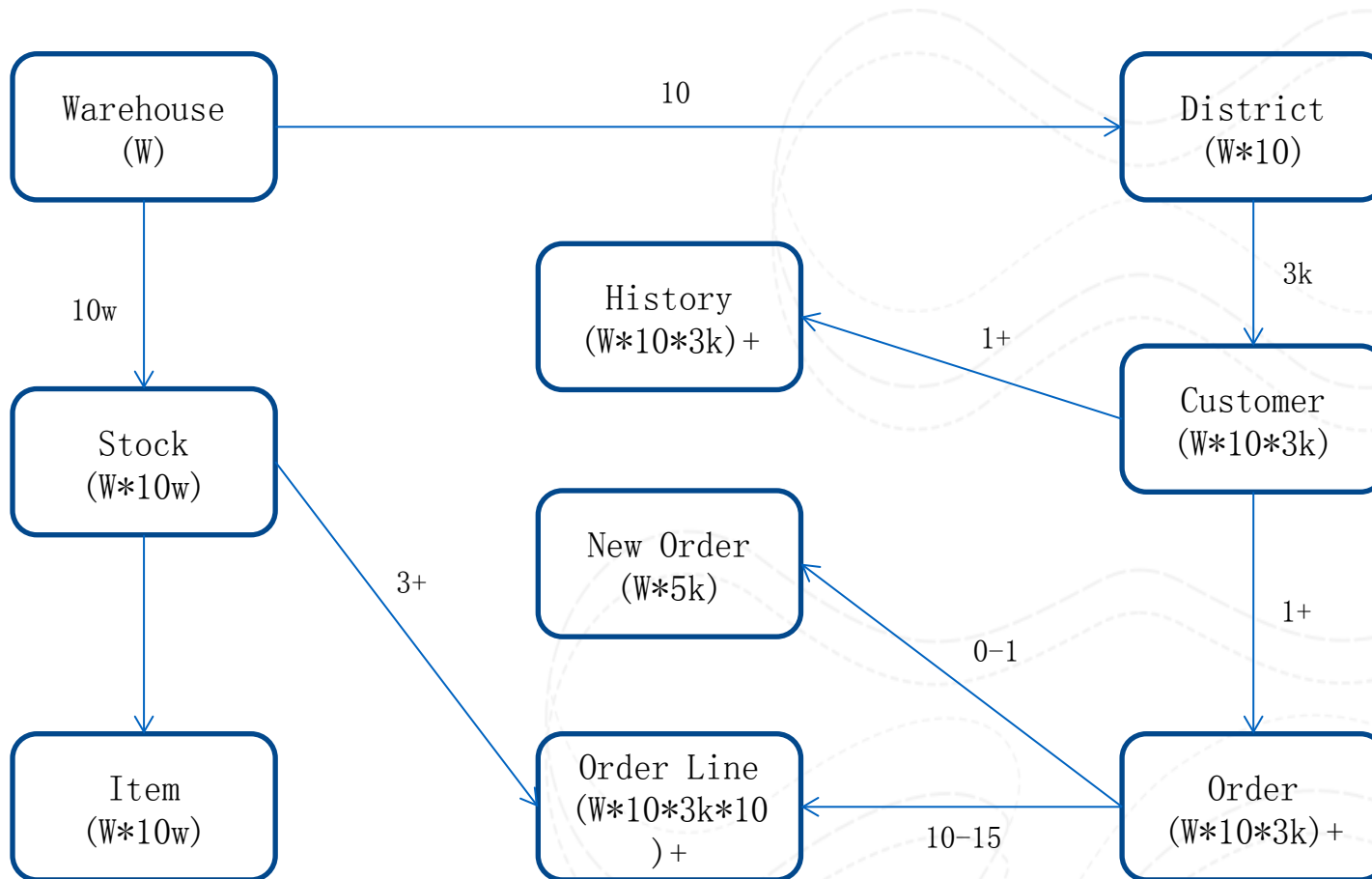


- 数据关系
  - 公司 -> 仓库(W)
  - 仓库 -> 销售点(10)
  - 销售点 -> 客户(3000)

W是生成测试数据时用户可以指定的



## • 数据关系



- 新建订单
- 支付
- 订单状态查询
- 发货
- 库存状态查询

## TPC-C 测试软件

- HammerDB : <http://www.hammerdb.com/document.html>

- OLTP 性能评测
- 使用sysbench测试阿里云RDS PostgreSQL性能  
( <https://yq.aliyun.com/articles/35517> )
- <https://github.com/digoal/sysbench>
- 网络延迟 性能评测
- PostgreSQL 网络延迟 瓶颈定量分析：  
<https://yq.aliyun.com/articles/35176>
- RDS 吞吐量 性能评测
- 通过并行COPY 或写大对象测试网络吞吐率

- 场景建模评测

- 创建测试表
- 生成测试数据
- 创建测试脚本
  - 变量与随机变量设置

```
\set varname expression
```

- 表达式中支持操作符 unary (-) or binary operators (+, -, \*, /, %)

```
\setrandom varname min max [ uniform | { gaussian | exponential } parameter ]  
\sleep number [ us | ms | s ]
```

## 例子

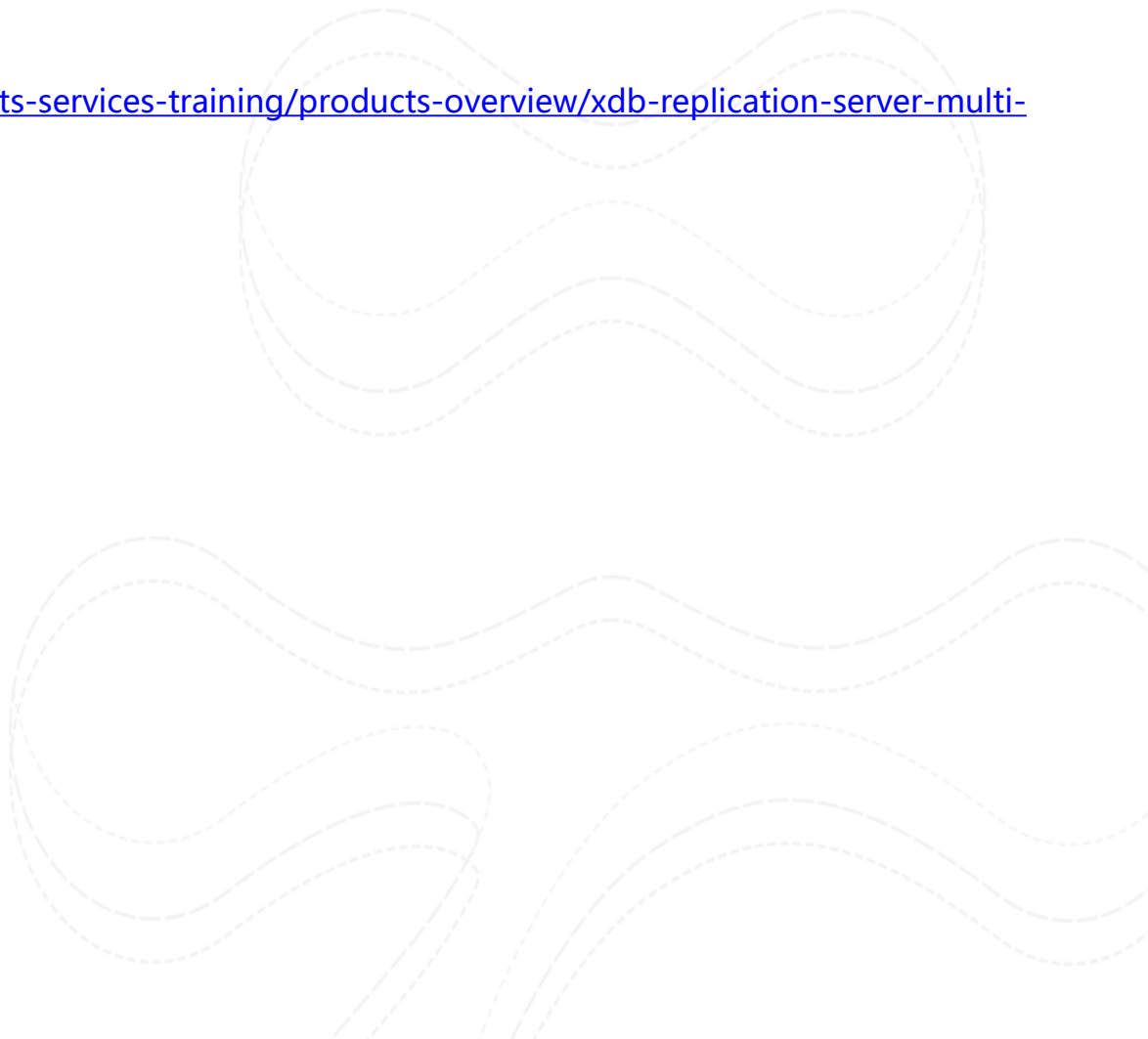
```
\set nbranches :scale
\set ntellers 10 * :scale
\set naccounts 100000 * :scale
\setrandom aid 1 :naccounts
\setrandom bid 1 :nbranches
\setrandom tid 1 :ntellers
\setrandom delta -5000 5000
BEGIN;
UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid;
SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
UPDATE pgbench_tellers SET tbalance = tbalance + :delta WHERE tid = :tid;
UPDATE pgbench_branches SET bbalance = bbalance + :delta WHERE bid = :bid;
INSERT INTO pgbench_history (tid, bid, aid, delta, mtime) VALUES
(:tid, :bid, :aid, :delta, CURRENT_TIMESTAMP);
END;
```

```
export PGPASSWORD=$RDS_PG_USER_PWD
pgbench -M prepared -f test.sql -n -r -P 5 -c 256 -j 8 -T 600 -h $RDS_PG_IP -p
$RDS_PG_PORT -U $RDS_PG_USER
```

- 从用户端PostgreSQL 迁移到 阿里云 PostgreSQL
- 方法1 在云端数据库创建与线下数据库一致的用户、数据库名
  - pg\_dump 依次导出线下数据库
  - pg\_restore 依次还原线下数据库
  - [https://help.aliyun.com/document\\_detail/rds/rds\\_qs\\_pg/rds\\_qs\\_pg\\_00007.html](https://help.aliyun.com/document_detail/rds/rds_qs_pg/rds_qs_pg_00007.html)
- 方法2 DTS
  - [https://help.aliyun.com/knowledge\\_list/8314900.html](https://help.aliyun.com/knowledge_list/8314900.html)
  - <https://dts.console.aliyun.com/>



- xDB
- <http://www.enterprisedb.com/products-services-training/products-overview/xdb-replication-server-multi-master>



- DTS
- [https://help.aliyun.com/knowledge\\_list/8314900.html](https://help.aliyun.com/knowledge_list/8314900.html)
- <https://dts.console.aliyun.com/>

- xDB
- 支持Oracle,SQL Server 到 EDB
- 支持Oracle,SQL Server 到 PostgreSQL
- 或反过来
- symmetricds
- Oracle GoldenGate

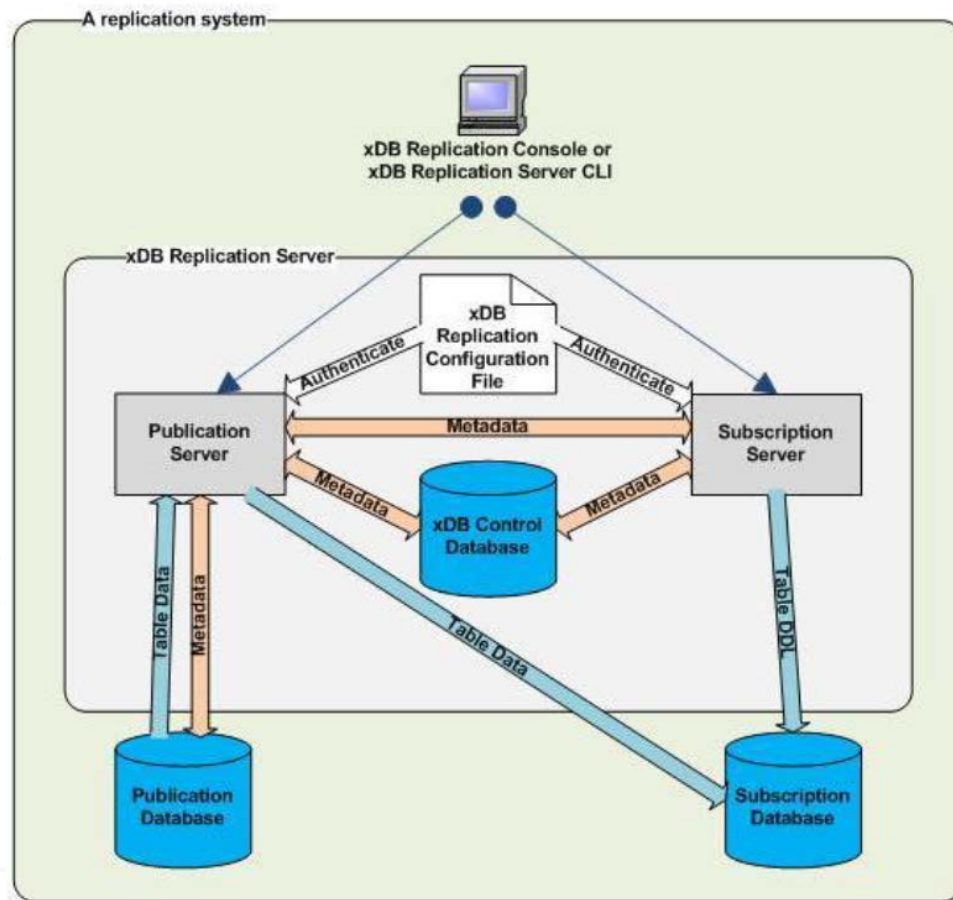


Figure 8 - xDB Replication Server - physical view (single-master replication system)

- 上云实践
- RDS PG使用最佳实践
- 阿里云RDS相关周边组件用法
- 数据同步



- 备份与恢复
- RT的优化
- 插件用法
- 参数最佳实践
- 安全与审计
- 规格升降级
- 数据分片的用法
- 读写分离
- 一主多备
- 异地容灾
- 诊断实践
- 事件通知

- 定时备份
- 创建任意时间点的临时实例

数据备份	临时实例	归档列表	备份设置
保留天数	7		
备份周期	星期一, 星期二, 星期三, 星期四, 星期五, 星期六, 星期日		
备份时间	16:00-17:00		
预计下次备份时间	2016-04-28 16:34:00		
日志备份	开启		
<a href="#">编辑</a>			

- 使用UDF
- 网络延迟 瓶颈定量分析
- <https://yq.aliyun.com/articles/35176>
- <https://yq.aliyun.com/articles/35517>

- 查询当前RDS PG版本支持的插件名

```
select setting from pg_settings where name = 'rds_available_extensions';
```

```
plpgsql, pg_stat_statements, btree_gin, btree_gist, chkpass, ci_text, cube, dblink, dict_int, earth  
distance, hstore, intagg, intarray, isn, ltree, pgcrypto, pgrowlocks, pg_prewarm, pg_trgm, postgr  
es_fdw, sslinfo, tablefunc, tsearch2, unaccent, postgis, postgis_topology, fuzzystrmatch, postgi  
s_tiger_geocoder, plperl, pltcl, plv8, plls, plcoffee, "uui d-  
ossp", zhparser, pgrouting, rdkit, pg_hint_plan, pgstattuple
```

- 创建插件
- create extension 插件名;
- 插件的用法，参考插件对应的手册进行操作



- auto\_explain 插件用法

```
alter role all set session_preload_libraries='auto_explain';  
alter role all set auto_explain.log_min_duration to '1s';
```

- <http://www.postgresql.org/docs/9.5/static/auto-explain.html>
- pg\_hint\_plan 插件的用法
- <https://yq.aliyun.com/articles/17212>

- 让所有用户可以使用pg hint plan插件。

```
alter role all set session_preload_libraries=pg_hint_plan;
```

- 重连数据库后即可使用
- 设置软只读用户

```
alter role 角色名 set transaction_read_only=off;
```

- 该用户重连数据库后，默认会开启只读事务。
- 设置适合RDS IO能力的参数

```
alter role all set random_page_cost=1.2;          ## 老版本未调整这个参数
```

- 设置使用异步提交

```
alter role all set synchronous_commit=off;      ## 降低IO瓶颈带来的影响
```

- 设置TCP心跳参数

```
alter role all set tcp_keepalives_idle=60;  
alter role all set tcp_keepalives_interval=10;  
alter role all set tcp_keepalives_count=10;
```

- 登陆RDS管理控制台，设置安全和审计
- 白名单为允许连接数据库的来源IP
- SQL审计
- 需要用户主动开启

您还未开启SQL审计，请 [立即开启](#)

数据安全性

白名单设置

SQL 审计

- default

0.0.0.0/0

注：IP白名单设置为0.0.0.0/0代表允许所有

# 规格升降级

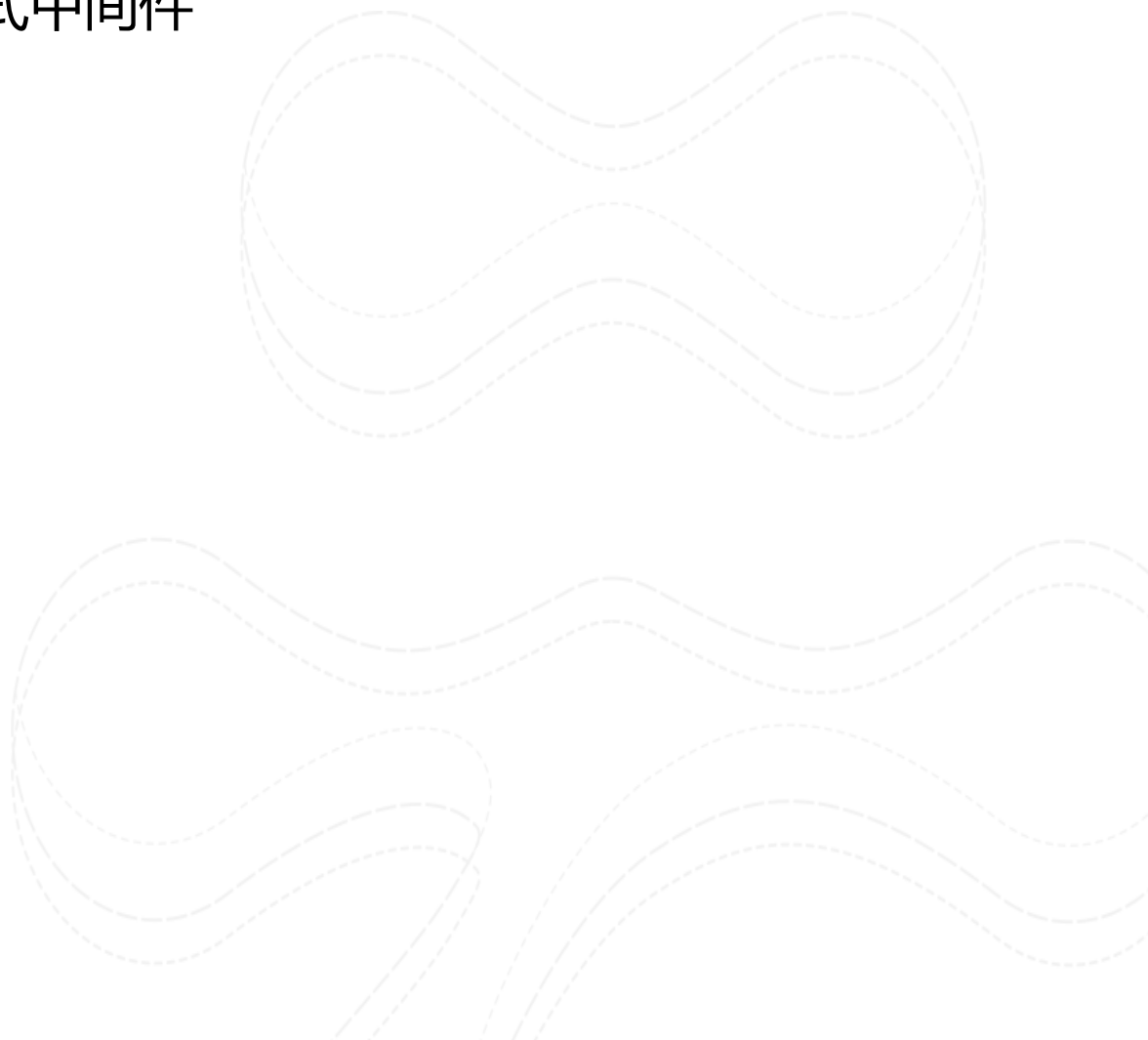
- 登陆RDS管理控制台
- 基本信息
- 变更配置



# 分布式(sharding)用法

- 用户在ECS上安装分布式中间件
- 根据中间件的手册配置
- 支持的中间件包括
  - plproxy
  - oneproxy for postgresql
  - pgpool-II
  - pg\_shard
  - citusdb

- 用户在ECS上安装分布式中间件
- 根据中间件的手册配置
- 支持的中间件包括
  - pgpool-II
  - oneproxy for postgresql



说明：RDS PG内部一主多备功能正在建设

- 用户自建RDS PG的逻辑备库
  - 用户可以使用RDS PG提供的alidecode插件，在异地创建RDS PG的逻辑备库
  - alidecode用法参考RDS PG手册
- 用户自建RDS PG的物理备库
  - <https://yq.aliyun.com/articles/7255>

说明：RDS PG内部异地容灾功能正在建设

- 用户自建RDS PG的逻辑备库
  - 用户可以使用RDS PG提供的alidecode插件，在异地创建RDS PG的逻辑备库
  - alidecode用法参考RDS PG手册
- 用户自建RDS PG的物理备库
  - <https://yq.aliyun.com/articles/7255>



- RDS管理控制台
- 设置报警阈值
- 设置监控间隔
- 查询慢SQL

## 性能优化

慢日志统计 | 慢日志明细

选择时间范围：2016-04-28 至 2016-04-28

总执行次数最多 ▼ 查询

时间	SQL 语句	总执行次数
		总执行次数最多的SQL
		总执行时间最多的SQL
		总逻辑读最多的SQL
		总物理读最多的SQL

## 监控与报警

监控 | 报警

### 监控项

磁盘空间使用率

IOPS使用率

连接数使用率

CPU使用率

- RDS管理控制台
- 日志管理

日志管理

错误日志

选择时间范围：  至

采集时间	日志内容
2016-04-28 17:24:35	DETAIL: last completed transaction was at log time 2016-04-28 17:24:32.548037+08

- auto explain 插件的使用

```
alter role all set session_preload_libraries='auto_explain, ...';  
alter role all set auto_explain.log_min_duration='1s';  
alter role all set auto_explain.log_analyze=on;  
alter role all set auto_explain.log_buffers=on;  
alter role all set auto_explain.log_timing=on;  
alter role all set auto_explain.log_triggers=on;  
alter role all set auto_explain.log_verbose=on;  
alter role all set auto_explain.log_nested_statements=on;
```

- 超过1秒的查询，记录当时的执行计划，在日志中可以追溯。

- debug
  - 打印SQL的parse, rewritten, plan日志，分析计划相关的问题，一般不需要设置。

```
debug_pretty_print  
debug_print_parse  
debug_print_plan  
debug_print_rewritten
```

- 设置客户端日志打印级别

```
set client_min_messages = debug5;
```

- 一键诊断（功能正在建设）
- 索引推荐（功能正在建设）



- [https://help.aliyun.com/document\\_detail/rds/OpenAPI-manual/RDS-OpenAPI-Invoke/request-structure.html](https://help.aliyun.com/document_detail/rds/OpenAPI-manual/RDS-OpenAPI-Invoke/request-structure.html)

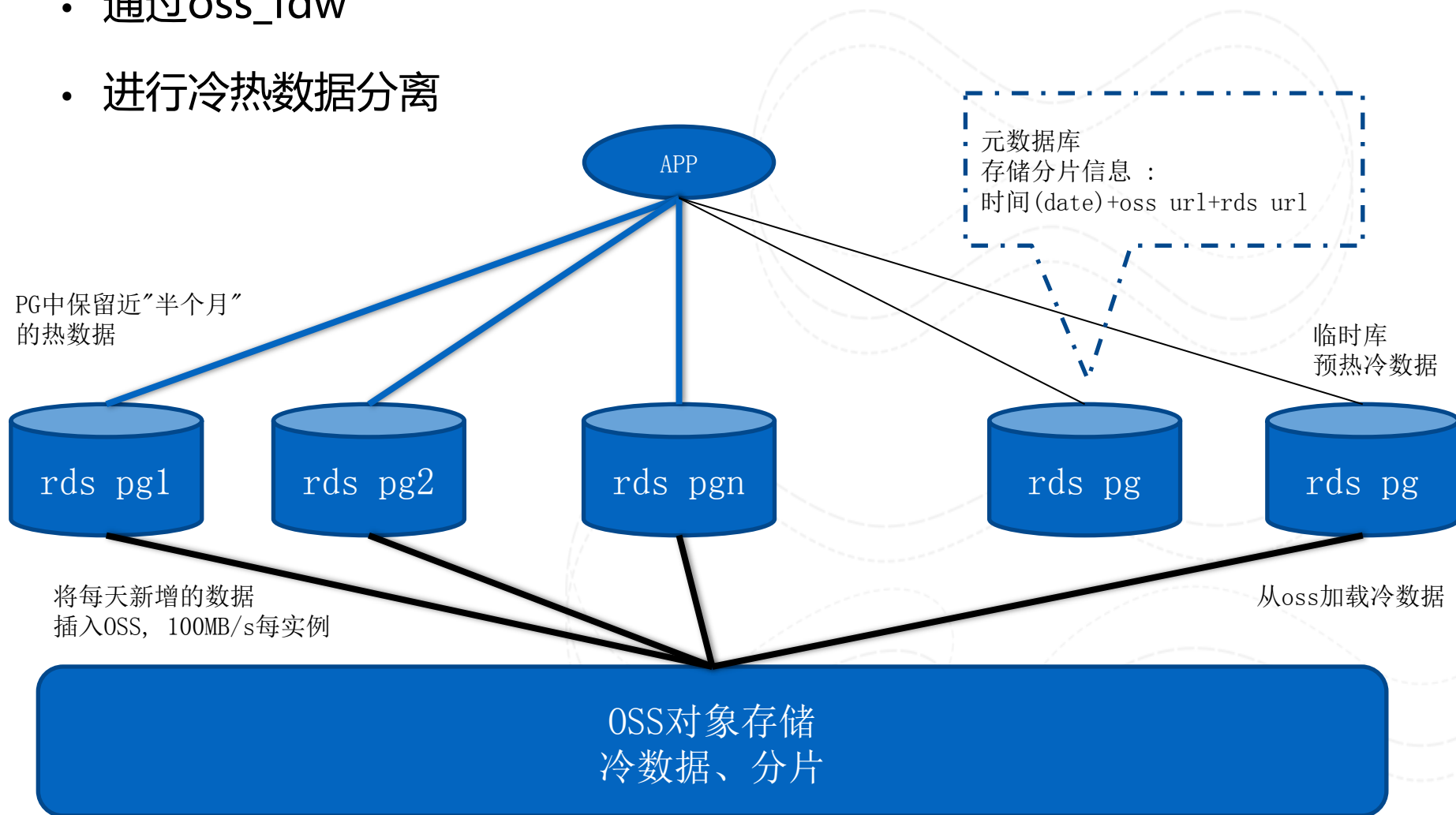
- 管控平台监控通知
- 设置好阈值，接收通知。
- 数据库可用性状态变化时自动通知

- 上云实践
- RDS PG使用最佳实践
- 阿里云RDS相关周边组件用法
- 数据同步



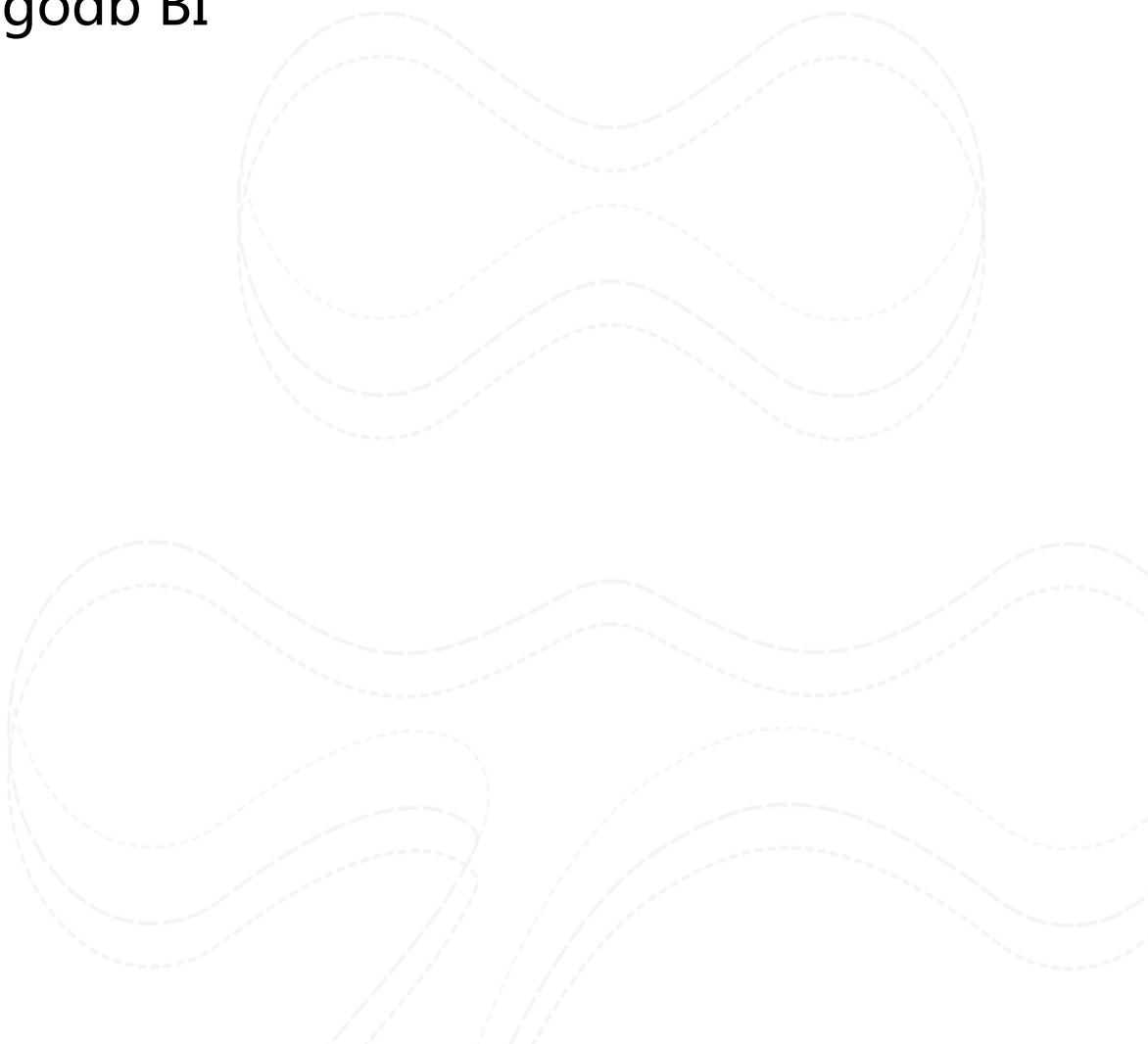


- 通过oss\_fdw
- 进行冷热数据分离



- <https://www.aliyun.com/>
- 例如
  - OSS
  - 用户可以将数据打包成CSV或其他格式的文件，上传到OSS，结合PostgreSQL oss\_fdw，用户可以实现线下数据批量上传到RDS PG。

- PostgreSQL 作为 mongodb BI
- mongo\_fdw



- 上云实践
- RDS PG使用最佳实践
- 阿里云RDS相关周边组件用法
- 数据同步



- 线上线下双向同步
- RDS PG到线下单向同步
- 线下到RDS PG单向同步



- 用户自建RDS PG的逻辑备库
  - 用户可以使用RDS PG提供的alidecode插件，在异地创建RDS PG的逻辑备库
  - alidecode用法参考RDS PG手册
- 用户自建RDS PG的物理备库
  - <https://yq.aliyun.com/articles/7255>

# 线下到RDS PG单向同步

- 线下数据库版本  $\geq 9.4$ 
  - 使用 logical decode
- 线下数据库版本  $< 9.4$ 
  - 使用 goldengate, symmetricds

# 线上线下双向同步

- 同上





# Thanks!



[yq.aliyun.com](http://yq.aliyun.com)

云栖社区，我们的IT江湖

云栖社区是由阿里云负责运营、阿里巴巴技术协会和阿里巴巴集团各技术团队提供内容支持的开放式技术社区