

阿里云 分析型数据库PostgreSQL版 常见问题

文档版本：20190827

法律声明

阿里云提醒您 在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 禁止： 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告： 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明： 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
<code>##</code>	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
<code>[]</code> 或者 <code>[a b]</code>	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
<code>{ }</code> 或者 <code>{a b}</code>	表示必选项，至多选择一个。	<code>swich {stand slave}</code>

目录

法律声明.....	I
通用约定.....	I
1 如何查看本地客户端IP来设置白名单?	1
2 如何查看表和数据库的数据量大小?	3
3 如何从其他数据源同步数据到 AnalyticDB for PostgreSQL?	4
4 如何获取更多技术信息?	5
5 如何选择实例规格?	6
6 磁盘锁定之后如何清理数据?	7
7 如何修改参数?	8
8 如何使用 AnalyticDB for PostgreSQL 列存和压缩功能?	9
9 实例创建失败或长时间处于创建中, 如何处理?	10
10 COPY操作是否需要superuser权限.....	11
11 为什么无法看到按量付费实例的账单? 何时能够开具发票?	12
12 如何终止异常SQL执行?	13
13 如何查看表的分布键定义和表的分区定义?	14
14 如何获得在线技术支持?	15
15 实例“节点”规格与原“计算组”规格的对应关系?	16

1 如何查看本地客户端IP来设置白名单?

问题描述

由于网络环境复杂多样，用户可能无法正确地找到本地客户端的IP地址来设置实例白名单。本文介绍如何查看本地客户端的IP。

下面以具体操作为例，描述详细的操作过程。

操作步骤

1. 将0.0.0.0/0添加到AnalyticDB for PostgreSQL实例的白名单，具体操作如下：

- a. 登录[云数据库AnalyticDB for PostgreSQL管理控制台](#)。
- b. 选择目标实例所在地域。
- c. 单击目标实例的 ID，进入实例基本信息页面。
- d. 在实例菜单栏中，选择数据安全性，进入数据安全性页面。
- e. 在白名单设置标签页中，单击 default 白名单分组后的修改，进入修改白名单分组页面。
- f. 删除组内白名单中的默认白名单 127.0.0.1，写入白名单地址0.0.0.0/0。



说明：

0.0.0.0/0 允许任何IP访问数据库，将会引入较高的安全风险，请尽快删除。

g. 单击确定，完成白名单设置。

2. 使用客户端连接到AnalyticDB for PostgreSQL实例，参见[#unique_4](#)下载安装psql客户端，使用如下连接语句连接数据库：

```
psql -h yourgpdbaddress.gpdb.rds.aliyuncs.com -p 3432 -d postgres -U gpdbaccount
```

其中，各个参数的定义如下：

- -h：指定主机地址。
- -p：指定端口号。
- -d：指定数据库（默认的数据库是 postgres）。
- -U：指定连接的用户。
- 可以通过 `psql --help` 查看更多选项。在 psql 中，可以执行 `\?` 查看更多 psql 中支持的命令。

3. 进入数据库后，在数据库的SQL命令行窗口中运行如下命令，查询客户端的IP地址。

```
select * from pg_stat_activity;
```

查询结果的CLIENT_ADDR字段即为客户端的IP地址。

```
postgres=> select * from pg_stat_activity;
 datid | datname | procpid | sess_id | usesysid | username |          current_qu
ery | waiting | query_start | client_port | application_name | backend_start
      | client_addr |          |          |          |          | xact_start
-----+-----+-----+-----+-----+-----+-----
-----+-----+-----+-----+-----+-----+-----
-----+-----+-----+-----+-----+-----+-----
10902 | postgres | 45909 | 254025 | 17239 | test123 | select * from pg_st
at_activity; | f | 2019-02-19 14:07:00.37242+08 | 2019-02-19 14:06:26.1704
67+08 | 100.11.34.5 | 31964 | psql | 2019-02-19 14:07:00.37242
+08 |
(1 row)

postgres=> □
```

4. 在AnalyticDB for PostgreSQL控制台中，将白名单0.0.0.0/0删除，输入上个步骤查询到的IP地址，即可正常访问数据库。

2 如何查看表和数据库的数据量大小?

假设表的模式为<schemaname>, 表名为<tablename>。

- 执行以下命令, 查询一张表的总大小 (单位为MB, 包含表的索引和数据) :

```
select pg_size_pretty(pg_total_relation_size('<schemaname>.<tablename>'));
```

- 执行以下命令, 查询表的数据大小 (单位MB, 不包括索引) :

```
select pg_size_pretty(pg_relation_size('<schemaname>.<tablename>'));
```

- 执行以下命令, 查询分区表所有分区的总大小 (单位MB, 包含表的索引和数据) :

```
select schemaname,tablename,round(sum(pg_total_relation_size(schemaname || '.' || partitiontablename))/1024/1024) "MB" from pg_partitions where schemaname='<schemaname>' and tablename='<tablename>' group by 1,2;
```

- 执行以下命令, 查询一个Schema下面的所有表的总大小 (单位MB, 包括索引和数据) :

```
select schemaname ,round(sum(pg_total_relation_size(schemaname || '.' || tablename))/1024/1024) "Size_MB" from pg_tables where schemaname='<schemaname>' group by 1;
```

- 执行以下命令, 查询每个数据库的大小 (单位MB) :

```
select datname,pg_size_pretty(pg_database_size(datname)) from pg_database;
```

3 如何从其他数据源同步数据到 AnalyticDB for PostgreSQL?

支持的 ETL 工具见下，可以同时参见更详细的同步方案列表[#unique_7](#)：

- **数据传输服务 (DTS)**：阿里云提供的实时数据同步服务，可以将其他数据源（RDS MySQL, ECS自建MySQL, PolarDB 等）实时同步数据到 AnalyticDB for PostgreSQL，构建实时数据仓库解决方案。
- **阿里云的数据集成服务(Data Integration)**：阿里云提供的 ETL 工具。在数据集成服务中，将 AnalyticDB for PostgreSQL 配置为一个 PostgreSQL 数据库，即可实现其他数据源（RDS、MaxCompute、TableStore 等）到 AnalyticDB for PostgreSQL 的数据同步。
 - 您可以直接从其他数据源读取数据，写入到 AnalyticDB for PostgreSQL 中。
 - 如果数据量较大，需要并发导入，则建议您先通过数据集成服务把数据从其他数据源导入到 OSS，再通过 OSS 外部表导入 AnalyticDB for PostgreSQL。
- **Pentaho Kettle 数据集成软件**：开源的 ETL 工具。
 - 支持将数据先通过 Kettle 导入到本地磁盘，再通过 COPY 或 OSS 导入到 AnalyticDB for PostgreSQL。
 - 也支持将 OSS 存储挂载为本地虚拟磁盘，通过 Kettle 导入到此磁盘，最后通过 AnalyticDB for PostgreSQL 的 OSS 外部表导入到 AnalyticDB for PostgreSQL 中。
- **Informatica软件**：商业化的 ETL 工具。
- **彩虹桥**：阿里云云市场中的 ETL 工具，提供商业化服务。
- **dbsync**：阿里云提供的开源数据库同步工具。
 - 支持从 MySQL、PostgreSQL 并发同步数据到 AnalyticDB for PostgreSQL。
 - 支持简单的数据转换。
 - 支持通过解析 Binlog，准实时地从 MySQL 同步数据到 AnalyticDB for PostgreSQL。
- 其他支持 Greenplum 的 ETL 工具。

4 如何获取更多技术信息?

1. 云栖社区 - AnalyticDB for PostgreSQL产品云栖号: [阿里云数据仓库云栖号](#)
2. 通过[工单系统](#)向我们咨询。
3. Greenplum Database开源社区官方资料: <http://www.greenplum.org>
4. 产品技术支持钉钉群 (群号: 11700737) : 技术专家在线支持, 新特性预告, 优惠活动发布
[强烈推荐!]



在线技术支持钉钉群

5 如何选择实例规格?

请参见[实例规格和选型指导](#)

6 磁盘锁定之后如何清理数据?

磁盘满判定条件

任何一个计算组磁盘空间达到限制或者主节点（Master）磁盘空间达到限制，就会锁定整个实例。

检查实例是否锁定

直接连接数据库，执行以下命令，检查实例是否被锁定：

```
show rds_force_trans_ro_non_sup;
```

根据返回结果，如果 `rds_force_trans_ro_non_sup` 的值为 `on`，表示实例已被锁定，数据库已处于只读状态。

如何清理数据

当实例因磁盘满被锁定时，AnalyticDB for PostgreSQL 仍然支持对数据表进行 `truncate/drop/grant` 操作。清理数据后，如果磁盘空间降到阈值以下，过大约5分钟实例会自动解锁。



注意：

实例锁定后，不支持 `delete` 操作。因为 `delete` 会写 `xlog`，使空间占用增大。

另外，您可以执行以下语句来查询表的大小：

```
select pg_size_pretty(pg_total_relation_size('test'));
```

7 如何修改参数?

AnalyticDB for PostgreSQL 用户可以修改连接会话内的参数配置（具体参考 Greenplum 的参数修改限制）。

您可以使用命令 `ALTER ROLE <用户名> SET <参数名> = <参数值>` 来修改参数。该参数修改只针对指定的用户生效。

如果要修改数据库实例级别参数，需要提交工单申请，或加入在线支持钉钉群向在线支持工程师申请。工单申请网址和在线支持钉钉群请参见[#unique_12](#)。

8 如何使用 AnalyticDB for PostgreSQL 列存和压缩功能?

在 AnalyticDB for PostgreSQL 中创建表时，默认使用行式存储，并且不启用压缩。如需使用列存和压缩功能，您必须在建表时指定列存和压缩选项。例如，您可以在建表语句中加入以下子句，来启用列存和压缩功能。

```
with (APPENDONLY=true, ORIENTATION=column, COMPRESSTYPE=zlib,  
      COMPRESSLEVEL=5, BLOCKSIZE=1048576, OIDS=false)
```

一般情况下，推荐您使用列存和压缩功能，尤其是在包含较多复杂查询，或需要降低存储成本的场景。



注意：

目前，AnalyticDB for PostgreSQL 仅支持 zlib 和 RLE_TYPE 压缩算法；如果指定了 quicklz 算法，会自动转换为 zlib。

9 实例创建失败或长时间处于创建中，如何处理？

实例创建后，一般在20分钟内能够正常运行。新建实例长时间处于创建中，一般是由于后端资源不足导致的，您可以[提工单](#)请求后端补充资源，或者更换其他可用区尝试购买。也可以登录产品的技术支持钉钉群，联系值班同学寻求帮助 [技术支持钉钉群链接](#)

10 COPY操作是否需要superuser权限

问题描述

在AnalyticDB for PostgreSQL上通过执行CREATE ROLE创建了一个普通用户gpuser01, 但是在执行COPY脚本的时候, 提示ERROR: must be superuser to COPY to or from a file。

如何操作才能将普通用户权限提升为superuser?

问题解答

- AnalyticDB for PostgreSQL不提供superuser权限, 更多信息参考[#unique_16](#)。
- 可以参考以下两个方法执行数据COPY到文件的操作。

```
psql -c 'copy xx to stdout' > file
```

或

```
cat file | psql -c 'copy xx from stdin'
```

11 为什么无法看到按量付费实例的账单？何时能够开具发票？

按量付费实例的月度账单在次月的2日输出。例如，2月份的账单，需要在3月2日之后才能查询到。输出月度账单后，才能开具发票，即发票也是在次月2日才能开具。

12 如何终止异常SQL执行?

若需要终止特定的SQL或会话来恢复系统状态，先通过`pg_stat_activity`系统视图查询当前的查询：

```
select current_query,procpid from pg_stat_activity ;
```

current_query	procpid
select current_query,procpid from pg_stat_activity; SELECT xxx	32584 32238



说明：

`current_query`为当前正在进行的查询，`procpid`为后台进程的PID。

终止非当前连接的SELECT：

```
SELECT pg_cancel_backend(pg_stat_activity.procpid)
FROM pg_stat_activity
WHERE procpid <> pg_backend_pid() and current_query like 'SELECT
%';--注意SELECT的大小写
```

```
pg_cancel_backend
-----
t
(1 rows)
```

或

```
SELECT pg_terminate_backend(pg_stat_activity.procpid)
FROM pg_stat_activity
WHERE procpid <> pg_backend_pid() and current_query like 'SELECT%';
--注意SELECT的大小写
```

```
pg_terminate_backend
-----
t
(1 rows)
```



说明：

`cancel` 和 `terminate` 仅适用于当前用户的查询或权限小于等于当前用户的查询。若提示"`ERROR : must be superuser or rds_superuser to signal other server processes`"，则说明后台有其他用户的连接，但不影响`cancel` 和 `terminate` 的执行。

13 如何查看表的分布键定义和表的分区定义?

如何查看分布键

- 使用psql查看

```
\d tblname
```

- 使用SQL查看

```
SELECT attname FROM pg_attribute WHERE attrelid='schemaname.tblname'::regclass and attnum in (SELECT unnest(attrnums) FROM pg_catalog .gp_distribution_policy t WHERE localoid='schemaname.tblname'::regclass);
```



说明:

- schemaname为模式名称。
- tblname为表名称。

如何查看分区定义

- 使用psql查看

```
\d+ tblname
```

- 使用SQL查看

```
SELECT pg_get_partition_def('schemaname.tblname'::regclass,true);
```

- schemaname为模式名称。
- tblname为表名称。

14 如何获得在线技术支持?

请加入如下钉钉群（群号：11700737），产品技术专家提供在线支持服务，同时定期发布优惠活动，及新特性预告！



在线技术支持钉钉群

15 实例“节点”规格与原“计算组”规格的对应关系？

AnalyticDB for PostgreSQL 在 2019年8月23日，实例规格定义变更，集群实例由“计算组”组成的规格定义改为由“节点”组成。本文主要用来描述“节点”和“计算组”的对应关系。

引入节点定义取代计算组定义，主要为了简化 MPP 集群的规格选型，遵循标准集群数据库的定义模式。一个节点只包含一个MPP数据分区，相对可以简化MPP集群的选型和使用，而一个计算组根据其规格可能会包含多个MPP数据分区。

计算组规格（不再使用）	节点	MPP数据分区数	资源合计	按量付费价格
1x2C SSD	2个 1核 SSD节点	2	2核/16GB内存/160GB用户存储/320GB总存储	1.76元/小时
1x16C SSD	16个 1核 SSD节点	16	16核/128GB内存/1280GB用户存储/2560GB总存储	14.08元/小时
4x4C SSD	4个 4核 SSD节点	4	16核/128GB内存/1280GB用户存储/2560GB总存储	14.08元/小时
2x2C HDD	2个 2核 HDD节点	2	4核/32GB内存/2TB用户存储/4TB总存储	4.60元/小时
4x4C HDD	4个 4核 HDD节点	4	16核/96GB内存/8TB用户存储/16TB总存储	18.40元/小时

按如上对应关系，节点和计算组在相同资源和分区数下，性能和容量保持一致，价格保持一致。